

М. ЭЙГЕН, Р. ВИНКЛЕР

ИГРА ЖИЗНИ

Перевод с немецкого
В. М. АНДРЕЕВА

под редакцией
М. В. ВОЛЬКЕНШТЕЙНА



МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ

1979

28.071
Э 30
УДК 577.3

LUDUS VITALIS

M. Eigen, R. Winkler

Sonderdruck aus
«Mannheimer Forum 73/74»
Erschienen in der
Studienreihe Boehringer Mannheim
Grafik: Erwin Poell, Heidelberg

Игра жизни. Эйген М., Винклер Р. Перевод с немецкого Андреева В. М. под ред. Волькенштейна М. В. Издательство «Наука», Главная редакция физико-математической литературы, 1979 г.

Книга посвящена одному из интереснейших вопросов современной теоретической биофизики — соотношению случайного и закономерного в возникновении упорядоченности живой материи на разных уровнях организации. Все вопросы рассмотрены на высоком научном уровне при ярком и образном изложении.

Табл. 4. Рис. 19. Библ. назв. 40.

Э 21005 — 122
053(02)-79 БЗ-10-22 — 79 2001040000

© Перевод на русский язык,
Главная редакция
физико-математической
литературы
издательства «Наука», 1979

ОГЛАВЛЕНИЕ

| | |
|--|----|
| ПРЕДИСЛОВИЕ РЕДАКТОРА | 4 |
| <i>Глава 1. ЧТО ТАКОЕ ЖИЗНЬ?</i> | 9 |
| <i>Глава 2. КАК ВОЗНИКАЕТ ИНФОРМАЦИЯ?</i> | 23 |
| <i>Глава 3. «ИГРЫ В БИСЕР»</i> | 37 |
| <i>Глава 4. ИГРАЕТ ЛИ БОГ В КОСТИ?</i> | 52 |
| <i>Глава 5. СТУПЕНИ К ДЕЙСТВИТЕЛЬНОСТИ</i> | 70 |
| ЛИТЕРАТУРА | 92 |

ПРЕДИСЛОВИЕ РЕДАКТОРА

Имя Манфреда Эйгена, иностранного члена Академии наук СССР, хорошо известно советскому читателю. Эйген был удостоен Нобелевской премии за создание и разработку новых методов химической кинетики — релаксационных методов. В последнее десятилетие научная деятельность Эйгена посвящена фундаментальным вопросам добиологической и биологической эволюции, вопросам возникновения и развития жизни на Земле. Ранее на русском языке были опубликованы первая работа Эйгена, трактующая эти вопросы [1], и ее сжатое изложение [2]. В данной книге, написанной Эйгеном вместе с его сотрудницей Рутхильд Винклер, обсуждаются те же проблемы. В отличие от работ [1, 2] здесь не приводится математический формализм теории самоорганизации биологически функциональных макромолекул, но дается лишь качественное описание этой теории. Относящиеся к ней тонкие и сложные вопросы проанализированы на основе игровых моделей с привлечением понятий современной теории информации. В дальнейшем Эйген и Винклер посвятили обширную монографию применениям теории игр в биологии и социологии [3].

Об игровой модели развивающейся Природы писал уже Гёте в 1809 г.: «Представьте себе природу, которая как бы стоит у игорного стола и неустанно выкрикивает: *au double!* (удваиваю! — М. В.), т. е. пользуясь уже выигранным, счастливо, до бесконечности продолжает игру сквозь все области своей деятельности. Камень, животное, растение — все после таких счастливых ходов постоянно снова идет на ставку, и кто знает, не является ли весь человек, в свою очередь, только ставкой на высшую цель?» (цитируется по [4]).

Гёте по существу указал в этих словах на то, что уже возникшая информация может сохраняться в ходе дальнейшего развития.

Игры, моделирующие явления жизни, описаны в книге популярно, ярко и убедительно. Однако в целом маленькая книжка Эйгена и Винклер требует от читателя серьезных усилий. В ней затронуты — зачастую бегло — очень важные вопросы естествознания, имеющие в конечном счете и общефилософское значение. Читая эту книжку, нужно размышлять, причем не обязательно во всем соглашаться с автором.

Задолго до Эйгена самоорганизацию биологических макромолекул рассматривал Кастлер [5]. Именно Кастлер впервые отметил различие между возникновением информации и выявлением уже имеющейся информации из маскирующих ее шумов. Очевидно, что для биологии существенно и то, и другое.

Можно оценить количество информации, содержащееся в молекуле ДНК или даже в организме в целом. Такие оценки имеются в работе Кастлера [5] и в монографии Блюменфельда [6]. Однако, как об этом говорится и в данной книге, для биологии существенно не столько количественная, сколько качественная характеристика информации — ее ценность. Впервые на это указал Шмальгаузен в своем посмертно изданном труде [7]. Проблема ценности информации непосредственно связана с рецепцией информации — ценность может определяться только по последствиям рецепции. Тем самым возможно лишь условное, относительное определение ценности. Можно определить ценность элемента информационного сообщения как степень его незаменимости. Ряд фактов свидетельствует о том, что ценность информации, определенная таким образом, возрастает в ходе биологического развития — и в филогенезе, и в онтогенезе. Вопросы ценности информации в связи с биологией рассмотрены в работах [8—10]. В книге [10] содержится также изложение и критический анализ теории Эйгена.

Теория эта — крупное событие в науке. Эйген разработал содержательную модель добиологической эволюции, позволяющую понять возникновение порядка из беспорядка. Для этой модели не потребовались какие-либо новые физические принципы — уже установленные понятия физики и теории информации

оказываются достаточными для моделирования и, тем самым, понимания фундаментальных биологических процессов. Конечно, Эйген далек от мысли, что предложенная им модель воспроизводит действительные события, происходившие на Земле, когда на ней зарождалась жизнь. Задача модельной теории Эйгена состоит в другом — в том, чтобы доказать возможность самоорганизации материи на основе существующих принципов физики. Эта задача решена Эйгеном. Книжка, предлагаемая вниманию читателя, также служит вкладом в решение этой задачи.

В последнее время Эйген опубликовал «трилогию», посвященную дальнейшему развитию теории и называемую «Гиперцикл. Принципы естественной самоорганизации» [12, 13]. Публикация эта завершена лишь недавно. В двух обширных сообщениях описаны интересные результаты. Приводим их краткое содержание:

1. Мишенью отбора и эволюции является квазивид, представленный распределением генотипически родственных реплицируемых единиц. В центре этого распределения находится копия (или вырожденный набор копий), отвечающая фенотипу с максимальной селективной ценностью.

2. Содержание информации в этой главной копии, выраженное числом символов (нуклеотидов), приходящимся на реплицируемую единицу, ограничено. Превышение этого порога содержания информации приведет к катастрофе, т. е. к распаду информации вследствие постоянного накопления ошибок.

3. Для достижения устойчивого количества информации, отвечающего нескольким тысячам нуклеотидов, необходима высокоразвитая система ферментативной репликации. Несколько тысяч нуклеотидов достаточно для кодирования нескольких молекул белков, как о том свидетельствуют современные РНК-фаги. Физические свойства, присущие нуклеиновым кислотам, допускают воспроизводимое накопление информации порядка не более, чем 50—100 нуклеотидов.

4. Для характеристики особого класса самореплицируемых реакционных сетей, гиперциклов, применимы топологические методы. Свойства гиперциклов оказываются достаточными для устойчивой интеграции информации, содержащейся во многих самореплицируемых единицах.

5. Эти свойства гиперциклов состоят в следующем: а) гиперцикл обеспечивает устойчивое и контролируемое сосуществование всех видов, соединенных циклическими связями; б) гиперцикл допускает когерентный рост всех его членов; в) гиперцикл конкурирует с любой отдельной репликативной единицей, не принадлежащей к циклу, независимо от того, является ли эта единица самостоятельной или частью другого гиперцикла или даже связанной с данным циклом «паразитной» связью; г) гиперцикл может увеличиваться или уменьшаться, если эти изменения дают селекционные преимущества; д) гиперциклы нелегко связываются в сети более высокого порядка.

6. Внутренние связи и кооперативные свойства гиперцикла могут развиваться для оптимального функционирования. «Фенотипические» преимущества, т. е. те вариации, которые дают прямые преимущества мутанту, немедленно стабилизируются. С другой стороны, «генотипические» преимущества требуют пространственного разделения для конкретной фиксации.

7. Отбор гиперцикла есть необратимое событие. В любой обычной дарвиновской системе мутанты, имеющие селективные преимущества, могут легко размножаться и стабилизироваться. Их рост независим от размера популяции. Для гиперциклов селективные преимущества всегда являются функциями численности популяции вследствие принципиально нелинейных свойств гиперциклов. Поэтому гиперцикл, однажды возникший, не может быть легко заменен новой системой, так как новый вид всегда возникает в форме одной или немногих копий.

Таким образом, гиперциклы, введенные в теории Эйгена и описанные также и в данной книге, образуют особый класс самоорганизующихся химических цепей. Эйген указывает, что такой тип функциональной организации может быть широко распространен и, возможно, играет роль в нейронных сетях или в социальных системах. С другой стороны, Эйген подчеркивает, что не следует фетишизировать гиперцикл. Он мог исчезнуть в ходе эволюции после того, как ферментативные системы с высокой точностью репродукции смогли индивидуализировать интегральную систему в форме живой клетки.

В обширной трилогии «Гиперцикл» проводится подробный математический анализ проблемы, на котором мы не можем здесь останавливаться.

В последней статье трилогии [13] модель гиперцикла применима уже к конкретному рассмотрению реалистического гиперцикла, фигурировавшего в добиологической и биологической эволюции. В этой содержательной работе Эйген и Шустер приходят к выводам о природе перехода от неживой к живой системе.

1. Появление макромолекул диктуется их структурной устойчивостью и содержанием мономеров. На ранней стадии наличествовали неопределенные белковоподобные цепи и немногие РНК-подобные полимеры, способные к репликации.

2. Состав первых полинуклеотидов диктовался концентрациями мономеров. Репродукция последовательностей зависит от точности копирования, которая выше у ГЦ-последовательностей. Репродуцируемые последовательности образовывали распределение квазивидов.

3. Возникали адапторы аминокислот. Первые аминокислоты связывались с адапторами в соответствии со своими концентрациями. Продукты трансляции состояли главным образом из Гли и Ала.

4. Становились доминирующими те продукты трансляции, которые катализировали репликацию собственных матричных РНК. Процесс триггерно усиливался некоторыми неинструментированными белками среды.

5. Мутанты доминирующих м-РНК интегрировались в цикл репродукции, если это давало преимущество. Строилась гиперциклическая организация, основанная на примитивном генетическом коде.

6. Эта организация развивалась далее, используя благоприятные фенотипические изменения. Для благоприятного отбора соответствующих генотипов оказалось необходимым пространственное разделение.

Таким образом, речь идет о живой развивающейся теории, само возникновение которой означает создание новой незаменимой информации.

Текст снабжен немногими примечаниями редактора. Что касается более подробных сведений о нуклеиновых кислотах и белках, то читатель найдет их в ряде книг [6, 10, 14, 15].

«Однако что же мы назовем простою игрою теперь, когда мы знаем, что из всех состояний человека именно игра и только игра делает его совершенным и сразу раскрывает его двойственную природу?»

Фридрих Шиллер. Письма об эстетическом воспитании человека. Письмо 15.— Собрание сочинений.— М.: Гос. изд. худ. лит., 1957, т. 6.

Глава 1

ЧТО ТАКОЕ ЖИЗНЬ?

В эссе «Что такое жизнь?», написанном в 1943 г., Эрвин Шредингер пришел к оптимистическому, хотя и не вполне успокоительному выводу: «Хотя современные физика и химия не могут объяснить эти процессы (происходящие в живом организме), нет никаких оснований сомневаться в возможности их научного объяснения». На основе знаний, накопленных при попытках дать определение таким фундаментальным понятиям, как пространство, время и материя, физики все дальше и дальше удалялись от вопроса «Что такое...?» То, что их интересует, есть не столько тавтология понятий, стоящая за определением, сколько отношения, с помощью которых положение вещей можно было бы свести к немногим фундаментальным величинам, однозначно характеризующим предписанными процедурами измерения.

Понятие *жизнь* охватывает множество явлений, имеющих очень различные степени сложности. Как известно, многие биологи именно в сложности и многообразии видят один из существенных признаков живого. Философов проблема *жизни* интересует прежде всего потому, что она касается происхождения и бытия человека. Но свойство «быть живым» без сомнения является общим для двух столь различных существ, как кишечная палочка и человек. Вряд ли общее определение может сообщить что-либо о крайне различных особенностях этих существ. Если мы хотим понять *жизнь* как особую форму организации мате-

В обширной трилогии «Гиперцикл» проводится подробный математический анализ проблемы, на котором мы не можем здесь останавливаться.

В последней статье трилогии [13] модель гиперцикла применима уже к конкретному рассмотрению реалистического гиперцикла, фигурировавшего в добиологической и биологической эволюции. В этой содержательной работе Эйген и Шустер приходят к выводам о природе перехода от неживой к живой системе.

1. Появление макромолекул диктуется их структурной устойчивостью и содержанием мономеров. На ранней стадии наличествовали неопределенные белковоподобные цепи и немногие РНК-подобные полимеры, способные к репликации.

2. Состав первых полинуклеотидов диктовался концентрациями мономеров. Репродукция последовательностей зависит от точности копирования, которая выше у ГЦ-последовательностей. Репродуцируемые последовательности образовывали распределение квазивидов.

3. Возникали адапторы аминокислот. Первые аминокислоты связывались с адапторами в соответствии со своими концентрациями. Продукты трансляции состояли главным образом из Гли и Ала.

4. Становились доминирующими те продукты трансляции, которые катализировали репликацию собственных матричных РНК. Процесс триггерно усиливался некоторыми неинструментированными белками среды.

5. Мутанты доминирующих м-РНК интегрировались в цикл репродукции, если это давало преимущество. Строилась гиперциклическая организация, основанная на примитивном генетическом коде.

6. Эта организация развивалась далее, используя благоприятные фенотипические изменения. Для благоприятного отбора соответствующих генотипов оказалось необходимым пространственное разделение.

Таким образом, речь идет о живой развивающейся теории, само возникновение которой означает создание новой незаменимой информации.

Текст снабжен немногими примечаниями редактора. Что касается более подробных сведений о нуклеиновых кислотах и белках, то читатель найдет их в ряде книг [6, 10, 14, 15].

«Однако что же мы назовем простою игрою теперь, когда мы знаем, что из всех состояний человека именно игра и только игра делает его совершенным и сразу раскрывает его двойственную природу?»

Фридрих Шиллер. Письма об эстетическом воспитании человека. Письмо 15.— Собрание сочинений.— М.: Гос. изд. худ. лит., 1957, т. 6.

Глава 1

ЧТО ТАКОЕ ЖИЗНЬ?

В эссе «Что такое жизнь?», написанном в 1943 г., Эрвин Шредингер пришел к оптимистическому, хотя и не вполне успокоительному выводу: «Хотя современные физика и химия не могут объяснить эти процессы (происходящие в живом организме), нет никаких оснований сомневаться в возможности их научного объяснения». На основе знаний, накопленных при попытках дать определение таким фундаментальным понятиям, как пространство, время и материя, физики все дальше и дальше удалялись от вопроса «Что такое...?» То, что их интересует, есть не столько тавтология понятий, стоящая за определением, сколько отношения, с помощью которых положение вещей можно было бы свести к немногим фундаментальным величинам, однозначно характеризваемым предписанными процедурами измерения.

Понятие *жизнь* охватывает множество явлений, имеющих очень различные степени сложности. Как известно, многие биологи именно в сложности и многообразии видят один из существенных признаков живого. Философов проблема *жизни* интересует прежде всего потому, что она касается происхождения и бытия человека. Но свойство «быть живым» без сомнения является общим для двух столь различных существ, как кишечная палочка и человек. Вряд ли общее определение может сообщить что-либо о крайне различных особенностях этих существ. Если мы хотим понять *жизнь* как особую форму организации мате-

рии, то мы должны помнить, что эта характеристика никоим образом не включает гигантские качественные и количественные различия в уровне организации низших и высших организмов.

Существует ли ясное разграничение живого и неживого? В настоящее время биологи ответят на этот вопрос скорее отрицательно. В самом деле, вряд ли можно четко определить эту границу. Например, вирусная частица в подходящей среде — используя метаболизм клетки хозяина — может размножаться, причем иногда столь необузданно, что весь организм хозяина в результате погибает. Таким образом, вирус обладает качествами, которые в других случаях можно найти только у *живых организмов* — существование которых, как известно, тоже всегда зависит от определенных условий среды. Но, с другой стороны, вирусы — это *частицы*, характеризующиеся определенной формой и составом и способные, как и неорганические вещества, образовывать кристаллическую решетку. Очевидно, что в такой форме они не подпадают ни под одно определение *живого*. Далее, частицы вирусов можно разложить на отдельные молекулярные компоненты и из них снова собрать инфекционные единицы — при этом теряется их первоначальная индивидуальность. *Жизнь и смерть* теряют здесь свой обычный смысл.

В дальнейшем, однако, мы вовсе не должны отказываться от ясности и определенности физической системы понятий. Мы попытаемся сделать наши определения такими, чтобы они относились лишь к системам с *заданной* функциональной организацией, и попытаемся однозначно определить условия существования рассматриваемого уровня организации. Сравнение с известными нам биологическими формами организации должно показать, в какой мере, следуя этим путем, нам удалось ограничить и однозначно охарактеризовать понятие *жизнь*.

В математике принято различать необходимые и достаточные условия. В нашем случае гораздо легче задать необходимые условия — при этом нужно заботиться лишь об обязательности, но не о полноте.

Из анализа доступных нам живых систем мы знаем, что во всех клетках происходит обмен веществ, или, точнее говоря, *метаболизм* (так как прежде всего имеются в виду превращения свободной энергии),

который является *необходимым* условием существования любой формы жизни *). Только постоянно используя приток свободной энергии, система может непрерывно обновляться и этим тормозить свое падение в состояние термодинамического равновесия, которое Эрвин Шредингер [1] метко назвал состоянием смерти. Характерный для процессов жизни динамический порядок может поддерживаться только за счет постоянной компенсации производства энтропии.

Следующим необходимым условием является способность к *самовоспроизведению*. Все молекулы и специфические упорядоченные надмолекулярные структуры, возникшие благодаря межмолекулярным взаимодействиям, имеют ограниченное время жизни из-за теплового движения. Чтобы не потерять накопленную в них информацию, они должны успевать до своего распада построить хотя бы одну идентичную копию, содержащую план строения и функционирования исходной структуры. Любое биологическое упорядочение направляется информацией **).

Инструментирование требует специфических взаимодействий. Конечные значения энергий взаимодействия и возмущения, создаваемые тепловыми флуктуациями, делают совершенно точное воспроизведение принципиально невозможным. *Всегда* существует определенный темп ошибок, или *мутаций*, — наличие этих ошибок является существенным условием возможности эволюционного прогресса.

Конечно, оба последних условия нужно считать необходимыми, только если рассматриваемая система не может постоянно возникать *de novo*. Только для эволюционирующей системы важно сохранять и развивать однажды достигнутое информационное состояние. В этом состоит непосредственная связь со *сложностью* живых систем.

В настоящее время общепризнано, что тремя существенными предпосылками образования живых структур являются *метаболизм, самовоспроизведение и*

*) Авторы называют метаболизмом, в отличие от обмена веществ, обмен веществ с окружающей средой, сопровождающийся изменением свободной энергии. В действительности обмен веществ всегда связан с таким изменением. (Прим. ред.)

**.) Смысл понятия «информация» обсуждается далее. (Прим. ред.)

мутабельность [2]. Эти свойства необходимы, но не достаточны. Количественное теоретическое рассмотрение динамики биологических систем реакций [3] показывает, что *не любой* вид самовоспроизведения и мутабельности может привести к возникновению систем, способных к неограниченной эволюции.

Одно из величайших научных достижений нашего столетия — выяснение молекулярного механизма переноса информации и наследования в живых организмах. Все живые существа используют одну и ту же универсальную схему для кодирования, переноса (и перевода) информации и биосинтеза. В этих процессах доминируют два класса молекул:

нуклеиновые кислоты как носители законодательного начала,

белки как носители исполнительного начала.

Белками определяется вся коррелированная во времени программа синтезов живой клетки. Информация, т. е. план строения клетки и, следовательно, всего организма, заложена в одной молекуле нуклеиновой кислоты. Копии этого плана строения передаются от поколения к поколению. Для дальнейшего нам не потребуются рассматривать детали химического строения и пространственной структуры этих молекул [4]. Сейчас для нас существенна только логика построения системы биологической переработки информации:

1. Нуклеиновые кислоты — это линейные цепные молекулы. Каждая субъединица (отдельный нуклеотид) представляет собой букву генетического алфавита. Этот алфавит состоит только из четырех букв, которые принято обозначать первыми буквами соответствующих химических названий

А, У (соответственно Т)*), Г, Ц.

Триплет, составленный из этих букв (например АУГ), является кодовым словом (кодоном) для элемента молекулы белка, т. е. для буквы белкового алфавита.

2. Белки — это тоже цепные молекулы. Их субъединицами являются аминокислотные остатки, соединенные друг с другом пептидными связями. В природных белках встречается около 20 различных аминокислот.

*) Элемент У в рибонуклеиновых кислотах (РНК) соответствует элементу Т в дезоксирибонуклеиновых кислотах (ДНК). Для переноса информации, т. е. для взаимодействия с элементом А, это различие не играет большой роли [4].

Это означает, что белковый алфавит содержит примерно двадцать различных букв — примерно столько же, сколько наш язык. Система соответствий «кодон — аминокислота», т. е. генетический код, является однозначной и универсальной: животные, растения и люди пользуются одной и той же схемой. Этот код, однако, «вырожден»: каждой аминокислоте, за немногими исключениями, соответствуют два, три, четыре или даже шесть кодонов. Это необходимо для того, чтобы перевод некоторой последовательности нуклеотидов, где могут встретиться все $4^3 = 64$ различных кодона, в последовательность аминокислот имел однозначный смысл. В природе передача информации происходит лишь в направлении от нуклеиновой кислоты к белку, но между различными типами нуклеиновых кислот (ДНК, РНК) информация может передаваться в обоих направлениях.

3. Полипептидная цепь молекулы белка определенным образом свернута в пространстве. Поэтому элементы из любых положений в последовательности могут в каком-то месте контактировать друг с другом. Таким образом возникает активный центр, т. е. группировка аминокислот, которые кооперативно выполняют свою функцию: например, специфичное узнавание и связывание субстрата, перенос протонов или — при помощи координационно связанного иона металла — обмен электронов и т. д. Оптимальные расстояния в активном центре поддерживаются определенной пространственной укладкой тех участков последовательности, которые располагаются между действующими группами. Это определяет минимальную длину полипептидной цепи, которая для хорошо адаптированного белка с четырьмя — пятью действующими группами составляет примерно сотню аминокислотных остатков. Так используется полное информационное содержание полипептидной цепи, т. е. окончательно оптимизированный эволюционный продукт представляет собой избранную, единственную в своем роде последовательность. сверх того, посредством изменений в укладке, обусловленных тем, что по ходу реакции возникают новые специфичные взаимодействия, каталитическая активность может регулироваться, иногда даже нацело выключаться или включаться. Итак, белки можно сравнить со сложными, регулируемые машинами. Действительно, они

Контролируют все химические превращения в клетке, т. е. все ее функционирование. Однако принцип строения, которым определяются эти достижения белков, совершенно не приспособлен для самовоспроизведения.

4. В отличие от белков нуклеиновые кислоты имеют сравнительно простое пространственное строение. Их

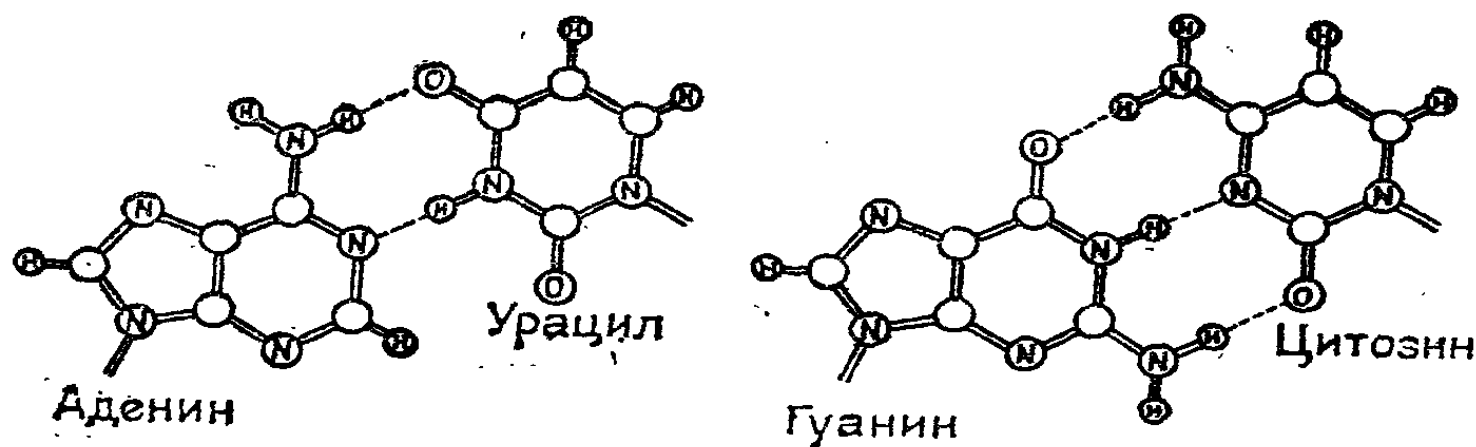


Рис. 1. Комплементарность оснований создается специфическими водородными связями — с одной стороны, между аденином (А) и урацилом (У) или тиминном (Т) и, с другой стороны, между гуанином (Г) и цитозином (Ц). В результате образуются геометрически почти идентичные пары оснований.

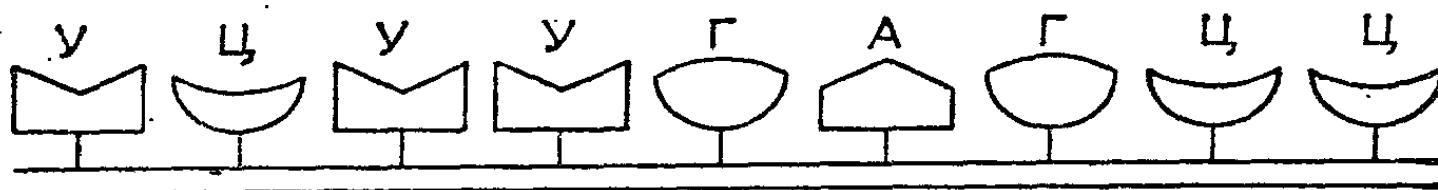
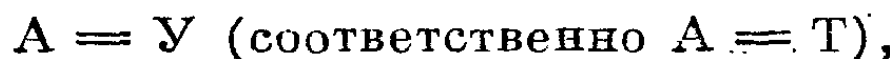


Рис. 2. В молекуле нуклеиновой кислоты отдельные элементы (нуклеотиды) соединены друг с другом в линейную последовательность. Каждое основание прикрепляется к остатку сахара (к рибозе или к дезоксирибозе), который, в свою очередь, висит на фосфатной группе. В основе полинуклеотидной цепи остатки сахара и фосфата непрерывно чередуются.

структурный мотив определяется простым принципом — попарными взаимодействиями звеньев нуклеиновой кислоты, а именно



Элементы А и У(Т), а также Г и Ц называются комплементарными друг другу. Предпочтительные взаимодействия приводят к тому, что молекулярная цепь свертывается таким образом, что комплементарные нуклеотиды оказываются спаренными и располагаются друг против друга (рис. 1—3).

Однако особенно стабильной должна быть двухцепочечная структура, состоящая из двух отдельных,

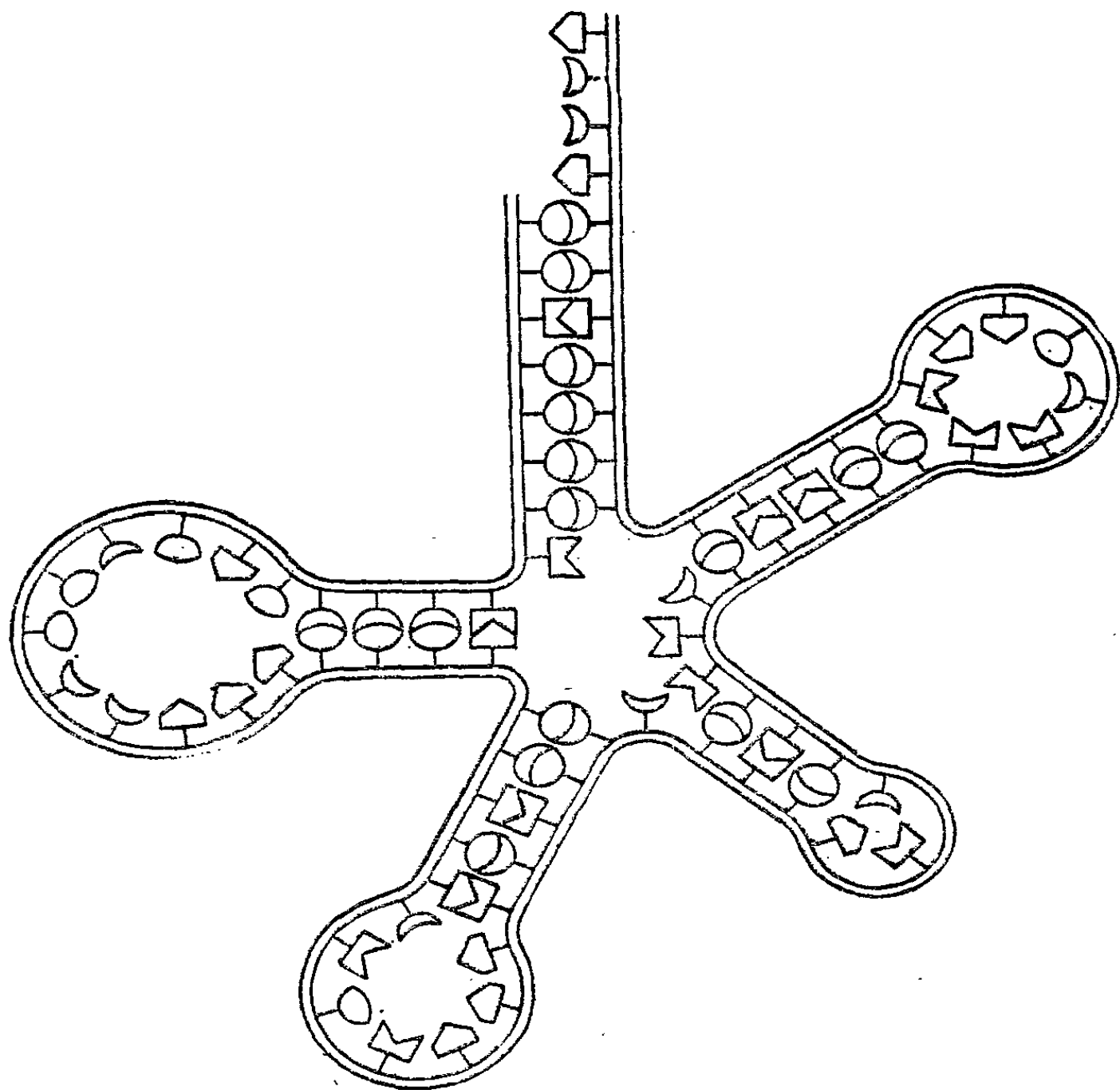


Рис. 3. Две цепи могут соединяться в двухнитевую структуру, скрепленную комплементарными взаимодействиями. Спираль ДНК [5] — это винтообразно скрученная двухцепочечная структура, имеющая очень большую стабильность. Молекулы РНК в большинстве случаев одноцепочечны, однако в них тоже имеются спиральные двухцепочечные участки, образующиеся в результате свертывания цепи. В качестве примера изображена двумерная укладка цепи тирозиновой т-РНК — структура «клеверного листа». В действительности эта молекула не плоская — она имеет определенную трехмерную структуру. Рентгеноструктурный анализ показал, что фенилаланиновая т-РНК имеет вид слегка изогнутой буквы Т (Rich et al.,— Science, 1973, v. 179, p. 285),

полностью комплементарных друг к другу цепей. Примером такой структуры является двойная спираль ДНК по Уотсону и Крику [5]. Узнавание комплемен-

тарных элементов является простым основным принципом воспроизведения нуклеиновых кислот. Нуклеиновые кислоты могут служить поэтому в качестве универсальных накопителей информации, в противоположность белкам. Любая последовательность имеет неотъемлемую способность инструктировать построение своей копии. Процесс копирования — как и при фотографической репродукции — проходит через промежуточную стадию негатива. Точность копирования ограничена. Она определяется разностями энергий взаимодействия для комплементарных и некомплементарных пар. Точность распознавания может быть значительно повышена с помощью специальных копирующих машин — т. е. опять-таки с помощью специально адаптированных белковых структур. На высших ступенях эволюции вероятность ошибки при воспроизведении становится значительно меньше, чем 10^{-9} . По сравнению с этим простое комплементарное взаимодействие в отсутствие копирующих ферментов дало бы вероятность ошибки от 1 до 10%.

5. Бесклеточная система, содержащая существенные компоненты — белки и нуклеиновые кислоты — и снабженная необходимым высокоэнергетическим строительным материалом, может репродуцировать свои составные части *in vitro*. Она проявляет все существенные функции живой клетки, но не является индивидуальным живым существом и не развивается как таковое [6].

Логическое построение этой молекулярной схемы биологической переработки информации в сжатом виде представлено на рис. 1—7 (рис. 4—6 иллюстрируют принцип строения белков. На модели рис. 6 показаны лишь пространственные положения аминокислотных остовов и расположения боковых цепей (R_i). С помощью рентгеноструктурного анализа можно определить положение каждого отдельного атома в этой огромной молекуле с точностью до 2 \AA). Многие детали — как пространственные структуры, так и динамические взаимосвязи — в настоящее время уже известны.

Насколько проста эта грандиозная концепция природы, настолько сложны ее проявления и многообразны те индивидуальные структуры, в которых она реализуется. Это неиссякающий источник информации для неограниченной эволюции: от анаэробной бакте-

рии, через дышащую эукариотическую клетку, ассимилирующую растительную клетку, через дифференцированные и, наконец, управляемые центральной нервной системой организмы к «осознающему себя» человеку, — отсюда эволюция может идти вперед уже на новой основе, управляемая разумом, и этой эволюции нигде не предвидится конца.

С биологической сложностью и, вместе с ней, с индивидуальной неповторимостью структур мы сталкиваемся уже на уровне макромолекул, что является непосредственным следствием именно той концепции, которая была описана выше. В специальной главе мы обсудим, что нужно понимать под информацией. Здесь же мы даем лишь некоторые примеры сложности и уникальности.

Цитохром с, транспортирующий электроны, — одна из самых маленьких белковых молекул, встречающихся

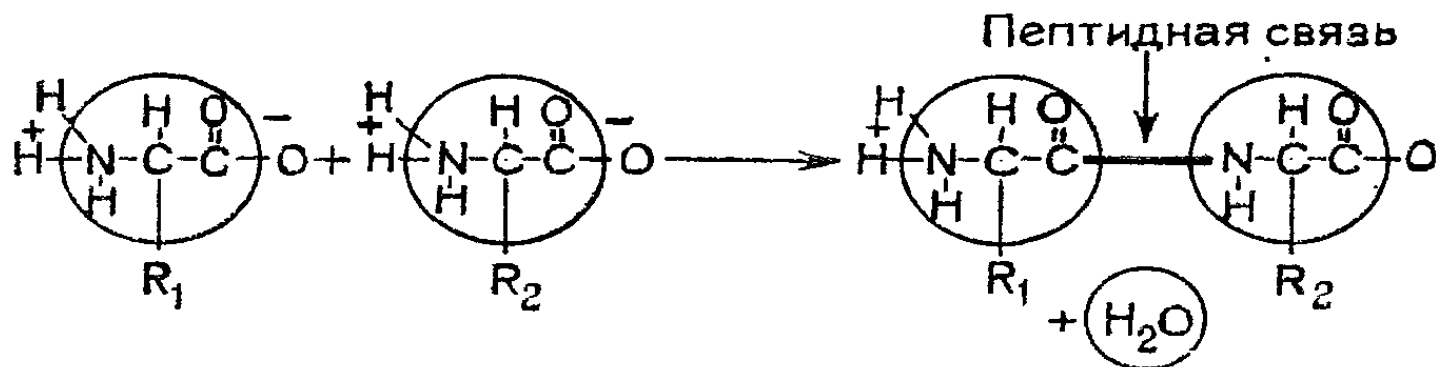


Рис. 4. Две аминокислоты соединяются в дипептид, отщепляя при этом молекулу воды. Индивидуальность аминокислотных блоков, выражающаяся в различиях боковых цепей R_1 и R_2 , остается при этом неизменной.

в природе. Его молекулярный вес равняется примерно 13 000, а полипептидная цепь состоит примерно из ста аминокислотных остатков.

Сколько различных последовательностей аминокислот могут иметь белковые молекулы такой длины? Этот вопрос напоминает старинную детскую песенку «Сколько звездочек на небе...?» Это «ужасно большое число» легко вычислить. Для каждого из ста положений в полипептидной цепи имеется двадцать возможностей. Итак, ответ гласит:

$$20 \cdot 20 \cdot 20 \dots = 20^{100} \approx 10^{130}.$$

Единица со ста тридцатью нулями!

Для реализации всех этих возможностей потребовалось бы в невообразимое число раз больше вещества,

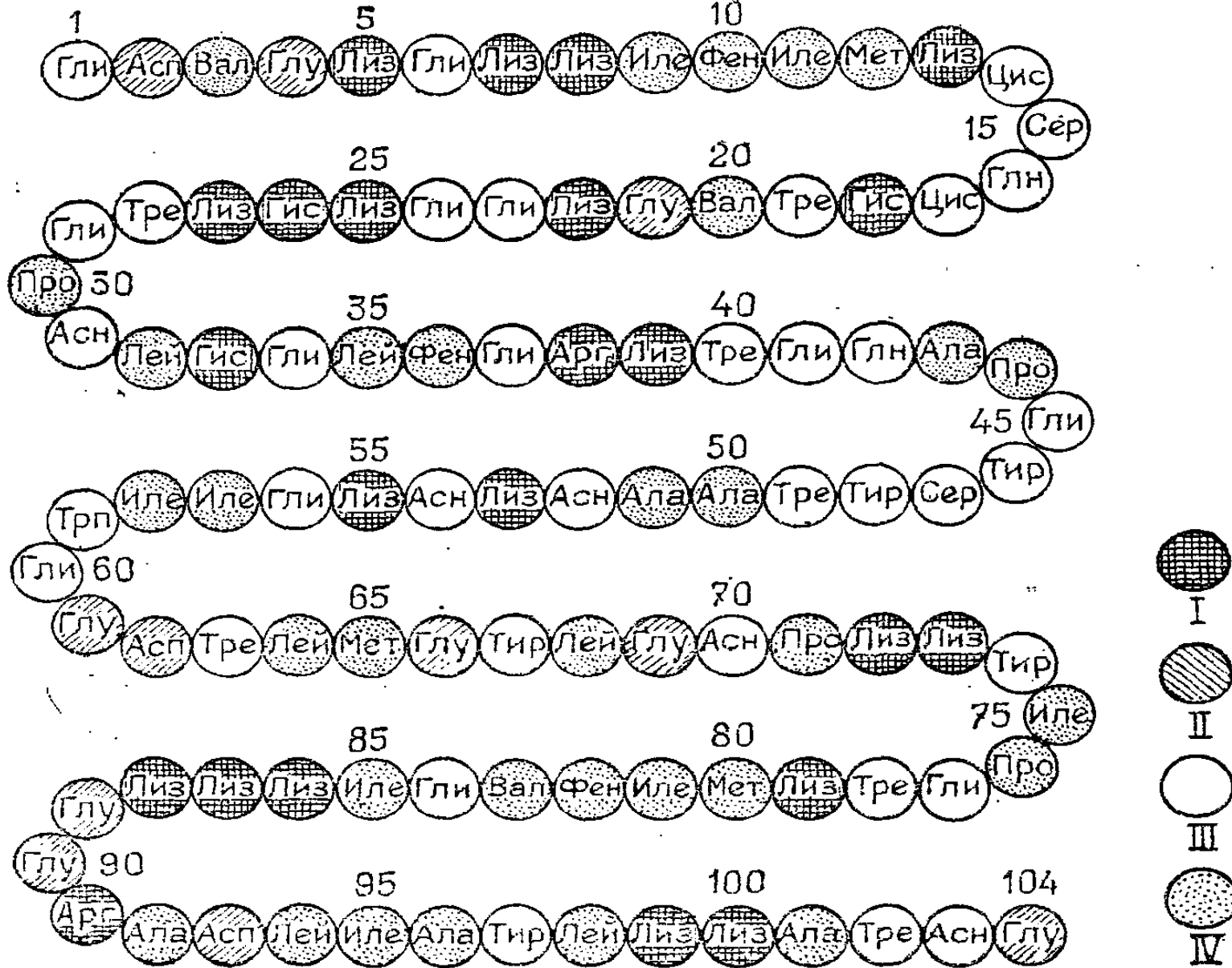


Рис. 5. Многократное повторение процесса, изображенного на рис. 4, приводит к образованию линейной полипептидной цепи (первичная структура белка). В качестве примера дана последовательность цитохрома с.

- Сокращенные обозначения аминокислотных блоков:
- | | |
|------------------------------|--------------------|
| Ала — аланин, | Лей — лейцин, |
| Арг — аргинин, | Лиз — лизин, |
| Асп — аспарагин, | Мет — метионин, |
| Асп — аспарагиновая кислота, | Про — пролин, |
| Вал — валин, | Сер — серин, |
| Гис — гистидин, | Тир — тирозин, |
| Гли — глицин, | Тре — треонин, |
| Глу — глутаминовая кислота, | Трп — триптофан, |
| Глн — глутамин, | Фен — фенилаланин. |
| Иле — изолейцин, | |

Кружками с различной штриховкой обозначены четыре группы боковых цепей: I — положительно заряженные, II — отрицательно заряженные, III — полярные и амбивалентные, IV — неполярные (гидрофобные). (Эта классификация не вполне свободна от произвола. Заряды кислотных и основных групп относятся к рН 7. К категории незаряженных, но полярных относятся аминокислоты с кислотными и основными функциональными группами. Большие участки боковых частей таких остатков могут иметь преимущественно гидрофобный характер. Глицин (R = H) также относится к этой группе, потому что его поведение в основном определяется полярной пептидной связью.)

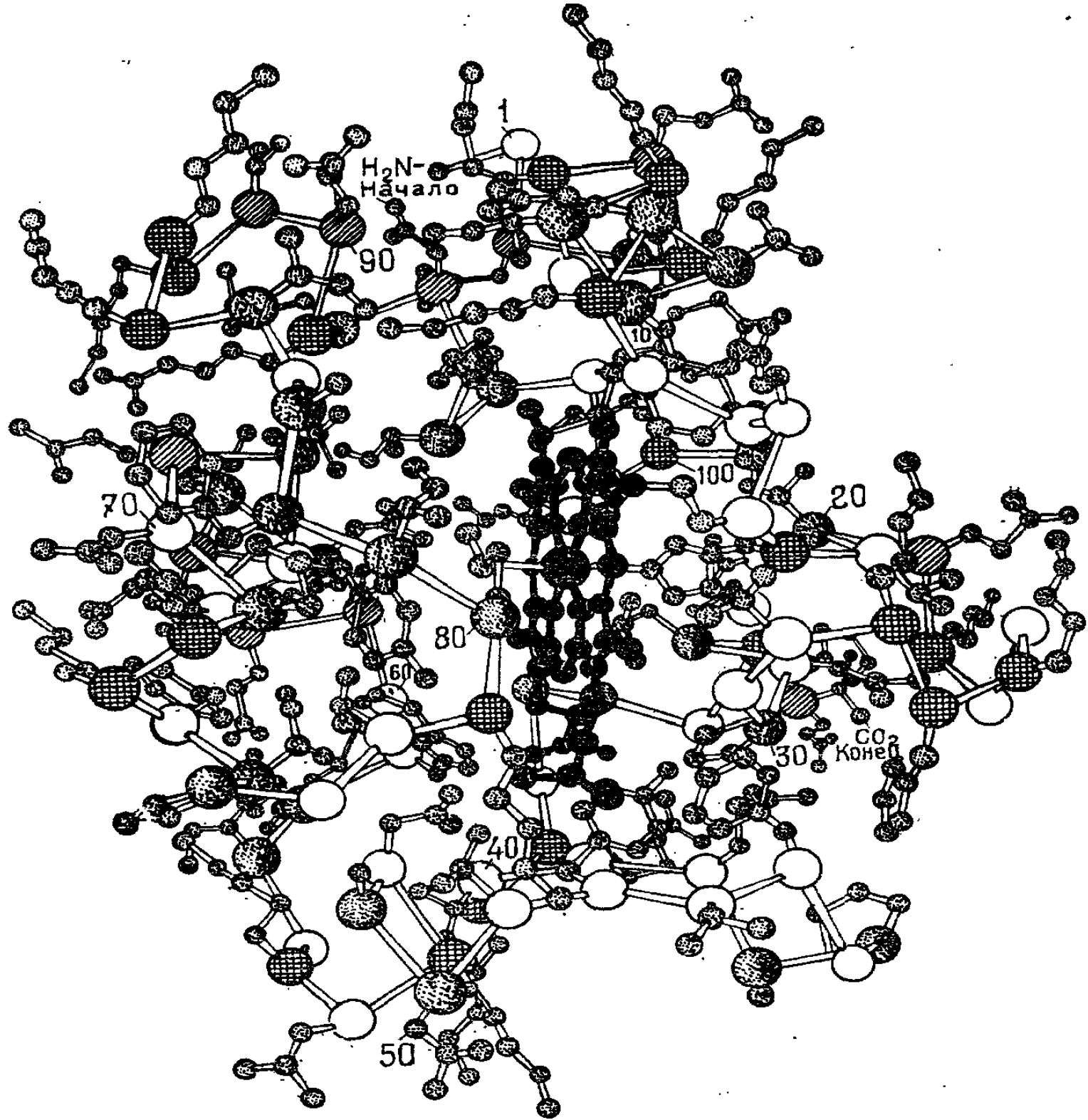


Рис. 6. Линейная полипептидная цепь пространственно свернута. Различают вторичную и третичную структуру. Вторичная структура возникает в результате образования мостиков водородных связей между группами $-\text{CO}-$ и $-\text{NH}-$. Особенно часто образуется «складчатая структура» (два антипараллельных участка полипептидной цепи связываются между собой мостиками $-\text{CO} \dots \text{NH}-$), и α -спираль (полипептидная цепь скручивается в спираль вследствие замыкания водородных связей между группами $-\text{CO}-$ и группами $-\text{NH}-$, удаленными от соответствующих групп $-\text{CO}-$ на четыре аминокислотных остатка). В цитохроме участки цепи 1 — 12 и 89 — 101 имеют спиральную структуру. Еще важнее пространственная, или третичная, структура, которая определяется взаимодействиями между боковыми цепями и создает активный центр. В активном центре цитохрома находится комплекс железа с порфирином, который играет роль донора (Fe^{2+}) или акцептора (Fe^{3+}) электронов. Вследствие строго определенной укладки полипептидной цепи активный центр абсолютно специфичен по отношению к субстрату. Эффективность железного комплекса поддерживается боковыми цепями аминокислот. Изменение пространственной структуры, привело бы к резкому изменению каталитической активности.

чем можно найти на всем «небесном своде». Все вещество во Вселенной, по оценкам физиков, эквивалентно «только» 10^{78} атомам водорода. Для сравнения: масса Земли соответствует 10^{51} атомам водорода. За

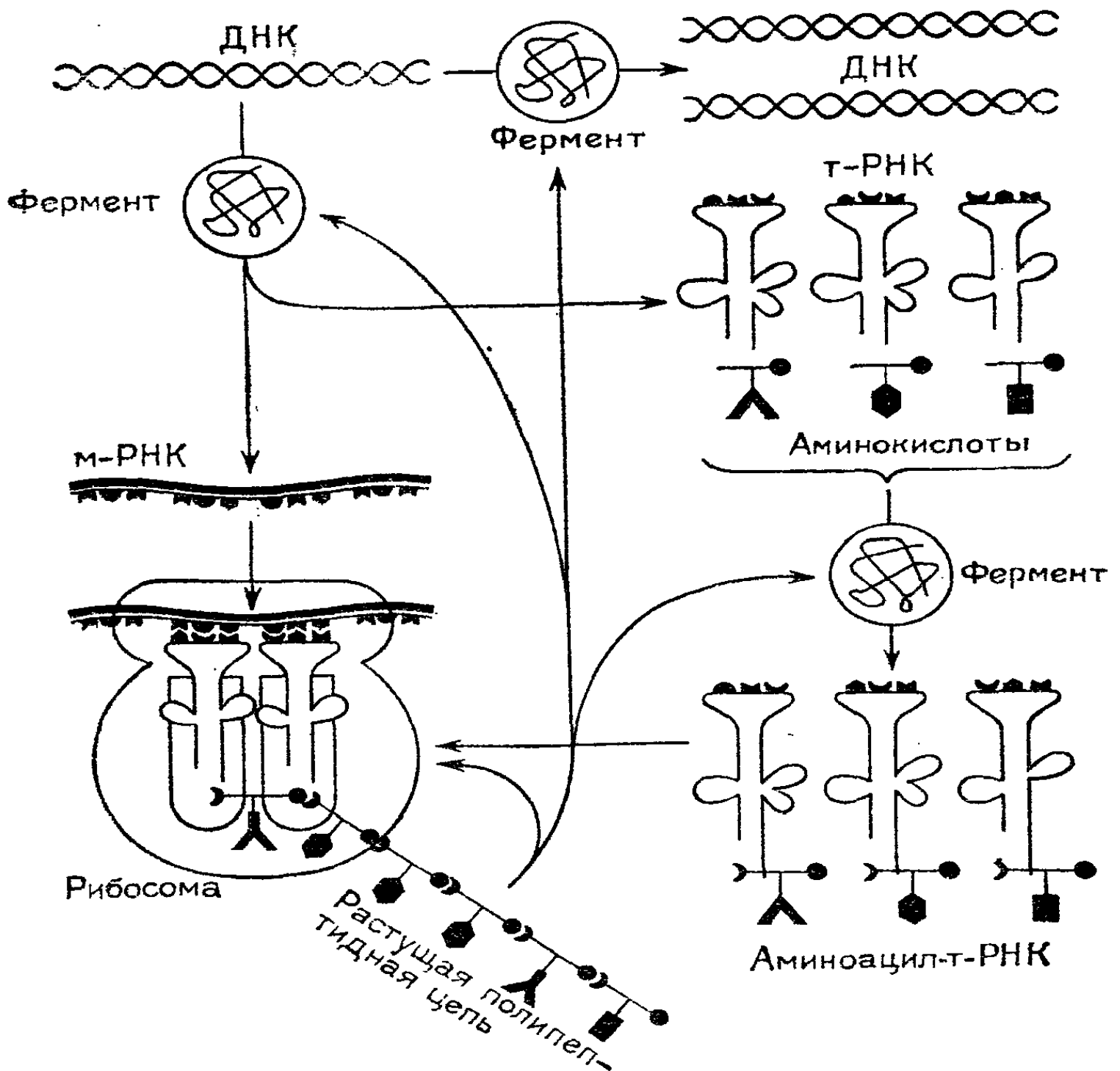


Рис. 7. Упрощенная схема цикла воспроизведения нуклеиновых кислот и биосинтеза белков. Ферменты, которые служат для репликации генетических сообщений (ДНК-полимеразы), для их транскрипции (ДНК-зависимые РНК-полимеразы) и трансляции (белки рибосом), а также для активации различных аминокислот (аминоацилсинтетазы), довольно хорошо изучены. Полная схема с обратными связями включает и другие белковые факторы регуляции и контроля.

этим сравнением с «астрономическими масштабами» скрывается некая цель. Мы хотим показать, что при таком расхождении в порядках величины (10^{130} , 10^{78} ,

10^{31}) не имеет смысла уходить во Вселенную, чтобы реализовать концепцию, для которой «нет места» на Земле.

При таком соотношении чисел в принципе мало что изменится, если мы введем в рассмотрение временную координату. Возраст Вселенной или, точнее, возраст такого состояния Вселенной, в котором могут разыгрываться реакции *органической* химии, составляет «только» 10^{17} секунд. Даже с помощью хорошо адаптированного ферментативного аппарата для разрыва или для образования одной пептидной связи потребуется время порядка 10^{-3} — 10^{-2} с. Если для решения задачи использовать все время существования Вселенной, то и в этом случае можно было бы испробовать лишь ничтожно малую долю из 10^{130} возможных последовательностей.

В этом числовом примере для наглядности мы взяли только одну, да и то маленькую молекулу белка. Молекула ДНК, в которой записана вся генетическая информация клетки кишечной бактерии, состоит из $4 \cdot 10^6$ элементов. Последовательность такого числа букв соответствует книге объемом в 1500 мелко набранных страниц. Число альтернативных последовательностей составляет здесь примерно $10^{1\ 000\ 000}$.

Конечно, именно в связи с этим возникает вопрос об уникальности подобных последовательностей и соответствующих структур. Что такое «уникальный»? Если это свойство означает «единственный в своем роде» не только в смысле *неповторимости* выбора, но и в смысле функционирования, то его следует отождествить со свойством быть *оптимальным*, т. е. *наилучшим*.

Эффективность молекулы белка можно определить. Можно непосредственно измерить скорости всех отдельных стадий ферментативного процесса — например, скорость связывания субстрата, скорость структурной аккомодации фермента, скорость превращения субстрата, скорость освобождения продуктов реакции — и, таким образом, установить, согласованы ли эти скорости друг с другом оптимальным образом. Времена, характеризующие продолжительность элементарных стадий реакции, обычно меньше 10^{-3} с, а часто даже меньше 10^{-6} с. Узнавание комплементарных единиц нуклеиновых кислот происходит за доли микросекунды. Встраивание такого элемента в макромолекулу требует

около 10^{-4} с. За такими реакциями удастся непосредственно следить лишь с помощью современных методов измерения, таких, как метод магнитного резонанса и релаксационные методы. Скорость химической реакции почти всегда достигает предельного значения, которое определяется законами физики. Во всех случаях обнаруживают оптимальное соотношение тенденции «как можно быстрее» и «настолько точно, как это нужно». Для создания такого оптимального функционирования не потребовалось участия никакого «демона Максвелла» — однако для нас, вторгающихся в микрокосмос со своим опытом и предубеждениями, сформированными окружающей средой, все выглядит так, как будто все это было «спланировано» каким-то гениальным конструктором. Впечатляющий пример дает упоминавшееся уже выше размножение кишечной палочки: весь план строения, т. е. «книга с 1500 страницами мелкого текста», последовательно прочитывается в течение двадцати минут, и одновременно выполняются все содержащиеся в ней предписания для синтезов. Кишечная палочка может репродуцироваться за двадцать минут. При этом информация, содержащаяся в каждой букве «плана строения», полностью перерабатывается за время, меньшее 10^{-3} секунды.

Подведем итог: макромолекулярные структуры, которые мы находим в ныне живущих организмах, являются уникальными — не столько в смысле неповторимости выбора из необозримого множества возможных альтернатив, сколько в смысле их оптимальной пригодности для выполнения соответствующей функции — «цели». Возникновение жизни, возникновение биологической информации — это проблема макромолекулярной самоорганизации.

КАК ВОЗНИКАЕТ ИНФОРМАЦИЯ?

Понятие информации имеет два аспекта, которые объединяются друг с другом только вопросом, помещенным в заглавии. Во-первых, это количественный аспект *абсолютной* информации, игнорирующий смысл и значение сообщения. Предполагается только, что сообщение отличается от других содержащимся в нем смыслом и что расположение символов в сообщении характерно для него. При передаче сообщения должно обеспечиваться сохранение расположения символов — в рамках допустимых отклонений. Этим аспектом занимается теория информации, или теория связи, развитая Шенноном [7], Винером [8] и Колмогоровым [9]. Она представляет собой математическую дисциплину — это прикладная теория вероятностей.

Второй аспект связан с оценкой сообщения, т. е. с его смыслом. Его называют также семантическим аспектом. Любая семантика зависит от известных соглашений, или, в более широком смысле, предполагает существование дополнительной информации (например, наличие определенной окружающей среды), которая локализуется вне сообщения и селективно взаимодействует с информацией, содержащейся в сообщении. Карл Фридрих фон Вайцзекер [10] выразил эту ситуацию посредством двух взаимно дополнительных тезисов:

1. Информацией является лишь то, что понимается.
2. Информацией является лишь то, что производит информацию.

Обратные утверждения неверны. Например, из второго тезиса нельзя сделать вывод, что информацию можно производить только из информации. Сохранение

около 10^{-4} с. За такими реакциями удается непосредственно следить лишь с помощью современных методов измерения, таких, как метод магнитного резонанса и релаксационные методы. Скорость химической реакции почти всегда достигает предельного значения, которое определяется законами физики. Во всех случаях обнаруживают оптимальное соотношение тенденции «как можно быстрее» и «настолько точно, как это нужно». Для создания такого оптимального функционирования не потребовалось участия никакого «демона Максвелла» — однако для нас, вторгающихся в микрокосмос со своим опытом и предубеждениями, сформированными окружающей средой, все выглядит так, как будто все это было «спланировано» каким-то гениальным конструктором. Впечатляющий пример дает упоминавшееся уже выше размножение кишечной палочки: весь план строения, т. е. «книга с 1500 страницами мелкого текста», последовательно прочитывается в течение двадцати минут, и одновременно выполняются все содержащиеся в ней предписания для синтезов. Кишечная палочка может репродуцироваться за двадцать минут. При этом информация, содержащаяся в каждой букве «плана строения», полностью перерабатывается за время, меньшее 10^{-3} секунды.

Подведем итог: макромолекулярные структуры, которые мы находим в ныне живущих организмах, являются уникальными — не столько в смысле неповторимости выбора из необозримого множества возможных альтернатив, сколько в смысле их оптимальной пригодности для выполнения соответствующей функции — «цели». Возникновение жизни, возникновение биологической информации — это проблема макромолекулярной самоорганизации.

КАК ВОЗНИКАЕТ ИНФОРМАЦИЯ?

Понятие информации имеет два аспекта, которые объединяются друг с другом только вопросом, помещенным в заглавии. Во-первых, это количественный аспект *абсолютной* информации, игнорирующий смысл и значение сообщения. Предполагается только, что сообщение отличается от других содержащимся в нем смыслом и что расположение символов в сообщении характерно для него. При передаче сообщения должно обеспечиваться сохранение расположения символов — в рамках допустимых отклонений. Этим аспектом занимается теория информации, или теория связи, развитая Шенноном [7], Винером [8] и Колмогоровым [9]. Она представляет собой математическую дисциплину — это прикладная теория вероятностей.

Второй аспект связан с оценкой сообщения, т. е. с его смыслом. Его называют также семантическим аспектом. Любая семантика зависит от известных соглашений, или, в более широком смысле, предполагает существование дополнительной информации (например, наличие определенной окружающей среды), которая локализуется вне сообщения и селективно взаимодействует с информацией, содержащейся в сообщении. Карл Фридрих фон Вайцзекер [10] выразил эту ситуацию посредством двух взаимно дополнительных тезисов:

1. Информацией является лишь то, что понимается.
2. Информацией является лишь то, что производит информацию.

Обратные утверждения неверны. Например, из второго тезиса нельзя сделать вывод, что информацию можно производить только из информации. Сохранение

количества или непрерывность информации исключали бы любую форму эволюции.

Если мы хотим понять, как возникает информация — и особенно биологическая информация, — то нам придется более подробно рассмотреть оба аспекта понятия информации. Ибо мы должны, собственно говоря, формулировать свой вопрос следующим образом: как возникает семантика? Нас интересует не столько расположение символов само по себе (ограниченное определенным распределением вероятностей), сколько функциональное значение, инструктированное этим расположением символов.

Пра-семантика биологической информации основана на функциональной упорядоченности, которая обеспечивает сохранение и дальнейшее развитие состояния жизни.

Абсолютная информация характеризуется множеством символов и их распределением. Символы должны быть представляемы сигналами, которые могут передаваться физическими способами. Для этого нужен однозначный код, т. е. однозначное соответствие между множеством символов и множеством элементарных знаков, или сигналов. На рис. 8 и 9 даны два примера: генетический код и телеграфный код. Они имеют совершенно различное происхождение. Первый (рис. 8) возник на раннем этапе существования Земли в результате самоорганизации вещества, второй (рис. 9) придуман только в нашем веке. Оба они, однако, построены по очень сходным логическим принципам. Они равномерны, т. е. все комбинации символов имеют одну и ту же длину: триплеты в генетическом коде и квинтеты в телеграфном коде. Код Морзе, наоборот, неравномерен. Добиться воспроизводимого функционирования такой схемы на молекулярном уровне было бы очень трудно из-за необходимости постоянно изменять «длину кадра».

В телеграфном коде все возможности используются взаимно однозначным образом. При двоичной системе квинтетный код дает $2^5 = 32$ комбинации. В немецком алфавите имеется 26 букв. Остаются шесть комбинаций, которые можно использовать для сигналов переключения. Генетический код состоит из $4^3 = 64$ триплетных комбинаций. Эта концепция была более подробно рассмотрена в первой главе. Использование в

2. Позиция

| | | 2. Позиция | | | | |
|------------|---|------------------|-----|------|------|---|
| | | у | ц | а | г | |
| 1. Позиция | у | Фен | Сер | Тир | Цис | у |
| | | Фен | Сер | Тир | Цис | ц |
| | | Лей | Сер | Терм | Терм | а |
| | | Лей | Сер | Терм | Трп | г |
| 1. Позиция | ц | Лей | Про | Гис | Арг | у |
| | | Лей | Про | Гис | Арг | ц |
| | | Лей | Про | Гли | Арг | а |
| | | Лей | Про | Гли | Арг | г |
| 1. Позиция | а | Иле | Тре | Аси | Сер | у |
| | | Иле | Тре | Аси | Сер | ц |
| | | Иле | Тре | Лиз | Арг | а |
| | | Мет ^а | Тре | Лиз | Арг | г |
| 1. Позиция | г | Вал | Ала | Асп | Гли | у |
| | | Вал | Ала | Асп | Гли | ц |
| | | Вал | Ала | Глу | Гли | а |
| | | Вал ^а | Ала | Глу | Гли | г |

3. Позиция

Рис. 8. Генетический код (значение штриховки см. на рис. 5). Терм — сигнал терминации, индекс «а» — сигнал инициации, если этот символ стоит в начале цистрона.

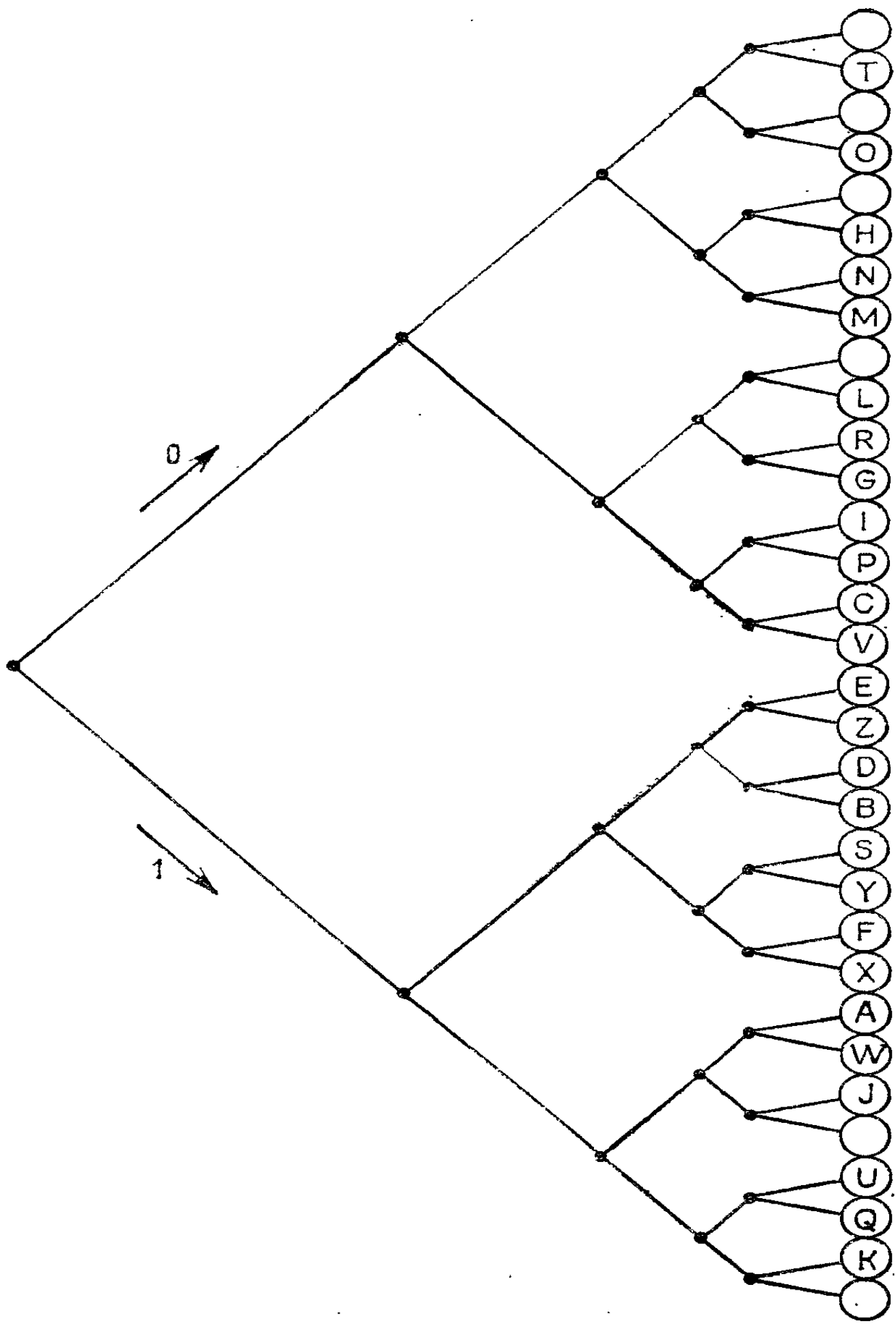


Рис. 9. Телеграфный код (см. Kindler Н.— Nova Acta Leopoldina, 1972, v. 37/1, p. 121).

качестве элементарных знаков кода не двух, а четырех нуклеотидов дает богатую палитру фенотипических возможностей выражения уже на уровне генных структур. Это проявляется прежде всего в высокой динамической гибкости двухцепочечных структур во время копирования. Использование только двух элементарных знаков в генетическом алфавите сделало бы необходимым — как и в случае телеграфного кода — введение квинтетов или даже секстетов. Передача информации, зависящая от взаимодействий между отдельными нуклеотидами, происходила бы слишком медленно в случае пяти или шести центров взаимодействия (комплементарных пар нуклеотидов на один кодон). Из кинетических исследований различных комбинаций нуклеотидов с определенным составом мы знаем [11], что триплетный код представляет собой оптимальное решение, найденное природой, не только с точки зрения его рациональности, но и с точки зрения динамической стабильности и гибкости.

Количество абстрактной информации в данном множестве символов можно выразить через распределение вероятностей всех возможных расположений. Обозначим через ν число элементарных знаков в данном множестве, а через λ — число различных классов знаков. Тогда число всех возможных расположений (последовательностей) будет λ^ν . Для белковой молекулы типа цитохрома (см. пример из первой главы) это число получилось порядка 10^{130} . Если бы все последовательности были равноценны, т. е. априорно равновероятны, то это число уже можно было бы использовать для задания количества информации. Действительно, в случае равномерного распределения вероятностей, чтобы встретить заданную последовательность, потребовалось бы перебрать и испытать все 10^{130} возможных последовательностей. При таком просмотре мы не располагаем, конечно, никакой информацией, которая могла бы вести нас к цели. «Красный свет» вспыхивает только тогда, когда появляется *совершенно* правильная последовательность. Таким «красным светом» может быть, например, тест на ферментативную активность, которая оказывается положительной только в том случае, когда реализована правильная последовательность.

Вероятности имеют свойство мультипликативности. Это значит, что вероятность одновременной реализации

нескольких не зависимых друг от друга событий, характеризующихся вероятностями p_1, p_2, \dots, p_k , равна произведению $p_1 \cdot p_2 \dots p_k$. С другой стороны, количество информации как мера множества должно иметь аддитивные свойства. Простейшая математическая функция, удовлетворяющая этим условиям, — это логарифм:

$$\log(a \cdot b) = \log a + \log b.$$

Целесообразно — особенно при кодировании двоичными элементарными знаками — пользоваться логарифмами при основании 2. В дальнейшем мы будем обозначать их символом ld . Для простого случая, когда все расположения имеют равные априорные вероятности, получим

$$H = \text{ld}(\lambda^v) = v \text{ld} \lambda.$$

Мера информации H , согласно Шеннону, называется энтропией*). В ней выражается эквивалентность реальной и потенциальной информации (т. е. информации, которая необходима для принятия правильного решения). Она является мерой неопределенности, или мерой множества возможных альтернативных решений. Для двоичных элементарных знаков в случае, когда различные расположения равновероятны, H просто равняется числу двоичных знаков (двоичных единиц, т. е. битов).

Состояние, для которого все возможные расположения имеют равные априорные вероятности, является состоянием с максимальной неопределенностью. Любое отклонение от такого равномерного распределения означает ограничение неопределенности. Если априорная вероятность i -го состояния равна p_i , то по Шеннону [7],

$$H = - \sum_{i=1}^n p_i \text{ld} p_i,$$

где $\sum_{i=1}^n p_i = 1$ и n — число всех возможных состояний

*) Связь информации или энтропии по Шеннону с термодинамической энтропией имеет не формальный, но физический характер. Эта связь рассмотрена в указанных в Предисловии монографиях Блюменфельда и автора этих строк. (Прим. ред.)

(например, λ^v). Таким образом, H — это среднее значение всех информационных вкладов $\text{ld } p_i$. Величины p_i являются одновременно весовыми множителями для этих отдельных вкладов. Если все они равны, то получается простое выражение $H = -\text{ld } p$; H обращается в нуль, если из n возможных расположений все, кроме одного, имеют вероятность нуль. При этом состояние с единичной вероятностью не будет иметь альтернатив, т. е. для его идентификации не требуется более никакой информации.

Примером такой ситуации может служить молекулярная структура, которая воспроизводимо образуется из стехиометрической смеси предшественников, специфично взаимодействующих друг с другом, причем тепловые флуктуации практически не могут помешать этому процессу. Для низкомолекулярных соединений этот случай реализуется очень часто. Другой пример — это пространственная укладка полипептидной цепи или цепи рибонуклеиновой кислоты, которая во многих случаях однозначно определяется последовательностью соответствующих элементов. Такой процесс укладки поэтому не включает в себя никакого изменения информации.

Рассмотрим теперь, чем могут быть обусловлены отклонения от равномерного распределения априорных вероятностей, — мы будем называть эти отклонения избыточностью. Имеются две принципиальные причины избыточности [12].

1. Символы различного сорта могут встречаться с различной частотой. Это означает, что в суммарном множестве символов v символы различного сорта представлены не в равных количествах v/λ .

2. Могут существовать предпочтительные последовательности символов. Специфичные взаимодействия между символами вызывают отклонения от случайного расположения. Эти взаимодействия не обязательно должны ограничиваться лишь ближайшими соседями. Так, например, в молекуле белка могут возникать очень сложные «условия соседства» по сравнению с линейной последовательностью (см. рис. 4—6).

Действие обоих этих факторов отчетливо выражено также в нашем языке. Здесь между ними имеется даже причинная связь: поскольку определенные комбинации букв встречаются чаще, чем другие, соответст-

Относительная частота букв в немецком языке

| Символ | p_i | $-p_i \lg p_i$ |
|----------|--------|----------------|
| 1 пробел | 0,1515 | 0,4125 |
| 2 E | 0,1470 | 0,4066 |
| 3 N | 0,0884 | 0,3093 |
| 4 R | 0,0686 | 0,2651 |
| 5 I | 0,0638 | 0,2532 |
| 6 S | 0,0539 | 0,2271 |
| 7 T | 0,0473 | 0,2082 |
| 8 D | 0,0439 | 0,1978 |
| 9 H | 0,0436 | 0,1969 |
| 10 A | 0,0433 | 0,1962 |
| 11 U | 0,0319 | 0,1585 |
| 12 L | 0,0293 | 0,1493 |
| 13 C | 0,0267 | 0,1397 |
| 14 G | 0,0267 | 0,1395 |
| 15 M | 0,0213 | 0,1184 |
| 16 O | 0,0177 | 0,1039 |
| 17 B | 0,0160 | 0,0959 |
| 18 Z | 0,0142 | 0,0873 |
| 19 W | 0,0142 | 0,0872 |
| 20 F | 0,0136 | 0,0843 |
| 21 K | 0,0096 | 0,0641 |
| 22 V | 0,0073 | 0,0521 |
| 23 U | 0,0058 | 0,0431 |
| 24 P | 0,0050 | 0,0382 |
| 25 A | 0,0049 | 0,0376 |
| 26 O | 0,0025 | 0,0219 |
| 27 J | 0,0016 | 0,0152 |
| 28 Y | 0,0002 | 0,0022 |
| 29 Q | 0,0001 | 0,0018 |
| 30 X | 0,0001 | 0,0017 |

$$\sum p_i = 1,0000$$

$$H = 4,1146 \text{ битов}$$

При равных частотах всех символов ($p_i = 1/30$) энтропия на один символ была бы $H = 4,9069$ битов.

вующие буквы используются с разными вероятностями. Дешифровка закодированных сообщений на языке с известной структурой основана на принципе такого сопоставления. Относительные частоты букв в немецком языке приведены в табл. 1.

Можно привести также много примеров часто встречающихся последовательностей определенных букв: употребительные окончания, как -en, -ig, -lich, или же помещение гласной между двумя согласными, и т. д. Множество всех избыточностей практически необозримо. Они распространяются на тончайшие особенности структуры языка и вовсе не ограничиваются ближним порядком в расположении букв. Значимость различных комбинаций букв также очень различна. Так, например, смысл искаженной телеграммы часто можно восстановить однозначно. С другой стороны, существуют ошибки, которые совершенно изменяют смысл, хотя они состоят в замене всего лишь одной буквы*).

В поисках связи между статистическим и семантическим аспектами информации мы прежде всего сталкиваемся с проблемой прироста информации (der Informationsgewinn). Здесь можно опираться непосредственно на теорию вероятностей.

Энтропия множества символов максимальна, когда все расположения символов равновероятны. Любое отклонение от равномерного распределения вероятностей означает ограничение неопределенности — количество информации, необходимое для полной идентификации ситуации или сообщения, уменьшается. Любое изменение распределения вероятностей в результате наблюдения или в результате того, что становятся известны какие-либо дополнительные условия, означает *прирост информации*.

Пусть

$$P = (p_1, p_2, \dots, p_n),$$

*) В одной книге, изданной в 1947 г., было написано: «Не считая вопросы приоритета существенными для истории науки, я все же должен отметить роль великого ученого...». Автора подвергли критике за недостаточное внимание к вопросам приоритета. Оправдываясь, автор сослался на опечатку. В рукописи, по его словам, значилось: «Но считая...» и т. д. (Прим. ред.)

где $p_i > 0$ и $\sum_{i=1}^n p_i = 1$, и $Q = (q_1, q_2, \dots, q_n)$, где $q_i \geq 0$

и $\sum_{i=1}^n q_i = 1$, — два распределения вероятностей для заданного (в обоих случаях идентичного) множества последовательностей символов, причем Q как условное распределение вероятностей получается из P в результате дополнительных наблюдений. Тогда прирост информации, по Реньи [13], будет равен

$$I(Q|P) = \sum_{i=1}^n q_i \text{ld} \left(\frac{q_i}{p_i} \right).$$

Если распределения вероятностей P и Q различны, то эта величина *всегда* положительна. Она обращается в нуль только тогда, когда наблюдение оставляет распределение P неизменным. Величина $I(Q|P)$ гораздо более определенно характеризует прирост информации, нежели разность энтропий $H(Q) - H(P)$, которая может быть как положительной, так и отрицательной. Последняя величина представляет собой разность двух средних значений, и поэтому ее информативность в отношении отдельного события очень ограничена. В величине $I(Q|P)$, напротив, результат наблюдения для каждого отдельного события сравнивается с вероятностью этого события до наблюдения, и лишь после этого делается усреднение по всем отдельным значениям. Таким образом, эта величина как среднее значение разностей более адекватно отражает истинный прирост информации, отнесенный к отдельному событию, чем разность средних значений. Учитывая важность этого понятия для эволюционного учения, дадим еще один пример, иллюстрирующий его отличие от более употребительного в теории информации понятия энтропии.

В заключительную стадию конкурса по архитектуре прошли три проекта. Имена авторов неизвестны жюри. Известно только, что среди них — архитектор А, который пользуется международным признанием и которому жюри охотнее всего доверило бы строительство. В одном из проектов по некоторым деталям вроде бы узнают «почерк» А. Конечно, полной уверенно-

сти в этом нет, так что распределение вероятностей того, что проект принадлежит A (результаты голосования), может быть примерно следующим:

$$p_1 = 0,8,$$

$$p_2 = 0,1,$$

$$p_3 = 0,1.$$

Эта неопределенность проявляется в конечной энтропии распределения:

$$H(P) = - \sum_{i=1}^3 p_i \text{ld } p_i \approx + 0,92 \text{ бита.}$$

Итак, проекту 1 присуждают первую премию. Но велико разочарование, когда выясняется, что автор проекта 1 вовсе не A .

После того, как автор проекта 1 становится известным, новое распределение вероятностей того, что определенный проект принадлежит A , будет таким:

$$q_1 = 0,$$

$$q_2 = 0,5;$$

$$q_3 = 0,5.$$

Теперь энтропия, выражающая субъективную неопределенность, даже возрастает:

$$H(Q) = - \sum_{i=1}^3 q_i \text{ld } q_i = 1 \text{ бит.}$$

Итак, энтропия не уменьшилась, а даже увеличилась. Раньше ошибочно думали, что знают больше. После обнаружения ошибки распределение вероятностей стало более равномерным, и, в сущности, именно это и нашло свое отражение в энтропии. С другой стороны, это конкретное событие обязательно дает прирост информации — и даже относительно большой прирост, потому что речь идет об исправлении ошибки:

$$I(Q|P) = \sum_{i=1}^3 q_i \text{ld } (q_i/p_i) \approx + 2,32 \text{ бита.}$$

Энтропия, как усредненная величина, в действительности очень мало отражает изменение информации, проистекающее из отдельных наблюдений, и это демонстрирует дальнейший ход конкурса.

Осталось два проекта, которым раньше уделялось не очень-то много внимания. Но теперь их изучают более детально. Результатом будет модифицированное распределение вероятностей (P') того, что определенный проект принадлежит A :

$$p'_2 = 0,8,$$

$$p'_3 = 0,2,$$

$$H(P') \approx 0,72 \text{ бита.}$$

Теперь остаются две возможности.

а) Предположение P' было верным, т. е. вскрытие конверта 2 показывает, что автором этого проекта является A . Тогда

$$q'_2 = 1,$$

$$q'_3 = 0.$$

б) Опять ошибка: A — автор не второго, а третьего проекта. В этом случае

$$q''_2 = 0,$$

$$q''_3 = 1.$$

Так как решение означает полную определенность, в обоих случаях энтропия будет равна нулю. Таким образом, эти два очень разных результата не отличаются по энтропии. Изменение энтропии в обоих случаях составляет

$$H = H(Q) - H(P) \approx -0,72 \text{ бита.}$$

Значение $I(Q'|P')$, напротив, дает возможность дифференцировать эти случаи. Случай а) означает лишь подтверждение гипотезы, т. е. сравнительно небольшой прирост информации:

$$\text{а) } I(Q'|P') \approx 0,32 \text{ бита.}$$

Случай б), наоборот, дает новые сведения. Соответст-

венно прирост информации получается много больше:

$$б) I(Q'' | P') \approx 2,32 \text{ бита.}$$

Распределение вероятностей совсем не исключало возможности ошибки (случай б). Нужно следить за тем, чтобы событию никогда не приписывалась нулевая вероятность, пока существует конечный интервал ошибок, т. е. пока ошибка еще возможна. Учет средних ошибок всегда приводит к конечным положительным значениям p_i . Это особенно важно для возникновения информации в природе, ибо все естественные элементарные процессы сопровождаются флуктуациями конечной величины.

В рассмотренном примере информация «возникает» только в мозгу наблюдателя. Отдельные наблюдения используются для того, чтобы последовательно сужать неопределенность, пока, наконец, не будет достигнуто исчерпывающее знание ситуации, которая сама по себе остается неизменной.

Не может ли то, что совершается вторично в мозгу «информируемого», происходить первично в системе, способной к самоорганизации и, следовательно, к последовательным сужениям (сначала произвольного) распределения вероятностей? Источником, производящим информацию, при этом может быть статистический флуктуационный процесс, т. е. «генератор шума». Такая система, несомненно, должна обладать известной способностью к фильтрации, или к отбору. Это означает, что появление «нестабильностей» делает невозможным возврат к первоначальному распределению вероятностей.

В своей монографии «Наука и теория информации» Леон Бриллюэн отметил, что рассмотрение семантического аспекта информации делает необходимым введение параметра ценности. Такой параметр должен иметь физическую интерпретацию. Оценивание равнозначно отбору, а способность к отбору как материальное свойство должна найти свое обоснование в динамических критериях. Ее нельзя объяснить, оставаясь в рамках одной лишь теории вероятностей. Возникновение и исчезновение, т. е. временность существования, является необходимой предпосылкой любой самоорганизации, основанной на отборе, а также предпосылкой возникновения информации. Но, кроме этого, нужны

еще особые закономерности, обусловленные природой системы.

Здесь перед нами встает фундаментальный вопрос: а может ли информация вообще возникать? Или же она лишь выявляется? Не сводится ли в конечном счете вся семантика к прасемаптике, и не определяется ли она в таком случае неотъемлемыми свойствами материи? На этот вопрос — коренной вопрос для настоящей работы — мы сможем дать ответ только в четвертой главе. Сначала нужно рассмотреть еще несколько предпосылок *).

*) Как указано в Предисловии, вопрос о смысле, содержании или ценности информации особенно существен для биологии и требует специального рассмотрения. Очень важен также вопрос о возникновении информации, поставленный авторами (см. далее). (Прим. ред.)

Глава 3

«ИГРЫ В БИСЕР»

Правило, случай и выбор относятся к существенным элементам игры. Многие игры целиком определяются диктатом правил, и потому их ход полностью закономерен. В других преобладает случай — мы говорим об азартных играх. Сочетание обоих элементов ведет к третьей категории игр, в которых на переднем плане стоит выбор, управляемый определенными критериями, т. е. оптимизирующая стратегия. К этой категории относятся наиболее интересные стратегические игры, такие, как шахматы и го. Очевидно, что здесь развитие игровой ситуации соответствует возникновению информации, обладающей семантикой, обусловленной правилами. По такому принципу «играет» и живая природа.

В дальнейшем мы будем постепенно развивать идею взаимодействия закона и случая, руководствуясь моделями — «играми в бисер». Мы будем называть их так не только потому, что шарики, необходимые для игры, могут быть и бисером, но прежде всего потому, что эти игры, основанные на простых правилах, на высшей стадии своего развития реализуют то представление о единстве Природы и Духа, которое изложено в «Игре в бисер» Германа Гессе:

«Эти правила, язык знаков и грамматика Игры суть не что иное, как высокоразвитая тайнопись, к которой причастны многие науки и искусства, особенно математика и музыка (соответственно музыковедение), и которая способна выразить и связать друг с другом содержание и результаты почти всех наук».

Для первой игры (рис. 10) нам понадобятся два ящика, по которым мы произвольно распределим оп-

еще особые закономерности, обусловленные природой системы.

Здесь перед нами встает фундаментальный вопрос: а может ли информация вообще возникать? Или же она лишь выявляется? Не сводится ли в конечном счете вся семантика к прасемаптике, и не определяется ли она в таком случае неотъемлемыми свойствами материи? На этот вопрос — корепной вопрос для настоящей работы — мы сможем дать ответ только в четвертой главе. Сначала нужно рассмотреть еще несколько предпосылок *).

*) Как указано в Предисловии, вопрос о смысле, содержании или ценности информации особенно существен для биологии и требует специального рассмотрения. Очень важен также вопрос о возникновении информации, поставленный авторами (см. далее). (Прим. ред.)

Глава 3

«ИГРЫ В БИСЕР»

Правило, случай и выбор относятся к существенным элементам игры. Многие игры целиком определяются диктатом правил, и потому их ход полностью закономерен. В других преобладает случай — мы говорим об азартных играх. Сочетание обоих элементов ведет к третьей категории игр, в которых на переднем плане стоит выбор, управляемый определенными критериями, т. е. оптимизирующая стратегия. К этой категории относятся наиболее интересные стратегические игры, такие, как шахматы и го. Очевидно, что здесь развитие игровой ситуации соответствует возникновению информации, обладающей семантикой, обусловленной правилами. По такому принципу «играет» и живая природа.

В дальнейшем мы будем постепенно развивать идею взаимодействия закона и случая, руководствуясь моделями — «играми в бисер». Мы будем называть их так не только потому, что шарики, необходимые для игры, могут быть и бисером, но прежде всего потому, что эти игры, основанные на простых правилах, на высшей стадии своего развития реализуют то представление о единстве Природы и Духа, которое изложено в «Игре в бисер» Германа Гессе:

«Эти правила, язык знаков и грамматика Игры суть не что иное, как высокоразвитая тайнопись, к которой причастны многие науки и искусства, особенно математика и музыка (соответственно музыковедение), и которая способна выразить и связать друг с другом содержание и результаты почти всех наук».

Для первой игры (рис. 10) нам понадобятся два ящика, по которым мы произвольно распределим оп-

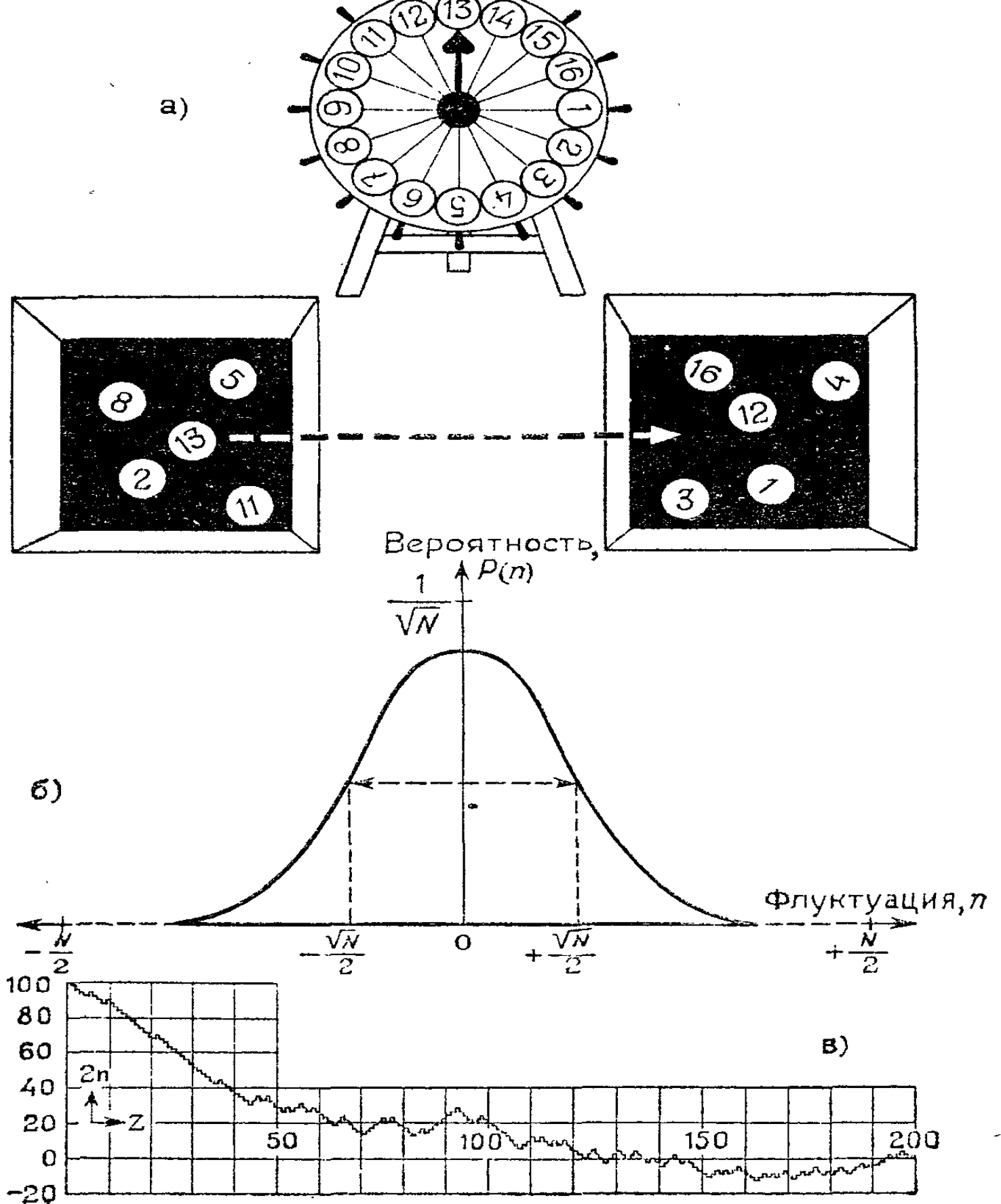


Рис. 10. Схема «урновой» игры. а) «Моментальный снимок» отдельного хода игры: шар 13 переносится из левого ящика в правый. б) Распределение вероятностей $P(n)$ того, что флуктуация будет иметь величину n . Это гауссово распределение, интеграл по которому (т. е. площадь под кривой) нормирован на единицу. Величина n задает отклонение от равномерного распределения ($N/2$ шаров в каждом ящике). в) Запись ходов одной игры по Кольраушу и Шредингеру (Physik. Zeitschr., 1926, Bd. 27, S. 307). По оси абсцисс отложено число ходов Z , по оси ординат — разность чисел шаров в обоих ящиках $2n$, $N=100$,

ределенное число (N) шаров. «Произвольно» означает, что начальное распределение шаров совершенно безразлично, например, можно положить их все в один ящик, а другой оставить пустым, или распределить их поровну по обоим ящикам. Важно, чтобы все шары были пронумерованы и чтобы каждый номер (от 1 до N) встречался только один раз. Кроме того, нам понадобится лотерейная машина — колесо счастья или барабан, с помощью которой можно было бы «вытянуть» с равной априорной вероятностью любое число от 1 до N . Теперь можно начинать игру. Правила игры очень просты: каждый раз, когда выбор падает на определенное число, шар с соответствующим номером следует вынуть из того ящика, где он находился, и переложить его в другой ящик.

Допустим, что продолжительность игры не очень мала — это выполняется для большинства лотерей. Игрока больше всего интересует выигрыш. Мы занялись этой игрой тоже лишь из-за ее результата, который полностью воспроизводим, — уже после сравнительно небольшого числа ходов в каждом ящике будет находиться в среднем $N/2$ шаров. Впрочем, здесь важна оговорка — *в среднем*, ибо если считать точно, то в одном из ящиков мы будем находить то большее, то меньшее число шаров, т. е. $N/2 \pm n$, а в другом ящике, соответственно, $N/2 \mp n$ шаров. Величина флуктуации n также имеет воспроизводимое среднее значение — она пропорциональна \sqrt{N} . Это означает, что *относительная* флуктуация убывает с ростом числа шаров N , хотя *абсолютная* флуктуация при этом растет. Чем больше шаров взято для игры, тем точнее будет выполняться равномерное распределение. При очень больших значениях N флуктуации будут едва заметны, и в пределе мы получим детерминистический закон распределения. Распределение вероятностей для n можно задать в явном виде. При достаточно больших N это распределение близко к гауссову, т. е. к колоколообразной кривой с полушириной порядка \sqrt{N} и относительной высотой порядка $1/\sqrt{N}$ (см. рис. 10).

Каждое состояние, на языке физиков — микросостояние, которое характеризуется перечислением всех номеров в каждом из ящиков, имеет вероятность 2^{-N} .

Впрочем, нас не так уж интересуют эти микросостояния. Мы гораздо больше хотим узнать что-нибудь о макросостояниях, которые характеризуются числами n . Существуют всего два макросостояния, которым соответствует только по одному микросостоянию, т. е. которые можно реализовать лишь одним распределением шаров по ящикам. Эти состояния характеризуются величиной

$$n = \pm N/2,$$

т. е. все шары находятся в одном ящике, второй ящик пуст. Естественно, что такие флуктуации реализуются очень редко, особенно при больших N . Напротив, состояние равномерного распределения $n = 0$ реализуется чаще всего, потому что его можно получить наибольшим числом способов размещения шаров. Энтропия достигает здесь максимума. Однако при больших N число способов реализации даже этого состояния очень мало по сравнению с числом всех состояний 2^N . Только примерно $(1/\sqrt{N})$ -я доля всех микросостояний соответствует состоянию с $n = 0$. Поэтому игра почти всегда будет заканчиваться с ненулевой, хотя и относительно малой, флуктуацией n , заключенной обычно в интервале между $+\sqrt{N}$ и $-\sqrt{N}$. Мы можем назвать *выигрышем* положительную флуктуацию и *проигрышем* — отрицательную. Такая игра обещает быть интересной только при малых N благодаря относительно большим флуктуациям. Для $N = 6$ можно пользоваться обычной игральной костью. В этом случае экстремальные флуктуации

$$n = \pm N/2$$

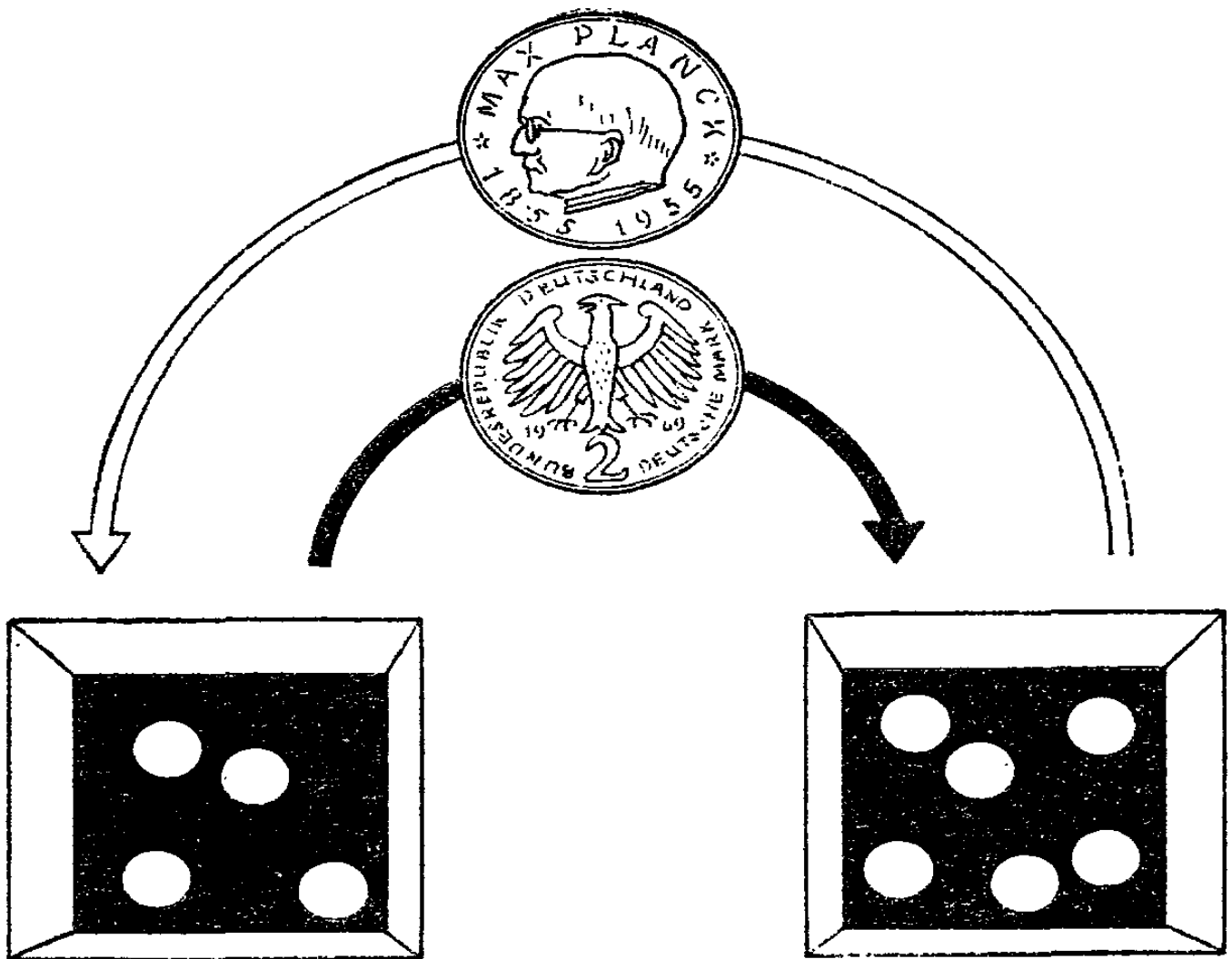
имеют вероятность $1/64$, таким образом, в среднем через 32 хода мы придем к одной из экстремальных ситуаций, которую можно считать концом игры.

В физике эта игра известна как «модель Эренфестов». Ее придумали голландские физики — супруги Эренфесты примерно на рубеже нашего века. Она сыграла существенную роль при выяснении природы флуктуаций в молекулярных распределениях. Не кто иные, как Кольрауш и Шредингер, имели даже терпение действительно «играть» в эту игру и опубликовать полученные результаты.

На рис. 10 приведена часть подобного ряда наблюдений.

Конечно, в случае физических процессов, доступных нашему наблюдению, мы имеем дело с очень большими числами, которые по порядку величины сравнимы с числом Авогадро ($\sim 10^{24}$). Для таких больших чисел флуктуации еле заметны. Законы равновесной термодинамики (например химический закон действия масс) можно поэтому с хорошим приближением считать детерминистическими. Можно считать, что два ящика в игре Эренфестов представляют собой различные состояния, между которыми идут взаимопревращения, например химическая реакция. Эта игра непосредственно обобщается на большое число различных состояний (или «ящиков») с различными вероятностями их заселенности. Такая игра совершенно аналогичным образом описывает установление равновесного распределения и флуктуационное поведение. Результаты справедливы для любых реакций, идущих вблизи равновесия, независимо от того, насколько сложен их механизм. Нумерация шаров существенна только для проведения игры. Она гарантирует эквивалентность всех шаров в статистическом процессе превращения. В остальном индивидуальность шаров нас не интересует — молекулы, находящиеся в одном и том же состоянии, неразличимы. Различимыми являются отдельные состояния — химические формы или индивидуальные макромолекулярные последовательности, а в случае нашей игры — ящики.

Чему может научить нас эта игра в связи с проблемой возникновения информации? Для любой материальной системы вблизи равновесия флуктуации имеют саморегулируемый характер. Этот самоконтроль является неотъемлемым свойством флуктуационного процесса, и поэтому равновесие является устойчивым состоянием. Распределение населенностей удерживается в рамках определенных флуктуационных границ энергетическим параметром, присущим системе. Эта игра ясно показывает, что возвращение системы в состояние равновесия тем более вероятно, чем более она отклонилась от равновесия. Экстремальные флуктуации, например вымирание сильно населенного в среднем состояния, едва ли будут реализоваться в пределах разумных пространственных и временных маш-



а)

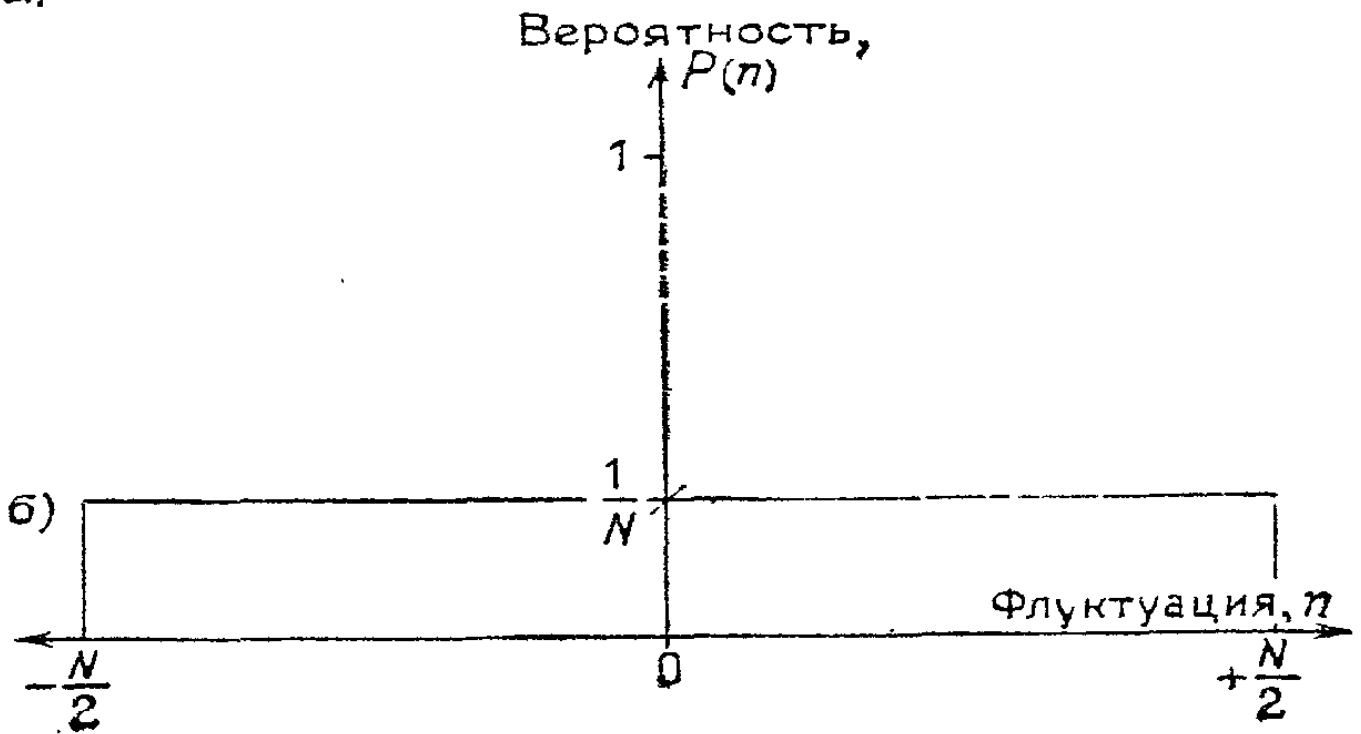


Рис. 11. «Орел» или «решка». Игра иллюстрирует неконтролируемый флуктуационный процесс. а) При выпадении «орла» шар переходит из левого ящика в правый, при выпадении «решки» — наоборот, из правого в левый. б) Распределение вероятностей теперь совершенно равномерно, т. е. любое отклонение n от равномерного распределения ($N/2$ в каждом ящике) одинаково вероятно.

табов. Даже в том случае, когда имеется большое число альтернативных состояний и вероятность того, что какое-то определенное состояние населено, очень мала, система будет беспорядочно пробегать через эти состояния снова и снова, так что для такой эргодической системы при усреднении во времени будут выполняться детерминистические законы физики. Для того, что мы называем возникновением информации, подобное поведение «противопоказано». Информация здесь не изменяется вследствие неизменности распределения вероятностей. Она зависит от характера внутренних взаимодействий, которые определяют усредненное поведение системы.

Вторая игра (рис. 11), которая будет служить контрпримером, иллюстрирует полностью недетерминированный флуктуационный процесс. Здесь можно обойтись сравнительно кратким разбором, потому что мы снова придем к выводу, что на этом пути, как и в первой игре, невозможно возникновение информации.

Нам снова понадобятся два ящика (в общем случае k ящиков) и некоторое количество шаров, которые следует распределить по разным ящикам. На этот раз мы будем исходить из равномерного распределения. Итак, в простейшем случае мы имеем два ящика, каждый из которых содержит сначала $N/2$ шаров. Для этой игры шары не нужно нумеровать и можно пользоваться лотерейной машиной, которая способна лишь выдавать решения типа «да — нет»; для монеты это будет «орел или решка», а для игральной кости — четные или нечетные числа. При выпадении «решки» будем переносить один шар из ящика 1 в ящик 2, при выпадении «орла» — наоборот. Каким здесь окажется результат?

Чтобы его найти, нужно долго играть. Вместо этого мы попытаемся его предсказать. Перед нами снова азартная игра, в которой отдельный результат непредсказуем. Но здесь утрачена существенная для первой игры способность к регулированию — зависимость вероятности переноса от населенности ящика. Результат одного бросания, «орел» или «решка», и, следовательно, вероятность определенного переноса

$$(1 \longrightarrow 2, \text{ или } 2 \longrightarrow 1)$$

полностью независимы от населенности ящиков, т. е. от предыстории флуктуаций. До тех пор, пока в ящике останется хоть один шар, эта вероятность неизменно равняется 50% для любого бросания.

Существует по теперешним временам не такая уж веселая история о человеке, который всегда брал с собой бомбу, когда ему приходилось летать. Он верил, что при этом увеличивается его безопасность: ведь две бомбы в одном самолете — событие куда менее вероятное, чем одна, его бомба. Этот человек перепутал, очевидно, правила обеих игр. Сотрудники органов безопасности вряд ли удовлетворились бы подобным объяснением.

Итак, когда при не зависимых друг от друга флуктуациях нет саморегулирования, *никакое* состояние не будет предпочтительнее любого другого. Это означает, что распределение вероятностей представляется прямоугольником (см. рис. 11): при усреднении по времени все состояния оказываются равноценными или равновероятными. Система беспорядочно «дрейфует» по всем возможным состояниям. В отличие от модели Эренфестов, экстремальное состояние здесь легко достижимо — в среднем примерно уже через N^2 бросаний. Если перейти к k ящикам (состояниям), то результат в принципе не меняется. Поскольку теперь имеется $k(k-1)$ возможностей переноса, то для проведения игры требуется несколько более хитроумная лотерейная машина. Как и раньше, система будет пробегать через все состояния в процессе случайного блуждания.

Вывод: если никакое состояние ничем не выделяется, то в процессе ненаправленного дрейфа не может породиться *информация*, понятие информации теряет здесь всякий смысл.

В *третьей игре* объединим Закон и Случай. Введем снова пронумерованные шары. На этот раз различия в номерах будут соответствовать истинному различию состояний. Для игры нам понадобится один ящик. Для хранения запаса шаров необходим еще один ящик, в котором каждый из N номеров должен быть представлен N шарами.

В игровой ящик кладут N шаров, причем каждый номер может быть представлен сначала только одним шаром. Правила игры снова очень просты: не загля-

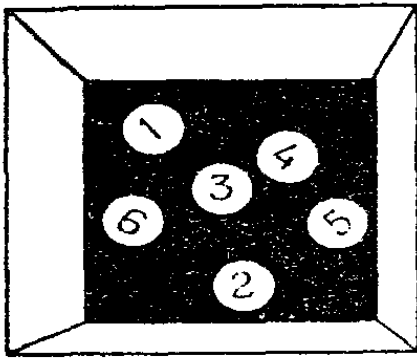
дывая в ящик, из него вынимают произвольно выбранный шар. Шары всегда должны быть хорошо перемешаны, чтобы шансы быть вынутым для каждого из них были одинаковы. Судьба вынутого шара может быть двойкой: он либо «удваивается», т. е. его кладут обратно в игровой ящик вместе со вторым шаром, имеющим тот же номер, либо его просто удаляют из игрового ящика и кладут обратно в запасной ящик. Если оба процесса чередовать строго попеременно, то общее число шаров в игровом ящике будет оставаться всегда постоянным ($=N$).

Чередование событий можно, конечно, предоставить случаю. Тогда число шаров в игровом ящике будет флуктуировать: оно может увеличиваться или, с равной вероятностью, уменьшаться. При этом рано или поздно оно обязательно упадет до нуля, в среднем это происходит уже через N^2 ходов. Этот случай иллюстрирует необратимую флуктуационную катастрофу, сравнимую с финансовым крахом игрока в рулетку.

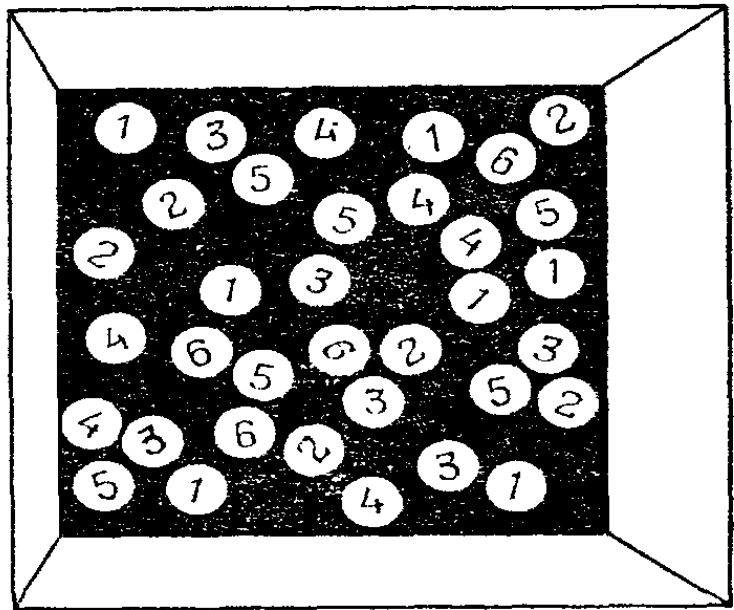
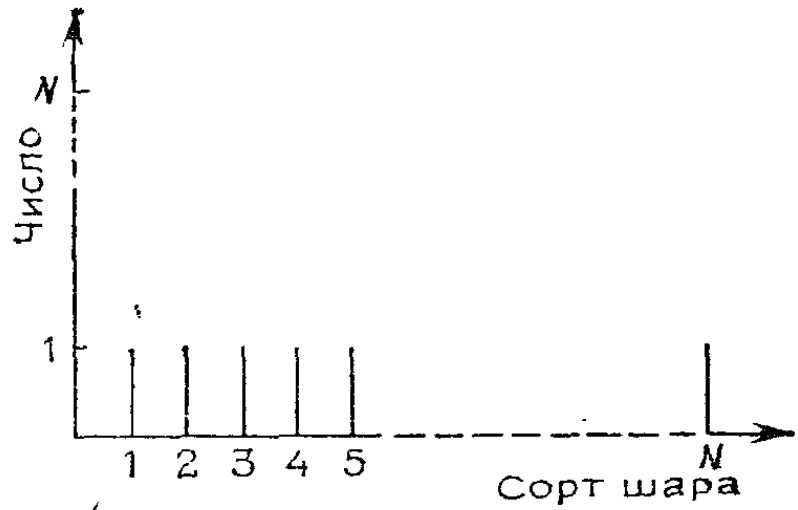
Число шаров с определенным номером будет флуктуировать, как и общее число шаров. Некоторые номера будут быстро вымирать, зато другие будут представлены несколькими шарами. То же самое будет происходить, конечно, и при строгом чередовании удвоения и удаления шаров, только в этом случае общее число шаров N в игровом ящике остается постоянным. Разберем подробнее этот более прозрачный случай. Номера, которые были представлены в игровом ящике только одним шаром и были удалены прежде, чем им удалось удвоиться, не имеют шансов вернуться в игру. Напротив, те номера, которые удвоились или даже размножились и приобрели в результате этого повышенную устойчивость — они не так легко могут стать жертвой флуктуационной катастрофы. Эта игра всегда приводит к неизбежному результату — «выживание» одного-единственного номера, который зато будет представлен N шарами.

Перед нами, очевидно, случай «отбора», хотя никто здесь специально не занимался отбором, следуя какому-либо критерию. Таким образом, отбор здесь является лишь следствием определенного флуктуационного поведения. Для рассмотренного варианта игры никак нельзя предсказать, какой именно номер будет ото-

Начальная ситуация

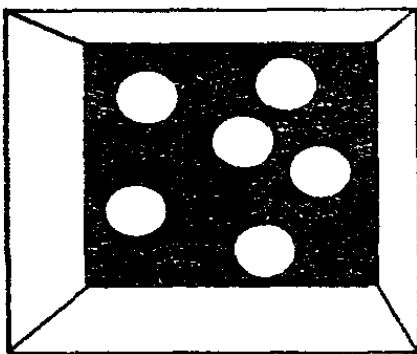


Игровой ящик



Ящик с запасом шаров

Конечная ситуация



Игровой ящик

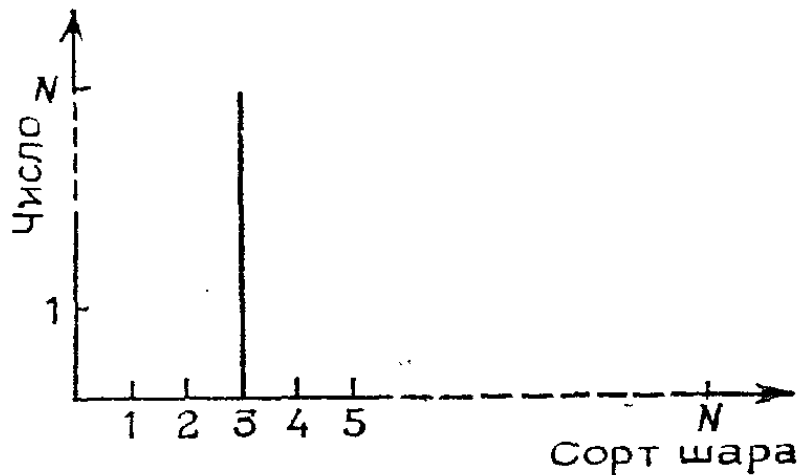


Рис. 12. Игра в селекцию. Если для каждого сорта шаров задать функцию $P(n)$, т. е. распределение вероятностей того, что соответствующий сорт будет представлен n шарами, то в конце игры для всех сортов, кроме одного, эти функции при всех значениях n будут равны нулю. Однако всегда будет существовать сорт, для которого $P(n)$ будет принимать ненулевые значения. Если нет ошибок, то $P(n)=0$ для всех $n < N$, но $P(n)=1$ для $n=N$ (δ -функция). При конечной вероятности ошибки эта кривая расширяется, но в общем случае она будет оставаться гораздо уже, чем распределение Гаусса, показанное на рис. 10.

бран. Все номера имеют в точности равные шансы. В начале игры распределение вероятностей строго равномерно. Для каждого номера вероятность выжить равняется $1/N$. Это распределение все время изменяется вследствие флуктуаций. Каждый раз, когда вымирает какой-либо номер, мы регистрируем прирост информации. Это продолжается до тех пор, пока, наконец, вероятности отбора для всех номеров, кроме одного, не обратятся в нуль. Этот единственный номер i отбирается, для него вероятность выживания p_i становится равной единице, т. е. превращается в достоверность. Вместо размытого распределения Гаусса в первой игре или совершенно выровненного прямоугольного распределения во второй игре здесь мы имеем дело с однозначным выбором (рис. 10—12). Рис. 12 иллюстрирует игру в селекцию. Два графика здесь характеризуют распределения шаров в начале и в конце игры. Распределение вероятностей было бы здесь малоинформативно, так как а priori любой сорт шаров может выжить с равной вероятностью. Но в действительности в конце игры всегда происходит выбор лишь одного сорта или, при конечной вероятности ошибки, лишь немногих сортов.

Дарвин характеризовал отбор как «выживание наиболее приспособленных». Часто задают вопросы, как охарактеризовать это состояние «наибольшей приспособленности» для живого существа и по каким критериям должна происходить такая «оценка». Если бы максимальная приспособленность определялась лишь при помощи самого факта выживания, то дарвиновский принцип выражал бы лишь тавтологию «выживание выживающих».

Только что рассмотренный вариант третьей игры демонстрирует именно такую ситуацию. Кроме самого факта выживания не имеется других критериев. «Выживание» здесь является абсолютно непредсказуемым результатом флуктуационного процесса. Дарвин, наверняка, не имел в виду этот случай, когда он формулировал свой принцип. Тем не менее, этот пример делает очевидным, что самоорганизация, основанная на отборе, является по существу следствием определенного физического поведения системы, когда флуктуации способны усиливаться. При этом отдельные флуктуации непредсказуемы. Они сводятся к элемен-

тарным событиям, т. е. к квантовым явлениям, подчиняющимся соотношениям неопределенностей Гейзенберга. В игре Эренфестов элементарный процесс тоже не детерминирован. Однако там непредсказуемость отдельных событий настолько ограничивается законом больших чисел, что для макроскопических систем вероятностные законы поведения превращаются в детерминистические. Последний же вариант игры допускает усиление флуктуаций, и поэтому микроскопическая неопределенность может здесь отображаться в макроскопических явлениях. Паскуаль Иордан указывал на возможность макроскопического отображения микроскопических событий еще на заре квантовой механики.

Только что описанная игра является лишь мысленным экспериментом, который в таком виде по *самой своей сути* совершенно непригоден для описания реальных физических систем. Процесс удвоения требует инструктирования на молекулярном уровне, который может реализоваться лишь посредством определенных селективных взаимодействий. Такие взаимодействия имеют конечную величину и всегда подвержены тепловым возмущениям, вследствие чего некоторая доля инструкций оказывается ошибочной. Для молекулярного процесса невозможен абсолютно правильный выбор между альтернативами. Таким образом, наша игра нереалистична. Чтобы верно отобразить действительность, какую-то долю удвоений мы должны заменить введением новых номеров. Это легко достигается изменением правил игры. Сохраним строгое чередование «рождения» и «смерти», но пусть теперь некоторая наперед выбранная часть всех редупликаций — например, в зависимости от вероятности ошибки, каждая десятая, сотая или миллионная часть, не будет происходить. Вместо этого система пополняется шаром с «новым» номером $> N$. Теперь, конечно, не может быть никакого устойчивого отбора, во всяком случае, пока сумма всех шаров может увеличиваться и уменьшаться с равной вероятностью. При постоянной сумме N из ящика в среднем удаляется столько же шаров, сколько в него попадает, однако одновременно из-за конечной вероятности ошибки в ящик постоянно приходят *новые* шары. Поэтому для каждого сорта (номера) шаров вероят-

ность удвоения должна быть несколько меньше, чем вероятность удаления из ящика. Это означает, что *каждый* номер должен через конечное время вымереть. Номера шаров, находящихся в ящике, блуждают по всему набору чисел. Никакого устойчивого отбора больше нет.

При конечной вероятности ошибки способность к отбору можно восстановить, если ввести распределение вероятностей *индивидуального* удвоения и удаления. Эти распределения вероятностей должны быть таковы, чтобы в среднем общее число шаров в ящике оставалось постоянным. В природе такая стационарность в общем случае устанавливается автоматически из-за ограниченности притока высокоэнергетических веществ (пищи). В эксперименте или в игре стационарность всегда можно поддерживать регулирующим воздействием (реакционную систему можно, например, разбавлять или концентрировать). Удельные скорости продукции каждого отдельного «вида» определяются разностью вероятностей удвоения и удаления, а также притоком и оттоком вследствие ошибок. Эти скорости можно выразить через функцию ценности W_i , которая содержит энергетические и временные параметры, характеризующие точность узнавания и трансляции, скорость копирования и время жизни. В зависимости от механизма реакций и степени их сопряжения функция ценности может иметь довольно сложный вид. Тем не менее ее физическое обоснование совершенно определено: единичным процессам соответствуют стохастические законы, макроскопическим приростам — детерминистическая теория. Из величины W_i вычитается средняя продуктивность $\bar{E}(t)$, зависящая от времени. Разность обеих величин определяет судьбу отдельного «вида»:

$$W_i - \bar{E}(t) > 0 \quad \text{означает выживание,}$$

$$W_i - \bar{E}(t) < 0 \quad \text{означает вымирание.}$$

Величина $\bar{E}(t)$ представляет собой переменное во времени пороговое значение. Изменяя его, система может постоянно приспосабливаться к новым ситуациям.

Стабильный отбор означает, что

$$\bar{E}(t) \longrightarrow W_{opt.}$$

Процесс копирования может ограничиваться единичными видами, доминирующими в популяции и окруженными «кометным хвостом» мутантов; эти виды будут поочередно сменять друг друга в процессе прогрессирующей оптимизации системы. Однако этот процесс может включать в себя и целый ансамбль, в котором отдельные виды будут сопряжены друг с другом посредством самых разнообразных связей. Каждый раз, когда среди мутантов возникает вид с большей селективной ценностью и оказывается способным выжить несмотря на флуктуации, которые на первой фазе угрожают его существованию, прежняя (мета-)стабильная популяция гибнет, и устанавливается новое стационарное распределение, в котором доминирует мутант, имеющий селективное преимущество. Несмотря на наличие «ошибочных копий», новое распределение все еще является очень острым. В каждый момент времени в системе представлена лишь *относительно* малая доля из общего числа всех возможных альтернатив (мутантов). Распределение числа мутаций аппроксимируется распределением Пуассона.

Итак, эволюция состоит из ряда неустойчивостей, при развитии которых распределения вероятностей необратимо изменяются, причем это сопровождается приростом информации *). Этот процесс оптимизации не обязательно соответствует простому монотонному росту функции ценности, и в то же время он не является ненаправленным флуктуационным процессом. Наша игра с шарами еще достаточно проста для того, чтобы в предельном случае больших чисел она имела однозначно определенный результат, соответствующий абсолютному максимуму W . «Наиболее приспособленный» здесь однозначно определяется посредством критерия ценности.

Это возвращает нас к фундаментальному вопросу, поставленному в конце предыдущей главы: может ли информация действительно возникать *de novo*, или же она только выявляется при помощи детерминистического в своей основе физического процесса.

Другими словами: фильтрующий процесс отбора выбирает определенные альтернативы из микроскопи-

*) Прирост информации не обязательно означает возникновение новой информации — речь может идти и о выявлении информации, уже имеющейся, но скрытой «шумами». (Прим. ред.)

ческих недетерминированных элементарных событий, усиливает их и выявляет в макроскопическом процессе эволюции. Прокладывает ли он дорогу к предопределенной цели? Или же он имеет такой характер, что только дорога определяет цель, и, следовательно, никакие исторические пути развития не являются априорно необходимыми? Только в последнем случае прирост информации, связанный с изменением распределения вероятностей, можно приписать истинному возникновению, а не выявлению семантической информации.

Глава 4

ИГРАЕТ ЛИ БОГ В КОСТИ?

Идея творения допускает две совершенно противоположные интерпретации:

1. «Начало» однозначно определяется взаимодействиями, присущими материи. Начальные условия определяют путь и цель эволюции.

2. «Начало» является случайным событием. При этом имеется в виду, что вероятность этого события практически равна нулю, но, с другой стороны, все-таки нельзя с уверенностью утверждать, что оно не может произойти. Чтобы такая флуктуация, реализовавшись однажды, могла проявиться в макроскопических масштабах, она должна обладать известной устойчивостью или даже способностью распространяться.

Первая интерпретация означала бы, что высоко развитые формы жизни, сходные с земными, должны встречаться во многих местах Вселенной. Кроме того, это означает, что если когда-либо в будущем удастся воссоздать в лаборатории начальные условия возникновения жизни на нашей планете — перспективы для этого более благоприятны, чем для полета к другим планетным системам, — то при этом с фатальной неизбежностью должна возникнуть жизнь с таким же кодом и с такими же механизмами, как и существующая ныне. В отношении этой гипотезы — если она правильна — мы можем лишь сказать, что физические законы, в соответствии с которыми мог бы произойти такой воспроизводимый процесс самоорганизации, следует еще найти. Многие физики склоняются именно к такому представлению. Они ответили бы отрицательно также на вопрос, стоящий в названии этой

главы, по крайней мере в отношении возникновения макроскопического феномена жизни.

Знаменитая фраза Эйнштейна «Бог не играет в кости» относилась к вопросу о принципиальной непредсказуемости элементарных событий, поставленному квантовой механикой. В письме от 4 декабря 1926 года Альберт Эйнштейн пишет Максу Борну:

«...Квантовая механика внушает большое почтение. Но внутренний голос говорит мне, что это не истинный Иаков. Теория дает много, но едва ли она подводит нас ближе к тайне Старика. Во всяком случае, я убежден, что Он не играет в кости...»

Английский математик Конуэй придумал оригинальную игру, которую он назвал «Жизнь» [15], потому что она превосходно имитирует рост, распад и изменения, происходящие в популяции живых организмов. Игра подробно описана в табл. 2. Простые правила рождения, выживания и гибели порождают

Таблица 2

Игра «Жизнь» по Конуэю

Игру проводят на плоской поверхности, разбитой на квадратные клетки, например, на доске для игры в го. Каждая клетка может либо оставаться пустой, либо может быть занята фишкой. Состояние зависит от занятости соседних клеток. Соседними считаются восемь клеток, непосредственно примыкающих к данной. Игра складывается из последовательности дискретных шагов, или ходов. Этим имитируется схема поколений. Переход от данного поколения к следующему происходит по определенным правилам, которые применяются одновременно ко всем клеткам доски. Правила:

1. Выживание. Фишка выживает в следующем поколении, если заняты две или три соседних клетки.

2. Гибель. Фишка удаляется с доски, если занято более трех или менее двух соседних клеток. В первом случае система слишком перенаселена, во втором — индивидуум слишком изолирован.

3. Рождение. Пустая клетка занимает фишкой, если заняты три, и только три, соседних с ней клетки.

Эти правила относятся всегда к состоянию, которое существует к началу «хода». Начальное расположение занятых клеток задается самим игроком (здесь заключается единственная возможность для игрока влиять на ход игры). Игра великолепным образом демонстрирует кооперативные изменения в популяции под действием отдельных ее составляющих.

Типичными конечными ситуациями являются гибель, неограниченный рост или стационарное поведение, представленное устойчивыми или осциллирующими конфигурациями. Некото-

рые примеры даны на рис. 13. Конуэй сначала предполагал, что из-за «правила перенаселения» неограниченный рост невозможен. Однако это предположение было опровергнуто открытием таких конфигураций, как «планерное ружье» — это незатухающий, пространственно фиксированный осциллятор («ружье»), который за каждый период испускает фигуру,двигающуюся по доске («планер»). Такие конфигурации, которые производят новые двигающиеся фигуры или уничтожают их, придают игре стратегический характер и делают ее одной из самых интересных стратегических игр.

необозримое многообразие взаимодействующих друг с другом популяционных структур с крайне интересными «функциональными» способностями. Однако весь ход игры подчиняется диктату заранее заданных правил отбора и однозначно определяется начальной конфигурацией. Не происходит ничего такого, что не было бы запланировано с самого начала «полностью информированным Создателем» или по крайней мере не было бы предсказуемо. Вся информация, которая выявляется в ходе игры, содержится уже в начальных условиях. Какой богатой в действительности может оказаться информация, функциональная семантика, показывают примеры рис. 13.

Какой бы интересной и захватывающей ни была бы эта игра, детерминистическое поведение не соответствует концепциям и экспериментальным данным молекулярной биологии — разве что мы захотим их игнорировать, ссылаясь на Создателя. Многие биологи склоняются скорее к другой крайности. Они верят в догму «абсолютного», «слепого» случая. В детерминистической модели постулируемые физические закономерности в большой степени покрыты мраком неизвестности, здесь же физическое объяснение становится вообще неуместным — действия Создателя полностью ускользают от нашего понимания. Флуктуация, находящаяся за пределами всякого ожидания, в принципе непредсказуема, хотя ее и нельзя исключить. Эту ситуацию можно сравнить с лотереей, только здесь «выигрыш» неизмеримо менее вероятен — вспомним примеры из первой главы, иллюстрирующие сложность и неповторимость биологических объектов. После этого ясно, что в действительности означает ссылка на абсолютный случай при объяснении «акта творения».

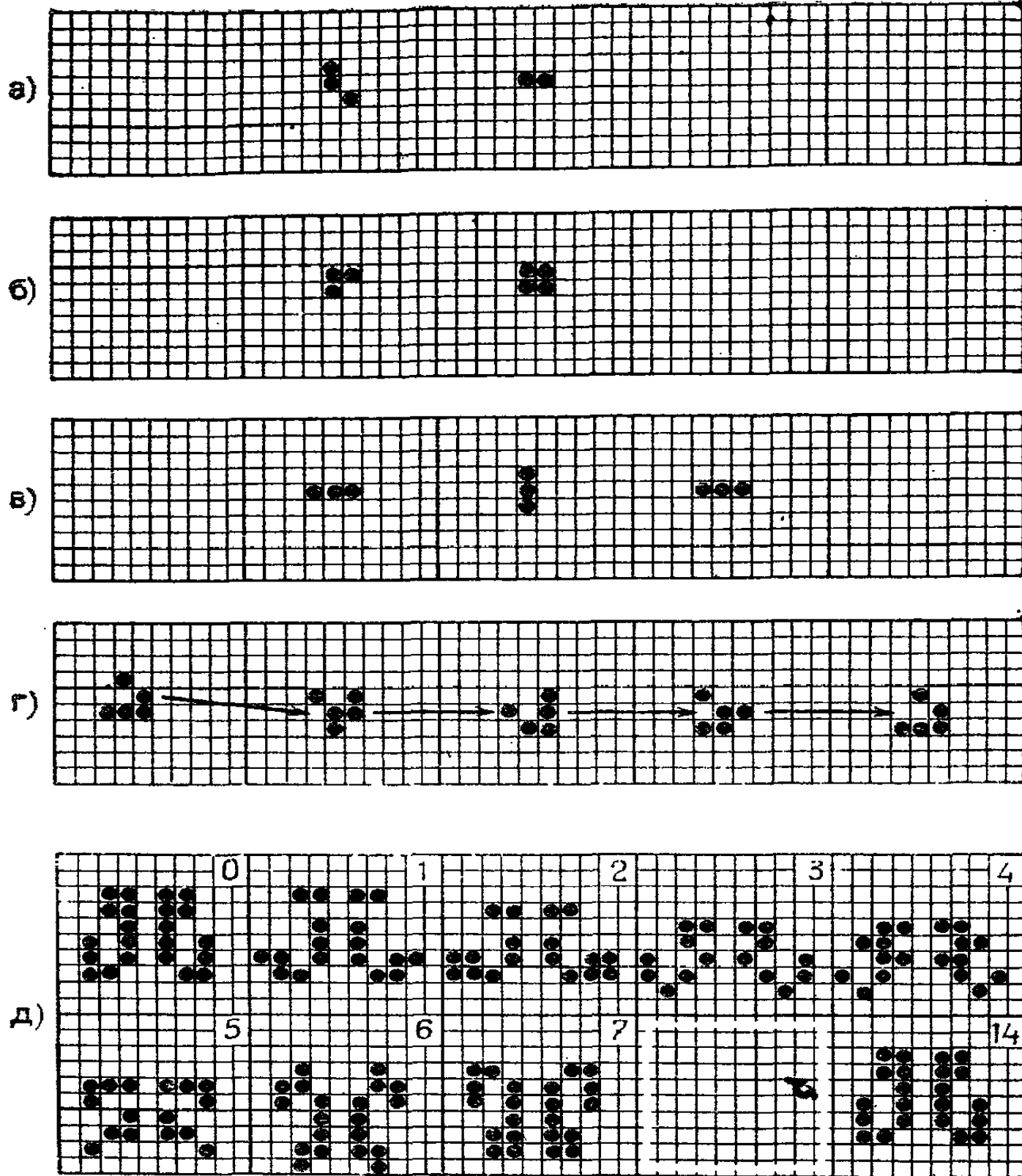


Рис. 13. а — д. Характерные конфигурации в игре «Жизнь» Конуэя. Временная последовательность конфигураций заменена пространственной последовательностью. Последовательные конфигурации «движущихся» фигур изображены на соседних участках доски. а) Триплет, который гибнет. б) Триплет, который переходит в устойчивый блок. в) Осциллирующий триплет «сигнальные огни». г) «Планер», или «парусник», — фигура, которая переходит сама в себя через каждые четыре поколения, сдвинувшись при этом вправо и вниз на один квадрат. д) «Тумблер» — осциллирующая фигура, которая через каждые семь поколений переворачивается, сдвинувшись при этом вниз на один квадрат, а через каждые 14 поколений снова возвращается к исходному состоянию. Необходимо отметить, что этот осциллятор, так же как и «сигнальные огни», не движется по полю как «планер», а все время остается в определенной области доски.

е)

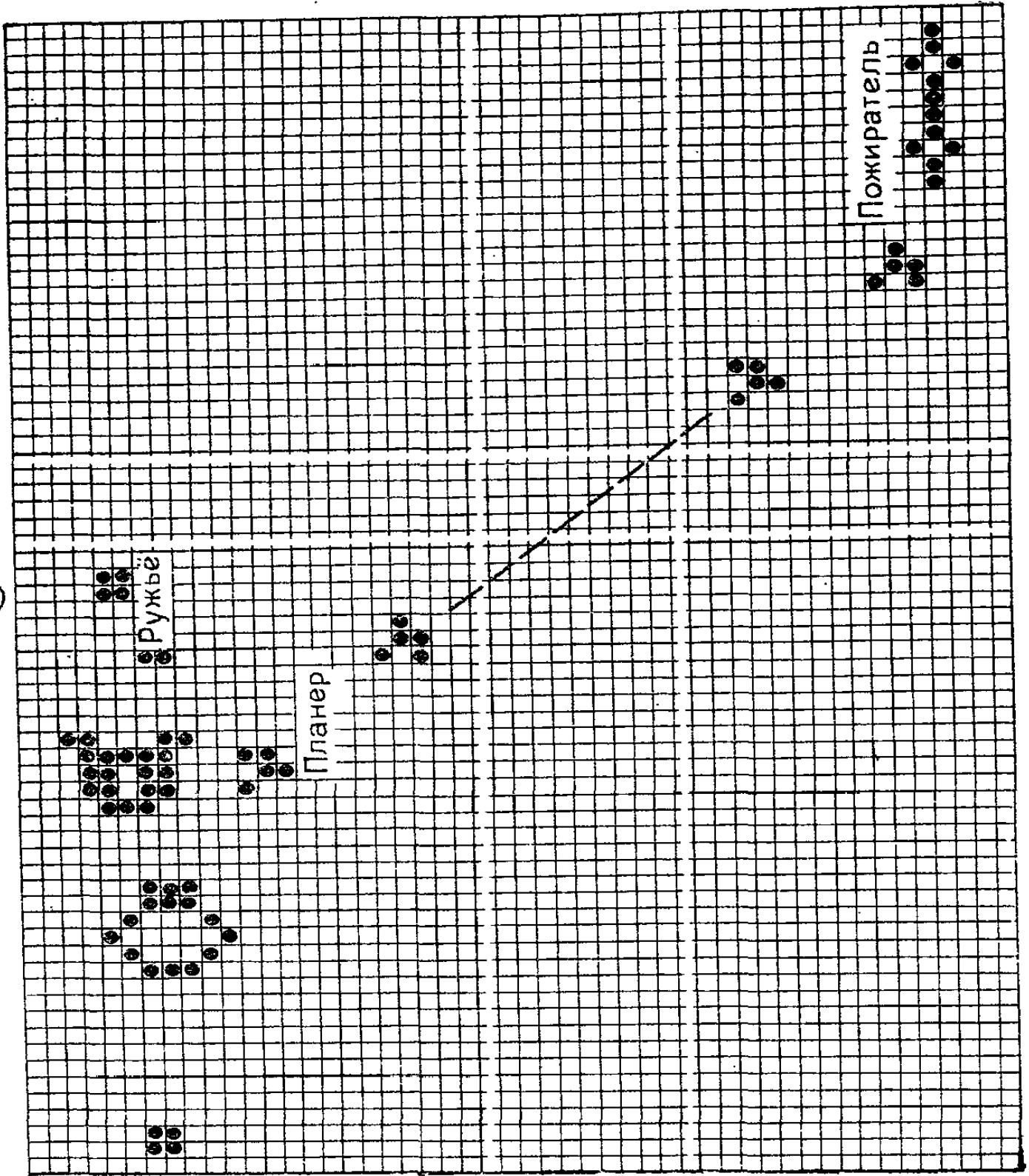


Рис. 13 е. «Планерное ружье» и «пожиратель». «Ружье» — это пространственно фиксированный осциллятор, который через каждые тридцать поколений возвращается в исходное состояние и в течение каждого периода испускает один «планер» (т. е. один раз за 30 поколений). «Планер» передвигается по доске и сталкивается с «пентадекатлоном» (это осциллятор с периодом 15 поколений). Последний «проглатывает» «планер», но сам при этом не изменяется.

Однако мы знаем, что существует еще и средний путь между обеими крайностями. Он тоже исходит из случайных событий, но только из таких, которые имеют конечную вероятность реализации.

Почти во всех лотереях один или несколько игроков выигрывают, хотя их нельзя установить заранее никаким хитроумным способом. Эта исходная ситуация, а именно то, что в лотерее всегда имеются различные выигрыши, уже является ограничением, которое сильно уменьшает неопределенность. В природе или в лабораторном эксперименте эта ситуация будет существенно изменять ожидания по сравнению с гипотезой абсолютного случая, не ограниченного никакими условиями. Если всегда имеется «победитель», то исход каждого эксперимента должен быть положительным, т. е. при воспроизведении условий, аналогичных тем, которые существовали в критическую фазу истории нашей планеты, снова должны были бы возникнуть живые структуры. В деталях — например в генетическом коде, в структурах ферментов и т. д. — они могли бы не иметь никакого сходства с известными нам структурами: они были бы лишь построены по тем же принципам и, таким образом, могли бы тоже оптимально приспособляться к окружающей среде. Но *точное* «случайное попадание» в истории, которое привело бы опять к современным формам жизни, невозможно. Крайне невероятно, чтобы во второй раз в каком-нибудь малом элементе объема в благоприятный момент собрался такой же набор молекул и воспроизвелся исторический процесс эволюции.

Вспомним, что третья «игра в бисер», описанная в предыдущей главе, всегда давала «победителя», при безошибочном воспроизведении даже в том случае, когда для всех конкурирующих видов шаров шансы были абсолютно равны. Конкурирующее поведение и выбор «победителя» — это свойства стационарной самовоспроизводящейся системы реакций вдали от равновесия. Учитывая эти свойства, можно указать *третий*, непрерывный путь к «Творению». При этом должны быть выполнены два условия:

1. Необходимо, чтобы существовал внутренний механизм отбора, который из любого случайного распределения отбирал «победителя», как в третьей «игре в бисер». «Выигрыш» здесь берется по отношению

к соответствующему распределению вероятностей и в общем случае означает не «главный выигрыш», а только «достаточно денег, чтобы можно было играть дальше» или даже лишь «остаться при своих». Победитель (в случае игры — доминирующий, или отобраный, тип шаров) постоянно порождает новых конкурентов вследствие ошибок при воспроизведении. При этом выигрыш становится все выше и выше, как на аукционе.

2. Непрерывный путь эволюции, характеризующийся последовательностью селективных ценностей, должен вести к вершине, т. е. к главному выигрышу. Это означает, что изменчивость должна быть достаточно велика, чтобы победитель всегда мог быть превзойден каким-нибудь конкурентом. Этот оптимизирующий путь к вершине не должен непременно быть монотонным ростом, но, с другой стороны, его «направление» в основном уже определяется дополнительными условиями отбора и точностью воспроизведения.

Давайте снова проиллюстрируем эту ситуацию игрой. Целью нашей *эволюционной игры* будет установление определенной последовательности символов в результате селективных изменений случайной последовательности. Будем пользоваться четырьмя сортами «бисера»: красным, зеленым, синим и желтым, которые можно соединять в цепочку. Кроме того, нам нужна игральная кость в виде тетраэдра, грани которого будут символизировать различные цвета. Предполагается, что эти цвета соответствуют четырем номерам рибонуклеиновых кислот — нуклеотидам А, У, Г и Ц.

Рассмотрим в качестве примера молекулу транспортной РНК (т-РНК — см. рис. 1 — 3), которая состоит из 80 нуклеотидов. В этом случае число возможных альтернативных последовательностей нуклеотидов составляет $4^{80} \approx 10^{48}$. Если бы мы имели тетраэдрическую кость для каждого из 80 положений, то чтобы случайно получилась правильная последовательность, потребовалось бы в среднем около 10^{48} бросаний всего набора костей.

Благодаря взаимодействию между комплементарными основаниями нуклеиновые кислоты имеют способность к самовоспроизведению. В присутствии поли-

меризующего фермента эта репликация даже в бесклеточной среде идет с относительно высокой скоростью и точностью (см. главу 5). Что же здесь нужно считать селективным преимуществом? При «переводе» генетической информации транспортные РНК служат адапторами. Поэтому их структуры должны быть устойчивыми и воспроизводимыми. Однако в водном растворе одноцепочечные РНК очень легко расщепляются — гидролизуются. Нуклеиновые кислоты предохраняются от этого тем, что они свертываются в двухцепочечные структуры. Такое свертывание требует наличия некоторой внутренней комплементарности, т. е. в различных частях последовательности должны существовать участки, комплементарные друг к другу. Далее, большое преимущество дает наличие определенной степени симметрии структуры относительно концов цепи. Ведь каждая молекула при репликации должна сначала произвести свою негативную копию. Но если молекула нуклеиновой кислоты имеет специальную структуру, дающую преимущество при ее узнавании и репликации полимеризующим ферментом, то это преимущество будет полностью использоваться лишь тогда, когда положительная и отрицательная копии обладают им в равной степени — во всяком случае, такой результат дает количественная теория. Только симметрично свернутые двухцепочечные структуры — например «шпилька», «клеверный лист» или многолепестковые формы — удовлетворяют этому условию вследствие самокомплементарности.

Однако пора вернуться к нашей игре!

Возьмем произвольную начальную последовательность и применим к ней правила игры, перечисленные

Т а б л и ц а 3

Игра «Эволюция РНК»

Для игры необходимы строительные элементы, которые можно было бы соединять в цепочку, например бусы. Больше всего подходят пластмассовые шарики с кнопочными разъемами. Эти элементы должны быть раскрашены в четыре цвета — красный, зеленый, синий и желтый, которые будут соответствовать четырем нуклеотидам А, У, Г и Ц. Каждый игрок получает цепочку из N элементов, которые сначала соединены между собой совершенно беспорядочно. Наконец, для игры нужна тетраэдри-

ческая кость, четыре грани которой будут соответствовать четырем различным цветам.

Цель игры состоит в том, чтобы, бросая кость и применяя определенные правила, получить как можно более совершенно «спаренную» последовательность. Начинают со случайной последовательности. Кость бросают поочередно — по кругу. Игрок заранее указывает положение в цепи и для этого избранным им положением бросает кость, за каждый «ход» только один раз. Затем соответствующий элемент заменяется на другой, согласно выпавшему цвету. Бросать кость нужно, однако, только для неспаренных положений. (Здесь в игру входит идея отбора. Каждый круг бросаний соответствует одному поколению. Спаренные области имеют «селективное преимущество», т. е. выживают).

Правила игры:

1. Стерическое правило. Различные области последовательности можно спаривать лишь путем изгибания цепочки в одной плоскости. При этом должны возникать петли. Пять элементов в такой петле не могут быть спарены по стерическим причинам. Может образоваться любая плоская форма: «шпилька», «клеверный лист», «ромашка» и т. д. Можно ввести дополнительное требование, чтобы концы цепи всегда находились рядом — друг против друга. Этим отбираются симметричные структуры, которые в природном процессе эволюции обладают определенным селективным преимуществом, так как у них облегчается взаимное копирование: $(+)$ -цепь \rightleftharpoons $(-)$ -цепь.

2. Правило комплементарности. Если в изогнутой цепочке два элемента с комплементарными цветами (красный — зеленый и синий — желтый) находятся друг против друга и одновременно выполнено третье правило, то они считаются «парой» и скрепляются между собой.

3. Правило кооперативности. Комплементарные элементы можно соединять в пару только тогда, когда при этом получается непрерывная последовательность пар, состоящая не менее чем из четырех красно-зеленых пар, или из двух красно-зеленых и одной сине-желтой пары, или же из двух сине-желтых пар. Для таких «устойчивых» пар больше не нужно бросать кость, они считаются отобранными.

Игра кончается после того, как пройдет заранее установленное число кругов, или «преждевременно», если кто-либо из игроков предъявит полностью спаренную структуру. В обоих случаях выигравшим считается тот, у которого будет наибольшее число очков. Каждая красно-зеленая пара (А — У) дает одно очко, каждая сине-желтая пара (Г — Ц) — два очка. Считаются только те пары, которые находятся в кооперативных участках. Эти правила точно соответствуют действительности: стабильность кооперативных Г — Ц-пар вдвое выше, чем А — У-пар. Те отдельные пары, которые располагаются за пределами кооперативных участков, при комнатной температуре нестабильны. Введение дополнительных правил кооперации и конкуренции может сделать эту игру значительно более уточненной и придать ей захватывающий «стратегический» характер.

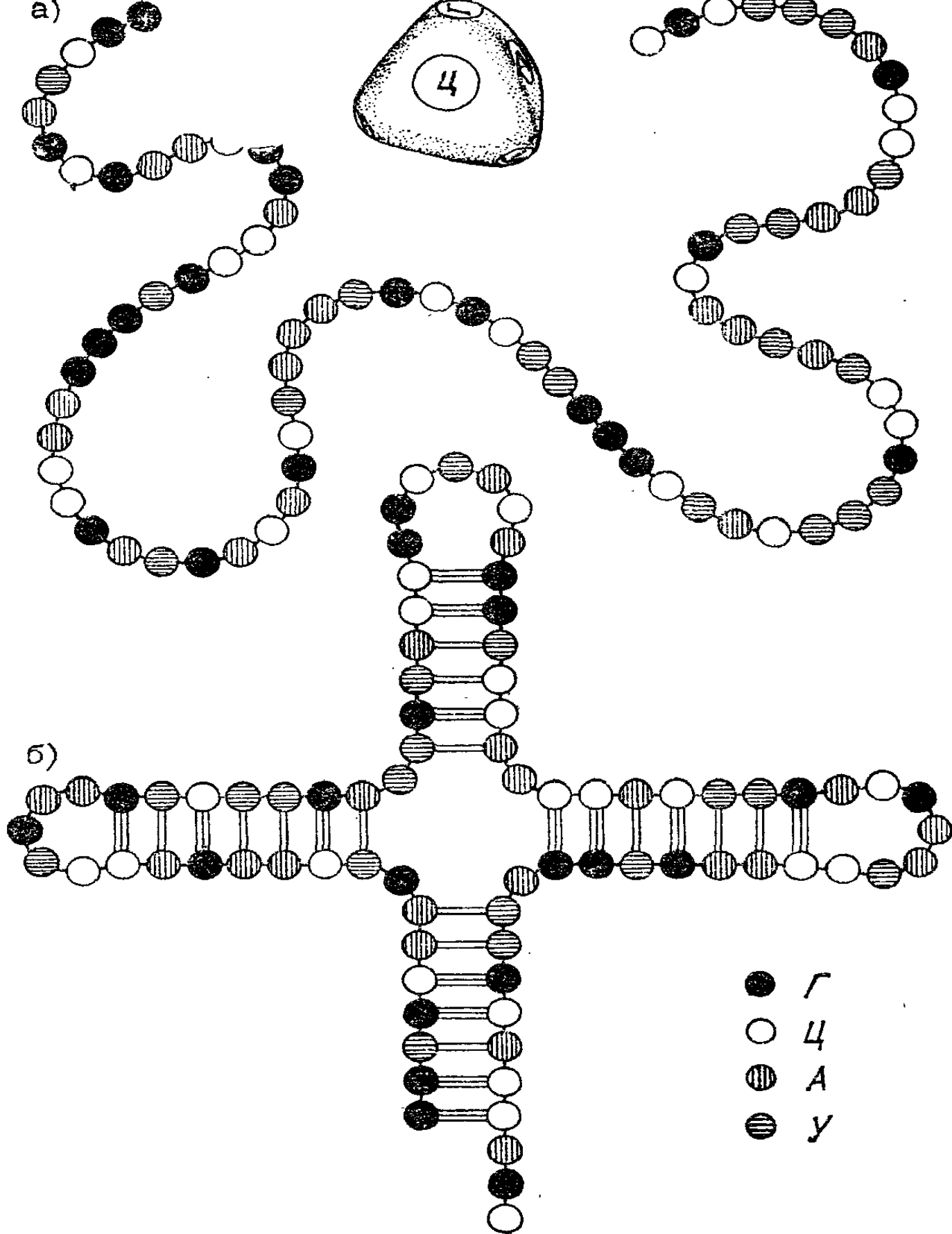


Рис. 14. Начальная и конечная фазы игры «Эволюция РНК». а) К началу игры имеется произвольная последовательность длиной около 80 нуклеотидов. С помощью тетраэдра можно имитировать мутации. Мутация отбирается, если она приводит к образованию комплементарной пары оснований (правила игры даны в табл. 3). б) «Клеверный лист» с оптимальным числом пар оснований. Теория игры (см. табл. 4) предсказывает возникновение таких структур в процессе эволюции.

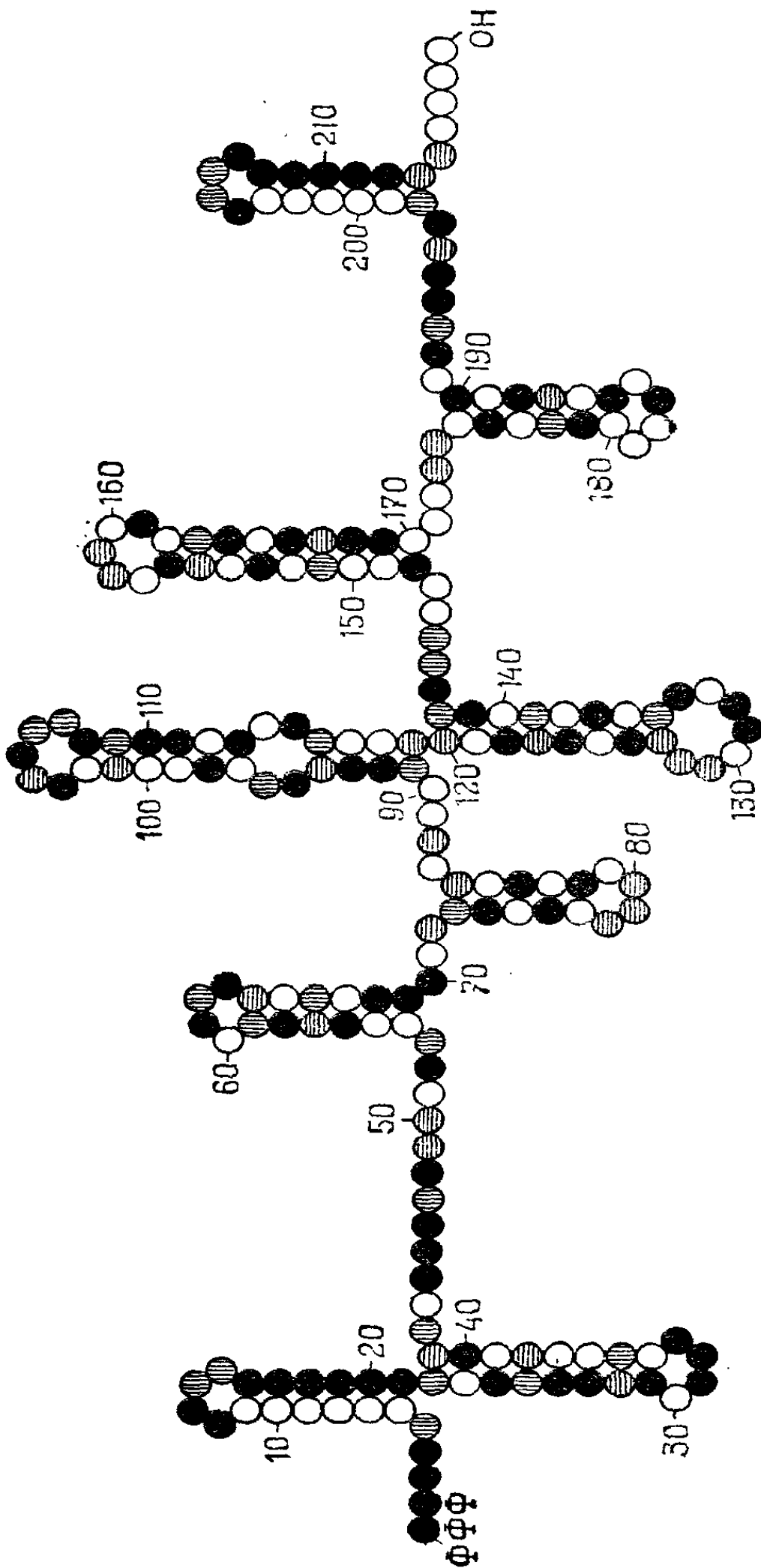


Рис. 15 а. Возможный способ спаривания в фаговой РНК.

б)

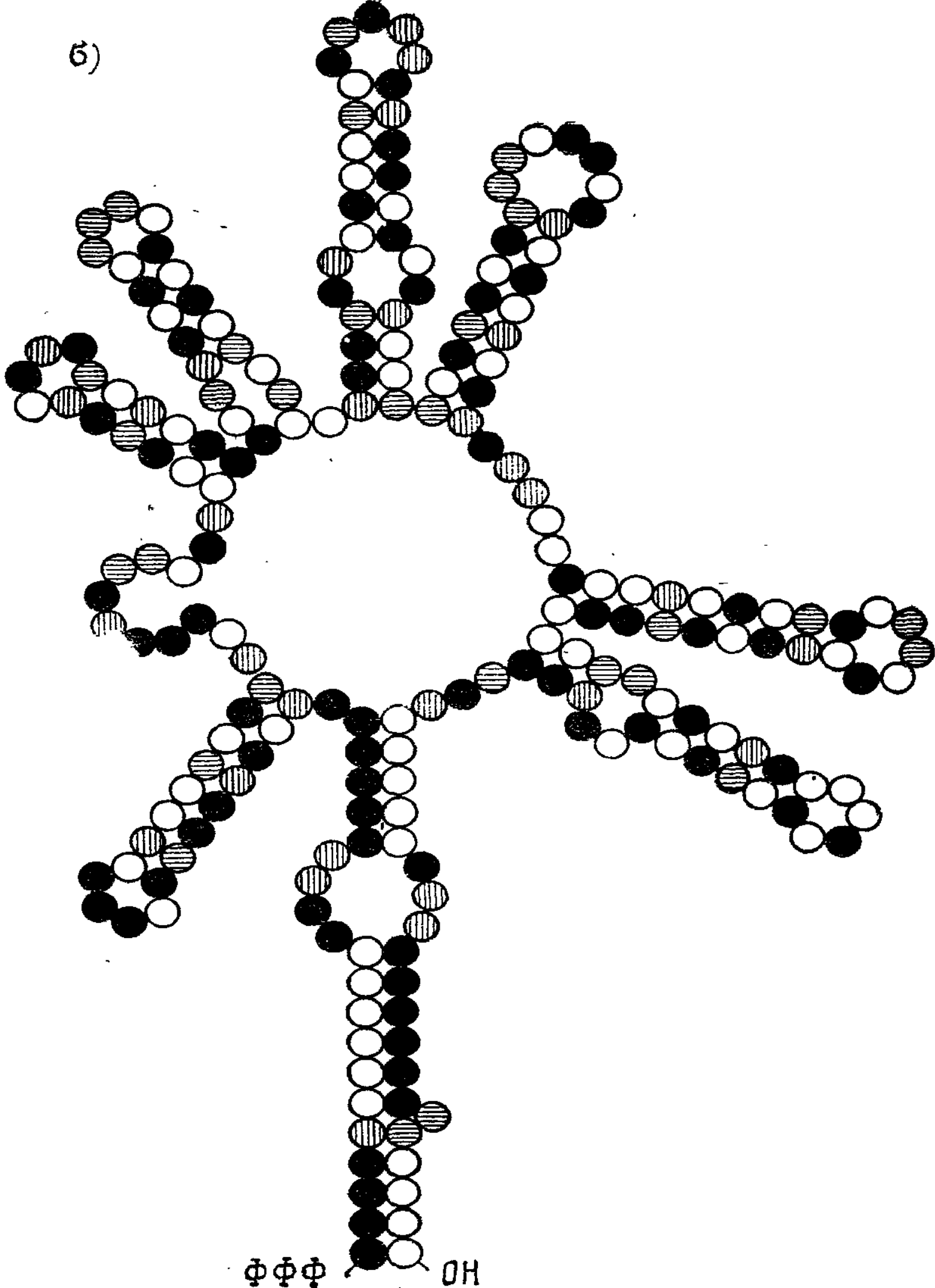


Рис. 15 б. Другой способ спаривания оснований в фаговой РНК.

в табл. 3. Эти правила основываются на количественных исследованиях олигонуклеотидов с различными последовательностями [11]. Таким образом, игра должна приводить к реалистическим выводам. В действительности это так и есть, насколько показывают проведенные до сих пор анализы последовательностей. Рис. 1—3 и 15 дают примеры экспериментально найденных последовательностей оснований в нуклеиновых кислотах.

На рис. 15а и б обозначены два альтернативных способа спаривания фаговой РНК. Фаг Q_{β} имеет специфичный фермент, который избирательно узнает инфекционную фаговую РНК. В пробирке под давлением отбора можно создавать «минимонстры» фаговой РНК. Они уже не инфекционны, но по-прежнему содержат область узнавания исходной молекулы РНК. Благодаря селекционным условиям преимущественно размножается мутант, имеющий максимальную скорость воспроизведения и стабильность. Молекула, изображенная на рисунке, была выделена Спигелманом и сотрудниками. Последовательности нуклеотидов в обеих комплементарных цепях были однозначно определены Миллсом, Крамером и Спигелманом (*Science*, 1973, v. 180, p. 916). Этот эксперимент (см. также рис. 16) точно имитируется эволюционной игрой, описанной на рис. 14 и в табл. 3 и 4. Реальная укладка цепи соответствует предсказаниям теории. По меньшей мере все функционирующие молекулы РНК (ср. также со структурой т-РНК, рис. 1—3 и 14), для которых существенны высокая стабильность и способность к воспроизведению, эволюционируют в природе согласно описанным критериям. Здесь приведены лишь вторичные структуры (укладка в плоскости). Как и в случае т-РНК, пространственная укладка будет приводить к дальнейшей стабилизации структуры.

Игра (табл. 3) правильно предсказывает вторичные структуры (укладки цепей в плоскости), которые из них можно вывести. Для транспортных РНК структура «клеверного листа» достигается быстрее, чем структура «шпильки», а для фаговой РНК, содержащей многие сотни оснований, «выигрывает» всегда структура «ромашки». В табл. 4 и на рис. 14 дан пример, показывающий ход игры. Оптимально спаренную

структуру транспортной РНК с 80 нуклеотидами в эволюционной игре можно получить уже после 80—100 бросаний. Для чисто азартной игры случайное совпадение, которое дало бы полностью спаренный, но в остальном совершенно произвольный «клеверный лист», должно было бы ожидаться в среднем только после $3 \cdot 10^{19}$ бросаний. Преимущество игры, направляемой отбором, становится тем больше, чем длиннее последовательность (см. табл. 4).

Таблица 4

Руководящие указания и примеры к игре «Эволюция РНК»
(см. также рис. 15)

Как легче всего добиться выигрыша в этой игре? Как и в игре Конуэя «Жизнь», важно начинать, имея благоприятную исходную конфигурацию, т. е. определенный тип укладки цепи. Его можно в большой степени определить самому, но чтобы досталась благоприятная исходная последовательность, нужна еще удача.

Рассмотрим цепочку из N элементов. Простейшей симметричной формой со спаренными концами будет шпилька. Она имеет длинную область спаривания, которая заканчивается петлей, — таким образом, для образования пар теряется только пять положений. Сначала может показаться, что это и есть оптимальная структура, поскольку слишком большое число петель должно, очевидно, уменьшать шансы на выигрыш. Максимальное число пар, которое может образоваться в фигуре с k петлями, равно

$$(N - 5k)/2.$$

Однако для исхода игры не столь важно, какое максимальное число пар можно образовать — более важным является, насколько быстро образуется возможно большее число пар. Во-первых, это зависит от удачи при бросаниях кости. Для того чтобы скомплементировать одну пару, в среднем требуется четыре бросания. Во-вторых, очень существенно, сколько пар можно образовать из случайной последовательности в самом начале игры. Здесь шпилька со сцепленными концами очень невыгодна, потому что она содержит лишь одну возможность расположения пар. Можно ожидать, что в начале игры будет иметься

в среднем только $\frac{1}{4} \cdot \frac{N-5}{2}$ пар. Если улыбнется счастье, то число пар будет несколько больше, при неудаче их будет меньше. Однако со средним поведением нигде нельзя продвигаться очень далеко. Секрет выигрыша состоит в том, что игру надо начинать с возможно большим положительным отклонением от среднего значения. Относительное отклонение тем больше, чем короче области спаривания. Положительные и отрицательные отклонения встречаются, однако, с равной вероят-

ностью. Для коротких областей нужно было бы иметь в своем распоряжении очень много альтернативных возможностей спаривания, чтобы выбрать из них наилучшую комбинацию с положительным отклонением. Победителем будет тот, кто сумеет найти оптимальный компромисс между требованием «не очень большого числа петель» и требованием «как можно более коротких последовательностей». Структура шпильки связывает игрока по рукам и ногам. Значительно более благоприятной является уже T -структура, в которой можно варьировать относительные длины всех трех плеч. Число элементов в каждом плече должно быть не меньше тех минимальных значений, которые указываются правилом кооперативности. При этом для заданной T -структуры ожидаемое значение начального числа пар снижается до $\frac{1}{4} \cdot \frac{N-10}{2}$ (по сравнению с $\frac{1}{4} \cdot \frac{N-5}{2}$ для структуры шпильки). Однако для больших значений N это несущественно. Выигрывается много больше благодаря тому, что теперь можно испробовать большое число возможных T -структур и выбрать из них ту, которая соответствует наибольшему положительному отклонению. Конечно, останавливаться на T -структурах нет никаких оснований — можно испробовать структуры трех- или четырехдольного клеверного листа или, наконец, структуру цветка со многими лепестками.

Где же граница?

Поскольку для образования одной кооперативной области в среднем необходимы три пары, максимальное число k_{\max} лепестков определяется следующим условием:

$$\frac{N - 5k_{\max}}{2} = (k_{\max} + 1) \cdot 3, \quad k_{\max} = \frac{N - 6}{11}.$$

Левая сторона этого уравнения дает максимальное число пар в структуре, содержащей k петель (лепестков). Правая сторона содержит утверждение, что каждый из k лепестков (а также стебель, т. е. всего $k+1$ область) в среднем должен содержать не менее чем три пары. Конечно, структура с k_{\max} лепестками снова крайне неблагоприятна, потому что она (как и шпилька) имеет лишь одну возможность расположения пар π , сверх этого, теряет $5k$ положений для образования петель. Поэтому можно ожидать, что оптимальной окажется некая промежуточная структура с $k < k_{\max}$ лепестками. Этот оптимум легко рассчитать. Для $N = 80$ (T -РНК) предпочтительной структурой является трех- или четырехлепестковый клеверный лист (для этого случая $k_{\max} = 6,7$). Для $N = 220$ («минивариант» РНК-фага Q_{β}) оптимум следует ожидать для $k = 9-10$ ($k_{\max} = 19,5$). Природа, очевидно, знала этот секрет, о чем свидетельствуют вторичные структуры реальных функционирующих молекул РНК, которые изображены на рис. 1 — 3 и 15. Дальнейшая «пространственная укладка» структуры в этой игре не учитывается — для этого потребовались бы более сложные правила. Однако при-

веденные здесь правила легко обобщить как в этом направлении, так и для того, чтобы они учитывали другие, более сложные, селективные преимущества.

Отличие от игры Конуэя состоит в том, что здесь участвует случай: во-первых, берется произвольная начальная последовательность, во-вторых, для спаривания используется бросание кости. Поэтому в этой игре отдельные последовательности и структуры не определяются детерминистическим образом, хотя при не очень больших значениях N число возможных альтернатив еще относительно невелико.

Эту эволюционную игру можно варьировать многими способами. Например, два игрока могут объединить свои усилия и, обмениваясь участками цепей — т. е. используя генетическую рекомбинацию, могут повысить скорость эволюции (при этом обмен будет одним из исходов при специальном бросании кости). Если ввести кодирование, то можно придумывать самые разнообразные игры с «переводом». Например, с помощью икосаэдра, двадцать грапей которого символизируют двадцать природных аминокислот, можно было бы «генерировать» белковые структуры. Наша фантазия так же мало ограничена, как и страсть природы к вариациям. Самопроизвольный ход эволюционной игры в природе, как и в игре «Жизнь» Конуэя, обеспечивается селекционными свойствами, которые имитировала третья «игра в бисер» (см. стр. 43). Эти свойства позволяют обходиться без игрока и наделяют систему способностью к самоорганизации. Предполагается лишь наличие качеств, общих для всех живых систем: метаболизм, самовоспроизведение и мутабельность. Математическое описание таких «природных» игр требует синтеза различных теорий:

1) теории вероятностей для описания прироста информации,

2) теории случайных процессов для описания «неустойчивостей»,

3) теории игр для определения оптимального пути эволюции.

Теория игр, или теория оптимизации, восходит к Джону фон Нейману [16]. Ее применение к естественной эволюции показывает, что нельзя определенно предсказать ни конечное состояние, ни путь к нему,

их можно только в известной степени ограничить с помощью критериев оптимизации. Для линейной системы оптимальный путь означает монотонное возрастание «селективной ценности». Правда, здесь случайное исходное распределение бесповоротно определяет путь. Вступив на эту горную тропу, монотонно ведущую все выше и выше, ее нельзя более покинуть, даже если имеется другой, более короткий и крутой, но местами спускающийся путь. И из всех возможных вершин можно достигнуть только одной, вовсе не обязательно самой высокой. Для этого случая наша эволюционная игра дает впечатляющий пример. Эта игра становится интересной только из-за того, что в ней происходит выбор между возможными альтернативами.

При нелинейном механизме условие монотонного роста функции ценности становится необязательным — в известных случаях система может перемещаться вверх и вниз. Однако связи, проявляющиеся в нелинейностях, накладывают иные ограничения на возможные пути эволюции.

Лишь состояние истинного равновесия является абсолютным экстремумом (максимумом энтропии или минимумом свободной энергии), независимо от пути: здесь микроскопически обратимый процесс всегда приводит к одному и тому же результату (это моделировала также первая «игра в бисер»). Для предельного случая больших чисел установление равновесия является детерминистическим процессом, который можно было бы вычислить заранее, зная параметры системы. Это не имеет места в случае эволюционного процесса оптимизации, который характеризуется соотношением $\bar{E}(t) \rightarrow W_{\text{opt}}$ и на путь которого наложены ограничения — дополнительные условия, заданные в виде неравенств.

Теперь мы можем ответить на фундаментальный вопрос, поставленный в конце предыдущей главы. Процесс эволюции в «малонаселенном информационном пространстве» из-за взаимодействия закона и случая действительно означает *возникновение информации*, а не только *выявление информации*, потому что он допускает альтернативы с различными целями. Что такое «малонаселенное информационное пространство»? Пространство — это множество всех возможных

расположений. «Малонаселенное» означает, что в этом пространстве число занятых состояний очень мало по сравнению с общим числом всех состояний. Это справедливо уже в случае множества возможных расположений для молекулы белка. Из 10^{130} возможных расположений аминокислот в полипептидной цепи в земных масштабах можно было бы реализовать лишь исчезающе малую долю. Это тем более справедливо для клетки в целом, которая является мельчайшей единицей живых организмов, или для объединения клеток — такого, как сеть нейронов в мозгу. Поэтому макромолекулярную эволюцию, морфогенез, основанный на избирательных межклеточных взаимодействиях, и, наконец, деятельность нашего духа, которая складывается из специфически отфильтрованного и скомбинированного друг с другом функционирования различных нервных клеток, нельзя больше описывать детерминистическим образом *).

Итак, Бог играет в кости?

Конечно! Однако и Он следует своим правилам игры.

Это дополнение в большой степени нейтрализует постановку вопроса, которая кажется сначала провокационной. Необходимо также уточнить, что Он — это Бог философов, синоним Природы, а не индивидуальный Бог, отражение человека, Бог сострадания, который находится за пределами всякой обсуждавшейся здесь тематики **).

*) Создание новой информации есть запоминание случайного выбора (Кастлер). В этом смысле модель Эйгена распадается на две части. Выбор «малонаселенного» участка фазового пространства («ящик Эйгена») означает возникновение информации. Но судьба макромолекул в этом «ящике», отбор цепей с максимальной селективной ценностью, есть выявление информации. (Прим. ред.).

***) Авторы употребляют слово «Бог» так же, как Эйнштейн, а до него Спиноза. Бог — это реально существующая Природа, познанием которой заняты естественные пауки. (Прим. ред.)

СТУПЕНИ К ДЕЙСТВИТЕЛЬНОСТИ

В предыдущих главах мы уделяли мало внимания реальным явлениям жизни. Нужно было понять принцип. Однако «жизнь» в полном ее смысле можно понять только в реальных проявлениях. Поэтому в последней главе мы попытаемся подняться по крутым ступеням к действительности.

Первая ступень: накопитель информации.

Ранние стадии возникновения жизни не наблюдаемы нами. В природе не обнаруживаются предшественники живых клеток. Ни филогенез, ни онтогенез, по видимому, не законсервировали для нас эти ступени. Конечно, можно высказывать догадки о происхождении отдельных клеточных органелл. Так, например, естественно задаться вопросом: нельзя ли рассматривать рибосомы, фабрики белков, как свидетелей доклеточного прошлого, содержащих сведения о начале трансляции, т. е. переводе генетической информации? Сходным образом сейчас думают, что митохондрии и хлоропласты — это остатки прокариотических клеток, симбиоз которых с другими клетками ознаменовал начало перехода к эукариотам. Позволят ли такие дедуктивные умозаключения реконструировать исторический ход доклеточной эволюции, покажет время. Однако возможен также индуктивный подход, и можно попытаться разрешить некоторые вопросы экспериментальным путем. Это, конечно, не означает, что кто-то сразу же серьезно возьмется за получение «синтетической» жизни в пробирке. Как однажды метко сказал Жак Моно [17], «великолепную симфонию живой природы не так-то просто исполнить нам — дилетантам».

Воспользуемся еще раз этим сравнением: симфония жизни постоянно исполняется перед нами природой. Нам нужно только присоединиться к ней и подыгрывать на том или другом инструменте. Индустрия

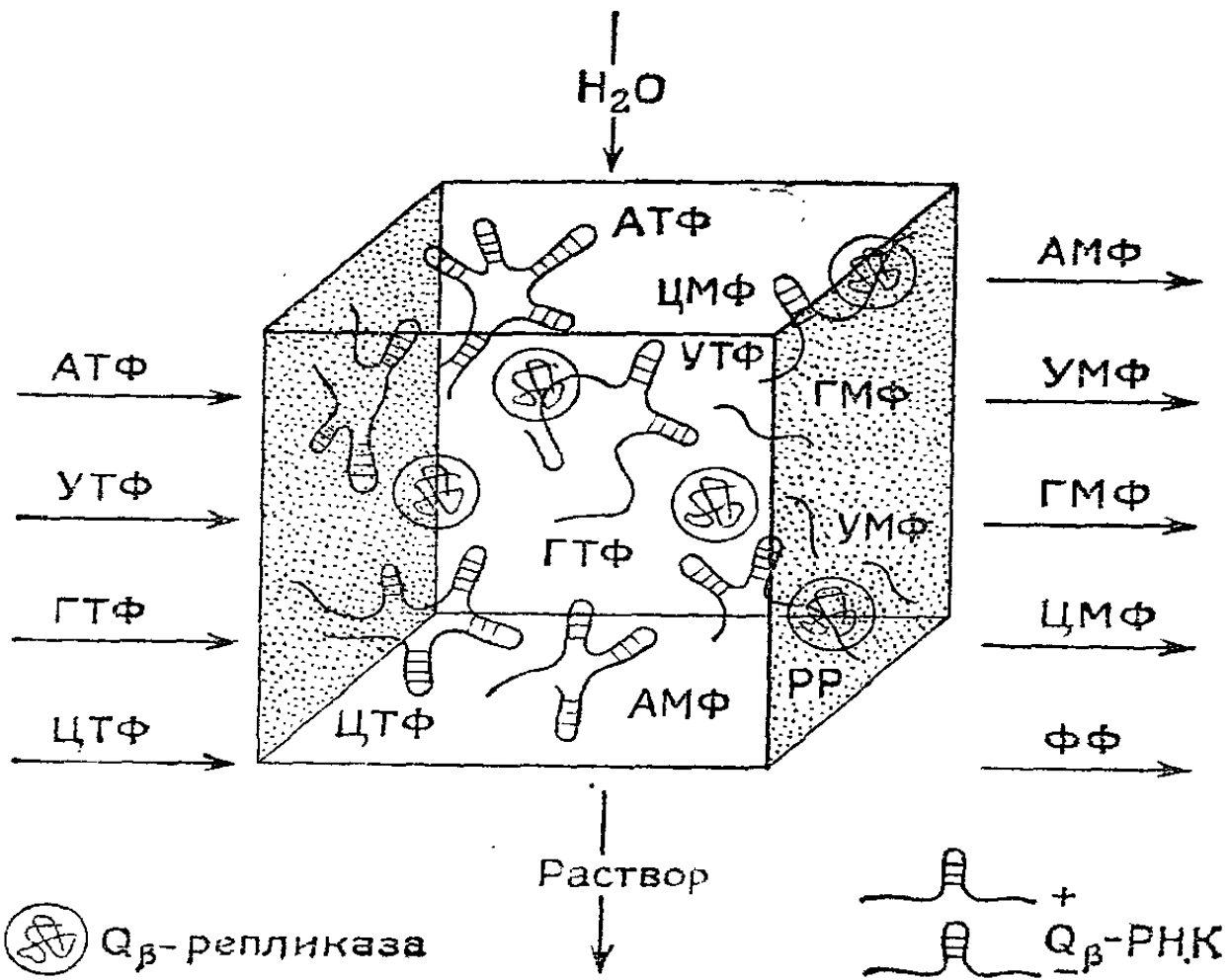


Рис. 16. «Эволюционная машина». Схема эволюционного эксперимента, разработанного в Геттингенском институте им. Макса Планка (Эйген, Зумпер, Кюпперс и Бибрихер). В реактор втекают регулируемые потоки высокоэнергетического строительного материала (нуклеозидтрифосфаты А, У, Г и Ц), а также растворитель, в котором содержатся необходимые кофакторы (буфер, соли и т. д.). Из реактора сквозь мембранный фильтр вытекают продукты синтеза (включая новообразованные макромолекулы) и продукты распада (нуклеозидмонофосфаты и пирофосфат). Реактор содержит активный фермент (репликазу) в необходимой концентрации. Мономеры индивидуально мечены радиоактивными изотопами. Макромолекулярный состав содержимого реактора анализируется через определенные интервалы времени. Эксперимент можно проводить либо в непрерывной, либо в периодической форме. В последней форме он во многом соответствует эволюционным опытам Спигелмана,

грампластинок давно открыла этот принцип. Пластинки типа «минус один» (Spiel-Mit) позволяют даже дилетанту стать «участником» выдающегося ансамбля.

На самом деле это тот тип экспериментов, которые можно реализовать уже сейчас, используя физический

и химический репертуары молекулярной биологии. При этом мы исходим из естественной системы и пытаемся лишь заменить один или несколько ее компонентов на их «дилетантские» варианты, которые, однако, имеют возможность «упражняться», т. е. в нашем случае эволюционировать. Прототипом таких экспериментов были опыты, поставленные Солом Спигелманом [6]. Схема эволюционного эксперимента представлена на рис. 16. В принципе такой эксперимент соответствует селекционной игре, которая описана в третьей главе, только здесь вместо бус взяты молекулы нуклеиновой кислоты. Под искусственным давлением отбора генетический материал фага порождает молекулярных «монстров», которые лучше адаптированы к своей новой ненатуральной среде, чем их «естественные» прародители.

Теория эволюции [3] позволяет количественно описывать эти опыты. Эксперименты подобного рода — как с естественными, так и с синтетическими нуклеиновыми кислотами — в настоящее время уже проводятся в ряде лабораторий, а в скором времени будут, несомненно, проведены и с белками в качестве фенотипических «мишеней».

Какие проблемы можно разрешить с помощью таких экспериментов? Химики прежде всего интересуют происхождение макромолекул и мономеров. Но это проблема чистой химии. Из лабораторных экспериментов и из анализа космического вещества мы знаем, что при подходящих условиях, т. е. в восстановительной атмосфере под действием ультрафиолетового излучения, электрических разрядов или крайних температурных колебаний, «в принципе» могут возникнуть все известные нам мономеры биологических макромолекул, причем в высокоэнергетической форме, так что они способны спонтанно объединяться с образованием полимеров. Итак, существование материального аккумулятора в достаточной мере оправдывается известными нам законами химии и физики.

Но как же возникла упорядоченность? Здесь существенную поддержку оказывают эволюционные эксперименты, упомянутые выше.

Масштаб воспроизводимой упорядоченности, т. е. емкость накопителя информации, ограничивается точностью узнавания символов!

Этот тезис имеет решающее значение для понимания первой ступени самоорганизации. Рассмотрим молекулу пуклейновой кислоты, состоящую из v единиц (см., например, рис. 15). При репликации каждая единица должна опознаваться своим комплементарным партнером. Допустим, что это происходит с точностью q . Вероятностный фактор q нормирован на единицу. Например, $q = 0,99$ означает, что в среднем из 100 процессов копирования только 99 происходит правильно. Для последовательности, состоящей из v единиц, точность, или «качество переноса информации», равняется

$$q \cdot q \cdot q \dots = q^v \equiv Q.$$

В действительности каждая из четырех букв А, У, Г, Ц, более того — каждая из 16 различных комбинаций пар букв характеризуется своим значением фактора точности q . Кроме того, его значение может зависеть от степени внутримолекулярного спаривания оснований. Однако для достаточно длинных цепей будет существовать постоянное среднее значение $\langle q \rangle$. Для дальнейших оценок важен тот факт, что значение $\langle q \rangle$ близко к единице.

Теперь существенный момент. Оптимизирующий процесс отбора характеризуется следующим пороговым соотношением для ценности:

$$W_m > \bar{E}_{k \neq m},$$

т. е. селективная ценность доминирующей последовательности (W_m) должна превышать среднюю продуктивность всех конкурентов ($\bar{E}_{k \neq m}$). При заданном распределении кинетических параметров это означает также, что существует критическое значение фактора качества Q_m . Обозначим это пороговое значение Q_{\min} . Оно однозначно определяется кинетическими параметрами отобранной последовательности m и средними значениями кинетических параметров всех ее конкурентов $k \neq m$. Последовательности, для которых $Q < Q_{\min}$, не могут быть отобраны воспроизводимым образом. Из связи между Q , $\langle q \rangle$ и v следует важное соотношение

$$v_{\max} = \frac{|\ln Q_{\min}|}{1 - \langle q \rangle}.$$

Итак, пороговое значение качества Q_{\min} определяет

максимальную длину v_{\max} последовательности, которая еще может воспроизводимо реплицироваться. Величина v_{\max} мало зависит от различий в кинетических параметрах, влияющих на значение Q_{\min} , поскольку Q_{\min} входит в выражение для v_{\max} только в виде логарифмического члена. Так, увеличение скорости образования отобранной последовательности в 10^6 раз привело бы лишь к шестикратному увеличению емкости информационного аккумулятора. На самом деле в ходе эволюции очень быстро произошла оптимизация скоростей. Эффективность катализаторов вряд ли заметно изменилась от кишечной палочки до человека. С другой стороны, достижимая информационная емкость очень сильно зависит от средней точности распознавания символов $\langle q \rangle$. Для неферментативного переноса информации (который зависит только от кооперативного взаимодействия комплементарных нуклеотидов), происходящего в оптимальных условиях, экспериментальное значение $\langle q \rangle$ равняется примерно 0,99 (в большинстве случаев меньше 0,99). Повышение точности до 0,99999, как это имеет место в экспериментах Спигелмана и обеспечивается там фагоспецифичным ферментом репликации, означает тысячекратное увеличение количества информации, которое может быть накоплено. Для кишечной палочки оно превышает миллион элементарных знаков, для высших организмов оно оказывается более миллиарда. Природе пришлось выдумать для этого уйму всяких приспособлений: кооперативную комплементарность нуклеотидов, регулируемое узнавание в активном центре реплицирующего фермента, контроль посредством специфичных ферментов репарации и т. д. — до избыточности последовательностей нуклеиновых кислот, которая увеличивается с размером генома.

Из этих соображений следуют два существенных вывода относительно строения накопителя информации:

1. В отсутствие ферментативных (или других аналогичных) факторов, повышающих качество узнавания, информационная емкость нуклеиновых кислот настолько мала, что на этой основе нельзя построить никакой эффективной исполнительной системы. Такая система должна содержать ряд факторов, необходимых для репликации, сопряжения (Kopplung) и трансляции. При линейном механизме комплементарной ре-

ликация конкуренция может привести лишь к отбору одной-единственной, относительно короткой последовательности.

2. После того как на более поздних этапах эволюции наконец были выполнены предпосылки, необходимые для высокой точности узнавания символов, скорость эволюции, основанной на простой репликации, неизбежно должна была становиться все меньше и меньше. Выход из этого тупика дало половое размножение, т. е. рекомбинационный обмен отдельных участков между гомологическими областями различных молекул нуклеиновой кислоты. Таким образом, сексуальность является необходимым следствием постоянного роста качества переноса информации. Ее появление переместило область воздействия эволюции с индивидуальной последовательности поколений на вид в его целостности. Франсуа Жакоб [18] указал на то, что расплатой за этот прогресс стала эволюционная необходимость смерти, причем запрограммированной, а не случайной смерти от несчастного случая. Когда генофонд вида постоянно перемещивается из-за процессов половой рекомбинации, вклад отдельного индивидуума должен оставаться ограниченным конечным интервалом времени, иначе он в конце концов сам уничтожит весь прогресс. Не сможет ли человек со своими способностями, далеко выходящими за рамки генетической программы, снова устранить смерть?

Наконец, хотелось бы разобрать еще один вопрос. Необходимым условием существования накопителя информации является способность к самовоспроизведению. Действительно ли ее можно было обеспечить лишь на основе комплементарного механизма узнавания? Ответ удалось получить лишь после проведения количественных исследований, и он оказался «да». При неодинаковых распространенностях различных мономеров, как это имеет место в нерегулируемых природных процессах, прямое самовоспроизведение привело бы к отбору гомополимеров, образованных повторением единиц только одного, наиболее распространенного типа. Лишь комплементарность гарантирует отбор смешанных, по меньшей мере двоичных последовательностей символов, наличие которых является необходимым условием возникновения схемы кодирования. С другой стороны, плюс-минус компле-

ментарность дает простейшую и наиболее эффективную систему, обеспечивающую выполнение этого условия.

Вторая ступень: трансляция.

Из соображений, изложенных выше, следует, что нуклеиновые кислоты своими силами не могут подняться на «следующий этаж», где они могли бы использовать свои выдающиеся способности накапливать информацию для создания эффективно функционирующей системы своей собственной репликации. Они не могут обойтись без помощи функционально действенных белков. Но в таком случае нельзя ли построить самовоспроизводящуюся систему из одних белков?

Для этого белки должны были бы использовать свой практически неограниченный каталитический потенциал. Например, из аминокислот можно было бы строить специфичные олигопептиды, а затем — в несколько приемов — соединять их друг с другом, пока не получатся длинные цепи — первичные структуры белков. Эта способность к самоорганизации отличается, однако, в двух отношениях от изначальной способности нуклеиновых кислот к комплементарному распознаванию:

1. Способность к самовоспроизведению здесь может появиться лишь у довольно сложной сети каталитических реакций, в которой имеются циклически замкнутые реакционные пути. О «комплементарности» здесь можно говорить только в широком смысле, например А дает В, В → С и т. д. до Х, который снова должен давать А (рис. 17).

2. Если каталитическое самовоспроизведение имеет место, то оно всегда является результатом специфических взаимодействий. Оно не является неотъемлемым свойством пространственно свернутых белковых структур. Например, в случае мутации данная каталитическая активность может быть полностью утрачена.

Оба эти свойства резко ограничивают способность эволюционировать. Сложные зацикленные сети реакций не могут освободиться от «предшественников», нет также простого способа реализовать селективное преимущество, ибо для этого благоприятная мутация всегда должна была бы вызывать серию других определенных мутаций, чтобы подействовать на свой собственный источник и посредством этого селективно усиливаться (см. рис. 17).

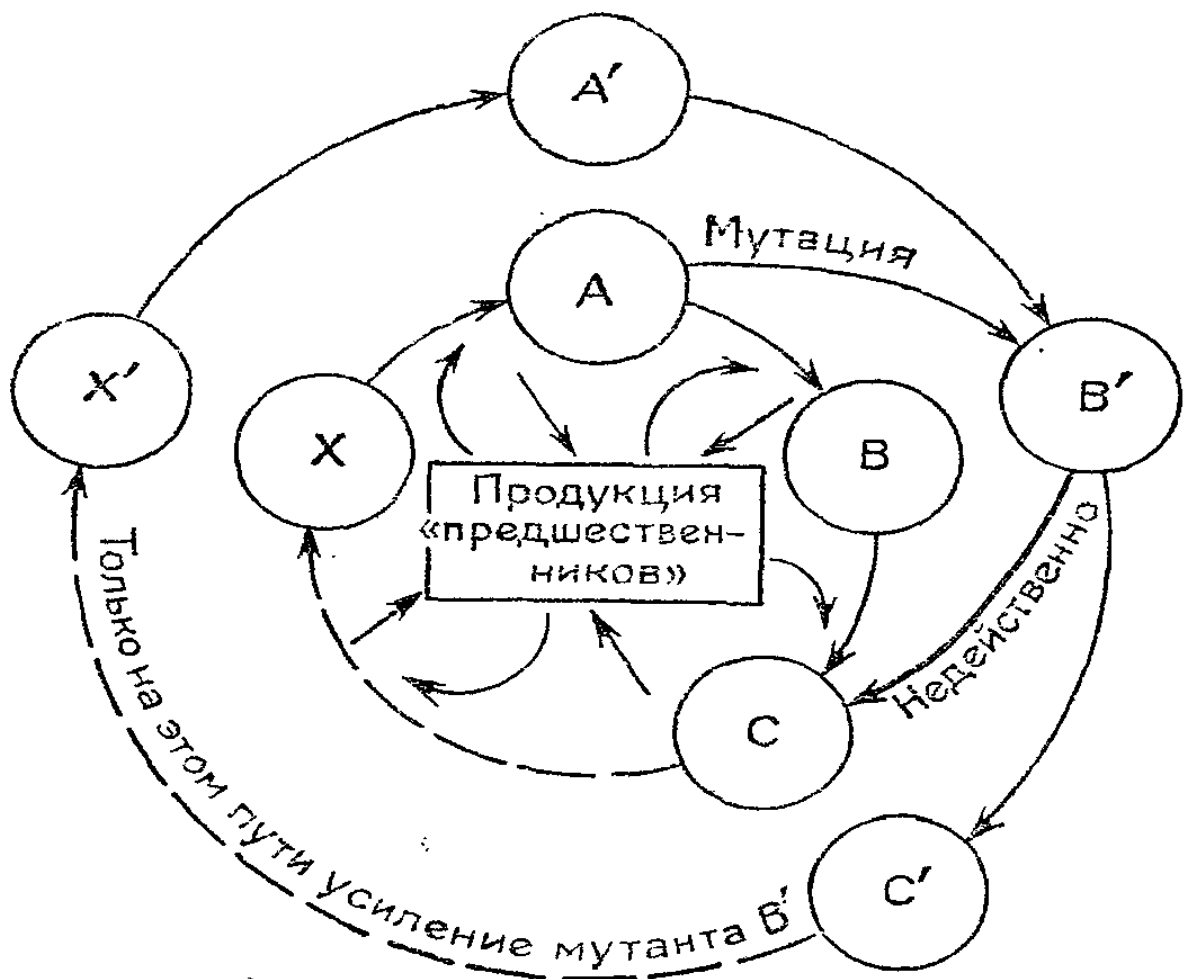


Рис. 17. Каталитический цикл. В каталитическом цикле вещество воспроизводится, проходя через ряд каталитически активных промежуточных стадий. Поскольку белки изначально не обладают способностью инструктировать образование идентичной или комплементарной последовательности и в то же время отдельные структуры могут быть каталитически активны, образование такого цикла являлось бы для белков единственной возможностью самовоспроизведения. Для этого некоторые компоненты цикла или сами по себе, или с помощью связанных с ними ферментов должны были бы катализировать синтез олигопептидов («предшественников») из активированных аминокислот. (В природе такие ферменты известны: например, один декапептид может синтезироваться непосредственно из аминокислот. «Предшественники» большего размера могли бы образоваться путем соединения олигопептидов друг с другом.) В цикле из олигопептидов каталитическим путем снова образуются каталитически активные структуры А, В, С, ..., Х, причем Х должен быть катализатором образования А. Поскольку цикл должен продуцировать «предшественников», число промежуточных стадий должно быть относительно велико.

В таком каталитическом цикле мутация дает преимущество только в том случае, если она будет избирательно способствовать своему собственному воспроизведению. Но это означает, что она должна вызывать серию новых определенных мутаций, которые приводят к образованию совершенно нового цикла (революция). Вероятность такой выгодной серии мутаций крайне мала. Поэтому белки не могут эволюционировать самостоятельно. Но их функциональные способности настолько велики, что при синтезе *de novo* всегда образуется какая-то доля функционально активных (хотя и мало адаптированных) структур.

Правда, в смеси полипептидных цепей со случайными последовательностями были найдены воспроизводимые каталитические свойства, которые были интерпретированы как самоорганизация. Но это обстоятельство только еще один раз демонстрирует большие функциональные способности белков. В каждом случайном распределении имеется достаточно большое число активных структур. Они адаптированы не оптимально, но тем не менее проявляют определенные каталитические качества. Воспроизводится не отдельная последовательность, а только паличие определенной функции, которая представлена большим набором каталитически активных структур. Но для эволюции, т. е. для оптимальной адаптируемости, последовательности — причем любые — должны были бы обладать способностью к самовоспроизведению,носящей характер неотъемлемого свойства полипептидных цепей. У белков подобные свойства по меньшей мере не обнаружены.

Итак, мы приходим к выводу, что генотип и фенотип с необходимостью должны быть представлены двумя различными классами молекул.

Как же могла возникнуть однозначная система трансляции, особенно если вначале не существовало никаких избирательных взаимодействий между соответствующими друг другу элементами обеих молекулярных классов?

Рассмотрим два алфавита, каждый из которых состоит из двух элементарных знаков, т. е. из двух букв: А, В и а, в. Наша система будет состоять из произвольных последовательностей букв, таких как

ААВАВВВАВАВВААА... и bababbbbbaaaabaab...

Допустим, что любые последовательности заглавных букв имеют способность к самовоспроизведению. Последовательности маленьких букв сами по себе такой способностью не обладают, для этого им требуются инструкции в виде последовательностей заглавных букв. Допустим, что последовательности маленьких букв способны зато выполнять любые функции, например, устанавливать соответствие маленьких и заглавных букв.

Вероятно, будет излишним еще раз напоминать, что здесь речь идет всего лишь об абстрактной моде-

ли. В природе вряд ли удалось бы реализовать систему функционирования, использующую только два класса единиц. Но эта модель легко обобщается на случай произвольного числа классов.

Для трансляции необходимо ввести однозначное соответствие между заглавными и маленькими буквами. Допустим, что такого соответствия, основанного, например, на изначальных взаимодействиях между А и а, между В и в (как это имеет место в случае нуклеиновых кислот и выражается в комплементарности АА', ВВ'), вначале не существует. Иначе проблема была бы тривиальной. Вместо этого допустим, что в наборе случайных последовательностей, состоящих из маленьких букв, для каждого из возможных соответствий

Аа, Аb, Ва, Вb

имеется по меньшей мере один «фактор узнавания», т. е. некоторая специальная последовательность маленьких букв (мы знаем, что такая ситуация действительно имеет место в случае белков). В случайном наборе этому условию могут удовлетворять многие различные последовательности, если только не предъявлять слишком высокие требования — не требовать оптимального функционирования.

Теперь разделим нашу систему на отсеки, которые достаточно велики для того, чтобы они содержали в среднем два «фактора узнавания». Всего может получиться десять различных вариантов:

| | | | | |
|----------|----------|----------|----------|----------|
| 1 | 2 | 3 | 4 | 5 |
| Аа Аа | Аа Аb | Аb Аb | Ва Ва | Ва Вb |
| 6 | 7 | 8 | 9 | 10 |
| Вb Вb | Аа Ва | Аb Вb | Аа Вb | Аb Ва |

Однозначная трансляция могла бы происходить только в двух последних отсеках из десяти. Однако она была бы устойчивой только в том случае, если бы соответствующие факторы узнавания также постоянно воспроизводились. Но для этого нужно, чтобы АВ-последовательности содержали информацию для факторов узнавания. При трансляции АВ-последовательностей

по правилам девятого или десятого отсека снова с равной вероятностью может образоваться любое из десяти возможных сочетаний факторов узнавания. Конечно, при этом имеется в виду только совпадение функций, а не индивидуальных последовательностей — вероятность последнего события много меньше. Из этого можно сделать еще один вывод: если количества последовательностей АВ- и аb-типа одинаковы, то только один из десяти отсеков, содержащих один и тот же набор факторов узнавания, будет содержать также информацию, необходимую для воспроизведения своего типа трансляции.

Итак, самоподдерживающаяся система трансляции будет всего в одном из

$$1/2 \cdot 10 \cdot 10 = 50 \text{ отсеков.}$$

Эти соображения легко обобщить. Если имеется не два, а λ различных классов букв, то теперь из λ^2 возможных сопоставлений между буквами можно образовать $1 \cdot 2 \cdot 3 \dots \lambda = \lambda!$ взаимно однозначных соответствий между двумя алфавитами. Таким образом, множитель $1/2$ из нашего примера в общем случае нужно заменить на $1/(\lambda!)$. Взаимно однозначные соответствия составляют, однако, лишь малую долю всех возможных соответствий, число которых (в нашем примере — десять) равняется «числу всех сочетаний с повторениями из λ^2 элементов по λ ». Некоторые числовые примеры создают впечатление, что эта модель дает ожидаемые вероятности для спонтанного старта трансляции. Для двух букв надо было просмотреть в среднем 50 отсеков заданного размера, чтобы найти одну самоподдерживающуюся устойчивую систему трансляции. Для четырех букв это число увеличивается до 10^6 , для восьми — до $2,5 \cdot 10^{15}$ и, наконец, для двадцати классов букв — до 10^{50} .

Вернемся к действительности. Что нужно понимать под двумя алфавитами, совершенно ясно: А, В, С, ... — это классы кодонов, которые надо сопоставить двадцати белковым единицам — аминокислотам а, b, с, ... Поскольку между кодоном и антикодом существует однозначная связь, этот код будет относиться также к соответствующим адапторам — транспортным РНК.

Факторы узнавания, происходящие из abc-алфавита, — это предшественники аминоацилсинтетаз, т. е.

ферментов, активирующих аминокислоты (а, в, с, ...) и переносящих их на соответствующие транспортные РНК (А, В, С, ...). С какой вероятностью в наборе случайных аминокислотных последовательностей может встретиться однозначная кодирующая активность, следует выяснить экспериментальным путем. Что такая активность возможна, показывает ее существование в природе. Если не предъявлять слишком высоких требований к избирательности узнавания, то можно надеяться, что кодирующая активность будет обнаруживаться достаточно часто. При этом должен возникать целый спектр молекул, различающихся по строению, как, например, в случае антител. Как выглядели первые адапторы — предшественники транспортных РНК, — пока совершенно неясно. Возможно, что это были олигонуклеотиды, например тетра-, пента- или гексануклеотиды, поскольку все различные олигонуклеотиды встречались, по-видимому, достаточно часто и в относительно больших количествах. Но возможно также, что они были составной частью первоначального генетического материала и представляли собой длинные полинуклеотидные последовательности, которые выполняли двойную функцию накопителей информации и адапторов. В настоящее время транспортные РНК на уровне генов представлены длинными последовательностями, от которых лишь после транскрипции «отрезаются» участки, соответствующие функционирующим молекулам адапторов.

Возникновение генетического кода и аппарата трансляции следует считать решающим шагом в молекулярной самоорганизации живых структур. Только благодаря этому стало возможным эволюционное развитие практически неограниченной информационной емкости на чисто функциональной основе. Модельные соображения показывают, что этот шаг предопределяется известными свойствами белков и нуклеиновых кислот, но кодовый ключ не обязательно должен быть предопределен с самого начала. С другой стороны, еще нельзя исключить, что существуют предпочтительные взаимодействия между отдельными аминокислотами и структурами антикодонов, которые склоняют природу в пользу определенного кода. Если такие специфические взаимодействия будут найдены — а на это может дать ответ лишь эксперимент, — то в соответствии с этим

придется модифицировать распределение вероятностей для различных вариантов кода, которое сначала считалось равномерным. Во всяком случае это облегчило бы старт процессов трансляции. Однако для однозначной фиксации кода требуется очень большая точность узнавания. До сих пор не получено никаких экспериментальных данных, которые указывали бы на существование таких достаточно избирательных взаимодействий.

С другой стороны, расчеты показывают, что в отсутствие специфичных взаимодействий должна была начаться трансляция, использующая «классифицирующий» код. Различение четырех классов аминокислот, например неполярных, полярных, положительно и отрицательно заряженных (они обозначены по-разному на рис. 4, 5, 6 и 8), — это минимальное требование для возникновения воспроизводимых исполнительных свойств. Восемь классов аминокислот дают уже почти неограниченные функциональные возможности. Возникновение такого «классифицирующего» кода имеет еще реализуемую вероятность. Однако 10^{50} вариантов уже нельзя разместить в земных масштабах. Это означает, что генетический код в известной нам сейчас дифференцированной форме никоим образом не мог возникнуть спонтанно — он должен был развиваться из «классифицирующего» кода. Таблица кода (см. рис. 8 и 9) на самом деле дает целый ряд оснований для такого предположения.

И если вряд ли когда-нибудь удастся реконструировать в лаборатории историческую случайность, проблеме возникновения кода в принципе все же можно, по-видимому, решить экспериментально.

Третья ступень: от гиперцикла к первичной клетке.

Схоластический вопрос «Что возникло раньше, курица или яйцо?» сводится к абсурду результатами молекулярной биологии. Подобная постановка вопроса, впрочем, и без того теряет смысл, когда события, находящиеся в причинной связи, соединены еще и обратной связью. Не спрашивают, где находится начало или конец окружности.

Взаимосвязь между самовоспроизводящимися «накопителями информации» и «исполнителями» с практически неограниченной шириной функционирования, которая установилась в результате спонтанного воз-

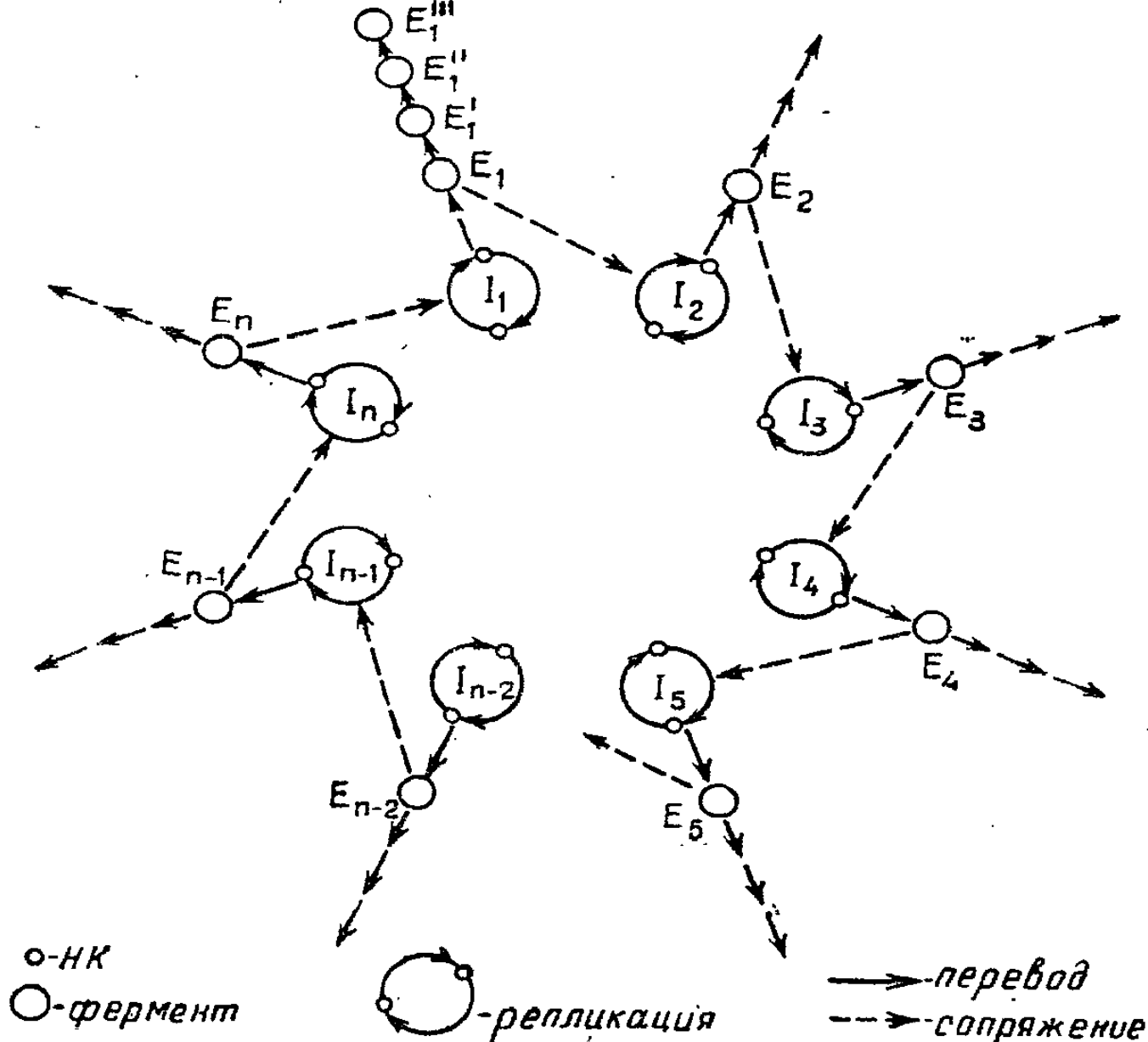


Рис. 18. Гиперцикл. Гиперцикл возникает, когда образуется циклическая цепочка связей между элементарными реакционными циклами (I_i). Элементарные циклы — это цепи нуклеиновой кислоты ограниченной длины (I_{+i}, I_{-i}), воспроизводящие друг друга по комплементарному механизму. В каждом элементарном цикле идет автономная репликация, т. е. плюс-цепь производит минус-цепь, а минус-цепь снова дает плюс-цепь. Эта репликация катализируется факторами связи (E_i) которые либо селективно повышают скорость и точность репликации в своем цикле, либо уменьшают в нем скорость распада. Каждый цикл I_i должен содержать информацию для такого фактора связи E_i , который действует на следующий элементарный цикл I_{i+1} . Кроме того, каждый цикл I_i может содержать информацию для факторов E'_i , выполняющих общие функции: это могут быть, например, факторы трансляции, репликазы или ферменты метаболизма. Только благодаря замыканию цикла (т. е. обратной связи E_n с I_1) элементарные циклы I_i , по отдельности конкурирующие друг с другом, вынуждаются к кооперации. I -циклы, связанные с гиперциклом независимыми факторами связи, а также «шунты» не допускаются. Это означает, что цикл I_i в гиперцикле, независимо от того, сколько белков он кодирует, всегда должен зависеть от общего фактора E_{i-1} и сам должен содержать информацию для другого фактора связи. Каждый I -цикл является поэтому координированно регулируемой генетической единицей, которую можно считать предшественником оперона.

никновения системы трансляции, открывает новые перспективы для процесса макромолекулярной самоорганизации. Ограничения возможностей, характерные для обоих классов молекул по отдельности, снимаются благодаря их кооперации. Правда, еще необходим отбор, т. е. избирательная репликация всей системы трансляции. Для этого все индивидуальные носители информации, входящие в эту систему и сначала конкурирующие друг с другом, должны объединиться в один «коллектив». Связи между отдельными нуклеиновыми кислотами должны образовать замкнутый цикл. Этим будет исключена любая дальнейшая конкуренция между членами цикла. Информация для белковых факторов связи, а также для ферментов трансляции должна содержаться в молекулах нуклеиновых кислот. Лишь таким путем весь коллектив сможет постоянно воспроизводиться как целое. *Гиперцикл*, изображенный на рис. 18, можно рассматривать как прототип такой самовоспроизводящейся системы реакций. Это название подразумевает, что индивидуальные молекулы РНК, циклически воспроизводящиеся на основе комплементарного инструктирования, сопряжены между собой дополнительной цепочкой связей, которая образует цикл.

Такой цикл реакций проявляет свойства, которые в других случаях мы находим только у живых систем.

Любой гиперцикл ведет себя как кооперативная единица с нелинейным законом роста. Если дополнительные условия ограничивают рост, например при постоянной скорости подвода высокоэнергетического строительного материала (пища), или если величина всей популяции поддерживается постоянной при помощи разбавления, то из-за нелинейности получается крайне жесткая конкуренция между различными циклами, и отбор приводит к результату «все или ничего». Это свойство позволяет системе эффективно использовать малые селективные преимущества и эволюционировать быстрее, чем это возможно для линейной системы реакций. Но прежде всего нелинейность приводит к фиксации универсального кода и универсальной хиральности («левого» или «правого» варианта) макромолекулярных структур. Это остается верным и для того случая, когда выбор конкретного варианта обязан лишь случайной флуктуации. Универсальность

следует из способности флуктуации усиливаться (это свойство демонстрировала третья «игра в бисер») — в нелинейном случае это ведет к результату «все или ничего». Кинетические параметры и, следовательно, селективные ценности нелинейной системы зависят от концентраций. Если концентрация из-за эффекта усиления хотя бы однажды вырастет до макроскопического уровня (в общем случае это означает усиление на 10—20 порядков), то никакой другой трансляционный коллектив, состоявший вначале лишь из нескольких копий, уже не сможет вырасти в присутствии господствующей системы. В нелинейной системе сосуществование различных вариантов кода исключается.

Из-за связей между индивидуальными носителями информации информационная емкость может достигнуть любого размера. Однако цикл закорачивает ненужные «петли», так что количество информации поддерживается на нужном уровне. При эволюции цикла происходит постоянное увеличение селективных ценностей всех последовательностей, связанных циклической цепочкой связей. Паразитные ветви, т. е. последовательности, которые не являются составной частью цикла, а только соединяются с ним посредством фактора связи, не допускаются гиперциклом. Напротив, такие ветви, которые являются частью какой-либо последовательности I_i , получающей пользу от цикла, и которым благоприятствует фактор связи E_{i-1} , принадлежащий циклу, будут сохраняться (см. рис. 18). Тогда только часть информации I_i будет использоваться для кодирования фактора связи E_i . Остаток можно употребить для кодирования общих функций: трансляции, полимеризации, метаболизма и т. д. Здесь уже намечается генная и оперонная структуры, которые мы находим в настоящее время в геноме бактериальной клетки (рис. 19).

Чтобы такие функции можно было селективно использовать и оптимизировать, члены гиперцикла должны образовывать комплекс или же собираться внутри какого-нибудь «отсека». Такая компартментализация является также защитой от «прореживания». Возможно, что уже само образование цикла было тесно связано со способностью к комплексообразованию или к образованию «отсеков». Однако длительное сохранение компартментализации возможно только в том слу-

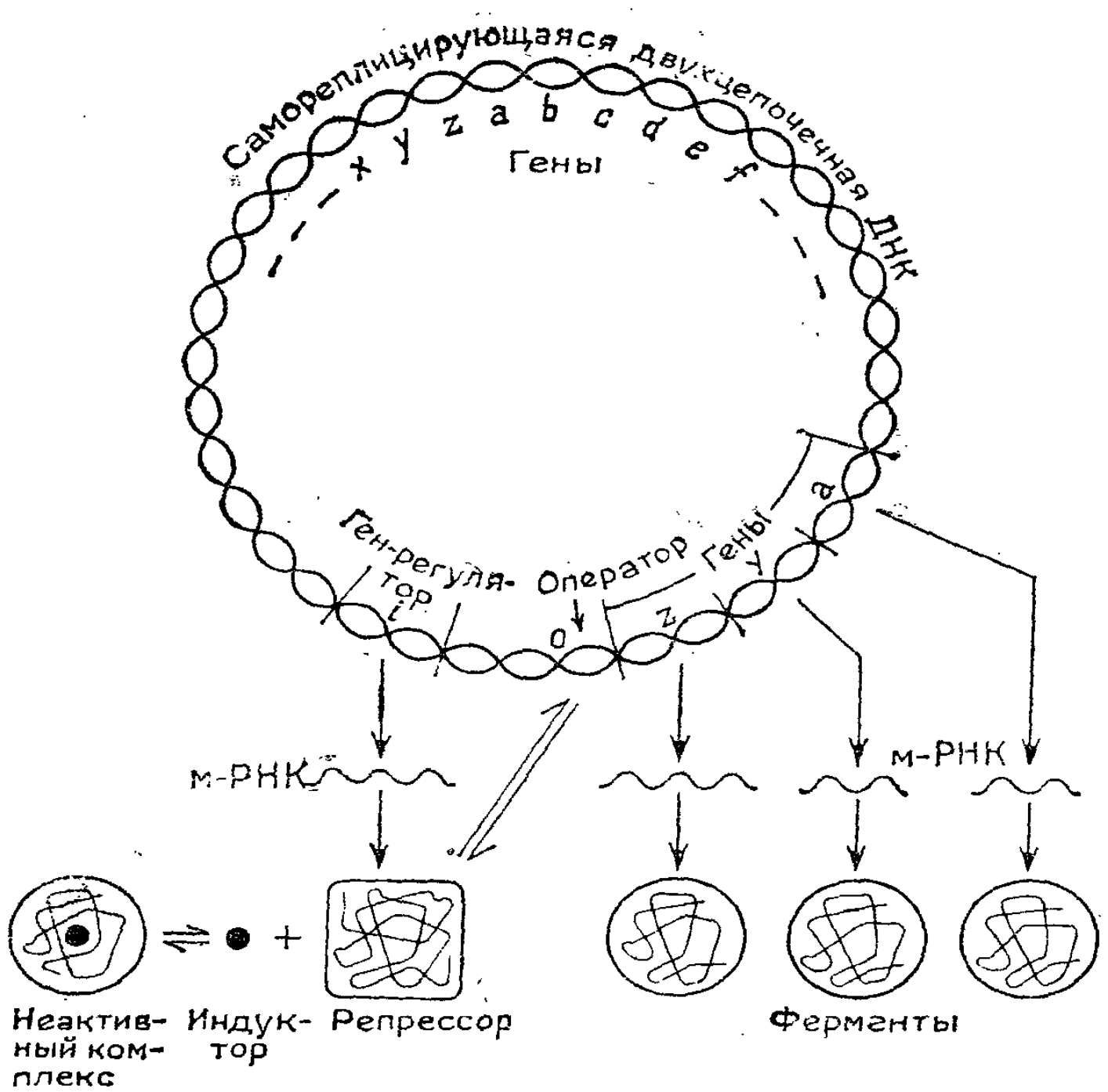


Рис. 19. Кольцевая структура хромосомы бактериальной клетки. Кишечные палочки и ряд фагов имеют кольцевые генетические карты. Рисунок схематически показывает линейное расположение генов а...z в кольцевой двухцепочечной молекуле ДНК. В опероне несколько генов объединяются с образованием координированно регулируемой единицы. Особенно хорошо исследован Лас-оперон (сначала Жакобом и Моно, а затем Гилбертом и Мюллер-Хиллом), который схематически изображен в увеличенном виде в нижней части рисунка. Он состоит из оператора (специфичной области распознавания на молекуле ДНК) и генов, кодирующих ферменты: β-галактозидазу, пермеазу и ацетилазу. При взаимодействии с репрессором — белковым комплексом, который кодируется пространственно удаленным регуляторным геном, чтение этих генов специфично блокируется. В отсутствие индуктора (лактоза или продукт катаболизма) репрессор прочно связан с оператором. Трансляция генетического сообщения всегда идет через промежуточную стадию м-РНК, которая переносит информацию.

чае, если она сочетается с индивидуализацией. Последняя может произойти под действием фермента «лигазы», который соединяет друг с другом свободные концы молекул РНК. Лигазы в настоящее время хорошо известны биохимикам. Таким образом, должны были возникнуть кольцевые генетические карты, которые мы теперь находим у кишечной палочки и у ряда вприсов. Некоторые особенности взаимного расположения отдельных генов также можно предсказать. После того как совершилась индивидуализация, замена РНК на более стабильную двухцепочечную ДНК будет давать селективное преимущество. Индивидуализированные виды размножаются, в сущности, по линейному механизму, который дает большие возможности для сосуществования, т. е. для параллельного развития многих индивидуумов.

Появление гиперцикла, несомненно, означало преодоление порога между неживым и живым. Однако должно быть ясным, что мы рассматриваем всего лишь модель. Правильность модели и ее следствий может быть проверена экспериментально. Правда, это вовсе не доказывает, что природа воспользовалась именно той возможностью, которая описывается моделью. Да мы и не пытаемся реконструировать историческую случайность. Мы хотим только показать, что за самоорганизацией живых структур стоит физический принцип, а именно: при выполнении определенных физических граничных условий возникновение и эволюция живых структур — несмотря на неопределенность индивидуального пути — представляют собой в принципе неизбежный процесс, причем образовавшиеся структуры всегда могут быть вытеснены еще лучше функционирующей системой. Условия, необходимые для старта, существовали в геологической истории нашей планеты, а, может быть, и во многих других местах Вселенной. Химия — наука, основанная на законах физики, дает все предпосылки для возможности образования структур, которые при наличии круговорота вещества и энергии будут постоянно воспроизводиться в силу присущих им взаимодействий и медленно изменяться вследствие статистических флуктуаций. Неограниченная самоорганизация живых структур зависит — как показывает пример гиперцикла — от наличия целого ряда специальных химических условий,

не говоря уже о необходимости метаболизма, самовоспроизведения и мутаций. Эволюция является «неограниченной» только тогда, когда рассматриваемая система может, благодаря своим внутренним качествам, самостоятельно достигнуть вышестоящего уровня организации.

Неотъемлемым свойством прогрессивной эволюции является ее направленность во времени. Это качество тесно связано с временным направлением роста энтропии при необратимых процессах. Возрастание энтропии нельзя объяснить, исходя лишь из законов механики дискретных материальных частиц, оно обосновывается статистической физикой. Это преимущественное направление времени особенно ярко проявляется в неустойчивостях, которые ведут к последовательным сужениям распределения вероятностей. Критерии устойчивости, сформулированные Гленсдорфом и Пригожиным [19] на основе термодинамической теории необратимых процессов, дают обоснование физическим закономерностям эволюции, которые наглядно обнаруживались на примерах наших селекционных и эволюционных игр, а также в реалистической модели гиперцикла. Эта закономерность ничего не меняет в индивидуальной *неопределенности* каждого исторического пути развития, которая обусловлена колоссальной и постоянно возрастающей в ходе эволюции сложностью системы. Исторический процесс включает в себя «возникновение», а не просто «выявление» информации. В материальных деталях он лишен всякой диалектической необходимости.

(N + 1)-я ступень: идея идей.

В сборнике «Симплициссимуса» за 1925 год мы находим забавную фантастическую историю про «биссов» — сатиру на развитие человечества в третьем тысячелетии.

Автор, Арнольд Хан, заставляет своего героя сделать потрясающее открытие:

«Наступил знаменательный 2703 год. В этом году Христофор Карцелес выступил перед человечеством со своей книгой «Метод методов». И тут начался скачок.

Что же такое было в «Методе методов»? Вы, методисты, говорит Карцелес, делите мир на проблемы и ищите к каждой проблеме свой метод — ключ. Я же делаю сам метод проблемой и ищу к нему ключ. Если

я найду этот ключ, то он откроет мне ларец, где лежат все ключи, которые когда-либо могут быть использованы для разгадывания и покорения мира... Я нашел его!»

Математики наших дней уже заняты поиском этого ключа. Идея машины Тьюринга — это не что иное, как мечта о «методе методов».

Что же касается «природы», то она давно нашла этот ключ. Метод самоорганизации — *отбор* — на высшей стадии своего развития, человеку, становится методом методов, идеей идей.

Индивидуализированная клетка представляет собой систему, на новом уровне селективно развивающую до совершенства все свойства, которые она в себе объединяет, и оптимально приспособляющуюся к своей среде. С ростом точности переноса информации растет и информационная емкость. Но чем больше эта емкость, тем многообразнее становятся возможности специализации. Из вида «праклетка» возникло множество различных видов клеток, конкурирующих или кооперирующихся друг с другом. Взаимодействие должно было стать организованным: клетки развили рецепторы, с помощью которых они получили возможность специфично — в расчете на определенную кооперацию — узнавать друг друга. Но выгоды, происходящие из кооперации, могут использоваться лишь в том случае, если новый план организации будет запечатлен также и в наследственности. Тогда все соматические типы клеток получают возможность всякий раз развиваться по программе, заложенной в одной-единственной зародышевой клетке.

Математические модели [20] клеточной дифференцировки и морфогенеза стали появляться только в последние годы и были испытаны на простых объектах: амебах [21] и гидридных организмах [22]. Они основываются на автокаталитических молекулярных механизмах усиления, очень похожих на уже известные нам механизмы эволюционного поведения на субклеточном уровне. Идея отбора полностью мобилизуется и осуществляется на уровне многоклеточных организмов. Результатом является кажущаяся, — впрочем, нет, реальная — «инструкция». Она основана на отборе индивидуумов, но прирост информации происходит только на уровне вида. Индивидуум выживает или

гибнет, сам он не «учится». Опыт идет только на пользу вида. Эволюция всей биосферы представляет собой величественный процесс накопления информации и образования памяти.

А нельзя ли было эту способность к обучению сделать полезной уже для индивидуального объединения клеток в течение его жизни? Для этого требовался новый уровень, на котором — как в генах — можно было бы селективно, под контролем функции ценности, накапливать опыт и производить его оценку. Этот «метод методов» должен был давать селективное преимущество везде, где он только возникал. Соответствующий эволюционный прорыв был, таким образом, неизбежен. Он начинается на уровне генов. Отбираются системы, у которых развились детекторы, с помощью которых можно реагировать на приходящие из окружающей среды раздражения, таким образом, чтобы повысить свои шансы на выживание. Примером может служить поведение грибка — фикомицета, который отвечает на световое раздражение ростом по направлению к свету (см. статью Макса Дельбрюка и сотрудников в последнем выпуске настоящего сборника).

В ходе эволюции возникают сети нервных клеток, которые *центрально* перерабатывают приходящие раздражения и превращают их в «успешные действия». С ростом сложности возникла необходимость — как и в случае молекулярного накопителя информации — уметь селективно отвечать на раздражения, т. е. делать выбор между альтернативами. Живое существо получило возможность независимо от генетической программы постоянно приспосабливаться к *непредвиденным* ситуациям. Только программируемость является здесь частью генетической программы. Для этого понадобилась также запрограммированная способность производить оценку: боль, удовольствие, страх, радость, вплоть до самых тонких эмоций. С помощью этой «вторичной» системы оценок избирательно накапливаются раздражения, идущие из окружающей среды. Их можно, наконец, воспроизводить даже абстрактно и проигрывать в новых комбинациях. Как и в случае мутаций, новые комбинации, возникшие благодаря флуктуациям, подвергаются оценке и отбору.

Уже появились первые модельные наброски теории самоорганизующихся сетей, применимой к централь-

ной нервной системе [23—25]. Они исходят из того самоорганизующегося поведения, которое демонстрировали нам селекционные и эволюционные игры. Информация возникает в результате изменения распределения вероятностей возбуждения и торможения нервных клеток под действием раздражений.

Направленность во времени, характерная для всех эволюционных процессов самоорганизации, является и здесь основой нашего внутреннего чувства времени. Недетерминированность и практическая неограниченность наших мыслей и идей являются следствием сложности схемы: она объединяет около 10 миллиардов нервных клеток, каждая из которых имеет около 10—100 тысяч контактов (синапсов) с другими клетками.

«Свободная воля» не была бы «свободной», если бы врожденный детерминизм не допускал никаких альтернатив, но, с другой стороны, она не была бы «волей», если бы выбор альтернатив предоставлялся лишь случаю, т. е. статистическим флуктуациям нервных процессов. Индивидуальная свобода — это работа фильтра, локализующегося внутри личности. Этот фильтр складывается из последственности, опыта и перманентной для системы способности производить оценки. На него подаются статистические флуктуации, управляемые раздражителями, а выходит из него «информация» в форме наших мыслей и идей *).

Мы еще очень далеки от полного понимания этого уровня игры игр — от понимания нас самих.

«Итак, напрасно было бы требовать от нас изложения всей истории и теории Игры в бисер: подобная задача сейчас не по плечу и куда более искушенным и достойным авторам. Решение ее — удел будущего, если, разумеется, к тому времени сохранятся как источники, так и духовные предпосылки. Еще того менее наш труд может служить учебником Игры — таковой вообще никогда не будет написан. Правила этой Игры игр усваиваются только обычным, предписанным путем, на что уходят годы, и никому из посвященных не придет на ум облегчать процесс их усвоения». (Герман Гессе: из введения к «Игре в бисер».)

*) Конечно, проблема свободы воли требует гораздо более глубокого анализа, как естественнонаучного, так и философского. (Прим. ред.)

ЛИТЕРАТУРА

К предисловию

1. *Эйген М.* Самоорганизация материи и эволюция биологических макромолекул.— М.: Мир, 1973.
2. *Эйген М.*— УФН, 1973, т. 109, с. 545.
3. *Eigen M., Winkler R.*— München, Zürich: Das Spiel, R. Piper and Co Verlag, 1975.
4. *Берг Л. С.* Труды по теории эволюции.— Л.: Наука, 1977, с. 60.
5. *Кастлер Г.* Возникновение биологической организации.— М.: Мир, 1967.
6. *Блюменфельд Л. А.* Проблемы биологической физики.— М.: Наука, 1977.
7. *Шмальгаузен И. И.* Кибернетические вопросы биологии.— Новосибирск: Наука, 1968.
8. *Волькенштейн М. В.*— Ж. общей биологии, 1976, т. 37, с. 483.
9. *Волькенштейн М. В., Чернавский Д. С.*— J. Soc. Biol. Structures, 1978, v. 1, p. 95; Изв. АН СССР сер. бпол., 1979, № 4.
10. *Волькенштейн М. В.* Общая биофизика.— М.: Наука, 1978.
11. *Eigen M., Schuster P.*— Naturwissenschaften, 1977, Bd. 64, № 11, S. 541.
12. *Eigen M., Schuster P.*— Naturwissenschaften, 1978, Bd. 65, № 1, S. 7.
13. *Eigen M., Schuster P.*— Naturwissenschaften, 1978, Bd. 65, № 7, S. 341.
14. *Волькенштейн М. В.* Молекулы и жизнь.— М.: Наука, 1965.
15. *Волькенштейн М. В.* Молекулярная биофизика.— М.: Наука, 1975.

К главам 1—5

1. *Шредингер Э.* Что такое жизнь?— М.: Атомиздат, 1972.
2. *Опарин А. И.* Происхождение жизни.— М.: Московский рабочий, 1924; Происхождение и начальное развитие жизни.— М.: Медицина, 1966.
3. *Эйген М.* Самоорганизация материи и эволюция биологических макромолекул.— М.: Мир, 1973.
4. *Уотсон Дж.* Молекулярная биология гена.— М.: Мир, 1979.
5. *Watson J. D., Crick F. H. C.*— Nature, 1953, v. 177, p. 964.

6. *Spiegelman S.* The Neurosciences. 2nd Study Program/Ed. F. O. Schmitt.— New York: The Rockefeller University Press, 1970.
7. *Шеннон К. Э.*— В сб.: Работы по теории информации и кибернетике.— М.: ИЛ, с. 243.
8. *Винер Н.* Кибернетика.— М.: Советское радио, 1968.
9. *Kolmogoroff A. N.*— C. R. Acad. Sci., 1939, v. 208, p. 2043.
10. *Von Weizsäcker C. F.* Die Einheit der Natur.— München: Carl Hanser Verlag, 1971.
11. *Eigen M., Pörschke D.* J. Mol. Biol. 1970, v. 53, p. 123.
12. *Gatlin L.* Information Theory and the Living System.— New York: Columbia University Press, 1972.
13. *Rényi A.* Wahrscheinlichkeitsrechnung.— Berlin, 1966.
14. *Бриллюэн Л.* Наука и теория информации.— М.: Физматгиз, 1960.
15. *Гарднер М.* Математические досуги.— М.: Мир, 1972, с. 458.
16. *Фон Нейман Дж., Моргенштерн О.* Теория игр и экономическое поведение.— М.: Наука, 1970.
17. *Monod J.* Zufall und Notwendigkeit.— München: R. Piper and Co Verlag, 1971.
18. *Jacob F.* Die Logik des Lebendigen.— Frankfurt: S. Fischer Verlag, 1972.
19. *Гленсдорф П., Пригожин И.* Термодинамическая теория структуры, устойчивости и флуктуаций.— М.: Мир, 1973.
20. *Thom R.* Stabilité Structurelle et Morphogénèse.— New York: W. A. Benjamin, Inc., 1972.
21. *Robertson A., Cohen M. H.*— Ann. Rev. Biophys. Bioeng., 1972, v. 1, p. 409.
22. *Gierer A., Meinhardt H.*— Kybernetik, 1972, Bd. 12, S. 30.
23. *Marr D.*— Proc. Roy. Soc., 1970.
24. *Wilson H. R., Cowan J.*— Biophys. J., 1972, v. 12, p. 1.
25. *Malsburg Ch.* Kybernetik, 1973.

M. Eigen, P. Schuster

THE HYPERCYCLE

A Principle of Natural Self-Organization

Springer-Verlag

Berlin Heidelberg New York 1979

М.Эйген, П.Шустер

ГИПЕРЦИКЛ

**принципы самоорганизации
макромолекул**

Перевод с английского

д-ра биол. наук В. М. АНДРЕЕВА

под редакцией

чл.-корр. АН СССР М. В. ВОЛЬКЕНШТЕЙНА

и проф. Д. С. ЧЕРНАВСКОГО

Москва «Мир» 1982

М. Эйген, известный ученый из ФРГ, лауреат Нобелевской премии, уже знаком советскому читателю по книге «Самоорганизация материи и эволюция биологических макромолекул» (М.: Мир, 1973). В настоящей книге, написанной совместно с австрийским ученым П. Шустером, М. Эйген развивает теорию гиперцикла как одного из принципов самоорганизации макромолекул. Последовательно рассмотрены отбор и эволюция РНК и ДНК, приведен математический анализ динамических систем применительно к проблеме возникновения жизни и обсуждена модель гиперцикла.

Предназначена для научных работников самых разных специальностей: биологов, химиков, физиков, математиков, философов, историков науки.

Редакция литературы по биологии

© by Springer-Verlag Berlin Heidelberg 1979.

Authorized translation from English language edition published by

Springer-Verlag Berlin—Heidelberg — New York.

2001000000

Э $\frac{21001-105}{041(01)-82}$ 105 — 82, ч. 1 © Перевод на русский язык, «Мир», 1982

ПРЕДИСЛОВИЕ РЕДАКТОРОВ ПЕРЕВОДА

«Гиперцикл» — третья книга М. Эйгена, переведенная на русский язык. В книге «Самоорганизация материи и эволюция биологических макромолекул» (М.: Мир, 1973) Эйген впервые сформулировал концепцию образования упорядоченных макромолекул из неупорядоченного вещества на основе матричной репродукции и последующего отбора. Таким образом, принцип Дарвина был применен к проблемам добиологической эволюции и происхождения жизни. Эта книга вызвала большой интерес и несомненно стимулировала дальнейшие работы в области физико-математического моделирования процессов добиологической и биологической эволюции. В небольшой книге «Игра жизни», написанной Эйгеном вместе с его сотрудницей Рутхильд Винклер (М.: Наука, 1979), те же вопросы рассматриваются популярно — с помощью остроумных игровых моделей.

В предлагаемой вниманию читателей монографии Эйген и Шустер впервые применили аппарат качественной теории динамических систем для анализа ранее развитой концепции. Нужно сказать, что метод «фазовых портретов» очень красив, и это ощущает каждый, кто с ним знакомится. Чувствуется, что и авторы испытали обаяние качественной теории. Справедливости ради следует отметить, что к проблемам эволюции и даже к вопросу происхождения жизни этот аппарат применялся и ранее. К сожалению, Эйген и Шустер не сопоставляют свои исследования с работами других ученых, в результате книга не полностью отражает современное состояние проблемы.

Математическая модель, разработанная авторами данной монографии, имеет главным образом иллюстративное значение. В действительности этим роль

математического подхода к проблеме возникновения жизни отнюдь не исчерпывается. Качественная теория динамических систем открывает возможности четкой постановки новых принципиально важных проблем и их решения.

Во-первых, основное понятие «отбор» при скрупулезном анализе оказывается неоднозначным. Обычно этот термин употребляется в смысле отбора наилучшего варианта, т. е. как достижение абсолютного оптимума; в этом случае результат отбора предопределен заранее. Однако в ходе эволюции и даже в самом начале ее (т. е. в процессе возникновения жизни) встречаются ситуации, при которых развитие может пойти по нескольким различным путям. Здесь следует говорить не об отборе, а о выборе одного из возможных вариантов; при этом достигается не абсолютный оптимум, а один из приемлемых. Результат процесса заранее не предопределен, он зависит от случайного выбора.

Метод фазового пространства позволяет наиболее четко поставить вопрос «Отбор или выбор» и даже ответить на него в рамках той или иной модели. В случае, если имеется несколько стационарных состояний, отделенных друг от друга сепаратрисами, мы имеем дело с выбором. Если же состояние только одно, то в этом (и только в этом) случае можно говорить об отборе. Вопрос о том, что же имело место в процессе происхождения жизни и дальнейшей эволюции — отбор или выбор — является одним из принципиальных. Ответ на него можно получить путем строгого математического анализа существующих моделей, выводов из них и сопоставления результатов с экспериментом.

Во-вторых, на основе динамической теории может быть четко поставлена и решена проблема возникновения биологической информации, увеличения ее количества и ценности. Отметим, что эта проблема связана с предыдущей, а именно: новая информация возникает в системе, только если происходит случайный выбор (Кастлер), а не отбор наилучшего варианта. В последнем случае можно говорить лишь

о реализации информации, заложенной в систему а priori, т. е. о выделении имеющейся информации из шума.

Из всего сказанного следует, что формулировка проблемы происхождения жизни на основе теории динамических систем открывает двери в новую и весьма интересную область. Авторы предлагаемой книги сделали в этом направлении первые шаги, как всегда в таком случае особенно трудные и важные.

*М. В. Волькенштейн
Д. С. Чернавский*

ПРЕДИСЛОВИЕ

Основой для этой книги послужила серия статей, опубликованных в журнале „Die Naturwissenschaften“ в 1977—78 гг. Разделение книги на три части отражает ее логическую структуру, которую можно резюмировать в виде трех тезисов.

А. Гиперцикл — это принцип естественной самоорганизации, обуславливающий интеграцию и согласованную эволюцию системы функционально связанных самореплицирующихся единиц.

Б. Гиперциклы — это новый класс нелинейных сетей реакций, обладающих уникальными свойствами и поддающихся унифицированному математическому описанию.

В. Гиперциклы могут возникать в распределении мутантов отдельного дарвиновского квазивида в результате стабилизации его дивергирующих мутантных генов. Зародившись, гиперциклы эволюционируют в сторону усложнения посредством процесса, аналогичного дупликации генов и специализации.

Чтобы пояснить смысл первого утверждения, можно обратиться к другому принципу самоорганизации материи — к дарвиновскому принципу естественного отбора. Этот принцип — как мы представляем его в настоящее время — является единственным понятным нам способом создания информации, будь то программа сложного живого организма, который представляет собой продукт эволюции менее сложных прародительских форм, или осмысленная последовательность букв, отбор которой можно имитировать в модельных эволюционных играх.

Естественный отбор — здесь акцентируется слово «естественный» — основан на самовоспроизведении.

Иными словами: если дана система самовоспроизводящихся единиц, которые строятся из материала, поступающего в ограниченном количестве из общего источника, то неизбежно возникает естественный отбор. Соответственно эволюционное поведение, управляемое естественным отбором, основано на самовоспроизведении с «шумом». Наличия этих физических свойств достаточно, чтобы стало возможным закономерное возникновение в высшей степени сложных систем, т. е. генерирование такой информации, как программа живого организма. Однако дарвиновскому механизму естественного отбора присущи количественные ограничения в объеме накапливаемой информации. Здесь-то и выходит на сцену гиперцикл. Гиперцикл — это тоже принцип самоорганизации, но он основан на других предпосылках и поэтому ведет к иным последствиям.

Теория дарвиновских систем, изложенная в части А, приводит к двум основным результатам.

а. Самовоспроизводящиеся единицы конкурируют в процессе отбора. Эта конкуренция может ослабляться для неродственных видов, уходящих в разные ниши. Тем не менее в пределах каждого распределения мутантов она должна быть достаточно эффективной, чтобы поддерживалась устойчивость дикого типа. Без такой конкурентной стабилизации его информация рассеялась бы.

б. Информационное содержание устойчивого дикого типа ограничено. Другими словами, количество информации не должно превышать некоего порога, величина которого обратно пропорциональна среднему темпу генерации ошибок (на символ). Далее, этот пороговый уровень зависит от логарифма превосходства дикого типа, т. е. от среднего селективного преимущества по отношению к мутантам суммарного (устойчивого) распределения. Распределение становится неустойчивым всякий раз, когда появляется нарушающий это условие мутант, имеющий преимущество над ранее устойчивым диким типом.

Эти свойства внутренне присущи дарвиновским системам. Они гарантируют эволюционное поведение, для которого характерны отбор и *стабильное* воспроизведение наиболее адаптированной самовоспроизводящейся единицы, а также замена ее любым мутантом, который адаптирован еще лучше. С другой стороны, эволюция такой системы ограничена определенным уровнем сложности — порогом максимального информационного содержания.

Из-за этого ограничения первыми самореплицирующимися единицами были, вероятно, относительно короткие цепи нуклеиновых кислот. Это единственный класс макромолекул, обладающих способностью к самовоспроизведению. Однако физические силы, определяющие точность саморепликации, имеют ограниченную специфичность. Увеличение точности могло произойти только благодаря наличию катализатора, причем катализатор — чтобы он мог подвергаться эволюционной адаптации — тоже должен был воспроизводиться. На этом этапе эволюции появилась необходимость в трансляции информации, унаследованной воспроизводящимся материалом.

Барьер был колоссальной высоты. Эволюция, по видимому, фактически остановилась. Требовалась машина, но чтобы создать ее, нужна была все та же машина. Даже примитивный аппарат трансляции должен был содержать, вероятно, не менее четырех адапторов для четырех разных аминокислот плюс соответствующее число ферментов и их мРНК. Количество информации, необходимое для такой системы, сравнимо с количеством информации в вирусе, содержащем одноцепочечную РНК. Однако такие вирусы могут использовать совершенный аппарат трансляции клетки-хозяина. Кроме того, они воспроизводятся с помощью высокоадаптированного ферментативного аппарата — конечного, т. е. оптимального продукта эволюции.

Геном РНК-содержащего фага не превосходит нескольких тысяч нуклеотидов, что как раз достаточно для кодирования небольшого числа (например, четырех) молекул белка. Как показано в части А,

этот предел обусловлен точностью копирования, которая может быть достигнута с помощью хорошо адаптированного фермента репликации. Любое дальнейшее увеличение информационного содержания потребовало бы таких совершенных механизмов, как коррекция с участием экзонуклеазы и лигазы, что стало доступно лишь ДНК-полимеразам на очень поздних стадиях эволюции. Как могла возникнуть даже примитивная система трансляции, если точность воспроизведения определялась исключительно физическими свойствами нуклеиновых кислот, не позволяющими воспроизводимо реплицировать отдельные нуклеотидные цепи длиной более 50—100 нуклеотидов? Количество информации, необходимое для системы трансляции, без которой не могло произойти увеличения точности, во много раз превышает то количество, которое может содержаться в отдельных самовоспроизводящихся цепях.

Гиперцикл — это средство объединения самовоспроизводящихся единиц ограниченной длины в новую устойчивую систему, способную к согласованной эволюции. Никакой другой способ организации — возникновение компартментов или образование нециклических сетей — не может гарантировать одновременного выполнения трех следующих условий:

- поддерживать конкуренцию в распределениях дикого типа всех самореплицирующихся единиц, чтобы сохранить их информацию;
- допускать сосуществование нескольких (в других условиях конкурирующих) единиц и их мутантных распределений и
- объединять эти единицы в систему, способную к согласованной эволюции, где преимущества одного индивида могут использоваться всеми членами системы, причем система как целое продолжает интенсивно конкурировать с любой единицей иного состава.

Наше утверждение логически следует из результатов части А:

Если искать физический механизм, гарантирующий непрерывную эволюцию аппарата трансляции,

то наличие гиперциклической организации является минимальным требованием. Необходимо, чтобы вовлеченные в гиперцикл носители информации были способны к саморепликации — однако одного этого недостаточно. При анализе условий гиперциклической организации сразу же видна эквивалентность этих условий предпосылкам дарвиновского отбора. Последний основан на самовоспроизведении, которое является одним из видов линейного автокатализа. Гиперцикл — это следующий, более высокий уровень в иерархии автокаталитических систем (как показано в части А). Он построен из автокатализаторов, или циклов воспроизведения, которые сочленены посредством циклического катализа, т. е. посредством еще одного автокатализа, наложенного на систему. Таким образом, гиперцикл основан на нелинейном автокатализе (например, второго или более высокого порядка).

Поскольку гиперциклы ведут себя «регулярно», их можно исследовать как конкретный класс сетей реакций. Такой общий анализ проводится в части Б (см. второе утверждение). Они проявляют уникальные физические свойства, которых лишены другие типы связей, и поэтому требуется унифицированное рассмотрение «абстрактного гиперцикла». Этот способ представления предмета сам по себе оправдывает учебную форму изложения.

С другой стороны, гиперцикл никоим образом не является лишь абстрактным продуктом нашего ума. Этот принцип еще сохраняется в цикле размножения РНК-содержащего фага, хотя здесь он прилагается только к замкнутому миру клетки-хозяина. Фаговый геном после трансляции производит фактор, становящийся субъединицей в репликационном комплексе, другие части которого поставляет клетка-хозяин. Этот фактор, кодируемый фагом, придает ферменту абсолютную фаговую специфичность. Поскольку фагоспецифичный репликационный комплекс отвергает все РНК клетки-хозяина, его наличие равносильно наложению петли обратной связи, автокаталитически усиливающей фаговый геном.

Наше утверждение относительно необходимости гиперциклической организации примитивного аппарата трансляции имеет природу «если — то» и еще не означает, что исторически именно так все и было. На самом деле могли происходить неожиданные экстраординарные события — флуктуации, которые не отражают никакой регулярности природы, но влияют на ход событий. Если мы собираемся показать, что эволюция протекала, подчиняясь определенному физическому принципу, то нам нужно искать этому исторические свидетельства, т. е. следы ранних организационных форм в современных организмах. Это сделано в части В, и наше третье утверждение относится к ней.

Некоторую информацию о своем происхождении дают нам транспортные РНК — ключевые вещества трансляции. По-видимому, они указывают на естественный путь разрешения проблемы инициации нелинейной сети реакций — проблемы нуклеации. Все члены этой сети являются потомками одной и той же родоначальной копии, тРНК-предшественника. Мутанты квазивидового распределения этого предшественника могли накопиться прежде, чем начал работать организационный принцип гиперцикла. Будучи близкородственными мутантами, все адапторы и молекулы мРНК, а также их продукты трансляции функционируют весьма сходным образом (как мишени и как исполнители). Поэтому они автоматически «вовлекаются» в организацию с многими перекрестными связями, включающую некий цикл. Как показано в части В, этот цикл может постепенно стабилизироваться путем эволюционного развития специфичности связей, причем все они могут быть типа «репликаза — ее мишень» (такой тип связи все еще используется РНК-содержащими фагами). Реалистичский гиперцикл можно исследовать экспериментально, например, при детальном изучении современного механизма трансляции.

Мы надеемся, что эта книга будет способствовать постановке адекватных вопросов, направленных на изучение проблем эволюции. Теория лишена

абсолютной ценности, если ее выводы нельзя проверить экспериментально. С другой стороны, теория должна давать нечто большее, чем простое объяснение экспериментальных фактов. Как сказал Эйнштейн: «Только теория может указать нам, какие эксперименты осмысленны».

В этом отношении наша книга написана не только для физика, который стремится к единообразному применению физических законов ко всей природе. Она адресована также химику, биохимику и биологу, чтобы подвигнуть их на постановку новых экспериментов, которые способствовали бы более глубокому пониманию жизни как проявления «регулярности природы» и ее происхождения.

Нашу работу в значительной степени стимулировали дискуссии с Фрэнсисом Криком, Стэнли Миллером и Лесли Оргелом; эти дискуссии были равносильны для нас «давлению отбора», заставлявшему искать бóльшую непрерывность в молекулярной эволюции. Особенно полезными были советы и замечания Кристофа Бибрихера, Ирвинга Эпштейна, Бернда Гутте, Дитмара Пёршке, Карла Зигмунда, Пола Вули и Роберта Вольфа. Рутхильд Винклер-Осватич предложила большинство иллюстраций и была всегда терпеливой и критичной собеседницей.

Спасибо им всем за помощь.

Геттинген,
6 ноября 1978 г.,

Манфред Эйген
Петер Шустер

1. Парадигма единства и разнообразия в эволюции

Почему существуют миллионы видов животных и растений несмотря на то, что молекулярный аппарат клетки в своей основе одинаков: один универсальный генетический код и макромолекулы определенной хиральности?

Современные генетики без колебаний немедленно ответят на первую часть этого вопроса. Разнообразие видов является результатом грандиозного ветвящегося процесса эволюции с мириадами отдельных шагов воспроизведения и мутаций. В процессе эволюции происходит отбор среди конкурентов, потребляющих общие пищевые ресурсы, хотя допускается также изоляция, смена ниши или даже взаимная толерантность и симбиоз — при наличии достаточно мягких селекционных ограничений. Дарвиновский принцип естественного отбора представляет собой принцип управления, дающий дифференцированную оценку генной популяции в отношении ее оптимальной приспособленности к среде. Строго говоря, он эффективен только при наличии подходящих граничных условий, которые в природе могут выполняться или не выполняться. В работах Фишера, Холдейна и Райта — основоположников главных направлений популяционной генетики — принципу естественного отбора была дана точная формулировка, показывающая его возможности и ограничения. Этот принцип опирается на общие свойства живых организмов, особенно на механизмы их воспроизведения. Последние зависят от ряда факторов, определяющих как генетическую однородность, так и неоднородность, которые были известны еще до установления детальных молекулярных механизмов наследственности (табл. 1).

Сознавая эту гетерогенность живого мира, действительно трудно понять причину его однородности на субклеточном уровне. Многие биологи просто объединяют все доклеточные эволюционные события и собирательно называют их «возникновением жизни». Конечно, если бы жизнь возникла в результате однократного колоссального акта творения и если бы

Таблица 1

Факторы естественного отбора (по Райту [1])

| Факторы генетической однородности | Факторы генетической неоднородности |
|---|---|
| Дупликация генов Агрегация генов Митоз Конъюгация Сцепление Ограничение размеров популяции Давление среды Кроссбридинг между подгруппами | Мутации генов Случайное деление агрегата Хромосомные aberrации Редукция (мейоз) Кроссинговер Гибридизация Индивидуальная адаптируемость Подразделение группы Локальная среда подгрупп |

этот акт — как уникальное и экстраординарное событие, выходящее за пределы всего статистически возможного с точки зрения физики, — в прошлом имел место только один раз, то мы могли бы удовлетвориться подобным объяснением. Но любая дальнейшая попытка понять «как?» была бы тщетной. Случайность нельзя свести ни к чему, кроме случайности.

Однако то, что нам известно о молекулярной тонкой структуре даже простейших существующих ныне клеток, никак не свидетельствует в пользу такого объяснения. Закономерности строения этой очень сложной структуры не оставляют сомнения в том, что первая живая клетка скорее всего сама была продуктом длительного процесса эволюции, который должен был состоять из многих отдельных — но не обязательно экстраординарных — этапов. В частно-

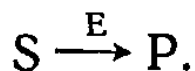
сти, генетический код можно представить как результат такого многошагового процесса эволюции [2], который, по-видимому, начался с однозначного кодирования лишь немногих из наиболее распространенных в первичных условиях аминокислот [3]. Хотя в коде не удается усмотреть совершенной логической структуры в отношении всех окончательных соответствий, он далек от случайного, и невольно приходит в голову мысль, что здесь «сработал» какой-то оптимизационный принцип. Его можно назвать принципом наименьшего изменения, потому что структура кода такова, что на уровне аминокислот последствия точечных мутаций сводятся к минимальным изменениям. Избыточные кодоны, т. е. триплеты, кодирующие одну и ту же аминокислоту, расположены в соседних положениях таблицы кода, а кодоны для аминокислот, сходных по своим физико-химическим свойствам, обычно различаются только по одному из трех положений, как правило, первому или третьему. Чтобы такая оптимизация могла произойти в процессе эволюции, требуется испытание многих возможностей путем проб и ошибок, в том числе целого ряда вырожденных вариантов кодирования. Итак, доклеточная эволюция характеризовалась, вероятно, такой же степенью ветвления, какую мы находим на уровне видов, если только она управлялась тем же дарвиновским механизмом естественного отбора.

Однако нам не известны никакие варианты генетического кода, даже если рассматривать его тонкую структуру. С одной стороны, было бы совершенно неудовлетворительным допущение, что оптимальные кодовые соответствия возникли по простой случайности, которая произошла всего однажды и как раз в нужный момент времени, не допуская никаких альтернатив, которые несомненно привели бы к ветвлению кода на варианты с различной тонкой структурой. С другой стороны, столь же невероятным было бы допущение, что исторический путь доклеточной эволюции был однозначно задан детерминированными физическими событиями,

Результаты наших исследований заставляют полагать, что дарвиновской эволюции видов предшествовал аналогичный многоэтапный процесс молекулярной эволюции, который привел к созданию уникального аппарата клетки, использующей универсальный код. Этот код окончательно установился не потому, что он был единственно возможным, а потому, что здесь работал своеобразный механизм отбора «раз и навсегда», причем этот процесс мог начаться с любых кодовых соответствий. Отбор «раз и навсегда» является следствием гиперциклической организации [4]. Тщательный анализ механизмов воспроизведения макромолекул позволяет думать, что наличие каталитических гиперциклов — это минимальное требование для возникновения макромолекулярной организации, способной накапливать, сохранять и обрабатывать генетическую информацию.

II. Что такое гиперцикл?

Рассмотрим последовательность реакций, в которой на каждом этапе продукты — с участием дополнительных реагентов или без них — подвергаются дальнейшим превращениям. Если в такой последовательности какой-либо из продуктов идентичен исходному реагенту какой-либо предшествующей стадии, то такая система подобна *реакционному циклу*, а цикл в целом — катализатору. В простейшем случае катализатор — это отдельная молекула, например фермент, превращающий субстрат в продукт:



Механизм, который стоит за этой формальной схемой, требует по меньшей мере трехчленного цикла (рис. 1). Более сложные реакционные циклы, выполняющие фундаментальные каталитические функции, показаны на рис. 2 и 3. Цикл Бете — Вайцзекера [5] (рис. 2) вносит существенный вклад в поддержание высокой скорости высвобождения энергии в массивных звездах. Он, если можно так сказать, застав-

ляет Солнце светить и, следовательно, является одной из наиболее важных *внешних* предпосылок существования жизни на Земле. Не меньшее значение, по-видимому, имеет цикл Кребса, или цикл лимонной кислоты [6], изображенный на рис. 3, хотя он касается *внутреннего* механизма живого. Этот цикл реакций опосредует и регулирует метаболизм углеводов и жирных кислот в живой клетке и выполняет также важные функции в анаболических (биосинте-

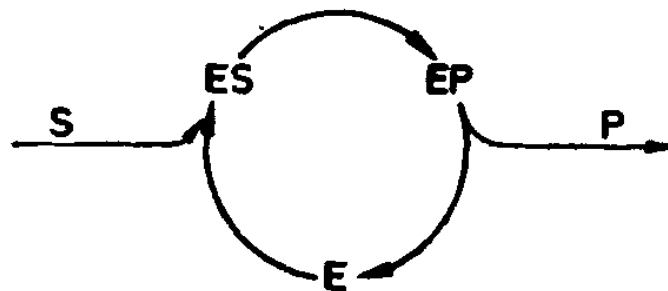


Рис. 1. Общий механизм ферментативного катализа по Михаэлису — Ментен. В этом процессе участвуют по меньшей мере три интермедиата: свободный фермент (E), комплекс фермент — субстрат (ES) и комплекс фермент — продукт (EP). Схема показывает, что каталитическое действие фермента эквивалентно циклическому воспроизведению интермедиатов при превращении субстрата (S) в продукт (P). Впрочем, она дает лишь формальное представление истинного процесса, который может включать в себя многоступенчатую активацию субстрата и индуцированные конформационные изменения фермента.

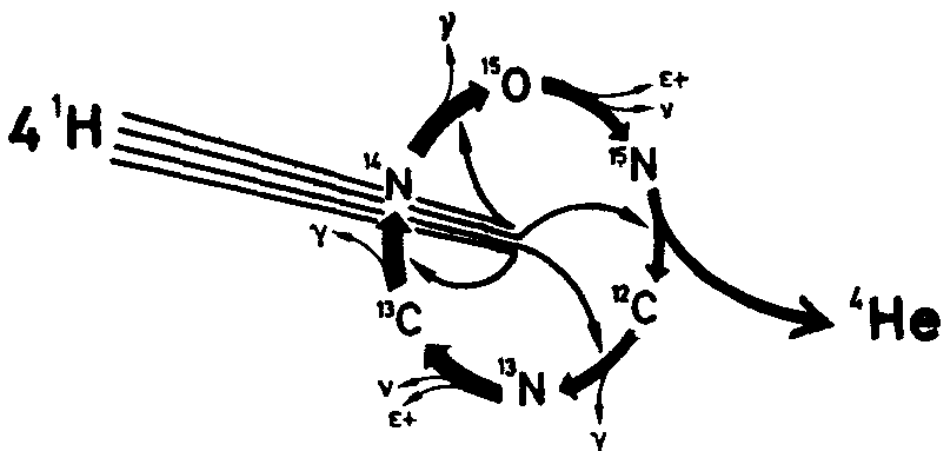


Рис. 2. Углеродный цикл, предложенный Бете и фон Вайцзекером, ответствен — по крайней мере частично — за высвобождение энергии в массивных звездах. Его компоненты — ^{12}C , ^{13}N , ^{13}C , ^{14}N , ^{15}O и ^{15}N — постоянно воспроизводятся в ходе реакционного цикла. Циклическая схема в целом представляет собой катализатор, который превращает четыре атома ^1H в один атом ^4He с высвобождением энергии в виде γ -квантов, позитронов (e^+) и нейтрино (ν).

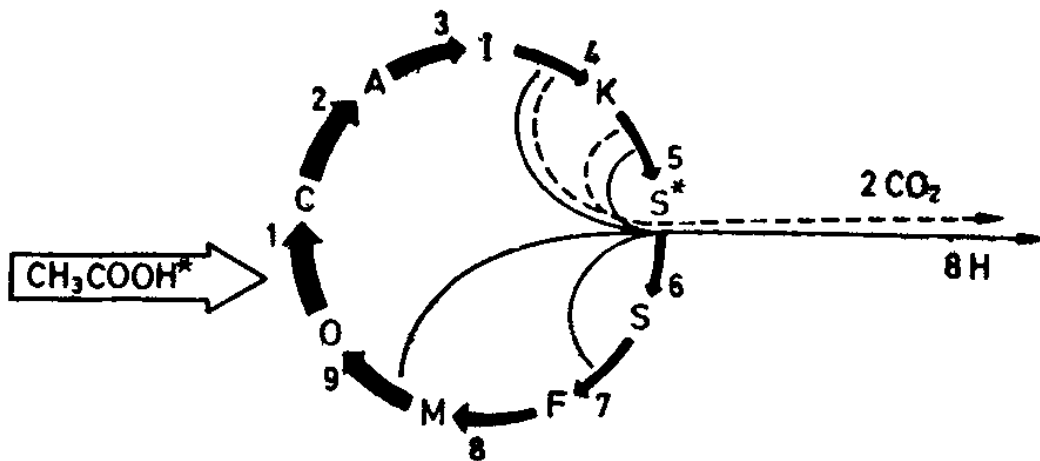


Рис. 3. Цикл трикарбоновых кислот, или цикл лимонной кислоты, — это общий каталитический аппарат биологического окисления «топливных» молекул. Полная схема цикла была предложена Кребсом, важный вклад в ее построение был внесен также Сеит-Дьёрди, Марциусом и Кнопом. Основные компоненты цикла: цитрат (С), *цис*-аконитат (А), изоцитрат (I), α -кетоглутарат (К), сукцинил-СоА (S^*), сукцинат (S), фумарат (F), 1-малат (M) и оксалоацетат (O). Ацетат вступает в цикл в активированной форме в виде ацетил-СоА (этап 1) и реагирует с оксалоацетатом и H_2O с образованием цитрата (С) и СоА ($+H^+$). Во всех превращениях участвуют ферменты и такие кофакторы, как СоА (этапы 1, 5, 6), Fe^{2+} (этапы 2, 3), NAD^+ (этапы 4, 5, 9), ТПФ, липоевая кислота (этап 5) и FAD (этап 7). На схеме не показаны дополнительные реагенты: H_2O (этапы 1, 3, 8), P_i и GDP (этап 6) и продукты реакций — H_2O (этап 2), H^+ (этапы 1, 9) и GTP (этап 6). В результате всех реакций цикла происходит полное окисление двух ацетильных углеродов до CO_2 (и H_2O). За один оборот цикла генерируется 12 макроэргических фосфатных связей, причем одна из них образуется в самом цикле (GTP, этап 6), а 11 — в результате окисления $NADH$ и $FADH_2$ [три пары электронов переносятся к NAD^+ (этапы 4, 5, 9) и одна пара — к FAD (этап 7)].

NB: Хотя цикл в целом функционирует как катализатор благодаря воспроизведению интермедиатов, он не похож на *каталитический цикл*, изображенный на рис. 4. Несмотря на то что каждый этап цикла катализируется ферментом, ни один из этих ферментов не образуется в самом цикле.

СоА — кофермент А, NAD — никотинамидадениндинуклеотид, GTP — гуанозинтрифосфат, FAD — флавинадениндинуклеотид, ТПФ — тиаминпирофосфат, GDP — гуанозиндифосфат, P_i — фосфат неорганический.

тических) процессах. В обеих схемах высокоэнергетическое вещество превращается в продукты, бедные энергией, при сохранении, т. е. циклическом воспроизведении существенных промежуточных продуктов. Оба цикла были предложены примерно в одно и то

же время (1937—1938 гг.), хотя между ними нет никакой причинной связи.

Однонаправленность циклического воспроизведения интермедиатов, конечно, предполагает, что система находится в состоянии, далеком от равновесия, и всегда связана с расходом энергии, часть которой диссипирует в окружающую среду. С другой стороны, установление равновесия в замкнутой системе приведет к тому, что все стадии будут находиться в состоянии детального равновесия. Каталитическое действие в такой замкнутой системе будет микроскопически обратимо, т. е. оно будет происходить с равной эффективностью в обоих направлениях.

Теперь, пользуясь простой итерацией, построим иерархии реакционных циклов и зададим их частные свойства. Это означает, что на следующем этапе мы рассматриваем реакционный цикл, в котором по меньшей мере один, но, возможно, и все интермедиаты сами являются катализаторами. Отметим, что эти интермедиаты, будучи катализаторами, теперь остаются неизменными во время реакции. Каждый из них образуется из потока высокоэнергетического строительного материала с использованием каталитической поддержки от предшествующего интермедиата (рис. 4). При большом числе интермедиатов такая система становится очень сложной по составу и поэтому вряд ли может встретиться в природе. Самым известным примером является четырехчленный цикл, связанный с матричной репликацией молекулы РНК (рис. 5). Такого рода механизм исследовался *in vitro*: использовали реакционную среду, содержащую нуклеозидтрифосфаты четырех типов в качестве высокоэнергетических строительных материалов, а также фаговую репликазу, которая играла роль постоянного фактора среды [7, 8] (более детальное описание дано Кюпперсом [9]). Каждая из двух цепей действует как матрица, инструктирующая синтез своей комплементарной копии,— как при фотографической репродукции.

Простейшим представителем этой категории реакционных систем является отдельный автокатализатор

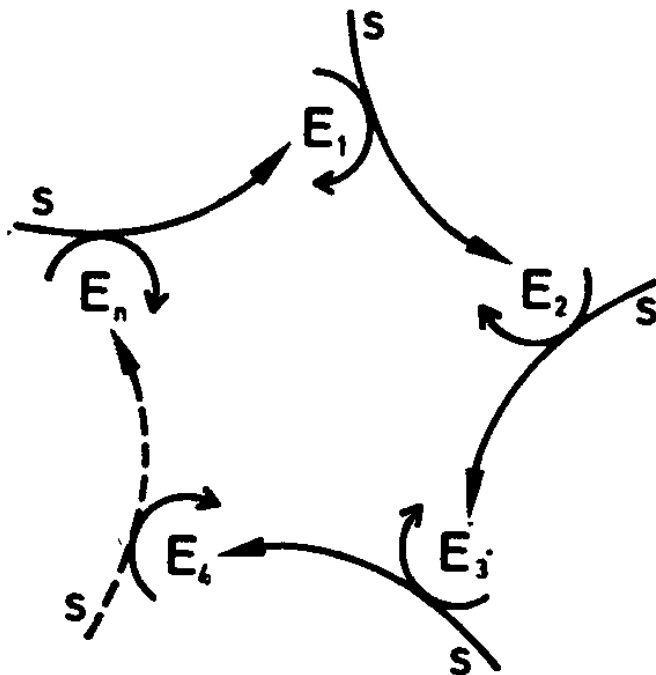
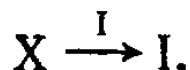


Рис. 4. Каталитический цикл представляет собой более высокий уровень организации в иерархии каталитических схем. Компоненты цикла $E_1 \rightarrow E_n$ сами являются катализаторами, которые образуются из каких-либо высокоэнергетических субстратов (S), причем каждый интермедиат E_i является катализатором для образования E_{i+1} . Каталитический цикл в целом эквивалентен автокатализатору, инструктирующему свое собственное воспроизведение. Для существования каталитического цикла достаточно, чтобы один из интермедиатов был катализатором для одной из последующих реакций.

или — в случае целого класса объектов I_i , несущих информацию, — самовоспроизводящаяся единица. Такой процесс формально может быть описан следующим образом:



В этой книге будут часто рассматриваться реакции такого типа; мы будем обозначать их символом $\textcircled{1}$.

Такой истинно самовоспроизводящейся формой является двухцепочечная ДНК (в отличие от одноцепочечной РНК) — обе ее цепи одновременно копируются полимеразой [10] (рис. 6). Эта же формальная схема применима и в случае прокариотической клетки, где все наследование осуществляется в основном независимо в каждой клеточной линии.

Кинетические уравнения как простых каталитических, так и автокаталитических систем при постоян-

ной концентрации субстрата содержат член первого порядка по концентрации катализатора. Кривые роста, однако, четко дифференцируют две эти системы. Количество продукта для простого каталитического процесса в указанных условиях растет во времени линейно, тогда как для автокаталитических систем характерен экспоненциальный рост.

Строго говоря, автокаталитическую систему уже можно называть гиперциклической, ибо она предста-

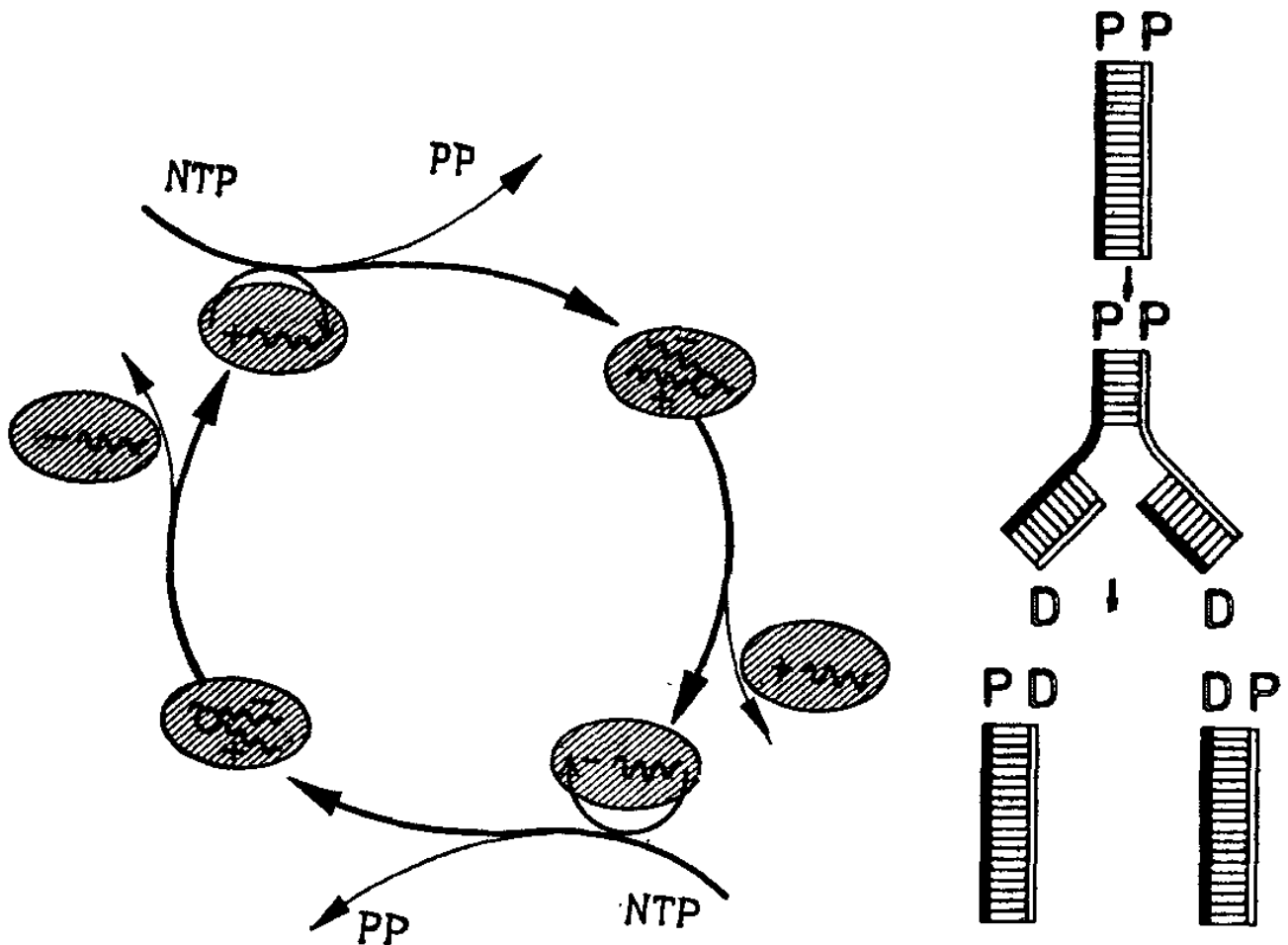


Рис. 5. Каталитический цикл, имеющий биологическое значение, — репликация одноцепочечной РНК. Интермедиаты (плюс- и минус-цепи) участвуют в цикле как матрицы для своего взаимного воспроизведения. Матричная функция эквивалентна избирательному катализу. Нуклеозидтрифосфаты (NTP) представляют собой высокоэнергетический строительный материал, а пирофосфат (PP) является отходом реакций цикла. Комплементарное инструктирование, механизм которого будет обсуждаться в связи с рис. 11, является существенно автокаталитической функцией, т. е. функцией самовоспроизведения.

Рис. 6. Истинный процесс самовоспроизведения (одночленный каталитический цикл) — репликация ДНК. Механизм этого чрезвычайно сложного процесса (см. рис. 12) обеспечивает связь каждой дочерней цепи (D) с соответствующей родительской цепью (P).

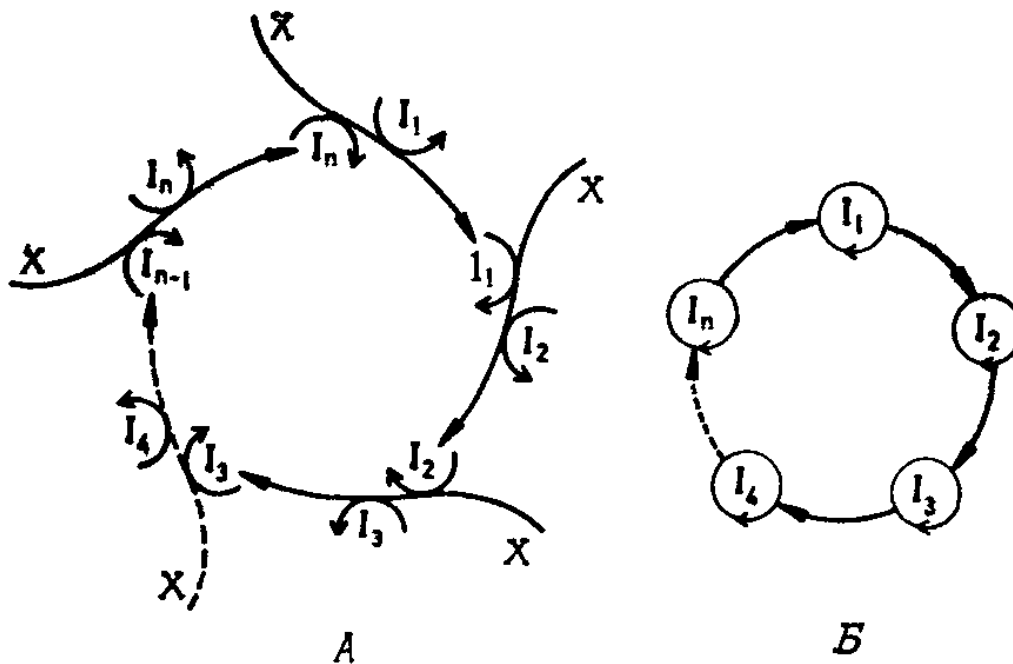


Рис. 7. Каталитический гиперцикл состоит из самоинструктурирующихся единиц I_i с двойными каталитическими функциями. В качестве автокатализатора или — в более общем случае — в качестве каталитического цикла интермедиат I_i способен инструктировать свое собственное воспроизведение, а кроме того, он оказывает каталитическое содействие воспроизведению следующего интермедиата (с использованием высокоэнергетического строительного материала X). Упрощенная схема (Б) показывает иерархию циклов.

вляет собой цикл катализаторов, которые сами являются циклами реакций. Мы, однако, ограничим использование этого термина такими системами, которые являются гиперциклическими по отношению к *каталитической функции*. В действительности это гиперциклы второй или более высокой степени, потому что они относятся к реакциям по меньшей мере второго порядка по концентрациям катализаторов.

Каталитический гиперцикл — это система, связывающая автокаталитические, или самовоспроизводящиеся единицы друг с другом посредством циклической связи. Такая система изображена на рис. 7. Интермедиаты $I_1—I_n$ как самовоспроизводящиеся единицы сами являются каталитическими циклами, например комбинациями плюс- и минус-цепей РНК, как показано на рис. 5. Однако процесс репликации как таковой должен прямо или косвенно стимулироваться дополнительными специфическими связями между различными репликативными единицами,

Реально такие связи могут обеспечиваться белками — продуктами трансляции предшествующих циклов воспроизведения РНК (рис. 8). Эти белки могут действовать как специфические репликазы или репрессоры или же как специфические защитные факторы, предохраняющие от деградации. Связи между самовоспроизводящимися единицами должны образовывать наложенный цикл — лишь тогда система в целом будет подобна гиперциклу. По сравнению с системами, показанными на рис. 4 и 5, гиперцикл является самовоспроизводящимся в более высокой степени.

Простейшим представителем систем этого класса снова является (квази)одношаговая система, т. е. усиленный автокатализатор. Мы сталкиваемся с такой системой в случае инфекции клетки РНК-

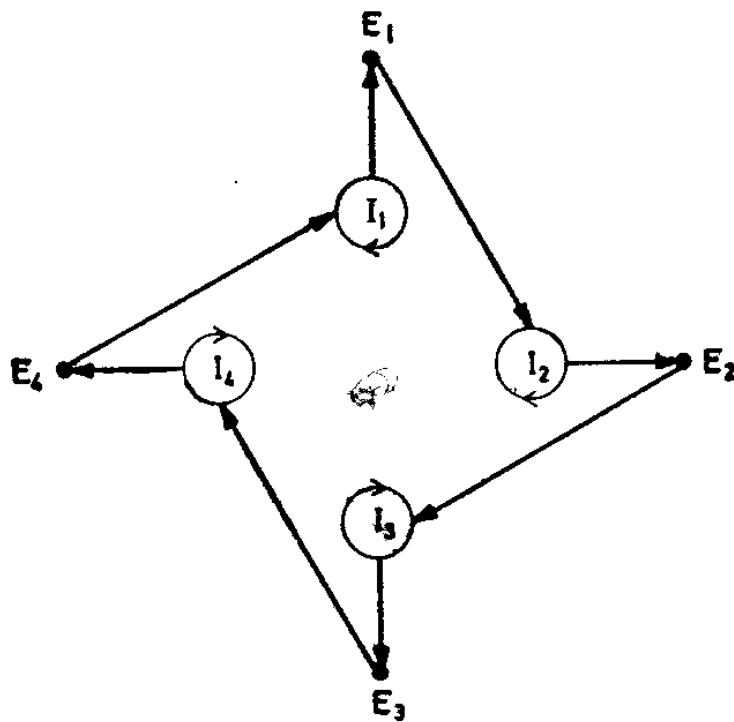


Рис. 8. Реалистическая модель гиперцикла второй степени, где носители информации I_i «выдают» инструкции двух типов: во-первых, для собственного воспроизведения и, во-вторых, для трансляции с образованием интермедиатов E_i второго типа, которые обладают оптимальными функциональными свойствами. Каждый фермент E_i катализирует воспроизведение следующего носителя информации I_i . Он может иметь и другие каталитические свойства, необходимые для осуществления процесса трансляции, метаболизма и т. д. В подобных случаях степень гиперциклической связи больше двух.

содержащим фагом (рис. 9). Если фаговая РНК (плюс-цепь) проникает в бактериальную клетку, то ее генотипическая информация транслируется аппаратом клетки-хозяина. Один из продуктов трансляции связывается с определенными факторами хозяина с образованием активного ферментного комплекса, который специфически реплицирует плюс- и минус-цепи фаговой РНК, действующие при воспроизведении как матрицы друг для друга [11]. Однако реп-

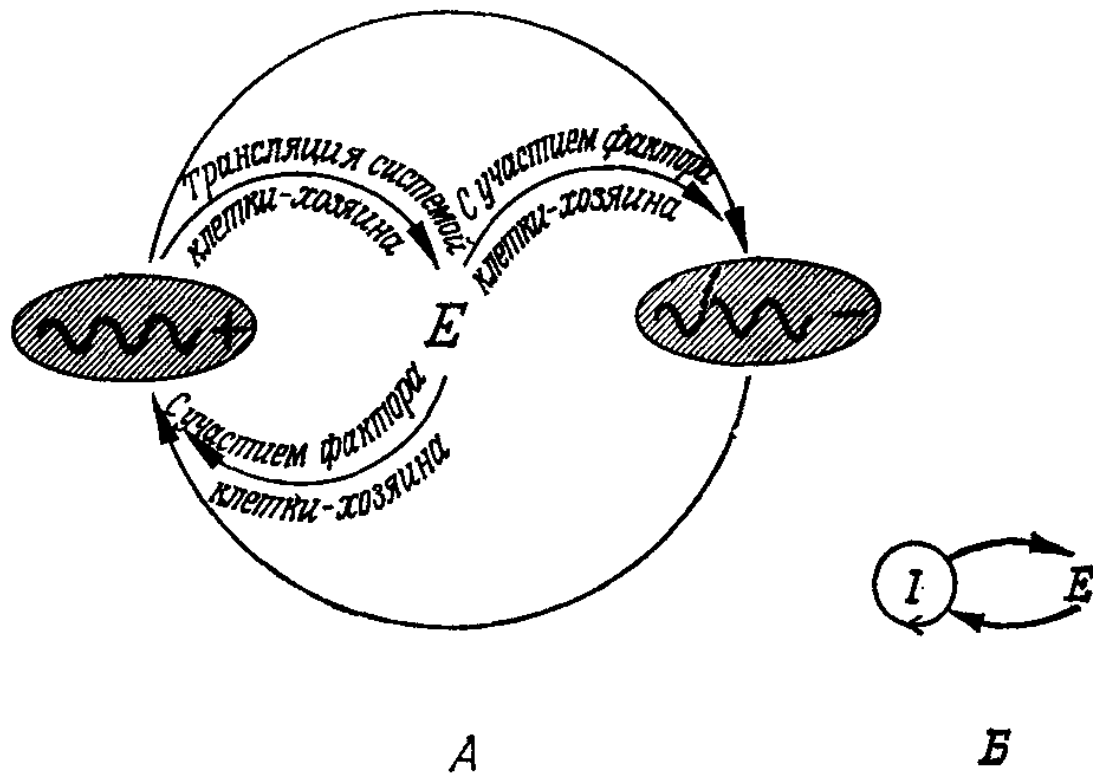


Рис. 9. Размножение РНК-содержащего фага в бактериальной клетке представляет собой простой гиперциклический процесс. Инфекционная плюс-цепь, используя аппарат трансляции клетки-хозяина, сначала инструктирует синтез белковой субъединицы (Е), которая связывается с другими белками клетки-хозяина с образованием фагоспецифичной РНК-репликазы. Этот репликазный комплекс узнает только определенные фенотипические особенности фаговой РНК, которые имеются как у плюс-, так и у минус-цепи благодаря симметрии определенных участков цепи РНК. В результате происходит вспышка синтеза фаговой РНК, которая — из-за гиперциклической природы — носит гиперболический характер (см. часть Б, рис. 17), пока не наступит насыщение по одному из интермедиатов или пока не истощатся запасы метаболитов клетки-хозяина. Схема Б показывает, что автокаталитическими или самоинструктивными свойствами может обладать всего один из интермедиатов при условии, что другие «партнеры» связаны с ним с помощью циклической обратной связи.



Катализатор E

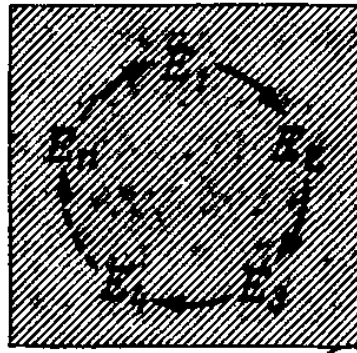
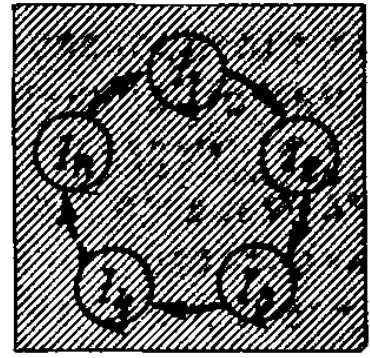
Автокатализатор I
(самореплицирующаяся
единица)Каталитический
гиперцикл H

Рис. 10. Иерархия циклических реакционных сетей очевидна из сопоставления этих схем (\rightarrow химическое превращение, \rightarrow каталитическое действие).

ликазный комплекс не размножает — в сколько-нибудь ощутимой степени — информационную РНК клетки-хозяина. В результате инфекции начинается *гиперболический* рост числа фаговых частиц, который в конце концов выходит на насыщение из-за ограниченности ресурсов клетки-хозяина.

Другой естественный гиперцикл может возникнуть в менделевских популяциях в начальной фазе видообразования — пока численности популяций малы. Для репродукции генов необходимо взаимодействие между двумя аллелями (M и F), т. е. между гомологичными участками в мужской и женской хромосомах, которые затем появляются в потомстве в перестроенной комбинации. Тот факт, что кинетические уравнения популяционной генетики менделеевских популяций [12] обычно не содержат гиперциклической нелинейности (которая ведет к гиперболическому, а не к экспоненциальному росту), объясняется насыщением, наступающим при сравнительно малых численностях популяций, когда скорость рождений (обычно) становится пропорциональной только численности самок.

Как видно из сравнения схем, показанных на рис. 10, гиперциклы представляют собой новый уровень организации. Этот факт проявляется в их уникальных свойствах. Не связанные друг с другом

самовоспроизводящиеся единицы гарантируют сохранение ограниченного количества информации, которое может передаваться от поколения к поколению. Это оказывается одной из необходимых предпосылок дарвиновского поведения, т. е. отбора и эволюции [13]. Подобно им, гиперциклы тоже селекционны, но, кроме того, они имеют интегрирующие свойства, которые делают возможной кооперацию между единицами, конкурирующими в других условиях. Более того, гиперциклы конкурируют даже еще более ожесточенно, чем дарвиновские виды, с любой самовоспроизводящейся единицей, не являющейся частью гиперцикла. Далее, они способны устанавливать глобальные формы организации благодаря свойству отбирать «раз и навсегда», которое не допускает сосуществования с другими гиперциклическими системами, если только они не стабилизированы связями более высокого порядка.

Простейший тип связи внутри гиперцикла — это прямая стимуляция или дерепрессия, которым в кинетических уравнениях соответствуют члены второго порядка для процессов образования. Члены для связей более высокого порядка могут войти тоже, и они, таким образом, определяют степень p гиперциклической организации.

Отдельные гиперциклы могут также связываться друг с другом с образованием иерархий. Это, однако, требует введения членов интерциклической связи, которые существенно зависят от степени организации. Чтобы установилось устойчивое сосуществование, для гиперциклов H_1 и H_2 , имеющих степени внутренней организации p_1 и p_2 , необходимы члены интерциклической связи степени $p_1 + p_2$.

Цель этой работы состоит в том, чтобы разработать детальную теорию той категории сетей реакций, которые мы окрестили *гиперциклами* [4]; и обсудить их значимость для биологической самоорганизации, особенно в связи с возникновением трансляции — события, которое можно считать решающим этапом в доклеточной эволюции.

III. Дарвиновские системы

III.1. Принцип естественного отбора

В физике известны законы, которые не могут быть сведены к более фундаментальным принципам. Как аксиомы они абстрагированы из опыта, причем любые следствия из них не противоречат эксперименту. Типичные примеры — первый и второй законы термодинамики.

Дарвиновский принцип естественного отбора не относится к категории первичных принципов. Согласно популяционной генетике [14], естественный отбор — это следствие очевидных фундаментальных свойств популяций живых организмов, подверженных определенным внешним воздействиям. Этот принцип позволяет четко определить смысл термина «наиболее приспособленный к условиям среды», не сводя его лишь к простой тавтологии «выживание выживающих». В применении к природным популяциям с их изменчивыми и обычно неизвестными граничными условиями принцип все-таки дает нам объяснение самого факта эволюции и филогенетических взаимоотношений между видами. В этом и состояла основная цель Чарлза Р. Дарвина [15] и его современника Альфреда Р. Уоллеса [16] — дать более удовлетворительное обоснование принципу эволюционного происхождения.

На самом деле большинство работ по генетике популяций в настоящее время связано с более практическими проблемами, когда рассматривают распространение генетической информации по менделевским популяциям и оставляют в стороне такие академические вопросы, как, например: действительно ли свойство быть живым является *необходимой* предпосылкой для селекционного и эволюционного поведения. Тот факт, что очевидные атрибуты живых организмов — метаболизм, самовоспроизведение, конечность продолжительности жизни и мутабельность — достаточны для объяснения селекционного и эволюционного поведения при наличии соответствующих

ограничений, заставил многих современных генетиков думать, что эти свойства характеризуют *лишь* феномен жизни и не могут быть присущи неживым системам [14]. Опыты в пробирках [7], результаты которых заставляют думать, что здесь естественный отбор и эволюция происходят *in vitro*, интерпретировались как наблюдение постбиологических явлений, а не как демонстрация типичного и специфического поведения материи. Необходимо отметить, что действие естественного отбора обнаруживается даже в модах лазерного излучения, и анализ механизма их усиления говорит о том, что здесь имеется не только формальная аналогия. И тем не менее никто не решился бы назвать моду лазерного излучения «живой».

Такие проблемы вряд ли привлекут внимание тех, кто занимается изучением свойств реальных живых организмов, однако они приобретают первостепенное значение в связи с проблемой *происхождения жизни*. Здесь нам действительно нужно выявить *необходимые* предпосылки, чтобы найти такие молекулярные системы, которые были бы способны к эволюционной самоорганизации. Та сложность, с которой мы сталкиваемся на уровне процесса макромолекулярной организации, требует, чтобы этот процесс управлялся такими же принципами отбора и эволюции, какие действуют в мире живого.

Этим проблемам были посвящены недавние исследования (loc. cit.), как теоретические, так и экспериментальные. Далее мы вкратце остановимся на некоторых ранее полученных результатах, касающихся дарвиновских систем.

III.2. Необходимые предпосылки возникновения дарвиновских систем

Что составляет молекулярную основу отбора и эволюции? Очевидно, что такое поведение не является глобальным атрибутом любой произвольной формы материи, а обусловлено наличием особых свойств, которые необходимо определить.

Для того чтобы система обладала способностью к самоотбору, должно выполняться важное условие — она должна стабилизировать определенные структуры за счет остальных. Критерии такой стабилизации имеют динамическую природу, потому что решение о том, какой вид должен быть отобран, зависит от распределения конкурентов, имеющихся в каждый момент времени. Другими словами, любая однажды отобранная структура не является статически устойчивой — она может стать неустойчивой, как только появятся другие, более «благоприятные» структуры или изменятся условия среды. Критерии оценки должны включать какой-нибудь элемент обратной связи, который приводил бы в соответствие ценность и динамическую устойчивость. Мутант с селективным преимуществом, однажды возникший вследствие некой флуктуации, должен иметь способность усиливаться в присутствии большого избытка менее приспособленных конкурентов. Таким образом, преимущество должно сопоставляться хотя бы с некоторыми из тех динамических свойств, которые ответственны за усиление. Лишь таким путем система может селекционно самоорганизоваться в отсутствие «внешнего селектора». Необходимое свойство обратной связи — это внутренне присущая способность к автокатализу, т. е. самовоспроизведение.

В работе [18], где был дан общий анализ с применением игровых моделей, мы указали те свойства материи, которые необходимы для дарвиновского поведения на молекулярном уровне. Эти свойства перечислены ниже.

1. *Метаболизм.* Как образование, так и разложение молекулярных видов должны быть независимы друг от друга и спонтанны, т. е. их движущей силой должно быть положительное сродство. Этого нельзя достичь ни в какой равновесной системе, где оба процесса взаимосвязаны микроскопической обратимостью, приводящей к устойчивому распределению всех конкурентов, присутствующих в системе. Сложность, т. е. наличие громадного числа альтернативных

структур в сочетании с пространственно-временными ограничениями, просто не допускает такого уравновешивания и требует постоянного разрушения и образования новых структур. Отбор может действовать только на промежуточные состояния, которые образуются из высокоэнергетических предшественников и разрушаются, превращаясь в низкоэнергетические отходы. Способность системы использовать необходимые для этой цели свободную энергию и вещества называется метаболизмом. Необходимость поддержания системы достаточно далеко от равновесия за счет постоянной компенсации производства энтропии впервые была ясно осознана Эрвином Шредингером [19].

2. *Самовоспроизведение.* Конкурирующие молекулярные структуры должны обладать внутренней способностью инструктировать свой собственный синтез. Можно показать, что такая автокаталитическая функция является необходимой для любого механизма отбора, включающего дестабилизацию популяции в присутствии единственной копии вновь возникшего благоприятного мутанта. Кроме того, самокопирование необходимо для сохранения информации, накопленной в системе за предыдущий период. Постоянная деградация — необходимая предпосылка для наличия свойств 1 и 3 — в противном случае привела бы к полной утрате информации.

3. *Мутабельность.* Точность любого процесса самовоспроизведения при конечной температуре ограничена из-за теплового шума. Действие шума особенно заметно, если копирование идет быстро — вернее, если энергия взаимодействия для каждого элементарного шага копирования лишь ненамного выше тепловой энергии. Следовательно, мутабельность всегда физически связана со способностью к самовоспроизведению. Более того, она (логически) необходима для эволюции. Ошибки копирования — это основной источник новой информации. Как мы увидим, для темпа мутаций существует пороговое значение, при котором скорость эволюции макси-

мальна, но которое не может быть превышено без потери всей информации, накопленной в ходе эволюции.

Лишь такие макромолекулярные системы, которые обладают всеми тремя необходимыми свойствами, могут быть носителями информации в неограниченном по существу эволюционном процессе. Перечисленные свойства должны наследоваться всеми членами соответствующего макромолекулярного класса, т. е. всеми возможными альтернативами, или мутантами данной структуры, и далее — они должны эффективно проявляться в широкой области концентраций — от одной копии до макроскопических количеств. «Предпосылка реализуемости» исключает системы со сложным составом и структурой, в которых указанные свойства проистекали бы из случайного сочетания молекулярных взаимодействий, а не из общего физического принципа. Рассмотрим, например, нуклеиновые кислоты в сравнении с белками. Основным, присущим всем нуклеиновым кислотам свойством является воспроизведение, основанное на силах физической природы, от которых зависит уникальная комплементарность четырех оснований. Белки же имеют гораздо более широкие функциональные возможности, они могут обладать, в частности, инструктирующими и репродуктивными свойствами. Однако каждая отдельная функция является следствием специфической укладки полипептидной цепи и не может быть атрибутом всего класса белков. Иногда она даже полностью утрачивается в результате одной мутации.

Чтобы материальные системы были способны к селекционной самоорганизации, они должны наследовать физические свойства, допускающие метаболизм, т. е. превращение высокоэнергетических реагентов в бедные энергией продукты, и самовоспроизведение («шумом»). Эти предпосылки являются необходимыми. При подходящих внешних условиях они оказываются также достаточными для селекционного и эволюционного поведения.

III.3. Динамика отбора

Простейшая система, удовлетворяющая указанным необходимым условиям, может описываться системой дифференциальных уравнений следующего вида [4] ($\dot{x} = dx/dt$; t — время):

$$\dot{x}_i = (A_i Q_i - D_i) x_i + \sum_{k \neq i} w_{ik} x_k + \Phi_i, \quad (1)$$

где i — индекс, которым помечены все различные самовоспроизводящиеся молекулярные единицы и, следовательно, характеризующий их индивидуальную (генетическую) информацию. Символом x_i мы обозначаем соответствующую популяционную переменную (или концентрацию). Физический смысл других параметров станет очевидным из обсуждения этого уравнения. Данная система уравнений прежде всего включает в себя те самовоспроизводящиеся единицы i , которые присутствуют в рассматриваемой системе и которые могут быть пронумерованы от 1 до N . Она может быть расширена так, чтобы в нее вошли и всевозможные мутанты, часть из которых появляется в ходе эволюции.

В этих уравнениях, описывающих открытую систему, метаболизм представлен спонтанным образованием ($A_i Q_i x_i$) и разложением ($D_i x_i$) молекулярного вида. «Спонтанность» означает, что обе реакции идут с положительным сродством и, следовательно, не являются взаимно обратимыми. Член A_i всегда содержит какую-либо стехиометрическую функцию $f_i(m_1, m_2, \dots, m_\lambda)$ концентраций высокоэнергетического строительного материала (λ классов), необходимого для синтеза молекулярного вида i , точная форма которой зависит от конкретных механизмов реакций. Этот высокоэнергетический строительный материал должен постоянно подводиться притоком вещества, а продукты реакций должны удаляться соответствующим оттоком (Φ_i). Член D_i , описывающий спонтанное разложение, линейно связан с x_i по обычному закону первого порядка. В более сложных системах как A_i , так и D_i могут включать в себя и другие функции концентраций, если соответствующие реакции катализируются ферментами или если между реакциями имеются дополнительные связи.

Самовоспроизведение — вторая предпосылка — проявляется в том, что член, описывающий образование, зависит от x_i . Прямая линейная зависимость является лишь простейшей формой автокатализа. Другие, более сложные, но все еще линейные механизмы, например комплементарное инструктирование или циклический катализ, можно рассматривать, как это будет показано, аналогичным образом. Нелинейный же автокатализ является главным объектом рассмотрения в этой части.

Мутабельность отражается в факторе качества Q_i , который может принимать любое значение между нулем и единицей. Этот фактор задает ту долю репродукций, которая происходит на данной матрице i с образованием ее точной копии. Конечно, имеется дополнительный член, относящийся к неточной репро-

дукции матрицы i , — $A_i(1 - Q_i x_i)$. Он отражает образование множества разнообразных «ошибочных копий», которые в большинстве случаев очень мало отличаются от вида i . Образование ошибочных копий i описывается соответствующими членами в кинетических уравнениях для его «родственников» k . В свою очередь, копия i получает вклады от этих «родственников» из-за ошибок в их репликации. Они учитываются суммой $\sum_{k \neq i} w_{ik} x_k$.

Параметр w_{ki} , задающий индивидуальный темп мутаций, обычно мал по сравнению с параметром скорости репродукции $A_i Q_i$ — его значение тем меньше, чем больше различие между i и k . Если присутствуют все виды и все их возможные мутанты учитываются индексами i и k (пробегающими все значения от 1 до N), то для ошибочных копий будет выполняться следующий закон сохранения:

$$\sum_i A_i (1 - Q_i) x_i = \sum_i \sum_{k \neq i} w_{ik} x_k \quad (2)$$

Наконец, член Φ_i , задающий индивидуальный поток, или транспорт, описывает добавление или удаление вида i любым способом, кроме химической реакции. Он нужен для того, чтобы учесть метаболические превращения (см. выше). В большинстве случаев каждый вид вносит вклад в суммарный поток Φ_i в соответствии со своим изобилием:

$$\Phi_i = \Phi_t \frac{x_i}{\sum_k x_k} \quad (3)$$

В эволюционных экспериментах суммарный поток можно подобрать так, чтобы обеспечивалось поддержание воспроизводимых глобальных условий — таких, как постоянство суммарной плотности популяции:

$$\sum_k x_k = \text{const} \equiv c_n \quad (4)$$

В этом случае поток Φ_t необходимо постоянно регулировать так, чтобы компенсировать избыточную общую продукцию, т. е.

$$\Phi_t = \sum_k A_k x_k - \sum_k D_k x_k \equiv \sum_k E_k x_k \quad (5)$$

Величину $E_i \equiv A_i - D_i$ мы называем «избыточной продуктивностью» матрицы i . Отметим, что ошибки не влияют на вид этой суммы вследствие уравнения сохранения (2).

Если, кроме того, отдельные потоки высокоэнергетического строительного материала также регулируются, задавая постоянные буферные уровни (m_1, \dots, m_λ) для каждого из λ классов, то стехиометрические функции $f_i(m_1, \dots, m_\lambda)$, входящие в выражения для кинетических параметров A_i , будут постоянны, и поэтому их явный вид можно не конкретизировать. Это ограни-

чение, при котором как количество неорганизованного материала, так и суммарное количество организованного материала поддерживаются на постоянном уровне путем регуляции потоков, мы будем называть ограничением «постоянной общей организации». Оно обычно выполняется в эволюционных экспериментах, например в проточном реакторе [9], или — в среднем — в опытах с серийными переносами [7]. Другое простое ограничение — это условие «постоянных потоков». В этом случае уровни концентраций являются переменными; они зависят от метаболизма при фиксированных притоках и оттоках. Оба ограничения приводят к тому, что система начинает проявлять ярко выраженное селекционное поведение при подходе к стационарному состоянию. Количественные результаты для обоих ограничений могут оказаться различными, но качественное поведение оказывается очень близким [4]. Поэтому здесь достаточно рассмотреть только один из двух предельных случаев. В природе характер ограничений может изменяться во времени, и, следовательно, они обычно не будут соответствовать никакой из этих простых крайних ситуаций — так же как в случае погодных условий обычно не выполняются простые термодинамические ограничения (например, постоянство давления, температуры и т. д.). Однако существенные принципы естественного отбора можно изучать только при контролируемых и воспроизводимых экспериментальных условиях.

Для ограничения постоянной общей организации кинетические уравнения (1) в сочетании с дополнительными условиями (2) — (5) принимают следующий вид:

$$\dot{x}_i = (W_{ii} - \bar{E}(t)) x_i + \sum_{k \neq i} w_{ik} x_k, \quad (6)$$

где величину

$$W_{ii} = A_i Q_i - D_i \quad (7)$$

можно назвать (внутренней) селективной ценностью, а величину

$$\bar{E}(t) = \sum_k E_k x_k / \sum_k x_k \quad (8)$$

— средней избыточной продуктивностью, которая является функцией времени. Только когда популяционные переменные $x_k(t)$ становятся постоянными, $\bar{E}(t)$ достигает стационарного значения, которое является метастабильным, потому что зависит от заселенностей спектра мутантов. Для постоянных (т. е. не зависящих от времени) значений W_{ii} и w_{ik} нелинейная система дифференциальных уравнений (6) может быть решена. Приближенные решения селекционной задачи были опубликованы ранее. В последние годы точное решение было получено Томпсоном и МакБрайдом [20] и независимо Джоисом, Энсом и Рангнекаром [21]. Выражения, полученные из точных решений и соответствующие второму порядку теории возмущений, находятся в

согласии с ранее опубликованными приближениями [4]¹. Следующее обсуждение основано на точных решениях Джонса и др. [21], которые дали элегантное количественное представление проблемы отбора.

III.4. Концепция квазивида

Отдельный вид не является независимой единицей из-за наличия связей. Сохранение суммарной численности популяции вынуждает все виды конкурировать друг с другом, в то время как мутабельность допускает еще некоторую кооперацию, особенно между близкородственными видами (т. е. видами i и k , для которых члены w_{ik} и w_{ki} не равны нулю).

Поэтому реорганизуем нашу систему следующим образом. Вместо того чтобы подразделять всю популяцию на N видов, введем новую систему N квазивида, популяционные переменные y_i которых являются линейными комбинациями исходных популяционных переменных x_i , причем их сумма, конечно, не изменяется:

$$\sum_{k=1}^N x_k = \sum_{k=1}^N y_k. \quad (9)$$

¹ Джонс и др. [21] отметили, что пренебрежение членом, соответствующим обратному потоку, $\sum_{k \neq i} w_{ik} x_k$, в работе [4]

незаконно при подходе к стационарному состоянию, потому что член $W_{mm} - \bar{E}(t)$ становится очень малым. Авторы ссылаются на уравнение II-49 из работы [4], где мутабельность намеренно не учитывалась (т. е. $Q = 1$), чтобы продемонстрировать природу решений, типичных для отбора. Однако они не обратили внимания на замечание, сделанное на стр. 482 в [4], что такое допущение может приближенно выполняться только в отношении доминирующего вида с четко выраженным селективным преимуществом, тогда как существование мутантов учитывается лишь членами w_{ik} . Приближения, полученные ранее (уравнения II-33а, II-43, II-59, II-69, II-72 в работе [4], см. также [22]), в действительности количественно согласуются с теми приближениями, которые можно получить из точных решений с помощью теории возмущений (уравнения (21) и (22) в [21] и (13), (18) и (19) из настоящей работы).

С другой стороны, мы хотели бы подчеркнуть, что считаем очень важным наличие точных решений, полученных Томпсоном и Мак-Брайдом [20] и Джонсом и др. [21]; это необычайно облегчает задачу создания связной и последовательной концепции квазивида.

Каким образом выполнить это новое подразделение, подсказывает структура дифференциальных уравнений (6). Оно соответствует аффинному преобразованию системы координат, хорошо известному из теории линейных дифференциальных уравнений. Для трансформированных популяционных переменных y_i получим новую систему уравнений:

$$\dot{y}_i = (\lambda_i - \bar{E}(t)) y_i. \quad (10)$$

Применение этой процедуры к нелинейным уравнениям (6) возможно благодаря тому, что член, ответственный за нелинейность, $\bar{E}(t)$, согласно уравнению (8), остается инвариантным при трансформации и теперь может быть выражен как среднее всех λ_i :

$$\bar{E}(t) = \frac{\sum_k \lambda_k y_k}{\sum_k y_k}. \quad (11)$$

Величины λ_i являются собственными значениями линейной динамической системы. Они, как и собственные векторы, которые связывают x_i с y_i , могут быть получены из матрицы, состоящей из коэффициентов W_{ii} и w_{ik} .

Физический смысл решений системы (10) очевиден. Любой квазивид (характеризующийся собственным значением λ_i и популяционной переменной y_i), собственное значение λ_i которого лежит ниже порога — среднего $\bar{E}(t)$ — будет вымирать. (Скорость изменения его численности отрицательна!) Соответственно численность каждого квазивида с λ_i выше порога будет расти. Далее, порог $\bar{E}(t)$ является функцией времени и — согласно уравнению (11) — будет расти, поскольку система благоприятствует квазивидам, которым соответствуют большие собственные значения. Это будет продолжаться до тех пор, пока система не достигнет стационарного состояния:

$$\bar{E}(t) \rightarrow \lambda_{\max}, \quad (12)$$

т. е. средняя продуктивность будет расти, пока не сравняется с максимальным собственным значением. К этому моменту все квазивиды, кроме одного,

а именно того, которому соответствует максимальное собственное значение, окажутся вымершими. Их популяционные переменные должны обратиться в нуль.

Итак, дарвиновский отбор и эволюция могут характеризоваться экстремальным принципом. Он определяет тип поведения репликативных единиц при заданных ограничениях отбора.

Здесь можно видеть аналогию, например, с процессом приближения к равновесию — фундаментальным типом поведения материальных систем при ограничивающем условии изоляции, которое характеризуется общим экстремальным принципом. Экстремальный принцип (12) связан с критериями устойчивости Пригожина—Глансдорфа [23]. Как оптимизационный принцип, он справедлив также для определенных классов нелинейных динамических систем [21]. Далее, справедливость решений квазилинейной системы (10) не ограничивается окрестностью стационарного состояния.

Каков же физический смысл квазивида? Вид в биологии — это группа особей, характеризующихся определенным фенотипом. На генотипическом уровне особи данного вида могут несколько различаться, но тем не менее их молекулы ДНК имеют весьма близкую структуру. Различаются же они последовательностью нуклеотидов. Рассматривая теперь такие молекулы — репликативные единицы, — мы используем именно эти различия в последовательности для того, чтобы дать определение (молекулярного) вида. Эти различия проявляются, конечно, и на уровне фенотипических свойств — таких, как скорость репликации, время жизни, темп мутирования и т. д.

Отдельный (молекулярный) вид не является, однако, истинным объектом отбора. Уравнение (10) говорит нам, что в результате отбора возникает скорее квазивид, т. е. организованная совокупность видов с определенным распределением вероятностей. Как таковой, он отбирается из всех других распределений. Под давлением отбора численности популяций всех квазивидов, кроме одного, в действительности должны

обращаться в нуль. Квазивид тесно связан с тем, что называется «диким типом» популяции.

Часто предполагают, что дикий тип — это стандартный генотип, который в распределении мутантов соответствует оптимально адаптированному фенотипу. Тот факт, что в случае фагового генома определение уникальной последовательности оказывается возможным, говорит в пользу этой точки зрения о доминировании стандартной копии. Однако более тщательное изучение распределения дикого типа фага Q_{β} (в лаборатории Ч. Вейсмана) [24] показало, что на самом деле лишь малая доля последовательностей идентична той, которая приписывается дикому типу; большинство же их представляет собой распределение копий с одной или несколькими ошибками, и только *среднее* по этому распределению аналогично последовательности дикого типа. Другими словами, стандартные копии могут составлять меньше (а иногда и много меньше) нескольких процентов всей популяции. Однако, хотя подавляющая часть популяции состоит из нестандартных типов, каждый индивидуальный мутант в этом распределении присутствует в очень малом количестве (по сравнению со стандартной копией). Полное распределение характеризуется *усредненной* последовательностью, которая идентична стандартной и, следовательно, определяет дикий тип. Квазивид, строгое определение которого было дано выше, представляет собой такое организованное распределение, характеризующееся одной (или большим числом) *усредненных* последовательностей. Типичные примеры распределения (относящиеся к РНК-содержащему фагу Q_{β}) приведены в табл. 2. Одна уникальная (усредненная) последовательность имеется только в том случае, если копия, идентичная стандарту, явно доминирует, т. е. если она имеет максимальную селективную ценность в пределах распределения. Мутанты, для которых W_{ii} очень близки к максимальным значениям, в среднем будут присутствовать в соответственно больших количествах (см. табл. 2). Их наличие приведет к

Таблица 2

Распространенность, или обилие, стандартной последовательности в распределении дикого типа определяется фактором качества Q_m и ее превосходством σ_m . При данном числе нуклеотидов v_m фактор качества может быть вычислен из среднего качества копирования буквы \bar{q}_m для нуклеотидов, которое определяется конкретным ферментативным механизмом считывания. Значения \bar{q}_m и σ_m определяют также максимальное число нуклеотидов v_{max} , которое не должно превышать стандартной последовательностью, в противном случае квазивидовое распределение становится неустойчивым. Представленные ниже данные относятся к последовательностям РНК, состоящим из 4500 нуклеотидов (фаг Q_β).

Значения, приведенные в темных треугольниках таблицы А, указывают относительные распространенности стандартной последовательности в распределении дикого типа (в процентах) согласно уравнениям (18) и (25). Отрицательные числа означают, что распределение неустойчиво. В светлых треугольниках приведены пороговые значения v_{max} для заданных \bar{q}_m и σ_m . Эти данные демонстрируют чувствительность v_{max} к параметру \bar{q}_m . Для хорошо адаптированного вида значение $1 - \bar{q}_m$ должно быть несколько выше, чем $1/v_m$ (например, $1 - \bar{q}_m = 0,0005$ для $v_m = 4500$ нуклеотидов, причем требуется, чтобы $\sigma_m \geq 10$).

| | | \bar{q}_m -среднее качество копирования буквы | | | |
|--------------------------|------------------|---|------------------|------------------|------------------|
| | | 0,9980 | 0,9990 | 0,9995 | 0,9998 |
| Превосходство σ_m | v_{max} | | | | |
| | 2 | Неустойчиво (<0) | Неустойчиво (<0) | Неустойчиво (<0) | Неустойчиво (<0) |
| | 20 | Неустойчиво (<0) | Неустойчиво (<0) | 5,8% | 37,6% |
| 200 | Неустойчиво (<0) | 0,6% | 10% | 40% | |
| | | 346 | 693 | 1386 | 3466 |
| | | 1498 | 2996 | 5991 | 14979 |
| | | 2649 | 5298 | 10597 | 26492 |

А

В таблице Б дан более реалистичский пример распределения квазивида, состоящего из $1 \cdot 10^9$ индивидов. Последовательности из 4500 нуклеотидов будут соответствовать 13 500 мутантов

Продолжение табл. 2

с одной ошибкой при допущении, что для каждого правильного нуклеотида (А, U, G или С) имеются три неправильные альтернативы. Опыты с РНК-репликазами показывают, однако, что замещения пурин → пурин и пиримидин → пиримидин происходят гораздо чаще, чем любые замещения перекрестного типа пурин ↔ пиримидин. Поэтому — чтобы быть ближе к реальности — мы примем, что для каждого нуклеотида имеется только одна неверная альтернатива. Соответственно кратность любой копии с k ошибками становится равной $\binom{v}{k}$. Вся совокупность 4500 различных мутантов с одной ошибкой была разбита на пять классов в соответствии с их вырожденными (средними) селективными ценностями.

| Класс мутанта | Вырожденность класса | Принятая относительная селективная ценность W_{kk}/W_{mm} | Численность популяции отдельного мутанта (×вырожденность) |
|----------------|----------------------|---|---|
| Без ошибок | 1 | 1 | $8,9 \cdot 10^7$ (×1) |
| Одна ошибка | | | |
| M_{1a} | 1 | 0,99 | $4 \cdot 10^8$ (×1) |
| M_{1b} | 4 | 0,9 | $5 \cdot 10^5$ (×4) |
| M_{1c} | 495 | 0,3 | $6,3 \cdot 10^4$ (×495) |
| M_{1d} | 2000 | 0,1 | $4,9 \cdot 10^4$ (×2000) |
| M_{1e} | 2000 | ~ 0 | $4,3 \cdot 10^4$ (×2000) |
| $\sum M_1$ | 4500 | | $2,2 \cdot 10^8$ (×1) |
| Много ошибок | | | |
| M_2 | ~ 10^7 | ~ 0 | < 30 (× 10^7) |
| $\sum M_{k>1}$ | ~ 2^{4500} | ~ 0 | $6,88 \cdot 10^8$ (×1) |

Б

В классе M_{1a} имеется всего один мутант, очень близкий к стандартному типу. Его селективная ценность W_{kk} отличается от селективной ценности стандартного типа только на 1%. Класс M_{1b} содержит четыре вырожденных мутанта, селективные ценности которых отличаются от стандартной на 10%, тогда как для 495 мутантов класса M_{1c} значения W_{kk} составляют 30% от W_{mm} . Большая группа из 2000 мутантов имеет на порядок меньшее значение W_{kk} и столько же мутантов вообще нежизнеспособно, т. е. они воспроизводятся со скоростью, несравненно меньшей, чем скорость воспроизведения стандарта. Далее, всем копиям с числом ошибок более двух были приписаны значения

Продолжение табл. 2

$W_{kk} \ll W_{mm}$. Хотя это допущение может быть нереалистичным, оно не имеет серьезных последствий для численностей популяций *индивидуальных* последовательностей, которые крайне малы из-за большого числа разнообразных ошибочных копий. Несмотря на это, сумма всех копий с многими ошибками является в этом примере наибольшей группой, за ней следует сумма всех копий с одной ошибкой. С другой стороны, стандартный тип по своей распространенности далеко превосходит все другие индивиды в распределении квазивида.

Был сделан другой вариант расчета, где относительная селективная ценность мутанта 1а с одной ошибкой была повышена с 0,99 до 0,9995. При этом основная часть распределения изменилась очень мало, а численность популяции мутанта 1а возросла до численности стандартного типа (т. е. обе они стали равными $8,4 \cdot 10^7$). Этот пример показывает границу применимости аппроксимаций, содержащихся в уравнениях (18) и (19), которые требуют выполнения неравенства $w_{km} \ll W_{kk} - W_{mm}$. Более строгий расчет дает численность популяции мутанта 1а, составляющую примерно 60% от численности стандартного типа. Даже при малых различиях селективных ценностей стандарт остается явно доминирующим видом. Лишь в том случае, когда мутант с одной ошибкой отличается от стандарта так мало, что выполняется неравенство $W_{mm} - W_{kk} \ll 1/v_m$, он может считаться вырожденным и, следовательно, неотличимым от стандарта.

тому, что последовательность дикого типа будет несколько размыта в определенных положениях. Если два близкородственных мутанта имеют (почти) одинаковые селективные ценности, они оба могут входить в квазивид с (почти) равными статистическими весами. Насколько близки друг к другу должны быть значения W_i для этих двух мутантов, чтобы они стали селекционно неразличимы, зависит от их «степени родства». Для отдаленных родственников соответствие должно быть гораздо более точным, чем для копий с одной ошибкой. Таким образом, специальному классу «обратимых нейтральных» мутантов может быть дано количественное определение. Конечно, существует второй, более широкий класс нейтральных мутантов, относящихся к различным квазивидам, которые вырождены по собственным значениям λ_i . Большинство этих нейтральных мутантов вымирает после своего появления, но некоторые из них могут распространиться в популяции и сосуществовать

с ранее установившимся квазивидом или даже вытеснить его. Это диффузионное распространение нейтральных мутантов можно анализировать только в рамках стохастической теории (см. ниже).

III.5. Реалистические аппроксимации

Выражение в явном виде для собственного значения отобранного квазивида может быть получено с помощью теории возмущений. Результат, соответствующий второму порядку теории возмущений, подобен выражению для W_{\max} , полученному ранее в работе [4] (уравнение II-33а):

$$\lambda_{\max} \approx W_{mm} + \sum_{k \neq m} \frac{\omega_{mk} \omega_{km}}{W_{mm} - W_{kk}}. \quad (13)$$

Здесь индекс m относится к тому (молекулярному) виду, который отличается наибольшей селективной ценностью. Эта аппроксимация годится только в том случае, если ни одно другое W_{kk} не оказывается слишком близким к этому значению и доминирующая копия m может считаться представительницей дикого типа. Табл. 2 показывает, насколько эффективным в действительности является это приближение для любой реально значимой системы. Чем больше информационное содержание, тем меньше индивидуальные значения ω . Тем самым эта аппроксимация выявляет очень важный факт: отбор (под давлением) является исключительно жестким по отношению к дальним родственникам (чем меньше значения ω_{mk} и ω_{km} , тем ближе могут быть W_{kk} и W_{mm} — без каких-либо ограничений относительно m). Однако отбор является мягким по отношению к очень близким родственникам. Они будут присутствовать в распределении и тогда, когда их селективная ценность W_{kk} много меньше W_{mm} (или даже равна нулю). Если вторым слагаемым в уравнении (13) можно пренебречь (см. значения в табл. 2), то экстремальный принцип [уравнение (12)] может быть выражен в следующей форме:

$$W_{mm} > \bar{E}_{k \neq m}, \quad (14)$$

или

$$Q_m > \sigma_m^{-1}, \quad (15)$$

где

$$\bar{E}_{k \neq m} = \sum_{k \neq m} E_k x_k / \sum_{k \neq m} x_k \quad (16)$$

представляет собой среднюю продуктивность всех конкурентов отобранного дикого типа m и

$$\sigma_m = \frac{A_m}{D_m + \bar{E}_{k \neq m}} \quad (17)$$

является параметром превосходства доминирующего вида.

В том же приближении могут быть вычислены относительные стационарные численности популяций, причем для доминирующей копии

$$\bar{x}_m / \sum_{k=1}^N \bar{x}_k = \frac{W_{mm} - \bar{E}_{k \neq m}}{E_m - \bar{E}_{k \neq m}} = \frac{Q_m - \sigma_m^{-1}}{1 - \sigma_m^{-1}}, \quad (18)$$

а для копии с одной ошибкой

$$\bar{x}_{1k} / \bar{x}_m = \frac{w_{km}}{W_{mm} - W_{kk}}. \quad (19)$$

Это выражение остается верным, пока $w_{km} \ll W_{mm} - W_{kk}$ (см. табл. 2).

Приближения более высокого порядка могут быть получены с использованием λ_{\max} независимо от того, в какой форме выражена эта величина. Уравнения (16)—(18) изменяются соответствующим образом.

Основываясь на этих приближениях, мы можем количественно характеризовать дарвиновское поведение макромолекулярных систем. Уравнение (13) показывает, в какой мере динамика отбора определяется индивидуальными свойствами доминирующего (стандартного) вида m , тогда как уравнения (18) и (19) указывают относительный вес стандартного вида и его мутантов (см. также табл. 2). Удивительно, насколько мала на самом деле доля стандартного вида в распределении дикого типа несмотря на то, что его физические параметры почти полностью определяют динамическое поведение распределения. Это дает квазивиду большую адаптационную и эволюционную гибкость и позволяет ему быстро реагировать на изменение среды.

Эти аппроксимации оказываются непригодны только в случае наличия двух или более доминирующих видов (см. табл. 2), которые являются (почти строго) нейтральными мутантами. Однако эти обратимые нейтральные мутанты могут быть объединены в селекционно неотличимый подкласс видов, который будет определять динамическое поведение как один доминирующий вид в соответствии с уравнениями, приведенными выше. Интересно отметить, что внутри квазивида нет отбора против обратимых нейтральных мутантов («обратными» являются такие мутанты, для которых w_{ik} и w_{ki} достаточно велики, чтобы гарантировать их воспроизводимое появление). Эти обратимые нейтральные мутанты, составляющие часть квазивида, необходимо отличать от неродственных нейтральных мутантов [w_{ik} и w_{ki} слишком малы, чтобы обеспечить их воспроизводимое появление согласно уравнению (18)]. Небольшое количество этих неродственных нейтральных мутантов может сосуществовать вследствие случайных флуктуаций [25]. Стохастическая теория показывает, что конкуренция между этими нейтральными видами представляет собой случайный дрейф, который ведет к вымиранию и недетерминированному «выживанию выживающего», а также к некоторому размножению и распространению новых

мутантов — явлению, которое генетики [26] называют недарвиновским поведением (т. е. выживанием без селективного преимущества). Следует понимать, что стохастическое поведение неродственных нейтральных мутантов, хотя оно и не было предсказано Дарвином и его последователями, не противоречит тем свойствам, которые ведут к детерминистическому дарвиновскому поведению. Напротив, отбор обратимых нейтральных мутантов находится в согласии с дарвиновским принципом, если его понимать правильно — как выводимый физический закон, который затем применяется к понятию квазивида.

III.6. Обобщения

При количественном представлении дарвиновской системы уравнениями (2) и (6) мы сделали ряд специальных допущений, касающихся структурных предпосылок и внешних ограничений. В этом разделе мы попытаемся установить, насколько далеко можно обобщить эти допущения без потери характерных особенностей дарвиновского поведения.

1. *Члены кинетических уравнений.* Линейные члены, описывающие автокаталитическое образование и разложение, протекающие как процессы первого порядка, могут быть заменены более общими выражениями. Для ферментативных механизмов самой общей является формула Михаэлиса — Ментен

$$\text{Скорость} \sim \frac{x_i}{1 + a_i x_i} \quad \text{или} \quad \frac{x_i}{1 + \sum_k a_k x_k},$$

которая заменяет простую зависимость от x_i . Было показано [4], что для автокаталитических механизмов этого типа отбор остается эффективным и при малых численностях популяций, т. е. в области, которая является критической для отбора. Насыщение не предотвращает размножения мутантов, имеющих преимущество, но может допустить определенное сосуществование, обусловленное переходом от экспоненциального роста к линейному при механизме без ограничений.

Вообще говоря, если порядок реакции определяется членом x_i^k , то дарвиновское поведение будет наблюдаться при

$$0 < k \leq 1.$$

Для $k = 0$ сосуществование приведет к лимитированному росту. При ограничении постоянных потоков такая ситуация может возникнуть для самовоспроизводящихся видов в том случае, если их скорость образования ограничена постоянной скоростью поступления высокоэнергетического строительного материала. Виды, которые потребляют разные (взаимно независимые) пищевые ресурсы, не будут конкурировать друг с другом. Независимые пищевые ресурсы будут представлять различные «ниши», что приведет к сосуществованию. Многочисленные вариации

различных видов часто обязаны своим существованием подобным механизмам, последствия которых находятся в полном согласии с дарвиновской теорией.

Для показателей $k > 1$ картина более детально проанализирована в части Б данной книги. В этом случае происходит крайне жесткий отбор, причем некоторые его следствия несовместимы со взглядами Дарвина, особенно в связи с проблемами наследования.

2. *Автокатализ*. Как обсуждалось в разд. II, прямое самовоспроизведение является лишь простейшим примером в классе линейных автокаталитических механизмов. В общем случае самовоспроизведение может осуществляться посредством циклического каталитического процесса. Например, фаги, содержащие одноцепочечную РНК, воспроизводятся путем взаимного ин-структирования с помощью двух комплементарных цепей. Кинетические уравнения для такого двухчленного каталитического цикла [4] дают два собственных значения для доминирующего вида:

$$\lambda_{1,2} = -\frac{D_+ + D_-}{2} \pm \sqrt{A_+ A_- Q_+ Q_- + \frac{1}{4} (D_+ - D_-)^2}. \quad (20)$$

Эти выражения основаны на таких же аппроксимациях, как и уравнение (13), — не учитываются члены, соответствующие мутациям. Одно из этих собственных значений, если оно положительно (т. е. $A_+ A_- Q_+ Q_- > D_+ D_-$), заменяет W_{mm} , относящееся к самовоспроизводящейся единице. Здесь кинетические параметры обеих цепей вносят равный вклад (среднее геометрическое) в селективную ценность. Оба они должны быть оптимизированы, чтобы имело место оптимальное функционирование. Везде, где важны фенотипические свойства цепей РНК, эквивалентность легче всего достигается при помощи структурной симметрии (см. тРНК, миди-вариант РНК Q_B [27]). Второе, всегда отрицательное собственное значение относится к процессу «уравновешивания» плюс- и минус-цепей, концентрации которых затем будут находиться в фиксированном отношении. Как только это уравновешивание достигнуто, плюс- и минус-цепи действуют как одна репликативная и конкурирующая единица.

Общее рассмотрение каталитических циклов показало [4, 20—22], что они сходны с простыми самоинструктирующимися системами. n -Членный цикл тоже характеризуется одним положительным собственным значением, которое соответствует селективной ценности отдельной самовоспроизводящейся единицы. В это собственное значение вносят вклад каталитические свойства всех членов, в простейшем случае — в виде среднего геометрического значений AQ , отсюда требование конечности значений AQ для всех членов цикла. Далее, n -членный цикл характеризуется $(n - 1)$ отрицательными собственными значениями, которые отражают внутреннее уравновешивание концентрационных отношений всех членов цикла.

3. *Мутации*. Основным источником мутаций, особенно на ранних стадиях эволюции, — это неверное копирование, т. е. включение

нуклеотида с некоплементарным основанием в процессе репликации. Опыты с фагом Q_{β} показывают, что вероятности ошибочной замены данного пурина или пиримидина его гомологом значительно отличаются от вероятностей ошибочной замены пурина одним из пиримидинов или *vice versa* [24]. При формальном рассмотрении с использованием параметров Q и ω отдельные виды мутаций — такие, как точечная мутация или мутация со сдвигом рамки вследствие делеции или вставки, — не различаются, хотя эти различия, конечно, важны при исследовании функциональных свойств мутантов. При формальном рассмотрении безразличны также причины мутаций — будь то неверное считывание при репликации, химически индуцированные изменения или радиационные повреждения. Иногда бывает необходимо некоторые мутации связывать скорее с членом, описывающим разложение (см. ниже), однако это не влияет на формальную структуру уравнений.

4. *Разложение.* Мутации, обусловленные внешними воздействиями (например, радиацией), следует учитывать в члене, описывающем разложение. Вообще говоря, разложение вида i может привести к появлению другого вида — k , который состоит только из фрагментов i . Такие процессы тоже не изменяют формальной структуры дифференциальных уравнений, если принять во внимание соответствующие уравнения сохранения.

5. *Внешние ограничения.* Явный вид решений селекционных уравнений зависит от наложенных ограничений. Ранее мы подробно обсудили случай постоянной общей организации. Сходные, хотя количественно другие, результаты получаются для ограничения постоянных потоков [4, 28, 29]. Внешние регулируемые параметры могут, конечно, включать в себя временные (например, любые периодические) изменения. Это может вести к механистическим преимуществам, но не изменяет существенных предпосылок и последствий дарвиновского поведения. Принцип экстремума (12) приобретает в этом случае следующий общий вид:

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} \bar{E}(t) dt = \lambda_m. \quad (21)$$

Дальнейшее обобщение принципа экстремума удалось получить Джонсу и др. [21].

Отбор квазивида при наличии конкурентов может рассматриваться также в растущей системе. Будет показано, что нестационарные системы можно рассматривать в общем виде после простой процедуры нормировки.

6. *Стохастическая теория.* Последнее обобщение имеет более принципиальный характер. Детерминистические кинетические уравнения, вообще говоря, описывают среднее поведение ансамблей, состоящих из большого числа индивидов. Элементарные же процессы могут описываться только с помощью вероятности реакций. Игровые модели, разработанные вместе с Винклер-Осватич [18], четко демонстрируют три основных типа поведе-

ния, которые могут рассматриваться стохастической теорией: а) внутренний самоконтроль флуктуаций, как это имеет место в окрестности устойчивого стационарного состояния и, в частности, при термодинамическом равновесии; б) самоусиление флуктуаций, характеризующее неустойчивость, и в) безразличие к флуктуациям, приводящее к случайному дрейфу.

В первом случае флуктуации важны только при малых численностях популяции и означают просто неопределенность в любом мгновенном микросостоянии. Для макросостояний, которые могут быть исследованы экспериментально, они дают ожидаемые значения в пределах вычислимых средних флуктуационных интервалов.

Во втором случае детерминистическое поведение ограничивается ответом на данную флуктуацию. Другими словами, этот детерминизм типа «если — то» предсказывает, что случится, если произойдет определенная флуктуация, и точность предсказания растет с увеличением флуктуации. Возникновение же самой флуктуации, однако, непредсказуемо, и эта микроскопическая неопределенность отображается на макроскопическом уровне через детерминистический в конечном счете процесс усиления. Этот случай особенно важен для дарвиновских систем. Большинство мутаций представляют собой флуктуации первого типа, т. е. они не дают никакого селективного преимущества и не угрожают стабильности дикого типа. После возникновения они исчезают детерминистическим образом, как и в случае равновесия. Однако существуют такие мутации, которые несут с собой селективное преимущество, и они имеют тенденцию усиливаться, т. е. порождают неустойчивость. Достигнут ли они успеха в завоевании доминирующего положения, зависит от величины их селективного преимущества и от размера флуктуации. Отдельная копия имеет довольно высокие шансы на вымирание, прежде чем она успеет воспроизвестись, особенно если ее значение W только ненамного больше, чем $\bar{E}(t)$. Стохастическая теория показывает, что для малых преимуществ, т. е. когда $(W_{m+1} - W_m) \ll W_m$, размер флуктуации должен увеличиться до определенной величины, т. е. до числа копий, которое соответствует величине $W_m / (W_{m+1} - W_m)$, прежде чем вероятность роста флуктуации станет больше, чем $1 - e^{-1}$. Другими словами, только мутанты, характеризующиеся заметными преимуществами, будут детерминистически влиять на эволюционное поведение. Близкие к нейтральным мутанты ведут себя стохастически — почти как истинно нейтральные мутанты — и, следовательно, соответствуют третьей категории игр, которые приводят к случайному дрейфу.

Нейтральные мутанты, возникающие с высокой частотой, конечно, являются частью дайного квазивида и как таковые несколько стабилизируются вследствие конечности темпа мутаций. Тем самым они используют флуктуационный ответ, что было характерно для первой категории. Поскольку подобная связь довольно слаба, относительные численности нейтральных родственников могут испытывать очень большие флуктуации. Неродственные (т. е. очень редкие) нейтральные мутанты, с другой

стороны, можно рассматривать как другой квазивид, собственные значения которого имеют ту же величину, что и у дикого типа. Большая часть этих нейтральных квазивидов должна вымирать, но если им удалось размножиться, то они могут стать устойчивыми и даже вытеснить прежний дикий тип («выживание выживающего»). Этот тип поведения может быть получен только из стохастической теории. Были сделаны соответствующие расчеты для распространения генов в менделевских популяциях, в частности Кимурой [25] и его школой.

Итак, эволюция является детерминистическим процессом в отношении своего прогрессивного характера. Всегда имеет место успешная конкуренция дикого типа с мутантами, имеющими меньше преимущества, сосуществование нейтральных или почти нейтральных близкородственных мутантов и рост численности нового квазивида, имеющего явное преимущество. Однако эволюция не является детерминистической в отношении временной последовательности появления мутантов, а также в отношении генетического дрейфа, вызванного наличием неродственных нейтральных мутантов. В действительности удается закрепиться лишь малой доле этих нейтральных квазивидов. Редкие нейтральные мутации могут, таким образом, быть более важными на поздних, нежели на ранних этапах эволюции, когда еще возможны многие изменения, дающие преимущество и возникающие со сравнительно высокой частотой.

III.7. Информационное содержание квазивида

В нашем подходе к молекулярной эволюции мы еще не встречались в явном виде с понятием генетической информации. Мы определили (молекулярный) вид как репликативную единицу с определенным информационным содержанием, выраженным в специфическом расположении молекулярных символов. Основываясь на сходстве между такими различными расположениями символов, мы пришли к понятию квазивида. Для вывода критериев отбора и эволюции оказалось достаточным, как и в популяционной генетике, отметить индивидуальные различия в генетической информации и связать их с ха-

ракторными динамическими свойствами, выраженными через селективные ценности W_{ii} .

С другой стороны, при таком подходе остаются без ответа вопросы такого типа: «Сколько информации может быть накоплено в данном квазивиде?» или «Где находится предел репродукционной способности квазивида и, следовательно, какова его эволюционная сила?» Поэтому теперь мы сформулируем более точно понятие информации.

В теории связи информационное содержание сообщения, состоящего из v_k символов, выражается следующим образом:

$$I_k = v_k i, \quad (22)$$

где i — среднее информационное содержание одного символа. По Шеннону, i можно связать с распределением вероятностей символов [30, 31]:

$$i = -K \sum_j p_j \ln p_j \quad (23)$$

причем

$$0 < p_j < 1 \quad \text{и} \quad \sum_j p_j = 1. \quad (24)$$

Константу K обычно берут равной $1/\ln 2$ для того, чтобы получить единицу «бит/символ». В общем случае алфавит состоит из нескольких классов символов (например, для нуклеиновых кислот $\lambda = 4$). Практическое использование уравнения (23) ограничивается такими ситуациями, когда известны априорные вероятности всех символов и когда число символов в сообщении достаточно велико, чтобы можно было использовать средние значения. Чтобы учесть все кооперативные эффекты или избыточности, влияющие на распределение вероятностей, может оказаться необходимым знать вероятности всех альтернативных λ^v комбинаций символов.

В данном случае нас интересует не столько статистическая априорная вероятность символа, сколько вероятность того, что данный символ правильно воспроизведется генетическим механизмом, какой бы ни была его сложность на разных уровнях организации. Эта вероятность относится к динамическому процессу переноса информации и поэтому должна определяться экспериментально из кинетических данных (везде, где это возможно; см. ниже). Обозначим вероятности правильного воспроизведения символов через q_j . Сообщение, состоящее из v_i (молекулярных) символов, будет правильно воспроизводиться с вероятностью, или фактором качества:

$$Q_i = \prod_{j=1}^{v_i} q_{ij} \equiv \bar{q}_i^{v_i}. \quad (25)$$

Хотя имеется всего λ классов символов, фактор качества для каждого символа в сообщении может зависеть от конкретного окружения символа, так что определение среднего геометрического \bar{q}_i может потребовать рассмотрения различных кооперативных эффектов, специфически связанных с сообщением « i ». Тем не менее для любого данного сообщения i среднее \bar{q}_i может быть определено, и оказывается, что для данного ферментативного аппарата воспроизведения эти средние можно использовать практически во всех случаях, когда сообщение достаточно велико. Более того, отдельные значения q обычно так близки к единице, что вместо среднего геометрического можно взять среднее арифметическое, т. е.

$$\left(\prod_{j=1}^{v_i} q_{ij} \right)^{1/v_i} \approx \frac{\sum_{j=1}^{v_i} q_{ij}}{v_i}, \quad \text{если } (1 - q_{ij}) \ll 1. \quad (26)$$

Уравнение (25) тогда выражает информационно-теоретический аспект воспроизведения, где \bar{q} , однако, относится к динамической, а не к статической вероятности. На численные значения \bar{q} могут влиять все особенности механизма воспроизведения символов, включая любую статистическую избыточность, которая уменьшает вероятность ошибки в процессе копирования. Природа изобрела искусные копирующие устройства — от комплементарного узнавания оснований до изоощренных ферментативных механизмов контроля и коррекции.

Генетическое воспроизведение — это непрерывно самоповторяющийся процесс, и этим он отличается от простой передачи информации через канал с шумом. При каждой отдельной передаче требуется не только простое восстановление смысла сообщения — при наличии определенной избыточности некоторая доля символов всегда могла бы воспроизводиться неточно. Необходимо также предотвратить постепенное накопление ошибок в последовательных циклах воспроизведения. Другими словами, некоторая доля абсолютно правильных копий дикого типа должна иметь возможность успешно конкурировать со всей совокупностью ошибочных копий. Только таким путем дикий тип может устойчиво сохраняться в распределении, иначе произойдет медленное рассеяние информации (теперь в ее семантическом смысле, т. е. под информацией понимается копия с оптимальным значением W_{ii}), пока в конце концов она не будет нацело потеряна,

Условие, которое гарантирует устойчивое сохранение информации,— это критерий отбора [уравнение (14) или (15)]. В общем виде этот критерий можно записать как

$$Q_m > Q_{\min} = \sigma_m^{-1}, \quad (27)$$

и теперь он применим к любому механизму воспроизведения, даже если σ_m нельзя выразить в такой простой форме, как для линейного механизма [см. уравнение (17)]. Если объединить критерий отбора, полученный из динамической теории, с информационным аспектом, который выражается уравнением (25), то мы получим важное пороговое соотношение для максимального количества информации, которое может содержаться в квазивиде:

$$v_{\max} = \frac{\ln \sigma_m}{1 - \bar{q}_m}. \quad (28)$$

Число молекулярных символов в самовоспроизводящейся единице ограничено, причем предельное значение обратно пропорционально средней вероятности ошибки на символ: $1 - \bar{q}_m$.

Имеется и другой способ формулировки этого важного соотношения. Ожидаемое значение ошибки в последовательности v_m символов $e_m = v_m(1 - \bar{q}_m)$ всегда должно оставаться ниже строго определенного порога

$$e_m < \sigma_m, \quad (29)$$

иначе информация, накопленная в эволюционном процессе, будет потеряна из-за катастрофы ошибок.

В табл. 2 приведены примеры, показывающие значимость уравнения (28). Порог не очень чувствителен к величине функции превосходства, но σ_m должна быть больше единицы (т. е. $\ln \sigma_m > 0$), чтобы гарантировать конечность значений для v_{\max} . На практике (см. ниже) $\ln \sigma_m$ обычно лежит между единицей и десятью. Формула (28) позволяет получить количественную оценку эволюционного потенциала, обеспечиваемого любым конкретным механизмом воспроизведения. Она показывает, например, что

вероятность ошибки, равная 1% (т. е. точность копирования символов равна 99%), достаточна лишь для того, чтобы правильно воспроизводить количество информации, не большее, чем несколько сотен символов (в зависимости от значения $ln\sigma_m$), или что для поддержания информационного содержания такого большого генома, как у *E. coli*, необходимо, чтобы ошибки происходили не чаще, чем одна на 10^6 — 10^7 нуклеотидов. Это соотношение поддается экспериментальной проверке, и ниже мы сообщим результаты соответствующих измерений. Уравнение (28) дает также количественное выражение того, что в популяционной генетике было названо «генетическим грузом» — понятие, важность которого была отмечена очень давно.

Результаты этого раздела можно резюмировать следующим образом. Любой механизм селекционного накопления информации характеризуется верхним пределом для числа букв, располагаемых в определенном порядке. Если этот предел превзойден, то порядок (т. е. эквивалент информации) при последовательных воспроизведениях будет постепенно исчезать. Устойчивость информации эквивалентна внутренней устойчивости квазивида. Она основывается на конкуренции в такой же степени, как и отбор квазивида. Однако два квазивида могут сосуществовать, тогда как превышение порога приводит к полной потере генетической информации. Итак, внутренняя устойчивость распределения квазивидов является более характерным атрибутом дарвиновского поведения, нежели их «борьба за существование».

IV. Порог ошибок и эволюция

IV.1. Моделирование на ЭВМ катастрофы ошибок

Физическое содержание порогового соотношения [уравнение (28)] может быть проиллюстрировано на примере небольшой машинной игры (табл. 3). Цель состоит в том, чтобы построить осмысленное сообщение из более или менее случайной последовательности

Таблица 8

Самокоррекция предложений является результатом эволюционной игры, пример которой дает эта таблица

Контрольное предложение:

TAKE ADVANTAGE OF MISTAKE

Это предложение было выбрано потому, что оно дает «селективно выгодную информацию» в отношении механизма эволюции. Его специальная форма допускает замыкание в цикл, если между отдельными словами вводятся функциональные связи (как это и будет сделано в части Б). При использовании кода, в котором каждая буква (и интервал между словами) представляется квинтетом бинарных символов, информационное содержание последовательности составляет $v_m = 125$ бит, допуская примерно $4 \cdot 10^{37}$ альтернатив. Это исключает чисто случайное появление информации. Последовательности букв, приведенные в этой таблице для ряда генераций, были выбраны как типичные для всей популяции последовательностей в памяти ЭВМ.

Исходная последовательность: ВАК GEVLNT GUPIF LESTKKM

Качество копирования буквы \bar{q}_m : 0,995

Селективное преимущество на бит: 10

| | |
|--------------|---------------------------|
| Генерация 1 | RAK GEVNNT GUPQF KESTKKM |
| Генерация 5 | NAK AEZ,NS GEPOF MESTMKU |
| Генерация 10 | VAKF ADV!NT.GE OF MISD!KE |
| Генерация 16 | TAKE ADVANTAGE OF MISTAKE |

(цель достигнута)

Первый пример показывает, что эволюция очень эффективна вблизи критического значения $(1 - \bar{q}) \approx 1/v_m$, которое при $v_m = 125$ дает $\bar{q}_m = 0,992$. Контрольное предложение обычно удается составить через 20 (± 6) генераций начиная со случайной последовательности букв для любого значения \bar{q} между 0,995 и 0,990. Эта высокая эффективность вблизи порога становится еще более очевидной, если сравнить эволюционный прогресс, достигнутый к данной генерации, для разных значений \bar{q}_m .

Продолжение табл. 3

Исходная последовательность: ВАК GEVLNT GUPIF LESTKQM
 Селективное преимущество на бит: 10

| \bar{q}_m | Лучшая последовательность через 11 генераций | Число ошибок |
|-------------|--|--------------|
| 0,999 | LAKD AEV,NTAGU AF KISTQKM | 9 |
| 0,995 | ТАКЕ ADVINT GE OF MISTAKE | 2 |
| 0,990 | ТАКЕВADVINTAGE OF MISXAKE | 3 |
| 0,985 | VATA ADBKMDI DHOD ?CSYBKE | 18 |

Аналогичная картина наблюдается для распада информации при $v_m > v_{\max}$. При вероятности ошибки $(1 - \bar{q}_m) = 1,5\%$ селективное преимущество на бит, равное 2,5, соответствует значению σ_m около 5. При таких условиях информация становится неустойчивой. Однако для малых селективных ценностей (как в данном примере) распад (или накопление) информации является сравнительно медленным процессом.

Исходное предложение: TAKE ADVANTAGE OF MISTAKE
 Качество копирования буквы \bar{q}_m : 0,985
 Селективное преимущество на бит: 2,5

| Число генераций | Лучшее предложение | Число ошибок |
|-----------------|---------------------------|--------------|
| 1 | TAKE ADVANTAGE OF MISTAKE | 0 |
| 5 | TAKF !DVALTAGE OF MISTAKE | 3 |
| 10 | TALF ADVALTACE OF MISTAKI | 5 |
| 20 | DAKE ADUAVEAGE OF MJUTAKE | 6 |
| 40 | TAKE ADVONTQCU OF MFST!ME | 7 |
| 71 | ТАКЕВ ?VALTAGI LV MIST!KE | 8 |

71 генерация для $\bar{q}_m = 0,97$

?АМЕВАДТАМОАСФНQEВА!STBМF 18

Сравнение двух последних строчек (относящихся к генерации 71) показывает, что при вероятности ошибки 3% ($\bar{q}_m = 0,97$) распад идет намного быстрее.

В другом примере (селективное преимущество на бит = 10) при $\bar{q}_m = 0,985$ порог еще не пройден, так что информация устойчива, как можно видеть из таблицы, приведенной ниже, где процесс начинается с правильного предложения.

Исходное предложение: TAKE ADVANTAGE OF MISTAKE

Качество копирования буквы \bar{q}_m : 0,985

Селективное преимущество на бит: 10

Продолжение табл. 3

| 71 генерация Выборка из 8 типичных предложений | Число ошибок |
|---|--------------|
| TAKE ADVANTAGE OF MISTAKE | 0 |
| TAKE ADVANTAGIPOF MISTAKE | 2 |
| TAKE ADVANTAGE OF MISTAKE | 0 |
| TBKE IDVANTAGE OF MISTAKE | 2 |
| SAKE ADVANTAGE OF MGSTAME | 3 |
| TAOE ADVANVAGE OF MISTAKE | 2 |
| TAKE ADVAVTAGE OF MISTAKE | 1 |
| TAKE ,DVANTAGE OF MISTAKE | 1 |

букв. Для этой цели в ЭВМ вначале закладывают набор из N случайных последовательностей и следующую программу:

- удалять из памяти любую последовательность после истечения определенного (среднего) времени жизни;
- воспроизводить любую последовательность, которая хранится в памяти, с характеристической скоростью;
- вносить случайные ошибки в репродуцируемые копии — опять же с некоторой выбранной средней скоростью замещения на символ.

Средние скорости для а) и б) подобраны таким образом, чтобы в памяти постоянно хранилось N предложений, причем каждое предложение имеет конечное время жизни. Следовательно, любая информация, приобретенная во время игры, может сохраниться лишь посредством верного воспроизведения имеющихся последовательностей. Увеличение информации,

с другой стороны, должно быть результатом селективной оценки различных мутантных последовательностей в соответствии с их смыслом или, лучше, в соответствии с их близостью к какому-либо смыслу. Эта оценка должна осуществляться посредством внутренних, зависящих от смысла свойств последовательностей,— свойств, которые должны влиять на скорости их воспроизведения и (или) удаления, как это отражается в селективной ценности.

При естественном отборе целью эволюции всегда является создание генотипа, соответствующего фенотипу с оптимальной селективной ценностью. Оценка, следовательно, производится через фенотипические, т. е. физические и химические свойства данного индивида, которые определяют скорость и качество воспроизведения, а также время жизни генотипа по отношению к среднему времени жизни его конкурентов, присутствующих в популяции. Подобным же образом и человеческая память способна связать с любой последовательностью букв репродуктивную ценность, которая связана с ее смыслом. Итак, в данной игре можно было бы оценивать любую последовательность букв в соответствии с ее смыслом, используя наше воображение. Конечно, ЭВМ может делать это только с помощью какой-либо конкретной программы, в которой эволюционирующие предложения сравниваются с одним или несколькими контрольными предложениями — «целями».

Допустим теперь, что каждая последовательность воспроизводится со скоростью, которая зависит от числа символов, совпадающих с символами (в соответствующих положениях) осмысленного целевого предложения. Используя бинарное кодирование, можно постулировать, что с каждым битом ближе к цели мы увеличиваем скорость воспроизведения, умножая ее на определенный множитель, а с каждым битом дальше от цели мы соответственно уменьшаем ее.

Информация о возможных целевых предложениях, которая должна вводиться в ЭВМ с самого начала, используется только для того, чтобы задать ЭВМ

схему оценки. Процедуру оценки легко можно было бы сделать более изощренной, чтобы она больше напоминала нашу мысленную оценку «смысла», вплоть до того, что конкретная цель, к которой в конце концов приводит оценка, не будет определена заранее. Однако эти детали механизма имитации вовсе не так уж важны для иллюстрации физического смысла пороговой формулы, поскольку мы можем дать определение внутренней процедуры оценки молекулярной эволюции — примером служит игровая модель, демонстрирующая эволюцию молекул тРНК [18]. Результаты машинного эксперимента, согласно табл. 3, можно резюмировать следующим образом.

При высоком качестве воспроизведения символов, например для предложения из 100 битов со средней вероятностью ошибки на символ

$$1 - \bar{q}_m \ll 10^{-2},$$

скорость эволюционного процесса очень мала даже при больших значениях σ_m , т. е. при больших селективных преимуществах. Максимальная скорость достигается, если выбрать средние вероятности ошибок $(1 - \bar{q}_m)$ того же порядка величины, что и $1/\nu_m$ (например, для 100 битов $\bar{q} \approx 0,99$).

При достаточно больших значениях σ_m (> 3) целевое предложение получается после такого числа генераций, которое по порядку величины соответствует эволюционному расстоянию между целью и начальной (более или менее случайной) последовательностью (например, 100 генераций). Однако, как только порог для $1 - \bar{q}_m = \ln \sigma_m / \nu_m$ будет превзойден, не может быть накоплено никакой новой информации, каким бы большим ни было выбранное селективное преимущество, приходящееся на один бит. Если начинать игру с почти правильного предложения, то информация распадается на случайную смесь букв, а не эволюционирует к правильной копии. Порог очень резкий, но скорость распада варьирует вблизи порога. Пороговое значение очень слабо зависит от величины σ_m , если только этот параметр не становится очень близким к единице. Превосходство

σ_m вычисляется из относительных селективных преимуществ, и, следовательно, требуются некоторые сведения о характере распределения ошибок (по отношению к соответствующей оптимальной копии). Это распределение, конечно, зависит от величин селективных преимуществ. Результаты машинного эксперимента очень близки к ожидаемому распределению ошибок, которое вблизи критического значения $1 - \bar{q}_m \approx 1/v_m$ (причем $\ln \sigma_m \approx 1$) приписывает почти одинаковые веса оптимальной копии, всем копиям с одной ошибкой (относительно оптимальной) и сумме всех копий с большим числом ошибок (где доминируют копии с двумя ошибками, а вклад копий с большим числом ошибок резко падает).

Для меньших селективных преимуществ (например, $W_{mm} - W_{kk} < 3$) это распределение сдвигается в пользу ошибочных копий и в ущерб (относительному) оптимуму, который для $\ln \sigma_m = 1$ уже составляет менее 10% от общего числа копий.

IV.2. Экспериментальные исследования с РНК-содержащими фагами

Насколько тривиальной может показаться эта игра — после того, как ее результаты рационализированы, — настолько уместна она в природе, определяя информацию, приобретенную на различных уровнях доклеточной и клеточной самоорганизации. Вейсман с сотрудниками провели эксперимент с фагом Q_β , результаты которого почти точно соответствуют исходу этой игры [32, 33].

С помощью направленного сайт-специфического мутагенеза была получена ошибочная копия фагового генома. Эта процедура заключается в синтезе *in vitro* минус-цепи фаговой РНК, содержащей в положении 39 от 5'-конца мутантный аналог основания N^4 -гидрокси-СМР вместо исходного УМР. Используя эту цепь в качестве матрицы и полимеризующий фермент Q_β -репликазы, можно было получить инфекционную плюс-цепь, в которой в положении 40 от 3'-конца (оно соответствует положению 39 от 5'-

конца в минус-цепи и локализовано в экстрацистронной области) остаток А замещен на G. Затем этой мутантной плюс-цепью были инфицированы сферопласты *E. coli* и получен урожай мутантных фаговых частиц, которых можно было изолировать из отдельных бляшек. Эксперименты с серийными переносами *in vivo* (инфекция *E. coli* целыми фаговыми частицами) и *in vitro* (кинетические исследования на изолированных цепях РНК с использованием Q_{β} -репликазы) позволили определить параметры скорости воспроизведения как для фага дикого типа, так и для мутанта-40, а также установить распределения их сателлитов. Метод «отпечатков пальцев» и анализ нуклеотидных последовательностей для ряда последовательных генераций показали, что в мутантной популяции происходят изменения, вызванные образованием ревертантов. Опыты с различными начальными распределениями фага дикого типа и мутанта обнаружили, что в процессе естественного отбора идет конкуренция между единственным доминирующим типом и мутантами. Количественные расчеты указывают на зависимость от данного селективного преимущества и от параметров распределения мутантной популяции. Селективное преимущество дикого типа по сравнению с данным конкретным мутантом равно

$$W_{\text{дикий тип}} - W_{\text{мутант}} \approx 2 \div 4,$$

в то время как скорость замещения равна примерно

$$1 - q \approx 3 \cdot 10^{-4}.$$

Это значение q определяется скоростью образования ревертантов и, следовательно, относится к следующим (комплементарным) замещениям:



Согласно уравнению (20), факторы качества для плюс- и минус-цепей вносят равный вклад в точность воспроизведения. Замещения $G \rightarrow A$ и $C \rightarrow U$, таким образом, эквивалентны. Возможно, они не очень сильно отличаются от замещений $A \rightarrow G$ и

$U \rightarrow C$, что обусловлено главным образом сходством «качаний» для взаимодействий GU и UG .

Поскольку реплицирующий фермент требует разворачивания матрицы, чтобы она могла связаться с активным центром, значения q , вероятно, не зависят от вторичной или третичной структуры матрицы. Опыты *in vitro* с миди-вариантом РНК Q_{β} [27] дают скорости замещения $C \rightarrow U$, которые согласуются со значениями, приведенными выше. Замещения пурин \rightarrow пиримидин и пиримидин \rightarrow пурин, по видимому, происходят гораздо реже и поэтому не вносят существенного вклада в величину \bar{q} .

Определить σ_m труднее, так как эта величина зависит от $\bar{E}_{k \neq m}$. Прежде всего обратим внимание на то, что модификация экстрацистронной области, которая не влияет ни на один белок, кодируемый фаговой РНК, существенно изменяет скорость репликации. Спигелман первый подчеркнул важность *фенотипических* свойств молекулы фаговой РНК в связи с механизмом репликации и отбора. Значение σ_m , приведенное выше, относится к определенному мутанту и его сателлитам. Другие мутации могут влиять на третичную структуру РНК Q_{β} иным образом, и, следовательно, соответствующие мутанты могут иметь другие скорости репликации. Более того, мутации в интрацистронных областях могут быть летальными и поэтому вообще не давать вклада в $\bar{E}_{k \neq m}$. Если считать измеренное значение характерным для большей части мутаций, то тогда для максимального информационного содержания получится величина, лишь слегка превышающая действительный размер генома фага Q_{β} , который составляет около 4500 нуклеотидов. Такое близкое соответствие кажется подозрительным, и мы сделали определенные оговорки. Однако эти оговорки относятся в основном к значению σ_m , которое входит в соответствующую формулу только в виде логарифма. При больших σ_m тоже был бы получен приемлемый предел для v_{\max} . Итак, приведенное значение в конечном счете может быть не так уж далеко от действительности.

Другая серия экспериментов, выполненных Вейсманом с сотрудниками [34], показывает, что в распределении дикого типа присутствует сравнительно небольшое количество стандартного фага. Эти данные заставляют думать, что $\sigma_m^{-1} \approx Q_m$ и что действительное число нуклеотидов на самом деле очень близко к пороговому значению v_{\max} [см. уравнение (18)].

Миди-вариант, использованный в экспериментах Миллса, Спигелмана и др. [7], состоит всего из 218 нуклеотидов и, следовательно, не так хорошо адаптирован к изменениям среды, как РНК Q_β . Он, конечно, оптимально адаптирован к специальной среде — «стандартной реакционной смеси», использованной в экспериментах *in vitro* (в которых не требовалось, чтобы молекулы РНК были инфекционны). Однако его реакция на изменения среды, например на добавление ингибитора репликации — этидиумбромиды, — оказалась довольно медленной. Мутант, полученный после 20 переносов, в каждом из которых могло происходить усиление в 10^5 раз, отличался от миди-варианта дикого типа всего в трех положениях и в новых условиях среды проявлял сравнительно небольшое селективное преимущество. Причиной медленной реакции является то, что 218 нуклеотидам при среднем качестве копирования отдельной буквы 0,9995 соответствуют значения Q , близкие к единице, так что последовательности дикого типа реплицируются очень точно — с образованием лишь сравнительно небольшого количества мутантов ($\leq 10\%$) из распределения ошибочных копий.

Замечательным результатом этих исследований с теоретической точки зрения является вовсе не тот факт, что пороговое соотношение выполняется как неравенство. Поскольку его вывод основывается на очень общих рассуждениях, любое серьезное несоответствие означало бы, что мы плохо понимаем, что такое дарвиновские системы. Для таких несоответствий нет никаких причин, потому что мы знаем достаточно хорошо молекулярный механизм репликации в таких «прозрачных» системах, как фаг Q_β . Поистине удивительным результатом является то,

что в действительности значение ν остается не только ниже порога ν_{\max} , но и почти совпадает с ним. Число нуклеотидов вполне могло бы лимитироваться не качеством копирования символа \bar{q} , а другими факторами, что привело бы к значениям ν , лежащим намного ниже порога, допускающего воспроизводимое накопление информации.

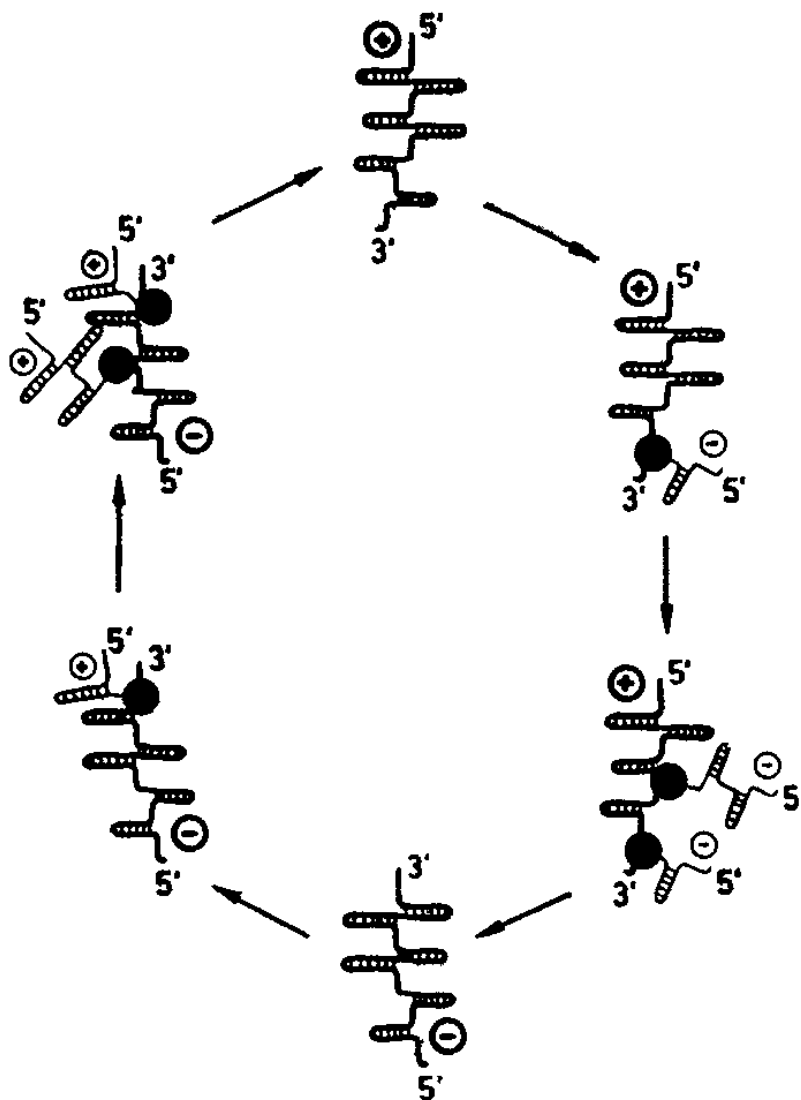


Рис. 11. Цикл репликации, который встречается у РНК-содержащих фагов, ведет к образованию одноцепочечных единиц РНК с высокоспецифичными вторичной и третичной структурами, которые играют роль фенотипических мишеней [11, 34]. Образование полных или частичных дуплексов из плюс- и минус-цепей (так называемых пар Хофшнайдера или Франклина соответственно) предотвращается немедленным закручиванием на себя новосинтезированной цепи. Используя миди-вариант фага Q_{β} , Спигелману и др. [35] удалось доказать существование эффектов ингибирования, которые обусловлены образованием дуплексов. Репликация одноцепочечных молекул основана на эффективном взаимодействии репликазы с плюс- и минус-цепями, что требует наличия определенной симметрии третичной структуры фенотипически важных участков.

Итак, мы вынуждены заключить, что эти РНК-содержащие фаги в ходе своей эволюции действительно стремились накопить максимально возможное количество информации, используя при этом также большие экстрацистронные (но фенотипически активные) области генома. Этот факт не исключает того, что при других (например, искусственных) условиях соревнование могли бы «выиграть» молекулы РНК гораздо меньшей длины — такие, как упомянутый выше миди-вариант, или что при других природных условиях существуют гораздо менее сложные жизнеспособные фаги. Еще важнее осознание того, что в природе не существует фагов с (одноцепочечной) РНК, геном которых был бы больше, чем примерно 10 000 нуклеотидов. Это заставляет полагать, что ферментативный механизм репликации РНК, особенно по отношению к дискриминации оснований А и G или U и C, достиг своего оптимума и не может быть далее улучшен. Никакая одноцепочечная РНК не способна поддерживать правильное воспроизведение информации, превышающей то количество, которое по порядку величины эквивалентно 1000—10 000 нуклеотидов (точное значение зависит от σ_m).

Конечно, с химической точки зрения могут существовать молекулы большей длины, но они не имеют эволюционной ценности. Более того, уравнение (20) предписывает, что требованиям селективного сохранения информации должны удовлетворять как плюс-, так и минус-цепи, хотя только одна из цепей обязана нести генетическую информацию, транслируемую аппаратом хозяина. Эти выводы относятся лишь к молекулам РНК, которые во время своей репликации действуют как одноцепочечные матрицы и для которых был предложен механизм репликации, изображенный на рис. 11 [11].

IV.3. Репликация ДНК

В случае двухцепочечных молекул, особенно ДНК, мы сталкиваемся с совершенно другой ситуацией. Такие молекулы обычно редуцируются в форме

двухцепочечных единиц, т. е. они могут считаться истинно самовоспроизводящими, по крайней мере в феноменологическом смысле (хотя передача информации и в этом случае основана на комплементарности нуклеотидов).

Рассмотрим вкратце, что известно [10] о воспроизведении таких двухцепочечных молекул ДНК (см. рис. 12).

1. Репликация является полуконсервативным процессом. Двухцепочечная ДНК копируется с образованием двух (совершенно идентичных) дуплексов, каждый из которых содержит одну родительскую цепь.

2. Репликация начинается в определенной точке роста и может распространяться в обоих направлениях. Раскручиванию двойной спирали способствуют так называемые расплетающие белки; некоторые из них были выделены и идентифицированы. Они увеличивают скорость раскручивания на три порядка, в результате репликационная вилка движется сравнительно быстро. В то же время необходимо снимать напряжение кручения, вызванное раскручиванием части молекулы. Существует предположение, что вращение участков молекулы вокруг фосфодиэфирной связи может происходить благодаря разрывам и воссоединениям цепи под действием эндонуклеаз и лигаз.

3. Репликация *каждой* из двух цепей происходит путем присоединения нуклеотидов в направлении $5' \rightarrow 3'$. ДНК-полимеразы способны присоединять мономеры только таким уникальным векторным путем, и этот процесс не может протекать на обеих цепях одновременно. Электронная микроскопия с разрешением около 100 Å выявила наличие одноцепочечных участков только на одной стороне репликационной вилки; это позволяет думать, что построение второй цепи начинается лишь после того, как возникает значительный промежуток, достаточный для продвижения в направлении $5' \rightarrow 3'$.

4. Репликация осуществляется короткими дискретными импульсами. У прокариот фрагменты,

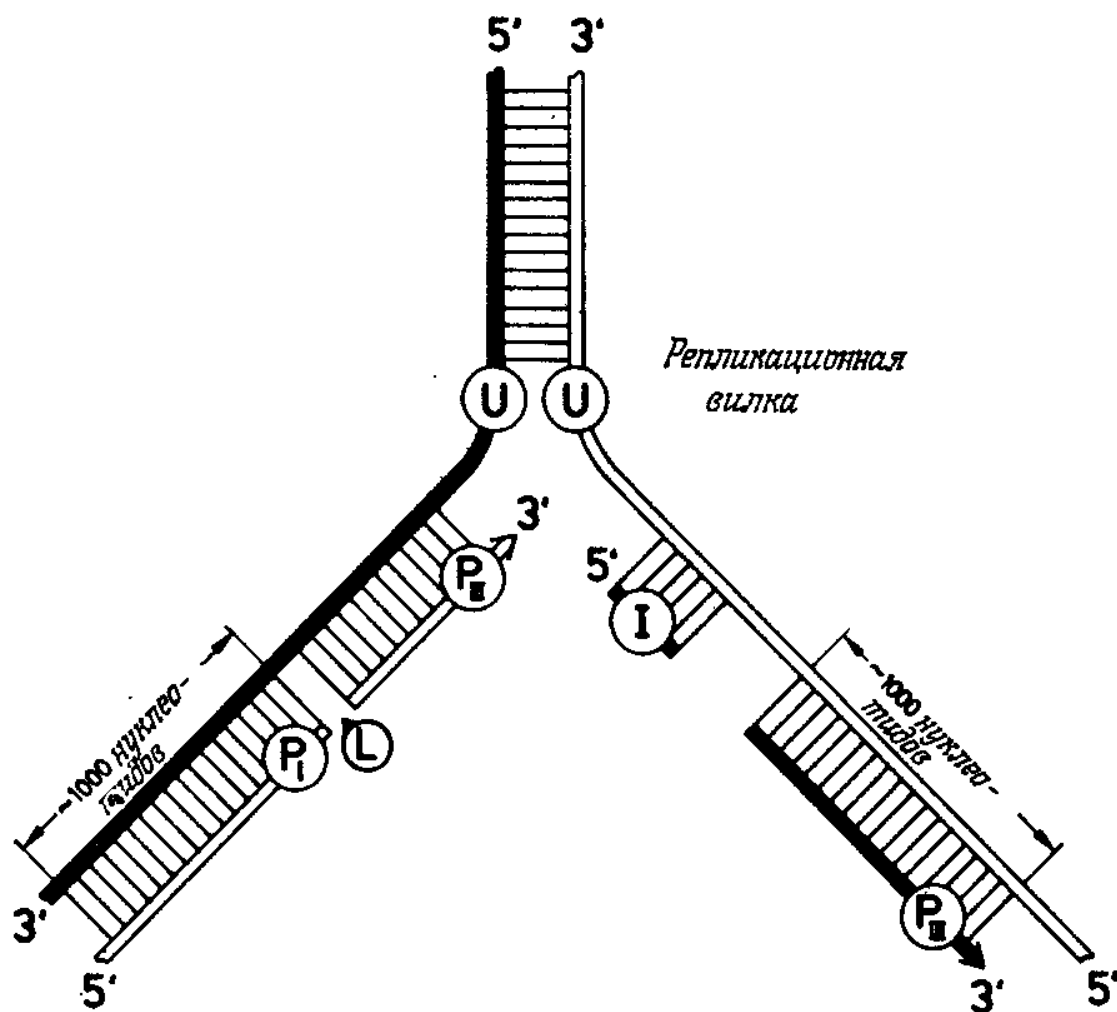


Рис. 12. Полуконсервативная репликация двухцепочечной ДНК — это крайне сложный процесс, состоящий из многих этапов, на которых протекают различные реакции и осуществляется контроль. Наиболее важные из этих этапов указаны на рисунке [10]. Обе дочерние цепи полимеризуются в направлении $5' \rightarrow 3'$. Раскручивание родительской двойной спирали производится расплетающим белком (U). Синтез новых фрагментов инициируется РНК-затравками, которые образуются с помощью ферментативной системы (I) и затем гидролизуются, по-видимому, при участии нуклеазной активности полимеразы-I (P_I). Считают, что рост цепи до образования фрагментов длиной 1000—2000 нуклеотидов в основном производится комплексом полимеразы-III (P_{III}). Эти новообразованные так называемые фрагменты Оказаки соединяются друг с другом с помощью лигазы (L). Неправильно спаренные нуклеотиды на 3'-конце фрагментов (и только они) вырезаются $3' \rightarrow 5'$ -экзонуклеазой, по всей вероятности, комплексом полимеразы-I (P_I), активность которого в основном связана с репарацией и заполнением пробелов. Другие процессы — например, репарация с помощью $5' \rightarrow 3'$ -экзонуклеазы, которая может удалять целые фрагменты ДНК, — на этой схеме не указаны, потому что они, по-видимому, не играют большой роли в синтезе новых цепей.

образующиеся за один импульс, имеют длину около 1000—2000 нуклеотидов. Синтез их инициируется заправкой, которой служат очень короткие фрагменты РНК. Участки реплицированной ДНК, располагающиеся вдоль обеих цепей позади репликационной вилки, соединяются затем друг с другом с помощью лигаз.

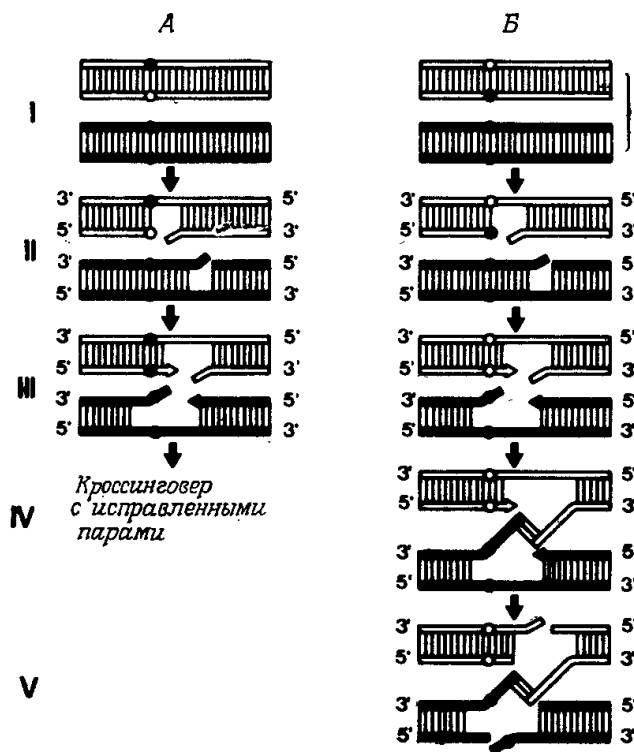
5. Различные функции, необходимые для репликации ДНК, были идентифицированы путем выделения отдельных ферментов и определения их активности. В частности, было охарактеризовано несколько полимерных комплексов (I, II, III), которые выполняют как полимеризующие функции, так и некоторые функции деградации. Здесь особенно интересна $3' \rightarrow 5'$ -эксонуклеазная активность полимеразы-I. Она обеспечивает избирательное отщепление неспаренного нуклеотида на $3'$ -конце растущей цепи. Поскольку рост цепи идет только в направлении $5' \rightarrow 3'$, эта эксонуклеазная функция позволяет корректировать новосинтезированные фрагменты. Ее максимальная активность составляет около 2% полимеразной активности. $3' \rightarrow 5'$ -эксонуклеазу следует отличать от $5' \rightarrow 3'$ -эксонуклеазы, которая также входит в состав комплекса ДНК-полимеразы-I и, вероятно, участвует в репарации с выщеплением. Она действует только на $5'$ -конец и расщепляет фосфодиэфирную связь в спаренной области, по-видимому, на расстоянии до 10 нуклеотидов от $5'$ -конца. Эта эксонуклеаза, следовательно, может выщеплять олигонуклеотиды, тогда как корректирующий $3' \rightarrow 5'$ -фермент удаляет только отдельные неспаренные нуклеотиды на конце растущей цепи.

Теперь мы можем понять, в чем состоит существенное различие между процессами репликации РНК и ДНК, которое выражается в различии средних факторов качества копирования символов для этих двух процессов. В случае репликации РНК точность передачи информации должна обеспечиваться в непрерывном процессе полимеризации. Как бы ни решала эту проблему РНК-репликаза, она достигает, по-видимому, предельного значения \bar{q} — между 0,9990 и 0,9999. Примерно та же точность могла бы

достигаться любым непрерывным механизмом полимеризации ДНК.

Исследование *in vitro* мутантных фаговых ДНК-полимераз, не обладающих 3' → 5'-экзонуклеазной активностью, показало, что эти ферменты сравнительно часто ошибаются — с частотой примерно 1 на каждую тысячу нуклеотидов. Сходные результаты получены для очищенной ДНК-полимеразы из *вируса птичьего миеобластога*. Однако есть данные и о меньших вероятностях ошибок. Например, для низкомолекулярных эукариотических ДНК-полимераз, не обладающих корректирующей экзонуклеазной активностью, были получены значения на порядок меньшие, чем в случаях, приведенных выше (одна ошибка на каждые 5—10 тысяч нуклеотидов) [36—39]. Появление фрагмента ДНК длиной 1000—2000 нуклеотидов во время полимеризации ДНК (в прокариотических клетках) может быть прямо связано с ограниченной точностью полимеразной функции. По-видимому, полимеразы не могут легко удлинять произведенный ею неправильно спаренный конец ([10], стр. 88) — правда, это наблюдалось в отсутствие экзонуклеаз. В то же время 3' → 5'-экзонуклеаза, если она имеется, будет опознавать несоответствие и вырезать неверный нуклеотид. Нет никаких оснований предполагать, что оптимальная разрешающая способность на этом этапе сильно отличается от той, которая характеризует процесс полимеризации. Итак, коррекция может снизить вероятность ошибки (в лучшем случае) еще на три порядка. Коррекция ошибок, с другой стороны, не может быть отложена до более поздней стадии, т. е. до завершения синтеза обеих цепей. Хотя системы репарации, использующие 5' → 3'-нуклеазные активности, существуют, они не могут установить, которая из двух цепей содержит неправильный член в неправильной паре [36]. Были предложены [40] и проверены экспериментально [41, 42] более детальные механизмы кинетической коррекции (см. обзор [43]).

Итак, можно сделать следующий вывод. Оптимальное среднее качество копирования символа для

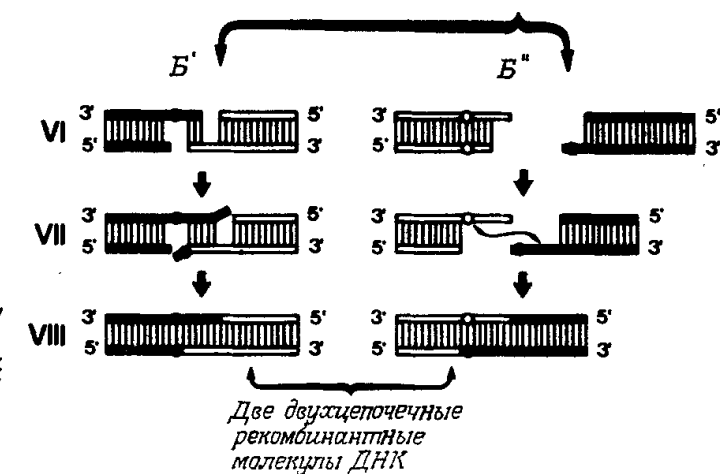


Две гомологичные молекулы ДНК

Эндонуклеаза делает надрез в одной цепи каждой молекулы, 5' → 3'-экзонуклеаза устраняет неверно спаренный концевой нуклеотид

ДНК-полимераза осуществляет синтез на одной стороне каждой разрезанной цепи, вытесняя два одноцепочечных хвоста

Если две такие одноцепочечные области являются гомологичными, они могут ассоциировать с образованием короткого двухцепочечного участка. Эндонуклеаза разрезает другие цепи с образованием одной рекомбинантной молекулы и двух молекулярных фрагментов с перекрывающимися концевыми последовательностями.



Экзонуклеаза опитцельяет нуклеотиды от одной из цепей в каждой из двух половин молекулы, обдлачая гомологичные области

Спаривание оснований дает двухцепочечный рекомбинант, пробелы заполняются с помощью ДНК-полимеразы и лигазы

Рис. 13. Генетическая рекомбинация позволяет обнаруживать ошибки в завершенных двухцепочечных ДНК. Эта модель первоначально была предложена для объяснения механизма кроссинговера. Ее можно использовать также для анализа механизма коррекции ошибок. Символ ● означает генетически правильный нуклеотид, символ ○ — ошибочный. Соответственно символ $\begin{matrix} \bullet \\ | \\ \bullet \end{matrix}$ всегда обозначает правильную комплементарную пару, символ $\begin{matrix} \bullet \\ | \\ \circ \end{matrix}$ — неправильную (некомплементарную) пару, а символ $\begin{matrix} \circ \\ | \\ \circ \end{matrix}$ — комплементарную, но ошибочную пару нуклеотидов, независимо от того, какой из четырех нуклеотидов, здесь затронут. Предположим, что надрезание цепи инициируется наличием некомплементарной пары (этап II). Тогда 3' → 5'-экзонуклеаза исправит ошибку: в 50% случаев образуется правильная комплементарная пара $\begin{matrix} \bullet \\ | \\ \bullet \end{matrix} \rightarrow \begin{matrix} \bullet \\ | \\ \bullet \end{matrix}$ (этап III), в то время как в других 50% случаев — неправильная комплементарная пара $\begin{matrix} \circ \\ | \\ \bullet \end{matrix} \rightarrow \begin{matrix} \circ \\ | \\ \circ \end{matrix}$. Неправильная (хотя и комплементарная) пара, однако, не фиксируется, как при простом механизме репарации. Рекомбинация с правильной копией (этапы VI, VII) восстанавливает исходную ситуацию (этапы I и VIII), при которой неправильным нуклеотидом занято только одно из четырех гомологичных положений. Следовательно, многократное повторение этого процесса может вести к постоянному снижению доли ошибок, а не к их 50%-ной фиксации. Эта схема показывает, что коррекция ошибок в завершенных цепях связана с кроссинговером. Данный механизм (из работы [45]) может быть реализован с помощью известных функций ДНК-полимеразы, что не исключает существования других гипотетических и даже еще более эффективных механизмов. Для полного понимания проблемы точности, связанной с необходимостью рассмотрения процессов вегетативного размножения, необходимы более детальные сведения о механизме, которые пока отсутствуют.

репликации ДНК достигает значения 0,999999 или несколько выше, что позволяет накапливать информацию вплоть до верхнего предела, эквивалентного 1—10 миллионам нуклеотидов (в зависимости от величины σ_m). Приятно отметить, что эта оценка совпадает с известными размерами геномов прокариотических клеток (например, *E. coli* — $4 \cdot 10^6$ пар оснований). И опять размер геномов любого организма вовсе не обязан быть равным этой предельной величине. Действительные размеры генома могут лимитироваться другими факторами, связанными, например, с упаковкой молекулы в случае ДНК-содержащих фагов, и т. д. Таким образом, как и в случае РНК-содержащих фагов могут реализоваться любые размеры генома ниже порога.

Для содержания генетической информации в прокариотической клетке существует верхний предел. Любой выход за пределы информационной емкости одноцепочечной молекулы (10^4 нуклеотидов) требует нового механизма репликации с участием двухцепочечных матриц и корректирующих ферментов. Точно так же новый предел (около 10^7 нуклеотидов), установленный механизмом репликации ДНК в прокариотической клетке, не мог быть превзойден, пока не появился новый механизм для дальнейшего уменьшения ошибок. Такой механизм, а именно генетическая рекомбинация, был изобретен природой на прокариотическом уровне. Однако потребовалось от двух до трех миллиардов (10^9) лет, прежде чем он достиг совершенства, чтобы вызвать новое увеличение количества генетической информации отдельных индивидов.

Процесс генетической рекомбинации, используемый всеми эукариотическими клетками, требует, чтобы два аллеля в их гомологичных положениях были идентичны. Поскольку вероятность ошибки для ферментативной репликации ДНК составляет менее 10^{-6} , нескорректированные ошибки крайне редки, и в четырех эквивалентных сайтах двух аллелей не может встретиться более одной ошибки. Следовательно, существует дальнейшая возможность обнару-

жения и исправления таких ошибок у рекомбинантов, даже если они содержатся в дуплексах, репликация которых была завершена ранее. Одна из возможных схем изображена на рис. 13. Однако механизм рекомбинации в настоящее время неизвестен достаточно подробно; не ясно также, сколько этапов в конечном счете ответственно за дальнейшее снижение вероятности ошибки. Известно лишь, что такое снижение было достигнуто, как показывает анализ эволюционного древа, и что это является важной предпосылкой увеличения генетической информационной емкости вплоть до уровня, характерного для человека.

IV.4. Первые репликативные единицы

Для обсуждения проблемы возникновения биологической информации мы должны начать с другого конца эволюционной шкалы и проанализировать те механизмы, которые привели к первым воспроизводимым генетическим структурам. Физические свойства, внутренне присущие нуклеотидам, дают возможность дискриминировать комплементарные и некомплементарные нуклеотиды с фактором качества \bar{q} , не превосходящим 0,90—0,99. Более детальный анализ, основанный на кинетических и равновесных исследованиях кооперативных взаимодействий между олигонуклеотидами, был опубликован ранее [4, 44]. В соответствии с известными различиями в свободных энергиях для дискриминации между комплементарными и некомплементарными парами оснований могло оказаться полезным наличие в среде больших количеств белков — каталитически активных, но не выполняющих никаких других специальных функций. Однако такие белковые предшественники в одних случаях будут благоприятствовать комплементарным, а в других случаях — некомплементарным взаимодействиям. Преимущество одних перед другими может определяться только неравенством свободных энергий для различных видов парных взаимодействий. Любое специфическое усиление комплементарных парных взаимодействий потребовало бы

конвергентной эволюции тех конкретных ферментов, которые благоприятствуют данному типу взаимодействий. Для этого сами ферменты должны стать частью самовоспроизводящейся системы, что в свою очередь требует эволюционного развития трансляции.

Первыми самовоспроизводящимися структурами нуклеиновых кислот с устойчивым информационным содержанием — при данных оптимальных значениях $\bar{q} = 0,90—0,99$ — были молекулы типа тРНК. Однако для любой воспроизводимой системы трансляции необходимо по крайней мере на порядок большее информационное содержание. Как мы знаем из анализа репликации РНК-содержащего фага, такому требованию удовлетворяют только оптимально адаптированные репликазы, которые не смогли бы эволюционно развиваться без совершенного механизма трансляции. Современные фаги являются поздними продуктами эволюции, чье существование обязано наличию такого механизма, без которого природе не удалось бы накопить так много информации в отдельной молекуле нуклеиновой кислоты. Следовательно, для молекулярной эволюции нуклеиновых кислот существовал барьер на уровне структур типа тРНК, подобный барьерам, характерным для более поздних этапов эволюции — т. е. для увеличения информационной емкости необходимо было возникновение какого-либо механизма нового типа.

Итак, тРНК или их предшественники, по-видимому, были «старейшими» репликативными единицами, которые начали накапливать информацию и отбирались как квазивиды т. е. как варианты одной и той же основной структуры.

Первым требованием была устойчивость к гидролизу. Игровая модель, подобная описанной в разд. IV.1, показала, что известная вторичная (и третичная) структура тРНК (см. рис. 14, А и Б) является прямым эволюционным следствием этого требования. Далее, симметрия этой структуры отражает оптимизацию ее репликативного механизма, который в соответствии с уравнением (20) предполагает эквивалентное поведение плюс- и минус-цепей, чтобы

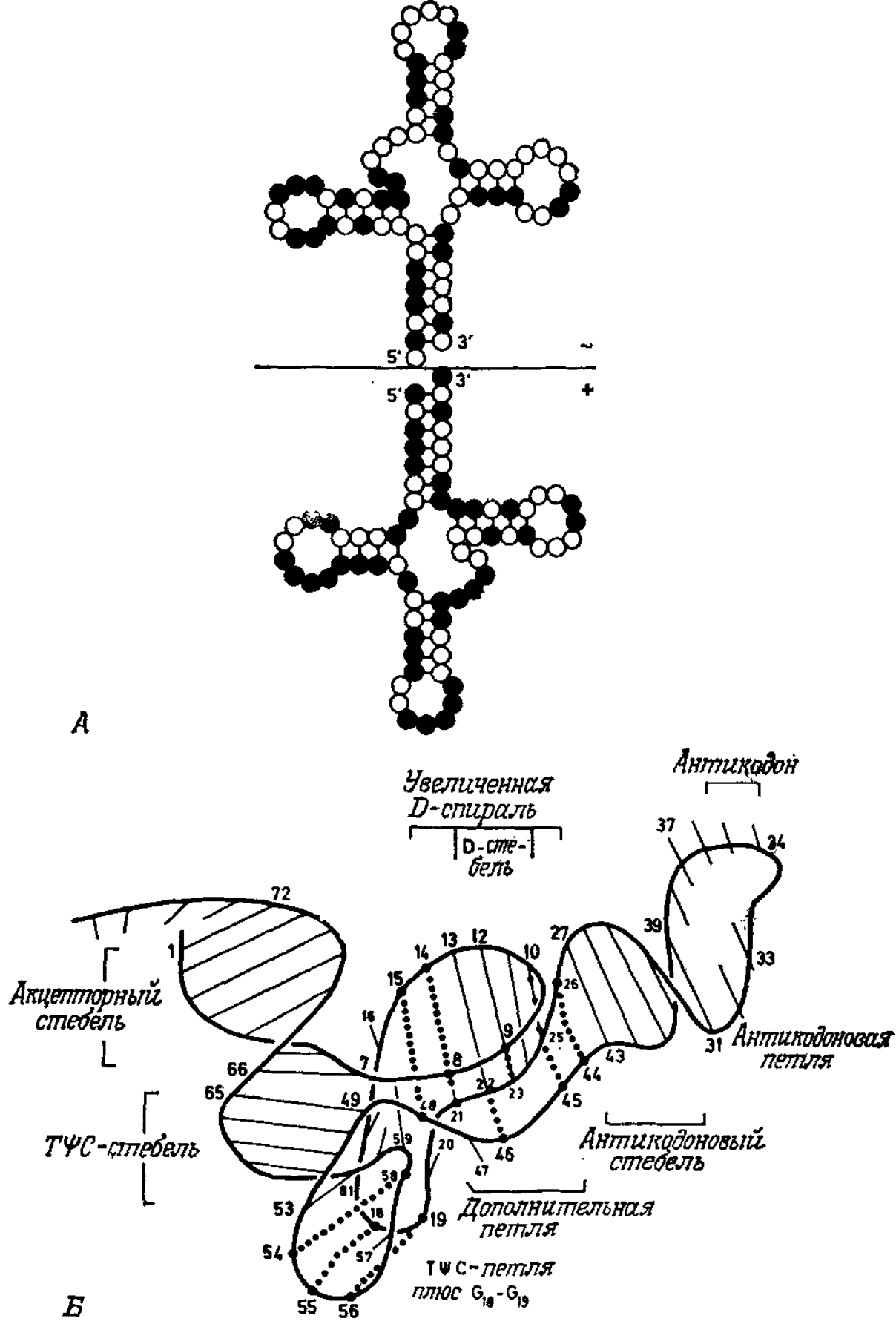


Рис. 14. Симметрия функциональных молекул РНК (рассмотрена на тРНК^{Phe}) способствует репликации одиночных цепей специфически адаптированными ферментами. Плюс- и минус-цепи симметричной структуры имеют одинаковые фенотипические особенности. Хотя в современных организмах транспортные РНК кодируются генотипически, их симметрия, возможно, еще отражает первичный механизм репродукции одноцепочечных РНК, при котором плюс- и минус-цепи одинаково важны. Эта симметрия наиболее очевидна на уровне вторичной структуры (А), хотя она проявляется также на уровне третичной структуры (Б) (из [46]).

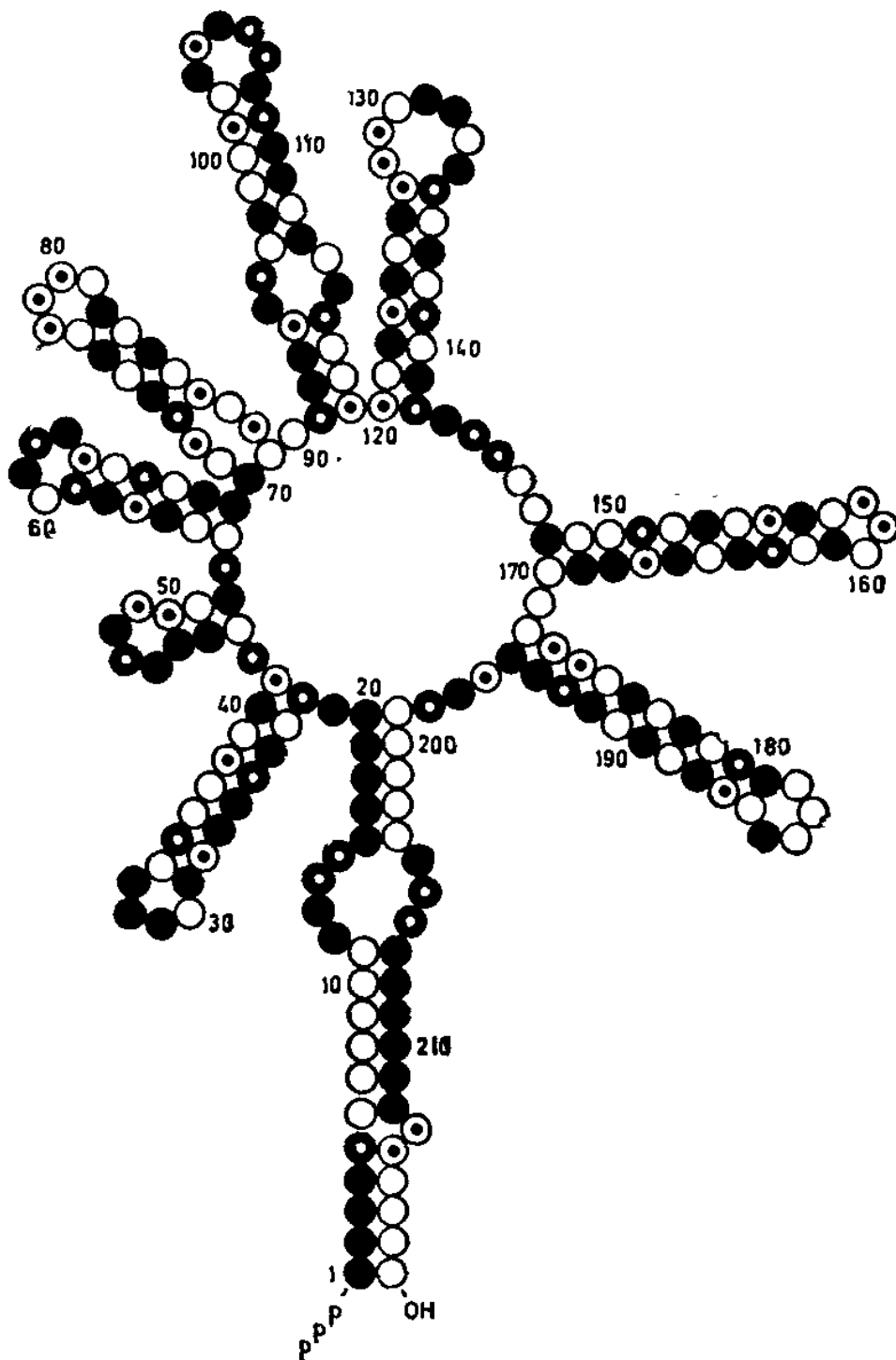


Рис. 15. Модель «цветка» для спигелмановского миди-варианта РНК фага Q_{β} (плюс-цепь). Требования симметрии менее важны в тех случаях, когда информация отображается в генотипах, которые воспроизводятся посредством стандартных механизмов полимеризации. Миди-вариант фага Q_{β} отобран *исключительно* по своей фенотипической информации — по критерию оптимальности структуры для узнавания ферментом Q_{β} -репликазой. Это свойство должно наследоваться как плюс-, так и минус-цепью. Симметрия структуры становится наиболее очевидной в модели «цветка», хотя эта конфигурация, по-видимому, не характерна для активной молекулы в естественных условиях. В соответствии с механизмом репликации одноцепочечных молекул (рис. 11) преимущество должна, по-видимому, иметь модель, предполагающая немедленное складывание цепи во время ее синтеза [27].

имело место оптимальное функционирование. Эту симметрию можно обнаружить на уровне РНК-содержащих фагов, особенно для вариантов, которые отобраны по свойству максимальной эффективности при репликации *in vitro* и не несут никакой лишней генетической информации (рис. 15).

IV.5. Необходимость возникновения гиперциклов

Цель этой работы состоит в том, чтобы показать, во-первых, что скачок в молекулярной эволюции, по-видимому, был вызван интеграцией нескольких самовоспроизводящихся единиц в кооперативную систему, и, во-вторых, что механизм, способный обеспечить такую интеграцию, мог принадлежать только к классу гиперциклов. К этому выводу можно прийти путем следующих рассуждений.

Информационное содержание первых репродуктивных единиц было ограничено $v_{\max} \leq 100$ нуклеотидами. Для построения системы трансляции требовалось несколько таких единиц, имеющих сходные функции, но разные специфичности. Указанная система могла возникнуть из одного квазивида, но эквивалентные партнеры должны были эволюционировать одновременно. Этого нельзя было достичь ни посредством сочленения их в одну самовоспроизводящуюся единицу большего размера (из-за порога ошибок), ни путем компартиментации (из-за сильной конкуренции между эквивалентными самовоспроизводящимися единицами внутри компартамента). Необходимо было образование функциональных связей между всеми самовоспроизводящимися единицами, имеющими следующие свойства:

1. Связь должна все-таки допускать конкуренцию каждой самовоспроизводящейся единицы с ее ошибочными копиями — иначе эти единицы не смогли бы сохранять свою информацию.

2. Связь должна «выключать» конкуренцию между теми самовоспроизводящимися единицами, которые интегрируются в новую функциональную систему, и допускать их кооперацию.

3. Интегрированная функциональная система должна быть способной успешно конкурировать с

Таблица 4

Основные этапы накопления информации в дарвиновских системах

| Вероятность ошибки на букву, $1 - \bar{q}_m$ | Превосходство, σ_m | Максимальное число букв, v_{\max} | Молекулярный механизм и биологический пример |
|--|---------------------------|-------------------------------------|--|
| $5 \cdot 10^{-2}$ | 2 | 14 | Репликация РНК в отсутствие ферментов ¹⁾ Предшественник тРНК, $v = 80$ |
| | 20 | 60 | |
| | 200 | 106 | |
| $5 \cdot 10^{-4}$ | 2 | 1 386 | Репликация одноцепочечной РНК с помощью специфических репликаз Фаг Q_{β} , $v = 4500$. |
| | 20 | 5 991 | |
| | 200 | 10 597 | |
| $1 \cdot 10^{-6}$ | 2 | $0,7 \cdot 10^6$ | Репликация ДНК с помощью полимераз, включая коррекцию экзонуклеазой <i>E. coli</i> , $v = 4 \cdot 10^6$ |
| | 20 | $3,0 \cdot 10^6$ | |
| | 200 | $5,3 \cdot 10^6$ | |
| $1 \cdot 10^{-9}$ | 2 | $0,7 \cdot 10^9$ | Репликация и рекомбинация ДНК в эукариотических клетках Позвоночные (человек), $v = 3 \cdot 10^9$ |
| | 20 | $3,0 \cdot 10^9$ | |
| | 200 | $5,3 \cdot 10^9$ | |

¹⁾ Убедительных данных о некатализируемой репликации РНК пока не получено, однако при катализе на поверхностях или при катализе протеинондами, которые не были специфически адаптированы (идея Фокса), вероятности ошибки могут соответствовать приведенным значениям.

любой другой, менее эффективной системой или единицей.

Эти три условия можно выполнить только путем установления циклической связи между самовоспроизводящимися единицами — другими словами, функциональная связь между автономными самовоспроиз-

водящимися единицами сама должна иметь самоусиливающуюся циклическую природу — иначе их суммарное информационное содержание не сможет сохраняться. Итак, гиперциклическая организация, по-видимому, является необходимой предпосылкой для нуклеации интегрированных самовоспроизводящихся систем с бóльшим информационным содержанием, чем требовалось для возникновения трансляции. Это утверждение следует из того, что будет показано в последующих разделах при более детальном анализе систем со связями.

Если нас спросят, в чем состоит особенность гиперциклов, мы ответим: «Они являются аналогами дарвиновских систем на следующем, более высоком уровне организации». Дарвиновское поведение — это основа для генерирования информации. Его предпосылкой является интеграция самовоспроизводящихся символов в самовоспроизводящиеся единицы, способные предохранять себя от накопления ошибок. То же требование сохраняется для интеграции самовоспроизводящихся и селективно устойчивых единиц в следующую, более высокую форму организации, чтобы в результате снова возникло селективно устойчивое поведение. Только циклическая связь — эквивалент автокаталитического усиления на данном уровне (см. разд. II) — способна привести к этой цели.

Результаты разд. IV суммированы в табл. 4, иллюстрирующей основные этапы накопления информации в дарвиновских системах, которым способствовали различные в отношении накопления механизмы воспроизведения. Эта таблица окажется полезной для обсуждения модели непрерывной эволюции от отдельных молекул до интегрированных клеточных систем, которая рассматривается в части В,

Для характеристики определенного класса самореплицирующихся сетей реакций — гиперциклов — используются топологические методы. Результаты показывают, что свойства гиперциклов достаточны для стабильного интегрирования информации, содержащейся в нескольких самореплицирующихся единицах. Среди исследованных каталитических сетей гиперциклическая организация оказывается необходимой предпосылкой для сохранения устойчивости информации и для ускорения ее дальнейшей эволюции. Методы, использованные в этой работе, привычные для математиков, изложены здесь довольно подробно, с тем чтобы логика рассуждения была понятна и нематематикам.

V. Конкретный пример

В части А этой трилогии о гиперциклах мы пришли к некоторым важным выводам об особенностях дарвиновских систем на молекулярном уровне.

1. Объектом отбора и эволюции является квазивид — распределение (генотипически) близкородственных репликативных единиц, сгруппированных вокруг копии (или вырожденного множества копий), соответствующей фенотипу с максимальной селективной ценностью.

2. Информационное содержание этой доминирующей копии — выраженное в числе символов v (нуклеотидов) на одну репликативную единицу —

ограничено: $v_m < \frac{\ln \sigma_m}{1 - \bar{q}_m}$, где $\sigma_m (> 1)$ — превосходство доминирующей копии, т. е. среднее селективное преимущество над видами остальной части распре-

ления, \bar{q}_m — среднее качество копирования символа. Превышение этого порога информационного содержания вызывает катастрофу ошибок, т. е. распад информации из-за постоянного накопления ошибок.

3. Для достижения устойчивого информационного содержания порядка нескольких тысяч нуклеотидов требуется крайне сложный аппарат ферментативной репликации. Такого количества нуклеотидов как раз достаточно, чтобы закодировать несколько белковых молекул, как у ныне существующих РНК-содержащих фагов. Физические свойства, внутренне присущие нуклеиновым кислотам, допускают воспроизводимое накопление информации в количестве, не превосходящем 50—100 нуклеотидов.

Последнее из этих трех утверждений не является бесспорным, если учесть, что наличие определенных факторов среды, например каталитических поверхностей или даже белковоподобных предшественников ферментов [47], может значительно сместить эти оценки. В действительности приведенные выше значения были получены из равновесных данных, а именно из свободных энергий для (кооперативных) комплементарных взаимодействий нуклеотидов по сравнению с некомплементарными. Тем не менее мы все же считаем, что они представляют собой верхние пределы, которые в природе могут достигаться только в присутствии подходящих катализаторов или путем отжига. Лабораторные эксперименты с полимеризацией на матрицах в отсутствие ферментов дают значительно меньшие значения. С другой стороны, катализаторы, имеющиеся в среде, не могут обеспечить точность распознавания символов, превышающую равновесную, если только они сами не становятся частью селективно оптимизирующейся системы. Для систематического благоприятствования взаимодействиям, дающим функциональные преимущества перед взаимодействиями, не дающими таких преимуществ, не существует другого пути, кроме поэтапной селективной оптимизации. Фаговые геномы могли эволюционировать в форме одноцепочечных молекул РНК только потому, что клетка-хозяин предоставляла

совершенный аппарат репликации и трансляции. Это постклеточные, а не доклеточные продукты эволюции. Нечто подобное величине информационного содержания генома таких вирусов, а именно воспроизводимая информация для набора ферментов, из которых мог бы возникнуть примитивный аппарат трансляции,— это как раз то, что, вероятно, требовалось в начале эволюции трансляции. Итак, важный вывод, вытекающий из части А, состоит в следующем. Для зарождения трансляции требуется интеграция нескольких репликативных единиц в кооперативную систему, чтобы обеспечить достаточное количество информации для построения аппарата трансляции и репликации. Только такой объединенный аппарат способен привести к дальнейшему возрастанию точности и, следовательно, к увеличению информационного содержания.

Как может происходить интеграция конкурирующих молекул, которая не сводится к их объединению в одну большую репликативную единицу — процессу, запрещенному наличием порогового соотношения для v_{\max} ? (Отметим, что интегрирующиеся единицы должны оставаться в конкурентных отношениях со своими мутантами, чтобы эволюционировать дальше и не потерять свою специфическую информацию.) Рассмотрим кратко три возможные альтернативы:

1. *Сосуществование.* Устойчивая взаимная толерантность самореплицирующихся единиц в отсутствие стабилизирующих взаимодействий возможна только для индивидов, принадлежащих к одному и тому же квазивиду. Квазивидовое распределение могло бы обеспечить благоприятные начальные условия для эволюции кооперативной системы, однако оно не способствует эволюции функциональных свойств. Связь, стабилизирующая квазивид, диктуется лишь генотипическими отношениями родства, которые обычно не совпадают с функциональными потребностями. Необходим набор селективно эквивалентных генотипов, дополняющих друг друга на *фенотипическом* уровне. Распределения квазивидов не удовлетворяют этим селекционным критериям.

2. *Компартментация.* Заключение дарвиновской системы в компартмент (отсек) тоже не решает эту проблему. Главным следствием компартментации является усиление конкуренции вследствие ограничения жизненного пространства и поступления питательных веществ. Следовательно, компартментация будет только еще более стабилизировать данный квазивид, имеющий селективное преимущество, а не способствовать эволюции эквивалентных партнеров согласно функциональным критериям, что требует генотипической дивергенции кооперирующихся партнеров. Однако компартментация может дать преимущество системе, в которой уже установилась устойчивая кооперация через функциональные связи (см. часть В). Более сложные компартменты типа современных живых клеток, содержащих только одну (или несколько) копий репликативной субъединицы вместе с аппаратом воспроизведения всего компартмента, требуют, конечно, чтобы среднее качество копирования символа \bar{q}_m согласовывалось с суммарным информационным содержанием в соответствии с формулой для v_{\max} . Другими словами, они подвержены таким же ограничениям, что и система, где все субъединицы объединены в одну длинную цепь.

3. *Функциональные связи.* Отбор функционально кооперирующихся партнеров может совершаться через функциональные связи, которые дают либо взаимное каталитическое усиление воспроизведения, либо структурную стабилизацию. Более подробное изучение таких связей и является основной целью этой работы.

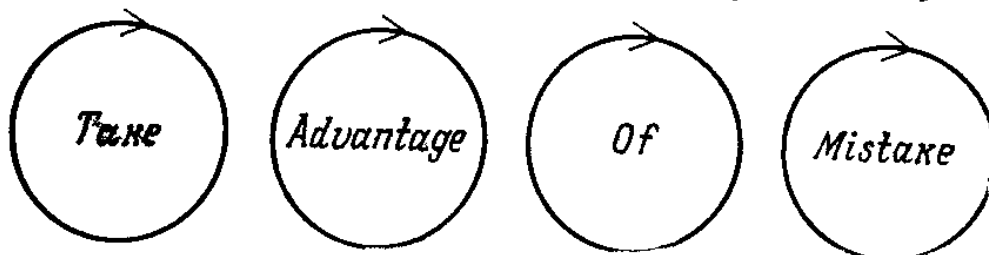
Поможем снова нашей интуиции — обсудим другой вариант машинной игры, рассмотренной в части А.

В первой части игры цель состояла в том, чтобы показать необходимость соответствия качества воспроизведения символов информационному содержанию того предложения, которое должно воспроизводиться. Теперь, во второй части, мы допустим, что величина среднего фактора качества \bar{q}_m недостаточна для стабильного воспроизведения всего предложения

Таблица 5

**Игра, изображающая конкуренцию
селективно эквивалентных репликативных единиц**

Целью игры является сохранение информации предложения



Каждое слово символизирует репликативную единицу. Все слова имеют равную селективную ценность. Селективное преимущество на бит равно 2,7. Каждая буква состоит из пяти двоичных цифр.

| | Вероятность мутации на цифру, (1-q), % | Число цифр, v | Фактор ка- чества слова, $Q = q^v$ | Ожидаемое значение ошибки, e |
|-----------|---|------------------|--|------------------------------------|
| TAKE | 3,15 | 20 | 0,53 | 0,63 |
| ADVANTAGE | 1,4 | 45 | 0,53 | 0,63 |
| OF | 6,3 | 10 | 0,53 | 0,63 |
| MISTAKE | 1,8 | 35 | 0,53 | 0,63 |

Поскольку между словами нет связи, каждая игра кончается *отбором одного слова*. Все слова являются вырожденными в отношении своих селективных ценностей, поэтому каждое из четырех слов имеет равные шансы стать «победителем». Из-за большой средней вероятности ошибки (~2% на бит) предложение в целом (125 бит) не является устойчивой репликативной единицей.

Ниже перечислены типичные исходы десяти игр. Символ X указывает, какое слово отбирается; при этом все другие элиминируются. Число указывает генерацию, после которой отбор заканчивается.

| Игра | TAKE | ADVANTAGE | OF | MISTAKE | Генерация |
|------|------|-----------|----|---------|-----------|
| 1 | X | | | | 12 |
| 2 | | | X | | 15 |
| 3 | | X | | | 19 |
| 4 | | X | | | 23 |
| 5 | X | | | | 10 |
| 6 | | | | X | 20 |
| 7 | | | X | | 9 |
| 8 | | X | | | 13 |
| 9 | | | X | | 22 |
| 10 | | X | | | 26 |

Распределение ошибок для отобранного слова ADVANTAGE.

Сплошная линия близка к распределению Пуассона $\frac{e^k e^{-e}}{k!}$, где $e = v(1 - \bar{q})$ — ожидаемое значение ошибки в слове ($v = 45$ бит).

(Ошибки относятся к одной двоичной цифре. Каждая неверная буква отличается от верной только в одной из своих пяти двоичных цифр.)

| 0 ошибок | 1 ошибка | 2 ошибки | 3 ошибки |
|---|---|------------------------|-----------|
| ADVANTAGE ADVANTAGE ADVANTAGE ADVANTAGE | ADVANTAG! | | |
| ADVANTAGE ADVANTAGE ADVANTAGE ADVANTAGE ADVANTAGE ADVANTAGE ADVANTAGE | ADVANDAGE ADVARTAGE ADVINTAGE ADFANTAGE ADZANTAGE AHVANTAGE AHVANTAGE | ADVANDAC | |
| ADVANTAGE ADVANTAGE | AHVANTAGE AFVANTAGE | ADVANDAIE ADVARXAGE | |
| ADVANTAGE | AFVANTAGE | ADVINTBGE | ITVANTAGU |

в форме репликативной единицы, но достаточна для копирования меньших единиц — таких, как отдельные слова. Это соответствует естественной ситуации на ранних этапах эволюции, когда физических взаимодействий между нуклеотидами могло оказаться достаточно для эволюции стабильных тРНК-подобных молекул (= отдельных слов), но они не допускали построения даже простейшего аппарата трансляции (= всего осмысленного предложения). Соответственно программа для ЭВМ составляется так, чтобы как раз могли воспроизводиться отдельные слова; при этом используются такие вероятности ошибок, которые позволяют избежать накопления последних.

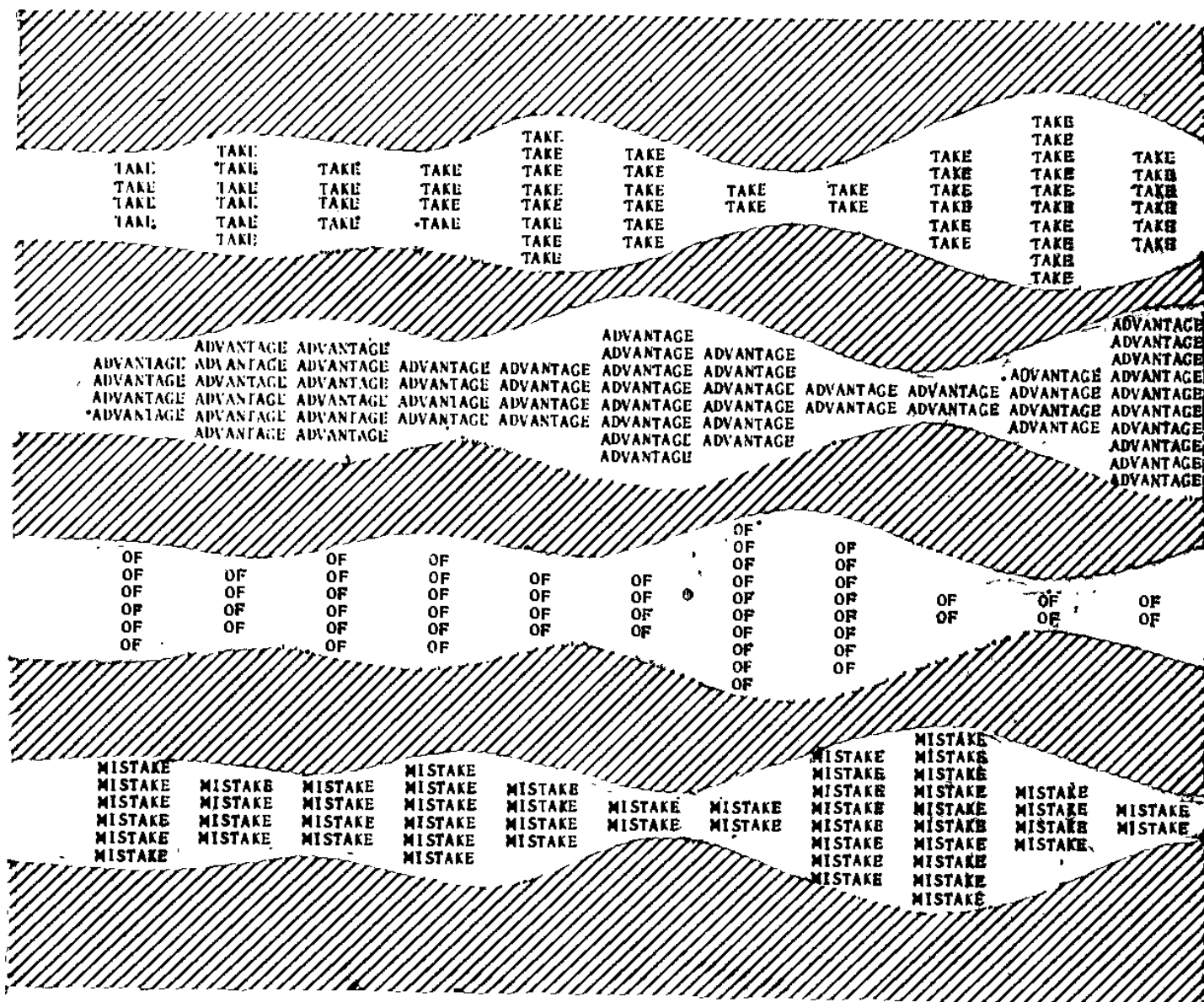
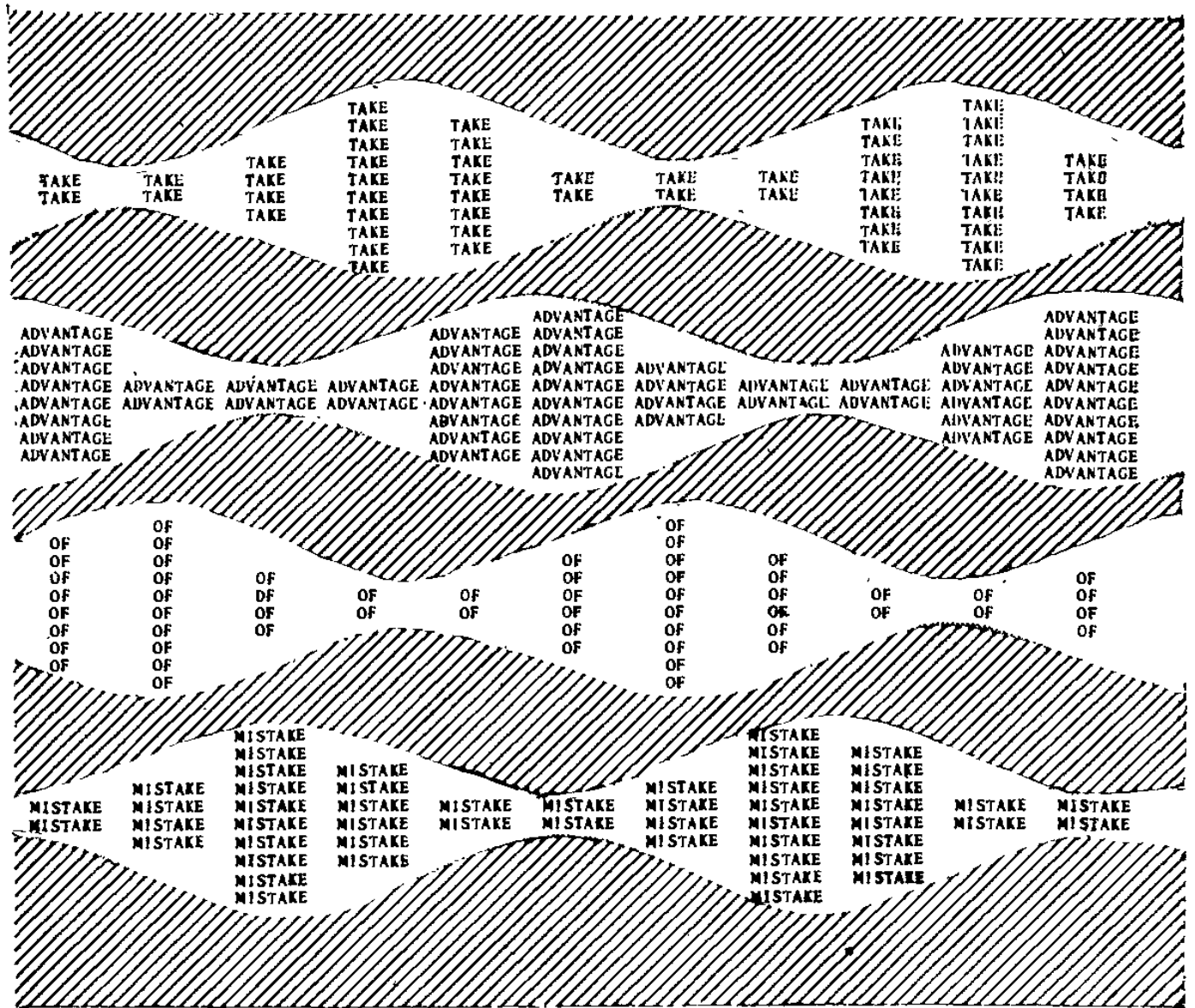


Рис. 16. Каждое слово предложения TAKE ADVANTAGE OF ницу. Информация предложения стабилизируется гиперциклически на диаграмме соответствует 10 копиям в памяти ЭВМ. Все мени; время изменяется дискретно вдоль горизонтальной оси. бания, причем от слова к слову сохраняется определенный рех слов устойчиво и относится к такому же типу, что и

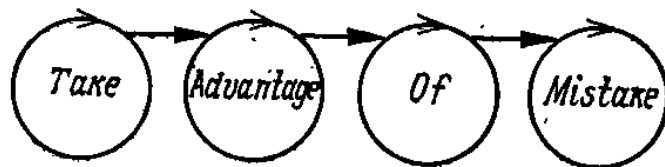
В первом варианте игры мы попытаемся получить простое сосуществование четырех слов. Для этого присвоим всем правильным *словам* в предложении одну и ту же селективную ценность, тогда как ошибка в любом слове будет делать его существование менее выгодным, чем правильного слова, что отразится в умножении селективной ценности на определенный множитель (на бит). Как и прежде, слова могут воспроизводиться, причем их суммарное число остается равным N . Этот вариант, однако, отличается от первоначальной игры тем, что теперь отдельные слова ведут себя как независимые репликативные



MISTAKE представляет собой самовоспроизводящуюся единичной связью между словами. Каждое напечатанное слово слова в вертикальном ряду относятся к одному моменту времени. После начального равномерного распределения слов возникают колебания по фазе. Распределение ошибок для каждого из четырех табл. 5.

единицы. Табл. 5 показывает некоторые типичные результаты такой игры. Несмотря на то что все слова имеют одну и ту же селективную ценность и конкурентоспособны по отношению к своим ошибочным копиям, предложение в целом неустойчиво. Конкуренцию может выиграть только одно из четырех слов, но никак нельзя предсказать, какое именно. Эту ситуацию можно характеризовать тавтологией «выживание выживающего». Термин «наиболее приспособленный» не характеризует здесь ничего, кроме самого результата соревнования.

В следующем варианте игры мы введем функциональную связь между соседними словами: данное слово катализирует воспроизведение следующего слова, если при этом образуется осмысленная последовательность



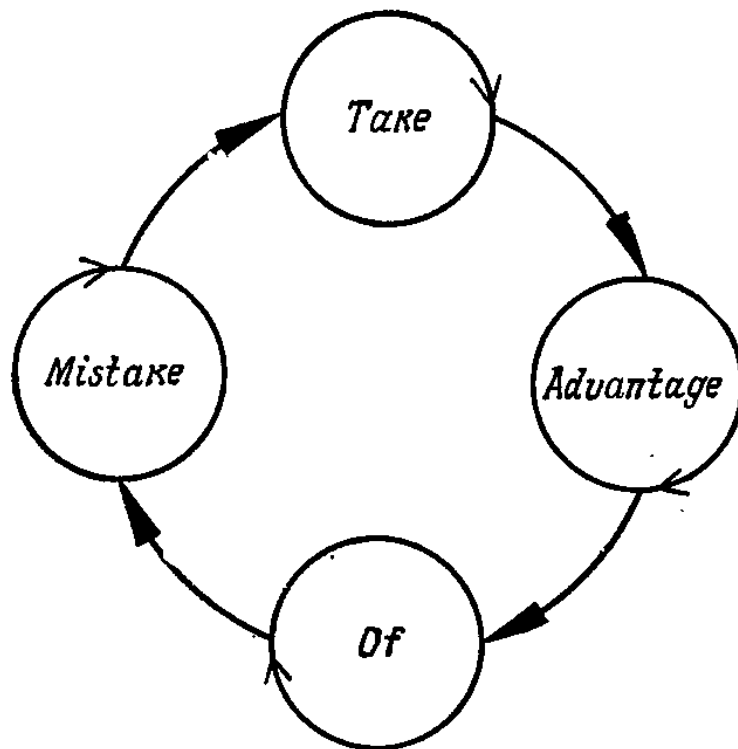
Сила связи пропорциональна численности популяции катализатора (т. е. числу копий данного слова в памяти ЭВМ). Другими словами, воспроизведение усиливается по следующему кинетическому закону:

$$k_1 x_1 \quad \text{и} \quad k_i x_i + k'_i x_i x_{i-1} \quad \text{для} \quad i = 2, 3, 4,$$

где x_i или x_{i-1} — соответствующие численности популяций, в данном случае популяций в памяти ЭВМ. Результатом этого варианта игры обычно является фиксация последнего слова в предложении, т. е. слова «mistake», в то время как все другие слова «вымирают». Только в том случае, когда связь относительно слаба и определенное значение k_i выбрано достаточно большим, соответствующее слово (i) может «вытеснить» все другие, т. е. происходит отбор против (существенно) независимых конкурентов. Тот результат, что последнее слово в последовательности получает всю выгоду от связи (когда члены связи преобладают), кажется удивительным. Можно было бы ожидать, что существует по крайней мере какая-то область устойчивости для всего предложения. Это верно для определенной величины численности популяций, если значения кинетических параметров располагаются в определенном порядке относительно положения слов в последовательности. Однако исследование особых точек, проведенное в разд. VII, показывает, что даже при этих специальных условиях численность лишь последнего члена в цепи будет расти пропорционально суммарной численности популяции, численность же всех других членов становится

существенно постоянной независимо от размеров суммарной популяции. Следовательно, в растущей популяции относительное обилие последнего члена резко изменяется, пока вся система снова не достигнет области, где устойчивость сохраняет только член с высоким обилием. В процессе молекулярной эволюции численности индивидов обычно испытывают именно такие резкие изменения — от одной мутантной копии до ощутимой величины (более чем) 10^9 копий. Итак, результат нашей игры оказывается характерным для реальных процессов, протекающих в природе.

Тот факт, что линейная связь — если она вообще работает — передает все преимущество последнему члену последовательности, дает четкое указание на возможное решение проблемы: связи должны образовывать замкнутый цикл



Тогда усиление, обусловленное связью, будет перемещаться по кругу через все слова в последовательности. Наше предложение на самом деле было выбрано так, что оно автоматически обеспечивает такое циклическое наложение из-за слова «mistake». Поскольку каждое слово является каталитическим циклом (т. е. самореплицирующейся единицей), вся система представляет собой гиперцикл второй степени

в соответствии с определением, данным в части А. Исход игры представлен на рис. 16. Видно, что все четыре слова имеют устойчивые стационарные численности, на которые наложены периодические вариации. Селективные ценности различных слов могут не быть одинаковыми — такая ситуация для любой реальной системы очень маловероятна. Далее, каждое слово представлено стабильным распределением мутантов. Если только какое-нибудь слово не стирается флуктуационной катастрофой (которая становится очень маловероятной при достаточно большом числе копий), численности популяций продолжают варьировать. Другими словами, информация всего предложения устойчива.

VI. Общая классификация динамических систем

VI.1. Определения

В следующих разделах мы займемся более строгим математическим анализом динамических систем, особенно таких, которые важны с точки зрения доклеточной самоорганизации. Чтобы определить, какие именно системы важны, мы должны обследовать различные классы сетей реакций, нециклических и циклических. Эволюционные процессы могут быть феноменологически описаны с помощью систем дифференциальных уравнений, как было показано на конкретном примере в части А. Тогда термин «динамическая система» будет относиться к полному многообразию всех интегральных кривых (решений) данной системы дифференциальных уравнений.

Рассмотрим общую динамическую систему, которая описывается n обыкновенными автономными дифференциальными уравнениями первого порядка:

$$\frac{dx_i}{dt} \equiv \dot{x}_i = \Lambda_i(x_1, \dots, x_n, k_1, \dots, k_m; B); \quad (30)$$

$$i = 1, 2, \dots, n.$$

В дальнейшем мы распространим наш анализ и на некоторые неавтономные системы, для которых $\Lambda_i = \Lambda_i(t)$.

Как и прежде, x_i — это популяционные переменные, которые обычно будут относиться к самореплицирующимся макромолекулярным ансамблям. Константы k_i ($i = 1, 2, \dots, m$) играют роль параметров и могут быть составлены из констант скоростей элементарных процессов, констант равновесия для обратимых стадий с быстро устанавливающимся равновесием и из концентраций таких молекул, которые служат (высокоэнергетическим) строительным материалом для синтеза макромолекул — при допущении, что эти концентрации поддерживаются постоянными и, следовательно, их можно считать не зависящими от времени. Оба набора значений x и k можно представить в виде вектор-столбцов в пространстве концентраций или соответственно в пространстве параметров:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} \quad \text{и} \quad \mathbf{k} = \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ \vdots \\ k_m \end{pmatrix}.$$

Буквой B мы обозначаем начальные условия для данного набора интегральных кривых; в нашем случае эти начальные условия являются набором начальных концентраций x_0 .

Согласно процедуре, использованной в части A, мы подразделяем функции Λ_i на три составляющих:

$$\Lambda_i = A_i - \Delta_i - \phi_i. \quad (31)$$

Величины A_i состоят из всех положительных вкладов в скорости химических реакций и представляют собой «усиление» переменных x_i , в то время как Δ_i включают все отрицательные кинетические члены, эквивалентные «разложению» макромолекулярного вида. Наконец, ϕ_i — это поток, который может либо разбавлять компонент i , либо поддерживать его концентрацию

на постоянном уровне в зависимости от внешних ограничений, наложенных на систему. Разность $A_i - \Delta_i$ может быть названа функцией чистого роста Γ_i . Для дарвиновской системы (см. часть А) Γ_i определяется как $W_{ii}x_i + \sum_{k \neq i} \omega_{ik}x_k$ и при суммировании по всем видам k от 1 до n получается выражение, подобное функции избыточного роста

$$E = \sum_{k=1}^n E_k x_k.$$

VI.2. Неограниченный рост

Снятие селекционных ограничений приводит к новой системе дифференциальных уравнений

$$\dot{x}_i = \Gamma_i(x, k, B), \quad (32)$$

описывающей ситуацию, которая впредь будет называться «неограниченным ростом». Эта терминология характеризует систему в целом; индивидуальные члены могут при этом распадаться или находиться в стационарном состоянии.

Допустим, что мы можем представить Γ_i в виде полинома, переменные которого суть концентрации x_i (полиномом можно аппроксимировать также иррациональные выражения или отношения полиномов). Тогда в Γ_i обычно можно выделить ведущие члены, которые будут доминировать в определенной области концентраций. Как правило, эти ведущие члены являются просто одночленами данной степени x_i . Они-то и определяют динамическое поведение системы.

Рис. 17¹ иллюстрирует простой случай $\dot{x} = kx^p$. Стандартные решения были нормированы так, чтобы $x(0) = 1$ и $\dot{x}(0) = 1$. Как указано в подписи к рисунку, все семейство интегральных кривых можно разде-

¹ В подписи к рис. 17 содержится ошибка, замеченная авторами данной книги (Eigen M., Schuster P. 1979. Naturwissenschaften, 66, 512): на самом деле кривая 4, для которой $\Gamma(x) = e^{1-x}$, не разделяет области А и В. Однако выводы авторов, касающиеся достаточно больших значений t , и все выводы, сделанные в тексте, остаются в силе. — Прим. перев.

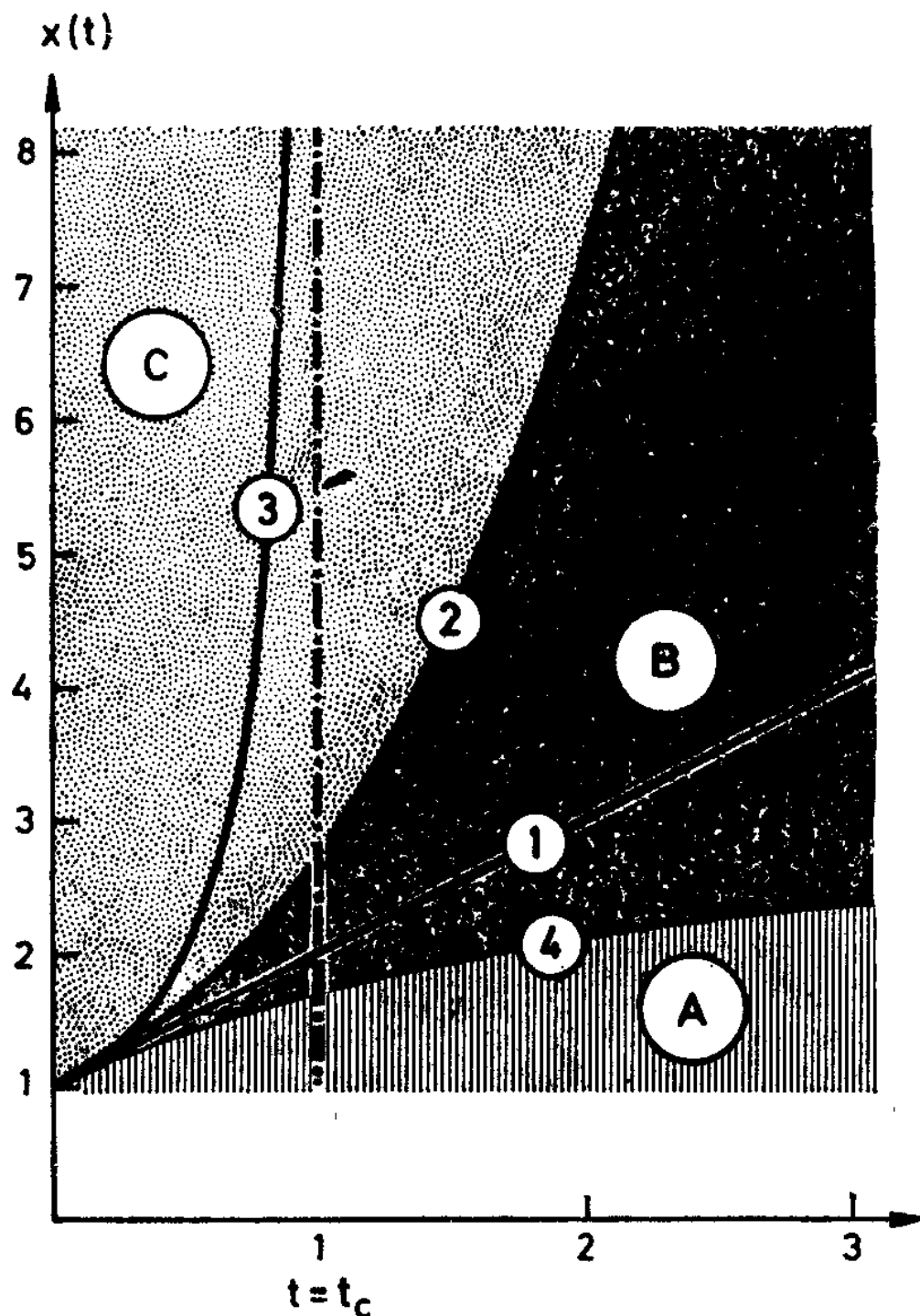


Рис. 17. Различные типы роста можно описать с помощью одночленных функций роста $\Gamma(x) = dx/dt$ (нормированных к $\Gamma = 1$ и $x = 1$ при $t = 0$). В области *A* нет функций роста, которые могут быть представлены простым одночленом $\Gamma = x^p$. В этой области все численности популяций $x(t)$ остаются конечными при $t \rightarrow \infty$. Пограничная кривая между областями *A* и *B* задается функцией роста $\Gamma(x) = e^{(1-x)}$ (кривая 4). Область *B* представлена всеми одночленами $\Gamma(x) = x^p$, где $-\infty < p < 1$. Кривая 1 — пример постоянной скорости роста ($p = 0$). Экспоненциальному росту ($p = 1$) соответствует пограничная кривая 2 между областями *B* и *C*. В то время как в области *B* численности популяций достигают бесконечности только через бесконечное время, в области *C* они имеют сингулярности в конечные моменты времени. Кривая 3, например, соответствует гиперболическому росту ($p = 2$, сингулярность при $t = t_c = 1$).

лить на три класса, располагающихся в разных областях диаграммы концентрация — время. Рассмотрим три типичных случая, которые представляют для нас особый интерес (см. табл. 6).

1. Кривая 1 характеризует систему с постоянной (положительной) скоростью роста. Популяционная переменная $x(t)$ растет линейно во времени. Эта кривая является также примером семейства кривых из области В, которые растут до бесконечности за бесконечное время. Обычным примером такого рода может служить реакция необратимого образования вещества при постоянных концентрациях реагента. Если популяции самовоспроизводящихся видов в экологических нишах, добывающие пищу из независимых источников, растут со скоростью, соответствующей постоянному притоку питательных веществ или постоянной скорости их образования, то они тоже являются примерами систем, скорость роста которых не зависит от размеров популяции.

2. Кривая 2 соответствует случаям, когда скорости роста линейно зависят от популяционной переменной, и соответствует экспоненциальному росту x в зависимости от t , типичному для дарвиновского поведения (часть А). Кроме того, кривая 2 разграничивает области В и С, т. е. функции, которые достигают бесконечности за бесконечное и за конечное время соответственно.

3. Наконец, кривая 3 представляет собой пример функции с сингулярностью при конечном значении времени $[t_c = (kx_0)^{-1}]$. В этом частном случае предполагается, что скорость роста пропорциональна квадрату популяционной переменной.

Всю область С можно охарактеризовать как область «гиперболического роста». Конечно, в любом реальном и конечном мире популяция не может расти до бесконечности — это обусловлено конечностью имеющихся ресурсов. Однако свойства, от которых зависит существование гипотетической сингулярности, будут все-таки приводить к поведению, совершенно не похожему на то, с чем мы сталкиваемся в дарвиновских системах,

Здесь мы можем дать более общее определение «степени» ρ функций роста, которое окажется полезным для классификации. Как и прежде, ρ_i — степень ведущего члена функции роста Γ_i . Тогда n -мерную динамическую систему можно характеризовать множеством значений ρ_i ($\rho_1, \rho_2, \dots, \rho_n$). В том случае, когда степени ρ_i одинаковы:

$$\rho_1 = \rho_2 = \dots = \rho_n \equiv \rho, \quad (33)$$

мы будем называть систему «чистой». В противном случае мы имеем дело со «смешанными» системами, которые можно классифицировать в соответствии с распределениями их значений ρ_i . Очевидно, что «чистые» системы исследовать значительно легче, чем «смешанные».

VI.3. Ограниченный рост и отбор

В действительности мы всегда имеем дело с системами, рост которых по той или иной причине ограничен. При экспериментальных исследованиях мы должны обеспечить воспроизводимость условий. Поэтому необходимо формализовать эти условия и включить их в теоретическое рассмотрение.

В термодинамике необратимых процессов мы предпочли бы выбрать такие селекционные ограничения, которые облегчают термодинамическое описание, например постоянные обобщенные силы или постоянные обобщенные потоки. Мы должны согласовать их с условиями отбора и эволюции, которые могут реализоваться в природе. Ограничения ϕ_i , использованные в уравнении (31), имеют слишком общий вид и поэтому непригодны для прямого анализа. Вообще говоря, можно выделить специфические и неспецифические селекционные ограничения. В первом случае ограничения действуют специфически на какой-либо один вид или на несколько видов, тогда как второй случай относится к регуляции общего потока ϕ . Тогда изменения всех популяционных переменных будут пропорциональны их текущим значениям x_i :

$$\phi_i = \frac{x_i}{\epsilon} \phi. \quad (34)$$

На практике неспецифические селекционные ограничения можно наложить на динамическую систему, введя непрерывный поток разбавления. Тем самым можно контролировать суммарную концентрацию $c = \sum x_i$. Соответствующее дифференциальное уравнение для c

$$\dot{c} = \sum_{j=1}^n \dot{x}_j = \sum_{j=1}^n \Gamma_j(\mathbf{x}) - \phi \quad (35)$$

удовлетворяет условию стационарности $\dot{c} = 0$, когда поток регулируется так, чтобы компенсировать результирующую избыточную продукцию:

$$\phi = \phi_0 = \sum_{j=1}^n \Gamma_j(\mathbf{x}). \quad (36)$$

Это селекционное ограничение, называемое ограничением «постоянной организации», было введено ранее и использовалось также в части А. Условие (36) будет часто применяться в последующих разделах для облегчения общего анализа селекционных процессов. Исследовались также другие типы ограничений [53]. Как будет видно из следующего раздела, важные особенности селекционных и эволюционных процессов довольно мало чувствительны к типу наложенных ограничений. (Хотя они, конечно, всегда отражаются на количественных результатах.)

Условие постоянной организации ведет к следующим дифференциальным уравнениям для динамической системы:

$$\dot{x}_i = \Gamma_i(\mathbf{x}) - \frac{x_i}{c_0} \sum_{j=1}^n \Gamma_j(\mathbf{x}), \quad i = 1, 2, \dots, n. \quad (37)$$

Здесь c_0 — стационарное значение суммарной концентрации, которое можно сохранять постоянным, поддерживая поток на уровне ϕ_0 .

Селекционное поведение трех простых функций роста ($p = 0, 1$ и 2 ; см. обсуждение в подписи к рис. 17) отражено в табл. 6.

1. Постоянная скорость роста — что соответствует линейному росту популяции во времени — при огра-

Таблица 6

Скорости роста и селекционное поведение
динамической системы $\dot{x} = \Gamma - \phi$ при селекционном
ограничении постоянной общей организации

| Неограниченный рост | | | Асимптотическое поведение при ограничении постоянной организации | |
|---------------------|----------------------------|------------------|--|--|
| р | интегральная кривая | тип роста | \bar{x} | тип селекционного поведения |
| 1 0 k | $x = x_0 + kt$ | Линейный | $\bar{x}_i = \frac{k_i c_0}{\sum k_j}$ | Сосуществование видов без отбора |
| 2 1 kx | $x = x_0 \exp(kt)$ | Экспоненциальный | $\bar{x}_k = c_0, \bar{x}_i = 0$ $k_k - k_i > 0, i \neq k$ | Конкуренция, ведущая к отбору глобально наиболее приспособленного вида |
| 3 2 kx ² | $x = x_0(1 - kx_0 t)^{-1}$ | Гиперболический | $\bar{x}_k = c_0, \bar{x}_i = 0$ $k=1, 2, \dots, n; i \neq k$ | Конкуренция, стремящаяся к локальной оптимизации, решение «раз и навсегда» |

ничении постоянной организации дает устойчивое сосуществование всех партнеров, присутствующих в системе. Рост численности мутантов, имеющих преимущество, сдвигает стационарные отношения, но система в целом остается устойчивой.

2. Линейная скорость роста, соответствующая экспоненциальному росту популяции, приводит к конкуренции и отбору «наиболее приспособленного». Мутанты, имеющие преимущество, после своего

появления дестабилизируют и замещают установившуюся популяцию.

3. Нелинейная скорость роста ($p > 1$), характерная для гиперболического роста, также ведет к отбору, более жесткому, чем в дарвиновской системе, о которой шла речь в п. 2. Однако мутанты с кинетическими параметрами, дающими преимущество, в общем не способны расти и дестабилизировать установившуюся популяцию, потому что селективная ценность является функцией численности популяции (например, для $p = 2$ $W \sim x$). Отсюда преимущество любой установившейся популяции с конечным значением x настолько велико, что оно едва ли может быть уменьшено одиночной мутантной копией, какой бы она ни была. Тогда отбор является решением, принятым «раз и навсегда». Здесь для сосуществования нескольких видов требуется кооперативная связь очень специального вида.

Приведенные примеры весьма типичны. Мы можем классифицировать системы в соответствии с их селекционным поведением либо как допускающие сосуществование, либо как конкурентные. В данной системе мы можем встретить более чем один тип поведения.

VI.4. Внутреннее уравнивание в растущих системах

Хотя условие постоянной организации значительно упрощает исследование динамической системы, при этом мы ограничиваемся системами с нулевым чистым ростом. В данном разделе мы попытаемся расширить круг рассматриваемых систем. Основная проблема состоит в том, чтобы найти, каким образом и при каких условиях можно делать предсказания о поведении растущих систем, основываясь на результатах, полученных из анализа соответствующих стационарных состояний. Для этой цели мы введем неспецифические селекционные ограничения [уравнение (34)], зависящие от времени:

$$\dot{x}_i = \Gamma_i(x) - \frac{x_i}{c(t)} \phi(t). \quad (38)$$

Функция $c(t)$ либо функция $\phi(t)$ может быть выбрана произвольно. Однако после этого вторая функция определяется следующим дифференциальным или соответственно интегральным уравнением:

$$\phi(t) = \sum_{i=1}^n \Gamma_i(\mathbf{x}) - \frac{dc}{dt}, \quad (39)$$

$$c(t) = c_0 + \int_0^t \left\{ \sum_{i=1}^n \Gamma_i(\mathbf{x}) - \phi(\tau) \right\} d\tau. \quad (40)$$

Теперь следует ввести нормированные популяционные переменные $\xi = \frac{1}{c} \mathbf{x}$. Тогда дифференциальные уравнения можно привести к следующему виду:

$$\dot{\xi}_i = \frac{1}{c(t)} \{ \Gamma_i(\mathbf{x}) - \xi_i \sum_j \Gamma_j(\mathbf{x}) \}. \quad (41)$$

Нетрудно убедиться, что $\dot{\xi}_i$ в явном виде не зависит от селекционного ограничения $\phi(t)$. Существует, однако, неявная зависимость через $c(t)$. Поэтому сделаем в нашем общем исследовании еще один шаг, рассмотрев несколько простых примеров. Допустим, что функции чистого роста $\Gamma_i(\mathbf{x})$ имеют одну и ту же степень λ по \mathbf{x} . Хотя это условие кажется очень жестким, мы увидим, что почти все наши основные модельные системы с ним согласуются, по крайней мере при определенных граничных условиях. Однородность по \mathbf{x} ведет к тому же условию, что и требование определенной степени p ($\lambda = p$) в системе с неограниченным ростом (см. разд. I.5). Теперь преобразование переменных становится тривиальным:

$$\Gamma_i(\mathbf{x}) = \Gamma_i(c\xi) = c^\lambda \Gamma_i(\xi), \quad (42)$$

и мы получаем следующие кинетические уравнения:

$$\dot{\xi}_i = c^{\lambda-1} \left\{ \Gamma_i(\xi) - \xi_i \sum_{j=1}^n \Gamma_j(\xi) \right\}. \quad (43)$$

Вид этих уравнений сразу же позволяет сделать два важных вывода. При $\lambda = p = 1$, т. е. для дарвиновской

системы, рассмотренной в части А, зависимость от c исчезает, и если использовать относительные популяционные переменные ξ_i , то не только поведение растущих и стационарных систем при $t \rightarrow \infty$, но и интегральные кривые будут идентичны.

Если $\lambda = \rho \neq 1$, то поведение ξ при $t \rightarrow \infty$ будет таким же, как и для стационарной системы при постоянной организации, если только $c(t)$ не обращается ни в нуль, ни в бесконечность. Итак, для всех реалистических систем с однородными функциями чистого роста Γ_i результаты исследования особых точек в ξ -пространстве, которые мы получим в следующем разделе, будут верны и для случая растущих популяций.

Последний результат можно распространить и на другие классы функций роста, как будет показано в разделе, посвященном исследованию особых точек. Внутреннее уравнивание чрезвычайно упрощает анализ сложных динамических систем. Во многих случаях результаты становятся идентичны или подобны тем, которые получены для стационарных условий. Если нас интересует селекционное поведение системы, то это именно те условия, которые имеет смысл рассматривать. В следующем разделе мы проанализируем более подробно различные динамические системы, находящиеся при этих условиях.

VII. Исследование особых точек самоорганизующихся сетей реакций

VII.1. Адекватный метод исследования

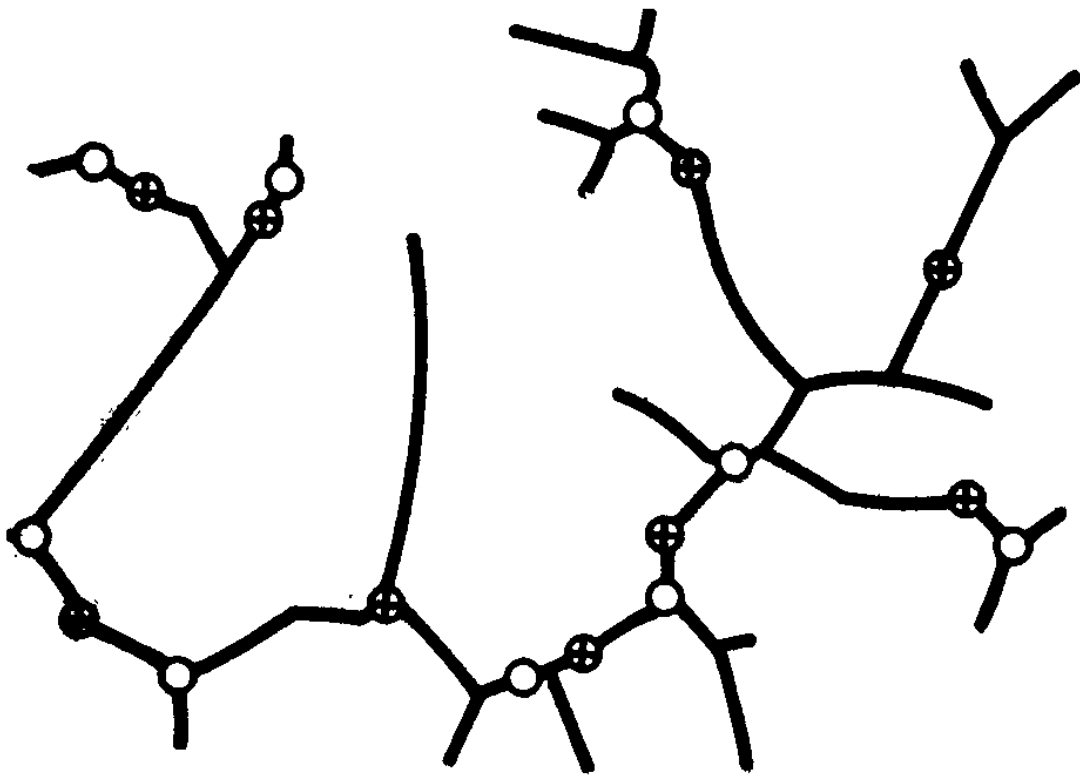
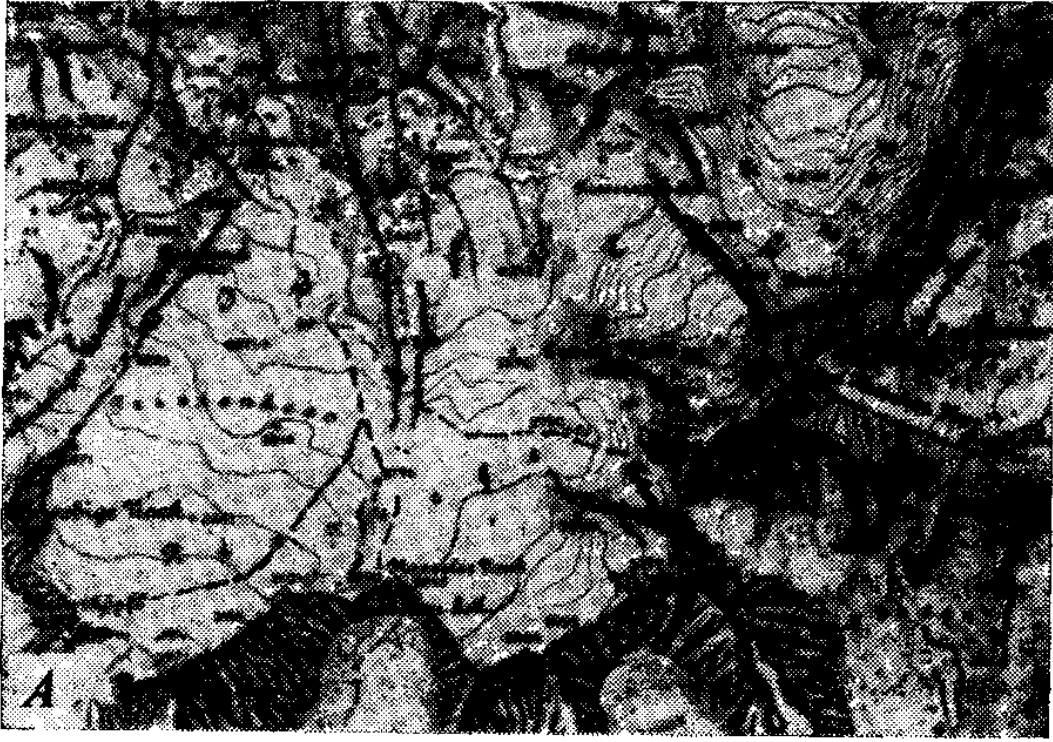
При анализе различных молекулярных процессов самоорганизации нас, естественно, больше интересует конечный результат отбора, чем детальное описание всего этого динамического процесса. Соответственно в этом разделе нам не потребуется вся информация, которая содержится в полной системе решений, удовлетворяющих системе дифференциальных уравнений. В качестве метода анализа мы выбираем исследование особых точек, потому что этот

метод наилучшим образом соответствует целям сравнительного анализа селекционного поведения. Лишь в некоторых случаях мы будем использовать также более тонкие методы, например исследование полных векторных полей.

В настоящее время исследование особых точек является обычным методом анализа асимптотического поведения динамических систем. Его изложение можно найти в соответствующих учебниках (см., например, [48]). Исследование особых точек применялось также для решения задач экономики, при анализе экологических моделей, а также химических реакций вдали от равновесия [49]. Представление о современном состоянии этой области можно получить, прочитав недавно опубликованный обзор [50].

VII.2. Топологические свойства

Пусть мы имеем карту горной страны (рис. 18) и хотим составить приблизительное представление об этом трехмерном ландшафте; такое представление дают нам линии уровня на двумерной карте. Это именно тот тип задач, с которыми имеет дело исследование особых точек. Ландшафт соответствует потенциальной поверхности, по которой движется динамическая система. В большинстве случаев полного знания этой потенциальной поверхности не требуется, и поэтому «карта особых точек» оказывается намного проще топографической карты, используемой нами для ориентации в незнакомой местности. «Карта особых точек» указывает положения лишь локально самых высоких и самых низких точек — таких, как горные вершины, перевалы и впадины, которые здесь называются источниками, седлами и стоками. Это особые точки потенциального поля. Часто на карту бывает необходимо нанести и гребни — линии, отделяющие долины друг от друга (рис. 18) и называемые поэтому «сепаратрисами». Карта особых точек, включающая сепаратрисы, позволяет предсказать, куда приведет траектория, которая начинается в данной точке на карте. Траектории — это линии наиболее



Б

Рис. 18. А. Топографическая карта дает абстрактное представление о местности. Линии (горизонталы) соединяют точки равной высоты. Показан район Восточных Альп (перепечатано из «Карты Австрии», масштаб 1 : 50 000, лист 177 (1962), с любезного разрешения Bundesamt für Eich und Vermessungswesen Abt. Landesaufnahme). Б. Карта особых точек, являющаяся дальнейшей абстракцией топографической карты А. ○ — источники, или вершины; ⊕ — седловые точки, ● — стоки; сплошными линиями обозначены сепаратрисы.

крутого спуска, по которым на местности будет следовать текущая вода. Однако гравитационное потенциальное поле у поверхности земли проще, чем те поля, с которыми мы сталкиваемся в самоорганизующихся динамических системах. В то время как вода, текущая по земле, всегда достигает стока — например, озера — самоорганизующиеся динамические системы могут проявлять более сложное поведение. Например, существует ситуация — предельные циклы, — когда (на языке нашей иллюстрации) вода не останавливается в определенной точке, а бесконечно циркулирует вдоль замкнутой кривой, которая определяется формой потенциального поля. Были описаны даже еще более странные ситуации, которые математики действительно называют «странными аттракторами», — они представляют собой нечто вроде непериодических орбит. Аттрактор — более общее понятие, нежели сток. В эту категорию включаются не только стоки, но и устойчивые замкнутые орбиты и непериодические орбиты.

На карте особых точек вся рассматриваемая поверхность может быть разделена на ряд областей, обычно называемых бассейнами, которые связаны с отдельными аттракторами. Границы этих областей являются сепаратрисами. Итак, из всех точек бассейна вода течет к одному и тому же аттрактору, который, конечно, должен находиться внутри этой области.

Теперь мы перейдем на более точный математический язык и охарактеризуем те величины и выражения, которые нам будут необходимы для дальнейшего обсуждения. Особые, или стационарные, точки динамической системы определяются как такие точки, в которых все концентрации или популяционные переменные x_i постоянны во времени. Следовательно, первые производные по времени обращаются в нуль:

$$\dot{x}_i = 0, \quad i = 1, 2, \dots, n; \quad (44)$$

тем самым определяются положения всех особых точек, принадлежащих данной динамической системе. Когда все случайные флуктуации популяционных

| Класс | | ω_1 | ω_2 | | |
|-------|--|--------------------|--------------------|-----------------------|----------------|
| 1 | | < 0 | < 0 | | Узловые |
| 1 | | < 0 | < 0 | $\omega_1 = \omega_2$ | Фокальные |
| 1 | | $-\alpha + i\beta$ | $-\alpha - i\beta$ | $\alpha, \beta > 0$ | Спиральные |
| 2 | | > 0 | < 0 | | Седловая точка |
| 3 | | > 0 | > 0 | | Источник |
| 4 | | > 0 | $= 0$ | | |
| 4 | | $= 0$ | < 0 | | |
| 4 | | $+i\beta$ | $-i\beta$ | | Центр |

Рис. 19. Символы, используемые для классификации различных особых точек. Класс 1 — устойчивые особые точки, или стоки. Класс 2 — седловые точки. Класс 3 — источники. Класс 4 — неустойчивые особые точки, в том числе точки, собственные значения которых имеют нулевые действительные части. Эти примеры относятся к двумерной динамической системе.

переменных полностью подавлены, интегрирование уравнений динамической системы, «стартовой» из особой точки, дает независимые от времени постоянные популяции. Реакция системы на малые изменения концентраций в окрестности данной особой точки является великолепной основой для классификации этих точек. Эта реакция системы может быть описана при помощи множества нормальных мод, характеризующихся обратными постоянными времени ω_k — собственными значениями системы линейных дифференциальных уравнений, которая является наилучшей аппроксимацией нелинейной системы в окрестности рассматриваемой точки (разд. VII.4). Соответственно можно выделить четыре основных класса особых точек:

1. *Устойчивые особые точки, или стоки*, т. е. локально наиболее низкие точки. Все собственные значения ω_k имеют отрицательные действительные части, и, следовательно, флуктуации по всем возможным

направлениям в пространстве концентраций компенсируются внутренней противодействующей силой. В химии стоки соответствуют химическим равновесиям в замкнутых системах и устойчивым стационарным состояниям в открытых термодинамических системах.

2. *Седловые точки*, для которых хотя бы одно направление является неустойчивым. Здесь по крайней мере одно значение ω_k должно иметь положительную действительную часть. Следовательно, небольшое возмущение или флуктуация в этом направлении приводят к возникновению силы, стремящейся увеличить флуктуацию. В результате динамическая система будет удаляться от седловой точки.

3. *Источник* — локально наиболее высокая точка. Он отличается от седла только тем, что неустойчив по всем направлениям. Все значения ω_k имеют положительные действительные части.

4. *Еще один класс особых точек*, которые не поддаются полному исследованию в рамках линейной теории. Некоторые частоты ω_k имеют нулевые действительные части, и их природа может меняться в зависимости от вклада нелинейных членов. Примером такого рода служат центры, которые характеризуются чисто мнимыми собственными значениями. Траектории в окрестности центра представляют собой многообразие концентрических орбит. С такими ситуациями мы встретимся в данной работе.

По истечении «достаточно большого» времени (т. е. времени, много большего, чем максимальная постоянная времени динамической системы) каждая реалистическая динамическая система (т. е. система без внешнего подавления флуктуаций) достигнет аттрактора. Следовательно, результат отбора будет всегда совпадать с аттрактором в пространстве концентраций.

Окончательный результат процесса отбора соответствует либо устойчивому стационарному состоянию, либо непрерывно и периодически изменяющемуся семейству состояний. В некоторых особенно редких ситуациях могут происходить, кроме того,

непериодические изменения в пределах определенного множества состояний. Для характеристики всех этих устойчивых или квазиустойчивых конечных ситуаций в дифференциальной топологии используют общий термин — «аттрактор» динамической системы, куда включаются устойчивые точки, замкнутые орбиты и аперриодические кривые. Внутри данного бассейна результатом процесса отбора является достижение одного и того же аттрактора, независимо от конкретных начальных условий.

VII.3. Адекватное пространство: симплекс концентраций

Концентрационные переменные, или численности популяций, образуют n -мерное открытое пространство $\mathbb{R}^n: \{x_1, x_2, \dots, x_n; -\infty < x_i < \infty, i=1, 2, \dots, n\}$, лишь часть которого имеет физический смысл:

$$\mathbb{X}^n \subset \mathbb{R}^n; \quad \mathbb{X}^n: \{x_1, x_2, \dots, x_n; x_i \geq 0, i=1, 2, \dots, n\}. \quad (45)$$

Все концентрационные переменные можно просуммировать; сумма представляет собой неотрицательную и конечную общую концентрацию c :

$$c = \sum_{i=1}^n x_i, \quad 0 \leq c < \infty,$$

которая используется для нормировки:

$$\xi_i = \frac{x_i}{c}; \quad 0 \leq \xi_i \leq 1; \quad \sum_{i=1}^n \xi_i = 1. \quad (46)$$

Благодаря свойствам своих переменных \mathbb{X}^n можно изоморфно отобразить на единичный симплекс S_n для каждого данного значения $c = c_0$. Соответствующее пространство будет обозначаться S^n :

$$c = c_0: \mathbb{X}^n \leftrightarrow S^n: \{\xi_1, \xi_2, \dots, \xi_n\}. \quad (47)$$

Единичный симплекс S_n — это правильный многогранник с n вершинами в соответствующем $(n-1)$ -

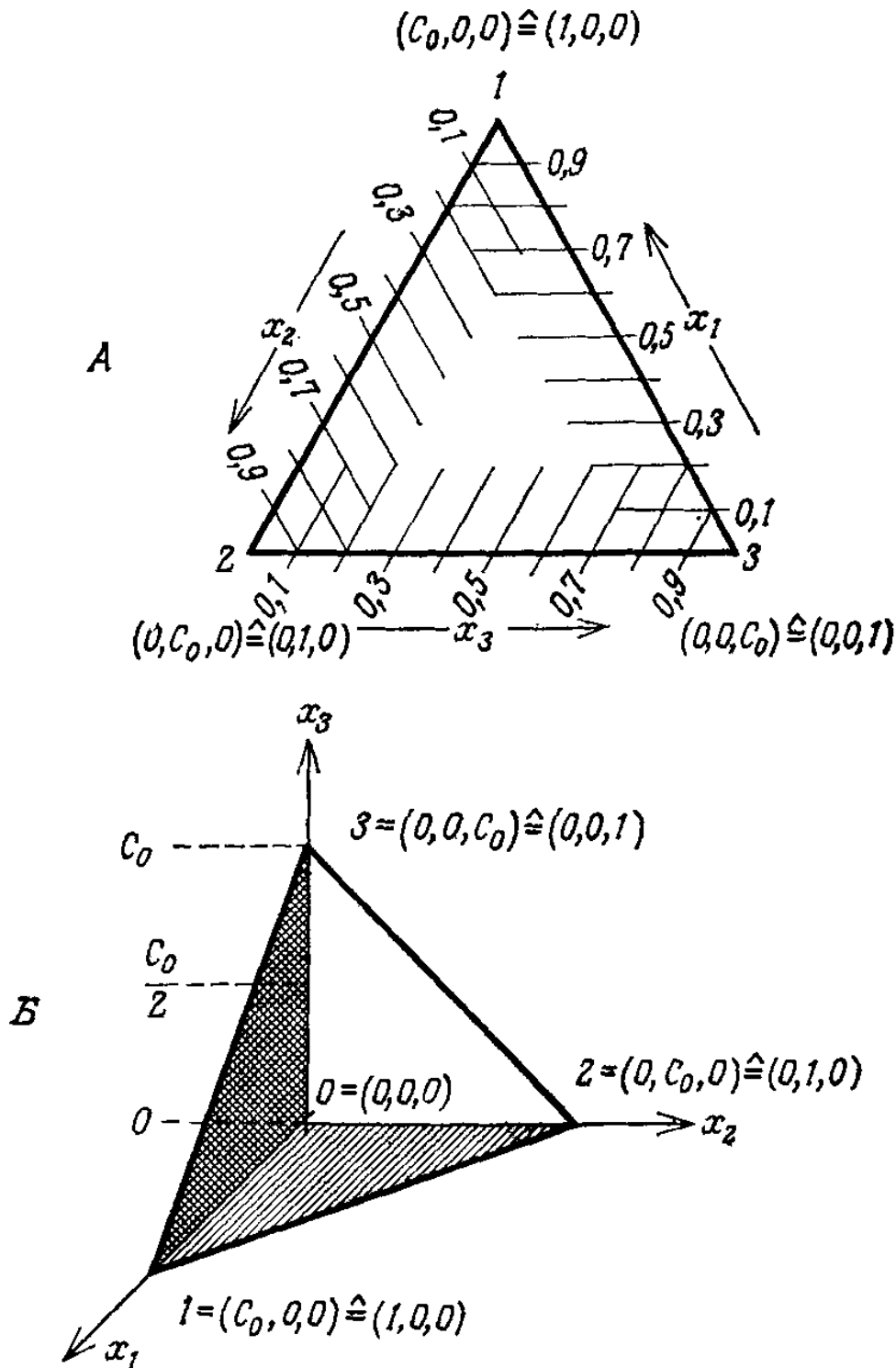


Рис. 20. Диаграмма А изображает симплекс S_3 , а диаграмма Б показывает, как этот симплекс вписывается в пространство концентраций, имеющих физический смысл. Для некоторых точек указаны соответствующие суммарные концентрации $c_0 = \sum_i x_i$ или координаты x_1, x_2, x_3 и ξ_1, ξ_2, ξ_3 (в скобках).

мерном подпространстве, определенном условием

$\sum_{i=1}^n \xi_i = 1$. Ребра симплекса имеют единичную длину

и представляют собой координатные оси для переменных ξ_i . В качестве иллюстрации на рис. 20

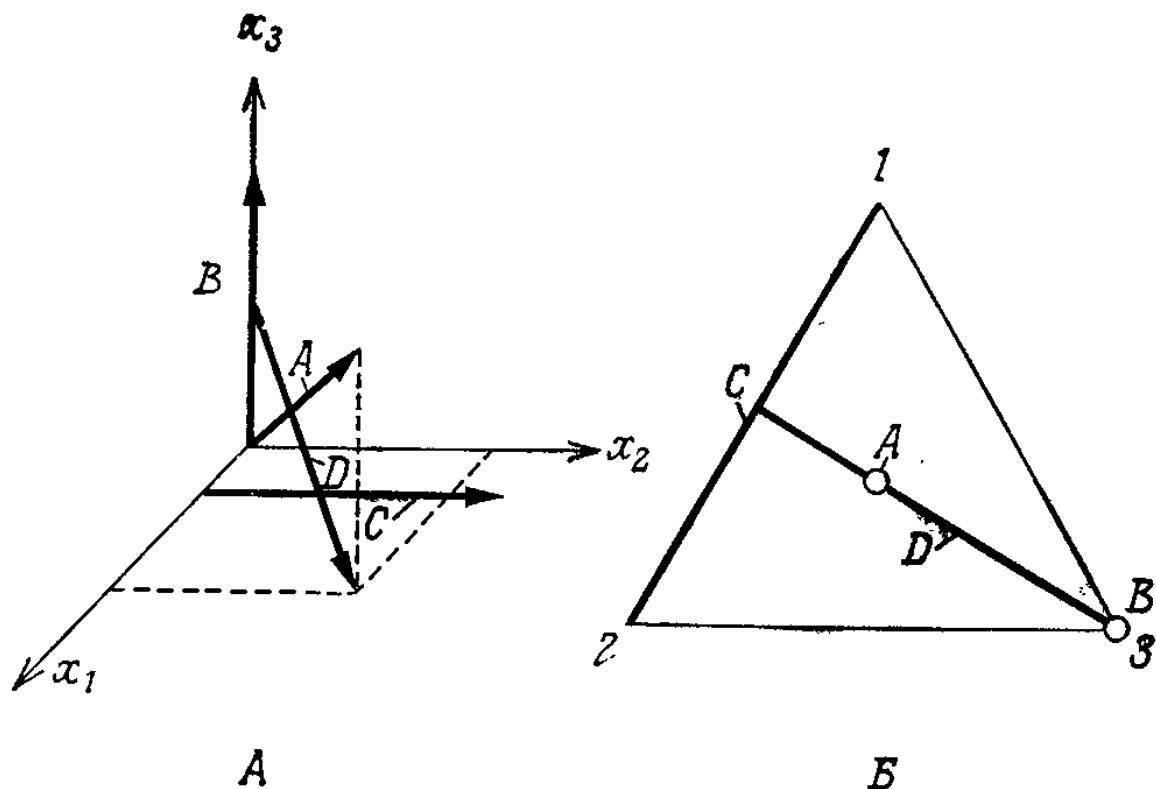


Рис. 21. Точки, положения которых удовлетворяют характерным зависимостям от суммарной концентрации c_0 в пространстве концентраций $\mathbb{X}_3(A)$ и на симплексе S_3 (B). $A = (c_0/3, c_0/3, c_0/3)$, $B = (0, 0, c_0)$, $C = (1, c_0 - 1, 0)$ и $D = (c_0 - 1, c_0 - 1, 2 - c_0)$. Стрелки на A показывают, куда движутся точки при увеличении суммарной концентрации (заметьте, что все множество точек, у которых все координаты пропорциональны c_0 , — например, A и B — отображаются в одну и ту же точку на S_3).

показан симплекс S_3 . Диаграммы на S_3 знакомы химикам по изображениям тройных систем. Вследствие уравнения (46) динамическая система на единичном симплексе потеряла одну степень свободы по сравнению с \mathbb{X}^n . Другими словами, переменные ξ_i из-за нормировки всегда относятся к фиксированному значению $c = c_0$, и тем самым вводится одна линейная зависимость между переменными.

Наконец, мы хотели бы подчеркнуть различие между картами на \mathbb{X}^n и S^n , которое становится очевидным при сравнении результатов, полученных для различных значений c_0 . Размеры симплекса S_n фиксированы вследствие нормировки, тогда как размеры области концентраций \mathbb{X}^n , имеющей физический смысл, варьируют с c_0 . Положения и нормальные моды особых точек в общем случае также будут зависеть от c_0 . Для полного описания асимптотиче-

ского поведения динамической системы необходимо построить карту особых точек, которые сами являются «функциями» суммарной концентрации c_0 . Положения многих особых точек, как мы увидим дальше, зависят от концентраций очень просто: их координаты пропорциональны c_0 . При изменении суммарной концентрации c_0 эти точки движутся вдоль прямых, проходящих через начало координат в X^n (см. рис. 21), и, следовательно, их образами являются отдельные точки в S_n . Соответственно карта особых точек в целом становится намного проще. Эта формальная зависимость карты особых точек от значения суммарной концентрации c_0 будет иметь особое значение при анализе растущих систем.

Высокосимметричная часть определенной $(n - 1)$ -мерной гиперплоскости, погруженной в n -мерное пространство концентраций, называется единичным симплексом. Пример симплекса в трехмерном пространстве дан на рис. 20. Единичный симплекс включает в себя всю область концентраций, имеющую физический смысл, и более всего пригоден для графического представления процессов отбора.

VII.4. Исследование нормальных мод

Начиная исследование общей системы линейных дифференциальных уравнений, мы сначала должны определить особые точки из условия $\dot{x}_i = 0$. Для прямого исследования динамической системы важно знать все особые точки в исследуемой области. Однако в общем случае этой информации недостаточно. Траектории n -мерной динамической системы часто заканчиваются в стоках. Однако могут существовать устойчивые замкнутые орбиты или странные аттракторы, о существовании которых можно судить на основании тщательного исследования природы областей, окружающих особые точки, и исследования векторных полей. Например, устойчивые предельные циклы в двух измерениях удается идентифицировать с помощью карт Пуанкаре. Информацию о природе особых точек можно получить в результате исследования нормальных мод.

Для этой цели динамическую систему линеаризуют в окрестности данной особой точки \bar{x} :

$$\dot{z}_i = \Lambda_i(\bar{x}) + \sum_{j=1}^n A_{ij} z_j + O(|z|^2). \quad (48)$$

Новые переменные z_i определяются следующим образом:

$$z_i = x_i - \bar{x}_i \quad \text{или} \quad z = x - \bar{x}. \quad (49)$$

Коэффициенты A_{ij} являются элементами матрицы Якоби (A), определенной в особой точке \bar{x} :

$$A_{ij} = \left(\frac{\partial \Lambda_i}{\partial x_j} \right)_{x=\bar{x}}. \quad (50)$$

Поскольку $\Lambda_i(\bar{x}) = 0$ по определению особой точки, линеаризованная система дифференциальных уравнений дается следующим выражением:

$$\dot{z} = A \cdot z. \quad (51)$$

Обратные постоянные времени, соответствующие нормальным модам, получаются теперь как собственные значения матрицы A. Собственные векторы ζ_i определяются в виде соответствующих линейных комбинаций концентрационных переменных:

$$A \cdot \zeta_j = \omega_j \cdot \zeta_j. \quad (52)$$

В общем случае ω_j являются комплексными величинами и определяют тип особой точки; наиболее важные типы были уже приведены на рис. 19.

Если матрица A не является сингулярной, устойчивая особая точка линеаризованной системы (51) почти во всех случаях соответствует устойчивой особой точке нелинейной системы [51]. Существуют, однако, некоторые важные исключения ($\text{Re } \omega_j = 0$): центр для линейной системы в нелинейном случае может стать спиральным стоком и *vice versa*. Примером поведения такого типа служит знаменитая модельная система Лотки — Вольterra [52]. В разд. VIII. 1 мы встретимся еще с одним примером — гиперциклом размерности $n = 4$.

Если для данной динамической системы получается несколько устойчивых особых точек, предельных циклов или других аттракторов, то желательно также определить бассейны, для которых аттракторы являются пределами траекторий при $t \rightarrow \infty$. Индивидуальные бассейны отделяются друг от друга сепаратрисами, которые в принципе можно определить интегрированием в обратном направлении ($t \rightarrow -t$), начиная с седловых точек и следуя линиям крутого спуска. Если для данной динамической системы известны все устойчивые особые точки и другие аттракторы, а также их бассейны, то мы можем предсказать результат процесса отбора, начинающегося с любой точки в данном пространстве концентраций.

В некоторых случаях мы можем получить $\operatorname{Re} \omega_j = 0$. Тогда линеаризация в окрестности особой точки не даст достаточной информации, и для полной характеристики необходимо вернуться к нелинейной динамической системе. Часто прямое исследование векторного поля в окрестности особой точки оказывается не очень трудным и дает нужные результаты.

Определение нормальных мод является существенной частью исследования особых точек. Оно представляет собой исследование траекторий динамической системы в малой окрестности особой точки. В большинстве случаев достаточно охарактеризовать устойчивость особой точки. Однако используемая при этом линейная аппроксимация иногда может не давать достаточной информации, и в таком случае требуются более тонкие методы анализа.

VII.5. Растущие системы

Из формулы (37) легко вывести дифференциальные уравнения для суммарной концентрации c :

$$\dot{c} = \sum_{j=1}^n \Gamma_j(\mathbf{x}) \left(1 - \frac{c}{c_0}\right), \quad (53)$$

где c_0 — стационарное значение суммарной концентрации, которая регулируется неспецифическим потоком ϕ_0 .

Очевидно, что это уравнение имеет особую точку при $c = c_0$. Собственное значение нормальной моды

$$\omega_c = -\frac{1}{c_0} \left(\sum_{j=1}^n \Gamma_j(\mathbf{x}) \right)_{c=c_0} \quad (54)$$

будет отрицательным, пока сумма всех Γ_i остается положительной. Итак, в точке $c = c_0$ мы имеем устойчивое стационарное состояние.

В некоторых системах карта особых точек, отражающая внутреннюю организацию этих систем, зависит также от суммарной концентрации c_0 . Теперь мы можем придать некоторый физический смысл нашему рассмотрению, ранее остававшемуся чисто математическим. Для этого допустим, что имеется

нестационарная динамическая система, которая начинает эволюционировать при $t = t_0$ с соответствующим начальным значением суммарной концентрации $c(t_0) = c_0$. Селекционные ограничения подбираются таким образом, чтобы суммарная концентрация $c(t)$ менялась медленно по сравнению с внутренними процессами в динамической системе, т. е. все изменения, обусловленные внешними процессами, происходят намного медленнее, чем изменения, обусловленные внутренней организацией системы. В каждый момент времени система будет находиться вблизи устойчивого решения (т. е. вблизи стока, устойчивой замкнутой орбиты или аттрактора другого вида). Когда приведенные выше условия выполнены, система подходит достаточно близко к асимптотическому решению, и процесс, зависящий от времени, может быть описан как последовательность стационарных решений с непрерывно изменяющейся суммарной концентрацией. Пользуясь более физическим языком, мы можем сказать, что динамическая система развивается при установившемся внутреннем равновесии. Как и следовало ожидать, анализ системы необыкновенно упрощается, если выполнено условие внутреннего уравнивания.

Внутреннее уравнивание в динамических системах с *однородными* функциями роста Γ_i легко исследовать, потому что в этом случае карта особых точек S_n не зависит от суммарной концентрации c_0 . При возрастании c_0 селекционное поведение не изменяется. Более того, в растущей однородной системе асимптотическое поведение не зависит от степени внутреннего уравнивания. Итак, в системах этого типа окончательный результат селекционного процесса будет одним и тем же независимо от того, установились ли во время роста внутренние равновесия. Существуют, однако, ситуации, когда концепция внутреннего уравнивания не может использоваться без тщательного исследования. При определенной критической суммарной концентрации $c = c_{кр}$ в карте особых точек могут произойти резкие изменения, например стоки могут стать неустойчивыми, устойчивые предельные циклы могут исчезнуть и т. д. Хорошо известная неустойчивость такого типа — это «бифуркация Хопфа» [58]. Внутренне уравновешенная динамическая система, которая приближается к такой точке с одной стороны, — например, растущая система, подходящая к критической концентрации со стороны меньших значений концентраций, — становится существенно неравновесной после того, как она пройдет критическую точку.

Для анализа динамических систем в окрестности подобных точек требуется специальный подход. Мы встретимся с такими примерами в разд. VII. Весьма общее исследование подобных ситуаций было проведено Томом [59]—имеется в виду его теория катастроф.

Конечно, с биофизической точки зрения такие сложные динамические системы более интересны. Ведь в самом деле, для появления организованных структур требуются резкие изменения, подобные упомянутым выше разрывностям в карте особых точек. Динамические системы, описывающие переходы между различными уровнями организации, с неизбежностью должны проходить через определенные критические стадии, или периоды. Для конкретности мы рассмотрим один важный пример из области самоорганизации биологических макромолекул: переход от множества независимых конкурентов к функциональной единице, состоящей из конкурирующих полинуклеотидов и белков. В соответствии с определением, данным в разд. 1.4, в конкурентной системе отбирается только один вид, и, следовательно, внутри S_n нет устойчивого аттрактора. С другой стороны, любая кооперативная система должна иметь такой аттрактор, иначе по крайней мере один из конкурирующих видов макромолекул вымрет через достаточно большое время. Следовательно, динамическая система, которая в принципе способна имитировать интересующее нас развитие от более хаотического к более организованному состоянию, должна содержать критическую неустойчивость при определенных значениях своих параметров.

VII.6. Анализ конкретных систем

a. Независимые конкуренты

Проиллюстрируем на конкретном примере, как проводится исследование особых точек. Возьмем задачу об отборе квазивида, о которой уже шла речь в части А. Результаты соответствующего математического исследования приведены в табл. 7. Координаты пространства концентраций даются нормальными переменными y_k ; собственные значения λ_k являются параметрами роста функций Γ_k . Анализ относится к данному распределению мутантов. Появление новых мутантов, дающих вклад в отобранный квазивид, будет изменять смысл концентрационных координат y_k , т. е. их связи с истинными концентрационными переменными x_k . Данные табл. 7 не нуждаются в пояснениях. В дальнейшем мы будем использовать их при сравнении трех функций роста $\Gamma_i = k_i x_i^p$, фигурирующих в табл. 6, т. е.:

Таблица 7

**Исследование особых точек для отбора квазивида
(см. часть А)**

Имеем следующее кинетическое уравнение:

$$\dot{y}_i = \lambda_i y_i - \frac{y_i}{c_0} \sum_{j=1}^n \lambda_j y_j \quad i = 1, 2, \dots, n.$$

Асимптотическое поведение определяется n особыми точками, расположенными в вершинах симплекса S_n .

$$\bar{y}_1 = \begin{pmatrix} c_0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad \bar{y}_2 = \begin{pmatrix} 0 \\ c_0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \bar{y}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ c_0 \end{pmatrix}.$$

Анализ нормальных мод для каждой особой точки \bar{y}_k дает спектр n значений $\omega_j^{(k)}$:

$$\begin{array}{ccc|c|ccc} \omega_j^{(1)} = \lambda_j - \lambda_1 & \omega_j^{(2)} = \lambda_j - \lambda_2 & & \dots & \omega_j^{(n)} = \lambda_j - \lambda_n \\ j=2, 3, \dots, n-1, n & j=1, 3, \dots, n-1, n & & & j=1, 2, \dots, n-2, n-1 \\ \omega_c^{(1)} = -\lambda_1 & \omega_c^{(2)} = -\lambda_2 & & & \omega_c^{(n)} = -\lambda_n \end{array}$$

Что касается степеней свободы симплекса S_n , то каждая особая точка \bar{y}_k имеет $n-1$ нормальных мод с обратимыми постоянными времени $\omega_j^{(k)}$, которые описывают внутреннюю организацию распределения, обусловленную конкуренцией между различными квазивидами. Далее, симплекс S_n имеет одну нормальную моду $\omega_c^{(k)}$, которая соответствует изменению суммарной концентрации c . Все внутренние моды $\omega_j^{(k)}$ равны разностям собственных значений λ . Следовательно, имеется только одна устойчивая особая точка для наибольшего собственного значения $\lambda_m > \lambda_j$, $j = 1, 2, \dots, n$, $j \neq m$. Это узловой сток, т. е. все значения $\omega_j^{(m)}$ различны и отрицательны. Соответственно квазивид с наименьшим собственным значением описывается источником — вследствие положительности значений ω_j . Оставшиеся $n-2$ особые точки являются седлами, потому что им соответствуют как положительные, так и отрицательные $\omega_j^{(k)}$.

- 1) постоянная скорость роста: $p=0$;
- 2) линейная скорость роста: $p=1$;
- 3) квадратичная скорость роста: $p=2$.

1. Первый случай дает одну устойчивую особую точку — фокальный сток внутри единичного симплекса S_n :

$$\bar{x} = \frac{c_0}{\sum_{j=1}^n k_j} \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}. \quad (55)$$

«Внутри» единичного симплекса означает, что для всех координат \bar{x} : $0 < \bar{x}_i < c_0$. (Отрицательное) собственное значение матрицы Якоби n -кратно вырождено:

$$\omega = - \frac{\sum_{j=1}^n k_j}{c_0}. \quad (56)$$

То же самое справедливо и для ω_c , которое относится к изменению суммарной концентрации c .

Результатом является устойчивое сосуществование всех видов.

2. Второй случай рассмотрим в табл. 7. Напомним, что имеется только одна устойчивая особая точка. Тот факт, что она расположена в вершине симплекса, указывает на коинкурентное поведение. Только одна из концентрационных координат узлового стока положительна ($=c_0$), все другие равны нулю. Как и в первом случае, карта не зависит от суммарной концентрации c_0 , а конечный результат не зависит от начальных условий.

3. Наконец, третий случай дает всего $2^n - 1$ особых точек, которые можно сгруппировать в три класса.

Первый класс включает n фокальных стоков, по одному в каждой вершине S_n .

$$\bar{x} = \begin{pmatrix} 0 \\ \vdots \\ x_k = c_0 \\ \vdots \\ 0 \end{pmatrix}, \quad \text{причем} \quad \begin{aligned} \omega_j^{(k)} &= -k_k c_0, \quad j = 1, 2, \dots, n-1 \\ \omega_c^k &= -k_k c_0 \end{aligned} \quad (57)$$

Других устойчивых особых точек нет. Расположение точек в вершинах единичного симплекса снова указывает на коинкурентное

поведение, позволяющее выжить лишь одному конкуренту, т. е. данная ситуация соответствует чистому состоянию. Однако в этом случае нелинейных скоростей роста результат конкуренции зависит от начальных условий, потому что имеется n устойчивых точек (в противоположность линейному автокатализатору, когда особая точка только одна). Это означает, что каждый из n конкурентов может решить спор в свою пользу — в зависимости от начальных численностей популяции. Как только победитель утвердился, любой из конкурентов уже не может легко вытеснить его. Поэтому мы называем эту ситуацию «отбором раз и навсегда». Как и в двух предыдущих случаях, карта особых точек не зависит от суммарной концентрации c_0 .

Два других класса особых точек включают в себя один источник внутри единичного симплекса (все координаты его конечны) и $(2^n - n - 2)$ седловых точек — по одной на каждом ребре и по одной на каждой грани (включая все возможные гиперграни) S_n . Оба класса особых точек соответствуют неустойчивостям. Мы не приводим их координат и нормальных мод — они могут быть получены простыми вычислениями. Вместо этого мы иллюстрируем типичное селекционное поведение растущих систем на нескольких примерах единичных симплексов размерности 3 (см. рис. 22).

Мы выбрали эти три сравнительно простых модельных случая, чтобы проиллюстрировать метод исследования особых точек и подчеркнуть те его свойства, на которые следует обратить внимание. Природа особой точки, в частности то, какое решение она дает — устойчивое или неустойчивое, — имеет первостепенное значение для проблем отбора и эволюции. Не менее важна локализация особых точек в единичном симплексе. Кооперативный отбор множества репликативных единиц требует, чтобы особая точка лежала внутри единичного симплекса S_k , относящегося к подпространству X^k , образованному концентрационными координатами k кооперирующихся единиц. С другой стороны, локализация стока на одной из вершин S_k характеризует конкуренцию, ведущую к отбору только одного из компонентов, в то время как локализации на ребрах, гранях или гипергранях указывают на частичную конкуренцию и отбор.

Построение аппарата трансляции, например, требует одновременного отбора нескольких репликативных единиц — предшественников различных генов. Ни одна из трех систем, рассмотренных выше, не

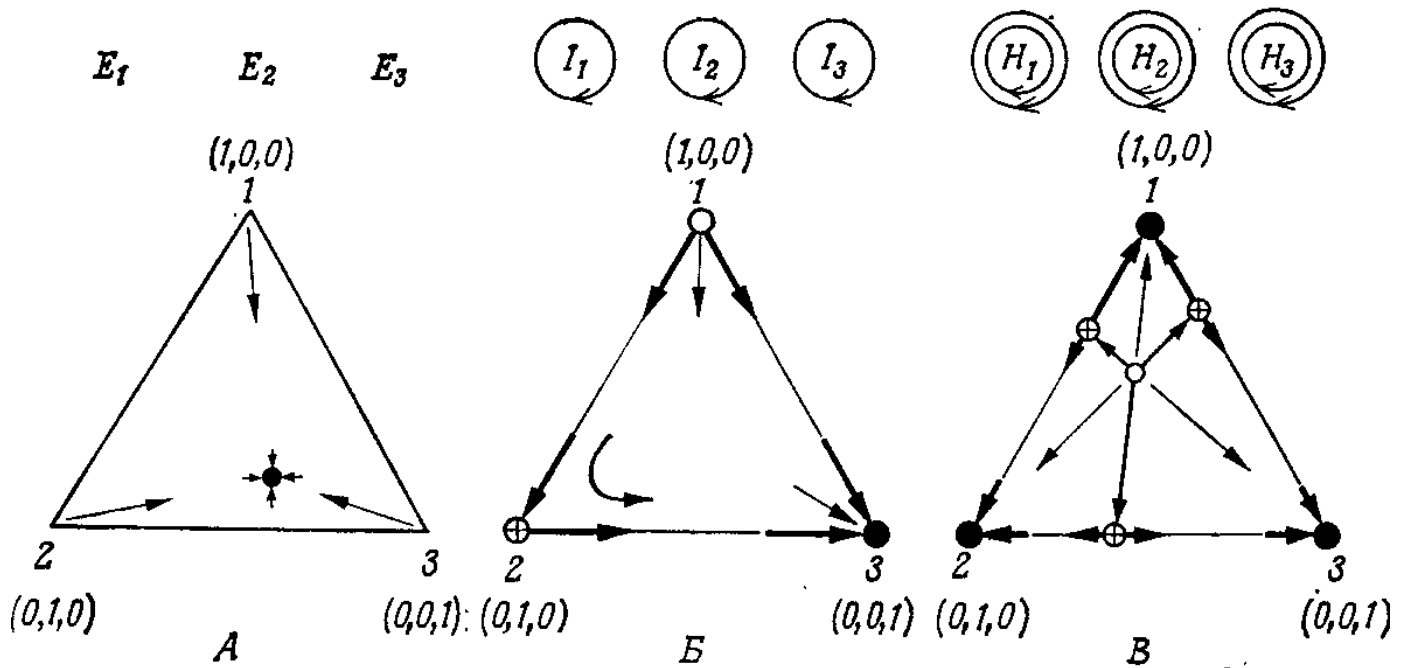


Рис. 22. Трехмерные карты особых точек для различных типов независимых конкурентов при ограничении постоянной организации. (Символы E , I и H были введены на рис. 10.)

А. Постоянная скорость роста ($p = 0$):

$$\dot{x}_i = k_i - \frac{x_i}{c_0} \sum_{j=1}^3 k_j$$

$$k_1 = 1; \quad k_2 = 2; \quad k_3 = 3.$$

На карте имеется фокус внутри единичного симплекса S_3 , что означает устойчивое сосуществование всех трех видов. Легко изобразить все многообразие траекторий — прямых линий, проходящих через каждую точку S_n и ведущих в устойчивый фокус.

Б. Линейная скорость роста ($p = 1$):

$$\dot{x}_i = k_i x_i - \frac{x_i}{c_0} \sum_{j=1}^3 k_j x_j; \quad k_1 = 1; \quad k_2 = 2; \quad k_3 = 3.$$

Единственное устойчивое асимптотическое решение системы — это чистое состояние, которое включает в себя только вид 3. За исключением двух ребер, $\overline{12}$ и $\overline{23}$, все траектории начинаются в точке 1 и кончаются в точке 3.

В. Квадратичная скорость роста ($p = 2$):

$$\dot{x}_i = k_i x_i^2 - \frac{x_i}{c_0} \sum_{j=1}^3 k_j x_j^2; \quad k_1 = 1; \quad k_2 = 2; \quad k_3 = 3.$$

Симплекс S_3 разбивается на три области, каждая из которых является бассейном устойчивой особой точки. Размеры бассейнов зависят от значений соответствующих констант скоростей. Поскольку наибольшей из них является k_3 , максимальные размеры бассейна имеет особая точка \bar{x}_3 .

удовлетворяет требованиям такого одновременного отбора. Первая система, по-видимому, допускает существование, но она не селекционна и поэтому не может эволюционировать к оптимальному функционированию. Вторая система допускает существование лишь в узких пределах распределения квазивида: она не толерантна к дивергенции генотипов, что требуется для облегчения фенотипической диверсификации. И наконец, третья система в высшей степени антикооперативна — настолько, что однажды установившийся вид подавляет в процессе отбора любого мутанта независимо от того, обладает ли он селективным преимуществом.

Следуя указаниям, вытекающим из сравнительного обзора в разд. V, мы проанализируем теперь более подробно ансамбли с функциональными связями. Эти ансамбли будут содержать репликативные единицы в целях сохранения генетической информации и в то же время они будут кооперативно стабилизироваться связями, которые делают функцию роста существенно нелинейной. Поэтому ожидаемые свойства системы со связями будут в какой-то мере сходны со свойствами, характерными для третьего примера независимых конкурентов.

б. Каталитические цепи

Самый прямой способ установления связи между всеми членами ансамбля — это построение цепи посредством связующих реакций, аналогично тому как мы связываем слова в предложения (рис. 23).

Члены кинетических уравнений, соответствующие этим связям, приводят к неоднородности функций чистого роста Γ_i для всех членов, кроме первого:

$$\dot{x}_1 = k_1 x_1 - \frac{x_1}{c_0} \left[k_1 x_1 + \sum_{j=2}^n (k_j x_j + k'_j x_j x_{j-1}) \right],$$

$$\dot{x}_i = k_i x_i + k'_i x_i x_{i-1} - \frac{x_i}{c_0} \left[k_1 x_1 + \sum_{j=2}^n (k_j x_j + k'_j x_j x_{j-1}) \right] \quad (58)$$

для $i = 2, 3, \dots, n$.

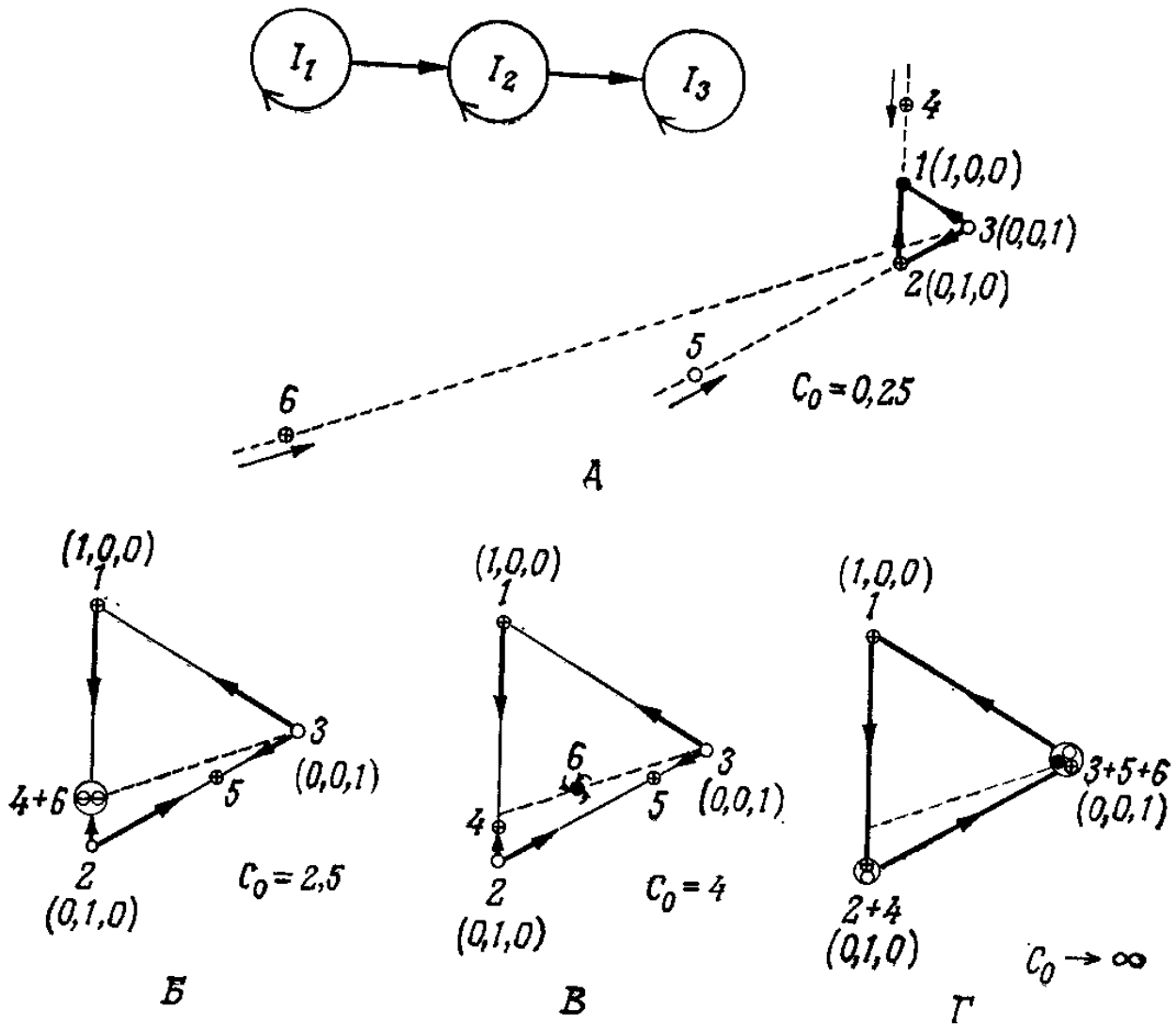


Рис. 23. Карты особых точек каталитической цепи самореплицирующихся единиц $\textcircled{1}_i$ при ограничении постоянной организации:

$$\Gamma_1 = k_1 x_1;$$

$$\Gamma_i = k_i x_i + k'_i x_i x_{i-1} \quad (\text{для } i = 2, 3);$$

$$k_1 = 3; \quad k_2 = 2; \quad k_3 = 1; \quad k'_2 = 2; \quad k'_3 = 1.$$

$$1 = \bar{x}_1; \quad 2 = \bar{x}_2 \dots 6 = \bar{x}_6,$$

При малых концентрациях (A) устойчивое решение соответствует отбору вида 1. Однако если два других вида еще не вымерли к тому моменту, когда суммарная концентрация достигает критического значения, то возникает новое стационарное состояние, в котором все три вида становятся устойчивыми (B). Дальнейшее увеличение суммарной концентрации (B) благоприятствует только виду 3, так что конечная ситуация (Gamma) отвечает отбору этого вида. Однако механизм данного отбора отличается от механизма отбора в случае независимых конкурентов.

Анализ особых точек для каталитических цепей размерности три

На карте особых точек имеется шесть особых точек с координатами и нормальными модами, приведенными ниже:

$$\bar{x}_1 = \begin{pmatrix} c_0 \\ 0 \\ 0 \end{pmatrix}$$

$$\bar{x}_2 = \begin{pmatrix} 0 \\ c_0 \\ 0 \end{pmatrix}$$

$$\bar{x}_3 = \begin{pmatrix} 0 \\ 0 \\ c_0 \end{pmatrix}$$

$$\omega_1^{(1)} = k_3 - k_1$$

$$\omega_1^{(2)} = k_1 - k_2$$

$$\omega_1^{(3)} = k_1 - k_3$$

$$\omega_2^{(1)} = k_2'c_0 - k_1 + k_2$$

$$\omega_2^{(2)} = k_3'c_0 - k_2 + k_3$$

$$\omega_2^{(3)} = k_2 - k_3$$

$$\bar{x}_4 = \begin{pmatrix} \frac{k_1 - k_2}{k_2'} \\ \frac{k_2'c_0 - k_1 + k_2}{k_2'} \\ 0 \end{pmatrix}$$

$$\bar{x}_5 = \begin{pmatrix} 0 \\ \frac{k_2 - k_3}{k_3'} \\ \frac{k_3'c_0 - k_2 + k_3}{k_3'} \end{pmatrix}$$

$$\bar{x}_6 = \begin{pmatrix} \frac{k_1 - k_2}{k_2'} \\ \frac{k_1 - k_3}{k_3'} \\ k \end{pmatrix}$$

$$\omega_1^{(4)} = \frac{(k_2'c_0 - k_1 + k_2)(k_2 - k_1)}{k_2'c_0}$$

$$\omega_1^{(5)} = k_1 - k_2$$

$$k = \frac{k_2'k_3'c_0 - k_3'(k_1 - k_2) - k_2'(k_1 - k_3)}{k_2'k_3'}$$

$$\omega_2^{(4)} = \frac{1}{k_2'} \{ k_2'k_3'c_0 -$$

$$\omega_2^{(5)} = \frac{(k_3'c_0 - k_2 + k_3)(k_3 - k_2)}{k_3'c_0}$$

$\omega_1^{(6)}$ и $\omega_2^{(6)}$ — собственные значения матрицы Якоби $A(x = \bar{x}_6)$

$$- k_3'(k_1 - k_2) - k_2'(k_1 - k_3) \}$$

Три особые точки \bar{x}_1 , \bar{x}_2 и \bar{x}_3 — лежат в вершинах единичного симплекса S_3 (см. рис. 23), и, следовательно, имеет место конкуренция независимо от природы особых точек. Положения трех других особых точек зависят (линейно) от суммарной концентрации c_0 . Две особые точки \bar{x}_4 и \bar{x}_5 — движутся вдоль ребер $\bar{12}$ и $\bar{23}$ симплекса, что указывает на частичную конкуренцию. Лишь особая точка \bar{x}_6 может перемещаться внутри S_3 , что означает кооперативный отбор всех членов цепи.

При низкой суммарной концентрации

$$c_0 < (k_1 - k_2)/k_2' \cdot (k_2 - k_3)/k_3' \text{ или } (k_1 - k_3)/k_3'$$

особые точки \bar{x}_4 , \bar{x}_5 или соответственно \bar{x}_6 находятся вне симплекса S_3 , т. е. вне той области пространства концентраций, которая имеет физический смысл (по крайней мере одна концентрационная координата отрицательна). При $c_0 \rightarrow 0$ координаты этих особых точек даже стремятся к бесконечности. Динамическая система асимптотически становится идентичной системе экспоненциально растущих (несвязанных) конкурентов, которая характеризуется особыми точками \bar{x}_1 , \bar{x}_2 и \bar{x}_3 .

Если $k_1 > k_2$, k_3 и c_0 лежит выше порогового значения, которое дается суммой $[(k_1 - k_2)/k_2'] + [(k_1 - k_3)/k_3']$, то особая точка \bar{x}_6 оказывается внутри единичного симплекса, что указывает на кооперативное поведение. Однако она не стремится ни к какой точке внутри S_3 , а перемещается к вершине 3.

Из-за отсутствия однородности карты особых точек будут иметь более сложный вид, чем в рассмотренных до сих пор случаях.

Чтобы эта процедура была понятной, начнем с трехмерной системы, а затем распространим анализ на многомерный случай. В табл. 8 приведена сводка необходимых соотношений для трехмерного случая, а также кратко охарактеризованы карты особых точек. В соответствии с этим анализом три члена ($I_1 - I_3$) линейной цепи самовоспроизводящихся единиц могут быть отобраны одновременно лишь при очень специальных условиях, а именно:

$$k_1 > k_2, k_3 \quad (59)$$

и

$$c_0 > \frac{k_1 - k_2}{k_2'} + \frac{k_1 - k_3}{k_3'} \quad (60)$$

Кажется очень мало вероятным, чтобы партнеры, которые оказались удовлетворяющими условию (59), продолжали удовлетворять ему на протяжении длительных периодов эволюции [это означало бы, что мутации, изменяющие соотношение (59), никогда не происходят]. Если бы они были способны к этому, система развивалась бы крайне асимметричным образом: с ростом c_0 увеличивалась бы только численность популяции последнего члена цепи — по крайней мере при селекционных ограничениях. Поскольку ясно, что это скоро привело бы к расхождению численностей популяций на целые порядки величины, можно сделать вывод, что такая система не способна стабилизировать совместное функционирование, поскольку она не может контролировать относительные численности популяций в большом интервале суммарных концентраций.

Это поведение иллюстрирует рис. 23, на котором представлено как бы несколько моментальных снимков непрерывного процесса в системе, растущей в состоянии, близком к внутреннему равновесию. Для концентраций c_0 ниже критического уровня, заданного уравнением (60), три особые точки — x_4 , x_5 , x_6 — расположены вне единичного симплекса (рис. 23, А). Если c_0 равно критическому значению, то особая точка x_6 достигает границы симплекса (рис. 23, Б) и с ростом c_0 движется внутри его. При этом она изменяет свою природу — теперь это устойчивая особая точка (рис. 23, В), которая в данном конкретном случае является спиральным стоком. (Более детальное исследование особых точек в случае неоднородных функций роста будет проведено в работе [53].) Рис. 23, Г иллюстрирует окончательную судьбу этой устойчивой особой точки, а именно миграцию в вершину 3. Тем самым система приходит к чистому состоянию $x_3 = c_0$.

Основные результаты, полученные для трехмерного случая, легко обобщаются для n -мерной системы. Роль вида 3 играет вид n , вместо шести имеется $2n$ особых точек. Самая интересная

особая точка — это \bar{x}_{2n} . Ее положение легко определить:

$$\bar{x}_{2n} = \begin{pmatrix} \frac{k_1 - k_2}{k_2'} \\ \frac{k_1 - k_3}{k_3'} \\ \vdots \\ \vdots \\ c_0 - \sum_{j=2}^n \frac{k_1 - k_j}{k_j'} \end{pmatrix}. \quad (61)$$

Особая точка \bar{x}_{2n} лежит внутри симплекса S_n в том и только в том случае, если константы скорости удовлетворяют соотношениям $k_1 > k_j$, $j = 2, 3, \dots, n$, и суммарная концентрация превосходит критическое значение

$$c_{\text{кр}} = \sum_{j=2}^n \frac{k_1 - k_j}{k_j'}. \quad (62)$$

Тогда \bar{x}_{2n} соответствует устойчивому стационарному состоянию. В этом состоянии все концентрации, кроме x_n , постоянны, и поэтому при больших суммарных концентрациях система приближается к чистому состоянию $x_n = c_0$.

Резюмируем поведение каталитических цепей.

1. Устойчивые стационарные состояния существуют только в том случае, если константы скорости и суммарная концентрация удовлетворяют определенным соотношениям:

$$k_1 > k_j; \quad j = 2, 3, \dots, n; \quad c_0 > \sum_{j=2}^n \frac{k_1 - k_j}{k_j'}.$$

Чтобы происходил отбор, элиминирующий другие нефункциональные единицы, на систему следует наложить селекционные ограничения, и отбор благоприятных мутантов не должен изменять требуемых неравенств, которым должны удовлетворять константы скорости.

2. Если условия п. 1 выполнены, то концентрации индивидуальных видов будут сравнимы по величине только в довольно узкой области суммарных

концентраций. С ростом c_0 последний член цепи I_n в (квази)стационарных условиях растет и, наконец, становится доминирующим.

Следовательно, каталитическая цепь вряд ли может служить системой, интегрирующей информацию.

в. Разветвленные системы

При эволюции систем со связями неизбежно будет происходить разветвление связей (рис. 24). Исследование особых точек таких разветвленных систем не обнаруживает каких-либо неожиданных новых особенностей. При очень малых суммарных концентрациях три вида ведут себя как независимые конкуренты. Теперь имеется два критических значения c_0 , при которых либо I_1 и I_2 , либо I_1 и I_3 сосуществуют. Какая именно из этих двух ситуаций реализуется, зависит от того, кому больше благоприятствует I_1 — I_2 или I_3 . Одна из двух особых точек оказывается устойчивым узлом, другая — седловой точкой. При более высокой суммарной концентрации устойчивая особая точка снова мигрирует по направлению к одной из вершин — 2 или соответственно 3. Такое поведение иллюстрирует рис. 24.

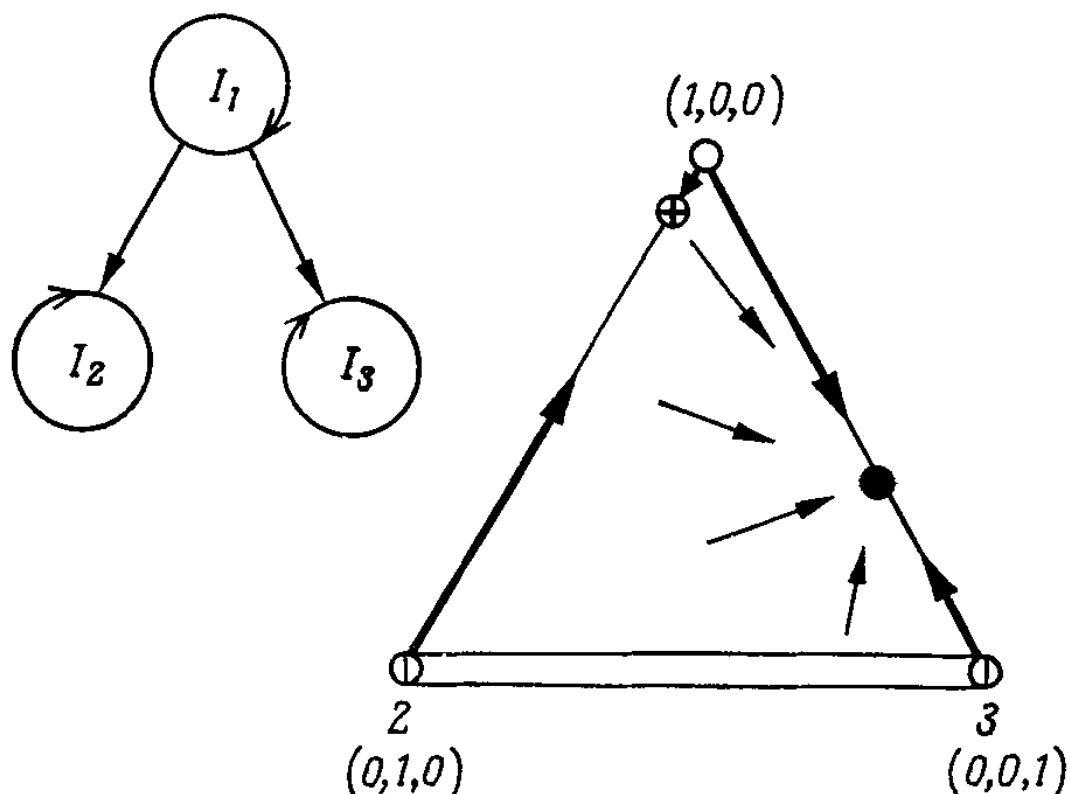


Рис. 24. Карта особых точек динамической системы, представляющей собой место разветвления в каталитической системе самореплицирующихся единиц (I_i) при ограничении постоянной организации.

$$\Gamma_1 = k_1 x_1; \quad \Gamma_i = k_i x_i + k'_i x_i x_1 \quad (\text{для } i = 2, 3),$$

$$k_1 = 3; \quad k_2 = k_3 = 0; \quad k'_2 = 1; \quad k'_3 = 2; \quad c_0 = 3,5,$$

Исследованию здесь трехмерную систему можно обобщить двумя способами:

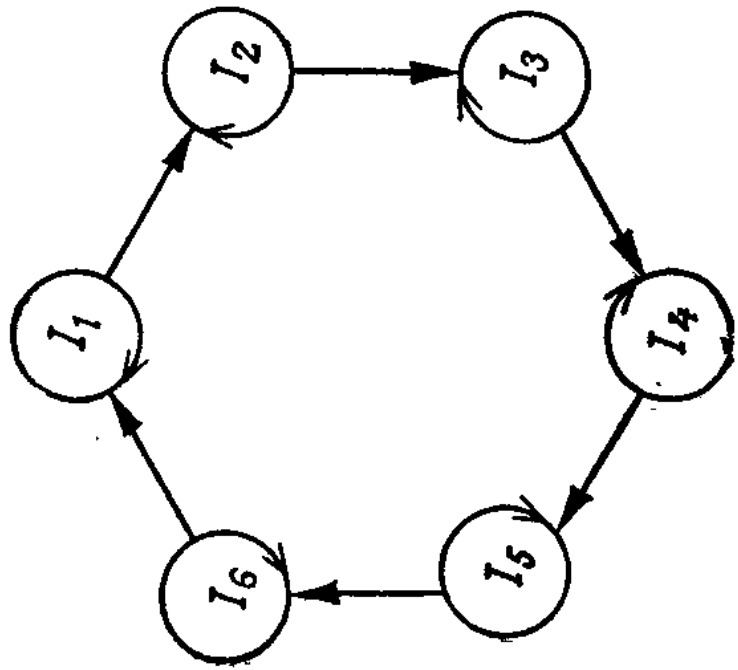
1. Из данной точки может начинаться более двух ветвей.
2. Отдельные ветви могут состоять из нескольких членов.

Исследование особых точек этих многомерных систем приводит в сущности к таким же результатам, как и в случае трех измерений. Их можно резюмировать следующим образом. Разветвленные системы самореплицирующихся единиц не являются устойчивыми на протяжении больших интервалов времени. Ветвь, рост которой наиболее эффективен, будет все больше доминировать, в то время как другие ветви будут исчезать. В конце концов останется только наиболее эффективная линейная цепь, и тем самым вся проблема сведется к динамической системе типа (58), которая уже рассматривалась в предыдущем разделе.

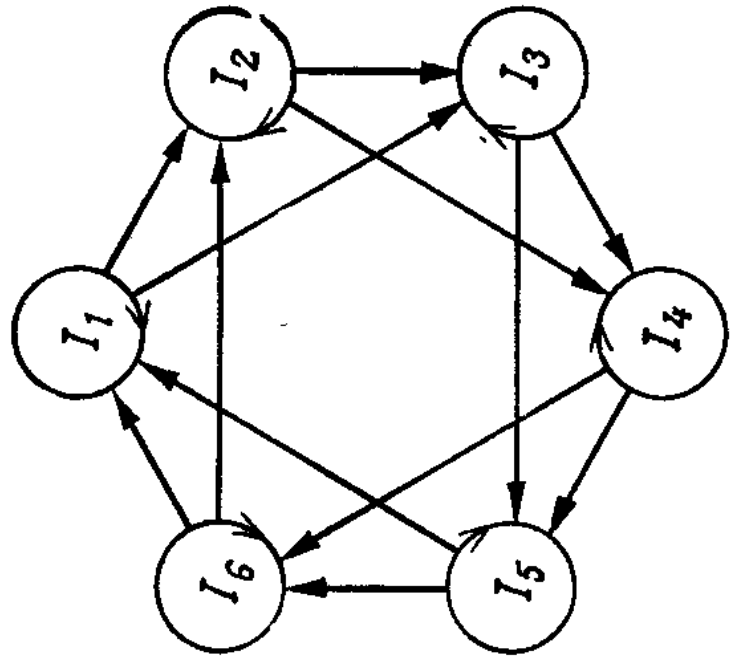
VII.7. Исследование особых точек гиперциклов

а. Классификация

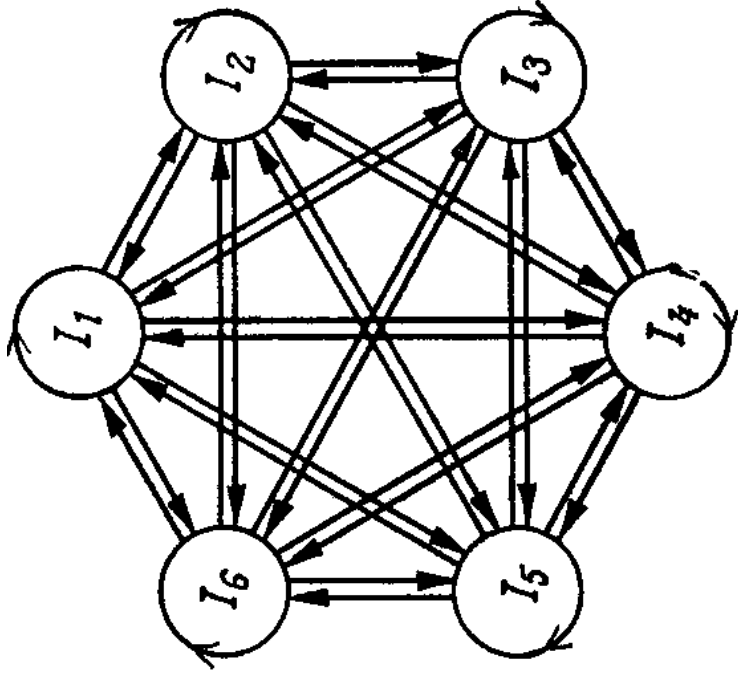
Как мы видели в части А, замыкание цикла в динамических системах приводит к появлению у системы в целом совершенно новых свойств. Множество молекул, которые образуются в замкнутом цикле химических реакций, эквивалентно катализатору. Цикл катализаторов в свою очередь имеет автокаталитические свойства (рис. 4), и его можно считать самореплицирующейся системой. Мы установили, что линейные или разветвленные связи между самореплицирующимися единицами не приводят к отбору объединенной системы с функциональными связями; теперь можно задать вопрос, не сопровождается ли замыкание цикла в цепи связей изменением характера селекционного поведения всего ансамбля? Есть основания ответить на этот вопрос утвердительно, поскольку мы знаем, что в открытых цепях реципиентом всех преимуществ связей всегда был последний член.



A



B



B

Рис. 25. Каталитические гиперциклы степени $p = 2$, $n = 6$ (A), степени $p = 3$, $n = 6$ (B) и степени $p = n = 6$ (B).

Общая классификация гиперциклов дана в части А. Простейшие представители этого класса сетей получаются в результате введения простой функциональной связи между самореплицирующимися единицами, как показано на рис. 25.

Данный раздел, посвященный гиперциклам, можно подразделить на три части. Сначала мы введем некоторые определения и критерии, полезные для классификации этого нового типа каталитических систем. Далее опишем результаты исследования особых точек для наиболее важных «чистых» типов гиперциклов. И наконец, рассмотрим один пример самоорганизующейся системы, которая представляет собой реалистический каталитический гиперцикл.

Прежде всего, гиперциклы отличаются от обычных каталитических циклов наличием нелинейных членов в выражениях для скоростей роста. В простых случаях функции Γ_i являются произведениями концентраций:

$$\Gamma_i = k_i \prod_{\lambda=1}^n x^{p_{\lambda i}}. \quad (63)$$

Показатели $p_{\lambda i}$ согласно (63), можно считать элементами матрицы \mathbf{P} . Индексы λ и i указывают, какая популяционная переменная x_λ в функции Γ_i должна возводиться в степень $p_{\lambda i}$. Следовательно, динамическая система полностью определяется матрицей показателей \mathbf{P} , вектором констант скоростей $\mathbf{k} = (k_1, \dots, k_n)$ и множеством начальных условий. Сначала мы рассмотрим только «чистые» случаи, которые характеризуются тем, что Γ_i являются однородными функциями. Требование однородности приводит к первому ограничению, налагаемому на элементы матрицы \mathbf{P} :

$$\sum_{\lambda=1}^n p_{\lambda i} = p; \quad i = 1, 2, \dots, n. \quad (64)$$

Сумма p теперь остается одной и той же для всех n дифференциальных уравнений и представляет собой степень функций роста, введенную в разд. V. Кроме условия однородности, мы потребуем, чтобы отдельные концентрации входили в Γ_i только в первой степени. Некоторые важные случаи зависимостей более высокого порядка будут рассмотрены ниже. Соответственно показатели $p_{\lambda i}$ имеют только два возможных значения: $p_{\lambda i} = \{0, 1\}$.

Наконец, введем циклическую симметрию в функцию чистого роста:

$$\Gamma_i = k_i x_i x_j x_k x_l \dots x_r,$$

$$j = i - 1 + n(\delta_{i1}); \quad l = i - 3 + n(\delta_{i1} + \delta_{i2} + \delta_{i3}); \dots \quad (65)$$

$$k = i - 2 + n(\delta_{i1} + \delta_{i2}); \quad r = i - p + n \sum_{\mu=1}^{p-1} \delta_{i\mu}.$$

Основные особенности реакций не зависят от предположения о циклической симметрии. Вместе с тем это предположение является разумным, если нет дополнительной информации о структурных различиях между кинетическими уравнениями для отдельных членов циклической системы. Теперь матрица P имеет общую форму простого вида. Ниже приведен конкретный пример — матрица P , где $n = 6$ и $p = 3$:

$$P(n=6, p=3) = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \quad (66)$$

Итак, гиперциклы с циклической симметрией и однородными функциями роста Γ_i полностью определяются значениями n и p и вектором k .

На рис. 25 показаны схематические диаграммы для трех гиперциклов с $n=6$ и $p=2, 3$ и 6 . Случай с $p=1$ следует исключить из общего класса каталитических систем, называемых гиперциклами, так как они относятся к категории систем с линейными скоростями роста Γ_i .

б. Общий анализ

Сводка результатов исследования особых точек гиперциклических систем дана в табл. 9.

Мы обсудим два случая, которые являются для нас наиболее важными

1. Простейший гиперцикл с $p=2$.
2. Гиперцикл, использующий каталитические связи между всеми членами, т. е. $p_{\lambda i} = 1$ для $i = 1, \dots, n$ и $\lambda = 1, \dots, n$, и, следовательно, $p=n$.

Таблица 9

Карта особых точек гиперцикла

Налагая на динамическую систему (65) ограничение постоянной организации, получаем

$$\dot{x}_i = k_i x_i x_j \dots x_l - \frac{x_i}{c_0} \sum_{r=1}^n k_r x_r x_s \dots x_t,$$

$$j = i - 1 + n \delta_{i1}, \dots, l = i - p + 1 + n \sum_{\mu=1}^{p-1} \delta_{i\mu};$$

$$s = r - 1 + n \delta_{r1}, \dots, t = r - p + 1 + n \sum_{\mu=1}^{p-1} \delta_{r\mu}; \quad p \leq n \quad (\text{T.9.1})$$

Исследование особых точек можно провести *аналитически* для любого n , если все константы скоростей одинаковы:

$$k_1 = k_2 = \dots = k_n = k. \quad (\text{T.9.2})$$

(Влияние вариаций отдельных констант скоростей на решения будет рассмотрено в разд. VIII.)

Результаты.

1. Одна особая точка, которую мы обозначаем \bar{x}_0 , всегда располагается в центре концентрационного симплекса.

2. n особых точек, $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n$, располагаются в вершинах симплекса S_n .

3. Во многих случаях существуют одно-, двух- или трехмерные многообразия особых точек или даже многообразия большей размерности, например состоящие из особых точек ребра, треугольники, тетраэдры или симплексы более высоких размерностей [54]. Эти многообразия всегда располагаются на границах соответствующих симплексов S_n . Например, ребра, образованные особыми точками, находятся на границах S_n , $n \geq 4$, треугольники — на границах S_n , $n \geq 6$, тетраэдры — на S_n , $n \geq 8$.

Анализ нормальных мод в окрестности центральной особой точки \bar{x}_0 , которая ответственна за кооперативный отбор, дает

$$\bar{x}_0 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{pmatrix} \cdot \frac{c_0}{n}; \quad \omega_j^{(0)} = - \frac{1 - \gamma^{-(p-1)}}{1 - \gamma} \left(\frac{c_0}{n} \right)^{p-1} \cdot k;$$

$$j = 1, 2, \dots, n - 1; \quad \gamma = e^{\frac{2\pi i}{n}};$$

$$\omega_n^{(0)} \equiv \omega_c^{(0)} = - \left(\frac{c_0}{n} \right)^{p-1} \cdot k. \quad (\text{T.9.3})$$

Продолжение табл. 9

Для $p = 2$ имеется n или $n - 1$ разных собственных значений, тогда как для $p = n$ все собственные значения равны $\omega_c^{(0)}$. Обычно встречается первый случай. Для $p = 2$ и четных n получаются $n - 1$ однократно вырожденных и одно дважды вырожденное собственное значение $\omega_c^{(0)} = \omega_{j=n/2}^{(0)} = -k(c_0/n)^{p-1}$, тогда как для нечетных n все собственные значения разные. Отрицательное $\omega_0^{(p)}$ снова означает, что динамическая система на симплексе S_n устойчива по отношению к флуктуациям суммарной концентрации c .

Первую систему мы назовем просто «элементарным гиперциклом», вторую — «компаунд-гиперциклом» в соответствии с его наиболее часто встречающейся физической реализацией в виде комплекса с кооперативным поведением.

в. Элементарный гиперцикл

С изменением размерности динамической системы наблюдаются интересные изменения природы особой точки в центре симплекса. Проанализируем более тщательно множества собственных значений для различных n , которые удобно представлять в виде векторов $\omega = \text{Re}\omega e_1 + i\text{Im}\omega e_2$ в комплексной гауссовой плоскости (рис. 26). Особая точка в центре при $n = 2$ является фокусом, при $n = 3$ — спиральным стоком, при $n = 4$ — центром. Для $n \geq 5$ мы получим седловые точки со спиральными компонентами в некоторых плоскостях. Эти характерные изменения природы особой точки напоминают бифуркацию Хопфа, несмотря на то что параметром в нашем случае является дискретно изменяющаяся величина — размерность n динамической системы. Как будет показано в ходе более общего анализа (разд. VIII), центральная особая точка является асимптотически устойчивой для $n = 2, 3$ и 4 . В случае более высокой размерности ($n \geq 5$) мы имеем более сложный аттрактор, а именно устойчивую замкнутую орбиту, или предельный цикл, который всегда остается внутри симплекса, никогда не достигая его границ. Для

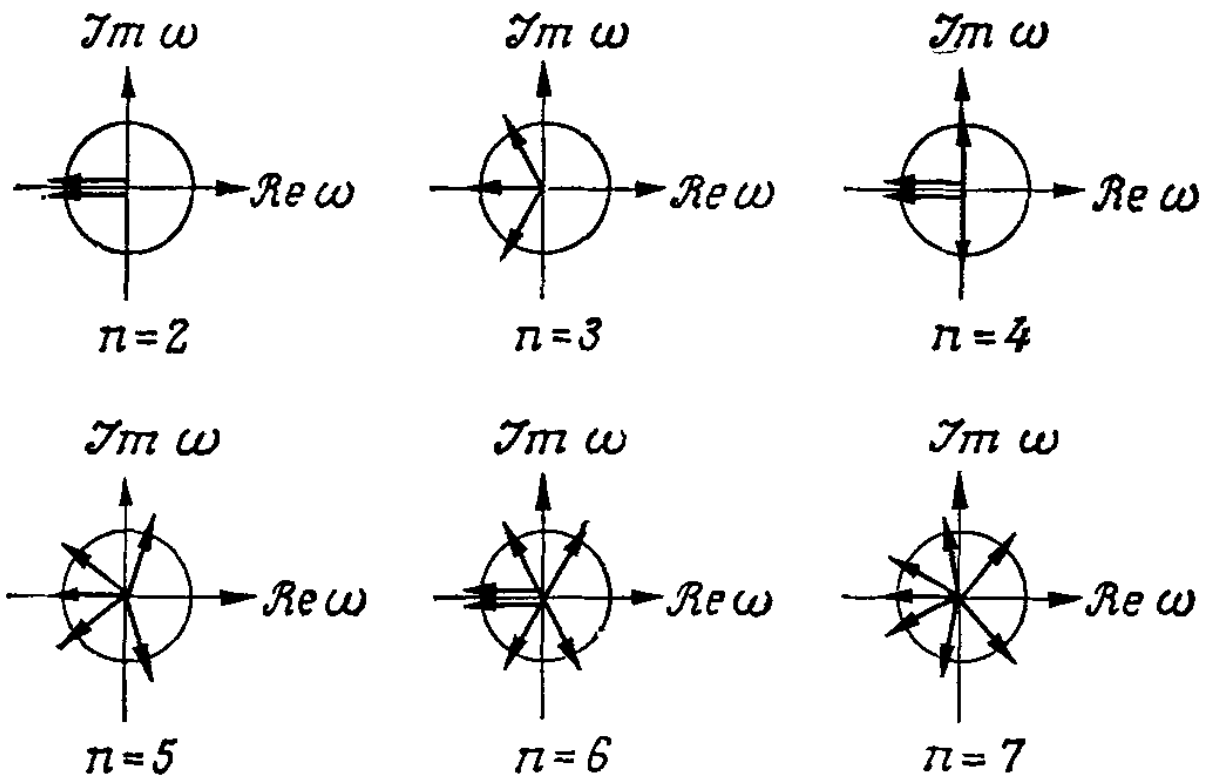


Рис. 26. Нормальные моды ω для центральной особой точки \bar{x}_0 в гиперциклах типа (65) с $p=2$ и размерностью n . $\text{Re } \omega$ и $\text{Im } \omega$ — соответственно действительная и мнимая части частоты ω .

последнего случая средние во времени концентрации x_i , равные

$$X_i(t) = \frac{1}{t} \int_0^t x_i(\tau) d\tau, \quad \bar{X}_i = \lim_{t \rightarrow \infty} X_i(t), \quad (67)$$

быстро приближаются к c_0/n (для одних и тех же k_i), т. е. точно к тому же значению, что и в случае устойчивых особых точек.

Для особых точек, расположенных в каждой из вершин $k(x_k = c_0)$ симплекса S_n , мы имеем одно положительное и $n-1$ нулевых значений ω_j^k . В разд. VIII мы проанализируем нелинейные вклады и идентифицируем эти особые точки как седловые. Следовательно, соответствующие асимптотические решения не будут давать вклада в селекционное поведение.

Конкретный пример — карта особых точек для гиперцикла с $p=2$ и $n=3$ — приведен на рис. 27.

В общем случае функции чистого роста для отдельных самореплицирующихся единиц, которые образуют динамическую

систему гиперцикла, содержат не только каталитические члены, но и члены роста первого порядка:

$$\Gamma_i = k_i x_i + k'_i x_i x_j. \quad (68)$$

Накладывая ограничение постоянной организации на динамическую систему с этими функциями роста, получим

$$\dot{x}_i = k_i x_i + k'_i x_i x_j - \frac{x_i}{c_0} \sum_k (k_k x_k + k'_k x_k x_l), \quad (69)$$

$$j = i - 1 + n\delta_{i1}, \quad l = k - 1 + n\delta_{k1}, \quad i = 1, 2, \dots, n.$$

С математической точки зрения каталитическая цепь (рис. 23) отличается от гиперцикла только одной константой скорости и получается из последнего, если положить $k'_1 = 0$. Следовательно, можно ожидать, что эти два типа динамических систем будут в чем-то сходны между собой. В соответствии с неоднородностью функций роста карты особых точек зависят от суммарной концентрации. При малых концентрациях обе системы становятся идентичны системе экспоненциально растущих независимых конкурентов (рис. 22). При высоких же концентрациях системы различаются. Динамическая система (69) асимптотически становится подобной соответствующему элементарному гиперциклу ($p = 2$).

В качестве конкретного примера снова рассмотрим систему размерности $n = 3$. Имеется семь особых точек: три из них сов-

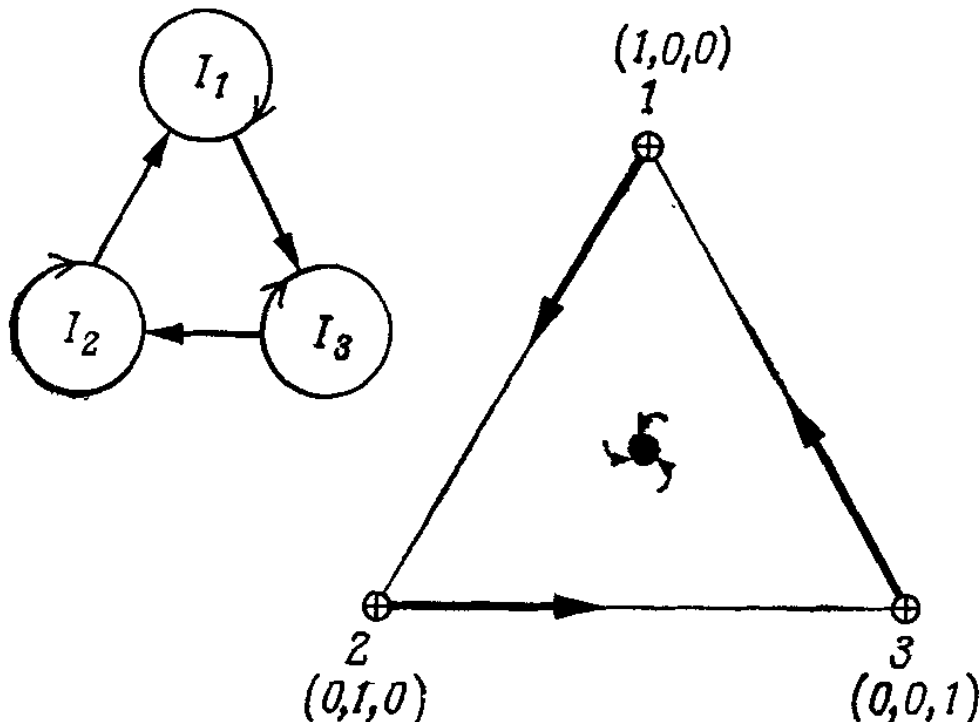


Рис. 27. Карта особых точек динамической системы (65), состоящей из самореплицирующихся единиц \hat{I}_i , которые образуют гиперцикл при ограничении постоянной организации. $\Gamma_i = k_i x_i x_j$; $j = i - 1 + n\delta_{i1}$; $n = 3$, $p = 2$, $k = (1, 1, 1)$.

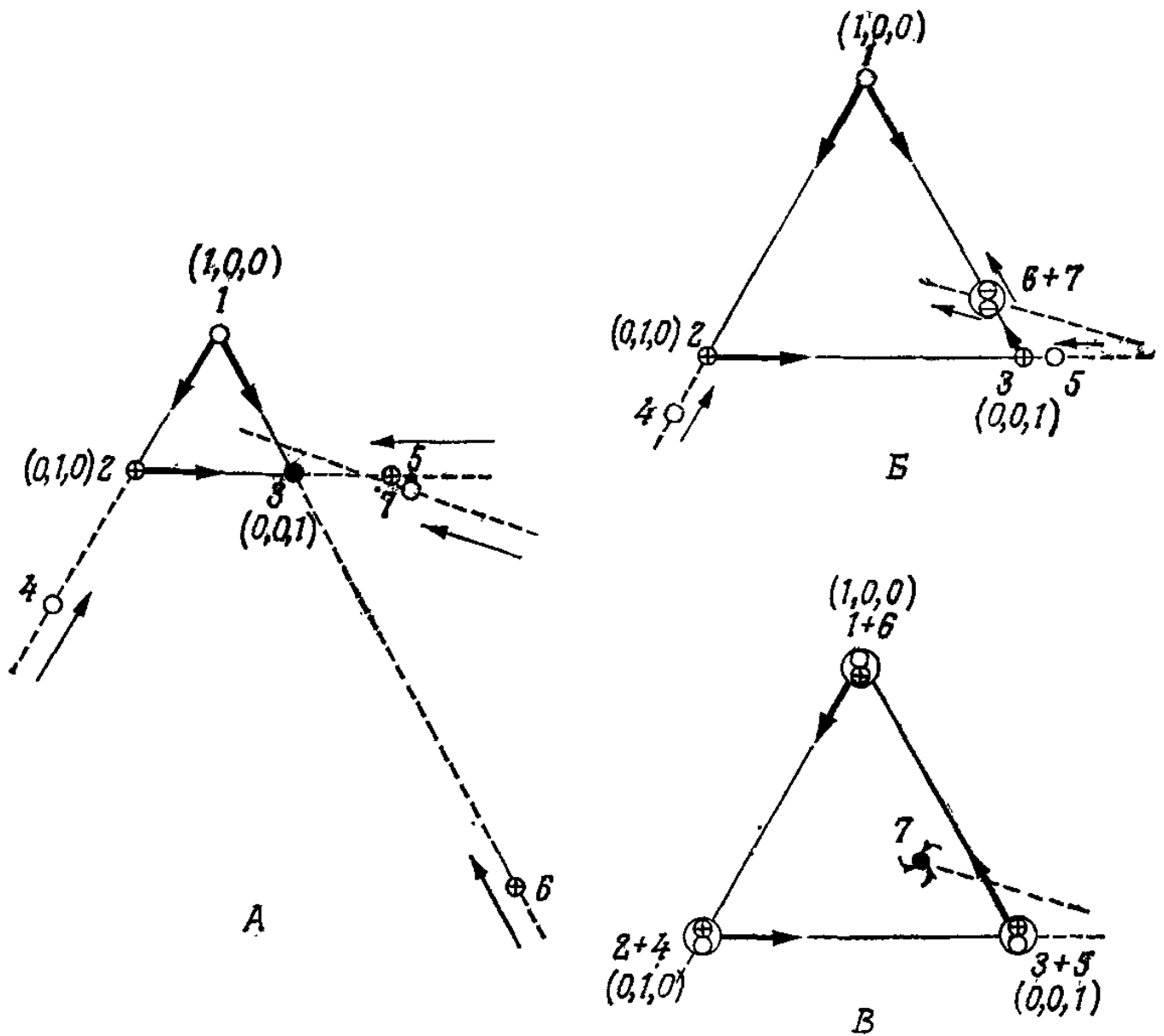


Рис. 28. Карта особых точек динамической системы (69), состоящей из самореплицирующихся единиц, которые образуют каталитический гиперцикл.

$$\Gamma_i = k_i x_i + k'_i x_i x_j, \quad j = i - 1 + n\delta_{i1};$$

$$n = 3, \quad p = 2, \quad \mathbf{k} = (1, 2, 3; 1, 2, 3).$$

$$A. c_0 = 0,5. \quad B. c_0 = 2,5. \quad B. \lim c_0 \rightarrow \infty;$$

$$1 = \bar{x}_1, \quad 2 = \bar{x}_2, \quad 3 = \bar{x}_3, \quad 4 = \bar{x}_{12}, \quad 5 = \bar{x}_{23}, \quad 6 = \bar{x}_{31}, \quad 7 = \bar{x}_0.$$

падают с вершинами симплекса S_3 , три другие лежат на ребрах, седьмая находится внутри S_3 .

Для определенного набора параметров \mathbf{k} были получены численные результаты, представленные на рис. 28. Как и для каталитической цепочки рис. 23, мы даем карты особых точек для трех разных значений суммарной концентрации c_0 : для нижнего и верхнего концентрационных пределов (A и B) и для критической точки (B).

$$c_0 = c_{кр} = k_3 (k_2'^{-1} + k_2'^{-1}) - k_1 k_1'^{-1} - k_2 k_2'^{-1}.$$

Рассмотрение развития динамических систем (58) и (69), близких к внутреннему равновесию, выявляет очень существенное различие между циклической и нециклической системами: циклическая система приводит к асимптотическому верхнему концентрационному пределу, который характеризуется постоянными относительными концентрациями отдельных видов, а открытая цепь при высокой суммарной концентрации приближается к чистому состоянию ($x_n = c_0$).

Резюмируя все развитие системы от нижнего до верхнего концентрационного предела, мы видим, что гиперцикл, который описывается динамической системой (69), представляет хороший пример самоорганизации. Начиная с конкуренции между отдельными видами растущая система приближается к конечному состоянию с динамической регуляцией чистой продукции всех членов. Этот внутренний контроль ведет к устойчивому стационарному состоянию или к состоянию с регулярными колебаниями популяционных переменных вблизи особой точки.

г. Компаунд-гиперцикл

Исследование случая $p = n$ дает простой общий результат: как и выше, внутри симплекса имеется одна особая точка. Вся граница симплекса, однако, состоит из неустойчивых особых точек, ребер из особых точек, плоскостей из особых точек и т. д. Поскольку инвариантная точка внутри симплекса (x_0) является фокусом при любых значениях n , все траектории, начинающиеся внутри симплекса, который представляет собой область, имеющую физический смысл, через достаточно большое время сойдутся к этой точке. Все собственные значения $\omega_j^{(0)}$, связанные с x_0 , одинаковы для данных k , c_0 и n . Их можно найти из формулы (3) в табл. 9, если положить $p = n$. Карта особых точек для компаунд-гиперцикла с $n = 3$ показана на рис. 29. Эти комплексы, таким образом, представляют прекрасные примеры регуляции относительных концентраций своих компонентов.

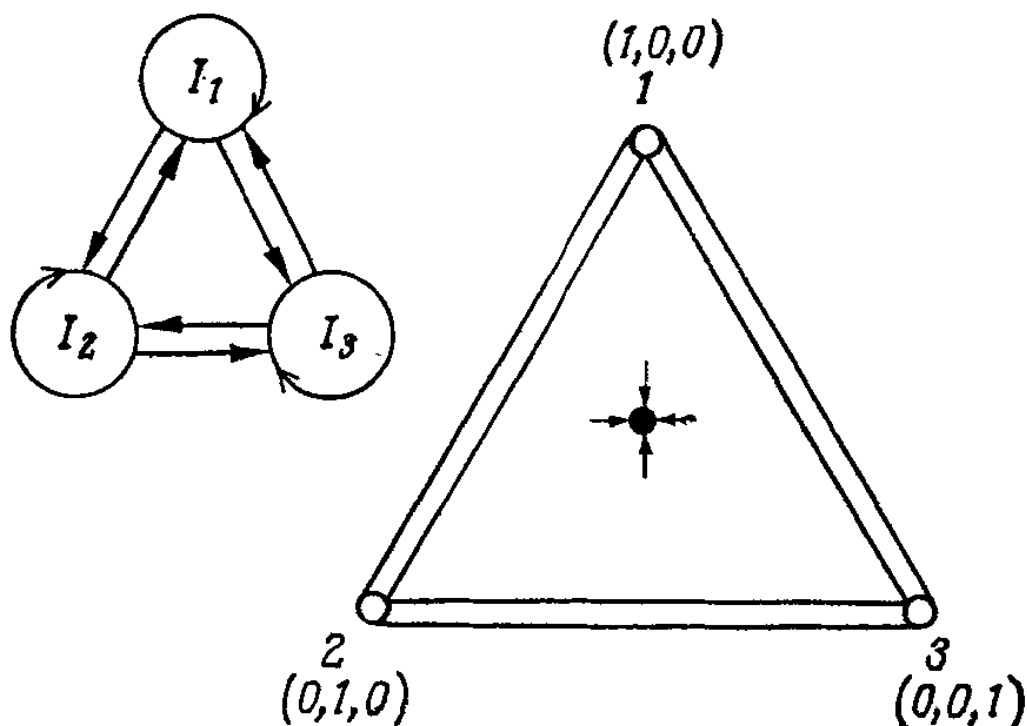


Рис. 29. Карта особых точек динамической системы (65), состоящей из самореплицирующихся единиц $\textcircled{I_i}$, которые образуют компаунд-гиперцикл при ограничении постоянной организации. $\Gamma_l = k_i x_1 x_2 \dots x_n$; $n = 3$, $p = 3$, $k = (1, 1, 1)$.

д. Сравнение различных гиперциклов

Характерным свойством гиперциклов является их способность интегрировать информацию. Действительно, простейшие члены этого класса являются наименее сложными динамическими структурами, способными предотвращать потерю информации из ансамбля функционально связанных самореплицирующихся единиц из-за элиминации некоторых его членов в результате селекционной конкуренции. С динамической точки зрения все виды гиперциклов эквивалентны по отношению к этому свойству. С другой стороны, менее сложные системы типа простых каталитических циклов (рис. 4) не могут интегрировать информацию, так как они не обладают *внутренней* способностью к самовоспроизведению (см. [4], с. 501 и далее).

Дальнейшее подразделение в иерархии гиперциклов может быть сделано в соответствии с их реализуемостью в природе, что будет рассмотрено в части В. Здесь для примера мы сравним элементар-

ный ($p=2$) и компаунд- ($p=n$) гиперциклы с точки зрения их физического воплощения. Элементарные гиперциклы в соответствии со своим законом роста требуют бимолекулярных столкновений макромолекул. К таким бимолекулярным процессам легко приводят различные механизмы; они следуют также из реалистических допущений о механизме репликации нуклеиновых кислот или о синтезе белка, инструктируемом мРНК (см. также разд. IX и часть В). Компаунд-гиперцикл требует, чтобы каждый партнер вносил вклад в скорость образования каждого компонента. Для реализации такого компаунд-гиперцикла необходимо либо мультимолекулярное столкновение, что крайне мало вероятно, либо образование промежуточного комплекса из n различных субъединиц, что очень невыгодно при низких концентрациях. Добиологические условия между тем характеризуются именно крайне низкими концентрациями *индивидуальных* макромолекул. Для эффективного начала эволюции через компаунд-гиперциклы потребовалось бы принять крайне высокие константы ассоциации, значительно превышающие величины, полученные экспериментально, а также ввести внутреннюю связь между этими константами и функциональную эффективность отдельных компонентов. Таким образом, компаунд-гиперцикл, вероятно, имеет меньше шансов служить предпосылкой для создания системы трансляции, нежели любой гиперцикл с меньшей степенью p . Однако на более поздних этапах доклеточной эволюции вероятность возникновения компаунд-гиперцикла могла бы стать выше.

Различные системы, состоящие из каталитически активных самовоспроизводящихся единиц, были изучены методом анализа особых точек. Результаты ясно указывают на необходимость гиперциклической связи. Только каталитические гиперциклы могут удовлетворять критериям интеграции информации, которые были перечислены в разд. IV. 5:

1. Селективная устойчивость каждого компонента из-за успешной конкуренции с ошибочными копиями.

2. Кооперативное поведение компонентов, объединенных в новую функциональную единицу.

3. Успешная конкуренция этой функциональной единицы с другими, менее эффективными системами.

VIII. Динамика элементарного гиперцикла

Поскольку гиперциклы являются адекватными системами предбиологической самоорганизации, имеет смысл провести более подробный анализ их динамического поведения. Для класса элементарных гиперциклов ($p=2$) может быть дано полное качественное описание вплоть до размерности $n=4$. Для больших размерностей, а также для гиперциклов с более сложной структурой топологический анализ можно облегчить, применяя численное интегрирование. Мы проиллюстрируем эти методы на примере элементарных гиперциклов, которые представляют все основные свойства гиперциклической самоорганизации¹.

VIII.1 Качественный анализ

Поскольку мы имеем дело с динамическими системами кооперирующихся компонентов, наибольший интерес представляют устойчивые аттракторы внутри области концентраций, имеющей физический смысл. Более конкретно, мы должны исследовать устойчивость таких особых точек, для которых отдельные собственные значения матрицы Якоби имеют нулевые действительные части. В разд. VII (табл. 9) мы встречались в основном с двумя случаями:

1. Нулевые собственные значения ($\omega_j^{(i)} = 0$, $j = 2, 3, \dots, n$ и $i = 1, 2, \dots, n$) для особых точек \bar{x}_i в вершинах симплексов S_n .

2. Чисто мнимые собственные значения ($\omega_{2,4}^{(0)} = \pm i$) для центральной особой точки четырехчленного гиперцикла на S_4^2 .

¹ Для частного случая гиперцикла, у которого член второго порядка $x_k x_{(k-1)}$ функции роста заменен членом $x_k \ln x_{k-1}$, можно получить аналитическое решение [21].

² Чисто мнимые собственные значения будут также у элементарных гиперциклов размерности $4k$, где k — целое число ≥ 2 . В этих многомерных примерах, однако, центральная особая точка является седлом независимо от природы вкладов более высокого порядка в чисто мнимые собственные значения.

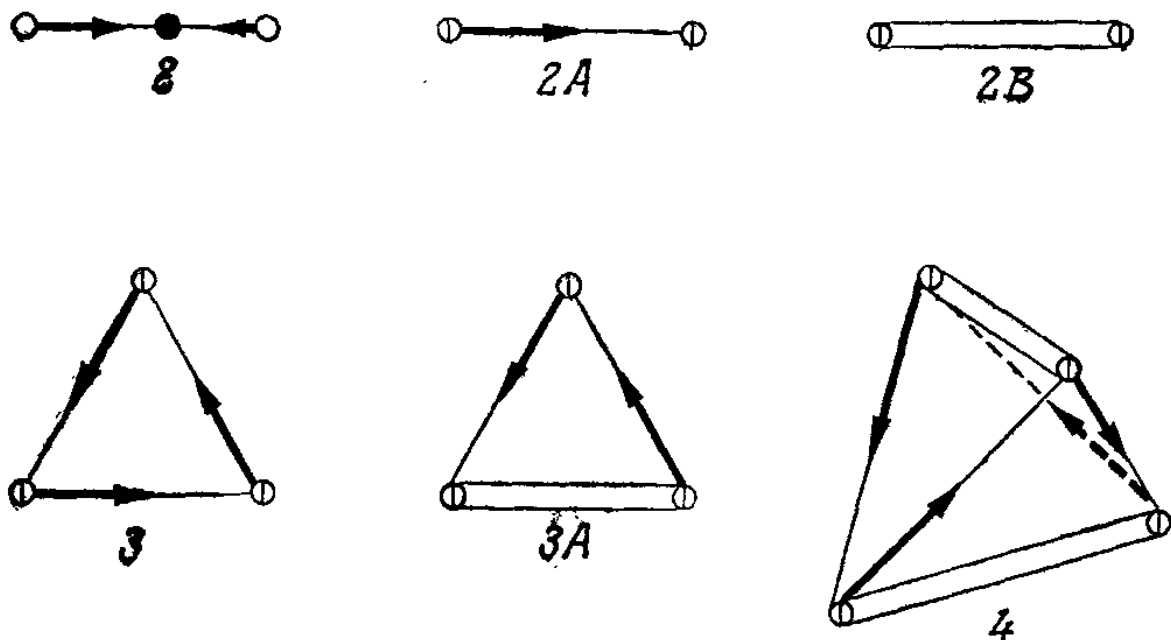


Рис. 30. Динамические системы, соответствующие элементарным гиперциклам размерности $n = 2, 3$ и 4 . Отдельные системы отображаются на симплексы S_n и могут быть разложены на глобально инвариантные динамические подсистемы (табл. 10). «Полные» подсистемы 2, 3 и 4 характеризуются ненулевыми значениями всех популяционных переменных и поэтому описывают развитие, происходящее внутри симплексов $S_n (IS_n: 0 < x_i < c_0, i = 1, 2, \dots, n)$. На границах симплексом $S_n (BS_n)$ одна или большее число популяционных переменных обращаются в нуль и получают динамические подсистемы меньшей размерности — такие, как «текущее ребро» 2А, «ребро особых точек» 2В и треугольник типа 3А (отметим, что динамическая система 2А локализуется на границах 3, 3А и 4, система 2В — на границах 3А и 4, а система 3А — на границах 4).

Перед тем как проводить общее доказательство устойчивости центральной особой точки в гиперциклах малой размерности $n \leq 4$, исследуем более детально топологию этих систем.

Динамические системы, соответствующие элементарным гиперциклам, могут быть представлены в виде совокупности нескольких подсистем, каждая из которых определяется на глобально инвариантном подпространстве. Множество точек или подпространство будет называться «глобально инвариантным» по отношению к данной динамической системе в том и только в том случае, если траектория, которая проходит через произвольную точку подпространства, никогда не покидает этого подпространства.

В частности, динамические системы на симплексах S_n можно подразделить на два класса: системы на границе (BS_n) и системы внутри симплекса (IS_n). Внутренняя часть симплекса (определение ее было дано выше) — это область, где ни одна популяционная переменная не обращается в нуль: $0 < x_i < 1, i = 1, 2, \dots, n$. Ясно, что динамические системы на IS_n наиболее интересны потому, что они описывают развитие интактных ги-

перциклов. В дальнейшем мы будем обозначать их номерами 2, 3, 4, ..., N.

На границе одна, две или большее число популяционных переменных обращаются в нуль. Следовательно, динамические системы на BS_n можно подразделить на динамические системы на симплексах меньшей размерности — на ребрах, гранях и гипергранях. Чтобы отличить эти системы от полных гиперциклов, мы будем использовать сокращенные обозначения **2A**, **2B**, **3A** и т. д. Все динамические системы, соответствующие элементарным гиперциклам размерности $n \leq 4$, показаны схематически на рис. 30. В качестве конкретного примера в табл. 10 представлено разложение четырехмерной системы на 11 подсистем.

Все динамические системы вплоть до размерности $n = 4$ можно анализировать методом Ляпунова (табл. 11). Для трех систем — 2, 3 и 4 — построены функции Ляпунова, и, следовательно, центральная особая точка представляет собой устойчивый аттрактор. Более того, бассейн этой особой точки распространяется на всю внутреннюю часть симплекса, что означает независимо от начального распределения популяционных переменных мы приходим к одному и тому же устойчивому набору

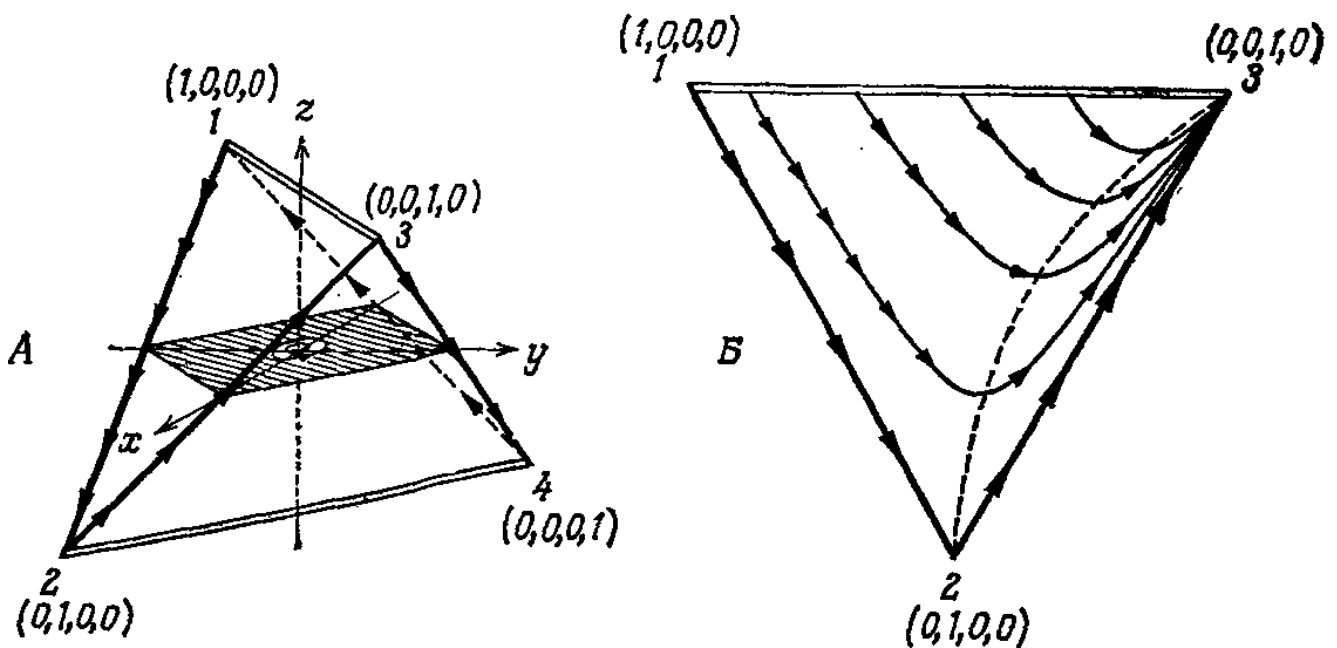


Рис. 31. Динамическая топология элементарного гиперцикла размерности $n = 4$. Динамическая система на симплексе состоит из системы 4 на внутренней части IS_4 и четырех эквивалентных систем типа 3A на равносторонних треугольниках (S_3), каждый из которых ограничен двумя текущими ребрами 2A и одним ребром особых точек 2B. А. Систему на внутренней части удобно описывать переменными x , y и z (см. табл. 11). В плоскости (x, y) (заштрихованная область на рисунке) лежит многообразие замкнутых concentрических траекторий, принадлежащих центру линеаризованной системы. Б. Динамическая система 3A. Каждая траектория начинается из какой-либо точки на ребре особых точек 13 и кончается в вершине 3. Штриховая линия соединяет все точки, в которых траектории параллельны ребру особых точек 13.

Таблица 10

Глобально инвариантные динамические подсистемы
элементарного гиперцикла размерности $n = 4$

| Символ | Условия | Динамическая система |
|--------|---|---|
| 4 | $\xi_1, \xi_2, \xi_3, \xi_4 > 0$ | $\dot{\xi}_i = \xi_i \xi_j - \xi_i \phi, \quad i = 1, 2, 3, 4,$ $j = i - 1 + n\delta_{i1} \text{ и}$ $\phi = \xi_1 \xi_2 + \xi_2 \xi_3 + \xi_3 \xi_4 + \xi_4 \xi_1$ |
| 3A | $\xi_4 = 0$ | $\dot{\xi}_1 = -\xi_1 \phi,$ $\dot{\xi}_i = \xi_i \xi_j - \xi_i \phi, \quad i = 2, 3,$ $j = i - 1 \text{ и}$ $\phi = \xi_1 \xi_2 + \xi_2 \xi_3$ |
| | $\xi_3 = 0, \xi_2 = 0$ или $\xi_1 = 0$ | Аналогично |
| 2A | $\xi_3 = \xi_4 = 0$ | $\dot{\xi}_1 = -\xi_1 \phi, \quad \phi = \xi_1 \xi_2,$ $\dot{\xi}_2 = \xi_1 \xi_2 - \xi_2 \phi$ |
| | $\xi_2 = \xi_3 = 0, \xi_1 = \xi_2 = 0$ или $\xi_4 = \xi_1 = 0$ | Аналогично |
| 2B | $\xi_2 = \xi_4 = 0,$ $\xi_1 = \xi_3 = 0$ | $\dot{\xi}_i = 0, \quad i = 1, 2, 3, 4$ |

Таблица 11

Функции Ляпунова [57] для основных гиперциклов
размерности $n = 2, 3$ и 4

Чтобы доказать устойчивость определенной особой точки \bar{x} динамической системы $\dot{x} = \Lambda(x)$, мы должны найти произвольную функцию $V(x)$, которая удовлетворяет следующим двум критериям:

$$(1) \quad V(\bar{x}) = 0 \quad \text{и} \quad V(x) > 0, \quad x \in U, \quad (\text{T.11.1})$$

т. е. функция обращается в нуль в особой точке и положительна в ее окрестности U . Таким образом, $V(x)$ достигает локального минимума в особой точке.

$$(2) \quad \dot{V}(x) \equiv \frac{dV}{dt} = \sum_{j=1}^n \left(\frac{\partial V}{\partial x_j} \right) \frac{dx_j}{dt} < 0, \quad x \in U, \quad (\text{T.11.2})$$

Продолжение табл. II

т. е. производная $V(x)$ по времени отрицательна в окрестности особой точки. По тривиальным причинам \dot{V} обращается в нуль в точке \bar{x} : $\dot{V}(\bar{x}) = 0$. Если для данной особой точки динамической системы такую функцию $V(x)$ удастся найти, то она называется (строгой) функцией Ляпунова, а точка \bar{x} является (асимптотически) устойчивой: траектория, проходящая через любую точку в окрестности \bar{x} , заканчивается в особой точке \bar{x} .

Функцию Ляпунова можно определить также при более слабом ограничении:

$$\dot{V}(x) \leq 0. \quad (\text{T.11.3})$$

Любая траектория, входящая в окрестность U точки \bar{x} , останется в этой окрестности. Приведем конкретный пример: стоки асимптотически устойчивы по Ляпунову в строгом смысле, тогда как центры устойчивы только в смысле слабого критерия (Т.11.3).

Для удобства мы будем использовать нормированные переменные ξ_i ; допустим также, что константы скоростей равны единице ($k_1 = k_2 = \dots = k_n = 1$). Теперь применим метод Ляпунова к основным гиперциклам.

Функция

$$V = \left(\frac{1}{n}\right)^n - \xi_1 \xi_2 \dots \xi_n \quad (\text{T.11.4})$$

имеет минимум и обращается в нуль в особой точке $\xi_i = \frac{1}{n}$; так, выполняется условие (Т.11.1). Производная V по времени может быть получена простым дифференцированием:

$$\begin{aligned} \dot{V} &= -\xi_1 \xi_2 \dots \xi_n (1 - nr); \\ r &= \sum_{k=1}^n \xi_k \xi_l, \quad l = k - 1 + n\delta_{k1}. \end{aligned} \quad (\text{T.11.5})$$

Теперь мы должны проверить критерий (Т.11.2) для систем с различными значениями n . Внутри симплекса S_n условие $\dot{V} < 0$ становится эквивалентным неравенству

$$r(\xi) < \frac{1}{n}. \quad (\text{T.11.6})$$

Найдем $r(\bar{\xi}_0) = \frac{1}{n}$, которое удовлетворяет уравнению $\dot{V}(\bar{\xi}_0) = 0$ ($\bar{\xi}_0$ — это центральная особая точка гиперцикла).

Для двумерной системы ($n = 2$) выполнение условия (Т.11.6) легко проверить:

$$\xi_1 = \xi; \quad \xi_2 = 1 - \xi \rightarrow r = 2\xi(1 - \xi) \leq \frac{1}{2}. \quad (\text{T.11.7})$$

Продолжение табл. II

Функция r является параболой с максимумом при $\xi = 1/2$. Итак, неравенство $r(\xi) < 1/2$ выполняется всюду, кроме особой точки $\xi = 1/2$, где $r = 1/2$. В этом случае V — это строгая функция Ляпунова и $\bar{\xi}_0$ асимптотически устойчива.

Для $n = 3$ ситуация аналогична. Неравенство (Т.11.6), $r < 1/3$, выполняется в каждой точке внутри симплекса S_3 , кроме особой точки $\bar{\xi}_0$, где $r = 1/3$. V снова является строгой функцией Ляпунова, и центральная особая точка $\bar{\xi}_0$ асимптотически устойчива.

Для четырехмерного случая проблема становится более сложной. Условие (Т.11.3) выполняется почти всюду на симплексе S_4 :

$$r = (\xi_1 + \xi_3)(\xi_2 + \xi_4) = s(1 - s), \quad 0 \leq s \leq 1. \quad (\text{Т.11.8})$$

Внутри симплекса мы имеем $0 \leq r \leq 1/4$, причем $r = 1/4$ тогда и только тогда, когда $s = 1/2$. Уравнение $s = 1/2$ определяет плоскость $\xi_1 + \xi_2 = 1/2$ (см. рис. 31, А и 34, Б). Очевидно, что V является лишь нестрогой функцией Ляпунова. Этот результат позволяет предположить, что центральная точка является по меньшей мере устойчивой. Для доказательства асимптотической устойчивости введем новые переменные x, y, z :

$$\begin{aligned} x &= -2(\xi_2 + \xi_3) + 1, & \dot{x} &= -(1 + z)(y - xz), \\ y &= 2(\xi_1 + \xi_2) - 1, & \dot{y} &= (1 - z)(x - yz), \\ z &= 2(\xi_1 + \xi_3) - 1, & \dot{z} &= z^3 - z + x^2 - y^2, \end{aligned}$$

которые сдвигают начало координат в центр симплекса S_4 , причем оси координат проходят теперь через середины ребер 23, 34 и 13 соответственно (см. рис. 31, А). Четвертая переменная, $\xi_4 = 1 - \xi_1 - \xi_2 - \xi_3$, исключается. Итак, ось z направлена перпендикулярно критической плоскости $\xi_1 + \xi_3 = 1/2$, через которую проходят оси x и y . В этой плоскости динамическая система упрощается до следующего вида: $\dot{x} = -y$, $\dot{y} = x$, $\dot{z} = x^2 - y^2$.

Производная z по времени обращается в нуль только вдоль двух прямых, $x = \pm y$, или $\xi_2 = \xi_4$ и $\xi_1 = \xi_3$ соответственно. Следовательно, в этой критической плоскости нет траекторий — за исключением особой точки $\bar{\xi}_0$ — и система проходит через нее за бесконечно малое время. Условие $\dot{V}(\xi(t)) < 0$ выполняется вдоль любой данной траектории почти в каждый момент времени — исключениями являются лишь те моменты, когда система проходит критическую плоскость $\xi_1 + \xi_3 = 1/2$. Вдоль всех траекторий $V(\xi(t))$ монотонно убывает с ростом t . V является строгой функцией Ляпунова, так что особая точка $\bar{\xi}_0$ асимптотически устойчива.

При более высоких размерностях, $n \geq 5$, $V(\xi)$ не является функцией Ляпунова, следовательно, этот метод не дает возможности сделать какие-либо предсказания об устойчивости центральной особой точки.

стационарных концентраций. Для данных динамических систем действительно характерно кооперативное поведение компонентов. Этот результат особенно важен для четырехмерной системы, где линейная аппроксимация, использованная при исследовании особых точек, показала наличие центра, окруженного многообразием концентрических замкнутых орбит в плоскости (x, y) (см. рис. 31, А), что не позволяет сделать определенных выводов об устойчивости.

Динамические системы на гранях симплексов (BS_n) определяют поведение неполных гиперциклов, т.е. каталитических гиперциклов, у которых отсутствует по крайней мере один из членов. В действительности эти системы описывают кинетику «вымирания» гиперцикла. Они имеют также некоторое значение на определенных этапах образования гиперцикла. На границах полных динамических систем вплоть до размерности 4 мы имеем два вида ребер — $2A$ и $2B$, а также грань $3A$ (рис. 30). Все три динамические системы можно исследовать прямым способом.

Ребро $2A$ соединяет два последовательных чистых состояния, или две вершины, которые мы обозначаем через i и j ($j = i + 1 - n\delta_{in}$). Как показано на рис. 32, вдоль ребра действует движущая сила в направлении $i \rightarrow j$. Итак, единственная траектория этой системы ведет от вершины i к вершине j . Соответственно мы будем называть систему $2A$ «текущим ребром». По мере приближения к вершине j движущая сила убывает по параболическому закону (рис. 32). Следовательно, линейный член в ряде Тейлора обращается в нуль в особой точке \bar{x}_j , и исследование особых точек не может дать нужных сведений о природе этой особой точки.

В элементарных гиперциклах вершины симплексов являются седловыми точками: вершина i устойчива по отношению к флуктуациям, направленным вдоль ребра $\bar{hi} (\delta x_h > 0, h = i - 1 + n\delta_{in})$, но неустойчива вдоль ребра $\bar{ij} (j = i + 1 - n\delta_{in})$. Итак, на границе любой полной динамической системы мы имеем замкнутую петлю — $\bar{12}, \bar{23}, \bar{34}, \dots, \bar{n1}$, вдоль которой система вращается в определенном направлении. Этот цикл не является единственной траекторией. Чтобы система могла перейти в следующее чистое состояние, в каждой вершине должны происходить флуктуации определенного вида. Существование этой петли отражает циклическую симметрию всей системы, а асимметрия в каждой отдельной вершине — принятую в нашей модели необратимость синтеза и деградации биополимеров.

В динамической системе $3A$ область значений переменных, имеющая физический смысл, ограничивается двумя последовательными текущими ребрами \bar{ij} и \bar{jk} ($j = i + 1 - n\delta_{in}$ и $k = j + 1 - n\delta_{in}$) и одним ребром особых точек \bar{ik} ($k \neq i - 1 + n\delta_{in}$). Траектории этой системы показаны на рис. 31, Б. Они начинаются из какой-либо точки на ребре особых точек и заканчиваются в вершине k , которая, таким образом, является единственным устойчивым аттрактором системы. Вид I_k , следова-

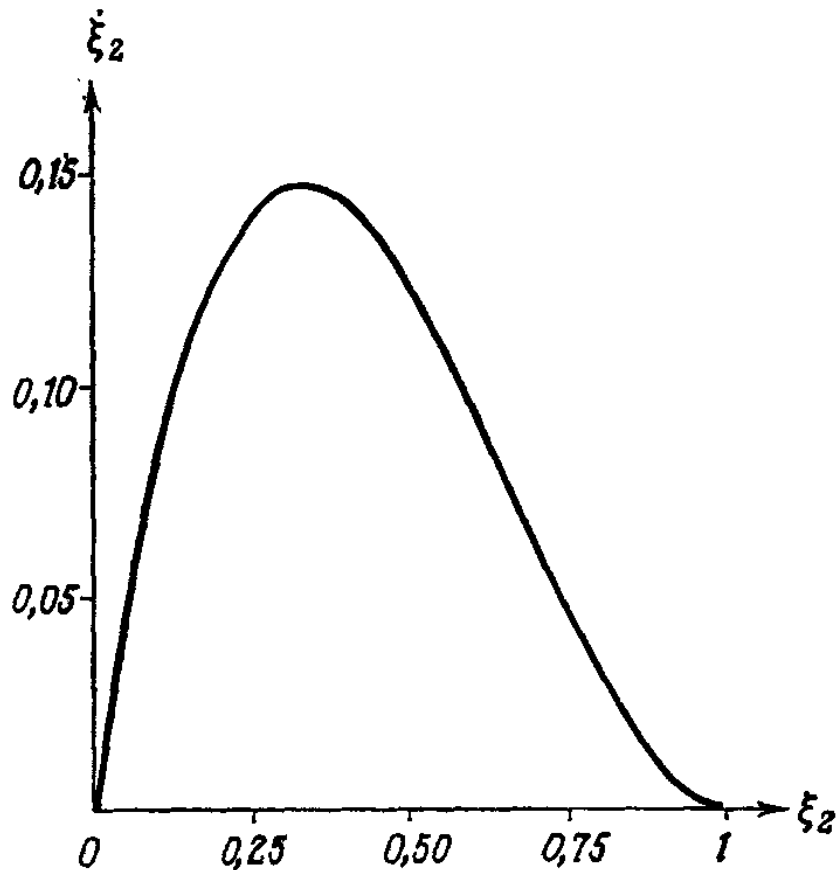


Рис. 32. «Текущее ребро» 2А. Касательный вектор $\dot{\xi}_2 = \xi_2(1 - \xi_2)^2$ положителен внутри всей области значений, имеющих физический смысл ($0 < \xi_2 < 1$), и обращается в нуль на обоих концах, которые являются двумя особыми точками системы: $\bar{\xi}_1 = (1, 0)$ и $\bar{\xi}_2 = (0, 1)$. Неравенство $\dot{\xi}_2 > 0$ означает, что ξ_2 возрастает в течение следующего бесконечно малого интервала времени. Следовательно, на 2А имеется только одна траектория, ведущая из $\bar{\xi}_1$ к $\bar{\xi}_2$, т. е. из вершины 1 ($\xi_2 = 0$) к вершине 2 ($\xi_2 = 1$). Динамическая система «течет» вдоль этого ребра. Отметим, что $d\dot{\xi}_2/d\xi_2$ обращается в нуль в особой точке $\bar{\xi}_2$ ($\xi_2 = 1$), вследствие чего собственное значение линеаризованной системы $\omega = 0$. Следовательно, исследование особых точек не позволяет установить характер устойчивости этой точки.

тельно, выживает и является остатком этого фрагмента гиперцикла.

Исследование границ основных гиперциклов можно обобщить на случай систем с большей размерностью. Полученные результаты дают возможность предсказать характер асимптотического развития неполных гиперциклов. После того как один из видов гиперцикла элиминируется каким-либо внешним событием, оставшаяся динамическая система становится неустойчивой и через достаточно большое время приходит в чистое состояние. Во всех случаях отбирается тот вид, который в гиперцикле ло-

кализован как раз перед разрывом. Другими словами, если i предшествует j ($j = i + 1 - \delta_{in}$), то вид I_i сохранится как последний компонент гиперцикла, разрушенного вымиранием компонента I_j . Это поведение вытекает из свойств каталитических цепей.

VIII.2. Численное интегрирование

Системы дифференциальных уравнений для элементарных гиперциклов с размерностями до $n \leq 12$ интегрировались с помощью стандартных численных методов. Соответствующие интегральные кривые $x(t)$ рассматривались в предыдущей работе [4], и их можно не приводить снова, так как здесь нас интересует другой аспект проблемы. Теперь наша цель — поиск устойчивых аттракторов внутри симплексов S_n , которые гарантируют кооперативное поведение компонентов. Соответствующее исследование многообразия траекторий достаточно просто.

Дифференциальные уравнения для траекторий получаются исключением явной временной зависимости из исходной

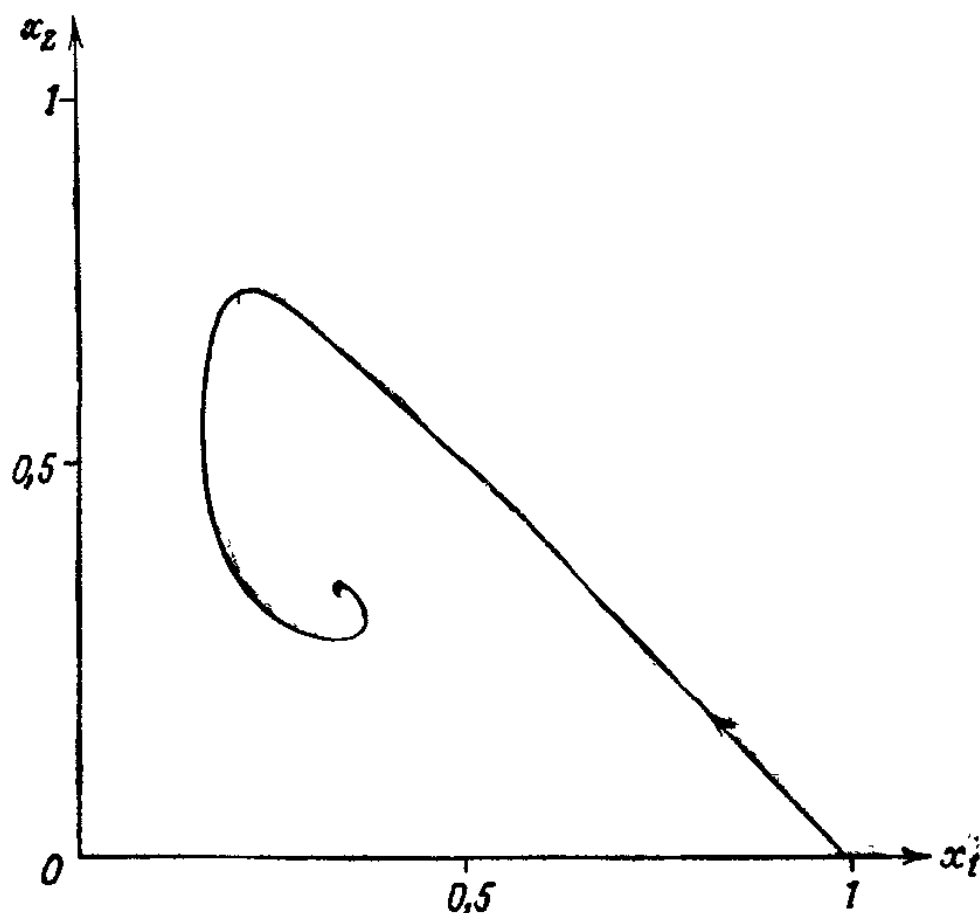


Рис. 33. Траектория динамической системы 3 для элементарного гиперцикла размерности $n = 3$ показана в виде проекции на плоскость (x_1, x_2) . Начальные условия: $x_1(0) = 0,98$, $x_2(0) = 0,01$.

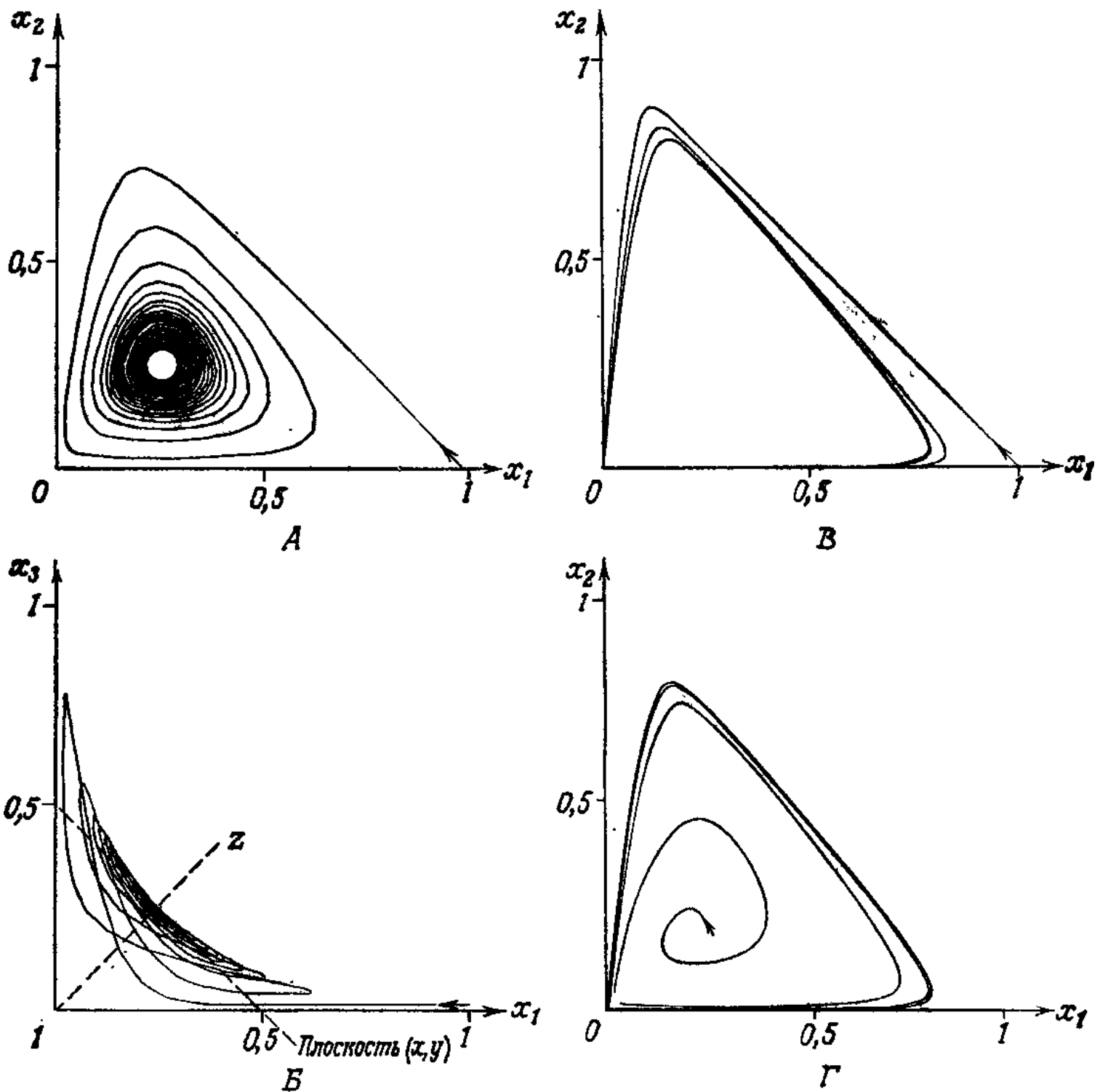
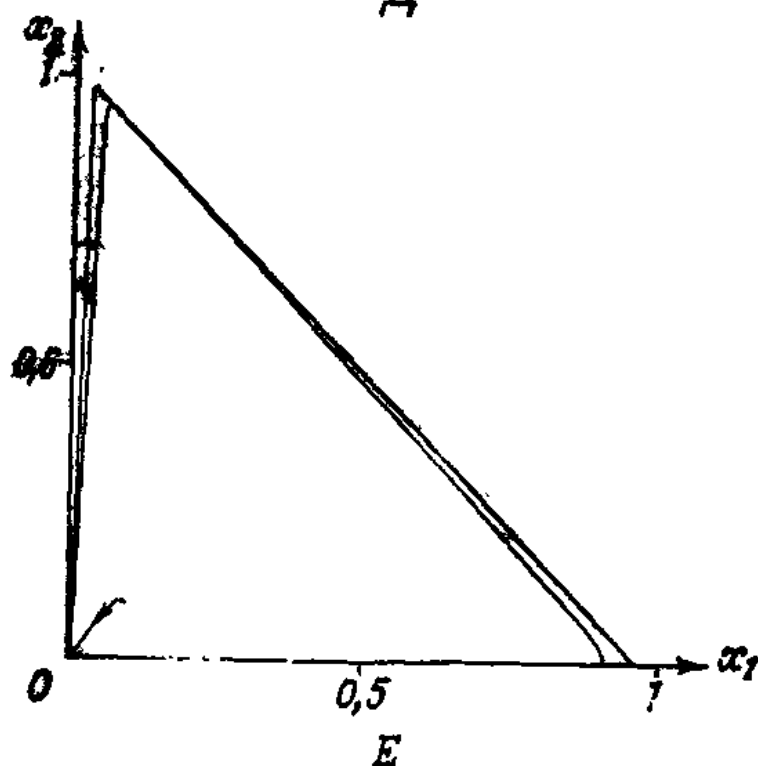
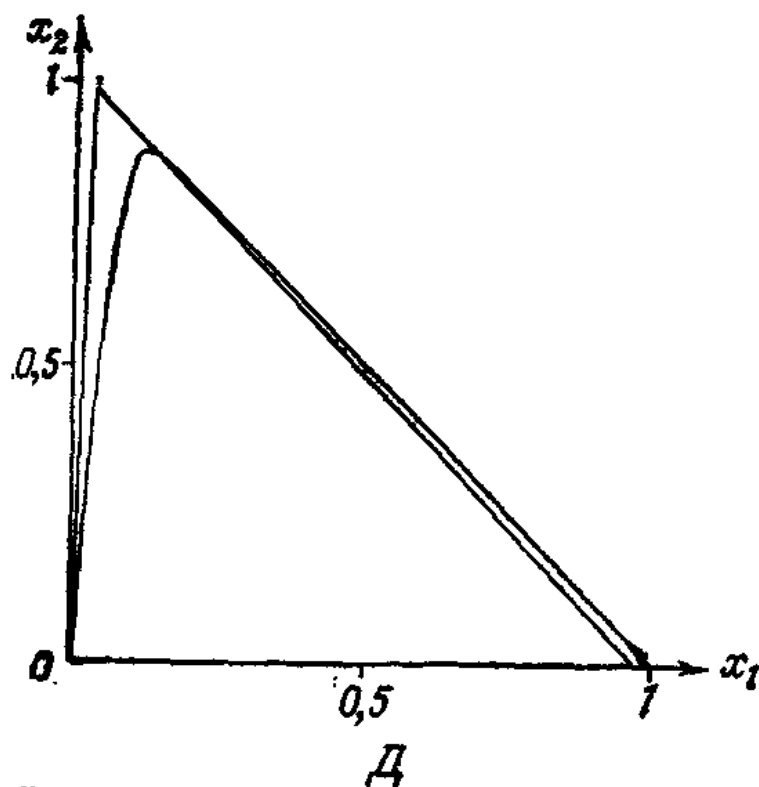


Рис. 34. Траектории динамических систем 4, 5 и 12 для элементарно. (1) $n = 4$: начальные условия — $x_1(0) = 0,97$, $x_2(0) = 0,03$. Траектория ведет по контуру и затухает. Б. Проекция траектории на плоскость (x_1, x_3) . [рис. 31] пересекает плоскость (x_1, x_3) по прямой $x_1 + x_3 = 1/2$. Траектория пересекает плоскость (x, y) только в отдельных точках (табл. 11). (2) $n = 5$: проекции на плоскость (x_1, x_2) . В. Начальные условия: $x_1(0) = 0,9999$, $x_2(0) = 0,0001$. Г. Начальные условия: $x_1(0) = 0,2004$, $x_2(0) = 0,7996$. Элементарная система приближается к одному и тому же предельному состоянию. (3) $n = 12$: проекции на плоскость (x_1, x_2) . Д. Начальные условия: $x_1(0) = 0,0848$, $x_2(0) = \dots = x_{12}(0) = 0,0152$. Траектория приближается к контуру 1—2, 2—3, ..., 11—12, 12—1.



ментарных гиперциклов размерности $n = 4, 5$ и 12 соответственно $x_3(0) = x_4(0) = 0,01$; показаны две проекции. А. Проекция спирали в центральную особую точку, причем колебания сильны. Плоскость центрального многообразия [плоскость (x, y)] на нее перпендикулярно ей проходит ось z . Отметим, что траектория не остается в ней на более продолжительное время. Начальные условия: $x_1(0) = 0,9996, x_2(0) = x_3(0) = x_4(0) = 0,1999$. Отметим, что динамическому циклу независимо от начальных условий (В или Г). Начальные условия: $x_1(0) = 0,9989, x_2(0) = \dots = x_{12}(0) = 0,0001$. Е. Начальное значение $x_1(0) = 0,0832$. Снова оба предельных цикла идентичны и проходят

динамической системы:

$$\begin{aligned} \frac{dx_2}{dx_1} &= \frac{\Lambda_2}{\Lambda_1} = f_2(x_1, x_2, \dots, x_n), \\ \frac{dx_3}{dx_1} &= \frac{\Lambda_3}{\Lambda_1} = f_3(x_1, x_2, \dots, x_n), \\ &\vdots \\ \frac{dx_n}{dx_1} &= \frac{\Lambda_n}{\Lambda_1} = f_n(x_1, x_2, \dots, x_n). \end{aligned} \quad (70)$$

Интегрирование этой новой $(n-1)$ -мерной динамической системы дает траектории (интегральные кривые):

$$\begin{aligned} x_2 &= g_2(x_1, x_2, \dots, x_n), \\ x_3 &= g_3(x_1, x_2, \dots, x_n), \\ &\vdots \\ x_n &= g_n(x_1, x_2, \dots, x_n). \end{aligned} \quad (71)$$

Итак, траектория — это кривая в n -мерном пространстве концентраций. Для графического представления мы будем использовать проекции этих кривых на плоскости (x_k, x_1) . Траектории для гиперциклов малой размерности ($n = 2, 3$ и 4) отражают уже известные свойства этих динамических систем. Случай $n = 2$ довольно тривиален: имеются только две орбиты, которые сходятся к центральному устойчивому фокусу (рис. 30). Траектории трехмерного гиперцикла ($n = 3$) являются спиралями, которые быстро сходятся к центральной особой точке (рис. 33). Этот вид траекторий соответствует сильно затухающим колебаниям интегральных кривых $x(t)$. Четырехчленный гиперцикл следует рассмотреть подробнее. Траектории снова идут по спирали в центр симплекса (рис. 34, А, Б). В отличие от трехмерного случая, направленная к центру сила намного слабее вращательного компонента. Соответственно сходимости к центральной особой точке крайне слабая. Проекция этой траектории на плоскость (x_1, x_3) прекрасно иллюстрирует полученный ранее результат: отсутствие траекторий в плоскости $x_1 + x_3 = 1/2$, $x_2 + x_4 = 1/2$. Действительно, как можно видеть на рис. 34, Б, траектории проходят вблизи седловидной изогнутой поверхности.

Для основных гиперциклов размерности $n \geq 5$ центральная особая точка представляет собой неустойчивое седло. На границе нет стока, и, следовательно, нужно ожидать наличия устойчивой замкнутой траектории. Однако соответствующие методы исследования пока не разработаны в достаточной степени и не позволяют доказать существование подобного аттрактора внутри симплекса. Поэтому мы должны полагаться на численные результаты.

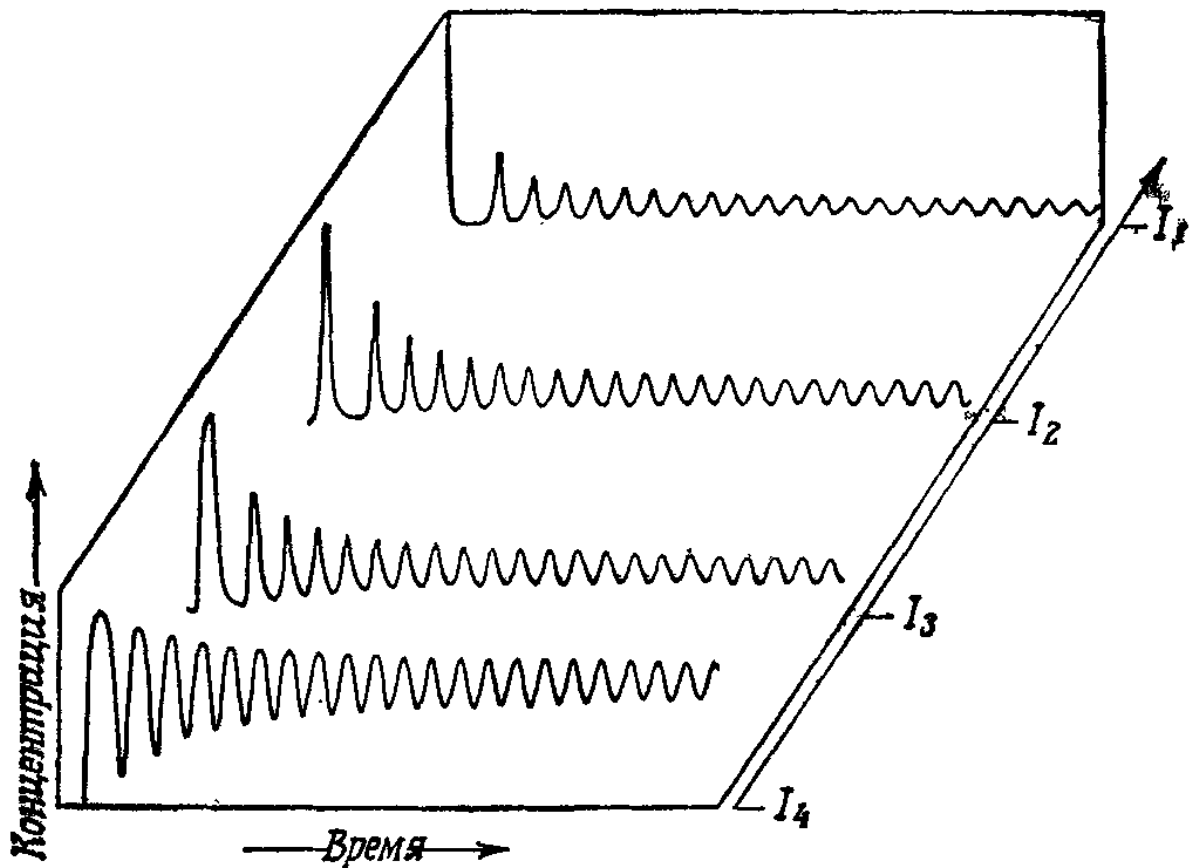


Рис. 35. Интегральные кривые динамической системы, соответствующей элементарному гиперциклу размерности $n = 4$ с неравными константами скоростей [$k_1 = 0,25$, $k_2 = 1,75$, $k_3 = 1,25$, $k_4 = 0,75$; начальные условия: $x_1(0) = 0,9997$, $x_2(0) = x_3(0) = x_4(0) = 0,0001$; полная концентрационная шкала = 1 единице концентрации, полная временная шкала = 1000 единиц времени]. Отметим, что концентрация I_1 (компонент, предшествующий самой быстрой стадии) оказывается наименьшей, а концентрация I_4 (компонент, предшествующий самой медленной стадии) — наибольшей.

Численное интегрирование действительно убедительно указывает на существование предельного цикла, или замкнутой траектории. Начиная движение из различных точек, очень близких к центру, грани, ребру или вершине симплекса, мы всегда через достаточно большое время приходим к одному и тому же предельному циклу. На рис. 34, $B - E$ показаны две типичные траектории для элементарных гиперциклов размерности $n = 5$ и $n = 12$. Как можно убедиться из сравнения этих двух рисунков, при увеличении n предельный цикл все ближе подходит к контуру $\overline{12}, \overline{23}, \dots, \overline{n1}$, о котором говорилось в предыдущем разделе. Следовательно, колебания отдельных концентраций становятся все более и более похожи на прямоугольные импульсы.

Использование численных методов позволяет также снять ограничение $k_1 = k_2 = \dots = k_n$. Были проведены вычисления для динамических систем с размерностями $n = 4$ и $n = 5$ и с произвольными значениями k . Оказалось, что общий характер интегральных кривых не изменился. Типичные примеры приведены на рис. 35 и 36. В обеих системах отдельные концентрации

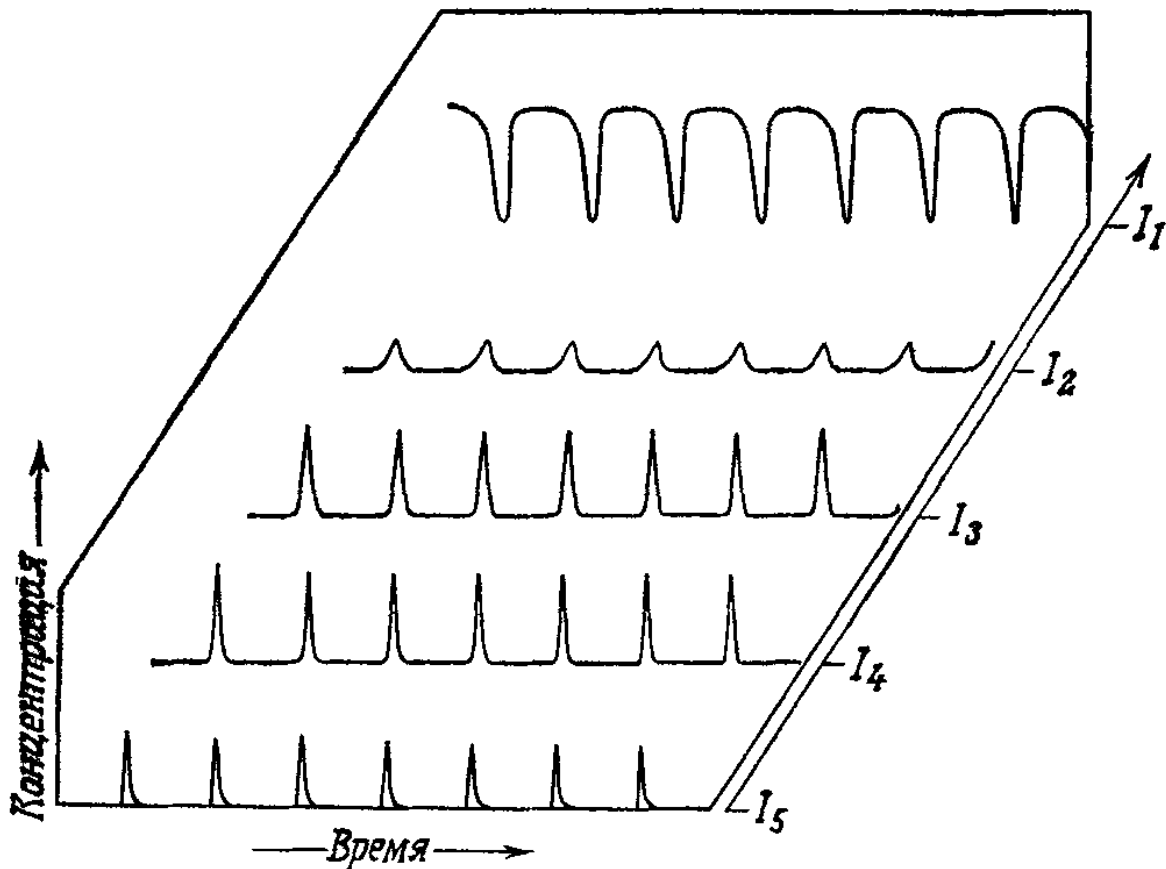


Рис. 36. Интегральные кривые динамической системы, соответствующей элементарному гиперциклу размерности $n = 5$ с неравными константами скоростей [$k_1 = 25/13$, $k_2 = 1/13$, $k_3 = 19/13$, $k_4 = 1$, $k_5 = 7/13$; начальные условия: $x_1(0) = 0,9996$, $x_2(0) = x_3(0) = x_4(0) = x_5(0) = 0,0001$; полная концентрационная шкала = 1 единице концентрации, полная временная шкала = 1000 единиц времени]. Отметим, что концентрация I_5 (компонент, предшествующий самой быстрой стадии) оказывается наименьшей, а концентрация I_1 (компонент, предшествующий самой медленной стадии) — наибольшей.

колеблются. Для $n = 4$ концентрационные волны затухают и динамическая система приближается к центральной особой точке. Ее координаты определяются следующими уравнениями:

$$\bar{x}_0 : \bar{x}_i^0 = \frac{k_i^{-1}}{\sum_{l=1}^n k_l^{-1}} c_0; \quad j = i + 1 - n\delta_{in}. \quad (72)$$

Пятичленные гиперциклы с неравными константами скоростей дают такие же незатухающие концентрационные импульсы, как и в случае систем с равными значениями k . Однако величина импульсов теперь не одинакова для всех компонентов. Средние по времени концентрации [определенные по формуле (67)] удовлетворяют уравнению (72), которое определяет положение (неустойчивой) центральной особой точки. Соответственно для тех видов, которые предшествуют стадии с относительно малой константой скорости, импульсы оказываются широкими, а для ви-

дов, предшествующих относительно быстрой реакционной стадии, — малыми по ширине и высоте. Итак, система регулирует концентрации своих компонентов таким образом, чтобы оптимизировать суммарную скорость продукции.

Гиперциклы большей размерности ($n \geq 5$) не остаются в устойчивых состояниях с постоянными стационарными концентрациями, а обнаруживают волнообразные колебания вокруг неустойчивой особой точки в центре. Тем не менее поведение компонентов является кооперативным, поскольку их концентрации регулируются динамикой всей системы, и ни одна популяционная переменная не обращается в нуль.

Динамические системы, соответствующие элементарным гиперциклам, имеют один и только один аттрактор внутри симплекса, бассейн которого распространяется на всю область положительных (ненулевых) концентраций всех компонентов. При малой размерности ($n \leq 4$) аттрактор является асимптотически устойчивой особой точкой, а именно фокусом для $n=2$ и спиральным стоком для $n=3$ и $n=4$. В системах с большей размерностью ($n \geq 5$) численное интегрирование убедительно указывает на существование устойчивого предельного цикла. Итак, все элементарные гиперциклы характеризуются кооперативным поведением компонентов.

Благодаря своим динамическим особенностям гиперциклы этого типа таят в себе множество еще не исследованных возможностей для самоорганизации (например, диссипативные структуры, если добавить сюда явления переноса). Они могут также играть важную роль в самоорганизации нервных сетей.

X. Гиперциклы с трансляцией

IX.1. Идеальные граничные условия и общие упрощения

Подходящий набор граничных условий можно реализовать в проточном реакторе [4, 9, 55, 56]. Концентрации всех низкомолекулярных соединений в нем (m_i , $i=1, 2, \dots, \lambda$) поддерживаются постоянными с помощью устройства, контролирующего потоки и в то же время снабжающего систему энергией. Концентрационные переменные относятся к макромоле-

кулярным видам, синтезируемым в реакторе, тогда как соответствующие параметры для всех других компонентов «стандартной реакционной смеси» не входят в явном виде в дифференциальные уравнения, но неявно входят в эффективные константы скоростей уравнения (30).

Из-за технических трудностей, а также по эвристическим причинам невозможно в явном виде учесть все элементарные стадии реакционного механизма. Вместо этого нам приходится использовать упрощенные схемы реакций, которые ведут к приемлемой «суммарной» кинетике. Такая стратегия обычна для химической кинетики. Например, кислотно-основные реакции в водных растворах обычно описываются феноменологическими уравнениями, которые не учитывают отдельных актов переноса протонов, а отражают только изменения в состоянии протонирования рассматриваемых молекул.

Кинетические уравнения, описывающие процессы полимеризации на матрице и трансляции, включают только численности популяций макромолекул, синтез которых завершен. Таким образом, инициация цепи и этапы ее роста в явном виде не рассматриваются. О правомочности этих аппроксимаций можно судить, сравнивая результаты с экспериментальными данными. В действительности тот тип «суммарной» кинетики, который мы здесь используем, достаточно хорошо известен (см. часть В).

IX.2. Кинетические уравнения

Каталитический гиперцикл, схематически представленный на рис. 37, состоит из двух наборов макромолекул: из n полинуклеотидов и n полипептидов. Репликация полинуклеотидов (I_i) катализируется полипептидами (E_i), которые в свою очередь являются продуктами трансляции полинуклеотидов. Гиперциклическая связь устанавливается с помощью динамических связей двух типов:

1. Каждый полинуклеотид I_i специфически транслируется в полипептид E_i . Очевидно, для трансляции

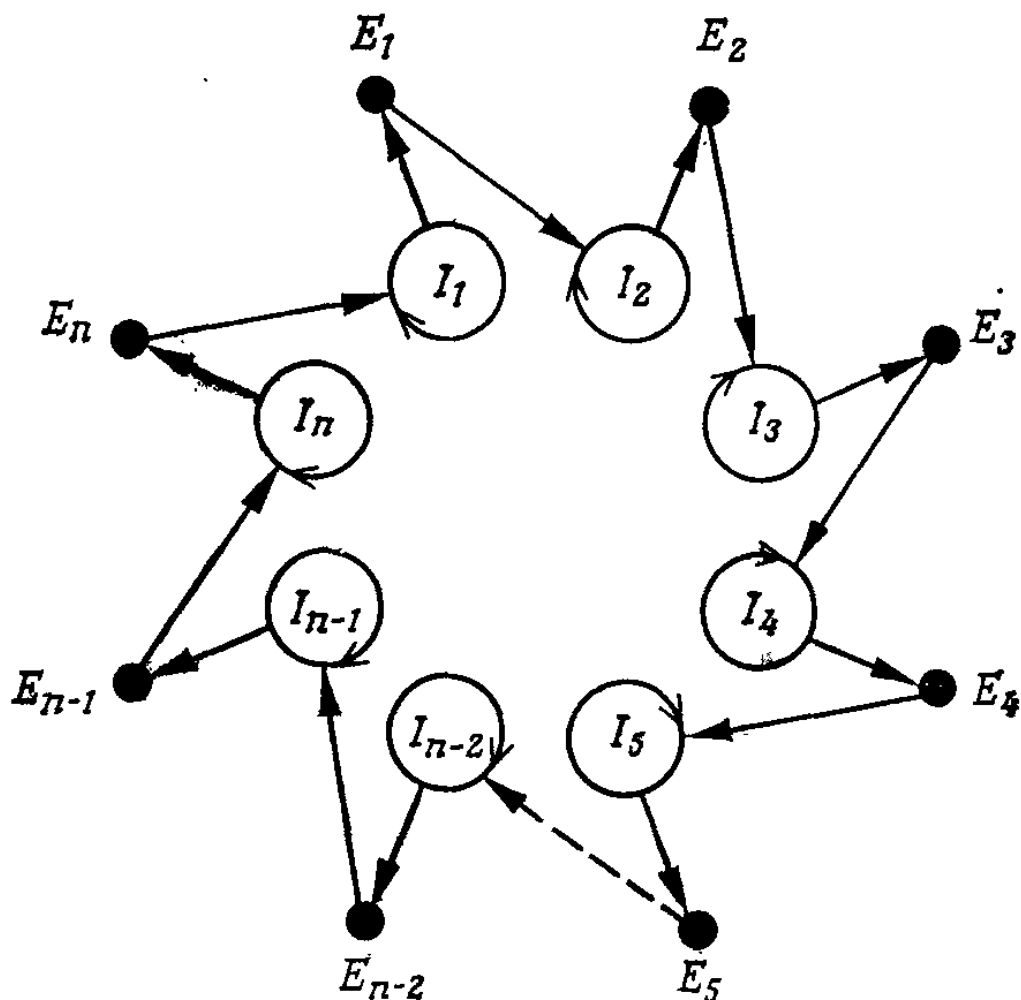


Рис. 37. Схематическое изображение гиперцикла с трансляцией. Размерность $2 \times n$, т. е. n полинуклеотидов и n полипептидов.

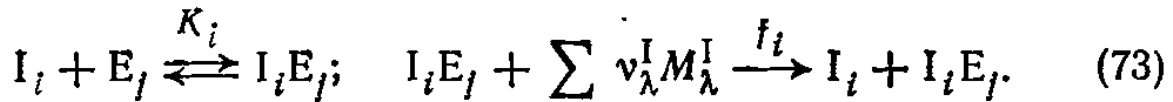
требуется наличие соответствующего аппарата, который включает в себя по меньшей мере некоторые из продуктов трансляции E_i и использует определенный генетический код.

2. Полинуклеотиды и полипептиды образуют специфические комплексы, которые также проявляют каталитическую активность при синтезе полинуклеотидных копий. Полипептиды могут быть специфическими репликазами или специфическими кофакторами белка с полимеразной активностью. Все вместе эти первичные белки обеспечивают наличие по меньшей мере двух функций: специфической репликаций и трансляции. В части В показано, как можно представить себе такую систему.

Связи между I_i и E_i должны быть такими, чтобы могло произойти замыкание цепи обратной связи (рис. 37). Математически циклическая симметрия вводится тем, что предполагается образование

специфического комплекса между ферментом E_j и полинуклеотидом I_i , причем $j = i + 1 - n\delta_{in}$.

Кинетика синтеза полинуклеотидов описывается уравнением Михаэлиса — Ментен, хотя мы не вводим предположения о том, что концентрации комплексов пренебрежимо малы:



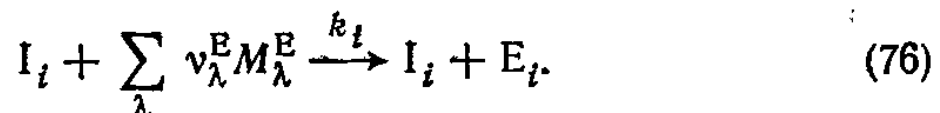
Четыре нуклеозидтрифосфата и их стехиометрические коэффициенты обозначаются соответственно через M_λ^I и v_λ^I , $\lambda = 1, 2, 3, 4$. Введем обозначения z_i для концентрации комплекса $I_i E_j$, x_j^0, y_i^0 — для полной концентрации полипептидов (E_j) и полинуклеотидов (I_i), x_j, y_i — для концентраций этих полимеров в свободном состоянии. Согласно закону сохранения масс

$$x_j^0 = x_j + z_i, \quad y_i^0 = y_i + z_i. \quad (74)$$

При быстром установлении равновесной концентрации комплекса концентрация z_i связана с полными концентрациями x_j^0 и y_i^0 следующим образом:

$$z_i = \frac{y_i^0 + x_j^0 + K_i}{2} \left(1 - \sqrt{1 - \frac{4y_i^0 x_j^0}{(y_i^0 + x_j^0 + K_i)^2}} \right). \quad (75)$$

Предполагается, что синтез полипептидов является неспецифичным, т.е. трансляция полинуклеотида I_i осуществляется с помощью общего «аппарата»:



Символами M_λ^E и v_λ^E обозначены активированные аминокислоты и их стехиометрические коэффициенты соответственно. Селекционные ограничения можно ввести посредством независимой регуляции полных концентраций обоих типов биополимеров (I и E). По аналогии с ограничением постоянной организации будем поддерживать обе суммы концентраций постоянными:

$$\sum_k y_k^0 = c_I^0; \quad \sum_k x_k^0 = c_E^0. \quad (77)$$

При всех этих условиях наша динамическая система, состоящая из $2n$ связанных дифференциальных уравнений, приобретает сле-

дующий вид:

$$\dot{y}_i^0 = f_i z_i - \frac{y_i^0}{c_1^0} \sum_{k=1}^n f_k z_k, \quad (78)$$

$$\dot{x}_i^0 = k_i y_i - \frac{x_i^0}{c_E^0} \sum_{k=1}^n k_k y_k, \quad i = 1, 2, \dots, n.$$

Нам достаточно обсудить два предельных случая:

1. При достаточно низких концентрациях концентрация z_0 комплекса $I_i E_j$ становится пропорциональной произведению концентраций полинуклеотида и полипептида:

$$y_i^0, x_j^0 \ll K_i; \quad z_i \sim \frac{1}{K_i} y_i^0 x_j^0; \quad z_i \ll y_i^0, x_j^0. \quad (79)$$

Если далее предположить, что процесс трансляции, реакция первого порядка, является быстрым по сравнению с репликацией, реакцией второго порядка (это допущение вполне обосновано, по крайней мере для низких концентраций полинуклеотидов), то концентрация полипептидов примет стационарное значение, которое можно включить в параметры скорости. Тогда образование полинуклеотидов будет описываться системой дифференциальных уравнений, типичной для элементарного гиперцикла размерности n .

2. При высоких концентрациях z_i становится равной меньшей из двух переменных¹:

$$z_i = \inf(y_i^0, x_j^0). \quad (80)$$

Соответственно мы получаем две возможные предельные ситуации:

$$K_i \ll y_i^0 \ll x_j^0: z_i \sim y_i^0, \quad (81)$$

$$K_i \ll x_j^0 \ll y_i^0: z_i \sim x_j^0. \quad (82)$$

В первом из этих двух случаев полинуклеотиды ведут себя как независимые конкуренты, тогда как полипептиды — из-за того что $y_i = y_i^0 - z_i \approx 0$ —

¹ \inf = infimum — математический термин, обозначающий минимальный член данного множества.

остаются в стационарных концентрациях. В естественных условиях, когда такие ограничения, как «постоянные полные концентрации», обычно не выполняются — по крайней мере, для предполагаемых малых значений y , — рост концентрации полинуклеотидов в результате привел бы к обращению концентрационных отношений $y:x$ и, следовательно, к выполнению условия (82). Вследствие этого начинают выполняться аппроксимации

$$z_i \approx x_i^0, \quad y_i \approx y_i^0, \quad (83)$$

что приводит к $2n$ -членному каталитическому циклу, но не к гиперциклу. Итак, в условиях насыщения, т. е. при высоких концентрациях компонентов, гиперцикл перестает вести себя как система с нелинейной скоростью роста. Как объединенная система, он имитирует свойства простого каталитического цикла, который эквивалентен автокатализатору или самовоспроизводящейся единице.

IX.3. Численные решения

Дифференциальные уравнения для каталитических гиперциклов с учетом в явном виде процесса образования комплекса между полинуклеотидами и полипептидами трудно исследовать аналитическими методами из-за иррациональности соответствующих выражений. В этих случаях численное интегрирование сопряжено с большими затратами времени, но тем не менее оно является единственным источником информации о свойствах этих динамических систем. Для иллюстрации динамики полинуклеотидно-полипептидных гиперциклов мы приведем графики решений, полученных с помощью ЭВМ, а также фазовые траектории. По сравнению с элементарными гиперциклами полинуклеотидно-полипептидные системы содержат новый класс параметров, а именно константы ассоциации комплексов K_i . Как и следовало ожидать из различий в кинетическом поведении при нижнем и верхнем концентрационных пределах, константы равновесия оказывают решающее влияние на

динамические свойства системы. С целью систематического исследования мы уменьшим число независимых параметров. Допущения здесь будут в основном те же, что и для элементарных гиперциклов: все константы скоростей для репликации полинуклеотидов $f_1 = f_2 = \dots = f_n = f$, для их трансляции в полипептиды $k_1 = k_2 = \dots = k_n = k$, а также все константы ассоциации $K_1 = K_2 = \dots = K_n = K$ предполагаются равными. Затем мы исследуем влияние K на свойства динамической системы при фиксированных значениях f и k для постоянного множества начальных концентраций. Для гиперциклов размерности $n \leq 4$ интегральные кривые через достаточно большое время приближаются к устойчивому стационарному состоянию. Отдельные концентрации могут испытывать затухающие колебания. Динамика этих систем в сущности такая же, как и для гиперциклов с большими значениями n и малыми константами равновесия.

Динамика гиперциклов большей размерности сложнее. Асимптотическое поведение системы изменяется с увеличением константы равновесия K . Ниже определенного критического значения $K_{кр}$ система сходится к устойчивым стационарным состояниям, тогда как при больших значениях K ($K > K_{кр}$) наблюдаются предельные циклы. По виду интегральных кривых и траекторий мы различаем четыре случая. Рассмотрим их в порядке увеличения константы равновесия K .

1. При малых значениях K динамическое поведение качественно такое же, как и для гиперциклов меньшей размерности. Интегральные кривые обнаруживают сильно затухающие колебания (рис. 38), а траектории быстро сходятся по спирали в центр — устойчивое стационарное состояние (рис. 39).

2. Общий тип динамического поведения в принципе такой же, как и в случае 1. Однако колебания затухают теперь слабо и приближение к стационарному состоянию становится крайне медленным (рис. 40, А, Б). Эта ситуация совершенно не похожа на случай 1, потому что члены, описывающие

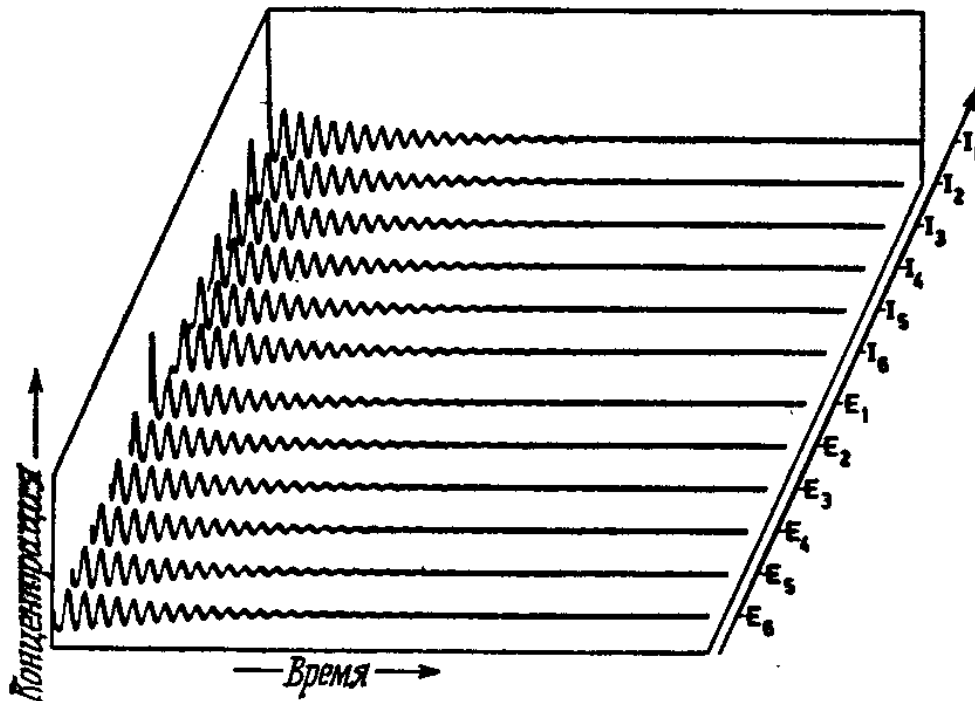


Рис. 38. Интегральные кривые динамической системы, соответствующей гиперциклу с трансляцией. Размерность 2×6 , $K = 0,25$; начальные условия: $y_1(0) = 5,0$, $y_2(0) = \dots = y_6(0) = 0,5$; $x_1(0) = \dots = x_6(0) = 1,0$; полная концентрационная шкала = 5 единицам концентрации, полная временная шкала = 1000 единиц времени. Значение константы равновесия K находится ниже критического уровня для бифуркации Хопфа; следовательно, наблюдаются затухающие колебания.

затухание, не появляются при исследовании нормальных мод, а требуют учета нелинейных вкладов. Феноменологически этот факт проявляется в том, что сначала амплитуды колебаний (почти) постоянны. Такая ситуация возникает, когда константа равновесия K немного меньше критического значения $K_{кр}$, т.е. $K = K_{кр} - \delta K$.

3. При значениях K , немного превышающих критическую константу равновесия ($K = K_{кр} + \delta K$), наблюдается интересное явление. Поведение динамической системы сначала во многом похоже на случай 2. Отдельные концентрации колеблются с относительно малыми амплитудами, но в противоположность случаю 2 амплитуды вначале немного растут. Однако вслед за этой фазой синусоидальных колебаний волны концентрации резко меняют свою форму и частоту (рис. 40, В, Г), а затем становятся очень похожими на прямоугольные импульсы, с которыми мы встречались при рассмотрении основных гиперциклов

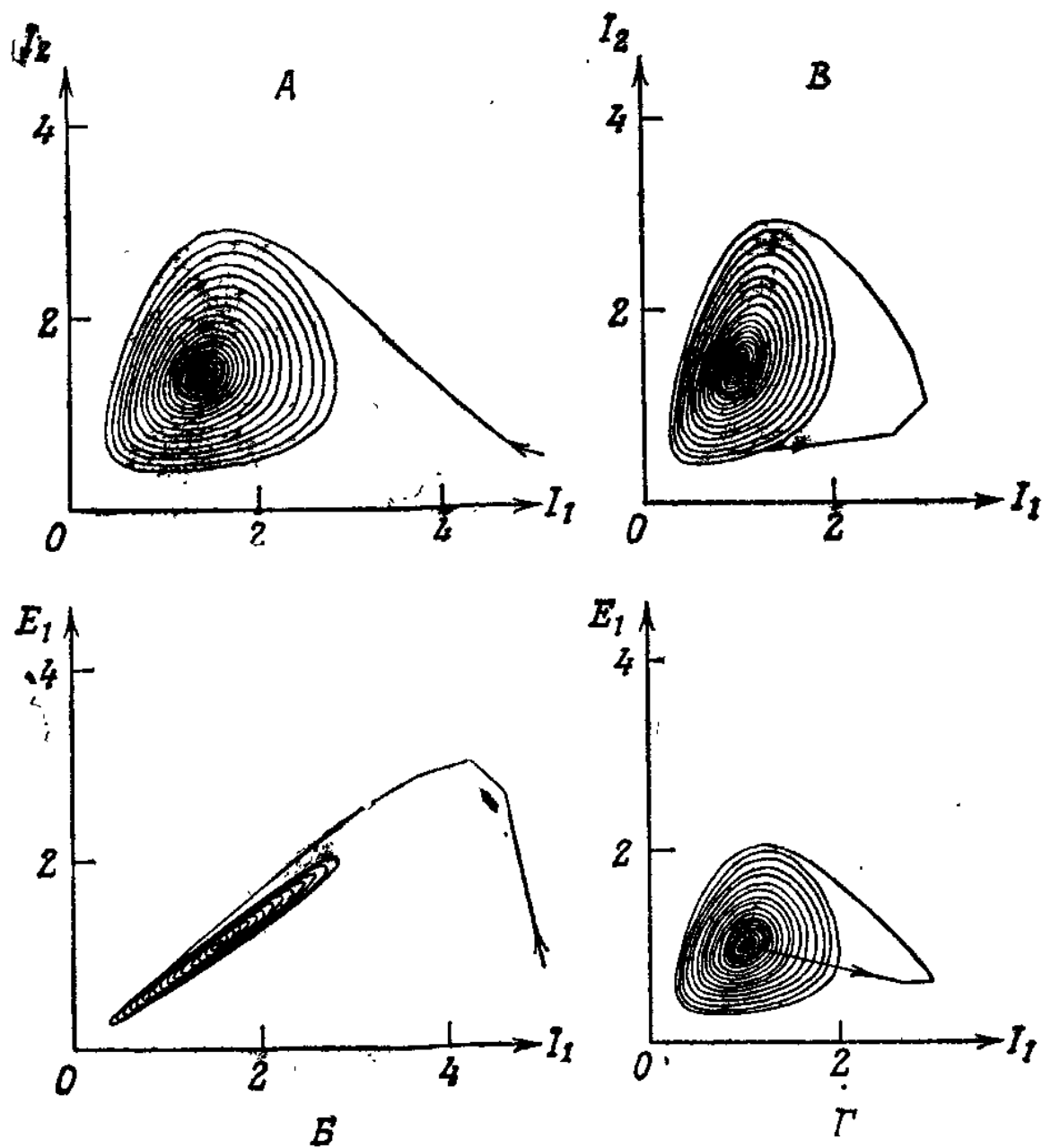
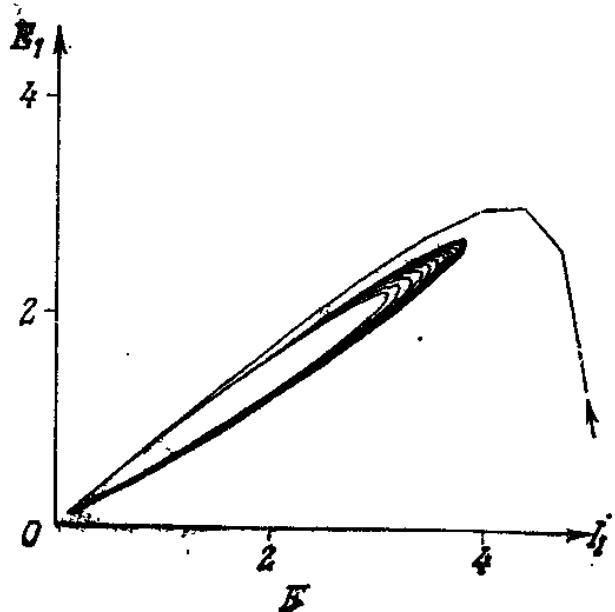
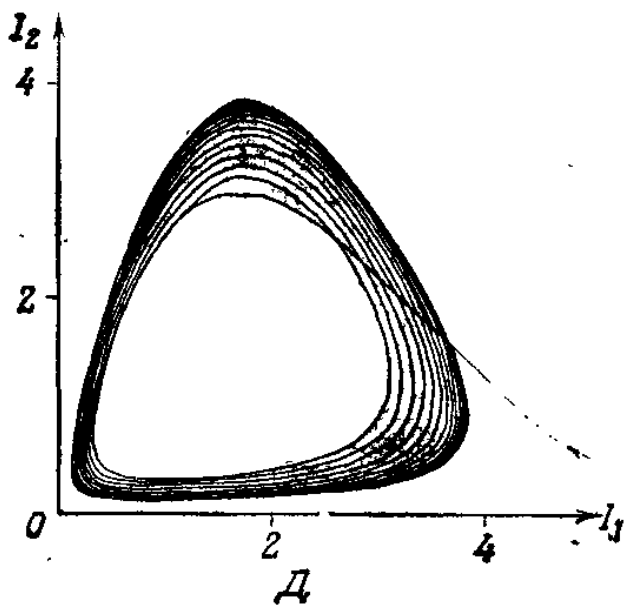
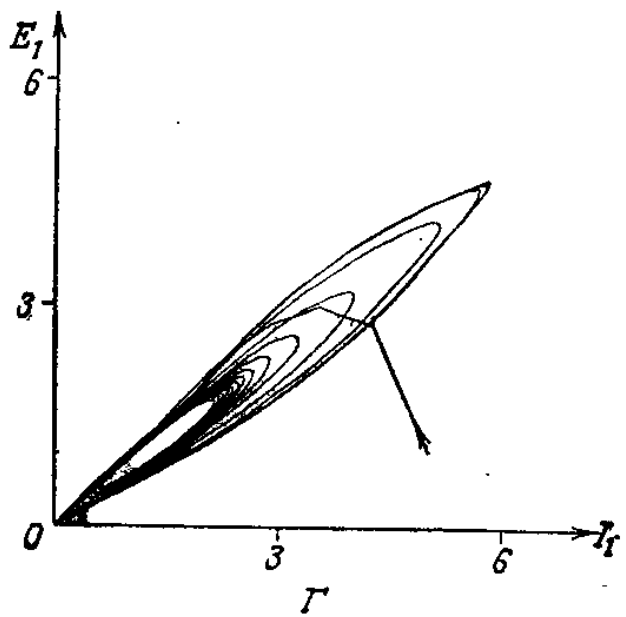
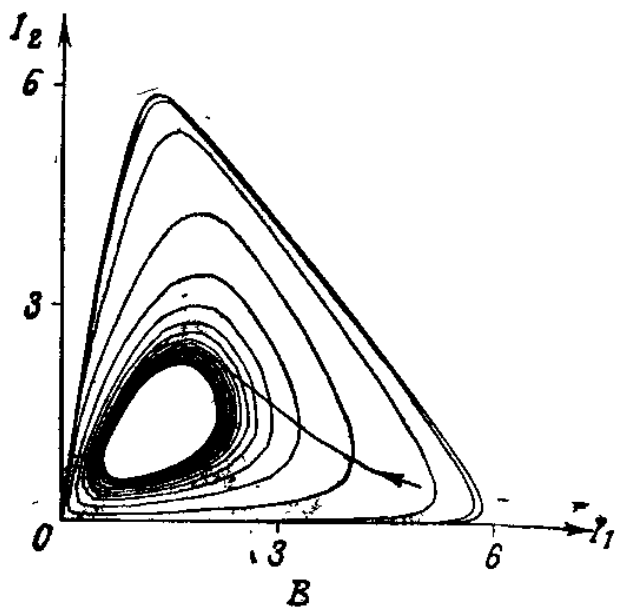
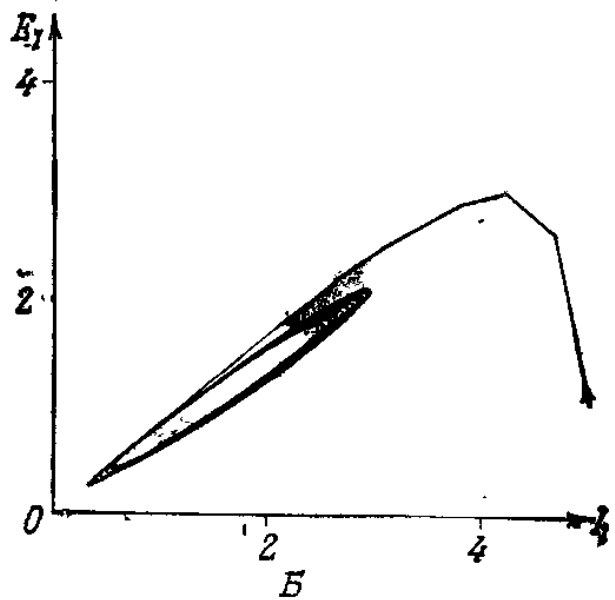
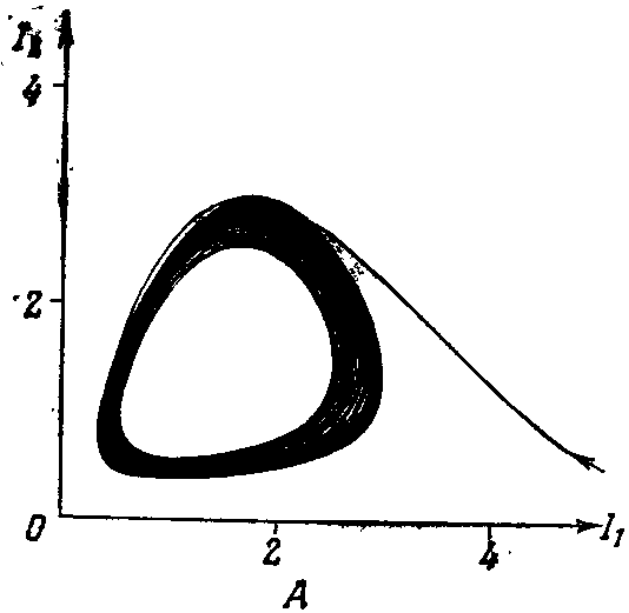


Рис. 39. Траектория динамической системы, соответствующей гиперциклу с трансляцией. Размерность 2×5 , $K = 1,0$; начальные условия: $y_1(0) = 5,0$; $y_2(0) = y_3(0) = y_4(0) = y_5(0) = 0,5$; $x_1(0) = x_2(0) = x_3(0) = x_4(0) = x_5(0) = 1,0$. А. Проекция траектории на плоскость (y_1, y_2) — на плоскость концентраций полинуклеотидов I_1 и I_2 . Б. Проекция на плоскость (y_1, x_1) — на плоскость концентраций полинуклеотида I_1 и его продукта трансляции — фермента E_1 . Отметим, что концентрация E_1 примерно пропорциональна концентрации I_1 , т. е. условие упрощения гиперцикла с трансляцией выполняется с хорошей точностью. В. Проекция на плоскость (x_1, y_2) — на плоскость концентраций полипептида E_1 и полинуклеотида I_2 , образование которого катализируется E_1 . Г. Проекция на плоскость (x_1, x_2) — на плоскость концентраций полипептидов E_1 и E_2 . Отметим, что K снова меньше критического значения для бифуркации Хопфа, и траектория стремится к центральной особой точке.



высокой размерности. Наконец, динамическая система приближается к предельному циклу.

4. При больших значениях K отдельные концентрации колеблются с возрастающей амплитудой, и динамическая система неуклонно приближается к предельному циклу (рис. 40, *Д*, *Е*). Такой тип изменения динамического поведения при непрерывном изменении параметра известен из литературы как бифуркация Хопфа [58]. Характерное замедление сходимости к асимптотическому решению, обнаруженное в случаях 2 и 3, было описано также и для других динамических систем; его обычно называют «критическим замедлением» при бифуркации Хопфа. В случае гиперциклов «замедление» вблизи критического значения K становится более выраженным с увеличением n . В пятичленном цикле ($n=5$) явления, соответствующие случаю 3, едва заметны. С другой стороны, каталитический цикл при $n=10$ имеет гораздо более длительный начальный период (о котором говорилось в п. 3), чем шестичленная система (рис. 41). Начальная фаза синусоидальных колебаний подобна метастабильному колебательному состоянию. Переход к конечному предельному циклу

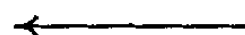


Рис. 40. Траектории динамических систем, соответствующих гиперциклам с трансляцией. *А* и *Б*. Размерность 2×5 , $K = 1,1$; начальные условия: $y_1(0) = 5,0$; $y_2(0) = y_3(0) = y_4(0) = y_5(0) = 0,5$; $x_1(0) = x_2(0) = x_3(0) = x_4(0) = x_5 = 1,0$; проекции на плоскости (I_1, I_2) и (I_1, E_1) соответственно; значение константы равновесия K немного ниже уровня бифуркации Хопфа, так что наблюдается очень медленная сходимости к устойчивой центральной особой точке. *В* и *Г*. Размерность 2×6 , $K = 0,2784$; начальные условия: $y_1(0) = 5,0$, $y_2(0) = \dots = y_6(0) = 0,5$; $x_1(0) = \dots = x_6(0) = 1,0$; проекции на плоскости концентраций (I_1, I_2) и (I_1, E_1) соответственно, значение константы равновесия K немного выше уровня бифуркации Хопфа, так что наблюдается метастабильный предельный цикл, а затем система сходится к устойчивому предельному циклу. *Д* и *Е*. Размерность 2×5 , $K = 1,2$; начальные условия: $y_1(0) = 5,0$, $y_2(0) = y_3(0) = y_4(0) = y_5(0) = 0,5$; $x_1(0) = x_2(0) = x_3(0) = x_4(0) = x_5(0) = 1,0$; проекции на плоскости (I_1, I_2) и (I_1, E_1) соответственно; значение константы равновесия K выше уровня бифуркации Хопфа, система непрерывно стремится к устойчивому предельному циклу. Отметим, что пропорциональность между E_1 и I_1 во всех трех случаях (*Б*, *Г*, *Е*) выполняется достаточно хорошо.

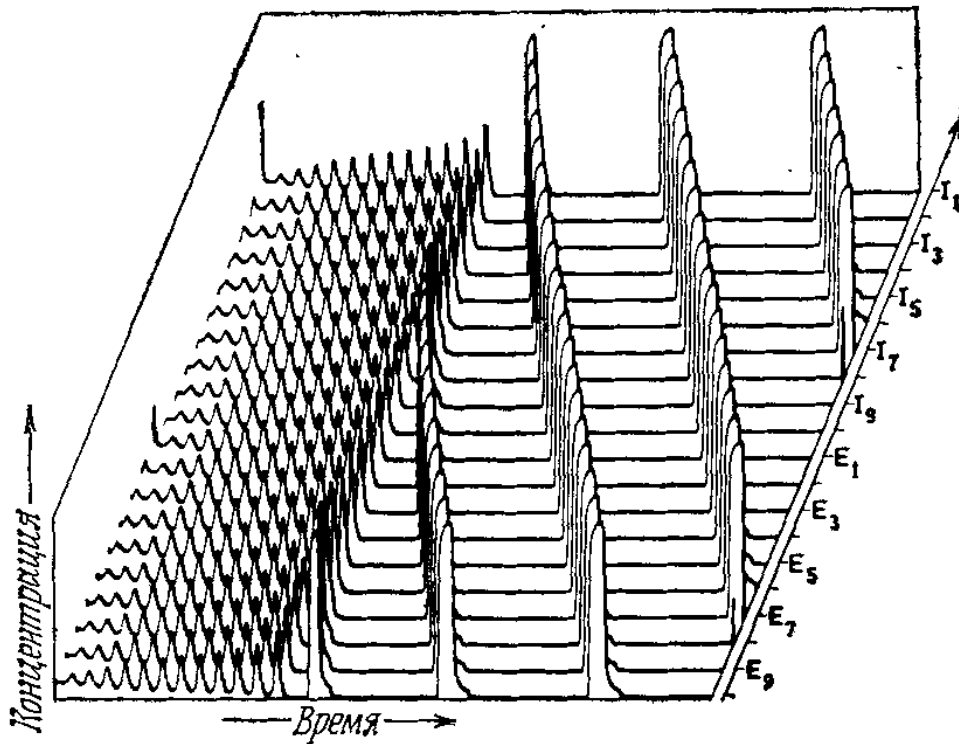


Рис. 41. Интегральные кривые динамической системы, соответствующей гиперциклу с трансляцией. Размерность 2×10 , $K = 0,026$; начальные условия: $y_1(0) = 5,0$; $y_2(0) = \dots = y_{10}(0) = 0,5$; $x_1(0) = \dots = x_{10}(0) = 1,0$, полная концентрационная шкала = 10 единицам концентрации, полная временная шкала = 1000 единиц времени. Выбранное значение константы равновесия немного выше критического уровня бифуркации Хопфа. Наблюдается метастабильное колебательное состояние, которое резко переходит в конечный предельный цикл с его характерными концентрационными волнами.

с ростом n становится более резким и четко выражен для десятичленного гиперцикла.

Все изученные полинуклеотидно-полипептидные гиперциклы имеют аттрактор внутри области значений концентраций, имеющей физический смысл. Они характеризуются кооперативным поведением своих компонентов. В зависимости от значения произведения полных концентраций (c_0^N и c_0^D) на константы ассоциации (K), а также от размера гиперцикла наблюдаются устойчивые особые точки или предельные циклы. При этом малые значения K «комплементарны» высоким концентрациям и *vice versa*. Асимптотическое поведение при нижнем и верхнем концентрационных пределах, полученное численным интегрированием, полностью согласуется с предсказаниями, основанными на анализе, проведенном в последнем

разделе. Одно из основных упрощений, касающееся квазистационарности полипептидного синтеза, может быть проверено непосредственно изучением проекций траекторий на плоскость (E_1, I_1) . В приближении стационарности должны получиться прямые. Как можно видеть из рис. 39, Б и 40, Б, Г, Е, пропорциональность двух концентраций примерно соблюдается, так что упрощенное рассмотрение, по-видимому, вполне оправдано. Целью численного анализа этого сложного реакционного механизма фактически была проверка эквивалентности сложного и элементарного гиперциклов в отношении их способности к самоорганизации. Таким образом, выводы, полученные на элементарных системах, справедливы также для всех типов реалистических гиперциклов с более сложной структурой (см. часть В).

Х. Сети гиперциклов

Х.1. Внутреннее уравнивание и конкуренция между гиперциклами

Концепция внутреннего уравнивания, введенная в разд. VI, кажется очень полезной для прямого анализа более сложных сетей, потому что она позволяет уменьшить число независимых переменных.

Сначала мы исследуем процесс уравнивания в элементарных гиперциклах. Для этого вычислим временные средние отдельных концентраций $X_i(t)$ [см. уравнение (67)] и сравним их с соответствующими интегральными кривыми $x_i(t)$ (рис. 42). Каким будет конечное состояние — стационарно инертным или колебательным, — не играет роли: временные средние $X_i(t)$ становятся практически постоянными после нескольких циклов. Поэтому допущение об установившемся внутреннем равновесии для гиперциклов кажется вполне оправданным приближением. Тем не менее в нескольких случаях мы проверим его справедливость.

Используя концепцию внутреннего уравнивания, можно вывести уравнение для чистой скорости

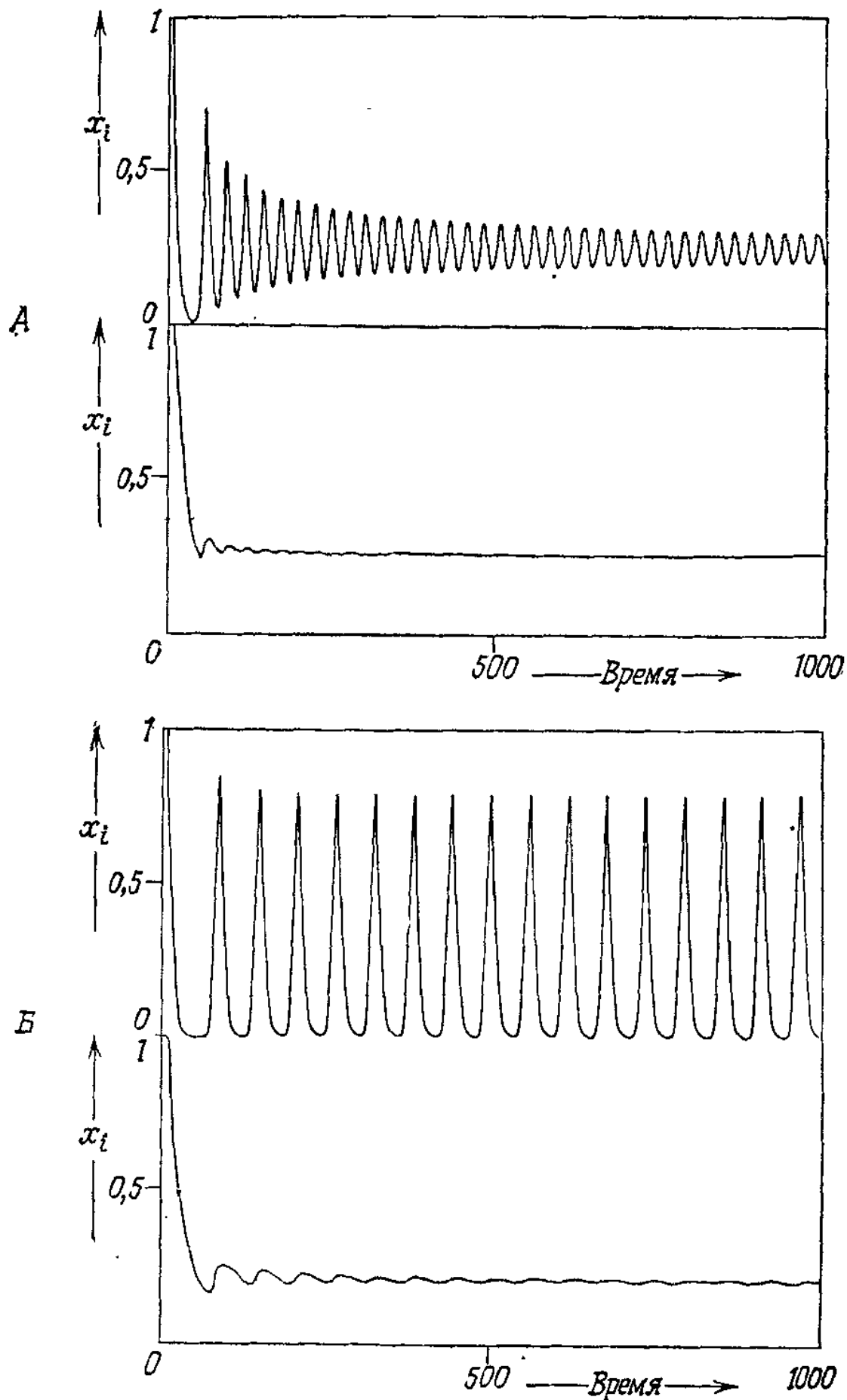


Рис. 42. Интегральные кривые динамических систем, соответствующих элементарным гиперциклам размерности $n = 4$ и $n = 5$ с равными константами скоростей и усредненными по времени концентрациями $X(t)$; *A* — $n = 4$, *B* — $n = 5$. Отметим, что $X(t)$ достигает \bar{X} через несколько периодов, т. е. в обоих примерах внутреннее равновесие устанавливается относительно быстро.

роста гиперцикла как целого:

$$\begin{aligned} \dot{c} &= \sum_{i=1}^n \dot{x}_i = \sum_{i=1}^n \Gamma_i(\mathbf{x}) = \sum_{i=1}^n k_i x_i x_j = \\ &= \frac{1}{\sum_{i=1}^n k_i^{-1}} c^2 \equiv \bar{k} c^2 = \Gamma(c), \end{aligned} \quad (84)$$

$$j = i + 1 - n\delta_{in}.$$

Итак, гиперциклы характеризуются квадратичными скоростями роста и следуют гиперболическому закону роста. Они представляют хорошие примеры того типа недарвиновского отбора «раз и навсегда», который обсуждался в разд. VI и VII.

Из выражения для \bar{k} [уравнение (84)] следует, что константа скорости для гиперцикла в целом будет иметь тот же порядок величины, что и константа скорости его самого медленного этапа.

При гипотетическом росте без ограничений гиперцикл растет до бесконечности за определенное критическое время (t_∞). В полностью уравновешенных системах эта неустойчивость наблюдается при

$$t_\infty^e = [kc(t=0)]^{-1}. \quad (85)$$

Результаты расчета для уравновешенных гиперциклов с использованием уравнения (85) можно сравнить со значениями из табл. 12, которые были получены численным интегрированием систем, далеких от внутреннего равновесия (t_∞^n). В полностью уравновешенных системах неустойчивость всегда возникает немного раньше: $t_\infty^e < t_\infty^n$. В целом эти численные различия имеют лишь второстепенное значение: общее поведение динамических систем и относительные значения t_∞ предсказываются правильно. Итак, допущение о внутреннем уравновешивании кажется хорошим приближением для большинства неуравновешенных систем.

Отбор среди гиперциклов, которые рассматриваются как целостные единицы, как правило, можнс

Таблица 12

**Неустойчивости в динамических системах для гиперциклов
при неограниченном росте**

| Размерность n | Граничные и начальные условия | | | Критические постоянные времени | |
|--------------------|-------------------------------|----------------------------------|--|---------------------------------|--|
| | константа скорости k | начальная концентрация $c(0)$ | начальное распределение ¹ $x(0)$ | в равновесии, t_{∞}^e | вдали от равновесия, t_{∞}^n |
| 2 | $1/2$ | 0,55 | (0,5; 0,05) | 3,64 | 5,0 |
| 3 | $1/3$ | 0,60 | (0,5; 0,05; 0,05) | 5,00 | 6,8 |
| 4 | $1/4$ | 0,65 | (0,5; 0,05; 0,05; 0,05) | 6,15 | 7,3 |

¹ Приведены распределения начальных концентраций, использованные при численном интегрировании неравновесной системы $x(0) = (x_1(0), x_2(0) \dots)$.

изучать при допущении внутреннего равновесия. Соответствующие динамические системы будут, конечно, идентичны системам, которые описывают независимых конкурентов, характеризующихся квадратичными скоростями роста. Конкуренцию между неуравновешенными гиперциклами исследовать труднее, поскольку возможно лишь численное интегрирование системы дифференциальных уравнений. В работе [53] был разобран пример, показывающий, что допущение внутреннего уравнивания является очень ценной аппроксимацией.

В качестве примера конкурирующих систем рассмотрим два гиперцикла H_A и H_B с n_A и n_B членами при ограничении постоянной организации.

В условиях внутреннего уравнивания система сводится к двум конкурентам с квадратичными скоростями роста. Напомним, что по результатам анализа особых точек гиперцикл H_A будет отбираться, когда его относительная начальная концентрация $c_A(0)$ превысит предельное критическое значение:

$$\lim_{t \rightarrow \infty} c_A(t) = c_0, \quad \text{если} \quad c_A(0) > \frac{k_B}{k_A + k_B} c_0. \quad (86)$$

В противном случае соревнование выигрывает гиперцикл H_B .

Поучительно рассмотреть еще один частный случай. Допустим, что в данном гиперцикле отдельные константы скоростей очень близки друг к другу, т. е. $k_1 \sim k_2 \sim \dots \sim k_n = \bar{k}_A$ и $k_{n+1} \sim k_{n+2} \sim \dots \sim k_{n+m} = \bar{k}_B$. Тогда константы скоростей для целых гиперциклов получаются следующим образом:

$$k_A = \frac{1}{n_A} \bar{k}_A \quad \text{и} \quad k_B = \frac{1}{n_B} \bar{k}_B. \quad (87)$$

Как мы видим, эти константы обратно пропорциональны числу членов гиперцикла, и, следовательно, меньшие циклы, по-видимому, будут иметь определенное селективное преимущество. Однако если мы допустим, что концентрации всех макромолекул примерно одинаковы (и равны \bar{x}), то невыгодность большего цикла точно компенсируется большим значением полной концентрации c :

$$c_A(0) = n_A \cdot \bar{x}, \quad c_B(0) = n_B \cdot \bar{x} \quad \text{и} \quad c_0 = (n_A + n_B) \bar{x},$$

$$\lim_{t \rightarrow \infty} c_A(t) = c_0, \quad \text{если} \quad \bar{k}_A > \bar{k}_B. \quad (88)$$

Таким образом, шансы на выживание для гиперциклов разных размеров или размерности n примерно одинаковы при условии равных начальных концентраций отдельных компонентов и равных констант скорости для этапов репликации.

Результаты, полученные для двух гиперциклов, легко обобщить на случай N независимых конкурентов.

X.2. Паразитная связь и каталитические сети

Циклически замкнутая каталитическая связь, соединяющая все активные члены $I_1 \dots I_n$ гиперцикла, может включать точки разветвления, что приводит к появлению внешних видов $I_{k \neq 1 \dots n}$, которые не являются составной частью кооперативной единицы. Мы будем называть эти внешние члены паразитами. Чтобы исследование можно было проводить в аналитической форме, допустим, что в цикле существует внутреннее равновесие (табл. 13). Было проведено исследование особых точек для двух динамических систем, описывающих гиперцикл с одночленным паразитом.

Первый пример — это каталитический гиперцикл и паразит, не способный к саморепликации (рис. 43, A). Выше определенного порогового уровня полной концентрации ($k_A c_0 > k$), как видно из

**Исследование особых точек гиперциклов
с паразитными связями**

Гиперцикл с присоединенной паразитной единицей I_x (рис. 43) описывается $n + 1$ дифференциальными уравнениями, из которых n идентичны уравнениям для изолированного гиперцикла. Для $(n + 1)$ -го дифференциального уравнения, которое определяет концентрацию паразитной единицы, получим: $[I_x] = x$, $[I_v] = x_v$,

$$\dot{x} = k_x x_v - \frac{x}{c_0} \phi \quad (\text{T.13.1})$$

для системы, изображенной на рис. 43, А, и

$$\dot{x} = k_x x_v x - \frac{x}{c_0} \phi \quad (\text{T.13.2})$$

для системы, изображенной на рис. 43, Б.

Если внутри гиперцикла установилось равновесие, то $n + 1$ дифференциальных уравнений сводятся к двум. Здесь оказывается удобным ввести новые константы скорости

$$k = k_x \cdot \frac{x_v}{c_A}; \quad c_A = \sum_{j=1}^n x_j \quad (\text{T.13.3})$$

и k_A , определенную уравнением (84).

(1) *Паразитная единица не является самовоспроизводящейся (см. рис. 43, А).*

$$\dot{c}_A = k_A c_A^2 - \frac{c_A}{c_0} (k_A c_A^2 + k c_A); \quad c_0 = c_A + x, \quad (\text{T.13.4})$$

$$\dot{x} = k c_A - \frac{x}{c_0} (k_A c_A^2 + k c_A).$$

Эта динамическая система имеет две особые точки:

$$\bar{x}_1: \bar{c}_A = \frac{k_A c_0 - k}{k_A}, \quad \bar{x} = \frac{k}{k_A}, \quad (\text{T.13.5})$$

$$\omega^{(1)} = - \frac{(k_A c_0 - k)^2}{k_A};$$

$$\bar{x}_2: \bar{c}_A = 0, \quad \bar{x} = c_0; \quad \omega^{(2)} = 0. \quad (\text{T.13.6})$$

Точка \bar{x}_1 устойчива при условии, что полная концентрация не равна критической $c_0 = k/k_A$.

Продолжение табл. 13

Анализ устойчивости \bar{x}_2 требует детального исследования членов более высокого порядка. Для точки $x = c_0 - \delta x$ находим

$$\dot{x} = \frac{(\delta x)^2}{c_0} (k - k_A c_0 + k_A \delta x). \quad (\text{T.13.7})$$

Отсюда следует, что

$$\dot{x} > 0 \quad \text{при} \quad c_0 < k/k_A$$

и

$$\dot{x} < 0 \quad \text{при} \quad c_0 > k/k_A.$$

Итак, особая точка \bar{x}_2 устойчива при концентрациях, меньших критического значения, и неустойчива при больших концентрациях. При низких концентрациях точка \bar{x}_1 находится вне области значений концентрации, имеющих физический смысл, точка $\bar{x} = c_0$ является единственным устойчивым стационарным состоянием, и гиперцикл разрушается паразитом. При больших концентрациях устойчивая особая точка \bar{x}_1 лежит на симплексе S_2 . Это означает, что гиперцикл и паразит сосуществуют.

(2) *Паразитная единица является самовоспроизводящейся (рис. 43, Б).*

$$\dot{c}_A = k_A c_A^2 - \frac{c_A}{c_0} (k_A c_A^2 + k c_A x), \quad c_0 = c_A + x,$$

$$\dot{x} = k c_A x - \frac{x}{c_0} (k_A c_A^2 + k c_A x). \quad (\text{T.13.8})$$

Система имеет две особые точки в вершинах S_2

$$\bar{x}_1: \bar{c}_A = c_0, \quad \bar{x} = 0, \quad \omega^{(1)} = (k - k_A) c_0, \quad (\text{T.13.9})$$

$$\bar{x}_2: \bar{c}_A = 0, \quad \bar{x} = c_0, \quad \omega^{(2)} = 0. \quad (\text{T.13.10})$$

Первая особая точка устойчива, если $k_A > k$. Для второй особой точки опять нужно исследовать члены более высокого порядка. При $x = c_0 - \delta x$ имеем

$$\dot{x} = \frac{k - k_A}{c_0} (\delta x)^2 (c_0 - \delta x). \quad (\text{T.13.11})$$

Теперь отсюда следует, что

$$\dot{x} > 0 \quad \text{при} \quad k > k_A$$

и

$$\dot{x} < 0 \quad \text{при} \quad k < k_A.$$

Итак, \bar{x}_2 устойчива, если выполняется неравенство $k > k_A$. Система конкурентна — это означает, что гиперцикл и паразитная единица не могут сосуществовать, за исключением особой ситуации, когда константы скоростей равны ($k = k_A$).

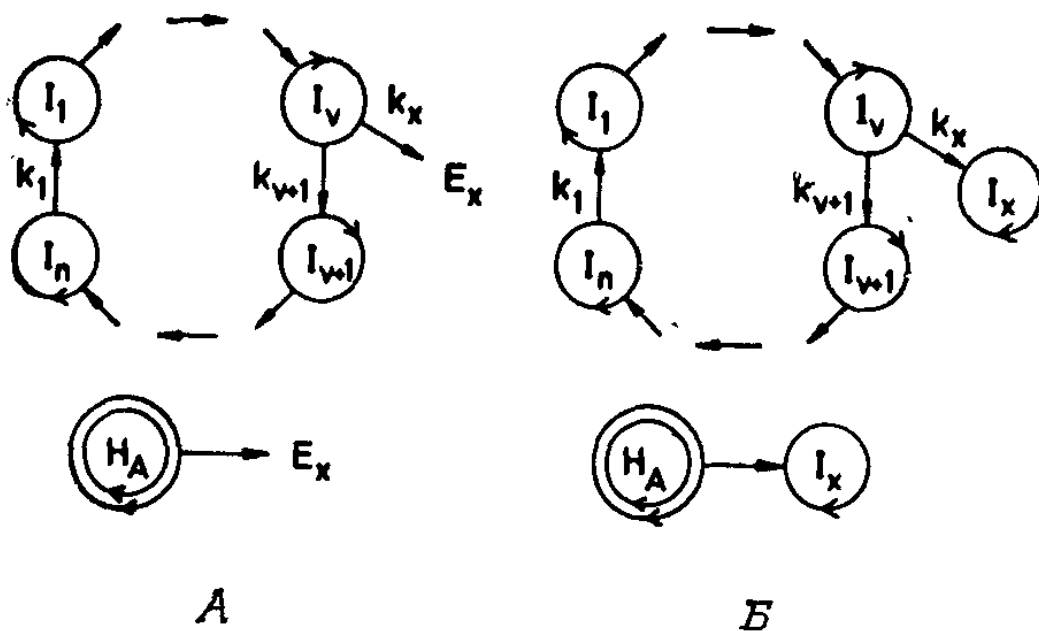


Рис. 43. Схематическое изображение гиперциклов с паразитными единицами. А. Паразитная единица не является самовоспроизводящейся. Б. Паразитная единица является самовоспроизводящейся. Предполагается, что разветвление происходит на уровне компонента. $\odot I_v$.

табл. 13, гиперцикл и паразит будут находиться в стационарном состоянии в ненулевой концентрации. Константа равновесия гиперцикла увеличивается с ростом c_0 , тогда как концентрация паразита остается постоянной. Следовательно, при достаточно большой концентрации паразит полностью теряет свое значение для динамики цикла. При малой полной концентрации ($k_A c_0 < k$) система становится неустойчивой. В рамках предположения о внутреннем равновесии паразит разрушает гиперцикл и в конце концов становится единственным оставшимся компонентом динамической системы.

Второй случай описывает развитие гиперцикла с самореплицирующимся паразитом (рис. 43, Б). Эта динамическая система характеризуется жестким отбором, зависящим от отношения констант скоростей k и k_A . При $k > k_A$ паразит разрушает гиперцикл, тогда как неравенство $k < k_A$ влечет за собой вымирание паразита. Возможно, представляет некоторый интерес рассмотрение этой динамической системы в явном виде на уровне отдельных полинуклеотидов. Из табл. 13 при условии установившегося внутренне-

го равновесия следует

$$k = k_x \frac{x_v}{c_A} = k_x \frac{k_{v+1}^{-1}}{\sum_i k_i^{-1}}. \quad (89)$$

Используя ранее полученное выражение

$$k_A = \left(\sum_i k_i^{-1} \right)^{-1},$$

находим

$$k < k_A \rightarrow \frac{k_x}{k_{v+1}} \cdot \frac{1}{\sum_i k_i^{-1}} < \frac{1}{\sum_i k_i^{-1}} \rightarrow k_x < k_{v+1}. \quad (90)$$

Итак, результат отбора полностью определяется отношением двух констант скоростей для реакционных этапов, находящихся на схеме гиперцикла сразу же за точкой разветвления, независимо от того, каковы значения других констант скоростей в гиперцикле.

Чтобы изучить влияние отклонений от равновесного распределения, было проведено численное интегрирование для динамических систем типа показанных на рис. 43. Фактически все полученные здесь выводы остаются в силе и для систем, далеких от равновесия.

Данные для одночленных паразитов можно распространить на произвольные цепи, используя результаты разд. VII.6. В общем случае судьба паразита тесно связана с развитием вида, прикрепленного к циклу: паразит всегда вымирает, когда концентрация вида, расположенного за точкой разветвления, приближается к нулю. Имеется один интересный частный случай: $k_x = k_{v+1}$. Дифференциальные уравнения для I_{v+1} и I_x идентичны, и, следовательно, соотношение между двумя видами всегда остается постоянным и равным своему начальному значению. Численное интегрирование для нескольких динамических систем этого типа показало, что в данном конкретном случае ($k_x = k_{v+1}$) все члены паразита, кроме вида I_x , будут вымирать.

Паразиты, имеющие вид цепочки, могут присоединиться к гиперциклу в двух местах, что приведет к образованию каталитической сети с точкой разветвления и точкой слияния. С помощью численного интегрирования мы показали, что системы такого типа неустойчивы: менее эффективная ветвь, т.е. ветвь с меньшими значениями констант скоростей k , вымирает, и остается только один простой гиперцикл.

Допуская произвольное расположение каталитических связей в системе самовоспроизводящихся макромолекул, мы гораздо чаще будем встречаться с

сильно разветвленными системами или сложными сетями, чем с регулярными гиперциклами. Поэтому, чтобы оценить вероятность образования гиперцикла, очень важно знать дальнейшее развитие таких систем. Для такого типа систем аналитические методы обычно непригодны, и мы должны, следовательно, полагаться на результаты численных методов.

Из набора интегральных кривых, полученных численным интегрированием дифференциальных уравнений для различных каталитических сетей, удалось сделать некоторые общие выводы. Как можно было ожидать, учитывая приведенные выше примеры, такие системы неустойчивы и распадаются на фрагменты меньших размеров. За исключением сложных динамических структур, обязанных своим существованием случайным совпадениям численных значений различных констант скоростей, возможные остатки каталитических сетей самореплицирующихся единиц — это только независимо растущие виды, каталитические цепи или каталитические гиперциклы. Итак, любая каталитическая сеть, состоящая из самореплицирующихся единиц с однородными членами связи, будет распадаться с образованием либо гиперцикла, который затем доминирует над другими фрагментами, либо конкурирующих динамических систем, которые не могут кооперативно эволюционировать.

Х.3. Иерархия связей между гиперциклами

В принципе гиперциклы можно связать друг с другом так, что образуется более высокоорганизованная система, если в кинетические уравнения ввести соответствующие каталитические члены. Рассмотрим два основных гиперцикла H_A и H_B и допустим, что гиперцикл H_A производит каталитический фактор роста для H_B и *vice versa*. Такой фактор роста может быть компонентом гиперцикла или его продуктом. Наш опыт подсказывает, что, по-видимому, взаимная каталитическая стимуляция будет приводить к кооперативному поведению.

Для упрощения прямого анализа допустим, что в обоих гиперциклах установилось внутреннее равновесие. Каталитические члены имеют третий порядок по концентрациям молекул ($k_A c_A^2 c_B$ и $k_B c_A c_B^2$ соответственно; см. также табл. 14). Следовательно, при достаточно больших концентрациях можно пренебречь функциями роста второго порядка для некатализируемой системы. Исследования особых точек недостаточно для изучения полученной динамической системы, потому что оно дает нулевые собственные значения для всех нормальных мод. Однако векторное поле легко поддается исследованию, так как система имеет только одну степень свободы. Как видно из рис. 44, два гиперцикла все еще конкурируют, несмотря на наличие каталитических факторов. Итак, введенный нами тип каталитической связи не является достаточным для кооперативного поведения.

Если повысить на единицу порядок каталитических членов, то в динамическую систему войдут члены роста четвертого порядка ($k_A c_A^2 c_B^2$, $k_B c_A c_B^2$). Анализируя векторное поле таким же образом, как и раньше (рис. 44), мы обнаруживаем устойчивую особую точку при конечных концентрациях обоих гиперциклов (см. также табл. 14). Итак, квадратичная связь достаточна для появления кооперативных отношений между каталитическими гиперциклами.

Представить физическую реализацию каталитической связи этого типа на уровне биологических макромолекул довольно трудно: наличие такого члена, как $k_A c_A^2 c_B^2$ или $k_B c_A c_B^2$, в суммарных кинетических уравнениях требует либо сложного многостадийного механизма, либо столкновения более чем двух макромолекул; обе эти возможности маловероятны¹. Поэтому мы склонны считать, что дальнейшее развитие по пути образования более сложных структур, состоящих из иерархически связанных

¹ Этих трудностей может не быть в случае искусственных динамических систем, где для введения каталитической связи используются технические устройства — например, электрические сети.

Таблица 14

Исследование особых точек каталитически
связанных гиперциклов H_A и H_B

(1) Члены связи третьего порядка.

$$\begin{aligned} \dot{c}_A &= k_A c_A^2 c_B - \frac{c_A}{c_0} \phi; & \phi &= (k_A c_A + k_B c_B) c_A c_B; \\ \dot{c}_B &= k_B c_A c_B^2 - \frac{c_B}{c_0} \phi; & c_0 &= c_A + c_B. \end{aligned} \quad (\text{T.14.1})$$

Исследование особых точек динамической системы дает

$$\bar{x}_1: \bar{c}_A = c_0, \quad \bar{c}_B = 0; \quad \omega = 0; \quad (\text{T.14.2})$$

$$\bar{x}_2: \bar{c}_A = 0, \quad \bar{c}_B = c_0; \quad \omega = 0. \quad (\text{T.14.3})$$

Система конкурентна. Исследование членов более высокого порядка (рис. 44) показывает, что \bar{x}_1 устойчива при $k_A > k_B$. С другой стороны, условие $k_A < k_B$ ведет к устойчивости \bar{x}_2 .

(2) Члены связи четвертого порядка.

$$\begin{aligned} \dot{c}_A &= k_A c_A^2 c_B^2 - \frac{c_A}{c_0} \phi; & \phi &= (k_A + k_B) c_A^2 c_B^2; \\ \dot{c}_B &= k_B c_A^2 c_B^2 - \frac{c_B}{c_0} \phi; & c_0 &= c_A + c_B. \end{aligned} \quad (\text{T.14.4})$$

Эта динамическая система характеризуется тремя особыми точками:

$$\bar{x}_1: \bar{c}_A = 0, \quad \bar{c}_B = 0; \quad \omega = 0; \quad (\text{T.14.5})$$

$$\bar{x}_2: \bar{c}_A = 0, \quad \bar{c}_B = c_0; \quad \omega = 0; \quad (\text{T.14.6})$$

$$\bar{x}_3: \bar{c}_A = \frac{k_A}{k_A + k_B} c_0; \quad \bar{c}_B = \frac{k_B}{k_A + k_B} c_0;$$

$$\omega = \frac{k_A^2 k_B^2}{(k_A + k_B)^3} c_0^3. \quad (\text{T.14.7})$$

Итак, \bar{x}_3 является устойчивой особой точкой, что указывает на кооперативное поведение двух связанных гиперциклов H_A и H_B при любых возможных условиях. Векторное поле, изображенное на рис. 44, показывает, что две другие особые точки, \bar{x}_1 и \bar{x}_2 , являются источниками.

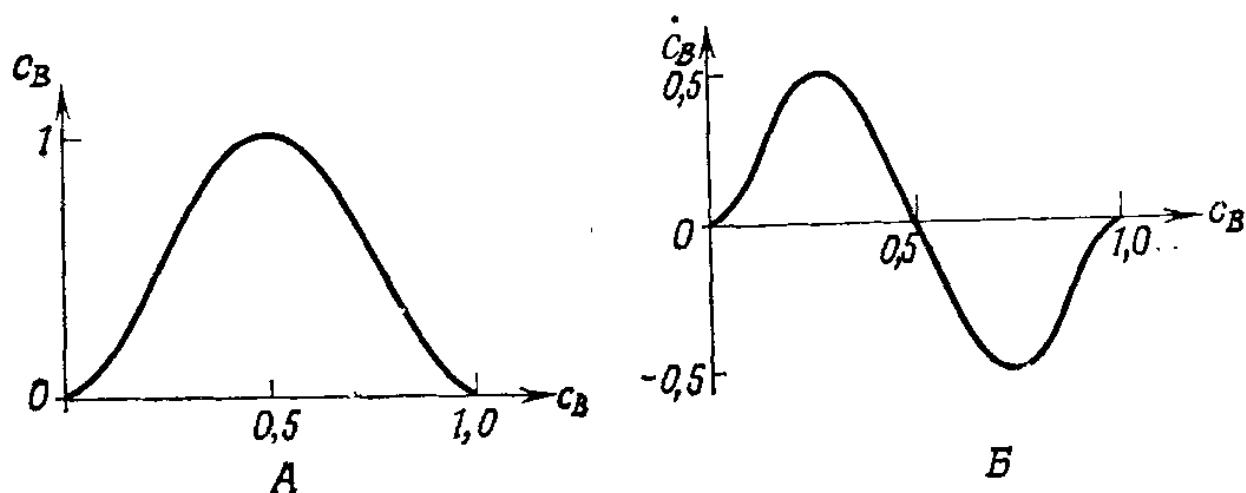


Рис. 44. Связь между гиперциклами. А. Члены каталитической связи — $k_A c_A^2 c_B$ и $k_B c_A c_B^2$. Касательный вектор положителен внутри области концентраций, имеющих физический смысл ($0 < c_B < 1$), за исключением двух особых точек; предполагается, что $k_B > k_A$, и, следовательно, отбирается гиперцикл c_B . Система конкурентна, несмотря на связь. Б. Члены каталитической связи — $k_A c_A^2 c_B^2$ и $k_B c_A^2 c_B^2$. Система содержит две неустойчивые особые точки в вершинах и устойчивую особую точку в центре ($c_A = c_B = 0,5$, так как $k_A = k_B$). Система кооперативна.

самореплицирующихся единиц, вряд ли может происходить посредством введения каталитических членов высокого порядка в систему, растущую в гомогенном растворе, — оно идет по пути индивидуализации уже существующих функциональных единиц. Это может быть достигнуто, например, пространственной изоляцией всех членов гиперцикла в компартменте. Образование прототипов современных клеток может служить одним из возможных механизмов, ведущих к индивидуализации гиперциклов. После того как изоляция завершилась, индивидуализованный гиперцикл может вести себя как простая репликативная единица. Следовательно, более вероятно, что гиперциклы являются промежуточным звеном в самоорганизации, а не ее конечной целью.

Выводы

Главной задачей части Б является абстрактное сравнительное изучение различных функциональных связей в самореплицирующихся системах. Используемые здесь методы обычны для дифференциальной топологии. Полные аналитические решения, за

исключением специальных случаев, обычно получить невозможно, так как соответствующие дифференциальные уравнения существенно нелинейны. Самовоспроизведение всегда ведет к зависимости скорости продукции от численности популяции соответствующего вида. Кооперация между различными видами через закодированные функциональные связи добавляет новые концентрационные члены, которые дают более высокий порядок зависимости скоростей от популяционных переменных.

Сравнительный анализ селекционного и эволюционного поведения не требует знания полных интегральных кривых. Обычно достаточно найти их конечные точки, для того чтобы решить, возможно ли устойчивое сосуществование всех партнеров в функциональном кооперативном ансамбле. Этой цели служит исследование особых точек в сочетании с методом Ляпунова и — в некоторых случаях — с более детальным изучением полного векторного поля. Результаты такого комбинированного анализа можно резюмировать следующим образом.

Функциональная интеграция ансамбля, состоящего из нескольких самореплицирующихся единиц, требует введения каталитических связей между всеми партнерами. Эти связи, наложенные на отдельные циклы репликации субъединиц, должны образовать замкнутый цикл, чтобы стабилизировать ансамбль путем взаимной регуляции всех популяционных переменных. Независимые конкуренты, которые при определенных пространственных условиях и на ограниченных интервалах времени могут сосуществовать в «нишах», а также каталитические цепи или разветвленные сети лишены способности к самоорганизации, типичной для гиперциклов.

Одного лишь сосуществования недостаточно для согласованного роста и эволюции всех партнеров ансамбля. Гиперцикл отличается следующими свойствами:

1. Он обеспечивает устойчивое и регулируемое сосуществование всех видов, объединенных циклической связью,

2. Он допускает согласованный рост всех своих членов.

3. Гиперцикл конкурирует с любой отдельной репликативной единицей, не принадлежащей циклу, независимо от того, является ли она независимой, представляет ли собой часть другого гиперцикла или даже присоединена к данному циклу паразитной связью.

4. Гиперцикл может увеличивать или уменьшать свои размеры, если такая модификация дает какое-либо селективное преимущество.

5. Гиперциклы с трудом сочленяются с образованием сетей более высокого порядка. Для взаимной стабилизации двух гиперциклов степени p необходимы члены связи степени $2p$.

6. Внутренние связи и кооперативные свойства гиперцикла могут эволюционировать по пути к оптимальному функционированию. «Фенотипические» преимущества, т. е. такие вариации, которые дают мутанту прямую выгоду, немедленно закрепляются. «Генотипические» же преимущества, которые благоприятствуют последующему продукту и поэтому лишь косвенно — репликативной единице, в которой произошла мутация, с другой стороны, требуют пространственной изоляции для конкурентной фиксации.

7. Отбор гиперцикла — это решение «раз и навсегда». В любой обычной дарвиновской системе мутанты, дающие селективное преимущество, легко могут размножиться и закрепиться. Их способность к росту не зависит от размера популяции. Для гиперциклов селективные преимущества всегда являются функциями численности популяции из-за существенно нелинейных свойств гиперциклов. Таким образом, однажды закрепившийся гиперцикл не может быть легко заменен никаким «новичком», поскольку новый вид всегда появляется сначала в количестве одной (или нескольких) копий.

Все эти свойства делают гиперцикл единственным в своем роде классом самоорганизующихся химических сетей. Это само по себе оправдывает более формальное изучение их свойств, что и было целью данной

части. Как было показано в части А, простые представители этого класса встречаются в природе. Возможно, этот тип функциональной организации широко распространен и играет какую-то роль в нервных сетях или в социальных системах. С другой стороны, мы не считаем, что гиперциклы — это какой-то фетиш. Их роль в молекулярной самоорганизации ограничена. Они позволяют интегрировать информацию, что было необходимо для возникновения трансляции.

Однако гиперцикл мог исчезнуть, как только появился ферментативный аппарат с высокой точностью репродукции, позволяющий интегрированной системе индивидуализоваться в форме живой клетки. Индивидуализованные репликативные системы имеют намного больший потенциал для дальнейшей диверсификации и дифференциации.

Существует много форм гиперциклической организации — от простой связи второго порядка до компунд-гиперцикла n -го порядка, в котором для каждого реакционного этапа требуется кооперативное действие всех членов. Хотя мы не знаем формы организации более простой, чем гиперцикл второго порядка, которая могла бы инициировать аппарат трансляции, мы хорошо сознаем сложность даже этой «простейшей возможной» системы. Поэтому нашей задачей в части В будет обоснование того, что реалистические гиперциклы действительно могут возникать из более простых предшественников, которые имелись в достаточных количествах в первичных условиях.

ЧАСТЬ В

РЕАЛИСТИЧЕСКИЙ ГИПЕРЦИКЛ

Предлагаемая модель «реалистического гиперцикла» тесно связана с молекулярной организацией примитивного аппарата репликации и трансляции. Гиперциклическая организация обуславливает селективную стабилизацию и эволюционную адаптацию всех гено- и фенотипических компонентов функционально связанного ансамбля. Она зарождается в форме молекулярного квазивида и усложняется в ходе эволюции путем мутаций и дупликаций генов. Ее первоначальная структура, по-видимому, отражается в установлении соответствия между кодонами и аминокислотами, в гомологии последовательностей тРНК, в двойных ферментативных функциях репликации и трансляции и в структурной и функциональной организации генома прокариотической клетки.

XI. Как могла зародиться трансляция?

«Общеизвестно, что возникновение белкового синтеза представляло собой очень трудную проблему. Мы имеем в виду не образование компонентов со случайной последовательностью, а синтез полипептидов, направляемый хотя бы в общих чертах матрицей — молекулой нуклеиновой кислоты, и имеющий такую природу, благодаря которой постепенно, шаг за шагом, в ходе эволюции сформировался современный генетический код, выражение которого сейчас требует наличия сложного аппарата активирующих ферментов, транспортных РНК, рибосом, различных факторов и т. д.».

Предмет нашего обсуждения трудно охарактеризовать более четко, чем этими вводными словами из

недавней статьи Крика, Бреннера, Клуга и Печеника [3].

Допустим на время, что примитивный аппарат репликации и трансляции, функционирующий с надлежащей точностью и адаптированный к достаточно богатому алфавиту молекулярных символов, возник в результате какого-то процесса, который далее не конкретизируется, например в результате самоорганизации или некоего акта творения, в природе или в лаборатории. Допустим, далее, что имеется среда, которая поставляет все активированные высокоэнергетические вещества, необходимые для синтеза таких макромолекул, как нуклеиновые кислоты и белки, для того чтобы репликация и трансляция могли быть спонтанными процессами, т. е. поддерживались бы положительным сродством. Мог ли такой ансамбль, каким бы путем он ни возник, продолжать эволюционировать как дарвиновская система? Другими словами, могла ли такая система неопределенно долго сохранять заложенную в нее первоначально информацию и совершенствовать ее дальше, пока не была бы достигнута максимальная функциональная эффективность?

Чтобы конкретизировать ситуацию, рассмотрим модель, изображенную на рис. 45. Плюс-цепи данного множества молекул РНК несут информацию для синтеза соответствующего числа молекул белков. Продукты трансляции могут выполнять по меньшей мере следующие функции: 1) Один белок играет роль РНК-полимеразы, подобной тем специфичным репликазам, которые связаны с различными РНК-содержащими фагами. Ее сайт узнавания адаптирован к специфической последовательности или структуре, имеющейся у всех плюс- и минус-цепей РНК; другими словами, она эффективно воспроизводит лишь те молекулы РНК, которые идентифицируются как члены определенного ансамбля. 2) Другие продукты трансляции функционируют как активирующие ферменты — они специфически узнают различные аминокислоты и присоединяют их к соответствующим адапторным РНК, каждая из которых не-

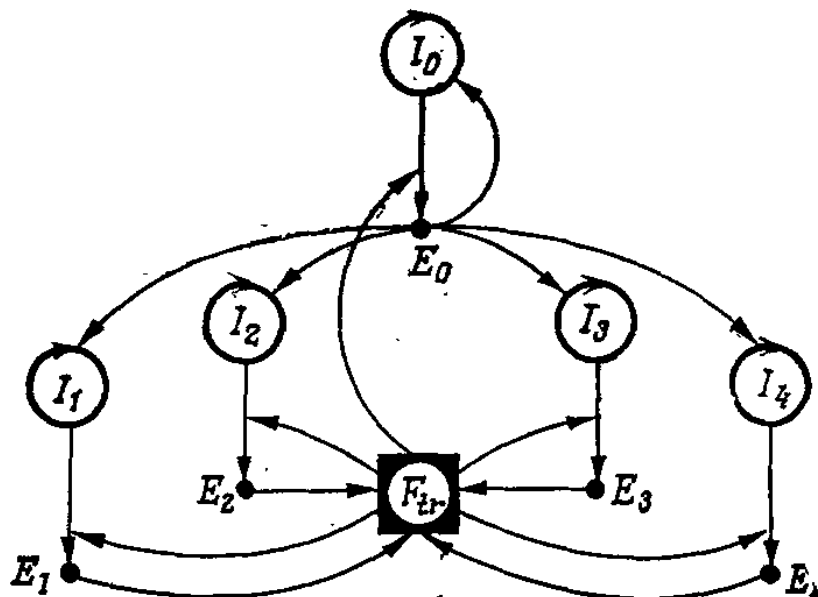


Рис. 45. Простейшая модель примитивной трансляции включает информационную РНК I_0 , которая кодирует репликазу E_0 , адаптированную к специфическому узнаванию последовательностей $I_0 - I_4$. Плюс-цепи $I_1 - I_4$ кодируют четыре синтетазы $E_1 - E_4$, а минус-цепи могут быть адапторами (тРНК) для четырех аминокислот. Такая система, хотя она и включает в себя все функции, необходимые для трансляции и самовоспроизведения, является неустойчивой из-за внутренней конкуренции. Если $I_0 - I_4$ не стабилизируются гиперциклической связью, то согласованная эволюция невозможна.

сет определенный антикодон. Число разных аминокислот и, следовательно, адапторов выбрано так, чтобы оно равнялось числу различных кодонов, встречающихся в информационных последовательностях, т. е. в плюс-цепях РНК, с тем, чтобы в результате получилась «замкнутая» система трансляции с определенным кодом. Она не обязательно должна включать полный генетический код в его современном виде, а может ограничиваться меньшим, но функционально достаточным числом аминокислот (например, четырьмя), используя определенные ограничения, наложенные на структуру кодонов, чтобы гарантировать однозначное считывание. В качестве адапторов можно использовать минус-цепи молекул РНК, или — если это условие окажется слишком жестким — адапторы можно ввести вместе с другими компонентами (например, с рибосомами) как постоянные факторы среды, подобные тем бактериальным факторам, которые обеспечивают репликацию фага и трансляцию в клетке-хозяине.

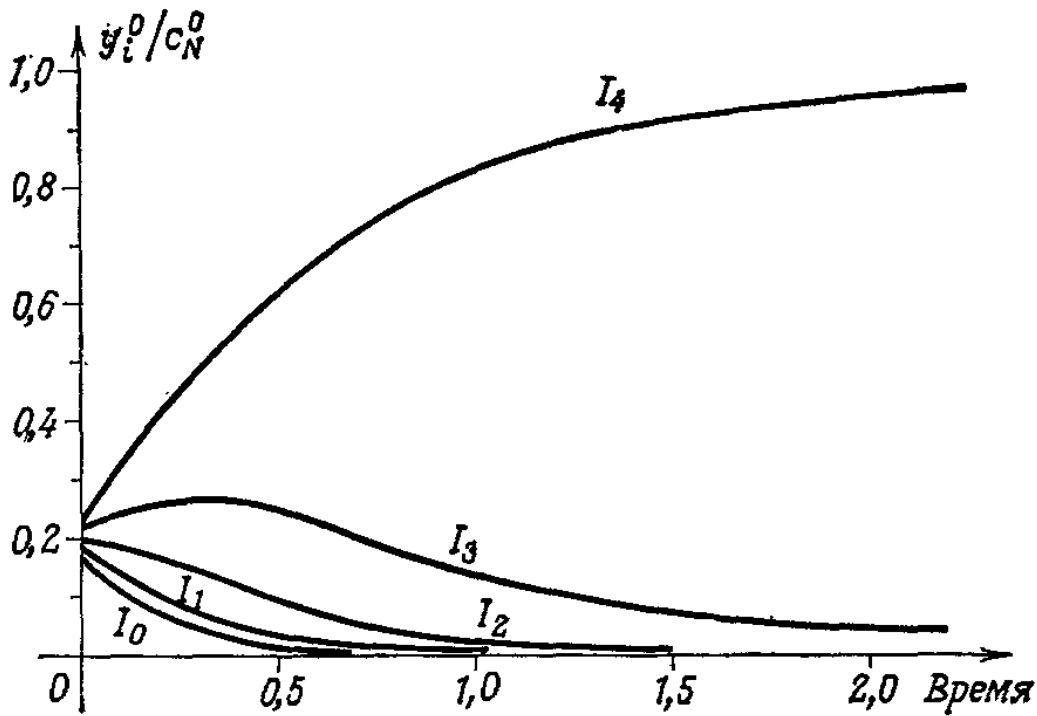


Рис. 46. Интегральные кривые для системы дифференциальных уравнений, имитирующих модель, изображенную на рис. 45. В этом частном случае предполагается, что начальные концентрации и автокаталитические константы скорости воспроизведения линейно растут при переходе от I_0 к I_4 , тогда как значения

других параметров — констант скоростей трансляции ($I_i \xrightarrow{k_i} E_i$), параметров, определяющих выбор аминокислот (вклады E_1, E_2, E_3, E_4 в F_{tr}), стабильностей фермент-субстратных комплексов ($I_i + E_0 \xrightleftharpoons{K_i} I_i E_0$) и т. д. — одинаковы для всех партнеров реакций. Кинетические кривые для относительных численностей популяций (y_i^0/c_N^0) отражают коикурентность поведения. Наиболее эффективно воспроизводящаяся матрица (I_4) в конце концов становится доминирующей ($y_4^0/c_N^0 \rightarrow 1$). Однако из-за того, что репликация (представленная E_0) и трансляция (вклады E_1, E_2 и E_3 в F_{tr}) прекращаются, I_4 тоже вымирает. Вся популяция обречена на гибель ($c_N^0 \rightarrow 0$).

С первого взгляда может показаться, что система, изображенная на рис. 45, в достаточной степени насыщена функциональными перекрестными связями: все I_i каталитически поддерживаются репликазой E_0 ; эта репликаза в свою очередь обязана своим существованием совместному действию ферментов трансляции $E_1 - E_4$, без которых она не смогла бы транслироваться с I_0 . Ферменты $E_1 - E_4$ для своего воспроизведения используют, конечно, эту же функцию трансляции, но поскольку они являют-

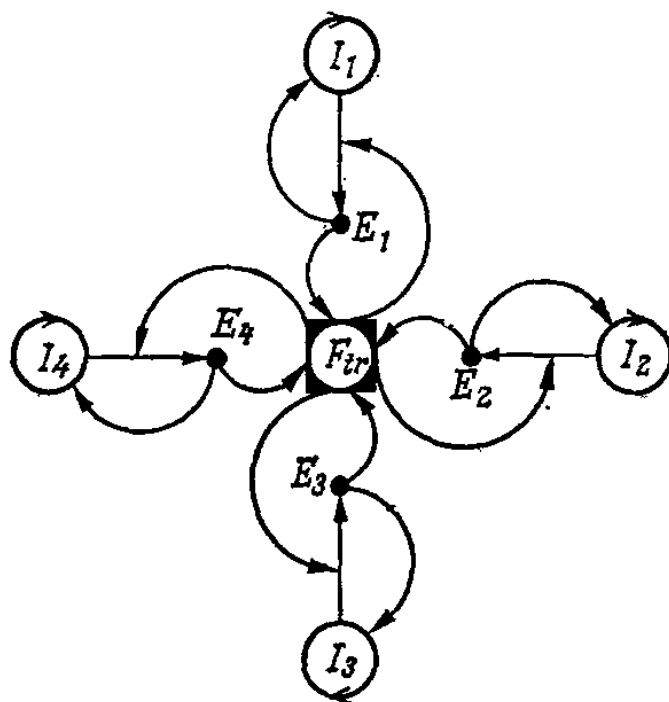


Рис. 47. В этой альтернативной модели примитивной репликации и трансляции предполагается, что ферменты $E_1 — E_4$ выполняют двойную функцию, т. е. что они играют роль специфичных репликаз для своих собственных мРНК и синтетаз при выборе четырех аминокислот. Судьба системы такая же, что и на рис. 45, поскольку информационные РНК сильно конкурируют друг с другом.

ся продуктами трансляции $I_1 — I_4$, то в конечном счете зависят также от E_0 или соответственно от I_0 .

Детальный анализ показывает, однако, что имеющихся связей недостаточно, чтобы гарантировать взаимную стабилизацию различных генотипических компонентов I_i . Общей репликазной функции E_0 и общей трансляционной функции F_{tr} во всех дифференциальных уравнениях соответствует один и тот же член. Поэтому эти уравнения сводятся к уравнениям для несвязанных конкурентов, умноженным на общую функцию времени $f(t)$. Из-за внутренней конкуренции система, которая сначала функционирует очень хорошо, обречена на распад. Типичный набор интегральных кривых, полученных численным интегрированием кинетических уравнений, приведен на рис. 46.

Другой пример такого же рода представлен на рис. 47. Здесь все мРНК производят свои специфические репликазы $E_1 — E_4$, которые выполняют также синтетазные функции (F_{tr}). И снова этой связи

через соответствующую трансляционную функцию оказывается недостаточно для стабилизации ансамбля. На наш вопрос — является ли наличие системы информационных РНК для репликационной и трансляционной функций и продуктов трансляции достаточным для их существования и эволюции — ответить следует так: если не введена связь определенного типа между различными репликативными компонентами I_i , то такие системы будут неустойчивы, несмотря на то что они обладают всеми свойствами, необходимыми для репликации и трансляции. Даже если бы все партнеры были селекционно эквивалентны (или почти эквивалентны) и, следовательно, могли бы сосуществовать некоторое время (в зависимости от размеров их популяций), они не смогли бы эволюционировать взаимно контролируемым образом и, следовательно, никогда не были бы способны оптимизировать свое функциональное взаимодействие. В конце концов они обязательно распадались бы, потому что случайная селекционная эквивалентность не может согласованно поддерживаться на протяжении длительных периодов эволюции, если она не подкреплена определенными связями.

Этот ответ не вызовет удивления, если учесть результаты части Б. Более тщательный анализ связей, обусловленных функционированием ферментов репликации и трансляции, не обнаруживает никакой гиперциклической структуры. Поэтому эти связи не могут обеспечить взаимную регуляцию величин популяционных переменных, которая необходима для согласованной эволюции членов организованной системы. Связи, имеющиеся в этих двух изученных системах, можно свести к двум общим функциям, которые, подобно факторам среды, действуют на всех партнеров совершенно одинаковым образом и, следовательно, не предоставляют никакой возможности для взаимного контроля.

Рассмотренные выше примеры иллюстрируют те положения, которые мы хотели подчеркнуть в данном разделе:

1. На ранних этапах эволюции, пока точность репликации и трансляции была низка и концентрации эффективно реплицирующихся единиц малы, гиперциклическая организация давала большое относительное преимущество над любой другой формой (структурной) организации (разд. XV).

2. Действительно, *удается* построить такие гиперциклические модели, которые можно считать реалистическими предшественниками аппарата воспроизведения и трансляции современных прокариотических клеток (разд. XVI).

Как представить возникновение трансляции, если допустить существование воспроизводимых молекул РНК, имеющих размеры тРНК, и наличие предпосылок для синтеза белков в примитивной форме, с использованием ограниченного числа (достаточно распространенных) аминокислот?

XII. Логика первичного кодирования

XII.1. Код RRY

Наиболее привлекательная спекулятивная модель возникновения матричного синтеза белков, предложенная совсем недавно [3], основана на целом ряде соображений, относящихся к проблеме связного считывания без запятых. Первичный код должен иметь определенную структуру рамок, иначе информация не сможет считываться согласованным образом. Случайные смещения фазы будут приводить к сдвигу рамки трансляции на каких-то участках информационной РНК и к искажению смысла. Поэтому авторы предлагают определенную последовательность оснований, которой должны придерживаться все кодоны. Или, другими словами, функционировать в качестве мРНК могут лишь те последовательности нуклеотидов, которые несут определенный паттерн («образ»). Единообразие паттерна может возникнуть благодаря инструкции, которую содержит выступающая антикодоновая петля тРНК, а также благодаря внутреннему самокопированию. Среди паттернов,

которые гарантируют неперекрывающееся считывание, авторы выбрали последовательность оснований пурин — пурин — пиримидин (или, в обычных обозначениях, RRY) как общую для всех кодонов, образующих сообщение. Этот выбор был продиктован регулярностью последовательности, которая была обнаружена в антикодоновой петле современных тРНК. Упомянутая последовательность такова: $3'NR\alpha\beta\gamma UY$, где $\alpha\beta\gamma$ — антикодон, N — любой из четырех нуклеотидов, а R и Y — пурин и пиримидин соответственно. Другой предпосылкой возникновения трансляции без рибосом является стабильность комплекса, образованного мРНК и растущей полипептидной цепью. Пептидил-тРНК не должна отделяться до того, как осуществится перенос следующей аминоацил-тРНК, т. е. пока не закончится трансляция всей мРНК, в противном случае синтезировались бы только функционально неэффективные белковые фрагменты. На основании известных данных о стабильности пар оснований очевидно, что простое взаимодействие кодон — антикодон не гарантирует необходимой стабильности комплекса мРНК — тРНК. Поэтому для обоснования модели существенны три вспомогательных допущения:

1. Структура антикодоновой петли адаптора (предшественника тРНК) такова, что если конкретизирован общий паттерн кодона, то РНК может всегда образовать пять пар оснований с мРНК. Таким образом, антикодоновой петле первичных тРНК приписывается следующая общая последовательность:



где YUR — антикодон.

2. Антикодоновая петля каждой первичной тРНК может принимать две разные конформации, изображенные на рис. 48. Обе они были описаны в более ранней статье Вёзе [60], который обозначил их FN и hf. (Обозначение FN образовано из первых букв фамилий Фуллера и Ходжсона [61] — ученых, кото-

рые предположили, что из семи неспаренных оснований в антикодоновой петле стопку образуют пять оснований, расположенных с 3'-конца; hf, по Вёзе, обозначает дополнительную конфигурацию, в которой стопку образуют пять оснований на 5'-конце петли.) Вёзе предположил, что переход между этими конфигурациями играет важную роль в синтезе белка на рибосомах, но указал также и на его возможное значение для прежних, более примитивных механизмов.

3. Какая из этих двух конфигураций действительно реализуется, зависит от того, с чем связана тРНК — с аминокислотой или с пептидной цепью. При переносе аминокислоты на пептидную цепь тРНК скачкообразно переходит из hf- в FH-конфигурацию (см. рис. 49).

Дополнительный, четвертый постулат, не абсолютно необходимый для обоснования модели, вводит взаимодействие между двумя соседними тРНК на мРНК, которое делает энергетически выгодной необходимую конфигурацию и вносит (дальнейший) вклад в стабилизацию комплекса полипептид — мРНК.

Рис. 49 показывает более детально, как может облегчаться синтез полипептидов, если учесть приведенные соображения. Растущая полипептидная цепь перемещается вдоль мРНК, используя в качестве «топлива» свободную энергию реакции переноса. Эта реакция может ускоряться общим неспецифическим катализатором, но такой механизм не требует

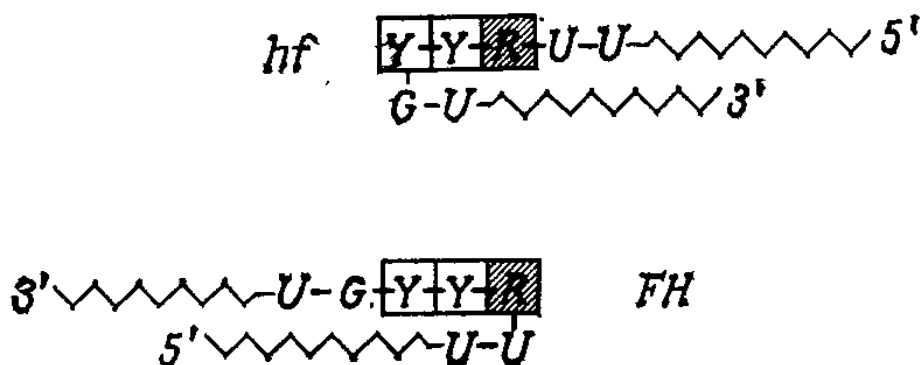


Рис. 48. Две возможные конфигурации антикодоновой петли тРНК (конфигурация FH — по Фуллеру и Ходжсону [61], hf — по Вёзе [60]). Антикодонный паттерн (в рамках) — по модели Крика и др. [3].

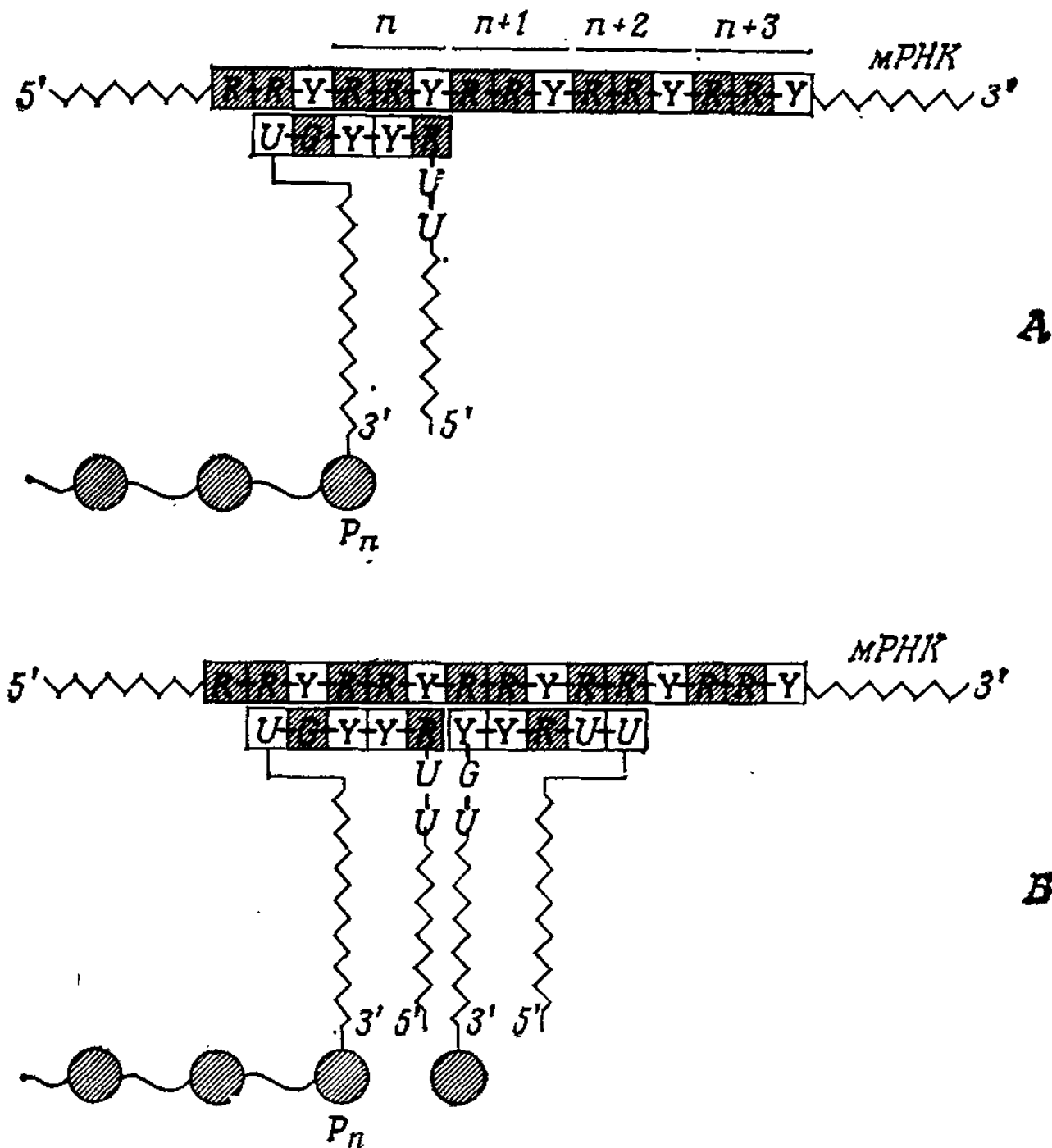
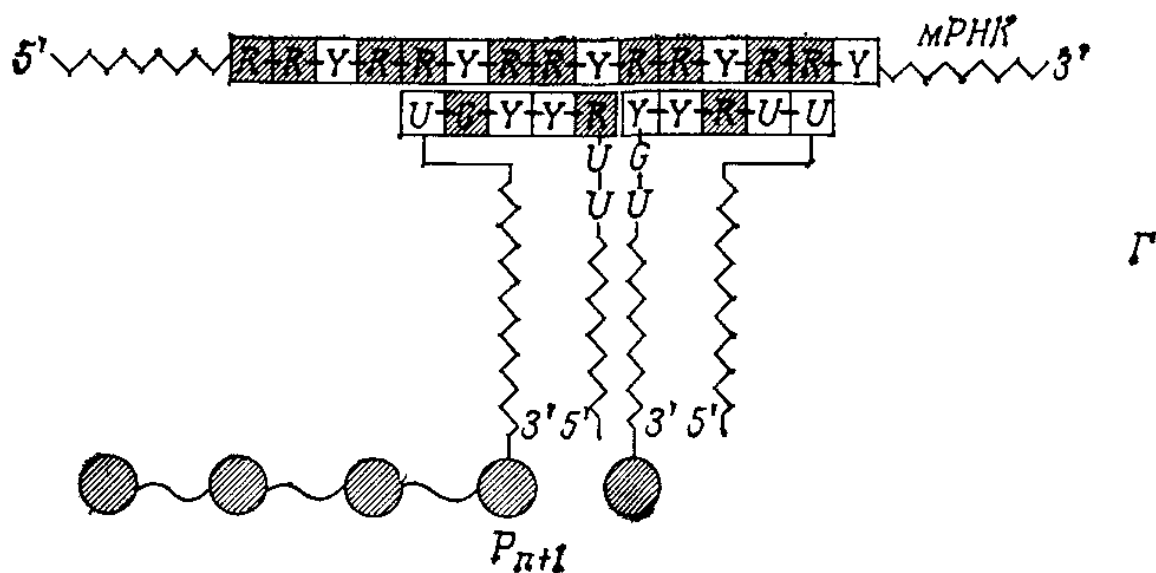
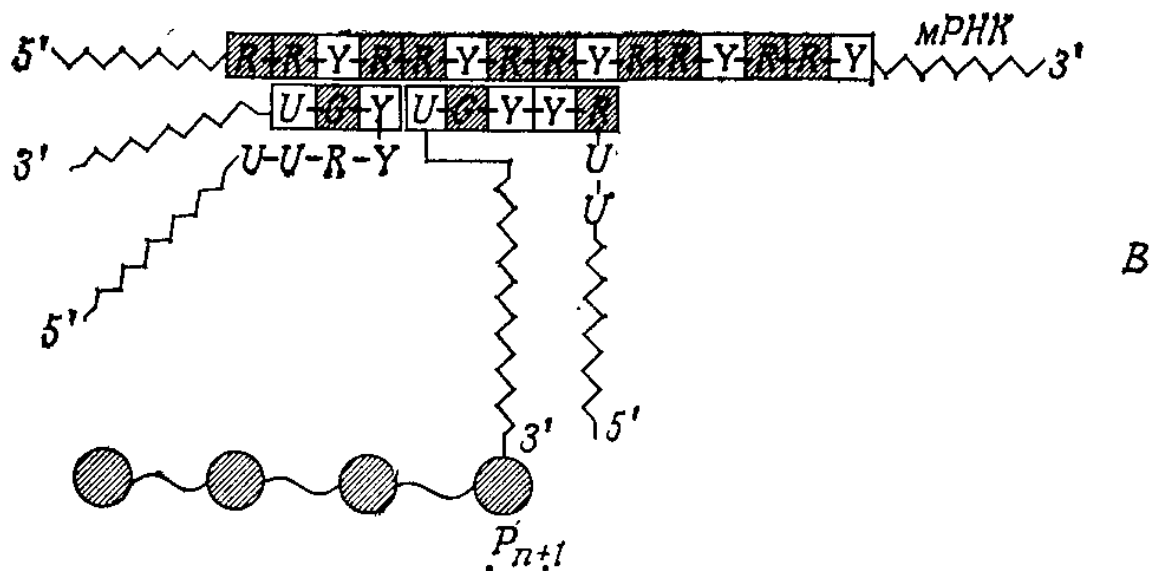


Рис. 49. Прimitивный механизм трансляции требует «липкого» которого растущая полипептидная цепь остается связанной с транспорт осуществляется скачками, при которых изменяется цепь всегда связана с мРНК с помощью пяти пар оснований, соседняя аминоксил-тРНК. Частичное перекрытие при спариванного триплетами оснований.

никакого сложного аппарата типа современных рибосом. Хотя взаимодействие между кодоном и антикодоном стабилизируется пятью парами оснований, существенно, что код в действительности использует триплеты оснований. Ранее уже подчеркивалось [62], что первичный код, использующий не триплеты оснований, не может объяснить возникновение современного кода.



взаимодействия между мРНК и пептидил-тРНК, в результате мРНК, пока не закончится трансляция. По Крику и др. [3], конформация тРНК ($FN \rightleftharpoons hf$). Новообразованная пептидная причем некоторую дополнительную стабилизацию обеспечивает вании оснований гарантирует связное чтение сообщения, зако-

Эта модель хорошо объясняет, как в отсутствие сложного аппарата трансляции может происходить

- спонтанное,
- последовательное,
- полное, т. е. не по частям, и
- воспроизводимое, т. е. при строгом сохранении заданной рамки кодонов,

считывание мРНК.

Такой код неразрывно связан с определенными структурными особенностями антикодоновой петли современных тРНК, и это позволяет допустить, что указанные молекулы являются потомками первых функционально организованных единиц. Четыре аминокислоты, кодируемые по этой модели, таковы:

| | | | |
|----------|--------------------------|----------|-----------|
| GG^U_C | GA^U_C | AG^U_C | AA^U_C |
| глицин | аспарагиновая кислота | серин | аспарагин |

Содержание некоторых из них в первичном бульоне было очень высоким [63].

С другой стороны, модель сталкивается и с некоторыми трудностями. В последовательности RRY имеется большой избыток пуринов и поэтому она не сможет легко складываться на себя с образованием стабильной структуры.

Вследствие этого такие последовательности

- очень легко гидролизуются (если они находятся в виде одиночных цепей),
- имеют повышенную тенденцию к образованию дуплексов (которые не могут легко реплицироваться с помощью примитивных механизмов),
- не имеют внутренней симметрии и
- производят минус-цепи с другим общим кодовым паттерном (т. е. 5' RYY).

ХИ.2. Код RNY

Прежде чем перейти к более детальному обсуждению вопросов, перечисленных выше, рассмотрим альтернативную модель, которая лишена этих конкретных недостатков. Предложение использовать общий кодовый паттерн RNY, где N — любой из четырех нуклеотидов A, U, G, C, тоже приписывается Крику и др. [3]. Однако авторы «разлюбили» свою модель из-за одного ее недостатка: если N является пиримидином, то антикодоновая петля с общей последовательностью 3' \textcircled{U} GYNRU \textcircled{U} в некоторых случаях может использовать лишь свои пять центральных

нуклеотидов для образования стабильных пар оснований с мРНК. Это возражение, однако, снимается на основании того факта, что код RNY может детерминировать восемь аминокислот, вследствие чего некоторые комбинации, не удовлетворяющие требованиям стабильности комплекса мРНК — пептидил-тРНК (см. ниже), можно исключить.

Каковы преимущества общего кодового паттерна RNY? Прежде всего, код RNY, как и его аналог RRY, является кодом без запятых. Более того, он симметричен по отношению к плюс- и минус-цепям. В направлении $5' \rightarrow 3'$ общая структура рамки как для плюс-, так и для минус-цепи будет $R_N^{N'} Y$, где N и N' комплементарны и расположены в цепях зеркально-симметричным образом. Такая же симметрия может возникнуть в пределах одной цепи; тем самым допускается образование симметричных петель вторичной структуры. Современные тРНК — это типичный пример молекул с (почти) симметричными петлями. Одноцепочечные фаговые геномы и их производные (например, миди-вариант РНК Q_{β}) тоже отличаются такими элементами симметрии. Здесь селективное преимущество симметричной структуры очевидно. Если молекулы воспроизводятся с помощью полимеразы, которая специфически узнает какую-то структурную особенность, то лишь при условии симметричной структуры плюс- и минус-цепи могут быть равно эффективными матрицами. Такое равенство эффективностей необходимо для отбора. Итак, симметрия тРНК, возможно, является реликтом той эпохи, когда эти молекулы еще должны были воспроизводиться автономно.

Образование петель увеличивает также устойчивость молекулы к гидролизу и создает предпосылки для правильного считывания. В открытой структуре, где большое число нуклеотидов остается неспаренными (например, в последовательностях типа RRY), репликация и трансляция могут начинаться в любом одноцепочечном участке последовательности, что приводит к образованию фрагментарных продуктов. В полностью двухцепочечной структуре начальными

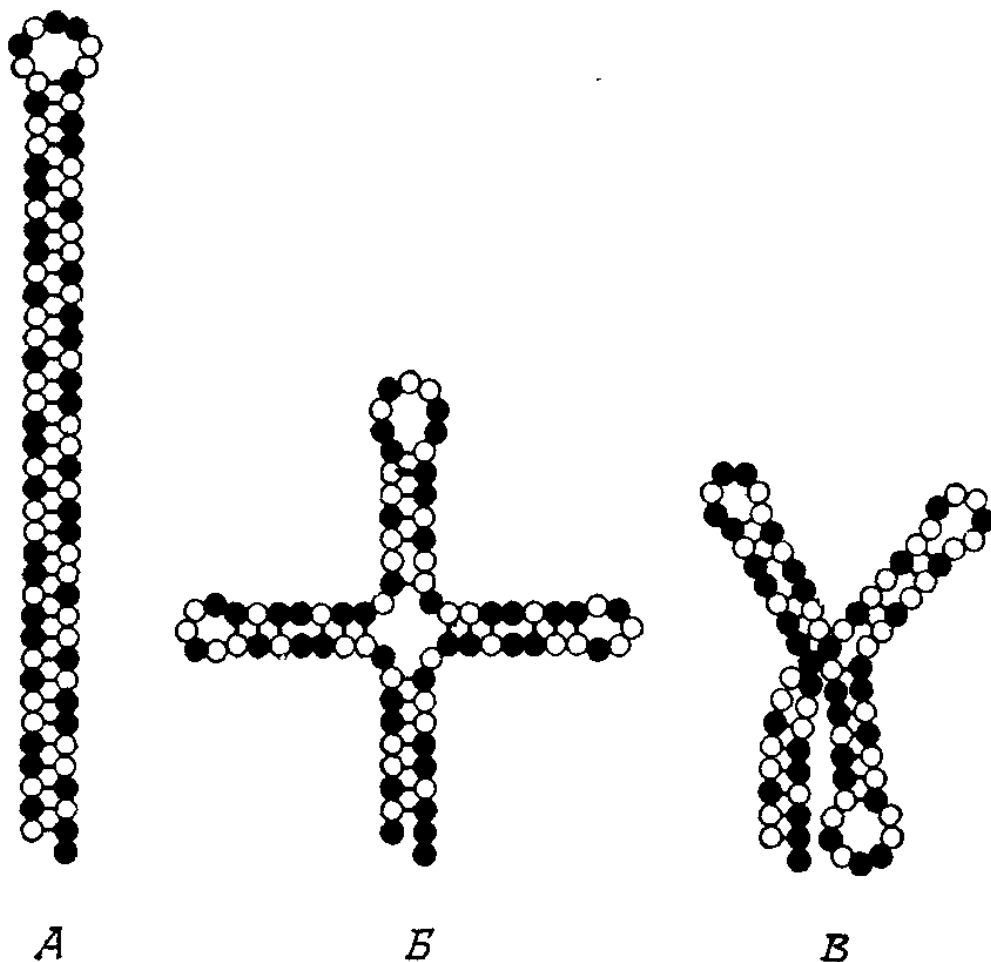


Рис. 50. Для симметричности вторичной структуры РНК необходима соответствующая внутренняя комплементарность (см. рис. 14 в части А) — только тогда плюс- и минус-цепи будут складываться одинаково. Такие структуры могут образовать кодонные паттерны типа RNY. Придумана игра, правила которой учитывают физические взаимодействия, характерные для олигонуклеотидов и тРНК [17, 18]. Эта игра показывает, какие структуры образуются с наибольшей вероятностью. Для образования шпильки необходимо, чтобы обе половинки молекулы были комплементарны (например, последовательность 5'-RRY может соединиться со своей минус-цепью, несущей паттерн 5'-RYY).

точками репликации и трансляции являются только открытые липкие концы. Таким способом может быть прочтена вся мРНК, при этом требуется лишь временное частичное развертывание матрицы, которое может стимулироваться взаимодействиями с растущей цепью. Симметрия давала бы и в этом случае преимущество, хотя ее наличие не обязательно (см. рис. 50).

Чисто логически код RNY кажется более привлекательным, чем код RRY, по трем причинам:

1. Селективное усиление молекул РНК должно быть эффективным как для плюс-, так и для минус-

цепей. Симметричные паттерны RNY в большей степени удовлетворяют этому требованию, чем последовательности RRY, которые отличаются от своих минус-цепей (RY \bar{Y}) и, следовательно, не могут с той же эффективностью узнаваться ферментами.

2. Из-за высокой сложности последовательностей вероятность отыскать именно те немногие из них, которые обладают свойствами, полезными для репликации и трансляции, очень мала. Если эти последовательности, будучи симметричными структурами, удовлетворяют требованиям п. 1, то как плюс-, так и минус-цепи могут быть кандидатами на роль молекул — носителей таких функций.

3. Эволюция аппарата трансляции с его различными тРНК и мРНК требует взаимной стабилизации всех репликативных молекул. Как будет показано ниже, гиперцикл может легче образоваться из квазивида, если этот квазивид вследствие своей симметрии обеспечивает наличие двух комплементарных функций.

Система может накапливать информацию и в конечном счете эволюционировать к большей сложности только в том случае, если она приемлет «дарвиновскую логику» селективной самоорганизации. Однако эта логика должна найти свое обоснование и свое выражение в материальных свойствах. Все, что могут узнавать компоненты в самом начале, — это естественная распространенность и сила взаимодействия. Это те свойства, с которыми мы должны иметь дело, чтобы понять, как возникла трансляция.

XIII. Физика первичного кодирования

XIII.1. Условия на старте

Самоорганизация как мультимолекулярный процесс требует, чтобы концентрация соответствующих мономеров и полимеров была достаточно высока. Таким образом, ее началу должен был предшествовать длительный этап предбиологического синтеза, во время которого накапливались все вещества, необходимые

для создания «очень обогащенного бульона». Мы не намерены задерживаться на этих процессах предбиологической химии; не собираемся мы также спорить и о деталях исторических граничных условий. Такие вопросы, как, например, образовался ли «бульон» в океанах, в водоемах или же в маленьких лужах или какую роль играли здесь поверхности раздела, крупнозернистые или пористые поверхности, могут быть важны только в том случае, если необходимо дать оценку абсолютным скоростям исторических процессов.

Здесь мы просто предположим, что когда началась самоорганизация, высокоэнергетические вещества всех видов имелись в изобилии, в том числе

- аминокислоты, различающиеся по степени своей распространенности,*
- нуклеотиды А, U, G, C,*
- полимеры обоих упомянутых классов, т. е. протеиноиды и тРНК-подобные вещества, имеющие более или менее случайные последовательности.*

«Менее случайные» в данном контексте означает наличие взаимодействий между ближайшими соседями и более сложных взаимодействий при укладке молекулы, которые ведут к преобладанию определенных структур, в то время как «более случайные» относится к тому, что исходно эти последовательности не предназначались для выполнения какой-либо функции: если такая способность и существовала с самого начала, то она могла появиться только случайно.

С другой стороны, мы не предполагаем наличия каких бы то ни было адаптированных белковых аппаратов — таких, как

- специфические полимеразы,*
- адаптированные синтетазы или какие-либо рибосомные структуры.*

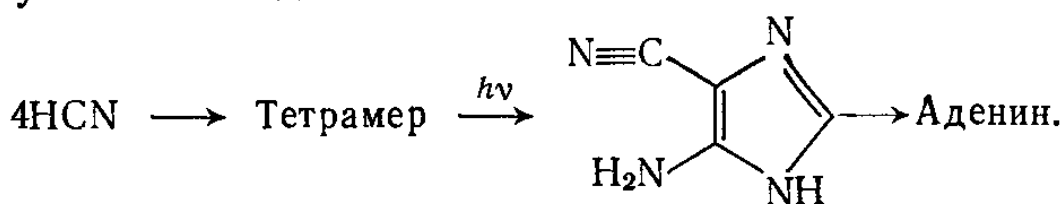
Это не означает, что неинструментированные, плохо адаптированные белковые катализаторы не могли содействовать зарождению репликации и трансляции. Однако эти белки, неспособные воспроизводиться и совершенствоваться посредством отбора, долж-

ны быть включены вместе с другими каталитическими поверхностями в категорию «постоянных факторов среды».

XIII.2. Распространенность нуклеотидов

Поскольку к самовоспроизведению способны лишь структуры, подобные нуклеиновым кислотам, важно более детально проанализировать, во-первых, их распространенности и, во-вторых, их взаимодействия друг с другом.

Нуклеотиды, а еще более их высокоэнергетические олигофосфатные формы образуются с большим трудом (если использовать возможные предбиологические механизмы), чем аминокислоты. Поэтому количественные данные об их относительной распространенности весьма скудны. Миллер и Оргел [63, с. 104] подчеркивают центральную роль адениновых нуклеотидов как в генетических процессах, так и при переносе энергии и связывают это с относительной легкостью их образования. Оро с сотрудниками [64] обнаружили, что в концентрированных водных растворах цианида аммония аденин может быть получен с выходом 0,5%, а Миллер и Оргел [63, с. 105] показали, что даже чистый цианистый водород в ходе реакции, катализируемой солнечным светом, дает важное промежуточное соединение:



Этот интермедиат может реагировать также с цианатом, мочевиной или с цианом с образованием гуанина. Менее известны механизмы синтеза пиримидинов. Удалось установить путь синтеза цитозина, используя сочетание цианата с цианоацетиленом,— последний образуется при электрическом разряде из смеси метана с азотом. Урацил является, по-видимому, продуктом гидролиза цитозина, и возможно, что его существование в первичных условиях было обязано именно этому источнику.

О распространенности пуринов и пиримидинов в первичных условиях можно сказать очень мало. Скорость матричной полимеризации пропорциональна концентрации включаемого мономера. Для комплементарного инструктирования требуются по меньшей мере два типа нуклеотидов, и их содержание в информационных последовательностях должно быть одинаковым. Поэтому включение менее распространенного нуклеотида всегда будет лимитирующим этапом, по крайней мере для элонгации цепи. Следовательно, большой избыток А над U в первичном распределении мономеров, который мог иметь место, лишь очень мало благоприятствовал бы синтезу сополимеров АU-типа по сравнению с сополимерами GC-типа, за исключением тех случаев, когда лимитирующим этапом является нуклеация олиго-А-затравок. Способность к репликативному росту лимитирована матричной функцией менее обильного члена комплементарной пары нуклеотидов. Если в первичных условиях распространенность G и C была промежуточной между распространенностями A и U, то сополимеры, богатые GC и богатые AU, вполне могли бы образовываться со сравнимыми скоростями.

Поэтому мы не можем далее придерживаться прежнего спекулятивного мнения, что первые кодоны выбирались исключительно из бинарного алфавита и состояли лишь из сополимеров AU.

XIII.3. Стабильность комплементарных структур

Более ценные указания, связанные с проблемой первых кодонов, могут быть получены из данных о стабильности пар оснований. Результаты исследования стабильности и скорости спаривания оснований с использованием различных комбинаций нуклеотидов подробно обсуждались в обзорах [4, 44]. Они обосновывают количественно то общепринятое мнение, что GC-пары значительно сильнее стабилизируют кооперативную стопку оснований, чем AU-пары.

Константа стабильности непрерывной и однородной олигомерной последовательности из n нуклеотид-

ных пар дается следующей формулой:

$$K_n = \beta s^n, \quad (91)$$

которая относится к линейной модели Изинга. β — это фактор кооперативности, который как для AU-, так и для GC-пар по порядку величины равен 10^{-3} ¹⁾, а s — константа стабильности отдельной пары в кооперативной стопке. Для гомополимера AU этот параметр примерно на порядок меньше, чем для гомополимера GC, или при грубой количественной оценке:

$$s_{AU} \approx 10, \quad \text{тогда как} \quad s_{GC} \approx 100.$$

Для случая, когда одна из комплементарных цепей может принимать ту конкретную конфигурацию стопки, которая реализуется в антикодоновой петле тРНК, получены бóльшие значения абсолютной стабильности, чем рассчитанные по формуле (91). Вероятно, фактор кооперативности β в этом случае иной. Однако Уленбек, Бэттер и Доти [65, 66] установили, что три- и тетрануклеотиды, комплементарные антикодоновой области тРНК и различающиеся одной парой AU, различаются по своим константам стабильности на один порядок — в полном согласии с приведенной выше оценкой. С этим также согласуется то, что найденные максимальные абсолютные значения констант стабильности относятся к взаимодействию двух тРНК, имеющих комплементарные антикодоны [67].

Данные, полученные для определенных коротких последовательностей, могут быть использованы по крайней мере для сравнения различных моделей репликации и трансляции и для суждения об их относительной значимости. Очевидно, что изолированные AU- или GC-пары нестабильны при любых реальных концентрациях. Для начала репликации необходим некий вспомогательный этап — образование затравки, и именно этот этап в первую очередь требует

¹ Такое соотношение формально верно как для спаривания оснований в пределах данной последовательности, так и для ассоциации двух комплементарных последовательностей, когда β имеет размерность M^{-1} .

участия ферментов. Современные фаговые РНК-репликазы также специфически адаптированы к паттерну последовательности нуклеотидов фагового генома. Элонгация цепи происходит путем кооперативного связывания очередного нуклеотида на вершине стопки пар оснований растущей цепи. Имеющиеся данные указывают, что GC-пара примерно в десять раз более стабильна, чем AU; это приводит к относительно большей точности копирования G и C, чем A и U. Если скорость репликации лимитируется образованием ковалентной связи в полинуклеотидном остове (а не спариванием оснований), то точность может зависеть от концентраций мономеров m_R и m_Y и от констант стабильности пар K_{RY} , K_{RR} и K_{YY} . Тогда точность воспроизведения для любого данного нуклеотида может быть получена из среднего геометрического точностей для обоих комплементарных процессов $R \rightarrow Y$, $Y \rightarrow R$:

$$q_{RY} = \frac{m_Y K_{RY}}{\sum_N m_N K_{RN}} \quad \text{и} \quad q_{YR} = \frac{m_R K_{YR}}{\sum_N m_N K_{YN}}, \quad (92)$$

где $K_{RY} \approx K_{YR}$ и суммирование идет по всем $N = A, U, G, C$. Для A и U или для G и C эти значения q равны, потому что ошибка может произойти как в плюс-, так и в минус-цепи. Если концентрации мономеров одинаковы, то константы стабильности определяют достижимую точность. Из этого следует, что G и C воспроизводятся значительно более точно, чем A и U. Однако отношение темпов ошибок при воспроизведении GC и AU в смешанных системах не совсем совпадает с (обратным) отношением соответствующих констант стабильности, главным образом из-за наличия «качающихся» (wobble) взаимодействий между G и U, которые являются основным источником ошибок воспроизведения — даже при репликации современных РНК-содержащих фагов [34].

Мы оценили значение q на основании различных данных по взаимодействиям между нуклеотидами в отсутствие ферментов. Эти оценки представлены в

табл. 15. Первые три строки в этой таблице относятся к случаю равных концентраций мономеров А, У, G, С. Это допущение может быть весьма далеким от реальности, и поэтому в трех следующих примерах

Таблица 15

Оценки точностей и вероятностей ошибок
для воспроизведения G и C по сравнению с А и U

| Концентрации мономеров | Константы стабильно- сти пар оснований | Точность q | | Вероятность ошибки $1-q$ | |
|---|---|--------------|------|-----------------------------|------|
| | | GC | AU | GC | AU |
| $m_A = m_G = m_C = m_U$ | $K_{RR} = K_{YY} = 1$ $K_{AC} = 1; K_{GU} = 10$ $K_{AU} = 10; K_{GC} = 100$ | 0,93 | 0,59 | 0,07 | 0,41 |
| $m_A = m_G = m_C = m_U$ | $K_{RR} \approx K_{YY} \ll 1$ $K_{AC} = 1; K_{GU} = 10$ $K_{AU} = 10; K_{GC} = 100$ | 0,95 | 0,67 | 0,05 | 0,33 |
| $m_A = m_G = m_C = m_U$ | $K_{RR} = K_{YY} \ll 1$ $K_{AC} = 1; K_{GU} = 5$ $K_{AU} = 10; K_{GC} = 100$ | 0,97 | 0,78 | 0,03 | 0,22 |
| $m_A = 10m_G$ $m_G = m_C$ $m_C = 10m_U$ | $K_{RR} \approx K_{YY} \ll 1$ $K_{AC} = 1; K_{GU} = 5$ $K_{AU} = 10; K_{GC} = 100$ | 0,93 | 0,81 | 0,07 | 0,19 |
| $m_A = 10m_G$ $m_G = m_C$ $m_C = 10m_U$ | $K_{RR} \approx K_{YY} \ll 1$ $K_{AC} = 1; K_{GU} = 5$ $K_{AU} = 10; K_{GC} = 100$ | 0,95 | 0,69 | 0,05 | 0,31 |
| $m_A = 10m_G$ $m_G = m_C$ $m_C = 10m_U$ | $K_{RR} = K_{YY} = 1$ $K_{AC} = 2; K_{GU} = 10$ $K_{AU} = 10; K_{GC} = 100$ | 0,86 | 0,25 | 0,14 | 0,75 |

оно снято. Можно поставить под сомнение правомочность использования данных по стабильности, которые были получены из опытов с олигонуклеотидами. Однако включение одного нуклеотида в процессе репликации сопровождается кооперативными взаимодействиями между парами оснований, и, следовательно,

относительные порядки величин, полученные для олигонуклеотидов, могут быть верны и для этого случая. Для вычисления q требуются лишь *относительные*, а не *абсолютные* значения стабильности.

Из различных оценок, представленных в табл. 15, следует такой вывод: G и C воспроизводятся со значительно большей точностью, чем A и U. В зависимости от превосходства отобранных последовательностей [σ ; см. уравнение (28), часть A] воспроизводимое информационное содержание GC-богатых последовательностей в ранних репликативных процессах ограничивается примерно 20—100 нуклеотидами, т. е. тРНК-подобными молекулами, тогда как для AU-богатых последовательностей оно вряд ли может превосходить 10—20 нуклеотидных остатков на репликативную единицу. Здесь следует подчеркнуть, что, вообще говоря, могли существовать более длинные последовательности *любого* состава. Однако они *не были* воспроизводимыми и поэтому не могли эволюционировать в соответствии с каким-либо функциональным требованием.

Из анализа экспериментальных данных о репликации фага в части A мы заключили, что даже хорошо адаптированные РНК-репликазы не позволяют построить *воспроизводимую* цепь длиной более 1000—10 000 нуклеотидов. Эта оценка соответствует действительному числу генов у РНК-содержащих фагов.

Мы можем теперь сделать окончательные выводы относительно первичных механизмов репликации: при условии *воспроизводимости* такой размер, как вся тРНК, могли иметь только GC-богатые полинуклеотиды. Таким образом, *GC-богатые последовательности могли выступать в роли первичных тРНК (адапторов) и воспроизводимых информационных РНК* — по крайней мере до тех пор, пока репликация не стала катализироваться умеренно адаптированными ферментами.

Такое же заключение можно сделать относительно начала *трансляции*. Как подчеркивали Крик и др. [3], стабильность комплекса пептидил-тРНК—мРНК очень важна для любой модели примитивной транс-

ляции. Исходя из приведенных выше данных константа стабильности комплекса, состоящего из пяти GC-пар, равна

$$K_{5GC} \approx 10^7 \text{ M}^{-1},$$

в то время как для пяти AU-пар она на пять порядков меньше:

$$K_{5AU} \approx 10^2 \text{ M}^{-1}.$$

Эти значения опять необходимо рассматривать как относительные: в действительности они могут быть несколько выше, если допустить наличие стэкинга или предположить, что мы имеем дело с тРНК современного типа; это, однако, не делает неверными рассуждения, основанные на относительных величинах.

Оценки можно сделать также, используя данные по временам жизни. Найдено, что измеренные константы скорости рекомбинации комплементарных олигонуклеотидов всегда имеют порядок величины

$$k_R \approx 10^6 \text{ M}^{-1} \cdot \text{с}^{-1}.$$

Если воспользоваться приведенными выше константами стабильности, времена жизни соответствующих комплексов будут равны

$$\tau_{5GC} \approx 10 \text{ с} \quad \text{и} \quad \tau_{5AU} \approx 10^{-4} \text{ с}.$$

Эти оценки снова могут стать несколько бóльшими по величине, если стабильности окажутся выше и если две соседние тРНК, связанные с информационной РНК, стабилизируют друг друга. Тогда времена жизни оказываются как раз достаточными для того, чтобы GC-богатые последовательности могли породить примитивную трансляцию. Эти времена жизни определенно слишком малы, если преобладают AU-пары. Теперь мы видим, что некоторые недостатки, свойственные коду RNY по сравнению с кодом RRY, обусловленные соотношением стабильностей, могут быть скомпенсированы использованием преимущественно G и C, по крайней мере в части положений R и Y. Четырехчленная GC-структура конечно более ста-

бильна, чем любая пятичленная структура, включающая более чем две пары АУ.

Вывод следующий:

Возникновению трансляции сильно благоприятствуют GC-богатые структуры — как в случае предшественников тРНК, так и для информационных РНК.

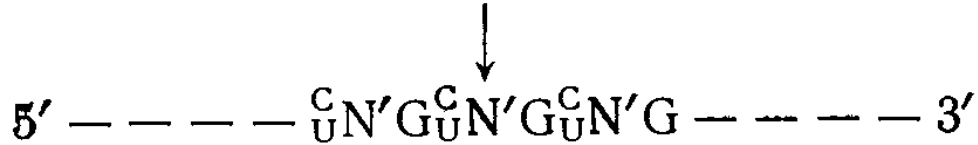
XIV. Код с GC-рамкой

XIV.1. Первые два кодона

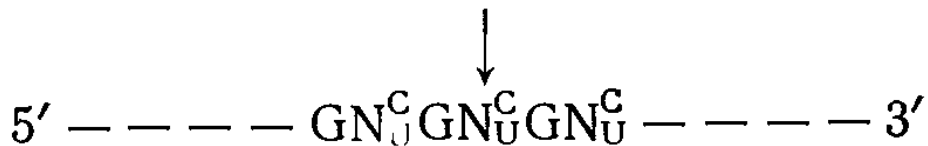
Если выводы из данных по стабильности объединить с рассуждениями Крика и др., то мы можем предсказать, какие кодоны были скорее всего самыми первыми.

Единственные достаточно длинные последовательности, которые могли точно воспроизводиться, — это, по-видимому, те, у которых преобладали остатки G и C. Тогда первыми кодонами были только комбинации этих двух остатков. Требование считывания без запятых исключает симметричные комбинации GGG/CCC и GCG/CGC. В этом легко убедиться, записав такие последовательности. Адапторы с правильными комбинациями в антикодоне могут связываться в различных перекрывающихся положениях. Если далее ввести кодонные комбинации, полученные из симметричных предшественников, то это приведет к еще более разрушительным последствиям. Тогда у нас останутся лишь две комплементарные пары комбинаций, а именно GGC/GCC и CCG/CGG (все комбинации читаются в направлении 5'→3'). С точки зрения симметрии они кажутся совершенно эквивалентными. Имеется, однако, небольшая асимметрия, связанная с «качанием» основания в третьем положении. Сравним последовательности мРНК, состоящие исключительно либо из кодонов CNG, либо из GNC. В первом случае «качающееся» основание G всегда находится в третьем положении кодона, во втором же случае — в первом положении, как в плюс-, так и в минус-цепях (если считывание всегда идет в направлении 5'→3'). Для репликации различные кодонные

положения неразличимы. Следовательно, каждый раз, когда встречается «качающееся» основание, в комплементарную цепь может быть внесена неопределенность, которая при трансляции в первом случае затрагивает первое, а во втором случае — третье положение кодона:



и



Воспроизведенные последовательности транслируются правильно только во втором случае, т. е. когда качания при взаимодействии с адаптором имеют место преимущественно в третьем, а не в первом положении кодона.

Другими словами, адаптор с антикодоном $3' \text{CNG}$ способен узнавать как $5'GN'C$, так и $5'GN'U$, а адаптор с антикодоном $3' \text{GNC}$ — только $5'CN'G$, но не $5'UN'G$. Эта аргументация может оказаться несостоятельной, если для связывания адаптора с мРНК требуется пять пар оснований, потому что тогда «качающиеся» положения могут выявляться не так четко. Тем не менее эта асимметрия в отношении первого и третьего положений кодона в современном генетическом коде существует и является твердо установленной¹. Относительно малых селекционных преимуществ обычно достаточно, чтобы повлиять на ход эволюции. Очевидно, Крик и др. предпочли модель RRY (или RNY) тоже на основании такой аргументации.

Теперь мы можем сделать однозначный выбор первых двух кодонов, а именно:



¹ Наша аргументация подкрепляется тем фактом, что в стационарном распределении G более устойчиво, чем C.

которые будут комплементарными, если их расположить антипараллельно. Этот выбор диктуется четырьмя требованиями:

- *стабильностью* взаимодействия между адаптором и мРНК;
- *точностью* репликации (оба требования подсказывают начинать с комбинации GC);
- считыванием *без запятых* при трансляции, для чего нужен несимметричный GC-паттерн;
- *согласованностью* трансляции, уменьшающей неоднозначность, которая обусловлена «*качаниями*» в третьем положении кодона.

Мы хотели бы подчеркнуть, что эти аргументы основаны исключительно на свойствах нуклеиновых кислот. Приятно отметить, что два кодона, GGC и GCC, в современном генетическом коде соответствуют двум простейшим аминокислотам — глицину и аланину, которые в экспериментах, моделирующих первичные условия, действительно образуются в количествах, далеко превосходящих все другие аминокислоты.

Можно возразить, что продукты трансляции, состоящие только из этих двух остатков, вряд ли могут быть хоть сколько-нибудь эффективными катализаторами. К этому вопросу мы вернемся в разд. XVI. Сейчас достаточно отметить, что трансляция на этой стадии еще не была свойством, необходимым для сохранения соответствующих мРНК. Первые GC-богатые последовательности отбирались исключительно благодаря своей структурной стабильности и способности правильно воспроизводиться. Этим требованиям одинаково хорошо могло удовлетворить много различных GC-богатых последовательностей, и все они могли оказаться отобранными как партнеры (более или менее вырожденные) одного квазивида. Наибольшее предпочтение отдавалось здесь симметричным структурам, потому что они удовлетворяли критериям стабильности одновременно для плюс- и для минус-цепей.

Среди стабильных структур, возникших, вероятно, благодаря матричной функции антикодоновых петель

и последующим дубликациям паттерна, могли оказаться кодирующие последовательности без запятых; тогда-то и зародилась трансляция. Если продукты трансляции дают хоть какое-нибудь преимущество в отношении стабильности или скорости воспроизведения своих информационных РНК, то они будут эволюционировать по дарвиновскому механизму и тем самым непрерывно изменять квазивидовое распределение. Прежде чем вернуться к рассмотрению такой стабилизации продуктами трансляции, задержимся еще немного на проблеме соотношения стабильности структуры и эффективности репликации, поскольку создается впечатление, что эти необходимые свойства основаны на противоречащих друг другу предпосылках.

XIV.2. «Апериодическая линейная GC-решетка»

Молекулу, подобную тРНК, с ее петлями, стабилизируемыми водородными связями, можно рассматривать как микрокристалл. Если она содержит длинные GC-участки, то ее внутренняя структура может быть совершенно инертной. Из кривых плавления петель тРНК или соответствующих олигонуклеотидов мы знаем, что непрерывная последовательность, состоящая только из четырех GC-пар, уже является очень стабильной. Каутс [68] исследовал олигонуклеотид, соответствующий добавочной петле тРНК_{1,2}^{Ser} из дрожжей, содержащий четыре GC-пары (этот олигонуклеотид был получен в результате частичного расщепления молекулы тРНК). Температура плавления его составляла $(84 \pm 1)^\circ\text{C}$, а $\Delta H = 44 \pm 4$ ккал·моль⁻¹. Это соответствует константе стабильности около $2 \cdot 10^5$ при 25° , что находится в хорошем согласии с приведенными выше значениями. Для отбора квазивида существенна не столько структурная стабильность цепей, сколько благоприятное сочетание структурной стабильности с эффективностью воспроизведения. Эффективная матричная функция требует быстрого локального расплетания петли, а этому препятствуют слишком длинные GC-участки. Однако

природные последовательности совсем не являются совершенными.

При высокой относительной концентрации мономеров А и ограниченной точности спаривания оснований GC-микрорешетки всегда будут сильно «загрязнены» остатками А, играющими роль дефектов в линейной GC-решетке. А ргіогі последовательность может быть любой — с низким или с высоким содержанием А, U, G или C. Отбираться, а затем воспроизводимо размножаться будет GC-богатая, но несовершенная последовательность. Если, например, каждое пятое положение в такой последовательности замещено остатками А или U, то спаренные участки в зависимости от внутренней комплементарности будут состоять в среднем не более чем из четырех пар GC (ср. с современными тРНК). Такие структуры легко могут локально плавиться, особенно если процессу репликации способствует белок, который в таком случае представляет собой наиболее примитивную форму репликазы.

Итак, отметим: *AU-дефекты в аperiodической GC-решетке дают селективное преимущество.*

Как сказал Томас Манн: «Жизнь избегает абсолютного совершенства»¹.

XIV.3. От GNC к RNY

Если в GC-богатых цепях отобранного квазивида имеется определенная концентрация дефектов А (и комплементарных дефектов U), то следующий этап в эволюции кода, по-видимому, можно считать предопределенным. Мутации могут возникнуть в любом из трех положений кодона, но их последствия будут совершенно различны. Замещение среднего основания в кодоне вызовет комплементарное замещение среднего основания в соответствующем антикодоне минус-цепи, и, следовательно, немедленно возникнут два новых кодона — GAC и GUC. Изменения в первом или третьем положении, с другой стороны, будут

¹ Th. Mann, *Der Zauberberg* (Т. Манн, Волшебная гора).

дополнены соответствующими изменениями в третьем (или соответственно в первом) положении минус-цепи и из-за эффекта «качания» приведут в конце концов к появлению только еще одного кодона. Кроме того, будет нарушена GC-рамка для считывания без запятых.

Требования стабильности сначала не допускают подстановки более чем одной AU-пары в участок, состоящий из пяти пар оснований, комплекса мРНК—тРНК. Следовательно, наиболее вероятно, что следующими возникнут кодоны 5'GAC и 5'GUC. Поскольку они являются мутантами пары 5'GGC/5'GCC, они могут присутствовать в достаточных количествах, входя в состав отобранного GC-богатого квазивида.

Однако, если на эти мутанты возлагается какая-либо функция при трансляции, они должны стать полностью эквивалентными доминирующему виду 5'GGC/5'GCC. Именно на этой стадии гиперциклическая стабилизация четырех кодонных адапторов (и информационных РНК, кодирующих факторы связи) становится абсолютно необходимым требованием. Без такого сочленения различные компоненты первичной системы трансляции хотя и могут сосуществовать некоторое время, но никогда не будут согласованно эволюционировать или оптимизировать свою кооперацию.

Четыре кодона способны кодировать четыре разные аминокислоты, которые создают основу для богатой палитры функций, так что образовавшиеся в результате белки могут стать эффективными факторами связи. Информационные РНК и тРНК как члены одного квазивида могли возникнуть из комплементарных цепей одного и того же вида РНК, выполняя, таким образом, обе функции.

С другой стороны, это могло существенно ограничить их дальнейшую эволюцию. Следовательно, мы должны допустить, что они возникли из общего предшественника, но дивергировали в различные последовательности из-за своих совершенно разных структурных и функциональных требований.

Согласно таблице современного генетического кода, GAC и GUC кодируют аспарагиновую кислоту и валин. Прежде чем обсуждать аминокислотный аспект более детально, рассмотрим вкратце несколько дальнейших этапов в эволюции на пути к более общему коду.

Требование высокой стабильности взаимодействия кодон — антикодон становится менее жестким по мере повышения адаптации продуктов трансляции. Наконец, допускаются взаимодействия с «качаниями», и код с GC-рамкой может эволюционировать к более общему коду с RY-рамкой. Все это вместе вводит в действие еще четыре аминокислоты. Первое замещение происходит еще тогда, когда остается в силе требование стабильности, которое уменьшает до минимума содержание AU-пар. Тем самым вводятся два кодона: 5'AGC (= серин) и 5'ACC (= треонин). В комплементарных им последовательностях изменяется третье положение кодона, в результате чего появляются 5'GCU и 5'GGU — вторые кодоны для аланина (GCC) и глицина (GGC). Вырождение, связанное с наличием взаимодействий с «качаниями» при воспроизведении этих последних кодонов, могло быть основной причиной появления кодонов AGC и ACC и детерминирования соответствующих аминокислот.

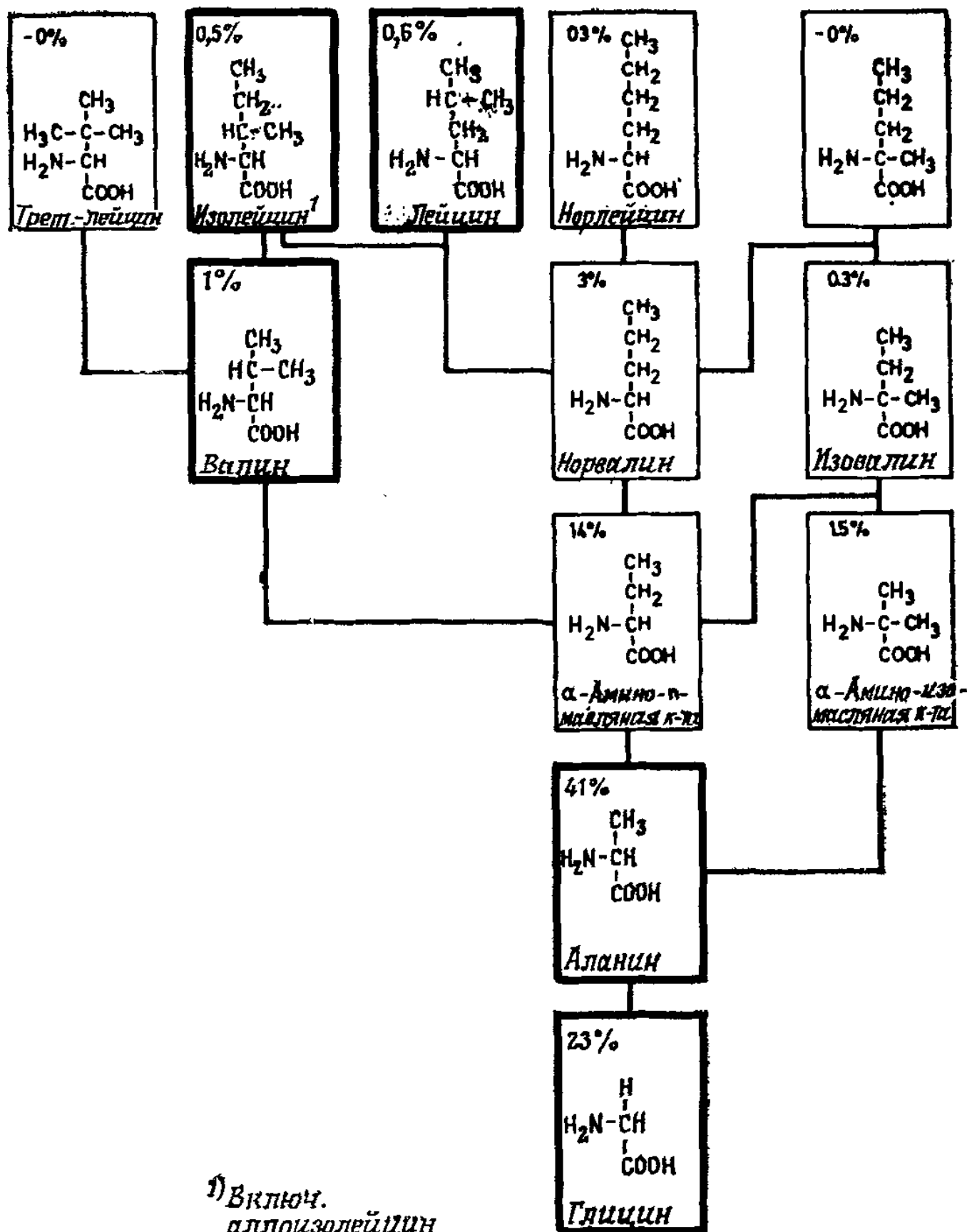
Появление в области кодона при эволюции ферментативного аппарата более одной пары AU привело к кодированию еще двух аминокислот: AA_C^U (аспарагина) и AU_C^U (изолейцина). Это исчерпало все возможности кодирования для кода RNY. Чтобы эволюция генетического кода могла продолжаться, необходимо было отказаться от ограничения непрерывных рамок. Теперь стала необходимой адаптация рибосомных предшественников.

XIV.4. Первичный алфавит аминокислот

Первичные распространенности различных аминокислот могут быть оценены очень надежно. Основные указания, необходимые для оценки вероятностей син-

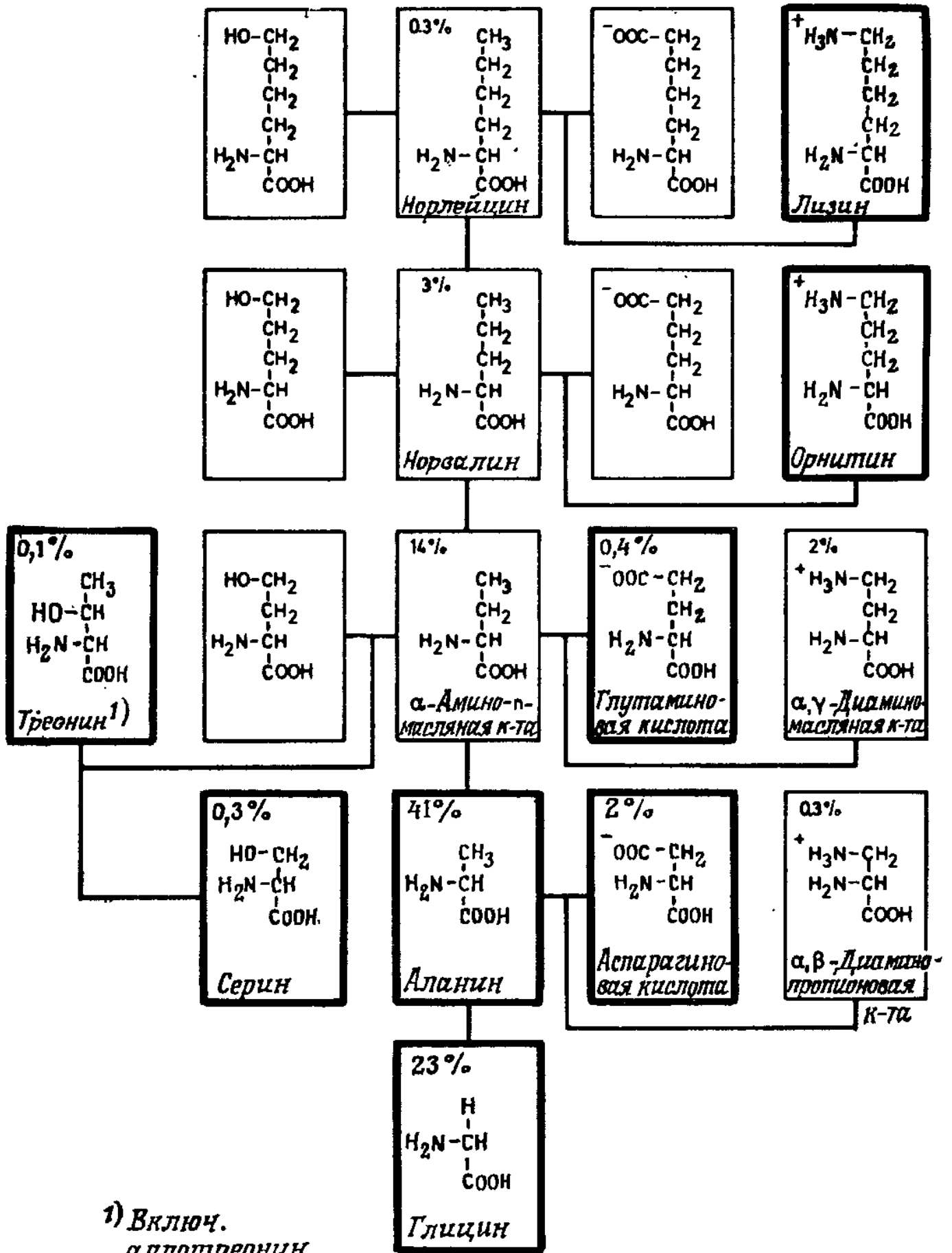
теза этих соединений в первичных условиях, дают уже их структура и состав. На рис. 51 приведена «родословная» первых 12 неполярных алифатических аминокислот, а также несколько ветвей, показывающих родственные отношения для простейших полярных боковых цепей. При рассмотрении этой схемы возникают интересные вопросы, касающиеся выбора Природой белкового алфавита.

Две простейшие аминокислоты — глицин и аланин — являются «природными». Видимо, требования гидрофобного взаимодействия легче было выполнить, добавив некоторые высшие гомологи — такие, как валин, лейцин и изолейцин. Возможно, этот выбор был случайным, а может быть, на него повлияли дискриминирующие взаимодействия с имеющимися адапторами. Среди полярных соединений мы обнаруживаем некоторые алифатические карбоновые кислоты (аспарагиновая и глутаминовая кислота), а также спирты (серин и треонин), но не соответствующие амины (α , β -диаминопропионовая кислота и α , γ -диаминомасляная кислота). Из двадцати «природных» аминокислот появляется только второй следующий гомолог (лизин), в то время как интермедиат (орнитин) представлен лишь в следовых количествах. Это может объясняться тем, что после активации второй аминогруппы происходит образование лактама или элиминация, которая обрывает полимеризацию. Более того, наличие второй аминогруппы может привести к разветвлению полипептидной цепи (хотя такой же аргумент можно привести и в отношении карбоксильных групп). Для первых функциональных компонентов наличие положительно заряженных боковых цепей вполне могло быть необязательным. Даже в воде современных морей концентрация Mg^{2+} достаточно высока (~ 50 мМ), чтобы вызвать заметное комплексообразование с карбоксильными группами. При восстановительных условиях в океанах могло быть растворено еще больше двухвалентных ионов (например, Fe^{2+}). Такие ионы металлов, связанные с карбоксильными группами и все-таки сохранившие свободные координирующие сайты, особенно



А

Рис. 51. Генеалогия первых алифатических аминокислот и не-Числа в левом верхнем углу каждого прямоугольника — это вычисленные условия по Миллеру [63] (т. е. молярный выход аминокислот, перечисленных в табл. 7—2, с. 87 упомянутой работы).



1) Включ. аллотреонин

Б

которые ветви для простейших полярных боковых цепей. относительные выходы соответствующих соединений в перданной аминокислоте, деленный на сумму выходов всех работы). Прямоугольники, соответствующие природным ами-

важны для облегчения взаимодействий между ранними белками и (отрицательно заряженными) полинуклеотидами. С этой точки зрения боковые цепи, содержащие отрицательно заряженные лиганды, кажутся более необходимыми, чем боковые цепи с положительными зарядами.

«Природные» аминокислоты, не представленные на рис. 51, имеют значительно более сложные боковые цепи, и поэтому их концентрации в первичном бульоне были сравнительно низки.

Предположения, основанные на данных о структуре и составе аминокислот, великолепно согласуются с результатами экспериментов, имитирующих предбиологический синтез указанных соединений (Миллер и др.; см. обзор в [63]). Выходы, полученные для природных аминокислот (как впрочем и для других ветвей генеалогического древа), примерно соответствуют величинам, ожидаемым из чисто химических данных (см. численные значения на рис. 51); в ходе этих экспериментов было выявлено также много других интересных деталей. Более того, эти результаты находятся в хорошем согласии с данными, полученными из анализа метеоритов [69, 70], которые отражают встречаемость аминокислот в межзвездном пространстве. Табл. 16 содержит сводку данных (взятых из работы [63]), имеющих отношение к предмету нашей дискуссии.

Первичный бульон без сомнения был очень богат глицином и аланином. В опытах Миллера эти аминокислоты образуются примерно в двадцатикратном избытке по сравнению с любой другой природной аминокислотой. Следующие два места на шкале распространенностей природных аминокислот занимают аспарагиновая кислота и валин, причем между ними и лейцином, глутаминовой кислотой, серином, изолейцином, треонином и пролином имеется большой разрыв.

Есть все основания предполагать, что соответствия между кодонами и аминокислотами в действительности возникали в порядке уменьшения распро-

страненностей последних. Если глицин и аланин по своим распространениям далеко превосходят все другие аминокислоты, то они вполне могли быть за-

Таблица 16

Выход природных аминокислот при имитации предбиологического синтеза и содержание их в метеорите Мерчисон. В первом столбце перечислены аминокислоты содержащиеся в белках, Во втором столбце приведены типичные результаты экспериментов Миллера (из обзора [63]). Эти данные были получены, когда через метан (336 мМ) в присутствии азота и воды пропускали искровой разряд. Суммарный выход аминокислот (включая те, которые не содержатся в белках) по углероду был равен 1,9%, соответствующие выходы глицина и аланина — 0,26 и 0,71%. Аналогичные результаты были получены и в других условиях, причем в иаибольшем количестве из природных аминокислот присутствовали обычно Gly, Ala, Asp и Val. Данные по содержанию аминокислот в метеорите получены Оро и др. [70] и Квенволденем и др. [69]. Во всех случаях D-изомеры составляли почти точно 50%. Другие ссылки можно найти в работе [63].

| Соединение | Выход, мкМ | мкг на 1 г метеорита |
|-----------------------|------------|----------------------|
| Глицин | 440 | 6 |
| Аланин | 790 | 3 |
| Аспарагиновая кислота | 34 | 2 |
| Валин | 19,5 | 2 |
| Лейцин | 11,3 | |
| Глутаминовая кислота | 7,7 | 3 |
| Серин | 5,0 | |
| Изолейцин | 4,8 | |
| Треонин ¹ | 1,6 | |
| Пролин | 1,5 | 1 |

¹ Включая аллотреонин.

кодированы раньше всех, сразу же после того, как сформировались химические механизмы активации.

В таком случае первыми предбиологическими полипептидами — как инструктированными, так и неинструктированными — были, вероятно, в основном, сополимеры Gly—Ala со случайными замещениями

другими аминокислотами, среди которых, по-видимому, встречались и такие, которые в конечном счете так и не стали кодироваться.

Соответствие между распространенностями природных аминокислот и очередностью их сопоставления первым четырем кодоном поразительно. Следует подчеркнуть, что наш выбор кодонов основан исключительно на аргументах, относящихся к структуре нуклеиновых кислот. Не только первые четыре кодона с GC-рамкой точно отвечают распространенностям соответствующих аминокислот, но и четыре следующих RNY-кодона сопоставляются именно с теми аминокислотами, которые — за исключением аспарагина — представлены в таблице Миллера со значительными выходами. Однако возникает вопрос — является ли соответствие AA_C^U аспарагину первичным или же (как подсказывает близкое сходство с кодоном лизина) этот кодон сначала отвечал одной из низших гомологичных диаминокислот, которые в таблице Миллера имеют обращенный порядок выходов по сравнению с аспарагиновой кислотой (α , γ -диаминомасляная кислота) и глутаминовой кислотой (α , β -диаминопропионовая кислота). Следует признать, что в отсутствие дополнительных данных все предположения по этому поводу могут носить только спекулятивный характер.

В этой связи важны также некоторые новые результаты, полученные из анализа аминокислотных последовательностей нуклеотид-связывающих ферментов, которые, по-видимому, существовали еще более $3 \cdot 10^9$ лет назад в доклеточных условиях [71, 72]. Эти данные позволяют предположить наличие последовательности-предшественника нуклеотид-связывающей поверхности, которая включала аминокислоты валин, аспарагиновую (и глутаминовую) кислоты, аланин и глицин, а кроме того, изолейцин, лизин и треонин (хотя эти данные в действительности относятся к более поздней стадии доклеточной эволюции, нежели та, которая обсуждается в этой работе).

XV. Гиперциклическая организация раннего аппарата трансляции

Любая модель эволюции раннего кода и аппарата трансляции должна удовлетворять условиям, в которых тРНК-подобные адапторы и предшественники генов (или информационных РНК) для различных ферментативных факторов могли бы не только сосуществовать, но и согласованно развиваться и эволюционировать к оптимальному функционированию. В частях А и Б было показано, что такая самоорганизация требует наличия циклически замкнутых реакционных связей между всеми партнерами, если только они не могут структурно интегрироваться в одну репликативную единицу. В этом разделе мы попытаемся показать, как реалистические модели кода связаны с гиперциклической организацией и как такие системы могут эволюционировать. Очевидная проблема, связанная с гиперциклом,— это проблема его возникновения. Предпосылкой гиперциклической организации является, по-видимому, наличие в достаточном количестве всех компонентов гиперцикла. Говоря более научным языком, для возникновения гиперцикла — сети реакций высокого порядка — требуется его «нуклеация» по какому-нибудь механизму высокого порядка.

Рассмотрим для сравнения простую репликативную единицу, численность которой растет по автокаталитическому закону первого порядка. В растворе, содержащем высокоэнергетический строительный материал в определенной концентрации, для начала процесса размножения нужна лишь одна копия. Такие эксперименты проводились с фаговой РНК или с ее неинфекционными вариантами [7, 8, 32, 34, 73]. Одной матричной цепи достаточно, чтобы в течение нескольких минут произвести большую популяцию идентичных копий (см. часть А).

Гиперцикл никогда не мог бы зародиться таким способом. Отдельная матричная копия не станет размножаться, пока в наличии не будет достаточно

большого числа ее специфических каталитических партнеров. Последние в свою очередь кодируются матрицами, которые сами по себе не могли бы размножаться без помощи своих продуктов трансляции. Размножение всех матриц в системе зависит от каталитической поддержки, но число катализаторов не сможет расти, если не размножаются матрицы. Насколько велика вероятность того, что нуклеация произойдет посредством какой-нибудь случайной флуктуации? Допустим, что имеется пробирка с 1 мл раствора. Константа скорости реакции между макромолекулами, лимитируемая диффузией, имеет порядок $10^8 \text{M}^{-1} \cdot \text{с}^{-1}$. Следовательно, в растворе должно содержаться по меньшей мере 10^8 идентичных копий партнера данной каталитической реакции, чтобы началось размножение матрицы с $\tau_{1/2}$, равным примерно 1 сут. Очень маловероятно, чтобы согласованное функционирование нескольких партнеров могло быть результатом случайного совпадения таких гигантских флуктуаций. Конечно, не исключено, что процессы размножения различных матриц имеют смешанный порядок, т. е. автокаталитические члены первого порядка (в отсутствие ферментов — это размножение матрицы без каталитического содействия других компонент) накладываются на члены катализируемой репликации второго порядка. Тогда гиперциклическая связь станет эффективной лишь после того, как концентрации вырастут до достаточно высокого уровня. Однако система не может знать заранее, которые из многих альтернативных последовательностей, размножающихся по закону ав. катализа первого порядка, будут давать полезную информацию для катализаторов, необходимых на более поздних этапах организации.

Существует только один путь решения этой проблемы:

Гиперциклу должна предшествовать система, имеющая высокую природную распространенность, из которой он возникает постепенно, путем мутаций и отбора.

Таким предшественником может быть квазивид, представленный распределением GC-богатых последовательностей. Численность всех членов устойчивого квазивида будет расти, пока их концентрации не станут достаточно велики. Как было показано в разд. XIV, некоторые GC-богатые последовательности могут дать начало трансляции путем установления соответствия между аминокислотами и определенными антикодонами. На этом этапе продукты трансляции на самом деле еще не являются необходимыми для сохранения системы, так что трансляцию пока можно считать игрой проб и ошибок. Если, однако, один из продуктов трансляции создаст преимущество для воспроизведения своей собственной мРНК, то эта мРНК может стать доминирующей в квазивидовом распределении.

Отдельный вид РНК может создать алфавит в лучшем случае из двух аминокислот, если как плюс-, так и минус-цепи действуют как адапторы для двух комплементарных кодонов (например, GGC и GCC). Если адапторные последовательности достаточно обильны, то с некой конечной вероятностью сосуществующие мутанты будут сопоставлять аминокислоты двум или даже четырем кодоном (включая GAC и GUC для аспарагиновой кислоты и валина), причем, по-видимому, снова используются как плюс-, так и минус-цепи. Все это может произойти еще на стадии квазивида.

Такая система, однако, может эволюционировать только в том случае, если различные виды РНК будут стабилизировать друг друга с помощью своих продуктов трансляции. Обсуждение деталей кодирования, например вопроса о том, могут ли плюс- и минус-цепи данного вида РНК эволюционировать одновременно и тем самым становиться двумя адапторами для комплементарных кодонов или же плюс-цепь в качестве мРНК кодирует фактор связи, в то время как в качестве адаптора действует только минус-цепь, мы отложим до разд. XVI. Здесь же мы исследуем такую проблему: как из квазивидового

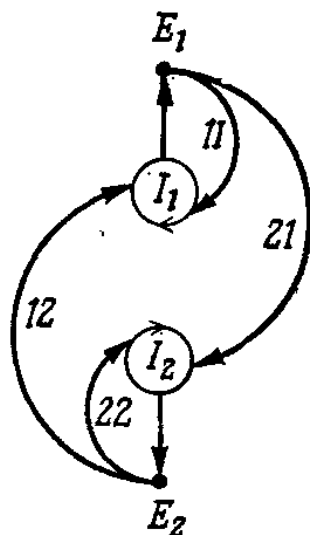


Рис. 52. Два мутантных гена I_1 и I_2 , кодирующих свои собственные репликазы E_1 и E_2 , могут создавать равноценные связи, приводящие к самоусилению [11, 22] и взаимному усилению [21, 12], благодаря своему близкому родству. Аналогичное поведение может быть свойственно репликазам современных РНК-содержащих фагов.

распределения может постепенно сформироваться гиперциклическая организация.

Рис. 52 иллюстрирует возможный путь решения этой проблемы. Допустим, что имеются два мутанта квазивида, присутствующих в достаточно высокой концентрации, плюс- и минус-цепи которых способны действовать как адапторы для (максимум) двух пар аминокислот (например, Gly/Ala и Asp/Val) и которые в то же время могут транслироваться в белок, состоящий из (максимум) четырех видов аминокислот. Если продукты трансляции обладают какой-либо каталитической функцией, благоприятствующей воспроизведению своих мРНК, то мы, вероятно, получим одну из ситуаций, представленных на рис. 52 или 53.

Обе информационные РНК, будучи близкородственными мутантами, кодируют два белка с близкими функциями. Если один из них — специфическая репликаза, то другой будет тоже репликазой, причем каждая из активностей будет усиливать как саму себя, так и вторую активность. Однако может иметь место и специфичность, потому что оба белка не обязательно должны узнавать обе последовательности одинаково хорошо, неродственные же последовательности они вообще не будут узнавать. Две «свои» последовательности имеют специфический сайт связыва-

ния для инициирования репликации. Различия в силе связывания для четырех возможных взаимодействий E_1 и E_2 с I_1 и I_2 могут быть небольшими. Однако эти различия (отмеченные на рисунках разной толщиной линий), какими бы малыми они ни были, будут иметь радикальные последствия, как видно из рассмотрения соответствующих диаграмм особых точек (табл. 17). Мы можем выделить четыре случая:

1. E_1 более благоприятствует I_1 , чем I_2 , а E_2 более благоприятствует I_2 , чем I_1 (рис. 53, А). *Следствие:* и I_1 , и I_2 гиперциклически усиливаются соответствующими ферментами, что ведет к сильной конкуренции. Выжить может только один из конкурентов, даже если они будут селекционно эквивалентны.

2. Как E_1 , так и E_2 более благоприятствуют I_1 , чем I_2 (рис. 53, Б).

Следствие: I_1 выходит победителем, а I_2 вымирает.

3. Как E_1 , так и E_2 более благоприятствуют I_2 , чем I_1 (рис. 53, В).

Следствие: теперь победителем является I_2 , а I_1 вымирает.

4. E_1 более благоприятствует I_2 , чем I_1 , E_2 более благоприятствует I_1 , чем I_2 (рис. 53, Г).

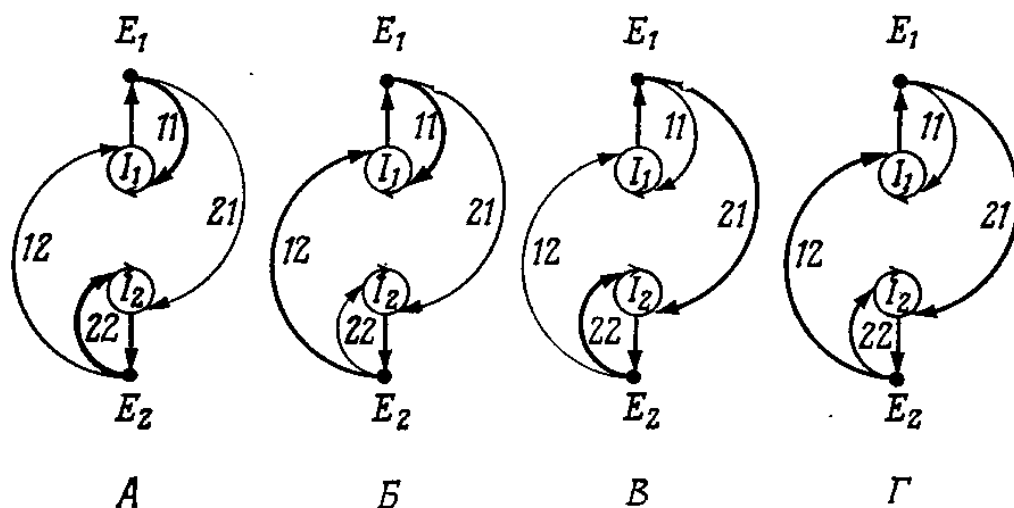


Рис. 53. Эволюционный принцип для гиперциклов иллюстрируется четырьмя возможными ситуациями, которые возникают благодаря связям между двумя мутантами, представленными на рис. 52. Толстые линии указывают, какие связи являются предпочтительными (каким бы малым ни было это предпочтение). Для устойчивости двучленного гиперцикла требуется, чтобы преобладало взаимное усиление (Г).

Исследование особых точек двучленного гиперцикла, представленного на рис. 52, проводилось с использованием упрощенных кинетических уравнений:

$$\dot{x}_i = \sum_{k=1,2} k_{ik} x_i x_k - \frac{x_i}{c} \sum_{l=1,2} \sum_{m=1,2} k_{lm} x_l x_m,$$

$$i = 1, 2; \quad x_1 + x_2 = c.$$

Эти уравнения дают три особые точки со следующими собственными значениями:

$$\bar{x}_1 = (c, 0); \quad \omega^{(1)} = (k_{21} - k_{11}) c,$$

$$\bar{x}_2 = (0, c); \quad \omega^{(2)} = (k_{12} - k_{22}) c,$$

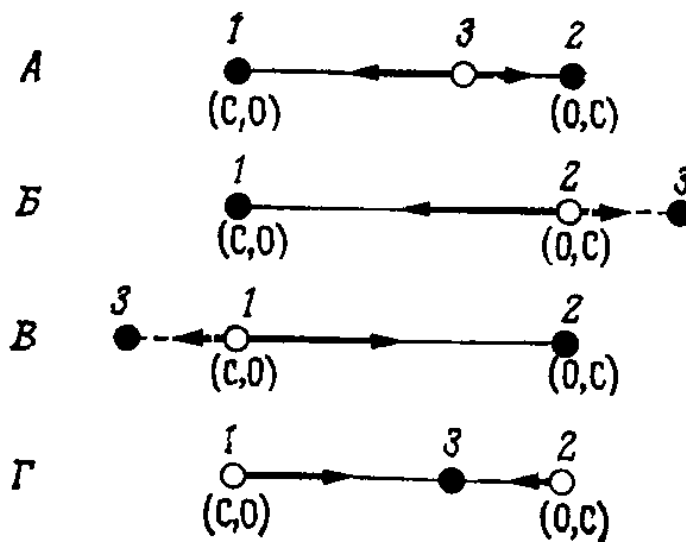
$$\bar{x}_3 = (k_{22} - k_{12}, k_{11} - k_{21}) \frac{c}{k_{11} - k_{21} + k_{22} - k_{12}};$$

$$\omega^{(3)} = \frac{(k_{11} - k_{21})(k_{22} - k_{12})}{k_{11} - k_{21} + k_{22} - k_{12}} c.$$

Можно выделить четыре случая:

- а) $k_{11} > k_{21}$; $k_{22} > k_{12}$ — конкуренция между I_1 и I_2 ;
- б) $k_{11} > k_{21}$; $k_{22} < k_{12}$ — отбор I_1 ;
- в) $k_{11} < k_{21}$; $k_{22} > k_{12}$ — отбор I_2 ;
- г) $k_{11} < k_{21}$; $k_{22} < k_{12}$ — гиперциклическая стабилизация I_1 и I_2 .

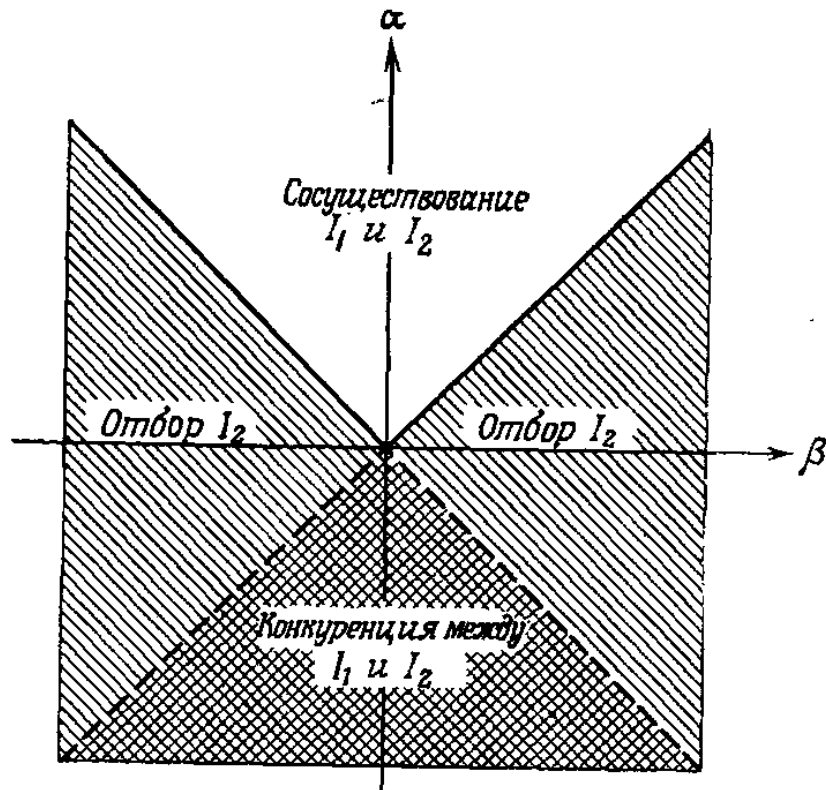
Диаграммы особых точек для этих четырех случаев (см. также часть Б):



Продолжение табл. 17

Единое представление можно получить, если ввести координаты

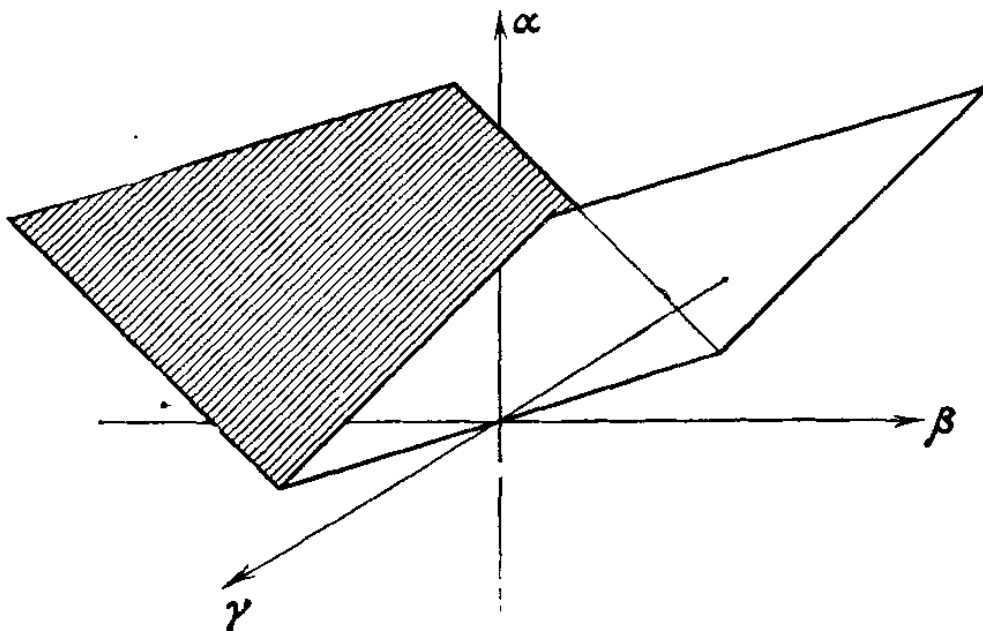
$$\alpha = k_{12} + k_{21} - k_{11} - k_{22} \quad \text{и} \quad \beta = k_{12} - k_{21} + k_{11} - k_{22}.$$



Если в кинетические уравнения кроме члена второго порядка ввести линейный автокаталитический член (при этом функция роста принимает следующий вид: $\Gamma_i = k_i x_i + \sum_{j=1,2} k_{ij} x_i x_j$), то область устойчивого гиперциклического сосуществования двух видов I_1 и I_2 будет представлять собой полупространство над согнутым листом, координаты которого в трехмерном пространстве параметров будут следующие:

$$\alpha = k_{12} + k_{21} - k_{11} - k_{22}, \quad \gamma = \frac{2}{c} (k_1 - k_2).$$

$$\beta = k_{12} - k_{21} + k_{11} - k_{22},$$



Следствие: имеет место взаимная гиперциклическая стабилизация I_1 и I_2 .

Важно отметить, что для такого поведения достаточно малых различий в силе связывания. Поэтому интересно знать, что будет в том случае, когда E_1 и E_2 ведут себя совершенно одинаково в отношении I_1 и I_2 . Здесь мы имеем полную беспристрастность, независимо от того, насколько различаются численности популяций x_i или y_i . I_1 или I_2 могут вымереть вследствие флуктуационной катастрофы, поскольку теперь нет взаимной стабилизации, как это было в случае 4. С другой стороны, флуктуации не усиливаются, и если численности популяций достаточно велики, то флуктуационная катастрофа практически никогда не произойдет. Совсем другим в этом отношении оказывается случай 1. Здесь система будет находиться в динамически сбалансированном состоянии только при условии строго одинаковых численностей популяций I_1 , I_2 , E_1 , E_2 . Небольшая флуктуация может нарушить равновесие и затем из-за самоусиления неизбежно приведет к отбору одного из двух видов. То же самое справедливо для любого ансамбля, в котором каждая мРНК благоприятствует только своей собственной репликазе (см. рис. 47). Чтобы навязать сосуществование, недостаточно связи, происходящей от общей функции трансляции,— все репликазы могут функционировать одновременно в качестве активирующих ферментов (используя свои сайты узнавания РНК). Как и в системе, показанной на рис. 45, здесь будет только один выживший, после чего трансляционный аппарат разрушится.

Точные условия образования гиперцикла выведены в табл. 17. Рисунки дают ясное представление об области устойчивости в пространстве обобщенных координат, зависящих от кинетических параметров.

Итак, мы получили эволюционный принцип для гиперциклов. Этот вид организации может возникнуть из отдельного квазивидового распределения, как только разовьются средства реакционной связи. Предпосылки сосуществования предшественников могут выполняться, как правило, только для близко-

родственных мутантов. Таким образом, возникновение гиперциклов требует предсуществования молекулярной дарвиновской системы, но ведет к совершенно новым последствиям. Этот эволюционный принцип работает даже при очень малых различиях в кинетических параметрах, и поэтому характеризуется чувствительностью к малым изменениям, вызванным мутациями. Если имеется квазивидовое распределение с эволюционирующими взаимодействиями между компонентами, то независимо от того, насколько слабы эти взаимодействия, гиперциклическая

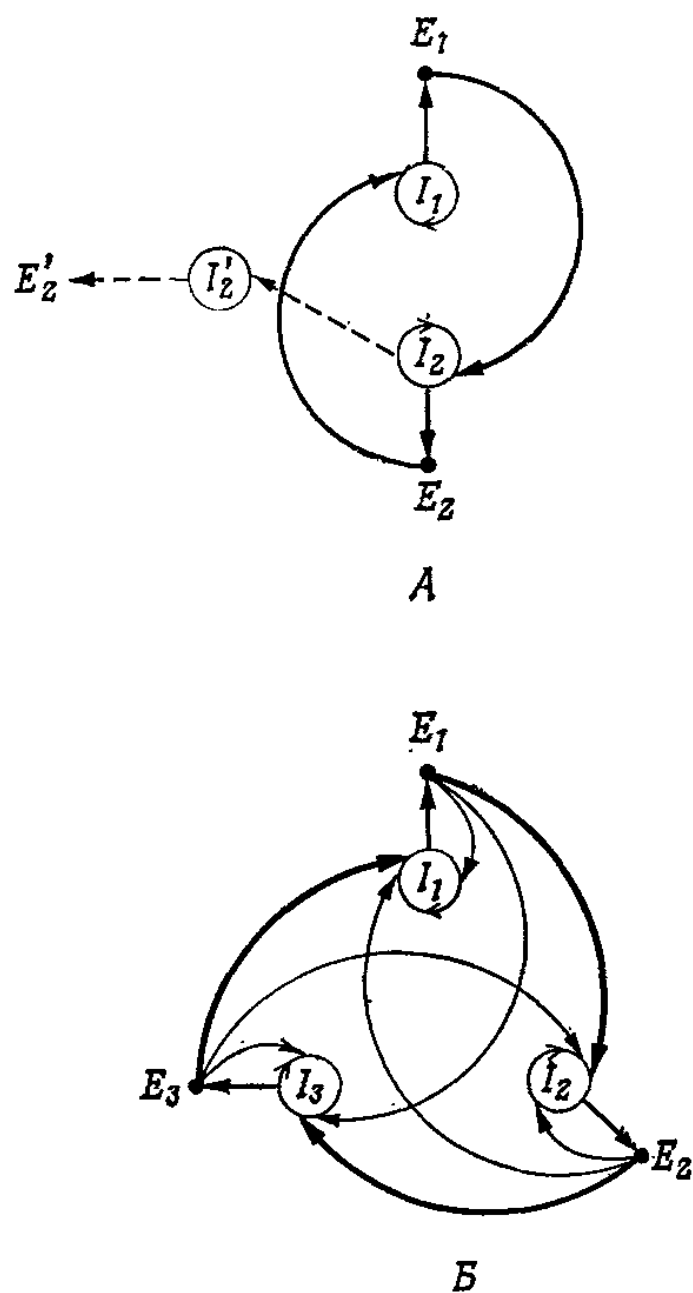


Рис. 54. Обобщение эволюционного принципа для гиперциклов. Связи должны удовлетворять критериям, полученным в табл. 17 и 18, т. е. взаимное усиление должно превалировать над самоусилением (см. толстые линии). А. Появляется I_2' — мутант I_2 . Б. Мутант (теперь I_3) включается в гиперцикл.

Обобщение эволюционного принципа для гиперциклов объясняется на примере перехода от двучленной системы к трехчленной (рис. 54).

Общие кинетические уравнения имеют тот же вид, что и уравнения, приведенные в табл. 17. Введение третьего члена I_3 (например, мутанта I_2) в двучленный гиперцикл, который был сначала устойчив, часто ведет к тому, что ранее устойчивая особая точка превращается в седловую. Это проявляется яснее всего, если допустить наличие циклической симметрии. При этом обозначения упрощаются

$$k_{11} = k_{22} = k_{33} = k_D; \quad k_{13} = k_{21} = k_{32} = k_+; \quad k_{12} = k_{23} = k_{31} = k_-$$

и получается следующая матрица кинетических коэффициентов:

$$K = \begin{vmatrix} k_D & k_- & k_+ \\ k_+ & k_D & k_- \\ k_- & k_+ & k_D \end{vmatrix}.$$

Особые точки и собственные значения таковы:

Вершины:

$$\bar{x}_1 = (c, 0, 0); \quad \omega_1^{(1)} = (k_- - k_D)c, \quad \omega_2^{(1)} = (k_+ - k_D)c;$$

\bar{x}_2, \bar{x}_3 — аналогично.

Ребра:

$$\bar{x}_4 = (0, k_D - k_-, k_D - k_+) \frac{c}{2k_D - k_+ - k_-};$$

$$\omega_1^{(4)} = \frac{k_- (k_D - k_-) + k_+ (k_D - k_+) + k_+ k_- - k_D^2}{2k_D - k_+ - k_-} c;$$

$$\omega_2^{(4)} = \frac{(k_D - k_+)(k_D - k_-)}{2k_D - k_+ - k_-} c.$$

\bar{x}_5, \bar{x}_6 — аналогично.

Внутренняя особая точка:

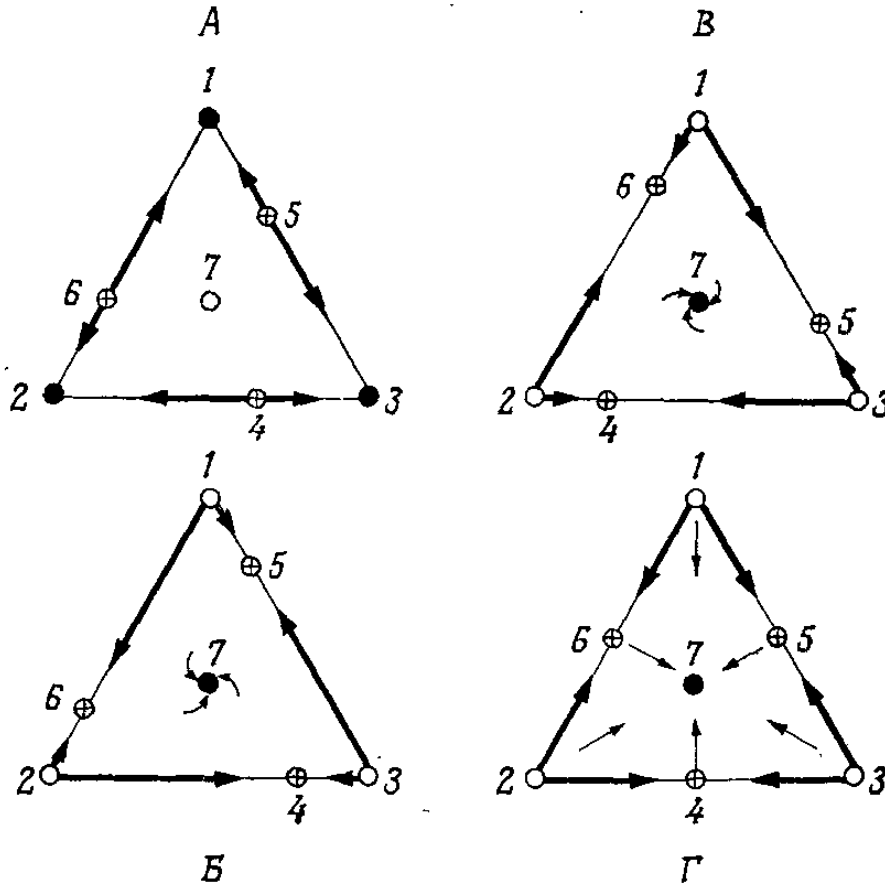
$$\bar{x}_7 = \left(\frac{c}{3}, \frac{c}{3}, \frac{c}{3} \right),$$

$$\omega_{1,2}^{(7)} = \left\{ 2k_D - k_+ - k_- \pm i \sqrt{3} (k_+ - k_-) \right\} \frac{c}{6}.$$

Снова особый интерес представляют четыре случая:

- а) $k_D > k_+, k_-$,
- б) $k_+ > k_- > k_D$,
- в) $k_- > k_+ > k_D$,
- г) $k_+ = k_- > k_D$,

которым соответствуют следующие диаграммы особых точек:



Большая величина диагональных членов ($k_D \gg k_+, k_-$) ведет к конкуренции — диаграмма А. В противном случае, т. е. для больших недиагональных элементов матрицы K , данные три вида проявляют кооперативное поведение. Направленне вращения вокруг спирального стока в центре симплекса определяется большей из двух констант k_+ и k_- . Когда $k_+ = k_-$, вращательная компонента отсутствует, и центральная особая точка является фокусом.

Пример, рассмотренный в этой таблице, является хорошей иллюстрацией усложнения гиперциклов в ходе эволюции. В отсутствие упрощающих предположений относительно констант скоростей анализ становится крайне сложным. Более детальное рассмотрение можно найти в работе [98], где дано обобщение на случай произвольной размерности.

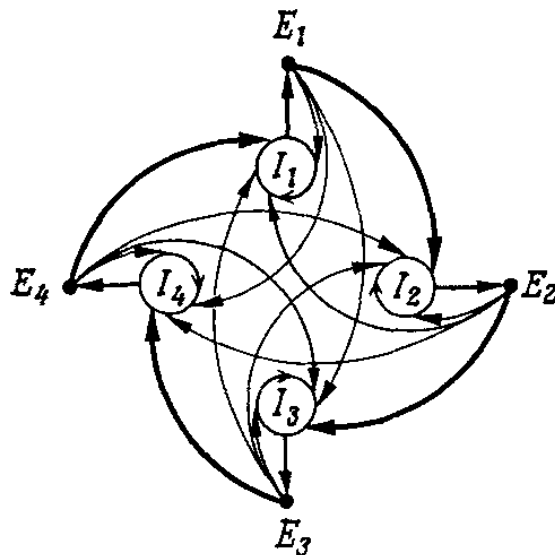


Рис. 55. «Реалистический» четырехчленный гиперцикл. Четыре информационных РНК I_1 — I_4 (мутанты одного общего предшественника) служат для кодирования четырех репликаз E_1 — E_4 , которые выполняют одинаковую функцию, но немного различаются по специфичности. Минус-цепи I_1 — I_4 могут одновременно служить в качестве адапторов аминокислот.

организация неизбежно возникает, раз уж такие взаимодействия появились.

Гиперцикл также будет неизбежно развиваться в сторону усложнения из-за мутаций (рис. 54, 55). Эволюционный принцип может быть обобщен по индукции таким образом, чтобы он был применим к любому n -членному гиперциклу. Тогда мутант I' может либо заменить I , либо вымереть, либо увеличить гиперцикл до $(n + 1)$ -членного размера (см. XVI. 10). Можно получить и более общие эволюционные критерии, как указано в табл. 18.

XVI. Десять вопросов, касающихся наших самых ранних молекулярных предков и тех следов, которые они оставили в биосинтетическом аппарате современных клеток

XVI.1. Один РНК-предшественник?

Этот вопрос касается сложности тех молекул, которые первыми начали воспроизводимо функционировать. Полинуклеотид из 100 остатков соответствует

сложности около 10^{60} альтернативных последовательностей. Если по соображениям стабильности мы ограничимся только сополимерами GC-типа (с включениями AU), то у нас останется еще 10^{30} возможных вариантов. Необходимой предпосылкой отбора одной или нескольких определенных последовательностей является точное самовоспроизведение. Оно неизбежно будет вести к дарвиновскому поведению с отбором одного квазивида: Отбор продуктов определяется просто их оптимальной селективной эффективностью, но структура продуктов зависит от исторического пути развития, который во многом определяется самокопированием олигонуклеотидных паттернов меньших размеров.

XVI.2. Что означает селективное преимущество для молекулы?

Селективная ценность определяется как оптимальная комбинация структурной стабильности и эффективности точной репликации. Она может быть выражена количественно через физические свойства молекулы в данной среде. Необходимым условием структурной стабильности, устойчивости к гидролизу и развития кооперативных свойств было увеличение длины молекулы. Небольшие олигонуклеотиды не могут укладываться с образованием стабильных структур и поэтому легко гидролизуются. Далее, они не обладают в достаточной степени способностью к комплексообразованию, чтобы обеспечить правильное копирование или трансляцию. Длина, с другой стороны, лимитируется скоростью репликации и точностью копирования. Было показано, что свойства GC-богатых последовательностей благоприятны для образования стабильных копий довольно большой длины. Сравнимы ли эти длины с размерами современных молекул тРНК, остается неясным. В молекулах тРНК была обнаружена гомология последовательностей [74], которая указывает на определенную степень самокопирования внутренних областей. Это самокопирование вполне могло произойти еще до того, как появилось

кодирование. Для возникновения трансляции требуются сильные взаимодействия между адапторами и информационными РНК, а эти взаимодействия не могут обеспечиваться молекулами небольшой длины.

Как только в результате появления трансляции система начинает воспроизводимо функционировать, селективная ценность достигает нового уровня. Тем не менее для каждой данной мРНК она должна выражаться через структурную стабильность и эффективность правильного воспроизведения. Однако теперь эти свойства зависят также от качества (и концентрации) продуктов трансляции. Поэтому специфическая связь, необходимая для гиперциклической организации, необходима также для любой системы, в которой продукты трансляции должны оцениваться при отборе и посредством этого эволюционировать. Такая связь имеет каталитическую или защитную природу.

XVI.3. Почему нужна гиперциклическая организация отдельных мутантных генов, а не один постоянно растущий геном?

Ответ на этот вопрос в основном был уже дан в части А. Для очень примитивного аппарата трансляции необходимо такое количество информации, какое соответствует геному современных РНК-содержащих фагов (или даже превосходит его). Информация фагового генома может сохраняться лишь с помощью фагоспецифического ферментативного комплекса, а извлечение этой информации основано на эффективности полного аппарата трансляции, предоставляемого клеткой-хозяином. Если согласиться с ответами, данными на первый и второй вопросы, то информация, необходимая для зарождения трансляции, должна создаваться на основе кооперации между несколькими мутантами — членами квазивидового распределения, а не просто в результате удлинения одной последовательности, для которой сначала не существовало бы никакого давления отбора.

Гиперциклическая стабилизация нескольких сосуществующих мутантов эквивалентна эволюции через дубликацию генов. Первоначально мутанты возникали как отдельные цепи, а не как ковалентно связанные дуплексы. Ограничения точности не допустили бы такого увеличения длины. Более того, вероятность получить необходимую комбинацию мутантов в одной цепи очень мала. Последовательности, состоящей из 100 остатков G и C, соответствует

100 мутантов с одной ошибкой,

4950 мутантов с двумя ошибками,

161 700 мутантов с тремя ошибками и т. д., т. е.

$$N_k = \binom{100}{k} \text{ мутантов с } k \text{ ошибками.}$$

Число цепей, содержащих n мутантных генов, отличающихся друг от друга в k определенных положениях (это различие может быть необходимой предпосылкой функционирования), равно

$$\binom{N_k + n - 1}{n} \approx \frac{N_k^n}{n!} \quad (\text{для } n \ll N_k),$$

что в случае $n = 4$, и $k = 3$ дает $3 \cdot 10^{19}$ альтернативных последовательностей. При наличии даже таких малых отклонений в реплицировавшихся генах вероятность найти копию с благоприятной комбинацией в одной гигантской цепи будет почти нулевой для любой популяции разумных размеров. Однако каждый из *изолированных* мутантных генов, содержащих три замещения, будет представлен в любой макроскопической популяции в достаточно большом количестве.

Наконец, не менее важно, что транспортные РНК — адапторы трансляции — во всяком случае должны были существовать в виде отдельных цепей. Эволюция объединенного генома потребовала бы с самого начала сложной регуляции на уровне транскрипции.

Естественный источник изолированных цепей РНК — это квазивидовое распределение. Все последовательности были сходны друг с другом, и таким же сходством, должно быть, обладали продукты их

трансляции. Как только один продукт трансляции приводил к возникновению фактора связи, эта функция появлялась и у всех других продуктов. В результате могла возникнуть и циклическая связь, необходимая для гиперциклической организации. Можно даже сказать, что гиперциклическая организация самым естественным образом вытекает из любой реалистической модели примитивной трансляции.

Дает ли современная организация генома клеток прокариот какие-либо указания на свою первоначальную структуру? Современные гены, конечно, гораздо больше примитивных информационных РНК. Удлинение генов, так же как и дупликация, давало преимущество всякий раз, когда постоянно растущая точность ферментативного аппарата допускала это. Продукты трансляции становились более сложными, создавая предпосылки для эволюции по пути создания более тонких полиферментных механизмов, использующих дифференцированные ферменты, которые произошли от общего предшественника. В примитивных системах отсутствовали механизмы рекомбинации, используемые клетками в настоящее время. Поэтому современная структура прокариотического генома могла сформироваться в результате удлинения изолированных генов, их дупликации и трипликации с образованием оперонов и их последующего отображения на ДНК, которая может использовать более совершенные средства воспроизведения, приводящие к образованию объединенного генома. Размеры современных оперонов хорошо соответствуют тем размерам, с которыми может работать совершенная РНК-репликаза (1000—10 000 нуклеотидов).

XVI.4. Необходимы ли при старте тРНК?

Этот вопрос можно сформулировать по-другому: «А почему бы адапторам не быть малыми олигонуклеотидами?»

Адапторы без информационных РНК бессмысленны. Короткие нуклеотидные последовательности не могут функционировать в качестве информационных

РНК. Декапептиды эквивалентны уже почти половине молекулы тРНК. Далее, короткие олигонуклеотиды могут быть нестабильны, так как они лишены третичной структуры. Простейшая симметричная структура, т. е. отдельная петля, стабилизированная четырьмя или пятью парами оснований, требует целых пятнадцати нуклеотидов. Бесферментное специфическое распознавание аминокислот, в котором участвуют одновременно антикодоновая петля и 3'-конец адаптора, возможно только при наличии более протяженных структур. То же самое верно для взаимодействий между двумя соседними адапторами, которые необходимы для стабилизации комплекса мРНК — пептидил-тРНК или для конформационного изменения (например, HF→fh), которое может облегчить перемещение растущей пептидной цепи вдоль информационной РНК. Маас и др. [99] недавно сообщили о таком конформационном изменении в антикодоновой петле тРНК, которое они зарегистрировали, наблюдая изменение флуоресценции основания У. Для фрагмента антикодоновой петли (т. е. для декануклеотида с последовательностью антикодоновой петли) этот эффект, по-видимому, отсутствует. Все это позволяет думать, что недостаточно длинные последовательности РНК не могут функционировать в качестве адапторов.

Тогда можно задать вопрос: «Что отличает адаптор от информационной РНК?» Их минимальная длина должна быть примерно одинаковой. Обе они должны иметь специфическую складчатую структуру, которая создает предпосылки воспроизводимого узнавания их факторами связи.

Поскольку каждая тРНК и каждая мРНК нуждаются в факторе связи (например, в репликазе), благоприятствующем их селекционной стабилизации, то двойные функции РНК-последовательностей совершенно необходимы. Итак, плюс- и минус-цепи данной РНК-последовательности могут использоваться совместно как информационная РНК и как адаптор.

XVI.5. Дают ли современные тРНК какие-либо указания на свое происхождение

Структурное сходство может быть либо следствием адаптации для достижения общей цели, либо указывать на происхождение от общего предка. Современные тРНК имеют много структурных соответствий [75]. Позволяют ли эти аналогии сделать вывод о том, что все тРНК произошли от общего предка? В соответствии с анализом Джукса [76], на этот вопрос можно ответить осторожным «да». Почему следует проявлять осторожность, можно показать на примере. Одна особенность, характерная для всех изученных до сих пор прокариотических и эукариотических тРНК,— это наличие последовательности ТψСГ в так называемой Т-петле — общем сайте узнавания при рибосомном контроле. Недавние исследования метанобразующих бактерий [77] показали, что у этих микроорганизмов, считающихся «наиболее древними ветвями из всех, которые встретились до сих пор в генеалогии бактерий», указанная последовательность отсутствует, а вместо нее у одной группы имеется последовательность ψψСГ, а у другой — УψСГ. Хотя эти данные не ставят под сомнение, а наоборот, подчеркивают тесные эволюционные связи данного класса микроорганизмов с другими прокариотами, они определенно показывают, что общие свойства могут согласованно приобретаться целыми классами. Это особенно характерно для тех молекул, которые производятся общим аппаратом, например рибосомой — местом синтеза всех белковых молекул.

Рис. 56 показывает соответствие последовательностей для четырех тРНК из *E. coli*, которые, как мы считаем, являются современными представителями первичных кодоновых адапторов. К сожалению, в нашем распоряжении не было последовательности аланин-специфичной тРНК для кодона GCC. Если сравнить этот вид тРНК, имеющей антикодон ^{54}UGC , с соответствующей тРНК для валина с антикодоном ^{54}UAC , то мы обнаружим большее сходство с последней тРНК, чем той, которая изображена на ри-

⁵ GCGGGAAUAGCUCAGDDGGD ¹⁰ AGAGCACGACCUUGCCAAGG ¹⁵ ²⁰ ²⁵ ³⁰ ³⁵ ⁴⁰
 GGGGGCAUAGCUCAGCDGGG AGAGCGCCUGCUUUGCACGC
 GGAGCGGSAGUUCAGDCGGDD AGAAUACCUGCCUGUC*ACGC
 GCGUCCGSAGCUCAGDDGGDDAGAGCACCCACCUUGACAUGG

⁴⁵ UCGGGGUCGCGAGTQCGAGUCUCGUUUUCCCGCUCSSA ⁵⁰ ⁵⁵ ⁶⁰ ⁶⁵ ⁷⁰ ⁷⁵
 AGGAGGUCUGCGGTQCGAUC CCGCGCGCUCSSACSSA
 AGGGGGUCGCGGGTQCGAGUCCCGQCCGUUCCGSSA
 UGGGGGUCGGUGGTQCGAGUCCACUCGGACGCACSSA

Рис. 56. Сравнение последовательностей тРНК для Gly, Ala, Asp и Val. К сожалению, последовательность тРНК^{Ala}, соответствующая кодону GCC, пока неизвестна. Возможно, сходство между тРНК для Gly и Ala было бы еще больше в случае правильной последовательности, которая соответствует антикодону GCC (о том, что это так, можно судить на основании сходства между двумя последовательностями для Ala и Val, которые относятся к антикодонам *UGC и *UAC соответственно). Эти последовательности показывают, что спаренные области состоят в основном из GC-пар и что близкое сходство говорит о родстве между Gly/Ala и Asp/Val (см. S в положении 8 для Asp и Val вместо U для Gly и Ala или вставку D между положениями 20 и 21 для Asp и Val). A — аденозин, *A = 2MA — N(2)-метиладенозин, C — цитидин, D — 5,6-дигидроуридин, G — гуанозин, *G = 7MG — N(7)-метилгуанозин, Q = Ψ — псевдоуридин, S — тиюридин, T — рибозилтимин, U — уридин, *U = 5AU — 5-оксидецетилуридин.

сунке (совпадают 57, а не 54 положения). Следовательно, правильная тРНК^{Ala} с антикодоном GGC может иметь не 44, а большее число совпадений с приведенной на рисунке глициновой тРНК. Если не принимать во внимание этот «физический недостаток», то можно сделать вывод, что 1) все приведенные последовательности совпадают в более чем в половине положений (33 совпадения, если включить «неправильный» Ala, или 41 совпадение для Gly, Asp и Val); 2) подгруппа Gly/Ala отличается несколькими особенностями от подгруппы Asp/Val (тиоуридин «S» вместо U в положении 8, вставка 5,6-дигидро-U между положениями 20 и 21; 3) все приведенные последовательности имеют явный избыток G и C по сравнению с A и U (или их производными), особенно в двухцепочечных участках.

Далее, сравнение с последовательностями других тРНК показывает, что эти особенности — хотя они определенно не редки и для большинства других тРНК — особенно сильно выражены для данной группы. Сходство так же велико, как и для разных адапторов одной и той же аминокислоты в одном и том же организме.

Один факт особенно красноречив. Если сравнить последовательности двух адапторов, имеющих комплементарные антикодоны (например, для Asp и Val), то совпадение между обеими плюс-цепями тРНК будет гораздо более полным, чем между одной плюс-цепью (если читать от 3' к 5') и другой минус-цепью (если читать от 5' к 3'). Действительно, если сравнить таким образом плюс- и минус-цепи одной и той же тРНК, то соответствие окажется лучшим. Эти соответствия являются выражением замечательной внутренней симметрии тРНК, при которой антикодон располагается почти точно в середине последовательности и тем самым образуется симметричная двумерная структура. Мы можем считать это свойство указанием на то, что тРНК возникла очень рано как независимая репликативная единица. Требование того, чтобы плюс- и минус-цепи имели сходную структуру, важно, лишь если они представляют собой независимые репликативные единицы, а не интегрированы структурно с образованием длинной геномной последовательности, как это имеет место теперь. Такая же ситуация характерна и для фаговых РНК или их вариантов, которые размножаются в виде отдельных репликативных единиц [78].

С другой стороны, адаптация тРНК к общему аппарату вызвала, вероятно, общие отклонения от необходимой вначале симметрии. Тот факт, что зеркальные копии плюс- и минус-цепей одной и той же тРНК обнаруживают больше симметричного сходства, чем зеркальные копии плюс- и минус-цепей тРНК с комплементарными антикодонами, позволяет думать, что обе тРНК эволюционно развились как мутанты одной и той же цепи, а не двух комплементарных цепей. Тогда можно сделать вывод, что

современные адапторы для кодонов GGC(Gly), GCC(Ala), GAC(Asp) и GUC(Val) возникли из одного квазивида как мутанты с одной ошибкой, происходящие от общего предка. Однако исходной симметрии было недостаточно (а почему ее должно быть достаточно?) для того, чтобы адапторные функции могли возникнуть у плюс- и минус-цепей данной РНК.

XVI.6. Как в информационной РНК могли возникнуть кодонные паттерны без запятых?

Первые мРНК были, вероятно, идентичны первым адапторам (или их комплементарным цепям). Действительно, функции адаптора и мРНК предполагают структурное соответствие. Каким бы ни был кодонный паттерн в последовательности мРНК, он должен иметь свою комплементарную реализацию в адапторе. Для примитивных систем такое требование легче всего выполнить, используя общий структурный паттерн для обоих типов молекул: чтобы первые адапторы были минус-цепями первых мРНК (если считать, что плюс-цепью по определению всегда является мРНК) и чтобы определенная симметрия структуры позволяла и плюс- и минус-цепям узнаваться одной и той же репликазой.

Первые длинные молекулы РНК были обогащены G и C, что является следствием отбора, основанного на критериях структурной стабильности и точности копирования. Молекулы с общим кодонным паттерном — таким, как GGC/GCC — требуют инструкции в виде затравки (например, могут использоваться катализаторы или имеющиеся открытые петли РНК) с последующей внутренней дубликацией. Это неизбежно приведет к образованию структур, содержащих по меньшей мере два кодонных паттерна с внутренней комплементарностью, например 5'GGC3' и 3'CCG5'.

Существует хороший пример, показывающий эффективность дубликации внутреннего паттерна при синтезе *de novo* и усиление последовательностей РНК фаговыми репликазами. Если Q_{β} -репликазу лишить всяких матриц, то она начинает «связывать» друг с

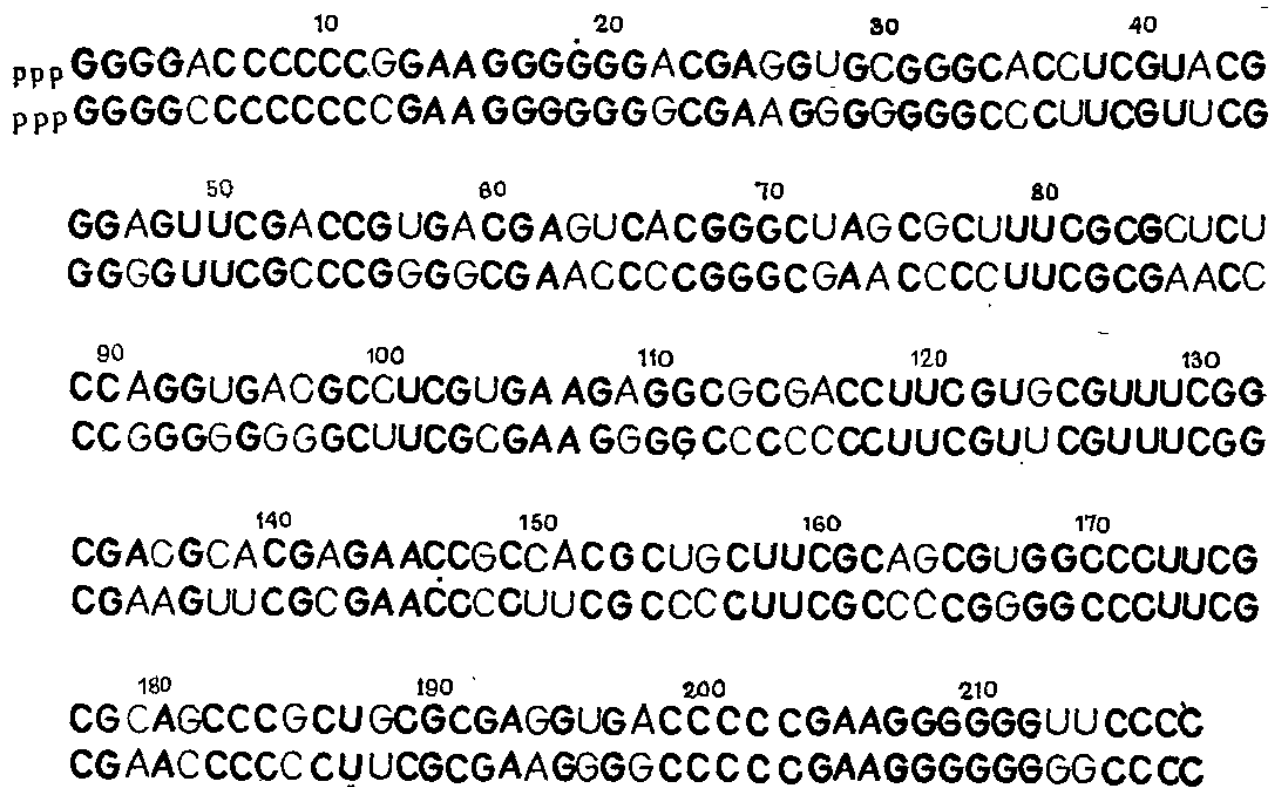


Рис. 57. Сопоставление последовательности миди-варианта Q_{β} (по Спигелману и др.) с искусственной последовательностью, состоящей из блоков CCC(C) и UUCG и из комплементарных блоков GGG(G) и CGAA. Совпадение в 169 положениях из 218 позволяет предположить, что миди-вариант — это продукт, образованный *de novo* ферментом Q_{β} -репликазой, которая имеет сайты узнавания для CCC(C) и UUCG (EF Tu). Кинетика синтеза *de novo* показывает, что в ферментативных сайтах узнавания образуется тетрамер, после чего происходит какое-то внутреннее самокопирование со случайными замещениями. Этот конкретный миди-вариант обычно выигрывает конкуренцию со всеми имеющимися последовательностями и поэтому является, вероятно, наиболее эффективной матрицей. Данный пример показывает, как в процессах примитивного копирования могут возникать однородные паттерны.

другом свои собственные затравки, которые она затем дублирует и избирательно усиливает, пока наконец не образуется однородная макроскопическая популяция последовательностей РНК с длиной несколько сотен нуклеотидов. В других условиях среды получают другие (но тоже однородные) популяции последовательностей [8]. Спигелман, Миллз и их сотрудники определили последовательности нескольких таких «миди-вариантов», которые всегда содержат специфический сайт узнавания для Q_{β} -репликазы [78]. Последующие эксперименты пролили свет на механизм этого синтеза *de novo*, показав, что небольшие

фрагменты, соответствующие последовательностям, которые узнаются ферментом, синтезируются как затравки, а затем внутренне дублируются и избирательно усиливаются. Более ранние исследования [22] показали, что фермент может узнавать, в частности, последовательности CCC(C) и UUCG. UUCG соответствует последовательности TψCG, общей для всех тРНК и специфически взаимодействующей, как известно, с рибосомным фактором элонгации EF Tu, который является субъединицей комплекса Q_β-репликазы. Сопоставление миди-варианта с последовательностью, состоящей исключительно из двух упомянутых олигонуклеотидов и комплементарных им GGG(G) и CGAA, выявляет совпадение в более чем 75% положений, что говорит об эффективности внутреннего копирования последовательностей-затравок (рис. 57).

Аналогичные соображения можно высказать и о первичных механизмах образования однородных паттернов. Если среди многих паттернов появился 5'GGC/5'GCC и, возможно, также 5'GAC/5'GUC, то эти паттерны в информационной РНК могли породить воспроизводимую трансляцию по механизму Крика и др. [3] и приобрести способность селективно усиливаться с помощью своих воспроизводимых продуктов трансляции.

XVI.7. Какими были первые функционально активные белки?

Простейший белок мог быть гомополипептидом, например полиглицином. Имел ли он какую-нибудь каталитическую активность? На этот вопрос могут и должны дать ответ эксперименты. Гетерогенные последовательности с достаточно большим числом остатков — скажем, 15—30 — способны к образованию β-слоя, имеющего активный центр, в котором концевая карбоксильная группа фиксирована в определенном положении вблизи концевой аминокислоты (рис. 58). Расстояние между ними варьирует с длиной цепи, так как в β-слое антипараллельные цепи закручены [79]. Концевая аминокислота имеет pK около 8, следовательно, каталитический центр

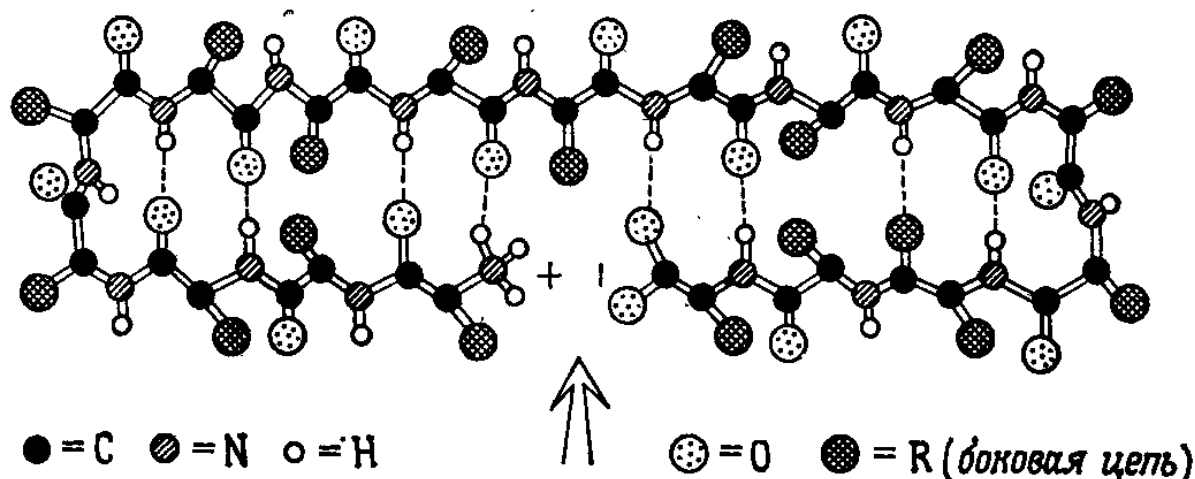


Рис. 58. Простой предшественник фермента изображен в виде β -структуры, состоящей из 15—25 аминокислот (для образования которой необходима мРНК длиной 45—75 нуклеотидов). В активный центр входит концевая аминогруппа, которая является очень эффективным донором протонов ($pK \sim 8$), концевая карбоксильная группа, которая служит акцептором протонов, и каталитически активная боковая цепь (например, Asp или Ser). Можно построить много альтернативных структур, но лишь в некоторых из них цепи будут закручены в спираль именно с тем шагом, который необходим для образования эффективного активного центра.

содержит по меньшей мере эффективную протонную донорно-акцепторную систему. Образованию β -структур благоприятствует чередование остатков Gly—Ala. Однако полипептиды, состоящие исключительно из остатков Gly и Ala, плохо растворимы, поэтому они могут присутствовать в ощутимых количествах лишь на поверхностях раздела.

Строение β -слоев изучали Чоу и Фасман [80], которые проанализировали рентгеноструктурные данные для 29 белков и выявили 459 β -изгибов. В числе трех остатков, чаще всего встречающихся в области β -изгиба во всех его четырех положениях, находятся Gly и Asp, тогда как в участках, расположенных за β -изгибом, преобладают гидрофобные остатки.

Важной предпосылкой каталитической активности является определенное пространственное расположение концевых групп. Для стабилизации воспроизводимого складывания может оказаться необходимым использование двух или более видов аминокислот. Давно известно, что β -слои являются важными элементами белковой структуры. Согласно Левитту

[81], существует весьма общий принцип стабилизации с их помощью активной конформации белков.

Большая распространенность глицина и аланина могла в основном определить характер первых белков, но для обеспечения растворимости более длинных последовательностей были необходимыми полярные боковые цепи. Наличие четырех видов аминокислот привело бы, конечно, к гораздо большей структурной гибкости. Если следующими двумя кандидатами были аспарагиновая кислота и валин, то могли образоваться глобулярные структуры, стабилизируемые гидрофобными взаимодействиями боковых цепей валина и аланина и сольбилизируемые карбоксильными боковыми цепями аспарагиновой кислоты. Этот остаток создает условия для образования специфических каталитических центров с участием двухвалентных ионов металлов.

Нашему воображению слишком трудно охватить колоссальное многообразие всех возможностей. В настоящее время проводятся эксперименты, цель которых состоит в исследовании различных структур в отношении их эффективности в дискриминации различных последовательностей РНК и их структурных особенностей. Результаты, полученные с рибонуклеазами [82], стимулируют поиски «минимальной структуры», способной специфически узнавать определенные последовательности РНК.

XVI.8. Необходимы ли при старте синтетазы?

В трехмерной структуре современных тРНК (см. часть А, рис. 14) антикодоновая петля фиксирована на значительном расстоянии от аминоацильного сайта. Такая структура адаптирована к функциональным потребностям современных молекул тРНК, которые определяются рибосомами и структурой синтетаз. С другой стороны, известно, что тРНК может совершать конформационные переходы, которые существенно изменяют ее форму и размеры. Риглер и др. [83] флуоресцентными методами исследовали времена жизни конформационных состояний, а также

времена вращательной релаксации и сделали вывод о существовании по меньшей мере трех конформационных состояний, быстро переходящих друг в друга. Аналогичные результаты получены Олсоном и др. [84], которые использовали методы, основанные на рассеянии лазерного излучения. Заселенность различных конформационных состояний сильно зависит от концентрации ионов магния. Важно отметить, что в условиях, которые соответствуют морской воде (~ 50 мМ Mg^{2+}), присутствует конформер, отличающийся от L-формы, обнаруженной при кристаллографических исследованиях, — по своей форме он гораздо ближе к цилиндру.

Все сказанное самым прямым образом связано с поставленным вопросом. Первичные ферменты состояли из весьма ограниченного числа аминокислотных остатков и поэтому не могли быть очень большими глобулярными структурами. Для того чтобы гарантировать однозначное соответствие аминокислоты антикодону, либо должны были существовать такие совершенные ферменты, как современные аминокислот-ацилсинтетазы, либо тРНК должна была иметь такую структуру, чтобы осуществлялся гораздо более тесный контакт между аминокислотным и антикодонным сайтами, чем в случае L-формы, с тем чтобы можно было одновременно контролировать оба сайта. В противном случае высокий темп мутаций на ранних этапах эволюции очень скоро нарушил бы любое однозначное случайное соответствие между двумя этими сайтами. С другой стороны, конформационный переход все-таки необходим, потому что механизм образования пептидной связи (см. рис. 48) требует достаточного пространственного разделения мРНК и растущей пептидной цепи. Данные, о которых шла речь, вызывают желание поразмыслить о таких возможностях. С другой стороны, при наличии структур типа изображенной на рис. 49, В кодирование первых аминокислот могло бы осуществляться даже без помощи ферментов. Структура самой тРНК заведомо достаточно сложна, чтобы стало возможным специфическое узнавание. Было отмечено [85], что четвер-

тое основание от 3'-конца (т. е. основание, следующее за 3'АСС) имеет какое-то отношение к антикодону. Первоначальные предположения об однозначной корреляции для всех тРНК в конечном счете не оправдались. Однако такая корреляция могла играть ключевую роль в первичных механизмах специфического распознавания аминокислот транспортными РНК. По имеющимся данным для *E. coli* и фага Т4, в положении, которое следует за 3'АСС, находятся такие нуклеотиды: U для Gly, A для Ala, G для Asp и A для Val. Для первичных адапторов было, несомненно, важным обеспечение однозначного соответствия с помощью дискриминирующих сайтов — свойство, которое могло быть частично утеряно в ходе эволюции. Это, конечно, чисто спекулятивные рассуждения, которые требуют экспериментальной проверки.

Вывод. Возможно, синтетазы не были необходимы на самых ранних этапах, а транспортные РНК в конечном счете оказались неудачной попыткой Природы создать ферменты из нуклеиновых кислот. Более эффективное распознавание могло эволюционно развиться из факторов связи, которые были предназначены для специфического узнавания тРНК-подобных структур.

XVI.9. Какими были первые ферменты?

Если синтетазы в действительности не являются необходимыми для зарождения трансляции (а это большое «если!»), то в качестве единственного абсолютного первичного условия согласованной эволюции трансляции у нас остаются только факторы связи — вероятно, репликазы. С помощью такой функции селективное преимущество, заключенное в продукте трансляции, может быть наиболее эффективно использовано для обратной связи с мРНК. Следовательно, специфические репликазы (которые все принадлежат к одному классу сходных белковых молекул) не только создают предпосылку для возникновения гиперциклической связи, но и оказываются

крайне важными для дальнейшей эволюции белков, так как только они могут сообщить информационной РНК, что именно является фенотипически выгодным, и осуществить отбор на генотипическом уровне, т. е. посредством интенсивного синтеза определенной мРНК. Как мы увидим в следующем разделе, такая селективная связь между генотипическим и фенотипическим уровнями лучше всего работает в сочетании с пространственным разделением, или компартментацией.

Далее следует ожидать появления каталитической поддержки для различных функций трансляции. Если репликазы установили определенную связь с тРНК-подобными информационными РНК (включая сюда и плюс-, и минус-цепи), то их способности к распознаванию вполне могут использоваться для осуществления синтетазной и «транслатазной» (т. е. прерибосомной) функций. Другими словами, генный дубликат репликазы вполне может быть предшественником мРНК для синтетазы, а также для такого фактора трансляции, как EF Tu, тем более что химические механизмы функционирования репликазы и фактора переноса очень сходны и в современных системах, по-видимому, осуществляются одними и теми же остатками.

Постепенно дивергирующие двойные функции, вероятно, представляли собой очень раннюю особенность механизмов репликации и трансляции, точно так же как дубликация генов была одной из движущих сил эволюции на ее более поздних этапах. Такие двойные функции оставили свой след в современных клеточных органеллах; вирусы также использовали их при своей постбиологической эволюции в клетке-хозяине. Геном фага Q_β кодирует только одну субъединицу своей репликазы, однако использует еще три фактора из клетки-хозяина, которые были идентифицированы как рибосомный белок S₁ и факторы элонгации EF Tu и EF Ts [87, 88].

Бибрихер [89] исследовал свойства этих факторов и установил, что они выполняют одновременно несколько функций при рибосомном контроле, используя свою приобретенную способность узнавать

молекулы тРНК. Он утверждает, что в клетке *E. coli* имеется предшественник фагоспецифичного β -фактора Q_{β} -репликазы, и это, по-видимому, действительно так. Используя иммунологические методы, Бибрихер смог идентифицировать белок, содержащий EF Tu и EF Ts, который ведет себя как предшественник Q_{β} -репликазы в незараженной клетке *E. coli* и который, по-видимому, участвует (неизвестным пока образом) в синтезе РНК *E. coli* [87]. Не исключено, что подобные соответствия будут обнаружены и в случае синтетаз. Создается впечатление, что как только возникала некая функция — например, способность узнавать определенные типы РНК, — Природа использовала эту функцию всюду, где только в ней появлялась необходимость (например, при специфической репликации, при транспорте рибосом и контроле над их функционированием, при активации аминокислот).

В некоторых отношениях образование РНК-содержащих фагов могло имитировать эволюцию первичных мРНК. Фаги используют максимально возможное число функций клетки-хозяина, за исключением одной — специфического узнавания своего собственного генома (т. е. связь посредством специфической репликации). Различные фаги (например, Q_{β} , MS 2, R 17) наследуют различные факторы узнавания [9], хотя все они происходят от общего предка в клетке-хозяине. В части А было также показано, что первая фаза цикла размножения РНК-содержащего фага эквивалентна простому гиперциклическому процессу усиления.

XVI.10. Почему же в конечном счете — клетки с объединенными геномами?

Гиперциклы способствуют увеличению информационного содержания благодаря функциональной интеграции системы мРНК, в которой длина отдельной репликативной единицы ограничена из-за конечной точности копирования. Увеличение информационного

содержания позволяет построить воспроизводимый аппарат репликации и трансляции, при наличии которого продукты трансляции могут эволюционировать к большей эффективности. Это дает возможность повысить точность, что в свою очередь приводит к увеличению информационного содержания каждой отдельной репликативной единицы и тем самым снова улучшает качество ферментов. Одновременно, как показано в разд. XV, гиперцикл сам эволюционирует к большей сложности, интегрируя все больше дифференцированных мутантных генов.

Увеличение информационного содержания не только дает ферменты лучшего качества — оно может позволить каждой репликативной единице наследовать информацию более чем для одного фермента. Тем самым двойные функции могут быть вычеркнуты из списка предпосылок эволюции на ее ранних этапах, т. е. дублицированные мРНК могут развиваться независимо, в соответствии с конкретными функциональными нуждами своих продуктов трансляции. Это могло послужить началом развития оперонных структур с механизмом контроля над одновременной репликацией нескольких структурных генов. Таким образом, из репликаз могли эволюционно развиться обычные полимеразы, связанные со специфическими контрольными факторами индукции или репрессии.

Теперь, когда мы осознали преимущества функциональной связи, которая, вероятно, является абсолютно необходимой для зарождения трансляции, уместно спросить: почему же функциональная связь в конечном счете была заменена полной структурной интеграцией всех генов — ведь геномы даже самых примитивных клеток структурно едины. Каковы ограничения гиперциклической организации и какие усовершенствования в нее можно внести?

В системе, регулируемой функциональными связями, можно выделить два класса мутаций. Мутации первого класса изменяют прежде всего фенотипические свойства самой информационной РНК и тем

самым модифицируют ее таргетную функцию¹ по отношению к специфической репликазе или фактору контроля. Эти мутации особенно важны на ранних этапах эволюции — роль фенотипических свойств РНК-структур очень велика. Такие таргетные мутации немедленно становятся селективно эффективными: благоприятные мутации будут фиксироваться, неблагоприятные — элиминироваться.

Второй класс мутаций, которые могут быть или не быть нейтральными по отношению к таргетной функции, относится к фенотипическим изменениям в продуктах трансляции. Чем более специализированы информационные РНК, тем специфичнее мутация может изменить характер функционирования продукта трансляции.

Будет ли отбор специфически благоприятствовать мутанту зависит только от таргетной функции; при этом неважно, в какую сторону изменится продукт трансляции — в благоприятную, в неблагоприятную или же его изменение будет нейтральным. Для более поздних стадий доклеточной эволюции самым обычным следствием мутации будет фенотипическое изменение продукта трансляции в сочетании с неизменной таргетной функцией. Затем мутант будет размножаться, но при этом не будет наблюдаться отбора ни в пользу мутанта, ни в пользу дикого типа (в случае, когда продукт трансляции мутанта окажется для последнего неблагоприятным). Единственное, что реально достигнуто, — это оценка системы в целом. Это может осуществляться путем пространственного разделения систем мРНК: с помощью ниш или даже еще более эффективно — посредством компартментации. Информационная РНК в данном компартменте может обогащать окружающую среду своими продуктами трансляции и конкурировать с РНК других компартментов, используя свою эффективность пролиферации. Частично этого можно достичь и просто

¹ Таргетная функция (target function) — это свойство нуклеиновой кислоты быть мишенью (target) при специфическом взаимодействии с данной макромолекулой, т. е. ее способность быть правильно узнаваемой. — *Прим. перев.*

путем пространственного разделения. Однако компартимент без гиперциклической организации вообще не будет работать. Усиленная конкуренция между всеми информационными РНК в ограниченном жизненном пространстве компартимента разрушит любое согласованное функционирование.

Компартимент может пролиферировать более эффективно, согласуя свое воспроизведение с редупликацией всего своего набора генов. Для этого, конечно, необходим достаточно сложный механизм контроля, работу которого можно облегчить интеграцией всех генов в одну гигантскую репликативную единицу. Такая индивидуализация всего компартимента требует высокой точности аппарата репликации. В части А мы сравнили информационное содержание на различных биологических уровнях с соответствующими (и наблюдаемыми) точностями репликации (см. табл. 4).

Индивидуализация компартимента, вероятно, связана с переходом от РНК-генов или оперонов к ДНК-геномам, так как достаточно высокую точность может гарантировать лишь механизм репликации ДНК. Новая индивидуализованная единица представляла собой интегрированную протоклетку. Предыдущей функциональной организации генов и их продуктов пришла на смену более совершенная структурно-функциональная организация. Более подробное изучение кольцевых генетических карт может еще обнаружить какие-то признаки фазы начала структурной организации, хотя рекомбинационные эпигенетические эффекты могли стереть многие следы.

Вследствие унификации и индивидуализации чистый рост при бесполом размножении клеток подчиняется автокаталитическому закону первого порядка (в отсутствие ингибирования). Дарвиновские свойства таких систем допускают селективную эволюцию и сосуществование большого числа разнообразных видов. Оказывается, что интегральное единство клетки имеет превосходство над более консервативной формой гиперциклической организации.

С другой стороны, последующая эволюция многоклеточных [90] организмов снова могла использовать аналогичные или другие формы гиперциклической организации (нелинейные сети), где новыми субъектами были клетки, и тем самым в некоторых отношениях она могла оказаться сходной с процессом молекулярной самоорганизации.

XVII. Реалистические граничные условия

Обсуждение «реалистического гиперцикла» было бы неполным без рассмотрения реалистических граничных условий. Мы остановимся на этом вопросе лишь вкратце не потому, что не признаем важности граничных условий для исторического процесса эволюции — возникновение жизни на нашей планете в конце концов есть историческое событие, — а потому, что понимаем, как мало в действительности мы можем сказать. В то время как *ранние стадии жизни* из-за эволюционной согласованности оставили хоть какой-то след в современных организмах, соответствующие свидетельства *ранних условий среды* отсутствуют. До сих пор в нашем обсуждении мы были, наверное, в какой-то мере несправедливы по отношению к экспериментам, имитирующим первичный безматричный синтез белка, которые были проведены Фоксом [91] и другими исследователями (см. обзор Дозе и Раухфусса [92]). Наша цель состояла в том, чтобы понять принципы ранних форм организации, которые сделали возможными самовоспроизведение, отбор и эволюционную адаптацию биосинтетических механизмов — какие свойственны современным живым клеткам. Белки не наследуют основных физических предпосылок для такой адаптационной самоорганизации — по крайней мере не в такой очевидной форме, как это свойственно нуклеиновым кислотам. С другой стороны, белки наследуют колоссальную функциональную емкость, в чем они далеко превосходят нуклеиновые кислоты. Поскольку белки в первичных условиях могут образоваться гораздо легче, наличие больших количеств разнообразных материалов,

обладающих каталитическими свойствами, было, наверное, существенным свойством среды. Исследования в этой области ясно показали, что в первичных условиях мог существовать вполне эффективный белковый катализ.

В этом отношении особого внимания заслуживают поверхности. Покрытые каталитически активным материалом, они могли представлять собой наиболее благоприятные места для протекания первичных процессов синтеза. Ограничение молекулярного движения в пределах плоскости очень сильно увеличивает эффективность столкновений, особенно если мы имеем дело с последовательностями реакций высокого порядка.

Л. Онзагер [93] подчеркивал, что в первичных условиях океаны, вероятно, почти сплошь были покрыты слоями отложившихся гидрофобных и гидрофильных веществ (см. также [94]). Возможно, в этих мультислоях существовали благоприятные условия для протекания предбиологических химических реакций. Ввиду очевидных преимуществ, которые дают поверхности, мы исследовали свойства гиперциклов при соответствующих граничных условиях среды.

В качестве простой модели рассмотрим систему, подобную той, которая схематически изображена на рис. 59. Синтез полимеров протекает только в поверхностном слое ($r = 0$), который имеет конечную емкость для связывания матриц и ферментов. Кинетические уравнения будут иметь тот же вид, что и для гомогенного раствора, только теперь мы должны в явном виде учесть диффузию. Мы вводим функцию роста, которая зависит от поверхностной концентрации репликативных молекул и ферментов. Предполагается, что диффузия внутри поверхностного слоя идет быстро и не лимитирует скорости реакций. Адсорбция и десорбция макромолекул рассматривается как реакция обмена между поверхностным слоем ($r = 0$) и слоем раствора, прилегающим к поверхности ($0 < r \leq 1$). Разложение может происходить на поверхности раздела и/или (только) в объеме раствора.

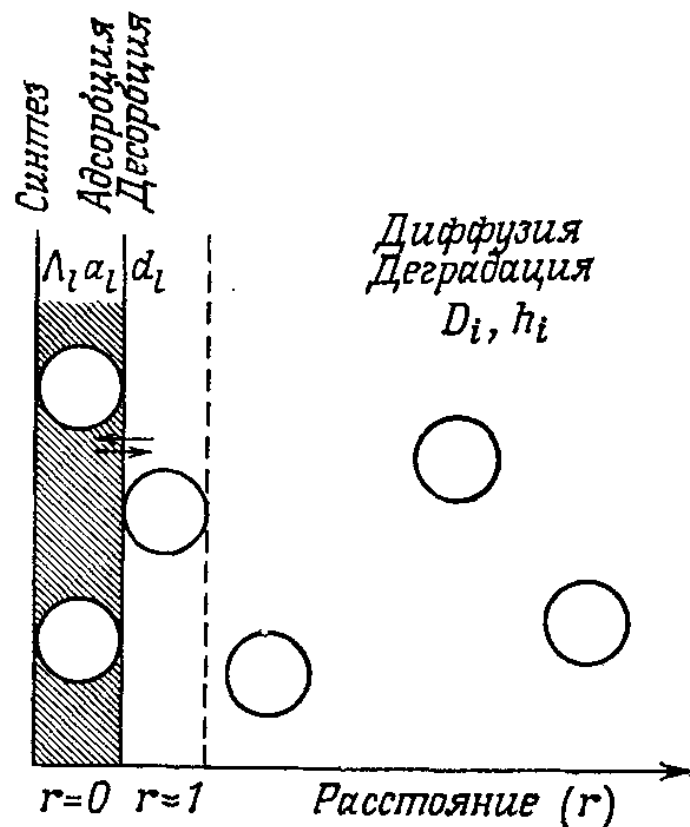
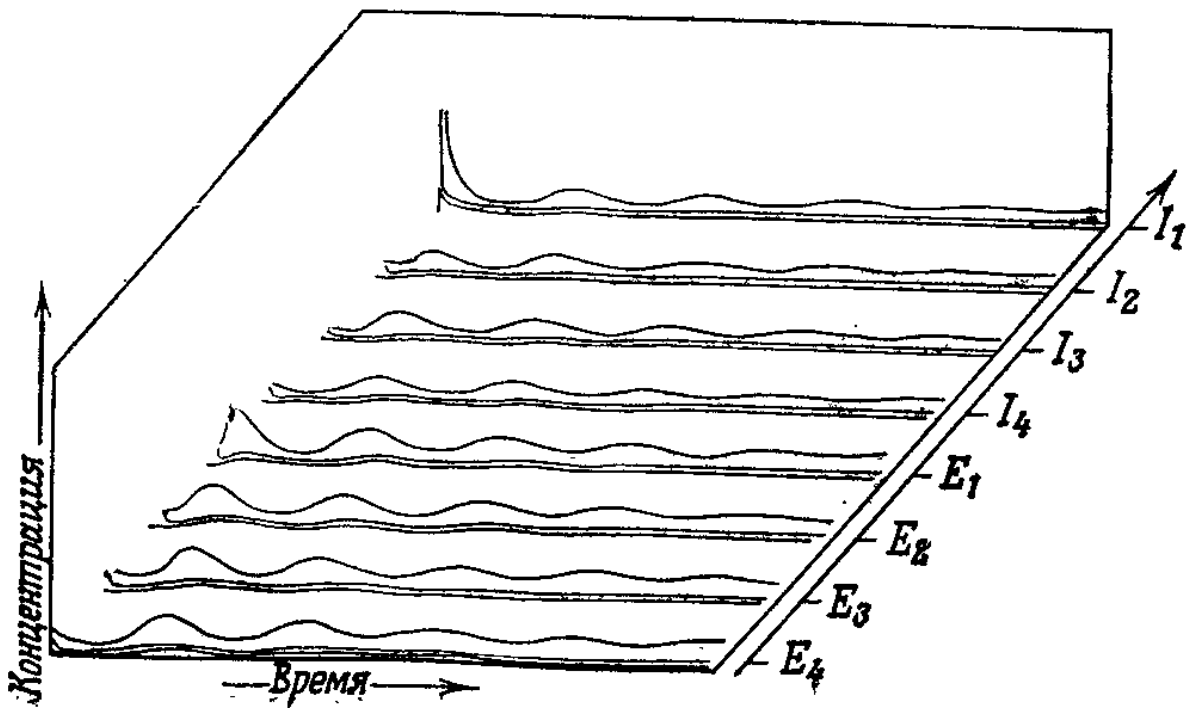
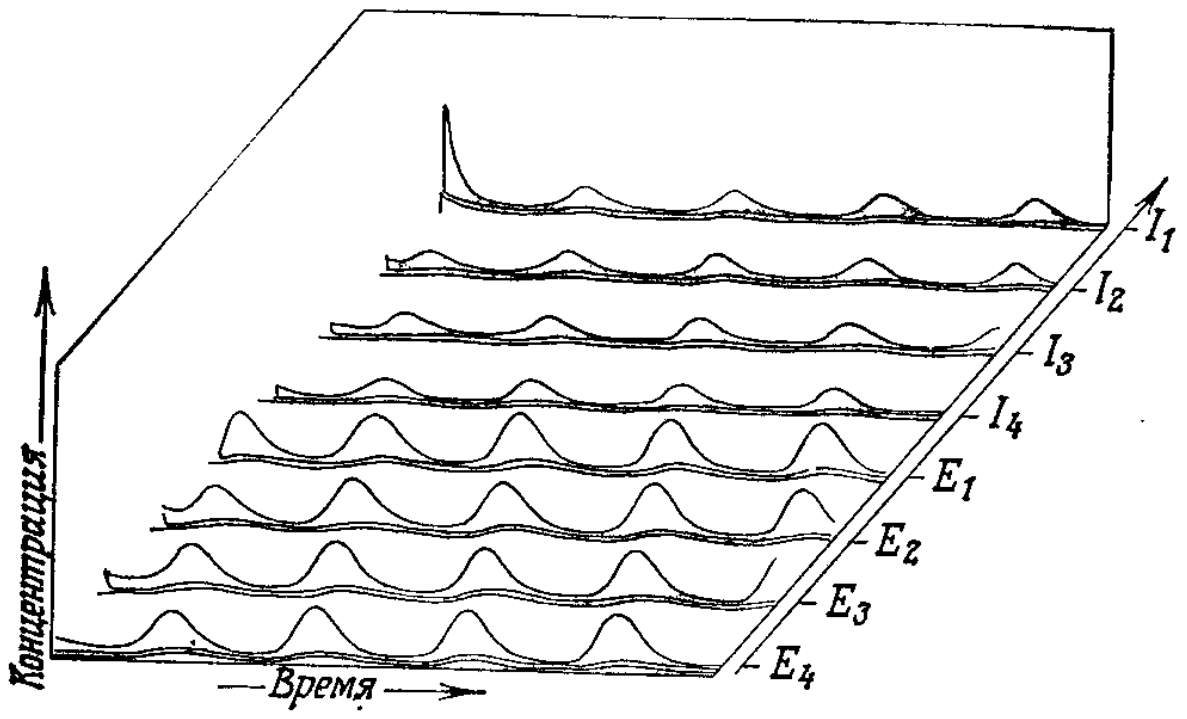
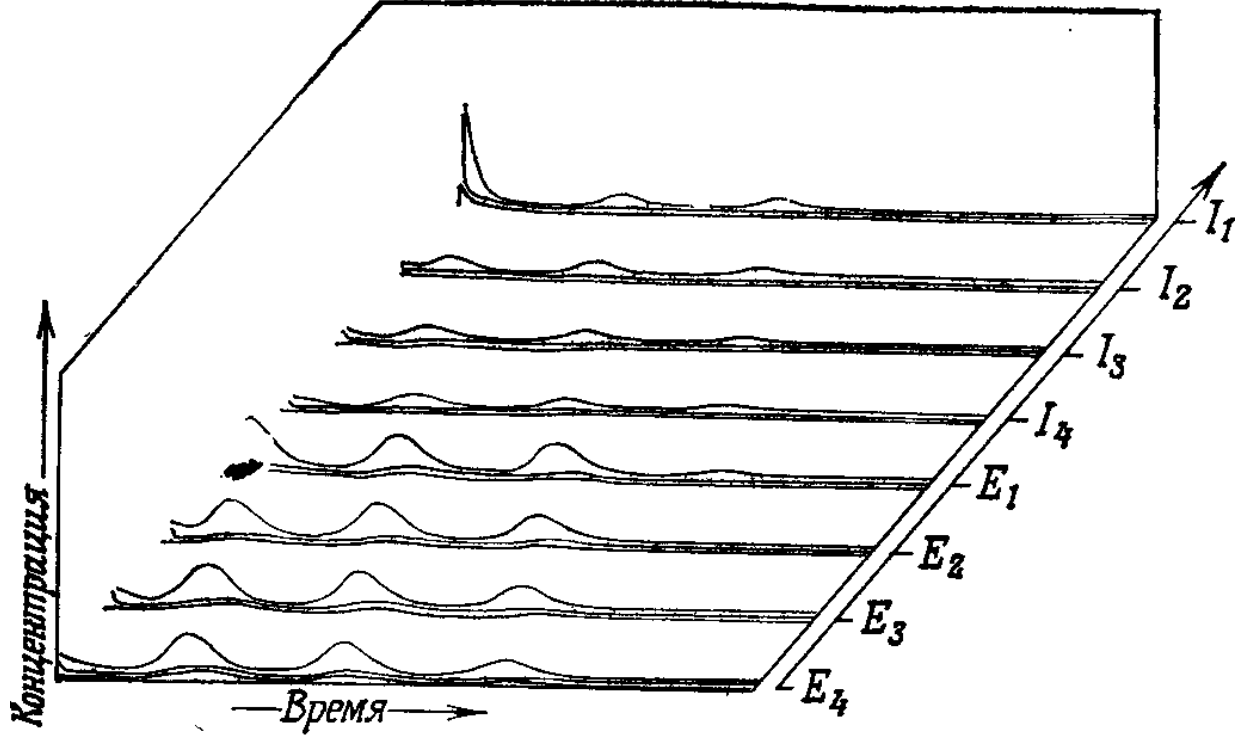


Рис. 59. Схематическое представление модели гетерогенной реакции с участием гиперциклической связи. Выделены три пространственные области: $r = 0$ — поверхность раздела, $r = 1$ — пограничный слой у поверхности раздела, $r > 1$ — объем раствора. На химические реакции, протекающие по схеме гиперцикла, накладывается диффузия к поверхности раздела и обратно.

Наконец, транспорт веществ к поверхности и от нее описывается соответствующим диффузионным членом.

В зависимости от предполагаемого механизма синтеза может возникнуть необходимость в рассмотрении независимых сайтов связывания для матриц и для ферментов. Мы использовали эту модель, чтобы получить какое-то представление о поведении гиперциклов с трансляцией (см. разд. IX в части Б). Численное интегрирование для нескольких наборов кинетических параметров проводилось по методу, описанному в литературе [95]. Можно выделить три основных результата, два из которых полностью аналогичны тем, которые получились для гиперциклов в гомогенном растворе.

А. При очень низких концентрациях полинуклеотидов и полипептидов или при больших значениях K_i [см. уравнения (73), (75) и (79)] поверхностные плотности полимеров не достигают стационарных



значений, а убывают либо монотонно, либо с затухающими колебаниями. Соответственно через некоторое время макромолекулы вымирают (рис. 60).

Б. Выше определенного порогового значения полной концентрации в системах с $n < 4$ наблюдается выход на предельный цикл. Такая ситуация аналогична случаю нижнего концентрационного предела в гомогенном растворе (рис. 61).

В. При достаточно больших концентрациях мы, наконец, получаем стационарное состояние:

$$\lim_{t \rightarrow \infty} \frac{\partial x_i}{\partial t} = 0, \quad \lim_{t \rightarrow \infty} \frac{\partial y_i}{\partial t} = 0,$$

$\bar{x}_i > 0, \bar{y}_i > 0, i = 1, 2, \dots, n$ (рис. 62),

где x_i и y_i — концентрации ферментов и информационных РНК соответственно, \bar{x}_i и \bar{y}_i — их стационарные значения, t — время.

←

Рис. 60—62. Интегральные кривые, полученные численным интегрированием системы дифференциальных уравнений в частных производных, которая соответствует модели, изображенной на рис. 59. В кинетические уравнения входит функция роста Λ_i , введенная в части Б, которая относится к четырехчленному гиперциклу с трансляцией (часть Б, разд. IX) и которая имеет ненулевые каталитические члены только на поверхности раздела ($r = 0$). Далее, уравнения учитывают адсорбцию и десорбцию (a_i, d_i , описывающие переходы частиц между состояниями $r = 0$ и $r = 1$), гидролиз (происходящий при $r \geq 1$) и диффузию в объем раствора ($r > 1$, т. е. обмен с пограничным слоем $r = 1$). Три кривые в каждой группе относятся к трем пространственным положениям: $r = 0$ (верхняя), $r = 1$ (средняя) и $r = 2$ (нижняя). Рисунки 60—62 различаются только тем, что они соответствуют разным значениям констант стабильности каталитически активных комплексов $I_i \times E_{i-1}$, которые имеют наибольшее значение (0,16) для рис. 60, промежуточное (0,06) — для рис. 61 и минимальное (0,04) для рис. 62. Баланс между образованием продуктов и их удалением из системы достаточен для того, чтобы можно было не вводить разбавление. Для ситуации, представленной на рис. 60, где все партнеры I_i и E_i вымирают, при выбранных нами значениях соответствующих однородным параметрам f_i, k_i — согласно части Б, разд. IX — a_i, d_i и D) автокаталитическое образование продуктов не может конкурировать с удалением их посредством диффузии и разложения. В двух остальных случаях на границе раздела устанавливается устойчивая гиперциклическая организация, при которой численности популяций либо колеблются (рис. 61), либо остаются стационарными (рис. 62).

В системах меньшей размерности ($n \leq 4$) наблюдалось лишь поведение типа А и В.

Эти модельные вычисления были дополнены исследованиями нескольких сходных проблем с использованием стохастических методов машинной имитации. Результаты снова выявили близкую аналогию между поведением гиперциклов на поверхностях раздела и в гомогенном растворе (как было подробно описано в части Б).

Учет реалистических граничных условий — это тот момент, на который делается особый акцент в статьях Куна [96]. Мы не спорим с допущением «структурированной среды», но мы и не знаем также, можно ли согласиться с постулированием строго определенных условий среды, пока не будут получены экспериментальные данные, которые покажут по меньшей мере полезность таких постулатов.

Наши модели никоим образом не ограничены случаем пространственной однородности (см. вычисления, приведенные выше). Логические предпосылки, лежащие в основе различных моделей, — а именно существование колоссального числа структурных вариантов, требующих естественного отбора, ограничение информационного содержания отдельных репликативных единиц из-за ограниченной точности или обязательное наличие функциональной связи для обеспечения согласованной эволюции всего ансамбля — на самом деле остаются в силе при допущении *любых* реалистических условий среды. Вывод Куна о том, что предложенный нами тип организации «ограничивается частным случаем пространственной однородности», не затрагивает существа дела. Кто стал бы сейчас утверждать, что жизнь могла возникнуть *только* в пористой среде, или на поверхностях раздела, или в мультислоях на поверхности океана, или в толще морской воды? Модели показывают, что она может зародиться — с большей или меньшей вероятностью — при любых перечисленных граничных условиях, но в том и только том случае, если выполняются определенные требования. Эти требования относятся к проблеме генерации и накопления ин-

формации и качественно не меняются при различных граничных условиях.

В основном то же самое можно сказать и в отношении временной однородности или неоднородности. В части Б было показано, что селективные критерии могут принимать особенно простую форму, если они применяются для условий стационарности. Поскольку в них входят относительные, а не абсолютные скорости реакций, они качественно остаются одинаковыми независимо от того, будет ли система расти, колебаться или находиться в стационарном состоянии.

Не вызывает сомнений, что отжиг является полезной процедурой в решении многих проблем, связанных с разделением фаз. Однако вопрос о том, будут ли тепловые флуктуации в равной степени полезны и для отбора более длинных полинуклеотидов, должен быть решен экспериментально.

Чтобы установить, могут ли температурные флуктуации способствовать отбору цепей с повышенным информационным содержанием, следует тщательно проанализировать относительные температурные коэффициенты всех соответствующих процессов. Температурный коэффициент гидролиза, по-видимому, максимален. *Инструктированная* репликация, как правило, вовсе не ускоряется при высоких температурах. Дело в том, что включающийся нуклеотид должен кооперативно связаться со своим комплементарным основанием в матрице, одновременно используя стэкинг-взаимодействие с предыдущим основанием растущей цепи, а выше точки плавления матрицы такое взаимодействие становится невозможным. Все эти соображения остаются в силе для любого вида окружающей среды, будь то водная фаза, поверхностный слой или компартмент в крупнозернистом или пористом материале.

Важным фактором повышения информационного содержания является то, что комплементарные взаимодействия сильнее некомплементарных. В общем случае дискриминация идет лучше при низких температурах, чем при высоких. Миллер и Оргел [63,

с. 126] делают следующий вывод из своих экспериментальных данных: «Мы не знаем, какова была температура первичного океана, но мы можем сказать, что нестабильность различных органических соединений и полимеров является убедительным доводом в пользу того, что жизнь не могла бы возникнуть в океане, если бы его температура не была ниже 25°C . Температура 0°C весьма способствовала бы этому процессу, а температура -21°C была бы даже еще лучше. При таких низких температурах большая часть воды на примитивной Земле находилась бы в виде льда, а жидкая морская вода имела бы только в экваториальных областях океана.

Есть и другое основание считать, что жизнь возникла при низких температурах — в океанах или в озерах. Все реакции, направляемые матрицами, которые, по-видимому, привели к возникновению биологической организации, идут только ниже температуры плавления соответствующей организованной полинуклеотидной структуры. В случае полинуклеотидно-моноклеотидных спиралей — это температуры от 0°C (или еще ниже) до примерно 35° .

Среду, в которой возникла жизнь, часто называют теплым разбавленным бульоном органических соединений. Мы думаем, что холодный концентрированный бульон являлся бы лучшей средой для возникновения жизни».

XVIII. Непрерывность эволюции

В этой последней части трилогии мы попытались показать, что гиперциклы действительно могут описывать реальные материальные системы, а не являются лишь умозрительными структурами.

Эволюция консервативна и поэтому, не считая случайных резких изменений, является, по-видимому, почти непрерывным процессом. Отбор фактически основан на неустойчивостях, вызванных появлением благоприятных мутантов, которые приводят к развалу ранее устойчивых распределений. Однако потомки обычно так близки к своим непосредственным предкам, что изменения выявляются очень постепен-

но. Предбиологическая эволюция не составляет исключения из этого правила.

Теперь кратко резюмируем, каковы, по нашему мнению, существенные стадии при переходе от неживого к живому (см. рис. 63).

1. Порядок появления первых макромолекул диктуется их структурной стабильностью и распространенностью составляющих химических компонентов. На ранней стадии, вероятно, имелись большие количества неких белковоподобных веществ и намного меньшие количества РНК-подобных полимеров. Однако РНК-подобные полимеры в силу своих физических свойств наследуют способность к самовоспроизведению, а это является необходимой предпосылкой для систематической эволюции.

2. Состав первых полинуклеотидов тоже диктуется распространенностью соответствующих химических компонентов. Ранние нуклеиновые кислоты вовсе не были однородным классом макромолекул — сюда входили L- и D-соединения и использовались различные типы эфирных связей — преимущественно 2'—5', но, кроме того, и 3'—5'. Воспроизводимость последовательностей зависит от точности копирования. Наиболее длинными воспроизводимыми последовательностями могут быть GC-богатые полимеры. С другой стороны, необходимы и замещения AU-типа. Они позволяют получить определенную структурную гибкость, которая благоприятствует быстрому воспроизведению. Воспроизводимые последовательности образуют квазивидовое распределение, которое характеризуется дарвиновским поведением.

3. Паттерны без запятых в распределении могут играть роль информационных РНК, в то время как цепи с открытыми комплементарными паттернами (возможно, минус-цепи мРНК) — роль адапторов. Первые аминокислоты сопоставлялись адапторам в порядке своей распространенности. Продукты трансляции были однородны, потому что они состояли в основном из глицина и аланина. То же самое должно быть справедливо для большинства неинструментированных белков.

*Первые
полинуклеотиды*

*GC-богатые
квазивиды*

*Зарождение кодирования,
продукты трансляции,
богатые Gly и Ala*

*Гиперциклическая фиксация
кода с GC-рамкой,
кодирование Gly, Ala, Asp
и Val, примитивные
репликазы*

*Эволюция гиперциклической
организации, код RNY,
репликазы, синтетазы,
предшественники рибосом,
эволюция кода,
пространственная
компартиментация*

*Полностью изолированные
гиперциклы, адаптированные
ферменты репликации
и трансляции, эволюция
функций метаболизма
и контроля, оперонная
структура, РНК
соответствует по длине
геному современных РНК-
содержащих вирусов*

*Протоклетка,
интегрированный геном-ДНК,
сложные ферменты,
механизмы контроля
считывания, дальнейшая
дарвиновская эволюция
допускает диверсификацию*

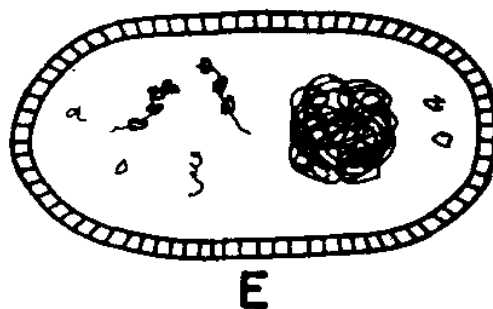
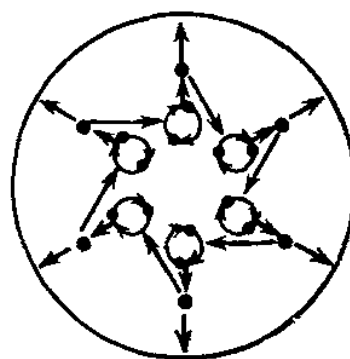
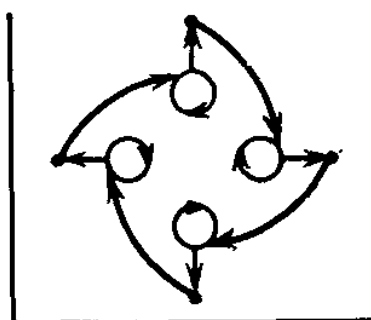
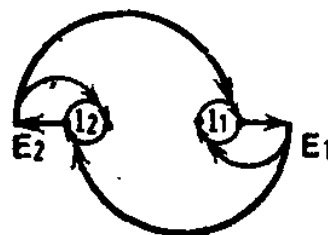
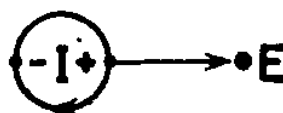
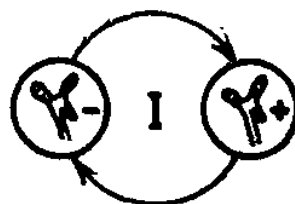


Рис. 63. Гипотетическая схема эволюции от отдельных макромолекул до интегрированных клеточных структур.

4. Если какой-нибудь из продуктов трансляции оказывает каталитическое содействие репликации своей собственной мРНК, то именно эта мРНК может стать доминирующей в распределении и вместе со своими близкородственными мутантами будет присутствовать в больших количествах. Этот процесс может быть инициирован каким-либо *неинструированным* белком среды, состав которого отражает относительные распространенности аминокислот, и, следовательно, этот белок может копировать свойства первичных *инструированных* белков.

5. Если мутанты доминирующей мРНК дают дополнительные преимущества, то они, согласно критериям гиперциклической эволюции, могут включиться в цикл воспроизведения. Таким образом может возникнуть гиперциклическая организация с несколькими кодонами. Такая гиперциклическая организация является предпосылкой для согласованного эволюционного развития аппарата трансляции. Включается все большее и большее число мутантов, и постоянно растущая точность воспроизведения обуславливает удлинение последовательностей. Объединенные предшественники посредством дубликации генов могут дать начало различным ферментативным активностям (репликазам, синтетазам, рибосомным факторам), которые затем дивергируют.

6. Сложная гиперциклическая организация может эволюционировать дальше только в том случае, если она эффективно использует благоприятные фенотипические изменения. Для селективного благоприятствования соответствующим генотипам необходимо пространственное разделение (путем образования компартментов или комплексов), при наличии которого может происходить отбор среди различных комбинаций мутантов. Рибосомы можно рассматривать как реликты такого комплексобразования.

Мы не знаем, на какой стадии такая система приобрела способность полностью интегрировать свое информационное содержание в виде одного гигантского генома. Для этого был необходим высоко совершенный ферментный аппарат, и роль хранителя

информации постепенно перешла к ДНК (это могло произойти уже на очень ранних этапах).

Этих кратких экскурсов в исторический процесс доклеточной эволюции, по-видимому, достаточно, чтобы показать, в каком направлении может пойти развитие, иницированное гиперциклической интеграцией самореплицирующихся молекулярных единиц, и как развивающаяся система может в конце концов конвергировать с образованием такой сложной организации, как прокариотическая клетка. Мы хотим подчеркнуть спекулятивный характер части В. Ранняя фаза самоорганизации оставила лишь некие следы, но не свидетельства, так что многие важные ее этапы все еще остаются неясными.

В наши намерения даже не входило раскрытие исторической правды. Для процесса, так сильно зависящего от случая — где недетерминированные микроскопические события, такие, как мутации, усиливаются и в конечном счете определяют ход макроскопического развития, — полная реконструкция истории вообще невозможна. «Нищета историцизма»¹ существует даже в биологии. С другой стороны, принципы, управляющие историческим процессом эволюции — даже в своих тонких деталях, — могут быть вполне доступны нашему пониманию. Тех следов, которые оставила история в современных системах, может оказаться достаточно, чтобы когда-нибудь мы смогли составить эти « n уравнений с n неизвестными».

В этой части мы хотели показать лишь то, что особый класс реакционных сетей, которые мы назвали гиперциклами, действительно является простейшей реалистической молекулярной организацией, обеспечивающей сосуществование функционально связанных самореплицирующихся единиц. Саморепликация необходима для сохранения информации. Следовательно, гиперцикл является простейшей системой, которая способна обеспечить эволюцию воспроизводимых функциональных связей. Он может воз-

¹) См. известную книгу Карла Поппера «Нищета историцизма» (Popper K., *The poverty of historicism*, London, 1957). — Прим. перев.

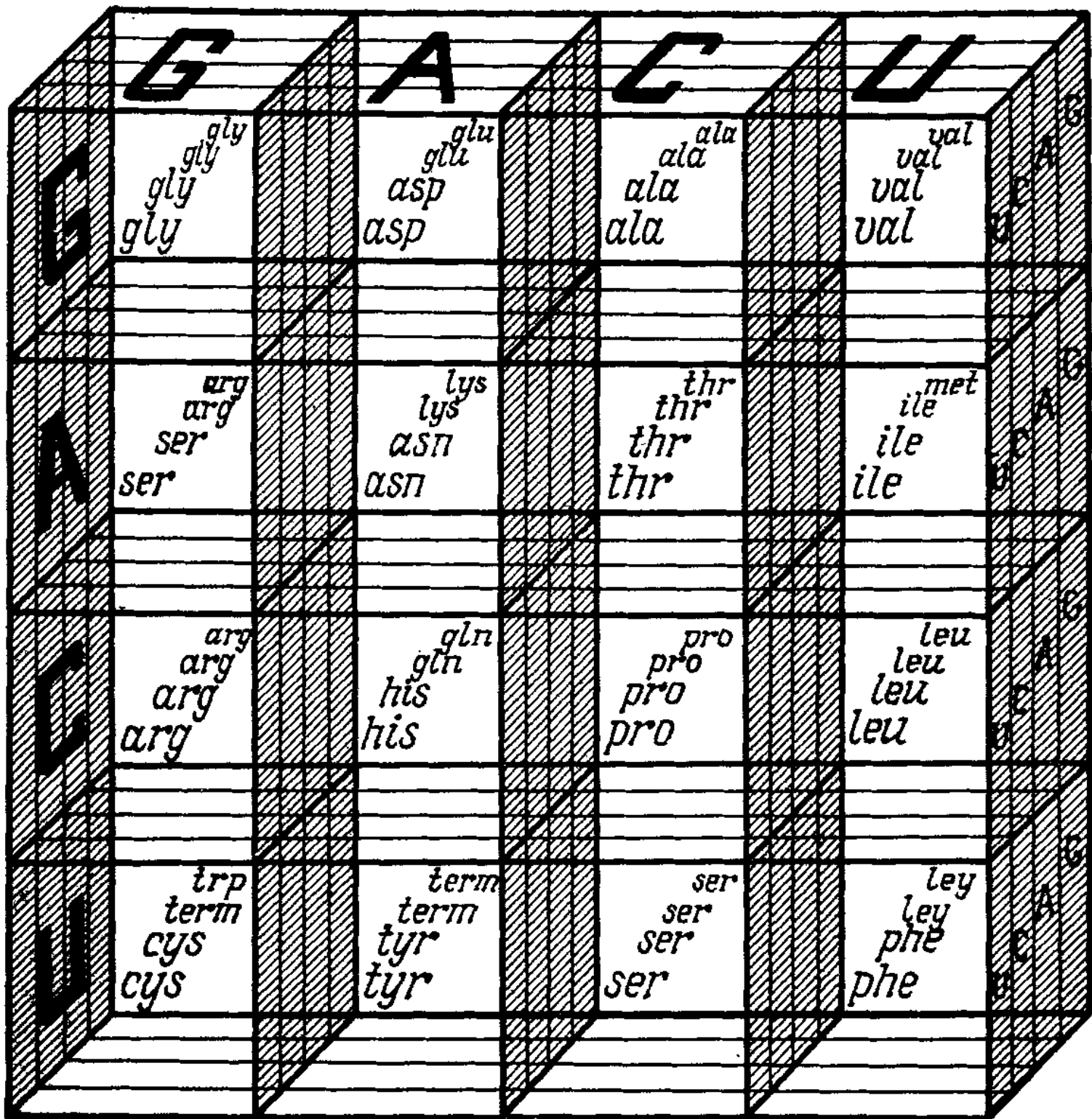


Рис. 64. Генетический код — это универсальный ключ для перевода генетической информации с законодательного языка нуклеиновых кислот на исполнительный язык белков. На рисунке три координаты геометрического пространства сопоставляются с позициями букв в триплетных кодонах. Четыре буквы алфавита нуклеиновых кислот расположены таким образом, чтобы кодоны для самых распространенных аминокислот попали в самую верхнюю часть куба.

никнуть из одной самореплицирующейся единицы и ее мутантов, т. е. из одного (молекулярного) квазивида. Гиперцикл неизбежно появлялся всякий раз, когда это разрешали условия, продиктованные Природой. И еще: «Если кто-нибудь может назвать более прекрасный треугольник, лежащий в основе вещей, мы будем приветствовать его не как соперника, но как друга правды»¹.

(Платон, «Тимей») [97]

¹ Вот это место в переводе С. С. Аверинцева: «Что ж, если кто-нибудь выберет и назовет нечто еще более прекрасное, предназначенное для того, чтобы создавать эти (четыре тела), мы подчинимся ему не как неприятелю, но как другу...». (Платон, Сочинения в трех томах, т. 3, часть I, с. 496. — М.: Мысль, 1971). — *Прим. перев.*

ЛИТЕРАТУРА

1. *Wright S.*, *Genetics*, 16, 97 (1931).
2. *Woese C. R.*, *The Genetic Code*, New York, Harper and Row, 1967.
3. *Crick F. C. R., et al.*, *Origins of Life*, 7, 389 (1976).
4. *Eigen M.*, *Naturwissenschaften*, 58, 465 (1971). [Имеется перевод: Эйген М. Самоорганизация материи и эволюция биологических макромолекул. — М.: Мир, 1976.]
5. *Bethe H.*, In: *Les Prix Nobel en 1967*, p. 135, Stockholm, 1969.
6. *Krebs H.*, In: *Nobel Lectures, Physiology or Medicine, 1942—1962*, p. 395, Amsterdam, Elsevier, 1964.
7. *Spiegelmann S.*, *Quart. Rev. Biophys.*, 4, 213 (1971); Haruna I, *Spiegelmann S.*, *Proc. Nat. Acad. Sci., USA*, 54, 579 (1975); Mills D. R., Peterson R. L., Spiegelmann S., *ibid.* 58, 217 (1967).
8. *Sumper M., Luce R.*, *ibid.* 72, 1750 (1975).
9. *Küppers B.-O.*, *Naturwissenschaften*, 66, 228 (1979).
10. *Kornberg A.*, *DNA Synthesis*, San Francisco, W. H. Freeman, 1974. [Имеется перевод: Корнберг А. Синтез ДНК. — М.: Мир, 1977.]
11. *RNA-Phages* (Zinder N. D., ed.), Cold Spring Harbor Monograph Series, Cold Spring Harbor Laboratory, 1975.
12. *Fischer R. A.*, *Proc. Roy. Soc. B*, 141, 510 (1953); Haldane J. B. S., *Proc. Camb. Phil. Soc.*, 23, 838 (1927); Wright S., *Bull. Am. Math. Soc.*, 48, 233 (1942).
13. *Eigen M.*, *Ber. Bunsenges. Physik. Chem.*, 80, 1059 (1976).
14. *Dobzhansky Th.*, *Genetics of the Evolutionary Process*, New York, Columbia Univ., Press, 1970.
15. *Darwin Ch.*, *of the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, Paleontological Society, 1854; *The Origin of Species*, Chapter 4, London, 1872. Everyman's Library, London: Dent and Sons, 1967. [Имеется перевод: Дарвин Ч. Происхождение видов, Сочинения, т. 3. — М. — Л.: Изд-во АН СССР, 1938.]
16. *Darwin Ch., Wallace A. R.*, *On the Tendency of the Species to Form Varieties and on the Perpetuation of the Species by Natural Means of Selection*, *J. Linn. Soc. (Zoology)*, 3, 45 (1858).
17. *Eigen M., Wankler R.*, *Ludus Vitalis, Mannheimer Forum 73/74*, Studienreihe Boehringer, Mannheim, 1973. [Имеется перевод: Эйген М., Винклер Р. Игра жизни. — М.: Наука, 1979.]
18. *Eigen M., Winkler-Oswatitsch R.*, *Das Spiel*, München, Piper, 1975.

19. *Schrödinger E.*, What is Life? Cambridge Univ. Press, 1944. [Имеется перевод: Шредингер Э. Что такое жизнь? — М.: Атомиздат, 1972.]
20. *Thompson C. J., McBride J. L.*, Math. Biosci., 21, 127 (1974).
21. *Jones B. L., Enns R. H., Rangnekar S. S.*, Bull. Math. Biol., 38, 15 (1976).
22. *Küppers B.-O.*, Dissertation, Göttingen, 1975.
23. *Glansdorff P., Prigogine I.*, Thermodynamic Theory of Structure, Stability and Fluctuations, New York, Wiley-Interscience, 1971. [Имеется перевод: Гленсдорф П., Пригожин И. Термодинамическая теория структуры, устойчивости и флуктуаций. — М.: Мир, 1973.]
24. *Sabo D., et al.*, to be published.
25. *Kimura M., Ohta T.*, Theoretical Aspects of Population Genetics, Princeton, New Jersey, Princeton Univ. Press, 1971.
26. *King J. L., Jukes T. N.*, Science, 164, 788 (1969).
27. *Kramer F. R. et al.*, J. Mol. Biol., 89, 719 (1974).
28. *Hoffmann G.*, Lecture at Meeting of the Senckenbergische Naturforscher Gesellschaft, April, 1974.
29. *Tyson J. J.*, In: Some Mathematical Questions in Biology (ed Levin S. A.) Providence, Rhode Island AMS Press, 1974.
30. *Shannon C. E., Weaver W.*, The Mathematical Theory of Communication, Urbana: Univ. of Illinois Press, 1949. [Имеется перевод: Шенион К. Э., в сб.: Работы по теории информации и кибернетике. — М.: ИЛ, 1963.]
31. *Brillouin L.*, Science and Information Theory, New York, Academic Press, 1963. [Имеется перевод: Бриллюэн Л. Наука и теория информации. — М.: Физматгиз, 1960.]
32. *Domingo E., Flavell H. A., Weismann Ch.*, Gene, 1, 3 (1976).
33. *Batschelet E., Domingo E., Weissmann Ch.*, *ibid.*, 1, 27 (1976).
34. *Weissmann Ch., Feix G., Stor H.*, Cold Spring Harbor Symp. Quant. Biol., 33, 83 (1968).
35. *Spiegelmann S.*, Lecture at the Symposium: Dynamics and Regulation of Evolving Systems, Schloss Elmau, May 1977.
36. *Hall E. W., Lehmann I. R.*, J. Mol. Biol., 36, 321 (1968).
37. *Battula N., Loeb L. A.*, J. Biol. Chem., 250, 4405 (1975).
38. *Chang L. M. S.*, *ibid.*, 248, 6983 (1973).
39. *Loeb L. A.*, In: The Enzymes, Vol. X, p. 173 (ed. P. D. Boyer), New York—London, Academic Press, 1974.
40. *Hopfield J. J.*, Proc. Nat. Acad. Sci., USA, 71, 4135 (1974).
41. *Englund P. T.*, J. Biol. Chem., 246, 5684 (1971).
42. *Bessmann M. J., et al.*, J. Mol. Biol., 88, 409 (1974).
43. *Jovin T. M.*, Ann. Rev. Biochem., 45, 889 (1976).
44. *Pörschke D.*, In: Chemical Relaxation in Molecular Biology, p. 191 (Pecht I, Rigler R., eds.), Heidelberg, Springer, 1977.
45. *Watson J. D.*, The Molecular Biology of the Gene, New York, Benjamin, 1970. [Имеется перевод: Уотсон Дж. Молекулярная биология гена. — М.: Мир, 1979.]
46. *Ladner J. E. et al.*, Proc. Nat. Acad. Sci. USA, 72, 4414 (1975).
47. *Fox S. W.* In: Protein Structure and Function, p. 126 (J. L. Fox, Z. Deyl, A. Blaser, eds.), New York, M. Dekker, 1976.

48. *Hirsch M. W., Smale S.*, Differential Equations, Dynamical Systems and Linear Algebra, New York, Academic Press, 1974.
49. *Glansdorff P., Prigogine I.*, Thermodynamic Theory of Structure, Stability and Fluctuations, London, Wiley Interscience, 1971. [Имеется перевод: Гленсдорф П., Пригожин И. Термодинамическая теория структуры, устойчивости и флуктуаций. — М.: Мир, 1973.]
50. *Grümm H. R.* (ed.), Analysis and Computation of Equilibria and Regions of Stability, IASA Conf. Proc., Vol. 8, Laxenburg, 1975.
51. *Coddington E. A., Levinson N.*, Theory of Ordinary Differential Equations, p. 321; New York, McGraw-Hill, 1955. [Имеется перевод: Коддингтон Э. А., Левинсон Н. Теория обыкновенных дифференциальных уравнений. — М.: ИЛ, 1958.]
52. *Volterra V.*, Mem. Acad. Lincei, 2, 31 (1926); *Lotka A. J.*, Elements of Mathematical Biology, New York, Dover, 1956.
53. *Eigen M. et al.*, forthcoming paper.
54. *Schuster P., Sigmund K., Wolff R.*, Bull. Math. Biol., 40, 743 (1978).
55. *Schuster P.*, Chemie in uns. Zeit, 6, 1 (1972); *Schuster P.* In: Biophysik, ein Lehrbuch, p. 688 (W. Hoppe et al., eds.), Berlin, Springer, 1977.
56. *Bhat R. K., Schneider F. W.*, Ber. Bunsenges. physik. Chem., 80, 1153 (1976).
57. *La Salle J., Lefschetz S.*, Stability by Lyapunov's Direct Method with Applications, New York, Academic Press, 1961.
58. *Marsden J. E., McCracken M.* (eds.), The Hopf Bifurcation and its Applications (Appl. Math. Sci., Vol. 19), New York, Springer, 1976. [Имеется перевод: Марсден Дж., Мак-Кракен М. Бифуркации Хопфа и их приложения. — М.: Мир, 1980.]
59. *Thom R.*, Stabilité Structurelle et Morphogenèse, New York, Benjamin, 1972.
60. *Woese C. R.*, Nature, 226, 817 (1970).
61. *Fuller W., Hodgson A.*, *ibid.*, 215, 817 (1967).
62. *Crick F. H. C.*, J. Mol. Biol., 38, 367 (1968).
63. *Miller S. L., Orgel L. E.*, The Origins of Life on Earth, Englewood Cliffs, N. J., Prentice Hall, 1973.
64. *Oró J., Kimball A. P.*, Biochem. Biophys. Res. Commun., 2, 407 (1960), *Oró J.*, Nature, 191, 1193 (1961).
65. *Lewis J. B., Doty P.*, *ibid.*, 225, 510 (1970).
66. *Uhlenbeck O. C., Baller J., Doty P.*, *ibid.*, 225, 508 (1970).
67. *Grosjean H., Söll D. G., Crothers D. M.*, J. Mol. Biol., 103, 499 (1976).
68. *Coutts S. M.*, Biochim. Biophys. Acta, 232, 94 (1971).
69. *Kvenvolden K. A. et al.*, Nature, 228, 923 (1970).
70. *Oró J. et al.*, *ibid.*, 230, 105 (1971); *Cromin J. R., Moore C. B.*, Science, 172, 1327 (1971).
71. *Rossmann M. G., Moras D., Olsen K. W.*, Nature, 250, 194 (1974).
72. *Walker G. W. R.*, Biosystems, 9, 139 (1977).
73. *Biebricher Ch., Eigen M., Luce R.*, in preparation,

74. Erhan S., Greller L. D., Rasco B., *Z. Naturforsch.*, **32c**, 413 (1977).
75. Dayhoff M. O., Atlas of Protein Sequence and Structure, Vol. 5, Suppl. 2, p. 271 (1976).
76. Jukes T. H., *Nature*, **246**, 22 (1973); Holmquist R., Jukes T. H., Pangburn S., *J. Mol. Biol.*, **78**, 91 (1973).
77. Fox G. E. et al., *Proc. Nat. Acad. Sci. USA*, **74**, 4537 (1977).
78. Uramer F. R. et al., *J. Mol. Biol.*, **89**, 719 (1974).
79. Chothia C., *ibid.*, **75**, 295 (1973).
80. Chou P. Y., Fasman G. D., *ibid.*, **115**, 135 (1977).
81. Levitt M., Chothia C., *Nature*, **261**, 552 (1976).
82. Gutte B., *J. Biol. Chem.*, **252**, 663 (1977).
83. Rigler R., personal communication; Rigler R., Ehrenberg M., Wintermeyer W. In: *Molecular Biology, Biochemistry and Biophysics*, Vol. 24, p. 219 (I. Pecht, R. Rigler, ed.), Berlin — Heidelberg — New York, Springer, 1977.
84. Olson T. et al., *J. Mol. Biol.*, **102**, 193 (1976).
85. Crothers D. M., Seno T., Söll D., G., *Proc. Nat. Acad. Sci. USA*, **69**, 3063 (1972).
86. Biebricher Ch., Druminski M., *ibid.* (submitted).
87. Wahba A. J. et al., *J. Biol. Chem.*, **249**, 3314 (1974).
88. Blumenthal T., Lauders T. A., Weber K., *Proc. Nat. Acad. Sci. USA*, **69**, 1313 (1972).
89. Biebricher Ch., in preparation.
90. Meinhardt H., in: *Synergetics*, p. 214 (H. Haken ed.), Berlin — Heidelberg — New York, Springer, 1977.
91. Fox S. W., Dose K., *Molecular Evolution and the Origin of Life*, San Francisco, Freeman, 1972. [Имеется перевод: Фокс С., Дозе К. Молекулярная эволюция и возникновение жизни. — М.: Мир, 1975.]
92. Dose K., Rauchfuss H., *Chemische Evolution und der Ursprung lebender Systeme*; Stuttgart, Wissensch. Verlagsges., 1975.
93. Onsager L. In: *Quantum Statistical Mechanics in Natural Sciences*, p. 1. (B. Kursunoglu ed), New York — London, Plenum Press, 1973.
94. Lasaga A. C., Holland H. D., Dwyer M. O., *Science*, **174**, 53 (1971).
95. Schmidt E., *Z. Ges. Eis- u. Kälteindustrie*, **44**, 163 (1937).
96. Kuhn H. In: *Synergetics*, p. 200 (H. Haken, ed.), Berlin — Heidelberg — New York, Springer, 1977.
97. Plato, Timaios. [Имеется перевод: Платон. Сочинения в трех томах, т. 3, часть I, Тимей (с. 455). — М.: Мысль, 1971.]
98. Hofbauer J., Schuster K., Sigmund K., Wolff R., *SIAM J. Appl. Math. C*, in press.
99. Urbanke W., Maass G., to be published.

ПРЕДМЕТНЫЙ

- Адаптор см. РНК транспортная
- Аминокислоты, анализ в метеоритах 213, 214
- комплексы с ионами металлов 209, 239, 240
- распространенности первичные 208—214
- Антикодон 186
- Аттрактор 103
- странный 103
- A_i 34, 91
- Бассейн аттрактора 103, 110
- Бифуркация Хопфа 112, 130, 158, 159, 161, 162
- Векторное поле 101, 111, 173, 174, 176
- Генетический груз 54
- Гиперциклы, взаимосвязи 172—175
- второй степени 24, 25, 126
- вымирание 143—145
- каталитические 18, 24, 27
- классификация 127—130
- компаунд 128, 130, 134, 135
- необходимость возникновения 77—79
- неполные (broken) 143, 144
- с диффузией 248—252
- — трансляцией 151
- элементарные 128—124
- Глобально инвариантное множество (подпространство, подсистема) 138
- G_i см. Функции чистого роста
- Дикий тип популяции 40, 44
- Дикого типа распределение 40, 41, 45, 61
- — последовательность 40, 43, 63
- Динамическая система 90
- — чистая 95
- — смешанная 95
- Дифференциальные уравнения отбора 34—36, 90—92, 96, 98
- Диффузия 216, 248
- ДНК, репликация 22, 23, 65—73, 78
- D_i , разложение молекулярного вида 34, 48
- Δ_i 91
- E_i см. Избыточная продуктивность
- Игровые модели 48, 49, 54—60, 83—90
- Избыточная продуктивность (E_i) 35
- Изинга модель 197
- Информационное содержание 50—54
- Источник 101—105
- Катастрофа ошибок 53, 54, 56, 81
- «Качания» (wobble, wobbling) 62
- в третьем положении кодона 204
- основания 202
- эффект 207
- «Качающиеся» взаимодействия 198, 208
- основания 203
- положения в мРНК 203

- Качества фактор (Q_i) 34, 51, 52, 61, 63, 64, 68, 69, 72, 78
 Качество копирования буквы среднее (\bar{q}_m) 41, 51, 52
 Квазивид 37—46, 50—54, 157—164, 240, 251
 Колебания 86, 87, 146—151, 157—164, 240, 251
 Компартиментация 83, 245, 246, 256
 Конкуренция между гиперциклами 163—167
 Концентрация суммарная (полная, общая) 106

 Ляпунова функция 139—142
 — — нестрогая (в слабом смысле) 141
 — — строгая (strict) 141

 Метаболизм 29, 31
 Михаэлиса — Ментен механизм 19, 46, 154
 Мутантная копия с многими ошибками 42, 43, 229
 — — — одной ошибкой 42, 43, 45, 229
 Мутации 15, 16, 32, 34, 35
 — нейтральные 43, 45, 49, 50

 Нормальные моды 109—111
 Нуклеозидтрифосфаты 23, 154

 Ограничение «постоянной общей организации» 35, 36, 96
 Оптимизационный принцип 17, 38, 39
 Особые точки, исследование 88, 100, 101
 — — карта 101—106
 — — классификация 104
 Ошибок темп 35, 61, 78, 198, 199

 Паразитная связь 167—172
 Полимеризация матричная 152

 Популяционная генетика 15, 16, 29, 50, 54
 Популяционные переменные (x_i) 34, 91
 — — нормированные (ξ_i) 99, 106
 Порог ошибок 53, 54, 78
 Поток (flow) индивидуальный (Φ_i) 34, 35
 — — ϕ_i 91, 92
 — суммарный (общий) (Φ_i) 35
 — — ϕ 95, 96
 Потоки (fluxes) обобщенные постоянные 95
 Превосходства параметр (σ_m) 41, 44, 53
 Предельный цикл 103, 106, 109, 110, 130, 146—151, 157, 160, 163
 Прокариотическая клетка 72, 78, 230, 232, 258
 Пространство параметров 91
 Проточный реактор 36, 151
 P, матрица показателей $P_{\lambda i}$ 127
 p_i , степень ведущего члена функции роста Γ_i 95
 $p_{\lambda i}$, показатели степени членов функции роста Γ_i 127

 Реакционная сеть циклическая 27
 Реакционный цикл 18
 — — каталистический 21—23
 Ребро особых точек 138, 143
 — текущее 138, 143, 144
 Рекомбинация генетическая 70—73
 Репарации механизм 67—73
 Репликационная вилка 66, 67
 Решение «раз и навсегда» 18, 97, 98, 165
 Рибонуклеаза 239
 Рибосома 179, 181, 232, 257
 РНК, модель цветка 76
 — репликация 21, 23, 26, 60—65, 78
 26, 60—65, 78
 — репликаза фагоспецифичная 26, 60—65, 78
 — самокопирование внутренних областей 227, 236

- Рибосома
 — синтез de novo 235—237
 — содержащий фаг Q_{β} 40, 41, 60—65, 78, 198, 243
 — — — размножение 26, 60—65
 — фага Q_{β} , мидвариант 62, 64, 76, 191, 236, 237
 Рост гиперболический 27, 93, 94, 97, 165
 — линейный 23, 46, 93, 94, 96, 97
 — неограниченный 92—95, 165
 — с ограничениями 95
 — экспоненциальный 23, 46, 93, 94

 Самовоспроизведение 22, 23
 Седло 101—105
 Селективная ценность (W_{ii}) 36, 38, 44, 45, 61, 84, 86, 90, 92
 Селекционные ограничения 15, 34—36, 48, 95—98
 Сепаратриса 101—103
 Серийные переносы 36, 61
 Силы обобщенные постоянные 95
 Симплекс концентраций 106—109
 — — внутренняя часть 138
 — — граница 138, 139
 Сингулярность кинетической кривой 93, 94
 Синтезные функции 183, 194, 239—243
 β -Слой 237, 238
 Собственные векторы 38, 110
 — значения 38, 44, 104, 110
 Степень полимеризации максимальная (v_{\max}) 41, 53, 55, 72, 78
 Сток 101—104
 Стопка оснований кооперативная 196
 Стохастическая теория 44, 48—50

 Таргетная функция 245
 Теория возмущений 44
 — катастроф 113
 Траектории 101, 109

 Уравновешивание внутреннее 98—100
 Устойчивость асимптотическая 130, 141

 Феноменологические уравнения 34, 90
 Фенотип 39, 58, 62, 80, 82, 118, 177, 244, 257
 Флуктуация 49, 103—105, 130, 143, 253
 Функция чистого роста (Γ_i) 92, 93, 96, 99, 111, 113, 127, 128
 — — — однородная 99, 100, 112, 127
 Φ_i , Φ_t , ϕ_i ϕ см. Поток

 Хиральность 15

 «Цель» (target) 58
 Цепи каталитические 118—124
 Цикл каталитический 22, 23
 — трикарбоновых кислот 20
 — углеродный 19

 Эволюционные эксперименты 35, 60—65, 215
 Эволюционный принцип для гиперциклов 219, 222—226
 Эволюция, исторический путь 15—18, 58
 — недарвиновская 45, 46, 165
 Экологические ниши 94, 245
 Экстрацистронные области 62
 Экстремальный принцип 38, 39
 Эукариотическая клетка 72, 73, 78

 Якоби матрица 110, 137

СОДЕРЖАНИЕ

| | |
|---|-----|
| Предисловие редакторов перевода | 5 |
| Предисловие | 8 |
| Часть А. Возникновение гиперцикла | 15 |
| I. Парадигма единства и разнообразия в эволюции | 15 |
| II. Что такое гиперцикл? | 18 |
| III. Дарвиновские системы | 29 |
| III. 1. Принцип естественного отбора | 29 |
| III. 2. Необходимые предпосылки возникновения дарвиновских систем | 30 |
| III. 3. Динамика отбора | 34 |
| III. 4. Концепция квазивида | 37 |
| III. 5. Реалистические аппроксимации | 44 |
| III. 6. Обобщения | 46 |
| III. 7. Информационное содержание квазивида | 50 |
| IV. Порог ошибок и эволюция | 54 |
| IV. 1. Моделирование на ЭВМ катастрофы ошибок | 54 |
| IV. 2. Экспериментальные исследования с РНК-содержащими фагами | 60 |
| IV. 3. Репликация ДНК | 65 |
| IV. 4. Первые репликативные единицы | 73 |
| IV. 5. Необходимость возникновения гиперциклов | 77 |
| Часть Б. Абстрактный гиперцикл | 80 |
| V. Конкретный пример | 80 |
| VI. Общая классификация динамических систем | 90 |
| VI. 1. Определения | 90 |
| VI. 2. Неограниченный рост | 92 |
| VI. 3. Ограниченный рост и отбор | 95 |
| VI. 4. Внутреннее уравнивание в растущих системах | 98 |
| VII. Исследование особых точек самоорганизующихся сетей реакций | 100 |
| VII. 1. Адекватный метод исследования | 100 |
| VII. 2. Топологические свойства | 101 |
| VII. 3. Адекватное пространство: симплекс концентраций | 106 |
| VII. 4. Исследование нормальных мод | 109 |
| VII. 5. Растущие системы | 111 |
| VII. 6. Анализ конкретных систем | 113 |

| | |
|--|-----|
| VII. 7. Исследование особых точек гиперциклов | 125 |
| VIII. Динамика элементарного гиперцикла | 137 |
| VIII. 1. Качественный анализ | 137 |
| VIII. 2. Численное интегрирование | 145 |
| IX. Гиперциклы с трансляцией | 151 |
| IX. 1. Идеальные граничные условия и общие упрощения | 151 |
| IX. 2. Кинетические уравнения | 152 |
| IX. 3. Численные решения | 156 |
| X. Сети гиперциклов | 163 |
| X. 1. Внутреннее уравнивание и конкуренция между гиперциклами | 163 |
| X. 2. Паразитная связь и каталитические сети | 167 |
| X. 3. Иерархия связей между гиперциклами | 172 |
| Выводы | 175 |
| Часть В. Реалистический гиперцикл | 179 |
| XI. Как могла зародиться трансляция? | 179 |
| XII. Логика первичного кодирования | 185 |
| XII. 1. Код RRY | 185 |
| XII. 2. Код RNY | 190 |
| XIII. Физика первичного кодирования | 193 |
| XIII. 1. Условия на старте | 193 |
| XIII. 2. Распространенность нуклеотидов | 195 |
| XIII. 3. Стабильность комплементарных структур | 196 |
| XIV. Код с GC-рамкой | 202 |
| XIV. 1. Первые два кодона | 202 |
| XIV. 2. «Апериодическая линейная GC-решетка» | 205 |
| XIV. 3. От GNC к RNY | 206 |
| XIV. 4. Первичный алфавит аминокислот | 208 |
| XV. Гиперциклическая организация раннего аппарата трансляции | 215 |
| XVI. Десять вопросов, касающихся наших самых ранних молекулярных предков и тех следов, которые они оставили в биосинтетическом аппарате современных клеток | 226 |
| XVI. 1. Один РНК-предшественник? | 226 |
| XVI. 2. Что означает селективное преимущество для молекулы? | 227 |
| XVI. 3. Почему нужна гиперциклическая организация отдельных мутантных генов, а не один постоянно растущий геном? | 228 |
| XVI. 4. Необходимы ли при старте тРНК? | 230 |
| XVI. 5. Дают ли современные тРНК какие-либо указания на свое происхождение? | 232 |
| XVI. 6. Как в информационной РНК могли возникнуть кодонные паттерны без запятой? | 235 |
| XVI. 7. Какими были первые функционально активные белки? | 237 |
| XVI. 8. Необходимы ли при старте синтетазы? | 239 |
| XVI. 9. Какими были первые ферменты? | 241 |

| | |
|--|-----|
| XVI. 10. Почему же в конечном счете — клетки с объединенными геномами? | 243 |
| XVII. Реалистические граничные условия | 247 |
| XVIII. Непрерывность эволюции | 254 |
| Литература | 261 |
| Предметный указатель | 265 |

УВАЖАЕМЫЙ ЧИТАТЕЛЬ!

Ваши замечания о содержании книги, ее оформлении, качестве перевода и другие просим присылать по адресу:
129820, Москва, И-110, ГСП
1-й Рижский пер., д. 2,
издательство «Мир»

MANFRED EIGEN

SELFORGANIZATION OF MATTER
AND THE EVOLUTION
OF BIOLOGICAL MACROMOLECULE

Die Naturwissenschaften
58. Jahrgang, Oktober 1971,
Heft 10

Springer-Verlag

Berlin • Heidelberg • New York

М. Эйген | Самоорганизация
материи
и эволюция
биологических
макромолекул

Перевод с английского
В. М. АНДРЕЕВА

Под редакцией и с предисловием
чл.-корр. АН СССР М. В. ВОЛЬКЕН



ИЗДАТЕЛЬСТВО «МИР»
Москва 1973

ОГЛАВЛЕНИЕ

| | |
|--|------------|
| Предисловие к русскому изданию | 5 |
| Список литературы | 9 |
| Глава I. Введение | 10 |
| § I. 1. «Причина и следствие» | 10 |
| § I. 2. Предпосылки самоорганизации | 13 |
| Глава II. Феноменологическая теория отбора | 36 |
| § II. 1. Понятие «информации» | 36 |
| § II. 2. Феноменологические уравнения | 39 |
| § II. 3. Селекционное ограничение | 47 |
| § II. 4. Селекционное равновесие | 55 |
| § II. 5. Фактор качества и распределение ошибок | 59 |
| § II. 6. Кинетика отбора | 63 |
| Глава III. Стохастический подход к отбору | 75 |
| § III. 1. Ограничения детерминистической теории отбора | 75 |
| § III. 2. Флуктуации вокруг равновесных состояний | 76 |
| § III. 3. Флуктуации в стационарном состоянии | 78 |
| § III. 4. Стохастические модели как цепи Маркова | 84 |
| § III. 5. Количественное рассмотрение трех прототипов отбора | 88 |
| Глава IV. Самоорганизация, основанная на комплементарном узнавании: нуклеиновые кислоты | 96 |
| § IV. 1. Истинное «самоинструктирование» | 96 |
| § IV. 2. Комплементарное инструктирование и отбор (теория) | 102 |
| § IV. 3. Комплементарное узнавание оснований (экспериментальные данные) | 108 |
| Глава V. Самоорганизация через циклический катализ: белки | 124 |
| § V. 1. Узнавание и ферментативный катализ | 124 |
| § V. 2. Самоорганизующиеся циклы ферментов (теория) | 128 |
| § V. 3. Могут ли белки воспроизводить себя? | 137 |

| | |
|---|-----|
| Глава VI. Самоупорядочение через закодированные каталитические функции | 143 |
| § VI. 1. Требования кооперации между нуклеиновыми кислотами и белками | 143 |
| § VI. 2. Самовоспроизводящийся гиперцикл | 145 |
| § VI. 3. О возникновении кода | 160 |
| Глава VII. Экспериментальное исследование эволюции | 172 |
| § VII. 1. Система Q β -репликазы | 172 |
| § VII. 2. Дарвиновская эволюция в пробирке | 174 |
| § VII. 3. Количественные исследования отбора | 178 |
| § VII. 4. Эксперименты типа «минус один» | 183 |
| Глава VIII. Заключение | 185 |
| § VIII. 1. Пределы теории | 185 |
| § VIII. 2. Концепция «ценности» | 185 |
| § VIII. 3. «Диссипация» и «возникновение информации» | 189 |
| § VIII. 4. Принципы отбора и эволюции | 193 |
| § VIII. 5. «Недетерминированная», но «неизбежная» | 197 |
| § VIII. 6. Можно ли объяснить жизнь на основе современных физических концепций? | 201 |
| Глава IX. Резюме | 203 |
| Список литературы | 219 |

М. ЭЙГЕН

Самоорганизация материи и эволюция биологических макромолекул

Редактор *Н. О. Фомина*
Художник *В. Е. Карпов*
Художественный редактор *Ю. Я. Максимов*
Технический редактор *Е. С. Герасимова*
Корректор *О. Ф. Иванова*

Сдано в набор 5/IV 1973 г. Подписано к печати 17/VIII 1973 г. Бумага тип. № 3
84×108¹/₃₂ = 3,38 бум. л. 11,34 усл. печ. л. Уч.-изд. л. 10,67
Изд. № 4/6856. Цена 76 коп. Зак. 605.

ИЗДАТЕЛЬСТВО «МИР», Москва, 1-й Рижский пер., 2

Ордена Трудового Красного Знамени
Ленинградская типография № 2 имени Евгении Соколовой
Союзполиграфпрома при Государственном комитете Совета
Министров СССР по делам издательств, полиграфии и книжной торговли
г. Ленинград, Л-52, Измайловский проспект, 29

В книге лауреата Нобелевской премии М. Эйгена излагается созданная им теория эволюции на молекулярном уровне. Эйген впервые строго показал возможность возникновения функциональных биологических макромолекул из хаотической смеси малых молекул. Тем самым сделана попытка раскрыть физическую сущность естественного отбора, и достигнут новый этап в понимании происхождения жизни.

Предназначена для научных работников самых разных специальностей: биологов, химиков, физиков, математиков, философов, историков науки.

Редакция биологической литературы

ПРЕДИСЛОВИЕ К РУССКОМУ ИЗДАНИЮ

Возникновение и становление молекулярной биологии, уже решившей ряд важнейших вопросов естествознания, таких, как проблема молекулярного строения и свойств гена, как проблема генетического кода, определили развитие молекулярной биофизики. Предметом молекулярной биофизики являются строение и физические свойства биологически функциональных молекул, прежде всего белков и нуклеиновых кислот. Биофизика, будучи неотъемлемой частью физики, а не вспомогательной биологической дисциплиной, ставит физические задачи, относящиеся к живым телам, т. е. исходит при изучении живой природы из общих законов, описывающих поведение вещества и поля, и из атомно-молекулярных представлений. Решение этих задач может достигаться и не физическими методами, как это было, например, в случае генетического кода.

Однако физика встречается с трудностями при рассмотрении свойственных живой природе процессов развития — как онтогенеза, так и филогенеза. На первый взгляд между обычной физикой и биологией имеются непреодолимые противоречия. Изолированная физическая система эволюционирует к состоянию наименьшей упорядоченности, характеризуемому максимальной энтропией. Напротив, в биологической системе в процессе ее развития создается все больший порядок — дарвиновская эволюция прослеживается от одноклеточных микроорганизмов до *Homo sapiens* с его мыслящим мозгом. Биологическое описание развития имеет зачастую финалистический характер — задается вопрос «для чего?», в то время как обычная физика каузальна — она ставит вопрос «почему?», по каким причинам реализуется то или иное явление. Эти противоречия приводят

к неовиталистическим концепциям, согласно которым современная физика не в состоянии объяснить биологические явления, имеющие специальный «биотонный» характер [1, 2], и делается вывод о необходимости создания некой принципиально новой физики.

Противоречия эти кажущиеся. Физические законы также могут формулироваться финалистически, если они исходят из вариационных принципов. Достаточно сослаться на принцип Гамильтона, Мопертюи, Ферма, Ле Шателье, на правило Ленца и т. д. Противоречие между биологией и термодинамикой снимается, так как живое тело — всегда открытая, а не изолированная система. Но физические подходы к рассмотрению таких систем весьма нетривиальны и физические проблемы, относящиеся к процессам развития, еще далеки от своего решения.

Манфред Эйген, удостоенный в 1967 г. Нобелевской премии по химии за создание и теоретическое обоснование новых релаксационных методов химической кинетики, написал работу, принципиальное значение которой как раз и состоит в физическом рассмотрении биологических процессов — отбора и эволюции. Эйген ограничивается моделированием добиологической эволюции макромолекул, но развитые им идеи и методы имеют и более общее, принципиальное значение.

Естественно, что теоретико-физическое исследование биологических проблем начинается с феноменологического рассмотрения, основанного на неравновесной термодинамике и теории информации. Опираясь на работы Пригожина и его школы [3, 4], Эйген четко показывает, что рассмотрение процессов развития принципиально невозможно в рамках линейной термодинамики, т. е. вблизи состояния равновесия. Модели Эйгена относятся к стационарному состоянию, удаленному от равновесия.

Применению теории информации в биологии посвящено множество работ. В этих работах, как правило, применялось лишь понятие количества информации. Оно оказывается полезным: даже простой перевод биологических закономерностей на язык теории информации весьма информативен. Так, информационные аспекты эволюционной теории были раскрыты в очень

важных и содержательных работах И. И. Шмальгаузена [5, 6].

Однако постепенно становилось ясно, что понятие количества информации мало что дает. Эйген справедливо подчеркивает в этой связи, что, будучи комплементарным энтропии, количество информации только ее и характеризует. В монографии Аптера [7] отмечается, что для биологии существенно не количество информации, а программа последующего развития. О том же писал Моно [8], который ввел понятие «телеономии», выражающее план развития, наличествующий в зиготе, в эволюционирующей популяции, в биосфере. Впрочем, Моно игнорирует основные особенности живых открытых систем, утверждая, что отличие живого организма от кристалла сводится лишь к количеству информации. Тем самым он никак не раскрывает физическое и философское содержание телеономии, и предложенная им натурфилософская конструкция повисает в воздухе.

Эйген с полной ясностью показывает, что для биологии важна ценность информации, а не ее количество. Информация обретает ценность в реальном физико-химическом процессе, и эта ценность должна быть выражена в измеримых физических величинах. Эйген предлагает теорию отбора и эволюции макромолекул, основанную на ценности информации, определяемой как селективная ценность. Селективная ценность выражается через конкретные кинетические параметры.

Внимательный анализ проблемы, проведенный Эйгеном, показывает, что для обеспечения селекции и эволюции необходима автокаталитическая система, которая была бы действительно подобна реально существующим нуклеиново-белковым гиперциклам. Ни нуклеиновые кислоты, ни белки, взятые по отдельности, не могут подвергаться отбору и эволюционировать. Может показаться, что Эйген не открыл здесь ничего нового. Но в действительности раскрыто физическое содержание процессов редупликации и синтеза биологических информационных макромолекул.

Общий методологический смысл работы Эйгена состоит в убедительной демонстрации неограниченных возможностей физики в трактовке процессов

добиологического и биологического развития. Тем самым сделан важный шаг на пути проникновения в наиболее актуальную область современной биологии. После установления основных закономерностей молекулярной биологии естествознание обращается к проблемам дифференцировки, эмбриогенеза и канцерогенеза. Их изучение только начинается. Тем большее значение приобретает труд Эйгена, в котором намечаются пути теоретико-физического исследования этих проблем, хотя сам автор об этом и не говорит.

Эйген с полным основанием приходит к важному выводу о достаточности современной физики для изучения биологических проблем. Нужна не новая физика, но новые применения физических понятий. При этом нет надобности вводить особые физические понятия специально для трактовки биологических явлений — так, как это сделал Гудвин [9].

Некоторые из положений, выдвинутых Эйгеном, дискуссионны. Необходимо, впрочем, сразу отвергнуть критику теории Эйгена, основанную на том, что его модель нереальна. Эйген справедливо указывает, что цикл Карно также представляет собой идеализацию. Дискуссионна промежуточная аргументация в связи с тем, что селективная ценность не всегда совпадает с ценностью информации (см. [10]). Это никак не умаляет значения работы Эйгена, основные выводы которой несомненно верны; это означает только, что теория Эйгена требует дальнейшего анализа и развития. Теория Эйгена — живая теория, которая может быть далее применена к исследованию явлений синтеза антител, клеточной дифференцировки, морфогенеза и т. д. Эйген пока что ограничился сопоставлением своей теории с «эволюцией в пробирке», реализованной в опытах Спигелмана.

Работа Эйгена необычайно богата идеями. Она целиком основывается на конкретных достижениях молекулярной биологии, ряд положений которой представлен в книге с исключительной ясностью и четкостью.

Уместно сравнить книгу Эйгена с классической книгой Шредингера «Что такое жизнь с точки зрения физики?» [11]. Труд Шредингера сыграл громадную стимулирующую роль в развитии молекулярной биологии

и биофизики. Но он был написан более четверти века назад. Сейчас Эйген имеет возможность опереться на молекулярную биологию и популяционную генетику, на неравновесную термодинамику и теорию информации. Его работа полностью опровергает любые попытки построения неовиталистических концепций. Можно не сомневаться в том, что в дальнейшем развитии биологии и биофизики работа Эйгена сыграет не менее важную роль, чем сыграла в свое время книга Шредингера.

Сокращенное и более популярное изложение работы Эйгена публикуется в «Успехах физических наук» [12].

Недавно появилась обстоятельная работа Куна «Самоорганизация молекулярных систем и эволюция генетического аппарата» [13], идейно связанная с трудом Эйгена. В этой работе дается качественная модель добиологической эволюции и начальных стадий филогенеза и приводятся разумные оценки времени, необходимого для этих процессов. Работа Куна показывает, что случайное возникновение сравнительно небольших макромолекул РНК могло необходимым образом привести к возникновению жизни в обозримые сроки.

М. Волькенштейн

СПИСОК ЛИТЕРАТУРЫ

1. *Elsasser W.*, The Physical Foundation of Biology, Pergamon Press, L., 1958.
2. *Вигнер Е.*, Этюды о симметрии, изд-во «Мир», М., 1971, стр. 160.
3. *Glansdorff P., Prigogine I.*, Thermodynamic Theory of Structure, Stability and Fluctuations, Interscience — Wiley, L., N. Y., 1971.
4. *Пригожин И., Николис Г.*, Усп. физ. наук, 109, вып. 3 (1973):
5. *Шмальгаузен И. И.*, Проблемы дарвинизма, изд-во «Наука», Л., 1969.
6. *Шмальгаузен И. И.*, Кибернетические вопросы биологии, изд-во «Наука», Сибирское отд., Новосибирск, 1968.
7. *Антер М.*, Кибернетика и развитие, изд-во «Мир», М., 1970.
8. *Monod J.*, Le hasard et la nécessité, Seuil, Paris, 1971.
9. *Гудвин Б.*, Временная организация клетки, изд-во «Мир», М., 1966.
10. *Волькенштейн М. В.*, Усп. физ. наук, 109, вып. 3 (1973).
11. *Шредингер Э.*, Что такое жизнь с точки зрения физики?, ИЛ, М., 1947; Атомиздат, 1972.
12. *Эйген М.*, Усп. физ. наук, 109, вып. 3 (1973).
13. *Kuhn H.*, Angew. Chemie, 84, No. 18, 838 (1972).

ВВЕДЕНИЕ

§ I. 1. «Причина и следствие»

Вопрос о возникновении жизни часто представляется как вопрос о «причине и следствии». Физические теории макроскопических процессов обычно содержат ответы на подобные вопросы, даже если соотношению между «причиной» и «следствием» дается статистическая интерпретация. Именно природой этого вопроса в значительной мере обусловлено мнение многих ученых, полагающих, что современная физика не дает убедительного объяснения существованию жизни, которая даже в своих простейших формах, по-видимому, всегда связана со сложными макроскопическими (т. е. мульти-молекулярными) системами, такими, как живая клетка.

Поразительные открытия молекулярной биологии привели к тому, что вышеупомянутый вопрос часто формулируют так: *что возникло раньше: белок или нуклеиновая кислота?* — современный вариант старой проблемы «курицы и яйца». В этой формулировке слово «раньше» обычно используют для определения причинной, а не временной связи, а слова «белок» и «нуклеиновая кислота» можно заменить словами «функция» и «информация». В такой форме этот вопрос, если отнести его к сложным взаимоотношениям нуклеиновых кислот и белков в современных живых клетках, некорректен и ведет к абсурду, потому что не может быть организованной «функции», если нет «информации», а эта «информация» приобретает смысл только через «функцию», которую она кодирует.

Такую систему можно сравнить с замкнутой петлей. Хотя очевидно, что линия, из которой образовалась петля, должна была где-то начаться, начальная точка теряет свое значение, как только круг замкнется. Взаи-

моотношения нуклеиновых кислот и белков в современных условиях соответствуют сложной иерархии «замкнутых петель» (рис. 1 и табл. 1).

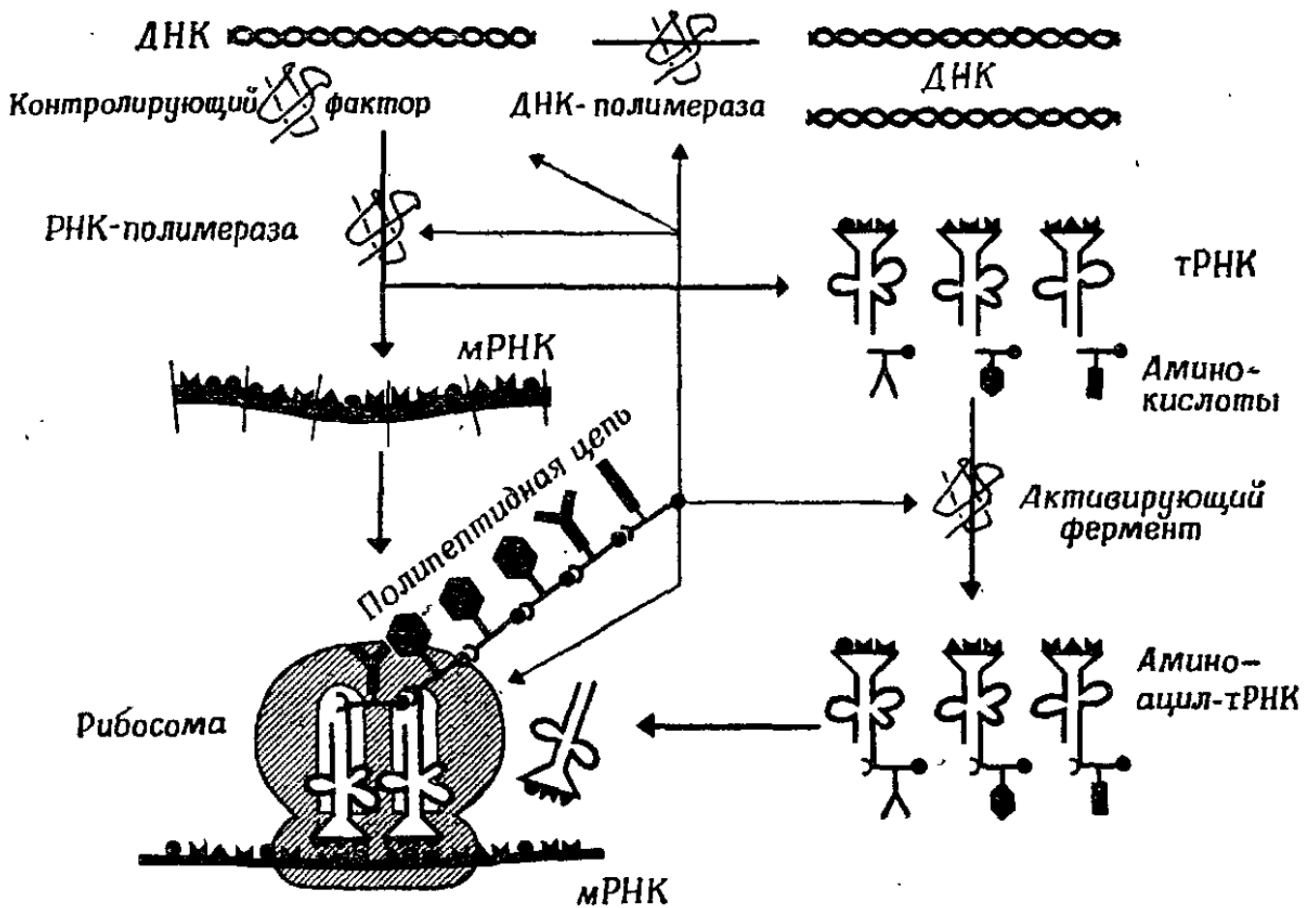


Рис. 1. Биосинтетический цикл самовоспроизведения клетки.

Таблица 1

**Нуклеиновые кислоты и белки тесно взаимодействуют
в цикле своего воспроизведения**

Важными функциональными связующими звеньями являются:

1. ДНК и ДНК-полимераза [1—3].

ДНК — стабильный источник информации — копируется с помощью фермента ДНК-полимеразы. Сначала считали, что функциональной единицей, катализирующей полимеризацию ДНК, является фермент Корнберга; но есть указания [3], что на самом деле репликацию ДНК в клетке ведет другой белковый комплекс, с большим молекулярным весом, который, по-видимому, прикреплен к мембране, тогда как фермент Корнберга выполняет репаративные функции.

2. Матричная, или информационная, РНК (мРНК) и РНК-полимераза [4].

Информация, содержащая инструкцию для синтеза белка, транскрибируется с ДНК на одноцепочечную, легче «читаемую» форму — матричную РНК. Это происходит с участием фермента

Продолжение табл.

РНК-полимеразы; РНК-полимераза из *E. coli* имеет молекулярный вес $\sim 5 \cdot 10^5$ и состоит из нескольких субъединиц, причем некоторые субъединицы являются специфичными регулируемыми факторами.

3. Транспортные РНК (тРНК) и аминоацилсинтетазы (активирующие ферменты) [5—8].

Для узнавания различных аминокислот соответствующими адапторами — молекулами транспортных РНК — требуется «второй код», который реализуется набором центров узнавания аминоацилсинтаз. Транспортные РНК — это молекулы, имеющие сравнительно малый молекулярный вес (около 60—80 нуклеотидов, ср. с рис. 2), строение которых (последовательность нуклеотидов) в ряде случаев известно. Аминокислота присоединяется к своей специфичной адапторной тРНК в результате энергетически сопряженной реакции, которая катализируется специфичным активирующим ферментом — аминоацилсинтазой. Эти ферменты представляют собой важное связующее звено между кодами нуклеиновых кислот и белков [8]. Их исследованием занимается несколько лабораторий.

4. Рибосомная РНК и белки [9, 10].

Синтез белка происходит на рибосоме, которая является комплексом РНК и белковых субъединиц с суммарным молекулярным весом около $2,7 \cdot 10^6$. Рибосома легко расщепляется на два фрагмента, которые седиментируют как 50S- и 30S-частицы. Меньший фрагмент содержит центр связывания мРНК, больший — каталитический центр для образования пептидной связи. Обе субчастицы участвуют в связывании аминоацил-тРНК, зависящем от матрицы мРНК. Недавно удалось разобрать субчастицы на отдельные молекулы белков и РНК, охарактеризовать их и реконструировать из них субчастицы.

5. Оперон, оператор, промотор и репрессор [11—13].

Транскрипция — очень строго регулируемый процесс. Его регуляция осуществляется путем индукции и репрессии. В регуляции участвуют белковые субъединицы (например, σ -фактор), которые кооперируют с РНК-полимеразой или входят в ее состав. Хорошо изученным примером генетического контроля служит репрессия и дерепрессия лактозного оперона. Репрессор — это белок с молекулярным весом 150 000 (4 идентичные субъединицы), который взаимодействует со специфичным репрессорным участком на молекуле ДНК (содержащим 10—20 пар оснований). Дерепрессия происходит путем образования комплекса репрессора с (низкомолекулярным) индуктором. Реакционные механизмы детально изучены. (Более подробное исследование проблем молекулярной биологии см. [14 и 15].)

Для решения проблемы подобного рода взаимоотношений между причиной и следствием необходима теория самоорганизации, которую можно было бы приме-

нить к молекулярным системам, или, точнее, к некоторым особым молекулярным системам, находящимся в среде с определенными свойствами. Можно полагать, что такой процесс молекулярной самоорганизации складывается из многих случайных событий, не имеющих какого-либо заданного в виде инструкции функционального значения. В действительности же важно установить, сколь вероятно то, что такие случайные события могут влиять на свой источник и становиться, таким образом, причиной некоего усиленного действия («следствия»). При определенных внешних условиях такие многократные взаимодействия между причиной и следствием могут привести к возникновению макроскопической функциональной организации, обладающей столь высокой способностью к самовоспроизведению, отбору и эволюции, что система может вырваться из условий, в которых она возникла, и преобразовать среду на собственную пользу.

§ 1.2. Предпосылки самоорганизации

1.2.1. Эволюция должна начинаться со случайных событий. «В начале» — каков бы ни был точный смысл этого понятия — по-видимому, имелся *молекулярный хаос*, и в гигантском многообразии химических соединений не было никакой функциональной организации. Таким образом, самоорганизация материи, которую мы связываем с «возникновением жизни», должна была *начаться* со случайных¹ событий. Из этого утверждения нельзя, однако, сделать вывод, что какие-либо, пусть даже примитивные, организмы, существующие в настоящее время, могли возникнуть в результате случайной сборки.

Несколько лет назад Е. Вигнер [16] написал статью «Вероятность существования самовоспроизводящейся системы», в которой он в неявной форме высказал подобное предположение. Суть его рассуждений сводится

¹ Термин «случайный» относится, конечно, к отсутствию функциональной организации, а не к отсутствию физических (т. е. атомных, молекулярных или даже надмолекулярных) структур.

к следующему. Допустим, что состояние «живого существа», а также состояния веществ, которыми оно питается, и состояние продуктов его метаболизма полностью заданы в квантовомеханическом смысле, т. е. могут быть описаны векторами в гильбертовом пространстве. Тогда процесс воспроизведения, т. е. взаимодействие организма с питательной средой, приводящее к редупликации организма, описывается преобразованием, в которое входит унитарная «матрица столкновения» S . Если предположить, что S — случайная матрица, то оказывается, что число уравнений, задающих это преобразование, значительно превосходит число неизвестных (т. е. компонентов векторов); при этом избыток равен разности $N^2R - (N + R + NR)$, где N и R — любые большие числа. Фактически для любого реального случая N и R настолько велики, что, как правильно заключает Вигнер, «было бы чудом», если бы уравнения преобразования выполнялись для какого-то набора неизвестных (утверждение, которое остается справедливым также и в том случае, если живому организму соответствует много альтернативных состояний).

Однако все это рассуждение основано на предположении, что S действительно является существенно случайной матрицей и, следовательно, взаимодействие не «инструктировано», т. е. вероятность любого данного состояния ничтожно мала по сравнению с большим числом возможных состояний. Этот результат может привести только к заключению, что никакое сложное состояние материи, которое мы теперь называем живым, не может возникнуть в результате случайной сборки. Наличие инструктирования на молекулярном уровне, означающее, что матрица преобразования S должна иметь очень специальный вид, может потребовать приспособления статистической механики к описанию процессов отбора и эволюции, но это вовсе не означает, «что современные законы и понятия квантовой механики необходимо модифицировать, прежде чем их можно будет применить к проблеме жизни» [16]¹.

¹ В мои намерения не входило обсуждать здесь определенные трудности применения квантовой механики для описания «макроско-

1.2.2. Инструкция требует информации. Я думаю, что именно Норберт Винер впервые предложил рассматривать информацию как новую физическую переменную. Мы имеем в своем распоряжении развитую теорию информации, начало которой положили работы Дж. фон Неймана, Н. Винера, К. Шеннона и других (литературу см. [17, 18]). Не можем ли мы воспользоваться этой теорией для решения интересующей нас проблемы самоинструирования?

Теория информации в современном понимании — это в основном *теория связи*. Она имеет дело с проблемами переработки, а не «генерации» информации. Она требует, чтобы информация была задана с самого начала во вполне определенной форме. Только после этого теория может сказать нам, как закодировать сообщение и как использовать избыточность; она может также сказать нам, как согласовать сообщение с возможностями перерабатывающих устройств, чтобы передать его через каналы с шумом, отфильтровать шум и восстановить сообщение с помощью устройств, сверяющих код; однако ей всегда требуется кто-то — обычно человек, чтобы с самого начала определить, что считать «информацией» и что «бессмыслицей».

Это обстоятельство отражено уже в самом определении *информации*. В простейшем случае, если мы хотим отобразить ситуацию с Z_1 событиями из Z_0 исходных событий, имеющих равную априорную вероятность, то для того, чтобы свести Z_0 возможностей к Z_1 , необходимо следующее количество информации:

$$I_1 = K \ln (Z_0/Z_1). \quad (I. 1)$$

Если мы хотим отобразить одно определенное событие, т. е. $Z_1 = 1$, нужное количество информации равно

$$I = K \ln Z_0. \quad (I. 2)$$

Это — определение. Оно выбрано таким образом, чтобы превратить совместные вероятности (которые

пических процессов». Эти трудности возникают и при рассмотрении других хорошо известных физических процессов, которые не связаны с феноменом жизни. На это обстоятельство обратил мое внимание Х. Примас.

всегда являются мультипликативными комбинациями вероятностей отдельных независимых событий) в аддитивные величины. Так, если Z_0 — число последовательностей длиной ν из λ различных единиц, то мы имеем

$$Z_0 = \lambda^\nu$$

и

$$I_{\lambda\nu} = K\nu \ln \lambda. \quad (I. 3)$$

Константа K выбрана в соответствии с двоичным кодом, т. е. $K = 1/\ln 2$. Следовательно, если $\lambda = 2$, то I равняется полному числу ν двоичных единиц («битов») в сообщении.

Таблица 2

Генетический код

| | | Второе положение | | | | |
|------------------|---|--|--------------------------|------------------------------|----------------------------|------------------|
| | | У | Ц | А | Г | |
| Первое положение | У | Фен Фен Лей Лей | Сер Сер Сер Сер | Тир Тир Терм. Терм. | Цис Цис Терм. Три | У Ц А Г |
| | Ц | Лей Лей Лей Лей | Про Про Про Про | Гис Гис Гян Гли | Арг Арг Арг Арг | У Ц А Г |
| | А | Иле Иле Иле Мет ¹⁾ | Тре Тре Тре Тре | Асн Асн Лиз Лиз | Сер Сер Арг Арг | У Ц А Г |
| | Г | Вал Вал Вал Вал ¹⁾ | Ала Ала Ала Ала | Асп Асп Глу Глу | Гли Гли Гли Гли | У Ц А Г |

Третье положение

Продолжение табл. 2

Каждая аминокислота (обозначенная начальными буквами) кодируется триплетом, составленным из оснований У, Ц, А, Г. С первого взгляда заметна существенная вырожденность в каждой клетке таблицы, указывающая на меньшее значение третьей буквы в триплете. (Подробности о замещениях в третьем положении — см. гипотезу «виляний» (гипотеза неоднозначного соответствия) Крика [22].) Заметно также некоторое сходство аминокислот в одном и том же столбце, что говорит о важном значении второй буквы. Триплеты, содержащие только У и А, дают большее число функциональных аминокислот (Фен, Лей, Тир, Иле, Асп, Лиз плюс 1 сигнал), чем триплеты, содержащие только Г и Ц (Про, Арг, Ала, Гли). Все эти факты могут содержать определенную информацию о возникновении кода и должны учитываться при построении различных теорий. Хотя код, по всей вероятности, универсален, следует оговорить, что эти соответствия (особенно «сигналы») относятся к *E. coli*.

Сокращенные обозначения аминокислот:

| | |
|-----------------------------|---|
| Ала — аланин | Мет — метионин ¹⁾ (формилметионин) |
| Арг — аргинин | Про — пролин |
| Асп — аспарагин | Сер — серин |
| Асп — аспарагиновая кислота | Тре — треонин |
| Вал — валин ¹⁾ | Три — триптофан |
| Гис — гистидин | Тир — тирозин |
| Гли — глицин | Фен — фенилаланин |
| Глн — глутамин | Цис — цистеин |
| Глу — глутаминовая кислота | Терм. — сигнал терминации |
| Иле — изолейцин | |

¹⁾ Триплеты АУГ и ГУГ означают «начало цепи», т. е. формил-Мет, если они стоят в начале цистрона; внутри цистрона они означают Мет или Вал соответственно.

Итак, количество информации в последовательности из n нуклеотидов ($\lambda = 4$) составляет $2n$ битов. Трансляция нуклеотидного кода в аминокислотный ($\lambda = 20$) требует, чтобы единицами кодирования служили по меньшей мере триплеты. Кодирование триплетными допускает также некоторую избыточность и сигналы старта и терминации. Генетический код (табл. 2), который мы рассмотрим более подробно в гл. VI, в настоящее время полностью выяснен, в основном благодаря работам Г. Корана, Х. Маттеи, М. Ниренберга, С. Очоа и их сотрудников (см. [19—21]).

Табл. 3 очень ярко демонстрирует *гигантскую информационную емкость биологических макромолекул*. Вероятность воспроизводимого нахождения при любых

Из v единиц с базисом λ можно составить
всего $N_{\lambda v} = \lambda^v$ последовательностей:

| Примеры | λ | v | $N_{\lambda v}$ |
|---|-----------|-----|-------------------|
| Небольшие субъединицы природных белков (с мол. весом 12 000) | 20 | 100 | 10^{130} |
| Полипептиды, которые кодируются только с помощью А и У (см. табл. 2) | 6 | 100 | 10^{78} |
| Цепи ДНК, кодирующие 33 аминокислоты | 4 | 99 | 10^{60} |
| АУ-сополимеры, кодирующие 33 аминокислоты, которым соответствуют кодоны класса АУ | 2 | 99 | 10^{30} |
| Олигопептиды, содержащие любые 12 из 20 природных аминокислот | 20 | 12 | $4 \cdot 10^{15}$ |
| Олигопептиды, содержащие 20 аминокислот, которым соответствуют кодоны класса АУ | 6 | 20 | $4 \cdot 10^{15}$ |

Для сравнения: число молекул белка с мол. весом 10^4 :

- а) предполагая плотную упаковку
- во всей Вселенной 10^{103}
 - в слое толщиной 1 м, покрывающем всю поверхность Земли $2 \cdot 10^{40}$
 - в 1 м^3 $6 \cdot 10^{25}$
- б) в 10^{-3} М растворе (что соответствует «бульону» с достаточно большой вязкостью)
- в Мировом океане 10^{42}
 - в пруду ($100 \times 100 \text{ м}$, глубина 10 м) $6 \cdot 10^{28}$
 - в луже (1 л) $6 \cdot 10^{20}$

Напомним, что молекулярные веса большинства природных белков и нуклеиновых кислот значительно больше, чем в приведенных примерах. Отметим далее, что возраст Земли равен «всего лишь» 10^{17} с, так что, даже предположив быстрый круговорот молекул белков, нельзя было бы перебрать все возможные последовательности (время жизни и продолжительность сборки каждой молекулы безусловно больше 1 с). Приведены только порядки величин. Вселенную представляем в виде сферы с диаметром 10 млрд световых лет; «плотной упаковкой» считается заполнение пространства с плотностью 1 г/см^3 .

разумных условиях (т. е. в объемах, имеющих разумные размеры), любой *заданной* последовательности из набора случайных последовательностей практически близка к нулю даже для относительно коротких цепей. А между тем такие последовательности не могут содержать сколько-нибудь значительного количества информации.

С другой стороны, вся информация, содержащаяся в цепочках ДНК хромосом весьма сложного живого существа, может превышать 10^{10} битов, что соответствует выбору одной возможности из $10^3 \cdot 10^9$ таких возможностей. Делались попытки сопоставить подобные числа с информацией, содержащейся в тех структурах и функциях, которые кодирует ДНК. В дальнейшем будет показано, что это невозможно, если не учитывать среду и историю данного существа. Информация, накопленная в процессе эволюции, — это «оцененная» информация, и число битов мало что говорит о ее функциональном значении.

Приведенное выше *определение информации* требует некоторой модификации, если мы имеем множество единиц с различными априорными вероятностями $p_1, p_2, \dots, p_\lambda$, как, например, в случае аминокислот в полипептидной цепи, или для букв в различных языках. Здесь средняя информация на единицу равна, по К. Шеннону [23, 24],

$$h = -K \sum_{i=1}^{\lambda} p_i \ln p_i, \text{ причем } \sum_{i=1}^{\lambda} p_i = 1 \quad (I. 4)$$

(символ h выбран по аналогии с H -функцией Больцмана). Как мы видим, любое ограничивающее условие (например, различные избыточности) снижает среднее количество информации на единицу по сравнению со случаем равных априорных вероятностей. В английском языке, например, это ограничение снижает среднее количество информации с 4,76 до 4,03 бита на букву. (Другие ограничения, например наличие предпочтительных последовательностей букв или слов, будут дополнительно изменять это число.)

Аналогия между шенноновским понятием информации и больцмановской статистической интерпретацией

энтропии (уравнения (I.2) и (I.4) очевидна, и всегда признавалось, что эта аналогия не только формальна [25]. В частности, это подчеркивал Э. Шредингер в своей замечательной книге «Что такое жизнь?» [26]. Если энтропия характеризует степень «незнания», то любое уменьшение «незнания» эквивалентно увеличению «знания», или «информации». Эта комплементарность между информацией и энтропией ясно показывает ограниченную применимость классической теории информации к проблемам эволюции. Везде, где информация имеет определенный смысл, например в языке (по соглашению) или в биологии (после того как эволюция привела к фиксации кода), эта теория имеет многочисленные применения. Однако от нее мало пользы, пока информация не приобрела своего «полного смысла» или пока имеется много путей создания новой информации. Здесь нам нужна новая переменная — «ценность», которая характеризовала бы уровень эволюции. Трактовка упорядоченности и неупорядоченности, указывающая лишь на их комплементарность, — на чем Э. Шредингер в своей книге сделал, пожалуй, слишком сильный упор — в этом случае недостаточна. На неадекватность современной теории информации в применении к биологическим проблемам указывал Л. Бриллюэн [18] в своей превосходной монографии «Наука и теория информации».

Мы видим, что теория информации в известной степени комплементарна классической статистической механике, по крайней мере в том отношении, что понятия энтропии и информации описывают степень незнания и знания. Для теории эволюции этих понятий недостаточно. Требуется уточнить, что мы называем знанием или незнанием. Для того чтобы охарактеризовать степень самоорганизации функциональной упорядоченности и определить направление эволюции, необходимо ввести понятие «ценности».

I.2.3. Информация возникает или приобретает ценность посредством отбора. В этом утверждении содержится один из существенных принципов биологии — дарвиновский принцип естественного отбора. Сам Дарвин считал его характерным свойством живого: «Сохранение благоприятных индивидуальных различий и

изменений и уничтожение вредных я назвал Естественным отбором или Переживанием наиболее приспособленных» ([27], стр. 328).

Впрочем, Дарвин ничего не утверждал относительно физической основы этого принципа: «Возражение, что наука до сих пор не пролила света на гораздо более высокие задачи о сущности и начале жизни, не имеет значения» ([27], стр. 659).

В генетике популяций, особенно в фундаментальных трудах великих школ Р. Фишера [28], Дж. Холдейна [29] и С. Райта [30], дарвиновский принцип получил математическую формулировку. Концепция «ценности», которая стоит за этим принципом, с трудом поддается объективной физической интерпретации, что вызвало необходимость в новой интерпретации, которая будет рассмотрена более подробно в конце книги (гл. VIII, § 5). Если мы хотим уничтожить разрыв между физикой и биологией, то необходимо разобраться в том, что такое «отбор» на языке точных молекулярных понятий, которые в конечном счете могут описываться квантовомеханической теорией. Мы должны вывести дарвиновский принцип из тех свойств материи, которые нам известны.

Чтобы приблизить понятие отбора к молекулярным свойствам и пояснить его смысл, рассмотрим несколько игр с белками и нуклеиновыми кислотами.

Для введения случайности воспользуемся игральными костями: икосаэдрической — для белков и тетраэдрической — для нуклеиновых кислот. Каждая грань кости будет соответствовать одной из двадцати природных аминокислот или одному из четырех нуклеотидов. Цель игры состоит в том, чтобы, бросая кость, получить заданную последовательность из v (например, 100) единиц: для нахождения каждого члена последовательности бросают кость и выпавшую цифру ставят на соответствующее место; если получившаяся последовательность не совпадает с нужной, весь процесс повторяют с самого начала.

Без какого-либо дополнительного правила отбора эта игра покажется довольно однообразной. Кроме тех случаев, когда v относительно мало, она тянулась бы

почти бесконечно. Мы уже видели, что для молекул белка, состоящих из 100 аминокислотных остатков, имеется примерно 10^{130} различных аминокислотных последовательностей, и нам пришлось бы бросать кость соответствующее число раз, чтобы получить заданную последовательность. Это просто еще один пример, иллюстрирующий вывод Вигнера о том, что ни одна белковая молекула с заданной структурой (и функцией) не могла бы возникнуть в результате случайной сборки.

Очень простая *модификация правил* позволит нам кончить игру за относительно короткое время. Введем отбор, придавая каждому правильно заполненному месту некое «селективное преимущество». В предельном случае это означало бы, что правильно заполненные места дальше в игре не участвуют. Поскольку в любой случайной полипептидной последовательности (в среднем) $N/20$ положений заняты правильными буквами, мы видим, что теперь для получения нужной последовательности будет достаточно в среднем $20(N - N/20) = 19N$ бросаний (т. е. 1900 при $N = 100$). Флуктуации достаточно велики, чтобы у каждого игрока были равные шансы на выигрыш. Игра все еще остается довольно скучной — единственным источником азарта мог бы быть выигрыш, — но мы ясно видим действие жесткого отбора на процесс проб и ошибок. Для получения правильной последовательности требуется примерно в 20 раз большее число проб, нежели в случае полностью инструктированной сборки.

Конечно, природа играет в гораздо более изощренные игры. «Селективное преимущество» обычно не является свойством отдельного аминокислотного или нуклеотидного остатка, кроме того, оно не является независимым от состояния других положений в цепи. Эти взаимосвязи делают игру более интересной, но они требуют большего числа «ходов» и определенных стратегий. Все еще довольно простой пример такой «стратегической» игры, на этот раз с нуклеиновыми кислотами, представлен в табл. 4. Этот пример показывает, почему при сборке тРНК природа предпочла в качестве вторичной структуры конфигурацию «клеверного листа»

(рис. 2). В случае нуклеиновых кислот «преимущество» обычно каким-то образом связано с наличием или отсутствием комплементарных пар оснований. В метаболизме нуклеиновых кислот комплементарность сильно

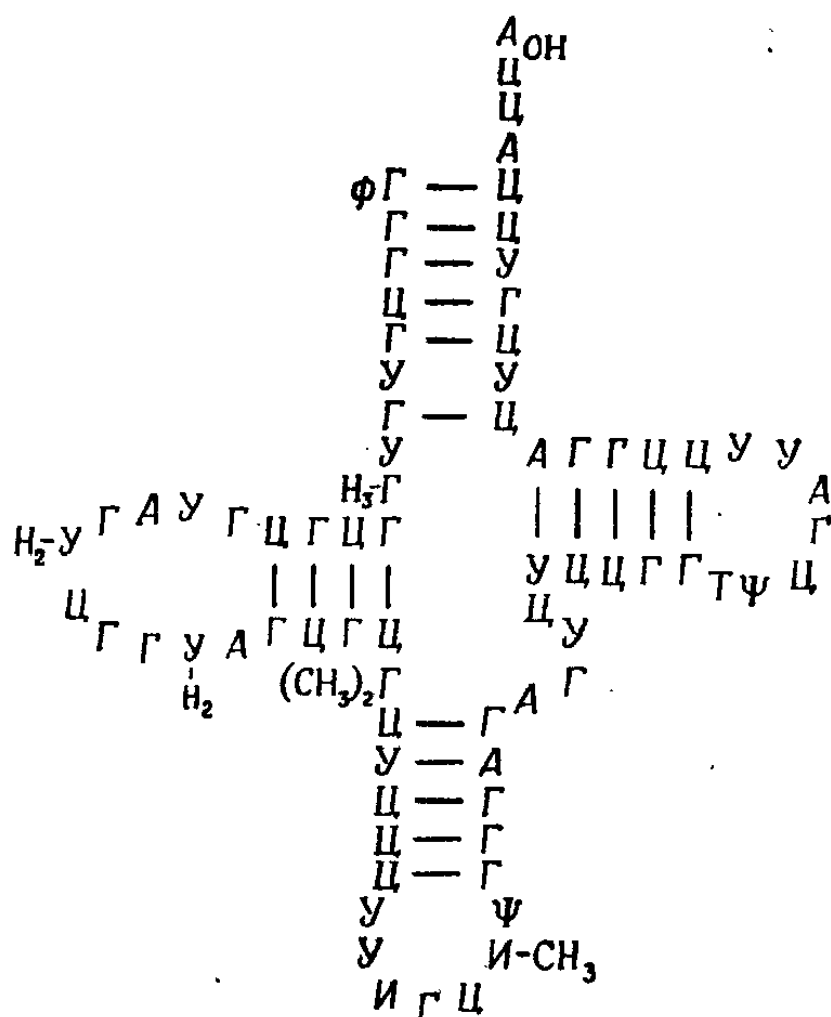


Рис. 2. «Клеверный лист» — модель тРНК.

Дана последовательность нуклеотидов аланинспецифичной тРНК из дрожжей (см. [5]). Необычные основания: Ψ — псевдоуридин, I — инозин, H_2-U — дигидроуридин, T — риботимидин, CH_3-G — 1-метилгуанозин, $(CH_3)_2-G$ — диметилгуанозин, CH_3-I — 1-метилюридин. Антикодон: 5' ИГЦ 3' (тРНК). Кодон (ГЦЦ) читается как 3' ЦЦГ 5' (мРНК).

влияет как на скорость синтеза, так и на скорость распада (см. гл. IV). В случае белков каталитические и регуляторные функции, зависящие от пространственной, т. е. третичной или четвертичной структуры, гораздо труднее связать с первичной структурой (последовательностью аминокислот). Все сведения, которыми мы располагаем относительно связи между структурой и функцией белков, получены в экспериментах лишь с очень немногими ферментами, детально изученными к настоящему времени (ср. [31]).

Таблица 4

Игра в тРНК или как делать клеверные листья

- Дано** Каждому игроку дается случайная последовательность из N единиц, принадлежащих к четырем классам А, У, Г, Ц, и тетраэдрическая кость, каждая грань которой соответствует одной из этих четырех букв.
- Цель** Бросая поочередно кость и заполняя определенное место в последовательности выпавшей буквой, каждый игрок пытается получить двухцепочечную структуру с максимальным числом пар АУ и ГЦ.
- Правила** Игра окончена, когда один из игроков объявит, что он получил «полную» структуру. Побеждает тот, кто наберет к этому времени максимальное число очков, причем за каждую ГЦ-пару засчитывается вдвое больше очков, чем за каждую АУ-пару. Пары можно образовывать только в том случае, если имеется непрерывная последовательность из не менее чем двух ГЦ-, или одной ГЦ- и двух АУ-, или же четырех АУ-пар (правило кооперативности). Соблюдая это условие, можно получать структуры любого типа (шпильки, скрепки, клеверные листья и т. п.). Для каждой петли в структуре следует оставлять неспаренными по меньшей мере 5 положений (ср. с моделью тРНК на рис. 2). Игроки бросают кость по очереди. Каждый игрок может бросить кость для любого положения по собственному выбору, однако он должен заранее объявить, для какого именно. Другое возможное ограничение: можно потребовать совпадения концов, т. е. потребовать, чтобы две концевые буквы образовали пару.

Замечание по процедуре. Может показаться удивительным, что победитель всегда получает структуру типа клеверного листа, похожую на известные вторичные структуры молекул тРНК (которые могут далее подвергнуться пространственной укладке).

«Секрет» игры состоит в том, чтобы с самого начала найти такую структуру, которая имеет максимальное число потенциальных пар оснований (включая некооперативные пары). Вероятность того, что в фиксированной структуре с n возможными парами окажется k и только k пар, можно рассчитать по формуле биномиального распределения

$$P_k = \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{n-k} \frac{n!}{k!(n-k)!}.$$

Среди всех фиксированных структур шпилька (имеющая только одну петлю) содержит максимальное число оснований, которые могут спариваться. Однако клеверный лист — структура более гибкая в том смысле, что он дает возможность испытать гораздо больше комбинаций, чем шпилька; это обусловлено тем, что отдельные листки можно сдвигать независимо от других, и мы с самого начала располагаем гораздо большим числом комбинаций. Это обстоятельство оказывается решающим. Вместе с тем при этом не может

Продолжение табл. 4

получиться слишком большое число листков, поскольку существует правило кооперативности. Для 80 нуклеотидов оптимум составляют 3—4 листка (+1 стебель). Для более длинных цепей он сдвигается к большим значениям.

Природа, видимо, играет в эту игру с давних времен. Необходимо отметить, что данное выше правило кооперативности соответствует константам стабильности, которые мы нашли для различных комбинаций оснований в олигонуклеотидах (см. гл. IV): для ГЦ-пары свободная энергия взаимодействия примерно вдвое больше, чем для АУ-пары.

Детали этой игры были разработаны Р. Винклером [51]. Количественные расчеты наиболее выгодных вторичных структур полинуклеотидных цепей были опубликованы ранее Дж. Фреско, Б. Олбертсом и П. Доти [52].

Главная цель этой работы — ввести в молекулярную динамику понятие «отбора» и связать его с уже известными нам молекулярными параметрами. Я сделал упор на относительно тривиальных примерах игр для того, чтобы показать, что правила отбора могут быть основаны на химических свойствах. Остается еще выяснить, как система могла бы использовать такие структурные преимущества и как из динамического поведения системы может следовать механизм оценивания.

1.2.4. Отбор происходит при наличии особых веществ и некоторых особых условий. *Какими свойствами должна обладать материя, чтобы могла начаться самоорганизация?*

Логически следует выделить несколько фаз эволюции, которые не вполне разделены во времени:

- 1) добиологическая «химическая» фаза;
- 2) фаза самоорганизации, приводящая к появлению реплицирующихся «индивидуумов»;
- 3) эволюция отдельных видов.

Долгое время биологи имели дело главным образом с третьей фазой, которая сама по себе состоит из многих стадий: от дифференциации и возникновения пола через появление нервных клеток, автономного регулирования и способов связи до самосознания и способности к логическому мышлению, которые представляют собой уникальные свойства человеческого разума. Я не хотел бы, чтобы у читателя создалось впечатление, что переходы между всеми этими стадиями менее

важны и менее ярко выражены, чем переходы между тремя стадиями, названными выше. Однако моя книга относится только ко второй фазе, к переходу от «неживого» к «живому». Как указывает Ф. Крик [32] в своем прекрасно написанном эссе «О молекулах и людях»: *«общеизвестно, как трудно определить понятие живого»*, потому что переход между живым и неживым не резкий. Таким образом, обращаясь к этому «этапу», мы должны начать с систем, которые заведомо не являются «живыми» (например, минералы), а кончить такими, из которых по крайней мере в результате развития может получиться то, что мы определенно называем «живым»: бактериальные клетки, растения, животные.

Самоорганизация зависит от определенных химических предпосылок, а также от специальных условий среды. Это не «просто» свойство материи. Предбиологическая фаза — это химия и как таковая «в принципе» описывается квантовой механикой. Необходимо, конечно, показать, что *условия на первобытной Земле благоприятствовали образованию нужных веществ*. Другая проблема состоит в том, что в отсутствие какой бы то ни было функциональной упорядоченности возникнет гораздо большее разнообразие, чем требуется в действительности. Сложность того, чему химия может дать объяснение, огромна, загляните только в «Бейльштейна»! Все, что *может* возникнуть в определенных условиях, *непрерывно* возникнет; первая фаза эволюции — это фаза дивергенции, и только посредством функциональной организации она может превратиться в фазу конвергентного воспроизведения и «оцененной» эволюции.

Детальное обсуждение проблем предбиологической «химической» эволюции выходит за рамки этой работы. Этим проблемам посвящена обширная литература, и существуют прекрасные обзоры, например монография А. И. Опарина [33] — одного из первых исследователей этой области, книга М. Кальвина [34], статьи С. Фокса [35], Поннамперумы [36] и других. Две недавние статьи Ф. Крика [38] и Л. Оргела [39] заслуживают особого упоминания, потому что в них дан критический обзор —

в свете экспериментальных результатов — условий возникновения второй фазы, рассмотрение которой является главной целью этой книги. Все авторы приходят к одному выводу, что основные строительные блоки биологических макромолекул — аминокислоты и высокоэнергетические фосфаты нуклеозидов (АТФ и его гомологи), а также многие другие «биохимические» соединения — могли образоваться, где это необходимо, и полимеризоваться в абиогенных условиях, например в восстановительной атмосфере и при наличии различных источников энергии. Говоря это, я не хочу создать впечатление, что все проблемы, касающиеся этой фазы эволюции, уже решены. Остаются многие вопросы, например, как возникла упорядоченная полимеризация, ведущая к образованию 3',5'-полинуклеотидов (эксперименты Г. Шрамма [40] и Л. Оргела [41] показали преобладание связей другого типа, например 2',5'-), или при каких условиях активированные аминокислоты будут полимеризоваться с образованием длинных полипептидных цепей, а не коротких олигомеров — эта проблема, вероятно, состоит в выборе подходящих каталитических условий (например, гетерогенного катализа), на что указывают недавние эксперименты А. Качальского и его группы [42]. Другая проблема — наличие достаточных количеств различных предшественников, необходимых для синтеза биологически активного материала. Решение этих проблем задаст работу еще целому поколению химиков, но это типично «химические» проблемы.

Сейчас для наших целей можно ограничиться предположением, что вещества типа высокоэнергетических фосфатов, активированных аминокислот и т. д., имелись в наличии и могли объединяться в макромолекулы, обладающие простыми каталитическими функциями, либо вследствие кооперативного действия определенных кислотных и основных групп, локализованных на боковых цепях аминокислот, либо вследствие усиления простых электронодонорных и электроноакцепторных свойств кофакторов, таких, как ионы металлов в различных валентных состояниях, либо (в случае нуклеиновых кислот) вследствие того, что они вели себя как простые матрицы.

Мы покажем, что наличие каталитических функций в сочетании с различными механизмами обратной связи, придающее системе способность к автокаталитическому росту, является одной из решающих предпосылок самоорганизации (см. гл. II).

Согласно классическому определению, катализатор увеличивает скорость прихода к равновесию, не сдвигая самого равновесия. Правда, известно, что большинство катализаторов влияет также на равновесие субстратов вследствие того, что свободная энергия их взаимодействия с субстратами не равна нулю. Однако при наличии истинного каталитического эффекта скорости прямой и обратной реакций изменяются в совершенно равной степени. Вследствие этого «автокаталитический рост» не может происходить в системе, которая находится в равновесии или вблизи равновесия. Это приводит к нашему второму вопросу: *в каких условиях среды может происходить самоорганизация?*

Фундаментальный ответ был дан Э. Шредингером [26], который писал: «Живая материя избегает прихода к равновесию». Равновесие (в изолированной системе) — это состояние с максимальной энтропией. Если мы удерживаем систему вдали от равновесия, мы должны постоянно компенсировать рост энтропии, т. е. «питать» систему свободной энергией или веществом, богатым энергией. Эта энергия используется для поддержания определенных реакций, которые не дают системе прийти в инертное или «мертвое» состояние равновесия. Справедливость этого утверждения очевидна, и заслуга Шредингера состоит в том, что он сформулировал его так ясно. Однако он понимал также недостаточность этого тезиса для объяснения того, «каким образом» упорядоченность поддерживается (другой) упорядоченностью и «каким образом» она сама возникла из неупорядоченности. Трудность возникает в связи с тем, что недостаточно разделить мир живого только на «упорядоченность» и «неупорядоченность».

Расширим еще немного рассмотрение термодинамического аспекта; так или иначе это понадобится для теории отбора, чтобы стоять с самого начала на твердой почве. Термодинамическая теория необратимых про-

цессов была развита И. Мейкснером [43], И. Пригожиным [44], С. де Гроотом [45] и другими. Она основана на соотношениях взаимности Онзагера [46], которые выражают микроскопическую обратимость, и, следовательно, применялись только вблизи от равновесия. Впоследствии П. Глансдорф и И. Пригожин [47] распространили эту теорию на системы, находящиеся вблизи стационарного состояния. Для нашего рассмотрения важна не столько сама энтропия, сколько ее производная по времени dS/dt , которая состоит из двух величин: потоков, направленных из внешней среды и во внешнюю среду, т. е. $d_e S/dt$, и производства энтропии внутри системы, т. е. $d_i S/dt$; последнюю величину мы обозначим σ . Для макроскопических систем (к которым применима формула Гиббса) σ можно выразить в виде суммы членов, каждый из которых является произведением потока J_i и обобщенной силы X_i (для неоднородных систем σ необходимо проинтегрировать по всему объему, но для простоты мы не будем здесь рассматривать такие системы). Сущность теории для систем, находящихся вблизи от равновесия (где выполняются линейные соотношения между потоками и силами), выражается тогда соотношением

$$\sigma = \sum_k J_k X_k \geq 0, \quad (I. 5)$$

т. е. энтропия всегда растет во времени для любой системы, находящейся вблизи от равновесия. В состоянии равновесия энтропия достигает максимума и σ обращается в нуль.

В табл. 5 суммированы применения теории к системам химических реакций. Каждая реакция характеризуется своей «степенью развития» ξ_i или ее производной

Таблица 5

Определения и формулы химической термодинамики

Сродство i -й реакции

$$A_i = - \sum_k \nu_{i k} \mu_k, \quad (1)$$

$$A_i = RT \left(\ln K_i - \ln \prod_k a_k^{\nu_{i k}} \right) \quad (2)$$

Стехиометрические коэффициенты:

$$\nu_{ik} \begin{cases} > 0 \text{ для продуктов реакции,} \\ < 0 \text{ для исходных реагентов.} \end{cases}$$

Пример: $4\text{O}_2 + \text{Nb} \rightleftharpoons \text{Nb}(\text{O}_2)_4$:

$$\nu_{\text{O}_2} = -4; \quad \nu_{\text{Nb}} = -1; \quad \nu_{\text{Nb}(\text{O}_2)_4} = +1.$$

Химический потенциал k -го вещества:

$$\mu_k = \mu_k^0 + RT \ln a_k, \quad (3)$$

μ_k^0 — стандартный химический потенциал; a_k — активность (ее можно заменить на концентрацию c_k , если раствор идеален).

Константа равновесия:

$$K_i = \prod_k \bar{a}_k^{\nu_{ik}} \quad (4)$$

\bar{a}_k относится к состоянию равновесия, где $A_i = 0$,

$$-RT \ln K_i = \sum_k \nu_{ik} \mu_k^0 \text{ [ср. с уравнением (2)].} \quad (5)$$

Степень развития (координата) реакции ξ_i :

$$d\xi_i = dn_k / \nu_{ik} \quad (n_k \text{ — число молей компонента } k). \quad (6)$$

При постоянных T и P :

$$dG = \sum_i A_i d\xi_i \quad (7)$$

(G — свободная энергия Гиббса; ξ_i сопряжена с A_i).

Скорость реакции:

$$V_i = d\xi_i / dt = \sum_k \varepsilon_{ik} \cdot A_k / T \quad (8)$$

(ε_{ik} — феноменологические коэффициенты).

Соотношения Онзагера:

$$\varepsilon_{ik} = \varepsilon_{ki} \quad (9)$$

Линеаризация:

$$V_i = \frac{1}{T} \sum_k \varepsilon_{ik} \sum_l \left(\frac{\partial A_k}{\partial \xi_l} \right)_{T, P} \delta \xi_l. \quad (10)$$

Матрицу (ε_{ik}) и тензор $\left(\frac{\partial A_k}{\partial \xi_l} \right)_{T, P}$ можно привести к диагональному виду. Трансформированные кинетические урав-

Продолжение табл. 5

нения принимают следующий вид:

$$V'_i = d\xi'_i/dt = \lambda_i \delta\xi'_i, \quad (11)$$

где $\delta\xi'_i$ — нормальная координата, $\lambda_i = -1/\tau_i$ — собственное значение, τ_i — время релаксации;

$$\delta\xi'_i = \delta\xi'_{i0} e^{-t/\tau_i} \quad (12)$$

— решение кинетического уравнения.

Производство энтропии:

$$\sigma = \frac{1}{T^2} \sum_{l, k} e_{l, k} A_l A_k \geq 0. \quad (13)$$

по времени (т. е. скоростью реакции V_i — скалярному потоку) и «средством» A_i (скалярной силой A_i/T). Приведенный в табл. 5 пример показывает, что вблизи от равновесия σ всегда можно выразить в виде квадратичной (положительно определенной) формы либо через величины ξ_i , либо через A_i . Линейное соотношение между потоками (скоростями реакций) и силами (средством) выполняется для любой системы реакций вблизи от равновесия, независимо от порядков этих реакций или от наличия каких-либо связей между различными реакциями. Решения системы линейных дифференциальных уравнений являются экспонентами с действительными и отрицательными аргументами; иными словами, любое отклонение концентраций от их равновесных значений будет экспоненциально спадать во времени:

$$(c_l - \bar{c}_l) = \sum_k a_{lk} e^{-t/\tau_k}. \quad (I. 6)$$

Приход к равновесию — это процесс «релаксации», характеризующийся спектром «времен релаксации» τ_k . На протяжении последних двадцати лет исследовано много таких релаксационных спектров для очень сложных систем реакций [48]. Периодические решения (экспоненты с комплексными показателями) или неустойчивые решения (экспоненты с положительным аргументом) не встречаются; этот факт следует из соотношений Онзагера (по которым матрица кинетических

коэффициентов симметрична) и знаков констант скоростей (из-за которых характеристическое уравнение является полиномом, все коэффициенты которого положительны) [49]. Различные корни — это отрицательные величины, обратные временам релаксации.

Аналогично уравнению (I.5), имеется критерий устойчивости, который выводится из классической формулы Эйнштейна для флуктуаций вблизи состояния равновесия [47]. Любая флуктуация вблизи устойчивого равновесия всегда дает отрицательное изменение энтропии

$$\delta_i S \leq 0. \quad (\text{I. 7})$$

Для флуктуации $\delta \xi_k$ вблизи устойчивого химического равновесия мы должны потребовать, чтобы

$$\delta_i S = \frac{1}{T} \sum A_k \delta \xi_k \leq 0 \quad (\text{I. 8})$$

или после разложения в ряд и диагонализации ($A_k = 0$ в равновесии)

$$\sum_k \frac{\partial A'_k}{\partial \xi'_k} (\delta \xi'_k)^2 \leq 0, \quad (\text{I. 9})$$

так как $\partial A'_k / \partial \xi'_k < 0$.

Тип решения, общий для всех систем вблизи истинного равновесия, означает, что отбор и эволюция не могут происходить в равновесных или почти равновесных системах даже при наличии нужных для этого веществ. В системах, находящихся вблизи от равновесия, автокатализ не будет приводить к росту, так как каталитическое ускорение в равной степени сказывается на прямой и на обратной реакциях.

Для систем в стационарном состоянии ограничений такого типа не существует. Как в равновесии, так и в стационарном состоянии концентрации постоянны, но эти состояния отличаются в другом аспекте. Соотношения взаимности Онзагера не выполняются в стационарном состоянии. Вследствие этого вблизи стационарного состояния могут возникнуть колебания.

И. Пригожин и П. Глансдорф [47] вывели соотношение, аналогичное уравнению (I.5), которое выполняется

вблизи стационарных состояний:

$$d_x \sigma = \sum_k J_k dX_k \leq 0, \quad (I.10)$$

т. е. изменение производства энтропии, обусловленное изменением сил X_k (например, изменение σ при постоянных потоках), всегда меньше нуля вблизи стационарного состояния и равно нулю в стационарном состоянии, т. е. при варьировании сил производство энтропии минимально в стационарном состоянии системы. (Необходимо отметить, что это соотношение неприменимо для $d_J \sigma$ или же для $d_x \sigma + d_J \sigma$.) Здесь снова имеет место критерий устойчивости, аналогичный уравнению (I.7). Согласно этому критерию, любая флуктуация вблизи устойчивого стационарного состояния будет проявляться в «избыточном производстве энтропии»

$$\delta_x \sigma = \sum_k \delta J_k \delta X_k \geq 0. \quad (I.11)$$

Эта величина положительна и вследствие (I.10), как и в случае производства энтропии вблизи равновесия, содержит только члены второго порядка; например, для химической системы

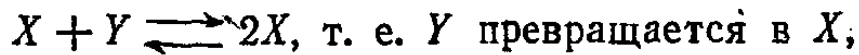
$$\delta_x \sigma = \frac{1}{T} \sum_k \delta V_k \delta A_k \geq 0. \quad (I.12)$$

Или: стационарное состояние *неустойчиво*, если возникает отрицательная флуктуация $\delta_x \sigma$. Как показывает табл. 6, такую нестабильность может проявить автокаталитическая система реакций (термин «автокаталитический» ниже будет интерпретироваться в самом широком смысле). Здесь изменение скорости, обусловленное флуктуацией концентрации, и соответствующее изменение сродства имеют противоположные знаки (заметим, что это верно только вдали от равновесия, где можно пренебречь скоростями «обратных» реакций). Мы покажем, что существование таких нестабильностей является предпосылкой для селективного роста и эволюции.

Пригожин и его сотрудники показали, что сочетание автокаталитического реакционного поведения с процессами

Таблица 6

Пример (простой) автокаталитической реакции



причем X является катализатором этой реакции ¹⁾

1. Вдали от равновесия

Скорость реакции:

$$V = \vec{k} [X] [Y]. \quad (1)$$

При постоянной концентрации $[Y]$: $\delta V \sim \delta [X]$.

Сродство:

$$A = RT \left(\ln K - \ln \frac{[X]}{[Y]} \right); \quad \delta A \sim -\delta [X]. \quad (2)$$

Избыточное производство энтропии:

$$T \delta_{x\sigma} \rightarrow \delta V \delta A \leq 0. \quad (3)$$

В отсутствие других процессов, которые обеспечивают стабильность, такая система не может прийти к устойчивому состоянию при постоянном подводе Y .

2. Вблизи от равновесия (флуктуация $\delta [X]$)

Скорость реакции

$$V = \vec{k} [X] [Y] - \overset{\leftarrow}{k} [X]^2 = \{ \vec{k} [Y] - \overset{\leftarrow}{k} [X] \} [X], \quad (4)$$

так как $\vec{k} [Y] \approx \overset{\leftarrow}{k} [X]$, член $\{ \}$ мал, а в равновесии равен нулю.

$$\delta V \approx -\overset{\leftarrow}{k} [X] \delta [X] \quad (5)$$

(пренебрегая членами второго порядка).

Сродство: как и выше.

Избыточное производство энтропии:

$$T \delta_{x\sigma} \geq 0 \quad (\text{устойчивое равновесие}). \quad (6)$$

¹⁾ Символом X здесь обозначено химическое соединение, и его необходимо отличать от обобщенных сил X_x . Индекс x при d или δ относится к изменению сил.

переноса может привести к своеобразному пространственному распределению реагирующих веществ. Такие распределения он назвал «диссипативными структурами», т. е. структурами, которые возникают в результате диссипации энергии, а не в результате действия консервативных молекулярных сил.

Пригожин [50] считает, что эти структуры были важны для создания функциональной упорядоченности в эволюции жизни. Они, несомненно, играли важную роль везде, где имелись благоприятные условия для их появления. Они могли оказывать некоторое влияние на ранний морфогенез, но я не думаю, что макроскопические пространственные структуры дают ключ к пониманию первых этапов биологической самоорганизации. Эти этапы могли происходить даже в бесструктурном «бульоне» при обязательном участии функциональных макромолекулярных структур, таких, как нуклеиновые кислоты и белки. Тип организации, который необходим вначале, — это не столько организация в физическом (т. е. геометрическом) пространстве, сколько функциональная упорядоченность в невероятном разнообразии химических соединений, возможно находящихся в гомогенной фазе. Нам нужна организация в другом «пространстве», которое может быть названо «информационным пространством». Эта упорядоченность тоже будет основана на принципе Пригожина — Глансдорфа, которому я придаю большое значение, но его использование потребует новых параметров и может вывести нас за пределы современной термодинамической теории.

ФЕНОМЕНОЛОГИЧЕСКАЯ ТЕОРИЯ ОТБОРА

§ II.1. Понятие «информации»

Упорядоченность в сложной системе реакций, происходящих с участием множества различных химических соединений, требует образования самовоспроизводящегося «функционального кода». Термин «функциональный код» означает наличие двух свойств: исполнительного и законодательного. Для проявления исполнительного свойства необходим аппарат, способный контролировать все реакции, идущие в системе; это может быть ансамбль взаимодействующих и саморегулирующихся катализаторов, состоящих предпочтительно из одного и того же материала, но имеющих практически неограниченные функциональные способности. Независимо от того, содержит ли первичная структура этого исполнительного аппарата инструкцию для его воспроизведения, или же она должна транслироваться из другого, законодательного источника, самоорганизация и дальнейшая эволюция коррелированного и воспроизводимого функционального поведения должна начинаться на уровне самовоспроизводящегося молекулярного кода.

Введем теперь понятие *носителей кода*. Тот факт, что мы знаем о существовании таких носителей кода, т. е. о нуклеиновых кислотах и белках, поможет нам создать полезную концепцию.

Введем множество, состоящее из N_v носителей информации:

$$i_{vk} \quad (k = 1, 2, \dots, N_v), \quad (\text{II. 1})$$

которые представляют собой последовательности из v единиц с базисом λ (например, $\lambda = 4$ для нуклеиновых кислот или $\lambda = 20$ для белков). Для простоты мы часто

будем иметь дело с таким однородным классом носителей информации, т. е. с последовательностями, имеющими одинаковую длину ν единиц.

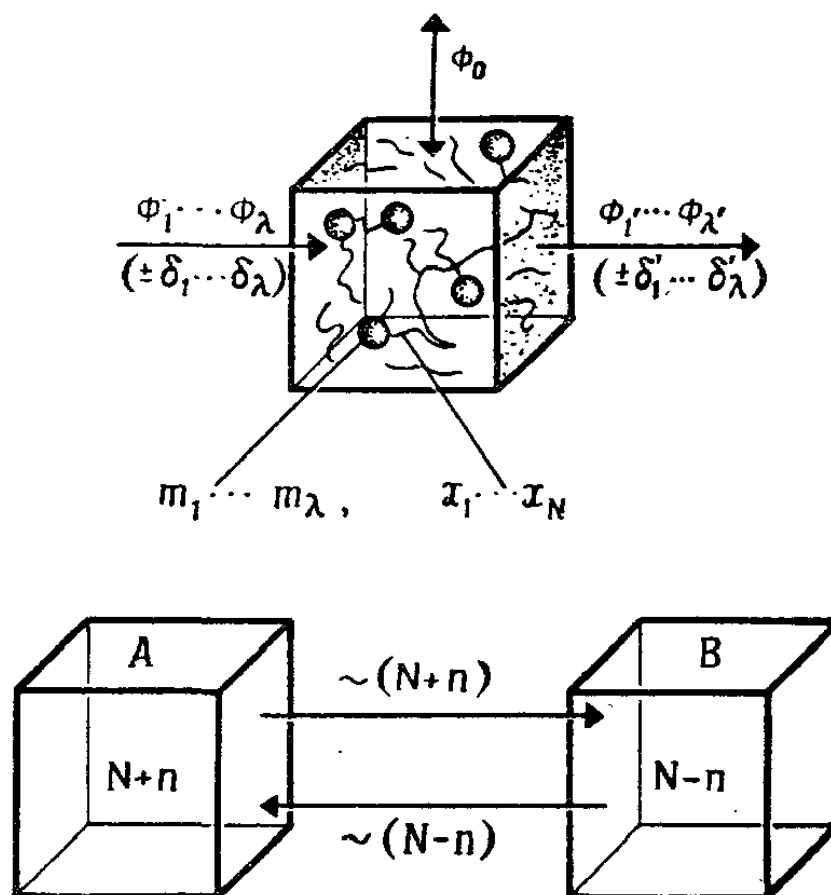


Рис. 3. «Информационный ящик».

Предполагается, что верхний ящик имеет полупроницаемые стенки, через которые могут втекать и вытекать высоко- и низкоэнергетические мономерные единицы. Внутри ящика происходят сборка и распад последовательностей (полимеров), представляющих собой «информацию». Как матричная сборка, так и распад могут катализироваться ферментами. Стационарное состояние может поддерживаться регулированием потоков $(\phi_1 \dots \phi_\lambda; \phi_1' \dots \phi_{\lambda'})$ или концентраций $(m_1 \dots m_\lambda; x_1 \dots x_N)$, которые контролируются потоком разбавления ϕ_0 . Для сравнения изображена равновесная система, распределенная между двумя ящиками A и B: $N_A = N_B$; $\pm n$ — флуктуации (обсужденные эренфестовской модели см. в § III. 2).

В классической теории информации «количество информации» в определенной последовательности $i_{\nu i}$ выражается числом битов:

$$I_\nu = \nu \ln \lambda / \ln 2 \text{ [битов]}. \quad (\text{II. 2})$$

Тогда N_ν представляет собой «структурную емкость» класса ν , т. е. число всех различных последовательностей с длиной ν и с базисом λ :

$$N_\nu = \lambda^\nu = 2^{I_\nu}. \quad (\text{II. 3})$$

Если последовательности могут иметь любую длину, ограниченную верхним пределом ν единиц, то мы получим

$$\sum_{\nu} N_{\nu} = \frac{\lambda^{\nu} - 1}{1 - \lambda^{-1}} \quad (\text{II. 4})$$

различных последовательностей. Это число может оказаться очень важным, если мы рассматриваем системы независимых конкурентов (с любой длиной от 1 до ν), где различные последовательности представляют собой определенные сообщения (с различными «селективными ценностями», см. ниже).

Система носителей информации может находиться внутри «ящика» конечного объема, имеющего проницаемые стенки, через которые могут втекать и вытекать высоко- и низкоэнергетические мономерные единицы (рис. 3). В ящике может находиться по $x_{\nu i}$ экземпляров каждого вида носителей информации (на единицу объема), и общая численность каждого класса составляет

$$n_{\nu} = \sum_{k=1}^{N_{\nu}} x_{\nu k}. \quad (\text{II. 5})$$

Для большинства из N_{ν} возможных носителей $i_{\nu k}$ их концентрации $x_{\nu k}$ будут равны нулю. Концентрации неорганизованных высокоэнергетических мономерных единиц 1, ..., λ в ящике равны m_1, \dots, m_{λ} , а их потоки в ящик — $\phi_1, \dots, \phi_{\lambda}$.

Число всех единиц — организованных и неорганизованных — составляет

$$M_0 = \sum_{k=1}^{\lambda} m_k + \sum_{\nu} \nu n_{\nu}. \quad (\text{II. 6})$$

«Степень организованности» D_0 в ящике, т. е. отношение числа организованных единиц к сумме организованных и неорганизованных единиц, равна

$$D_0 = \frac{\sum_{\nu} \nu n_{\nu}}{M_0}. \quad (\text{II. 7})$$

Полная численность данного класса в объеме V практически во всех реальных случаях крайне мала по сравнению с информационной емкостью этого класса N_v :

$$n_v \cdot V \ll N_v \quad (\text{II. 8})$$

(например, для $v = 100$, $\lambda = 20$ величина $n_v \cdot V$ может быть меньше или равна 10^{20} , тогда как $N_v = 10^{130}$; см. табл. 3). Из этого вытекает важное следствие, что при равновероятном распределении в отсутствие самоинструирования ожидаемое значение числа цепей с любой данной последовательностью практически равно нулю. Далее для тех последовательностей, которые образовались случайно, вероятность найти еще один *случайно* образовавшийся экземпляр с той же последовательностью снова практически равна нулю. Эта «ненасыщенность» информационной емкости очень важна для процедуры оптимизации в эволюции.

Наша следующая задача — придать определенные динамические свойства носителям информации и построить теорию отбора. Эта теория должна содержать вывод параметра, который выражает «селективное преимущество» через молекулярные свойства.

Для рассмотрения связанных систем можно представить себе «информационное пространство», которое задается множеством популяционных переменных x_{vi} и функциональными отношениями между ними. Для квазилинейных систем (истинно линейные системы неспособны «отбирать») это может привести к введению «нормальных координат», так же как в случае линейных релаксационных процессов.

§ II.2. Феноменологические уравнения

Рассмотрим класс носителей информации i_{vi} , каждый из которых имеется в числе x_{vi} экземпляров на единицу объема

$$0 \leq x_{vi} < n_v \quad (n_v \cdot V \ll N_v). \quad (\text{II. 9})$$

Отбор в дарвиновском смысле должен зависеть от динамических свойств системы, которые представлены

скоростями сборки и распада носителей информации. Следовательно, мы должны начать с кинетических уравнений для образования и распада макромолекул, которые описывали бы эволюционное поведение.

Для систем, далеких от равновесия, мы не можем ожидать простой линейной зависимости между потоками и силами. Таким образом, невыгодно начинать с формализма феноменологических уравнений, используемых в линейной области термодинамической теории необратимых процессов (т. е. с простых соотношений между скоростью и сродством). Поскольку было показано, что для отбора требуется совершенно особое реакционное поведение, может оказаться полезным применение терминологии химической кинетики, которая дает в более явном виде класс и порядок соответствующих реакций.

Введем три гипотезы, которые окажутся необходимыми условиями отбора:

1. Система должна быть открытой и далекой от внутреннего равновесия. Для того чтобы не дать системе прийти к равновесию, мы должны постоянно снабжать ее свободной энергией, например, в виде богатых энергией мономеров (таких, как АТФ и его аналоги по азотистым основаниям, или таких, как активированные аминокислоты). С другой стороны, распад может давать низкоэнергетические продукты. Таким образом, движущей силой обеих реакций — образования и распада носителей информации — является положительное сродство; для этих двух процессов нет соотношения типа «микроскопической обратимости», которое имелось бы вблизи от равновесия.

2. Скорость образования должна превышать скорость распада и иметь по меньшей мере тот же порядок по x_{vi} . Так как распад — это обычно реакция (по меньшей мере) первого порядка по x_{vi}^1 , то реакция образо-

¹ В реакциях распада, катализируемых ферментами, скорость реакции может стать независимой от концентрации субстрата, если фермент насыщен; однако такие реакции вовсе не делают неверным высказанное утверждение. Концентрация субстрата всегда проходит через «область ненулевого порядка» при своем росте, а конкуренция в условиях насыщения все еще зависит от популяционных переменных отдельных видов (пример дан в гл. VII).

вания должна быть «автокаталитической». Если бы скорость образования была меньшего порядка по x_{vi} , чем скорость распада, система не обладала бы внутренней способностью к росту, которая необходима для отбора против менее эффективных конкурентов. В таком случае все i_{vi} могли бы расти только до постоянного уровня, на котором скорости их образования сравнялись бы со скоростями распада. Подобная система несла бы в себе всю бесполезную информацию предшествующих мутаций, которая в конце концов блокировала бы дальнейшую эволюцию.

3. Вследствие условия $n_v \cdot V \ll N_v$ неинструктированным образованием какого-либо отдельного носителя информации можно полностью пренебречь.

В гл. IV—VI будет показано, что к автокаталитическому поведению способны многие различные классы реакций, но лишь некоторые определенные типы могут обеспечить эволюционное поведение. Одно очень важное уточнение по отношению к макромолекулярным носителям информации заключается в том, что нужно различать «самоинструктирующее» и «общее автокаталитическое» поведение. Процесс может быть автокаталитическим в том смысле, что продукт реакции стимулирует свое собственное образование — возможно, посредством какого-либо каталитического реакционного цикла, в котором участвует несколько промежуточных соединений (см. гл. V). Тогда скорость образования данного i_{vi} может быть пропорциональна не его концентрации x_{vi} , а концентрации $x_{vj} \neq i$ какого-то другого носителя информации, причем отношение x_{vj}/x_{vi} в конце концов достигает постоянного значения. В этом случае автокатализ является особым свойством какого-то отдельного ансамбля. Для того чтобы появились общие «матричные» свойства, необходимы «самоинструктирующиеся» носители информации. Это означает, что *любая* последовательность единиц инструктирует *свое собственное* воспроизведение. Если в процессе репродукции возникает ошибка, то *ошибочная копия* будет воспроизводиться дальше. Такое поведение типично для нуклеиновых кислот, тогда как «общее автокаталитическое» поведение может проявляться у белков.

Теперь можно написать в общем виде феноменологическое кинетическое уравнение для каждого носителя информации¹ ($\dot{x} \equiv dx/dt$):

$$\dot{x}_i = (\mathcal{F}_i - \mathcal{R}_i) x_i + \sum_{l \neq i} \varphi_{il} x_l. \quad (\text{II. 10})$$

Первый и второй члены в этом уравнении определяют скорость *самоинструментированного* образования носителя информации i_{vi} и скорость его удаления (посредством распада, разбавления и т. д.). Третий член включает в себя все другие каналы образования (не спонтанного), которое происходит в результате неточного копирования других последовательностей, близких к i_{vi} . Члены \mathcal{F}_i и \mathcal{R}_i — общие кинетические параметры, которые могут включать в себя несколько различных членов. Частная форма уравнения (II. 10) выбрана такой потому, что она выражает требование наличия «неотъемлемых» автокаталитических свойств, но не подразумевает, что реакция обязательно представляет собой реакцию первого порядка. Кинетический параметр \mathcal{F}_i , безусловно, является функцией концентраций мономеров (m_1, \dots, m_λ); кроме того, \mathcal{F}_i и \mathcal{R}_i могут зависеть еще от x_i или от x_k — популяционных переменных других видов.

Мы можем конкретизировать \mathcal{F}_i и \mathcal{R}_i , пользуясь чисто феноменологическими параметрами (которые могут включать в себя сумму членов по отдельным реакциям)²

$$\mathcal{F}_i = k_0 \mathcal{A}_i Q_i; \quad \mathcal{R}_i = k_0 \mathcal{D}_i + \varphi_{0i}. \quad (\text{II. 11})$$

Здесь k_0 — *общая* константа скорости первого порядка с размерностью с^{-1} . Она дает стандартное (или пороговое) значение констант для всех конкурентных процессов образования. При этом в размерность остальных параметров \mathcal{A}_i , Q_i и \mathcal{D}_i уже не входит время.

¹ В дальнейшем мы будем опускать индекс v , кроме тех случаев, когда необходимо выделить соответствующий класс носителей информации.

² Эти конкретные символы выбраны для того, чтобы подчеркнуть общий вид параметров скорости и качества, которые могут быть сложными функциями концентраций и могут состоять из нескольких членов. [\mathcal{F} (formation) — образование; \mathcal{A} (amplification) — усиление; Q (quality) — качество; \mathcal{R} (removal) — удаление; \mathcal{D} (decomposition) — разложение. — Прим. перев.]

Произведение $\mathcal{A}_i Q_i$ характеризует индивидуальную величину и форму параметров скорости образования \mathcal{F}_i . Назовем величину \mathcal{A}_i «фактором усиления»; $k_0 \mathcal{A}_i$ представляет собой фактически константу скорости; эта константа показывает, с какой быстротой направляется синтез матрицей i_i . Таким образом, $k_0 \mathcal{A}_i$ учитывает все процессы образования (за единицу времени), которые инструктируются матрицей i_i , независимо от того, ведут ли они к образованию точных копий i_i или мутантов. Затем мы вводим «фактор качества» Q_i , который говорит нам, какая доля Q_i этих процессов ведет к образованию точной копии i_i . Доля $(1 - Q_i)$ всех процессов «копирования», направляемых i_i , отвечает образованию мутантов i_j , которые частично все еще сходны с оригиналом i_i , но содержат различные ошибки, имеющие определенное распределение вероятности. Эти мутанты обычно характеризуются кинетическими параметрами $\mathcal{F}_j \leq \mathcal{F}_i$, но иногда может возникнуть копия, обладающая каким-либо преимуществом ($\mathcal{F}_j > \mathcal{F}_i$).

Определения \mathcal{A}_i и Q_i чисто феноменологические. Можно подсчитать число процессов копирования (в единицу времени), инструктированных данной матрицей, и можно определить также — по последовательности единиц или другими способами — число правильных копий i_i . Таким образом, оба параметра имеют физический смысл. Это проявляется также, если рассматривать суммарную скорость образования

$$\sum_{k=1}^N \mathcal{F}_k x_k + \sum_{k=1}^N \sum_l \varphi_{kl} x_l = k_0 \sum_{k=1}^N \mathcal{A}_k x_k, \quad (\text{II. 12})$$

где правая часть уже не содержит параметров качества, потому что в суммарную продукцию входит как доля точных копий Q_i , так и доля ошибочных (или мутантных) копий $(1 - Q_i)$.

Последний член в уравнении (II. 10) отражает тот факт, что любая ошибка означает образование близкой копии. Этот член может быть очень мал или им даже можно пренебречь для некоторых избранных видов, но он может быть важен для постоянного образования «свиты ошибочных копий» от отобранной главной копии.

Из условия сохранения — в отсутствие спонтанного неинструктированного синтеза — получим выражение для суммарной продукции ошибочных копий:

$$k_0 \sum_{k=1}^N \mathcal{A}_k (1 - Q_k) x_k = \sum_{k=1}^N \sum_l \Phi_{kl} x_l. \quad (\text{II. 13})$$

Произведения $\mathcal{A}_i Q_i$ содержат также стехиометрические функции $f_i(m_1, \dots, m_\lambda)$, описывающие зависимость скорости образования от концентраций мономерных (богатых энергией) единиц. Их точный вид зависит от конкретного механизма полимеризации (см. гл. IV), особенно на этапе, лимитирующем скорость. Если концентрации m_1, \dots, m_λ забуферены — условие, которое выбирается для большинства эволюционных экспериментов (см. гл. VII), — то f_i можно включить в $\mathcal{A}_i Q_i$ в качестве постоянного множителя. Иными словами, можно написать

$$\mathcal{A}_i Q_i = f_i(m_1, \dots, m_\lambda) \mathcal{A}'_i Q'_i \quad (\text{II. 14})$$

(и, может быть, также соответствующие средние для различных ошибочных копий). $\mathcal{A}'_i Q'_i$ не содержат больше концентрационных членов, зависящих от m_1, \dots, m_λ , но могут еще зависеть от концентрации x_i или от любой x_k . ($f_i(m_1, \dots, m_\lambda)$ можно пронормировать, например, к начальным или конечным условиям.)

С помощью соотношения (II. 11) для \mathcal{P}_i мы отделили индивидуальный член \mathcal{D}_i , описывающий разложение (который, как и член $\mathcal{A}_i Q_i$, множится на общую константу скорости k_0), от члена Φ_{0i} , который описывает разбавление и который в кинетическом уравнении множится на x_i («пропорциональное» разбавление). В самом простом случае мы можем опустить индекс i ; в этом случае Φ_0 будет связан с общим потоком «разбавления» Φ_0^1 следующим образом:

$$\Phi_0 = \frac{\Phi_0^1}{\sum_{k=1}^N x_k}. \quad (\text{II. 15})$$

¹ Φ_0 может быть также отрицательным, если происходит «концентрирование» системы. Однако обычно член Φ_0 будет использоваться для компенсации роста.

Тогда общая скорость удаления будет равна

$$\sum_{k=1}^N \mathcal{R}_k x_k = k_0 \sum_{k=1}^N \mathcal{D}_k x_k + \phi_0. \quad (\text{II. 16})$$

Как уже упоминалось, параметры \mathcal{D}_i могут быть функциями концентраций x_i или x_k , но здесь мы не можем выделить какой-либо параметр «качества», так как разложение носителей информации любого вида ведет к образованию бесполезных продуктов («отбросов»), если не водить сложных репаративных механизмов.

Появление мутантов, вызванное не ошибками копирования, а какими-либо другими процессами, можно формально включить в те же три параметра \mathcal{A}_i , \mathcal{Q}_i и \mathcal{D}_i . Тогда уравнение (II.10) примет вид

$$\dot{x}_i = k_0 [\mathcal{A}_i \mathcal{Q}_i - \mathcal{D}_i] x_i + \sum_{l \neq i} \phi_{il} x_l - \phi_{0i} x_i. \quad (\text{II. 17})$$

Это феноменологическое кинетическое уравнение описывает вообще любую систему реакций, которая обладает следующими характерными свойствами:

а) метаболизмом, представленным двумя суммарными членами $k_0 \sum \mathcal{D}_k x_k$ и $k_0 \sum \mathcal{A}_k x_k$, которые описывают превращение высокоэнергетических веществ в низкоэнергетические;

б) способностью к самовоспроизведению, как показывает вид кинетического уравнения. Предполагается, что члены, описывающие образование и распад, пропорциональны x_i и $\mathcal{F}_i > \mathcal{D}_i$ для $0 < x_i < n$, какова бы ни была дальнейшая зависимость \mathcal{F}_i и \mathcal{D}_i от концентрации.

в) мутабельностью, представленной фактором качества $\mathcal{Q}_i < 1$.

Оказывается, что первая часть уравнения (II.17) ответственна за способность к отбору и что \mathcal{A}_i , \mathcal{Q}_i и \mathcal{D}_i являются решающими феноменологическими параметрами. Даже для сложных «живых» существ отбор определяется этими параметрами, в которые могут входить сложные концентрационные члены, обусловленные «внутренними» связями, и которые могут зависеть от многих переменных внешней среды.

Однако в такой форме уравнение (II.17) еще не описывает процесс отбора. Оно определяет некую сегрегацию, обусловленную пороговым свойством

$$\mathcal{A}_i Q_i \geq \mathcal{D}_i. \quad (\text{II. 18})$$

Те носители информации, которые находятся выше порога ($\mathcal{A}_i Q_i > \mathcal{D}_i$), будут увеличиваться в числе, а те, которые находятся ниже порога, ($\mathcal{A}_i Q_i < \mathcal{D}_i$), будут вымирать.

Если пренебречь второй частью уравнения (II.17), то его решение можно записать в следующем общем виде:

$$x_i(t) = x_i^0 \exp \left\{ k_0 \int_0^t (\mathcal{A}_i Q_i - \mathcal{D}_i) dt \right\}. \quad (\text{II. 19})$$

Для постоянных \mathcal{A}_i , Q_i и \mathcal{D}_i это решение представляет собой вещественную экспоненту с положительным или отрицательным аргументом. Если какая-либо из этих реакций приближается к равновесию, то соответствующее решение принимает вид экспоненты с отрицательным аргументом, как показано в разд. I.2.4. Это получается с помощью разложения в ряд обратимых (в общем случае нелинейных) членов кинетического уравнения, после чего переменной становится отклонение «степени развития реакции» от равновесного значения (см. табл. 5).

Нелинейные системы могут проявить гораздо более резкое «сегрегационное» поведение. Например, дифференциальное уравнение типа

$$\dot{x}_i = (\mathcal{A}_i Q_i - \mathcal{D}_i) x_i, \text{ где } (\mathcal{A}_i Q_i - \mathcal{D}_i) = a_i + b_i x_i, \quad (\text{II. 20})$$

имеет особую точку при конечном t , если b_i и $(a_i + b_i x_i^0)$ больше нуля (a_i может быть даже меньше нуля). При $x_i^0 = x_i(t=0)$ мы имеем

$$x_i(t) = x_i^0 \frac{a_i e^{a_i t}}{a_i + b_i x_i^0 (1 - e^{a_i t})}, \quad (\text{II. 21})$$

т. е. $x_i(t)$ стремится к бесконечности при

$$t \rightarrow a_i^{-1} \ln \left(1 + \frac{a_i}{b_i x_i^0} \right). \quad (\text{II. 22})$$

Если a_i отрицательно, то его абсолютное значение должно быть меньше $b_i x_i^0$; в противном случае вид решения меняется. Если $a_i \ll b_i x_i^0$, решение сводится к простой гиперболе

$$x_i(t) = \frac{x_i^0}{1 - b_i x_i^0 t} \quad (\text{II. 23})$$

с особой точкой при $t = 1/(b_i x_i^0)$.

Эти и подобные им решения описывают резкое селекционное поведение (см. гл. VI) и оказываются очень важными для начала эволюции.

Я намеренно назвал описанное выше поведение «сегрегационным», а не «селекционным», потому что оно ведет только к разделению системы на две части, выделяемые пороговым свойством ($\mathcal{A}_i Q_i \cong \mathcal{D}_i$). Если мы хотим интерпретировать «отбор» как действие некоего экстремального принципа, то нам понадобятся не только способности к росту отдельных компонентов, но и «внешнее» селекционное ограничение, чтобы вызвать в системе настоящую конкуренцию за выживание.

§ II. 3. Селекционные ограничения

Можно придумать много внешних ограничений и внутренних связей, которые сделали бы систему реакций более конкурентной (см. проблемы «борьбы» в книге В. Вольтерра [53]). Однако есть две простые процедуры, которые послужат нам общей основой для теории отбора. Обе эти процедуры можно связать с теорией Пригожина и Глансдорфа [47], которая рассматривает химические реакции вблизи стационарного состояния.

Потребуем, чтобы система находилась в стационарном состоянии. В термодинамической теории (см. гл. I) такие системы можно рассматривать либо при постоянных реакционных силах, либо при постоянных реакционных потоках.

Подобным же образом при рассмотрении информационного ящика, введенного в § II. 1, мы можем поддерживать постоянной либо общую организацию (и тем

самым некое суммарное сродство, см. табл. 5), либо суммарный поток единиц (который определяется потоками мономерных единиц внутрь и наружу, а также суммарными реакционными потоками внутри ящика).

Точнее, первое условие означает, что суммарное число организованных и неорганизованных единиц, а также «степень организованности» в ящике должны поддерживаться постоянными. Физически этому можно способствовать, если забуферить концентрации $m_1 \dots m_\lambda$ богатых энергией мономеров и регулировать (посредством ϕ_0) суммарный поток таким образом, чтобы суммарное число носителей информации оставалось постоянным.

Тогда в случае постоянной общей организации налагаются следующие ограничения:

$$m_1 \dots m_\lambda = \text{const}, \quad (\text{II. 24})$$

вследствие чего $f_i(m_1 \dots m_\lambda) = \text{const}$ и могут быть включены в $\mathcal{A}_i Q_i$,

$$\sum_{k=1}^N x_k = \text{const} = n \quad (\text{II. 25})$$

(или, если мы рассматриваем различные классы v , $\sum_v v n_v = \text{const}$).

Поток разбавления ϕ_0 нужно регулировать таким образом, чтобы он компенсировал суммарную избыточную продукцию:

$$\phi_0 = k_0 \sum_{k=1}^N [\mathcal{A}_k - \mathcal{D}_k] x_k. \quad (\text{II. 26})$$

Альтернативное условие отбора допускает, чтобы содержимое ящика варьировало, но тогда приток мономерных, богатых энергией единиц ($\phi_1 \dots \phi_\lambda$), а также реакционный поток, т. е. суммарная скорость сборки и распада носителей информации (включая и отток бедных энергией продуктов разложения), должны быть постоянными:

$$\phi_1 \dots \phi_\lambda = \text{const} \quad (\text{II. 27})$$

или, если мономеры получаются из одного и того же источника,

$$\phi_M = \sum_{k=1}^{\lambda} \phi_k = \text{const}$$

и

$$k_0 \sum_{k=1}^N \mathcal{A}_k x_k = k_0 \sum_{k=1}^N \mathcal{D}_k x_k = \phi_M = \text{const.} \quad (\text{II. 28})$$

Экспериментальное осуществление и теоретическое рассмотрение последнего случая более трудны. Некоторые процессы в природе могут протекать в условиях, близких к этим. В течение какого-то времени энергия может поставляться с постоянной скоростью (например, солнечная энергия), так что количество богатого энергией вещества может изменяться таким образом, чтобы скорость продукции оставалась постоянной. (Увеличение констант скоростей компенсируется уменьшением концентрации мономерных единиц.) Аналогично число носителей информации может возрасти до такого уровня, когда их распад будет регулироваться скоростью их образования. Строгое соблюдение этого условия, однако, требует сложных методов контроля. Для эволюционных экспериментов легче поддерживать условия постоянной общей организации (см. опыты С. Спигелмана с фагом Q β , описанные в гл. VII). Можно построить «эволюционные машины», которые будут автоматически контролировать и поддерживать определенные условия; можно также представить себе другие ограничения, включающие различные комбинации из двух приведенных выше условий.

Замечание: Хотя эта детализация определенных условий важна для понимания принципов эволюции и количественной оценки экспериментальных данных, совершенно не обязательно, чтобы какой-либо реальный эволюционный процесс в природе происходил при этих специальных условиях, так же как ни одна паровая машина не работает в точности при условиях термодинамического равновесия в цикле Карно.

Вернемся теперь к феноменологическим уравнениям и перепишем их с должным учетом двух различных селекционных ограничений:

а) Постоянная общая организация

Введем определения:

$$E_i = \mathcal{A}_i - \mathcal{D}_i - (\text{избыточная}) \text{ продуктивность,} \quad (\text{II. 29})$$

$$\bar{E} = \frac{\sum_{k=1}^N E_k x_k}{\sum_{k=1}^N x_k} - \text{«средняя продуктивность»,} \quad (\text{II. 30})$$

$$W_i^0 = \mathcal{A}_i Q_i - \mathcal{D}_i - \text{«селективная ценность».} \quad (\text{II. 31})$$

Все эти величины относятся к ограничению постоянной общей организации.

Заменив φ_0 согласно формулам (II. 15) и (II. 26), мы можем переписать уравнение (II. 17) в виде

$$\dot{x}_i = k_0 [W_i^0 - \bar{E}] x_i + \sum_{l \neq i} \varphi_{il} x_l. \quad (\text{II. 32})$$

Член $\left(\sum_{l \neq i} \varphi_{il} x_l \right)$, соответствующий «исправлению ошибок» или «обратному потоку информации», обычно оказывается пренебрежимо малым для любой отобранной главной копии i_m . Такая главная копия, однако, всегда будет сопровождаться «кометным хвостом» ошибочных копий, постоянное присутствие которых обусловлено главным образом членом $\varphi_{im} x_m$ (индексом m обозначена отобранная главная копия). Уравнение (II.32) существенно нелинейно — даже для постоянных значений \mathcal{A}_i , Q_i и \mathcal{D}_i — вследствие того, что каждая популяционная переменная x_i входит в выражение для средней продуктивности \bar{E} . Таким образом, все уравнения связываются этим членом, который дает скользящее и саморегулирующееся пороговое значение, отражающее самоорганизацию системы. Будет расти численность только тех носителей информации, «селективные ценности» W_i^0 которых находятся выше порога \bar{E} . Вследствие своего роста они постоянно сдвигают порог \bar{E} в сторону все больших значений, пока не достигается оптимум \bar{E} , равный максимальной селективной ценности всех имеющихся видов:

$$\bar{E} \rightarrow W_m^0 \quad (\text{II. 33})$$

или, в более общем случае (например, для осциллирующих систем), для любого периода колебаний Δt

$$\int_t^{t+\Delta t} \bar{E} dt \rightarrow \int_t^{t+\Delta t} W_m^0 dt \quad \text{или} \quad \langle \bar{E} \rangle \rightarrow \frac{1}{\Delta t} \int_t^{t+\Delta t} W_m^0 dt. \quad (\text{II. 34})$$

В зависимости от типа реакционной системы (т. е. от связей между различными компонентами) селективная ценность W_m^0 будет принадлежать единственному носителю информации (или вырожденному классу), если параметры A_m , Q_m и D_m — константы, или она может относиться к каталитическому циклу, и тогда она будет функцией концентраций всех участников цикла, или же она может относиться даже к целой иерархии реакционных циклов, что выражается в появлении концентрационных членов высшего порядка. Для нелинейных систем конечное значение W_m^0 для любого вида носителей информации может зависеть от начальных концентраций, как, например, показывает равенство (II. 21). Важно отметить, что индекс m в W_m^0 относится к виду с «максимальной» селективной ценностью по сравнению со всеми конкурентами, присутствующими в конечной фазе. Далее, соотношения (II. 33) или (II. 34) являются приближениями лишь для того случая, когда членами, описывающими «исправление ошибок», можно пренебречь. В противном случае W_m^0 можно заменить на

$$W_m^0 + \frac{1}{k_0} \sum_{l \neq m} \varphi_{ml} \frac{\bar{x}_l}{\bar{x}_m},$$

где второй член представляет собой средний «обратный поток» информации от мутантов (он может также включать в себя члены, описывающие исправление).

Когда вклад членов, выражающих исправление, мал, соотношение (II. 33) в следующем приближении можно записать так:

$$\bar{E} \rightarrow W_m^0 + \frac{1}{k_0^2} \sum_{l \neq m} \frac{\varphi_{ml} \varphi_{lm}}{W_m^0 - W_l^0}. \quad (\text{II. 33a})$$

Это приближение пригодно до тех пор, пока второй член мал по сравнению с любой разностью $(W_m^0 - W_{l \neq m}^0)$, что

может всегда выполняться, если Q_m достаточно близко к единице.

Таким образом, мы охарактеризовали отбор неким экстремальным принципом по аналогии с термодинамическим равновесием. Однако там мы имеем «истинный» максимум энтропии или минимум свободной энергии, тогда как здесь мы имеем дело лишь с «оптимумами», т. е. с относительными максимумами, которые получаются, если варьировать состав популяции при известных ограничивающих условиях. Мы можем также назвать состояние, характеризующееся критерием (II. 33) или (II. 34), «селекционным равновесием», но необходимо помнить, что здесь мы имеем дело с метастабильным состоянием равновесия. Оно стабилизирует информацию, необходимую для воспроизведения того носителя информации, который может считаться «наиболее приспособленным» в популяции, но только до тех пор, пока во флуктуирующем распределении ошибок (ср. $Q < 1$ и $nV \ll N$) не возникнет «селективного преимущества», характеризующегося неравенством $W_{m+1}^0 > W_m^0$. Как только появляется новая копия (или ансамбль) такого рода i_{m+1} (ср. стохастический подход в гл. III), прежнее равновесие разрушается и система идет к новому (метастабильному) состоянию равновесия, которое характеризуется другим значением \bar{E} ; все это в целом соответствует процедуре оптимизации. Если W_m^0 не зависит от x_i , то изменение \bar{E} (при постоянных условиях окружающей среды) соответствует монотонному росту

$$W_m^0 < W_{m+1}^0 < \dots < W_{\text{opt}}^0 \quad (\text{II. 35})$$

Конечное состояние является «оптимальным состоянием», т. е. максимумом при ограничивающих условиях, заданных в виде неравенств. Система может тогда достигать лишь некоторых определенных состояний из полного набора N возможностей; она привязана к определенному пути, по которому вынуждена «карабкаться».

Если существуют внутренние связи, выражаемые концентрационными зависимостями параметров W_i^0 , то весь процесс оптимизации становится более сложным. Максимум селективной ценности в популяции не имеет

«абсолютного» значения, потому что он не относится теперь к данному распределению концентраций x_i . В такой системе любое изменение распределения x_i означает «изменение окружающей среды». Оптимизация здесь относится к «дифференциальному» процессу; она может даже сопровождаться общим уменьшением (возможно, всех) селективных ценностей (например, вследствие того, что один из отобранных носителей информации вызывает «загрязнение среды»). Она может также состоять в использовании большего количества информации (чтобы приспособиться к изменившемуся окружению). В общем процедура оптимизации в эволюции не обязательно должна представлять собой простое монотонное изменение селективных ценностей. Каким бы ни было конечное состояние, \bar{E} относится здесь к максимальному значению W^0 всех конкурентов, присутствующих в конечной фазе. Те виды, которые принадлежат к кооперативной системе, достигнут одного и того же значения W_m^0 , и это равенство можно использовать для вычисления «равновесного» распределения избранных видов по аналогии с «законом действия масс».

Важно отметить, что различие между «отбором» (происходящим в данном наборе населенных состояний при фиксированных условиях среды) и «эволюцией» (дальнейшей процедурой оптимизации по отношению к изменениям популяции и среды) является абстракцией. Если рассматривать весь процесс эволюции как некую игру, то эта абстракция позволяет использовать механизм отбора как исполнительный инструмент для вычисления состояния игры согласно определенным правилам и, таким образом, заменить игрока. Эта абстракция приблизительно выполняется лишь для таких систем, где $nV \ll N$ и Q близко к 1. Она требует также экспериментальных условий, которые допускают, чтобы отбор происходил за сроки, небольшие по сравнению со временем эволюционного изменения.

б) Постоянные общие потоки организованных и неорганизованных единиц

Система феноменологических уравнений имеет несколько более сложный вид, чем при ограничениях

первого типа. Рассмотрим поэтому упрощенный пример, которому присущи, однако, все существенные особенности общего случая. Предположим, что все носители информации (включая ошибочные копии) имеют примерно один и тот же общий состав и скорость их образования описывается одной и той же (средней) стехиометрической функцией $f(m_1 \dots m_\lambda)$. Далее предположим, что имеется постоянный и однородный приток ϕ_M богатых энергией мономеров (в постоянном соотношении). Тогда ограничения, введенные равенствами (II.27) и (II.28), дают

$$f(m_1 \dots m_\lambda) = \frac{\phi_M}{N \sum_k \mathcal{A}'_k x_k} \quad (\text{II. 36})$$

и кинетические уравнения могут быть написаны по аналогии с уравнением (II.32) (пренебрегая членами «обратного потока»)

$$\dot{x}_i = \frac{k_0 \mathcal{D}_i}{\bar{P} + 1} [W_i^F - \bar{P}] x_i. \quad (\text{II. 37})$$

Здесь \bar{P} — снова средняя «продуктивность», однако она определяется как

$$\bar{P} = \frac{\bar{\mathcal{A}'}}{\bar{\mathcal{D}}} - 1, \quad (\text{II. 38})$$

где

$$\bar{\mathcal{A}'} = \frac{\sum_{k=1}^N \mathcal{A}'_k x_k}{\sum_{k=1}^N x_k} \quad \text{и} \quad \bar{\mathcal{D}} \quad \text{определяется соответственно.}$$

Как продуктивность, так и селективная ценность

$$P_i = \frac{\mathcal{A}'_i}{\mathcal{D}_i} - 1; \quad W_i^F = \frac{\mathcal{A}'_i Q_i}{\mathcal{D}_i} - 1 \quad (\text{II. 39})$$

относятся здесь к ограничению постоянных потоков.

Если не существует общей стехиометрической функции $f(m_1 \dots m_\lambda)$, то уравнение (II.36) можно все-таки получить формально, используя подходящие средние наряду с индивидуальными стехиометрическими членами.

Случай постоянных потоков особенно интересен в связи с применением принципа Пригожина — Глансдорфа, который обсуждался в конце гл. I. Этот принцип относится к изменению сил при постоянных потоках. Из-за наличия скользящего порога система снова выбирает максимальную селективную ценность среди имеющейся популяции. Здесь можно показать, что каждая мутация, ведущая к дальнейшему росту «селективной ценности», соответствует отрицательной флуктуации производства энтропии, указывая на неустойчивость существующего стационарного состояния. При условии постоянных потоков эволюция соответствует последовательности таких нестабильностей, когда доминирующий вид i_m вымирает, уступая место новому виду i_{m+1} , который имеет конечное (положительное) селективное преимущество $(W_{m+1}^F - W_m^F)$.

§ II.4. Селекционное равновесие

Мы назвали состояние с максимальной продуктивностью для данной популяции «селекционным равновесием». В этом «равновесии» происходят флуктуации распределения ошибок, и оно метастабильно по отношению к возникновению видов с селективными преимуществами. Тем не менее, как и в химической термодинамике, мы можем вывести «константы равновесия» из условий (II.33) или (II.34) или соответственно из их аналогов при ограничении постоянных потоков.

а) Постоянная общая организация

Средняя продуктивность \bar{E} записывается как

$$\bar{E} = \frac{E_m x_m + \sum_{k \neq m} E_k x_k}{\sum x_k} = \bar{E}_{k \neq m} + \frac{x_m}{n} (E_m - \bar{E}_{k \neq m}), \quad (\text{II. 40})$$

где среднее значение остаточной продуктивности определяется как

$$\bar{E}_{k \neq m} = \frac{\sum_{k \neq m} E_k x_k}{\sum_{k \neq m} x_k}, \quad (\text{II. 41})$$

если принять

$$\sum_{k=1}^N x_k = \text{const} = n \quad \text{и} \quad \sum_{k \neq m} x_k = n - x_m. \quad (\text{II. 42})$$

Условие равновесия (II.33) тогда дает нам в первом приближении «равновесную фракцию» отобранного вида ($Q \geq 1$, если пренебречь членами «обратного потока», ср. II. 6, в)

$$\frac{\bar{x}_m}{n} = \frac{W_m^0 - \bar{E}_{k \neq m}}{E_m - \bar{E}_{k \neq m}}. \quad (\text{II. 43})$$

Критерий отбора можно записать в виде

$$W_m^0 > \bar{E}_{k \neq m}.$$

Это выражение дает физическое определение дарвиновскому термину «наиболее приспособленный».

Отметим, что относительное выживание \bar{x}_m/n не пропорционально Q_m , но стационарная ошибочная доля $(1 - \bar{x}_m/n)$ пропорциональна $(1 - Q_m)$, т. е.

$$1 - \frac{\bar{x}_m}{n} = \frac{A_m}{E_m - \bar{E}_{k \neq m}} (1 - Q_m). \quad (\text{II. 44})$$

Если бы Q_m равнялось единице, то W_m^0 стало бы равно E_m и \bar{x}_m достигла бы n . Это был бы крайний случай отбора, совершенно бесполезный, однако, для дальнейшей эволюции. «Ценность», которую приобретет такая система, ограничивается выбором из относительно узкого (случайного) набора всего лишь n видов. Мы видим также, что термин «ценность» не имеет смысла, если мы не определим, ценность «для чего». Ценность для отбора при специальных условиях уже отличается от более общей ценности для оптимальной эволюции. В последнем случае нам необходим фактор качества Q_m , достаточно большой, чтобы обеспечить выживание

$$Q_m > Q_{\min} = \frac{\bar{A}_{k \neq m} + D_m - \bar{D}_{k \neq m}}{A_m}, \quad (\text{II. 45})$$

но вместе с тем как можно более малый, чтобы обеспечить возникновение возможно большего числа мутан-

тов, за счет которых мог бы происходить дальнейший «прогресс». Система должна удовлетворять очень важному условию отбора (II.45), чтобы сохранить уже накопленную информацию. Таким образом, влияние малой величины $Q_m (> Q_{\min})$ проявляется двояким образом: во-первых, она приводит к большему разнообразию мутантов, среди которых система может отбирать, и, следовательно, позволяет прийти к более высокому конечному оптимальному значению W_m , а во-вторых, увеличивает скорость эволюции.

Селекционное равновесие в системах связанных реакций влечет за собой выживание целых ансамблей носителей информации. Здесь мы должны решать систему алгебраических уравнений. «Селективные ценности» и «продуктивности» могут стать очень сложными выражениями, содержащими все A -, Q - и D -параметры связанной системы (примеры будут даны в гл. V и VI). В случае линейных систем мы можем преобразовать переменные и представить все ансамбли посредством «нормальных координат», которые ведут себя аналогично концентрациям отдельных самоинструктирующихся видов. Для некоторых нелинейных систем равновесные соотношения также можно рассчитать в явном виде. Очевидно, для таких систем термин «равновесие» имеет много общего с тем, что мы обычно называем «химическим равновесием», потому что оно связывает концентрации нескольких, а иногда даже многих компонентов. Однако различие состоит в том, что при истинном равновесии партнеры превращаются друг в друга, тогда как при селекционном равновесии партнеры собираются из одного резервуара, а разлагаются в другой резервуар, не превращаясь друг в друга; но они делают это согласованным образом, и это обеспечивает сохранение фиксированных пропорций, пока поддерживаются фиксированные условия среды.

б) Постоянный поток информации

По аналогии с равенством (II.43) мы можем вычислить равновесное отношение отобранного вида, пользуясь определениями, введенными в разд. II, 3, б, и

получим (снова, как приближение для $Q \leq 1$)

$$\frac{\bar{x}_m}{\phi_M/k_0\mathcal{D}_m} \equiv \frac{\bar{x}_m}{\bar{n}} \equiv \frac{W_m^F - \bar{P}_{k \neq m}}{P_m - \bar{P}_{k \neq m}}. \quad (\text{II. 46})$$

Здесь \bar{x}_m нормирована не к сумме $\sum_{k=1}^N x_k$, которая не инвариантна, как в случае постоянной общей организации, а к аналогичной (при данном \mathcal{D}_m) постоянной величине

$$\frac{\phi_M}{k_0\mathcal{D}_m} \equiv \frac{\sum_{k=1}^N \mathcal{D}_k x_k}{\mathcal{D}_m} \equiv \bar{n}. \quad (\text{II. 47})$$

Здесь поток ϕ_M — сохраняющаяся величина и $\phi_M/k_0\mathcal{D}_m$ является аналогом общей концентрации. Критерий отбора снова дается неравенством $W_m^F > \bar{P}_{k \neq m}$ по аналогии с селекционным равновесием при постоянной организации.

Селективные преимущества могут быть введены через любой из трех параметров \mathcal{A} , Q и \mathcal{D} , которые определяют селективную ценность. Соответственно можно различать три предельных случая (табл. 7), включающих репрессию, дерепрессию, специфичное промотирование, переваривание, которое блокируется специфичной защитой, и т. д.

Таблица 7

Критерии отбора (предельные случаи по отношению к \mathcal{A}_m , Q_m и \mathcal{D}_m)

$$1. \mathcal{D}_m = \mathcal{D}_{k \neq m}, \text{ т. е. } \phi_M/\mathcal{P}_m = \sum_k x_k = n.$$

Постоянные силы и постоянные потоки

$$\frac{\bar{x}_m}{n} = \frac{\mathcal{A}_m Q_m - \bar{\mathcal{A}}_{k \neq m}}{\mathcal{A}_m - \bar{\mathcal{A}}_{k \neq m}};$$

отбор, если $\mathcal{A}_m Q_m > \bar{\mathcal{A}}_{k \neq m}$.

Продолжение табл. 7

$$2. \mathcal{A}_m = \bar{\mathcal{A}}_{k \neq m}.$$

а) Постоянные силы

$$\frac{\bar{x}_m}{n} = \frac{(\bar{\mathcal{D}}_{k \neq m} - \mathcal{D}_m) - (1 - Q_m) \bar{\mathcal{A}}_{k \neq m}}{\bar{\mathcal{D}}_{k \neq m} - \mathcal{D}_m};$$

отбор, если $\bar{\mathcal{D}}_{k \neq m} > \mathcal{D}_m + \bar{\mathcal{A}}_{k \neq m} (1 - Q_m)$.

б) Постоянные потоки

$$\frac{\bar{x}_m}{n} = \frac{Q_m \bar{\mathcal{D}}_{k \neq m} - \mathcal{D}_m}{\bar{\mathcal{D}}_{k \neq m} - \mathcal{D}_m};$$

отбор, если $Q_m \bar{\mathcal{D}}_{k \neq m} > \mathcal{D}_m$.

$$3. \mathcal{A}_m Q_m = \bar{\mathcal{A}}_{k \neq m}.$$

а) Постоянные силы

$$\frac{\bar{x}_m}{n} = \frac{\bar{\mathcal{D}}_{k \neq m} - \mathcal{D}_m}{(\mathcal{A}_m - \bar{\mathcal{A}}_{k \neq m}) + (\bar{\mathcal{D}}_{k \neq m} - \mathcal{D}_m)};$$

отбор, если $\bar{\mathcal{D}}_{k \neq m} > \mathcal{D}_m$.

б) Постоянные потоки

$$\frac{\bar{x}_m}{n} = Q_m \frac{\bar{\mathcal{D}}_{k \neq m} - \mathcal{D}_m}{\bar{\mathcal{D}}_{k \neq m} - Q_m \mathcal{D}_m};$$

отбор, если $\bar{\mathcal{D}}_{k \neq m} > \mathcal{D}_m$.

§ II.5. Фактор качества и распределение ошибок

Совершенно очевидно, что «селективная ценность» как динамическое свойство зависит от кинетических параметров, таких, как \mathcal{A} и \mathcal{D} . Менее очевидна роль фактора качества Q , который описывает точность воспроизведения. Для простых моделей можно в явном виде задать связь Q с такими молекулярными параметрами, как q_i — точность узнавания отдельных единиц, которая выражается через измеримые свободные энергии (или параметры активации) взаимодействия при «правильном» (комплементарном) и «неправильном» (некомплементарном) спаривании отдельных единиц. Обычно при таком узнавании имеет место кооперативное взаимодействие, в котором существенна природа ближайших

соседних пар (и, возможно, последующих). Это значительно увеличивает число возможных комбинаций. Пример неферментативного узнавания оснований нуклеиновых кислот А, У, Г и Ц ($\lambda = 4$) рассмотрен в гл. IV.

Если имеются всего два различных значения стабильности — одно для всех (вырожденных) комплементарных пар и другое для всех (вырожденных) некомплементарных пар, — то соотношение между феноменологическим параметром Q и молекулярным (однобуквенным) параметром узнавания q становится очень простым. Соответствующие формулы для такого случая представлены в табл. 8. Их можно обобщить для моделей, в которых имеется несколько невырожденных классов единиц (например, нуклеотиды).

Таблица 8

Фактор качества Q для модели «двух состояний»

Вероятность точного воспроизведения отдельной единицы: q .

Вероятность того, что произойдет одна из $(\lambda - 1)$ -кратно вырожденных ошибок: $(1 - q)$.

Математическое ожидание ошибки для последовательности из v единиц:

$$\varepsilon = v(1 - q). \quad (1)$$

Вырожденности:

| | | |
|--------------------------------|-------------------------------|-----|
| 1 | перестановка с 0 дефектов, | |
| $v(\lambda - 1)$ | перестановок с 1 дефектом, | (2) |
| $\binom{v}{k} (\lambda - 1)^k$ | перестановок с k дефектами, | |

$$\text{сумма: } \sum_{k=0}^v \binom{v}{k} (\lambda - 1)^k = \lambda^v \quad \text{возможных последовательностей (ср. табл. 4).} \quad (3)$$

Вероятность возникновения копии, не содержащей ошибок:

$$Q = Q_{v0} = q^v \rightarrow e^{-\varepsilon}. \quad (4)$$

Распределение вероятностей для возникновения последовательностей с k и только с k дефектами:

$$Q_{vk} = q^{v-k} (1 - q)^k \binom{v}{k} \quad (\text{биномиальное распределение}), \quad (5)$$

для $k(1 - q) \ll 1$ оно аппроксимируется распределением Пуассона:

$$Q_{vk} = \frac{\varepsilon^k \cdot e^{-\varepsilon}}{k!}. \quad (6)$$

Продолжение табл. 8

Сумма ошибочных копий:

$$\sum_{k=1}^v q^{v-k} (1-q)^k \binom{v}{k} = 1 - q^v = 1 - Q_{v0}. \quad (7)$$

Вероятность возникновения одного определенного мутанта, содержащего k ошибок в определенных положениях:

$$P_{vk} = \frac{Q_{vk}}{\binom{v}{k} (\lambda - 1)^k} = \frac{q^{v-k} (1-q)^k}{(\lambda - 1)^k} \rightarrow \left(\frac{q^{-1} - 1}{\lambda - 1} \right)^k e^{-e}. \quad (8)$$

Скорость продукции данного мутанта (например, с каким-то селективным преимуществом) пропорциональна P_{vk} . Распределения вероятностей (5), (6) и (8) показаны на рис. 4.

Минимальное значение Q_{v0} для стабильного отбора определено уравнением (II.45). При данном значении фактора узнавания q критерий стабильного отбора приобретает вид

$$v < \frac{|\ln Q_{\min}|}{|\ln q|} \quad (9)$$

или, если $(1-q) \ll 1$,

$$v < \frac{|\ln Q_{\min}|}{1-q}; \quad (10)$$

тем самым определяется предельное количество информации I_{\max} , которое может правильно воспроизводиться и сохраняться в системе.

Смысл Q как фактора качества иллюстрирует график распределения ошибок (рис. 4). Чем меньше ожидаемое значение ошибки, тем круче становится соответствующая кривая Q , т. е. тем выше фактор «качества». Важный критерий эволюции, уравнение (II.45), связывающее «разброс» кинетических параметров с минимумом фактора качества Q_{\min} , определяет максимальное количество информации v_{\max} , которое может сохраняться в процессе воспроизведения при фиксированном значении параметра узнавания q [ср. с уравнением (10) в табл. 8]:

$$v_{\max} = \frac{|\ln Q_{\min}|}{1-q}.$$

Вследствие этого элементарный механизм узнавания (т. е. q) должен улучшаться с ростом информационного содержания I_v носителя информации в ходе эволюции.

На высших ступенях молекулярной эволюции требуются очень сложные контрольные механизмы, чтобы увеличить точность узнавания отдельных букв (например, ферментативная проверка кода) и для того, чтобы значения параметров \mathcal{A} и \mathcal{D} могли варьировать в достаточно широких пределах (см. гл. IV—VI).

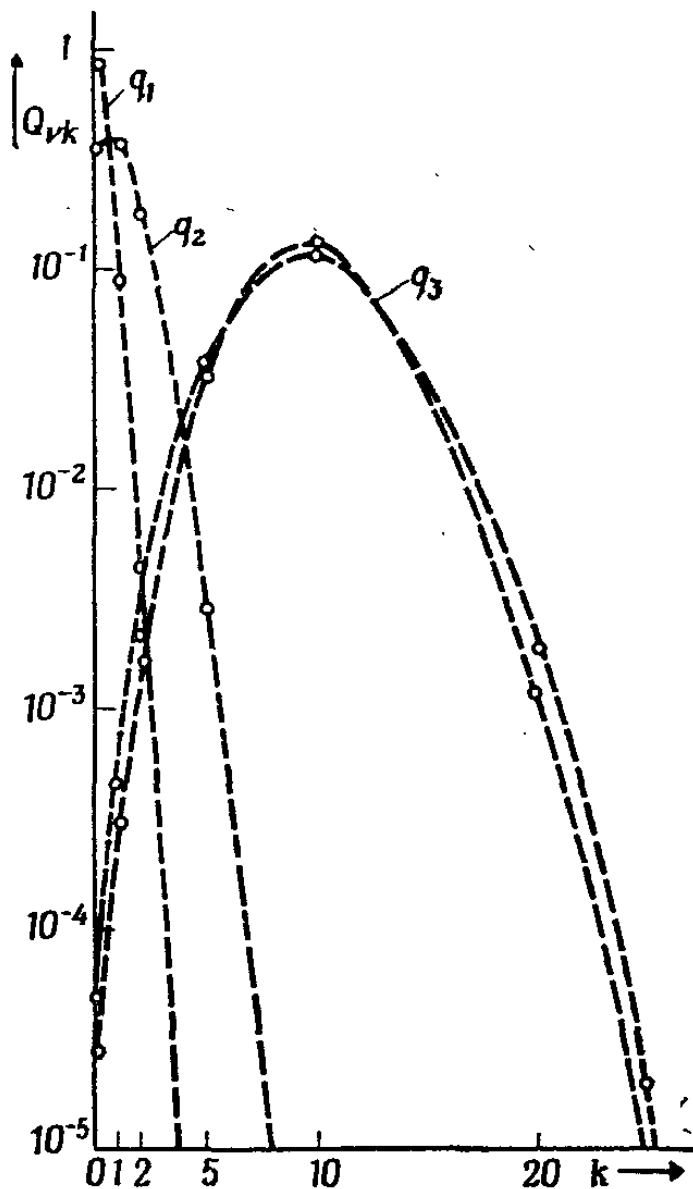


Рис. 4. Распределения вероятностей появления ошибочных копий. Q_{vk} вычислены по уравнению (5) в табл. 8. Значения параметров:

| | | |
|-------------|-------------|---------------------|
| $\nu=100$ | $q_1=0,999$ | $\varepsilon_1=0,1$ |
| $\lambda=4$ | $q_2=0,990$ | $\varepsilon_2=1$ |
| | $q_3=0,900$ | $\varepsilon_3=10$ |

Две кривые для q_3 даны для сравнения биномиального распределения с распределением Пуассона [уравнения (5) и (6)]. Для q_1 и q_2 при малых значениях k эти кривые совпадают (в пределах точности графика). Однако, когда $k \rightarrow \nu$, возникают большие различия. Кривые имеют максимумы при $k=\varepsilon$. В линейной шкале эти распределения убывали бы крайне резко. Распределение P_{vk} из уравнения (8) на графике не изображено, потому что оно слишком резко убывает с ростом k . Например, для q .

| k | P_{vk} |
|-----|---------------------|
| 0 | $3,7 \cdot 10^{-1}$ |
| 1 | $1,2 \cdot 10^{-3}$ |
| 2 | $4,2 \cdot 10^{-6}$ |

Цель этого обсуждения состояла в том, чтобы показать значение фактора качества Q ; не только для отбора в данной популяции, но и для скорости дальнейшей эволюции и для конечного оптимального состояния. В популяции, характеризующейся сравнимыми значениями кинетических параметров, система выбирает высокие значения Q , а это, по-видимому, неблагоприятно для скорости дальнейшей эволюции. С другой стороны, высокие значения Q делают возможным образование видов с большим информационным содержанием, что

в конечном счете оказывается благоприятным для дальнейшей эволюции. Мы видим, что процесс оптимизации может иногда совмещать противоречивые требования. Например, специфичное узнавание субстрата ферментом требует высокой константы стабильности для фермент-субстратного комплекса, но слишком высокие константы стабильности ограничивают скорость оборотов фермента. Эта проблема оптимальной эволюции, возникающая в результате обобщения концепции ценности, обсуждается в гл. VIII. Уже здесь мы видим, что «ценность» всегда требует спецификации оцениваемого свойства и, чем выше уровень эволюции, тем более необходима эта спецификация.

Существенный результат этого обсуждения состоит в том, что I_{\max} — максимальное количество информации $v_{\max} \frac{\ln \lambda}{\ln 2}$, которое может сохраняться при воспроизведении, — должно зависеть от точности узнавания элементарной единицы. Эта величина, определяемая изменением параметров \mathcal{A} и \mathcal{D} , входит только как логарифмический член и поэтому должна оказывать ограничивающее влияние только в случае малых изменений \mathcal{A} и \mathcal{D} (т. е. $\mathcal{A}_i \approx \bar{\mathcal{A}}_{k \neq i}$ и $\mathcal{D}_i \approx \bar{\mathcal{D}}_{k \neq i}$).

§ II.6. Кинетика отбора

Феноменологические уравнения с рассмотренными типами ограничений всегда представляют собой системы нелинейных дифференциальных уравнений. Явные решения, конечно, зависят от специального вида уравнений, который в свою очередь определяется конкретным механизмом реакции. Некоторые механизмы для самоорганизующихся систем, содержащих белки и нуклеиновые кислоты, будут обсуждаться в гл. IV—VI. Здесь мы будем иметь дело только с некоторыми прототипами решений для постоянных параметров W_i , чтобы охарактеризовать процесс отбора. Рассмотрим три случая в порядке возрастания сложности:

а) Постоянная общая организация

$$\bar{E}_{k \neq i} \text{ переменна; } Q_i \approx 1, \text{ т. е. } W_i^0 \approx E_i.$$

Точное решение системы дифференциальных уравнений (II.32) может быть дано только в том случае, если полностью пренебречь членами $\sum_{l \neq i} \varphi_{il} x_l$, т. е. потоком информации в мутантные копии и из них ($Q_i = 1$). В этом случае система уравнений упрощается:

$$\dot{x}_i = k_0 [W_i^0 - \bar{E}] x_i \quad (\text{II. 48})$$

и имеет следующее решение [$x_i^0 = x_i(t=0)$]:

$$x_i(t) = \frac{x_i^0 n \exp(k_0 W_i^0 t)}{\sum_{k=1}^N \frac{x_k^0 E_k}{W_k^0} \exp(k_0 W_k^0 t)}, \quad (\text{II. 49})$$

где любое W_k^0 можно также заменить на E_k .

Это решение можно получить, исходя из неявной формы, которая получается при интегрировании (II.48):

$$x_i(t) = x_i^0 \frac{\exp(k_0 W_i^0 t)}{\exp \left\{ k_0 \int_0^t \bar{E}(\tau) d\tau \right\}}. \quad (\text{II. 50})$$

Интегральный член исчезает для любого отношения (x_k/x_i), которое можно ввести в выражение для \bar{E} , если записать его следующим образом:

$$\bar{E} = \frac{x_i}{n} \sum_{k=1}^N \frac{x_k}{x_i} E_k. \quad (\text{II. 51})$$

Тогда кинетическое уравнение (II.48)

$$\dot{x}_i = k_0 W_i^0 x_i - k_0 \frac{\sum_{k=1}^N x_k^0 E_k \exp[k_0 (W_k^0 - W_i^0) t]}{n x_i^0} x_i^2 \quad (\text{II. 52})$$

представляет собой специальную форму дифференциального уравнения Бернулли

$$\dot{x} + g(t)x + f(t)x^2 = 0 \quad (\text{II.53})$$

с хорошо известным решением [54]

$$\frac{1}{x(t)} = e(t) \int \frac{f(t)}{e(t)} dt; \quad e(t) = \exp \left[\int g(t) dt \right]. \quad (\text{II. 54})$$

Подстановка

$$g(t) = \text{const} = -k_0 W_i^0 \quad (\text{II. 55})$$

$$f(t) = \frac{k_0}{n x_i^0} \sum_{k=1}^N x_k^0 E_k \exp [k_0 (W_k^0 - W_i^0) t] \quad (\text{II. 56})$$

дает решение (II. 49).

Это решение описывает в явном виде процедуру отбора. При $t = 0$ каждое x_i равно своему начальному значению x_i^0 . При $t \rightarrow \infty$ сумма экспонент может быть представлена ее наибольшим членом, который принадлежит виду с наибольшей «селективной ценностью»: W_m^0 . Этот вид отбирается. Его концентрация стремится к стационарному значению

$$\bar{x}_m = \frac{W_m^0}{E_m} n, \quad (\text{II. 57})$$

которое равно n , если $Q_m = 1$ (т. е. $W_m^0 = E_m$). Все другие виды должны в конце концов исчезнуть в соответствии с уравнением

$$x_i(t) = n \frac{x_i^0}{x_m^0} \frac{W_m^0}{E_m} \exp [(W_i^0 - W_m^0) k_0 t]. \quad (\text{II. 58})$$

Прежде чем x_m достигнет своего «доминантного» уровня, некоторые $x_i(t)$ могут сначала увеличиться в числе и пройти через максимум и лишь после этого исчезнуть. На рис. 5 приведен пример с четырьмя конкурирующими видами.

Этот подход все еще может дать полезное решение для главной копии, если Q_m не точно равно, но близко к единице. Когда отобранная главная копия достигнет доминирующего уровня, она будет конкурировать преимущественно с мутантами, возникающими вследствие неточного воспроизведения. Изложенный подход не описывает поведения этих мутантов, для которых в соответствии с равенством (II. 13) необходимо учитывать

дополнительные члены (особенно $\varphi_{im}x_m$) в кинетических уравнениях. Их концентрации, следовательно, не упадут до нуля. Для отобранной главной копии выражения (II.49) и (II.57) дают хорошее приближение, пока $\bar{E}_{k \neq m} \ll W_m^0$. Здесь мы не уточняем вида ошибочных копий, а просто имеем в виду, что какое-то число

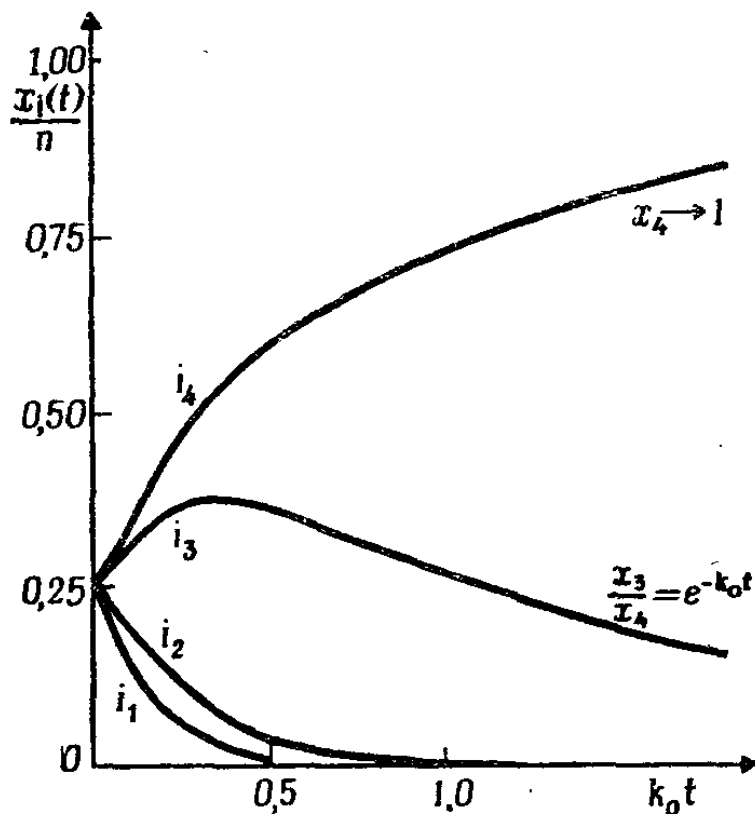


Рис. 5. Отбор в системе четырех конкурирующих видов по уравнению (II.49).

$$W_1^0 = 1; W_2^0 = 4; W_3^0 = 9; W_4^0 = 10.$$

ошибок существует. (Если $W_m^0 = E_m$, т. е. $Q_m = 1$, система не может развиваться дальше.) Прежде чем рассматривать более подробно взаимодействие между главной копией и мутантами, мы можем заняться специальным случаем, когда можно дать простое решение для $Q_i \leq 1$.

б) Постоянная общая организация или постоянные потоки

$$E_{k \neq m} \text{ или } \bar{P}_{k \neq m} \text{ постоянны.}$$

Этот случай описывает либо конкуренцию между двумя видами m и k , либо конкуренцию с целым вырожденным классом видов $k \neq m$, в котором $\bar{E}_{k \neq m}$ или

$P_{k \neq m}$ постоянны. Мы видим также, что этот случай представляет собой хорошее приближение для отбора в системе видов, которые не вырождены по $E_{k \neq m}$ или $P_{k \neq m}$, но дают довольно постоянное распределение вокруг средних значений $\bar{E}_{k \neq m}$ и $\bar{P}_{k \neq m}$. В то время как отобранный вид имеет другую скорость воспроизведения с $W_m^0 > \bar{E}_{m \neq k}$ или $W_m^F > \bar{P}_{k \neq m}$.

В случае постоянной общей организации решение для отобранного вида следующее:

$$x_m(t) = x_m^0 \frac{\exp[(W_m^0 - \bar{E}_{k \neq m})k_0 t]}{1 + \frac{x_m^0}{x_m} \{ \exp[(W_m^0 - \bar{E}_{k \neq m})k_0 t] - 1 \}}, \quad (\text{II. 59})$$

где

$$\bar{x}_m = n \frac{W_m^0 - \bar{E}_{k \neq m}}{E_m - \bar{E}_{k \neq m}} \quad [\text{ср. с равенством (II. 43)}]$$

— «равновесное» значение x_m (в соответствии с определением § II. 4), которое достигается при $t \rightarrow \infty$. Опять все «независимые» виды с $W_{k \neq m}^0 < W_m^0$ в конце концов вымирают, тогда как сумма по всем мутантам (вырожденным в соответствии с допущением $\bar{E}_{k \neq m} = \text{const}$ доходит до стационарного уровня [ср. с равенством (II. 44)].

Существует еще одно ограничение по отношению к Q_m : оно не должно быть слишком мало, чтобы можно было пренебречь повторным появлением вида i_m из мутантов (в результате исправления ошибок при последующих воспроизведениях). Допустимый диапазон значений Q_m зависит здесь от количества информации, т. е. от числа единиц v в цепи. Это же ограничение имеет место также для «равновесного» значения, согласно уравнению (II. 43). В других случаях уравнение (II. 43) служит хорошим приближением для конечной фазы отбора в пределах одной популяции, где имеет место практически только конкуренция между мутантами (и главной копией). Здесь среднее значение $\bar{E}_{k \neq m}$ действительно может быть аппроксимировано константой,

так что это решение воспроизводит правильное «равновесное» значение x_m (и сохраняет силу также в окрестности селекционного «равновесия»). С другой стороны, решение (II.49) (первый случай) не приводит к правильному равновесному значению, за исключением ситуации, когда $\bar{E}_{k \neq m} \ll W_m^0$; но это хорошее приближение для начальной фазы отбора, когда число мутантов мало по сравнению с числом неродственных копий. Решения, подобные (II.59), но с $Q = 1$, неоднократно обсуждались в литературе по генетике популяций [55]. Известный случай — уравнение Росса для распространения малярии [56]. А. Лотка дал также общие математические критерии устойчивости решений таких уравнений [57].

Явное решение задачи о конкуренции между «двумя видами» можно получить также в случае ограничения *постоянных потоков*, предположив, по крайней мере, одну и ту же стехиометрию $f(m_1 \dots m_\lambda)$ для всех видов (такое же допущение мы делали раньше в связи с выражениями (II.36) и (II.46). В кинетическом уравнении (II.37) члены $\bar{P}_{k \neq m}$, W_m^F и $n\bar{A} = n\bar{D} \equiv \phi_M$ являются константами (в предположении, что образование и распад подчиняются кинетике первого порядка), в то время как члены $n\bar{A}'$ и, следовательно, $(1 + \bar{P}) \equiv \bar{A}'/\bar{D}$ являются линейными функциями x_m . Индивидуальные кинетические уравнения, имеющие вид

$$\dot{x}_m = \frac{A + Bx_m}{C + Dx_m} x_m, \quad (\text{II. 60})$$

где

$$A = \phi_M \mathcal{D}_m (W_m^F - \bar{P}_{k \neq m}), \quad C = \frac{\phi_M}{k_0} (1 + \bar{P}_{k \neq m}),$$

$$B = -k_0 \mathcal{D}_m^2 (P_m - \bar{P}_{k \neq m}), \quad D = -B/k_0 \mathcal{D}_m,$$

можно проинтегрировать и получить следующие альтернативные решения:

$$\frac{[x_m(t)]^{\alpha_m}}{\dot{x}_m - x_m(t)} = \frac{[x_m^0]^{\alpha_m}}{\bar{x}_m - x_m^0} \exp\{k_0 \mathcal{D}_m [1 - \alpha_m] t\}$$

или

$$\frac{[x_m(t)]^{\alpha_m}}{[x_m(t)]^{\alpha_m} + [\bar{x}_m - x_m(t)]} = \frac{[x_m^0]^{\alpha_m} \exp\{k_0 \mathcal{D}_m [1 - \alpha_m] t\}}{[\bar{x}_m - x_m^0] + [x_m^0]^{\alpha_m} \exp\{k_0 \mathcal{D}_m [1 - \alpha_m] t\}}, \quad (\text{II. 61})$$

где

$$\alpha_m = \frac{1 + \bar{P}_{k \neq m}}{1 + W_m^F} \quad \text{или} \quad (1 - \alpha_m) = \frac{W_m^F - \bar{P}_{k \neq m}}{1 + W_m^F}. \quad (\text{II. 62})$$

Это решение ведет себя очень сходно с решением (II. 59). При $t \rightarrow 0$ численность «отобранного» вида растет по закону

$$x_m(t) = x_m^0 \exp\left\{ \frac{k_0 \mathcal{D}_m}{1 + \bar{P}_{k \neq m}} (W_m^F - \bar{P}_{k \neq m}) t \right\} \quad (\text{II. 63})$$

(приближение для $x_m(t) \ll \bar{x}_m$), тогда как при $t \rightarrow \infty$ $x_m(t)$ стремится к «равновесному значению»

$$\bar{x}_m = \bar{n} \frac{W_m^F - \bar{P}_{k \neq m}}{P_m - \bar{P}_{k \neq m}} \quad [\text{ср. с (II. 46)}].$$

в) Постоянная общая организация

$\bar{E}_{k \neq m}$ переменено; приблизительный учет продукции ошибок: $Q < 1$ (но не $\ll 1$).

По мере роста численности отобранного вида воспроизводимо возникающие мутанты главной копии i_m дают все больший вклад в общую продукцию. Эти мутанты можно разделить на классы согласно числу дефектных положений (по сравнению с главной копией), т. е. $\sum_k i_{1k}$, $\sum_k i_{2k}$ и т. д. Для последовательностей длиной v единиц число копий в каждом классе дается биномиальными коэффициентами, причем для λ типов единиц каждый дефект в данном положении $(\lambda - 1)$ -кратно вырожден; так, мы имеем $\binom{v}{l} (\lambda - 1)^l$ различных копий в классе с l дефектами. Частота продукции отдельных

копий i_{lk} убывает с увеличением l : например, в рамках простейшей модели, предполагающей равные q -факторы для узнавания отдельных единиц (ср. с табл. 8), эта частота пропорциональна

$$\frac{q^{v-l}(1-q)^l}{(\lambda-1)^l}.$$

Соответственно каждый класс дефектов будет давать вклад в восстановление главной копии, причем этот вклад убывает с увеличением числа дефектов.

Для того чтобы получить некоторое количественное представление о влиянии мутантов на отбор, рассмотрим упрощенный случай — влияние единичных дефектов. Это приближение пригодно лишь до тех пор, пока математическое ожидание ошибки $v(1-q)$ остается значительно меньше единицы, т. е. $0 \ll Q < 1$.

Тогда мы должны различать три класса кинетических уравнений ($W \equiv W^0$ всегда относится к постоянной организации).

а) Для одной главной копии i_m

$$\dot{x}_m = k_0 (W_m - \bar{E}) x_m + k_0 \sum_{k=1}^{v(\lambda-1)} \frac{E_{1k} - W_{1k}}{\beta_{1k}} x_{1k}. \quad (\text{II. 64a})$$

б) Для $v(\lambda-1)$ мутантов i_{1i} , имеющих один однобуквенный дефект,

$$\dot{x}_{1i} = k_0 (W_{1i} - \bar{E}) x_{1i} + k_0 \frac{E_m - W_m}{\beta_{mi}} x_m. \quad (\text{II. 64b})$$

в) Для $(n - x_m - \sum_k x_{1k})$ независимых конкурентов i_{ji}

$$\dot{x}_{ji} = k_0 (W_{ji} - \bar{E}) x_{ji}. \quad (\text{II. 64c})$$

Третью категорию независимых конкурентов можно считать также мутантами главной копии, имеющими $j \geq 2$ дефектов. Для достаточно длинных последовательностей это число возможных конкурентов так велико, что вероятность случайно найти любую данную копию практически равна нулю. Предполагается, что их

возникновением в результате мутаций главной копии можно пренебречь. Это относится и к тем мутантам главной копии, которые имеют дефекты только в двух положениях. Далее, приближение требует (до тех пор, пока независимые конкуренты дают существенный вклад в решение), чтобы

$$\sum_k x_{2k} \ll \sum_{j>2} \sum_k x_{jk}. \quad (\text{II. 65})$$

Факторы β в уравнениях (II. 64a) и (II. 64b) в общем случае связывают продукцию данного мутанта с общей продукцией дефектов, которая равна

$$\mathcal{A}_i (1 - Q_{i0}) = (E_i - W_i).$$

В случае одинаковых значений q с $Q_{i0} = q^v$ и равных скоростей продукции ошибок β_{mi} , например, сводится к

$$\beta_{mi} = v(\lambda - 1), \quad (\text{II. 66})$$

так как доля $1/\{v(\lambda - 1)\}$ всех ошибок, возникших при воспроизведении главной копии, будет соответствовать продукции любой однодефектной копии.

Для того чтобы получить самосогласованное приближение, мы должны пренебречь всеми мутациями, кроме возникновения единичных дефектов в главной копии и восстановления главной копии из однодефектных копий. Это заставляет допустить, что $W_{ji} \approx E_{ji}$, $(E_{1k} - W_{1k}) = \mathcal{A}_{1k}(1 - q)$ и $\beta_{1k} = 1$. Последнее предположение, т. е. замена Q на q для всех однодефектных копий, кажется на первый взгляд несколько нереалистичным, так как оно допускает только один сорт ошибки, а именно тот, который ведет к восстановлению главной копии. Однако пренебрежение другими ошибками при воспроизведении однодефектных копий (которые могли бы дать копии с двумя дефектами) согласуется с этим приближением и влияет только на малые поправочные члены. Для любой из отдельных копий i_{1k} различие между E_{1k} и W_{1k} пренебрежимо мало; однако в этом приближении надо учитывать сумму всех этих членов в уравнении (II. 64a).

Для главной копии мы получим следующую зависимость концентрации от времени ($x_m^0 = x_m(t=0)$):

$$\frac{x_m(t)}{n} = W'_m x_m^0 \left\{ x_m^0 \left[E_m + \sum_{k=1}^{v(\lambda-1)} A_k E_{1k} \right] + \right. \\ \left. + \sum_{k=1}^{v(\lambda-1)} B_k \exp \{k_0 (W_{1k} - W'_m) t\} + \right. \\ \left. + \sum_{jk} C_{jk} \exp \{k_0 (W_{jk} - W'_m) t\} \right\}^{-1}, \quad (\text{II. 67})$$

где

$$A_k = - \frac{E_m - W_m}{\beta_{mk} (W_{1k} - W'_m)}, \\ B_k = \frac{W'_m}{W_{1k}} E_{1k} \left(x_{1k}^0 + x_m^0 \frac{E_m - W_m}{\beta_{mk} (W_{1k} - W'_m)} \right), \quad (\text{II. 68}) \\ C_{jk} = \frac{W'_m}{W_{jk}} E_{jk} x_{jk}^0.$$

W'_m можно выразить как

$$W'_m \approx W_m + \frac{E_m - W_m}{W_m - \bar{E}_{1k}} \left\langle \frac{E_{1k} - W_{1k}}{\beta_{1k}} \right\rangle, \quad (\text{II. 69})$$

где символ $\langle \rangle$ означает «равновесное среднее», взятое по всем $v(\lambda-1)$ однодефектным копиям. Можно видеть, что W'_m сводится к W_m при $Q_m \rightarrow 1$. При $t \rightarrow \infty$ решение стремится к равновесному отношению [ср. с уравнением (II. 43)]

$$\frac{\bar{x}_m}{n} = \frac{W'_m - \bar{E}_{1k}}{E_m - \bar{E}_{1k}}, \quad (\text{II. 70})$$

где любой остаточный член пренебрежимо мал в рамках настоящего приближения, которое требует, чтобы главная копия отличалась достаточно большим селективным преимуществом: $W'_m > W_{1k}$.

Насколько мала разница между W'_m и W_m , можно видеть из модели, предполагающей одинаковую точ-

ность узнавания единиц, которая дает

$$W'_m - W_m \approx \frac{\nu(1-q)^2}{\lambda-1} \frac{\mathcal{A}_m \bar{\mathcal{A}}_{1k}}{W_m - \bar{E}_{1k}}, \quad (\text{II. 71})$$

где $(1-q) \ll 1/\nu$. (Отметим, что в этой модели средние $\bar{\mathcal{A}}_{1k}$ и \bar{E}_{1k} занимают место прежних средних $\bar{\mathcal{A}}_{k \neq m}$ и $\bar{E}_{k \neq m}$.)

Решения для однодефектных копий i_{1i} и для независимых конкурентов i_{ji} связаны с решением для главной копии следующими соотношениями:

$$\frac{x_{1i}(t)}{x_m(t)} = \left\{ \frac{x_{1i}^0}{x_m^0} + \frac{E_m - W_m}{\beta_{mi}(W_{1i} - W'_m)} \right\} \times \\ \times \exp[k_0(W_{1i} - W'_m)t] - \frac{E_m - W_m}{\beta_{mi}(W_{1i} - W'_m)}, \quad (\text{II. 72})$$

$$\frac{x_{ji}(t)}{x_m(t)} = \frac{x_{ji}^0}{x_m^0} \exp\{k_0(W_{ji} - W'_m)t\}. \quad (\text{II. 73})$$

Для получения этих явных решений, относящихся к отдельным видам, было сделано предположение, что прежде чем главные копии перерастут конкурентов, между ними и мутантами устанавливается «равновесие», так что в (малом) поправочном члене уравнения (II. 64a) отношение

$$\sum_{k=1}^{\nu(\lambda-1)} x_{1k}/x_m$$

можно заменить (постоянным) равновесным отношением. При этом кинетическое уравнение для главной копии упрощается:

$$\dot{x}_m = k_0[W'_m - \bar{E}]x_m. \quad (\text{II. 74})$$

Здесь W'_m отличается от W_m только членом, который остается малым, пока сумма $\sum_k x_{1k}$ мала по сравнению с x_m (т. е. $(1 - Q_m) \ll 1$). Дальнейшая процедура аналогична случаю «а». Легко находим выражение для

x_{1i}/x_m , причем

$$\frac{d}{dt} \left(\frac{x_{1i}}{x_m} \right) = \frac{\dot{x}_{1i}x_m - \dot{x}_m x_{1i}}{x_m^2}$$

берем из уравнения (II.64); формулу для x_{ji}/x_m находим аналогичным образом. Эти отношения используются для выражения \bar{E} в форме (II.51), и после подстановки в (II.74) для x_m получается дифференциальное уравнение типа Бернулли. Интегрирование можно провести по аналогии со случаем «а».

До сих пор наше обсуждение ограничивалось теоретическим рассмотрением самоотбирающихся реакционных систем, но не каких-либо реальных приложений. Для них удовлетворительные решения всегда можно получить с помощью ЭВМ. Предыдущее рассмотрение показывает нам, до какой степени мы можем использовать простые приближения. На самом деле простая линейная форма первичных кинетических уравнений применима лишь для очень немногих систем реакций. В общем случае нам придется учитывать различные взаимодействия, сначала между «информационными» и «функциональными» молекулами, а затем также между индивидуальными членами этих классов. Это может привести к возникновению *циклов или сетей реакций*, причем поведение каждого отдельного участника реакции будет описываться нелинейными кинетическими уравнениями (например, типа уравнения Михаэлиса — Ментен или даже более сложными). Эти более реалистичные системы будут детально рассмотрены и сопоставлены с экспериментальными результатами в гл. IV—VI. Мы увидим, что характерные свойства процесса отбора, вытекающие из теории, изложенной в этой главе, четко воспроизводятся в более сложных системах, хотя явные решения могут обнаруживать важные качественные отличия — возможны колебания различных типов или наличие особых точек, дающих очень резкий отбор. Мы увидим также, что эти качественные различия окажутся очень важными в связи с обсуждением проблемы возникновения самоорганизующихся «живых» систем.

СТОХАСТИЧЕСКИЙ ПОДХОД К ОТБОРУ

§ III. 1. Ограничения детерминистической теории отбора

До сих пор мы считали, что отбор представляет собой детерминированный процесс. Феноменологические уравнения четко выявляют, какой именно вид носителей информации в данной популяции будет отобран. Любой мутант, обладающий каким-либо селективным преимуществом ($W_i > \bar{E}$), неизбежно будет распространяться в популяции.

Имеются два обстоятельства, которые существенно лимитируют такое детерминистическое описание отбора:

1. Элементарный процесс, ведущий к возникновению какого-либо конкретного мутанта, существенно недетерминирован. Автокаталитическое усиление ведет к макроскопическому отображению случайных микроскопических событий¹.

2. Процесс роста численности сам по себе подвержен статистическим флуктуациям. Поскольку этот рост начинается с отдельных копий, такие флуктуации необходимо принимать во внимание. Они могут значительно модифицировать результаты детерминистической теории, которые справедливы только для средних по большому числу копий данного носителя информации.

Существует также дополнительная трудность, возникающая из-за того, что определенные стационарные состояния — в отличие от истинных равновесных состояний — метастабильны. Эти состояния не могут стабилизироваться, и поэтому для поддержания их в течение длительного времени необходима регуляция. Ввиду

¹ Насколько мне известно, П. Иордан [59] был первым, кто обратил внимание на «усиление» элементарных событий, которые подвержены квантовомеханической неопределенности.

x_{1i}/x_m , причем

$$\frac{d}{dt} \left(\frac{x_{1i}}{x_m} \right) = \frac{\dot{x}_{1i}x_m - \dot{x}_m x_{1i}}{x_m^2}$$

берем из уравнения (II.64); формулу для x_{ji}/x_m находим аналогичным образом. Эти отношения используются для выражения \bar{E} в форме (II.51), и после подстановки в (II.74) для x_m получается дифференциальное уравнение типа Бернулли. Интегрирование можно провести по аналогии со случаем «а».

До сих пор наше обсуждение ограничивалось теоретическим рассмотрением самоотбирающихся реакционных систем, но не каких-либо реальных приложений. Для них удовлетворительные решения всегда можно получить с помощью ЭВМ. Предыдущее рассмотрение показывает нам, до какой степени мы можем использовать простые приближения. На самом деле простая линейная форма первичных кинетических уравнений применима лишь для очень немногих систем реакций. В общем случае нам придется учитывать различные взаимодействия, сначала между «информационными» и «функциональными» молекулами, а затем также между индивидуальными членами этих классов. Это может привести к возникновению *циклов или сетей реакций*, причем поведение каждого отдельного участника реакции будет описываться нелинейными кинетическими уравнениями (например, типа уравнения Михаэлиса — Ментен или даже более сложными). Эти более реалистичные системы будут детально рассмотрены и сопоставлены с экспериментальными результатами в гл. IV—VI. Мы увидим, что характерные свойства процесса отбора, вытекающие из теории, изложенной в этой главе, четко воспроизводятся в более сложных системах, хотя явные решения могут обнаруживать важные качественные отличия — возможны колебания различных типов или наличие особых точек, дающих очень резкий отбор. Мы увидим также, что эти качественные различия окажутся очень важными в связи с обсуждением проблемы возникновения самоорганизующихся «живых» систем.

СТОХАСТИЧЕСКИЙ ПОДХОД К ОТБОРУ

§ III. 1. Ограничения детерминистической теории отбора

До сих пор мы считали, что отбор представляет собой детерминированный процесс. Феноменологические уравнения четко выявляют, какой именно вид носителей информации в данной популяции будет отобран. Любой мутант, обладающий каким-либо селективным преимуществом ($W_i > \bar{E}$), неизбежно будет распространяться в популяции.

Имеются два обстоятельства, которые существенно лимитируют такое детерминистическое описание отбора:

1. Элементарный процесс, ведущий к возникновению какого-либо конкретного мутанта, существенно недетерминирован. Автокаталитическое усиление ведет к макроскопическому отображению случайных микроскопических событий¹.

2. Процесс роста численности сам по себе подвержен статистическим флуктуациям. Поскольку этот рост начинается с отдельных копий, такие флуктуации необходимо принимать во внимание. Они могут значительно модифицировать результаты детерминистической теории, которые справедливы только для средних по большому числу копий данного носителя информации.

Существует также дополнительная трудность, возникающая из-за того, что определенные стационарные состояния — в отличие от истинных равновесных состояний — метастабильны. Эти состояния не могут стабилизироваться, и поэтому для поддержания их в течение длительного времени необходима регуляция. Ввиду

¹ Насколько мне известно, П. Иордан [59] был первым, кто обратил внимание на «усиление» элементарных событий, которые подвержены квантовомеханической неопределенности.

всех этих фактов нам придется пересмотреть проблему отбора с точки зрения вероятностной теории. Мы увидим, что стохастический¹ подход приведет к важным модификациям (детерминистической) феноменологической теории.

§ III. 2. Флуктуации вокруг равновесных состояний

Чтобы охарактеризовать различие между флуктуациями вблизи стационарного состояния и вблизи устойчивого равновесия, начнем это обсуждение с рассмотрения классического примера флуктуаций вблизи состояния равновесия — с эренфестовской модели урн (см. также рис. 3).

Имеются две урны и некоторое (большое) число — скажем, $2N$ — шаров, произвольно распределенных между этими урнами. Шары пронумерованы от 1 до $2N$.

Игра состоит в том, что случайно выбирается номер (можно, например, тянуть жребий, бросать кость или использовать любую другую лотерейную процедуру), после чего шар с соответствующим номером переносится из одной урны в другую. Если эту процедуру повторить достаточно много раз, то результатом будет — независимо от начального распределения — равномерное распределение $2N$ шаров по двум урнам.

Эту модель придумали П. и Т. Эренфесты [62], а впоследствии ее стохастические аспекты рассматривали Д. тер Хаар и К. Грин [63], М. Кац [64], М. Клейн [65] и другие. К. Кольрауш и Э. Шредингер [66] проверили эту модель экспериментально. Стохастическое рассмотрение модели (например, у М. Каца) выявляет следующие ее особенности:

1. Равновесное состояние, хотя оно и подвержено флуктуациям, обладает устойчивостью. В среднем каждая урна будет содержать N шаров.

¹ «Стохастическая» теория — это распространение теории вероятностей на динамические проблемы. От «стохастика» — догадка. Прекрасный обзор дан А. Рамакришнаном в «Энциклопедии физики», III/2 [60].

2. Вблизи равновесного состояния будут происходить флуктуации: одна урна будет содержать $N + n$, другая — $N - n$ шаров, причем n может принимать любое значение от $-N$ до $+N$. По аналогии с теоремой Больцмана мы можем охарактеризовать рассматриваемую модель функцией распределения

$$H = (N + n) \ln (N + n) + (N - n) \ln (N - n) \quad (\text{III. 1})$$

или для $n \ll N$

$$H = \frac{2n^2}{N} + \text{const}, \quad (\text{III. 2})$$

откуда видно, что флуктуации распределены симметрично по отношению к $n = 0$.

3. Вероятность найти $(N + n)$ шаров в одной и $(N - n)$ шаров в другой урне равна

$$P_0(n) = 2^{-2N} \frac{(2N)!}{(N + n)! (N - n)!} \rightarrow \frac{e^{-\left(\frac{n}{\sqrt{N}}\right)^2}}{\sqrt{\pi N}}. \quad (\text{III. 3})$$

Это распределение вероятностей стационарно по отношению к стохастическим уравнениям, т. е. оно не зависит от времени, тогда как n постоянно флуктурует. Распределение вероятностей симметрично по отношению к $n = 0$ (распределение Гаусса), причем его ширина пропорциональна \sqrt{N} . Большие флуктуации, порядка $n = N$, крайне маловероятны. В самом деле,

$$\frac{P_0(n = N)}{P_0(n = 0)} = 2^{-2N} \sqrt{\pi N}. \quad (\text{III. 4})$$

4. Ту же величину имеет и $\tau(n = 0) / \tau(n = \pm N)$ — отношение «времен возврата» (т. е. средних интервалов времени, через которые повторяются данные макросостояния). Время возврата минимально при $n = 0$.

Из всего этого следуют важные выводы. Равновесие является «устойчивым» состоянием. Флуктуации обладают свойством саморегуляции; чем больше отклонение в одном направлении, тем больше вероятность его обращения, т. е. восстановления равновесия. Средние флуктуации пропорциональны \sqrt{N} , поэтому они не существенны при больших N . Отношение времен возврата [см. уравнение (III, 4)] показывает, насколько редко

происходят на самом деле большие флуктуации, если N — большое число. Эта модель имела историческое значение для понимания природы необратимого процесса по сравнению с флуктуацией [67].

§ III.3. Флуктуации в стационарном состоянии

В предыдущем примере две урны представляли собой два эквивалентных состояния, между которыми в системе $2N$ эквивалентных шаров устанавливалось устойчивое равновесие. Давайте изменим нашу модель: вместо двух урн возьмем только одну, но теперь при вытягивании жребия предусмотрим не один, а два билета для каждого шара: белый и черный. Каждый раз, когда мы вытащим белый билетик, мы будем добавлять еще один шар в урну, а вытащив черный билетик, будем удалять один шар из урны. В сущности, если считать шары эквивалентными, нам больше не нужна их нумерация. Впрочем, если нас интересует эволюция отдельных видов, можно сохранить пронумерованные шары; при этом определенные номера будут дублироваться всякий раз, когда вытаскивается соответствующий белый билетик. Мы должны обеспечить также, чтобы каждому шару в урне всегда соответствовали ровно два билета (белый и черный); поэтому удаление или добавление шара всегда должно сопровождаться удалением или добавлением двух соответствующих билетиков.

Эта модель представляет собой типичную задачу о стационарном состоянии, где вероятности образования и распада равны и пропорциональны имеющемуся числу частиц. Детерминистическое уравнение снова дало бы распределение, не зависящее от времени. Однако это распределение метастабильно вследствие того факта, что флуктуации скоростей добавления и удаления не зависят друг от друга, и поэтому здесь нет саморегуляции. Этот пример тесно связан с задачей химической кинетики, которую мы обсудим ниже: для этого сформулируем ее в более реалистических терминах.

Вспомним «информационный ящик», введенный в § II.1. Этот ящик имеет полупроницаемые стенки, че-

рез которые могут проходить мономерные единицы в низко- или высокоэнергетическом состояниях, тогда как все макромолекулярные носители информации и весь аппарат репликации (ферменты и т. д.) остаются внутри ящика. Условия внутри ящика благоприятствуют синтезу макромолекул; однако здесь возможен только матричный синтез, т. е. репликация. *Сформулируем три задачи*, которые иллюстрируют стохастические аспекты отбора и будут количественно рассмотрены в следующем параграфе:

1. Дано большое число N различных последовательностей (для простоты — равной длины), причем каждая последовательность имеется лишь в одном экземпляре; образование новых последовательностей может происходить только путем матричного синтеза, причем редупликация является точной, т. е. ошибки отсутствуют ($Q_k = 1$). Все кинетические факторы образования и распада \mathcal{F}_k и \mathcal{R}_k равны, т. е.

$$W_i = E_i = W_k = E_k = 0 \quad (i, k = 1, 2, \dots, N).$$

(Равенство $W_i = E_i$ имеет место, поскольку $Q_i = 1$.) Согласно детерминистической теории, такая система находилась бы в стационарном состоянии, и в ней ничего не происходило бы. Это, очевидно, не может быть верным для каждой отдельной копии. Это может быть приблизительно верным для всего множества копий, по крайней мере на протяжении некоторого интервала времени, если только мы не различаем отдельные копии (все они вырождены по отношению к кинетическим свойствам). Наш вопрос заключается в том, какова истинная судьба суммарного содержимого ящика, а также отдельных классов носителей информации. (Необходимо подчеркнуть, что мы начинаем с N различных копий, причем все они могут представлять собой различные «сообщения».)

2. Условия те же, что и в первой задаче, но на этот раз допускается возможность ошибок при воспроизведении, т. е. $Q_k < 1$. Однако для всех копий будем, как и прежде, требовать, чтобы $W_i = W_k$ ($i, k = 1, 2, \dots, N$). Теперь из-за ошибок в процессе репликации могут

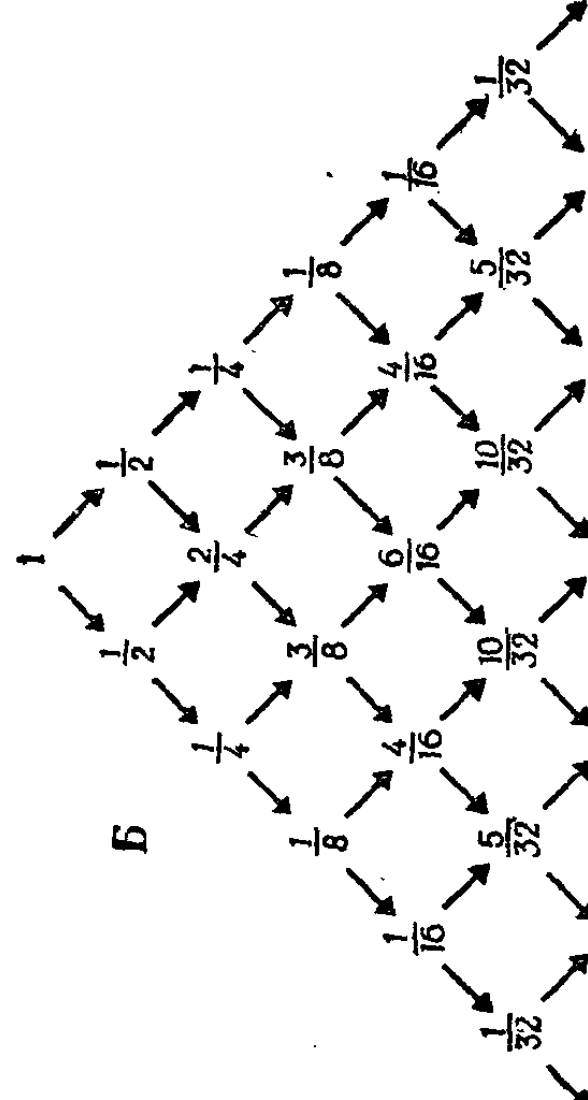
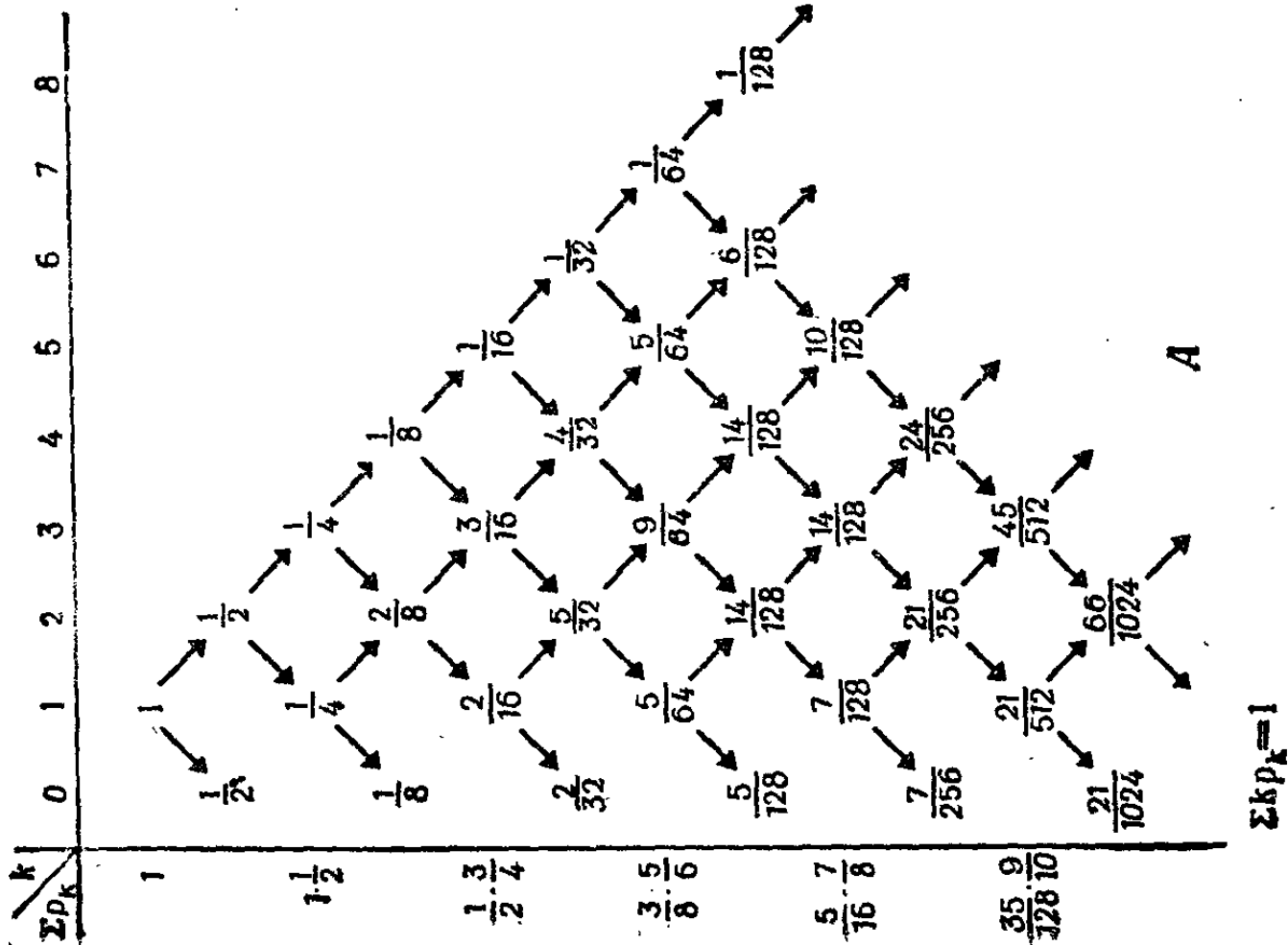


Рис. 6. Графическое представление вероятностей для модели стационарного состояния.

Числа k по горизонтали означают избыточность каждой копии. Отметим, что здесь представлены лишь вероятности, а не эволюция во времени. Каждый шаг означает изменение численности популяции на единицу; вероятности добавления и удаления одинаковы. Временные интервалы между переходами уменьшаются с ростом k — они обратно пропорциональны k . Такая форма представления данных выбрана для того, чтобы продемонстрировать асимметрию случайного блуждания, которая обусловлена «вымиранием» при $k=0$.

Если бы этой асимметрии не было, т. е. если бы каждое последующее состояние получалось из предыдущего с вероятностью $\frac{1}{2}$ (допуская также отрицательные значения k), то получился бы «треугольник Паскаля», в котором члены каждого горизонтального ряда представляют собой биномиальные коэффициенты, деленные на сумму всех коэффициентов, отвечающих этому ряду. Можно видеть, что вымирание изменяет вероятности несимметричным образом, так что в обоих треугольниках совпадают только правые границы.

Допустим, что мы хотим предсказать вероятность вымирания. Начнем с любого числа k и рассмотрим случайное блуждание в соответствующем треугольнике с вершиной в начальной точке. Так как вероятности перехода для каждого направления равны, то потребуются в среднем k^2 шагов, чтобы достичь состояния вымирания (т. е. $k=0$). Так как временные интервалы убывают обратно пропорционально k , можно предсказать, что для вымирания потребуется в среднем время, заключенное между k -кратным и k^2 -кратным элементарным временем $1/\mathcal{F}$. Искомое время должно быть ближе к k/\mathcal{F} , но больше его. При $t = k^2/\mathcal{F}$ вымирание должно наступить почти наверняка. Кратчайшее время, за которое в этой модели можно достичь нулевого состояния (начиная с k), дается выражением

$$\frac{1}{\mathcal{F}t} \left(\sum_{i=1}^k \frac{1}{i} \right) \rightarrow \frac{1}{\mathcal{F}t} (C + \ln k) \text{ для больших } k,$$

где C — постоянная Эйлера:

$$C = - \int_0^{\infty} e^{-t} \ln t \, dt \approx 0,5772.$$

возникать новые виды последовательностей. Нас снова интересует, как меняется во времени суммарное и индивидуальное информационное содержание.

3. Начиная с условий задачи 2, допустим, наконец, что кинетические параметры образования и распада, т. е. W_i, W_k ($i, k = 1, 2, \dots, N$), могут быть различны. Без регулирования потоков эта система может находиться в стационарном состоянии только вначале, когда средняя общая скорость образования равна средней общей скорости распада. Однако стационарное состояние можно поддерживать, если регулировать приток мономерных единиц и (или) растворителя в соответствии с теми ограничениями, которые были определены в гл. II. Здесь нас интересует эволюционное поведение индивидуальных видов, особенно если они возникают как одиночные мутантные копии, имеющие селективное преимущество $W_i > \bar{E}$.

Прежде чем начать математическое рассмотрение, попытаемся представить себе, какого поведения системы можно ожидать. В первой задаче система, очевидно, «замкнута» по отношению к добавлению любой новой информации (которая может быть введена только посредством ошибок воспроизведения). Вместе с тем всякий раз, когда отдельный избыточный носитель информации разлагается, не успев редуцироваться, происходит потеря информации. Это будет случаться постоянно, поэтому количество индивидуальной избыточной информации будет все время уменьшаться. Если суммарное число носителей информации очень велико, такие «отрицательные» флуктуации могут компенсироваться «положительными» флуктуациями (т. е. размножением других копий). Таким образом, информационное содержание будет сначала сужаться до нескольких (или даже до одной) весьма избыточных последовательностей, после чего *вся популяция в конце концов вымрет*. Такая судьба популяции в целом неизбежна, потому что флуктуации образования и распада происходят независимо друг от друга. Как показывает рис. 6 (обсуждение см. в подписи), полное вымирание может произойти даже за относительно короткое (по сравнению с соот-

ветствующими временами возврата для равновесных флуктуаций; см. § III. 2) время.

Поскольку мы не различаем отдельных носителей информации, мы можем ожидать, что точно такая же судьба постигнет всю популяцию во втором примере. Кинетические параметры для всех видов одинаковы и суммарное образование точно компенсируется суммарным распадом. Однако для отдельного носителя информации (т. е. индивидуального информационного содержания) эволюционное поведение значительно отличается от поведения в первом случае. Если факторы усиления A_k точно равны факторам разложения D_k , но факторы качества Q_k меньше единицы, каждый вид должен вымереть, потому что для каждого вида $W_k = A_k Q_k - D_k < 0$. В качестве компенсации эта система постоянно производит новую информацию посредством ошибок в процессе копирования. Она «беспорядочно дрейфует в информационном пространстве» до тех пор, пока, как и в первом случае, вся популяция не будет стерта «флуктуационной катастрофой».

Только в третьем случае мы можем ждать устойчивого и воспроизводимого поведения. Здесь система будет отбирать вид с максимальным W_k (допуская, что вообще имеется вид, для которого $W_k > 0$). До тех пор пока этот вид существует только в небольшом числе копий, он будет подвергаться опасности вымирания из-за флуктуаций. Однако, чем больше растет его численность, тем устойчивее он будет становиться и в конце концов станет доминировать в популяции в соответствии с детерминистическими уравнениями. Возникает интересный вопрос: имеется ли «точка, откуда нет возврата» в случайном блуждании, изображенном на рис. 6, если вероятности благоприятствуют росту? Каждому летчику знакома такая точка на взлетной полосе; когда он ее прошел, он должен взлететь. По аналогии мы можем искать такую «критическую» точку на оси концентраций, которая — если она пройдена мутантом — не оставляет возможности для «возврата».

Количественные ответы на все эти вопросы можно получить только из количественного теоретического рассмотрения.

§ III.4. Стохастические модели как цепи Маркова

Стохастическое исследование различных эволюционных моделей в настоящее время продолжается. Сейчас мы можем дать принципиальные ответы на большинство сформулированных выше вопросов, используя результаты, полученные ранее для некоторых простых линейных моделей.

А. Бартоломей [68] недавно изложил стохастическую теорию стационарных линейных процессов «рождения и гибели». Аналогичные проблемы рассматривали другие авторы: например, случай простой автокаталитической (прямой) реакции рассмотрен М. Дельбрюком еще в 1940 г. [69]. Существует обзор литературы по применениям теории случайных процессов к задачам химической кинетики (см. Д. Мак-Кварри [70]).

Следующее обсуждение будет основано на элегантном изложении Бартоломея [68], в котором он использовал метод Q -матрицы Дуба [71]. Элементы метода Дуба и процедура Бартоломея суммированы в табл. 9 и 10.

Таблица 9

Метод Q -матрицы Дуба для случайных процессов,
по А. Бартоломею [68]

I. Требования

«Функция переходных вероятностей» $p_{ij}(t)$ для перехода из состояния S_i в S_j ($i, j = 0, 1, 2, \dots$) определяется как условная вероятность

$$p\{x_{t_0+t} = S_j \mid x_{t_0} = S_i\}, \quad (1)$$

т. е. вероятность того, что случайная переменная x_t будет иметь значение S_j в момент времени $t_0 + t$, если она имела значение S_i в момент времени t_0 . Процесс $\{x_t, 0 \leq t < \infty\}$ называется стационарной цепью Маркова, если вероятности перехода удовлетворяют следующим условиям:

$$p_{ij}(t) \geq 0 \quad (i, j = 0, 1, 2, \dots), \quad (2)$$

$$\sum_j p_{ij}(t) = 1 \quad (i = 0, 1, 2, \dots), \quad (3)$$

$$\sum_j p_{ij}(s) p_{jk}(t) = p_{ik}(s+t) \quad (i, k = 0, 1, 2, \dots), \quad (4)$$

$$\lim_{t \rightarrow 0} p_{ij}(t) = \delta_{ij} = \begin{cases} 1 & \text{для } i = j, \\ 0 & \text{для } i \neq j. \end{cases} \quad (5)$$

Продолжение табл. 9

II. Определение матрицы $Q = (q_{ij})$

$$q_{ii} = \lim_{t \rightarrow 0} \frac{p_{ii}(t) - 1}{t} \equiv \frac{dp_{ii}}{dt} \quad (t = 0), \quad (6)$$

$$q_{ij} = \lim_{t \rightarrow 0} \frac{p_{ij}(t)}{t} \equiv \frac{dp_{ij}}{dt} \quad (t = 0). \quad (7)$$

Требования I и определения II используются для построения следующей системы дифференциальных уравнений ($\dot{p} = dp/dt$):

III. а) Прямая система

$$\dot{p}_{ik}(t) = q_{kk}p_{ik}(t) + \sum_{j \neq k} q_{jk}p_{ij}(t). \quad (8)$$

б) Обратная система

$$\dot{p}_{ik}(t) = q_{ii}p_{ik}(t) + \sum_{j \neq i} q_{ij}p_{jk}(t). \quad (9)$$

Прямая система описывает, что происходит в последнем временном интервале ($t \rightarrow 0$) до перехода; обратная система описывает, что происходит в первом временном интервале после перехода. Это отражено в суммах: в случае прямой системы варьируют конечное состояние, в случае обратной системы варьируют начальное состояние.

IV. Общее решение

Для конечного числа состояний и при данных начальных условиях, например при условии (5), единственное решение для обеих систем может быть дано в матричной форме, как показано Дубом [71]:

$$P(t) = e^{tQ}. \quad (10)$$

$P = (p_{ij})$ и $Q = (q_{ij})$ — матрицы, e^{tQ} получается последовательным суммированием членов разложения экспоненты в ряд

$$1 + tQ + t^2Q^2/2! + \dots \quad (11)$$

Для определения величин $p_{ij}(t)$ и соответствующих q_{ij} некоторым свойствам детерминистических уравнений дается вероятностная интерпретация (ср. табл. 10).

Таблица 10

Линейные процессы рождения и гибели как цепи Маркова, по А. Бартоломею [68]

Дано детерминистическое уравнение

$$\dot{x} = (\mathcal{F} - \mathcal{R})x \quad (1)$$

и его решение

$$x(t) = x(t=0) e^{(\mathcal{F} - \mathcal{R})t}, \quad (2)$$

где \mathcal{F} и \mathcal{R} — константы.

Продолжение табл. 10

Для малого интервала времени между t_1 и $t_1 + \Delta t_1$ результирующее изменение Δx_1 в популяции равно

$$\frac{\Delta x_1}{x_1} = \mathcal{F} \Delta t_1 - \mathcal{R} \Delta t_1 + 0(\Delta t_1), \quad (3)$$

где $0(\Delta t_1)$ включает в себя все бесконечно малые высших порядков. Δt_1 выбирается достаточно малым, чтобы в этом временном интервале имели конечное значение лишь вероятности образования или распада одной копии, которые даются выражениями

$$\mathcal{F} x_1 \Delta t_1 + 0(\Delta t_1) \quad \text{или} \quad \mathcal{R} x_1 \Delta t_1 + 0(\Delta t_1). \quad (4)$$

Тогда численность популяции может измениться только на плюс или минус единицу; таким образом, разрешены только переходы $S_i \rightarrow S_{i-1}$ или $S_i \rightarrow S_{i+1}$ ($i = 1, 2, \dots$), и эти переходы могут происходить только в результате одного элементарного события (множественные рождения и гибели, дающие в сумме изменение на ± 1 , исключаются). Далее, переход $S_0 \rightarrow S_1$ имеет нулевую вероятность, так как, достигнув состояния S_0 , система «вымирает». Используя разложение $p_{ij}(t)$ в ряд

$$p_{ij}(\Delta t_1) = p_{ij}(0) + \dot{p}_{ij}(0) \Delta t_1 + \ddot{p}_{ij}(0) \frac{(\Delta t_1)^2}{2} + \dots,$$

строим Q -матрицу согласно уравнениям (6) и (7) в табл. 9 и после сравнения с выражениями (4) в настоящей таблице получим

$$p_{i, i-1}(\Delta t_1) = q_{i, i-1} \Delta t_1 + 0(\Delta t_1), \quad (5)$$

$$p_{i, i+1}(\Delta t_1) = q_{i, i+1} \Delta t_1 + 0(\Delta t_1), \quad (6)$$

$$q_{i, i-1} = i\mathcal{R}; \quad q_{i, i+1} = i\mathcal{F}; \quad q_{ii} = -i(\mathcal{F} + \mathcal{R}) \quad (7)$$

($i = 0, 1, 2, \dots$ — состояния популяции).

Теперь можно построить «прямые» и «обратные» уравнения, приведенные в табл. 9:

$$\dot{p}_{ik}(t) = -k(\mathcal{F} + \mathcal{R}) p_{ik}(t) + (k-1)\mathcal{F} p_{i, k-1}(t) + (k+1)\mathcal{R} p_{i, k+1}(t), \quad (8)$$

$$\dot{p}_{ik}(t) = -i(\mathcal{F} + \mathcal{R}) p_{ik}(t) + i\mathcal{F} p_{i+1, k}(t) + i\mathcal{R} p_{i-1, k}(t). \quad (9)$$

Чтобы решить эти уравнения, введем «производящую функцию вероятностей»:

$$\phi_i(s, t) = \sum s^k p_{ik}(t). \quad (10)$$

Из $\partial\phi/\partial s$ и $\partial\phi/\partial t$ получим

$$\partial\phi_i/\partial t - (s-1)(\mathcal{F}s - \mathcal{R}) \partial\phi_i/\partial s = 0 \quad (11)$$

с дополнительным уравнением

$$\frac{dt}{-1} = \frac{ds}{(s-1)(\mathcal{F}s - \mathcal{R})}, \quad (12)$$

Продолжение табл. 10

интегрирование которого для двух случаев $\mathcal{F} \neq \mathcal{R}$ и $\mathcal{F} = \mathcal{R}$ ведет к общему решению уравнения в частных производных для производящей функции:

1. $\mathcal{F} \neq \mathcal{R}$

$$\phi_i(s, t) = \left[\frac{(\mathcal{R}e^{(\mathcal{F}-\mathcal{R})t} - \mathcal{R}) - (\mathcal{R}e^{(\mathcal{F}-\mathcal{R})t} - \mathcal{F})s}{(\mathcal{F}e^{(\mathcal{F}-\mathcal{R})t} - \mathcal{R}) - (\mathcal{F}e^{(\mathcal{F}-\mathcal{R})t} - \mathcal{F})s} \right]^i. \quad (13)$$

Разложение по степеням s дает значения коэффициентов при s^k , которые, по формуле (10), равны соответствующим вероятностям:

$$p_{i k}(t) = \sum_{n=0}^{k \text{ или } i} (-1)^n \binom{i}{n} \binom{i+k-n-1}{k-n} \mathcal{F}^{k-n} \mathcal{R}^{i-n} \times \\ \times [e^{(\mathcal{F}-\mathcal{R})t} - 1]^{i+k-2n} [\mathcal{F}e^{(\mathcal{F}-\mathcal{R})t} - \mathcal{R}]^{-i-k+n} [\mathcal{R}e^{(\mathcal{F}-\mathcal{R})t} - \mathcal{F}]^n. \quad (14)$$

Верхний предел суммирования равен k , если $0 < k < i$, и равен i , если $k \geq i$. Мы имеем

$$\lim_{t \rightarrow 0} p_{i k}(t) = \begin{cases} 1 & (k = i), \\ 0 & (k \neq i). \end{cases} \quad (15)$$

Математическое ожидание i равно

$$\varepsilon_i(t) = \left(\frac{\partial \phi_i}{\partial s} \right)_{s=1} = i e^{(\mathcal{F}-\mathcal{R})t}, \quad (16)$$

т. е. совпадает с результатом детерминистической теории; дисперсия i равна

$$\sigma_i^2(t) = \left(\frac{\partial^2 \phi_i}{\partial s^2} \right)_{s=1} + \varepsilon_i(t) - \varepsilon_i^2(t) = \\ = i \frac{\mathcal{F} + \mathcal{R}}{\mathcal{F} - \mathcal{R}} e^{(\mathcal{F}-\mathcal{R})t} [e^{(\mathcal{F}-\mathcal{R})t} - 1]. \quad (17)$$

2. $\mathcal{F} = \mathcal{R}$. Решение аналогично:

$$\phi_i(s, t) = \left[\frac{\mathcal{F}t - (\mathcal{F}t - 1)s}{(1 + \mathcal{F}t) - \mathcal{F}ts} \right]^i; \quad (18)$$

$$p_{i k}(t) = \sum_{n=0}^{k \text{ или } i} (-1)^n \binom{i}{n} \binom{i+k-n-1}{k-n} (\mathcal{F}t)^{i+k-2n} \times \\ \times (\mathcal{F}t - 1)^n (\mathcal{F}t + 1)^{-i-k+n}; \quad (19)$$

$$\varepsilon_i(t) = \left(\frac{\partial \phi_i}{\partial s} \Big|_{s=1} \right) = i; \quad (20)$$

$$\left(\frac{\partial^2 \phi_i}{\partial s^2} \Big|_{s=1} \right) = i(2\mathcal{F}t + i - 1). \quad (21)$$

Рассматриваемый процесс, детерминистический аспект которого выражен уравнением (II. 10) с $\varphi_{i1} = 0$, \mathcal{F}_i и $\mathcal{R}_i = \text{const}$, представляет собой стационарную цепь Маркова

$$\{x_t, 0 \leq t < \infty\}. \quad (\text{III. 5})$$

Время (t) является непрерывным параметром; x_t — случайная популяционная переменная — принадлежит дискретному счетному множеству состояний $S_0, S_1 \dots$. Переход из одного состояния (S_i) в другое (S_j) описывается вероятностью перехода $p_{ij}(t)$. Это вероятность того, что система, находящаяся в состоянии S_i при $t = 0$, достигает состояния S_j в момент времени t . Знание $p_{ij}(t)$ позволяет определить математическое ожидание и дисперсию для населенности любого состояния во время t , если заданы начальные условия для $t = 0$. Вероятности перехода для линейного процесса рождения и гибели, вычисленные в табл. 10, дают основу для дальнейшего обсуждения.

§ III. 5. Количественное рассмотрение трех прототипов отбора

Случай 1. Вернемся назад к первой из трех задач о стационарном состоянии (см. § III. 3) и спросим, какова вероятность выживания или вымирания системы, которая определяется параметрами $\mathcal{F}_k \equiv \mathcal{F} = \mathcal{R}_k \equiv \mathcal{R}$ (для всех видов k , имеющих в наличии). Если при $t = 0$ вся популяция состояла из N видов, вероятность вымирания (для $\mathcal{F} = \mathcal{R}$) следует из выражения (19), приведенного в табл. 10:

$$p_{N_0}(t) = \left[\frac{\mathcal{F}t}{1 + \mathcal{F}t} \right]^N \quad (\text{III. 6})$$

$$\rightarrow e^{-N/\mathcal{F}t} \quad \text{для } \mathcal{F}t \gg 1. \quad (\text{III. 7})$$

Эта вероятность близка к единице для $\mathcal{F}t \gg N$. Уже при $\mathcal{F}t = N$ мы имеем

$$p_{N_0}(t = N/\mathcal{F}) = 1/e, \quad (\text{III. 8})$$

т. е. вероятность вымирания популяции больше $1/3$. При $\mathcal{F}t = N^2$ имеем

$$p_{N_0}(t = N^2/\mathcal{F}) = 1 - 1/N \quad (\text{III. 9})$$

— вымирание почти достоверно.

Математическое ожидание величины популяции остается равным N независимо от t [см. выражение (20) в табл. 10]. Однако дисперсия растет со временем. При $t = N/\mathcal{F}$ она становится равной

$$\sigma_N^2(t = N/\mathcal{F}) = 2N^2. \quad (\text{III. 10})$$

Итак, мы нашли, как и предполагали ранее (см. подпись к рис. 6), что стационарная популяция, состоящая при $t = 0$ из N видов, имеет среднюю продолжительность жизни, заключенную между N/\mathcal{F} и N^2/\mathcal{F} . При $t = N^2/\mathcal{F}$ вымирание почти достоверно.

Посмотрим теперь, какова судьба N различных видов, если вначале каждый из них имелся только в одном экземпляре.

Согласно выражению (19) в табл. 10, вероятность того, что единственная копия размножится в количестве k экземпляров, равна

$$p_{1k}(t) = \frac{1}{(\mathcal{F}t)^2} \frac{1}{(1 + 1/\mathcal{F}t)^{k+1}} = \frac{1}{\mathcal{F}t(1 + \mathcal{F}t)} \left[\frac{\mathcal{F}t}{1 + \mathcal{F}t} \right]^k \quad (\text{III. 11})$$

(для $k \geq 1$), и вероятность того, что этот вид вымрет, равна

$$p_{10}(t) = \frac{\mathcal{F}t}{1 + \mathcal{F}t}. \quad (\text{III. 12})$$

Просуммировав уравнения (III. 11) по всем $k \geq 1$ и добавив к сумме уравнение (III. 12), получим

$$\sum_{k=0}^{\infty} p_{1k}(t) = \frac{1}{1 + \mathcal{F}t} + \frac{\mathcal{F}t}{1 + \mathcal{F}t} = 1 \quad (\text{III. 13})$$

в согласии с уравнением (3) в табл. 9.

Для любого $k \geq 1$, $p_{1k}(t)$ как функция t проходит через максимум при

$$\mathcal{F}t = \frac{k-1}{2}. \quad (\text{III. 14})$$

При этом p_{1k} достигает значения

$$(p_{1k})_{\max} = \left(\frac{2}{ek}\right)^2 \quad (\text{III. 15})$$

(e — основание натуральных логарифмов). Если нас интересует, например, в какое время и для какого k максимальное p_{1k} будет равно $1/N$, мы получим

$$k = \frac{2}{e} \sqrt{N} \quad (\text{III. 16})$$

или

$$t = \frac{1}{\mathcal{F}} \left(\frac{\sqrt{N}}{e} - \frac{1}{2} \right). \quad (\text{III. 17})$$

Будем рассматривать $p_{1k}(t)$ как функцию k . Для $\mathcal{F}t \gg 1$ получается асимптотическое выражение

$$p_{1k}(t) \rightarrow \frac{e^{-k/\mathcal{F}t}}{(\mathcal{F}t)^2} \quad (k \geq 1). \quad (\text{III. 18})$$

Это распределение, т. е. $p_{1k}(t)$ как функция k (в логарифмической шкале) при данном t_1 , показано для

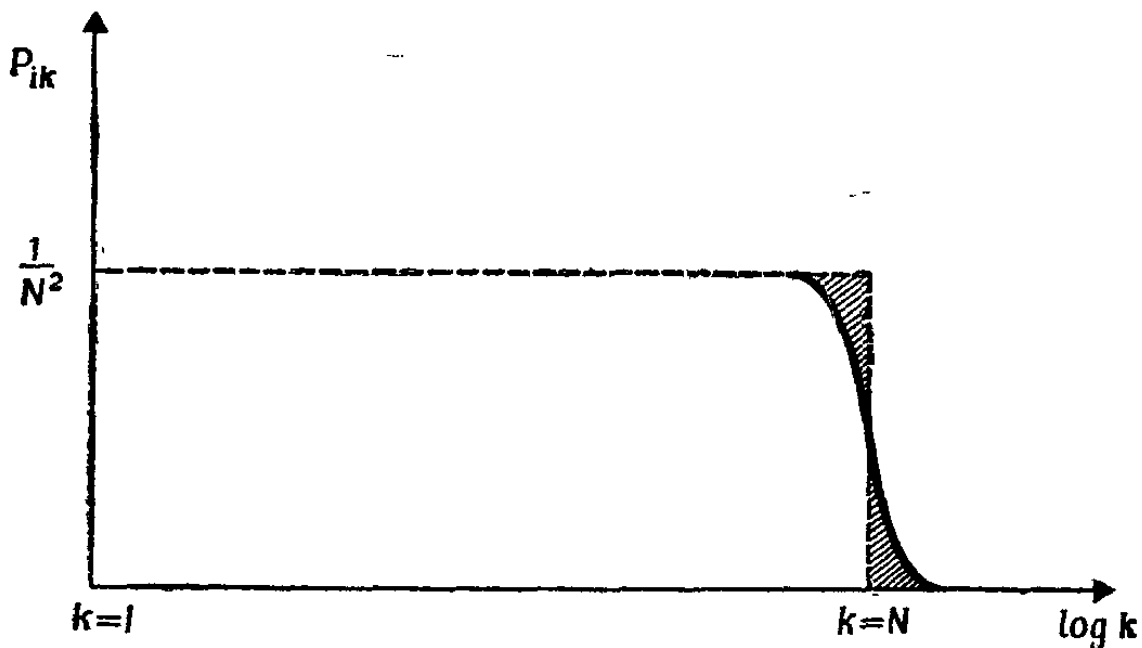


Рис. 7. Распределение вероятностей для «выживших» при $t_1 = N/\mathcal{F}$, согласно (III. 18).

$t_1 = N/\mathcal{F}$ на рис. 7. Как можно видеть, вероятности почти не зависят от k — $p_{1k}(N/\mathcal{F}) = 1/N^2$, пока k не достигнет N по порядку величины, после чего начинается экспоненциальный спад. В то же время вероятность выми-

рания равна

$$p_{10} = 1 - 1/N. \quad (\text{III. 19})$$

Учитывая, что

$$\int_0^{\infty} e^{-k/N} dk = N, \quad (\text{III. 20})$$

распределение вероятностей можно аппроксимировать константой $1/N^2$ в интервале от $k = 1$ до N , причем при $k = 0$ происходит скачок до значения $1 - 1/N$. Итак, начиная с одной копии при $t = 0$ имеются примерно равные шансы для любой степени усиления k от 1 до N . Так как вначале ($t = 0$) мы имели N различных отдельных копий, вероятность, что одна из них усилится до численности $> N/2$, уже равна 50% (т. е. $N \sum_{k=N/2}^{\infty} p_{1k} = 0,5$) при $t = N/\mathcal{F}$. В то же время большинство других копий вымирает, так как

$$p_{10} = 1 - 1/N.$$

Математическое ожидание $\varepsilon = \sum_{k=0}^{\infty} kp_k$ (p_k находятся по формуле III. 11) для каждого вида снова остается равным единице и не зависит от времени, тогда как дисперсия для каждого вида растет; например, для $\mathcal{F}t = N$ она равна

$$\sigma_1^2(t = N/\mathcal{F}) = 2N. \quad (\text{III. 21})$$

В заключение отметим:

Разнообразие носителей информации сужается до нескольких, или даже до одного, весьма избыточного вида, после чего все носители информации вымирают. Такое селекционное поведение, вызванное автокаталитической природой процесса образования, представляет собой типичный случай «выживания выживших». В группе, где все индивидуумы имеют равные шансы на выживание, нет другого критерия отбора, кроме его исхода, т. е. самого факта выживания.

Этот тип чисто «стохастического» выживания в стационарном состоянии нереален, потому что фактор качества никогда не может быть точно равным единице.

Его следует отличать от отбора дарвиновского типа, для которого можно сформулировать оптимизационные критерии выживания.

Случай 2. Во второй задаче, предложенной в § III.3, необходимо качественно различать решения, относящиеся ко всей популяции и к отдельным индивидуумам. Для всей популяции мы можем положить $\mathcal{F} = \mathcal{R}$ (не различая отдельные виды). Таким образом, поведение популяции в целом будет точно таким, как описано выше, т. е. произойдет полное вымирание за время, по порядку величины превосходящее N^2/\mathcal{F} .

Для каждого отдельного вида, однако, необходимо положить $\mathcal{F}_k < \mathcal{R}_k$, поскольку $Q_k < 1$; при этом снова предполагается, что все отдельные виды вырождены по отношению к \mathcal{F}_k , \mathcal{R}_k или Q_k . Тогда выражение (14) в табл. 10 показывает, что p_{10} приближается к 1 для $t \gg 1/(\mathcal{R}_k - \mathcal{F}_k)$:

$$p_{10}(t) \rightarrow 1 - (1 - \mathcal{F}_k/\mathcal{R}_k) e^{-(\mathcal{R}_k - \mathcal{F}_k)t}. \quad (\text{III. 22})$$

Математическое ожидание для каждой копии уменьшается от 1 до 0:

$$\varepsilon_i(t) = e^{-(\mathcal{R}_k - \mathcal{F}_k)t}, \quad (\text{III. 23})$$

а дисперсия стремится к

$$\sigma_1^2(t) \rightarrow \frac{\mathcal{R}_k + \mathcal{F}_k}{\mathcal{R}_k - \mathcal{F}_k} e^{-(\mathcal{R}_k - \mathcal{F}_k)t}. \quad (\text{III. 24})$$

Каждый отдельный вид вымирает. Ни один вид не сможет выжить на протяжении времени, сравнимого с временем жизни «выживших» в первом примере. Вместо этого будет наблюдаться постоянный дрейф информационного содержания вследствие ошибок воспроизведения. За свое время жизни (т. е. за время, меньшее чем N^2/\mathcal{F}) система успеет просканировать большой объем информации, не задерживаясь на воспроизведении какой-либо определенной копии в течение времени, существенно превышающего $\frac{1}{\mathcal{R}_k - \mathcal{F}_k}$. Однако, как сказано выше, система в целом будет иметь ту же судьбу, что и в случае 1.

Случай 3. Устойчивый и предсказуемый отбор может наблюдаться только в том случае, если кинетические

параметры изменяются в конечных пределах. Однако даже в этом случае мы имеем область «неопределенности» для отбора. Во-первых, для $\mathcal{F} > \mathcal{R}$ получим математическое ожидание для растущей популяции

$$\varepsilon_i(t) = ie^{(\mathcal{F} - \mathcal{R})t} \quad [\text{ср. также с уравнением (III.23)}]$$

в согласии с детерминистической теорией. Более того, для больших i при $t = 0$ флуктуации имеют величину порядка \sqrt{i} (i — начальное значение). Рассмотрим теперь равновесное распределение с $W_m = \bar{E}$. Виды, обладающие конечным «селективным преимуществом» $W_{m+1} > W_m$, имеют ненулевые, но не стопроцентные шансы быть отобранными. Согласно уравнению (14) в табл. 10, вероятность вымирания вида, если при $t = 0$ он был представлен k копиями, равна

$$p_{k0}(t) = \left[1 - \frac{\mathcal{F}_m - \mathcal{R}_m}{\mathcal{F}_m - \mathcal{R}_m \exp[(\mathcal{R}_m - \mathcal{F}_m)t]} \right]^k, \quad (\text{III.25})$$

при $t \rightarrow \infty$

$$\lim_{t \rightarrow \infty} p_{k0}(t) = \left(\frac{\mathcal{R}_m}{\mathcal{F}_m} \right)^k. \quad (\text{III.26})$$

Этот результат дает точный ответ на последний из вопросов § III.3. Если в предварительно уравновешенной популяции (т. е. в такой, что $W_m = \bar{E} = 0$) возникнет мутант ($m+1$) с селективным преимуществом (т. е. $W_{m+1} > W_m$), то одна первоначально возникшая копия может в конце концов погибнуть с вероятностью $\mathcal{R}_m/\mathcal{F}_m$. Вероятность ее выживания равна $1 - \mathcal{R}_m/\mathcal{F}_m$, и эта вероятность увеличивается с ростом избыточности (k) мутантных копий. Однако в линейной модели нет истинной «точки, откуда нет возврата». Вероятность вымирания для $\mathcal{F}_{m+1} > \mathcal{R}_{m+1}$ по уравнению III.26 монотонно убывает с ростом k и приближается к нулю лишь асимптотически при больших k .

Однако мы можем определить пороговое значение вероятности как «точку полувозврата» $k_{1/2}$ согласно условию

$$\left(\frac{\mathcal{R}_m}{\mathcal{F}_m} \right)^{k_{1/2}} = 0,5 \quad (\text{III.27})$$

или как соответствующую «точку релаксации» $k_{1/e}$

$$k_{1/e} = \frac{1}{\ln(\mathcal{F}_m/\mathcal{R}_m)}, \quad (\text{III. 28})$$

что при условии $(\mathcal{F}_m - \mathcal{R}_m) \ll \mathcal{R}_m$ дает

$$k_{1/e} = \frac{\mathcal{R}_m}{\mathcal{F}_m - \mathcal{R}_m}. \quad (\text{III. 29})$$

При этом же условии $\lim_{t \rightarrow \infty} p_{k0}$ может быть выражен как

$$\left(\frac{\mathcal{R}_m}{\mathcal{F}_m}\right)^k \rightarrow e^{-k \frac{\mathcal{F}_m - \mathcal{R}_m}{\mathcal{R}_m}}. \quad (\text{III. 30})$$

Табл. 11 дает некоторые значения для точек релаксации $k_{1/e}$ при различных $\mathcal{F}_m/\mathcal{R}_m$. Мы видим, что *мелкие селективные преимущества редко приводят к выживанию и доминированию, как это предсказывает детерминистическая теория. Этот факт еще раз подчеркивает случайный характер процессов отбора. Стохастическое выражение для частоты появления одного определенного мутанта следует умножить на $(1 - \lim_{t \rightarrow \infty} p_{10})$, чтобы получилось выражение для вероятности его макроскопического появления.*

Сделанные здесь выводы относятся лишь к линейным растущим системам. В настоящее время производятся расчеты для систем с ограничениями постоянных сил или потоков, а также для (реальных) нелинейных ра-

Таблица 11

Стохастический порог для выживания мутанта с селективным преимуществом $\mathcal{F}_i > \mathcal{R}_i$

| $\mathcal{F}_i/\mathcal{R}_i$ | $(\mathcal{F}_i - \mathcal{R}_i)/\mathcal{R}_i$ | $k_{1/e}$ |
|-------------------------------|---|-----------|
| 2 | 1 | 1,44 |
| 1,3 | 0,3 | 3,82 |
| 1,1 | 0,1 | 10,5 |
| 1,01 | 10^{-2} | 10^2 |
| 1,00001 | 10^{-5} | 10^5 |

ствующих систем, которые оказываются особенно интересными в связи с «нуклеацией» живых систем (см. гл. VI).

Стохастическое рассмотрение, которое существенно опирается на линейную модель рождения — гибели Бартоломея и которое сейчас распространяется на истинные стационарные состояния, приводит к некоторым важным модификациям детерминистической феноменологической теории эволюции. Оно не только подчеркивает случайную природу элементарных процессов, но и очень ясно демонстрирует, что определенные утверждения, выведенные из детерминистической теории, должны быть модифицированы для того, чтобы они правильно описывали существенные особенности эволюционных процессов.

САМООРГАНИЗАЦИЯ, ОСНОВАННАЯ НА КОМПЛЕМЕНТАРНОМ УЗНАВАНИИ: НУКЛЕИНОВЫЕ КИСЛОТЫ

§ IV. 1. Истинное «самоинструирование»

Теория отбора — имеющая довольно общий характер — до сих пор детально рассматривалась только для простых квазилинейных систем. «Квазилинейной» мы можем назвать любую систему, описываемую уравнением (II.32), например простое выражение, которое получится, если в (II.32) отбросить члены, отвечающие обратному потоку:

$$\dot{x}_i = k_0 (W_i^0 - \bar{E}) x_i, \quad (\text{IV. 1})$$

и в котором «селективная ценность» W_i является константой. Отметим, что ни исходное кинетическое уравнение для роста без дополнительных ограничений, ни его окончательная форма, в которой учитываются ограничения, налагаемые отбором, не может быть действительно линейным дифференциальным уравнением. Член, описывающий «образование» в исходном уравнении, содержит «стехиометрическую функцию» $f_i(m_1 \dots m_\lambda)$, т. е. функцию концентраций (в общем случае переменных) высокоэнергетических мономеров, точный вид которой зависит от механизма матричного процесса полимеризации, умноженную на автокаталитический член x_i . Только если высокоэнергетические мономеры забуферены согласно условию постоянной общей организации, селективную ценность можно считать постоянной в случае простого «самоинструированного» процесса. То же условие, однако, вводит в дифференциальное уравнение функцию E , которая зависит от концентраций (т. е. содержит x_i). Таким образом, «квазилинейность» может относиться только к члену $W_i^0 x_i$, означая, что W_i^0 те-

перь уже не зависит от x_i или от каких-либо x_k (которые в системе кинетических уравнений являются переменными).

При этих условиях уравнение (IV.1) описывает простейший случай истинного «самоинструктирования», т. е. образование специфичной последовательности i инструктируется самой матрицей i .

Как можно себе представить наличие процессов самоинструктирования такого типа — в качестве некоего общего явления — в природе? Мы знаем, конечно, много разных автокаталитических процессов, когда определенный продукт реакции участвует в своем собственном образовании. Однако здесь мы требуем большего: *любой* продукт процесса полимеризации, т. е. *любая* определенная последовательность, должна инструктировать образование *своей собственной* копии.

Простым примером может служить образование полирибоадениловой кислоты (поли-рибо А) при низких рН. Известно, что при рН ниже 4 поли-рибо А образует двухцепочечную спиральную структуру вследствие специфического спаривания протонированных остатков аденина. В отличие от структуры, постулированной Уотсоном и Криком, эта двойная спираль содержит параллельно ориентированные цепи, идущие от 3'-конца к 5'-концу. Кроме того, протонирование оснований в сущности нейтрализует отрицательные заряды фосфатных групп остова, вследствие чего эта структура более устойчива при низкой ионной силе. В других отношениях эта спираль ведет себя как структура Уотсона — Крика. Такой вывод можно сделать на основании подробных термодинамических и кинетических исследований, выполненных в нашей лаборатории [73]. Сходное явление наблюдается в случае поли-рибо Ц, где спаривание требует наличия как протонированных, так и непротонированных групп, так что существование двойной (и тройной) спирали ограничено очень узким диапазоном рН.

Во всяком случае легко можно представить себе «самоинструктирующуюся» матрицу, которая направляет синтез *идентичной* копии; такая система описывается кинетическими уравнениями типа (IV.1) с постоянными селективными ценностями W_i^0 . Проанализируем теперь

более детально параметры \mathcal{A}_i , Q_i и \mathcal{D}_i , которые определяют процессы отбора и самоорганизации в системе.

Точный вид фактора усиления \mathcal{A}_i зависит от механизма матричной полимеризации. Некоторые механизмы, в основу которых были положены стохастические модели, обсуждались Дж. Гиббсом [74].

Простейшая модель дает один и тот же временной интервал для включения каждой единицы в полимерную цепь. Если редупликация каждой цепи должна быть закончена, прежде чем может быть начата новая цепь, то постоянная времени редупликации должна расти с длиной цепи. Она пропорциональна v (числу единиц в цепи) в начале процесса и становится пропорциональной v^2 , когда процесс полимеризации приближается к равновесию. Здесь вероятности прямой и обратной реакций становятся равными, так что реакция напоминает процесс простой одномерной диффузии (где время перемещения пропорционально квадрату расстояния). Для наших целей мы можем не рассматривать системы, близкие к равновесию, и это исключает возможность, что зависимость от длины цепи будет сильнее, чем линейная.

С другой стороны, существуют механизмы, которые дают значительно более слабую зависимость от длины цепи:

а) Процесс может быть кооперативным и требовать определенного времени (и определенной длины цепи) для нуклеации, которое может оказаться гораздо больше, чем время «распространения». Тогда — вплоть до некоторой «кооперативной длины» — время редупликации будет определяться временем нуклеации и, следовательно, не будет зависеть от длины. Такое кооперативное поведение имеет место, например, при узнавании оснований в процессе образования (двойной) спирали олиго-рибо А, где «длина нуклеации» равна 3, а «кооперативная длина» — примерно 30 парам оснований (при комнатной температуре).

б) Новая цепь может быть начата как на матрице, так и на реплике, прежде чем реплика будет закончена. Как хорошо известно, это явление имеет место при ферментативном считывании одноцепочечной матрицы (на-

пример, рибосомами или РНК-полимеразами), и вполне возможно, что оно происходит также при неферментативной репродукции одиночной цепи. Дж. Гиббс [74] показал, что в таком процессе множественной репродукции могут возникать «пробки» — как в уличном движении. В отличие от случая (а) зависимость от длины цепи будет здесь проявляться только для сравнительно коротких последовательностей, исчезая по достижении некоторой определенной длины.

Если объединить случаи «а» и «б», то можно получить в результате очень слабую (слабее линейной) зависимость от длины цепи, которая если и благоприятствует воспроизведению коротких последовательностей, то лишь в очень слабой степени. Могут существовать, однако, очень резкие различия в скорости для различных длинных последовательностей, потому что складывание цепи и образование петель может создать набор участков, на которых матричные эффективности будут сильно различаться. Далее, если в конкуренции участвуют различные высокоэнергетические мономерные единицы, концентрации которых различны, то единица, имеющая наибольшую концентрацию, будет включаться в цепь с большей вероятностью (поскольку это влияет на A_i и Q_i), в результате чего могут получиться очень однородные последовательности. Если в репликации участвуют ферменты¹, то в узнавании (как и в случае «нуклеации») могут участвовать специфичные участки матричной последовательности (возможно, что участвуют оба конца цепи, как в случае Q β -репликазы; см. гл. VII).

Фактор Q_{i0} накладывает более жесткие ограничения на длину цепи, чем фактор усиления, вследствие степенной зависимости

$$Q_{i0} = \prod_{k=1}^{\lambda} q_k^{v_{ik}}, \quad (\text{IV.2})$$

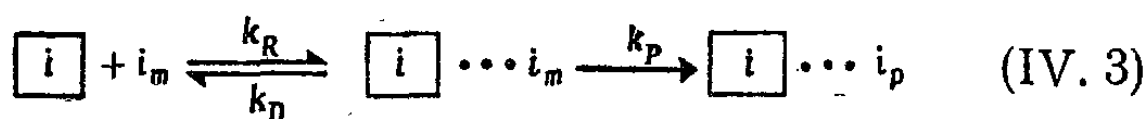
где индекс k относится к различным единицам 1, ..., λ . Если бы факторы узнавания отдельных единиц q_i были

¹ Процессы, катализируемые ферментами, можно рассматривать как «квазилинейные», если фермент (например, репликаза) представляет собой постоянный «фактор среды», т. е. не является частью «эволюционирующей» системы (см. гл. VII).

равны, то можно было бы применить простую модель, которая обсуждалась в табл. 8 (см. § II. 5).

Как было показано, максимальное число единиц, которое может воспроизводимо копироваться, ограничено неопределенностью узнавания отдельных единиц и (слабо) зависит также от «разброса» кинетических коэффициентов. Вследствие этого, когда «единичный» (т. е. отвечающий одной единице) фактор качества равен, например, 0,99 (т. е. вероятность ошибки равна 1%), длина воспроизводимых последовательностей ограничена числом единиц v_i порядка 100 (или даже меньше, если «разброс» $\mathcal{A}_i - \bar{\mathcal{A}}_{k \neq i}$ и $\mathcal{D}_i - \bar{\mathcal{D}}_{k \neq i}$ близок к нулю).

Единичный фактор качества q можно выразить через свободные энергии парного взаимодействия, если узнавание является «равновесным» процессом. Рассмотрим такую стадию самоинструктированного включения мономера:



\boxed{i} представляет собой единицу матрицы, i_m — высокоэнергетическую мономерную единицу и i_p — единицу, включенную в полимерную копию; k_R , k_D и k_P — константы скорости процессов спаривания, диссоциации пар и включения единиц (полимеризации). «Равновесное» узнавание означает, что

$$k_P \ll k_D. \quad (\text{IV. 4})$$

Измерения, описанные ниже, показывают, что кооперативное спаривание (т. е. распространение «нуклеированной» области) происходит с большой скоростью (10^6 — 10^7 s^{-1}), и можно принять, что k_D больше 10^5 s^{-1} (для ГЦ) или 10^6 s^{-1} (для АУ).

Предполагая «равновесное узнавание», можно записать

$$q_i = \frac{[\text{Вероятность пары } ii]}{[\text{Сумма вероятностей для всех пар } i]} = \frac{m_i K_{ii}}{\sum_{k=1}^{\lambda} m_k K_{ik}}, \quad (\text{IV. 5})$$

где m_i или m_k — концентрации мономерных единиц и K_{ii} или K_{ik} — соответствующие (кооперативные) константы стабильности пары. Если равновесность узнавания не полная, то константы стабильности можно заменить на подходящие константы стационарного состояния. Константы стабильности K_{ik} можно выразить также через свободные энергии образования пары:

$$K_{ik} = \exp(-\Delta G_{ik}/RT). \quad (\text{IV. 6})$$

Только в том случае, когда концентрации всех мономерных единиц забуферены до одного и того же значения, все m_k сокращаются и получается простое выражение

$$1/q_i = \sum_{k=1}^{\lambda} \exp[(\Delta G_{ii} - \Delta G_{ik})/RT]. \quad (\text{IV. 7})$$

Как будет показано, для процессов неферментативного узнавания очень трудно получить значения единичных q -факторов, которые были бы существенно выше 0,99 (что соответствует $\Delta G_{ik} \approx 3$ ккал/моль).

Наконец, факторы разложения \mathcal{D}_k менее важны в свете настоящего обсуждения. Скорость образования отобранного вида должна превосходить скорость разложения ($W_i > 0$). Что касается \mathcal{D} -факторов, они (для неферментативных процессов) проявляют те же тенденции, что и другие факторы, т. е. не благоприятствуют длинным цепям. Если не образуется защитная макромолекулярная оболочка, длинные цепи будут гидролизоваться легче, чем короткие.

Имеются и другие аспекты, которые могут оказаться важными при обсуждении других механизмов. Однако они имеют меньшее значение по сравнению со следующим заключением об эволюции «самоинструктирующихся» кодовых систем.

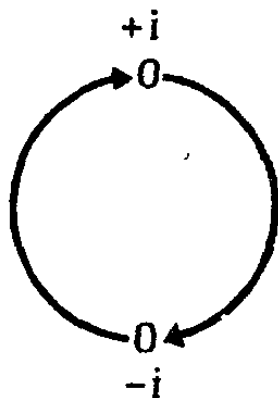
Очень маловероятно, чтобы все мономеры имелись с самого начала в равных количествах. Если, с другой стороны, одна из единиц — например, адениновый нуклеотид — имеется в большом избытке, то это так сильно будет благоприятствовать ее включению при конкуренции с другими единицами, что может привести к доминированию очень однородных полимерных

последовательностей (например, поли-рибоА). Такие однородные последовательности не могут кодировать, и вследствие этого такая система давала бы малые преимущества для дальнейшей эволюции самоорганизации, не говоря уже о других ее недостатках (см. ниже).

В настоящее время мы не находим в природе простого «самоинструктирующегося» кода. В принципе он мог бы существовать, но при этом он находился бы в очень невыгодном положении по сравнению с системой, использующей «комплементарное инструктирование». Такая система даже при наличии одного доминирующего мономера, имеющегося в большом избытке, немедленно стала бы накапливать другую, комплементарную единицу и строить, таким образом, смешанные системы, которые являются предпосылкой для возникновения кода.

§ IV.2. Комплементарное инструктирование и отбор (теория)

Простая форма уравнения (IV.1) не может использоваться для описания «комплементарного инструктирования». Процесс комплементарного копирования представляет собой чередование синтезов «положительных» и «отрицательных» копий, которые мы будем обозначать $+i$ (плюс-цепь) и $-i$ (минус-цепь). Коллектив ($\pm i$) можно представить в виде циклического графа



Каждый такой коллектив описывается двумя кинетическими уравнениями:

$$\begin{aligned} \dot{x}_{+i} &= k_0 (\mathcal{A}_{+i} Q_{+i} x_{-i} - \mathcal{D}_{+i} x_{+i}), \\ \dot{x}_{-i} &= k_0 (\mathcal{A}_{-i} Q_{-i} x_{+i} - \mathcal{D}_{-i} x_{-i}) \end{aligned} \quad (\text{IV.8})$$

и два собственных значения λ будут решениями характеристического уравнения

$$\begin{vmatrix} -(\mathcal{D}_{+i} + \lambda) & \mathcal{A}_{+i} \cdot \mathcal{Q}_{+i} \\ \mathcal{A}_{-i} \cdot \mathcal{Q}_{-i} & -(\mathcal{D}_{-i} + \lambda) \end{vmatrix} = 0, \quad (\text{IV.9})$$

$$\lambda_{1,2} = -\frac{\mathcal{D}_{+i} + \mathcal{D}_{-i}}{2} \pm \frac{1}{2} \sqrt{4\mathcal{A}_{+i}\mathcal{Q}_{+i}\mathcal{A}_{-i}\mathcal{Q}_{-i} + (\mathcal{D}_{+i} - \mathcal{D}_{-i})^2}. \quad (\text{IV.10})$$

Эти собственные значения можно приписать определенным «нормальным координатам» реакции [75] — концентрационным параметрам y_i , которые представляют собой линейные комбинации x_{+i} и x_{-i} . Одно из значений λ всегда отрицательно, другое может быть положительным, если

$$\mathcal{A}_{+i}\mathcal{Q}_{+i}\mathcal{A}_{-i}\mathcal{Q}_{-i} > \mathcal{D}_{+i}\mathcal{D}_{-i}.$$

Соотношение между переменными x и y можно записать в векторной форме

$$\vec{y}_i = M_i \vec{x}_i; \quad \vec{x}_i = M_i^{-1} \vec{y}_i, \quad (\text{IV.11})$$

где матрица M_i^{-1} обратна матрице M_i , а M_i строится из компонент собственных векторов. Рассмотрим для простоты (более прозрачный) случай, когда $\mathcal{D}_{+i} = \mathcal{D}_{-i}$ (для большинства экспериментов при постоянных силах и регулируемых потоках это является хорошим приближением). Для этого случая матрицы имеют следующий вид:

$$M_i = \begin{pmatrix} 1 & -\sqrt{\frac{\mathcal{A}_{+i}\mathcal{Q}_{+i}}{\mathcal{A}_{-i}\mathcal{Q}_{-i}}} \\ +\sqrt{\frac{\mathcal{A}_{-i}\mathcal{Q}_{-i}}{\mathcal{A}_{+i}\mathcal{Q}_{+i}}} & 1 \end{pmatrix};$$

$$M_i^{-1} = \frac{1}{2} \begin{pmatrix} 1 & +\sqrt{\frac{\mathcal{A}_{+i}\mathcal{Q}_{+i}}{\mathcal{A}_{-i}\mathcal{Q}_{-i}}} \\ -\sqrt{\frac{\mathcal{A}_{-i}\mathcal{Q}_{-i}}{\mathcal{A}_{+i}\mathcal{Q}_{+i}}} & 1 \end{pmatrix}. \quad (\text{IV.12})$$

В том же приближении собственные значения равны

$$\lambda_{1,2} = \pm \sqrt{\mathcal{A}_{+i}\mathcal{A}_{-i}\mathcal{Q}_{+i}\mathcal{Q}_{-i}} - \mathcal{D}_i. \quad (\text{IV.13})$$

Физическая интерпретация состоит в том, что каждый реакционный цикл имеет два действительных собственных значения, которые определяют кинетическое поведение. Одно из них всегда отрицательно. Оно описывает релаксационный процесс «уравновешивания» при образовании плюс- и минус-цепей:

$$y_{1i}(t) = y_{1i}^0 \exp \left[- \left(\sqrt{\mathcal{A}_{+i}\mathcal{A}_{-i}\mathcal{Q}_{+i}\mathcal{Q}_{-i}} + \mathcal{D}_i \right) k_0 t \right]. \quad (\text{IV. 14})$$

Релаксация приводит к постоянному отношению x_{+i} и x_{-i} . Второе собственное значение положительно, если член, описывающий усредненное значение образования $\sqrt{\mathcal{A}_{+i}\mathcal{A}_{-i}\mathcal{Q}_{+i}\mathcal{Q}_{-i}}$, превосходит член, описывающий распад $\sqrt{\mathcal{D}_{+i}\mathcal{D}_{-i}} = \mathcal{D}_i$ (аналогично положительному W в уравнении IV.1). Оно выражает автокаталитический рост $(\pm i)$ -коллектива (для $\mathcal{D}_{+i} = \mathcal{D}_{-i} \equiv \mathcal{D}_i$)

$$y_{2i}(t) = y_{2i}^0 \exp \left[\left(+ \sqrt{\mathcal{A}_{+i}\mathcal{A}_{-i}\mathcal{Q}_{+i}\mathcal{Q}_{-i}} - \mathcal{D}_i \right) k_0 t \right]. \quad (\text{IV. 15})$$

Эта часть решения важна для отбора. Матрицы M_i и M_i^{-1} позволяют перейти от «нормальных» к истинным концентрациям или наоборот; например, в том случае, когда $\mathcal{D}_{+i} = \mathcal{D}_{-i}$

$$\begin{aligned} y_{1i} &= x_{+i} - \sqrt{\frac{\mathcal{A}_{+i}\mathcal{Q}_{+i}}{\mathcal{A}_{-i}\mathcal{Q}_{-i}}} x_{-i}, \\ x_{+i} &= \frac{1}{2} \left[y_{1i} + \sqrt{\frac{\mathcal{A}_{+i}\mathcal{Q}_{+i}}{\mathcal{A}_{-i}\mathcal{Q}_{-i}}} y_{2i} \right], \\ y_{2i} &= \sqrt{\frac{\mathcal{A}_{-i}\mathcal{Q}_{-i}}{\mathcal{A}_{+i}\mathcal{Q}_{+i}}} x_{+i} + x_{-i}, \\ x_{-i} &= \frac{1}{2} \left[- \sqrt{\frac{\mathcal{A}_{-i}\mathcal{Q}_{-i}}{\mathcal{A}_{+i}\mathcal{Q}_{+i}}} y_{1i} + y_{2i} \right]. \end{aligned} \quad (\text{IV. 16})$$

Так как $y_{1i} \rightarrow 0$ при $t \rightarrow \infty$, «уравновешенное» отношение x_{+i}/x_{-i} равно

$$\frac{\bar{x}_{+i}}{\bar{x}_{-i}} = \sqrt{\frac{\mathcal{A}_{+i}\mathcal{Q}_{+i}}{\mathcal{A}_{-i}\mathcal{Q}_{-i}}}. \quad (\text{IV. 17})$$

Для равных скоростей образования плюс- и минус-цепей это отношение равно единице. Если процесс начинается с уравновешенного отношения (при $t \equiv 0$), то имеет ме-

сто только одно решение, т. е. рост ($y_{1i}^0 = 0$). В более обычном случае (см. эксперименты с фагом Q β , описанные в гл. VII) процесс начинается с одной копии, например с плюс-цепи ($x_{-i}^0 = 0$). Тогда изменение во времени каждого вида складывается из обоих решений, пока не достигается постоянное отношение, после чего решение роста доминирует.

Введем теперь в кинетические уравнения давление отбора (ограничение постоянной общей организации):

$$\begin{aligned} \dot{x}_{+i} &= k_0 \left[\left(\mathcal{A}_{+i} Q_{+i} \frac{x_{-i}}{x_{+i}} - \mathcal{D}_{+i} \right) - \bar{E} \right] x_{+i}, \\ \dot{x}_{-i} &= k_0 \left[\left(\mathcal{A}_{-i} Q_{-i} \frac{x_{+i}}{x_{-i}} - \mathcal{D}_{-i} \right) - \bar{E} \right] x_{-i}. \end{aligned} \quad (\text{IV. 18})$$

Сразу видно, что эти уравнения относятся к тому общему типу, который обсуждался в гл. II, поскольку отношение (x_{+i}/x_{-i}) или обратное отношение не исчезает с ростом t .

Обозначим это отношение $x_{+i}/x_{-i} \equiv z_i$. Из (IV.18) немедленно следует дифференциальное уравнение для изменения z во времени:

$$\begin{aligned} \dot{z}_i &= \frac{\dot{x}_{+i}x_{-i} - \dot{x}_{-i}x_{+i}}{x_{-i}^2} = \\ &= k_0 [\mathcal{A}_{+i} Q_{+i} + (\mathcal{D}_{-i} - \mathcal{D}_{+i})z_i - \mathcal{A}_{-i} Q_{-i} z_i^2]. \end{aligned} \quad (\text{IV. 19})$$

В результате интегрирования получаем довольно громоздкое выражение

$$z_i(t) = \frac{\bar{z}_i (z_i^* + e^{-a_i t}) - 2\beta_i e^{-a_i t}}{z_i^* - e^{-a_i t}}, \quad (\text{IV. 20})$$

где

$$z_i^* = \frac{z_i^0 + \bar{z}_i - 2\beta_i}{z_i^0 - \bar{z}_i}, \quad \beta_i = \frac{\mathcal{D}_{-} - \mathcal{D}_{+}}{2\mathcal{A}_{-} Q_{-}}$$

и

$$a_i = k_0 \sqrt{4\mathcal{A}_{+}\mathcal{A}_{-}Q_{+}Q_{-} + (\mathcal{D}_{+} - \mathcal{D}_{-})^2},$$

которое при $t = 0$ дает начальное значение z_i^0 и при $t \rightarrow \infty$ уравновешенное отношение

$$\bar{z}_i = \sqrt{\frac{\mathcal{A}_+ Q_+}{\mathcal{A}_- Q_-}} + \beta_i^2 + \beta_i. \quad (\text{IV. 21})$$

Подстановка $z_i(t)$ в (VI. 18) приводит к неоднородному дифференциальному уравнению, которое можно проинтегрировать.

Общее поведение системы с комплементарной инструкцией аналогично поведению «самоинструктирующейся» системы, которое обсуждалось выше. Каждый коллектив, несущий информацию, теперь состоит из двух компонентов и может быть представлен двухкомпонентным вектором $\vec{x}_i = \begin{pmatrix} x_{+i} \\ x_{-i} \end{pmatrix}$. Прежняя величина $W_i^0 = \mathcal{A}_i Q_i - \mathcal{D}_i$ теперь заменяется матрицей

$$\begin{pmatrix} -\mathcal{D}_{+i} & \mathcal{A}_{+i} \cdot Q_{+i} \\ \mathcal{A}_{-i} \cdot Q_{-i} & -\mathcal{D}_{-i} \end{pmatrix}.$$

Эта матрица имеет два собственных значения, одно из которых относится к уравновешиванию коллектива ($\pm i$), а другое описывает свойство конкурентного роста этого коллектива. Следовательно, это собственное значение и соответствующая ему нормальная координата (y_{2i}) входят в уравнение для отбора. После уравновешивания коллектива мы можем заменить нормальную координату y_{2i} просто суммой $x_{+i} + x_{-i} = y_i^*$, оба члена которой пропорциональны y_{2i} , и затем записать в обычной форме уравнение отбора при постоянных силах

$$\dot{y}_i^* = k_0 (W_i^0 - \bar{E}) y_i^*, \quad (\text{IV. 22})$$

где $W_i^0 = \sqrt{\mathcal{A}_{+i} Q_{+i} \mathcal{A}_{-i} Q_{-i}} - \mathcal{D}_i$ (предполагая, что $\mathcal{D}_+ = \mathcal{D}_- = \mathcal{D}_i$) и

$$n\bar{E} = \sum_{k=\pm 1}^{\pm N} E_k x_k; \quad E_k = \mathcal{A}_k - \mathcal{D}_k,$$

где суммирование распространяется на все $+$ и $-k$.

При $W_i^0 = \bar{E}$ получим селекционное равновесие, в котором (снова при условии $\mathcal{D}_{+i} = \mathcal{D}_{-i}$)

$$\frac{\bar{y}_{2i}}{n} \frac{(1 + \sqrt{\mathcal{A}_+ \mathcal{A}_- / \mathcal{Q}_+ \mathcal{Q}_-})}{2} = \frac{y_i^*}{n} = \frac{W_i^0 - \bar{E}_{k \neq i}}{\bar{E}_i - \bar{E}_{k \neq i}}, \quad (\text{IV. 23})$$

где

$$\bar{E}_i = \frac{E_{+i} \sqrt{\mathcal{A}_{+i} \mathcal{Q}_{+i}} + E_{-i} \sqrt{\mathcal{A}_{-i} \mathcal{Q}_{-i}}}{\sqrt{\mathcal{A}_{+i} \mathcal{Q}_{+i}} + \sqrt{\mathcal{A}_{-i} \mathcal{Q}_{-i}}} = \frac{\bar{z}_i E_{+i} + E_{-i}}{1 + \bar{z}_i}.$$

(В случае $\mathcal{D}_{+i} \neq \mathcal{D}_{-i}$ это выражение сохраняет свой общий вид, но W_i^0 и E_i должны быть вычислены согласно уравнению (IV. 10) при помощи полных матриц преобразования M_i и M_i^{-1} .)

Можно видеть, что до тех пор, пока $W_i^0 > \bar{E}_{k \neq i}$, отбор (\pm)-коллектива происходит даже в том случае, если для одной из копий (например, для $-i$) значение E меньше, чем $\bar{E}_{k \neq i}$. Различные коллективы i_i снова подвергаются отбору. При отсутствии других связей между различными i_i выживает только один коллектив i_m (вместе с его «кометным хвостом» ошибок), т. е. тот, который имеет максимальную селективную ценность $W_m^0 > 0$ (или вырожденная группа коллективов). Селективная ценность коллектива содержит среднее геометрическое параметров $\mathcal{A}_m \mathcal{Q}_m$ для плюс- и минус-цепей. Этот интересный результат показывает, что параметры воспроизведения обеих цепей одинаково важны. (Отметим, что среднее арифметическое было бы эквивалентно члену, «лимитирующему скорость».) В настоящем случае более медленный компонент дорастает до большей стационарной концентрации и тем самым достигает скорости быстрого компонента.

Важная особенность «комплементарного инструктирования» состоит в том, что даже при большом избытке одного мономера система всегда должна накопить по меньшей мере два различных мономера, которые затем встречаются у отобранного вида в почти равных количествах. Тогда вследствие копирования ошибок система будет всегда образовывать смешанные последовательности. Это служит предпосылкой для возникновения

любого кода, который может приобрести «смысл» («представляя» информацию, которая поддается оценке) в том случае, если каждая из смешанных последовательностей дает селективное преимущество для своего собственного воспроизведения. Как мы увидим дальше, бинарные кодовые системы с самого начала имеют определенные преимущества, не препятствуя в то же время переходу к высшей (например, четверной) форме кода, если это почему-либо выгодно. Вследствие конкурентной природы отдельных ($\pm i$)-ансамблей количество информации, которое может храниться, ограничено емкостью одного-единственного класса носителей (т. е. главной копии с ее постоянным «кометным хвостом» ошибочных копий) или емкостью вырожденной группы. Следовательно, ограничения, налагаемые на длину, которые определяются в основном факторами качества Q_{+i} и Q_{-i} , являются важными. Информацию о комплементарном узнавании оснований нуклеиновых кислот можно получить из экспериментальных данных.

§ IV.3. Комплементарное узнавание оснований (экспериментальные данные)

IV.3.1. Образование одиночной пары. Комплементарное инструктирование основано на том, что пары образуются только между А и У или между Г и Ц соответственно. Что же лежит в основе такого специфического взаимодействия, которое гарантирует точное считывание, трансляцию и усиление генетических сообщений во всем органическом мире, от фага до человека?

Биохимик ответил бы, конечно: «Специфичные ферменты», физико-химик, возможно, предпочел бы сказать: «Специфичные силы». Правы и тот и другой, потому что мы достаточно хорошо знаем некоторые ферменты, но мы знаем также, что существует специфичное взаимодействие без участия ферментов, например гибридизация комплементарных цепей нуклеиновых кислот.

Структуры специфичных комплементарных пар, предложенные Ф. Криком и Дж. Уотсоном в их эпохальной статье [1], показаны на рис. 8. Однако, взглянув на

рис. 9, можно немедленно убедиться в том, что одних только водородных связей недостаточно для объяснения очевидного выбора природы.

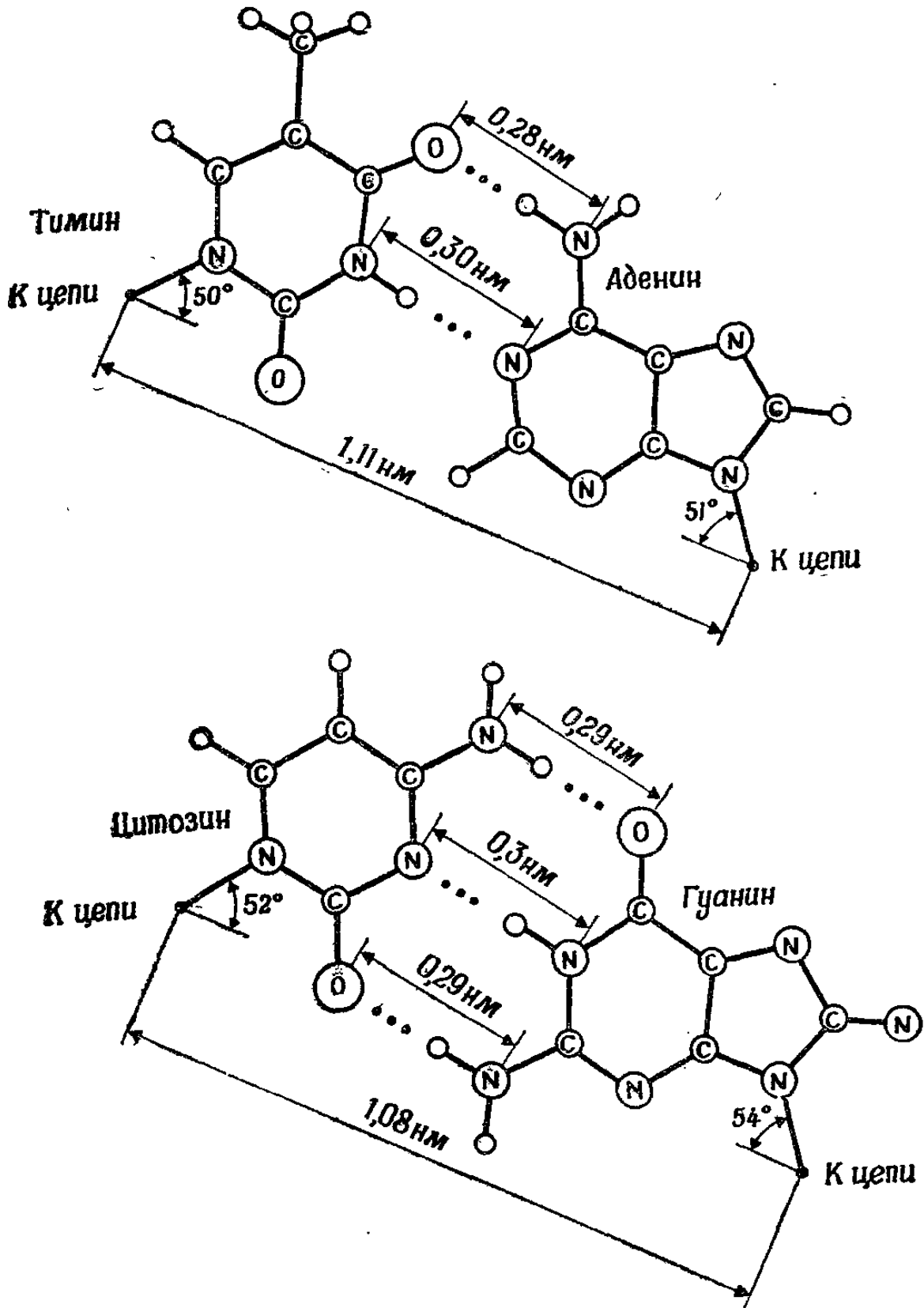


Рис. 8. Комплементарные пары оснований.

Во-первых, существует различие в геометрии разных пар (рис. 9), так что изоморфизм двух уотсон-криковских пар определенно будет давать преимущество

в отношении образования однородных двухцепочечных структур, содержащих все четыре нуклеотида, и особенно в отношении эволюционной адаптации общего поли-

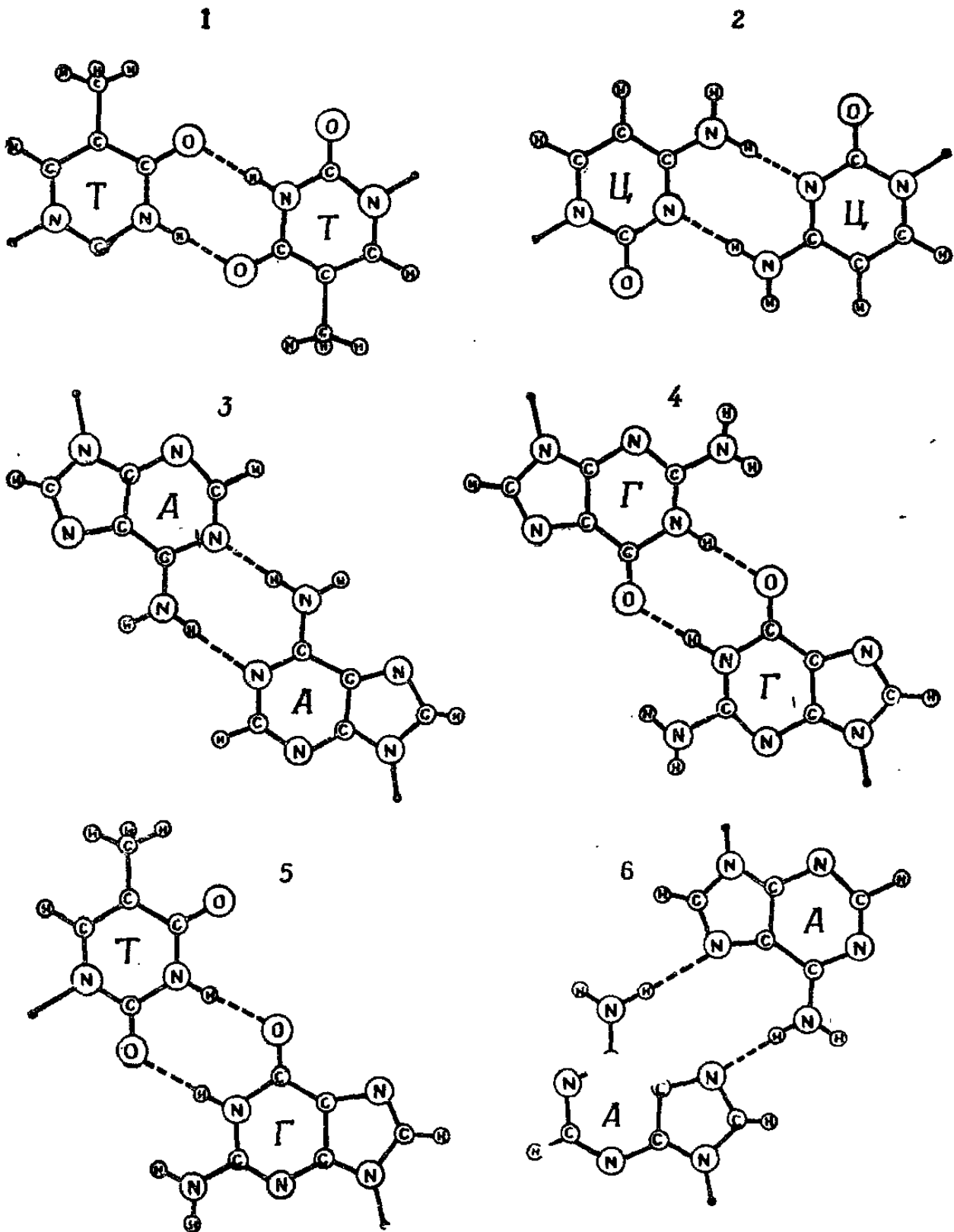


Рис. 9. «Некомплементарные» комбинации оснований.

меризующего фермента. С другой стороны, существует возможность того, что эволюция началась с двухбуквенного кода, содержащего, например, только А и У. Наконец, до того как ферменты приобрели специфическую

адаптацию, возможно, существовал какой-то тип взаимодействия кодон — антикодон с «комплементарностью» по иным параметрам. Для того чтобы ответить на эти вопросы, необходимы соответствующие эксперименты.

Некоторые такие эксперименты уже проделаны. Например, имеются количественные данные о свободных энергиях образования одиночной пары для различных комбинаций оснований. Попытки провести такие измерения, пользуясь в качестве растворителя водной средой, окончились разочарованием: оказалось, что ни одна из пар не стабильна. Даже при самых высоких концентрациях довольно хорошо растворимых нуклеотидов нельзя было заметить образования одиночных пар. И мы знаем почему: если водородные связи между такими полярными группами, как NH, NH₂, OH и CO или N, являются единственным источником стабилизации, то полярные молекулы H₂O будут слишком сильно с ними конкурировать, сольватируя все доступные дипольные группы. Следовательно, для опытов, где исследуется *относительная* устойчивость, лучше пользоваться неполярным растворителем.

Проведены измерения в различных неполярных растворителях с нуклеозидами, которые становятся растворимыми в этих растворителях при замещении различных положений в рибозе неполярными группами (это не мешает основаниям образовывать водородные связи). Очень большая работа в этом направлении, в частности изучение инфракрасных спектров различных замещенных нуклеозидов, проведена А. Ричем и его сотр. [76, 77] в Массачусетском технологическом институте. Аналогичную работу опубликовали Э. Кюхлер и И. Деркош из Венского университета [78]. Диэлектрические измерения образования пар были проведены Т. Функом в нашей лаборатории [79], частично в сотрудничестве с Р. Хопманом и Ф. Эггерсом. Они определили также кинетические параметры по данным релаксационных измерений [80].

Все эти результаты в общем согласуются друг с другом. Комплементарные пары АУ и ГЦ самые сильные по сравнению с другими альтернативами. В табл. 12 приведены некоторые значения констант стабильности, определенные по данным диэлектрических измерений.

Можно видеть, что пара АУ по крайней мере в десять раз более стабильна, чем АА или УУ, тогда как ГЦ является самой прочной парой среди всех комбинаций, содержащих Г или Ц. Пара ГЦ гораздо стабильнее, чем пара АУ. Поскольку Г обнаруживает также значительное «самоспаривание», то определение констант стабильности любых гетеропар, содержащих гуанин, за исключением пары ГЦ, становится невозможным. Следует отметить, что некоторые пары могут образоваться разными способами и что константы равновесия, приведенные в табл. 12, представляют собой суммарные значения.

Таблица 12

Константы стабильности для спаривания

| CCl_4 | У | А | Ц | Г |
|------------------------|--------------------|--------------------|----------------|---------------------------------|
| C_6H_6 | | | | |
| У | 45 15 | 550 | < 50 | < 10^3 |
| А | 150 | 22 8 | < 50 | < 10^3 |
| Ц | < 28 | < 28 | 50 28 | > 10^4 |
| Г | < $1,2 \cdot 10^3$ | < $1,2 \cdot 10^3$ | $3 \cdot 10^4$ | $\sim 10^3$ $1,2 \cdot 10^5$ |

$K_{\text{ass}} [M^{-1}]$ при 25 °С (2'-3'-5'-0—замещенные рибонуклеозиды в неполярных средах).

Все пары образуются очень быстро. Константы скорости показывают, что каждое столкновение ведет к образованию пары и что времена жизни пар в общем меньше одной микросекунды.

Относительно большая устойчивость пар АУ и ГЦ, вероятно, обусловлена одной сильно поляризованной водородной связью. Единственная комбинация ХН ... У такого типа, общая для пар АУ и ГЦ, но отсутствующая

у. их конкурентов, — это водородная связь между довольно кислой группой NH (рК ~ 9,5) и азотом кольца (рК ~ 4,5). Большую стабильность такой связи можно объяснить на основе квантовой механики. Итак, эта важная предпосылка возникновения кода в принципе объясняется физической теорией.

В итоге можно сказать, что, судя по всем имеющимся данным, пары образуют преимущественно те основания, которые мы теперь называем «комплементарными». «Узнавание» — процесс очень быстрый, но вследствие этой быстроты абсолютные стабильности очень низки — так низки, что в полярных средах ощутимого образования пар не происходит. Такие низкие стабильности не могут обеспечить очень точное узнавание, необходимое для чтения кода в длинных последовательностях. Для того чтобы получить больше информации об узнавании оснований при последовательном считывании кода, следует изучить кооперативные взаимодействия в олигомерах и полимерах.

IV. 3.2. Кооперативные взаимодействия в олиго- и полинуклеотидах. Мы уже подчеркивали, что в водных средах отдельные пары оснований слишком нестабильны, чтобы их можно было обнаружить современными методами. С другой стороны, хорошо известно, что комплементарные полимерные цепи образуют очень устойчивые двойные спирали, которые «плавятся» только при высоких температурах. Форма кривых плавления указывает на сильное кооперативное поведение. Источник кооперативности — относительно сильное стопочное взаимодействие между соседними парами оснований, а также хелатные эффекты, возникающие вследствие «замораживания» степеней свободы после образования спирали. Очевидно, что эти кооперативные эффекты сильно увеличивают «инструктивные» способности нуклеиновых кислот.

Прямой путь исследования кооперативных явлений — строить полимер шаг за шагом, начиная с моонуклеотида, и изучать, как изменяются его термодинамические и кинетические свойства с увеличением длины цепи. Большим преимуществом является относительная простота термодинамического и кинетического анализа

конформационных изменений олигомеров, особенно то, что спаривание оснований происходит у них по типу «все или ничего». Трудности такой работы на олигомерах связаны с получением материала: приходится проводить расщепление полимерных образцов, а затем разделять, собирать и очищать продукты гидролиза. Различные группы в нашей, а также в других¹ лабораториях занимались этим в течение нескольких лет.

Первые кинетические исследования были выполнены в дипломной работе Д. Пёршке [72, 73] на олигомерных адениловых кислотах (длина цепи от 2 до 10 мономеров). Как уже говорилось, олиго- и поли-рибоА образуют двуспиральные структуры в кислой области ($\text{pH} \sim 4$), и это дало хорошую модель для наших первых исследований. Затем Д. Пёршке [82] изучил образование двойных и тройных спиралей из олиго-рибоА и олиго-рибоУ при нейтральном pH (длина цепи до 18 мономеров). Эта работа включала в себя термодинамические (фазовые диаграммы и кривые плавления, которые регистрировались с помощью спектрофотометрии в ультрафиолете), а также кинетические исследования с использованием проточных и релаксационных (температурный скачок) методов.

Распространение этих исследований на ГЦ-олигомеры встретилось с большими трудностями, обусловленными агрегацией этих молекул до более сложных структур. С. Поддеру (во время его пребывания в Гёттингене) удалось исследовать спаривание тетрамеров



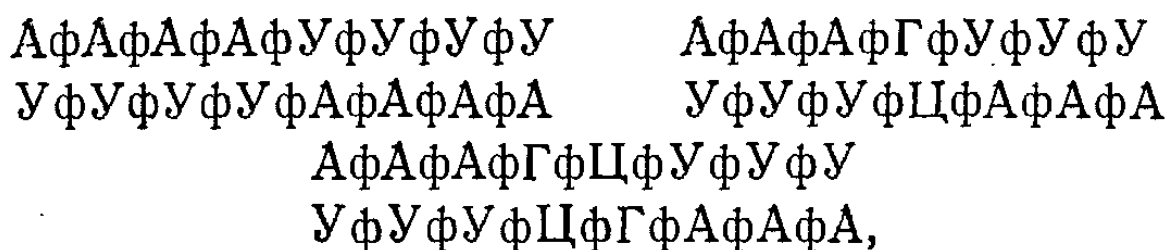
а также спаривание ГфГ и ГфГфГ с поли-рибоЦ.

Г. Маас и объединенная группа из лабораторий Гёттингена и Брауншвейг-Штокгейма (Д. Ризнер, Р. Рёмер, С. Коутс) в сотрудничестве с группой из лаборатории Х. Цахау в Мюнхене изучали термодинамику плавления и конформационную кинетику тРНК [84] (рис. 2) и ее

¹ Например, Р. Болдуин (Стэнфорд), Д. Крозерс (Йель), Н. Дэвидсон (Пасадена), П. Доти (Гарвард), Дж. Фреско (Принстон), Б. Зимм (Ла Хойя).

фрагментов с известной последовательностью [85], которые они получали, расщепляя нативную молекулу в определенном положении по методу, описанному Х. Цахау [86].

Очень ценные данные о специфичном спаривании различных последовательностей получены в лаборатории П. Доти (Гарвардский университет) при исследовании олигомерных сополимеров [87, 88], таких, как



а также различных олигомеров, спаренных с открытыми комплементарными последовательностями тРНК (табл. 13).

Таблица 13

Константы стабильности для спаривания триплетов и квадруплетов оснований (три- и тетрануклеотидов) с открытыми областями тРНК (преимущественно, антикодонами) [87, 88]

| Формилмет-тРНК AA УАЦ УЦ | K_{ass}^1 | Тир-тРНК | K_{ass} |
|-----------------------------|--------------------|---------------------------|-----------|
| АУГ (правильный кодон) | 1 200 ± 200 | УАЦ (правильный кодон) | 700 |
| АУГА | 13 500 | УАЦА | 90 000 |
| АУГУ | 1 400 | УАУ («виляние» в 3') | 700 |
| АУГЦ | 900 | УАУА | 37 000 |
| АУГГ | 1 000 | Фен-тРНК ²⁾ | |
| ГУГ («виляние» в 5') | 1 200 | АУ ААГ* УЦ* | |
| ГУГА | 9 800 | УУЦ (правильный кодон) | 900 |
| ГУГУ | 1 000 | УУЦА | 10 000 |
| | | УУУ («виляние» в 3') | 300 |
| | | УУУА | 1 000 |

1) $K_{ass}(M^{-1})$ измерены в водном растворе (1,0 M NaCl, 10 mM MgCl₂, 10 mM фосфатный буфер, pH 7) при 0° C. При значениях $K < 400M^{-1}$ ассоциация недостоверна. А* — N (6) диметиладенин; У* — псевдоурацил (ψ); Г* — 2-О-метилгуанин; Ц* — 2-О-метилцитозин.

2) Личное сообщение О. Уленбека.

Во время визитов в Гёттинген два сотрудника П. Доти — Ф. Мартин и О. Уленбек — проводили совместно с Д. Пёршке дальнейшие термодинамические и кинетические исследования упомянутых олигомеров [89]. Ниже приводятся результаты и (предварительные) выводы из всех этих работ.

Равновесие спираль — клубок. Кривые «плавления» двойных спиралей зависят от длины цепи: наклон кривых увеличивается с длиной цепи. График зависимости обратной «температуры плавления» (температуры полуперехода спираль \rightarrow статистический клубок) от $1/(\nu-1)$ (ν — число единиц, отвечающее максимальному числу пар оснований в двухцепочечной спиральной структуре) представляет собой прямую линию. Наклон этой прямой пропорционален $1/\Delta H$; отрезок, отсекаемый ею на координатной оси, дает $1/T_\infty = \frac{\Delta S}{\Delta H}$ — величину, обратную температуре плавления «бесконечно длинной» спирали. «Бесконечно» означает: длинную по сравнению с «кооперативной длиной», в которую входит около 30 пар оснований. Следовательно, T_∞ можно определить из опытов с длинными полинуклеотидами. Для того чтобы получить этот результат, необходимо ввести поправку на одноцепочечное стопочное взаимодействие, степень развития которого различна при разных температурах плавления (как можно определить из опытов с одноцепочечными полимерами). Когда длина цепи мала, можно считать, что переход от спирали к клубку происходит по закону «все или ничего». Пусть $\sigma_i s$ — константа стабильности для каждой пары оснований, примыкающей к непрерывной последовательности $(i-1)$ пар. Тогда, помня, что «параметр нуклеации» σ_i близок к единице, если i больше некоторого числа n — «длины нуклеации», можно написать выражение для константы равновесия перехода типа «все или ничего», в котором участвует $\nu > n$ пар оснований:

$$K_\nu = \bar{\sigma} s^\nu, \quad \bar{\sigma} = \sigma_1 \sigma_2 \dots \sigma_n. \quad (\text{IV. 24})$$

Кинетические данные (см. ниже) показывают, что длина нуклеации n при комнатной температуре равна трем для АУ-спиралей и двум для ГЦ-спиралей при тем-

пературах выше 50°C . Это эквивалентно утверждению, что σ_n приближается к единице для $n \geq 3$. Тогда кривые плавления показывают, что для $n \geq 3$ имеется постоянный (и отрицательный) инкремент ΔH_s и ΔS_s , связанный с каждой парой оснований, причем

$$\ln s = -\frac{\Delta H_s}{RT} + \frac{\Delta S_s}{R}. \quad (\text{IV. 25})$$

Тот факт, что исправленная $1/T_m$ пропорциональна $1/(v-1)$, а не $1/v$, показывает, что постоянный инкремент не имеет места для $n \leq 3$, что, в частности, параметр нуклеации обнаруживает температурную зависимость, противоположную таковой для s , так что $\bar{\sigma}s$ почти не зависит от температуры ($\Delta H_{\sigma s} \sim 0$). Хотя $\bar{\sigma}s$ не является константой стабильности для первой пары оснований (она равна σ_1s), физическая интерпретация состоит в том, что существенная часть ΔH_s происходит за счет стопочного взаимодействия (отметим относительно большой инкремент ΔH , обусловленный стопочным взаимодействием, для каждой отдельной цепи), а не от образования водородных связей между комплементарными основаниями. Значение ΔH для первой пары оснований должно быть, таким образом, довольно мало. Эта пара очень нестабильна ($\sigma_1s < 1$) из-за конкуренции с молекулами растворителя (значения констант см. в табл. 13). Для более длинных цепей модель «все или ничего» неприменима. В этих случаях необходимо вычислить статистическую сумму системы с помощью вероятностных методов, что и было сделано Дж. Эпплкви-стом и В. Дамлом [90] для модели «криво застегнутой молнии» («staggering zipper»). Д. Пёршке [91] обработал экспериментальные данные в соответствии с этой моделью на ЭВМ.

Кинетика и механизм спаривания оснований. Пользуясь кинетическими данными, можно пролить свет на механизм кооперативного спаривания оснований. Для коротких спиралей однозначно подтверждается переход по типу «все или ничего». Релаксационный спектр состоит только из одной временной постоянной, которая относится к процессу второго порядка типа «все или ничего», т. е. к рекомбинации обеих (комплементарных)

цепей с образованием двойной спирали, имеющей максимальное число пар оснований. Только в случае более длинных цепей наблюдается дополнительный спектр постоянных времени первого порядка ($\tau \ll 1$ мкс), который отражает изменения в конформациях с разомкнутыми концами или в зигзагообразных («staggered») конформациях. Процесс второго порядка характеризуется большими скоростями рекомбинации, которые, однако, лежат ниже предела диффузионного контроля, указывая на порог нуклеации ($k_R: 10^6 - 10^7 M^{-1}c^{-1}$). Значения этих констант скорости почти одинаковы для любой длины цепи больше трех (но меньше кооперативной длины). Они обнаруживают небольшую, но значимую зависимость от температуры, которая приводит к отрицательным кажущимся энергиям активации для АА-спиралей (рН ~ 4), АУ-спиралей (нейтральная область, $T < 50^\circ C$), а также для ГЦ-спиралей при более высокой температуре ($T > 50^\circ C$).

Абсолютные значения этих параметров «активации» позволяют совершенно непосредственно определить длину нуклеации. Если бы образование первой пары оснований лимитировало скорость¹, то энергия активации (для процесса столкновения) была бы положительной. Поскольку мы знаем, что первая пара нестабильна ($\sigma_{1s} < 1$ означает, что пара диссоциирует быстрее, чем образуется), следует ожидать наличия барьера нуклеации. Если этот барьер представляет собой только первую пару, т. е. если реакция «застегивания молнии» начинает распространяться, как только образуется стабильное ядро, состоящее из двух пар, то константа скорости должна иметь следующий вид: $k_R = \sigma_{1s} k_{12}$, где k_{12} — константа скорости (первого порядка) образования второй пары (следующей за первой). Так как инкремент энергии активации для k_{12} должен быть положительным и σ_{1s} имеет почти нулевой инкремент ΔH , а σ_{1s} включает в лучшем случае малое отрицательное значение ΔH , то полная кажущаяся энергия активации должна быть близка к нулю, но не к экспериментально

¹ Наблюдаются только те процессы, которые ведут к полным спиралам с максимальным числом пар оснований.

наблюдаемому значению 10 ккал/моль (например, для олигомеров АУ). Это относительно большое отрицательное значение совместимо только с выражением $k_R = \sigma s^2 k_{23}$, описывающим процесс, в котором стадией, лимитирующей скорость, является образование третьей пары (k_{23}). Таким образом, для последовательностей АУ стабильное ядро представляют собой спаренные триплеты оснований. Тогда константа скорости распространения (k_{23}) будет равна примерно $10^7 - 10^8 \text{ с}^{-1}$ в согласии с наблюдаемыми временами релаксации $\leq 1 \text{ мкс}$ для конфигураций с разомкнутыми концами. Олигомеры ГЦ при комнатной температуре могут образовывать стабильные «двухпарные ядра», так как их константы стабильности «s» значительно выше. Их константы рекомбинации дают положительные энергии активации, чего и следует ожидать, учитывая большие значения энергетических параметров (по сравнению с АУ). В кривых плавления они «маскируются» параметрами стопочного взаимодействия одиночных цепей (которые тоже выше).

Эти результаты подтверждаются независимыми определениями констант скорости диссоциации. Вследствие микроскопической обратимости

$$k_D = \frac{k_{23}}{s^{\nu-2}} \quad (\text{так как } k_R/k_D = \bar{\sigma} s^\nu). \quad (\text{IV. 26})$$

Поскольку в знаменателе стоит $s^{\nu-2}$, энергии активации положительны и включают в себя большой инкремент $(\nu - 2)\Delta H_s$.

Ожидаемые значения и соответствующее уменьшение абсолютных значений k_D с увеличением длины цепи (по порядку величины) наблюдаются в действительности.

IV. 3.3. Выводы об узнавании. В табл. 13 собраны данные, полученные из равновесных и кинетических исследований, выполненных на олиго- и полирибонуклеотидах. Самое интересное явление — это предпочтение, отдаваемое триpletу, причем не только по логически очевидным причинам (т. е. из-за того, что это дает возможность кодировать более чем 20 символов), а скорее вследствие *механистических совпадений*. Кодоны, состоящие из менее чем трех букв, были бы очень нестабильны

(во всяком случае для А и У). Кодоны, состоящие из более чем трех букв, особенно в случае Г и Ц, были бы слишком «липкими». Время жизни пары кодон — антикодон не должно превышать несколько миллисекунд, чтобы создать возможность для оптимальной адаптации ферментов с соответствующими числами оборотов. Тот же тип оптимизации — компромисс между стабильностью и скоростью — всегда находят в случае фермент-

Таблица 14

Средние равновесные и кинетические параметры кооперативного спаривания оснований, полученные из релаксационных исследований с олигорибонуклеотидами (с длиной цепи от 3 до 18) в водной среде ¹⁾

| Кооперативная пара | σ, M^{-1} | s | ΔH ²⁾ , ккал/моль | $k_R, M^{-1} c^{-1}$ | $\Delta H (k_R),$ ккал/моль | k_{23} или k_{12}, c^{-1} |
|--------------------|------------------|---------------|--------------------------------------|----------------------|-----------------------------|-------------------------------|
| ... А ... | $\sim 10^{-3}$ | 10 | -11 | 10^6 | -9 | 10^7 |
| ... У ... | $\sim 10^{-3}$ | От 100 до 200 | -15 | 10^6 | от +5 до +7 | 10^7 |
| ... Г ... | | | | | | |
| ... Ц ... | | | | | | |

¹⁾ Данные экстраполированы к 0 °С и относятся к рН 7 и ионной силе 0,05 М (какодилат натрия) для АУ-пар и 0,1 М (фосфатный буфер, рН 7,2) для ГЦ-пар. Измерения Д. Пёршке [72, 73, 82, 91] и С. Поддера [76].

²⁾ Экстраполировано к одиночным цепям без стопочного взаимодействия (реально измеренные значения, относящиеся к одиночным цепям со стопочным взаимодействием, значительно ниже).

З а м е ч а н и е. Для констант скорости и равновесия даны только порядки величин, поскольку точные значения зависят от специальных сочетаний пар (кооперативность) и от условий опыта (ионная сила и т. д.).

Времена жизни для различных областей можно оценить, пользуясь значениями k_{23} (или k_{12}) и s (или σs^N соответственно).

Данные табл. 13 относятся в основном к спариванию в антикодоне, но включают, возможно, и другие открытые области. Эти значения выше, чем в табл. 14, вероятно, из-за стерической стабилизации петли антикодона. Следовательно, значения σ из этих данных экстраполировать нельзя. Ошибочное спаривание внутри области, в которой основания спарены, должно дать более низкие константы стабильности (находящиеся «на уровне шумов»), чем те, которые можно получить, пользуясь константами для ошибочного спаривания на концах (см. табл. 13).

субстратных взаимодействий. Любой выигрыш в стабильности означает снижение скорости диссоциации комплекса, которая должна регулироваться в соответствии с числом оборотов, чтобы не стать этапом, лимитирующим скорость всего процесса.

Можно отметить еще одно интересное свойство триплета. Константы стабильности для комплементарных триплетов, полученные на олигомерных двойных спиральях с длиной цепи более чем $\nu = 4$, оказываются значительно ниже тех, которые были определены непосредственно (в лаборатории П. Доти) по взаимодействию триплетов или квадруплетов оснований с открытыми областями тРНК. То же самое имеет место для параметров активации при спаривании Г с Ц, начиная с дублетной пары (см. положительные значения). По-видимому, в коротких открытых областях существуют более сильные стопочные взаимодействия, потому что пары оснований могут при этом расположиться так, чтобы получилось наиболее выгодное стопочное перекрывание. Часть этой энергии взаимодействия теряется, когда рассматриваемая область вытягивается в скрученную спиральную структуру, где на расположение пар оснований накладываются большие ограничения. (Такие же «стерические» ограничения относятся к «виляющей» паре ГУ, которая может образоваться только на одном из концов триплета.)

Если использовать данные, приведенные в табл. 14, можно сделать некоторые заключения относительно значений единичных факторов качества, которые определяют точность комплементарного инструктирования. Используя уравнение (IV.5), модифицированное для случая комплементарного взаимодействия, например для АУ (А — элемент матрицы),

$$q_{AU} = \frac{m_U K_{AU}}{\sum_{X=AUGC} m_X K_{AX}}, \quad (\text{IV. 27})$$

получим значения, которые даже для оптимальных условий с трудом могут достигнуть 0,99. Точное определение потребовало бы более полной информации о константах

стабильности для «ошибочных пар» внутри комплементарной области, и для различных комбинаций ближайших соседей. Для АУ эти величины, вероятно, значительно ниже 0,99, для ГЦ они могут быть — в определенных условиях — несколько выше: пара ГЦ всегда по меньшей мере в 10—50 раз более стабильна, чем пара АУ (в зависимости от ближайших соседей). Эти значения дают верхние пределы. Они предполагают, что в процессе комплементарного узнавания достигается равновесие, до того как мономер включится в полимерную цепь, и относятся к тому случаю, когда концентрации всех мономеров забуферены примерно до одного и того же значения. Такие условия вряд ли существуют в природе. Можно сделать вывод, что значения Q будут значительно отклоняться от единицы уже для относительно коротких цепей. Согласно уравнению (II.45), для воспроизводимого образования носителя кода Q должно оставаться выше определенного порогового значения. Таким образом, воспроизводимое образование нуклеиновых кислот с определенными последовательностями оснований без участия катализаторов было возможно только для относительно коротких цепей, содержащих, по-видимому, не больше 30—100 звеньев (для $q = 0,99$ и $v = 100$ получим $Q \approx 1/e$; см. табл. 8). Другая трудность — это *механизм редупликации*. Для сколько-нибудь точного узнавания комплементарных оснований требуются низкие температуры. При таких температурах образующиеся двойные спирали стабильны. Таким образом, для диссоциации спирали (что является непременным условием многократного матричного действия данной цепи) необходимы большие градиенты температуры. Поэтому большие различия в скоростях (которые существуют и в настоящее время, когда процессы репликации инструктируются ферментами) не могли бы возникнуть как следствие индивидуальной вторичной структуры.

Если, наконец, поставить вопрос:

Могут ли нуклеиновые кислоты организовать самореплицирующуюся и эволюционирующую единицу без содействия катализаторов?

То следует ответить:

Вследствие комплементарных взаимодействий каждый коллектив, состоящий из положительной и отрицательной цепи, обладает способностью к самоинструктированию. При благоприятных условиях среди них мог бы происходить отбор отдельных коллективов с определенными последовательностями. Однако эти последовательности — если они вообще воспроизводимо образуются — содержат очень малое количество информации ($v < 100$). Так как различные коллективы должны конкурировать друг с другом, такая система (без помощи самоинструктирующихся катализаторов) не смогла бы достигнуть такого уровня организации, на котором возможно согласованное функционирование.

САМООРГАНИЗАЦИЯ ЧЕРЕЗ ЦИКЛИЧЕСКИЙ КАТАЛИЗ: БЕЛКИ

§ V.1. Узнавание и ферментативный катализ

Прежде чем обсуждать вопрос о том, могут ли белки образовать самоинструктирующиеся системы, отметим некоторые из их свойств.

1. Как показано во введении (гл. I), имея двадцать видов аминокислот, можно образовать бесчисленное множество различных последовательностей (табл. 3), лишь небольшая часть которых могла бы образоваться случайно. С другой стороны, последовательности аминокислот не обладают внутренне присущей им способностью инструктировать образование идентичных или комплементарных последовательностей. Во всех тех случаях, когда нечто похожее на такую способность, казалось бы, имеется, как, например, в «складчатом слое», или в β -структурах, стабилизированных солевыми мостиками, или в других простых регулярных структурах с комплементарным расположением определенных аминокислот, как в коллагене, это оказывается следствием весьма специфического расположения аминокислот, а вовсе не свойством, присущим самим мономерам, как это было в случае нуклеиновых кислот. Такие специфические «инструктирующие» расположения лишены чрезвычайно важного свойства — способности мутировать. Ошибки здесь не смогли бы воспроизводиться. Кроме того, полипептидные цепи проявляют сильную тенденцию к свертыванию с образованием специфичной пространственной структуры, что очень мешало бы любому процессу прямого копирования.

2. *Пространственная укладка*, с другой стороны, лежит в основе способности белков узнавать специфичные структуры. Этой же уникальной особенностью

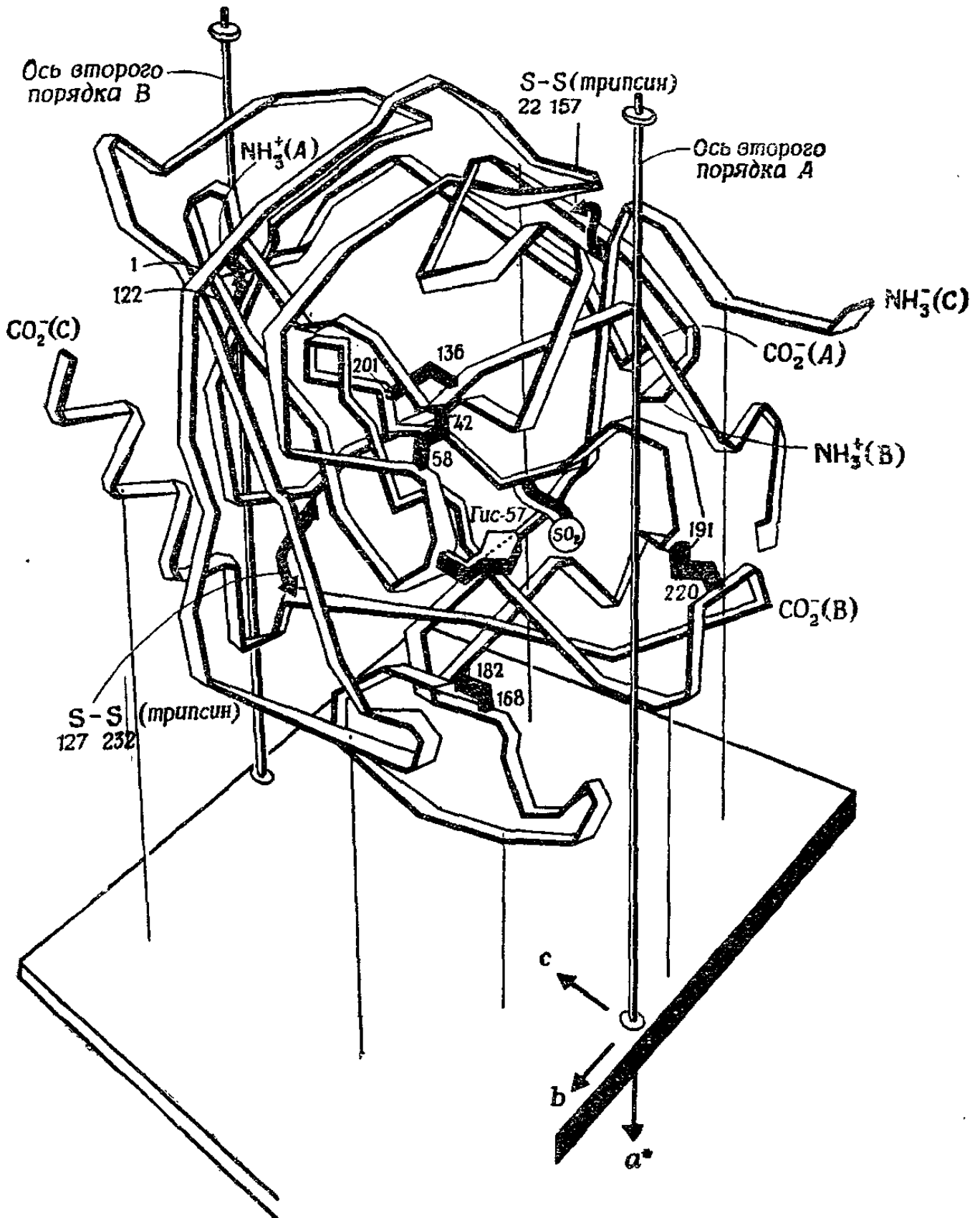


Рис. 10. Схематическое изображение конформации полипептидных цепей в α -химотрипсине [92].

определяются каталитические свойства ферментов при условии, что узнающие группы имеют, кроме того, согласованные каталитические функции. Для примера на рис. 10 и 11 изображен активный центр химотрипсина. Точное пространственное расположение функциональных

групп было установлено лишь недавно методом рентгеноструктурного анализа [92].

Это прекрасный пример того, как функция может зависеть от очень сложной структуры, благодаря которой группы, находящиеся в совершенно различных местах

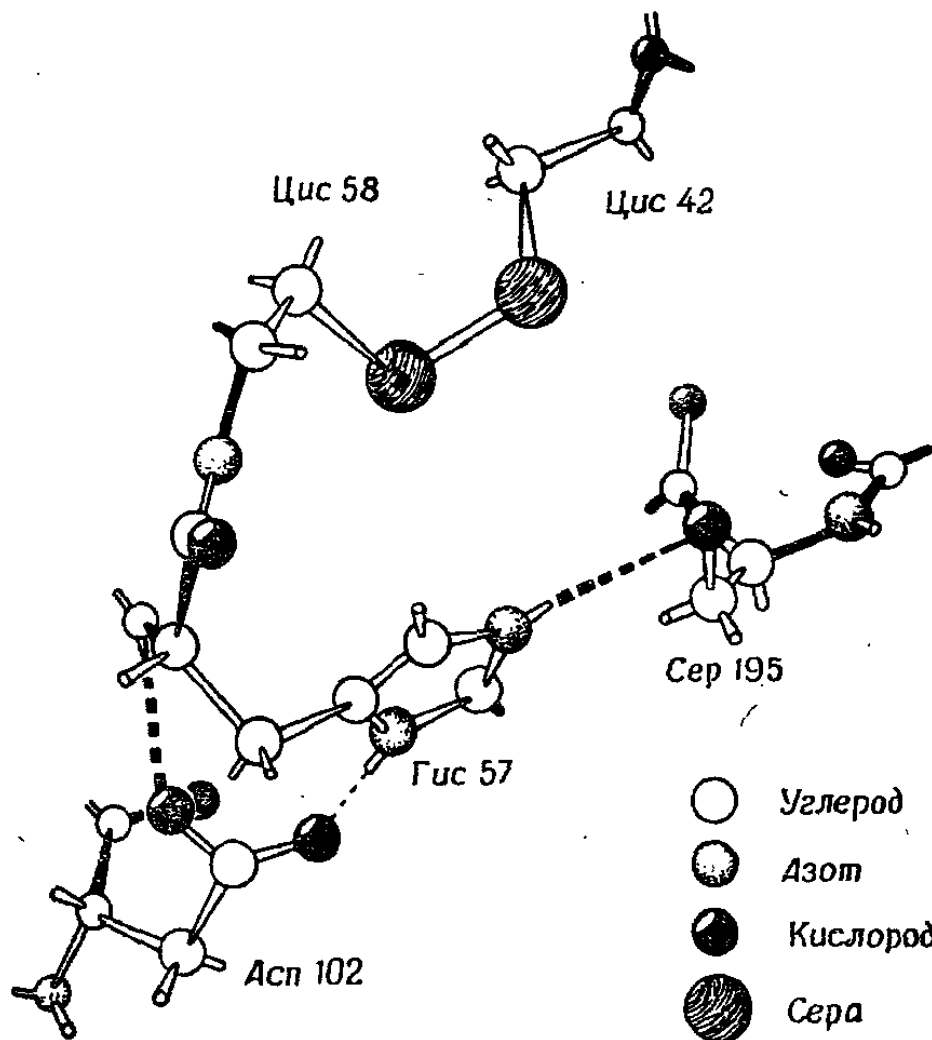


Рис. 11. Конформация некоторых боковых цепей аминокислот в активном центре α -химотрипсина, которая демонстрирует, что «узнавание» белками — результат специфичной пространственной укладки, а не «внутренне присущее» аминокислотам свойство [92].

последовательности аминокислот, оказываются рядом в строго определенном пространственном расположении. Колоссальное разнообразие специфичных участков узнавания демонстрируется также огромным числом различных антител, которые способны связывать любой гаптен, даже если этот гаптен никогда не контактировал с антителом в процессе эволюции. Далее, в лабораторных экспериментах было показано, что при случайном синтезе полипептидов возникает много каталитических функций, иногда очень специфичных (например, химотрипсино-

подобная функция у полипептидов с беспорядочной структурой [93]). Эти продукты не образуются воспроизводимым образом. Даже если некоторые функции возникают воспроизводимо, их носителями оказываются совершенно различные и неродственные структуры.

3. *Специфичное узнавание* определенной макромолекулы ограничивается узнаванием коротких последовательностей или пространственных (третичных) структур, имеющих относительно малую протяженность. Так, некоторые ферменты (например, папаин [94]) расщепляют пептиды с определенными последовательностями аминокислот, тогда как другие ферменты синтезируют специфичные последовательности. Ф. Липман [95] и его группа недавно показали, что полный синтез грамицидина S (циклический декапептид) происходит при помощи одного только фермента, без участия нуклеиновых кислот. Этот фермент с молекулярной массой 280 000, по-видимому, представляет собой сложную систему, состоящую из нескольких субъединиц. Во всяком случае, он функционирует в точности как «белковая матрица», соединяя АТФ-активированные аминокислоты в строго определенную последовательность

D-Фен — L-Про — L-Вал — L-Орн — L-Лей,

которая остается прикрепленной к белку тиоэфирной связью до тех пор, пока два законченных пентапептида не соединяются с образованием циклической молекулы грамицидина S. Таким образом, необходимо помнить, что специфичная и очень точная инструкция для образования белка может содержаться в самом белке, без участия кода нуклеиновых кислот. Такая инструкция, однако, пригодна только для образования относительно коротких последовательностей (например, пентапептидов). Тем не менее, имея в виду эту способность, можно представить себе систему ферментов, которые производят олигопептиды и затем специфично соединяют их друг с другом, пока после ряда стадий не получатся полные молекулы белка, способные, возможно, катализировать свое собственное воспроизведение.

4. По-видимому, в такой системе возможно очень точное регулирование. Поскольку для действия фермента необходимо точное пространственное расположение отдельных групп, свойства фермента могут сильно изменяться в результате конформационных изменений, вызванных связыванием индукторов или же взаимодействием с другими белками. Модели подобного регулирования функции ферментов были впервые предложены Ф. Жакобом и Ж. Моно [11]. Специальные механизмы были рассмотрены Ж. Шанжё, Ж. Моно и Дж. Уайманом [96], а также Д. Кошландом, Дж. Немети и Д. Филмером [97] и проверены исследованиями кинетики различных ферментативных реакций (например, К. Киршнер [98, 99] использовал релаксационные методы для исследования глицеральдегидфосфатдегидрогеназы). Было показано, что эти системы ферментов могут обладать всеми свойствами, присущими электронным регулирующим устройствам [73]. Поэтому любая самовоспроизводящаяся система, если она возникла, может обладать самыми сложными контрольными функциями.

§ V.2. Самоорганизующиеся циклы ферментов (теория)

V.2.1. Каталитические сети. Исходя из свойств, описанных в § 1, можно построить «каталитическую сеть» (рис. 12). Некоторые входящие в нее белки обладают способностью катализировать конденсацию ограниченного числа аминокислот в цепи с определенными последовательностями (например, до определенных пептидов); другие такие «ферменты» узнают данные концевые последовательности этих олигопептидов и соединяют их в более длинные цепи, так что в конце концов могут возникнуть цепи любой длины. Ферменты, обладающие такими каталитическими функциями, обычно полифункциональны. Они могут узнавать определенные последовательности, принадлежащие различным полипептидным цепям (имеющим различную длину), причем их действие зависит также от конкретной третичной структуры субстрата и доступности узнаваемого участка,

который не должен быть спрятан внутри свернутой полипептидной цепи.

Допустим теперь, что каждый катализатор в этой сети синтезируется с помощью другого катализатора. Активация данной каталитической контрольной функции происходит посредством определенных сшивок или разрывов цепей. Подобные процессы сейчас хорошо известны: например, активация, в результате которой из

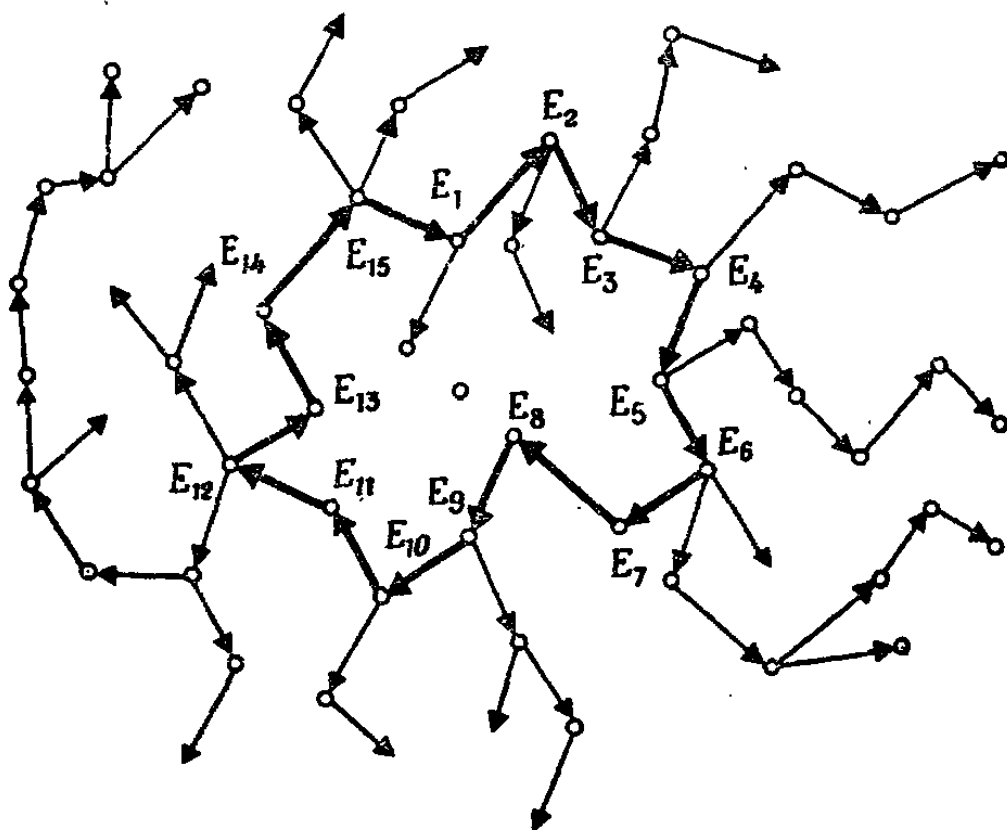
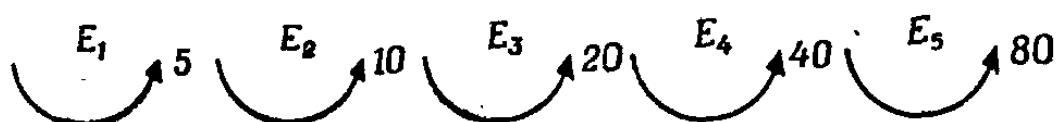


Рис. 12. Каталитическая сеть, в которой имеется замкнутая петля E_1, \dots, E_{15} .

трипсиногена или химотрипсиногена образуется соответственно трипсин или химотрипсин, происходит путем ферментативного разрыва пептидной связи недалеко от одного конца цепи. Таким образом может возникнуть сильно разветвленная каталитическая сеть, как показано на рис. 12. Для того чтобы эта сеть стала самовоспроизводящейся, по крайней мере некоторые из этих ферментов должны быть полифункциональны (см. разветвление на рис. 12), потому что каждому ферменту для своего самовоспроизведения требуется более одного фермента. Например, если можно опознать последовательности из максимум 5 аминокислот, то для синтеза цепи с 80 пептидными связями потребуется по меньшей

мере 5 ферментов, чтобы степень полимеризации увеличилась до 80, например, так:



Чем больше ферментов в этой сети, тем больше вероятность найти замкнутую петлю. Только такое замыкание кольца делает систему автокаталитической и, следовательно, гарантирует самовоспроизведение. Если петля достаточно велика, все вспомогательные функции, такие, как синтез большого числа различных олигопептидов и цепей-предшественников, могут легко локализоваться в ветвях.

V. 2.2. Самовоспроизводящаяся петля и ее варианты. Сосредоточим теперь внимание на тех ферментах, которые образуют замкнутую петлю, и пронумеруем их от

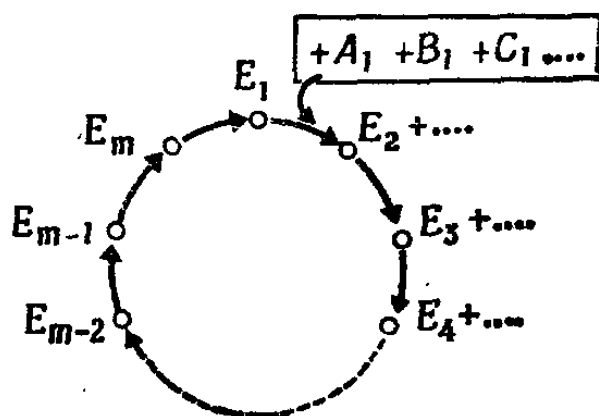


Рис. 13. Представление каталитического цикла в виде графа.

E_1 до E_n , как поступает химик-органик, когда отмечает «хромофоры» в сложной ароматической структуре. Представим эту петлю циклическим графом (рис. 13). Дифференциальные уравнения для скоростей реакций будут в общем случае нелинейны. Однако для простоты можно рассматривать линейную аппроксимацию, которая соответствует ранее разобранному случаю забуференных концентраций субстратов. Даже если это не столь реалистично, как в случае мономерных единиц — ведь субстраты для циклического пути в сети являются в основном полимерными предшественниками, — такие условия в принципе можно реализовать. Более того, основные выводы будут верны также и для нелинейного случая.

Система кинетических уравнений является обобщением систем (IV.8) или (IV.18). Для каждого цикла с m членами в отсутствие ограничений отбора имеем

$$\begin{array}{l} \dot{x}_1 = \mathcal{F}_1 x_m - \mathcal{R}_1 x_1 \\ \dot{x}_2 = \mathcal{F}_2 x_1 - \mathcal{R}_2 x_2 \\ \vdots \\ \dot{x}_m = \mathcal{F}_m x_{m-1} - \mathcal{R}_m x_m \end{array} \quad \left| \begin{array}{l} \mathcal{F}_i = k_0 \mathcal{A}_i Q_i \\ \mathcal{R}_i = k_0 \mathcal{D}_i \end{array} \right. \quad (\text{V. 1})$$

Матрице кинетических коэффициентов

$$\begin{pmatrix} -\mathcal{R}_1 & 0 & \cdot & 0 & \mathcal{F}_1 \\ \mathcal{F}_2 & -\mathcal{R}_2 & \cdot & \cdot & 0 \\ 0 & \mathcal{F}_3 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & 0 & \mathcal{F}_m & -\mathcal{R}_m \end{pmatrix} \quad (\text{V. 2})$$

отвечает характеристическое уравнение

$$\prod_{k=1}^m (\mathcal{R}_k + \lambda) = \prod_{k=1}^m \mathcal{F}_k, \quad (\text{V. 3})$$

имеющее m корней — собственных значений системы. Так как все \mathcal{R}_i и \mathcal{F}_i положительны, получается полином m -й степени по λ , в котором все члены, содержащие λ^i при $i \geq 1$, имеют положительные коэффициенты. Однако для постоянного члена ($\sim \lambda^0$) получается выражение

$$\prod \mathcal{R}_k - \prod \mathcal{F}_k, \quad (\text{V. 4})$$

которое отрицательно при $\prod \mathcal{F}_k > \prod \mathcal{R}_k$. В этом случае в полиноме имеется одна переменная знака, т. е. все члены, кроме последнего, положительны. Согласно правилу Декарта, здесь имеется *одно* положительное и $(m - 1)$ отрицательных (возможно, комплексных) собственных значений.

По аналогии с уравнением (IV.13) при специальном условии $\mathcal{R}_1 = \mathcal{R}_2 \dots \mathcal{R}_m \equiv \mathcal{R}$ положительное собственное значение равно

$$\lambda = + \sqrt[m]{\prod_{k=1}^m \mathcal{F}_k - \mathcal{R}} = \hat{\mathcal{F}} - \mathcal{R}. \quad (\text{V. 5})$$

В отсутствие распада (то есть когда все члены \mathcal{R} равны нулю) решения сводятся просто к корням, содержащим абсолютное значение среднего геометрического $\widehat{\mathcal{F}}$, умноженное на корень m -й степени из единицы, по формуле

$$e^{ik2\pi/m} = \cos(k2\pi/m) + i \sin(k2\pi/m), \quad (\text{V. 6})$$

$$k = 0, 1, \dots, (m - 1).$$

После «уравновешивания» (т. е. после затухания $m - 1$ нормальных решений, соответствующих отрицательным собственным значениям) соотношение плотностей популяций различных членов цикла становится постоянным. Эти «равновесные соотношения» таковы (при условии равных \mathcal{R})¹:

$$\frac{\bar{x}_{i-1}}{\bar{x}_i} = \frac{\sqrt[m]{\prod_{k=1}^m \mathcal{F}_k}}{\mathcal{F}_i} = \frac{\widehat{\mathcal{F}}}{\mathcal{F}_i}, \quad (\text{V. 7})$$

$$\begin{aligned} \frac{\bar{x}_i}{\sum_{k=1}^m \bar{x}_k} &= \\ &= \frac{1}{1 + \frac{\mathcal{F}_{i+1}}{\widehat{\mathcal{F}}} + \frac{\mathcal{F}_{i+1}\mathcal{F}_{i+2}}{\widehat{\mathcal{F}}^2} + \dots + \frac{\mathcal{F}_{i+1}\mathcal{F}_{i+2}\dots\mathcal{F}_{i+m-1}}{\widehat{\mathcal{F}}^{m-1}}} \\ &= \frac{1}{1 + \sum_{s=1}^{m-1} \prod_{k=1}^s (\mathcal{F}_{i+k}/\widehat{\mathcal{F}}^s)}. \end{aligned} \quad (\text{V. 8})$$

В уравнении (V. 8) индексы меняются циклическим образом, т. е. $i + k = i + k - m$ для $i + k > m$

¹ Для $\mathcal{R}_1 \neq \mathcal{R}_2 \neq \dots \neq \mathcal{R}_m$ соответствующие выражения гораздо более сложны, но их легко получить при помощи рекуррентных формул, начиная с тождества

$$1 = \frac{x_m}{x_1} \cdot \frac{x_1}{x_2} \cdot \dots \cdot \frac{x_{m-1}}{x_m}$$

и «условия равновесия»

$$\mathcal{F}_1 \frac{\bar{x}_m}{\bar{x}_1} - \mathcal{R}_1 = \mathcal{F}_2 \frac{\bar{x}_1}{\bar{x}_2} - \mathcal{R}_2 = \dots = \mathcal{F}_m \frac{\bar{x}_{m-1}}{\bar{x}_m} - \mathcal{R}_m.$$

($k = m - i + 1, m - i + 2 \dots m - 1$). Физическая интерпретация полученного результата следующая: *каждый цикл имеет одно нормальное решение, представляющее собой автокаталитический рост всего «коллектива», и $(m - 1)$ нормальных решений, представляющих собой релаксационные явления, т. е. «уравновешивание» внутри цикла. Итак, замкнутая каталитическая петля эквивалентна системе с самоинструментированием или с комплементарным инструментированием.*

Прежде чем перейти к анализу поведения системы, содержащей несколько конкурирующих циклов под давлением отбора, рассмотрим различные типы релаксационных сетей, чтобы получить более детальное представление о предпосылках отбора.

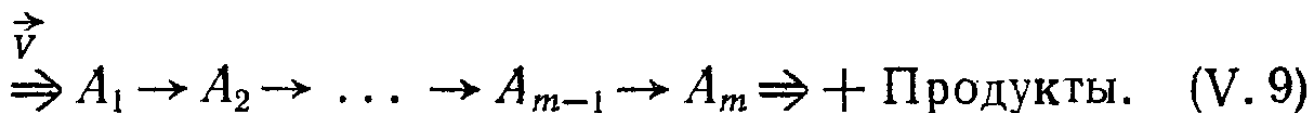
Для селективного самовоспроизведения реакционной сети представляются необходимыми два условия:

1) система должна содержать «замкнутую петлю» реакций;

2) связи между реакционными состояниями должны быть каталитическими.

Чтобы иллюстрировать первое условие, можно рассмотреть общую открытую (квазилинейную) цепь реакций

+ Реагенты



Снова предположим, что кроме A_i , составляющих цепь, все другие реагенты имеются в больших (и, следовательно, постоянных) концентрациях, так что их можно не рассматривать явно. При постоянном реакционном потоке \vec{V} различные состояния будут заполнены до стационарных уровней и вся система будет в конце концов «перетекать через край». Система будет «производить» различные состояния при благоприятных условиях, но она не будет «воспроизводить» себя, т. е. скорости продукции данных состояний не зависят от их населенности. Система — если даже некоторые или все A_i каталитически активны, т. е. участвуют в реакции не расходуясь, — не имеет важной способности к автокаталитическому росту.

Рассмотрим теперь такую же циклическую систему, как и на рис. 13, в которой, однако, E_i не катализаторы, а простые участники и продукты реакции. Система может быть проточной, и концентрации всех участников реакции (A_i, B_i, C_i, \dots), кроме E_i , снова будут постоянны. Такой цикл будет воспроизводить себя. Однако, поскольку ни один из E_i не является катализатором, система придет в стационарное состояние, где каждый из E_i после своего образования будет снова расходоваться. Несколько таких циклов не будут конкурировать друг с другом; между ними установятся стационарные отношения. Такие циклы существуют в биологических системах, и различные этапы реакции также обычно катализируются ферментами. Однако *эти ферменты* не воспроизводятся именно данным циклом, их производит какой-то другой «контур», который является частью самоподдерживающегося цикла живого существа в целом.

Математическое исследование позволяет получить здесь более ясную картину. В отсутствие других реакций разложения при некаталитическом превращении E_i в E_{i+1} скорость исчезновения E_i равна скорости появления E_{i+1} . Таким образом, в матрице кинетических коэффициентов (V.2) все $-\mathcal{R}_i$ заменяются на $-\mathcal{F}_{i+1}$.

Тогда характеристическое уравнение дает только $(m - 1)$ собственных значений, которые все отрицательны; например,

$$\text{для } \mathcal{F}_1 = \mathcal{F}_2 \dots \mathcal{F}_m \equiv \mathcal{F}: \lambda = \mathcal{F} \left(\sqrt[m]{1} - 1 \right)$$

$$\text{или для } m = 4: \lambda_1 = 0; \lambda_2 = -2\mathcal{F}; \lambda_{3,4} = -\mathcal{F} (1 \pm i).$$

Отметим, что для нециклической реакции член в верхнем правом углу матрицы (V.2) должен отсутствовать. Для приведенного примера характеристическое уравнение сводится к $(1 + \lambda/\mathcal{F})^m = 0$, что дает $\lambda_{1,2,3,4} = -\mathcal{F}$.

Для селективного самовоспроизведения существенно, чтобы участники цикла A_i не расходовались в реакции, т. е. чтобы они были катализаторами. Однако цикл остается способным к конкуренции и самовоспроизведению, даже если не все его этапы являются каталитическими. Другими словами, не обязательно, чтобы на ка-

ждой стадии продукт реакции катализировал следующую стадию, не расходуясь при этом. Для того чтобы цикл стал автокаталитическим, достаточно, чтобы в нем производился только один такой катализатор. Это легко видеть из предыдущего рассмотрения (см. матрицу V. 2). В некаталитическом цикле мы должны были бы заменить все $-\mathcal{R}_i$ на $-\mathcal{F}_{i+1}$, потому что члены цикла расходуются в тех же реакциях, в которых возникают соответствующие продукты. Если, однако, первый член в цикле, и только он один, является катализатором E_1 , который образуется из E_m , но не расходуется при образовании E_2 , то тогда верхний левый член ($-\mathcal{R}_1$) в матрице (V. 2) становится равным нулю.

В этом случае мы получим уравнение

$$\lambda \prod_{k=2}^m (\mathcal{F}_k + \lambda) = \prod_{k=1}^m \mathcal{F}_k, \quad (\text{V. 10})$$

которое имеет действительный положительный корень, например при $m = 2$ и $\mathcal{F}_1 = \mathcal{F}_2 \equiv \mathcal{F}$:

$$\lambda (\mathcal{F} + \lambda) = \mathcal{F}^2,$$

$$\lambda_1 = \frac{\mathcal{F}}{2} (\sqrt{5} - 1), \quad (\text{V. 11})$$

$$\lambda_2 = -\frac{\mathcal{F}}{2} (\sqrt{5} + 1).$$

V. 2.3. Конкуренция между различными циклами: отбор. Если имеется ряд различных независимых (несвязанных) каталитических циклов, причем каждый цикл «внутренне уравновешен», то система характеризуется нормальной координатой y_{i+} , отвечающей положительному собственному значению. Так как в уравновешенной системе все другие нормальные решения релаксировали,

мы можем заменить y_{i+} на $y_i^* = \sum_{k=1}^{m_i} x_{ik}$. Тогда уравнение отбора при постоянных силах для каждого цикла принимает общую форму

$$\dot{y}_i^* = k_0 (W_i - \bar{E}) y_i^*, \quad (\text{V. 12})$$

где селективная ценность W_i определяется положительным собственным значением цикла (например, для $\mathcal{R}_{i1} = \mathcal{R}_{i2} \dots \equiv \mathcal{R}_i$)

$$W_i = \hat{\mathcal{F}}_i - \mathcal{R}_i \quad (\text{V. 13})$$

и $\hat{\mathcal{F}}_i$, согласно уравнению (V. 5). Отбор может происходить только при условии, что $\mathcal{F}_i > \mathcal{R}_i$ или — в случае

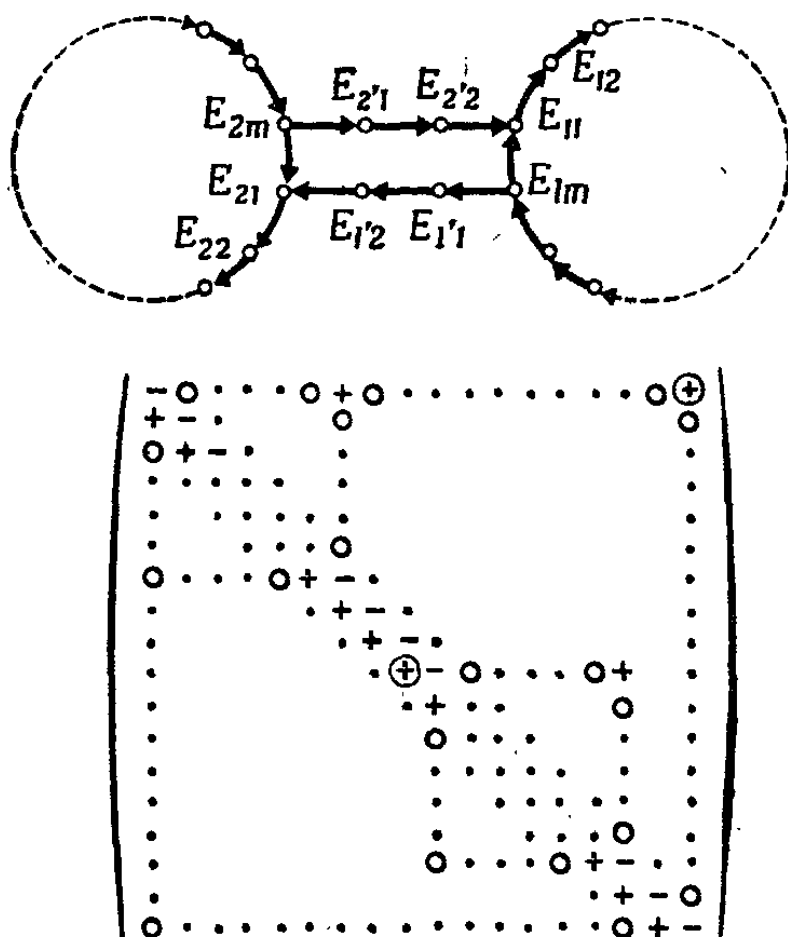


Рис. 14. Представление в виде графа циклов, связанных петлей обратной связи.

Система содержит три цикла: циклы 1 и 2 плюс большая петля, включающая в себя оба эти цикла. Вследствие обратной связи циклы не конкурируют, а стабилизируют друг друга. Матрица кинетических коэффициентов всей системы описывает реакционное поведение. Столбцы соответствуют видам, а строки — реакциям. Два элемента в кружках \oplus представляют собой точки сопряжения петель (через реакции, катализируемые $E_{2/2}$ и $E_{1/2}$). Если бы эти элементы обратились в нуль, то матрица распалась бы и две получившиеся при этом матрицы описывали бы два независимых (конкурирующих) цикла. Области, выделенные точками, соответствуют трем замкнутым петлям. Отрезки диагонали вне циклов 1 и 2 соответствуют ответвлениям.

различных \mathcal{R}_{ik} — если $\hat{\mathcal{F}}_i > \mathcal{R}_i$, где символ Λ означает среднее геометрическое.

По аналогии с уравнением (IV. 23) мы могли бы также вычислить концентрации для отобранного цикла. Независимые циклы ведут себя как отдельные самоинструк-

тирующиеся или комплементарные коллективы, если исходные реакции могут быть описаны системой линейных уравнений.

Циклы, однако, не обязательно должны быть независимы друг от друга. Вследствие наличия разветвлений (т. е. вследствие полифункциональности некоторых ферментов) они могут быть связаны друг с другом, как показано, например, на рис. 14. Если написать кинетические уравнения для всех реагентов, независимо от того, к какому циклу они принадлежат, но пронумеровать их в соответствии с их принадлежностью к определенному циклу, то их свойства легко выводятся из матрицы кинетических коэффициентов (рис. 14).

Типичное свойство этих систем состоит в том, что отобранный цикл (с максимальным W_i) несет с собой все связанные с ним ответвления и циклы при условии, что связь возникает в отобранном цикле. Ответвление от цикла может быть большим недостатком (если оно не несет никакой функции, полезной для воспроизведения цикла), потому что оно создает бесполезный балласт (или даже несет вредную функцию), который ограничивает возможность воспроизведения и тем самым может снизить селективную ценность W_i . Мы будем называть эти ветви «паразитными».

§ V.3. Могут ли белки воспроизводить себя?

Результаты, приведенные в предыдущих параграфах, подсказывают отрицательный ответ на этот вопрос.

Белки имеют два свойства, которые на первый взгляд делают их даже более пригодными, чем нуклеиновые кислоты, для инициирования самоорганизации:

1) значительно большую точность узнавания определенных субстратов благодаря наличию у белков третичной структуры,

2) большее количество информации, содержащееся в многочленном цикле (с ветвями), по сравнению с информационной емкостью одной цепи, имеющей ограниченную длину.

Недостаток состоит в том, что отдельный белок не может воспроизвести длинную цепь из одиночных звеньев

и, кроме того, способность к специфичному узнаванию не представляет собой неотъемлемое свойство «любой» цепи — напротив, это уникальное свойство лишь «определенных» цепей или же редкое совпадение специальных функциональных свойств различных цепей. Вследствие этого белки, которые катализируют свое собственное воспроизведение посредством специфичных циклов, не будут автоматически воспроизводить мутантов, возникающих в результате ошибочного копирования, даже если эти мутанты обладали бы каким-то преимуществом.

Поскольку белки в качестве самовоспроизводящихся молекул обладают как преимуществами, так и недостатками, этот вопрос следует проанализировать более детально.

Было показано, что независимые каталитические циклы являются «самоинструктирующимися», подобно способным к самоинструктированию или комплементарному инструктированию отдельным молекулам, таким, как молекулы нуклеиновых кислот (которым в отличие от белков это свойство присуще изначально). Хотя в настоящее время мы не знаем, существует ли в природе такая самоподдерживающаяся сеть белков — за исключением системы биосинтеза некоторых антибиотиков типа грамицидина S (хотя и в этом случае ферменты, участвующие в биосинтезе, генетически закодированы), — во всяком случае можно представить себе возможность искусственного построения подобной сети.

Какова вероятность самопроизвольного образования таких циклов? Если взять только функцию катализа образования пептидных связей между различными аминокислотами, то можно сказать, что в наборе случайных последовательностей определенная их доля будет всегда проявлять такую каталитическую активность. Весь процесс образования белковоподобных веществ со случайными последовательностями, таким образом, уже является автокаталитическим. Это оказывается существенной предпосылкой для эволюции живых существ, и важно отметить, что это может происходить без инструктирования, осуществляемого нуклеиновыми кислотами.

Однако для эволюционного поведения случайного автокаталитического роста недостаточно. Система может

улучшаться только в результате использования селективных преимуществ, а это требует спецификации последовательностей. В гл. I было показано, что вероятность случайного появления нескольких точно совпадающих последовательностей слишком мала, чтобы иметь какое-либо значение. С другой стороны, только оптимально адаптированные ферменты (какими они нам известны в настоящее время) представлены уникальными последовательностями. Система вполне может начать с далеко не оптимальных качеств функционирования, и для этого может оказаться достаточно спецификации относительно немногих стратегических положений в последовательности. Специфичная функция может проявляться — в гораздо большей степени, чем у современных ферментов, — у относительно широкого класса различных последовательностей, у которых совпадает только ограниченное число звеньев.

Допустим, что имеется достаточно большое число функциональных белков, которые катализируют образование других белков из предшественников. Пусть каждому из этих белков поставлена в соответствие некоторая точка. Соединим линиями те точки, которые каталитически связаны. Предположим, что для каждой из этих точек имеется одна и та же априорная вероятность p того, что она будет мишенью для каталитической активности другой данной точки. Таким образом, для непрерывной петли, включающей k точек, соответствующая вероятность равна p^k . Общее число k -членных замкнутых петель равно $\frac{n!}{k(n-k)!}$. Выражение $\frac{n!}{(n-k)!}$ дает число всех возможных комплексов, отличающихся выбором и последовательностью элементов (т. е. это число всех перестановок из n элементов по k).

Поскольку мы рассматриваем циклические расположения, эти перестановки k -кратно вырождены (начало и конец цикла могут находиться в любом из k положений, например DABC и ABCD и т. д.). Следовательно, вероятность возникновения любого k -членного цикла равна

$$P_k = \frac{p^k n!}{k(n-k)!} \quad (\text{V. 14})$$

При больших n и $(n - k)$ можно применить формулу Стирлинга

$$n! \simeq n^n e^{-n} \sqrt{2\pi n} \quad (\text{V. 15})$$

и получить выражение для максимального значения P_k (из условия $\partial P_k / \partial k = 0$)

$$p = \frac{1}{n - k_m} \exp \left\{ \frac{2n - 3k_m}{2k_m (n - k_m)} \right\}, \quad (\text{V. 16})$$

откуда следует, что k_m близко к n для любой $p > \frac{e}{n}$ и приближается к единице для $p < 1/n$.

Следовательно, можно сделать вывод, что при больших n (когда $p \gg 1/n$) будут образовываться большие петли со многими соединениями, в которых будут участвовать почти все каталитически активные белки.

Вся эта процедура, предполагающая одинаковые априорные вероятности существования каталитической связи, может оказаться сомнительной, потому что при этом допускается грубое упрощение очень сложной ситуации, для точного описания которой необходима гораздо более детальная информация, чем та, которой мы располагаем сейчас. Специфическая корреляция определенных последовательностей с определенными репродуктивными функциями вполне возможна, но эти функции не будут представлять собой неотъемлемое свойство структуры, как это было в случае нуклеиновых кислот, где данная последовательность всегда индуцирует воспроизведение самой себя. У белков такая априорная корреляция между последовательностью и репродуктивной функцией отсутствует, так что выводы о воспроизводящихся петлях остаются качественно верными. Каждая связь, конечно, характеризуется своими кинетическими параметрами, выраженными в селективных ценностях W_i . Вероятность существования данной связи уменьшается с увеличением W_i . Таким образом, в действительности следует использовать понятие плотности распределения вероятностей $p(W)$ и определить вероятность $p(W)dW$ существования связи в интервале селективных ценностей от W до $W + dW$. Такая спецификация требуется в том случае, когда различные петли, характеризующиеся различными селективными

ценностями W_i , подвергаются действию отбора. В этом случае каждая петля с бóльшим значением W_i окажется связанной со многими (более протяженными) петлями, имеющими меньшие селективные ценности, будучи лишена возможности «освободиться» от них в результате отбора.

Один из крупных недостатков простых воспроизводящихся петель состоит в том, что они не могут освобождаться в процессе отбора от паразитных связей, функционирующих менее эффективно.

Каталитические циклы имеют еще один, в эволюционном аспекте даже еще более серьезный недостаток. Предположим, что имеется независимый цикл, который отбирается среди своих конкурентов.

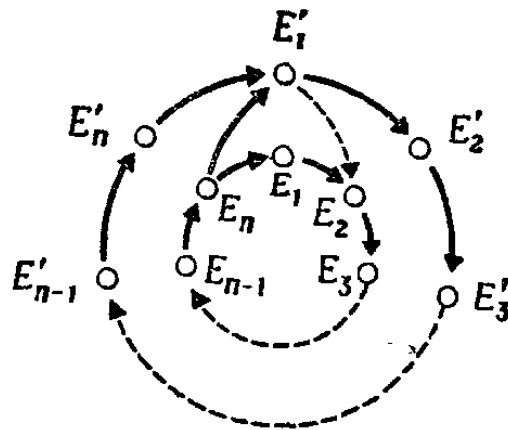


Рис. 15. Воспроизведение мутантов в каталитическом цикле.

Казалось бы, этот цикл (рис. 15) может эволюционировать в результате возникновения мутаций, т. е. вследствие неточного воспроизведения одного из ферментов, участвующих в цикле. Пусть E'_1 — такой мутант. Тогда того условия, что E'_1 — «лучший» катализатор, чем E_1 , недостаточно для эволюции, потому что $E_2 \dots E_n$ снова воспроизведутся, но это снова приведет к E_1 , а не к E'_1 . Чтобы цикл улучшился, должна произойти цепь специфичных мутаций $E'_1 \rightarrow E'_2 \dots E'_n \rightarrow E'_1$, т. е. должен образоваться новый специфичный цикл. Но вероятность того, что произойдет цепь заданных событий, гораздо меньше, чем вероятность появления «какого-нибудь» k -членного цикла (для рис. 15: из равенства (V. 14) видно, что p^k мало по сравнению с P_k).

До сих пор мы описывали каталитические сети при помощи линейных кинетических уравнений, хотя при общих условиях (например, когда концентрации некоторых участников реакции непостоянны) уравнения должны быть нелинейными. Более того, на поведение, которое описывается кинетическими уравнениями (V.1), могут накладываться нелинейные эффекты. Возникает вопрос, придется ли при учете нелинейных эффектов внести какие-либо качественные изменения в существенные выводы. Мы безусловно увидим качественное изменение решений (см., например, следующие параграфы). Однако основной вывод о том, что такие системы не могут эволюционировать из-за отсутствия у белков внутренне присущей им способности инструктировать свое самовоспроизведение или комплементарное воспроизведение, остается в силе и даже приобретает еще большее значение для нелинейных каталитических цепей, способность которых к воспроизведению опять-таки является результатом наличия целого набора уникальных макромолекулярных последовательностей. Такая система не может легко использовать «селективные преимущества», потому что она нагружена слишком большим количеством «информации с низкой селективной ценностью». Линейное приближение селекционного поведения достаточно хорошо иллюстрирует это положение.

САМОУПОРЯДОЧЕНИЕ ЧЕРЕЗ ЗАКОДИРОВАННЫЕ КАТАЛИТИЧЕСКИЕ ФУНКЦИИ

§ VI. 1. Требование кооперации между нуклеиновыми кислотами и белками

Результаты наших исследований нуклеиновых кислот и белков как прототипов молекул, передающих информацию, заставляют сделать следующие выводы:

1. *Нуклеиновые кислоты* обеспечивают выполнение одного важного условия самоорганизации, а именно существования комплементарного инструктирования как основы присущего им селективного самовоспроизведения и образования кода, использующего алфавит из четного числа единиц (например, двоичного или четверичного кода). Точность узнавания не настолько велика, чтобы допускать накопление большого — и вместе с тем воспроизводимого — количества информации в отдельных цепях. Это потребовало бы относительно больших свободных энергий взаимодействия между отдельными комплементарными единицами, что в свою очередь сделало бы код «липким» и привело бы к невозможности считывать его с разумной скоростью. Благодаря кооперативности взаимодействия единиц узнавание в системе кодон — антикодон может быть весьма избирательным и тем не менее занимает всего несколько микросекунд. Вследствие комплементарного взаимодействия внутри одной цепи могут образоваться характерные одноцепочечные структуры (мишени для узнавания ферментами). Каталитические активности, если они вообще существуют, слишком низки, чтобы привести к возникновению связи между конкурирующими носителями информации, и, следовательно, не могут вызвать существенного увеличения информационного содержания.

2. *Белки*, с другой стороны, обладают именно этим свойством, т. е. колоссальным функциональным

До сих пор мы описывали каталитические сети при помощи линейных кинетических уравнений, хотя при общих условиях (например, когда концентрации некоторых участников реакции непостоянны) уравнения должны быть нелинейными. Более того, на поведение, которое описывается кинетическими уравнениями (V.1), могут накладываться нелинейные эффекты. Возникает вопрос, придется ли при учете нелинейных эффектов внести какие-либо качественные изменения в существенные выводы. Мы безусловно увидим качественное изменение решений (см., например, следующие параграфы). Однако основной вывод о том, что такие системы не могут эволюционировать из-за отсутствия у белков внутренне присущей им способности инструктировать свое самовоспроизведение или комплементарное воспроизведение, остается в силе и даже приобретает еще большее значение для нелинейных каталитических цепей, способность которых к воспроизведению опять-таки является результатом наличия целого набора уникальных макромолекулярных последовательностей. Такая система не может легко использовать «селективные преимущества», потому что она нагружена слишком большим количеством «информации с низкой селективной ценностью». Линейное приближение селекционного поведения достаточно хорошо иллюстрирует это положение.

САМОУПОРЯДОЧЕНИЕ ЧЕРЕЗ ЗАКОДИРОВАННЫЕ КАТАЛИТИЧЕСКИЕ ФУНКЦИИ

§ VI. 1. Требование кооперации между нуклеиновыми кислотами и белками

Результаты наших исследований нуклеиновых кислот и белков как прототипов молекул, передающих информацию, заставляют сделать следующие выводы:

1. *Нуклеиновые кислоты* обеспечивают выполнение одного важного условия самоорганизации, а именно существования комплементарного инструктирования как основы присущего им селективного самовоспроизведения и образования кода, использующего алфавит из четного числа единиц (например, двоичного или четверичного кода). Точность узнавания не настолько велика, чтобы допускать накопление большого — и вместе с тем воспроизводимого — количества информации в отдельных цепях. Это потребовало бы относительно больших свободных энергий взаимодействия между отдельными комплементарными единицами, что в свою очередь сделало бы код «липким» и привело бы к невозможности считывать его с разумной скоростью. Благодаря кооперативности взаимодействия единиц узнавание в системе кодон — антикодон может быть весьма избирательным и тем не менее занимает всего несколько микросекунд. Вследствие комплементарного взаимодействия внутри одной цепи могут образоваться характерные одноцепочечные структуры (мишени для узнавания ферментами). Каталитические активности, если они вообще существуют, слишком низки, чтобы привести к возникновению связи между конкурирующими носителями информации, и, следовательно, не могут вызвать существенного увеличения информационного содержания.

2. *Белки*, с другой стороны, обладают именно этим свойством, т. е. колоссальным функциональным

разнообразием, очень высокой способностью к узнаванию и не менее высокой специфичностью (примером может служить чрезвычайно высокая способность антитела к узнаванию определенного антигена). Посредством каталитических связей они могут соединять вместе много носителей информации и таким образом создавать очень большую информационную емкость. Узнавание, однако, не является неотъемлемым свойством элементов последовательности (т. е. аминокислотных остатков), а возникает в результате особого пространственного расположения определенных остатков в активном центре. Эта структура может обратимо изменяться и модифицировать свою узнающую (и каталитическую) способность, в силу чего у нее возникают свойства, поддающиеся регулированию (например, посредством «аллостерической» активации). Белки могут также обнаруживать общее автокаталитическое поведение с отбором вследствие циклических связей. Однако это не является свойством, внутренне присущим всем белковым молекулам, а представляет собой уникальное свойство отдельных видов белка. Следовательно, белки не могут легко использовать селективные преимущества, появляющиеся у («фенотипических») мутантов, и поэтому не могут обеспечить выполнения существенного условия эволюции, которое обеспечивают нуклеиновые кислоты.

3. *«Линейные» реакционные системы* не могут сочетать в себе все свойства, необходимые для зарождения самоорганизующейся системы. Если различные виды носителей информации самовоспроизводятся и подвергаются отбору, то отобранная информационная емкость будет ограничена информационной емкостью одного-единственного вида. Такие системы нуждаются в каталитических связях для увеличения своей информационной емкости. Если же воспроизведение обусловлено только (линейной) циклической каталитической связью, то в этом случае может воспроизводиться большое количество информации, но такая система не может с помощью отбора освободиться от «паразитных» ветвей. Только нелинейные системы обладают всеми свойствами, необходимыми для начала самоорганизации (дальнейшие аргументы приведены ниже), и допускают

дальнейшую эволюцию до такого уровня, когда система может выйти за рамки специальных условий, необходимых для ее возникновения. (Уровень сложности таков, что вероятность получить его путем случайной сборки ничтожно мала.)

4. Сочетание *комплементарного инструктирования с каталитической связью* приводит к нелинейному селекционному поведению. Мы должны найти простейший способ связывания функций нуклеиновых кислот и белков, чтобы воспроизвести тот тип эволюционного поведения, который может привести к структурам и функциям живой клетки. Мы не претендуем на то, чтобы объяснить исторический ход эволюции. Мы лишь хотим попытаться установить ее минимальные предпосылки и несколько прояснить физические принципы эволюционного процесса. Независимо от своей конкретной структуры любая эволюционирующая система должна использовать как кодирующие свойства нуклеиновых кислот, так и каталитическую способность белков. А для этого необходимо возникновение аппарата трансляции. Итак, мы должны рассмотреть два вопроса, которые тесно зависят друг от друга:

а) Как может возникнуть *код и соответствующий аппарат трансляции*? В настоящее время этот аппарат использует адапторы в виде молекул тРНК и распознающие ферменты — аминоксилсинтетазы.

б) Каким образом система такого рода, представляющая собой ансамбль нуклеиновых кислот и белков, может превратиться в стабильную самовоспроизводящуюся единицу, способную к дальнейшей эволюции?

Начнем со второго вопроса, потому что ответ на него пригодится при решении первого вопроса.

§ VI.2. Самовоспроизводящийся гиперцикл

VI.2.1. Модель. Рассмотрим простую модель, изображенную на рис. 16. Она состоит из ряда нуклеотидных последовательностей (или, лучше, (\pm) -коллективов) I_i с ограниченной длиной цепи. Эти коллективы не должны содержать больше информации, чем требуется для кодирования одной или двух каталитически

активных полипептидных цепей, которые мы обозначим E_i . Кружок вокруг I_i условно представляет способность нуклеотидного коллектива к комплементарному инструктированию, а каждый такой коллектив состоит из

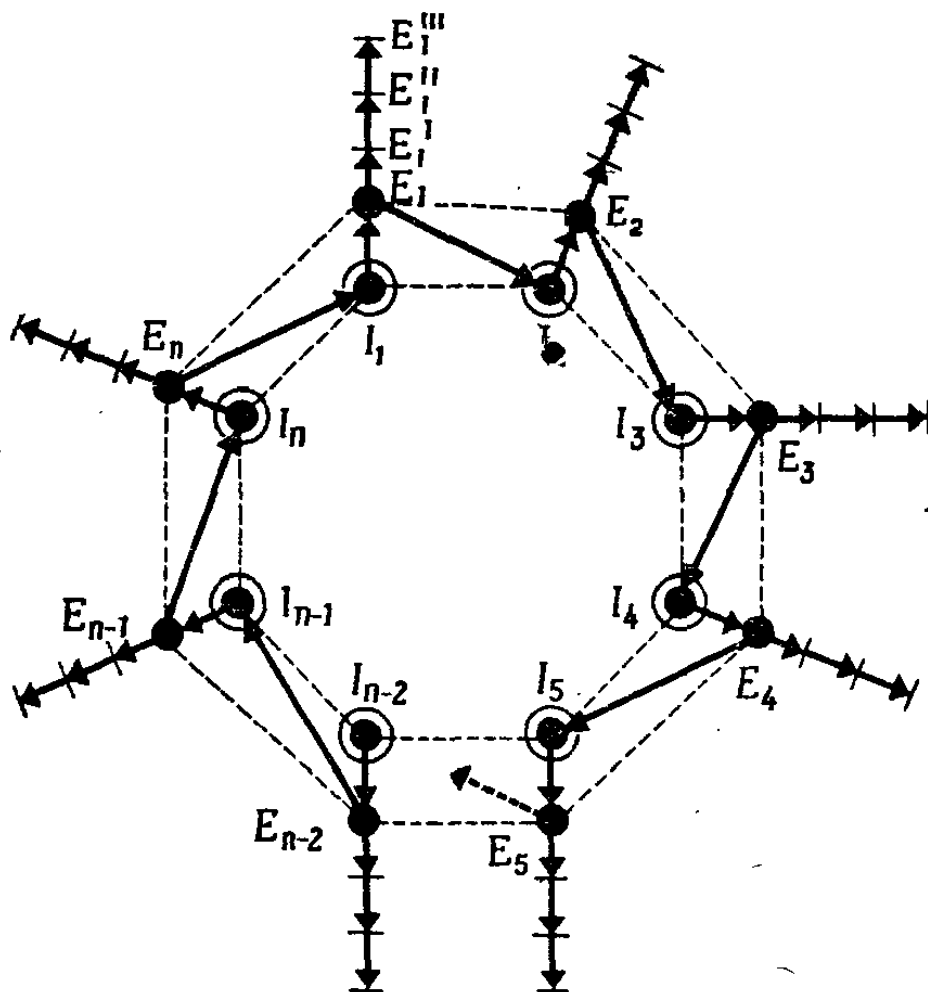


Рис. 16. Самоинструктирующийся каталитический гиперцикл.

I_i — носители информации, т. е. комплементарные цепи РНК. Каждый черный кружок обозначает самоинструктирующее свойство коллектива I_i , содержащего две комплементарные цепи. E_i (кодируемый I_i) — каталитическая функция. Каждая ветвь E_i может выполнять несколько функций (например, полимеризацию, трансляцию, регулирование), одна из которых должна обеспечивать связь с носителем информации I_{i+1} (например, катализировать образование I_{i+1} путем специфичного узнавания). Коонтур, включающий в себя все связи, должен быть замкнут, т. е. должна существовать такая E_n , которая катализирует образование I_1 . Гиперцикл описывается системой нелинейных дифференциальных уравнений.

«положительной» и «отрицательной» цепей, взаимно воспроизводящих друг друга. Этот процесс воспроизведения специфично катализируется предшествующей полипептидной цепью E_{i-1} .

Полипептидная цепь E_{i-1} кодируется нуклеотидной цепью I_{i-1} . Наличие транслирующей системы обеспечивает достаточно точную передачу информации от

I_i к E_i . Только часть информации, хранящейся в каждом I_i , должна использоваться для кодирования функции, которая усиливает образование следующего носителя информации; вся остальная информация может использоваться для кодирования общих ферментативных функций, таких, как трансляция, полимеризация, регуляторные функции и т. д. Далее каждый из E_i , функция которого состоит в том, чтобы специфично усиливать образование следующего носителя информации, может (но не должен) представлять собой специфичную полимеразу (как, например, Q β -репликаза, см. гл. VII). E_i может быть также специфичным индуктором (или дерепрессором), действующим на общую полимеразу. Важно, чтобы весь «гиперцикл» был замкнут, т. е. чтобы существовал такой E_n , который действует на I_1 . Таким образом, система представляет собой «циклическую иерархию», в которой многие циклические (комплементарные) коллективы нуклеотидов связаны вместе посредством ферментативного «гиперцикла». Наличие этой вторичной замкнутой петли важно потому, что в противном случае различные I_k стали бы не кооперировать, а конкурировать и вытеснять друг друга.

Эта система имеет ряд свойств, которые ниже обсуждаются более подробно:

1. Каждый цикл (подобно системам, рассмотренным в гл. IV и V) способен к автокаталитическому росту.

2. Независимые циклы конкурируют и подвергаются отбору.

3. Вследствие нелинейности отбор очень жесткий, и если имеются особые точки, то действие отбора происходит практически по типу «все или ничего».

4. Обладая этими селективными свойствами, система способна, во-первых, использовать очень малые селективные преимущества (вероятность возникновения которых должна быть стохастически значима) и, во-вторых, очень быстро эволюционировать; отобранная система не будет толерантна к возникновению независимых конкурентов, поэтому код и хиральность будут универсальны.

5. Циклические связи обеспечивают информационную емкость, приспособленную к требованиям системы. Тем не менее репликационная длина отдельной кодирующей

единицы (см. оценку ν в табл. 8) должна быть достаточно мала, чтобы обеспечить воспроизводимость.

6. Система может эволюционировать, т. е. совершенствоваться, путем использования селективных преимуществ. «Генотипические» мутации, т. е. изменения в I_i , могут немедленно использоваться E_{i-1} и не должны ждать, пока произойдет коррелированная серия мутаций для того, чтобы распространиться по циклу, как это было необходимо для «линейных» каталитических циклов (см. гл. V). Селективные преимущества могут проявляться через репрессию, дерепрессию или промотирование.

7. В системе происходит отбор против паразитных ветвей, если селективные ценности последних ниже, чем для членов цикла. Паразитные ветви с высокими селективными ценностями не дадут возможности циклу пройти стадию нуклеации, если они имеются с самого начала. Однако если эти ветви появляются после того, как цикл прошел стадию нуклеации, то вследствие нелинейного селекционного поведения они не имеют шансов расти. Цикл может сжаться, т. е. сократить число своих членов, если это дает какое-либо селективное преимущество. Условия одновременного существования нескольких связанных циклов очень ограничены.

8. Существует только один тип ветвей, которые могут сосуществовать с циклом; это ветви, селективные ценности которых точно совпадают с селективной ценностью цикла. Точное совпадение возможно только в том случае, если ветвь и кодирующий участок I_i внутри цикла используют один и тот же промотор, локализованный в E_{i-1} . Это автоматически приведет к генной и оперонной структуре кодовой системы. В области ветвей у системы могут развиваться функции «общего пользования» (например, функции, выполняемые полимеразами, системой трансляции, регулирующими факторами, метаболическими ферментами).

9. Вскоре после нуклеации система должна пространственно обособиться от среды. Только компартиментализованные системы могут использовать функциональные ветви (возникшие вследствие мутаций) исключительно на свою собственную пользу (одновременно создавая

возможность для эволюции ветвей). При помощи того же механизма система спасается от «загрязнений», вызванных неблагоприятными мутациями в ветвях.

10. Система, обособившаяся от среды, может «индивидуализироваться» путем соединения своих кодирующих единиц в стабильную цепь, например с помощью (эволюционирующей) лигазы, и воспроизводить всю эту цепь как индивидуальную единицу. В такой цепи (которая при участии лигаз будет циклической) гены, происходящие от данной единицы I_i , должны располагаться в соседних положениях. Однако матрица, кодирующая связывающий фактор и происходящая из I_{i-1} , может находиться на довольно значительном расстоянии.

Последние два пункта, очевидно, не представляют собой неотъемлемые свойства цикла. Они показывают, что цикл не является «тупиком» в отношении дальнейшей эволюции. Он способен использовать любые преимущества, что приводит его к такому уровню сложности, который близок к сложности живой клетки. Или, другими словами, лишь те системы, которые сумели компартиментализоваться и обособиться, получили в конце концов шансы на выживание.

VI.2.2. Теоретическое рассмотрение. Теория циклической системы, описанной выше, разработана нами вместе с П. Шустером. Вычисления сделаны на ЭВМ. Подробности этой работы будут опубликованы позднее [100]. В этом параграфе кратко излагаются лишь некоторые предварительные результаты.

Рассмотрим сначала простой предельный случай нелинейных кинетических уравнений, дающий конструктивное представление о том, какой тип решений следует ожидать. Если белки E_i находятся в квазиравновесии со своими кодирующими единицами I_i , то мы можем не рассматривать в явном виде скорости их образования и можем считать, что их концентрации пропорциональны y_k — концентрациям инструктирующих кодовых единиц I_k (которые здесь мы считаем просто «самоинструктирующимися»). Если, далее, из пары концентраций E_{i-1} и I_i ни одна не имеется в большом избытке, и, следовательно, можно принять, что концентрация комплекса между E_{i-1} и I_i пропорциональна произведению обеих

концентраций (соответствуя области второго порядка в модели Михаэлиса — Ментен), то кинетические уравнения для образования кодирующих единиц — без огра-

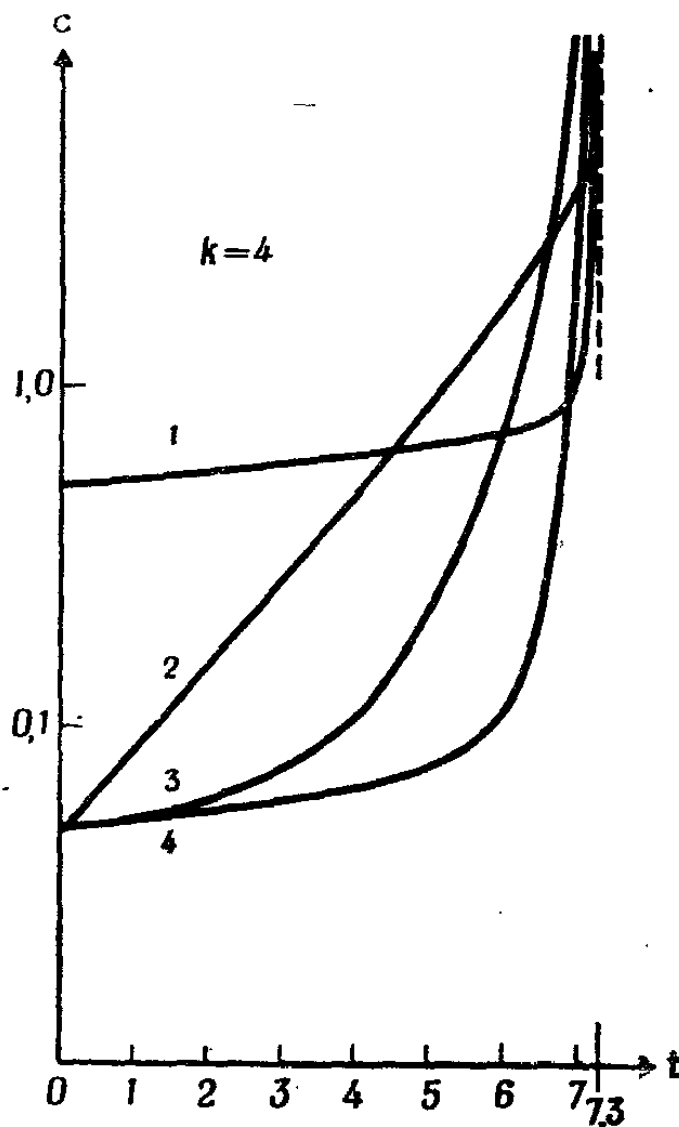


Рис. 17. Рост четырехчленного гиперцикла (4 носителя информации кодируют 4 фермента сопряжения E_i).

Предполагается, что при $t=0$ имеется 10-кратный избыток I_1 по сравнению с I_2 , I_3 и I_4 . Скорости образования для всех четырех членов сводятся к простому выражению второго порядка $x_i x_{i-1}$. Разложение не учитывается. Ось времени трансформирована и относится к константе скорости $k=1$ (или $t=kt'$). Существование особой точки при $t=7,3$ приводит к очень резкому отбору, если имеется несколько конкурирующих циклов. (Вычисления П. Шустера [100].)

ничений, налагаемых отбором, — можно записать следующим образом:

$$\dot{y}_i = (\mathcal{F}_i + \mathcal{F}'_i y_{i-1} - \mathcal{R}_i) y_i. \quad (\text{VI. 1})$$

(Снова предполагается, что концентрации всех высокоэнергетических мономеров забуферены. Члены, описывающие «обратный поток», также не учитываются.)

Решение этого уравнения — как и уравнения (II. 21) — имеет особую точку при конечном t (при условии, что

$\mathcal{F}_i + \mathcal{F}'_i y_{i-1} > \mathcal{R}_i$). Это иллюстрирует рис. 17. Интересно отметить, что для циклов с тремя и более членами могут возникнуть колебания. Эти колебания еще более выражены при наличии налагаемых отбором ограничений (см. рис. 18). Двучленные циклы не дают колебаний; при трех членах колебания есть, но они затухают; при четырех и более членах затухание исчезает. Если начать с независимых от времени средних (соответствующих внутреннему равновесию), то колебания возникают после малого возмущения. Независимые от времени средние можно получить из стационарного решения (внутреннее «равновесие»)

$$\frac{\langle y_i \rangle}{\sum_k \langle y_k \rangle} = \frac{1 - \sum_k \frac{b_{i+1} - b_k}{b_k} / \sum_k \langle y_k \rangle}{a_{i+1} \sum_k \frac{1}{a_k}}, \quad (\text{VI. 2})$$

где сумма распространяется на все k членов цикла, $a_i = \mathcal{F}'_i$ и $b_i = \mathcal{F}_i - \mathcal{R}_i$ согласно уравнению (VI. 1). Символ $\langle \rangle$ для осциллирующих концентраций означает среднее по времени.

При увеличении числа k кодирующих единиц осциллирующее поведение проявляется в виде волн с пиками характерной формы, которые оббегают вокруг цикла. Форма этих пиков зависит как от k , так и от относительных скоростей образования и распада отдельных кодирующих единиц (см. рис. 18—21). Селекционное поведение конкурирующих циклов зависит не только от среднего кинетического параметра цикла, но и от относительных скоростей отдельных этапов и их распределения между различными этапами, а также от k — числа кодирующих единиц в цикле (рис. 22—24). Держать пари о результате такой конкуренции оказывается так же трудно и почти так же увлекательно, как играть на скачках. Обсуждение деталей реакционных механизмов выходит за рамки этой работы (это можно найти в другом месте). То же самое справедливо для поведения связанных циклов, свойства которых во многом напоминают социальное поведение.

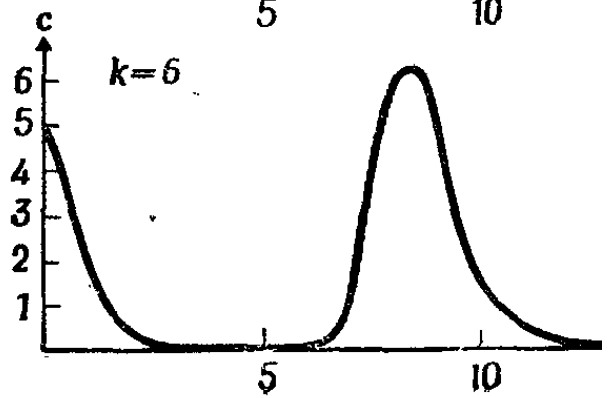
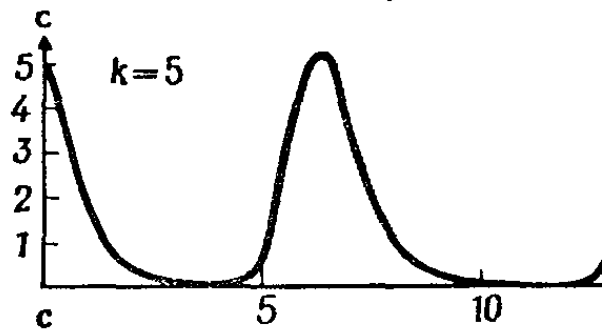
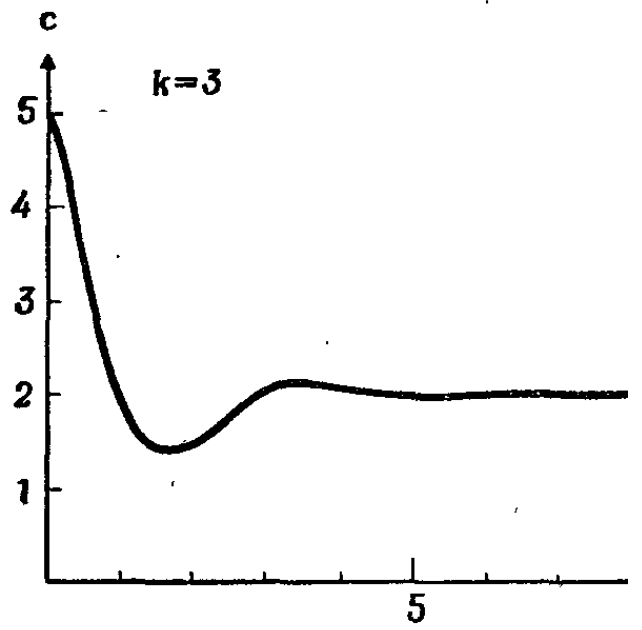
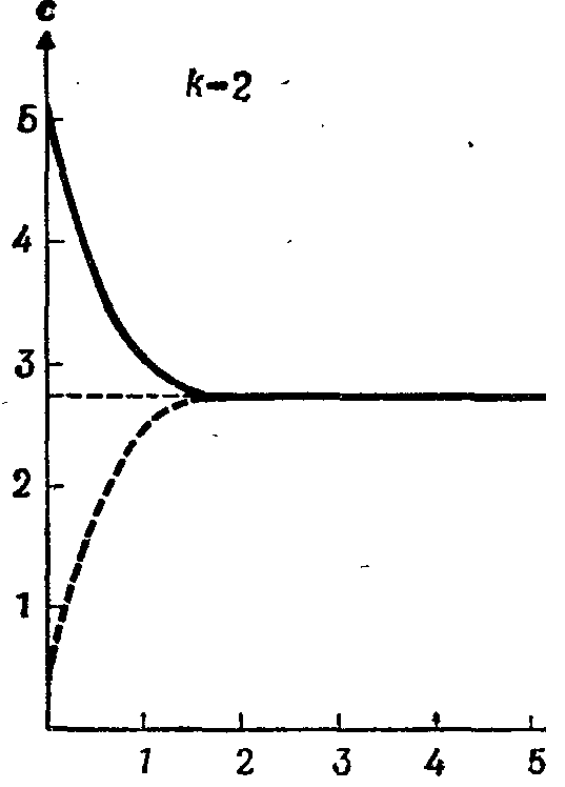
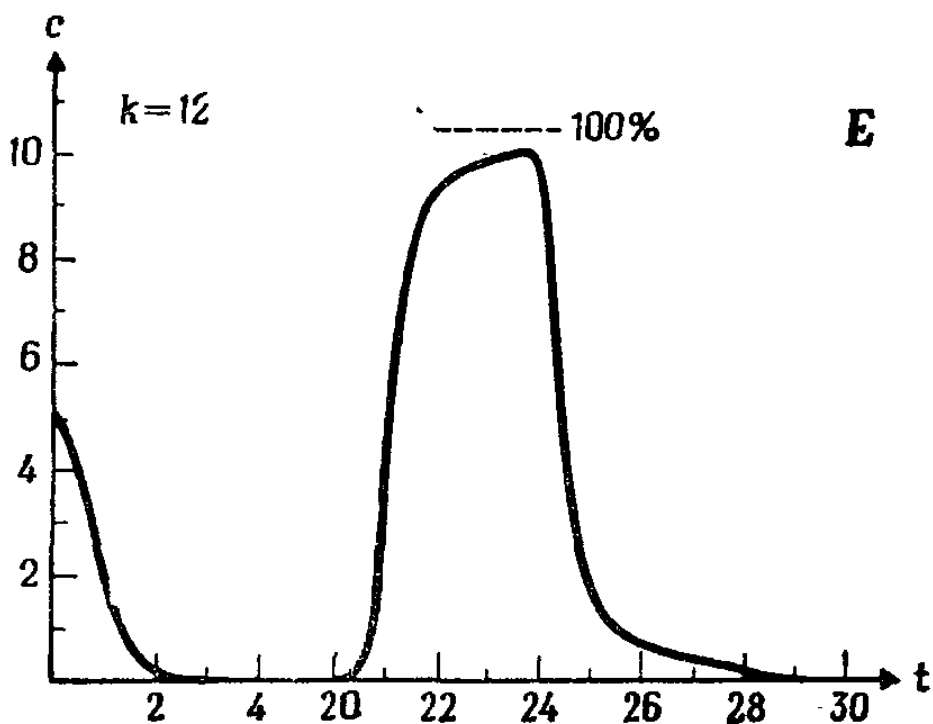
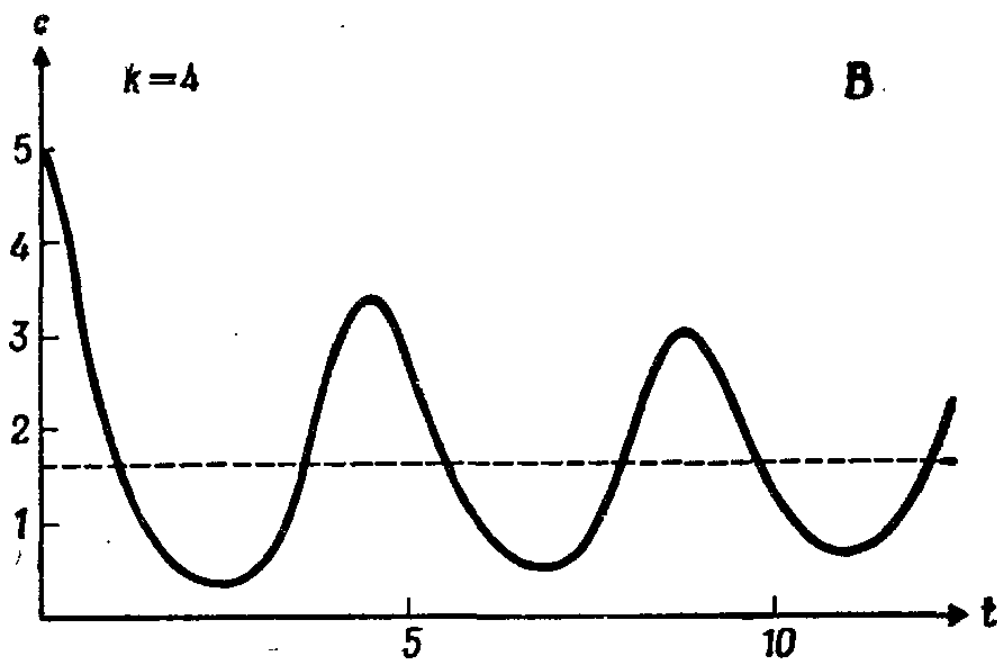


Рис. 18. Решения, описывающие отбор k -членных гиперциклов при ограничении постоянной общей организации ($k = 2, 3, 4, 5, 6, 12$).

Скорость образования снова описывается простым выражением второго порядка — одинаковым для всех членов цикла (ср. рис. 17), скорость «удаления» — членом первого порядка, чтобы выполнялось условие $\sum_k x_k = \text{const}$. Показаны решения только для одного члена цикла. Для

$k \leq 3$ существуют постоянные «равновесные» значения; для $k = 3$ система приближается к селекционному равновесию, совершая затухающие колебания; для $k > 3$ возникают устойчивые колебания. Это можно показать, если, отходя от постоянного распределения, при $t = 0$ создать малое возмущение. При этом устанавливается колебательный режим. (Вычисления П. Шустера [100].)



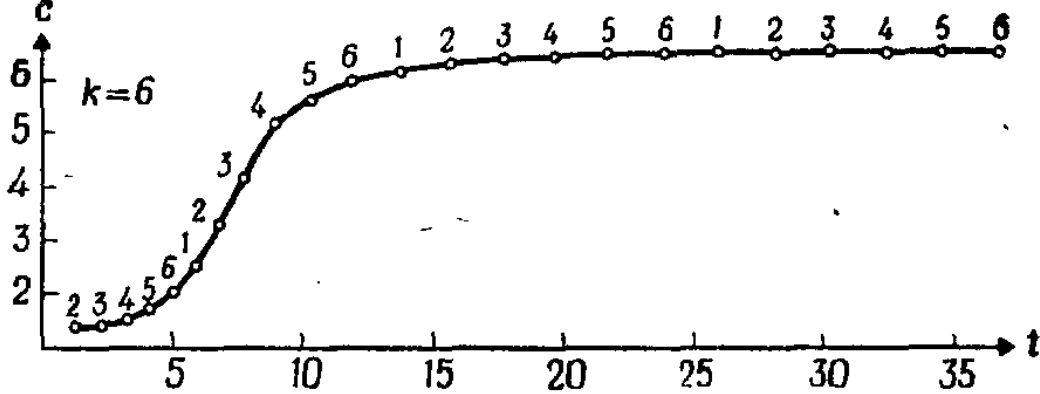


Рис. 19. Изменение амплитуды колебаний в гиперцикле [100].
 Колебания возникают из постоянного стационарного распределения при
 $t=0$; $x_1^0=1,35$; $x_k^0 > 1=1,25$.

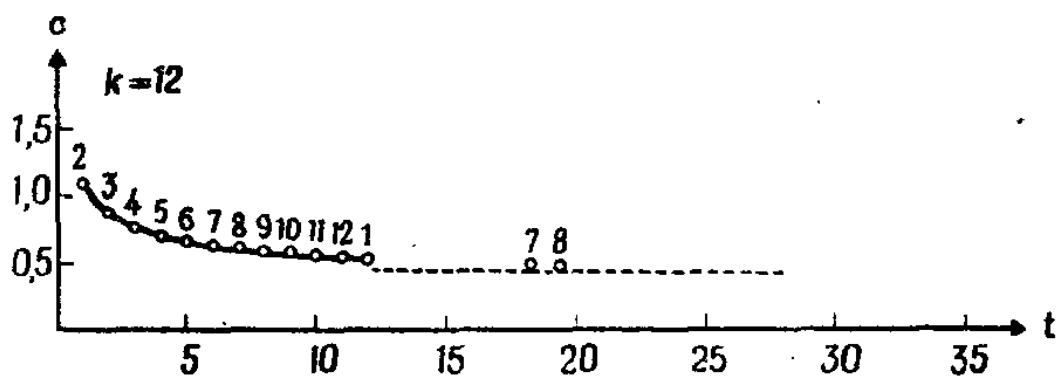


Рис. 20. Снижение амплитуды в линейном каталитическом цикле
 (см. гл. V) для сравнения с рис. 19.
 При наложении ограничений, обусловленных отбором, концентрации всегда
 приходят к постоянному стационарному уровню (см. [100]).

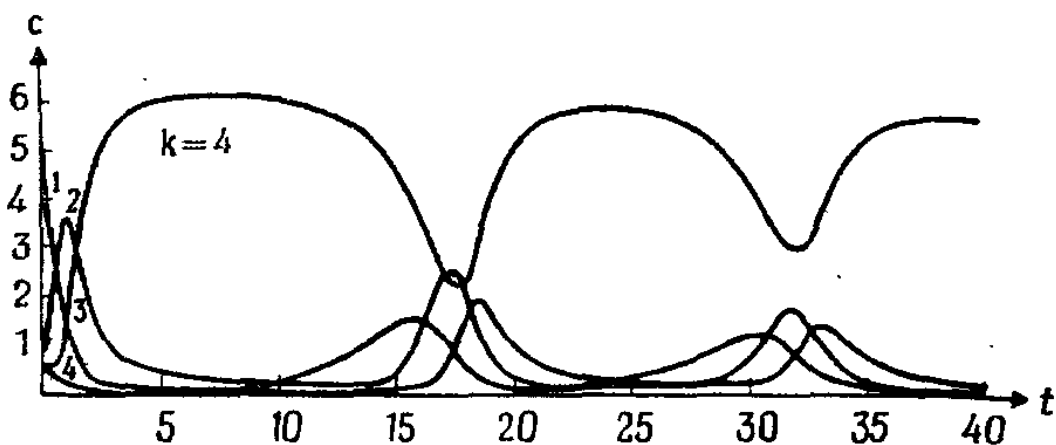


Рис. 21. Стационарные колебания в уравновешенном четырехчленном гиперцикле с несимметричным распределением скоростей [100].
 Константа скорости образования 4-го члена в 10 раз меньше, чем для всех
 других членов гиперцикла.

Более полное решение должно учитывать следующие осложнения:

1. Система E_i , хотя она и связана с I_i , имеет свои характерные скорости роста. Следовательно, необходимо ввести два набора концентрационных переменных: x_i — для ферментов E_i и y_i — для кодирующей системы I_i .

2. Скорость редупликации носителей кода не пропорциональна произведению $x_{i-1}y_i$. Приближение Михаэлиса — Ментен (в котором предполагается большой избыток субстрата над ферментом) не удовлетворительно. Концентрации комплексов между E_{i-1} и I_i , обозначаемые z_i , следует вычислять на основании закона действия масс. Их можно представить в виде

$$z_i = \sqrt{x_{i-1}y_i} \operatorname{tg} \frac{\alpha}{2} = \frac{x_{i-1} + y_i + K_i^{-1}}{2} (1 - \cos \alpha), \quad (\text{VI.3})$$

причем

$$\sin \alpha = \frac{2\sqrt{x_{i-1}y_i}}{x_{i-1} + y_i + K_i^{-1}}, \quad (\text{VI.4})$$

где x_{i-1} и y_i — полные концентрации E_{i-1} или I_i соответственно (независимо от того, свободны они или связаны в комплексе) и K_i — константа стабильности комплекса между E_{i-1} и I_i .

3. Скорость образования I_i включает два члена: линейный член и член, пропорциональный z_i . Всегда имеется какое-то воспроизводимое образование I_i , происходящее без специфического содействия E_{i-1} . Если бы мы имели только член второго порядка, нуклеация цикла была бы очень маловероятным процессом.

4. Связанные системы кинетических уравнений можно записать следующим образом¹:

$$\begin{aligned} \dot{y}_i &= (\mathcal{F}'_{I_i} - \mathcal{D}_{I_i}) y_i + \mathcal{F}''_{I_i} \cdot z_i, \\ \dot{x}_i &= \mathcal{F}_{E_i} \cdot y_i - \mathcal{D}_{E_i} \cdot x_i, \end{aligned} \quad (\text{VI.5})$$

где z_i дается уравнением (VI.3).

¹ Различение концентраций свободных [($y_i - z_i$) или ($x_i - z_i$) соответственно] и связанных (z_i) молекул не изменило бы общего вида этих уравнений, потому что в них входят члены, пропорциональные как y_i (или x_i), так и z_i .

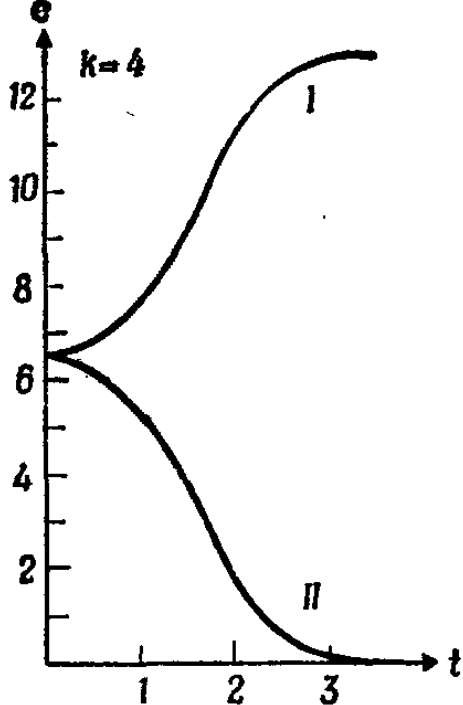


Рис. 22. Конкуренция между двумя 4-членными гиперциклами, у одного из которых константы скорости образования на 10% меньше, чем у другого: $\mathcal{F}'_{11} = \mathcal{F}'_{12} = \mathcal{F}'_{13} = \mathcal{F}'_{14} = 1,0$; $\mathcal{F}'_{21} = \mathcal{F}'_{22} = \mathcal{F}'_{23} = \mathcal{F}'_{24} = 0,9$ (ср. [100]).

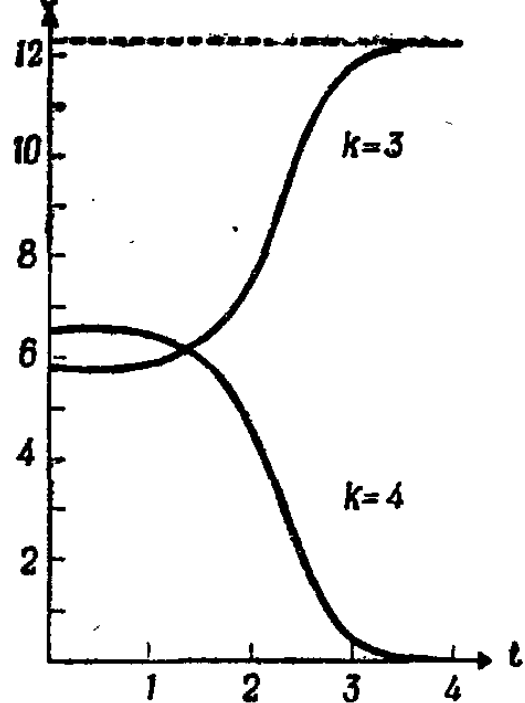


Рис. 23. Конкуренция между трех- и четырехчленными циклами, имеющими одинаковые индивидуальные кинетические параметры, но различающимися по начальным концентрациям [100]. Трехчленный цикл «побеждает».

Для трехчленного цикла $\sum_k x_k^0 = 5,76$; для четырехчленного — $\sum_k x_k^0 = 6,50$. Отметим, что концентрация каждого вида молекул, участвующих в четырехчленном цикле, примерно на 15% меньше, чем концентрация молекул в трехчленном цикле.

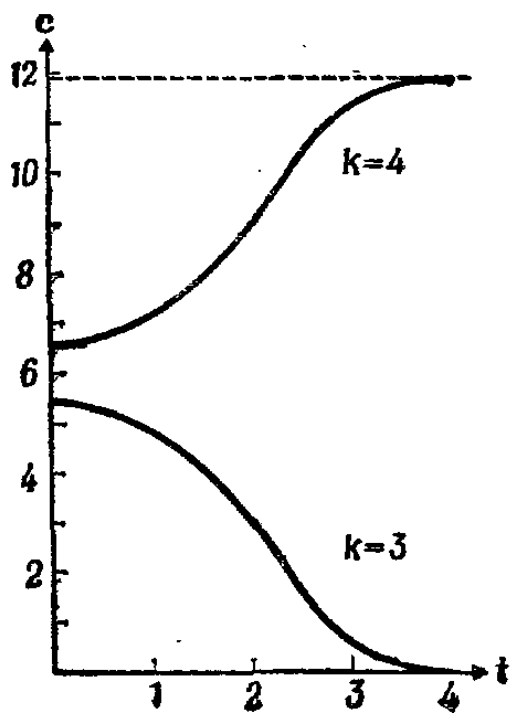


Рис. 24. Конкуренция между трех- и четырехчленными циклами, которые имеют одинаковые индивидуальные кинетические параметры, но различаются по начальным концентрациям [100].

Для трехчленного цикла $\sum_k x_k^0 = 5,4$; для четырехчленного $\sum_k x_k^0 = 6,5$. Отметим, что здесь

побеждает четырехчленный цикл, хотя концентрация каждого вида входящих в него молекул и в этом случае меньше, чем концентрация молекул в трехчленном цикле.

Решение (полученное при ограничениях постоянной общей организации) имеет область второго порядка — как обсуждалось выше — с совершенно аналогичными свойствами; оно выполняется при условии

$$x_{i-1} + y_i \ll K_i^{-1}.$$

Если одна из концентраций превышает K_i^{-1} , решения становятся квазилинейными и колебания исчезают, но это происходит в общем случае уже после того, как система подвергалась резкому отбору при прохождении через нелинейную область. Дальнейшие подробности см. [100].

Выводы теории, относящиеся к селекционному поведению конкурирующих циклов, можно суммировать следующим образом. Под действием ограничений, налагаемых отбором, различные гиперциклы конкурируют друг с другом. Выживает только одна система, характеризующаяся максимальной функцией ценности, которая имеет очень сложный вид и может быть выражена через параметры скорости и качества, а также через средние концентрации членов цикла. Концентрации отдельных членов могут осциллировать, пока система проходит через нелинейную область. Отбор очень жесткий и этим объясняется единственность кода и хиральности. Всякий раз, когда зарождается цикл, выбирая определенный код и аппарат трансляции (см. § VI. 3) — а он должен это сделать, чтобы воспроизводить свои функциональные свойства, — жесткий отбор приводит к универсальному использованию этого конкретного кода, потому что новые циклы не могут сосуществовать с первым, после того как он стабилизировался. То же самое верно для хиральности. После того как полимеризующие функции избрали данную стереохимическую конфигурацию, этот выбор будет сохраняться и эволюционировать до совершенства, которое требует однородной стереоспецифичности. Хотя а priori обе конфигурации равновероятны, одна из них, которая в результате флуктуаций случайно имелась в наличии в момент нуклеации, впоследствии всегда остается предпочтительной вследствие нелинейного усиления. Для

нелинейных систем с селекционным поведением типа «все или ничего» только один вид может победить. Это не обязательно относится к тем функциям, которые не участвуют в нелинейной связи. Например, олигопептиды, которые не производятся биосинтетическим аппаратом, использующим закодированные сообщения, могут содержать также D-аминокислоты (например, грамицидин S, см. гл. V).

Нелинейность является также причиной отбора против паразитных ветвей, которые не могут расти после того, как произошла нуклеация стабильного гиперцикла, если только они не являются частью какой-либо единицы, кодируемой этим циклом. Колебательное поведение цикла добавляет новые особенности — прежде всего в отношении жесткости отбора. Различные члены цикла воспроизводятся в виде волн, оббегающих вокруг цикла, и отбор заканчивается после того, как эти волны сделают всего несколько оборотов (рис. 22). Такое поведение делает затруднительным сосуществование связанных гиперциклов — если только не возникают связи совершенно особого типа (см. [100]).

Возникновение гиперцикла зависит от наличия аппарата трансляции (см. § VI. 3), причем процесс нуклеации основывается на статистике того же типа, что и для линейных циклов [см. § V. 3, уравнения (V. 14) — (V. 16)]. Однако в отличие от циклических белковых сетей гиперцикл может эволюционировать и тем самым оптимизировать свое функционирование. Во-первых, он не разветвлен, как линейная сеть, рассмотренная в гл. V; во-вторых, он может использовать селективные преимущества, появляющиеся внутри цикла, в той мере, в какой они создают преимущество для воспроизведения генотипической мутации. Такая эволюция может также приводить к сжатию цикла до оптимальных размеров. Эти размеры должны обеспечивать достаточно большую информационную емкость, включая все вспомогательные функции, такие, как полимеризация, трансляция и регулирование (а позже — и более сложные метаболические функции). Определенное преимущество дает объединение различных единиц I_R в коллектив, имеющий относительно большую информационную емкость. Отдельная

единица может кодировать только две функции: *узнавание* — для того чтобы обеспечить циклическую связь с помощью системы E_R — и (иногда) *вспомогательную функцию* (т. е. трансляцию или полимеризацию). Если все функции, которыми обладает данный цикл, должны быть закодированы в одной непрерывной молекуле нуклеиновой кислоты, то с самого начала необходима очень большая точность узнавания (т. е. очень хорошо адаптированные ферменты), иначе все это количество информации не сможет стабильно воспроизводиться; информация будет утекать до тех пор, пока вся полезная информация не будет утрачена.

С другой стороны, если цикл уже выработал достаточно точную систему узнавания, то наличие лигазы, сшивающей различные единицы в одну воспроизводящую цепь, дает определенные селективные преимущества. Это могло быть тем моментом, когда эволюция структур из ДНК стала давать преимущества по сравнению с одноцепочечными структурами из РНК. «Обособление» гиперцикла, который тем самым становится истинной «самовоспроизводящейся» системой, следует рассматривать в связи с «компартиментализацией». Ни «обособление», ни «компартиментализация» не принадлежат к числу неотъемлемых свойств гиперцикла (в отличие от других свойств, упоминавшихся выше). Однако, если они появляются после нуклеации, они могут создать селективное преимущество и, таким образом, представляют собой неизбежное следствие эволюции гиперцикла. Преимущества состоят в использовании мутаций. Мутации, особенно мутация, затрагивающая вспомогательную функцию, оказавшаяся благоприятной, будет использоваться всей «диспергированной» системой и, таким образом, не будет специально благоприятствовать воспроизведению мутанта, если только она не возникла внутри отдельного «отсека». В последнем случае мутация будет благоприятствовать только одной этой системе, которая после обособления будет успешно конкурировать со своими предшественниками. То же самое справедливо для неблагоприятных мутаций, которые, если они возникают в отдельном «отсеке», будут оказывать действие (или даже разрушать) только на этот «отсек» и, таким

образом, исчезнут, тогда как в противном случае они могли бы засорить всю систему.

Подходящими «отсеками» могли бы быть коацерваты, впервые описанные А. И. Опариным [101], или микросферы, состоящие из липидов, или протеиноиды (имеющие в основном гидрофобные боковые цепи); последние, как было показано С. Фоксом [102] и другими, могут спонтанно образовываться в условиях, которые могли способствовать конденсации аминокислот и которые, вероятно, существовали на примитивной Земле.

Липидные микросферы встречаются очень часто, как было показано различными авторами. Их можно воспроизводимо получить и использовать в лабораторных экспериментах [103—106].

§ VI. 3. О возникновении кода

Возможность существования гиперцикла, рассмотренного в § VI. 1 и VI. 2, зависит, очевидно, от наличия кода и аппарата трансляции. Такой код потребовался бы для любой модели, которая использует соответствие между инструктирующей способностью, присущей нуклеиновым кислотам, и функциональным потенциалом трехмерных белковых структур. Проблема возникновения уникального кода (существование которого является фактом) носит, таким образом, более общий характер, нежели вопрос о существовании какой-либо конкретной модели для самоподдерживающегося каталитического гиперцикла.

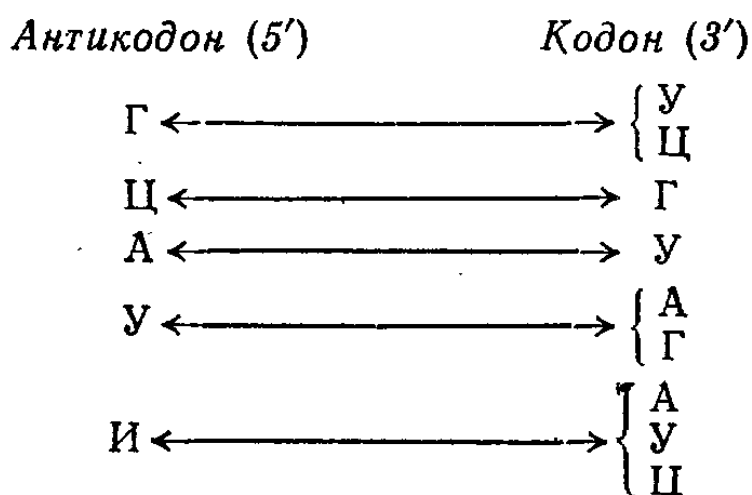
Не содержит ли современная таблица генетического кода (табл. 2) какого-либо намека на свое происхождение? Эта проблема тщательно анализировалась рядом авторов, и в частности в монографии А. Вёзе [107]. Можно напомнить следующие важные факты:

1. Все, что нам сегодня известно, указывает на универсальность кода.

2. Маловероятно, что современный триплетный код развился из предшествующего дублетного или синглетного кода [108]. Как указал Ф. Крик [108], изменение величины кодона повлекло бы за собой полную потерю информации, накопленной к тому времени, если только

при этом расстояние между кодонами в сообщении не оставалось неизменным или же если не существовало простой процедуры перевода информации из старой последовательности кодонов в новую. Это не исключает возможности того, что отдельные единицы в различных положениях кодона могут иметь разный вес в смысле своего информационного содержания.

3. Гипотеза виляний, предложенная Криком [22], дает объяснение сильной вырожденности кода в отношении третьего положения в кодоне (3'-конец). Она объясняет вырожденность кода вырожденностью в комплементарности оснований на 5'-конце антикодона и 3'-конце кодона



что приводит к эквивалентности У и Ц, а также А и Г (или А, У и Ц) в третьем положении кодона.

4. Среднее положение в кодирующем триплете, по-видимому, играет большую роль в определении природы соответствующей аминокислоты (гидрофобная, полярная или заряженная). Эта закономерность, возможно, имеет значение для возникновения кода, основанного на каких-то специфических взаимодействиях между нуклеотидами и аминокислотами [109].

5. Код, по-видимому, отражает принцип оптимизации, как указал И. Рехенберг, который провел формализованное исследование этого вопроса [110]. Для большинства кодонов изменение одной из трех букв приводит к минимальному изменению природы аминокислот; таким образом, это тоже может отражать упомянутую закономерность в отношении среднего положения. Аминокислоты снова были разделены на четыре класса: гидрофобные,

полярные, положительно заряженные и отрицательно заряженные. Сюда можно добавить функциональные корреляции, например кислотные или основные функции боковых цепей аминокислот, а также структурные соответствия, например сходство Фен-Тир и т. д. Любой принцип оптимизации имел бы особое значение для случайного возникновения кода (см. ниже) и оказывал бы влияние на выбор вариантов в процессе нуклеации. Он привел бы также к тому, что информационное содержание последовательности оснований стало бы до некоторой степени инвариантным (если говорить о классах аминокислот) по отношению к перекрывающемуся считыванию кода (в эпоху плохо адаптированных ферментов перекрывания, вероятно, случались довольно часто).

6. Восемь кодонов, состоящих из А и У, дают гораздо большее разнообразие функций, чем восемь триплетов из Г и Ц (табл. 2). Ф. Липман [111] впервые обратил внимание на этот факт и на его возможную связь с простым механизмом возникновения кода. В пользу этой точки зрения имеются следующие аргументы:

а) вероятно, очень большая распространенность А по сравнению с У, Ц и Г в условиях примитивной Земли, что приводит к большей распространенности АУ-пар по сравнению с ГЦ-парами;

б) большая стабильность ГЦ-пар по сравнению с АУ-парами, что позволяет впоследствии ГЦ замещать АУ, если это дает селективное преимущество;

в) недавно найденное [112] относительно высокое содержание АУ-пар в рибосомной РНК митохондрий и хлоропластов, которая, возможно, не подвергалась сильному давлению отбора (в соответствии с гипотезой, что обе эти клеточные органеллы представляют собой прокариотические включения в клетках эукариотов);

г) и, наконец, прагматический аргумент; с точки зрения статистики, нуклеация любого кода происходит легче всего в том случае, когда число различных классов единиц сведено к минимуму.

Как могло возникнуть определенное соответствие между аминокислотами и кодонами или антикодонами?

Без сомнения, простейшее объяснение состояло бы в том, что между этими двумя системами единиц суще-

ствуется специфическое взаимодействие. Было предложено множество моделей такого взаимодействия [113]: отдельный кодон — аминокислота; аминокислота, входящая в щель между двумя спаренными триплетами кодон — антикодон; аминокислота, которая опознается большой сложно свернутой адапторной молекулой — предшественником тРНК [38, 39]; предполагали также, что такие предшественники аминоацил-тРНК имели повышенную стабильность (Л. Оргел, личное сообщение). Ценность любой такой модели зависит исключительно от наличия подтверждающих ее экспериментальных данных, а они до сих пор очень скудны.

Очевидно, можно значительно повысить специфичность узнавания аминокислот нуклеотидами, если не ограничиваться взаимодействием отдельных единиц, т. е. аминокислотой и кодоном или антикодоном. Одной из причин того, почему тРНК является сравнительно большой молекулой, может быть либо необходимость достаточно характерной третичной структуры, обеспечивающей узнавание ее каким-то ферментом, либо то, что сама протяженная структура обладает способностью к узнаванию, как это свойственно молекуле фермента. Обе эти возможности будут обсуждаться в связи со «случайными» моделями. Остается еще одна трудность: как предотвратить случайное соответствие, потому что все перестановки в области антикодона могут, по-видимому, произойти без изменения адапторной структуры.

Бессмысленно развивать модель, основанную лишь на гипотетическом, до сих пор экспериментально не установленном взаимодействии. Во всяком случае, для однозначно определенного начала потребовались бы очень специфические взаимодействия (q , близкое к единице), но очень сомнительно, существуют ли вообще достаточно сильные взаимодействия, обеспечивающие прямое соответствие между антикодонами и аминокислотами¹. Таким образом, встает законный вопрос:

¹ Имеется еще другое соображение: если бы существовало такое изначальное взаимодействие между аминокислотами и кодонами, то можно было бы предположить, что ферменты эволюционируют, используя это взаимодействие, но это привело бы к возможности обратного считывания от белка к нуклеиновой кислоте (что

Если специфические взаимодействия отсутствуют, то не могло ли однозначное кодовое соответствие появиться из случайных комбинаций аминокислот с антикодонами?

Любое специфическое взаимодействие между кодоном (или между любой неразрывно связанной с кодоном структурой) и аминокислотой — когда бы оно ни существовало — может повысить вероятность в остальном ненаправленного начала трансляции. Нас опять-таки интересует не столько конкретная (спекулятивная) модель, сколько оценка вероятностей для случайного возникновения (или нуклеации) возможных предшественников ныне известной адапторной системы узнавания. Другими словами, мы хотим знать, насколько сложной должна быть система, чтобы в ней с ненулевой вероятностью могла возникнуть однозначная трансляция, предполагая, что никакое специфическое взаимодействие не влияет на выбор кодового соответствия.

Предположим, что имеются только такие взаимодействия, существование которых экспериментально обосновано. Например, мы знаем, что способностью к очень специфичному узнаванию тРНК или подобных ей структур обладают трехмерные белковые структуры. Мы знаем также, что аминокислоты могут быть активированы (например, при помощи АТФ) и присоединены к нуклеотидной последовательности, но мы не знаем никакого специфического и изначального взаимодействия между аминокислотами и антикодонами, которое удовлетворительно реализовалось бы без помощи ферментов.

Простейшая модель «случайного» начала трансляции основывается на существенно равных априорных вероятностях различных соответствий между аминокислотами и кодонами или антикодонами. Таким образом, любая из аминокислот $a, b, c \dots$ может а priori связаться с любым из адапторов $A, B, C \dots$, так что соответствие $a - A$,

противоречит «центральной догме» молекулярной биологии). Хотя есть ферменты, допускающие обращение транскрипции (РНК \rightarrow ДНК), нет никаких данных (и их даже трудно представить), что такое обращение считывания может иметь место для трансляции, хотя, если бы такой механизм существовал, он давал бы известные преимущества.

b — B и т. п. отражает только конечный исход (к которому ретроспективно и приспособлены обозначения). Рассмотрим три модели, к которым относится это допущение и для которых поэтому возможен общий подход:

1. Аминокислота узнается гретичной структурой (например, щелью) полинуклеотида, похожего на предшественник тРНК. Антикодон локализуется в открытой петле; но он может не участвовать в узнавании (или не определять его) и фиксации аминокислоты на адапторе, так что а ргіогі данной аминокислоте может быть сопоставлен любой триплет.

2. Модель, аналогичная первой, с той разницей, что аминокислота заменена олиго- или полипептидом (второй адаптор), концевая аминокислота которого должна быть активирована. Такое взаимодействие полинуклеотида с полипептидом может быть гораздо более тесным и специфичным, чем взаимодействие с отдельной аминокислотой, но, как и прежде, для каждой данной взаимодействующей структуры полипептида и полинуклеотида можно произвести любую подстановку концевой аминокислоты и доступного антикодона.

3. Модель, сходная со второй моделью, в которой, однако, полипептиды обладают ферментативной активностью (как предшественники аминокислотсинтетаз) — специфично узнают свободную аминокислоту (или олигопептид), которая должна присоединяться к данной адапторной структуре; эта модель снова допускает любые возможные комбинации антикодон — аминокислота.

В любой из этих моделей мы имеем λ единиц, т. е. аминокислоты a, b, c, \dots , которые нужно поставить в соответствие кодомам A, B, C, \dots с помощью адапторов A', B', C' (комплементарных $A, B, C \dots$). Имеется всего λ^2 возможных соответствий, например, для $\lambda = 2$: aA', aB', bA' и bB' . Различные соответствия возможны потому, что A' , например, согласно нашему допущению, представляет собой целый класс адапторов, которые все имеют один и тот же антикодон, соответствующий A , но в остальном могут взаимодействовать совершенно по-разному с различными аминокислотами или активирующими ферментами. Предполагается, что все соответствия имеют равные априорные вероятности. Рассмотрим

теперь элемент объема¹, в котором имеется λ таких соответствий. Тогда вероятность (P) найти данный набор соответствий будет обратно пропорциональна числу всех возможных наборов, которое в свою очередь равно числу всех сочетаний с повторениями из λ^2 элементов по λ :

$$V_r = \binom{\lambda^2 + \lambda - 1}{\lambda}.$$

Сюда входят также те комбинации, в которых все соответствия относятся к одному типу, например $a - A'$, а также те, в которых данный адаптор соединен со многими различными аминокислотами: $a - A'$, $b - A'$, $c - A' \dots$, или наоборот; короче говоря, допускается любой набор соответствий. Это весьма крайнее (и, по видимому, не очень реалистичное) допущение, но оно сделано для того, чтобы получить нижний предел для вероятностей, так что любое отклонение может только усилить нашу аргументацию.

Среди всех соответствий имеется $\lambda!$ уникальных, т. е. таких, в которых данная аминокислота соединена только с одним (анти-) кодоном и наоборот. Таким образом, вероятность найти элемент объема, где имеется какое-либо *уникальное* соответствие, равна

$$P_{\text{ass}} \sim \frac{\lambda!}{V_r} = \frac{(\lambda!)^2 (\lambda^2 - 1)!}{(\lambda^2 + \lambda - 1)!}. \quad (\text{VI. 6})$$

Такой элемент объема начнет — возможно, с помощью катализатора — однозначно транслировать нуклеотидные последовательности в последовательности аминокислот, но только до тех пор, пока в рассматриваемом элементе объема поддерживается данная конкретная «флук-

¹ Величина этого элемента объема подобрана исходя из условия, что в нем имеется λ соответствий. В этом объеме, конечно, находится гораздо большее число полипептидных и полинуклеотидных последовательностей. Они, однако, не участвуют в фиксации аминокислот на адапторе. Важно лишь следующее соответствие для данной последовательности должно быть специфичным или, лучше, специфичные соответствия должны быть определяющими. В противном случае данная комбинация не будет однозначной и на все будет накладываться довольно значительный «шум».

туация» соответствий. Чтобы стабилизировать этот тип трансляции, мы должны найти среди *нуклеотидных последовательностей* такие, которые после трансляции усиливают использование того же кода. Только такой ансамбль нуклеотидов представлял бы собой стабильный и воспроизводимый источник информации для кода и аппарата трансляции (состоящего из определенного набора адапторов и активирующих ферментов). Для того чтобы отбор благоприятствовал именно данному ансамблю, подавляя другие конкурирующие системы, особенно те, которые не однозначны и поэтому всегда дают некоторое воспроизведение нонсенов, в данной системе должен возникнуть самоусиливающийся гиперцикл, описанный в первых двух параграфах этой главы¹.

Вероятность найти такой набор нуклеотидов, который будет усиливать определенную функцию трансляции, зависит от тех же предпосылок, что и вероятность найти набор белков, выполняющий эту функцию. Если полинуклеотиды каким-то образом транслируются, им будет соответствовать некий набор случайных полипептидных последовательностей; таким образом, мы должны исходить из тех же предположений, как и в том случае, когда мы искали λ специфичных «координаторов», однако теперь не будет вырожденности $\lambda!$, потому что система должна усиливать одно из λ возможных однозначных соответствий. Если концентрации случайных нуклеотидных последовательностей близки к концентрациям случайных полипептидов, вероятность найти определенный набор в данном элементе объема снова будет равна величине, обратной «числу всех сочетаний с повторениями» из λ^2 элементов по λ .

¹ Одна возможность состоит в том, что адапторы с самого начала были достаточно протяженными нуклеотидными структурами, которые могли осуществлять двойные функции:

- а) действовать как специфичные адапторы, имея петлю антикодона, и могли специфично узнаваться активирующими ферментами;
- б) специфичная информация (I_i), содержащаяся в их последовательностях, кодировала бы ферменты E_i , которые являлись членами гиперцикла. Однако на этом этапе нельзя сделать определенного вывода о сложности системы, в которой могла происходить нуклеация; в нее могли входить также адапторы с короткой цепью, распространенность которых а priori высока,

Тогда совместная вероятность приближенно равна

$$P \sim \frac{(\lambda!)^3 [(\lambda^2 - 1)!]^2}{[(\lambda^2 + \lambda - 1)!]^2}. \quad (\text{VI. 7})$$

Эту закономерность иллюстрируют следующие примеры:

| | | | | |
|-----------|-------------------|---------------------|--------------------|--------------------|
| λ | 2 | 4 | 8 | 20 |
| P | $2 \cdot 10^{-2}$ | $1,6 \cdot 10^{-6}$ | $4 \cdot 10^{-16}$ | $5 \cdot 10^{-50}$ |

Совместная вероятность содержит также множитель, описывающий отношение концентраций полинуклеотидных и полипептидных последовательностей. Далее, сделаны совершенно нереалистичные допущения, например одинаковые априорные вероятности для всех последовательностей, которые приводят к тому, что эта формула дает лишь грубую оценку некоторых величин.

Есть еще один важный пункт: процедура состоит в том, чтобы найти вероятность существования и воспроизведения определенной *функции* (т. е. соответствия аминокислота — кодон) в популяции случайных полипептидов, но не в том, чтобы найти вероятность совместного нахождения определенных *последовательностей*. Функция может быть представлена очень большим числом различных полипептидных последовательностей, таким большим, что в действительности эти последовательности можно найти практически в любой случайной популяции (см. эксперименты С. Фокса [102] и его сотрудников, которые показали, что в любом случайном наборе полипептидов можно обнаружить функцию, моделирующую активность химотрипсина). Назовем эту вероятность p — независимо от ее специального вида. Тогда та же вероятность наличия той же *функции* (после трансляции) будет относиться и к популяции случайных нуклеотидных последовательностей (имеющих равные концентрации). В первом случае мы все же имеем $\lambda!$ однозначных выборов для соответствий, тогда как второй выбор должен совпадать с первым. Совместная вероятность тогда становится равной $\lambda!p^2$. Мы ищем вероятность совпадения

функций, а не последовательностей. Если бы мы искали вероятность наличия определенной нуклеотидной последовательности, которая после трансляции дала бы полипептидную последовательность, точно совпадающую с данной (с которой начался определенный тип трансляции), то эта вероятность была бы порядка 10^{-130} (для 100 аминокислот, принадлежащих к 20 классам), подтверждая аргументацию Вигнера (см. гл. I).

Однако в подобных оценках таится достаточно ловушек, чтобы заставить нас отказаться от более подробного анализа, пока не накопится больше экспериментальных данных о каталитических функциях полипептидов со случайными последовательностями. Основной довод — что определенная каталитическая специфичность не является уникальным свойством одной или нескольких данных последовательностей, а встречается довольно часто в любой случайной популяции достаточно большого размера — может быть проверен экспериментально (см. гл. VII). Даже не имея таких данных, мы можем оценить, при какой степени сложности однозначная трансляция, начинающаяся со случайных флуктуаций, становится совершенно невероятной.

В отношении вероятности случайного начала трансляции можно сделать следующие выводы (см. также численные примеры, приведенные после уравнения VI.7). Бинарная система трансляции может возникнуть, по-видимому, очень легко, но с функциональной точки зрения двух единиц (или классов единиц) было бы недостаточно для обеспечения специфичности. Сделано предположение, что четыре класса единиц — минимум, необходимый для начала процедуры оптимизации в эволюции кода. Вероятность нуклеации четырехбуквенной трансляции еще имеет приемлемую величину. Вероятность, отвечающая восьмибуквенному коду, находится, по-видимому, на пределе того, что представляется осуществимым при разумных концентрациях в пределах Земли и в масштабах времени, на протяжении которого происходила ранняя эволюция (оно, вероятно, значительно меньше 10^9 лет $\sim 3 \cdot 10^{16}$ с). Восмибуквенный код могла бы обеспечить система АУ. Далее, восьми аминокислот, по-видимому, достаточно для построения функционально

специфических последовательностей любого типа. Не обязательно даже начинать, имея только 8 аминокислот, — можно иметь 8 (или меньше) классов функционально родственных молекул. То же самое справедливо для инструктирующего кода, который мог бы начаться с вырожденных классов и развиваться дальше в соответствии с некоей процедурой оптимизации [110]. Поэтому вполне возможно, что тот конкретный код, который мы находим теперь, возник из случайной флуктуации; следовательно, нет необходимости предполагать существование очень специфического (прямого или непрямого) взаимодействия между кодоном и аминокислотой. Если это верно, то любая независимо эволюционирующая система («где-то» во Вселенной или «когда-нибудь» в лаборатории) могла бы использовать другой код, но он был бы основан на таких же принципах. Далее, с самого начала теперешний код мог быть и не единственным; однако *нелинейная жесткая процедура отбора гарантирует его универсальность*.

С другой стороны, следует подчеркнуть, что в настоящее время единственно правильным может быть утверждение: «Нельзя исключить возможность того, что ...». Поэтому единственный смысл сделанных выше оценок состоит в том, чтобы найти, при какой степени сложности случайное начало становится слишком маловероятным. Оргел выдвинул интересную идею о том, что в процессе эволюции происходила не нуклеация всего словаря (например, из 4 или 8 букв), а поэтапный или непрерывный переход к образованию системы трансляции, который начался с 1—2 предпочтительных соответствий между адаптором и аминокислотой. Наличие таких «изначальных» соответствий всегда будет увеличивать вероятность нуклеации самовоспроизводящейся функциональной сети.

Наконец, можно задать вопрос: как понимать такое «случайное начало» трансляции с точки зрения физики?

Здесь снова надо прибегнуть к критерию ценности в теории отбора, который аналогичен принципу Пригожина — Глансдорфа в нелинейной необратимой термодинамике. В самоорганизующейся системе с селекционным поведением (что определяется известными свойствами реакционной системы и наложением дополнительных внешних ограничений) появление нового вида или

ансамбля с большей селективной ценностью всегда будет приводить к неустойчивости, т. е. к разрушению прежнего стационарного состояния и построению нового стационарного состояния, в котором доминирует вид или ансамбль с наибольшей селективной ценностью. «Успех» нового вида подчинен определенным ограничениям и может корректно описываться стохастической теорией.

Мы приходим к следующему заключению.

Нуклеиновые кислоты обеспечивают выполнение необходимого условия самоорганизации. Однако, чтобы создать высокую информационную емкость, они нуждаются в каталитически активном связывающем факторе, имеющем большую способность к узнаванию. «Информация» приобретает свой смысл только посредством функциональной корреляции. Любая флуктуация, в присутствии потенциальных связывающих факторов ведущая к однозначной трансляции, и ее усиление посредством образования каталитического гиперцикла создает колоссальное селективное преимущество и вызывает распад прежнего стационарного состояния с некоррелированным самовоспроизведением.

Вследствие такой нестабильности нуклеация этой функциональной корреляции (мы можем назвать ее возникновением жизни) оказывается неизбежным событием — если благоприятные условия существования потока свободной энергии поддерживаются в течение достаточно длительного времени. Это первичное событие не уникально. В любом случае код станет универсальным вследствие нелинейной конкуренции.

ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ ЭВОЛЮЦИИ

Ценность теоретической модели определяется тем, в какой мере она поддается экспериментальной проверке; ценность общей теории зависит от того, в какой мере она может направить такую работу и определить ясные и воспроизводимые условия для сравнительных исследований. Далее хороший эксперимент позволяет сделать выбор между возможными альтернативами, обычно путем исключения неверных.

Эволюционные эксперименты в пробирке пока еще очень немногочисленны, потому что необходимые для этого методы и объекты, т. е. строго определенные виды молекул, стали доступны только в последние годы. Остроумный и простой по замыслу прямой модельный эксперимент такого рода провели С. Спигелман и его группа. Мы обсудим его более подробно, поскольку он типичен для экспериментов именно такого рода, который предлагает наша теория (обзор литературы см. [114]).

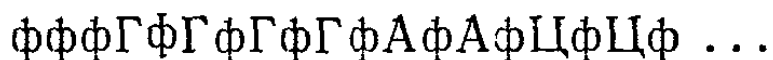
§ VII.1. Система Q β -репликазы

История Q β началась с утверждения, под которым — в то время, когда его высказал С. Спигелман [115], — подписались бы очень немногие из его коллег-биохимиков. Это утверждение состояло в том, что фаг Q β использует специфичный фермент репликации, опознающий только РНК фага Q β . В ответ на весь проявленный скептицизм Спигелман представил высокоочищенный и хорошо охарактеризованный фермент, способный катализировать воспроизведение инфекционной вирусной РНК в бесклеточной среде. Тот факт, что в бесклеточной системе действительно синтезировалась РНК, содержа-

шая все инструкции, был продемонстрирован в следующем классическом эксперименте [116]. Раствор, не содержащий клеток, подвергали серийному разведению (оставляя на каждом этапе достаточно времени для воспроизведения), причем раствор в конечном разведении содержал менее 10^{-15} доли исходного числа природных фаговых матриц, т. е. практически не содержал ни одной такой матрицы, и все-таки этот образец обладал такой же инфекционностью, как и исходный. Более того, использование чувствительного к температуре мутанта исключало возможность того, чтобы источником информации служило что-либо, кроме молекул РНК. Носителем этой информации является только плюс-цепь; поэтому для воспроизведения инфекционных молекул необходим индукционный период, в течение которого должны накопиться комплементарные (неинфекционные) минус-цепи. Воспроизведение всей популяции может инициировать даже одна матричная молекула, и в таком случае оно приводит к образованию клона идентичных потомков. Это было показано в опыте (см. Р. Левисон и С. Спигелман [117]), в котором образец разбавляли и разливали в пробирки. Благодаря синхронной инициации синтеза можно было идентифицировать пробирки, получившие 0, 1, 2 или более молекул матрицы — в строгом согласии с распределением Пуассона. Опыты по «обрубанию» показали, что для узнавания ферментом необходимы участки, распределенные по всей длине молекулы, включая обе концевые группы: ни одну из двух половинок цепи по отдельности репликаза не узнавала. Так как плюс- и минус-цепи реплицируются при помощи одного и того же фермента, можно ожидать определенной симметрии в распределении комплементарных областей по длине цепи. Любые «внутренние» комплементарные участки одной цепи должны зеркально отражаться в комплементарной цепи, следовательно, они, вероятно, располагаются в 3'- и 5'-половинках симметрично.

Эту очень интересную структурную проблему можно решить путем анализа последовательности, который проводится (и частично завершен) в лаборатории Вейсмана [118]. Определением последовательности занималась также группа Спигелмана [119], которая показала, что 3'- и

5'-концы действительно в какой-то мере комплементарны. 5'-конец плюс-цепи имеет строение



Минус-цепь тоже заканчивается группой фффГ на 5'-конце и имеет более длинную последовательность пуринов. Из этого следует, что 3'-конец плюс-цепи имеет область, комплементарную 5'-концу; из-за комплементарности с минус-цепью она должна содержать много урацила и цитозина и также кончатся цитозином.

Открытие системы Q β может иметь как теоретическое, так и практическое значение. «Неожиданное структурное разнообразие и тонкость» (цитируется по С. Спигелману) отдельных молекул РНК делает понятным, каким образом силы отбора могли вмещаться во взаимодействие между нуклеиновыми кислотами и белками и направлять доклеточную эволюцию. В практических целях можно использовать специфичные участки узнавания: а) в сочетании с дегенерировавшей неинфекционной РНК, что может препятствовать фаговой инфекции; б) в качестве специфичных индукторов синтеза РНК в других, соответственным образом модифицированных системах.

§ VII.2. Дарвиновская эволюция в пробирке

Получение очищенной Q β -репликазы дало возможность поставить ряд интереснейших экспериментов, в которых определенный вид молекул подвергали давлению отбора посредством «серийных переносов» и заставляли их, таким образом, проделывать «эволюцию в пробирке». Эксперименты такого рода начинают с того, что берут стандартную реакционную смесь (см. [121]): 0,25 мл раствора содержат 10^{-1} М трис-НСl, рН 7,4, $2 \cdot 10^{-2}$ М MgCl₂, $3 \cdot 10^{-3}$ М ЭДТА, по 200 мкмоль АТФ, УТФ, ГТФ, ЦТФ (УТФ помечен ³²P в α -положении, причем 4000 расп/мин соответствует 1 мкг синтезированной РНК) и 40 мкг Q β -репликазы (очищенной путем центрифугирования в градиентах CsCl и сахарозы). Подробное описание процедуры определения нуклеотидного состава, седимента-

ционного анализа, а также различных анализов на ферментативную активность см. [120, 121].

Инкубируя эту смесь с вирусной РНК (в данном случае РНК чувствительного к температуре мутанта *ts-1*), инициируют синтез фермента и затем готовят

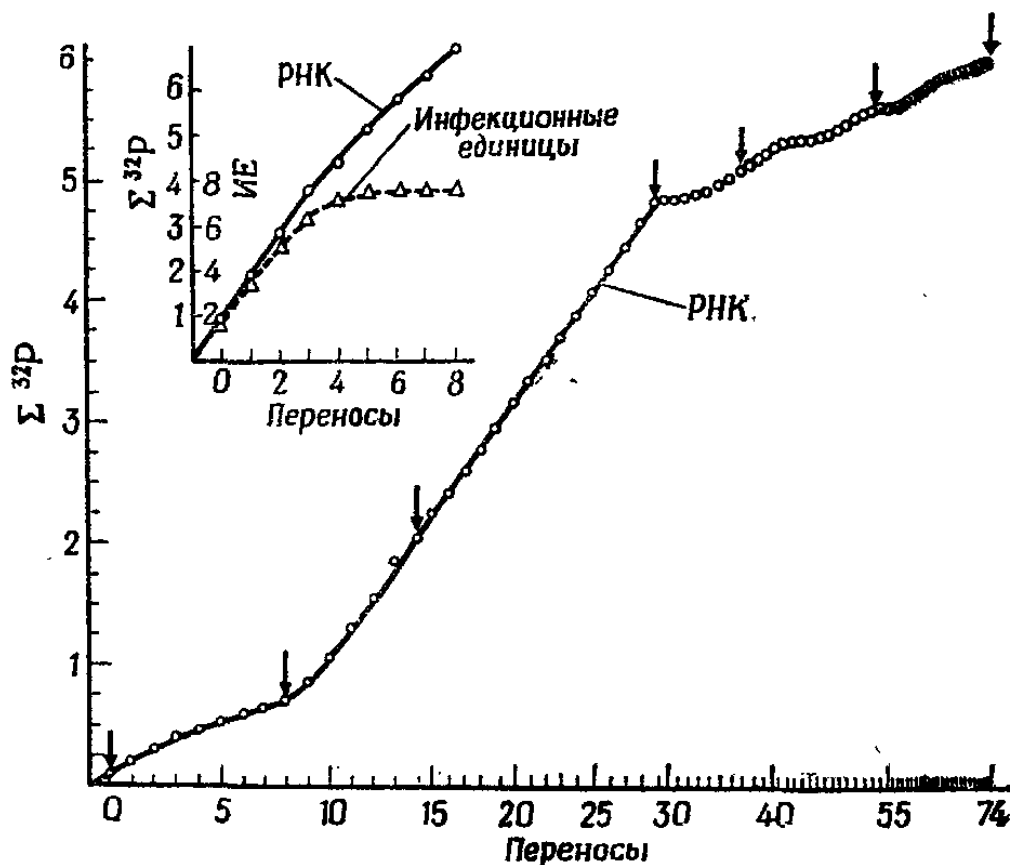


Рис. 25. Эксперимент с серийными переносами РНК фага Q β [121].

Описание в тексте. Стрелки над переносами 0, 8, 14, 29, 37, 53 и 73 указывают, что в этих случаях от 0,01 до 0,1 доли продукта использовали в качестве заправки для вспомогательной реакции, продукт которой подвергали седиментационному анализу в градиенте сахарозы. Время инкубации: 20 мин (переносы 0—13), 15 мин (переносы 14—29), 10 мин (переносы 30—38), 7 мин (переносы 39—52) и 5 мин (переносы 53—74). Результаты показывают, что биологически компетентная РНК перестает появляться после 4-го переноса. По оси ординат — радиоактивность (расп/мин на 0,25 мм) $\times 10^{-5}$ (для большого графика) и $\times 10^{-4}$ (для малого).

серию разбавлений, перенося каждый раз после определенного времени инкубации 0,02 мл реакционной смеси в 0,25 мл свежего стандартного раствора. Первая реакция инициировалась 0,2 мкг РНК *ts-1*, которую инкубировали в течение 20 мин при 35 °С. Затем время инкубации снижали от 20 мин (переносы 1—13) до 15 мин (переносы 14—29), до 10 мин (переносы 30—38), до 7 мин (переносы 39—52) и, наконец, до 5 мин (переносы 53—74), после чего получали конечный продукт. При каждом переносе брали 0,02 мл смеси для подсчета импульсов и

еще 0,02 мл для инициирования реакции в следующей пробирке. В переносах 0, 8, 14, 29, 37, 53 и 73 некоторое количество продукта использовалось для инициирования синтеза РНК, которую брали затем для седиментационного анализа в градиенте сахарозы.

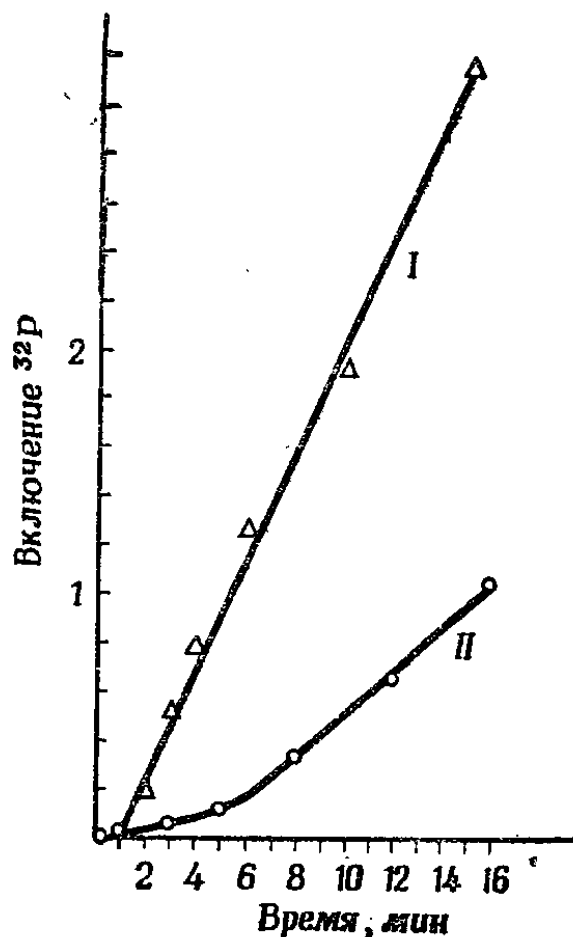
Анализ продуктов дал следующую информацию о реплицированных молекулах РНК (см. рис. 25): инфекционность была утрачена после 4-го переноса; молекулярные массы РНК-матриц по мере переносов более или менее постоянно снижались, пока после 74-го переноса не получился конечный продукт, утративший около 83% своей исходной генетической информации. Из 3600 нуклеотидных остатков родительской молекулы сохранилось только 550. Параллельно с уменьшением молекулярного веса наблюдается увеличение скорости включения ^{32}P , так что к 74-му переносу скорость включения на один нуклеотид возросла в 2,6 раза по сравнению с начальной скоростью. Это прямо показывают исследования кинетики включения нуклеотидов в условиях насыщения. Рис. 26 служит примером: увеличение скорости в линейной области сопровождается уменьшением периода индукции, в течение которого количество минус-цепей должно возрасти вплоть до «уравновешенного» отношения плюс- и минус-цепей. В условиях специальных селекционных ограничений можно было наблюдать дальнейшее увеличение скорости, сопровождающееся уменьшением длины цепи до 180 нуклеотидных остатков. Эта фракция «мини-монстров» исследуется также в лаборатории Л. Оргела в Ла Хойя, Калифорния (личное сообщение С. Спигелмана).

На первый взгляд результаты этих экспериментов отражают только тривиальный «эволюционный» ответ на данное предписание: «размножаться как можно быстрее». Молекулы РНК, освобожденные от необходимости быть инфекционными, приспособляются к таким «райским» условиям, отбрасывая всю информацию, которая не нужна для быстрой репликации. Однако они не могут сократить процесс репликации, просто обрывая цикл, прежде чем будет достигнуто место терминации, потому что для узнавания необходимы различные части молекулы, в частности оба ее конца. На самом деле сокра-

щение временных интервалов между переносами не вынуждает к такому поведению. Реакция не обрывается вследствие переноса, и более вероятно, что молекула фермента, которая начала процесс репликации на матрице, закончит свое дело, а не отделится от нее, чтобы искать новую матрицу.

Рис. 26. Сравнение кинетики синтеза 74-го варианта (кривая I) и исходной РНК мутанта ts фага Q β (кривая II) [121].

Иницированы две стандартные реакции (см. текст) в 0,25 мл смеси при 35 °С. В одной реакции затравкой служила одноцепочечная «вариантная» РНК (74-й перенос), очищенная гель-фильтрацией, в другой — РНК мутанта ts фага Q β (в обоих случаях количество затравки выше уровня насыщения). В указанные моменты времени брали пробы по 0,02 мл и анализировали их на включение 32 P-УТФ (расп/мин/0,02 мл) $\cdot 10^{-3}$.



В сущности в эксперименте с серийными переносами создается приближение к условиям «постоянных реакционных сил», или постоянной «общей организации».

Концентрации, которые были в исходной стандартной реакционной смеси, — по крайней мере, концентрации высокоэнергетических мономеров АТФ, УТФ, ГТФ и ЦТФ — восстанавливаются при каждом переносе, поэтому средство реакций образования в среднем находится на постоянном уровне. Сокращение промежутков времени между отдельными переносами позволяет компенсировать растущую скорость репликации. Однако, как показывает рис. 25, в опыте Спигелмана все-таки был некоторый дрейф стационарного состояния. Он вызван в основном тем, что селекционные ограничения выбирались

интуитивно (что совершенно оправдано, поскольку эксперименты ставили с целью получить качественные результаты). Более строго заданные селекционные ограничения, обеспечивающие сохранение стационарного состояния при постоянных условиях, могут влиять на скорость эволюции; такие условия потребовались бы для количественных выводов.

Качественный вывод состоит в том, что система всегда благоприятствует виду, имеющему максимальную селективную ценность. В «райских» условиях экспериментов *in vitro* инфекционность является не условием, а помехой для быстрого воспроизведения. Этот пример ясно показывает, что, хотя селективная ценность всегда определяется параметрами скорости и узнавания A , Q и D , эти параметры могут резко меняться при изменении условий среды. В последующих работах С. Спигелман и сотрудники сообщили о выделении целого ряда мутантов, которые адаптировались к различным модификациям условий отбора *in vitro*. В число свойств, которые можно было встроить в эти «варианты», входит устойчивость к аналогам нормальных рибонуклеозидтрифосфатов или к таким ингибиторам, как этидийбромид. Оказалось возможным также производить отбор на виды с повышенным молекулярным весом, фиксируя фермент на мембране и благоприятствуя тем самым (длинным) цепям, которые лучше адсорбируются. Эксперименты с «голоданием» по одному из оснований, например цитозину, не дали мутантов, бедных этим основанием: фермент сумел приспособиться к изменившимся условиям и повысить эффективность включения цитозина. Все эти эксперименты свидетельствуют о колоссальной структурной и функциональной вариабельности и адаптивной способности одноцепочечных РНК и об их возможной роли на ранних стадиях эволюционного процесса.

§ VII.3. Количественные исследования отбора

Важные данные можно получить из количественных экспериментов по эволюции. Хотя в этом направлении еще предстоит много работы, некоторую информацию можно извлечь из уже опубликованных данных.

«Селективное преимущество» всегда относится к воспроизведению всего вида или ансамбля. Если скорость воспроизведения зависит от длины цепи, то просто утрата ненужной информации, приводящая к укорочению цепи — без увеличения «характеристической» скорости воспроизведения (которая относится к отдельной единице), — уже может оказаться «селективным преимуществом». Однако в условиях насыщения, когда скорость также становится независимой от концентрации матрицы, такая зависимость от длины цепи исчезнет. Термин «насыщение» означает, что практически все молекулы фермента E связаны с матрицами I_i , т. е. $\sum_k x_{EI_k} \approx x_{E_0}$.

Такие условия явно обладают рядом преимуществ при кинетических исследованиях, и Спигелман использовал их в своих опытах. Если скорость воспроизведения не зависит от концентрации матриц, число матричных молекул (определяемое по включению ^{32}P) растет линейно во времени (рис. 26). Наличие периода индукции, заметного на рис. 26, указывает на различие в параметрах скорости и (или) связывания плюс- и минус-цепей. Если обозначить плюс- и минус-цепи I_+ и I_- соответственно, а их концентрации x_{I_+}/x_{I_-} , то кинетические уравнения для условий, указанных в подписи к рис. 26, будут иметь следующий вид:

$$\begin{aligned} \dot{x}_{I_+} &= \mathcal{F}_+ x_{EI_-}, \\ \dot{x}_{I_-} &= \mathcal{F}_- x_{EI_+}, \end{aligned} \quad (\text{VII. 1})$$

где EI_+ и EI_- — комплексы фермент — матрица. Процедура, аналогичная выводу уравнения Михаэлиса — Ментен, приводит к выражению

$$x_{EI_+} + x_{EI_-} = x_{E_0} \frac{K_+ x_{I_+} + K_- x_{I_-}}{1 + K_+ x_{I_+} + K_- x_{I_-}} \quad (\text{VII. 2})$$

(E_0 — суммарное количество свободного и связанного фермента).

K_+ и K_- — константы стабильности комплексов фермент — матрица или их аналоги для стационарного состояния. Если затравкой служили одни только плюс-цепи, то начальный наклон относится только к

к включению в минус-цепи:

$$\text{«Начальный наклон»} \sim \mathcal{F}_- x_{E_0}. \quad (\text{VII. 3})$$

После накопления достаточно большого числа минус-цепей достигается постоянное «равновесное» отношение $\bar{x}_{I_+}/\bar{x}_{I_-}$.

Принимая, что

$$d/dt (\bar{x}_{I_+}/\bar{x}_{I_-}) \approx 0$$

или

$$\dot{x}_{I_+}/\bar{x}_{I_+} \approx \dot{x}_{I_-}/\bar{x}_{I_-} \quad (\text{VII. 4})$$

и используя уравнение (VII. 1), получаем

$$\bar{x}_{I_+}/\bar{x}_{I_-} = \sqrt{K_+ \mathcal{F}_+ / K_- \mathcal{F}_-} \quad (\text{VII. 5})$$

или для конечного наклона, для тех условий, когда ^{32}P включается и в плюс- и в минус-цепи:

$$\text{«Конечный наклон»} \sim \frac{\mathcal{F}_+ \sqrt{K_- \mathcal{F}_-} + \mathcal{F}_- \sqrt{K_+ \mathcal{F}_+}}{\sqrt{K_- \mathcal{F}_-} + \sqrt{K_+ \mathcal{F}_+}} x_{E_0}. \quad (\text{VII. 6})$$

Последнее выражение при $K_+ = K_-$ сводится к

$$\sqrt{\mathcal{F}_+ \mathcal{F}_-} \cdot x_{E_0},$$

а для $\mathcal{F}_+ = \mathcal{F}_- \equiv \mathcal{F}$ к

$$\mathcal{F} x_{E_0}$$

(ср. с результатами § IV. 2, которые, однако, не относятся к ферментативному воспроизведению в области насыщения).

Из экспериментальных данных следует вывод, что плюс-цепь образуется быстрее, чем минус-цепь (каждая использует в качестве матрицы комплементарную цепь). Различие в наклонах можно объяснить различиями в значениях \mathcal{F} и K (а не только K). Если различие обусловлено только параметром скорости, то \mathcal{F}_+ может быть в 100 раз больше, чем \mathcal{F}_- . Для количественного расчета необходимы, однако, более подробные экспериментальные данные. Конечный «вариант» (после 74 переносов) имеет более высокую среднюю скорость обра-

зования. Конечный наклон возрастает в 2,6 раза (по сравнению с конечным наклоном исходной РНК фага Q β). Хотя имеются веские основания считать, что это изменение в основном обусловлено ростом параметров скорости, экспериментальные данные не позволяют точно вычислить отдельные \mathcal{F} и K . Уменьшение периода индукции может быть обусловлено как ростом скоростей, так и присутствием плюс- и минус-цепи в начальной фазе. (Отметим, что периода индукции не должно быть, если обе цепи с самого начала представлены в своем «уравновешенном» соотношении.) Имеющиеся в настоящее время данные нельзя считать достаточными для обоснования вывода, что увеличение наклона в 2,6 раза (это определенно означает увеличение средней скорости копирования отдельных единиц) означает 15-кратное увеличение скорости воспроизведения индивидуальных «вариантных» молекул РНК (по сравнению с исходной РНК фага Q β). Необходимо не только получить более детальные кинетические данные, но и доказать, что укорочение молекулы линейно связано с уменьшением времени ее полного воспроизведения, что маловероятно, если матрицу могут считывать одновременно несколько молекул фермента, и что должно зависеть также от концентраций. Тем не менее скорость воспроизведения отдельных молекул должна заметно расти, иначе включение ^{32}P — как показано на рис. 26 — не возрастало бы с числом переносов, несмотря на уменьшение времени инкубации в 4 раза.

Следует также подчеркнуть, что сделанные выше оценки относятся к самой простой модели взаимодействия фермент — матрица, предполагающей стационарное состояние, и были даны только для того, чтобы продемонстрировать возможности получения более полной информации — путем количественного исследования кинетики и механизмов.

Необходимо сделать еще одно замечание о *механизме отбора*. Основываясь на обсуждении в гл. II, можно было бы утверждать, что в области насыщения механизм отбора не действует. Процесс перестает быть «автокаталитическим», если x_i уже не пропорциональны x_i , а являются постоянными. Однако это было бы верно

только в том случае, если бы мы рассматривали репликацию идентичных матриц. Если же имеется несколько конкурирующих последовательностей I_i , каждая из которых может образовать комплекс с ферментом, характеризующийся константой стабильности

$$K_i = \frac{x_{EI_i}}{x_{I_i} x_E}, \quad (\text{VII. 7})$$

то мы получим

$$\sum_k x_{EI_k} = x_{E_0} \frac{\sum_k K_k x_{I_k}}{1 + \sum_k K_k x_{I_k}} \quad (\text{VII. 8})$$

или

$$x_{EI_i} = x_{E_0} \frac{K_i x_{I_i}}{1 + \sum_k K_k x_{I_k}}$$

и

$$\dot{x}_{I_i} = \frac{x_{E_0}}{1 + \sum_k K_k x_{I_k}} \mathcal{F}_i K_i x_{I_i} \quad (\text{VII. 9})$$

для $K_i \equiv K$ (VII. 9) сводится к

$$\dot{x}_{I_i} = \frac{\mathcal{F}_i x_{E_0} x_{I_i}}{x_{I_i} + K^{-1} + \sum_{k \neq i} x_{I_k}}. \quad (\text{VII. 10})$$

Физическая интерпретация следующая: даже в области насыщения, которая теперь определяется условием $\sum_k K_k x_{I_k} \geq 1$, определенный мутант I_i , появляющийся в небольшом числе копий, воспроизводится в соответствии с механизмом экспоненциального отбора ($\dot{x}_{I_i} \sim x_{I_i}$), до тех пор пока концентрация отобранного вида не достигнет уровня, означающего доминирование: $x_{I_i} \geq K^{-1} + \sum_{k \neq i} x_{I_k}$.

¹ Для простоты здесь имеется в виду механизм «самоузнавания». Аналогичный основной результат получается и для комплементарного узнавания.

§ VII.4. Эксперименты типа «минус один»

Эволюционные эксперименты, прототипом которых служит описанный выше эксперимент с РНК фага Q β , могут действительно дать количественное описание эволюционных процессов на молекулярном уровне. Цифры, однако, будут иметь смысл только в том случае, если для реакции будут подобраны хорошо определенные и воспроизводимые условия и будут наложены постоянные селекционные ограничения. Если молекулярные процессы представляют собой сложную картину реакций с участием нескольких одновременно эволюционирующих видов, то механизм может стать безнадежно сложным.

Как следует проводить эксперименты *in vitro* в подобном случае?

Существует возможность, которую мне хочется назвать подходом «минус один». «Минус один» — так называют музыкальную запись, когда произведение, в исполнении которого должны принимать участие n музыкантов, записывается в исполнении $(n - 1)$ музыкантов. Недостающую партию исполняет (у себя дома) музыкант — обычно дилетант, который таким образом получает возможность играть в сопровождении оркестра.

Предлагаемый эволюционный эксперимент основан именно на этом принципе (как и эксперимент Спигелмана с Q β -репликазой): все виды, кроме одного, присутствуют в окончательной форме. Недостающий вид, эволюцию которого из случайных предшественников намечено проследить, варьирует в каждом эксперименте (число экспериментов должно быть больше числа участвующих в них видов, чтобы можно было провести также анализ связей, поскольку суммарный процесс не является просто суммой отдельных процессов). Общую скорость эволюции такой системы можно тогда оценить из совокупности всех данных.

Очевидно, что эти эксперименты требуют постоянных и воспроизводимых реакционных условий, причем ограничения, налагаемые отбором, также должны поддерживаться постоянными, иначе результаты окажутся несопоставимыми. Можно представить себе автоматическое устройство, которое поддерживало бы концентрации

мономеров, полимеров и ферментов на постоянных уровнях, причем регулирующее воздействие на систему (непрерывное разбавление или определенные серийные переносы) оценивалось бы по результатам автоматического анализа включения ^{32}P (и/или других меток). Обособленная система, поддерживающая концентрации мономеров, полимерной РНК и ферментов, может быть реализована с помощью полупроницаемых стенок, сделанных из миллиметровых фильтров, и т. д. Самой интересной — но и трудной — частью таких экспериментов было бы включение в них бесклеточного синтеза белка.

Наконец, можно сделать вывод, что такие эксперименты *in vitro* позволяют, по-видимому, проверить различные модели возникновения кода и эволюции молекулярного аппарата трансляции.

ЗАКЛЮЧЕНИЕ

§ VIII. 1. Пределы теории

Что объясняет эта теория? Она дает общий принцип отбора и эволюции на молекулярном уровне, основанный на критерии устойчивости стационарных состояний в (нелинейной) термодинамической теории. Эволюция представляется неизбежным событием, если задано присутствие определенного вещества с определенными автокаталитическими свойствами и если поддерживается такая величина потока (свободной) энергии, которая необходима для компенсации стационарного производства энтропии. Эта теория дает количественную основу для постановки лабораторных экспериментов по эволюции и обработки их результатов.

Что в принципе может объяснить эта теория? Она может дать правила построения простых молекулярных моделей, соответствующих возможным предшественникам «живых» клеток. Были исследованы четыре такие модели; удалось показать, что лишь одна из них удовлетворяет всем требованиям, обеспечивающим эволюцию к нынешнему состоянию клеточной формы жизни.

Чего никогда не объяснит эта теория? Она никогда не даст точного описания исторического пути эволюции. Это «никогда» — следствие стохастической природы процессов, участвующих в эволюции, а также огромнейшего числа возможных вариантов. То же самое относится к предсказанию будущего развития за пределами некоторых временных границ. Поэтому: «О чем нечего сказать, о том следует молчать» [122].

§ VIII.2. Концепция «ценности»

Когда я читал лекцию на эту тему в Вейцмановском институте, мой друг Шнейор Лифсон спросил меня: «Новые концепции обычно сопровождаются введением новой

мономеров, полимеров и ферментов на постоянных уровнях, причем регулирующее воздействие на систему (непрерывное разбавление или определенные серийные переносы) оценивалось бы по результатам автоматического анализа включения ^{32}P (и/или других меток). Обособленная система, поддерживающая концентрации мономеров, полимерной РНК и ферментов, может быть реализована с помощью полупроницаемых стенок, сделанных из миллиметровых фильтров, и т. д. Самой интересной — но и трудной — частью таких экспериментов было бы включение в них бесклеточного синтеза белка.

Наконец, можно сделать вывод, что такие эксперименты *in vitro* позволяют, по-видимому, проверить различные модели возникновения кода и эволюции молекулярного аппарата трансляции.

ЗАКЛЮЧЕНИЕ

§ VIII. 1. Пределы теории

Что объясняет эта теория? Она дает общий принцип отбора и эволюции на молекулярном уровне, основанный на критерии устойчивости стационарных состояний в (нелинейной) термодинамической теории. Эволюция представляется неизбежным событием, если задано присутствие определенного вещества с определенными автокаталитическими свойствами и если поддерживается такая величина потока (свободной) энергии, которая необходима для компенсации стационарного производства энтропии. Эта теория дает количественную основу для постановки лабораторных экспериментов по эволюции и обработки их результатов.

Что в принципе может объяснить эта теория? Она может дать правила построения простых молекулярных моделей, соответствующих возможным предшественникам «живых» клеток. Были исследованы четыре такие модели; удалось показать, что лишь одна из них удовлетворяет всем требованиям, обеспечивающим эволюцию к нынешнему состоянию клеточной формы жизни.

Чего никогда не объяснит эта теория? Она никогда не даст точного описания исторического пути эволюции. Это «никогда» — следствие стохастической природы процессов, участвующих в эволюции, а также огромнейшего числа возможных вариантов. То же самое относится к предсказанию будущего развития за пределами некоторых временных границ. Поэтому: «О чем нечего сказать, о том следует молчать» [122].

§ VIII.2. Концепция «ценности»

Когда я читал лекцию на эту тему в Вейцмановском институте, мой друг Шнейор Лифсон спросил меня: «Новые концепции обычно сопровождаются введением новой

константы. А у Вас?» Прежде чем ответить на этот вопрос, я позволю себе провести различие между двумя видами концепций. Одни я буду называть «новой физикой», а другие — «новыми», но выводимыми из прежних «концепциями».

До сих пор мы лишь дважды испытали наступление «новой физики». Это дало нам две мировые фундаментальные постоянные: постоянную Планка, фигурирующую в квантовомеханическом соотношении неопределенностей, и скорость света, возведенную в ранг фундаментальной мировой константы только с появлением теории относительности. Такое вполне могло бы случиться и в третий раз, поскольку, как однажды пошутил Гейзенберг, мы все-таки пользуемся тройственной системой: сантиметр — грамм — секунда. Термин «новая физика» означает отказ от всеобщей применимости ранее принятых фундаментальных принципов под влиянием экспериментальных фактов, полученных при ясных и вполне определенных условиях и тем не менее не согласующихся с результатами прежней теории.

С другой стороны, «новая концепция» второго рода не обесценивает ни один из ранее установленных принципов; она лишь затрагивает какой-то новый аспект, и ее можно вывести из уже известных принципов. Такая ситуация возникает, когда существуют некоторые экспериментальные факты, не получившие объяснения — скорее по причине недостаточно глубокого осмысливания и отсутствия нужного подхода, нежели из-за нарушения какого-либо фундаментального закона. Превосходным примером служит статистическая концепция термодинамики, выдвинутая Больцманом после того, как стало ясно, что материя состоит из молекул и атомов, к которым должны быть применимы законы механики Ньютона (позже уступившие место квантовомеханическим законам). Единственная трудность состояла в многочисленности этих частиц (например, 10^{24}), для описания каждой из которых нужно задать три пространственные координаты и три проекции момента. Прорыв в этой проблеме наступил с введением статистических методов, которые дали возможность получать функции распределения и характеризовать макроскопические состояния усреднен-

ными величинами, такими, как температура. (Квантово-механические правила усреднения, установленные позднее, оказались даже проще классических.) Довольно быстро выяснилось, что статистическая концепция термодинамики требует введения новой (хотя и выводимой из прежних понятий) величины; эта новая величина выражает количество «информации», теряемой при процедуре усреднения по всем (Z) состояниям, между которыми может быть распределена энергия. *Величиной, характеризующей эту «убыль знания», является энтропия*, и ее физический смысл ясен из больцмановского соотношения, которое (для микроканонического ансамбля) может быть записано в простой форме:

$$S = k \ln Z.$$

Если энтропия описывает «убыль информации», обусловленную тем, что Z микросостояний представляется одним (усредненным) числом, то соотношение того же типа могло бы быть использовано и для описания самой «информации», поскольку «информация» характеризуется единственным выбором из Z возможных а priori равновероятных вариантов (см. уравнение (I. 2)).

Подобным же образом, если речь идет о микросостояниях с различными априорными вероятностями, то средняя (нормированная) величина информации может быть описана по аналогии с больцмановской H -функцией — по формуле Шеннона (I. 4).

Несогласованность знаков для таких концептуальных определений «энтропии» и «информации» была замечена с самого начала. Так, П. Тэйт, близкий друг Кельвина, в статье, опубликованной в 1868 г., выражал сожаление по поводу того, что Клаузиус выбрал положительный знак для энтропии, которую он фактически считал «отрицательной» величиной (см. [18], стр. 116).

Больцмановская константа, фигурирующая в равенстве (III. 1), не является фундаментальной мировой постоянной. Ее физический смысл вытекает из сложившихся понятий температуры и теплоты. С тем же успехом ее значение могло быть установлено в соответствии с шенноновской концепцией информации и представлено в таком случае безразмерной величиной, например $1/\ln 2$.

С другой стороны, тот факт, что с энтропией в физику ввели «новую концепцию», становится очевидным из аксиоматического обоснования термодинамики, данного Каратеодори и другими [123, 124].

Концепция, с которой мы встречаемся в теории отбора, имеет аналогичную природу, и это дает возможность ответить на вопрос, поставленный Шнейором Лифсоном. Объяснение основных принципов эволюции как самоорганизации на молекулярном уровне не требует «новой физики»; скорее оно опирается на принципы, выводимые из уже известных законов и связывающие макроскопические явления с динамическим поведением на элементарном уровне. Упомянутая концепция выражается с помощью введения понятия ценности — параметра, который следует ассоциировать с понятием информации. Каким образом это делается?

Среди различных состояний ансамбля «информацию» может представлять собой каждое определенное состояние, а не только одно-единственное состояние или определенное подмножество состояний. Введение параметра ценности (который может меняться практически непрерывно), связанного с каждым информационным состоянием, позволяет нам развить *общую теорию, которая рассматривает возникновение или самоорганизацию «ценной» информации*, объединяя тем самым дарвиновскую теорию эволюции с классической теорией информации, а также — после применения этой концепции к самоорганизации на молекулярном уровне — обеспечивая количественную основу для молекулярной биологии.

«Селективная ценность» и «средняя избыточная продуктивность» являются величинами, выводимыми из ранее известных соотношений; они связаны с размерностью времени, даже если привести их к безразмерному виду введением некоторой общей константы скорости (k_0).

В чем же физический смысл этих величин?

Рассмотрим макромолекулярную цепь, представляющую собой последовательность мономерных единиц по крайней мере двух типов. Предполагается, что все возможные последовательности имеют в точности одинаковое энергетическое содержание (в реальном случае это может выполняться **лишь с более или менее хорошим**

приближением). С термодинамической точки зрения все эти состояния неразличимы, или «вырождены». Для всех цепей данного класса с одинаковой длиной образование из мономерных единиц, а также распад на низкоэнергетические фрагменты характеризуются одним и тем же суммарным сродством. Однако если механизмы реакций таковы, что в каждом случае промежуточные соединения различны, то скорости реакций внутри данного класса могут варьировать. Существуют три феноменологических параметра, которые характеризуют «селективную ценность» каждой отдельной последовательности по отношению к ее воспроизведению: скорости образования и распада, физически определяемые по их «свободным энергиям активации», а также фактор «качества», который можно связать с возможностью ветвления реакции в промежуточном состоянии (в котором происходит «инструктирование»). Все другие возможные влияния (обусловленные средой) являются вторичными в том смысле, что они могут воздействовать лишь через посредство этих трех параметров. Определенная комбинация этих трех факторов, зависящая от конкретных дополнительных ограничений, в стационарном состоянии определяет селективную ценность.

§ VIII. 3. «Диссипация» и «возникновение информации»

Принцип Пригожина — Глансдорфа [47] служит связующим звеном между теорией отбора и термодинамикой необратимых процессов.

Стационарное состояние при постоянных потоках характеризуется минимумом производства энтропии. Если построить график (внутренней) энтропии как функции времени, то должна получиться линейная зависимость (рис. 27). Стационарно производимую энтропию можно было бы определить, например, измеряя поток тепла в термостат и в то же время поддерживая реакцию систему при постоянных внутренних условиях. В таком стационарном состоянии доминирует «отобранная» последовательность (или коллектив), которая определяется

из условия

$$\bar{P} = W_{\max}. \quad [\text{ср. с (II. 37)}]$$

Предположим, что произошла (стохастически значимая) флуктуация, состоящая в возникновении мутанта, обладающего большей селективной ценностью, нежели

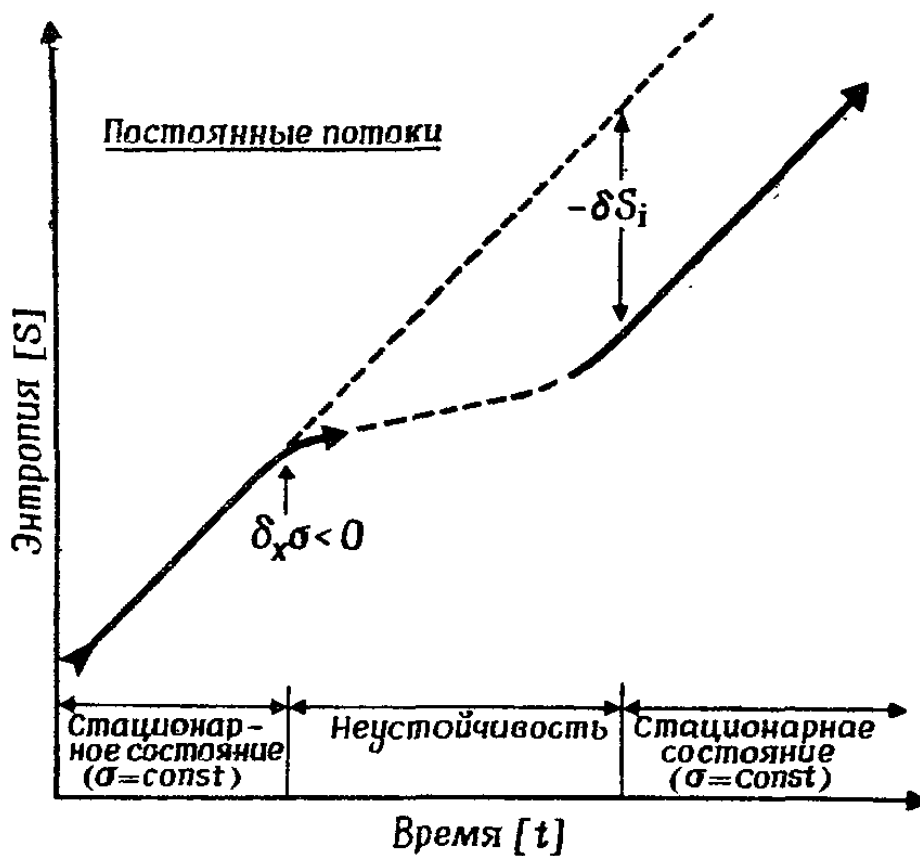


Рис. 27. Зависимость энтропии от времени для процесса отбора при постоянных общих потоках единиц.

Возникновению мутанта с селективным преимуществом (W^F мутанта $>$ W^F главной копии) соответствует отрицательная флуктуация производства энтропии, которая приводит к неустойчивости и распаду стационарного состояния. Преведняя главная копия вымирает, тогда как мутант достигает доминирующего уровня. Поскольку энергии обоих видов равны, равны и наклоны прямых, описывающих производство энтропии σ в обоих стационарных состояниях. Постоянная разность между графиками энтропии в стационарном состоянии отражает увеличение упорядоченности, соответствующее повышению общего уровня организации.

копия, отобранная прежде. Это явление эквивалентно отрицательной флуктуации производства энтропии (т. е. эквивалентно возрастанию средней скорости, связанному с уменьшением суммарного сродства вырожденного класса). По Пригожину и Глансдорфу, такая отрицательная флуктуация должна привести к разрушению существовавшего стационарного состояния, которое не может поддерживаться при неизменных потоках извне. Таким образом, согласно термодинамической теории,

эволюционное поведение при постоянных потоках характеризуется возникновением неустойчивостей. Если после восстановления стационарного состояния сравнить два отобранных вида — старый и новый, то мы увидим, что произошло изменение «оцененной» информации, которое выражается в росте упорядоченности. На рис. 27 отрицательная флуктуация производства энтропии проявляется в отклонении кривой вправо, причем это отклонение вследствие неустойчивости растет до тех пор, пока система не придет к новому стационарному состоянию. Поскольку внешние потоки поддерживаются неизменными, а мутантные копии обладают тем же средством, что и их предшественники, начальный наклон кривой восстанавливается при достижении нового стационарного состояния. Результирующая постоянная разность абсолютных значений энтропии S_i (расстояние между сплошной линией и параллельной ей пунктирной) в точности равна той разнице энтропий, которая обусловлена возросшей внутренней упорядоченностью, выражаемой «степенью организованности» (т. е. долей всех тех единиц в системе, которые организованы в полимерные последовательности). Никакого нарушения второго начала термодинамики здесь не происходит. Тем не менее термодинамическое описание не отражает того, что почти вся однородная популяция (или некоторые из ее важных составных частей) должна полностью смениться другой, чтобы произвести новую (более «ценную») информацию и тем самым уменьшить внутреннюю энтропию.

Подобный баланс становится более запутанным, если мы допустим, что различные последовательности имеют различные свободные энергии и, следовательно, обладают различным средством в отношении процессов образования и распада. Если в результате этого изменятся селективные ценности, то это отразится в эволюционном поведении. Это проявится и в балансе энтропии при неизменных потоках, поскольку суммарное и индивидуальное средство уже не будет совпадать. Однако решающим параметром для отбора остается все же «селективная ценность», если селекционные ограничения для конкурирующих последовательностей задаются постоянством

суммарной концентрации полимеров и мономеров, а не «постоянными средними значениями сродства». При таких «буферных» условиях не существует дефицита свободной энергии и, следовательно, с точки зрения отбора никакая «экономичность» не имеет значения. Уже отсюда ясно, что эта концепция отбора выходит за рамки термодинамики необратимых процессов. Здесь интерес представляет индивидуальное *информационное* содержание и его «ценность» по отношению к процессу воспроизведения, а не баланс энтропии вообще. Описанный выше пример вырожденных последовательностей был избран для иллюстрации того факта, что «селективная ценность» как движущая сила эволюции представляет собой новую переменную, которая связана с термодинамикой необратимых процессов, но не укладывается в ее современные рамки. Что касается использования свободной энергии, особенно в более сложных воспроизводящихся системах, то об этом следует сделать еще одно замечание.

Нередко приходится читать, что основной принцип эволюции состоит в экономном использовании свободной энергии. Но это не всегда правильно. Там, где свободная энергия поставляется в избытке, система будет использовать любое селективное преимущество, сколько бы оно ни «стоило». Однако если расход энергии становится одним из вторичных ограничений, поскольку он влияет на селективную ценность, система будет реагировать соответствующим образом. Решающим фактором эволюции является использование *информации*, ассоциированной с высокой «селективной ценностью», а не экономия расходования свободной энергии.

«Информация» в данном контексте (т. е. в применении к самоорганизующимся материальным системам) — нечто большее, нежели просто «структурный коррелят функции». Она описывает некоторые детали, утраченные при статистическом усреднении. Она относится к отдельным, феноменологически различимым состояниям ансамбля, в котором общее число возможных состояний может значительно превышать число состояний, которые заселены или могут быть заселены. Такая информация имеет значение только в том случае, если она способна

сохраняться на протяжении динамических процессов образования и распада. Она может подразумевать существование предшествующей информации, из которой она произошла, но если проследить ее истоки, то мы будем вынуждены сказать, что «первичная» информация представляет собой функцию, которая обеспечивает свое собственное воспроизведение и изменение к состоянию большей «безопасности», которому соответствует большая избыточность. В таком случае она может также включать в себя вторичную информацию, не нужную для воспроизведения, в качестве инструктивного коррелята некоторой функции, ограниченной во времени.

«Селективная ценность» характеризует исполнительные качества носителей информации, определяющие их шансы на выживание и сохранение наиболее устойчивого (или наиболее приспособленного) состояния.

§ VIII.4. Принципы отбора и эволюции

Чем больше информационное содержание, тем более оправдано подразделение эволюции на два процесса:

- 1) отбор среди заселенных альтернативных состояний;
- 2) эволюция отобранных состояний.

Оба эти процесса сливаются в один, если структурная емкость столь низка, что заселены *все возможные* альтернативные состояния. Однако число возможных вариантов обычно чрезвычайно велико по сравнению с числом состояний, которые могут быть заселены, и лишь при этих условиях понятие информации оказывается полезным. Все информационное содержание генома человека, например, хранится примерно в 10^{10} нуклеотидах, но эволюция при этом сделала выбор одного или нескольких вариантов из $4^{10^{10}}$ возможных (среди которых есть, конечно, и вырожденные варианты).

Между наиболее благоприятными условиями для отбора и для эволюции может существовать некоторое различие, так что для оптимизации всего процесса требуется некоторый компромисс, подобно тому как сама селективность представляет собой некое равновесие между точностью узнавания (следствием высокой точности

узнавания является «липкость») и динамической гибкостью.

Отбор при постоянных селекционных ограничениях является процессом, при котором средняя продуктивность оптимизируется путем достижения высшей селективной ценности, имеющейся в популяции носителей информации при данных условиях среды.

Такой процесс характеризуется следующим оптимизационным принципом:

$$\bar{P} \text{ (либо } \bar{E}) \rightarrow W_m,$$

который для нелинейных систем может быть заменен более общим интегральным соотношением (II.34). Здесь W_m означает *относительный* максимум в популяции конкурентов. \bar{E} приближается к указанному значению, удовлетворяя при этом определенным ограничениям (процедура оптимизации).

Процесс отбора обычно связан с экономизацией количества информации. В фиксированных условиях среды те последовательности, которые могут обойтись наименьшим числом единиц для хранения всей необходимой информации (т. е. те, которые не несут никакой бесполезной информации), будут обычно проявлять наивысшую селективную ценность, при которой скорость образования A_i , время жизни $1/D_i$ и параметр точности Q_i достигают наибольшего возможного значения. Для сохранения информации отобранного состояния требуется некая минимальная пороговая точность (Q_{\min}), определяемая соотношением (II.45).

При оптимальном отборе требуемая точность переноса информации должна быть приведена в соответствие с количеством передаваемой информации.

С другой стороны, эволюция происходит быстрее всего при наименьшем возможном значении Q , но, поскольку она основана на отборе, значение Q должно превышать порог Q_{\min} . Изменчивость уменьшается с ростом Q , и это будет всегда благоприятствовать эволюции тех систем, для которых Q близка к Q_{\min} .

Эволюция представляет собой процедуру дальнейшей оптимизации при наличии определенных ограничений, накладываемых критериями отбора.

Например, в «линейной» системе с постоянными параметрами «ценности» при экстремальных ограничениях постоянной общей организации эволюционный путь ограничивается монотонным ростом W_m для всех последовательно отбираемых носителей информации:

$$W_{m1} < W_{m2} < W_{m3} < \dots < W_{opt}.$$

Это исключает все пути эволюции, проходящие через минимум W_m , среди которых могут находиться пути, приводящие к более высоким конечным значениям W_m . Ограничения, каковы бы они ни были, уменьшают число возможных выборов и тем самым, скорее всего, препятствуют наилучшему выбору. Можно сделать заключение, что для систем с большой и ненасыщенной информационной емкостью в общем случае справедливо неравенство

$$W_{opt} < W_{max},$$

характеризующее конечную разность, которую Жак Моно [125] назвал бы различием между тем, что «есть», и тем, что «должно быть». W_{max} представляет собой абсолютный максимум, достижение которого потребовало бы изначального присутствия или доступности «всей» информации. Если построить «информационное» пространство, координаты которого представляют все возможные «информации», то поведение в процессе эволюции точки, изображающей систему, будет описываться фундаментально неэргодической траекторией, отражающей неизбежность эволюционного процесса. Именно эта особенность определяет преимущественное направление времени для всех живых систем. Этот эффект связан с однонаправленным возрастанием энтропии во всех необратимых процессах, но выражен еще более явно.

Для нелинейных систем абсолютное значение W_m , вообще говоря, не обязательно возрастает в процессе эволюции. Дело в том, что наличие внутренних связей между носителями информации эквивалентно изменению окружающей среды. Более того, вторичные изменения (например, загрязнение среды) могут вызвать всеобщее снижение всех селективных ценностей. Тогда эволюция

все еще характеризуется последовательностью значений W_m , каждое из которых характеризует вид или коллектив с оптимальными параметрами; но такая последовательность, однако, не обязательно должна быть монотонно возрастающей. Она даже может быть убывающей в связи с изменениями окружающей среды, но тогда, как правило, система будет вынуждена эволюционировать в сторону большего информационного содержания. Для того чтобы пережить изменения в окружающей среде или изменить окружающую среду на свою пользу, требуется дополнительная информация. Таким образом: *эволюция может быть связана как с возрастанием селективной ценности, так и с использованием большего количества информации.*

Влияние второго эффекта будет превалировать по крайней мере на высших уровнях молекулярной организации, поскольку в сложной системе возможные связи становятся столь многочисленными, что большое число мутаций может привести к общему снижению селективной ценности; это может компенсироваться только появлением таких мутантов, которые способны противостоять изменениям окружающей среды. Термины «хорошо» или «плохо» приобретают смысл, как только отдельные носители информации начинают взаимодействовать друг с другом и тем самым увеличивают или уменьшают свою «ценность».

Эволюцию на молекулярном уровне можно считать некой игрой, в которой разум игрока заменен селективным «инстинктом», призванным содействовать выживанию среди хаотически проявляющихся воздействий внешнего мира. Поэтому мы считаем, что теория игр, созданная Джоном фон Нейманом [126] и ныне ставшая весьма развитой отраслью науки [127], является ключом к любому дальнейшему обобщению теории эволюции¹.

¹ Специальные математические методы для решения оптимизационных задач, возникающих в молекулярной биологии, были предложены И. Рехенбергом [110]. Рассмотрение общих явлений эволюции с точки зрения теории игр было предложено Р. Левонтином [128], а также Ф. Ворбертоном [129]. М. Кимура [130] и Д. Мак-Кей [131] подчеркивали тот факт, что «информация» может возникнуть из «шума».

§ VIII. 5. «Недетерминированная», но «неизбежная»

Тот факт, что «отбор» и «эволюцию» — по аналогии¹ с «равновесием» в термодинамике — можно охарактеризовать экстремальными принципами, позволяет физически обосновать и количественно сформулировать дарвиновскую теорию. В такой форме теория Дарвина уже не просто описывает некий исторический путь, но представляет собой выводимый из физики закон, управляющий общими процессами самоорганизации материи. Исследование эволюции высших форм жизни, особенно «разумных» форм регулирования, потребует, однако, введения дополнительных принципов. Таким образом, мы сознательно ограничиваемся здесь областью молекулярной биологии, т. е. процессами самоорганизации на молекулярном уровне.

Тот факт, что мы имеем физическую характеристику концепции «ценности», может модифицировать нашу интерпретацию Дарвина. Среди современных биологов широко распространена точка зрения, состоящая в том, что дарвиновский принцип (выражаясь словами Уоддингтона [132]) — это всего лишь «либо трюизм, либо тавтология». Гюнтер Стент [133] в своей книге «Наступление золотого века» писал: «Как известно сейчас каждому, выживание наиболее приспособленных — это не более чем тавтология: «выживание выживших», а поэтому определение «неприспособленный» — это не объективно научное, а лишь субъективное суждение о ценности».

Такие утверждения были бы правильны только в том случае, если термин «ценность» — или как бы мы это ни назвали — применялся бы просто для обозначения результата события, в остальном полностью недетерминированного. Поэтому важно провести вероятностный анализ концепции «ценности», которая была введена сначала в (детерминистическую) феноменологическую теорию. Такой анализ необходимо провести отдельно для «отбора» (среди данного множества носителей информации) и для эволюции (которая использует процедуру

¹ Помимо аналогии, отметим и различия: а) различие между «оптимизацией» и «максимизацией»; б) отсутствие микроскопической обратимости при стационарном состоянии; в) ограниченную применимость полных дифференциалов,

отбора для приближения к оптимальной ценности). Основные идеи вероятностного анализа разобраны в гл. III, и один из полученных там результатов имеет большое значение для затронутого здесь вопроса. Речь идет о точном воспроизведении ($Q = 1$) множества n различных последовательностей, которые все вырождены по W_i . Вероятностный анализ показывает, что такая система — благодаря своему автокаталитическому поведению — будет всегда (либо с большой вероятностью) сужать свое информационное содержание до одной последовательности, которая, однако, будет представлена растущим числом экземпляров вплоть до достижения избыточности в n копий. Это истинный пример «выживания выживших», ибо не существует способа предсказать, *какая именно* матрица выживет — физически все они неразличимы; просто происходит флуктуация, которая усиливается вследствие присущего системе механизма воспроизведения. Этот пример представляет собой, однако, нереалистичный особый случай, поскольку Q никогда не может быть точно равно единице. Если сохраняется условие полной вырожденности и общее равновесие между образованием и распадом, то селективная ценность каждой отдельной последовательности должна стать отрицательной (когда $\mathcal{F}_i = \mathcal{D}_i$, но $Q_i < 1$), и, следовательно, из множества полностью вырожденных носителей информации не может произойти *никакого* отбора стабильной информации. *Устойчивый* отбор вида с максимальным значением W_i произойдет в стационарном состоянии только в том случае, если селективные ценности различных носителей информации будут распределены на каком-то интервале значений.

Этот процесс опять-таки не является полностью детерминистическим, хотя величина флуктуаций уменьшается с возрастанием числа отобранных копий — как и для любого саморегулирующегося случайного процесса. Поскольку отбор обычно начинается с небольшого числа мутантных копий или даже с одного-единственного мутанта, флуктуации имеют огромное значение.

С проблемой полной вырожденности, обсуждавшейся выше, тесно связано явление «случайного дрейфа». Это явление имеет место, когда несколько видов, возникших

вследствие «нейтральных» мутаций [134], оказываются вырожденными в отношении своей селективной ценности. В литературе его часто называют «недарвиновской» эволюцией [135]. Этим эффектом пренебрегали при ранних оценках темпов эволюции, основанных на анализе аминокислотных последовательностей белков организмов, находящихся на различных филогенетических уровнях. Эффект этот несомненно следует учитывать при таких оценках, однако вряд ли удачно называть его «недарвиновским». «Нейтральные» мутанты и их «случайный дрейф» вполне вписываются в рамки более абстрактной концепции отбора, изложенной в этой статье.

Более жесткие ограничения на детерминированность накладывает эффект «копирования ошибок» в процессе воспроизведения или другие типы мутаций, лежащие в основе оптимизационной процедуры в эволюции. Хотя единицы, занимающие различные положения, нельзя считать вполне эквивалентными в отношении мутаций и возникающие мутанты все еще сходны с главной копией, между «причиной» и «эффектом» в мутационном процессе уже нет никакой связи или эта связь незначительна, и весь процесс представляется случайным. Кстати, элементарные физические процессы, приводящие к мутациям, сами существенно недетерминированы в силу их квантовомеханической природы [59]. Процессы автокаталитического отбора отфильтровывают и усиливают мутантов, обладающих высокой селективной ценностью, снижая тем самым недетерминированность — в той мере, в какой здесь затрагивается принцип ценности. Однако недетерминированность все-таки сохраняется в отношении выбора индивидуальных копий и затем отображается на макроскопическом уровне. Вследствие этого невозможно с достаточной точностью ни проследить прошлую эволюцию, ни предсказать будущую эволюцию за пределами известных временных границ. Более того, именно недетерминированность делает невозможным достижение абсолютного максимума селективной ценности.

На высших уровнях эволюции, особенно на уровне генетики популяций, наиболее обещающим подходом к проблеме отбора был такой, при котором за отправной пункт принимали сам факт выживания. Это направление

связано с именами Р. Фишера [28], Дж. Холдейна [29] и С. Райта [30] (термин «селективная ценность» заимствован из их работ). Именно успех этого направления привел некоторых биологов к тавтологической реинтерпретации дарвиновского принципа отбора¹.

Тем не менее, если мы можем связать «выживание» с физически объективной «ценностью» (которая представляет собой весьма специфическое сочетание параметров скорости и взаимодействия), то принцип отбора уже перестает быть *тривиальной* тавтологией или трюизмом. Любой принцип, как только становится ясным его логическое содержание, может показаться более или менее очевидным, потому что логика представляет собой раскрытие тавтологий или соответствий.

Мы можем прийти теперь к заключению, что эволюцию жизни, если ее основывать на выводимых из физики принципах, следует рассматривать как *неизбежный* процесс, несмотря на ее недетерминированный ход (см. ниже). Модели, рассмотренные в гл. IV—VI, и эксперименты, обсуждавшиеся в гл. IV, VI и VII, показывают, что эволюция не только неизбежна «в принципе», но и достаточно вероятна в пределах реалистичных временных интервалов. Она требует соответствующих условий окружающей среды (которые выполнены отнюдь не повсюду) и их поддержания. Такие условия существовали на Земле и должны существовать на многих других планетах во Вселенной. Временных ограничений на дальнейшее продолжение эволюционного процесса не существует — необходимо лишь снабжение энергией. Таким образом, любые предсказания, связанные с временными ограничениями, якобы «присущими» эволюции жизни, в итоге зависят от наших знаний о доступных космических источниках энергии и таким образом должны быть тесно связаны с проблемами космологии (для которых пока не найдено определенных решений).

¹ Заслуживают упоминания работы Вольтерра [53] и Лотки [58], относящиеся к формальной математической трактовке частных задач о росте в условиях «конкуренции» и «борьбы», но работы эти меньше связаны с общей проблемой генерирования «информации» в макромолекулах.

Абстрактная формулировка принципов отбора и эволюции, изложенная в данной работе, конечно, не включает в себя предположения, что эволюция действительно происходила в условиях крайних и абстрактных селекционных ограничений, отвечающих стационарным состояниям. Мы уже подчеркивали аналогию с равновесной термодинамикой. В свое время абстрактная трактовка цикла Карно, которым определяется максимально возможный коэффициент полезного действия паровой машины, привела к пониманию принципов (равновесной) термодинамики, хотя ни один тепловой двигатель никогда не работал и не мог бы работать при равновесных условиях. Эволюционный прогресс был бы крайне замедлен (можно было бы указать и на социологические аналогии), если бы всегда имели место экстремальные ограничения.

Однако эти условия вскрывают сами принципы, дают нам возможность анализировать модели, указывают путь к постановке воспроизводимых экспериментов и в итоге могут привести нас к реконструкции определенных эволюционных событий. Итак, мы теперь готовы к обсуждению решающего вопроса.

§ VIII.6. Можно ли объяснить жизнь на основе современных физических концепций?

Простое «да» может поставить нас в трудное положение, ибо нас могут попросить доказать справедливость нашего утверждения, например, с помощью полной индукции. Поэтому будет, пожалуй, разумнее «вывернуть вопрос наизнанку»: если кто-нибудь заявит, что физика не может дать объяснения жизни, предложить ему доказать свое утверждение или, еще лучше, опровергнуть его, приведя контрпример.

Найдется ли такой контрпример? Это очень сильно зависит от того, придем ли мы к общему согласию относительно определения жизни.

А. И. Опарин [136] однажды предложил следующий перечень «свойств живого»: метаболизм, способность к самовоспроизведению, мутабельность.

Этими свойствами может обладать машина типа «самовоспроизводящегося автомата» Дж. фон Неймана

[137]. Такая машина определенно имела бы «метаболизм». В окружении электрических розеток или баков с горючим эта машина нашла бы достаточно «пищи» и выполняла бы работу любого типа. Особенностью этого автомата является, конечно, не только способность к самовоспроизведению в соответствии с определенной программой, но также и способность к воспроизведению самой этой программы, т. е. любая машина-потомок способна воспроизводить ошибки, которые могли бы привести к возникновению «генотипических» мутантов, обладающих селективным преимуществом. Но можем ли мы назвать такой робот «живым»? Вероятно нет, ибо прародитель этого робота должен быть создан человеком; и поэтому его придется назвать «искусственным».

Другими примерами являются воспроизводящиеся макромолекулярные циклы, рассмотренные в гл. IV—VI. Они могут зародиться сами, но мы не можем приписывать свойство «быть живым» чему-то менее сложному, чем каталитические гиперциклы, описанные в гл. VI, которым присущи примерно десять свойств (см. § VI. 2), в том числе три из перечня Опарина.

Существование такого цикла зависит только от: а) некоторых химических свойств материи (определенных в гл. IV—VI), которые, по крайней мере в принципе, можно объяснить квантовомеханической теорией, а также б) наличия определенных физических условий, которые, по всей вероятности, существовали на Земле.

Мы должны сделать вывод, что для обоснования биологии не требуется никакой «новой физики», но вместе с тем мы видим, как мало дает нам такое заключение. Переход от одиночной макромолекулы к каталитическому гиперциклу или «живой» клетке определенно менее драматичен, чем переход от одиночной клетки к сознательному и разумному человеческому существу. Для понимания различных этапов этого перехода потребуются, вероятно, так же мало «новой физики» и так же много новых (выводимых из физики) «концепций», как потребовалось для первого шага.

Как сказал полвека назад Витгенштейн [138]: «Решение жизненных проблем люди замечают по исчезновению этих проблем»,

РЕЗЮМЕ

Настоящая работа представляет собой одновременно и обзор, и оригинальное исследование. Она предназначена для физиков и биологов. Физиков необходимо ознакомить с некоторыми фактами, хорошо известными биологам (или биохимикам), чтобы показать, на каких предпосылках следует строить теорию явлений жизни и почему следует обсуждать именно данные, а не какие-то иные модели. С другой стороны, биологам необходимо показать, что жизненные процессы подчиняются физическим законам, которые поддаются количественной формулировке. Поэтому здесь подробно обсуждаются многие привычные для физика законы и решения, часто в приближенной форме, выбранной для большей «прозрачности».

В центре внимания находится вопрос, который в явном виде сформулирован в заключительной главе: *«Можно ли объяснить жизнь на основе современных физических концепций?»*

Ответ в той мере, в какой его вообще можно выразить в одном тезисе, таков: те биологические процессы и явления, которые достаточно хорошо исследованы в настоящее время, не дают никакого повода считать, что физика в ее современном виде не в состоянии объяснить их, хотя — как и в случае макроскопических явлений в неживом мире — детальность описания имеет пределы, обусловленные не фундаментальными законами, а только сложностью самих явлений. Это отнюдь не означает, что в биологических явлениях знакомые нам основные принципы физики не могут выступать в особой форме, присущей только этим явлениям. В качестве примера можно прежде всего назвать выводимое из физики поня-

тие ценности, характерное для теории возникновения информации и доминирующее в оптимизирующем эволюционном процессе, или же характерное для эволюционного процесса преимущественное направление во времени, которое обусловлено критериями устойчивости термодинамической теории необратимых процессов и превращает эволюцию в существенно «несводимое» событие.

В гл. II дана феноменологическая формулировка эволюционной теории. Теория Дарвина оказывается оптимальным принципом, вытекающим из определенных физических предпосылок, а вовсе не «несводимым» феноменом, который относится только к биосфере. Критерий устойчивости Пригожина — Гландорфа связывает эту теорию с термодинамической теорией стационарных состояний. Такие понятия, как «ограничения, накладываемые отбором» и «селективная ценность», можно сформулировать количественно и физически объективно, если принять определенные динамические условия (например, постоянные «потоки» или «силы»). Фигурирующее здесь понятие «ценности» дает основу для такой теории информации, которая включает в себя описание создания информации. «Информация» здесь является «молекулярным» свойством, которое объясняется динамической теорией материи; она «оценивается» по своей способности к самовоспроизведению. Информация дает возможность отличать друг от друга энергетически вырожденные состояния, характеризуя тем самым «первичную» самоорганизацию материи с чисто функциональной точки зрения.

В гл. III исследуются пределы применимости феноменологической теории, которые обусловлены недетерминированностью событий. Теория случайных процессов подтверждает выводы феноменологической теории относительно поведения средних величин в случае большого числа частиц, но она уводит далеко за пределы феноменологической теории. Поскольку эволюция всегда происходит из элементарных событий, которые «усиливаются» процессом роста и тем самым «отображаются» макроскопически, то по сравнению с замкнутыми равновесными системами в эволюционирующей системе

существует значительно более сильная неопределенность как в отношении индивидуальных структур, так и в отношении «исторического» хода событий. Теория случайных процессов позволяет также ликвидировать один предрассудок, распространенный среди биологов и опирающийся именно на эту «неопределенность» элементарных процессов: будто дарвиновский принцип «выживание наиболее приспособленных» представляет собой лишь тривиальную тавтологию — «выживание выживших». Подобная интерпретация была бы оправдана лишь в том случае, если «наибольшая приспособленность» — как результат случая — определялась бы только самим фактом «выживания». Это особый случай, который едва ли реализуется в естественных условиях. Для его реализации необходимо, во-первых, абсолютно точное, т. е. абсолютно безошибочное, воспроизведение, и, во-вторых, полное вырождение всех селективных ценностей, т. е. неразличимость динамических свойств конкурирующих видов. При этих условиях случайная флуктуация могла бы вследствие автокаталитической природы воспроизведения усилиться настолько, что это привело бы к сохранению только ее носителей и результат отбора был бы непредсказуем. Однако, поскольку энергии взаимодействия имеют некую конечную величину, процессы воспроизведения отягчены известной частотой возникновения ошибок. Кроме того, селективные ценности различных воспроизводящихся видов в общем различны. Но в таком случае отбор всегда будет реализоваться как некоторый процесс оптимизации, причем понятие «наиболее приспособленный» соответствует максимальной ценности при дополнительных ограничивающих условиях, имеющих форму неравенств, т. е. оптимуму. Флуктуации играют большую роль, потому что мутанты, обладающие каким-либо преимуществом и поэтому сохраняемые отбором, появляются сначала в одном-единственном экземпляре. Такие флуктуации имеют тем большее значение, чем меньше «селективные преимущества», благоприятствующие их отбору.

В гл. IV—VI рассматриваются конкретные модели реакционных систем. Можно показать, что *возникновение «жизнеспособных» структур* (возможных предше-

ственников одноклеточных микроорганизмов) связано с особыми условиями механизма отбора, которые не могут обеспечить по отдельности ни нуклеиновые кислоты, ни белки. Процесс воспроизведения, лежащий в основе механизма отбора, должен уже сам по себе иметь нелинейную природу. Как нуклеиновые кислоты, так и белки могут воспроизводиться в соответствии с «квазилинейным» реакционным механизмом. Но при этом возникают системы, которые в случае нуклеиновых кислот содержат слишком мало, а в случае белков — слишком много информации. «Слишком мало» означает, что различные конкурирующие друг с другом последовательности не в состоянии набрать и удерживать достаточное количество информации для кодирования функций, которым благоприятствует отбор. «Слишком много» информации означает, напротив, что вероятность благоприятной мутации слишком мала или что система не может освободиться от груза «паразитных» связей. Напротив, система, которая содержит как нуклеиновые кислоты, так и белки, может в процессе отбора использовать функциональные преимущества тех и других. Свойства, определяющие преимущества, таковы:

а) присущая нуклеиновым кислотам способность к самоинструированию, которая делает возможным воспроизведение не только каждого существующего информационного состояния, но и любого дальнейшего изменения этого состояния;

б) гигантская функциональная емкость белков (специфичное узнавание, катализ и регулирование), которая необходима для установления связи и корреляции между отдельными реакциями при построении упорядоченных функциональных единиц.

Такая (нелинейная) связь между нуклеиновыми кислотами и белками приводит к образованию иерархии реакционных циклов, которая обладает существенными свойствами «живой» системы и «открыта» для дальнейшей эволюции вплоть до живой клетки. Возникновение такого самовоспроизводящегося «гиперцикла» зависит от образования однозначной системы кодирования. Во второй половине гл. VI обсуждаются предпосылки возникновения кода с однозначным соответствием.

Наконец, в гл. VII описаны эволюционные эксперименты с РНК фага Q β , впервые выполненные С. Спигелманом. Теория, развитая в гл. II и IV, дает основу для проведения воспроизводимых измерений и для получения из них количественных выводов.

Результаты настоящей работы можно подытожить следующим образом:

1. Детальный анализ механизмов воспроизведения нуклеиновых кислот и белков не дает никаких оснований для гипотезы о существовании каких-то сил или взаимодействий, присущих только явлениям жизни.

2. Каждая отдельная система, возникающая в результате мутаций и отбора, непредсказуема в отношении своей структуры; тем не менее неизбежным результатом всегда является процесс эволюции — это закон. Появление мутации с селективным преимуществом соответствует неустойчивости, которую можно объяснить при помощи принципа Пригожина — Глансдорфа, выведенного для стационарных необратимых термодинамических процессов. Поэтому оптимизирующий процесс эволюции в принципе неизбежен, хотя выбор конкретного пути не детерминирован.

3. Наконец, оказывается, что процесс возникновения жизни связан с появлением ряда свойств, причем все эти свойства поддаются однозначному физическому обоснованию. Предварительные условия для появления этих свойств, по-видимому, выполнялись шаг за шагом, так что «возникновение жизни», как и эволюцию видов, нельзя представить в виде однократного акта творения.

Благодарности. Многие идеи, высказанные в этой работе, нельзя, конечно, считать новыми. Однако работа предназначена одновременно для физиков и для биологов, и то, что может показаться тривиальным одному, не покажется тривиальным другому. У меня есть еще одно оправдание для такой длинной работы, а именно то, что целое всегда представляет собой нечто большее, чем просто сумма отдельных идей.

На высказанные здесь соображения по теории отбора сильно повлияла беседа за завтраком с Ф. Криком. Большим стимулом послужили также эволюционные экс-

перименты С. Спигелмана и фундаментальные идеи И. Пригожина по «необратимой термодинамике». П. Шустер (ныне работающий в Венском университете) провел (и ведет) большую вычислительную работу на ЭВМ по самоорганизующимся каталитическим циклам, подробное изложение которой будет опубликовано отдельно [100]. Нас очень забавляло держать пари с ЭВМ об исходе эволюционной конкуренции — ЭВМ всегда выигрывала. Р. Торнли исправил мой английский; кроме того, был еще один человек, который имел терпение все выслушивать.

Спасибо им всем!

СПИСОК ЛИТЕРАТУРЫ

1. *Watson J. D., Crick F. H. C.*, Nature, 177, 964 (1953).
2. *Kornberg A.*, Science, 163, 1410 (1969).
3. *De Lucia P., Cairns J.*, Nature, 224, 1164 (1969).
4. Transcription of Genetic Material, Cold Spring Harbor Symp. Quant. Biol., 35 (1970).
5. *Holley R. W., Apgar J., Everett G. A., Madison J. T., Marquissee M., Merrill S. H., Penswick J. R., Zamir A.*, Science, 147, 1462 (1965). (Р. Холли, Д. Ангар, Г. Эверетт, Д. Медисон, М. Маркуиси, С. Мэррил, Дж. Пенсуик, А. Замир, в сб. «Синтез и структура нуклеиновых кислот», изд-во «Мир», М., 1966, стр. 94.)
6. *Zachau H. G., Dütting D., Feldmann H.*, Angew. Chem., 78, 392 (1966).
7. *Rajbhandari V. L., Stuart A., Faulkner R. D., Chang S. H., Khorana H. G.*, Cold Spring Harbor Symp. Quant. Biol., 31, 425 (1966).
8. *Yarus M., Berg P.*, J. Mol. Biol., 42, 171 (1969).
9. *Nomura M.*, in: Scientific American, December 1969 and The Neurosciences, 2nd Study Program (ed. by F. O. Schmitt), New York, The Rockefeller University Press, 1970, p. 913.
10. *Garret R. A., Wittmann H. G.*, in: Aspects of Protein Biosynthesis (ed. by T. B. Anfinsen), New York, Academic Press, 1971.
11. *Jacob F., Monod J.*, J. Mol. Biol., 3, 318 (1961).
12. *Gilbert K., Müller-Hill B.*, Proc. Nat. Acad. Sci. USA, 56, 1891 (1966), 58, 2415 (1967).
13. *Riggs A. D., Bourgeois S., Cohn M.*, J. Mol. Biol., 53, 401 (1970).
14. *Watson J. D.*, The Molecular Biology of the Gene, New York, W. A. Benjamin, Inc. 1970. (Имеется перевод 1-го издания: Дж. Уотсон, Молекулярная биология гена, изд-во «Мир», М., 1967.)
15. Molekularbiologie (ed. by Th. Wieland and G. Pfeleiderer), Frankfurt: Umschau-Verlag, 1969.
16. *Wigner E.*, in: The Logic of Personal Knowledge, Essays presented to Michael Polanyi on his seventieth birthday 11th March 1961 (ed. by E. Shils), Glencoe Ill.: Free Press 1961. (Имеется перевод: Е. Вигнер в сб. «Этюды о симметрии», изд-во «Мир», М., 1971, стр. 160.)
17. *Wiener N.*, Cybernetics (German Translation: Kybernetik). Düsseldorf-Wien, Econ Verlag, 1963. (Н. Винер, Кибернетика, 2-е изд., изд-во «Советское радио», М., 1968.)
18. *Brillouin L.*, Science and Information Theory, New York, Academic Press, 1963. (Л. Бриллюэн, Наука и теория информации, Физматгиз, М., 1960.)

перименты С. Спигелмана и фундаментальные идеи И. Пригожина по «необратимой термодинамике». П. Шустер (ныне работающий в Венском университете) провел (и ведет) большую вычислительную работу на ЭВМ по самоорганизующимся каталитическим циклам, подробное изложение которой будет опубликовано отдельно [100]. Нас очень забавляло держать пари с ЭВМ об исходе эволюционной конкуренции — ЭВМ всегда выигрывала. Р. Торнли исправил мой английский; кроме того, был еще один человек, который имел терпение все выслушивать.

Спасибо им всем!

СПИСОК ЛИТЕРАТУРЫ

1. *Watson J. D., Crick F. H. C.*, Nature, 177, 964 (1953).
2. *Kornberg A.*, Science, 163, 1410 (1969).
3. *De Lucia P., Cairns J.*, Nature, 224, 1164 (1969).
4. Transcription of Genetic Material, Cold Spring Harbor Symp. Quant. Biol., 35 (1970).
5. *Holley R. W., Apgar J., Everett G. A., Madison J. T., Marquissee M., Merrill S. H., Penswick J. R., Zamir A.*, Science, 147, 1462 (1965). (Р. Холли, Д. Апгар, Г. Эверетт, Д. Медисон, М. Маркуиси, С. Мэррил, Дж. Пенсуик, А. Замир, в сб. «Синтез и структура нуклеиновых кислот», изд-во «Мир», М., 1966, стр. 94.)
6. *Zachau H. G., Dütting D., Feldmann H.*, Angew. Chem., 78, 392 (1966).
7. *Rajbhandari V. L., Stuart A., Faulkner R. D., Chang S. H., Khorana H. G.*, Cold Spring Harbor Symp. Quant. Biol., 31, 425 (1966).
8. *Yarus M., Berg P.*, J. Mol. Biol., 42, 171 (1969).
9. *Nomura M.*, in: Scientific American, December 1969 and The Neurosciences, 2nd Study Program (ed. by F. O. Schmitt), New York, The Rockefeller University Press, 1970, p. 913.
10. *Garret R. A., Wittmann H. G.*, in: Aspects of Protein Biosynthesis (ed. by T. B. Anfinsen), New York, Academic Press, 1971.
11. *Jacob F., Monod J.*, J. Mol. Biol., 3, 318 (1961).
12. *Gilbert K., Müller-Hill B.*, Proc. Nat. Acad. Sci. USA, 56, 1891 (1966), 58, 2415 (1967).
13. *Riggs A. D., Bourgeois S., Cohn M.*, J. Mol. Biol., 53, 401 (1970).
14. *Watson J. D.*, The Molecular Biology of the Gene, New York, W. A. Benjamin, Inc. 1970. (Имеется перевод 1-го издания: Дж. Уотсон, Молекулярная биология гена, изд-во «Мир», М., 1967.)
15. Molekularbiologie (ed. by Th. Wieland and G. Pfeleiderer), Frankfurt: Umschau-Verlag, 1969.
16. *Wigner E.*, in: The Logic of Personal Knowledge, Essays presented to Michael Polanyi on his seventieth birthday 11th March 1961 (ed. by E. Shils), Glencoe Ill.: Free Press 1961. (Имеется перевод: Е. Вигнер в сб. «Этюды о симметрии», изд-во «Мир», М., 1971, стр. 160.)
17. *Wiener N.*, Cybernetics (German Translation: Kybernetik). Düsseldorf-Wien, Econ Verlag, 1963. (Н. Винер, Кибернетика, 2-е изд., изд-во «Советское радио», М., 1968.)
18. *Brillouin L.*, Science and Information Theory, New York, Academic Press, 1963. (Л. Бриллюэн, Наука и теория информации, Физматгиз, М., 1960.)

19. *Nishimura S., Jones S. D., Khorana H. G.*, J. Mol. Biol. **13**, 302 (1965).
20. *Nirenberg M. W., Matthaei J. H.*, Proc. Nat. Acad. Sci. USA, **47**, 1588 (1961). (*М. Ниренберг и Г. Маттеи*, в сб. «Молекулярная генетика», ИЛ, М., 1963, стр. 69.)
21. *Ochoa S.*, in: Horizons in Biochemistry (ed. by M. Kasha and B. Pullman), New York, Academic Press 1962 (*С. Очоа*, в сб. «Горизонты биохимии», изд-во «Мир», М., 1964, стр. 120.)
22. *Crick F. H. C.*, J. Mol. Biol., **19**, 548 (1966).
23. *Shannon C. E.*, Bell System Tech. J., **30**, 50 (1951).
24. *Shannon C. E., Weaver W.*, The Mathematical Theory of Communication, Urbana Univ. of Illinois Press, 1949. (*К. Э. Шеннон*, в сб. «Работы по теории информации и кибернетике», стр. 243, ИЛ, М., 1963.)
25. *Dutta M.*, Physics Today, **21**, 75 (1968).
26. *Schroedinger E.*, What is Life? Cambridge University Press, 1944. (*Э. Шредингер*, Что такое жизнь?, 2-е изд., Атомиздат, М., 1972.)
27. *Darwin C.*, The Origin of Species, chapter 4, London 1872, Reprinted by Collier MacMillan Ltd., London, 1959, p. 91. (*Ч. Дарвин*, Сочинения, т. 3, Изд-во АН СССР, М. — Л., 1938.)
28. *Fisher R. A.*, Proc. Roy. Soc., **B141**, 510 (1953).
29. *Haldane J. B. S.*, Proc. Camb. Phil. Soc., **23**, 838 (1927).
30. *Wright S.*, Evolution in Mendelian Populations, Genetics, **16**, 97 (1931); см. также Bull. Am. Math. Soc., **48**, 233 (1942).
31. *Perutz M.*, Nature, **228**, 726 (1970).
32. *Crick F. H. C.*, Of Molecules and Men. Seattle — London, Univ. of Washington Press, 1966.
33. *Опарин А. И.*, Происхождение жизни, изд-во «Московский рабочий», М., 1924. (Эта работа переиздана в качестве приложения в книге: *Дж. Бернал*, Возникновение жизни, изд-во «Мир», М., 1969, стр. 250.) *Опарин А. И.*, Жизнь, ее природа, происхождение и развитие, «Наука» М., 1968.
34. *Calvin M.*, Chemical Evolution, Oxford, Clarendon Press, 1969. (*М. Кальвин*, Химическая эволюция, изд-во «Мир», М., 1971.)
35. *Fox S. W.*, Science, **132**, 200 (1960); Naturwissenschaften, **56**, 1 (1969).
36. *Ponnamperuma C., Gabel N. W.*, Space Life Sci., **1**, 64 (1968).
37. *Ponnamperuma C., Sagan C., Mariner R.*, Nature, **199**, 222 (1963).
38. *Crick F. H. C.*, J. Mol. Biol., **38**, 367 (1968).
39. *Orgel L.*, J. Mol. Biol., **38**, 381 (1968).
40. *Schramm G.*, in: The Origins of Prebiological Systems (ed. by S. W. Fox), New York, Academic Press, 1965, p. 229. (Происхождение предбиологических систем, изд-во «Мир», М., 1966, стр. 303.)
41. *Schneider-Bernloehr H., Lohrmann R., Sulston J., Orgel L. E., Miles H. T.*, J. Mol. Biol., **37**, 151 (1968).
42. *Paecht-Horowitz M., Berger J., Katchalsky A.*, Nature, **228**, 636, (1970).
43. *Meixner J.*, Ann. Phys., **43**, 470 (1943); Kolloid. Z., **134**, 3 (1953).

44. *Prigogine I.*, Etude Thermodynamique des Processus Irréversibles, Liège. Desoer, 1947.
45. *De Groot S. R.*, Thermodynamics of Irreversible Processes, Amsterdam, North Holland Publ. Co., 1952. (С. Р. де Гроот, Термодинамика необратимых процессов, ГИТЛ, М., 1956.)
46. *Onsager L.*, Phys. Rev., 37, 405 (1931); 38, 2265 (1931).
47. *Glansdorff P.*, *Prigogine I.*, Thermodynamic Theory of Structure, Stability and Fluctuations, New York, Wiley Interscience, 1971.
48. *Eigen M.*, in: Les Prix Nobel en, 1967, Stockholm, 1969, pg. 151.
49. *Jost W.*, Z. Naturforsch, 20, 159 (1947); Z. phys. Chem., 105, 317 (1950).
50. *Prigogine I.*, in: Theoretical Physics and Biology (ed. by M. Mairois), Amsterdam, North Holland Publ. Co., 1969.
51. *Winkler R.*, *Eigen M.*, готовится к печати.
52. *Fresco J. R.*, *Alberts B. M.*, *Doty P.*, Nature, 188, 98 (1960).
53. *Volterra V.*, Leçons sur la théorie mathématique de la lutte pour la vie. Paris, 1931.
54. *Kamke E.*, Differentialgleichungen, Lösungsmethoden und Lösungen. Leipzig, 1944. (Э. Камке, Справочник по обыкновенным дифференциальным уравнениям, изд. 4-е, испр., изд-во «Наука», М., 1971; Э. Камке, Справочник по дифференциальным уравнениям в частных производных первого порядка, изд-во «Наука», М., 1966.)
55. *Haldane J. B. S.*, The Cause of Evolution, Longman, Green and Co., 1932. (Дж. Б. С. Холдейн, Факторы эволюции, ГИБМЛ, М. — Л., 1935.)
56. *Ross R.*, The Prevention of Malaria (1911).
57. *Lotka A. J.*, Amer. J. Hygiene, January Suppl. (1923).
58. *Lotka A. J.*, Elements of Mathematical Biology, New York, Dover Publications, 1956.
59. *Jordan P.*, Die Physik und das Geheimnis des organischen Lebens, Braunschweig, F. Vieweg u. Sohn 1945.
60. *Ramakrishnan A.*, in: Encyclopedia of Physics, vol. VIII/2, p. 524 (ed. by S. Flügge), Berlin, Springer, 1959.
61. *Bartlett M. S.*, Stochastic Processes, Cambridge, 1955. (М. Бартлетт, Введение в теорию случайных процессов, ИЛ, М., 1958.)
62. *Ehrenfest P.*, *Ehrenfest T.*, Phys. Z., 8, 311 (1907).
63. *ter Haar D.*, *Green C. D.*, Proc. Phys. Soc. London, A66, 153 (1953).
64. *Kac M.*, Amer. Math. Monthly, 54, 369 (1947).
65. *Klein J. M.*, Phys. Rev., 103, 17 (1956).
66. *Kohlrusch K. W. F.*, *Schrödinger E.*, Phys. Z., 27, 306 (1926).
67. *Münster A.*, in: Encyclopedia of Physics, vol. VIII/2, p. 218 (ed. by S. Flügge), Berlin, Springer, 1959.
68. *Bartolomay A. F.*, Bull. Math. Biophys., 20, 97 (1958); 20, 175 (1958); 21, 363 (1959).
69. *Delbrück M. J.*, Chem. Phys., 8, 120 (1940).
70. *McQuarrie D. A.*, Stochastic Approach to Chemical Kinetics, London, Methuen, 1967.
71. *Doob J. L.*, Stochastic Processes, New York, J. Wiley, 1953. (Дж. Л. Дуб, Вероятностные процессы, ИЛ, М., 1956.)

72. *Eigen M., Pörschke D.*, J. Mol. Biol., **53**, 123 (1970); cf. also part II ([91]) and part III (в печати).
73. *Eigen M.*, Nobel Symposium 5 on Fast Reactions and Primary Processes in Chemical Kinetics (ed. by S. Claesson), Stockholm, Almquist and Wiksell Intersci. Publ., 1968.
74. *Gibbs J.*, Advances in Chemical Physics (ed. by K. E. Shuler), pg. 185. New York, Interscience, 1969.
75. *Eigen M., De Maeyer L.*, in: Technique of Organic Chemistry, vol. VIII/2 (ed. by A. Weissberger), New York, Wiley Interscience, 1963.
76. *Kyogoku Y., Lord R. C., Rich A.*, Science, **154**, 518 (1966); J. Amer. Chem. Soc. **89**, 496 (1967); Proc. Nat. Acad. Sci. **57**, 250 (1967).
77. *Voet D., Rich A.*, Progr. Nucleic Acid Res. Molec. Biol., **10**, 183 (1970).
78. *Küchler E., Derkosch J.*, Z. Naturforsch., **21b**, 209 (1966).
79. *Funck Th.*, неопубликованные данные.
80. *Eggers F., Funck Th., Hopmann R.*, неопубликованные данные.
81. *Pörschke D.*, Diplom-thesis, Göttingen, 1966.
82. *Pörschke D.*, Dissertation, Göttingen-Braunschweig, 1968.
83. *Podder S. K.*, Biochemistry, **10**, 2415 (1971).
84. *Römer R., Riesner D., Coutts S. M., Maass G.*, Europ. J. Biochem. **15**, 77, 85 (1970).
85. *Römer R., Riesner D., Maass G., Wintermayer W., Thiebe R., Zachau H. G.*, FEBS Letters **5**, 15, 23 (1969).
86. *Philippsen P., Thiebe R., Wintermayer W., Zachau H. G.*, Biochem. Biophys. Res. Commun., **33**, 922 (1968).
87. *Uhlenbeck O. C., Baller J., Doty P.*, Nature, **225**, 508 (1970).
88. *Lewis J. B., Doty P.*, Nature, **225**, 510 (1970).
89. *Martin F., Pörschke D., Uhlenbeck O.*, неопубликованные данные.
90. *Applequist J., Damle V.*, J. Amer. Chem. Soc., **87**, 1450 (1965).
91. *Pörschke D.*, Biopolymers, **10**, 1989 (1971).
92. *Blow D. M., Birktoft J. J., Hartley B. S.*, Nature, **221**, 337 (1969).
93. *Fox S.*, Naturwissenschaften, **56**, 1 (1969).
94. *Berger A., Schechter I.*, Phil. Trans. Roy. Soc., London, **B257**, 249 (1970).
95. *Lipmann F.*, in: O. Meyerhoff Symposium, July 1970, Heidelberg.
96. *Monod J., Wyman J., Changeux J. P.*, J. Mol. Biol., **12**, 88 (1965).
97. *Koshland D. E., Nemethy G., Filmer D.*, Biochemistry, **5**, 365 (1966).
98. *Kirschner K., Eigen M., Bittman R., Voight B.*, Proc. Nat. Acad. Sci. (Wash.), **56**, 1661 (1966).
99. *Kirschner K.*, Curr. Topics Microbiol. Immunol., **44**, 124 (1968).
100. *Eigen M., Schuster P.*, готовится к печати.
101. *Опарин А. И.*, in: The Origin of Prebiological Systems and of their Molecular Matrices (ed. by S. W. Fox), New York, Academic Press, 1965, p. 331. (Происхождение предбиологических систем, изд-во «Мир», М., 1966, стр. 335.)

102. Fox S. W., in: The Origin of Prebiological Systems and of their Molecular Matrices (ed. by S. W. Fox), New York, Academic Press 1965, p. 361. (Происхождение предбиологических систем, изд-во «Мир», М., 1966, стр. 362.)
103. Huang C., Biochemistry 8, 344 (1969).
104. Träuble H., Grell E., Neurosci. Res. Program Bull. 1971.
105. Цофина Л. Е., Либерман Е. А., Бабаков А. В., Nature, 212, 681 (1966).
106. Delbrück M., in: The Neurosciences, 2nd Study Program, New York, The Rockefeller Univ. Press 1970, p. 677.
107. Woese C. R., The Genetic Code, New York, Harper and Row 1967.
108. Crick F. H. C., The Genetic Code. (Описание оригинального frameshift — эксперимента С. Бреннера и Ф. Крика), Scientific American, October 1962, см. также [35]. Ф. Крик, Генетический код (I), в сб. «Структура и функция клетки», изд-во «Мир», М., 1964, стр. 9.)
109. Crothers D., личное сообщение на семинаре в Гёттингене, 1970.
110. Rechenberg I., Dissertation, Technical University, Berlin, 1970.
111. Lipmann F., in: The Origin of Prebiological Systems and their Molecular Matrices, New York, Academic Press, 1965, p. 259. (Происхождение предбиологических систем, изд-во «Мир», М., 1966, стр. 261.)
112. Edelman M., Verma I. M., Littauer U. Z., J. Mol. Biol., 49, 67 (1970).
113. Woese C. R., Bio Science, 20, 471 (1970).
114. Spiegelman S., in: The Neurosciences, 2nd Study Program (ed. by F. O. Schmitt), New York, The Rockefeller Univ. Press, 1970.
115. Haruna I., Nozu K., Ohtaka Y., Spiegelman S., Proc. Nat. Acad. Sci. USA, 50, 905 (1963); 54, 579 (1965).
116. Spiegelman S., Haruna I., Holland I. B., Beaudreau G., Mills D. R., Proc. Nat. Acad. Sci. USA, 54, 919 (1965).
117. Levisohn R., Spiegelman S., Proc. Nat. Acad. Sci. USA, 60, 866 (1968).
118. Weber H., Weissmann Ch., Chimia, 24, 35 (1970).
119. Bishop D. H. L., Mills D. R., Spiegelman S., Biochemistry, 7, 3744 (1968).
120. Levisohn R., Spiegelman S., Proc. Nat. Acad. Sci. USA, 63, 805 (1969).
121. Mills D. R., Peterson R. I., Spiegelman S., Proc. Nat. Acad. Sci. USA, 58, 217 (1967).
122. Wittgenstein L., Tractatus Logico-Philosophicus, 7, London, Routledge and Kegan Paul, 1922.
123. Carathéodory C., Math. Ann., 67, 355 (1909).
124. Falk G., Jung H., in: Encyclopedia of Physics, vol. III/2, p. 119 (ed. by S. Flügge), Berlin, Springer, 1959.
125. Monod J., Le Hasard et la Nécessité, Paris, Seuil, 1970.
126. Von Neumann J., Zur Theorie der Gesellschaftsspiele, Math. Ann., 100, 295 (1928).
127. Von Neumann J., Morgenstern O., Theory of Games and Economic Behaviour, Princeton, Univ. Press, 1953. (Дж. фон Нейман,

- О. Morgenstern, Теория игр и экономическое поведение, изд-во «Наука», М., 1970; см. также С. И. Засуховицкий, Л. И. Авдеева, Линейное и выпуклое программирование, изд. 2-е переработ. и доп., изд-во «Наука», М., 1967.)
- Lewontin R. C., J. Theoret. Biol., 1, 382 (1961).
- Warburton F. E., J. Theoret. Biol., 16, 78 (1967).
- Kimura M., Genet. Res. Camb., 2, 127 (1961).
- MacKay D. M., in: Information, Mechanism and Meaning, chapter 12, pg. 132. Boston, M. I. T. Press 1969.
- Waddington C. H., The Nature of Life, London, Unwin Books, 1961.
- Stent G., The Coming of the Golden Age. Garden City N. Y., The Natural History Press, 1969.
- Kimura M., Nature, 217, 624 (1968).
- King J. L., Jukes T. H., Science, 164, 788 (1969).
- Опарин А. И., Жизнь, ее природа, происхождение и развитие, М., 1960. (Имеется 2-е дополн. изд-е, изд-во «Наука», М., 1968).
- Von Neumann J., Theory of self-reproducing automata, Urbana, Univ. Illinois Press, 1966, Дж. фон Нейман, Теория самовоспроизводящихся автоматов, изд-во «Мир», М., 1971).
- Wittgenstein L., Tractatus Logico-Philosophicus, 6. 521, London, Routledge and Kegan Paul, 1922.

