

ПРИРОДА БИОХИМИЧЕСКОГО КОДА *)

В. Ю. Гаврилов, Ю. Н. Зограф

За последние годы достигнуты чрезвычайно крупные успехи в понимании процессов синтеза белков в живой клетке и выяснении роли специфических биополимеров — нуклеиновых кислот — в этих процессах. Наиболее крупными достижениями в этой области являются осуществление в лабораторных условиях биосинтеза искусственных аналогов нуклеиновых кислот, выяснение основных этапов синтеза белка, выяснение строения отдельных типов нуклеиновых кислот и функций, выполняемых ими. В самое последнее время крупные успехи достигнуты в исследовании природы биохимического кода, т. е. в выяснении закономерностей, связывающих тонкие детали строения нуклеиновых кислот с расположением остатков аминокислот в синтезируемых клеткой белках.

Настоящий обзор написан с целью познакомить читателя-физика с основными фактами и идеями в этой исключительно быстро развивающейся области естествознания. По необходимости он носит несколько популярный характер и ни в коей мере не претендует на исчерпывающее изложение затронутых вопросов.

В клетке имеются два типа нуклеиновых кислот — дезоксирибонуклеиновая кислота (ДНК) и рибонуклеиновая кислота (РНК), причем последняя имеет несколько существенно различных по своему строению и по своим функциям фракций. В настоящее время общепризнано (хотя и встречает еще возражения со стороны отдельных специалистов), что ДНК содержится в ядрах клеток, а РНК имеется как в ядрах, так и в цитоплазме. Накоплено множество фактов, доказывающих, что ДНК содержит в себе генетическую информацию, определяющую передачу от поколения к поколению признаков, характерных для данной клетки или для всего многоклеточного организма в целом. Каждый из этих фактов, взятый в отдельности, можно в какой-то мере оспаривать, но вся их совокупность в целом не оставляет сомнения в специфической генетической роли ДНК в клетке.

В 1953 г. Ватсон и Крик, проанализировав данные рентгеноструктурного анализа натриевой соли дезоксирибонуклеиновой кислоты и используя ряд стереохимических соображений, предложили свою модель строения ДНК^{1, 2}. Согласно их предположениям ДНК представляет собой двойную спираль диаметром 18—20 Å и с шагом 34 Å (рис. 1), состоящую из двух противоположно направленных сахаро-фосфатных нитей, включающих в себя остатки фосфорной кислоты и молекулы сахара — дезоксирибозы. К молекулам сахара прикреплены молекулы пуриновых и пиримидиновых оснований, обращенные внутрь двойной спирали и связанные

*) Доклад на семинаре Института физических проблем им. С. И. Вавилова 10 мая 1962 г.

между собой относительно непрочными водородными связями (рис. 2). При достаточно большой влажности плоскости, в которой лежат пары оснований, прикрепленных к сахарам двух противоположных нитей спирали, перпендикулярны к ее оси.

Разноименное направление сахаро-фосфатных нитей ДНК обусловлено несимметричностью строения молекулы сахара, и спиральная структура ДНК может быть обеспечена только при условии, что обе нити имеют противоположное направление. Факт разноименности направлений двух нитей был установлен экспериментально с помощью биохимических методов³.

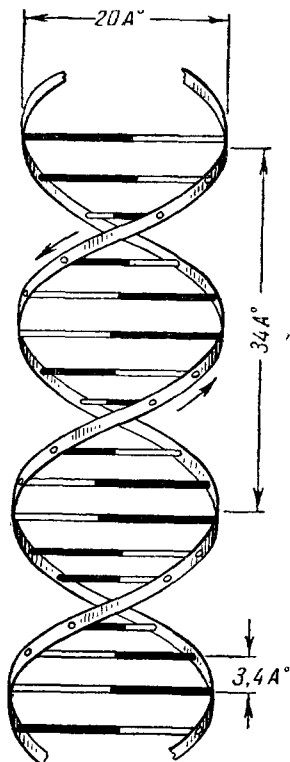


Рис. 1. Схема двухнитевой спиральной структуры ДНК по модели Ватсона и Крика.

Размеры молекулы приведены для ДНК в форме В (при большой относительной влажности).

Дезоксирибонуклеиновая кислота (ДНК) содержит четыре различных основания: два пуриновых основания — аденин (А) и гуанин (Г)—и два пиримидиновых основания — тимин (Т) и цитозин (Ц). Элементарное звено ДНК, состоящее из остатка фосфорной кислоты, сахара и прикрепленного к нему основания, называется нуклеотидом. Из стереохимических соображений Ватсон и Крик пришли к выводу, что основания, принадлежащие различным нитям молекулы, связаны водородными связями в строго определенные пары: А связывается двумя водородными связями с Т, а Г — тремя водородными связями с Ц. Такая модель великолепно согласуется с экспериментальными данными⁴ и объясняет, почему в ДНК всегда количество А равно количеству Т, а количество Г равно количеству Ц (так называемое правило Чаргаффа⁵). Состав ДНК заметно различается от вида к виду. Для микроорганизмов, например, отношение $A+T/G+C$ варьирует от 0,35 до 2,84; у высших растений это отношение колеблется от 1,08 до 1,78; у беспозвоночных — от 1,27 до 1,94, а у позвоночных — только в пределах от 1,27 до 1,50⁶.

Кроме указанных четырех оснований, в ДНК встречаются и другие, из которых наиболее существенны 5-метилцитозин, заменяющий часть цитозина в ДНК ряда организмов, и 5-оксиметилцитозин, замещающий весь цитозин у некоторых фагов. Последовательность нуклеотидов в ДНК неизвестна⁷, но известно, что для каждого вида она специфична и, в частности, вероятности нахождения рядом двух заданных нуклеотидов различны для ДНК разных видов³. Молекулярный вес и, соответственно, длина выделенной молекулы ДНК сильно зависят от способа ее выделения⁸. Хотя обычно молекулярный вес ДНК в выделенных препаратах колеблется от 5 до $15 \cdot 10^6$, недавно из фага Т2 были получены молекулы ДНК длиной 52 мк, что соответствует молекулярному весу около $110 \cdot 10^6$ ⁹. Имеются указания на то, что в некоторых физиологических состояниях клетки ДНК внутри нее может разделяться, по крайней мере частично, на отдельные нити, т. е. переходить из двухнитевого состояния в однонитевое¹⁰. Следует отметить, что при разрыве водородных связей две нити спирали ДНК не могут быть просто вынуты друг из друга и механизм расплетения двух таких нитей в настоящее время совершенно не ясен. Большинство организмов содержат двухнитевую ДНК, однако имеются и исклю-

чения. Так, например, фаг ФХ174 имеет однонитевую ДНК¹¹. Для таких ДНК правило Чаргаффа, естественно, не выполняется.

Правило дополнительности оснований различных нитей ДНК, т. е. возникновение водородных связей только между определенными парами дополнительных оснований (между А и Т и между Г и Ц), приводит к тому, что последовательность оснований в одной нити ДНК полностью определяет последовательность оснований в другой, дополнительной к ней нити. Исходя из этого, Ватсон и Крик^{12,13} предположили, что

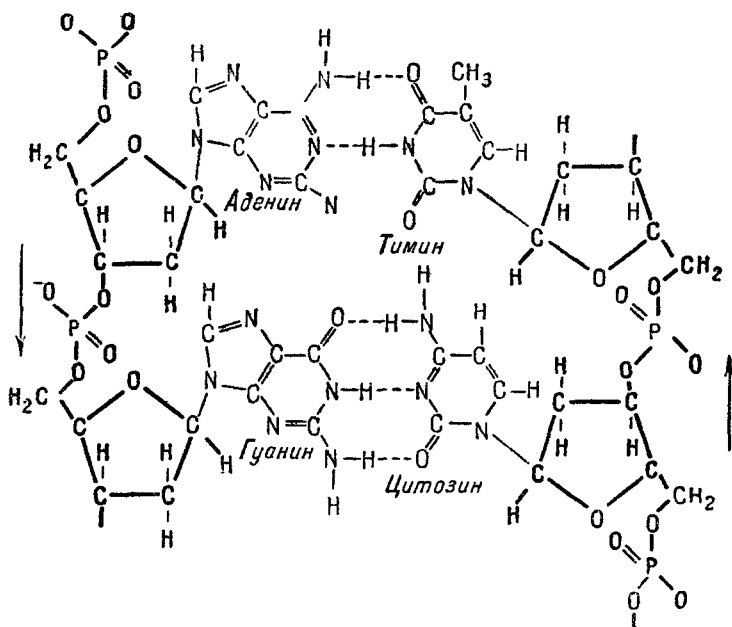


Рис. 2. Строение участка молекулы ДНК.

Плоскости оснований должны быть повернуты перпендикулярно к плоскости чертежа. Сахаро-фосфатные нити выделены жирным шрифтом. Пунктирными линиями показаны водородные связи между основаниями различных нитей.

ДНК может воспроизводиться посредством редупликации, когда каждая из нитей исходной молекулы служит матрицей для построения одной из двух дочерних молекул ДНК. При этом при разрыве слабых водородных связей между основаниями происходит разделение нитей, а затем каждое освободившееся основание нити взаимодействует с дополнительным к нему основанием свободного нуклеотида, что обуславливает полимеризацию дополнительной нити со специфической последовательностью нуклеотидов. Такое свойство ДНК позволило предположить, что именно так происходит точное удвоение генетической информации, содержащейся в молекуле ДНК, и передача этой информации от поколения к поколению в ходе деления клеток. Мезельсон и Сталь¹⁴, исследуя распределение метки, внесенной в ДНК, при воспроизведении ДНК в бактериальных клетках, показали, что новые молекулы ДНК содержат одну старую и одну заново синтезированную нити, т. е. что синтез ДНК происходит по схеме редупликации Ватсона и Крика. Они же показали, что при нагревании изолированной ДНК приблизительно до 100° С ее нити разделяются (денатурация). В дальнейшем Доти¹⁶⁻¹⁸ обнаружил, что при определенных условиях охлаждения такие разделенные нагреванием нити ДНК вновь объединяются в двойную спираль (так называемая ренатурация ДНК). Доти даже получил биологически активные «гибридные»

молекулы ДНК, взятой от двух близких видов бактерий, искусственно разделенной и вновь реконструированной в двухнитевые молекулы.

Крупнейшим достижением биохимии явился осуществленный в 1958 г. Корнбергом вне клетки биосинтез молекул ДНК из нуклеозидтрифосфатов при участии выделенного им из бактерий специфического фермента — полимеразы ДНК¹⁹⁻²². При этом, независимо от того, из клеток каких организмов был выделен этот фермент, синтезировалась ДНК подобная той, которая была введена в такую систему в качестве затравки. С помощью этого фермента можно получить, не вводя в систему ДНК-затравку, искусственные аналоги ДНК — двухнитевые синтетические А—Т и Г—Ц-кополимеры. При этом поли-А—Т состоит из двух нитей, в каждой

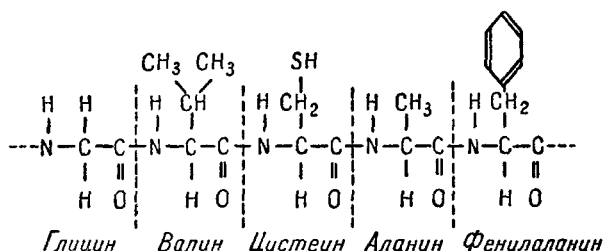


Рис. 3. Участок белковой цепи.

Пептидные связи образованы между атомами С и N соседних аминокислотных остатков, разделенных пунктиром.

то синтезированная ДНК получается двухнитевой. Поэтому можно считать, что процессы расхождения нитей и полимеризации комплементарной нити разделены, но конкретный физико-химический механизм каждого из них не ясен.

Как в ядре, так и в цитоплазме имеется большое количество другой нуклеиновой кислоты — рибонуклеиновой (РНК). В химическом отношении она отличается от ДНК тем, что в ее состав входит другой сахар — рибоза и вместо тимина она содержит урацил (У). В клетках имеется несколько различных типов РНК — низкомолекулярная растворимая РНК (sРНК), состоящая примерно из 50—100 нуклеотидов, и высокополимерные РНК с молекулярными весами порядка миллионов. Несмотря на то, что эти молекулы, вероятно, однонитевые, они обладают некоторой упорядоченностью. Предполагается, что в них имеются петлевидные спиральные участки, в которых нуклеотиды соединены водородными связями^{23,24}.

Специфические свойства каждой из клеток определяются прежде всего составом ее белков. Белками, в частности, являются все ферменты клетки. Количество типов различных белков в клетках очень велико, даже в небольших бактериальных клетках их число превышает несколько сот. Белки представляют собой полипептидные цепи, состоящие из нескольких сотен, а иногда из нескольких десятков остатков аминокислот, связанных между собой пептидными связями (рис. 3). Они имеют сложную вторичную структуру (конфигурация полипептидной цепи, определяемая системой водородных связей между аминокислотными остатками) и третичную структуру (расположение цепи в пространстве, определяемое в основном дисульфидными связями между остатками содержащей серу аминокислоты — цистеина). Многие белки состоят из нескольких отдельных полипептидных цепей, связанных дисульфидными связями. Полипептидная цепь белка имеет определенное направление, так как пептидная связь образуется между атомами С и N соседних аминокислот, т. е. несимметрична. Рядом авторов и, в частности Криком²⁵, была высказана гипотеза, подтверждающаяся экспериментальными данными²⁶, о том, что первич-

из которых регулярно чередуются А и Т, а поли-Г—Ц содержит одну нить, целиком состоящую из Г, и другую, целиком состоящую из Ц. Было показано, что синтез ДНК в такой системе происходит на одной нити введенной затравки, достраивающей себе дополнительную нить.

Если в качестве затравки взята однонитевая ДНК

ная структура, т. е. последовательность аминокислотных остатков в полипептидной цепи белка, определяет его вторичную и третичную структуры, а следовательно, и его функциональные свойства. Основная масса белков синтезируется в цитоплазме, причем в основном синтез происходит в специальных структурных образованиях клетки — рибонуклеопротеидных гранулах, так называемых рибосомах. Рибосомы очень богаты РНК и содержат РНК двух категорий — с молекулярными весами $5 \div 6 \cdot 10^5$ и $1,1 \div 1,3 \cdot 10^6$ ^{27, 28}. Строение рибосом в настоящее время неизвестно. При малых концентрациях Mg^{++} в среде имеются два типа рибосом — рибосомы с константой седиментации 30s с молекулярным весом около $0,95 \cdot 10^6$ и рибосомы с константой седиментации 50s и молекулярным весом $1,85 \cdot 10^6$. При увеличении концентрации Mg^{++} наблюдается объединение таких рибосом в частицы с большими константами седиментации — в рибосомы 70s и 100s ²⁹:



Синтез белка происходит в 70s-рибосомах и, по-видимому, в 100s-рибосомах, в то время как остальные конфигурации рибосомных частиц не активны.

Начиная с 1954—1956 гг. была проделана огромная работа по выяснению всей схемы белкового синтеза. Основой для нее явились исследования группы Замечника и Хогланда ³⁰⁻³⁴, показавшей, что этот процесс происходит через несколько стадий. Сначала под действием определенных ферментов происходит активация аминокислот с образованием специфических соединений — аминоациладенилатов. При этом в процессе обязательно участвует обладающее макроэргическими связями соединение — аденозинтрифосфорная кислота (АТФ). В ходе процесса от нее отщепляется неорганический пирофосфат, а оставшая часть АТФ соединяется с аминокислотой. На следующей стадии происходит присоединение активированной аминокислоты к молекуле sРНК с освобождением аденозинмонофосфорной кислоты. Аминокислота при этом присоединяется к определенному концу sРНК, причем все молекулы sРНК имеют на этом конце одну и ту же последовательность трех нуклеотидов: Ц—Ц—А. Эти первые две стадии проходят при участии специфических активирующих ферментов. При этом молекулы sРНК явно специфичны по отношению к аминокислотам, но существуют указания, что у самых различных организмов набор sРНК один и тот же. Активирующие ферменты также специфичны по отношению к аминокислотам. По-видимому, для каждой аминокислоты имеется своя sРНК и свой активирующий фермент. Последняя стадия этого процесса, происходящая внутри рибосом — выстраивание молекул sРНК с прикрепленными к ним аминокислотами на определяющей строение белка матрице и соединение аминокислот пептидными связями, т. е. уже непосредственное формирование полипептидной цепи белка. При этом sРНК освобождается и вновь выходит в раствор. Конкретный механизм этого этапа синтеза пока неясен. Химия в настоящее время еще не знает подобных явлений синтеза полимерных цепей на матрицах, и какие-либо этапы подобного типа синтеза совершенно не изучены. Интересно отметить, что объединение аминокислот в пептидную цепь происходит только в определенном направлении последовательно, и от матрицы отрывается только полностью сформировавшаяся полипептидная цепь белка. Так, синтез пептидной цепи гемоглобина начинается с N-конца ³⁵ и продолжается в течение 1,5 минуты, т. е. за одну секунду полимеризуются примерно две аминокислоты ³⁶.

Рядом исследователей давно уже высказывалось мнение, что строение нуклеиновых кислот определяет свойства синтезируемых клеткой белков.

Весьма убедительным экспериментом, обосновывающим эту точку зрения, явился, например, эксперимент, проведенный Шраммом и Шустером на вирусе табачной мозаики³⁷.

В состав вирусов, паразитирующих на клетках животных или растений, всегда входит некоторое количество белков и нуклеиновая кислота — как правило, ДНК у вирусов животных и РНК у вирусов растений. Сами по себе вирусы неспособны размножаться, и процесс их размножения происходит лишь внутри зараженной ими клетки. При заражении вирусом клетки в ней начинает синтезироваться характерная для вируса нуклеиновая кислота и ряд специфических белков, либо непосредственно входящих в состав новых вирусных частиц, либо необходимых в качестве ферментов, регулирующих процессы синтеза новых для данной клетки биохимических соединений.

Вирус табачной мозаики, паразитирующий на листьях табака и вызывающий характерное их заболевание, по своему строению представляет собой самый простой из изученных в настоящее время вирусов. В его состав входит только одна молекула РНК, содержащая примерно 6500 нуклеотидов, и 2200 молекул белка, причем все эти молекулы одинаковы и каждая из них состоит из полипептидной цепи, содержащей 158 аминокислотных остатков, последовательность которых полностью определена³⁸. Обработывая вирусные частицы азотистой кислотой, Шрамм и Шустер показали, что азотистая кислота является химическим мутагеном, т. е. способна вызывать у вируса новые передающиеся от поколения к поколению свойства. Известно, что азотистая кислота действует на пуриновые и пиримидиновые основания в РНК, причем переводит в некоторых точках молекулы аденин в гуанин или цитозин в урацил. Вместе с тем непосредственно на белки она не воздействует. При действии азотистой кислоты на РНК вируса табачной мозаики Цугитой и Френкель-Конратом³⁹ был обнаружен целый ряд мутантов, отличающихся тем, что в молекулах белка вируса вместо какого-либо аминокислотного остатка в пептидной цепи оказывался остаток другой аминокислоты, т. е. изменение в РНК привело к изменению в аминокислотных остатках, образующих белок.

По мере развития представлений об этапах синтеза белков в клетке концепция о связи между составом белков и последовательностью оснований в нуклеиновой кислоте, несущей на себе генетическую информацию, столкнулась с рядом трудностей. Не было видно прямой связи между ДНК в ядре клетки и синтезируемыми в рибосомах белками. Входящая в состав рибосом высокополимерная РНК по суммарному составу оснований может быть совершенно отлична от состава оснований ДНК и чрезвычайно мало варьирует от клетки к клетке*).

В 1960 г. это затруднение было снято в связи с открытием особого типа высокополимерной РНК-информационной или «messenger»-РНК (mРНК), которая синтезируется в ядре на ДНК как на матрице, выходит затем в цитоплазму и участвует в качестве матрицы в синтезе белка в рибосомах. Эта mРНК составляет лишь незначительную часть всей РНК клетки, обычно не более 2—3%. Она отличается от остальной РНК тем, что обладает малым временем жизни — порядка нескольких минут. Можно думать, что каждая молекула такой РНК обеспечивает в рибосоме синтез одной молекулы белка, а затем распадается. При импульсном введении в клетку меченых нуклеотидов почти вся метка оказывается именно в mРНК. Таким методом и был открыт синтез mРНК в заражен-

*) Тем не менее, Белозерским и Спириным была обнаружена некоторая корреляция между нуклеотидными составами ДНК и суммарной РНК у микроорганизмов (А. Н. Белозерский, А. С. Спиринов, *Nature* 182, 111 (1958)).

ных фагом клетках бактерий, причем ее нуклеотидный состав оказался соответствующим нуклеотидному составу ДНК фага³⁹⁻⁴¹. Впоследствии это же было сделано и для некоторых других объектов^{42,43}. Оказалось, что такая РНК включается в 70s- и 100s-рибосомы⁴³ и одновременно в клетке содержат ее в себе около 10% рибосом⁴⁴. Именно мРНК, а не сами рибосомные частицы определяют специфичность синтезируемого белка. Действительно, при заражении бактерии фагом синтез новых специфических белков, определяемых фаговой ДНК, происходит в старых бактериальных рибосомах, существовавших еще до заражения, но с участием новой мРНК, соответствующей по нуклеотидному составу ДНК фага, но не ДНК бактерии⁴⁵. Такая мРНК имеет молекулярный вес около $3 \cdot 10^5$.

Не только валовый состав оснований мРНК соответствует составу ДНК, но мРНК обладает также последовательностью оснований, дополнительной к последовательности оснований ДНК, так как при совместной ренатурации в пробирке соответствующих мРНК и предварительно разделенных нагреванием ДНК получаются «гибридные» комплексы, состоящие из одной нити ДНК и одной нити мРНК⁴⁶. Комплексы, состоящие из ДНК и мРНК, правда еще невыясненного строения, выделялись также и непосредственно из клеток⁴⁷⁻⁴⁸.

Из различных организмов были выделены системы ферментов, осуществляющие синтез РНК при участии ДНК в качестве матрицы⁴⁹⁻⁵³. Нуклеотидный состав и последовательность нуклеотидов в такой искусственно синтезированной мРНК полностью соответствует составу и последовательности ДНК затравки. При этом синтез РНК может происходить при участии как двухнитевой, так и однонитевой ДНК затравки.

Таким образом, по-видимому, матрицей при синтезе мРНК служит каждая из нитей ДНК и одновременно синтезируются две нити мРНК, дополнительные к обеим нитям ДНК⁵³.

Предположение о связи между строением ДНК и белков, синтезируемых клеткой, привело Гамова^{54, 55} в 1954 г. к формулировке задачи о биохимическом коде, т. е. задачи о том, как последовательность оснований в нуклеиновой кислоте (ДНК или определяемой ею РНК) определяет последовательность аминокислот в белке. При этом Гамов еще предполагал тогда, что белки синтезируются непосредственно на ДНК. Из всех встречающихся в природе аминокислот во всех организмах содержатся в белках только 20, указанные в табл. I.

Таблица 1

Аминокислоты, непосредственно участвующие
в формировании полицептидной цепи белка

Аминокислота	Сокращение	Аминокислота	Сокращение
Аланин	ала	Лейцин	лей
Аргинин	арг	Изолейцин	илей
Аспарагиновая кислота	асп	Лизин	лиз
Аспарагин	аспN	Метионин	мет
Валин	вал	Пролин	про
Гистидин	гис	Серин	сер
Глицин	гли	Тирозин	тир
Глутаминовая кислота	глу	Треонин	тре
Глутамин	глуN	Триптофан	три
		Фенилаланин	фен
		Цистеин	цис

Другие аминокислоты, встречающиеся в природе, либо вообще не включаются в белки, либо входят в состав только небольшого количества специфических белков (оксипролин в коллагене, тирозин-*О*-сульфат в фибриногене и т. п.). Впервые Крик предположил, что такие «иррегулярные» аминокислоты образуются в результате модификации некоторых из соответствующих 20 «регулярных» аминокислот уже после включения последних в полипептидную цепь белка, которая строится только из 20 «регулярных» аминокислот. Таким образом, последовательность из четырех различных оснований в нуклеиновой кислоте должна определять последовательность из 20 различных аминокислотных остатков в белке. Так как в ДНК (и РНК) имеются четыре различных основания, в нуклеиновой кислоте число различных специфических областей, состоящих из n нуклеотидов, равно 4^n , если на вид специфических областей не наложено каких-либо ограничений. Необходимость закодировать такими специфическими областями 20 аминокислот приводит к тому, что должно выполняться неравенство $4^n \geq 20$. Поэтому специфические области в нуклеиновой кислоте должны состоять минимум из трех нуклеотидов (триплет), если, конечно, размер специфических областей один и тот же для всех аминокислот.

Вообще говоря, специфические области в нуклеиновой кислоте, определяющие соседние аминокислоты в полипептидной цепи белка, могли бы перекрываться, т. е. иметь один или несколько общих нуклеотидов. В случае перекрытия соседних специфических областей в нуклеиновой кислоте должна наблюдаться корреляция между соседними аминокислотами в белке. Такая корреляция может быть мала только в том случае, если размер перекрытия много меньше размера самой специфической области. Вопрос о корреляции соседних аминокислот в белках исследовался Ичасом, Гамовым и Ричем^{56,57}, проанализировавшими закон распределения соседних аминокислотных остатков у ряда различных белков, для которых известна последовательность остатков аминокислот по крайней мере у части полипептидной цепи. Они показали, что никакой корреляции между соседними остатками не наблюдается. Их выводы позднее были подтверждены Бреннером⁵⁸, и можно считать доказанным, что если код триплетен и универсален, т. е. един для всех организмов, то перекрытия между соседними триплетами быть не может.

Неперекрываемость соседних специфических областей следует и из анализа распределения аминокислотных остатков, измененных в цепи белка вируса табачной мозаики после того, как его РНК была подвергнута действию мутагенных факторов. Такой анализ, проведенный в 1960—1961 гг. Френкель-Конратом и Цугитой^{59,60} и Витманом⁶¹, показал, что в большинстве случаев мутация приводит к изменению в полипептидной цепи белка лишь одного аминокислотного остатка, не затрагивая соседних. Даже в тех случаях, когда в молекуле белка были изменены одновременно два аминокислотных остатка, они оказывались расположенными в различных частях молекулы.

На универсальность кода указывает вероятная видовая неспецифичность sРНК, участвующей в синтезе белка. Кроме того, универсальность кода следует из работы Суеки⁶², исследовавшего корреляцию между аминокислотным составом всех белков и нуклеотидным составом ДНК у различных бактерий. Обнаруженная для ряда аминокислот корреляция оказалась одной и той же для всех исследованных 11 видов бактерий.

Распространенность аминокислот в белках заметно отличается от случайной^{63,64}. Некоторые аминокислоты, например лейцин или аланин, встречаются примерно в десять раз чаще, чем такие, как цистеин или метионин. Такое превосходство одних аминокислот над другими

может быть вызвано либо тем, что в нуклеиновой кислоте одних специфических областей заведомо больше, чем других, либо за счет соответствия одним аминокислотам большего количества различных специфических областей. В последнем случае, когда одной аминокислоте могут соответствовать несколько различных специфических областей, код называется вырожденным.

Идея о том, что последовательность аминокислот в белке определяется последовательностью оснований в ДНК, сразу вызвала появление ряда конкретных схем кодирования. В частности, в работах Гамова^{54-56, 64} указывалось на то обстоятельство, что все 64 возможных триплета можно разбить на 20 групп, каждая из которых соответствует одной «регулярной» аминокислоте и отличается от остальных групп нуклеотидным составом. В этом случае одной аминокислоте может соответствовать несколько триплетов, различающихся только порядком нуклеотидов в них. Кроме этого, был предложен и ряд других схем кодирования^{56, 65}. Во всех этих схемах не предполагалось, что могут быть незначимые комбинации нуклеотидов, которые не соответствуют никакой аминокислоте.

Почти сразу же после появления первых идей о биохимическом коде выявились трудности, связанные с тем, что если имеется непрерывающийся код, то необходимо еще задать и способ чтения записи, ибо нет однозначного считывания записанной информации. Действительно, пусть информация записана в одонитевой последовательности нуклеотидов (например, если мы имеем дело с мРНК), у которой, естественно, имеется выделенное направление

... 123456789 ...

Тогда в случае непрерывающегося кода эту запись, вообще говоря, можно прочесть несколькими способами, например в случае триплетного кода — тремя способами:

... 123 — 456 — 789 ... ,

... 1 — 234 — 567 — 89 ... ,

... 12 — 345 — 678 — 9 ...

Однозначный результат может быть получен либо при наличии особых «запятых» между триплетами, либо при наличии начала отсчета и считывании сообщения, начиная с такой начальной точки последовательно триплет за триплетом. Ичас⁶⁶ предполагал даже, что может происходить считывание всех трех сообщений, получаемых из данного при сдвиге, и что белок образуется из трех полипептидных цепей, получающихся при считывании всех трех сообщений, закодированных в одной последовательности нуклеотидов. Однако это явно невозможно, так как во многих белках число входящих в их состав аминокислотных остатков не кратно трем.

Для получения однозначного результата при считывании информации Криком, Гриффитом и Оргелем⁶⁷ был предложен так называемый «код без запятых». В коде такого типа аминокислотам соответствуют только некоторые, «значащие», триплеты, такие, которые при перекрытии не дают значащих же триплетов. Например, триплет типа ААА не может быть значащим, так как перекрытие двух таких триплетов дает тот же триплет ААА. Из четырех различных оснований можно построить различными способами ровно 20 триплетов, удовлетворяющих этому условию. При таком коде считывание информации может происходить одновременно на всей молекуле нуклеиновой кислоты, а не последовательно,

начиная с ее конца. Дальнейшие исследования не подтвердили эту схему.

Дополнительные трудности в понимании однозначности считывания связаны с тем, что ДНК имеет двухнитевую структуру. Дело в том, что двухнитевая структура не имеет выделенного направления, так как нити в ней направлены в противоположные стороны. Поэтому, если для считывания информации с мРНК на белок направление считывания задано и неоднозначность можно устранить заданием начальной точки, для считывания информации с ДНК необходимо выделить направление этого считывания. При отсутствии выделенного направления считывания информации с ДНК можно достигнуть однозначности либо за счет идентичности информации в обеих нитях⁶⁸, либо за счет отсутствия информации в одной из нитей⁶⁹. Однако, если в одонитевой последовательности (мРНК) существует начальная точка, обуславливающая однозначность считывания информации, то такая начальная точка должна существовать и в ДНК. Такая начальная точка на ДНК может выделить нить, с которой должна считываться информация, выделяя одновременно и направление считывания и разделяя участки ДНК, ответственные за синтез различных полипептидных цепей.

Все указанные выше затруднения привели к тому, что несмотря на значительное число теоретических работ, посвященных проблеме биохимического кода начиная с 1954 г., никакой ясной схемы кодирования, имеющей убедительные преимущества перед другими, по существу, не появилось. Да и сами гипотезы о существовании такого рода кода нуждались в экспериментальных доказательствах. Первым прямым экспериментальным доказательством такого рода явилась доложенная на V Международном биохимическом конгрессе, проходившем в Москве в августе 1961 г., работа Ниренберга и Маттеи^{70,71}. Они выделили центрифугированием из кишечной палочки фракцию рибосом и смешали ее с надосадочной жидкостью, содержащей, в частности, sРНК и активирующие ферменты. Введя в эту систему аминокислоты и полирибоуридилловую кислоту (поли-У), полученную по методу Очоа и содержащую в качестве основания только урацил, они обнаружили появление полипептидов, состоящих из одного только сорта аминокислоты — полифенилаланина. Позднее ими было показано, что процесс синтеза полифенилаланина проходит через все обычные стадии синтеза белка — происходит активация фенилаланина, присоединение его к соответствующей sРНК и затем синтез полифенилаланина в рибосомах с участием поли-У в качестве матрицы⁷². Одновременно был доказан и одонитевый характер матрицы при синтезе белка в рибосомах, так как при добавлении в систему полиадениловой кислоты (поли-А), образующей двухнитевый спиральный комплекс с поли-У, синтез полифенилаланина подавляется.

Из работы Ниренберга и Маттеи следует, что специфическая группа нуклеотидов в мРНК, кодирующая фенилаланин, состоит из УУУ... Это исключает возможность существования «кода без запятых» Крика, Гриффита и Оргеля. То, что такого рода комбинация оказалась осмысленной, заставило предположить, что для однозначности считывания информации на нуклеиновой кислоте должны быть особые точки, определяющие начало считывания, и считывание должно происходить последовательно целыми специфическими группами нуклеотидов, начиная с такой точки. Указание на наличие особой начальной точки позволило Крику сформулировать гипотезу о типе кода. Он предположил, что на ДНК имеются точки начала отсчета, разделяющие области ДНК, ответственные за синтез разных полипептидных цепей, и что считывание происходит последовательно, начиная с начальной точки, целыми специфическими

группами нуклеотидов, причем величина специфических групп нуклеотидов одинакова для всех аминокислот. Он предположил также, что код вырожден, т. е. одной аминокислоте могут соответствовать несколько различных специфических групп нуклеотидов, и что существуют незначительные комбинации нуклеотидов, не соответствующие никаким аминокислотам. Для выяснения этой гипотезы им вместе с Бреннером, Барнетом и Ваттс-Тобин⁷³ были исследованы мутации, возникающие под действием акридино в специальном участке ДНК бактериофага Т4. Бактериофаги — бактериальные вирусы — содержат внутри белковой оболочки одну большую молекулу ДНК и небольшое число растворимых белков. При заражении бактерии фагом (Т-четные фаги паразитируют на кишечной палочке) внутрь бактерии проникает вся ДНК фага и лишь незначительная часть его белков. После заражения в бактерии начинается интенсивный синтез специфических РНК и белков, а на 6—7-й минуте синтез новой фаговой ДНК. Примерно через полчаса после заражения в бактерии накапливается около 100 (у некоторых фагов до нескольких сотен) зрелых фаговых частиц, растворяется ее оболочка (так называемый лизис бактерии) и вышедшие наружу фаги способны заражать новые бактерии. Если заразить одну бактерию одновременно двумя фагами, различающимися по некоторым генетическим признакам, то внутри бактерии происходит рекомбинация генетического материала этих фагов, при которой молекулы ДНК обоих фагов оказываются разорванными в одном и том же месте и часть молекулы одного фага оказывается сращенной с дополняющей ее до целой молекулы частью молекулы ДНК другого фага. По-видимому, такой разрыв может произойти на любом участке молекулы ДНК. Среди потомства исходных фагов получаются фаги, обладающие смешанным генетическим материалом и объединяющие в себе соответствующие признаки обоих исходных фагов. Очевидно, что при такого рода процессе два близко расположенных на молекуле ДНК одного из исходных фагов генетических участка будут разъединены между собой и попадут в молекулы ДНК различных вновь образующихся фагов с меньшей вероятностью, чем два удаленных друг от друга участка.

Экспериментальное выяснение частот рекомбинаций определенного типа и сравнение этих частот между собой дают возможность построить довольно детальные генетические карты фагов, на которых отмечено, в какой последовательности следуют вдоль молекулы ДНК фага генетические участки, ответственные за формирование того или иного признака, и даны расстояния между такими участками в условных единицах обратной частоты рекомбинаций соответствующего типа.

Если высевать фаг при большом разведении в слой с бактериями, то потомство каждого отдельного фага будет «выедать» бактерии, т. е. лизировать их, что обуславливает возникновение на этом слое бактерий «пустых» пятен, или «бляшек». Фаг Т4 стандартного («дикого») типа растет на различных штаммах кишечной палочки, в частности на штаммах В и К12. Фаг, у которого некоторая область генетического материала — локус *r*II — повреждена, на штамме К12 не растет, а на штамме В образует бляшки *r*-типа, отличающиеся от бляшек фага дикого типа большими размерами. Бензером⁷⁴ было показано, что локус *r*II состоит из двух независимо функционирующих участков — цистронов А и В, которые отвечают за синтез двух разных белков. Однако оба эти белка определяют один и тот же признак, так что поражение любого из этих двух цистронов приводит к *r*-типу. Еще ранее предполагалось⁷⁵, что при действии на ДНК специфических химических соединений — акридино в результате встраивания акридина в структуру ДНК при последующей

редупликации происходит либо удаление одного из нуклеотидов, либо добавление нуклеотида в полинуклеотидной цепи. При этом не исключена возможность, что удаляются или добавляются сразу несколько нуклеотидов. Крик указал, что если его предположения о типе кода верны, то удаление нуклеотида должно приводить к тому, что, начиная с этого места, чтение всех специфических групп сдвинуто на одно основание вправо (рис. 4, а). Добавление нуклеотида соответственно должно привести к сдвигу считывания на одно основание влево (рис. 4, б). Это должно сказываться на том, что, начиная с аминокислоты, определяемой специфической группой нуклеотидов, в которой произошло добавление или удаление нуклеотида, дальнейшая последовательность аминокислот в белке, определяемом этим цистроном, полностью изменена. При этом последовательность аминокислот в белке изменяется по-разному в зависимости от того, произошел ли сдвиг влево или вправо. Полученный белок существенно отличается от белка фага дикого типа, и с этим связано изменение поведения мутантного фага. Если существуют незначащие комбинации нуклеотидов, то не исключена возможность появления в некотором месте ДНК такой незначащей группы при каком-либо сдвиге. Возникновение незначащей группы нуклеотидов должно привести к обрыву в этом месте полипептидной цепи белка.

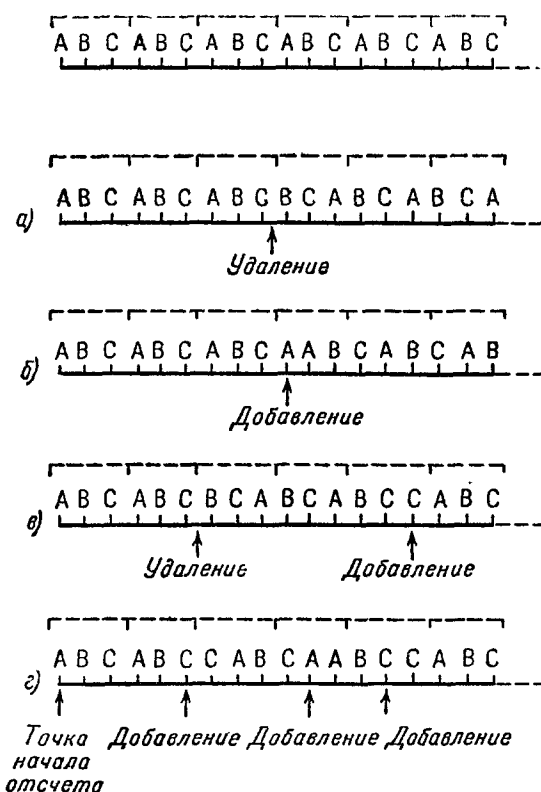


Рис. 4. Влияние добавления и удаления оснований в нуклеиновой кислоте на считывание информации.

Буквы А, В и С обозначают первый, второй и третий нуклеотиды триплета и могут представлять собой различные нуклеотиды. Принимается, что код триплетен. Считывание происходит слева направо целыми триплетами.

этом же цистроне, по концепции Крика должно приводить к сдвигу считывания за первой мутацией вправо на одно основание, а за второй — влево, также на одно основание. Очевидно, что в этом случае изменение считываемой информации должно происходить только на участке между этими мутациями (рис. 4, в). В результате аминокислотная последовательность должна измениться только между аминокислотами, определяемыми мутированными участками ДНК, в то время как остальная часть полипептидной цепи белка такого мутированного фага должна совпадать с полипептидной цепью белка фага дикого типа. Если наличие такого измененного участка полипептидной цепи не влияет существенно на функционирование белка, то появление второй мутации, восстанавливающей с некото-

рого места аминокислотную последовательность в белке, должно приводить к возврату фага к виду, лишь мало отличающемуся от дикого, так называемому псевдодикому типу.

Действительно, Криком, Бреннером, Барнетом и Ваттс-Тобин⁷⁵ было обнаружено существование в первой четверти цистрона В локуса *r*III фага T4 обратных мутаций, возвращающих фаг при наличии прямой акридиновой мутации к псевдодикому типу и не совпадающих по своему положению на генетической карте с прямой мутацией. Такие обратные мутации называются супрессорами, так как их появление подавляет действие прямой мутации. Наличие супрессоров было обнаружено также в локусе *h*III того же фага T4 Джинксом⁷⁶. Используя мутацию, обозначенную ими *FCO* и возникшую под действием одного из акридинов (профлавина), Крик с сотрудниками нашли 18 возникших спонтанно супрессоров этой мутации, расположенных в области, занимающей примерно $\frac{1}{10}$ цистрона. При этом нужно иметь в виду, что каждая из таких мутаций, взятая сама по себе, т. е. в том случае, когда она получена не в качестве обратной мутации *FCO*, а непосредственно на диком типе фага, опять-таки переводит фаг в *r*-тип, что вполне согласуется с представлениями Крика. Для каждой такой мутации, взятой сама по себе, можно образовать ее супрессоры, вновь переводящие фаг в псевдодиккий тип. Такие мутации, следовательно, являются супрессорами супрессоров мутации *FCO*. Крик и его сотрудники получили целый ряд мутаций подобного рода как спонтанно, так и под действием акридинов (рис. 5). Кроме того, они нашли целый ряд мутаций, подавляющих действие взятых отдельно супрессоров супрессоров (т. е., иными словами, получили супрессоры супрессоров супрессоров *FCO*) (рис. 5).

Пусть мутация *FCO* заключается в добавлении одного нуклеотида в полинуклеотидную последовательность ДНК фага. Тогда мутации-супрессоры должны заключаться в удалении одного нуклеотида, а супрессоры супрессоров — опять-таки в добавлении одного нуклеотида и т. д. Из гипотезы Крика следует, что все двойные мутанты, обе мутации которых возникли при удалении (или добавлении) нуклеотидов, должны быть *r*-типа. Действительно, все двойные мутанты, в которых путем рекомбинаций объединены мутация *FCO* и какой-либо из супрессоров ее супрессоров, имеют *r*-тип; *r*-тип имеют также и такие двойные мутанты, в которых объединены два разных супрессора мутации *FCO*, а также и те, которые обладают двумя супрессорами разных супрессоров мутации *FCO*.

Ограниченность области расположения супрессоров данной мутации объясняется существованием незначущих комбинаций нуклеотидов. Действительно, если при сдвиге считывания между удаленным и добавленным основаниями возникает незначущая комбинация, которая не соответствует никакой аминокислоте, то это приводит к разрыву полипептидной цепи белка и к отсутствию синтеза целой, функционально активной молекулы белка. Поэтому супрессоры некоторой мутаций могут быть расположены только в некотором ограниченном участке генетического материала вокруг этой мутации. Соответственно, по гипотезе Крика, не любая мутация, являющаяся удалением основания, может быть супрессором мутации, вызванной добавлением основания. Возврат фага к дикому или псевдодикому типу должен наблюдаться только у тех двойных мутантов, у которых в области, в которой сдвинуто считывание, при таком сдвиге не возникает незначущих комбинаций нуклеотидов. Естественно, что для сдвига вправо и влево величина такого рода области может быть совершенно различна. В случае, если в двойном мутанте будут объединены мутации, одна из которых условно соответствует добавлению

дикий тип, а в девяти случаях был получен *r*-тип (табл. II). Это опять-таки прекрасно подтвердило сделанные Криком предположения о характере кода.

Таблица II
Типы двойных мутантов

Скрещиваемые мутанты ↓	41	<i>FCO</i>	40	42	38	63	38
1	ω	ω	ω		ω		ω
86		ω	ω	ω	ω	ω	
9	<i>r</i>	ω	ω	ω	ω		ω
82	<i>r</i>		ω	ω	ω	ω	
21	<i>r</i>	ω			ω		ω
88	<i>r</i>	<i>r</i>			ω	ω	
87	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>			ω

Указаны типы двойных мутантов, получающихся при рекомбинации мутантов разного знака, обозначающих строки и столбцы. ω —дикий или псевдодикий тип, *r* обозначает *r*-тип. Подчеркнуты те двойные мутанты, которые использовались для выделения супрессоров.

Из гипотезы Крика следует, что если всем аминокислотам в молекуле ДНК соответствуют группы по *n* нуклеотидов, то при удалении (или добавлении) из молекулы нуклеиновой кислоты ровно *n* нуклеотидов, хотя бы они и были расположены в разных местах цистрона, считывание информации сдвинуто лишь между первым и *n*-м удаленными нуклеотидами, в то время как за последним удаленным нуклеотидом оно восстанавливается (рис. 4, 2). Поэтому если в области сдвига считывания не возникает незначущих комбинаций, то такой *n*-кратный мутант должен обязательно обладать псевдодиким типом. Были исследованы шесть таких тройных мутантов, в которых были объединены три мутации одного и того же знака. Все эти тройные мутанты обладали диким или псевдодиким типом. Таким образом, было показано, что код т р и п л е т е н, если, конечно, под действием профлавина каждый раз действительно удаляется или добавляется одно основание, а не несколько рядом стоящих.

Если же мутация получается при удалении или включении в ДНК нескольких оснований, то число нуклеотидов в кодовой единице больше, но кратно трем. Последнее, впрочем, мало вероятно.

По концепции Крика при удалении точки, определяющей начало считывания с данного цистрона, должно происходить как бы сшивание двух областей ДНК, определяющих синтез двух разных молекул белка, и в результате должен происходить синтез одной молекулы вместо двух. Это было проверено Криком и др.⁷³ с помощью делеции (выпадения целого куска генетического материала), затрагивающей частично оба цистрона А и В локуса λ II. Обычно эти цистроны функционируют независимо и повреждение одного из них не сказывается на функционировании другого. Было обнаружено, что акридиновые мутации в левом конце цистрона А, если они комбинируются с такой делецией, нарушают функционирование цистрона В, т. е. при наличии такой делеции эти цистроны оказываются связанными.

Таким образом, изложенные выше генетические эксперименты полностью подтвердили гипотезу Крика о характере кода. Код действительно оказался триплетным (или, в крайнем случае, длина кодовой ячейки кратна трем нуклеотидам), неперекрывающимся и вырожденным. Его вырожденность видна из того, что длина участка цистрона, на котором еще обнаруживаются супрессоры данной мутации, в отдельных случаях достигает нескольких десятков триплетных ячеек. Следовательно, несмотря на то, что при образовании исходной мутации начало отсчета сдвинуто и смысл следующих за этим сдвигом триплетов получается совершенно иным, чем у немутированной молекулы, на протяжении нескольких десятков триплетов часто не встречается ни одного незначущего. Такая ситуация может иметь место только при значительном вырождении кода.

Классический эксперимент Крика, Барнета, Бреннера и Ваттс-Тобин вскрыл характер кода и, в частности, подтвердил наличие в геноме специальных точек, определяющих начало считывания каждой генетической «фразы», которой соответствует один из синтезируемых белков. Однако такого типа эксперимент в принципе не мог дать сам «словарь», т. е. определить, какая конкретная комбинация из четырех возможных оснований соответствует каждому из аминокислотных остатков в цепи белка. Такая задача, или во всяком случае значительная ее часть, была решена биохимиками.

Нуклеотидный состав триплетов можно найти, исследуя включение аминокислот в белок в системе рибосом и надосадочной жидкости при добавлении к этой системе полирибонуклеотидов различного состава. Полирибонуклеотиды синтезируются из рибонуклеозиддифосфатов с помощью фермента полинуклеотидфосфорилазы, выделенного впервые Очоа⁷⁷. Нуклеотидный состав синтезированных таким способом полирибонуклеотидов зависит не от нуклеотидного состава затравки, которой в этом случае могут служить любые обрывки полирибонуклеотидов и даже динуклеотиды, а определяется только относительными концентрациями различных рибонуклеозиддифосфатов в среде. При этом последовательность нуклеотидов в цепи синтезированного таким способом полирибонуклеотида совершенно случайна. Если к ферменту Очоа добавить только один сорт рибонуклеозиддифосфата, то образуется полинуклеотид, содержащий только один вид основания, например поли-У. Аналогично в системе Очоа можно получить и полинуклеотиды, содержащие только два или только три вида оснований, например поли-УЦ или поли-УГЦ и т. д.

В конце 1961 и начале 1962 гг. Очоа⁷⁸⁻⁸¹ и параллельно Ниренберг и Маттеи⁸² провели серию исследований по изучению состава полипептидов, возникающих в системе рибосом и надосадочной жидкости при введении в нее различных полирибонуклеотидов. При изучении включения аминокислот в белок в присутствии полирибонуклеотидов, состоящих только из одного сорта оснований, оказалось, что поли-У стимулируют включение фенилаланина, а поли-А и поли-Ц не стимулируют включения никакой аминокислоты. Это означает, что фенилаланину соответствует триплет УУУ, а триплеты ЦЦЦ и ААА не соответствуют никаким аминокислотам. Было обнаружено также, что поли-УА стимулирует включение фен, илей, тир, лей, лиз, аспN; поли-УЦ — включение фен, про, лей, сер; поли-УГ — фен, гли, три, лей, вал, цис; поли-УАЦ — фен, сер, тир, лей, илей, лиз, аспN, про, тре, гис; поли-УЦГ — фен, гли, ала, арг, три, вал, цис, лей; поли-УАГ — фен, тир, глу, мет, вал, асп, аспN (табл. III). Зная, что фенилаланину соответствует триплет УУУ, можно, варьируя состав полирибонуклеотида и сравнивая включение

различных аминокислот с включением в белок фенилаланина, определить нуклеотидный состав триплетов, соответствующих определенным аминокислотам.

Таблица III

Относительная стимуляция включения аминокислот
в белок полирибонуклеотидами

Аминокислота	Отношение включения в белок фенилаланина к включению данной аминокислоты при добавлении указанных полинуклеотидов (усреднено по нескольким опытам)					
	УЦ 5 : 1	УА 5 : 1	УГ 5 : 1	УАЦ 6 : 1 : 1	УГЦ 6 : 1 : 1	УАГ 6 : 1 : 1
ала	—	—	—	—	31	—
арг	—	—	—	—	30	—
асп	—	—	—	—	—	41
аспN	—	15	—	15	—	20
вал	—	—	5	—	5	4
гис	—	—	—	29	—	—
гли	—	—	24	—	40	—
глу	—	—	—	—	—	64
лей	5	7	8	4	4	—
илей	—	5	—	6	—	—
лиз	—	32	—	46	—	—
мет	—	—	—	—	—	23
про	13	—	—	29	—	—
сер	4	—	—	64	—	—
тир	—	4	—	4	—	5
тре	17	—	—	11	—	—
три	—	—	20	—	24	—
цис	—	—	5	—	4	—

Так как последовательность оснований в искусственном полирибонуклеотиде случайна, очевидно, что, например, в случае поли-УЦ вероятность того, что один и тот же триплет содержит три урацила, относится к вероятности того, что триплет содержит два урацила и один цитозин, как

$$\frac{C_y^3}{C_y^2 C_{ц}} = \frac{C_y}{C_{ц}},$$

где C_y и $C_{ц}$ — относительные содержания в полимере урацила и цитозина.

Кополимер урацила и цитозина — поли-УЦ — стимулирует включение в белок не только фенилаланина, но и серина, лейцина, треонина и пролина. В случае поли-УЦ, в котором отношение $У : Ц = 5 : 1$, отношение включения в получившийся полипептид фенилаланина к серину близко к 5. Так как фенилаланин кодируется триплетом УУУ, то отсюда видно, что серину должен соответствовать триплет, состоящий из двух У и одного Ц. При этом порядок расположения У и Ц в триплете остается неизвестным. Таким методом были найдены триплеты, соответствующие всем 20 аминокислотам (табл. IV). Пока не известны все триплеты, соответствующие аминокислотам, так как включение аминокислот сравнивается с включением фенилаланина и, соответственно, определены только триплеты, содержащие У. Однако уже сейчас ясно, что существуют значащие триплеты (найлены пока ААА и ЦЦЦ) и что код вырожден, так как некоторым аминокислотам соответствует несколько триплетов. Например, включение в белок лейцина стимулируется как поли-УЦ, так и поли-УГ, так что лейцину должны соответствовать по крайней мере два триплета, один из которых содержит только У и Ц, а другой — только У и Г.

Таблица IV

Триплеты, соответствующие аминокислотам

Аминокислота	Триплеты (порядок нуклеотидов неизвестен)			Аминокислота	Триплеты (порядок нуклеотидов неизвестен)		
	1	2	3		1	2	3
фен ала арг асп аспN	УАГ УАА, УАЦ	УУУ УГЦ УГЦ		илей лиз мет про сер тир тре	УАГ УАЦ, УЦЦ	УУА УАА УЦЦ УУЦ УУА	УГГ УГЦ
вал гис гли глу глуN лей	УАЦ УАГ УГЦ*) УУА	УУГ УГГ УУЦ, УУГ	УГЦ	три цис		УГГ УУГ	

1—триплеты, определенные только в работах Очоа; 3—триплеты, определенные только в работе Ниренберга и Маттеи; 2—триплеты, определенные как Очоа, так и Ниренбергом и Маттеи.

*) Триплет, соответствующий глутамину, идентифицирован Очоа по обнаруженному Цугитой⁸⁷ замещения глуN → вал в белке ВТМ при обработке его РНК азотистой кислотой.

Для определения включения аминокислоты в белок в систему добавляются все аминокислоты, одна из которых содержит радиоактивную метку. Включение этой аминокислоты определяется по радиоактивности кислотонерастворимой фракции получившегося белка. Поэтому данные, полученные на основании анализа относительного содержания в полипептидах различных аминокислотных остатков, сами по себе могут быть в ряде случаев не вполне достоверны и нуждаются в проверке. Точность таких данных может быть увеличена при проведении нескольких параллельных опытов с применением полинуклеотидов с разными соотношениями входящих в них оснований.

Некоторая дополнительная информация относительно нуклеотидного состава триплетов, соответствующих аминокислотам, проверяющая сведения, найденные биохимическим путем, может быть получена и из данных по замещениям аминокислот в белке оболочки вируса табачной мозаики (ВТМ) в опытах типа описанных выше опытов Шрамма и Шустера.

Витманом⁶¹, а также и другими авторами^{59,60} найдены следующие замены аминокислот в белке ВТМ при обработке РНК ВТМ азотистой кислотой:

про → сер,	тре → сер,
про → лей,	илей → вал,
сер → лей,	глу или глуN → вал,
сер → фен,	глу или глуN → гли,
лей → фен,	асп или аспN → ала,
тре → ала,	асп или аспN → гли,
тре → мет,	асп или аспN → сер.
тре → илей,	

Большинство этих переходов согласуются с данными относительно нуклеотидного состава триплетов, соответствующих этим аминокислотам, полученными Очоа и Ниренбергом и Маттеи. Это, в частности, подтверждает универсальность кода. Выпадает только замещение треонина на метионин, но следует иметь в виду, что пока биохимическим путем определена только часть всех триплетов, соответствующих аминокислотам.

Так как при обработке азотистой кислотой в РНК могут происходить только изменения А на Г и Ц на У, то, как указал Гирер⁸³, все 64 триплета можно разбить на восемь групп по восемь триплетов в каждой (так называемые октеты) таким образом, что при действии азотистой кислоты может происходить преобразование триплетов друг в друга только внутри каждого октета (рис. 6).

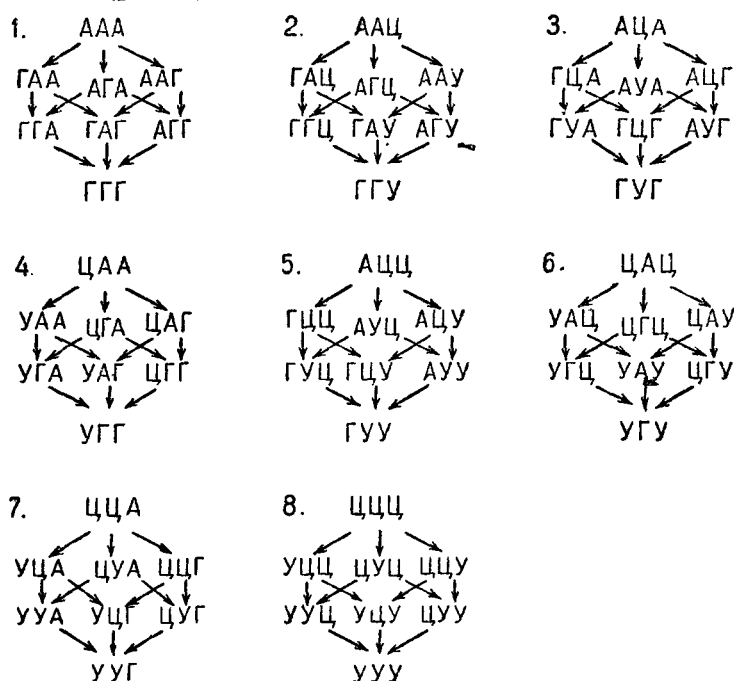


Рис. 6. Схема октетов⁸³.

В случае невырожденного кода, когда каждая аминокислота кодируется только одним триплетом, замещения аминокислот должны разбиваться в соответствии со схемой октетов на независимые группы. Попытка Витмана⁶¹ сопоставить наблюдаемые им замещения со схемой октета показала вырожденность кода еще до открытия этого факта Очоа, так как он обнаружил переходы между аминокислотами, которые должны лежать в разных октетах. При неполных данных относительно нуклеотидного состава триплетов и замещений аминокислот, имеющихся в настоящее время, пока еще нельзя произвести однозначного распределения аминокислот по октетам (за исключением октета 8).

Таким образом, в настоящее время определен не только общий характер биохимического кода, но и нуклеотидный состав ряда значащих и незначащих триплетов. Однако до сих пор не определен порядок следования друг за другом нуклеотидов в триплетах, соответствующих тем или иным аминокислотам. Этот порядок в принципе не может быть определен с помощью искусственных полирибонуклеотидов, полученных по методу Очоа, так как в процессе их биосинтеза отдельные нуклеотиды распреде-

ляются вдоль цепи полимера вполне случайным образом. Ситуация была бы совершенно другой, если бы удалось получить путем химического синтеза полирибонуклеотиды с заданным расположением оснований вдоль цепи. Однако эта задача пока не решена.

Представляется возможным определение порядка нуклеотидов в триплетах при исследовании замены участка полипептидной цепи белка у двойных акридиновых мутантов, полученных в опытах типа опыта Крика и его сотрудников. Действительно, когда фаг обладает как прямой мутацией, так и ее супрессором, то, как уже говорилось, на участке молекулы ДНК между мутацией и ее супрессором происходит сдвиг считывания генетической информации. Это проявляется в том, что аминокислота в измененном участке полипептидной цепи белка определяется триплетом, получающемся как перекрытие двух триплетов, соответствующих двум соседним аминокислотам в белке фага исходного дикого типа. Если известна последовательность аминокислот в белке исходного типа и в белке двойного акридинового мутанта и если известен нуклеотидный состав всех триплетов, соответствующих определенным аминокислотам, то очевидно, что можно получить ряд данных, позволяющих определить порядок нуклеотидов в триплетах. Однако для этого необходимо знать, какой именно белок соответствует определенному участку генетического материала, в котором находятся супрессоры акридиновых мутаций. Кроме того, этот белок надо уметь выделить и очистить от даже небольших примесей других белков. Только после этого в нем можно будет определить последовательность аминокислот как в белке исходного типа, так и в белках двойных мутантов.

Может быть, трудность, связанную с отсутствием метода синтеза полинуклеотидов с известным порядком оснований в них, можно обойти, если, как это предложил К. С. Михайлов, на конец полирибонуклеотида типа поли-У посадить известную группу нуклеотидов или даже один нуклеотид и определить концевые аминокислоты в полипептиде, синтезируемом на таком полинуклеотиде в системе Ниренберга и Маттеи. К сожалению, однако, заранее неизвестно, участвуют ли концевые группы полирибонуклеотидов в синтезе белка в такой системе.

Единственным биосинтезированным полинуклеотидом с известным чередованием оснований является полидезокси-АТ. Этот кополимер, синтезируемый ДНК-полимеразой, выделенной Корнбергом, обладает регулярно чередующейся последовательностью А и Т. Синтезированный на нем, как на матрице, с помощью РНК-полимеразы кополимер полирибо-АУ также обладает регулярно чередующейся последовательностью А и У. В системе Ниренберга и Маттеи такой полирибонуклеотид должен стимулировать включение аминокислот, кодируемых триплетами с вполне определенной последовательностью нуклеотидов — АУА и УАУ, если такие триплеты не являются незначимыми. Триплет УАУ соответствует либо тирозину, либо лейцину, либо изолейцину. Из трех триплетов, состоящих из У и 2А, по крайней мере два соответствуют аминокислотам — аспарагину и лизину. Это позволяет надеяться на положительный результат такого эксперимента, который может выяснить порядок нуклеотидов в двух триплетах.

Таким образом, видны пути определения порядка нуклеотидов в триплетах и есть все основания надеяться, что весьма скоро биохимический код, связывающий последовательность аминокислотных остатков в белке с последовательностью оснований на участке молекулы ДНК, определяющим строение этого белка, будет полностью раскрыт.

Помимо выяснения типа биохимического кода и точного значения отдельных входящих в него триплетов за последнее время достигнуты

значительные успехи в выяснении сложного механизма, управляющего считыванием этого кода. Этот механизм является составной частью системы управления белковым синтезом в клетке и, по-видимому, прежде всего влияет на синтез мРНК. Так, например, Хесиным и Шемякиным недавно выяснено, что при заражении фагом бактерии мРНК, синтезируемая уже в бактерии на попавшей в последнюю ДНК фага, различна на разных стадиях развития фага. В первые минуты после заражения фагом бактерии мРНК синтезируется на одних участках ДНК, а затем на других ⁸⁴.

Тонкие генетические эксперименты, проведенные в последние годы Жакобом и Моно, показали, что существуют случаи, когда кроме структурного гена, служащего матрицей для синтеза мРНК и определяющего последовательность аминокислотных остатков в соответствующем белке, имеется еще ген-оператор, расположенный рядом со структурным геном и, по-видимому, непосредственно осуществляющий управление синтезом мРНК. Кроме того, существует еще и ген-регулятор, который может лежать в другой части генома. Этот ген управляет синтезом какого-то специального вещества (возможно, опять-таки одной из фракций РНК), называемого репрессором, которое влияет на функционирование гена-оператора ⁸⁵.

Исключительно интересной и пока еще не выясненной является проблема «точки», отделяющей одну осмысленную «фразу» на молекуле ДНК, соответствующую структурному гену, от другой. Этот вопрос тесно связан с выяснением механизма действия фермента, осуществляющего считывание информации с ДНК на мРНК, — полимеразы РНК. В настоящее время механизмы действия ферментов, осуществляющих синтез на матрице, совершенно не изучены.

Известны три процесса синтеза на матрице: редупликация ДНК, синтез мРНК на ДНК и синтез белка на мРНК. Все эти процессы характеризуются тем, что матрицей служит отдельная нить нуклеиновой кислоты и синтез происходит последовательно в определенном направлении, задаваемом матрицей. О действии такого рода ферментов ничего не известно, и можно лишь высказать некоторые гипотезы, отнюдь не претендующие на достоверность. Возможно, что отдельная молекула фермента прикрепляется к матрице и движется вдоль нее. Это позволяет объяснить, в частности, почему всегда происходит синтез целой новой молекулы и не наблюдается отрыва от матрицы только частично синтезированной молекулы. Может быть, что при синтезе мРНК на ДНК фермент — полимеразы РНК — прикрепляется к особой «точке», разделяющей соседние независимо функционирующие участки ДНК (цистроны), и следует вдоль нити ДНК до тех пор, пока не достигнет следующей точки. Эти точки, существование которых показано Криком, возможно, являются специфическими последовательностями нескольких нуклеотидов, образованных из тех же четырех оснований, или содержат в себе какие-либо основания, редко встречающиеся в молекуле. При синтезе мРНК на ДНК, возможно, происходит синтез обеих нитей мРНК, дополнительных к обеим нитям ДНК (неизвестно, происходит ли синтез обеих нитей мРНК одновременно или последовательно). Однако наиболее вероятным представляется, что только одна из дополнительных нитей мРНК может служить матрицей при синтезе белка и, следовательно, только эта нить мРНК может включаться в рибосомы. Назначение второй нити пока неясно. Возможно, что эти дополнительные нити мРНК различаются концевыми группами, обусловленными характером «точек» на ДНК.

Недавно Бергом ⁸⁶ было показано, что по крайней мере в специальной системе, состоящей из выделенных из клетки рибосом, надосадочной

жидкости и ряда ферментов, в синтезе белка могут принимать участие только мРНК, синтезированные на двухнитевой молекуле ДНК; мРНК, синтезированная на денатурированной ДНК, не увеличивает включения аминокислот в белки рибосом. На основании этих данных можно думать, что точка на ДНК нормально функционирует лишь в двухнитевом состоянии, хотя синтез мРНК и происходит отдельно на каждой нити.

В настоящее время во всем мире проводятся широкие исследования деталей системы регуляции белкового синтеза, и можно ожидать, что в этом направлении уже ближайшие годы принесут новые крупные успехи.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

1. J. D. Watson, F. H. C. Crick, *Nature* **171**, 737 (1953).
2. J. D. Watson, F. H. C. Crick, *Proc. Roy. Soc. (London)* **A223**, 80 (1954).
3. J. Josse, A. D. Kaiser, A. Kornberg, *J. Biol. Chem.* **236**, 864 (1961).
4. R. Langridge, H. R. Wilson, C. W. Hooper, M. H. F. Wilkins, L. D. Hamilton, *J. Molec. Biol.* **2**, 19 (1960).
5. E. Chargaff, C. F. Crampton, R. Lipschitz, *Nature* **172**, 289 (1953).
6. N. Sueoka, *J. Molec. Biol.* **3**, 31 (1961).
7. E. Chargaff, в сб. «*Biol. Structure and Function*», Vol. 1, London — N. Y.: Acad. Press, 1961, стр. 67.
8. C. Levinthal, P. F. Davison, *J. Molec. Biol.* **3**, 674 (1961).
9. J. Cairns, *J. Molec. Biol.* **3**, 756 (1961).
10. J. K. Setlow, R. B. Setlow, *Proc. Natl. Acad. Sci. US* **46**, 791 (1960).
11. R. Sinsheimer, *J. Molec. Biol.* **1**, 43 (1959).
12. J. D. Watson, F. H. C. Crick, *Nature* **171**, 964 (1953).
13. J. D. Watson, F. H. C. Crick, *Cold Spring Hrb. Symp. Quant. Biol.* **18**, 123 (1953) (перевод в сб. «Проблемы современной цитофизиологии», М., ИЛ, 1956).
14. M. Meselson, F. W. Stahl, *Proc. Natl. Acad. Sci. US* **44**, 671 (1958).
15. J. Marmur, D. Lane, *Proc. Natl. Acad. Sci. US* **46**, 453 (1960).
16. P. Doty, J. Marmur, J. Eigner, C. Schildkraut, *Proc. Natl. Acad. Sci. US* **46**, 461 (1960).
17. J. Marmur, P. Doty, *J. Molec. Biol.* **3**, 585 (1961).
18. C. L. Schildkraut, J. Marmur, P. Doty, *J. Molec. Biol.* **3**, 595 (1961).
19. I. R. Lehman, M. J. Bessman, E. S. Simms, A. Kornberg, *J. Biol. Chem.* **233**, 163 (1958).
20. M. J. Bessman, I. R. Lehman, E. S. Simms, A. Kornberg, *J. Biol. Chem.* **233**, 171 (1958).
21. J. Adler, I. R. Lehman, M. J. Bessman, E. S. Simms, A. Kornberg, *Proc. Natl. Acad. Sci. US* **44**, 641 (1958).
22. A. Kornberg, *Science* **131**, 1503 (1960) (перевод в сб. «Современные проблемы биохимии», М., ИЛ, 1961).
23. P. Doty, H. Boedteker, J. R. Fresco, R. Haselkorn, M. Litt, *Proc. Natl. Acad. Sci. US* **45**, 482 (1959).
24. А. С. Спирин, *J. Molec. Biol.* **2**, 436 (1960).
25. F. H. C. Crick, *Symp. Soc. Exptl. Biol.* **12**, 138 (1958) (перевод в сб. «Биологическое воспроизведение макромолекул», М., ИЛ, 1960).
26. H. K. King, *Sci. Progr.* **49**, 703 (1961).
27. B. D. Hall, P. Doty, *J. Molec. Biol.* **1**, 111 (1959).
28. C. G. Kurland, *J. Molec. Biol.* **2**, 83 (1960).
29. A. Tissières, J. D. Watson, D. Schlessinger, B. D. Hollingworth, *J. Molec. Biol.* **1**, 221 (1959).
30. P. C. Zamecnik, E. B. Keller, *J. Biol. Chem.* **209**, 337 (1954).
31. M. B. Hoagland, *Biochim. biophys. acta* **16**, 288 (1955).
32. M. B. Hoagland, E. B. Keller, P. C. Zamecnik, *J. Biol. Chem.* **218**, 345 (1956).
33. L. I. Hecht, M. L. Stephenson, P. C. Zamecnik, *Proc. Natl. Acad. Sci. US* **45**, 505 (1959).
34. L. I. Hecht, P. C. Zamecnik, M. L. Stephenson, J. F. Scott, *J. Biol. Chem.* **233**, 954 (1958).

35. J. Bishop, J. Leahy, R. Schweet, *Proc. Natl. Acad. Sci. US* 46, 1030 (1960).
36. H. M. Dintzis, *Proc. Natl. Acad. Sci. US* 47, 247 (1961).
37. H. Schuster, G. Schramm, *Zs. Naturforsch.* 13b, 697 (1958).
38. A. Tsugita, D. T. Gish, J. Young, H. Fraenkel-Conrat, C. A. Knight, W. M. Stanley, *Proc. Natl. Acad. Sci. US* 46, 1463 (1960).
39. E. Volkin, L. Astrachan, *Virology* 2, 149; 433 (1956).
40. Э. Фолькин, Л. Астрахан, в сб. «Химические основы наследственности», М., ИЛ, 1960, стр. 556.
41. M. Nomura, B. D. Hall, S. Spiegelman, *J. Molec. Biol.* 2, 306 (1960).
42. M. Yčas, W. S. Vincent, *Proc. Natl. Acad. Sci. US* 46, 804 (1960).
43. F. Gros, H. Hiatt, W. Gilbert, C. G. Kurland, R. W. Risenbrough, J. D. Watson, *Nature* 190, 581 (1961).
44. R. W. Risenbrough, A. Tissières, J. D. Watson, *Proc. Natl. Acad. Sci. US* 48, 430 (1962).
45. S. Brenner, F. Jacob, M. Meselson, *Nature* 190, 576 (1961).
46. B. D. Hall, S. Spiegelman, *Proc. Natl. Acad. Sci. US* 47, 137 (1961).
47. S. Spiegelman, B. D. Hall, R. Strock, *Proc. Natl. Acad. Sci. US* 47, 1135 (1961).
48. H. M. Schulman, D. M. Bonner, *Proc. Natl. Acad. Sci. US* 48, 53 (1962).
49. S. B. Weiss, *Proc. Natl. Acad. Sci. US* 46, 1020 (1960).
50. J. Hurwitz, A. Bresler, R. Diringier, *Biochem. Biophys. Res. Commun.* 3, 15 (1960).
51. S. Ochoa, D. P. Burma, H. Kröger, J. D. Weill, *Proc. Natl. Acad. Sci. US* 47, 670 (1961).
52. J. J. Furth, J. Hurwitz, M. Goldman, *Biochem. Biophys. Res. Commun.* 4, 362 (1961).
53. M. Chamberlin, P. Berg, *Proc. Natl. Acad. Sci. US* 48, 81 (1962).
54. G. Gamow, *Biol. Med. Danske Vid. Selskab.* 22, № 3 (1954).
55. G. Gamow, *Nature* 173, 318 (1954).
56. G. Gamow, A. Rich, M. Yčas, *Advances in Biol. Med. Phys.* 4, 23 (1956) (перевод в сб. «Вопросы биофизики», М., ИЛ, 1957).
57. М. Ичас, в сб. «Теория информации в биологии», М., ИЛ, 1960, стр. 72.
58. S. Brenner, *Proc. Natl. Acad. Sci. US* 43, 687 (1957).
59. A. Tsugita, H. Fraenkel-Conrat, *Proc. Natl. Acad. Sci. US* 46, 636 (1960).
60. A. Tsugita, H. Fraenkel-Conrat, *J. Molec. Biol.* 4, 73 (1962).
61. H. G. Wittmann, *Naturwiss.* 48, 730 (1961).
62. N. Sueoka, *Proc. Natl. Acad. Sci. US* 47, 1141 (1961).
63. G. Gamow, *Biol. Med. Danske Vid. Selskab.* 22, № 8 (1954).
64. G. Gamow, M. Yčas, *Proc. Natl. Acad. Sci. US* 41, 1011 (1955).
65. A. Dounce, *Enzymologia* 15, 251 (1952).
66. M. Yčas, *Naturwiss.* 43, 197 (1956).
67. F. H. C. Crick, J. S. Griffith, L. E. Orgel, *Proc. Natl. Acad. Sci. US* 43, 416 (1957).
68. Е. В. Чавчаниядзе, *Биофизика* 3, 391 (1958).
69. S. W. Golomb, L. R. Welth, M. Delbruck, *Biol. Med. Danske Vid. Selskab.* 23, № 9 (1958).
70. М. В. Ниренберг, Дж. Маттеи, V Международный биохим. конгресс, Москва, 1961.
71. M. W. Nirenberg, J. H. Matthaei, *Proc. Natl. Acad. Sci. US* 47, 1558 (1961).
72. M. W. Nirenberg, J. H. Matthaei, O. W. Jones, *Proc. Natl. Acad. Sci. US* 48, 104 (1962).
73. F. H. C. Crick, L. Barnett, S. Brenner, R. J. Watts-Tobin, *Nature* 182, 1227 (1961).
74. С. Бензер, в сб. «Химические основы наследственности», М., ИЛ, 1960, стр. 56.
75. S. Brenner, L. Barnett, F. H. C. Crick, A. Orgel, *J. Molec. Biol.* 3, 121 (1961).
76. J. S. Jinks, *Heredity* 16, 153; 241 (1961).
77. С. Очоа, Л. Хелпел, в сб. «Химические основы наследственности», М., ИЛ, 1960, стр. 500.
78. P. Lengyel, J. S. Speyer, S. Ochoa, *Proc. Natl. Acad. Sci. US* 47, 1936 (1961).
79. J. S. Speyer, P. Lengyel, C. Basilio, S. Ochoa, *Proc. Natl. Acad. Sci. US* 48, 63 (1962).

80. P. Lengyel, J. S. Speyer, C. Basilio, S. Ochoa, Proc. Natl. Acad. Sci. US 48, 282 (1962).
 81. J. S. Speyer, P. Lengyel, C. Basilio, S. Ochoa, Proc. Natl. Acad. Sci. US 48, 441 (1962).
 82. R. G. Martin, J. H. Matthaei, O. W. Jones, M. W. Nirenberg, Biochem. Biophys. Res. Commun. 6, 410 (1962).
 83. А. Гирер, V Международный биохим. конгресс, Москва, 1961. Симпозиум III.
 84. М. Ф. Шемякин, Р. Б. Хесия, ДАН СССР (1962) (в печати).
 85. F. Jacob, J. Monod, J. Molec., Biol. 3, 318 (1961).
 86. W. B. Wood, P. Berg, Proc. Natl. Acad. Sci. US 48, 94 (1962).
 87. A. Tsugita, Protein, Nucleic Acid, Enzyme (Tokyo) 6, 385 (1961).
-