

§ 9.6. ГЕНЕТИЧЕСКИЙ КОД

Расшифровка генетического кода была проведена синтетическим, а не аналитическим путем. Аналитический путь еще закрыт, так как пока не удалось установить первичную структуру матрицы — ДНК или мРНК, соответствующую синтезируемой белковой цепи с известной последовательностью аминокислотных остатков.

Путь к решению проблемы был впервые найден в работе Ниренберга и Маттеи, доложенной на 5-м Международном биохимическом конгрессе в Москве в 1961 г. [103]. При введении синтетических полирибонуклеотидов в бесклеточную систему аминокислоты включаются в полипептидную цепь. Бесклеточная система содержит рибосомы, набор тРНК, АТФ, все необходимые ферменты, но не содержит ДНК и мРНК. Система приготавливалась из разрушенных клеток *E. coli*. Центрифугированием выделялись рибосомная и надосадочная фракции. Рибосомы отмывались, надосадочная жидкость, содержавшая ацилирующие ферменты и тРНК, диализовалась против специального буфера, включавшего стабилизирующий систему меркаптоэтанол. К смеси этих двух очищенных фракций добавлялась система, генерирующая АТФ, и 20 аминокислот. В бесклеточную систему вводились РНК или синтетические полирибонуклеотиды и изучалось включение меченных C^{14} аминокислот во фракцию, нерастворимую в трихлоруксусной кислоте, т. е. в полипептиды. Оказалось, что в присутствии РНК метка включается значительно больше, чем в отсутствие РНК. Включение подавляется рибонуклеазой — ферментом, разрушающим РНК, а также пуромицином и хлорамфениколом. Было установлено, что Поли-У стимулирует включение C^{14} -фенилаланила в полипептид. В отсутствие Поли-У на 1 мг белка приходилось 44 радиоактивных импульса в минуту, в его присутствии — 39 800 импульсов.

Таким образом, удалось «обмануть» биологическую систему — вместо природной мРНК рибосомы взаимодействовали с синтетической цепью Поли-У. Этот полимер кодировал фенилаланин. Если код триплетен, то остатку Фен отвечает кодоч УУУ. Воспользовавшись полицитидиловой кислотой Поли-Ц, Ниренберг и Маттеи нашли, что она стимулирует включение Про. Таким образом, для Про был установлен кодон ЦЦЦ.

Работа с полимерами, имеющими регулярные последовательности, например с ... АУАУАУАУ ..., оказалась невозможной, так как они образуют двуспиральные структуры, лишенные матричной активности [104]. Неактивен также Поли-Г. Полинуклеотиды со случайной спиральной структурой не связываются с рибосомами. Далее (см. [1, 5, 7]) изучалось действие сополимеров известного состава, но с неизвестной последовательностью

нуклеотидов. В таком полимере известна частота появления триплетов определенного состава. Сополимер А—У в пропорции 1 : 5 стимулирует включение Фен, Лей, Иле и в меньших количествах Асн и Лиз. Если 100 — доля триплета ЗУ, то доля 1А 2У — 20 ($3У/1А\ 2У = 5$), 2А 1У — 4 ($3У/2А\ 1У = 25$) и 3А — 0,8 ($3У/3А = 125$). Сравнивая эти значения со степенями включения (для Фен она принимается равной 100), можно установить состав кодонов для названных аминокислотных остатков. (табл. 9.2).

Таблица 9.2

**Кодирующие триплеты, полученные на основании опытов
с Поли-АУ (1:5) [3]**

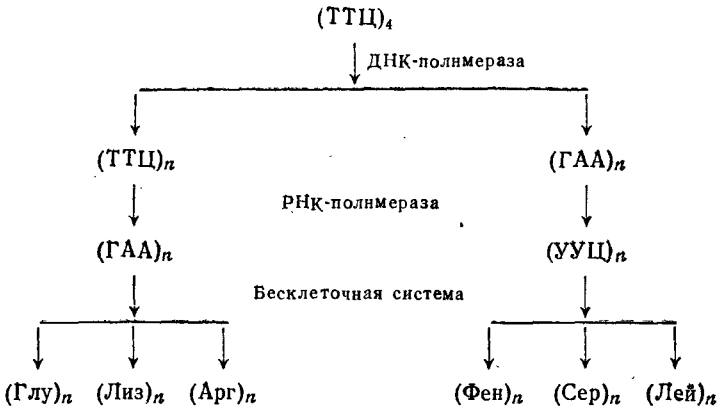
Аминокислотный остаток	Вычисленные доли триплетов				Сумма долей	Включение аминокислоты
	3А	2А 1У	1А 2У	3У		
Асн	—	4	—	—	4	6,6
Иле	—	4	20	—	24	20
Лей	—	—	20	—	20	15
Лиз	0,8	4	—	—	4,8	3,1
Тир	—	—	20	—	20	25
Фен	—	—	—	100	100	100

Так были установлены триплеты ЗУ для Фен, 1А 2У для Тир, 2А 1У и 1А 2У для Иле, 1А 2У для Лей, 2А 1У для Асн и 3А и 2А 1У для Лиз. В дальнейшем последний триплет оказался не кодирующим Лиз, остальные подтвердились.

Однако эти опыты еще не дают полной расшифровки кода. Неизвестно, какой из трех кодонов АУУ, УАУ или УУА кодирует Тир и т. п. Подробная история исследования кода изложена в монографии Ичаса [5]. Полная расшифровка кода была получена в работах Ниренберга, применившего в последующих опытах уже не полинуклеотиды, а тринуклеотиды известного строения [105]. В системах образуются комплексы тринуклеотид — тРНК — АК (аминокислота). Синтез полипептида в этих условиях не идет, но, поскольку тринуклеотид имитирует кодон, образование комплекса позволяет прочесть его. Для этого нужно изучить все тРНК, которые последовательно связываются с мечеными C^{14} аминокислотами. Так были исследованы все 64 триплета и установлено, с какими аминокислотами они связаны. Окончательную расшифровку, подтверждающую эти результаты, провел Корана (см. [106, 107]). Он синтезировал олигодезоксирибонуклеотиды, содержащие от 8 до 12 звеньев в цепи, с известной последовательностью. Они представляли собой по-

вторяющиеся триплеты (например, $(ТТЦ)_4$). Далее такой олигомер применялся в качестве матрицы *in vitro* для синтеза ДНК-подобного полимера в системе, содержавшей нуклеозидтрифосфаты Т, Ц, Г, А, нужное количество ионов и ДНК-полимеразу (ср. стр. 537). При этом синтезировалась двойная спираль, подобная ДНК. Обе цепи спирали содержали комплементарные, повторяющиеся n раз триплеты. Каждая из цепей затем служила матрицей для синтеза полирибонуклеотида с помощью РНК-полимеразы.

Таким образом, Корана получил две цепи, имитирующие мРНК с известной последовательностью повторяющихся триплетов. Обе цепи вводились в бесклеточную систему, и определялось включение аминокислоты в полипептидную фракцию по методу Ниренберга. Эти элегантные опыты позволили проверить шесть кодонов в одном многостадийном синтезе в соответствии со схемой



Полимер $(ГАА)_n$ содержит кодоны ГАА, ААГ и АГА, а полимер $(УУЦ)_n$ — кодоны УУЦ, УЦУ и ЦУУ. Так как функциональность гомополимеров $(ААА)_n$, $(ЦЦЦ)_n$, $(ГГГ)_n$ и $(УУУ)_n$ уже определена, для проверки остальных 60 кодонов нужно провести 10 таких многостадийных синтезов.

Приведем установленный Ниренбергом с сотрудниками и Кораной с сотрудниками кодонно-аминокислотный словарь (табл. 9.3), а также обратный аминокислотно-кодонный словарь (табл. 9.4).

Нуклеотиды расположены в последовательности букв А, Ц, Г, У, совпадающей с латинским алфавитом.

Кодоны ГУГ (Вал) и АУГ (Мет) кодируют указанные аминокислоты в середине цепи мРНК. Вместе с тем они служат

Таблица 9.3

Кодонно-аминокислотный словарь

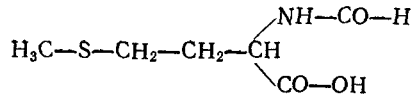
1. ААА	Лиз	17. ЦАА	Гли	33. ГАА	Глу	49. УАА	—
2. ААЦ	Аси	18. ЦАЦ	Гис	34. ГАЦ	Асп	50. УАЦ	Тир
3. ААГ	Лиз	19. ЦАГ	Гли	35. ГАГ	Глу	51. УАГ	—
4. ААУ	Аси	20. ЦАУ	Гис	36. ГАУ	Асп	52. УАУ	Тир
5. АЦА	Тре	21. ЦЦА	Про	37. ГЦА	Ала	53. УЦА	Сер
6. АЦЦ	Тре	22. ЦЦЦ	Про	38. ГЦЦ	Ала	54. УЦЦ	Сер
7. АЦГ	Тре	23. ЦЦГ	Про	39. ГЦГ	Ала	55. УЦГ	Сер
8. АЦУ	Тре	24. ЦЦУ	Про	40. ГЦУ	Ала	56. УЦУ	Сер
9. АГЦ	Арг	25. ЦГА	Арг	41. ГГА	Гли	57. УГА	—
10. АГЦ	Сер	26. ЦГЦ	Арг	42. ГГЦ	Гли	58. УГЦ	Цис
11. АГГ	Арг	27. ЦГГ	Арг	43. ГГГ	Гли	59. УГГ	Три
12. АГУ	Сер	28. ЦГУ	Арг	44. ГГУ	Гли	60. УГУ	Цис
13. АУА	Иле	29. ЦУА	Лей	45. ГУА	Вал	61. УУА	Лей
14. АУЦ	Иле	30. ЦУЦ	Лей	46. ГУЦ	Вал	62. УУЦ	Фен
15. АУГ	Мет	31. ЦУГ	Лей	47. ГУГ	Вал	63. УУГ	Лей
16. АУУ	Иле	32. ЦУУ	Лей	48. ГУУ	Вал	64. УУУ	Фен

Таблица 9.4

Аминокислотно-кодонный словарь

1. Ала	ГЦА, ГЦЦ, ГЦГ, ГЦУ	12. Лиз	ААА, ААГ
2. Арг	АГА, АГГ, ЦГА, ЦГЦ, ЦГГ, ЦГУ	13. Мет	АУГ
3. Аси	ААЦ, ААУ	14. Про	ЦЦА, ЦЦЦ, ЦЦГ, ЦЦУ
4. Асп	ГАЦ, ГАУ	15. Сер	АГЦ, АГУ, УЦА, УЦЦ, УЦГ, УЦУ
5. Вал	ГУА, ГУЦ, ГУГ, ГУУ	16. Тир	УАЦ, УАУ
6. Гис	ЦАЦ, ЦАУ	17. Тре	АЦА, АЦЦ, АЦГ, АЦУ
7. Гли	ГГА, ГГЦ, ГГГ, ГГУ	18. Три	УГГ
8. Гли	ЦАА, ЦАГ	19. Фен	УУЦ, УУУ
9. Глу	ГАА, ГАГ	20. Цис	УГЦ, УГУ
10. Иле	АУА, АУЦ, АУУ		
11. Лей	ЦУА, ЦУЦ, ЦУГ, ЦУУ, УУА, УУГ		

инициаторами синтеза цепи, кодируя на ее N-конце формилметионин



Если искусственная матрица этих кодонов не содержит, то синтезируются цепи, начинающиеся с произвольного звена и имеющие поэтому различные N-концы. Напротив, при наличии ГУГ или АУГ образуются стандартные формилметиониновые N-концы. Корана показал, что Поли-УГ образует цепь Формилмет — (Цис — Вал)_n. Однако в природных белках формилметионин не содержится. В бесклеточной системе *E. coli* образуются

цепи с N-концами Формилмет — Ала, Формилмет — Сер —, но не Формилмет — Мет —. Естественные белки *E. coli* имеют обычно на N-концах Мет, Ала, Сер. Можно думать, что в живых системах действуют два фермента, один из них отщепляет Формилмет от цепи, другой — формильную группу от Мет (см. статьи Циндера и др. в [107]). Кодоны УАА («охра»), УАГ («янтарь») и УГА не кодируют никаких аминокислот, на них белковая цепь обрывается. Это — терминальные кодоны. Корана установил, что синтез полипептида не идет на матрицах (ГАУА)_n и (ГУАА)_n. Последние содержат много нефункциональных триплетов.

Установленный таким образом код подтверждается множеством фактов (см. [5, 7]).

Еще до расшифровки кода Суеока исследовал корреляцию состава тотального белка бактерий с составом их ДНК [108] Суеока нашел для 16 аминокислотных остатков линейную зависимость их содержания в белках от содержания ГЦ в ДНК. При этом аминокислотные остатки разделяются на три группы. Для остатков первой группы их содержание растет с увеличением Г + Ц и коэффициент корреляции *b* велик. Для остатков второй группы их содержание практически не зависит от концентрации Г + Ц и *b* близко к нулю. Для остатков третьей группы их содержание убывает с увеличением Г + Ц. Эти результаты хорошо объясняются кодовым словарем, что показано в табл. 9.5 [1, 109].

Универсален ли генетический код? Действуют ли аналогичным образом одинаковые кодоны в различных организмах? Ответ на этот вопрос положительный. Поли-У стимулирует включение Фен в полипептидную цепь в бесклеточных системах, полученных из клеток млекопитающих и водорослей. То же относится к другим синтетическим полинуклеотидам (см. [5]). Три-нуклеотидная техника Ниренберга была применена к бесклеточным системам, полученным из клеток амфибии *Xenopus laevis* и морской свинки, и привела к тем же результатам [110]. Меняется, по-видимому, лишь относительное участие различных кодонов для одной и той же аминокислоты, но кодовый словарь остается тем же, что и для *E. coli*.

Универсальность кода в сущности доказывается размножением фагов и вирусов в клетках. Вирусная ДНК или РНК использует биосинтетический аппарат клетки и синтезирует свои белки. Белок закодирован нуклеиновой кислотой вируса, а «бесклеточная система» — иная.

Методы, развитые Кораной, позволили решить проблему синтеза гена. Корана провел химический синтез последовательности дезоксирибонуклеотидов, комплементарной к известной последовательности рибонуклеотидов в Ала-тРНК дрожжей [111].

Была получена двуспиральная ДНК, т. е. ген, кодирующий синтез этой тРНК. Синтез первой цепи ДНК осуществлялся присоединением олигонуклеотидов к растущей последовательности при участии ферментов ДНК-лигаз и АТФ в качестве источника энергии. Вторая цепь получалась в редупликационном синтезе (см. стр. 585).

Таблица 9.5

Корреляция содержания аминокислотных остатков в белке
с содержанием Г + Ц в ДНК

Аминокислотный остаток	Коэффициент корреляции	Кодоны	Число Г + Ц в кодонах	Средняя доля Г + Ц в кодонах
Группа 1, $b > 0$				
Ала	0,164	ГЦА, ГЦЦ, ГЦГ, ГЦУ	2, 3, 3, 2	0,83
Арг	0,089	АГА, АГГ, ЦГА, ЦГЦ ЦГГ, ЦГУ	1, 2, 2, 3 3, 2	0,72
Гли	0,051	ГГА, ГГЦ, ГГГ, ГГУ	2, 3, 3, 2	0,83
Про	0,024	ЦЦА, ЦЦЦ, ЦЦГ, ЦЦУ	2, 3, 3, 2	0,83
Среднее	0,082		2, 39	0,79
Группа 2, $b \approx 0$				
Вал	0,008	ГАУ, ГУЦ, ГУГ, ГУУ	1, 2, 2, 1	0,50
Тре	0,000	АЦА, АЦЦ, АЦГ, АЦУ	1, 2, 2, 1	0,50
Лей	-0,006	ЦУА, ЦУЦ, ЦУГ, ЦУУ УУА, УУГ	1, 2, 2, 1 0, 1	0,39
Гис	-0,010	ЦАЦ, ЦАУ	2, 1	0,50
Сер	-0,017	АЦГ, АГУ, УЦА, УЦЦ УЦГ, УЦУ	2, 1, 1, 2 2, 1	0,50
Среднее	-0,005		1, 41	0,47
Группа 3, $b < 0$				
Мет	-0,024	АУГ	1	0,33
Фен	-0,040	УУЦ, УУУ	1, 0	0,17
Тир	-0,047	УАЦ, УАУ	1, 0	0,17
Глу	-0,052	ГАА, ГАГ	1, 2	0,50
Глн		ЦАА, ЦАГ	1, 2	0,50
Асп	-0,053	ГАЦ, ГАУ	2, 1	0,50
Асн		ААЦ, ААУ	1, 0	0,17
Лиз	-0,084	ААА, ААГ	0, 1	0,17
Иле	-0,098	АУА, АУЦ, АУУ	0, 1, 0	0,11
Среднее	-0,057		0, 83	0,28

Расшифровка генетического кода — крупнейшее достижение молекулярной биологии, биохимии и биофизики. От постановки задачи до ее решения прошло лишь 10 лет — срок очень малый.

Установление кода выдвинуло новые проблемы.

Имеет ли генетический словарь физический, молекулярный смысл или корреляция между кодонами и аминокислотами совершенно случайна? Что можно сказать об эволюции кода в этой связи? Какие факторы влияют на чтение кода, на процессы транскрипции и трансляции? Что и как искажает код? Каковы физико-химические причины мутаций?

Код есть программа трансляции. Чрезвычайно важно понять условия реализации этой программы. Молекулярные механизмы ответственны за кинетику взаимодействия антикодонов с кодонами, за действие рибосом и ацилирующих ферментов. Сегодня мы почти ничего не знаем об этих механизмах и еще не располагаем физической теорией чтения кода (см. § 9.8).

§ 9.7. ФИЗИЧЕСКИЙ СМЫСЛ ГЕНЕТИЧЕСКОГО КОДА

Генетически кодируется только первичная структура белка. Однако биологические функции определяются его пространственным строением. Первичная структура и пространственное строение однозначно связаны (см. § 4.9). Тем самым, генетически закодированы пространственное строение и биологическая функция белка. В то же время естественный отбор идет не по первичной, а по пространственной структуре — по биологическому поведению.

Различные мутационные замещения по-разному сказываются на строении белка. Из изложенного в § 4.6 следует, что мутации, сильно изменяющие гидрофобность аминокислотного остатка, должны сильнее сказываться на биологических свойствах белка, чем мутации, мало меняющие гидрофобность. Первый тип мутаций более опасен для существования вида, чем второй. Можно думать, что генетический код построен природой таким образом, чтобы обеспечить преимущество мутациям первого типа. Нужно выявить эти особенности кода.

Такая формулировка проблемы содержалась в работах [112], где было показано, что код обладает большой надежностью по отношению к неблагоприятным заменам полярных аминокислотных остатков на неполярные и наоборот. Между аминокислотами и кодонами имеется осмысленная корреляция (см. также [113]).

Прежде всего следует установить, какой характер имеют мутационные аминокислотные замещения. Мы уже видели, что внутренние аминокислотные остатки в гемоглобине и миоглобине варьируют, но остаются неполярными (см. стр. 232).