





MANFRED EIGEN

SELFORGANIZATION OF MATTER  
AND THE EVOLUTION  
OF BIOLOGICAL MACROMOLECULES

*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. ВОЛЬКЕНШТЕЙНА

ИЗДАТЕЛЬСТВО «МИР»  
Москва 1973

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, характеризующему 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. *Glandsdorff 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),

## ГЛАВА I

### ВВЕДЕНИЕ

#### § 1.1. «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).

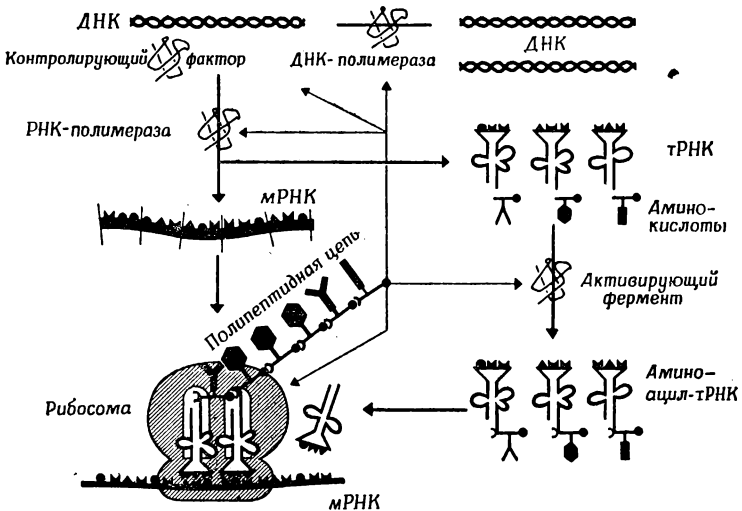


Рис. 1. Биосинтетический цикл самовоспроизведения клетки.

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 регулирующими 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 аминоацилсинтетаз. 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 — аминоацилсинтетазой. 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,  $\sigma$ -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 случайных<sup>1</sup> 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

---

<sup>1</sup> 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]<sup>1</sup>.

---

<sup>1</sup> 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 самоинструктирования?

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 отобрать 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:

$$I_1 = K \ln (Z_0/Z_1). \quad (\text{I. 1})$$

Если мы хотим отобрать одно определенное событие, т. е.  $Z_1 = 1$ , нужное количество информации равно

$$I = K \ln Z_0. \quad (\text{I. 2})$$

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 всех последовательностей длиной  $\nu$  из  $\lambda$  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  $\nu$  двоичных единиц («битов») в сообщении.

Таблица 2

## Генетический код

Второе 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, Асн, 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 <sup>1)</sup> (формилметионин)
Арг — аргинин	Про — пролин
Асн — аспарагин	Сер — серин
Асп — аспарагиновая кислота	Тре — треонин
Вал — валин <sup>1)</sup>	Три — триптофан
Гис — гистидин	Тир — тирозин
Гли — глицин	Фен — фенилаланин
Глн — глутамин	Цис — цистеин
Глу — глутаминовая кислота	Терм. — сигнал терминации
Иле — изолейцин	

<sup>1)</sup> 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 триплетами 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

Таблица 3

Из  $v$  единиц с базисом  $\lambda$  можно составить  
всего  $N_{\lambda v} = \lambda^v$  последовательностей:

Примеры	$\lambda$	$v$	$N_{\lambda v}$
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  $1 \text{ м}^3$   $6 \cdot 10^{25}$
- б) в  $10^{-3} \text{ M}$  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  $1 \text{ г/см}^3$ .

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],

$$h = -K \sum_{i=1}^{\lambda} p_i \ln p_i, \text{ причем } \sum_{i=1}^{\lambda} p_i = 1 \quad (\text{I. 4})$$

(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».

**1.2.3. 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  $\nu$  (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  $\nu$  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 кости 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

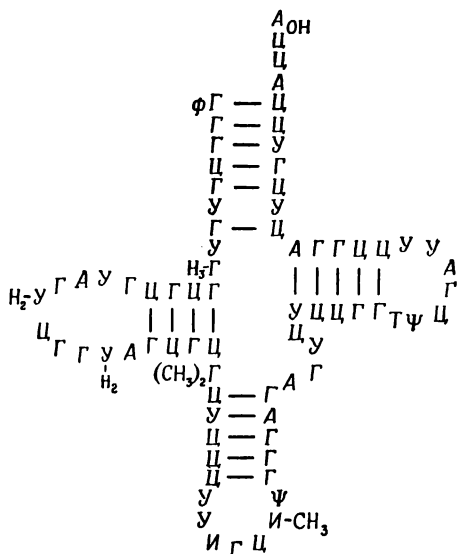


Рис. 2. «Клеверный лист» — модель тРНК.

Дана последовательность нуклеотидов аланинспецифичной тРНК из дрожжей (см. [5]). Необычные основания:  $\Psi$  — 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

$$P_k = \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{n-k} \cdot \frac{n!}{k!(n-k)!}.$$

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. Гландорф 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  $\sigma$ . Для макроскопических систем (к которым применима формула Гиббса)  $\sigma$  можно выразить в виде суммы членов, каждый из которых является произведением потока  $J_i$  и обобщенной силы  $X_i$  (для неоднородных систем  $\sigma$  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  $\sigma$  обращается в нуль.

В табл. 5 суммированы применения теории к системам химических реакций. Каждая реакция характеризуется своей «степенью развития»  $\xi_i$  или ее производной

Таблица 5

#### Определения и формулы химической термодинамики

*Сродство  $i$ -й реакции*

$$A_i = - \sum_k \nu_{ik} \mu_k, \quad (1)$$

$$A_i = RT \left( \ln K_i - \ln \prod_k a_k^{\nu_{ik}} \right) \quad (2)$$

Стехиометрические коэффициенты:

$$\nu_{i\ k} \begin{cases} > 0 \text{ для продуктов реакции,} \\ < 0 \text{ для исходных реагентов.} \end{cases}$$

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:

$$\mu_k = \mu_k^0 + RT \ln a_k, \quad (3)$$

$\mu_k^0$  — стандартный химический потенциал;  $a_k$  — активность (ее можно заменить на концентрацию  $c_k$ , если раствор идеален).

Константа равновесия:

$$K_l = \prod_k \bar{a}_k^{\nu_{l\ k}} \quad (4)$$

$\bar{a}_k$  относится к состоянию равновесия, где  $A_l = 0$ ,

$$-RT \ln K_l = \sum_k \nu_{l\ k} \mu_k^0 \text{ [ср. с уравнением (2)].} \quad (5)$$

Степень развития (координата) реакции  $\xi_l$ :

$$d\xi_l = dn_k / \nu_{l\ k} \quad (n_k \text{ — число молей компонента } k). \quad (6)$$

При постоянных  $T$  и  $P$ :

$$dG = \sum_i A_i d\xi_i \quad (7)$$

( $G$  — свободная энергия Гиббса;  $\xi_l$  сопряжена с  $A_l$ ).

Скорость реакции:

$$V_l = d\xi_l / dt = \sum_k e_{l\ k} \cdot A_k / T \quad (8)$$

( $e_{l\ k}$  — феноменологические коэффициенты).

Соотношения Онзагера:

$$e_{l\ k} = e_{k\ l}. \quad (9)$$

Линеаризация:

$$V_l = \frac{1}{T} \sum_k e_{l\ k} \sum_l \left( \frac{\partial A_k}{\partial \xi_l} \right)_{T, P} \delta \xi_l. \quad (10)$$

Матрицу  $(e_{l\ k})$  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,  $\tau_i$  — время релаксации;

$$\delta\xi'_i = \delta\xi'_{i0} e^{-t/\tau_i} \quad (12)$$

— решение кинетического уравнения.

Производство энтропии:

$$\sigma = \frac{1}{T^2} \sum_{i,k} e_{ik} A_i A_k \geq 0. \quad (13)$$

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  $\sigma$  всегда можно выразить в виде квадратичной (положительно определенной) формы либо через величины  $\xi_i$ , 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:

$$(c_i - \bar{c}_i) = \sum_k a_{ik} e^{-t/\tau_k}. \quad (I. 6)$$

Приход к равновесию — это процесс «релаксации», характеризующийся спектром «времен релаксации»  $\tau_k$ . 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. Гландорф [47] вывели соотношение, аналогичное уравнению (I.5), которое выполняется

вблизи стационарных состояний:

$$d_x \sigma = \sum_k J_k dX_k \leq 0, \quad (\text{I.10})$$

т. е. изменение производства энтропии, обусловленное изменением сил  $X_k$  (например, изменение  $\sigma$  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»

$$\delta_x \sigma = \sum_k \delta J_k \delta X_k \geq 0. \quad (\text{I.11})$$

Эта величина положительна и вследствие (I.10), как и в случае производства энтропии вблизи равновесия, содержит только члены второго порядка; например, для химической системы

$$\delta_x \sigma = \frac{1}{T} \sum_k \delta V_k \delta A_k \geq 0. \quad (\text{I.12})$$

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**

$X + Y \rightleftharpoons 2X$ , т. е.  $Y$  превращается в  $X$ ,

причем  $X$  является катализатором этой реакции <sup>1)</sup>

**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*

$$V = \vec{k} [X] [Y] - \overleftarrow{k} [X]^2 = \{ \vec{k} [Y] - \overleftarrow{k} [X] \} [X], \quad (4)$$

так как  $\vec{k} [Y] \approx \overleftarrow{k} [X]$ , член  $\{ \}$  мал, а в равновесии равен нулю.

$$\delta V \approx -\overleftarrow{k} [X] \delta [X] \quad (5)$$

(пренебрегая членами второго порядка).

*Сродство:* как и выше.

*Избыточное производство энтропии:*

$$T\delta_x\sigma \geq 0 \quad (\text{устойчивое равновесие}). \quad (6)$$

<sup>1)</sup> Символом  $X$  здесь обозначено химическое соединение, и его необходимо отличать от обобщенных сил  $X_i$ . Индекс  $x$  при  $d$  или  $\delta$  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.

## ГЛАВА II

# 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:

$$i_{vk} \ (k = 1, 2, \dots, N_v), \quad (\text{II. 1})$$

которые представляют собой последовательности из  $v$  единиц с базисом  $\lambda$  (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  $v$  единиц.

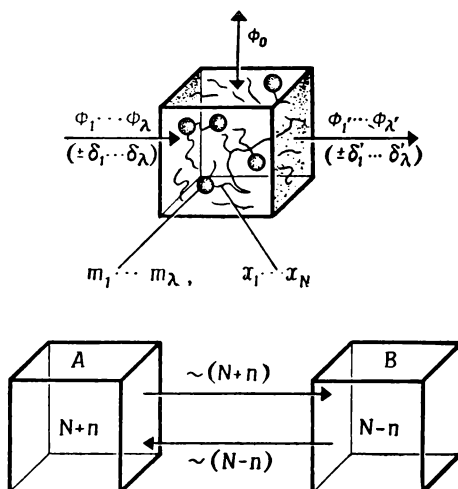


Рис. 3. «Информационный ящик».

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  $\Phi_0$ . Для сравнения изображена равновесная система, распределенная между двумя ящиками A и B:  $N_A = N_B$ ;  $\pm n$  — флуктуации (обсуждение эренфестовской модели см. в § III. 2).

В классической теории информации «количество информации» в определенной последовательности  $i_{vi}$  выражается числом битов:

$$I_v = v \ln \lambda / \ln 2 \text{ [битов]}. \quad (\text{II. 2})$$

Тогда  $N_v$  представляет собой «структурную емкость» класса  $v$ , т. е. число всех различных последовательностей с длиной  $v$  и с базисом  $\lambda$ :

$$N_v = \lambda^v = 2^{I'_v}. \quad (\text{II. 3})$$

Если последовательности могут иметь любую длину, ограниченную верхним пределом  $\nu$  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  $\nu$ ), 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  $N_{\nu}$  возможных носителей  $i_{\nu k}$  их концентрации  $x_{\nu k}$  будут равны нулю. Концентрации неорганизованных высокоэнергетических мономерных единиц 1, ...,  $\lambda$  в ящике равны  $m_1, \dots, m_{\lambda}$ , 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 равномерном 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-

<sup>1</sup> 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 информации<sup>1</sup> ( $\dot{x} \equiv dx/dt$ ):

$$\dot{x}_i = (\mathcal{F}_i - \mathcal{R}_i) x_i + \sum_{l \neq i} \varphi_{il} x_l. \quad (\text{II. 10})$$

Первый и второй члены в этом уравнении определяют скорость *самоинструктированного* 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 ( $m_1, \dots, m_\lambda$ ); кроме того,  $\mathcal{F}_i$  и  $\mathcal{R}_i$  могут зависеть еще от  $x_i$  или от  $x_h$  — 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)<sup>2</sup>

$$\mathcal{F}_i = k_0 \mathcal{A}_i Q_i; \quad \mathcal{R}_i = k_0 \mathcal{D}_i + \varphi_{0i}. \quad (\text{II. 11})$$

Здесь  $k_0$  — *общая* константа скорости первого порядка с размерностью  $\text{с}^{-1}$ . 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.

<sup>1</sup> В дальнейшем мы будем опускать индекс  $v$ , кроме тех случаев, когда необходимо выделить соответствующий класс носителей информации.

<sup>2</sup> 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:

$$k_0 \sum_{k=1}^N \mathcal{A}_k (1 - Q_k) x_k = \sum_{k=1}^N \sum_l \varphi_{kl} x_l. \quad (\text{II. 13})$$

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  $m_1, \dots, m_\lambda$  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  $m_1, \dots, m_\lambda$ , 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  $\mathcal{R}_i$  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  $\varphi_{0i}$ , 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  $\varphi_0$  будет связан с общим потоком «разбавления»  $\phi_0^1$  следующим образом:

$$\varphi_0 = \frac{\phi_0}{\sum_{k=1}^N x_k}. \quad (\text{II. 15})$$

---

<sup>1</sup>  $\varphi_0$  может быть также отрицательным, если происходит «концентрирование» системы. Однако обычно член  $\varphi_0$  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  $\mathcal{A}_i$ ,  $Q_i$  и  $\mathcal{D}_i$ . Тогда уравнение (II.10) примет вид

$$\dot{x}_i = k_0 [\mathcal{A}_i Q_i - \mathcal{D}_i] x_i + \sum_{l \neq i} \phi_{li} x_l - \phi_{0i} x_i. \quad (\text{II. 17})$$

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  $Q_i < 1$ .

Оказывается, что первая часть уравнения (II.17) ответственна за способность к отбору и что  $\mathcal{A}_i$ ,  $Q_i$  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 ( $\mathcal{A}_i Q_i \geq \mathcal{D}_i$ ). 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 Гландорфа [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  $\phi_0$ ) 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  $\phi_0$  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 $\beta$ , 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  $\varphi_0$  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  $\Delta t$

$$\int_t^{t+\Delta t} \bar{E} dt \rightarrow \int_t^{t+\Delta t} W_m^0 dt \quad \text{или} \quad \langle \bar{E} \rangle \rightarrow \frac{1}{\Delta t} \int_t^{t+\Delta t} W_m^0 dt. \quad (\text{II. 34})$$

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  $\mathcal{A}_m$ ,  $\mathcal{Q}_m$  и  $\mathcal{D}_m$  — 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

$$W_m^0 + \frac{1}{k_0} \sum_{l \neq m} \Phi_{ml} \frac{\bar{x}_l}{\bar{x}_m},$$

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:

$$\bar{E} \rightarrow W_m^0 + \frac{1}{k_0^2} \sum_{l \neq m} \frac{\Phi_{ml} \Phi_{lm}}{W_m^0 - W_l^0}. \quad (\text{II. 33a})$$

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  $\phi_M$  богатых энергией мономеров (в постоянном соотношении). Тогда ограничения, введенные равенствами (II.27) и (II.28), дают

$$f(m_1 \dots m_\lambda) = \frac{\phi_M}{k_0 \sum_k^N \mathcal{A}'_k x_k} \quad (\text{II. 36})$$

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.

$$1 - \frac{\bar{x}_m}{n} = \frac{\mathcal{A}_m}{E_m - \bar{E}_{k \neq m}} (1 - Q_m). \quad (\text{II. 44})$$

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

$$Q_m > Q_{\min} = \frac{\bar{\mathcal{A}}_{k \neq m} + \mathcal{D}_m - \bar{\mathcal{D}}_{k \neq m}}{\mathcal{A}_m}, \quad (\text{II. 45})$$

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  $\mathcal{A}$ -,  $Q$ - и  $\mathcal{D}$ -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  $Q \leq 1$ )

$$\frac{\bar{x}_m}{\phi_M/k_0\mathcal{D}_m} \equiv \frac{\bar{x}_m}{\bar{n}} = \frac{W_m^F - \bar{P}_{k \neq m}}{P_m - \bar{P}_{k \neq m}}. \quad (\text{II. 46})$$

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

$$\frac{\phi_M}{k_0\mathcal{D}_m} = \frac{\sum_{k=1}^N \mathcal{D}_k x_k}{\mathcal{D}_m} \equiv \bar{n}. \quad (\text{II. 47})$$

Здесь поток  $\phi_M$  — 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.  $\mathcal{D}_m = \mathcal{D}_{k \neq m}$ , т. е.  $\phi_M/\mathcal{D}_m = \sum_k x_k = n$ .

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) элементарных пар и другое для всех (вырожденных) неэлементарных 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,

$$\text{сумма: } \sum_{k=0}^v \binom{v}{k} (\lambda - 1)^k = \lambda^v \quad \begin{array}{l} \text{возможных} \\ \text{последовательностей} \end{array} \quad (3)$$

(ср. табл. 4).

*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:

$$P_{vk} = \frac{Q_{vk}}{\binom{v}{k} (\lambda - 1)^k} = \frac{q^{v-k} (1-q)^k}{(\lambda - 1)^k} \rightarrow \left( \frac{q^{-1} - 1}{\lambda - 1} \right)^k e^{-\varepsilon}. \quad (8)$$

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).

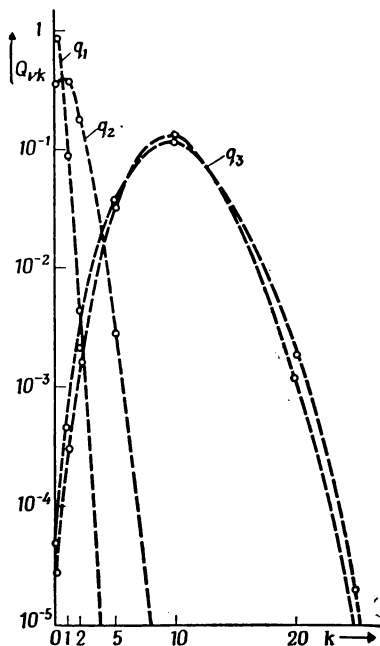


Рис. 4. Распределения вероятностей появления ошибочных копий.  $Q_{vk}$  вычислены по уравнению (5) в табл. 8. Значения параметров:

$v=100$	$q_1=0,999$	$\varepsilon_1=0,1$
$\lambda=4$	$q_2=0,990$	$\varepsilon_2=1$
	$q_3=0,900$	$\varepsilon_3=10$

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  $k \rightarrow v$ , 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  $Q_i$  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

$\bar{E}_{k \neq i}$  переменна;  $Q_i \approx 1$ , 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]

$$\frac{1}{x(t)} = e(t) \int \frac{\dot{f}(t)}{e(t)} dt; \quad e(t) = \exp \left[ \int g(t) dt \right]. \quad (\text{II. 54})$$

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

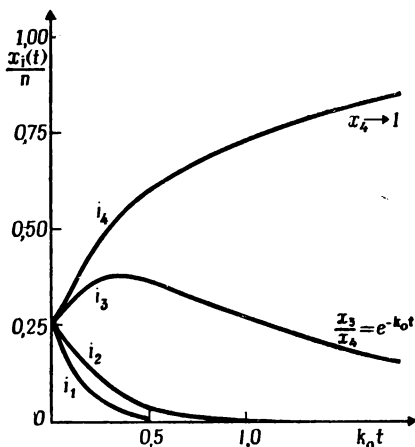


Рис. 5. Отбор в системе четырех конкурирующих видов по уравнению (II.49).

$$W_1^0 = 1; W_2^0 = 4; W_3^0 = 9; W_4^0 = 10.$$

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 выродившимся классом видов  $k \neq m$ , в котором  $\bar{E}_{k \neq m}$  или

$R_{k \neq m}$  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  $R_{k \neq m}$ , но дают довольно постоянное распределение вокруг средних значений  $\bar{E}_{k \neq m}$  и  $\bar{P}_{k=m}$ , в то время как отобранный вид имеет другую скорость воспроизведения с  $W_m^0 > \bar{E}_{m \neq k}$  или  $W_m^0 > \bar{P}_{k \neq m}$ .

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

$$\bar{x}_m = n \frac{W_m^0 - \bar{E}_{k \neq m}}{E_m - \bar{E}_{k \neq m}} \text{ [ср. с равенством (II. 43)]}$$

— «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  $n\bar{\mathcal{A}} = n\bar{\mathcal{D}} \equiv \phi_M$  являются константами (в предположении, что образование и распад подчиняются кинетике первого порядка), в то время как члены  $n\bar{\mathcal{A}}'$  и, следовательно,  $(1 + \bar{P}) \equiv \bar{\mathcal{A}}'/\bar{\mathcal{D}}$  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:

$$\frac{[x_m(t)]^{\alpha_m}}{\dot{x}_m - x_m(t)} = \frac{[x_m^0]^{\alpha_m}}{\dot{x}_m - x_m^0} \exp \{k_0 \mathcal{D}_m [1 - \alpha_m] t\}$$

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

$E_{k \neq m}$  переменна; 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  $\lambda$  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  $i_{jl}$

$$\dot{x}_{jl} = k_0(W_{jl} - \bar{E})x_{jl}. \quad (\text{II. 64c})$$

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  $\beta$  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  $\beta_{mi}$ , 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 ( $x_m^0 = x_m(t=0)$ ):

$$\begin{aligned} \frac{x_m(t)}{n} = & W'_m x_m^0 \left\{ x_m^0 \left[ E_m + \sum_{k=1}^{v(\lambda-1)} A_k E_{1k} \right] + \right. \\ & + \sum_{k=1}^{v(\lambda-1)} B_k \exp \{k_0 (W_{1k} - W'_m) t\} + \\ & \left. + \sum_{jk} C_{jk} \exp \{k_0 (W_{jk} - W'_m) t\} \right\}^{-1}, \quad (\text{II. 67}) \end{aligned}$$

где

$$\begin{aligned} A_k &= - \frac{E_m - W_m}{\beta_{mk} (W_{1k} - W'_m)}, \\ B_k &= \frac{W'_m}{W_{1k}} E_{1k} \left( x_{1k}^0 + x_m^0 \frac{E_m - W_m}{\beta_{mk} (W_{1k} - W'_m)} \right), \quad (\text{II. 68}) \\ C_{jk} &= \frac{W'_m}{W_{jk}} E_{jk} x_{jk}^0. \end{aligned}$$

$W'_m$  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:

$$\begin{aligned} \frac{x_{1i}(t)}{x_m(t)} = & \left\{ \frac{x_{1i}^0}{x_m^0} + \frac{E_m - W_m}{\beta_{mi}(W_{1i} - W'_m)} \right\} \times \\ & \times \exp[k_0(W_{1i} - W'_m)t] - \frac{E_m - W_m}{\beta_{mi}(W_{1i} - W'_m)}, \end{aligned} \quad (\text{II. 72})$$

$$\frac{x_{ji}(t)}{x_m(t)} = \frac{x_{ji}^0}{x_m^0} \exp\{k_0(W_{ji} - W'_m)t\}. \quad (\text{II. 73})$$

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.

## ГЛАВА III

### 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 событий<sup>1</sup>.

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

---

<sup>1</sup> 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 стохастический<sup>1</sup> 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.

---

<sup>1</sup> «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 флуктуирует. Распределение вероятностей симметрично по отношению к  $n = 0$  (распределение Гаусса), причем его полуширина 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 несущественны 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 сформируем 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



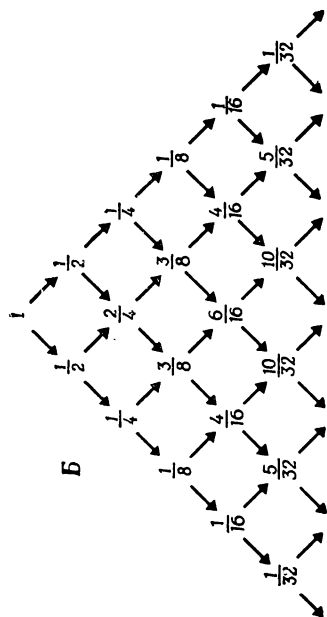
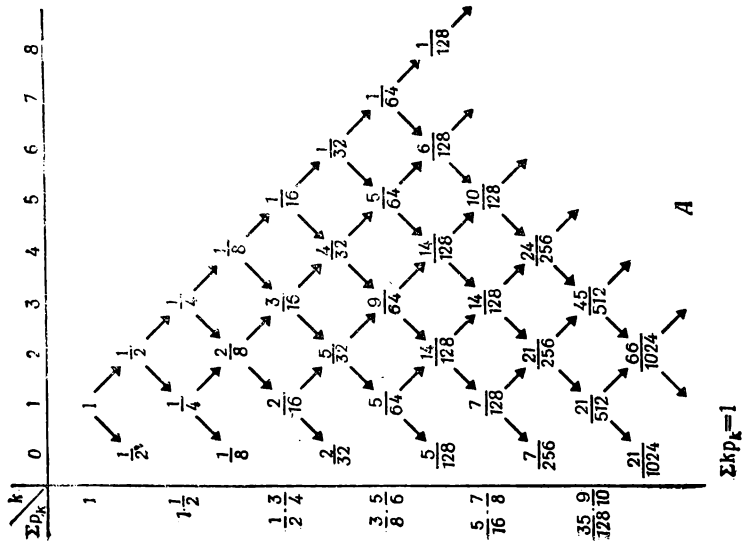


Рис. 6. Графическое представление вероятностей для модели стационарного состояния.

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 потребуется 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

$$\frac{1}{\mathcal{F}} \left( \sum_{i=1}^k \frac{1}{i} \right) \rightarrow \frac{1}{\mathcal{F}} (C + \ln k) \text{ для больших } k,$$

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  $W_i > E$ .

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  $\mathcal{A}_k$  точно равны факторам разложения  $\mathcal{D}_k$ , но факторы качества  $Q_k$  меньше единицы, каждый вид должен вымереть, потому что для каждого вида  $W_k = \mathcal{A}_k Q_k - \mathcal{D}_k < 0$ . 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 элегантно 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

$$p\{x_{t_0+t} = S_j | x_{t_0} = S_i\}, \quad (1)$$

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. Бартоломео [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  $\Delta x_1$  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.  $\Delta t_1$  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  $\pm 1$ , 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

$$e_i(t) = \left( \frac{\partial \phi_i}{\partial s} \right)_{s=1} = ie^{(\mathcal{F}-\mathcal{R})t}, \quad (16)$$

т. е. совпадает с результатом детерминистической теории; дисперсия  $i$  равна

$$\sigma_i^2(t) = \left( \frac{\partial^2 \phi_i}{\partial s^2} \right)_{s=1} + e_i(t) - e_i^2(t) = \\ = i \frac{\mathcal{F} + \mathcal{R}}{\mathcal{F} - \mathcal{R}} e^{(\mathcal{F}-\mathcal{R})t} [e^{(\mathcal{F}-\mathcal{R})t} - 1]. \quad (17)$$

2.  $\mathcal{F} = \mathcal{R}$ . Решение аналогично:

$$\phi_i(s, t) = \left[ \frac{\mathcal{F}t - (\mathcal{F}t - 1)s}{(1 + \mathcal{F}t) - \mathcal{F}ts} \right]^i; \quad (18)$$

$$p_{i \ k}(t) = \sum_{n=0}^{k \text{ или } i} (-1)^n \binom{i}{n} \binom{i+k-n-1}{k-n} (\mathcal{F}t)^{i+k-2n} \times \\ \times (\mathcal{F}t - 1)^n (\mathcal{F}t + 1)^{-i-k+n}; \quad (19)$$

$$e_i(t) = \left( \frac{\partial \phi_i}{\partial s} \right)_{s=1} = i; \quad (20)$$

$$\left( \frac{\partial^2 \phi_i}{\partial s^2} \right)_{s=1} = i(2\mathcal{F}t + i - 1). \quad (21)$$



Рассматриваемый процесс, детерминистический аспект которого выражен уравнением (II. 10) с  $\varphi_{il} = 0$ ,  $\mathcal{F}_i$  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

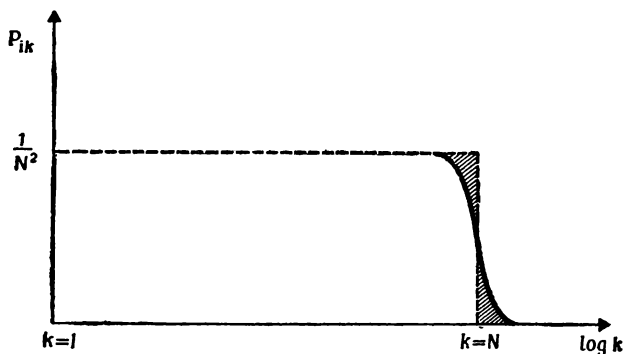


Рис. 7. Распределение вероятностей для «выживших» при  $t_1 = N/\mathcal{F}$ , согласно (III. 18).

$t_1 = N/\mathcal{F}$  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  $\varepsilon = \sum_{k=0}^{\infty} k p_k$  ( $p_k$  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

$$e_i(t) = ie^{(\mathcal{F}-\mathcal{R})t} \quad [\text{ср. также с уравнением (III. 23)}]$$

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.



## ГЛАВА IV

### САМООРГАНИЗАЦИЯ, ОСНОВАННАЯ НА КОМПЛЕМЕНТАРНОМ УЗНАВАНИИ: НУКЛЕИНОВЫЕ КИСЛОТЫ

#### § IV.1. Истинное «самоинструктирование»

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 «самоинструктированного» 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  $x_h$  (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  $\mathcal{A}_i$  и  $Q_i$ ), в результате чего могут получиться очень однородные последовательности. Если в репликации участвуют ферменты<sup>1</sup>, то в узнавании (как и в случае «нуклеации») могут участвовать специфичные участки матричной последовательности (возможно, что участвуют оба конца цепи, как в случае Q $\beta$ -репликазы; см. гл. VII).

Фактор  $Q_{i0}$  накладывает более жесткие ограничения на длину цепи, чем фактор усиления, вследствие степенной зависимости

$$Q_{i0} = \prod_{k=1}^{\lambda} q_k^{v_{ik}}, \quad (\text{IV. 2})$$

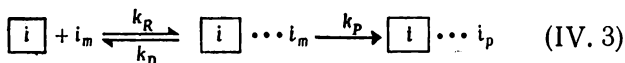
где индекс  $k$  относится к различным единицам  $1, \dots, \lambda$ . Если бы факторы узнавания отдельных единиц  $q_i$  были

<sup>1</sup> 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 самоинструкированного 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 ( $10^6$ — $10^7$  с<sup>-1</sup>), и можно принять, что  $k_D$  больше  $10^5$  с<sup>-1</sup> (для ГЦ) или  $10^6$  с<sup>-1</sup> (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 гидролизываться 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.

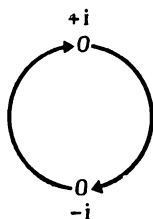
Очень маловероятно, чтобы все мономеры имелись с самого начала в равных количествах. Если, с другой стороны, одна из единиц — например, адениновый нуклеотид — имеется в большом избытке, то это так сильно будет благоприятствовать ее включению при конкуренции с другими единицами, что может привести к доминированию 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  $\lambda$  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  $\lambda$  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

$$\vec{y}_i = \mathbf{M}_i \vec{x}_i; \quad \vec{x}_i = \mathbf{M}_i^{-1} \vec{y}_i, \quad (\text{IV.11})$$

где матрица  $\mathbf{M}_i^{-1}$  обратна матрице  $\mathbf{M}_i$ , а  $\mathbf{M}_i$  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:

$$\mathbf{M}_i = \begin{bmatrix} 1 & -\sqrt{\frac{\mathcal{A}_{+i}\mathcal{Q}_{+i}}{\mathcal{A}_{-i}\mathcal{Q}_{-i}}} \\ +\sqrt{\frac{\mathcal{A}_{-i}\mathcal{Q}_{-i}}{\mathcal{A}_{+i}\mathcal{Q}_{+i}}} & 1 \end{bmatrix};$$

$$\mathbf{M}_i^{-1} = \frac{1}{2} \begin{bmatrix} 1 & +\sqrt{\frac{\mathcal{A}_{+i}\mathcal{Q}_{+i}}{\mathcal{A}_{-i}\mathcal{Q}_{-i}}} \\ -\sqrt{\frac{\mathcal{A}_{-i}\mathcal{Q}_{-i}}{\mathcal{A}_{+i}\mathcal{Q}_{+i}}} & 1 \end{bmatrix}. \quad (\text{IV.12})$$

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  $\mathbf{M}_i$  и  $\mathbf{M}_i^{-1}$  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 ( $y_{1t}^0 = 0$ ). В более обычном случае (см. эксперименты с фагом Q $\beta$ , описанные в гл. VII) процесс начинается с одной копии, например с плюс-цепи ( $x_{-t}^0 = 0$ ). 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):

$$\begin{aligned}\dot{x}_{+t} &= k_0 \left[ \left( \mathcal{A}_{+t} Q_{+t} \frac{x_{-t}}{x_{+t}} - \mathcal{D}_{+t} \right) - \bar{E} \right] x_{+t}, \\ \dot{x}_{-t} &= k_0 \left[ \left( \mathcal{A}_{-t} Q_{-t} \frac{x_{+t}}{x_{-t}} - \mathcal{D}_{-t} \right) - \bar{E} \right] x_{-t}.\end{aligned}\quad (\text{IV. 18})$$

Сразу видно, что эти уравнения относятся к тому общему типу, который обсуждался в гл. II, поскольку отношение ( $x_{+t}/x_{-t}$ ) или обратное отношение не исчезает с ростом  $t$ .

Обозначим это отношение  $x_{+t}/x_{-t} \equiv z_t$ . Из (IV. 18) немедленно следует дифференциальное уравнение для изменения  $z$  во времени:

$$\begin{aligned}\dot{z}_t &= \frac{\dot{x}_{+t}x_{-t} - \dot{x}_{-t}x_{+t}}{x_{-t}^2} = \\ &= k_0 [\mathcal{A}_{+t} Q_{+t} + (\mathcal{D}_{-t} - \mathcal{D}_{+t}) z_t - \mathcal{A}_{-t} Q_{-t} z_t^2].\end{aligned}\quad (\text{IV. 19})$$

В результате интегрирования получаем довольно громоздкое выражение

$$z_t(t) = \frac{\bar{z}_t (z_t^* + e^{-\alpha_i t}) - 2\beta_i e^{-\alpha_i t}}{z_t^* - e^{-\alpha_i t}}, \quad (\text{IV. 20})$$

где

$$z_t^* = \frac{z_t^0 + \bar{z}_t - 2\beta_t}{z_t^0 - \bar{z}_t}, \quad \beta_t = \frac{\mathcal{D}_{-t} - \mathcal{D}_{+t}}{2\mathcal{A}_{-t} Q_{-t}}$$

и

$$\alpha_i = k_0 \sqrt{4\mathcal{A}_{+t}\mathcal{A}_{-t}Q_{+t}Q_{-t} + (\mathcal{D}_{+t} - \mathcal{D}_{-t})^2},$$

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 «самоинструктирующей» 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 уравниванию 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 уравнивания 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  $\mathbf{M}_i$  и  $\mathbf{M}_i^{-1}$ .)

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 по-чему-либо выгодно. Вследствие конкурентной природы отдельных ( $\pm i$ )-ансамблей количество информации, которое может храниться, ограничено емкостью одноединственного 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 специфичного 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.

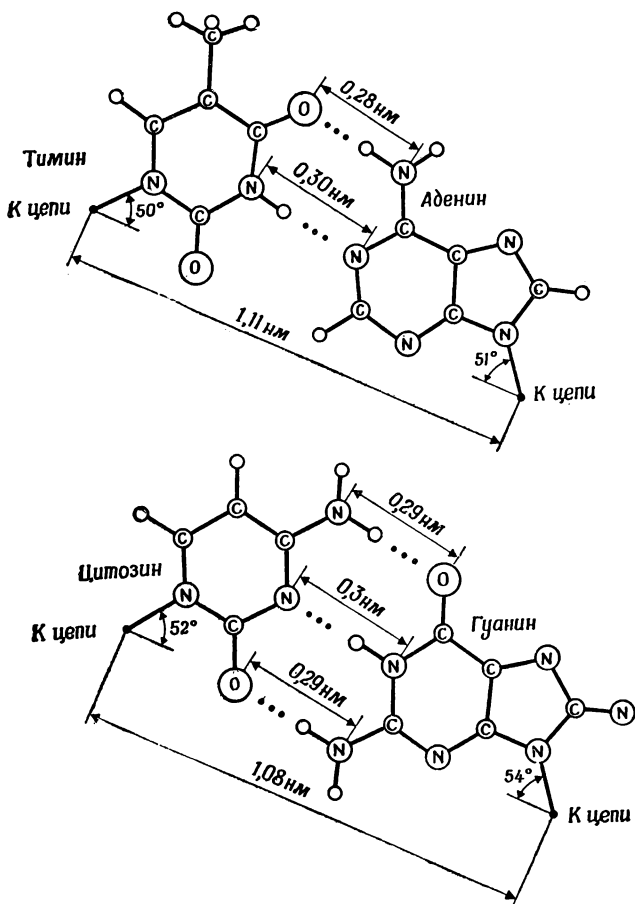


Рис. 8. Комплементарные пары оснований.

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-

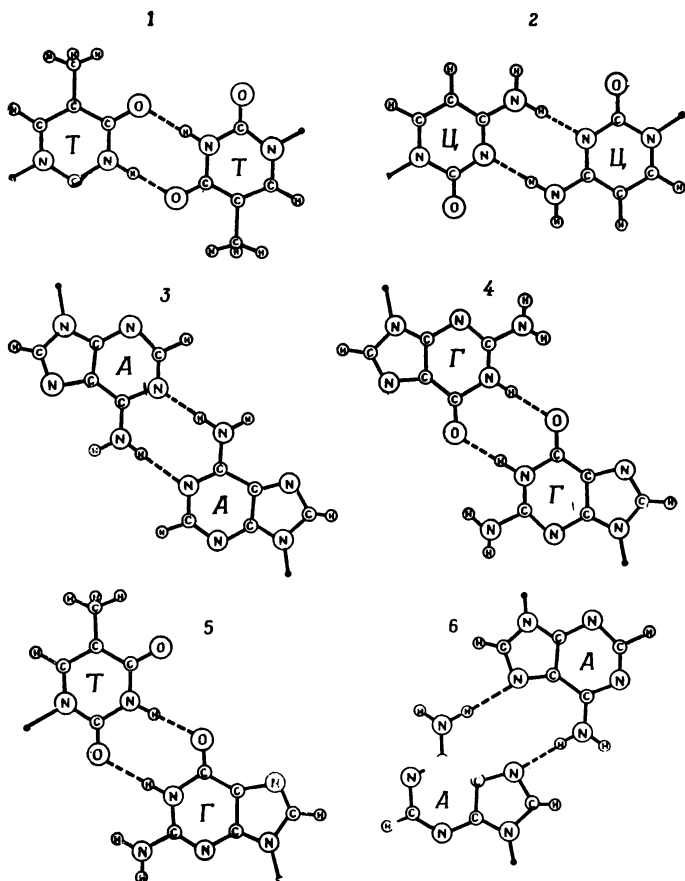


Рис. 9. «Некомплементарные» комбинации оснований.

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  $\text{NH}$ ,  $\text{NH}_2$ ,  $\text{OH}$  и  $\text{CO}$  или  $\text{N}$ , являются единственным источником стабилизации, то полярные молекулы  $\text{H}_2\text{O}$  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

$\begin{array}{c} \text{CCl}_4 \\ \text{C}_6\text{H}_6 \end{array}$	У	А	Ц	Г
У	45 15	550	< 50	< $10^3$
А	150	22 8	< 50	< $10^3$
Ц	< 28	< 28	50 28	> $10^4$
Г	< $1,2 \cdot 10^3$	< $1,2 \cdot 10^3$	$3 \cdot 10^4$	$\sim 10^3$ $1,2 \cdot 10^5$

$K_{\text{асс}} [M^{-1}]$  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 до-вольно кислой группой  $\text{NH}$  ( $\text{pK} \sim 9,5$ ) и азотом кольца ( $\text{pK} \sim 4,5$ ). 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 других<sup>1</sup> 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  $\text{pH}$  (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 встретилося с большими трудностями, обусловленными агрегацией этих молекул до более сложных структур. С. Поддеру (во время его пребывания в Гёттингене) удалось исследовать спаривание тетрамеров



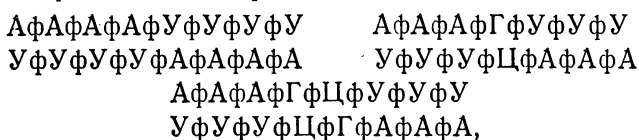
а также спаривание ГфГ и ГфГфГ с поли-рибоЦ.

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

<sup>1</sup> 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[UAЦ]УЦ	$K_{\text{ass}}^{1)}$	Тир-тРНК	$K_{\text{ass}}$
АУГ (правильный кодон)	1 200 $\pm 200$	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 <sup>2)</sup>	
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)  $K_{\text{ass}}(M^{-1})$  измерены в водном растворе (1,0 М NaCl, 10 мМ MgCl<sub>2</sub>, 10 мМ фосфатный буфер, рН 7) при 0° С. При значениях  $K < 400 M^{-1}$  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 перехода спираль  $\rightarrow$  статистический клубок) от  $1/(v-1)$  ( $v$  — 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,  $T_\infty$  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»  $\sigma_i$  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  $v > n$  пар оснований:

$$K_v = \bar{\sigma} s^v, \quad \bar{\sigma} = \sigma_1 \sigma_2 \dots \sigma_n. \quad (\text{IV. 24})$$

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  $\sigma_n$  приближается к единице для  $n \geq 3$ . Тогда кривые плавления показывают, что для  $n \geq 3$  имеется постоянный (и отрицательный) инкремент  $\Delta H_s$  и  $\Delta S_s$ , 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 ( $\Delta H_{\bar{\sigma}s} \sim 0$ ). Хотя  $\bar{\sigma}s$  не является константой стабильности для первой пары оснований (она равна  $\sigma_1s$ ), физическая интерпретация состоит в том, что существенная часть  $\Delta H_s$  происходит за счет стопочного взаимодействия (отметим относительно большой инкремент  $\Delta H$ , обусловленный стопочным взаимодействием, для каждой отдельной цепи), а не от образования водородных связей между комплементарными основаниями. Значение  $\Delta H$  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 ( $k_R: 10^6 - 10^7 \text{ M}^{-1} \text{ с}^{-1}$ ). 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 ( $\text{pH} \sim 4$ ), АУ-спиралей (нейтральная область,  $T < 50^\circ \text{C}$ ), а также для ГЦ-спиралей при более высокой температуре ( $T > 50^\circ \text{C}$ ).

Абсолютные значения этих параметров «активации» позволяют совершенно непосредственно определить длину нуклеации. Если бы образование первой пары оснований лимитировало скорость<sup>1</sup>, то энергия активации (для процесса столкновения) была бы положительной. Поскольку мы знаем, что первая пара нестабильна ( $\sigma_1 s < 1$  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:  $k_R = \sigma_1 s k_{12}$ , 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  $\sigma s$  имеет почти нулевой инкремент  $\Delta H$ , а  $\sigma_1 s$  включает в лучшем случае малое отрицательное значение  $\Delta H$ , то полная кажущаяся энергия активации должна быть близка к нулю, но не к экспериментально

<sup>1</sup> 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

$$k_D = \frac{k_{23}}{s^{v-2}} \text{ (так как } k_R/k_D = \sigma s^v \text{)}. \quad (\text{IV. 26})$$

Поскольку в знаменателе стоит  $s^{v-2}$ , энергии активации положительны и включают в себя большой инкремент  $(v - 2)\Delta H_s$ .

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 триплету, 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 <sup>1)</sup>**

Кооперативная пара	$\sigma, M^{-1}$	$s$	$\Delta H$ <sup>2)</sup> , 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 ...						
... Г ...	$\sim 10^{-3}$	От 100	-15	$10^6$	от +5	$10^7$
... Ц ...		до 200			до +7	

<sup>1)</sup> 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. Подлера [76].

<sup>2)</sup> 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  $\sigma s^N$  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  $\sigma$  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 спиралах с длиной цепи более чем  $v = 4$ , 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),

$$q_{AU} = \frac{m_Y K_{AU}}{\sum_{X=AU\Gamma C} m_X K_{AX}}, \quad (\text{IV. 27})$$

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 самоинструктрованию. 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.*

## ГЛАВА V

# 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  $\beta$ -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

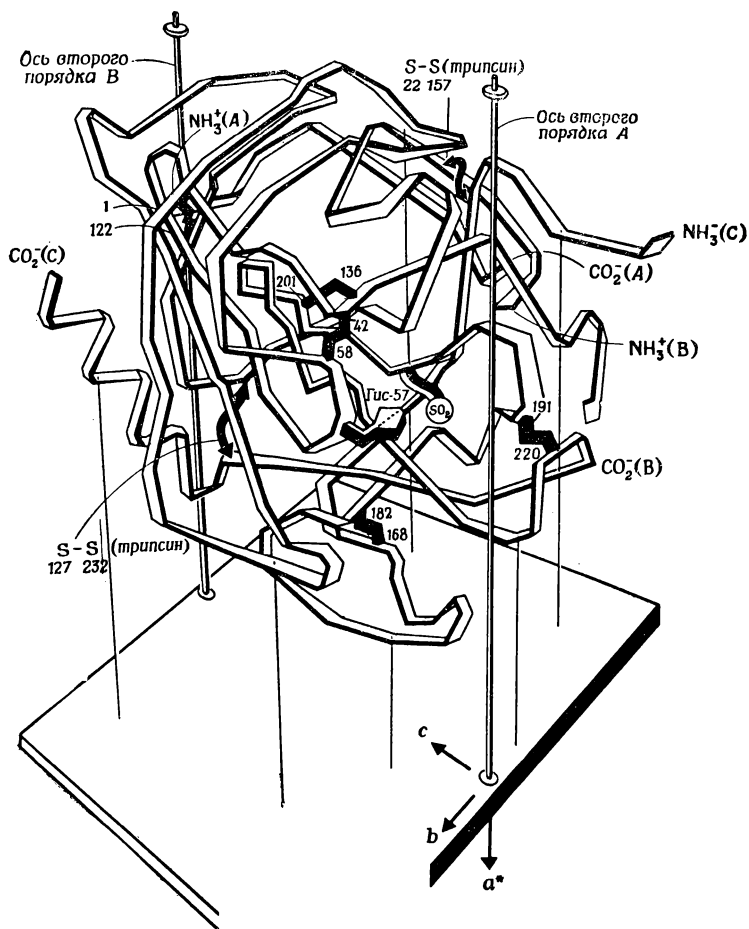
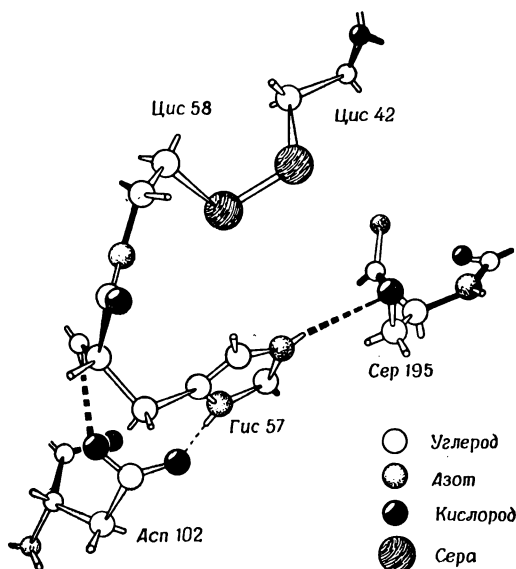


Рис. 10. Схематическое изображение конформации полипептидных цепей в  $\alpha$ -химотрипсине [92].

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 рентгено-структурного 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  $\alpha$ -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, химотрипсина-

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 самими 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 пентапептидов); 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

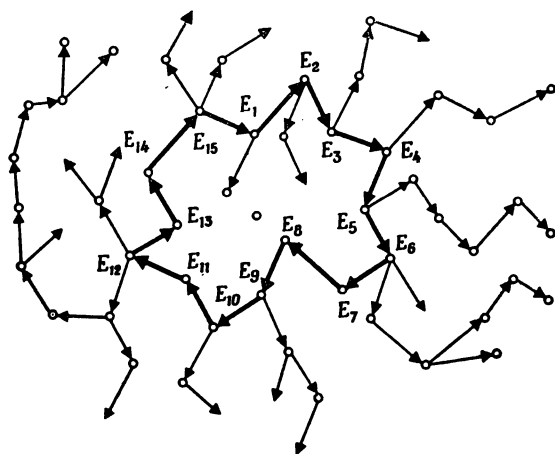
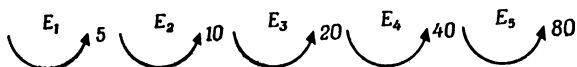


Рис. 12. Каталитическая сеть, в которой имеется замкнутая петля  $E_1, \dots, E_{15}$ .

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

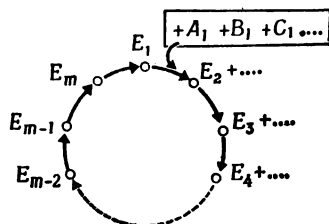


Рис. 13. Представление каталитического цикла в виде графа.

$E_1$  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

$$\left. \begin{aligned} \dot{x}_1 &= \mathcal{F}_1 x_m - \mathcal{R}_1 x_1 \\ \dot{x}_2 &= \mathcal{F}_2 x_1 - \mathcal{R}_2 x_2 \\ &\vdots \\ \dot{x}_m &= \mathcal{F}_m x_{m-1} - \mathcal{R}_m x_m \end{aligned} \right| \begin{aligned} \mathcal{F}_i &= k_0 \mathcal{A}_i Q_i \\ \mathcal{R}_i &= k_0 \mathcal{D}_i \end{aligned} \quad (\text{V. 1})$$

Матрице кинетических коэффициентов

$$\begin{pmatrix} -\mathcal{R}_1 & 0 & . & 0 & \mathcal{F}_1 \\ \mathcal{F}_2 & -\mathcal{R}_2 & & & 0 \\ 0 & \mathcal{F}_3 & . & . & . \\ . & . & . & . & 0 \\ 0 & . & 0 & \mathcal{F}_m & -\mathcal{R}_m \end{pmatrix} \quad (\text{V. 2})$$

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  $\lambda$ , в котором все члены, содержащие  $\lambda^i$  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  $\hat{\mathcal{F}}$ , 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  $\mathcal{R}$ )<sup>1</sup>:

$$\frac{\bar{x}_{l-1}}{\bar{x}_i} = \frac{\sqrt[m]{\prod_{k=1}^m \mathcal{F}_k}}{\mathcal{F}_i} = \frac{\hat{\mathcal{F}}}{\mathcal{F}_i}, \quad (\text{V. 7})$$

$$\begin{aligned} \frac{\bar{x}_i}{\sum_{k=1}^m \bar{x}_k} &= \\ &= \frac{1}{1 + \frac{\mathcal{F}_{i+1}}{\hat{\mathcal{F}}} + \frac{\mathcal{F}_{i+1}\mathcal{F}_{i+2}}{\hat{\mathcal{F}}^2} + \dots + \frac{\mathcal{F}_{i+1}\mathcal{F}_{i+2} \dots \mathcal{F}_{i+m-1}}{\hat{\mathcal{F}}^{m-1}}} = \\ &= \frac{1}{1 + \sum_{s=1}^{m-1} \prod_{k=1}^s (\mathcal{F}_{i+k}/\hat{\mathcal{F}}^s)}. \end{aligned} \quad (\text{V. 8})$$

В уравнении (V. 8) индексы меняются циклическим образом, т. е.  $i+k = i+k-m$  для  $i+k > m$

<sup>1</sup> Для  $\mathcal{R}_1 \neq \mathcal{R}_2 \neq \dots \mathcal{R}_m$  соответствующие выражения гораздо более сложны, но их легко получить при помощи рекуррентных формул, начиная с тождества

$$1 = \frac{x_m}{x_1} \cdot \frac{x_1}{x_2} \dots \frac{x_{m-1}}{x_m}$$

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 самоинструктированием или с ком-плементарным инструктированием.*

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.

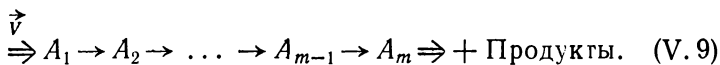
Для селективного самовоспроизведения реакционной сети представляются необходимыми два условия:

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,

$$\text{для } \mathcal{F}_1 = \mathcal{F}_2 \dots \mathcal{F}_m \equiv \mathcal{F}: \quad \lambda = \mathcal{F} \left( \sqrt[m]{1} - 1 \right)$$

$$\text{или для } m=4: \quad \lambda_1 = 0; \quad \lambda_2 = -2\mathcal{F}; \quad \lambda_{3,4} = -\mathcal{F}(1 \pm i).$$

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  $\mathcal{F}_1 = \mathcal{F}_2 \equiv \mathcal{F}$ :

$$\begin{aligned} \lambda(\mathcal{F} + \lambda) &= \mathcal{F}^2, \\ \lambda_1 &= \frac{\mathcal{F}}{2}(\sqrt{5} - 1), \\ \lambda_2 &= -\frac{\mathcal{F}}{2}(\sqrt{5} + 1). \end{aligned} \quad (\text{V. 11})$$

**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

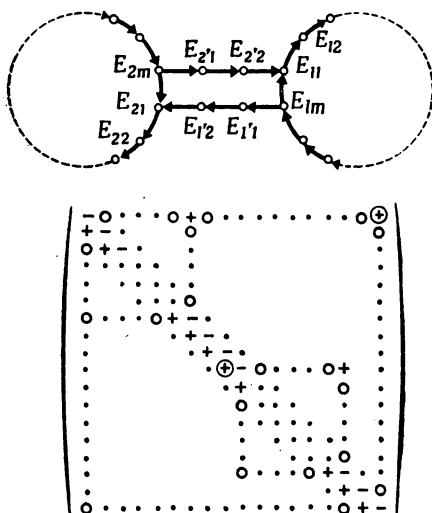


Рис. 14. Представление в виде графа циклов, связанных петель обратной связи.

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  $\hat{\mathcal{F}}_i > \hat{\mathcal{R}}_i$ , где символ  $\wedge$  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 большим 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.

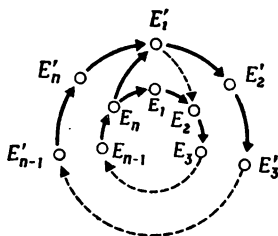


Рис. 15. Воспроизведение мутантов в каталитическом цикле.

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 специфических мутаций  $E'_1 \rightarrow E'_2 \dots E'_n \rightarrow E'_1$ , т. е. должен образоваться новый специфический 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.

## ГЛАВА VI

# 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 воспроизводящегося — 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  $E_i$ . Кругок вокруг  $I_i$  условно представляет способность нуклеотидного коллектива к комплементарному инст-руктированию, а каждый такой коллектив состоит из

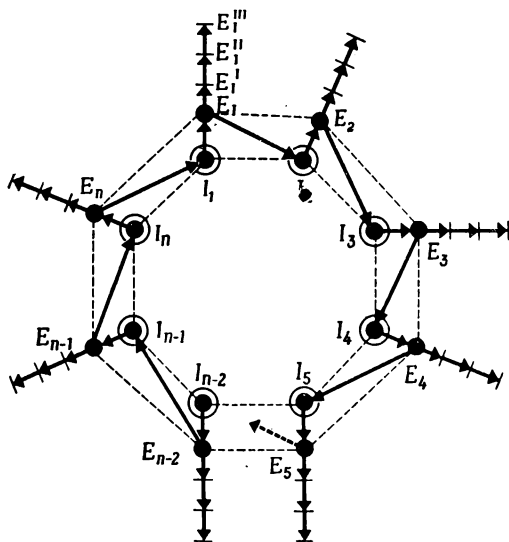


Рис. 16. Самоинструктирующий каталитический гиперцикл.

$I_i$  — 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 специфического узнавания). Контур, 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 $\beta$ -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  $\nu$  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 компартментализованные 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-

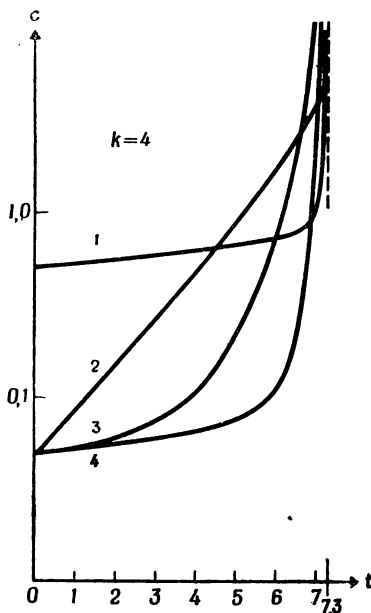


Рис. 17. Рост четырехчленного гиперцикла (4 носителя информации кодируют 4 фермента сопряжения  $E_i$ ).

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

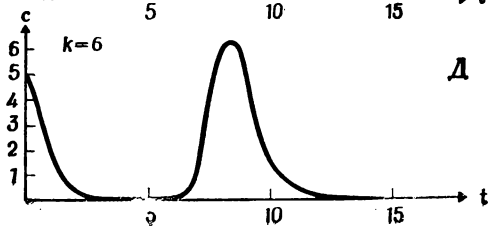
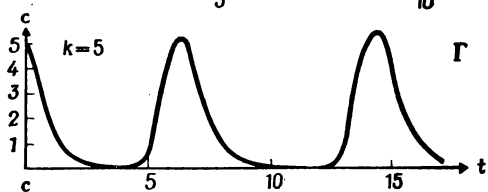
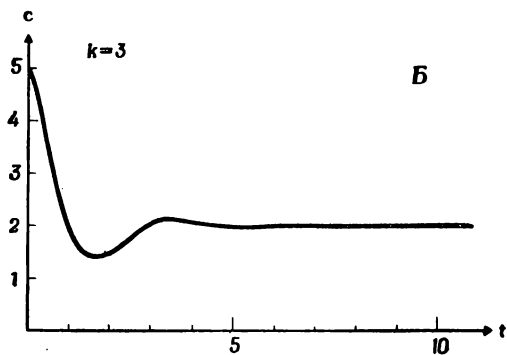
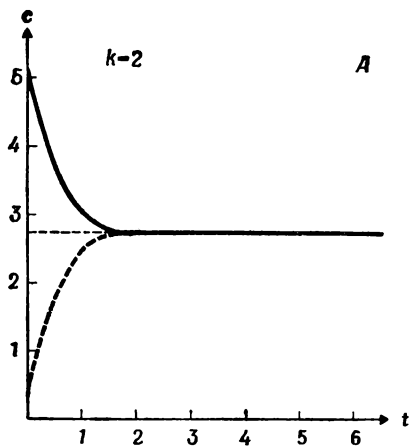
$\mathcal{F}_i + \mathcal{F}'_i y_{i-1} > \mathcal{R}_i$ ). Это иллюстрирует рис. 17. Интересно отметить, что для циклов с тремя и более членами могут возникнуть колебания. Эти колебания еще более выражены при наличии налагаемых отбором ограничений (см. рис. 18). Двучленные циклы не дают колебаний; при 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.

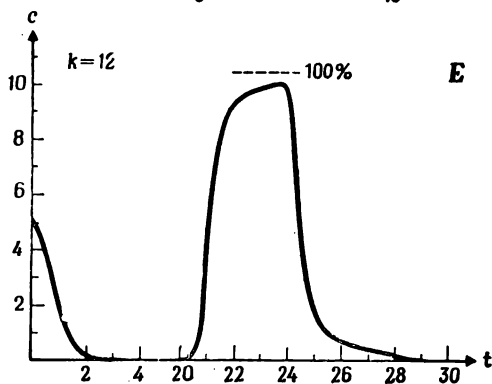
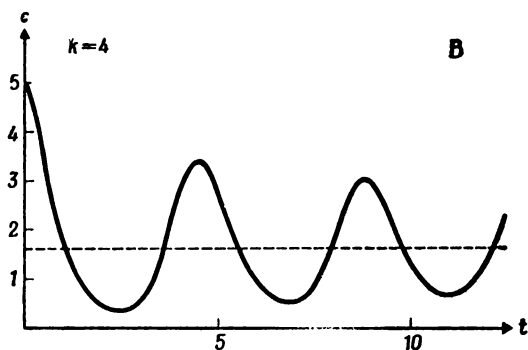




**Рис. 18.** Решения, описывающие отбор  $k$ -членных гиперциклов при ограничении постоянной общей организации ( $k = 2, 3, 4, 5, 6, 12$ ).

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, отправляясь от постоянного распределения, при  $t=0$  создать малое возмущение. При этом устанавливается колебательный режим. (Вычисления П. Шустера [100].)



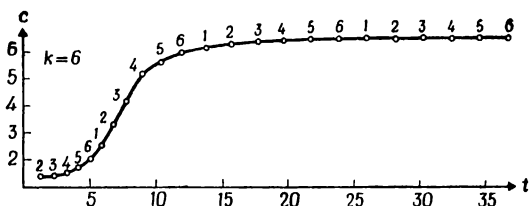


Рис. 19. Изменение амплитуды колебаний в гиперцикле [100].  
Колебания возникают из постоянного стационарного распределения при  $t=0$ ;  $x_1^0=1,35$ ;  $x_k^0 > 1=1,25$ .

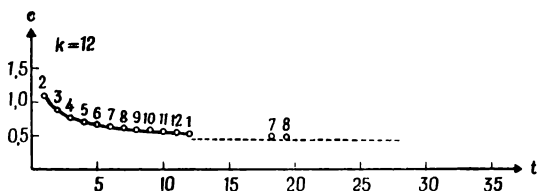


Рис. 20. Снижение амплитуды в линейном каталитическом цикле (см. гл. V) для сравнения с рис. 19.  
При наложении ограничений, обусловленных отбором, концентрации всегда приходят к постоянному стационарному уровню (см. [100]).

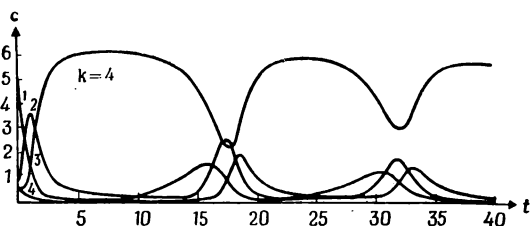


Рис. 21. Стационарные колебания в уравновешенном четырехчленном гиперцикле с несимметричным распределением скоростей [100].  
Константа скорости образования 4-го члена в 10 раз меньше, чем для всех других членов гиперцикла.

Более полное решение должно учитывать следующие усложнения:

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 образом<sup>1</sup>:

$$\begin{aligned} \dot{y}_i &= (\mathcal{F}'_{I_i} - \mathcal{D}_{I_i}) y_i + \mathcal{F}''_{I_i} \cdot z_i, \\ \dot{x}_i &= \mathcal{F}_{E_i} \cdot y_i - \mathcal{D}_{E_i} \cdot x_i, \end{aligned} \quad (\text{VI.5})$$

где  $z_i$  дается уравнением (VI.3).

<sup>1</sup> 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  $z_i$ .

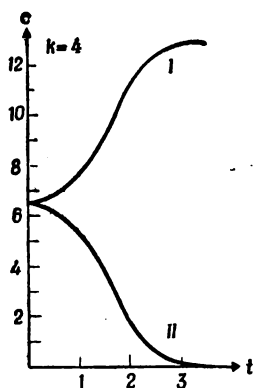


Рис. 22. Конкуренция между двумя 4-членными гиперциклами, у одного из которых константы скорости образования на 10% меньше, чем у другого:  $\mathcal{F}'_{11} = \mathcal{F}'_{12} = \mathcal{F}'_{13} = \mathcal{F}'_{14} = 1,0$ ;  $\mathcal{F}'_{21} = \mathcal{F}'_{22} = \mathcal{F}'_{23} = \mathcal{F}'_{24} = 0,9$  (ср. [100]).

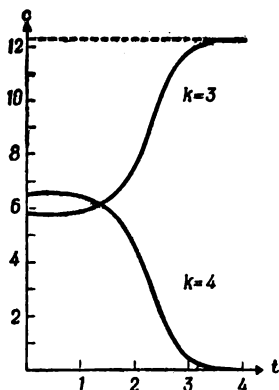


Рис. 23. Конкуренция между трех- и четырехчленным циклами, имеющими одинаковые индивидуальные кинетические параметры, но различающимися по начальной концентрациям [100]. Трехчленный цикл «побеждает».

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.

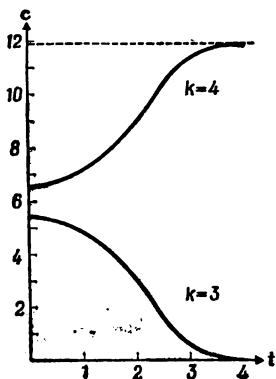


Рис. 24. Конкуренция между трех- и четырехчленным циклами, которые имеют одинаковые индивидуальные кинетические параметры, но различаются по начальным концентрациям [100].

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 имела 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  $I_h$  в коллектив, имеющий относительно большую информационную емкость. Отдельная

единица может кодировать только две функции: *узнавание* — для того чтобы обеспечить циклическую связь с помощью системы  $E_k$  — 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 воспроизводящуюся 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.

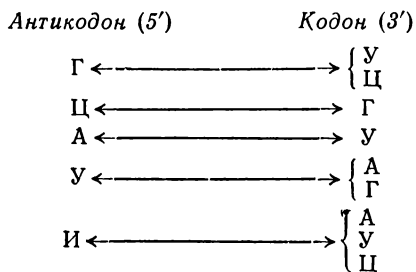
Не содержит ли современная таблица генетического кода (табл. 2) какого-либо намека на свое происхождение? Эта проблема тщательно анализировалась рядом авторов, и в частности в монографии А. Вёзе [107]. Можно напомнить следующие важные факты:

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 кодовами 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 специфичное 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 специфичные взаимодействия ( $q$ , близкое к единице), но очень сомнительно, существуют ли вообще достаточно сильные взаимодействия, обеспечивающие прямое соответствие между антикодонами и аминокислотами<sup>1</sup>. Таким образом, встает законный вопрос:

<sup>1</sup> 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 priori 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 аминоацилсинтетаз) — 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  $\lambda$  единиц, т. е. аминокислоты  $a, b, c, \dots$ , которые нужно поставить в соответствие кодонам  $A, B, C, \dots$  с помощью адапторов  $A', B', C'$  (комплементарных  $A, B, C \dots$ ). Имеется всего  $\lambda^2$  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 объема <sup>1</sup>, в котором имеется  $\lambda$  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  $\lambda^2$  элементов по  $\lambda$ :

$$V_r = \binom{\lambda^2 + \lambda - 1}{\lambda}.$$

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-

---

<sup>1</sup> Величина этого элемента объема подобрана исходя из условия, что в нем имеется  $\lambda$  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 главы<sup>1</sup>.

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  $\lambda$  специфичных «координаторов», однако теперь не будет вырожденности  $\lambda!$ , потому что система должна усиливать одно из  $\lambda$  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  $\lambda^2$  элементов по  $\lambda$ .

---

<sup>1</sup> 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:

$\lambda$	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). Восьмибуквенный 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 специфичного (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.*

## ГЛАВА VII

### 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]).

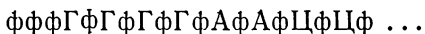
#### § VII.1. Система Q $\beta$ -репликазы

История Q $\beta$  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 $\beta$  использует специфичный фермент репликации, опознающий только РНК фага Q $\beta$ . 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 $\beta$  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 вмешаться 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 $\beta$ -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*-HCl, pH 7,4,  $2 \cdot 10^{-2}$  М MgCl<sub>2</sub>,  $3 \cdot 10^{-3}$  М ЭДТА, по 200 ммкмоль АТФ, УТФ, ГТФ, ЦТФ (УТФ помечен <sup>32</sup>P в  $\alpha$ -положении, причем 4000 расп/мин соответствует 1 мкг синтезированной РНК) и 40 мкг Q $\beta$ -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

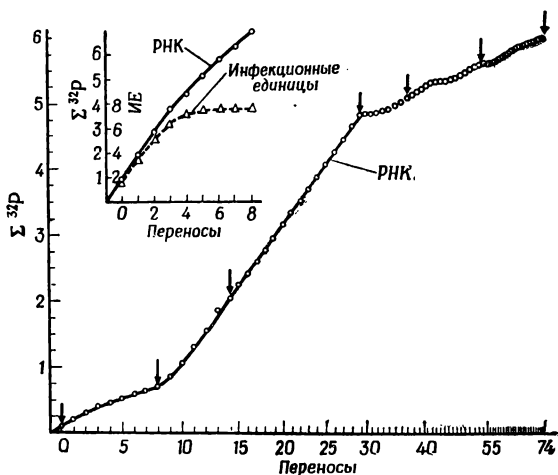


Рис. 25. Эксперимент с серийными переносами РНК фага Q $\beta$  [121].

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 затравки 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.

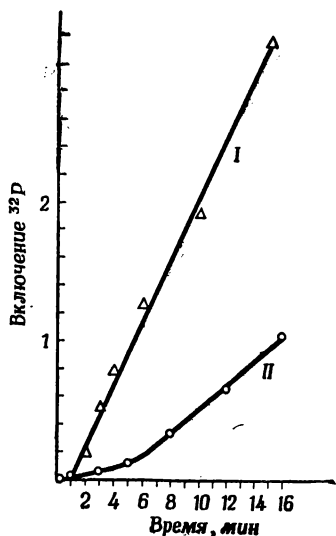
Анализ продуктов дал следующую информацию о реплицированных молекулах РНК (см. рис. 25): инфекционность была утрачена после 4-го переноса; молекулярные массы РНК-матриц по мере переносов более или менее постоянно снижались, пока после 74-го переноса не получился конечный продукт, утративший около 83% своей исходной генетической информации. Из 3600 нуклеотидных остатков родительской молекулы сохранилось только 550. Параллельно с уменьшением молекулярного веса наблюдается увеличение скорости включения  $^{32}\text{P}$ , 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 возрасть 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-

шение 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 $\beta$  (кривая II) [121].

Иницированы две стандартные реакции (см. текст) в 0,25 мл смеси при 35 °C. В одной реакции затравкой служила одноцепочечная «вариантная» РНК (74-й перенос), очищенная гель-фильтрацией, в другой — РНК мутанта ts фага Q $\beta$  (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  $^{32}\text{P}$ ) 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:

$$\begin{aligned}\dot{x}_{I_+} &= \mathcal{F}_+ x_{EI_-}, \\ \dot{x}_{I_-} &= \mathcal{F}_- x_{EI_+},\end{aligned}\tag{VII. 1}$$

где  $EI_+$  и  $EI_-$  — комплексы фермент — матрица. Процедура, аналогичная выводу уравнения Михаэлиса—Ментен, приводит к выражению

$$x_{EI_+} + x_{EI_-} = x_{E_0} \frac{K_+ x_{I_+} + K_- x_{I_-}}{1 + K_+ x_{I_+} + K_- x_{I_-}}\tag{VII. 2}$$

( $E_0$  — 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  $^{32}\text{P}$  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 $\beta$ ). 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 определено означает увеличение средней скорости копирования отдельных единиц) означает 15-кратное увеличение скорости воспроизведения индивидуальных «вариантных» молекул РНК (по сравнению с исходной РНК фага Q $\beta$ ). 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  $^{32}\text{P}$  — 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:  $x_{I_i} \geq K^{-1} + \sum_{k \neq i} x_{I_k}$ <sup>1</sup>.

<sup>1</sup> 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 $\beta$ , 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 $\beta$ -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  $^{32}\text{P}$  (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.

## ГЛАВА VIII

### 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 «инструкирование»). 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 реакционную систему при постоянных внутренних условиях. В таком стационарном состоянии доминирует «отобранная» последовательность<sub>i</sub> (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

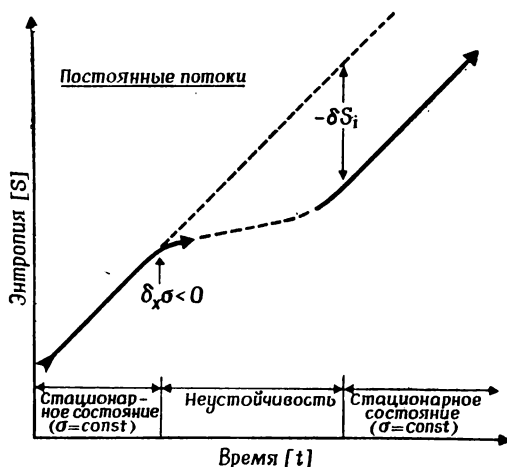


Рис. 27. Зависимость энтропии от времени для процесса отбора при постоянных общих потоках единиц.

Возникновение мутанта с селективным преимуществом ( $W^F$  мутанта  $> W^F$  главной копии) соответствует отрицательной флуктуации 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  $\sigma$  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 сродством, 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 сродством 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 сродство 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  $\mathcal{A}_i$ , время жизни  $1/\mathcal{D}_i$  и параметр точности  $\mathcal{Q}_i$  достигают наибольшего возможного значения. Для сохранения информации отобранного состояния требуется некая минимальная пороговая точность ( $\mathcal{Q}_{\min}$ ), 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  $\mathcal{Q}$ , но, поскольку она основана на отборе, значение  $\mathcal{Q}$  должно превышать порог  $\mathcal{Q}_{\min}$ . Изменчивость уменьшается с ростом  $\mathcal{Q}$ , и это будет всегда благоприятствовать эволюции тех систем, для которых  $\mathcal{Q}$  близка к  $\mathcal{Q}_{\min}$ .

*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:

$$W_{m1} < W_{m2} < W_{m3} < \dots < W_{\text{opt}}.$$

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

$$W_{\text{opt}} < W_{\text{max}},$$

характеризующее конечную разность, которую Жак Моно [125] назвал бы различием между тем, что «есть», и тем, что «должно быть».  $W_{\text{max}}$  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 эволюции<sup>1</sup>.

---

<sup>1</sup> 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. Левантином [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 аналогии<sup>1</sup> 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

<sup>1</sup> 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, нереалистический 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  $\mathcal{F}_i = D_i$ , 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 отбора<sup>1</sup>.

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).

---

<sup>1</sup> 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 зарождаться 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 потребуется, 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»,

## ГЛАВА IX

### 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 предположениях 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 называть 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 «потoki» 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 пред-  
рассудок, 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 реализовываться 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 самоинструктированию, 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 $\beta$ , 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., Merril S. H., Penswick J. R., Zamir A.*, Science, **147**, 1462 (1965). (Р. Холли, Д. Аппар, Г. Эверетт, Д. Медисон, М. Маркисси, 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. Marois), 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. Барплетт*, 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., Hormann 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. Моргенштерн, 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.)
128. *Lewontin R. C.*, J. Theoret. Biol., **1**, 382 (1961).
129. *Warburton F. E.*, J. Theoret. Biol., **16**, 78 (1967).
130. *Kimura M.*, Genet. Res. Camb., **2**, 127 (1961).
131. *MacKay D. M.*, in: Information, Mechanism and Meaning, chapter 12, pg. 132. Boston, M. I. T. Press 1969.
132. *Waddington C. H.*, The Nature of Life, London, Unwin Books, 1961.
133. *Stent G.*, The Coming of the Golden Age. Garden City N. Y., The Natural History Press, 1969.
134. *Kimura M.*, Nature, **217**, 624 (1968).
135. *King J. L., Jukes T. H.*, Science, **164**, 788 (1969).
136. *Опарин А. И.*, Жизнь, ее природа, происхождение и развитие, М., 1960. (Имеется 2-е дополн. изд-е, изд-во «Наука», М., 1968).
137. *Von Neumann J.*, Theory of self-reproducing automata, Urbana, Univ. Illinois Press, 1966, Дж. фон Нейман, Теория самовоспроизводящихся автоматов, изд-во «Мир», М., 1971).
138. *Wittgenstein L.*, Tractatus Logico-Philosophicus, 6. 521, London, Routledge and Kegan Paul, 1922.

## О Г Л А В Л Е Н И Е

Предисловие к русскому изданию . . . . .	5
Список литературы . . . . .	9
<b>Глава I. Введение . . . . .</b>	<b>10</b>
§ I.1. «Причина и следствие» . . . . .	10
§ I.2. Предпосылки самоорганизации . . . . .	13
<b>Глава II. Феноменологическая теория отбора . . . . .</b>	<b>36</b>
§ 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
<b>Глава III. Стохастический подход к отбору . . . . .</b>	<b>75</b>
§ 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
<b>Глава IV. Самоорганизация, основанная на комплементарном узнавании: нуклеиновые кислоты . . . . .</b>	<b>96</b>
§ IV.1. Истинное «самоинструктирование» . . . . .	96
§ IV.2. Комплементарное инструктирование и отбор (теория) . . . . .	102
§ IV.3. Комплементарное узнавание оснований (экспериментальные данные) . . . . .	108
<b>Глава V. Самоорганизация через циклический катализ: белки . . . . .</b>	<b>124</b>
§ 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. Требование 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 $\beta$ -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  
84×108<sup>1</sup>/<sub>32</sub> = 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

**ВЫЙДЕТ В СВЕТ В 1973 г.**

**Руттен М. ПРОИСХОЖДЕНИЕ ЖИЗНИ.**  
Перевод с английского, 25 изд. л.

В книге рассмотрены современные представления о возможности возникновения жизни на Земле и в других областях Вселенной путем постепенных превращений неживой природы, без вмешательства божественных и других сверхъестественных сил. Большая ценность этой книги по сравнению со всеми другими на эту тему заключается в том, что автор, шаг за шагом описывая те происходившие на Земле с самых ранних периодов ее геологической истории события, на фоне которых могла возникнуть жизнь, привлекает данные буквально из всех областей естествознания. Весь этот огромный материал изложен последовательно, логично и доступно, что делает книгу совершенно незаменимой для специалистов по самым разным дисциплинам.

Предназначена для биологов всех специальностей, для геологов, палеонтологов, физиков, химиков, астрономов, философов, историков науки, для преподавателей высшей и средней школы, для лекторов (в частности, может быть широко использована для пропаганды научного атеизма).

