



УДК 576.1+551.7

## О ПРОИСХОЖДЕНИИ ГЕНЕТИЧЕСКИ КОДИРУЕМОГО СИНТЕЗА БЕЛКОВ

© 2021 г. С. П. Коваленко<sup>#</sup>

Институт микробиологии НАН Беларуси, Беларусь, 220141 Минск, ул. акад. В.Ф. Купревича, 2

Поступила в редакцию 14.12.2020 г.

После доработки 28.12.2020 г.

Принята к публикации 30.12.2020 г.

Каким образом в живых организмах синтезируются белки, имеющие определенную аминокислотную последовательность, перестало быть тайной после расшифровки генетического кода шесть десятилетий назад. Однако происхождение системы генетически кодируемого синтеза белков до сих пор остается предметом гипотез, экспериментально обоснованных лишь частично. На основе обзора работ преимущественно последнего десятилетия автор сформулировал гипотезу, в значительной степени конкретизирующую отдельные этапы формирования системы трансляции и генетического кода. Гипотеза базируется на концепции эволюционной редукции неоднозначности первоначального кода. Отличительные положения гипотезы: 1) появление аптамеров, обогащающих протоклетку редкими каталитически активными аминокислотами, и эволюционное превращение их в АРСазо-подобные рибозимы, что привело к возникновению операционного кода; 2) участие в формировании его и генетического кода сразу большого числа аминокислот (в том числе неканонических); 3) распознавание антикодонной петлей прото-тРНК *всего* кодона мРНК с самого начала кодируемого синтеза белков; 4) коэволюция операционного и генетического кодов, устранившая их первоначальную неоднозначность.

*Ключевые слова:* белки, пептиды, аминокислоты, аминоацилирование, генетический код, операционный код, РНК, тРНК, мРНК, АРСазы, ПТЦ, рибозимы

DOI: 10.31857/S0132342321060129

### ВВЕДЕНИЕ

Белки – основной компонент живой клетки, составляющий около половины ее сухой массы. Их роль в жизнедеятельности разнообразна: белки катализируют биохимические процессы, образуют различные структуры, выполняют двигательную функцию, переносят кислород и молекулы других полезных для организма веществ, защищают его от инфекций, выступают рецепторами и физиологическими регуляторами. Для выполнения каждой из этих функций необходимо, чтобы молекула соответствующего белка имела специфическую для нее последовательность аминокислот. Но копировать свои молекулы белки неспособны.

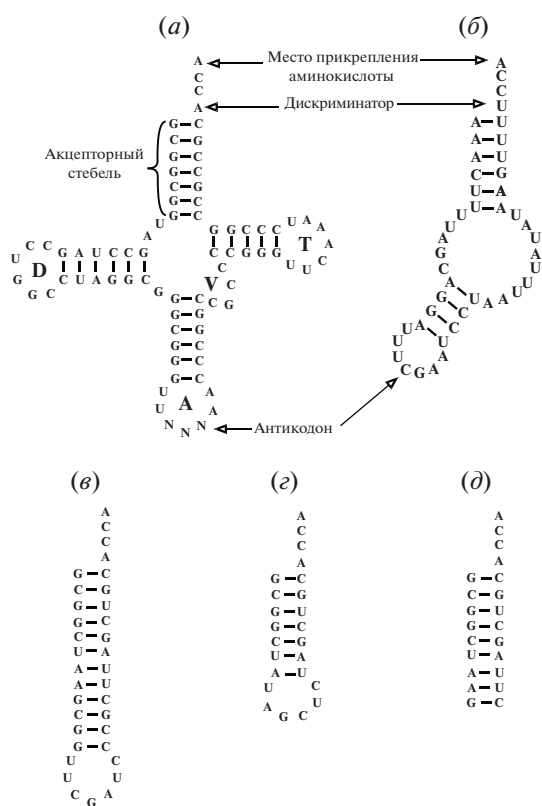
Ровно 60 лет назад благодаря работам М. Ниренберга [1] и С. Очоа [2] по расшифровке генетического кода стало ясно, как живые организмы синтезируют белки с определенной последовательностью составляющих их аминокислот. Син-

тез белковых молекул по своей сути – раскодирование генетической информации, заключенной в ДНК. Этот процесс называется трансляцией, т.е. переводом информации с нуклеотидного языка на язык аминокислотной последовательности в молекулах белков. Он происходит на рибосомах с помощью небольших молекул тРНК способных нести определенную аминокислоту и узнавать определенную тройку нуклеотидов (кодон) в молекулах мРНК, переносящих информацию с ДНК. Но чтобы нести присущую ей аминокислоту, молекула тРНК должна быть узнана особым ферментом – аминоацил-тРНК-синтетазой (АРСазой), – который и присоединит к ней эту аминокислоту (см., например, работу Х. Хартмана и Т. Смита [3]).

Опознав нужную аминокислоту, АРСаза активирует ее с помощью предварительно захваченной молекулы АТФ, синтезируя аминоациладенилат. На втором этапе АРСаза узнает определенную разновидность тРНК по структуре ее акцепторного стебля (двуспирального концевой участка – “черешка” этой трилистниковой молекулы) и присоединяет аминокислоту к ее 3'-концу сложноэфирной связью (рис. 1а). В редких случаях

Сокращения: АРСаза – аминоацил-тРНК-синтаза; рАРСаза – рибозим, предшественник АРСазы; ПТЦ – пептидилтрансферазный центр.

<sup>#</sup> Автор для связи: (эл. почта: spkovalenko2008@mail.ru).



**Рис. 1.** Вторичные структуры молекул тРНК и их заменителей: (а) – гипотетическая тРНК<sup>Phe</sup>; (б) – митохондриальная тРНК<sup>Arg</sup> нематоды *R. culicivora*; (в) – шпилькообразная миниспираль, содержащая акцепторный стебель тРНК<sup>Ala</sup> и петлю Т; (г) – микроспираль с акцепторным стеблем тРНК<sup>Ala</sup>; (д) – дуплекс, содержащий акцепторный стебель тРНК<sup>Ala</sup>. Буквы внутри петель обозначают их названия.

форма молекулы тРНК бывает иной (рис. 1б), но место присоединения аминокислоты остается тем же.

Так действуют АРСазы класса II [3], самые древние. АРСазам класса I недостаточно лишь узнавания акцепторного стебля, они проверяют еще и антикодонную петлю тРНК. Они имеют более сложную структуру и возникли, по-видимому, на более поздней стадии эволюции [3].

“Заряженная” аминокислотой молекула тРНК отправляется на рибосому и находит там свое место на молекуле мРНК за счет комплементарного спаривания своего антикодона с кодоном. Со следующим кодоном молекулы мРНК спаривается другая тРНК, несущая другую аминокислоту. Затем эти аминокислоты в пептидилтрансферазном центре (ПТЦ) рибосомы “сшиваются” именно в том порядке, в каком молекулы их тРНК последовательно (по мере перемещения мРНК внутри рибосомы) располагались на молекуле мРНК. Так генетическая программа, записанная

последовательностью нуклеотидов ДНК, определяет последовательности аминокислот во всех белках, а значит, их структуру и функцию.

В процессе раскодирования генетической информации можно выделить четыре ключевых участника: АРСазы, тРНК, мРНК и ПТЦ. У современных организмов их гораздо больше. Это компоненты рибосомы (три молекулы РНК и 70 белков), ~40 разновидностей молекул тРНК, 20 типов АРСаз (по числу аминокислот) и около десятка вспомогательных белков. Таким образом, для раскодирования генетической информации требуется очень сложная система (система трансляции), состоящая из многих компонентов, причем отсутствие любого из них нарушает ее функционирование. Поэтому ее назвали “неупрощаемой сложностью”.

Многочисленные попытки научного объяснения происхождения генетического кодирования синтеза белков предпринимались, начиная с довольно умозрительной, но поистине пророческой статьи Ф. Крика [4]. Однако только десятилетия усилий многих лабораторий обеспечили базу для создания гипотез ([3, 5–9] и др.), в той или иной степени экспериментально обоснованных.

Генетически кодируемый синтез белков мог возникнуть лишь в довольно развитой самовоспроизводящейся системе. Такой системой считают “мир РНК” в связи с тем, что молекулы РНК способны как к хранению и передаче наследственной информации, так и к выполнению каталитических функций. Формирование генетического кода происходило постепенно. Появление в предбиологической системе каждого нового компонента из перечисленных выше участников кодируемого синтеза белков должно было приносить ей какое-либо селективное преимущество. Иначе этот компонент не мог бы закрепиться эволюцией. Иногда для дальнейшего прогресса появления лишь одного компонента бывает недостаточно (а появление сразу двух – маловероятно). В таких случаях помогает экзаптация – привлечение элементов, эволюционно сложившихся для выполнения других функций, не связанных с трансляцией, обычно менее сложных [5, 6]. Например, рибозим, выполняющий какую-то каталитическую функцию, может продолжать свою работу, но часть его копий отвлекается для содействия вновь появившемуся компоненту системы. А при небольшом изменении этой копии в результате мутации бывший рибозим становится более полезным помощником новичку, порой даже утрачивая свою первоначальную функцию. Он становится участником другого процесса, экзаптируется в него. Происходит “переквалификация” биополимеров, этапы которой Ю. Вольф и Е. Куинин [6] назвали дубликацией и субфункционализацией.

В нашем предыдущем обзоре [10] описана сеть физико-химических процессов, приводящая к возникновению протоклеток, которые были только пассивным вместилищем генетического материала (РНК). Он активно эволюционировал, но практически не оказывал влияния на синтез компонентов мембран. Поэтому такие протоклетки еще нельзя было считать организмами, даже примитивными. Был сделан вывод, что превращение их в протоорганизмы, в которых генотип полностью осуществляет формирование фенотипа, могло произойти лишь после возникновения генетически кодируемого синтеза белков – более универсальных катализаторов и основных материалов клеточных структур. Настоящая статья – логическое продолжение предыдущего обзора.

### ФОРМИРОВАНИЕ БЕЛОК-СИНТЕЗИРУЮЩЕГО АППАРАТА ПРОТОКЛЕТОК

Исследования абиотических физико-химических процессов обеспечили довольно ясное представление об ассортименте органических веществ на молодой Земле, среди них – формальдегид и муравьиная кислота, жирные кислоты с небольшой длиной молекул (в пределах 8–10 атомов углерода), аминокислоты и гидроксикислоты, все четыре основания РНК и низкомолекулярные сахара. Доказана высокая вероятность абиотического образования нуклеотидов [11–13], гомохиральных олигонуклеотидов [14], полиэфиров, пептидов и депсипептидов [15–17], а также полимолекулярных комплексов – жирнокислотных мицелл и пузырьков [18], пузырьков и трубочек, образуемых липидоподобными пептидами [19, 20].

Популяция олигонуклеотидов состояла из молекул различной длины. Постоянное воздействие гидролиза и УФ-лучей расщепляло их цепочки. Наиболее устойчивыми к этим факторам, а потому доминирующими, были дуплексы, возникающие в процессе репликации, и шпилькообразные молекулы, состоящие из двуспиральных стеблей и соединяющих обе спирали петель из неспаренных нуклеотидов (рис. 1а, 1б). Такие шпильки дали начало многим видам РНК: рибозимам, аптамерам, тРНК, рибосомным РНК (в том числе ПТЦ).

Среди олигонуклеотидов, даже иногда довольно коротких, находились рибозимы, катализирующие разнообразные реакции: аминокислотирование [21, 22], лигирование [23] и фосфорилирование олигонуклеотидов [24], их матричную репликацию [25, 26], синтез пептидов [21, 22], образование гликозидной связи между основаниями РНК и рибозой [27, 28] и др. В зависимости от размеров и особенностей структуры молекул рибозимы могут ускорять спонтанные реакции в разной степени. Например, самый маленький из рибозимов 5'-GUGGC-3', выделенный в лабора-

тории М. Яруса [21], ускоряет спонтанную реакцию аминокислотирования в 25 раз, а 90-мерный рибозим Х. Суги – в  $1.6 \times 10^5$  раз [29], превосходя даже белковые ферменты. Но большинство рибозимов уступают соответствующим ферментам по активности примерно на 3 порядка [30].

Рибозимный мир не мог бы успешно эволюционировать, если бы не была решена проблема гомохиральности нуклеотидов. Короткие гомохиральные олигонуклеотиды, обладающие рибозимной активностью, могли возникнуть благодаря стереохимически избирательному воздействию монтмориллонита [14] и последующему удлинению. Однако этого было недостаточно для образования более сложной метаболической сети.

Поэтому одним из первых рибозимов (после лигаз и примитивных полимераз, обеспечивающих тиражирование относительно коротких олигонуклеотидов) должен был возникнуть D-рибозим, способный узнавать D-рибозу (скорее всего – в активированной форме) и присоединять ее к основаниям РНК. К сожалению, такой рибозим пока не получен. Но, с одной стороны, получены рибозимы, катализирующие реакцию 5-фосфо-D-рибозо-1-пирофосфата (pRpp) с основаниями РНК [27, 28] *in cis*, т.е. субстрат был присоединен к молекуле рибозима, и ему не требовалось узнавать pRpp среди множества других реагентов. Он узнавал только основания РНК. С другой стороны, в живых организмах существуют олигонуклеотиды (рибосвитчи), узнающие pRpp [31]. Объединение этих двух способностей в одном рибозиме дало бы желаемый результат, что, видимо, и осуществила Природа.

Первым следствием этого кардинального решения проблемы гомохиральности было вытеснение из абиотической среды не только L-нуклеотидов, но также заодно всех ксенонуклеотидов и их полимеров. Ведь катализируемый синтез гораздо эффективнее спонтанного абиотического, производящего все перечисленные формы нуклеотидов. Вторым следствием формирования мира РНК, основанного на D-рибозе, по мнению К. Тамуры [32], могла стать селекция L-аминокислот посредством аминокислотирования тРНК. В его экспериментах D-аминокислоты присоединялись D-рибозимами к молекулам тРНК в 4 раза медленнее, чем L-аминокислоты. Как выяснилось путем стереометрических измерений [33], затруднения были вызваны пространственной ориентацией D-аминокислот в активном центре рибозима. Так, аминокислоты D-рибозимы (далее ААР) при формировании белок-синтезирующего аппарата решили проблему гомохиральности белковых молекул.

Из перечисленных выше четырех ключевых компонентов системы трансляции центральную роль в формировании кодируемого синтеза белков

многие исследователи приписывают тРНК. Ее считают наиболее древней разновидностью РНК.

Современные молекулы тРНК имеют вторичную структуру, похожую на клеверный лист, где “черешок” – акцепторный стебель, а средний “листочек” – антикодонная петля (рис. 1а). Остальные два “листочка” (петли D и T) сближаются, взаимодействуют и образуют “локоть”, место сгиба, поэтому молекула приобретает L-образную форму. Однако и в экспериментах, и даже в природе вполне успешно действуют и более простые молекулы РНК. Например, в митохондриях многих многоклеточных обнаружены тРНК разной формы. Рис. 1б показывает вторичную структуру аргининовой тРНК митохондрий нематоды *Romanomermis culicivorax*. У нее есть акцепторный стебель и стебель с антикодонной петлей, а стебли с петлями D и T отсутствуют. Она имеет форму бумеранга и состоит всего из 44 нт. Причем у того же организма тРНК других аминокислот отличаются от нее. Сериновая тРНК утратила только D-петлю, а лейциновая – T-петлю [34]. И это вовсе не единичный случай.

Еще интереснее опыты П. Шиммеля. Синтезировав более десятка шпилькообразных мини- и микроспиралей из разных тРНК, содержащих акцепторный стебель и петлю T, он продемонстрировал их аминоацилирование соответствующими АРСазами класса II из *Escherichia coli*. Например, установлено, что АРСазе Ala достаточно узнать пару G ~ U акцепторного стебля аланиновой тРНК, чтобы присоединить к ней Ala [35]. Даже часть молекулы тРНК<sup>Ala</sup> (рис. 1в–1д), содержащая ее акцепторный стебель, безошибочно ацилировалась этой АРСазой [36]. Кроме того, в опытах без использования малой рибосомной субъединицы выяснилось, что такие аминоацилированные шпильки взаимодействуют с ПТЦ большой субъединицы не хуже, чем тРНК [37]. На основании своих работ П. Шиммель сделал выводы, что в акцепторном стебле тРНК существует код, узнаваемый АРСазами (названный операционным кодом), и что синтез белков в прото-ПТЦ сначала был некодируемым (не нуждался в мРНК и малой рибосомной субъединице).

Перечисленные факты заставляют усомниться в том, что клеверообразная форма тРНК была обязательно необходима при формировании системы трансляции, и достаточно было более простых, а значит – более абиотически доступных форм, вроде показанных на рис. 1в и 1г. По крайней мере, так считают многие исследователи [3, 7, 9, 38–40]. Это возможно при некодируемом синтезе белков. Но при кодируемом синтезе, как доказала И. Агмон [41], 3'-концы шпилек, связанных с соседними кодонами мРНК, слишком далеко расходятся друг от друга. Видимо, для кодируемого

синтеза были пригодны лишь изогнутые молекулы шпилек (например, как на рис. 1б).

Происхождение L-образной структуры из первичных шпилек трактуется по-разному. Одни исследователи считают современную форму тРНК продуктом димеризации шпильки ([38, 39, 42, 43] и ссылки в них), другие [9, 40, 44–47] развивают гипотезу слияния трех шпилек из 31 нт с последующими двумя симметричными делециями. Обе группы авторов обосновывают свою точку зрения данными исследований структур тРНК в филогенетическом плане. Вторая группа даже вычислила последовательность первичной тРНК, тРНК<sup>Pri</sup>, которая на 84% совпала с глициновой тРНК архей [45] (ее вторичная структура представлена на рис. 1а). Авторы заключили, что все остальные тРНК произошли от нее, а первой аминокислотой, включенной в генетический код, был Gly. Но обе стороны все-таки едины во мнении, что в формировании системы трансляции участвовали шпильки, а не L-образные молекулы.

Без участия рибозимов короткие пептиды могли синтезироваться спонтанно через промежуточную стадию депсипептидов [16, 17]. В этих пептидах доминировали Gly и Ala, самые распространенные. Содержание остальных аминокислот уступало им на 1–2 порядка величины. Такие пептиды не могли выполнять других функций, кроме структурных. Если в их состав попадала одна из полярных аминокислот, ее остаток играл роль “головы”, остальная часть цепочки была неполярным “хвостом”. Эти липидоподобные пептиды, как указывалось в предыдущем обзоре [10], служили довольно прочным остовом мембран протоклеток. Их комплексы с жирными кислотами имели большие преимущества над чисто жирнокислотными мембранами. Но абиотический синтез пептидов медленный: он требует нескольких циклов увлажнения/высушивания для превращения депсипептидов в чисто аминокислотные олигомеры. Скорость его намного уступает каталитическому синтезу.

Уже 5-звенный ААР М. Яруса катализирует образование пептидных связей [21]. Он сначала аминоацилирует 3'-конец тетрауклеотида 5'-GCCU-3', частично комплементарного самому себе. Аминокислота (в опытах – Phe) присоединяется к 2'-ОН рибозы. Далее она может мигрировать на 3'-ОН той же рибозы. Рибозим продолжает аминоацилировать ставший свободным 2'-ОН, затем новые молекулы Phe присоединяются уже к NH<sub>2</sub>-группам аминокислот, занявших оба гидроксильных рибозы. Так возникают пептидные связи. В опытах обнаружено возникновение пентапептидов.

Однако при действии этого рибозима не происходит перемещение пептидной цепочки на другую молекулу (рис. 2, слева). Синтезируемый пептид удлиняется на аминном конце молекулы,

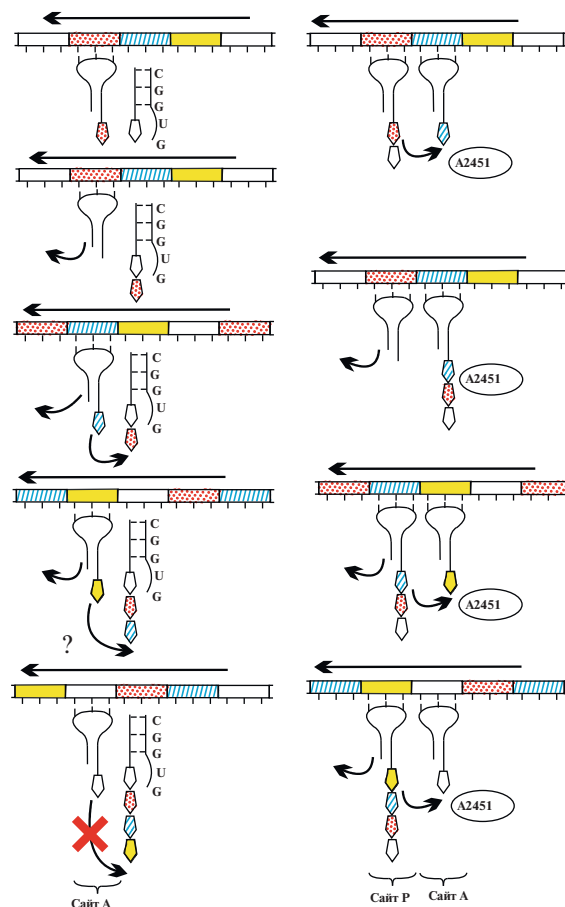
в то время как пептидилтрансферазная реакция удлиняет карбоксильный конец. Подобные рибозимы вряд ли могли обеспечить строгий порядок присоединения разных аминокислот. Для демонстрации невозможности синтеза длинных пептидов с определенной последовательностью аминокислот с помощью таких ААР мы изобразили на рис. 2 (слева) гипотетическую ситуацию, если бы в рибосоме каталитический центр содержал упомянутый 5-нт (или ему подобный) рибозим.

Допустим, что в проторибосому поступали активированные аминокислоты не в виде свободного аминоксиладенилата, как в опытах М. Яруса и соавт. [21], а в виде входящего в состав прототРНК на ее 3'-конце. Несмотря на постоянное движение молекулы мРНК, обеспечивающее смену в сайте А молекул протот-тРНК, растущая молекула пептида остается неподвижной и прочно связанной с рибозимом. Смещается только место присоединения к ней новой аминокислоты. Оно все больше отдаляется от активного центра рибозима (что, по-видимому, сначала замедляет дальнейший прирост цепи, а позже делает его невозможным). Но при пептидилтрансферазной реакции прирост пептидной цепочки происходит в одном и том же месте, не удаляясь от активного центра (рис. 2, справа).

В современной трансляции молекула тРНК с присоединенной к ней аминокислотой входит в сайт А ПТЦ. На  $\text{NH}_2$ -группу ее аминокислоты в результате пептидилтрансферазной реакции перебрасывается пептид с находящейся в сайте Р другой тРНК. Затем это смещение пептидной цепочки компенсируется возвратом ее снова в сайт Р вместе с тРНК и связанным с ней водородными связями кодоном мРНК, а из сайта Р удаляется дезацилированная тРНК. Перемещение тРНК, несущей удлинившийся пептид, из сайта А в сайт Р называется транслокацией. Она может осуществляться при отсутствии малой субъединицы и факторов элонгации [48, 49], что, вероятно, происходило на этапе некодируемого синтеза белков.

В случае с 5-нт ААР (или подобным) невозможна такая же слаженность перемещения мРНК и тРНК (если бы таковые участвовали в этом процессе) с постепенным нарастанием пептидной цепи. Причина, по-видимому, заключается в том, что этот пептид нарастает со своего *N*-конца, в то время как *C*-конец прочно связан с неподвижным рибозимом. В пептидилтрансферазной реакции пептид нарастает с *C*-конца, который постоянно движется. То он переносится на аминокислоту вновь прибывшей в сайт А молекулы тРНК, то возвращается вместе с нею на прежнее место в сайт Р. И эти перемещения согласованы с движением молекулы мРНК.

Механизм пептидилтрансферазной реакции, получивший название “протонный челнок”,



**Рис. 2.** Схемы удлинения молекул пептида при катализе 5-нт ААР (слева) и ПТЦ рибосомы (справа). Многоцветная лента, движущаяся влево (направление указано длинной стрелкой), обозначает молекулу мРНК, шпильки – молекулы тРНК, цветные пятиугольники – аминокислоты. Расцветка каждой аминокислоты соответствует расцветке соответствующего ей кодона на молекуле мРНК. Активные центры слева представлены нуклеотидами G и U, справа – нуклеотидом A2451. Слева при каждом акте присоединения новой аминокислоты к пептиду его конец все более отдаляется от активного центра. Его вхождение в активный центр возможно лишь при изгибе цепочки. Это происходит все реже по мере ее удлинения. Вопросительный знак означает проблематичность реакции, красный крест – ее невозможность. Схема демонстрирует невозможность кодирования синтеза длинных пептидов рибозимами, подобными 5-нт рибозиму М. Яруса, даже при встраивании их в белок-синтезирующий аппарат протоклеток.

описан в работах [50–52]. Он еще не вполне изучен, но основные участники каталитического процесса выявлены. Главную роль в нем играет аденозин в позиции 2451 большой рибосомной субъединицы. Особенно важна его группа 2'-ОН, отсутствие или блокирование которой ингибирует образование пептидной связи. Считается, что протон атакующей  $\text{NH}_2$ -группы аминоксил-тРНК, находящейся в сайте А, через несколько проме-



жуточных участников (в их числе – 2'-ОН A2451) переходит на 3'-О концевой аденозина тРНК в сайте Р. В результате пептидилный остаток отрывается от тРНК сайта Р и присоединяется к лишней протона аминокислоте аминокислоты в сайте А.

Факт активного участия группы 2'-ОН A2451 – убедительный аргумент, почему Природа выбрала именно РНК, а не другие нуклеиновые кислоты для катализа этой важнейшей реакции. Трео-НК и пиранозил-РНК вообще не имеют свободной группы 2'-ОН, т.к. она участвует в фосфатных мостиках, а у арабино-НК 2'-гидроксил повернут в противоположную сторону от активного центра. Так что наиболее вероятные кандидаты на роль заместителя РНК непригодны для катализа пептидилтрансферазной реакции.

A2451 сохраняется на протяжении всей эволюции в ПТЦ всех трех царств, как показало исследование их нуклеотидных последовательностей С. Фариаса с соавт. [53]. Это доказывает ключевую роль A2451 в биосинтезе белков.

Рибозимы с пептидилтрансферазной активностью были получены методом искусственной эволюции, причем их структура оказалась похожей на структуру природного ПТЦ [54]. Позже была синтезирована рибосомная РНК длиной 615 нт [55] – в 5 раз меньше, чем в большой субъединице бактериальных рибосом. Она совмещала способности к взаимодействию с тРНК и к проведению пептидилтрансферазной реакции.

Исследовав большую группу олигонуклеотидов, А. Йонат с соавт. [56] обнаружили, что некоторые из них, близкие по структуре рибонуклеиновым участкам ПТЦ, склонны к спонтанной димеризации. Эти олигонуклеотиды характеризуются локтеобразной вторичной структурой, встречающейся довольно часто у функциональных молекул, в том числе у тРНК. В результате димеризации образуется структура, как бы опоясывающая значительное пустое пространство – “карманы”, служащие для связывания и последующего перемещения ацилированных тРНК. Авторы считают, что прото-ПТЦ образовался именно таким путем.

С. Фариас с соавт. ([53, 57, 58] и ссылки в них) обнаружили, что последовательности некоторых участков молекул тРНК, сложенные в определенном порядке, гомологичны последовательности ПТЦ бактерии *Thermus thermophilus* на 50.53%. В последней из серии их работ [53] установлено, что этот ПТЦ составлен из олигонуклеотидных последовательностей (18–50 нт), которые принадлежали, согласно их расчетам, молекулам глициновой, пролиновой, фенилаланиновой, тирозиновой и глутаминовой прото-тРНК. Авторы заключили, что части молекул этих пяти

предшественниц соответствующих тРНК могли дать начало первичному ПТЦ.

Существование разных тРНК, когда еще не было ни ПТЦ, ни синтеза белков, можно объяснить тем, что они (точнее – их предшественники, прото-тРНК) выполняли другие полезные функции, например, рибозимов или аптамеров. Так считают многие ученые (например, [59, 60] и ссылки в них). Показано [60], что в бескислородной среде РНК (в том числе тРНК) катализируют реакции переноса электронов, служащие предпосылкой для метаболической активности в пребиотических средах.

Позже прото-тРНК были экзаптивированы в систему трансляции и вместе с прото-ПТЦ и ААР, неспецифически “заряжающими” прото-тРНК (вместо АРСаз) аминокислотами, образовали белок-синтезирующий аппарат протоклеток. Аминоацилированные прото-тРНК, хаотически входя одна за другой в прото-ПТЦ, обеспечили некодируемый синтез пептидов.

## ОПЕРАЦИОННЫЙ КОД – ПРОМЕЖУТОЧНЫЙ ЭТАП

Экспериментально установлено, что рибозимная активность может быть усилена аминокислотами и пептидами, действующими как кофакторы [5, 61, 62]. Молекулы РНК не столь богаты функциональными группами, как аминокислоты. А. Кун и А. Радвани ([63] и ссылки в ней) выбрали из Атласа каталитических сайтов данные об активных центрах разнообразных ферментов (гидролаз, трансфераз, оксидоредуктаз, лиаз, лигаз, изомераз), сосредоточившись на тех случаях, когда каталитическое действие производит одна аминокислота, без помощи других. Они установили, что чаще всего к таким аминокислотам (в убывающем порядке) относятся Glu, Asp, Lys, His, Arg, Tyr и Ser. У них в боковых радикалах содержатся следующие группы: карбоксильная, гидроксильная (ароматическая или алифатическая), аминогруппа и имидазольный цикл. Авторы предполагают, что эти аминокислоты и в составе олигонуклеотида, выполняющего вместо полипептида функции связывания субстрата и его позиционирования, могут катализировать реакцию как кофакторы. Подтверждением их предположения явились работы ([64] и ссылки в ней), показавшие резкое усиление каталитической активности дезоксирибозимов при включении в их структуру химическими методами имидазольных циклов, гуанидо- и аминогрупп.

Еще ранее установлено, что гистидин, по-видимому, за счет своего имидазольного кольца, усиливает активность дезоксирибозима, расщепляющего РНК. Другие аминокислоты никакого действия не оказывали [61]. Многие пептиды, име-

ющие в своем составе Lys и Arg, также повышают активность рибозимов ([62] и ссылки в ней). Каталитической активностью обладают даже короткие пептиды (Val-Asp, Ala-Asp [65], Ala-His [66], Ser-His [67] и His-Phe-Phe [68]), содержащие перечисленные аминокислоты. Для протоклеток особенно высокую ценность представляли именно они. Однако их тогда было на 1–2 порядка величины меньше по сравнению с Gly и Ala, если судить по аминокислотному составу метеоритов [69] и данным экспериментов, моделирующих абиотические условия [70].

Как протоклеткам, так и бесклеточным самореплицирующимся ансамблям рибозимов, изолированным в минеральных порах, для обретения селективного преимущества было необходимо захватывать и удерживать редкие каталитические аминокислоты, не позволяя им быстро покидать свой компартмент вследствие диффузии. Важность захвата и накопления редких аминокислот подчеркивается во многих гипотезах [5–8, 63].

Первым этапом на этом пути могло стать появление олигонуклеотидов, способных специфически захватывать каталитически активные аминокислоты. Такие захватчики, аптамеры, для восьми аминокислот действительно получены (см. работу [71] и ссылки в ней).

Простое депонирование ценных аминокислот большими молекулами РНК, даже обладающими несколькими центрами связывания, довольно дорого обходилось ансамблю, хотя и предоставляло селективное преимущество над другими. Большой выигрыш для него дала эволюция аптамеров в специфические ААР, присоединяющие узнаваемые аминокислоты ковалентной связью к своей (*in cis*) или другой (*in trans*) молекуле. Это можно рассматривать в качестве второго этапа формирования операционного кода. Такое соединение прочнее, а главное – при этом возникают случаи присоединения каталитической аминокислоты именно к тому олигонуклеотиду, рибозимная активность которого данной аминокислотой будет усилена. Этот этап тоже мог принести пользу ансамблю за счет ускорения полезной реакции. Но ценная аминокислота затрачивалась и на ацилирование посторонних молекул.

Третий этап – дальнейшая эволюция каждого из бывших аптамеров в направлении узнавания *только* того рибозима, который нуждается именно в данной аминокислоте. Это резко повысило конкурентоспособность ансамбля. Узнавание, скорее всего, происходило по концевому участку шпильки, куда присоединялась аминокислота [63] (аналогу акцепторного стебля современной тРНК). Был ли этот участок полностью двуспиральным или содержал неспаренные нуклеотиды (вроде теперешних ССА) – вопрос второстепенный. Главное – появилось соответствие между струк-

турой акцепторного стебля и присоединяемой к нему аминокислотой, которое и стало впоследствии операционным кодом. Бывшие аптамеры стали предшественниками АРСаз (рАРСазами), а усиливаемые аминокислотами рибозимы превратились в прото-тРНК благодаря постепенному эволюционному преобразованию.

Здесь важно подчеркнуть, что из-за несовершенства первичных структур, еще не прошедших эволюционную подгонку, могли быть нарушения процесса трансляции. Во-первых, существовали неспецифические ААР, которые могли ацилировать прото-тРНК не свойственной для нее аминокислотой. Это – ошибка белок-синтезирующего аппарата. Во-вторых, могло существовать несколько разновидностей прото-тРНК, имеющих одинаковые акцепторные стебли, но разные петли. Это создает вырожденность только генетического кода. В-третьих, рАРСазы могли узнавать похожие по структуре или физико-химическим свойствам аминокислоты одинаково и присоединять их к одной и той же разновидности прото-тРНК. Результат этого случая – неоднозначность обоих кодов. Анализируя первые результаты расшифровки генетического кода, классики молекулярной биологии, особенно К. Вёзе [72], считали, что первичный генетический код был неоднозначным, а аппарат кодируемого синтеза белков допускал много ошибок. Последующая эволюция должна была постепенно редуцировать эту неоднозначность и уменьшить вред от ошибок трансляции. Однако внимание ученых сосредоточилось лишь на второй части вопроса (“минимизации ошибок” при трансляции и в результате мутационных замен нуклеотидов). Проблему редукции неоднозначности кода лишь недавно осветил М. Барбьери (см. работу [73] и ссылки в ней), развивая эту концепцию в отношении и других биологических кодов (нейронного и языкового). К сожалению, достаточной степени конкретизации концепция все же не получила. Попытки ее развития в этом направлении изложена ниже.

Предшественники тРНК (шпильки) эволюционировали в направлении эффективного встраивания в “карманы” прото-ПТЦ, что ускоряло синтез пептидов. Этот синтез был неcodируемым, но в то время важна была не столько последовательность аминокислот в пептидах, сколько их аминокислотный состав. Пептиды, обогащенные каталитическими аминокислотами, давали преимущество ансамблю или протоклетке.

На этом этапе стало возможным массовое образование протоклеток на основе прежних самореплицирующихся ансамблей РНК, компартиментализованных в порах минералов. Ранее абиогенно синтезируемые пептиды были довольно однообразны по аминокислотному составу. По-

лярных аминокислот было мало, чтобы обеспечить достаточное количество материала для синтеза липидоподобных пептидов, образующих прочные мембраны (об их роли см. наш предыдущий обзор [10]). Обогащение отдельных ансамблей редкими аминокислотами резко изменило ситуацию к лучшему.

Имеются экспериментальные доказательства описанных процессов. Например, описаны рибозимы, способные активировать аминокислоты для аминокислотирования. 114-Мерный рибозим КК13, полученный в лаборатории М. Яруса [74], активирует свободные аминокислоты своей 5'-концевой трифосфатной группой, отщепляя пирофосфат и присоединяя их к оставшемуся фосфату. Как показала С. Мюллер [75], активированную аминокислоту с 5'-конца КК13 другой рибозим может затем перенести на свой 3'-концевой нуклеотид. Тем самым он освобождает КК13, предоставляя ему возможность вновь зарядить 5'-концевой фосфат энергией. Однако КК13 не способен узнавать АТР подобно АРСазам. Зато таким свойством обладают аптамеры, вроде описанного Дж. Шостаком [76] 40-мерного олигонуклеотида. При эволюции могла бы возникнуть гибридная форма, обеспечивающая селективное преимущество. КК13 мог так же фосфорилироваться, например, при реакции с диамидофосфатом [77]. К сожалению, КК13 не обладает специфичностью, активируя любые аминокислоты. Но во многих опытах обнаружена способность ААР узнавать определенные аминокислоты, хотя и не свободные, а активированные. 90-Мерный рибозим 77 М. Яруса узнает Phe-аденилат [78], рибозимы Х. Суги — цианометилловые эфиры Phe [29] или Gln [79].

Х. Суга с соавт. [29] получили 90-мерный рибозим (позже укороченный до 57 нт без потери функций), очень похожий на Phe-АРСаза. Он обладает следующими чертами АРСаз: узнает в качестве субстрата L-Phe; узнает молекулу тРНК по ее дискриминатору и двум цитозинам на 3'-конце; аминокислотировать исключительно 3'-гидроксил концевого аденозина тРНК; способен аминокислотировать не только молекулу тРНК, но и ее часть — миниспираль, содержащую только акцепторный стебель и петлю Т (как в опытах П. Шиммеля).

Специфичность узнавания L-Phe рибозимом достаточно высока: D-Phe узнавался в 5 раз хуже, остальные аминокислоты не узнавались, кроме родственного L-Тур. Отличие этого рибозима от АРСаз состоит только в неспособности активировать аминокислоту.

Х. Суга исследовал расположение в молекуле этого рибозима сайтов узнавания тРНК и Phe. Они оказались буквально рядом: связывающий Phe триплет UUA и триплет GGU, комплементарный дискриминатору и двум неспаренным цито-

зинам. Еще один триплет UUA находится пятью нуклеотидами раньше. Таким образом, концевой аденозин тРНК и активированная аминокислота оказываются в непосредственной близости друг от друга и реагируют.

Все ранее выделенные селекцией *in vitro* АРСазо-подобные рибозимы узнают на 3'-конце тРНК только нуклеотид-дискриминатор и два соседних неспаренных цитозинового нуклеотида, что роднит их с АРСазами класса II. Однако в недавней работе Х. Суги [80] получен рибозим из природного рибосвитча, узнающего определенную тРНК по *антикодону* и петле D. К нему была добавлена рандомическая последовательность, и рибозим селективировали *in vitro* на способность узнавать и присоединять активированный Phe к той же тРНК. Полученный рибозим уже больше похож на АРСазы класса I.

Таким образом, экспериментально доказано, что любая функция АРСаз может быть выполнена рибозимами и аптамерами. Но совместить эти функции в одном рибозиме за два десятилетия исследователям не удалось. Однако у Природы был в распоряжении неизмеримо больший срок для создания рАРСаз, которое положило начало операционному коду.

Современный операционный код очень сложен. Многие авторы лишь для упрощения утверждали, что он определяется только тремя первыми парами нуклеотидов в акцепторном стебле тРНК и что он не вырожден. В реальности на специфичность взаимодействия тРНК с соответствующей ей АРСазой могут влиять любые из семи пар нуклеотидов акцепторного стебля, неспаренный нуклеотид (дискриминатор), а также нуклеотиды, находящиеся вне акцепторного стебля, вроде аденина в позиции 20 (в D-петле). Возможных комбинаций лишь нуклеотидов акцепторного стебля почти на 3 порядка больше числа кодонов в генетическом коде. За миллиарды лет этот потенциал разнообразия реализовался, хотя только частично.

Д. Граур с соавт. [81] исследовали ~5000 тРНК из 21 вида архей и 102 видов бактерий и построили деревья детерминант, содержащихся в этих молекулах и соответствующих каждой аминокислоте, присоединяемой к их акцепторному стеблю. Они показали, что только для немногих аминокислот комплект нуклеотидных детерминант прост у всех изученных видов архей. Остальные аминокислоты в клетках разных видов архей, а тем более бактерий, определяются сложными и разнообразными наборами детерминант.

Высокая вырожденность операционного кода не позволяет разработать такую же четкую таблицу, как для генетического кода. Табл. 1 представляет собой попытку наглядно изобразить операционный код для изученных архей по данным



Таблица 1. Детерминанты операционного кода тРНК архей, узнаваемые АРСазами класса I и класса II

№	Амино-кислота	Узнаваемые нуклеотиды акцепторного стебля тРНК							
		73	72/1	71/2	70/3	69/4	68/5	67/6	66/7
1	His	C							
2	Cys	U					Py/A		
3	Thr	U					Pu/Py		
4	Arg	G			G или U/N				
5	Arg	G			C/N		N/не C		
6	Arg	G	N/A		C/N		N/C		
7	Arg	G	N/G		C/N	G/N	N/C	C/N	N/U
8	Asn	G				N/Pu		N/не G	
9	Asn	G		N/Py		N/Py	N/не A	N/C	
10	Asn	G				N/Pu	N/не C	N/G	He W-C
11	Asp	G		N/Py		N/Py	N/не A	N/G	
12	Lys	G		N/G		N/Py			
13	Lys	G	G/N		C/N	A/N	N/C		
14	Lys	G	G/N		C/N	G/N	N/C	C/N	N/G
15	Ser	G				N/Pu	N/C	N/G	
16	Ser	G				N/Pu	N/не C	N/G	W-C
17	Ser	G		N/Py		N/Py	A/N		
18	Ala	A		U/G					
19	Arg	A	N/G	Py/N		N/Py	N/G или C	C/C	
20	Arg	A	N/G	Py/N		N/Py	N/G	He C/N	
21	Arg	A	N/G	Py/N	W-C	N/Py	N/C	He C/N	
22	Arg	A	N/G	Py/N	U/не A	N/Py	N/C	He C/N	
23	Arg	A	N/G	G/N	N/C		N/Pu	N/A	
24	Arg	A	N/G	G/N	N/C	N/C	N/Py	N/C	
25	Gln	A	N/A или C	Py/N		Pu/N			
26	Gln	A	N/A	Py/N		C/N	N/C	N/C	
27	Glu	A	N/G	G/N	N/U		N/Pu		
28	Glu	A	N/G	G/N	N/Py		N/Py	N/Pu	
29	Glu	A	N/G	G/N	N/U	N/C	N/Py	N/C	
30	Gly	A		N/C	N/Pu				
31	Ile	A	N/A или C	Py/N		C/N	N/C	N/Pu	
32	Ile	A	N/G	Py/N		N/Pu	N/U	A или C/N	
33	Ile	A	N/G	Py/Pu	W-C	N/Py	N/C	He C/A	
34	Ile	A	N/G	Py/Pu	W-C	N/Py	N/C	He C/C	N/C
35	Leu	A	N/G	G/N	N/Pu				
36	Leu	A	N/G	Py/N	N/A	N/Pu	N/G или C	A или C/N	
37	f-Met	A	N/A или C	Py/N		C/N	N/Pu		
38	Met	A	N/G	Py/N		N/Pu	N/A	A или C/N	
39	Met	A	N/G	Py/N	G/не C	N/Py	N/C	He C/N	
40	Met	A	N/G	Py/Pu	C/G	N/Py	N/C	He C/N	
41	Met	A	N/G	G/N	N/C		N/Pu	N/G	
42	Met	A	N/G	G/N	N/Py	N/G	N/Py	N/C	
43	Phe	A		N/C	G /C				
44	Phe	A		N/G	W-C	N/G		C/N	

Таблица 1. Окончание

№	Аминокислота	Узнаваемые нуклеотиды акцепторного стебля тРНК							
		73	72/1	71/2	70/3	69/4	68/5	67/6	66/7
45	<i>Pro</i>	A		N/G	W-C			G/N	
46	<i>Pro</i>	A		N/G	W-C	N/C		C/N	
47	<i>Trp</i>	A	N/G	Py/N		N/Pu		G/N	
48	<i>Trp</i>	A	N/G	Py/N	N/G	N/Pu	N/G или C	A или C	
49	<i>Tyr</i>	A	N/C	G/N					
50	<i>Tyr</i>	A	N/C	Py/N		C/N	N/C	N/C	
51	<i>Val</i>	A	N/G	Py/N		N/Py	N/U		
52	<i>Val</i>	A	N/G	Py/N		N/Py	N/G или C	C/G	
53	<i>Val</i>	A	N/G	Py/Pu	U/A	N/Py	N/C	He C/N	

Примечание: курсивом выделены названия аминокислот, распознаваемых АРСазами класса I; N – любой нуклеотид, Pu – пурин, Py – пиримидин, W-C – комплементарная пара по Уотсону–Крику. Незаполненные ячейки – пары, не принимающие участия в узнавании тРНК АРСазами.

упомянутых авторов [81]. В нее ради упрощения не включены тРНК, встречающиеся лишь у немногих видов, а также такие, которые у разных видов ацилируются разными аминокислотами.

Хотя современный операционный код сложен для большинства аминокислот, наличие нескольких аминокислот с простым набором детерминант подсказывает, что его первоначальный вариант был гораздо проще. Поскольку археи – древнейшая из сохранившихся форм жизни, о первичном операционном коде целесообразно судить именно по ним. Впоследствии он усложнился в процессе эволюции благодаря возникновению новых рАРСаз и узнаваемых ими прото-тРНК, т.е. при редукции неоднозначности. В разделе “Козволюция операционного и генетического кодов” обсуждена роль операционного кода в формировании генетического кода и их совместная эволюция.

### ПЕРЕХОД ОТ НЕКОДИРУЕМОГО К КОДИРУЕМОМУ СИНТЕЗУ БЕЛКОВ

Для воссоздания картины появления и эволюции генетически кодируемого синтеза белков прежде всего следует установить, какие аминокислоты были доступны на древней Земле. На основании опытов С. Миллера и его последователей, а также изучения химического состава метеоритов сложилось мнение об ограниченности ассортимента абиотически синтезированных аминокислот. Однако некоторые ученые (работа [82] и ссылки в ней) предостерегали от поспешных выводов, считая, что упомянутые опыты не должны приниматься как неоспоримый критерий абиотичности аминокислот, поскольку не вполне адекватно отражают условия на древней Земле. Кроме того, возможные пути абиотического синтеза аминокислот

еще не полностью изучены. И действительно появляется все больше экспериментальных доказательств, что и другие аминокислоты могли образоваться в условиях, отличающихся от таковых в упомянутых опытах, но не противоречащих возможным на заре жизни. В работах лаборатории Дж. Сазерленда [83] исследованы пути абиотического синтеза 12 аминокислот, входящих в состав белков (кроме Ile, Lys, серосодержащих и ароматических). Видимо, аминокислоты, не содержащиеся в белках, авторов не интересовали. Поэтому они не обратили внимания, что вещество 38, полученное в их экспериментах, в результате реакции Штрекера превращается в 2,4-диаминомасляную кислоту [10]. А она в абиотических условиях могла играть важную роль, входя в состав положительно заряженных пептидов и, возможно, заменяя канонические аминокислоты [3].

В опытах, подобных экспериментам С. Миллера, но с добавлением сероводорода [70], обнаружено 10 белковых аминокислот, в том числе Ile и Met, отсутствующие в продуктах Дж. Сазерленда. His получен из эритрозы [84], хотя и с невысоким выходом. А в опытах, моделирующих условия Космоса [85, 86], использующих в качестве источников углерода полиароматические углеводороды, выявлено 13 аминокислот, включая His, Phe, Trp, Arg и Ile. Поскольку в реакционную смесь не добавляли источников серы, нельзя исключить также возможность образования серосодержащих аминокислот в условиях Космоса. В метеоритах различных классов найдены 15 аминокислот, т.е. все белковые, кроме амидов, Cys, Trp и Arg [69]. Обнаружены также 2,3-диаминопропионовая и 2,4-диаминомасляная кислоты [87].

В свете приведенных данных деление аминокислот на “ранние” и “поздние” выглядит неубедительно. К “поздним” аминокислотам, напри-

мер, еще недавно относили Cys. Но группе М. Поунера [88] удалось найти пребиотически вероятный путь синтеза этой аминокислоты. Поэтому делить аминокислоты целесообразно лишь на более и менее распространенные. Однако благодаря аптамерам, накапливающим концентрацию последних в прото клетках, и это различие в значительной мере стирается. Принять участие в начале кодируемого синтеза пептидов могла любая аминокислота, даже неканоническая. Эту же точку зрения обосновывают А. Нестеров-Мюллер с соавт. [89] в недавней оригинальной, но слишком уж умозрительной гипотезе “слияния четырех независимо сосуществовавших протокодов”. Даже они сами признают, что им неясно, как могли возникнуть эти протокоды.

К моменту начала формирования системы трансляции прото клетки имели мембраны из липидообразных пептидов и жирных кислот. Пептиды синтезировались абиогенно из наиболее распространенных аминокислот с включением иногда более редких. Прочность таких мембран была выше, чем из одних только жирных кислот, но не очень высокой.

П. Шиммель [90] предположил, что первые синтезы пептидов, осуществляемые примитивным ПТЦ, были не кодируемыми. Еще до появления рАРСаз существовали неспецифические ААР. Они присоединяли любую аминокислоту (даже неканоническую) к любому олигонуклеотиду, в том числе к прото-тРНК. Состав синтезируемых пептидов отражал статистику распространенности аминокислот. Появление рАРСаз обеспечило синтез пептидов с более высоким содержанием редких аминокислот. Последовательность их была случайной, но это не мешало лучшему выполнению функций. Поскольку в их состав чаще входили положительно и отрицательно заряженные аминокислоты, сцепление между пептидами было сильнее, и мембраны стали прочнее. Кроме того, эти пептиды могли быть в качестве кофакторов рибозимов эффективнее, чем отдельные аминокислоты. Например, пептиды, в составе которых имеются два остатка Asp или Glu, могли бы усилить активность рибозимов-РНК-полимераз за счет координирования иона магния [6]. Это ускорило бы репликацию олигонуклеотидов в прото клетках, синтезирующих такие пептиды.

В период не кодируемого синтеза, вероятно, интенсивно шла эволюция белок-синтезирующего аппарата прото клеток. Примитивный ПТЦ и шпилькоподобные носители аминокислот подгонялись относительно друг друга. В результате скорость встраивания двух шпилек в карманы ПТЦ и сближения их 3'-концов возрастала, ускоряя синтез. Значит, длина цепочек увеличивалась, их уже можно было называть полипептидами, воз-

можно, даже белками. Поскольку в синтез вовлекались и шпильки, ацилированные неспецифическими ААР, в молекулах полипептидов чередовались неполярные и каталитические полярные аминокислоты. Это позволило полипептидам образовывать складки и спиральные структуры, что могло обеспечить ферментативную активность. М. Хехт [91, 92] показал, что значительная часть рандомических полипептидов проявляет ее и даже способна заменять утраченные ферменты у ауксотрофных мутантов бактерий. Следовательно, не кодируемый белковый синтез приносил хотя бы небольшую пользу прото клеткам и способствовал вытеснению ими других форм, не способных к такому синтезу.

Для начала кодируемого синтеза необходимо было возникновение молекул мРНК. Этот вопрос наиболее подробно рассмотрен в гипотезе Х. Бернарда и У. Тэйта [7]. Авторы взяли на вооружение полузабытые работы прошлого века, однако не утратившие своей актуальности [93, 94]. В дальнейшем изложении будут описаны эти работы вместе с дополнительными аргументами.

В прото клетках реплицировались функциональные молекулы РНК – рибозимы (ПТЦ в их числе), аптамеры и прото-тРНК. Все они сразу же после репликации и разъединения дуплекса образовывали частично двуспиральные структуры, что предотвращало их быстрый гидролиз, в отличие от молекул со случайными нуклеотидными последовательностями, возникающих при синтезе *de novo*. Молекулы мРНК для своего функционирования, наоборот, не должны были обладать сильно выраженной тенденцией к сворачиванию в двуспиральные участки. Значит, они были бы уязвимы для гидролиза и нуждались в более надежном убежище, чем примитивные мембраны, образованные случайно возникающими пептидами и жирными кислотами. Но после формирования системы не кодируемого синтеза пептидов с разнообразным составом аминокислот условия внутри прото клеток стали благоприятными и для несвернутых РНК. Можно предположить также, что молекулы мРНК имели тенденцию к спариванию с одноцепочечными участками (петлями) шпилек, в том числе прото-тРНК. Так оба партнера взаимно защищались от гидролиза и УФ-лучей. Это свойство и определило первую функцию, которую стали выполнять несвернутые цепочки мРНК. Возможно, это было просто комплементарное взаимодействие ее с антикодонными петлями прото-тРНК, облегчающее встраивание последней в прото-ПТЦ. Видимо, мРНК выполняла сначала лишь роль стабилизатора [7], удерживая две молекулы прото-тРНК вместе до образования пептидной связи. Это и подтвердили упомянутые эксперименты. И. Йонак и И. Рыхлик [93] показали, что молекула поли(А) помогает укрепиться в сайте Р ПТЦ молекуле ли-

зиновой тРНК (имеющей антикодон UUU), в то время как поли(C) и поли(U,C) влияния не оказывали. А. Гнирке и К. Нирхаус [94] обнаружили аналогичное действие поли(U) на тРНК<sup>Phe</sup> (с антикодоном AAA), причем поли(A) в этом случае была неэффективна.

Удачно встроенные в прото-ПТЦ с помощью мРНК шпильки, заряженные аминокислотами, создали условия для первого кодируемого образования пептидной связи. Поскольку при некодируемом синтезе процесс транслокации уже был отлажен, начавшийся кодируемый синтез был продолжен, если со следующим кодоном мРНК успела связаться новая шпилька. И если синтезируемые пептиды были в какой-то степени полезны протоклетке, она получала селективное преимущество перед другими, а молекула мРНК закреплялась эволюцией.

В недавней работе лаборатории З. Бертона [9] описан вероятный сценарий начала кодируемого синтеза белка. По их мнению, существовал большой набор миниспиралей длиной 31 нт. Среди них одна была шпилькой с петлей D, вторая – с петлей T. Первая узнавала кодон с центральным нуклеотидом С благодаря его спариванию с нуклеотидом G, находящимся в ее петле. Вторая узнавала кодон UUU, комплементарный триплету AAA ее петли. Другие миниспиралы, участвующие в процессе, обеспечивали узнавание остальных кодонов первичных мРНК. Но дальнейшее изложение авторами процесса трансляции вызывает вопросы. Во-первых, по их мнению, все миниспиралы были ацилированы только одной аминокислотой – Gly. Причина такой избирательности аминоацилирования авторами не указана. Во-вторых, не объяснены причины последующего вымирания всех миниспиралей, кроме первых двух. Именно из этих выживших миниспиралей (две шпильки с петлей T и одна с петлей D), как считают авторы, сформировалась L-образная молекула тРНК.

Описанный З. Бертоном процесс начала кодируемого синтеза пептидов в целом вполне вероятен. Он совпадает с изложенным выше и в разделе “Операционный код – промежуточный этап”, которое уточняет, что в качестве миниспиралей в нем могли участвовать специфически ацилированные рАРСазами рибозимы, т.е. они были экзапированы в процесс трансляции. Это уточнение отвечает на вопрос о специфичности ацилирования миниспиралей. Только они были ацилированы не глицином, а *разными* аминокислотами, среди которых преобладали хотя и более редкие, зато более ценные аминокислоты. Кроме специфично аминоацилирующих рАРСаз, в ацилировании миниспиралей продолжали также участвовать и неспецифические ААР. А вымирания миниспиралей не требовалось, они эволюци-

онировали вместе с неспецифическими ААР в сторону специфического узнавания друг друга и превращению тем самым в новые прото-тРНК.

Таким образом, практически общепринятая точка зрения, что глицин был первой аминокислотой, вовлеченной в генетически кодируемый синтез пептидов, может быть неверной, как об этом прямо заявил В. Ма [8]. Он считает, что первыми были не простейшие, а именно сложные аминокислоты, которые могли служить кофакторами рибозимов. Похожие взгляды высказывались и другими исследователями [5, 63, 95].

В первых кодируемых синтезах полипептидов могли участвовать все аминокислоты, присоединяемые рАРСазами или даже неспецифичными ААР к шпилькам, *способным встраиваться* в прото-ПТЦ. Участие в них неспецифичных ААР создавало помехи, которые были позже преодолены. Но, с другой стороны, оно было полезным для синтеза длинных полипептидов, о чем будет подробно сказано ниже.

#### КОЭВОЛЮЦИЯ ОПЕРАЦИОННОГО И ГЕНЕТИЧЕСКОГО КОДОВ

Итак, рАРСазы создали пул специфически ацилированных аминокислотами шпилек, которые благодаря содействию мРНК стали удобно встраиваться в прото-ПТЦ. Там происходила пептидилтрансферазная реакция, “нанизывающая” аминокислоты одна на другую в том порядке, в котором различные прото-тРНК входили в прото-ПТЦ. А этот порядок определялся образованием комплементарных связей с кодонами молекулы мРНК, т.е. он был генетически кодируемым. Участвовала ли в этом процессе какая-либо молекула РНК, предшественница малой рибосомной субъединицы, – неизвестно. На первом этапе ее участия могло и не потребоваться. Попробуем разобраться, какие шпильки и аминокислоты могли быть первыми участницами этого процесса.

Для конкретности изложения используем тРНКом археи *Pyrococcus furiosus* [81], считающейся видом, наиболее близким к первоначальным формам жизни [44, 45]. Табл. 2 показывает структуру акцепторных стеблей всех ее тРНК. Заметим, что нуклеотидный состав их очень однообразен: содержание пар G : C превышает 95.9%. Всегда ли он был таким или его так сформировала эволюция – не ясно. Скорее всего, он изменился за 4 млрд лет. Поэтому будем отслеживать эту эволюцию пошагово, при замене одного нуклеотида в молекуле на каждом этапе описываемых событий, предположительно происходивших в то далекое время. Рис. 3 отражает эту последовательность эволюционных изменений акцепторных стеблей и антикодонов тРНК различных аминокислот. Таким образом, формирование генетиче-

Таблица 2. Операционный код у археи *Pyrococcus furiosus* (по данным Д. Граура с соавт. [81])

№	Аминокислота	Антикодон(ы)	Нуклеотиды акцепторного стебля тРНК							
			73	72/1	71/2	70/3	69/4	68/5	67/6	66/7
1	His	GUG	C	C/G	G/C	G/C	C/G	C/G	C/G	C/G
2	<i>Cys</i>	GCA	U	C/G	G/C	G/C	C/G	C/G	C/G	U/A
3	Thr	GGU, UGU	U	C/G	G/C	G/C	G/C	G/C	C/G	C/G
4	Thr	CGU	U	C/G	G/C	G/C	C/G	C/G	C/G	U/A
5	<i>Arg</i>	CCU	G	C/G	C/G	U/A	G/C	G/C	C/G	C/G
6	<i>Arg</i>	GCG	G	C/G	G/C	G/C	G/C	G/C	C/G	C/G
7	<i>Arg</i>	CCG, UCG	G	C/G	C/G	C/G	G/C	G/C	C/G	C/G
8	<i>Arg</i>	UCU	G	C/G	C/G	C/G	G/C	G/C	G/C	C/G
9	Asn	GUU	G	C/G	G/C	G/C	C/G	G/C	G/C	C/G
10	Asp	GUC	G	C/G	G/C	G/C	G/C	C/G	C/G	C/G
11	Lys	UUU, CUU	G	C/G	C/G	C/G	G/C	G/C	C/G	C/G
12	Ser	UGA	G	C/G	G/C	G/C	C/G	C/G	C/G	U/A
13	Ser	GCU	G	C/G	G/C	G/C	C/G	G/C	C/G	C/G
14	Ser	CGA, GGA	G	C/G	G/C	G/C	C/G	C/G	C/G	C/G
15	Ala	GGC, CGC, UGC	A	C/G	C/G	U/G	G/C	G/C	C/G	C/G
16	<i>Gln</i>	CUG, UUG	A	U/A	C/G	G/C	G/C	G/C	G/C	C/G
17	<i>Glu</i>	UUC, CUC	A	C/G	G/C	G/C	G/C	G/C	C/G	C/G
18	Gly	UCC, CCC, GCC	A	C/G	G/C	C/G	C/G	A/U	C/G	C/G
19	<i>Ile</i>	GAU	A	C/G	C/G	C/G	G/C	G/C	G/C	C/G
20	<i>Leu</i>	CAA, CAG, GAG, UAA, UAG	A	C/G	G/C	C/G	C/G	C/G	C/G	C/G
21	<i>f-Met</i>	CAU	A	U/A	C/G	G/C	C/G	C/G	C/G	G/G
22	<i>Met1</i>	CAU	A	C/G	C/G	C/G	G/C	G/C	G/C	C/G
23	<i>Met2</i>	CAU	A	C/G	G/C	G/C	C/G	C/G	C/G	C/G
24	Phe	GAA	A	C/G	C/G	C/G	C/G	G/C	C/G	C/G
25	Pro	CGG, UGG, GGG	A	C/G	C/G	C/G	C/G	G/C	G/C	C/G
26	<i>Trp</i>	CCA	A	C/G	C/G	C/G	C/G	C/G	G/C	C/G
27	<i>Tyr</i>	GUA	A	G/C	G/C	G/C	C/G	G/C	C/G	C/G
28	<i>Val</i>	CAC, UAC, GAC	A	C/G	C/G	C/G	G/C	G/C	G/C	C/G

Примечание: курсивом выделены названия аминокислот, узнаваемых APCазами класса I.

ского кода можно будет увидеть в тесном взаимодействии с эволюцией операционного кода.

Хотя в табл. 2 представлены данные по всем существующим сейчас аминокислотам, необходимо внести небольшие коррективы, учитывая условия ранней Земли. Вероятно, Asn и Gln отсутствовали из-за полного гидролиза цианогрупп их предшественников до Asp и Glu. Вряд ли существовали fMet и Trp, абиотический синтез которых не доказан. Так что можно не учитывать их в дальнейшем изложении. Могли существовать и включаться в пептиды также неканонические аминокислоты, о которых сейчас ничего не известно.

Логично предположить, что первичный операционный код был основан на простых наборах де-

терминант. Вполне возможно, что первые рAPCазы узнавали лишь нуклеотид-дискриминатор прототРНК (или, вероятнее, триплет из дискриминатора и двух нуклеотидов С, как APCазо-подобный рибозим X. Суги [29]). Некоторые могли различать, например, еще 2–3 нуклеотидные пары в акцепторном стебле. В то же время, как и рибозим X. Суги, узнающий не только Phe, но и родственный ему Tyr, первые рAPCазы могли не делать различий между родственными аминокислотами (в том числе и теми, которые сейчас отсутствуют в белках). Действительно, в табл. 2 видно, что участки из первых трех пар нуклеотидов в основном сходны у родственных аминокислот: Asp и Glu, Ser и Thr, Lys и Arg, а также почти всех гидрофобных аминокислот.

Аминокислота	Анти-кодон(ы)	Акцепторный стебель				
		73	72/1	71/2	70/3	69/4
His	NUG	C	C/G	G/C	G/C	C/G
Cys	NGU	U	C/G	G/C	G/C	C/G
Ser		U	C/G	G/C	G/C	C/G
Thr		U	C/G	G/C	G/C	C/G
Arg	NCG	G	C/G	C/G	C/G	G/C
Lys		G	C/G	C/G	C/G	G/C
Asp	NUC	A	C/G	G/C	G/C	G/C
Glu		A	C/G	G/C	G/C	G/C
Ala	Не определены					
Gly						
Pro						
Остальные						

Аминокислота	Анти-кодон(ы)	Акцепторный стебель				
		73	72/1	71/2	70/3	69/4
His	NUG	C	C/G	G/C	G/C	C/G
Cys	NGU	U	C/G	G/C	G/C	C/G
Ser		U	C/G	G/C	G/C	C/G
Thr		U	C/G	G/C	G/C	C/G
Arg		G	C/G	C/G	C/G	G/C
Lys	NCG, PyCU	G	C/G	C/G	C/G	G/C
Asp		A	C/G	G/C	G/C	G/C
Glu	NUC	A	C/G	G/C	G/C	G/C
Ala		A	C/G	C/G	C/G	C/G
Gly		A	C/G	C/G	C/G	C/G
Pro	NGC	A	C/G	C/G	C/G	C/G
Остальные		A	C/G	C/G	C/G	C/G

Аминокислота	Анти-кодон(ы)	Акцепторный стебель				
		73	72/1	71/2	70/3	69/4
His	NUG	C	C/G	G/C	G/C	C/G
Cys	NGU	U	C/G	G/C	G/C	C/G
Ser		U	C/G	G/C	G/C	C/G
Thr		U	C/G	G/C	G/C	C/G
Arg	PyCU, NCG	G	C/G	C/G	C/G	G/C
Lys	PyUU	G	C/G	C/G	C/G	G/C
Asp	NUC	A	C/G	G/C	G/C	G/C
Glu		A	C/G	G/C	G/C	G/C
Ala	NGC	A	C/G	C/G	C/G	C/G
Gly		A	C/G	C/G	C/G	C/G
Pro		A	C/G	C/G	C/G	C/G
Остальные	NAN	A	C/G	C/G	C/G	G/C

Аминокислота	Анти-кодон(ы)	Акцепторный стебель				
		73	72/1	71/2	70/3	69/4
His	NUG	C	C/G	G/C	G/C	C/G
Cys	NGA, PuCU	U	C/G	G/C	G/C	C/G
Ser		U	C/G	G/C	G/C	C/G
Thr	NGU	U	C/G	G/C	G/C	C/G
Arg	PyCU, NCG	G	C/G	C/G	C/G	G/C
Lys	PyUU	G	C/G	C/G	C/G	G/C
Asp	NUC	A	C/G	G/C	G/C	G/C
Glu		A	C/G	G/C	G/C	G/C
Ala	NGC	A	C/G	C/G	C/G	C/G
Gly	NCC	A	C/G	G/C	C/G	C/G
Pro	NGG	A	C/G	C/G	C/G	C/G
Остальные	NAN	A	C/G	C/G	C/G	G/C

Аминокислота	Анти-кодон(ы)	Акцепторный стебель				
		73	72/1	71/2	70/3	69/4
His	NUG	C	C/G	G/C	G/C	C/G
Cys	NCA	U	C/G	G/C	G/C	C/G
Ser	NGA, PuCU	G	C/G	G/C	G/C	C/G
Thr	NGU	U	C/G	G/C	G/C	C/G
Arg	PyCU, NCG	G	C/G	C/G	C/G	G/C
Lys	PyUU	G	C/G	C/G	C/G	G/C
Asp	NUC	A	C/G	G/C	G/C	G/C
Glu		A	C/G	G/C	G/C	G/C
Ala	NGC	A	C/G	C/G	C/G	G/C
Gly	NCC	A	C/G	G/C	C/G	C/G
Pro	NGG	A	C/G	C/G	C/G	C/G
Остальные	NAN	A	C/G	C/G	C/G	G/C

Аминокислота	Анти-кодон(ы)	Акцепторный стебель				
		73	72/1	71/2	70/3	69/4
His	NUG	C	C/G	G/C	G/C	C/G
Cys	NCA	U	C/G	G/C	G/C	C/G
Ser	NGA, PuCU	G	C/G	G/C	G/C	C/G
Thr	NGU	U	C/G	G/C	G/C	C/G
Arg	PyCU, NCG	G	C/G	C/G	C/G	G/C
Lys	PyUU	G	C/G	C/G	C/G	G/C
Asp	NUC	G	C/G	G/C	G/C	G/C
Glu		A	C/G	G/C	G/C	G/C
Ala	NGC	A	C/G	C/G	U/G	G/C
Gly	NCC	A	C/G	G/C	C/G	C/G
Pro	NGG	A	C/G	C/G	C/G	C/G
Остальные	NAN	A	C/G	C/G	C/G	G/C

Рис. 3. Последовательные этапы коэволюции операционного и генетического кодов. Примечание: курсивом выделены названия аминокислот, распознаваемых АРСазами класса I; N – любой нуклеотид, Pu – пурин, Py – пиримидин. Объяснения – в тексте.

кислот. Подытоживая сказанное, можно предложить следующую гипотезу.

Наибольшей ценностью для протоцитов обладали каталитически активные аминокислоты – полярные. Поэтому эволюционно закрепились рАРСазы, специфически ацилирующие такими аминокислотами рибозимы, ставшие впоследствии прото-тРНК. А поскольку каждому виду рибозимов для усиления каталитической активности необходим был определенный тип аминокислоты, быстро произошло разделение полярных аминокислот на группы по типу каталитической активности. Таких групп аминокислот было

четыре: 1) циклическая положительно заряженная (His), 2) нециклические положительно заряженные (Arg и Lys), 3) отрицательно заряженные (Asp и Glu), 4) незаряженные полярные (Ser, Thr, Cys).

Тогда число специфически присоединяемых к прото-тРНК групп аминокислот было бы равно четырем, хотя число самих аминокислот – восемь. Следовательно, операционный код был неоднозначным. His присоединялся к прото-тРНК с дискриминатором C. Ser, Thr и Cys, составив одну группу, получили дискриминатор U, Arg и Lys – G, а Asp и Glu – A. Такое распределение, вероятно,



было случайным: просто потому, что рибозимы, усилимые соответствующими аминокислотами, обладали именно данными дискриминаторами. Антикодоны в петлях этих рибозимов тоже могли быть иными, чем на рис. 3а. В гипотезе использованы антикодоны из табл. 2 (для аминокислот His, Thr, Arg, Asp и Glu), т.е. реально существующие. Если Ser, Thr и Cys узнавались рАРСазой и присоединялись к одной прото-тРНК, то и антикодон для всех был одинаков. То же самое справедливо для Arg и Lys.

Рассматривая табл. 2, заметим, что у архей *P. furiosus* все сказанное почти точно так и выполняется. Отклонений от гипотезы два: у Ser дискриминатор G, а не U, у Asp – G, а не A, хотя все остальные детерминанты (три пары акцепторного стебля) идентичны у родственных аминокислот. Однако логично предположить, что замена дискриминатора могла произойти на более позднем этапе, как это наблюдается в ряде случаев у современных архей. Например, *Methanopyrus kandleri* и *Haloferax volcanii* присоединяют Arg и Lys к тРНК, обладающим дискриминатором A вместо G, у большинства бактерий тРНК<sup>Gly</sup> имеют дискриминатор U (вместо A), а тРНК<sup>Gln</sup> – G (вместо A) [81].

Возможен альтернативный вариант для Asp и Glu: они могли узнаваться рАРСазой, которая бы различала, кроме дискриминатора G, вторую и третью нуклеотидные пары. Эти пары у тРНК обеих аминокислот одинаковы, но отличаются от пар тРНК Arg и Lys, имеющих тот же дискриминатор. Тогда шпильки с дискриминатором A были бы вообще не заняты.

Остальные аминокислоты первоначально (при неcodируемом синтезе, возможно – и на первых этапах кодируемого синтеза) присоединялись только неспецифическими ААР к любым шпилькам РНК, в том числе к шпилькам, обычно специфически ацилируемым полярными аминокислотами. При неcodируемом синтезе больших проблем не возникало, просто синтезируемые пептиды разбавлялись неполярными аминокислотами. Как отмечалось выше, могла быть даже польза от чередования в длинных полипептидах полярных и неполярных аминокислот. Да и при кодируемом синтезе, хотя бы в начальный период, неспецифичность ацилирования шпилек была полезной: благодаря ей не прерывался синтез полипептидной цепочки до полного считывания информации с мРНК, т.к. шпильки заполняли все кодоны мРНК. Ведь пока еще немногочисленные шпильки, ацилируемые рАРСазой, не покрывали всю протяженность мРНК своими антикодонами (для этого требовалось с учетом вобблинга как минимум 16 вариантов специфически аминокислоты ацилируемых прото-тРНК). Длинные полипептиды все же полезнее, чем короткие

обрывки, состоящие только из каталитических аминокислот.

Но если синтезируемый полипептид обладал какой-то полезной активностью, замена нужной аминокислоты на присоединенную неспецифическим ААР (обычно неродственную) приводила к утрате его функции. Если специфическое аминокислотное ацилирование преобладало, и “испорченных” белковых молекул в протоклетке было немного, все равно такая протоклетка проигрывала тем, которые избавились от вмешательства неспецифических ААР. Поэтому эволюция способствовала превращению неспецифических ААР в рАРСазы, специфически узнающие некаталитические аминокислоты (сначала – в виде группы) и акцепторные стебли шпилек, ранее не имевшие рАРСаз. Следовательно, шпильки, специфично ацилируемые рАРСазой, становилось все больше, и они захватывали новые кодоны.

Необходимость распознавать *отдельные* некаталитические аминокислоты, особенно Gly и Pro, которые сильно влияли на конфигурацию пептидной цепочки, появилась рано. Специфические для них (и, возможно, для Ala) рАРСазы, по видимому, отличали эти небольшие молекулы от более крупных с объемными боковыми гидрофобными радикалами. Если Ala, Gly и Pro узнавались одной рАРСазой, то и антикодон был на всех один. То же самое произошло и для аминокислот с объемными гидрофобными боковыми радикалами. Так, специфически аминокислоты ацилируемые прото-тРНК захватили шесть секторов (из 16) таблицы генетического кода, который был неоднозначным, как и операционный. На рис. 3 показана последовательность его формирования и постепенного захвата всех кодонов.

Исходная ситуация представлена на рис. 3а. Полярные аминокислоты образовали четыре группы, каждая из которых специфически присоединяется своей рАРСазой к тому рибозиму, чью каталитическую активность она усиливает. Остальные присоединяются неспецифическими ААР к любым шпилькам.

Рис. 3б показывает изменения в группе диаминокислот. При репликации прото-тРНК возникла мутантная копия с заменой G на U в антикодоне. Обе копии ацилируются той же рАРСазой. За счет мутации произошел первый захват новых кодонов мРНК, но не одной аминокислотой, а группой родственных аминокислот. Другой захват произошел группой из трех аминокислот в результате появления рАРСазы, узнающей Ala, Gly, Pro и одну из существовавших шпилек с дискриминатором A и антикодоном NGC.

На рис. 3в изображено разделение Arg и Lys или аминокислот, заменявших их в абиотический период. В антикодоне копии прото-тРНК, возникшей на предыдущем этапе, C был заменен на U.

Мутация произошла (возможно, не одна и значительно позже) и в рАРСаза, в результате чего один ее потомок стал узнавать только Arg и обе его прото-тРНК, а другой – только Lys и мутантную форму прото-тРНК. Так произошел первый акт редукции неоднозначности кода и заодно захват лизином нового сектора таблицы генетического кода.

Появилась рАРСаза для аминокислот с объемными радикалами. Она присоединяла всю их группу к прото-тРНК с дискриминатором А, что ознаменовало полный переход от неспецифического к специфическому аминоацилированию. Хотя предстоял еще длительный период преобразования неспецифических ААР в специфические рАРСазы, которое обеспечило постепенную редукцию неоднозначности обоих кодов.

Следующий этап был богат изменениями (рис. 3е). Ser и Cys отделились от Thr. Одна за другой произошли две мутации в антикодоне. Обе мутантные прото-тРНК (одна с заменой G на C, другая с заменой U на A) достались Ser и Cys, захватившим сразу два новых сектора. Но один из них был поделен с Arg. Thr сохранил прежний сектор. Акцепторный стебель прото-тРНК<sup>Thr</sup> мутировал: 4-я пара стала G/C. Значит, потом рАРСаза образовала две различающиеся копии. Одна из них перестала узнавать Thr, но узнавала прежний акцепторный стебель, другая перестала узнавать Ser и Cys, но начала узнавать измененный акцепторный стебель.

Не менее драматичные события произошли в группе Ala, Gly и Pro: они полностью размежевались. В антикодоне несущих их прото-тРНК произошли одна за другой две замены: в одной мутантной форме G был заменен на C, в другой – наоборот, соседний C заменен на G. У первой мутантной прото-тРНК вторая пара акцепторного стебля стала G/C, и ее стал узнавать измененный вариант рАРСазы, продолжающий присоединять к ней Gly. Вторую мутантную по антикодону форму стала узнавать рАРСаза Pro. Прото-тРНК с прежним антикодоном ацилировалась только Ala. Но в ее акцепторном стебле изменилась четвертая пара (нестандартная пара G ~ U, вероятно, возникла позже). Так что рАРСаза тоже должна была измениться и образовать три мутантных копии, чтобы различать эти изменения и более специфично узнавать ранее не различаемые аминокислоты.

Пятый этап привел к обособлению Cys. Замена G на C в антикодоне прото-тРНК позволила Cys захватить правый верхний угол таблицы генетического кода. На рис. 3д указан захват четырех кодонов, хотя в действительности могло быть иначе (лишь два верхних, занимаемых в настоящее время). Но не исключен и изображенный вариант. Позже Cys мог уступить нижние кодоны

своему селеновому аналогу или возникшему биосинтетическим путем триптофану. Cys мог также потерять кодон четвертого столбца, общий с Ser.

Наконец, мутация в антикодоне позволила разделиться Asp и Glu (рис. 3е). Мутантная форма прото-тРНК претерпела еще и мутацию дискриминатора. К ней стала присоединяться только Asp, а за Glu сохранились прежние рАРСаза и прото-тРНК. Это размежевание могло произойти уже после установления кодируемого синтеза белковых АРСаз и с их помощью, тем более что данные аминокислоты сейчас обслуживаются АРСазами разных классов.

Разделяться на классы рАРСазы тоже могли, как АРСазо-подобные рибозимы Х. Суги [29, 80], но, с точки зрения данной гипотезы, это не обязательно.

В результате описанных процессов большинство аминокислот захватили кодоны, занимаемые ими в настоящее время. Не описаны только распределение кодонов между крупными аминокислотами и захват верхнего квадранта третьего столбца таблицы генетического кода. Однако если существовал рибозим, усиливаемый тирозином, то первых специфических рАРСаз могло возникнуть не четыре, а пять, и этот квадрант был бы захвачен сразу.

Вполне вероятно, что мутации в антикодонах прото-тРНК намного опережали изменения рАРСаз, т.е. захват всех кодонов таблицы генетического кода произошел очень быстро. Более того, эти мутации (если они не нарушали рибозимную активность шпилек) могли произойти даже до появления соответствующих рАРСаз, которые с самого начала аминоацилировали 2–3 шпильки с одинаковыми акцепторными стеблями, но разными петлями (об этом говорилось выше). В таком случае сразу же после появления рАРСаз для всех некаталитических аминокислот специфически аминоацилируемыми прото-тРНК могли быть захвачены не шесть, а все 16 секторов таблицы генетического кода.

Размежевание родственных аминокислот шло медленнее, т.к. нужное изменение рАРСаз происходило, вероятно, в результате нескольких мутаций. В тексте и на рис. 3 эти события объединены в один этап лишь для краткости изложения. Поэтому родственные аминокислоты могли длительное время кодироваться в виде единых групп (может быть, даже до появления белковых АРСаз, которые лучше могли различать нюансы структуры родственных аминокислот).

Но и в состоянии еще не полной однозначности операционный и генетический коды обеспечивали синтез белков, обладающих структурой, достаточно успешно сохраняющей выполняемую ими функцию. Из-за включения в каждую позицию белковой молекулы то одной, то другой род-

ственной аминокислоты активности отдельных молекул с одинаковой функцией различались, но в допустимых пределах. Их функция не утрачивалась. А эта вариабельность активности была мотивом, стимулирующим редукцию неоднозначности.

Дальнейшая эволюция кодируемого синтеза белков происходила уже при активном участии белковых ферментов. Они обеспечили и размежевание крупных (или даже *всех?*) аминокислот, и добавление амидов с частичным вытеснением His и Lys, и решение проблем инициации и терминации белкового синтеза, и превращение шпилек в L-образные тРНК.

### ЗАКЛЮЧЕНИЕ

Представленная в данном обзоре гипотеза происхождения системы трансляции базируется на концепции эволюционной редукции неоднозначности первоначального кода. В ней использованы некоторые положения ранее высказанных гипотез [3–9]. Однако они значительно конкретизированы и дополнены.

Эволюционное превращение аптамеров, обогащающих протоклетки и бесклеточные ансамбли каталитическими аминокислотами, в рАРСады положило начало операционному коду. Он был неоднозначным из-за несовершенства узнавания рАРСадами отдельных аминокислот, что обусловило присоединение целой группы родственных аминокислот к одной и той же разновидности прото-тРНК. Выбор этой разновидности определялся ее каталитическими свойствами, усиливаемыми присоединяемыми к ней аминокислотами. Следовательно, причина выбора не связана ни с распространенностью аминокислот, ни с их взаимодействием с петлями шпилек. Значит, операционный код возник как случайность. А поскольку петли этих специфически аминокислизируемых прото-тРНК впоследствии стали антикодонными, то случайностью стал и первичный генетический код, т.е. прав Ф. Крик [4], характеризовавший его как “замороженную случайность”.

С другой стороны, последующее размежевание родственных аминокислот (т.е. редукция неоднозначности обоих кодов) происходило путем замен единичных нуклеотидов в антикодонах прото-тРНК. Хотя сами мутационные замены нуклеотидов носили случайный характер, кодоны родственных аминокислот располагались в таблице генетического кода не произвольно, а упорядоченно, занимая места в соседних секторах или даже в том же секторе, как Asp и Glu, Ile и Met, Leu и Phe. Следовательно, такая упорядоченность, вопреки мнению А. Куна и А. Радвани [63], не противоречит теории “замороженной случайности”, а сосуществует с ней. В то же время

именно эта упорядоченность обусловила явления “минимизации ошибок трансляции” и “минимизации вреда, наносимого мутациями”, лежащие в основе гипотезы [96]. Их следует рассматривать как побочный результат *редукции неоднозначности* обоих кодов.

Последовательность описанных выше этапов условна, многие события могли происходить в другом порядке. Например, не исключено, что крупные неполярные аминокислоты могли раньше начать специфически присоединяться к определенной шпильке. Ведь обогащение ими синтезируемых белков могло давать селективные преимущества протоклеткам благодаря лучшей связи таких белков с липидной частью мембран, что увеличило бы их прочность. Главное – описанная эволюция постепенно редуцировала неоднозначность обоих кодов параллельно и взаимосвязано.

Принципиальное отличие гипотезы от ранее опубликованных состоит в том, что она не нуждается в разделении аминокислот на ранние и поздние и не оставляет ни одну из них в стороне от процесса формирования кодируемого синтеза белков. При наличии неспецифических ААР больше десятка аминокислот, существовавших на заре жизни, не могли ждать своей очереди (согласно чьей-то гипотезе) для включения в кодируемый синтез белков. Любая из них (*и не одна!*) могла быть в числе первых участниц этого процесса.

Второе отличие – гипотеза предусматривает взаимодействие *всего* антикодона прото-тРНК со *всем* кодоном молекулы мРНК (с учетом вобблинга), начиная с самого раннего этапа формирования кода. Она не нуждается в предположениях, что первоначально прото-тРНК узнавали почему-то только второй нуклеотид кодона, а потом, якобы с повышением точности узнавания, “научились” узнавать первый и т.д. (гипотеза “2-1-3” [97]). Во-первых, неясно, что мешало взаимодействию с антикодоном не только второго, но и соседних нуклеотидов кодона. Во-вторых, сомнительно, что комплементарное взаимодействие только между двумя нуклеотидами достаточно прочно, чтобы удерживать тРНК, пока она встроится в карман ПТЦ и дождетса там завершения пептидилтрансферазной реакции. Ведь эти процессы не мгновенны, поэтому приход в сайт А аминокислизированной тРНК с некомплементарным антикодоном не приводит сразу же к пептидилтрансферазной реакции. Такая тРНК успеваеет покинуть сайт до ее начала. Задерживается только тРНК, образующая *несколько* водородных связей с кодоном мРНК.

Избыток связей (при взаимодействии тетраплетов) замедлил бы белковый синтез, поэтому эволюция выбрала триплеты.

Если бы гипотетические рибозимы, усиливаемые полярными аминокислотами, имели другие дискриминаторы и петли, все равно сформировался бы генетически кодируемый синтез белков. Но таблица генетического кода выглядела бы иначе. Описанная здесь схема возникновения и эволюции кода привела к *существующей* таблице. Однако возможны и другие схемы, дающие тот же результат.

В связи с затруднениями, которые испытывает гипотеза мира РНК (проблематичность ее самокопирования и легкая гидролизуемость) в новых работах ([98–101] и ссылки в них) приводятся все более убедительные аргументы, что миру РНК могли предшествовать миры на основе других информационных полимеров. Возможно, скоро будет экспериментально доказано, что некоторые из них легче, чем РНК, без участия ферментов образуются и самовоспроизводятся. Они также могут сочетать генетические и каталитические способности (быть ксенозимами) и обеспечивать не только достаточно точную репликацию своих молекул и, вероятно, транскрипцию с них молекул РНК, но и аминокислотирование и создание пептидной связи, как рибозимы М. Яруса. Но по причинам, указанным в разделе “Формирование белок-синтезирующего аппарата протоклеток”, способность осуществлять пептидилтрансферазную реакцию у них вряд ли будет обнаружена. Не исключено также, что их молекулы (подобно трео-НК и пиранозил-РНК) обладают меньшей гибкостью, чем РНК, и не пригодны для адапторной функции. Эти недостатки и обусловили в ходе эволюции их замену миром РНК, т.к. лишь на ее основе стало возможным формирование генетически кодируемого синтеза белков.

#### СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Статья не содержит описания исследований, выполненных с участием людей или использованием животных в качестве объектов.

#### КОНФЛИКТ ИНТЕРЕСОВ

Автор заявляет об отсутствии конфликта интересов.

#### СПИСОК ЛИТЕРАТУРЫ

1. Nirenberg M.W., Matthaei J.H. // Proc. Natl. Acad. Sci. USA. 1961. V. 47. P. 1588–1602. <https://doi.org/10.1073/pnas.47.10.1588>
2. Lengyel P., Speyer J.F., Ochoa S. // Proc. Natl. Acad. Sci. USA. 1961. V. 47. P. 1936–1942. <https://doi.org/10.1073/pnas.47.12.1936>
3. Hartman H., Smith T.F. // Life (Basel). 2014. V. 4. P. 227–249. <https://doi.org/10.3390/life4020227>
4. Crick F.H.C. // J. Mol. Biol. 1968. V. 38. P. 367–369. [https://doi.org/10.1016/0022-2836\(68\)90392-6](https://doi.org/10.1016/0022-2836(68)90392-6)
5. Szathmáry E. // Proc. Natl. Acad. Sci. USA. 1993. V. 90. P. 9916–9920. <https://doi.org/10.1073/pnas.90.21.9916>
6. Wolf Y.I., Koonin E.V. // Biol. Direct. 2007. V. 2. 14. <https://doi.org/10.1186/1745-6150-2-14>
7. Bernhardt H.S., Tate W.P. // Biol. Direct. 2010. V. 5. 16. <https://doi.org/10.1186/1745-6150-5-16>
8. Ma W. // Biol. Direct. 2010. V. 5. 65. <https://doi.org/10.1186/1745-6150-5-65>
9. Lei L., Burton Z.F. // Life (Basel). 2020. V. 10. 21. <https://doi.org/10.3390/life10030021>
10. Коваленко С.П. // Биоорг. химия. 2020. Т. 46. С. 486–504. [Kovalenko S.P. // Russ. J. Bioorg. Chem. 2020. V. 46. P. 675–691.] <https://doi.org/10.1134/S1068162020040093>
11. Powner M.W., Gerland B., Sutherland J.D. // Nature. 2009. V. 459. P. 239–242. <https://doi.org/10.1038/nature08013>
12. Roberts S.J., Szabla R., Todd Z.R., Stairs S., Bučar D.-K., Šponer J., Sasselov D.D., Powner M.W. // Nat. Commun. 2018. V. 9. 4073. <https://doi.org/10.1038/s41467-018-06374-z>
13. Becker S., Thoma I., Deutsch A., Gehrke T., Mayer P., Zipse H., Carell T. // Science. 2016. V. 352. P. 833–836. <https://doi.org/10.1126/science.aad2808>
14. Jheeta S., Joshi P.C. // Life (Basel). 2014. V. 4. P. 318–330. <https://doi.org/10.3390/life4030318>
15. Chandru K., Mamajanov I., Cleaves H.J., Jia T.Z. // Life (Basel). 2020. V. 10. 6. <https://doi.org/10.3390/life10010006>
16. Forsythe J.G., Petrov A.S., Millar W.C., Yu S.-S., Krishnamurthy R., Grover M.A., Hud N.V., Fernández F.M. // Proc. Natl. Acad. Sci. USA. 2017. V. 114. P. E7652–E7659. <https://doi.org/10.1073/pnas.1711631114>
17. Frenkel-Pinter M., Haynes J.W., Martin C., Petrov A.S., Burcar B.T., Krishnamurthy R., Hud N.V., Leman L.J., Williams L.D. // Proc. Natl. Acad. Sci. USA. 2019. V. 116. P. 16338–16346. <https://doi.org/10.1073/pnas.1904849116>
18. Black R.A., Blosser M.C. // Life (Basel). 2016. V. 6. 33. <https://doi.org/10.3390/life6030033>
19. Vauthey S., Santoso S., Gong H.Y., Watson N., Zhang S.G. // Proc. Natl. Acad. Sci. USA. 2002. V. 99. P. 5355–5360. <https://doi.org/10.1073/pnas.072089599>
20. Childers W.S., Ni R., Mehta A.K., Lynn D.G. // Curr. Opin. Chem. Biol. 2009. V. 13. P. 652–659. <https://doi.org/10.1016/j.cbpa.2009.09.027>
21. Turk R.M., Chumachenko N.V., Yarus M. // Proc. Natl. Acad. Sci. USA. 2010. V. 107. P. 4585–4589. <https://doi.org/10.1073/pnas.0912895107>
22. Yarus M. // Philos. Trans. R. Soc. Lond. B. Biol. Sci. 2011. V. 366. P. 2902–2909. <https://doi.org/10.1098/rstb.2011.0139>

23. *Vlassov A.V., Kazakov S.A., Johnston B.H., Landweber L.F.* // *J. Mol. Evol.* 2005. V. 61. P. 264–273. <https://doi.org/10.1007/s00239-004-0362-7>
24. *Moretti J.E., Müller U.F.* // *Nucleic Acids Res.* 2014. V. 42. P. 4767–4778. <https://doi.org/10.1093/nar/gkt1405>
25. *Wochner A., Attwater J., Coulson A., Holliger P.* // *Science.* 2011. V. 332. P. 209–212. <https://doi.org/10.1126/science.1200752>
26. *Horning D.P., Joyce G.F.* // *Proc. Natl. Acad. Sci. USA.* 2016. V. 113. P. 9786–9791. <https://doi.org/10.1073/pnas.1610103113>
27. *Lau M.W.L., Unrau P.J.* // *Chem. Biol.* 2009. V. 16. P. 815–825. <https://doi.org/10.1016/j.chembiol.2009.07.005>
28. *Martin L.L., Unrau P.J., Müller U.F.* // *Life (Basel).* 2015. V. 5. P. 247–268. <https://doi.org/10.3390/life5010247>
29. *Saito H., Watanabe K., Suga H.* // *RNA.* 2001. V. 7. P. 1867–1878. <https://doi.org/10.1017/S1355838201013164>
30. *Janzen E., Blanco C., Peng H., Kenchel J., Chen I.A.* // *Chem. Rev.* 2020. V. 120. P. 4879–4897. <https://doi.org/10.1021/acs.chemrev.9b00620>
31. *Knappenberger A.J., Reiss C.W., Strobel S.A.* // *eLife.* 2018. V. 7. e36381. <https://doi.org/10.7554/eLife.36381>
32. *Tamura K.* // *J. Mol. Evol.* 2019. V. 87. P. 143–146. <https://doi.org/10.1007/s00239-019-09897-1>
33. *Ando T., Takahashi S., Tamura K.* // *Nucleic Acids Res.* 2018. V. 46. P. 11144–11152. <https://doi.org/10.1093/nar/gky909>
34. *Jühling T., Duchardt-Ferner E., Bonin S., Wöhnert J., Pütz J., Florentz C., Betat H., Sauter C., Mörl M.* // *Nucl. Acids Res.* 2018. V. 46. P. 9170–9180. <https://doi.org/10.1093/nar/gky593>
35. *Hou Y.M., Schimmel P.* // *Nature.* 1988. V. 333. P. 140–145. <https://doi.org/10.1038/333140a0>
36. *Francklyn C., Schimmel P.* // *Nature.* 1989. V. 337. P. 478–481. <https://doi.org/10.1038/337478a0>
37. *Sardesai N.Y., Green R., Schimmel P.* // *Biochemistry.* 1999. V. 38. P. 12080–12088. <https://doi.org/10.1021/bi991126f>
38. *Di Giulio M.* // *J. Theor. Biol.* 2019. V. 480. P. 99–103. <https://doi.org/10.1016/j.jtbi.2019.07.020>
39. *Di Giulio M.* // *J. Mol. Evol.* 2020. V. 88. P. 228–233. <https://doi.org/10.1007/s00239-020-09927-3>
40. *Root-Bernstein R., Kim Y., Sanjay A., Burton Z.F.* // *Transcription.* 2016. V. 7. P. 153–163. <https://doi.org/10.1080/21541264.2016.1235527>
41. *Agmon I.* // *Int. J. Mol. Sci.* 2018. V. 19. 4021. <https://doi.org/10.3390/ijms19124021>
42. *Tanaka T., Kikuchi Y.* // *Viva Orig.* 2001. V. 29. P. 134–142.
43. *Tamura K.* // *Life (Basel).* 2015. V. 5. P. 1687–1699. <https://doi.org/10.3390/life5041687>
44. *Pak D., Root-Bernstein R., Burton Z.F.* // *Transcription.* 2017. V. 8. P. 205–219. <https://doi.org/10.1080/21541264.2017.1318811>
45. *Pak D., Du N., Kim Y., Sun Y., Burton Z.F.* // *Transcription.* 2018. V. 9. P. 137–151. <https://doi.org/10.1080/21541264.2018.1429837>
46. *Kim Y., Opron K., Burton Z.F.* // *Life (Basel).* 2019. V. 9. 37. <https://doi.org/10.3390/life9020037>
47. *Burton Z.F.* // *J. Mol. Evol.* 2020. V. 88. P. 234–242. <https://doi.org/10.1007/s00239-020-09928-2>
48. *Monro R.E.* // *Nature.* 1969. V. 223. P. 903–905. <https://doi.org/10.1038/223903a0>
49. *Pestka S.* // *J. Biol. Chem.* 1969. V. 244. P. 1533–1539. <https://doi.org/10.1073/pnas.64.2.709>
50. *Tamura K.* // *J. Biosci.* 2011. V. 36. P. 921–928. <https://doi.org/10.1007/s12038-011-9158-2>
51. *Świderek K., Marti S., Tuñón I., Moliner V., Bertran J.* // *J. Am. Chem. Soc.* 2015. V. 137. P. 12024–12034. <https://doi.org/10.1021/jacs.5b05916>
52. *Zhang X., Jiang Y., Mao Q., Tan H., Li X., Chen G., Jia Z.* // *Molecules.* 2017. V. 22. 571. <https://doi.org/10.3390/molecules22040571>
53. *Prosdocimi F., Zamudio G.S., Palacios-Pérez M., Farias S.T., José M.V.* // *Life (Basel).* 2020. V. 10. 134. <https://doi.org/10.3390/life10080134>
54. *Zhang B., Cech T.R.* // *Chem. Biol.* 1998. V. 5. P. 539–553. [https://doi.org/10.1016/S1074-5521\(98\)90113-2](https://doi.org/10.1016/S1074-5521(98)90113-2)
55. *Hsiao C., Lenz T.K., Peters J.K., Fang P.-Y., Schneider D.M., Anderson E.J., Preeprem T., Bowman J.C., O'Neill E.B., Lie L., Athavale S.S., Gossett J.J., Trippe C., Murray J., Petrov A.S., Wartell R.M., Harvey S.C., Hud N.V., Williams L.D.* // *Nucleic Acids Res.* 2013. V. 41. P. 3373–3385. <https://doi.org/10.1093/nar/gkt023>
56. *Davidovich C., Belousoff M., Bashan A., Yonath A.* // *Res. Microbiol.* 2009. V. 160. P. 487–492. <https://doi.org/10.1016/j.resmic.2009.07.004>
57. *Farias S.T., Rego T.G., Jose M.V.* // *FEBS Open Bio.* 2014. V. 4. P. 175–178. <https://doi.org/10.1016/j.fob.2014.01.010>
58. *Farias S.T., Jose M.V.* // *Progr. Biophys. Mol. Biol.* 2020. V. 153. 28e34. <https://doi.org/10.1016/j.pbiomolbio.2020.02.006>
59. *Bhattacharyya S., Varshney U.* // *RNA Biol.* 2016. V. 13. P. 810–819. <https://doi.org/10.1080/15476286.2016.1195943>
60. *Hsiao C., Chou I.C., Okafor C.D., Bowman J.C., O'Neill E.B., Athavale S.S., Petrov A.S., Hud N.V., Wartell R.M., Harvey S.C., Williams L.D.* // *Nat. Chem.* 2013. V. 5. P. 525–528. <https://doi.org/10.1038/nchem.1649>
61. *Roth A., Breaker R.R.* // *Proc. Natl. Acad. Sci. USA.* 1998. V. 95. P. 6027–6031. <https://doi.org/10.1073/pnas.95.11.6027>
62. *Robertson M.P., Knudsen S.M., Ellington A.D.* // *RNA.* 2004. V. 10. P. 114–127. <https://doi.org/10.1261/rna.5900204>
63. *Kun Á., Radványi A.* // *BioSystems.* 2018. V. 164. P. 217–225. <https://doi.org/10.1016/j.biosystems.2017.10.006>

64. Wang Y., Liu E., Lam C.H., Perrin D.M. // *Chem. Sci.* 2018. V. 9. P. 1813–1821.  
<https://doi.org/10.1039/c7sc04491g>
65. Shimizu M. // *J. Biochem.* 1995. V. 117. P. 23–26.  
<https://doi.org/10.1093/oxfordjournals.jbchem.a124715>
66. Shimizu M. // *J. Biochem.* 1996. V. 119. P. 832–834.  
<https://doi.org/10.1093/oxfordjournals.jbchem.a021318>
67. Wieczorek R., Adamala K., Gasperi T., Politelli F., Stano P. // *Life (Basel)*. 2017. V. 7. 19.  
<https://doi.org/10.3390/life7020019>
68. Kurbasic M., Garcia A.M., Viada S., Marchesan S. // *Molecules*. 2021. V. 26. 173.  
<https://doi.org/10.3390/molecules26010173>
69. Cobb A.K., Pudritz R.E. // *Astrophys. J.* 2014. V. 783. 140.  
<https://doi.org/10.1088/0004-637X/783/2/140>
70. Parker E.T., Cleaves H.J., Dworkin J.P., Glavin D.P., Callahan M., Aubrey A., Lazcano A., Bada J.L. // *Proc. Natl. Acad. Sci. USA*. 2011. V. 108. P. 5526–5531.  
<https://doi.org/10.1073/pnas.1019191108>
71. Yarus M., Widmann J.J., Knight R. // *J. Mol. Evol.* 2009. V. 69. P. 406–429.  
<https://doi.org/10.1007/s00239-009-9270-1>
72. Woese C.R. // *Proc. Natl. Acad. Sci. USA*. 1965. V. 54. P. 71–75.  
<https://doi.org/10.1073/pnas.54.6.1546>
73. Barbieri M. // *BioSystems*. 2019. V. 185. 104024.  
<https://doi.org/10.1016/j.biosystems.2019.104024>
74. Kumar R.K., Yarus M. // *Biochemistry*. 2001. V. 40. P. 6998–7004.  
<https://doi.org/doi:10.1021/bi010710x>
75. Xu J., Appel B., Balke D., Wichert C., Müller S. // *ChemBioChem*. 2014. V. 15. P. 1200–1209.  
<https://doi.org/10.1002/cbic.201300741>
76. Huang Z., Szostak J. // *RNA*. 2003. V. 9. P. 1456–1463.  
<https://doi.org/10.1261/rna.5990203>
77. Gibard C., Bhowmik S., Karki M., Kim E.-K., Krishnamurthy R. // *Nat. Chem.* 2018. V. 10. P. 212–217.  
<https://doi.org/10.1038/NCHEM.2878>
78. Illangasekare M., Yarus M. // *Proc. Natl. Acad. Sci. USA*. 1999. V. 96. P. 5470–5475.  
<https://doi.org/10.1073/pnas.96.10.5470>
79. Lee N., Suga H. // *RNA*. 2001. V. 7. P. 1043–1051.  
<https://doi.org/10.1017/S1355838201010457>
80. Ishida S., Terasaka N., Katoh T., Suga H. // *Nat. Chem. Biol.* 2020. V. 16. P. 702–709.  
<https://doi.org/10.1038/s41589-020-0500-6>
81. Shaul S., Berel D., Benjamini Y., Graur D. // *RNA*. 2010. V. 16. P. 141–153.  
<https://doi.org/10.1261/rna.1745910>
82. Rodin A.S., Szathmáry E., Rodin S.N. // *Biol. Direct*. 2009. V. 4. 4.  
<https://doi.org/10.1186/1745-6150-4-4>
83. Patel B.H., Percivalle C., Ritson D.J., Duffy C.D., Sutherland J.D. // *Nat. Chem.* 2015. V. 7. P. 301–307.  
<https://doi.org/10.1038/NCHEM.2202>
84. Shen C., Yang L., Miller S.L., Oro J. // *J. Mol. Evol.* 1990. V. 31. P. 167–174.  
<https://doi.org/10.1007/BF02109492>
85. Sandford S.A., Nuevo M., Bera P.P., Lee T.J. // *Chem. Rev.* 2020. V. 120. P. 4616–4659.  
<https://doi.org/10.1021/acs.chemrev.9b00560>
86. Chen Y.-J., Nuevo M., Yih T.-S., Ip W.-H., Fung H.-S., Cheng C.-Y., Tsai H.-R., Wu C.-Y.R. // *Mon. Not. R. Astron. Soc.* 2008. V. 384. P. 605–610.  
<https://doi.org/10.1111/j.1365-2966.2007.12687.x>
87. Meierhenrich U.J., Muñoz Caro G.M., Bredehöft J.H., Jessberger E.K., Thiemann W.H.-P. // *Proc. Natl. Acad. Sci. USA*. 2004. V. 101. P. 9182–9186.  
<https://doi.org/10.1073/pnas.0403043101>
88. Foden C.S., Islam S., Fernández-García C., Maugeri L., Sheppard T.D., Powner M.W. // *Science*. 2020. V. 370. P. 865–869.  
<https://doi.org/10.1126/science.abd5680>
89. Nesterov-Mueller A., Popov R., Seligmann H. // *Life (Basel)*. 2021. V. 11. 4.  
<https://doi.org/10.3390/life11010004>
90. Schimmel P., Henderson B. // *Proc. Natl. Acad. Sci. USA*. 1994. V. 91. P. 11283–11286.  
<https://doi.org/10.1073/pnas.91.24.11283>
91. Patel S.C., Bradley L.H., Jinadasa S.P., Hecht M.H. // *Protein Sci.* 2009. V. 18. P. 1388–1400.  
<https://doi.org/10.1002/pro.147>
92. Fisher M.A., McKinley K.L., Bradley L.H., Viola S.R., Hecht M.H. // *PLoS One*. 2011. V. 6. e15364.  
<https://doi.org/10.1371/journal.pone.0015364>
93. Jonák J., Rychlík I. // *Biochim. Biophys. Acta*. 1970. V. 199. P. 421–424.  
[https://doi.org/10.1016/0005-2787\(70\)90084-5](https://doi.org/10.1016/0005-2787(70)90084-5)
94. Gnirke A., Nierhaus K.H. // *J. Biol. Chem.* 1986. V. 261. P. 14506–14514.  
[https://doi.org/10.1016/S0021-9258\(18\)66898-X](https://doi.org/10.1016/S0021-9258(18)66898-X)
95. Belousoff M.J., Davidovich C., Zimmerman E., Caspi Y., Wekselman I., Rozenszajn L., Shapira T., Sade-Falk O., Taha L., Bashan A., Weiss M.S., Yonath A. // *Biochem. Soc. Trans.* 2010. V. 38. P. 422–427.  
<https://doi.org/10.1042/bst0380422>
96. Massey S.E. // *J. Mol. Evol.* 2006. V. 62. P. 809–810.  
<https://doi.org/10.3390/life5021301>
97. Freeland S.J., Wu T., Keulmann N. // *Orig. Life Evol. Biosph.* 2003. V. 33. P. 457–477.  
<https://doi.org/10.1023/A:1025771327614>
98. Schuster G.B., Cafferty B.J., Karunakaran S.C., Hud N.V. // *J. Am. Chem. Soc.* 2021. V. 143. P. 9279–9296.  
<https://doi.org/10.1021/jacs.0c13081>
99. Anderson B.A., Fauché K., Karunakaran S.C., Yera-bolu J.R., Hud N.V., Krishnamurthy R. // *Chem. Eur. J.* 2021. V. 27. P. 4033–4042.  
<https://doi.org/10.1002/chem.202004397>
100. Fialho D.M., Karunakaran S.C., Greeson K.W., Martinez I., Schuster G.B., Krishnamurthy R., Hud N.V. // *J. Amer. Chem. Soc.* 2021. V. 143. P. 13525–13537.  
<https://doi.org/10.1021/jacs.1c02287>
101. Wang Y., Wang Y., Song D., Sun X., Zhang Z., Li X., Li Z., Yu H. // *J. Am. Chem. Soc.* 2021. V. 143. P. 8154–8163.  
<https://doi.org/10.1021/jacs.1c02895>



## On the Origin of Genetically Coded Protein Synthesis

S. P. Kovalenko<sup>#</sup>

<sup>#</sup>*E-mail: spkovalenko2008@mail.ru*

*Institute of Microbiology of the National Academy of Sciences of Belarus, ul. Surganova 52, Minsk, 220040 Republic of Belarus*

How proteins with a specific amino acid sequence synthesized in living organisms is no longer a mystery after the deciphering of the genetic code six decades ago. However, the origin of the system of genetically coded protein synthesis is still the subject of hypotheses, experimentally justified only partially. Based on the review of works mainly of the last decade, the author formulated a hypothesis that largely concretizes certain stages of the formation of the translation system and the genetic code. The hypothesis based on the concept of evolutionary reduction of the ambiguity of the primordial code. The distinctive positions of the hypothesis: (i) the emergence of aptamers, enriching the protocell by the rare catalytically active amino acids, and their evolutionary transformation in aaRS-like ribozymes which resulted in the operational code, (ii) the participation at once of large number of amino acids (including non-canonical) in the formation of operational and genetic codes, (iii) the recognition of the whole codon of mRNA by the anticodon loop of proto-tRNA from the very beginning of encoded protein synthesis, (iv) the co-evolution of the operational and genetic codes that eliminated their initial ambiguity.

*Keywords: proteins, peptides, amino acids, aminoacylation, genetic code, operational code, RNA, tRNA, mRNA, aaRSs, PTC, ribozymes*