

весия любой катализатор (фермент) в равной мере ускоряет и прямую и обратную реакции. Рассмотрим ту же автокаталитическую реакцию вдали от равновесия. Если $k_- [C] \ll k_+ [B]$, то

$$v = k_+ [B][C] \quad (1,50)$$

и при $[B] = \text{const}$

$$\delta v = k_+ [B] \delta [C], \quad (1,51)$$

т. е. по-прежнему

$$\delta A = -RT \frac{\delta [C]}{[C]}. \quad (1,48)$$

Имеем

$$T \delta_x \sigma = \delta A \delta v = -RT k_+ \frac{[B]}{[C]} (\delta [C])^2 \leq 0. \quad (1,52)$$

Стационарное состояние неустойчиво в случае автокаталитической реакции. Показано, что такие нестабильности лежат в основе роста и эволюции. Пригожин и его сотрудники установили, что сочетание автокаталитических реакций и процессов транспорта может привести в условиях, далеких от равновесия, к образованию упорядоченных *диссипативных* структур, определяемых диссипацией энергии, а не консервативными молекулярными силами [21, 26]. Эйген построил на этой основе молекулярную теорию самоорганизации биополимеров, их естественного отбора и эволюции [13].

§ 1.4. ТЕОРИЯ ИНФОРМАЦИИ И БИОЛОГИЯ

Феноменологическое рассмотрение живых систем, естественно, пользуется понятиями теории информации. Существование и развитие клетки и организма, являющихся открытыми и саморегулируемыми системами, требует генерации и передачи информации по каналам прямой и обратной связи.

Теория связи применяет количественную меру уже имеющейся информации, т. е. меру количества сведений, содержащихся в некотором сообщении. Такая теория исследует способы кодирования сообщения, использования избыточности (см. далее), оптимизации, преодоления «шумов» и т. д.

Количество информации при выборе P_1 событий из P_0 равновероятных событий определяется как

$$I = K \log P_0/P_1, \quad (1,53)$$

где K — численный коэффициент. Если $P_1 = 1$, то

$$I = K \log P_0. \quad (1,54)$$

При таком определении получается аддитивность информации при перемножении независимых вероятностей. Так, результат бросания одной игральной кости дает количество информа-

ции $K \log 6$, результат бросания двух костей дает вдвое большее количество информации $2K \log 6 = K \log 36$. Для количественных расчетов нужно определить основание логарифмов и значение K . Пользуясь основанием 2 и принимая $K = 1$, имеем

$$I = \log_2 P_0. \quad (1,55)$$

Такое определение количества информации отвечает задачам с некоторым числом n различных независимых выборов, каждый из которых сводится к двоичной альтернативе. При этом $P_0 = 2^n$ и, согласно (1,55),

$$I = \log_2 2^n = n \text{ бит.}$$

Единицей информации служит один *бит*, отвечающий одному выбору из двух равновероятных возможностей (т. е. $P_0 = 2$). При бросании кости $I = \log_2 6 = 2,585$ бит.

Сколько бит информации содержит произвольное трехзначное число? Первая цифра имеет 9 различных значений — от 1 до 9, вторая и третья — по 10 значений — от 0 до 9. Имеем

$$I = \log_2 9 + 2 \log_2 10 = 2,64 + 2 \cdot 3,32 = 9,28 \text{ бит.}$$

Принятое определение информации соответствует двоичной системе исчисления, в которой любое число записывается в виде степеней числа 2 посредством цифр 0 и 1. Одна десятичная единица дает 3,32 бит, т. е. двоичная запись числа требует в среднем в 3,32 раза больше цифр, чем десятичная.

Если исходные события не равновероятны и распределение вероятностей задано, то количество информации при выборе одного или нескольких событий меньше, чем в случае равновероятности. Задание распределения вероятностей означает наличие некоторой предварительной информации о получаемом сообщении. Эта предварительная информация вычитается из получаемой.

Допустим, что имеется сообщение, содержащее G последовательных ячеек — скажем, текст из G последовательных букв. В каждой из G ячеек может находиться одна из M букв (для русского языка $M = 32$). В сообщении содержится N_1 букв А, N_2 букв Б и т. д. Имеем

$$G = \sum_{j=1}^{I=M} N_j.$$

Вероятность появления данной буквы равна

$$p_j = N_j/G,$$

причем

$$\sum_{j=1}^M p_j = 1.$$

Общее число P сообщений, т. е. последовательностей из G букв M -буквенного языка, равно

$$P = \frac{G^M}{\prod_{j=1}^M N_j!}.$$

Вводя натуральные логарифмы, получаем информацию, содержащуюся в одном сообщении:

$$\begin{aligned} I &= K \ln P = K \left[\ln(G^M) - \sum_{j=1}^M \ln(N_j!) \right] \approx \\ &\approx K \left(G \ln G - \sum_{j=1}^M N_j \ln N_j \right). \end{aligned}$$

Здесь предполагается, что $G, N_j \gg 1$ и применима теорема Стирлинга. Следовательно,

$$I \approx -KG \sum_{j=1}^M \frac{N_j}{G} \ln \frac{N_j}{G} = -KG \sum_{j=1}^M p_j \ln p_j. \quad (1,56)$$

Мы получили формулу Шеннона. Переходя к двоичному определению информации, примем, что

$$K = (\ln 2)^{-1}.$$

Тогда

$$I = -G \sum_{j=1}^M p_j \log_2 p_j.$$

Количество информации, приходящейся на одну букву, равно

$$i = I/G = - \sum_{j=1}^M p_j \log_2 p_j.$$

В действительности настоящий языковой текст содержит меньшее количество информации, так как в нем имеются добавочные ограничения. Вероятности последовательного появления букв взаимосвязаны — они образуют *цепь Маркова*, т. е. вероятность появления данной буквы зависит от того, какие буквы ей предшествовали. Так, в русском языке появление гласной вслед за гласной гораздо менее вероятно, чем после согласной. Любая корреляция событий уменьшает количество информации, содержащееся в сообщении о них.

Все изложенное существенно для теории и практики связи. Но теория информации имеет и прямое физическое содержание. Подобно количеству информации, энтропия выражается

как логарифм термодинамической вероятности — числа микросостояний, реализующих данное макросостояние, т. е.

$$S = k \ln W, \quad (1,57)$$

где $k = 1,38 \cdot 10^{-16}$ эрг·град⁻¹ — постоянная Больцмана.

Сопоставим формулы (1,54) и (1,57). Представление количества информации в виде

$$I = k \ln P \quad (1,58)$$

позволяет выражать информацию в энтропийных единицах. Имеет ли, однако, это реальный смысл? Уменьшение числа микросостояний системы означает увеличение количества информации о системе и одновременно уменьшение энтропии. Таким образом, информация эквивалентна негэнтропии — энтропии со знаком минус. Пусть число микросостояний убывает от P_0 до P_1 . Количество получаемой при этом информации составляет

$$I = k \ln (P_0/P_1) = S_0 - S_1,$$

т. е. I равно уменьшению энтропии или увеличению негэнтропии. Негэнтропийное толкование информации было предложено Сциллардом [27].

При смешении N_1 молекул газа 1 и N_2 молекул газа 2 энтропия образовавшейся смеси больше энтропии разделенных газов S_0 на величину энтропии смешения

$$\Delta S = S - S_0 = -kN(p_1 \ln p_1 + p_2 \ln p_2) > 0,$$

где $N = N_1 + N_2$, $p_1 = N_1/(N_1 + N_2)$, $p_2 = N_2/(N_1 + N_2)$. Это равносильно уменьшению количества информации на ту же величину. Значит, исходная информация составляла

$$I_0 = -kN(p_1 \ln p_1 + p_2 \ln p_2).$$

Полученное выражение совпадает с формулой Шеннона. Один бит информации эквивалентен $k \ln 2$, т. е. $\sim 10^{-16}$ эрг·град⁻¹ — весьма малой величине.

В таком определении нет ничего субъективного. Получение информации о физической системе требует реальных физических воздействий на нее. Нельзя получить информацию о событиях в адиабатически изолированной системе. Любкой опыт, дающий информацию о некоторой системе, приводит к увеличению энтропии системы или ее окружения, причем, в соответствии со вторым началом, это увеличение энтропии всегда превышает количество полученной информации. Информация оплачивается энтропией, так как наблюдение связано с увеличением энтропии, с необратимым процессом.

Жидкость обладает большей энтропией, чем кристалл, так как ее состояние может быть реализовано большим числом спо-

собою распределения молекул, чем состояние кристалла. При кристаллизации жидкости мы получаем информацию о расположении молекул в узлах кристаллической решетки. Жидкость нужно охладить, что невозможно в адиабатическом процессе. Для ее охлаждения необходим холодильник, температура и энтропия которого будут возрастать. Возрастание энтропии холодильника превысит получаемую информацию о кристалле, выраженную в энтропийных единицах.

Анализируя второе начало термодинамики, Максвелл предложил фантастический мысленный опыт. Представим себе сосуд с газом, разделенный на две части перегородкой с дверцей. Дверцей управляет «демон», находящийся в одной половине сосуда. Он пропускает в свою половину быстрые молекулы и выпускает из нее медленные. В результате возникает разность температур между двумя половинами сосуда, что противоречит второму началу термодинамики, так как демон не затрачивает никакой энергии.

Бриллюэн рассмотрел этот парадокс с позиций теории информации [28]. Для того чтобы увидеть молекулу, демон должен ее осветить, т. е. располагать источником излучения, не находящимся в равновесии с окружающей средой. Такой источник имеет меньшую энтропию, чем среда, т. е. обладает негэнтропией, которую демон превращает в информацию. Полученная информация используется для понижения энтропии газа, т. е. для увеличения негэнтропии. Расчет полного изменения энтропии показывает, что энтропия системы в целом (газ и источник света) не уменьшится, а возрастет, и, следовательно, второе начало справедливо.

Как уже говорилось (стр. 16), Эльзассер утверждал, что развитие многоклеточного организма из зиготы несовместимо с законами физики. Он считал возрастание запаса информации при морфогенезе, при росте организма, «биотонным». Эти представления об эпигенезе в развитии противостоят концепции преформизма, согласно которой вся информация о будущем организме уже заключена в зиготе.

Оценить количество информации, содержащееся в клетке и в организме, трудно. Можно считать, что информация (т. е. упорядоченность), обеспечивающая биологическую функциональность, полностью определяется атомной структурой [29, 30]. Данков и Каствлер оценили количество информации, требуемое для отбора одного атома из всего набора атомов в клетке и задания его положения с точностью, определяемой тепловыми колебаниями при температуре тела [30]. По их оценке на один атом приходится 24,5 бит, на одну клетку *Esherichia coli* — порядка 10^{10} бит, на одну зиготу млекопитающего — порядка 10^{15} бит, на человеческий организм — порядка 10^{25} бит (т. е.

10^9 эрг/град ~ 24 кал/град в энтропийной мере). Если считать эти расчеты имеющими реальный смысл, то, несомненно, $I_2 \equiv I_{\text{орг}} \gg I_{\text{зиг}} \equiv I_1$. Эльзассер считает недоступным физическому пониманию возрастание количества информации с I_1 до I_2 , т. е. соответствующую убыль энтропии. Однако в действительности здесь нет трудности.

Изменение количества информации в открытой системе можно представить в виде, подобном выражению (1,16),

$$\Delta I = \Delta_i I + \Delta_e I.$$

Член $\Delta_i I < 0$ выражает уменьшение I вследствие внутренних процессов, для которых $\Delta_i S > 0$. В развивающемся организме $\Delta_e I > 0$. Имеем

$$I_2 = I_1 + \Delta_e I + \Delta_i I.$$

Так как $I_2 \gg I_1$, очевидно, что $\Delta_e I \gg I_1 - |\Delta_i I|$. Информация черпается открытой системой из окружающей среды, энтропия которой возрастает. Организм, растущий из зиготы, подобен в этом смысле кристаллу, растущему из зародыша, помещенного в насыщенный раствор. В обоих случаях возрастание упорядоченности, возрастание количества информации, перекрывается увеличением энтропии окружающей среды — холодильника при кристаллизации. Концепция Эльзассера виталистична; неявным образом предполагается несоблюдение второго начала в живой природе. Равен, подвергший критике идеи Эльзассера, считает, что во всех клетках организма содержится одна и та же генетическая информация. Развитие организма означает не увеличение количества информации, но увеличение избыточности, т. е. многократное ее повторение [31]. Равен трактует зиготу как канал связи, причем родительские организмы служат источником информации, а вырастающий организм — ее приемником. Развитие сводится к декодированию информации. Равен исходит из возможности абсолютной оценки количества информации в зиготе и организме. В действительности, как показал Аптер [32] (см. также [33, 57]), такая оценка всегда относительна и условна. Тождество генов в клетках организма не означает избыточности. Развитие есть результат взаимодействия различных частей эмбриона, информация содержится не только в хромосомах, но и во всех внутри- и межклеточных взаимоотношениях. Концепции преформизма и эпигенеза в сущности непригодны для описания развития, которое нельзя свести к увеличению или сохранению количества информации. Задача состоит не в таком описании, но в выяснении сущности развития, его физической природы, его атомно-молекулярных основ.

Информационные подходы, оперирующие понятием количества информации, существенны в теоретической биологии и био-

физике не потому, что они решают контroversу преформизм—эпигенез, в сущности лишенную содержания. Их смысл в другом. Во-первых, перевод биологических концепций на язык теории информации позволяет сделать эти концепции более ясными и строгими и вложить в них новое содержание. Такую работу провел Шмальгаузен, раскрывший некоторые информационные аспекты эволюционной теории [34, 35]. Во-вторых, исследование клетки, организма как сложной саморегулируемой системы требует информационного подхода, поскольку управление, регуляция, осуществляется посредством прямых и обратных информационных потоков. Задача биофизики в этой области состоит в выяснении материальной природы запасаания, кодирования, передачи и рецепции информации в живых системах. Проблема генетического кода (см. ниже гл. 9) имеет ясный информационный смысл. Вопросы о количестве генетической информации, кодируемой в молекулах ДНК, о помехоустойчивости кода, о механизмах реализации информационной программы вполне содержательны.

Информация в организме передается молекулами и ионами, она имеет химическую природу. Рецепция информации сводится к узнаванию молекул молекулами, она определяется их взаимодействиями. Именно в таком смысле организм есть химическая машина, и в этом его отличие от «роботов», созданных руками человека. В современных кибернетических устройствах сигналы имеют электрическую или магнитную природу, в организме сигналами служат молекулы.

Кастлер ввел в рассмотрение молекулярной информации принцип сигнатуры [29]. Взаимодействие молекул означает передачу информации. При этом обычно передается не вся информация, содержащаяся в молекуле, но лишь некоторая ее часть, именуемая *сигнатурой*. Сигнатурой молекулы служат те ее особенности, благодаря которым она участвует в данной реакции. Понятия однозначности и многозначности молекулярных сигнатур, надежности их действия имеют реальный физический смысл.

Всякая передача информации не идеальна — в ходе передачи происходит частичное разрушение информации вследствие помех или ошибок. Такие помехи или ошибки называются *шумами*. Рассмотрим передачу генетической информации (см. [36]).

Количество информации в генетическом сообщении равно

$$I \leq I_0 - I_n,$$

где I_0 — исходное количество информации, I_n — «шумовые» потери вследствие ошибок в процессе удвоения хромосом в зародышевой или соматической клетках, а также ошибок в рецепции генетической информации, т. е. в синтезе белка (см. ниже

гл. 9). I_n можно выразить через вероятность прочтения или перехода от «буквы» сорта i к «букве» сорта j (кодирование), которую мы обозначим через $p_i(j)$. Вероятность появления «буквы» i обозначим через $p(i)$. Тогда

$$I = I_0 + \sum_{i,j} p(i) p_i(j) \log_2 p_i(j).$$

Допустим, что на генетическую систему оказывается некоторое воздействие, характеризуемое мерой λ . Например, λ есть доза ионизирующего излучения. Имеем

$$\frac{dI}{d\lambda} = \log_2 e \sum_{i,j} \left\{ p(i) \frac{dp_i(j)}{d\lambda} + p(i) \ln p_i(j) \frac{dp_i(j)}{d\lambda} + p_i(j) \ln p_i(j) \frac{dp_i(j)}{d\lambda} \right\}.$$

Производная $dI/d\lambda$ возрастает до бесконечности за счёт второго члена в правой части уравнения, если $p(i)$ и $dp_i(j)/d\lambda$ отличны от нуля, а одна из вероятностей $p_i(j)$ стремится к нулю. Так как

$$\sum_j p_i(j) = 1,$$

то в идеальном случае, когда одна из величин $p_i(j) = 1$, остальные равны нулю. Отсюда следует $dI/d\lambda \rightarrow \infty$, и система оказывается неустойчивой. Это означает, что неустойчива система, состоящая из совокупности организмов, содержащих тождественные гены. В реальной совокупности должен быть шум, т. е. генетические сообщения в разных изогенных организмах не тождественны друг другу. В них должны возникать мутации. Согласно принципу Данкова [30], оптимум, достигаемый естественным отбором, вовсе не означает создания идеальной системы без шумов. При передаче сообщения борьба с шумами производится посредством избыточной информации. Так, терпящий бедствие корабль многократно передает сигнал SOS, чтобы увеличить вероятность его рецепции. Принцип Данкова утверждает, что оптимальное количество избыточной информации в эволюции вида и особи равно не тому количеству, которое уничижает все ошибки, но тому, которое минимизирует сумму стоимости ошибок и стоимости их исправления. Допускается столько ошибок, сколько совместимо с дальнейшим существованием, и используется минимальное количество избыточной информации. Тем самым достигается некое стационарное состояние. Эти соображения можно, очевидно, сформулировать на языке термодинамики, перейдя к энтропийной мере информации.

Сказанным, в общем, исчерпываются возможности теории информации, оперирующей только понятием об ее количестве. Очевидно, из адекватности количества информации и негэнтро-

пии следует, что при объяснении биологических процессов понятие количества информации не может дать большего, чем понятие энтропии. Здесь не рассматривается генерация информации и ее инструктирующее значение, определяющее ту или иную биологическую функцию носителя информации. При исследовании развивающейся и эволюционирующей системы необходимо ввести понятие *ценности информации* для реализации конкретного процесса, эквивалентное ее программирующему, инструктирующему, значению. Ценность информации выражает ее содержание, тогда как количество информации не имеет отношения к ее содержанию. Содержание можно оценить лишь применительно к определенным физическим процессам.

Эйген впервые предпринял попытку дать количественное физическое определение ценности информации, содержащейся в биологически функциональных макромолекулах в процессе их естественного отбора [13]. *Селективная ценность* по Эйгену есть безразмерный кинетический параметр, характеризующий автокаталитический синтез макромолекулярных цепей, который происходит за счет цепей с меньшей ценностью, подвергающихся деструкции. Процесс протекает в стационарной открытой системе, далекой от равновесия. Теория Эйгена согласуется с теорией возникновения диссипативных структур Пригожина и Гландорфа и описывает некий начальный этап добиологической эволюции.

Теория Эйгена, имеющая важное значение для биофизики, основывается на реальных свойствах биологических макромолекул. Изложение, анализ, уточнение и дальнейшее развитие этой теории возможны лишь после детального ознакомления с биологическими макромолекулами. Обсуждение теории Эйгена выходит за рамки этой книги.

§ 1.5. КООПЕРАТИВНОСТЬ

Построение физики живых систем требует исследования своеобразных микроскопических, атомно-молекулярных свойств. Клетка, организм — гетерогенные конденсированные системы, построенные из квазикристаллических тел (надмолекулярные структуры) и жидкостей, из больших и малых молекул. Надмолекулярные структуры в организме высокоупорядочены и представляют собой преимущественно линейные и двумерные системы. В качестве линейных систем укажем на миофибриллу мышцы, аксон нервной клетки, условно двумерных — различные клеточные и внутриклеточные мембраны, β -формы белковых структур. Говоря о квазикристалличности такого рода структур, мы имеем в виду именно их высокую упорядоченность, выражающуюся в ряде случаев в периодичности строения. Вместе