Глава 9

НЕРАВНОВЕСНАЯ ТЕРМОДИНАМИКА В БИОЛОГИИ

§ 9.1. Информация и энтропия

Изучая структуру и свойства биополимеров, мы до сих пор, как правило, ограничивались рассмотрением равновесных состояний. Пока речь идет о готовой структуре, это законно. Но возникновение структуры, т. е. биосинтез и самосборка, является совокупностью неравновесных процессов, протекающих необратимо. Тем более это относится к функционированию биологической системы как целого, к процессам эволюционного и индивидуального развития.

Как мы видели, организм представляет собой своего рода химическую машину, которая работает благодаря прямым и обратным молекулярным связям. Молекулярная сигнализация служит для передачи информационных сообщений. Соответственно общая физическая трактовка биологических систем основывается на теории информации, с которой неразрывно связана термодинамика.

Теория информации вводит меру *количества информации.* Допустим, что имеется P_0 различных равновероятных событий. Так, при бросании монеты $P_0 = 2$, при бросании кости $P_0 = 6$. Чем больше P_0, тем больше неопределенность до получения сообщения о событии и тем больше количество информации (далее называемое просто информацией) при получении сообщения. В начальной ситуации, до бросания монеты или кости, информация равна нулю, $I_0 = 0$, в конечной ситуации $I_1 \neq 0$. Очевидно, что мера информации должна быть связана с P_0. Естественно потребовать, чтобы информация была аддитивной для независимых событий (скажем, при бросании двух костей).

Таким образом, если имеются два набора событий P_{01} и P_{02}, так что полное число событий есть

$$P_0 = P_{01}P_{02} \tag{9.1}$$

(при бросании двух костей $36 = 6 \cdot 6$), то должно быть

$$I(P_{01}P_{02}) = I(P_{01}) + I(P_{02}). \tag{9.2}$$

Это соотношение удовлетворяется единственным решением

$$I = K \ln P_0, \tag{9.3}$$

причем константа K произвольна. Тем самым произвольно и основание логарифма. Обычно пользуются двоичной системой с основанием 2. Если образовать все возможные «слова» или последовательности двух чисел 0 и 1 длины n, то имеется $P = 2^n$ возможностей. Потребуем, чтобы

$$I = K \ln P = Kn \ln 2 = n, \tag{9.4}$$

т. е.

$$K = 1/\ln 2 = \log_2 e \tag{9.5}$$

и

$$I = \log_2 P. \tag{9.6}$$

При этом информация исчисляется в *битах*. Так, при $P_0 = 2$, $P_1 = 1$ (бросание монеты)

$$I = \log_2 P_0 - \log_2 P_1 = 1 - 0 = 1 \text{ бит.}$$

Сколько бит 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

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

Принятое определение информации соответствует двоичной системе, 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.

Допустим, что имеется сообщение, содержащее N последовательных ячеек,— текст из N букв. В каждой из N ячеек может находиться одна из M букв (в русском языке $M = 32$). В сообщении содержится N_1 букв А, N_2 букв Б и т. д. Имеем

$$N = \sum_{j=1}^{M} N_j. \tag{9.7}$$

Вероятность появления данной буквы

$$p_j = N_j/N, \quad j = 1,\ 2,\ \ldots,\ M, \tag{9.8}$$

причем

$$\sum_{j=1}^{M} p_j = 1. \tag{9.9}$$

Общее число различных последовательностей из N букв M-буквенного языка, т. е. число возможных различных сообщений, равно

$$P = \frac{N!}{\prod_{j=1}^{M} N_j!}. \tag{9.10}$$

Информация в одном сообщении равна (пользуемся формулой

Стирлинга, дающей хорошие результаты при $N_j > 100$)

$$I = K \ln P =$$

$$= K \left[\ln (N!) - \sum_{j=1}^{M} \ln (N_j!) \right] \approx K \left(N \ln N - \sum_{j=1}^{M} N_j \ln N_j \right), \qquad (9.11)$$

или

$$I = -KN \sum_{j=1}^{M} p_j \ln p_j, \qquad (9.12)$$

и информация, приходящаяся на букву,

$$i = I/N = -K \sum_{j=1}^{M} p_j \ln p_j. \qquad (9.13)$$

Мы получили *формулу Шеннона* — более общее выражение для информации, соответствующее последовательности событий, обладающих неодинаковыми вероятностями p_j. При этом, если $K = 1/\ln 2$, информация выражается в битах, если $K = \kappa = 1{,}38 \times \times 10^{-23}$ Дж/К (постоянная Больцмана), I выражается в Дж/К, т. е. в энтропийных единицах. Назовем величину

$$S = -\kappa \sum_{j} p_j \ln p_j \qquad (9.14)$$

энтропией. Далее мы увидим, что эта величина действительно есть физическая энтропия.

Покажем, что изменение неопределенности ведет к выигрышу информации. Перейдем от распределения вероятностей $P = (p_1, p_2, \ldots, p_M)$ к распределению $Q = (q_1, q_2, \ldots, q_M)$. Спрашивается, как изменяется при этом информация? При изменении вероятности события с p до 1 изменение информации равно

$$\Delta I = \kappa \ln \frac{1}{p} = -\kappa \ln p,$$

и при изменении $p \to q$

$$\Delta I = \kappa \ln (q/p). \qquad (9.15)$$

При изменении всего распределения вероятностей $P \to Q$ изменение информации равно сумме парциальных изменений ΔI, умноженных на конечные вероятности q_j:

$$\Delta I = \kappa \sum_{j} q_j \ln \frac{q_j}{p_j}. \qquad (9.16)$$

Эта величина всегда положительна, лишь при $Q = P$ изменение информации $\Delta I = 0$. Докажем это. При любых x кроме $x = 1$

$$\ln x > 1 - 1/x.$$

Следовательно,

$$\sum_{j} q_j \ln \frac{q_j}{p_j} > \sum_{j} q_j \left(1 - \frac{p_j}{q_j} \right) = \sum_{j} q_j - \sum_{j} p_j = 0.$$

Таким образом, знание о переходе $P \to Q$ уменьшает неопределенность и дает положительный выигрыш информации.

Пусть p_j — вероятности нахождения системы в состояниях с энергией E_j. Имеем

$$E = \sum_j p_j E_j. \tag{9.17}$$

Ищем максимум величины S/κ (9.14) при одновременном выполнении условий (9.9) и (9.17). Воспользуемся методом неопределенных множителей Лагранжа. Умножим (9.9) на $\alpha - 1$ и (9.17) на β, прибавим эти выражения к (9.14) и варьируем полученную сумму. Приравняв вариацию нулю, получим

$$\delta\left[\frac{1}{\kappa} S - (\alpha - 1)\sum_j p_j - \beta \sum_j p_j E_j\right] = 0. \tag{9.18}$$

Дифференцируя по p_j и приравнивая производную нулю, находим

$$-\ln p_j - 1 - (\alpha - 1) - \beta E_j = 0,$$

откуда

$$p_j = \exp(-\alpha - \beta E_j). \tag{9.19}$$

Подставляя (9.19) в (9.14), находим максимальную энтропию

$$\frac{1}{\kappa} S_{\max} = \alpha \sum p_j + \beta \sum p_j E_j = \alpha + \beta E. \tag{9.20}$$

С другой стороны, из (9.9) и (9.19) следует

$$1 = \sum_j p_j = e^{-\alpha} \sum_j e^{-\beta E_j} = e^{-\alpha} Z,$$

т. е.

$$e^{\alpha} = Z, \quad \alpha = \ln Z. \tag{9.21}$$

Получаем из (9.20)

$$E - \frac{1}{\kappa\beta} S_{\max} = -\frac{1}{\beta} \ln Z. \tag{9.22}$$

Это хорошо известное уравнение термодинамики и статистической физики. Здесь

$$Z = \sum_j \exp(-\beta E_j) \tag{9.23}$$

— *статистическая сумма*, $S_{\max}$ — равновесная энтропия, $\beta^{-1} = \kappa T$, T — абсолютная температура. Выражение (9.19) есть *функция распределения Больцмана*

$$p_j = \frac{\exp(-E_j/\kappa T)}{\sum_j \exp(-E_j/\kappa T)}. \tag{9.24}$$

Средняя энергия системы E и равновесная энтропия S (опускаем индекс max) — функции объема V. Выражение (9.22) дает *свободную энергию Гельмгольца*

$$E - TS = -\kappa T \ln Z = F. \tag{9.25}$$

Свободная энергия Гиббса есть

$$G = F + pV = E + pV - TS = H - TS, \tag{9.26}$$

где H — энтальпия, p — давление.

Энергия и энтропия выражаются известными формулами:

$$E = \kappa T^2 \frac{d \ln Z}{dT}, \tag{9.27}$$

$$S = \kappa T \frac{d \ln Z}{dT} + \kappa \ln Z. \tag{9.28}$$

Таким образом, информационная энтропия (9.14) эквивалентна термодинамической (9.28). Один бит информации соответствует $\kappa \ln 2 = 10^{-23}$ Дж/К, т. е. очень малой термодинамической величине. Эта эквивалентность имеет реальный физический смысл — за полученную информацию нужно платить увеличением энтропии. Любое измерение связано с возрастанием энтропии окружающей среды. Энтропийная цена бита $\kappa \ln 2$ есть его минимальная стоимость. При бросании монеты получается один бит информации, но выделение энтропии вследствие нагревания монеты при ее ударе о пол много больше $\kappa \ln 2$. Монета может быть и сколь угодно большой.

Эквивалентность информации в битах и энтропии в Дж/К подобна эквивалентности массы и энергии по закону Эйнштейна

$$m = E/c^2.$$

Переводной множитель $c^{-2} \approx 1 \cdot 10^{-21}$ с^2/см^2 здесь также очень мал.

Оценим стоимость одного бита информации в единицах работы.

Допустим, что имеется идеальный газ, содержащий N молекул при давлении p и температуре T. В результате флуктуации объем газа уменьшается с V до $V - \delta V$. Работа, расходуемая на такое уменьшение, есть $W = p\delta V$. Вычислим выигрыш информации. Каждая молекула с вероятностью 1 находится в объеме V и с вероятностью $1 - \delta V/V$ в объеме $V - \delta V$. Для N молекул вероятность равна $(1 - \delta V/V)^N$. Выигрыш информации равен

$$\Delta I = K \ln \frac{1}{(1 - \delta V/V)^N} \approx KN \frac{\delta V}{V} \quad (\text{при } N \gg 1).$$

Следовательно,

$$\frac{W}{\Delta I} = \frac{p\delta V}{KN\delta V} V = \frac{pV}{KN} = \frac{\kappa T}{K}.$$

Работа, затрачиваемая на единицу информации, пропорциональна температуре, при которой определяется информация. Если ΔI

исчисляется в битах, то $K = 1/\ln 2$ и получение одного бита информации требует затраты энергии, равной $\kappa T \ln 2$, что при $T = 300$ К составляет $2 \cdot 10^{-21}$ Дж. Это — нижняя оценка затраты энергии.

Возможна грубая, условная оценка количества информации, содержащейся в живом организме. По Блюменфельду основное количество информации в человеческом организме определяется упорядоченным расположением аминокислотных остатков в 7 кг белков, чему соответствует $3 \cdot 10^{25}$ остатков. Это дает $1{,}3 \cdot 10^{26}$ бит. Другие вклады значительно меньше: 150 г ДНК, содержащимся в человеческом организме, отвечает $6 \cdot 10^{23}$ бит, упорядоченному расположению 10^{13} клеток — $4 \cdot 10^{14}$ бит и упорядоченному расположению 10^{8} молекул биополимеров в клетке — всего лишь $2{,}6 \cdot 10^{9}$ бит. Белковая информация очень мала в термодинамической мере: $1{,}3 \cdot 10^{26}\, \kappa \ln 2 = 1{,}3 \cdot 10^{3}$ Дж/К ≈ 300 кал/К. В энтропийных единицах упорядоченность живого организма заведомо мала, она значительно меньше упорядоченности куска горной породы той же массы уже потому, что организм содержит жидкости.

В этой оценке не учитывается *избыточность* (т. е. повторность) информации. Количество неизбыточной информации в организме много больше, чем в куске горной породы. Мы вернемся к этим вопросам в § 17.8.

Вся компьютерная техника связана с передачей и перекодировкой информационных сообщений. В принципе возможны компьютеры, построенные на молекулярно-биологической основе — использующие молекулы белков и нуклеиновых кислот для запасания, передачи и перекодировки информации. Это ведь и реализуется в живой природе. Построение искусственных молекулярных, а не транзисторных компьютеров — дело будущего.

Мы не рассматривали здесь важные проблемы, связанные с созданием новой информации и ее *рецепцией*.

В наших рассуждениях предполагалось, что информация кем-то или чем-то воспринимается. Однако молча принималось, что все, на что способен рецептор,— это отличить одну букву от другой. Для применений информационных представлений в теории связи такое предположение необходимо и достаточно. Процесс рецепции как физическое явление при этом не учитывается.

Как мы увидим (§ 17.8), применение теории информации в биологии требует анализа последствий рецепции сообщения. Учет этих последствий означает, что нужно изучить сам акт рецепции как необратимый, неравновесный процесс перехода рецепторной системы из менее устойчивого состояния в более устойчивое. При этом происходит запоминание информации.

Что касается *создания новой информации*, то, как это показал Кастлер, оно означает запоминание случайного выбора. Поясним это простым примером. Вы оставляете чемодан в автоматической камере хранения на вокзале и выбираете произвольный шифр, чтобы эту камеру открыть. Этот набор цифр запоминается или записывается. Тем самым создана новая информация.

Новая информация создается в каждом акте полового размножения. Это — случайный выбор, никакими детерминистическими законами природы не предусмотрено, что именно данная пара произведет потомство. Новая особь несет новую информацию — рекомбинацию родительских геномов.

Теория запоминания информации — ее рецепции и создания — не разработана. При обсуждении запоминания не имеет смысла более говорить об эквивалентности информации и энтропии. Дело в том, что мы не располагаем пока определением энтропии для систем, далеких от равновесия, для процессов запоминания. Равновесная же система, естественно, ничего не помнит. Термодинамика систем, далеких от равновесия и обладающих долговременной памятью, еще не построена.

§ 9.2. Неравновесные процессы

Все изложенное относилось к равновесным системам, в которых энтропия максимальна. Живые системы неравновесны. Перейдем к рассмотрению неравновесных систем и процессов.

Рассмотрим систему, состоящую из двух подсистем, скажем, газ, объем которого разделен на две части. Предполагаем, что в обеих подсистемах заданы распределения вероятностей p_i' и p_i''. Соответствующие энтропии равны

$$S' = -\kappa \sum_i p_i' \ln p_i', \quad S'' = -\kappa \sum_i p_i'' \ln p_i''. \tag{9.28}$$

По-прежнему

$$\sum_i p_i' = 1, \qquad \sum_i p_i'' = 1 \tag{9.29}$$

и

$$E' = \sum_i p_i' E_i', \quad E'' = \sum_i p_i'' E_i''. \tag{9.30}$$

Считаем суммарную энергию постоянной:

$$E' + E'' = E. \tag{9.31}$$

Энтропия аддитивна:

$$S' + S'' = S. \tag{9.32}$$

Имеем

$$\frac{\partial S}{\partial E'} = \frac{\partial S'}{\partial E'} + \frac{\partial S''}{\partial E'} = \frac{\partial S'}{\partial E'} + \frac{\partial S''}{\partial (E - E'')} = \frac{\partial S'}{\partial E'} - \frac{\partial S''}{\partial E''} \tag{9.33}$$

и, согласно (9.20),

$$\frac{\partial S}{\partial E'} = \kappa(\beta' - \beta'') = \frac{1}{T'} - \frac{1}{T''}. \tag{9.34}$$