

УДК 573.554

ПРОИСХОЖДЕНИЕ ГЕНЕТИЧЕСКОГО КОДА И ТРАНСЛЯЦИИ В РАМКАХ СОВРЕМЕННЫХ КОНЦЕПЦИЙ ПРОИСХОЖДЕНИЯ ЖИЗНИ

Обзор

© 2022 Л.Г. Кондратьева¹, М.С. Дьячкова², А.В. Гальченко^{3*}

¹ Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова РАН, 117997 Москва, Россия

² Институт общей генетики им. Н.И. Вавилова РАН, 119991 Москва, Россия

³ Российский университет дружбы народов (РУДН), 117198 Москва, Россия; электронная почта: gav.jina@gmail.com

Поступила в редакцию 26.05.2021

После доработки 10.11.2021

Принята к публикации 08.12.2021

Происхождение генетического кода и системы трансляции, возможно, является центральной и самой трудной проблемой в изучении происхождения жизни и одной из самых трудных во всей эволюционной биологии. Существует большое количество гипотез возникновения и развития современных генетических систем, затрагивающих происхождение и раннюю эволюцию генетического кода, а также возникновение репликации и трансляции. Наиболее широко известные гипотезы рассмотрены в данном обзоре. Однако ни одна из этих гипотез не описывает без пробелов и допущений все этапы ранней эволюции генетических систем. Гипотеза РНК-мира является главенствующей на сегодняшний день научной идеей о ранней эволюции биологических и пребиологических объектов. Главное её преимущество заключается в том, что она предлагает в качестве первых живых систем РНК как самодостаточные, с точки зрения воспроизведения, молекулы, которые способны функционировать как каталитический компонент системы и в то же время – как матричный. Однако есть и существенные недостатки. В частности, до сих пор не открыта и не получена экспериментально рибозимная процессивная полимеразы. Учитывая взаимную потребность белков и нуклеиновых кислот в современном мире, многие авторы предлагают сценарии ранней эволюции на основе коэволюции этих двух классов органических молекул. Подобные гипотезы постулируют, что для репликации нуклеиновых кислот было необходимо возникновение трансляции, в отличие от мира РНК, где появлению трансляции предшествовала эра самореплицирующихся РНК. И хотя такие сценарии менее экономичны, с эволюционной точки зрения, так как требуют одномоментного появления и эволюции сразу двух классов органических молекул, а также синхронизации по времени появления репликации и трансляции, большим их преимуществом является то, что они предлагают развитие сразу гораздо более точной и процессивной белковой репликации.

КЛЮЧЕВЫЕ СЛОВА: эволюция, зарождение жизни, трансляция, генетический код, РНК-мир, проген, белковый мир, липидный мир.

DOI: 10.31857/S0320972522010043

ВВЕДЕНИЕ

Земля сформировалась ~4,5 миллиарда лет назад, и вскоре после этого, ~4 миллиарда лет назад [1], сложившиеся на планете условия при-

вели к появлению особого состояния материи, которое мы называем жизнью. По определению NASA, жизнь – это «самоподдерживающаяся химическая система, способная к дарвиновской эволюции» [2]. Для реализации данной способности появился первобытный генетический код, который обеспечил трансляцию нуклеиновых кислот в белки. Для объяснения потенциальных механизмов зарождения генетических систем было предложено большое количество различных гипотез, каждая из которых делает акцент на различных событиях, приводящих к возникновению функциональной трансляции и самоподдерживающейся системы. Они описывают как происхождение генетического кода,

Принятые сокращения: aaRS – аминоксил-тРНК-синтетаза; DRT – прямое кодирование аминокислот с РНК-матрицы; C-DRT – цепь, комплементарная DRT-матрице; GADV – модель эволюции белкового мира; GARD – модель репликации мицеллярного автокатализа; PR – шкала растворимости аминокислот в пиридине (PolarRequirement); Self-гARS – самоаминоацилирующие рибозимы; SART – самоаминоацилирующаяся рибозимная матрица.

* Адресат для корреспонденции.

т.е. возникновение соответствия между аминокислотами и триплетами нуклеотидов, так и происхождение трансляции, т.е. полимеризации аминокислот на основе закодированной в нуклеиновых кислотах информации, и репликации – воспроизведения генетической информации.

Трансляционная система даже в самых простых современных клетках (например, паразитических и эндосимбиотических бактериях и археях, таких как *Candidatus Carsonella ruddii*, *Mycoplasma* или *Nanoarchaeon*) чрезвычайно сложна [3]. В основе трансляции лежит большой комплекс (рибосома), состоящий по меньшей мере из трёх молекул РНК и 60–80 белков (хотя у облигатных эндосимбионтов может быть меньше 35 рибосомных белков [4]), взаимодействующих с другими компонентами системы трансляции: полным набором тРНК для 20 аминокислот (~40 видов тРНК с учётом присутствия изоакцепторных тРНК у всех видов), 20 аминоксил-тРНК-синтетазами (aaRS), а также факторами трансляции.

При этом ключевой проблемой в определении того, как возникла современная трансляционная система, является следующий парадокс: высокая точность трансляции вряд ли могла быть достигнута без сложных, высокоразвитых молекул РНК и белков, при этом сложный белковый механизм не мог бы развиваться без точной системы трансляции. Этот парадокс рассматривается в цикле Дарвина–Эйгена (Darwin–Eigen cycle) [3].

ЦИКЛ ДАРВИНА–ЭЙГЕНА

Теория Eigen [5] постулирует существование фундаментального предела, ограничивающего размер генома и точность репликации в самореплицирующихся системах (порог Эйгена): если произведение частоты ошибок и информационной ёмкости системы ниже порога Эйгена, происходит стабильное наследование; если данная величина выше пороговой, мутационный процесс неизбежно приводит к катастрофическому вымиранию системы [5]. Иными словами, уровень ошибок репликационного аппарата лимитирует максимальный размер генома (количество закодированной информации). Для стабильного поддержания репликации, что является необходимым начальным условием биологической эволюции, система обязана непрерывно демонстрировать значения, не выходящие за пределы указанного интервала, в том числе на самых ранних этапах своего существования. Максимальный размер реплицируемой после-

довательности без коррекции ошибок, лимитируемый порогом Эйгена, крайне незначителен. В рамках теории Эйгена увеличение информационной ёмкости и расширение генома могло быть достигнуто благодаря увеличению точности репликации. Однако эволюция высокоточной системы репликации, обеспечивающей коррекцию ошибок, требовала бы значительно большего объёма информации, чем та, которая могла быть закодирована в последовательности самореплицирующейся системы в данный момент эволюционной траектории. Это создаёт, по меньшей мере, видимость парадокса возникновения и эволюции живых систем, описываемого следующим образом: для кодирования древнего механизма репликации была необходима некоторая минимальная сложность генома, в то время как концепция предела ошибок существенно ограничивала его размер. Таким образом, на первый взгляд усложнение генома не представлялось возможным благодаря существованию порога Эйгена [6]. Одним из возможных вариантов разрешения кажущегося парадокса является модель гиперциклов, предложенная Eigen и Schuster [7]. В гиперцикле РНК и ферменты копируются следующим образом: РНК I_i кодирует фермент E_i ($i = 1, 2, \dots, n$); ферменты циклически катализируют репликацию РНК таким образом, что E_i катализирует репликацию I_{i+1} , а E_n катализирует репликацию I_1 . Таким образом, в гиперцикле пептиды совместно с РНК формируют целостную систему кооперативно взаимодействующих макромолекул. Согласно данной модели, ферменты могли способствовать повышению точности репликации, в результате чего количество наследуемой информации в примитивной системе возрастало. Модель гиперциклов имеет и недостатки: как было показано в ряде работ [8–10], гиперциклы демонстрируют слабую устойчивость по отношению к мутациям, которые могут приводить к разрушению структуры гиперцикла. Другим потенциальным решением парадокса Эйгена является комбинация естественного отбора и генетического дрейфа. Случайные события рекомбинации или дупликации частей генома, приводящие к умеренному увеличению его размера, могут закрепляться в малых популяциях благодаря генетическому дрейфу. При условии достаточной точности механизма репликации подобные небольшие изменения в размере генома не уменьшают значительным образом приспособленность системы, однако предоставляют генетический материал для эволюции новых адаптивных функций. Среди таких функций могут быть механизмы, повышающие точность репликации, что отодвигает систему от «обрыва

РНК-МИР

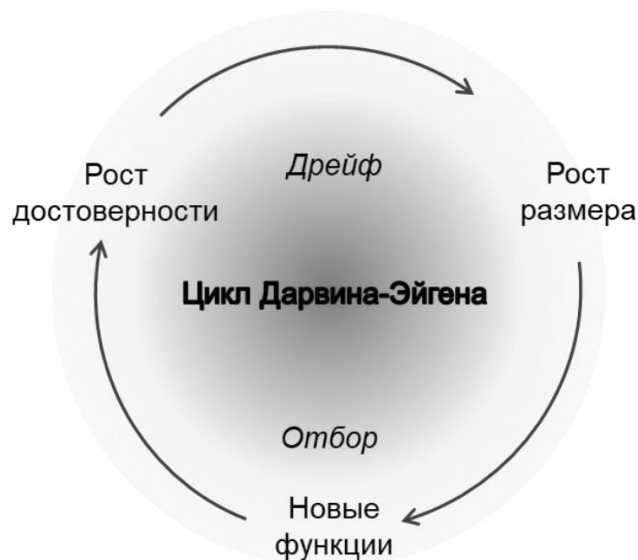


Рис. 1. Цикл Дарвина–Эйгена (Darwin–Eigen cycle). Модифицировано по [3]

Эйгена» и обеспечивает необходимый потенциал к дальнейшему увеличению объёма кодируемой информации. Таким образом, цикл повторяет себя в спиральной прогрессии, приводя к непрерывному увеличению размера генома [6]. Данная система с положительной обратной связью получила название цикла Дарвина–Эйгена [3] (рис. 1).

Основной проблемой рассматриваемой теории остаётся вопрос о первоначальной инициации цикла Дарвина–Эйгена, т.е. как именно была достигнута наименьшая сложность генома, необходимая для приемлемой репликации. Даже в наименее точных современных системах, таких как вириды (точность механизмов репликации составляет всего $\sim 10^{-2}$ [11], т.е. в среднем одна ошибка на 100 нуклеотидов), репликация катализируется сложными белковыми полимеразми, которые, в свою очередь, синтезируются в результате трансляции при участии сложного и многокомпонентного рибосомного аппарата. Koopin и Novozhilov [12] предложили следующее решение кажущегося парадокса первоначальной инициации цикла Дарвина–Эйгена: трансляция в древних самореплицирующихся системах могла обеспечиваться компонентами небелкового происхождения, а инициация цикла произошла в условиях мира РНК. При этом, помимо гипотезы РНК-мира, существуют другие теории, описывающие в качестве родоначальных ДНК, нуклеопротеиды, белки или предполагающие их коэволюцию. Рассмотрим возникновение аппарата трансляции и репликации в рамках этих теорий.

Гипотеза РНК-мира является наиболее широко принятой идеей возникновения жизни на Земле. Rich [13] в 1962 г. впервые предложил наличие энзиматической активности у РНК, затем Crick [14] и Orgel [15] в 1968 г. предположили, что исходной молекулой для возникновения жизни могла являться РНК, а Gilbert [16] ввёл удачный термин «мир РНК» в 1986 г. после того, как была доказана энзиматическая активность РНК [15]. Классическая версия зарождения жизни в рамках РНК-мира заключается в первоначальном синтезе и случайной полимеризации нуклеотидов, которые привели к образованию пулов олигомеров реплицирующихся нуклеиновых кислот; рекомбинационные события привели к образованию более длинных олигомеров, которые складывались в структуры различной сложности и образовывали ферментативно активные рибозимы. По мере увеличения сложности полинуклеотидов возникали первые РНК-репликазы, образовывались оболочки, и это привело к появлению протоклеток со способностью к эволюции [17].

Свидетельством в пользу РНК-мира стало открытие разнообразных рибозимов, которые могли осуществлять всю основную «работу» на заре зарождения жизни [18–21]. Обнаруженные рибозимные активности могли бы свидетельствовать о том, что именно РНК играли роль биологических катализаторов в рамках гипотезы РНК-мира. И несмотря на то что природные РНК-репликазы до сих пор не были найдены, существуют искусственно созданные рибозимы со свойствами РНК-полимеразы дистрибутивного типа [22, 23].

Известно, что самыми консервативными существующими генами являются гены трансляции [24]. Рибосомные белки, тРНК и aaRS, а также белковые факторы, которые участвуют в инициации, элонгации и терминации, обладают удивительно консервативными структурами и универсальны для всех видов [25, 26]. Общая коровая рРНК, которая является главным функциональным компонентом рибосомы, консервативна на протяжении всей филогенетической истории [27, 28], при этом сохраняется как последовательность, так и вторичная, и третичная структуры [24].

При этом в теории РНК-мира есть ряд нерешённых проблем [29–31]. Так, аргументом против теории является то, что РНК слишком сложна, чтобы возникать *de novo* в пребиотической среде. В частности, есть сложности с абиотическим синтезом пиримидиновых рибонуклеотидов [30, 32, 33], нестабильностью молекул РНК

и их высокой чувствительностью к рН и наличию ионов магния [29]. Кроме того, для существования и эволюции РНК-мира существует ряд ограничений по физическим условиям: необходимо наличие РНК-адсорбирующих поверхностей, регулярных циклов подсушивания/увлажнения, нагревания/охлаждения, замораживания/оттаивания, а также механизмов защиты РНК от космического излучения [34]. Существует и проблема хиральности при репликации РНК [35]. Нуклеотиды и их предшественники в пребиотическом мире существовали, вероятно, как рацемическая смесь право- и левовращающих молекул, и включение противоположно вращающего нуклеотида в цепь нарушало бы репликацию.

Надёжность репликации является ещё одной проблемой РНК-мира: поскольку точность самосборки РНК низкая, количество ошибок увеличивается по мере удлинения цепей, и циклы повторяются. Накопление ошибочных молекул в конечном итоге могло бы привести к гибели всей системы [30]. Большой проблемой является также низкая процессивность предполагаемой древней репликации молекул РНК. Кроме того, при рассмотрении самопроизвольной репликации матриц скорость их образования сопоставима со скоростью их деградации [36]. Часть парадоксов, связанных с накоплением ошибок, может быть разрешена, если предположить, что РНКовая РНК-полимераза состояла из нескольких частей, каждая из которых относительно короткая, а сборка осуществлялась за счёт спаривания. Это могло бы повысить вероятность безошибочной репликации фрагментов, а значит и образования полноценных комплексов и, по-видимому, могло сделать более эффективным отбор. Однако подобные полимеразы не были обнаружены или получены экспериментально.

Появление смысловых цепочек большого размера также сложно объяснить в рамках РНК-мира [30]. Этот момент можно проиллюстрировать на примере случайного происхождения рибозимной РНК-полимеразы: появление нескольких РНК-полимеразных рибозимов потенциально возможно, однако они будут занимать лишь незначительную долю от огромного количества вероятных последовательностей РНК [36]. Таким образом, формирование значимых функциональных полинуклеотидов случайным образом кажется маловероятным [30].

Кроме того, катализ является относительно редким свойством РНК, а также РНК имеет слишком ограниченный каталитический репертуар [29]. Несмотря на достижения в области пребиотической химии, до сих пор не удалось

продемонстрировать надёжную и непрерывную саморепликацию РНК из исходных молекул. Изолированная РНК может оказаться недостаточным условием для того, чтобы катализировать собственную репликацию, и может нуждаться в наличии либо других молекул, либо определённых условий окружающей среды [11].

Тем не менее теория РНК-мира является наиболее принятой гипотезой, в рамках которой предложены различные модели эволюции ключевых компонентов репликации и трансляции: возникновения генетического кода, образования молекул мРНК, тРНК, aaRS, рибосом, а также самого механизма репликации РНК.

Происхождение репликации в рамках гипотезы РНК-мира. Неферментативная матричная полимеризация нуклеотидов. Согласно гипотезам неферментативной репликации, удвоение молекулы РНК могло проходить с помощью достраивания праймера активированными (например, имидазолом) мономерами [37, 38] или олигомерами [39]. Механизм, использующий активированные мономеры, имеет много общего с современными универсальными биологическими процессами ферментативной репликации РНК и ДНК. Экспериментально показано, что скорость присоединения активированных мономеров к праймеру значительно выше, чем лигирование праймера к олигонуклеотиду [40]. При этом олигонуклеотиды могут долгое время находиться на матрице, и низкая скорость лигирования в таком случае не будет проблемой. Считается, что эффективное копирование длинных матриц РНК может быть достигнуто за счёт гибридного процесса, в котором олигонуклеотиды выступают в роли праймеров на нескольких сайтах на матрице, а присоединяющиеся мономеры достраивают места пробелов, после чего происходит лигирование всех фрагментов [38, 41]. При этом при рассмотрении неферментативной репликации необходимо отметить, что при таком механизме могут образовываться не только правильные связи между нуклеотидами (3'–5'), но и связи между «неправильными» гидроксилами рибозы (2'–5'). Кроме того, низкая скорость такой репликации сопоставима со скоростью деградации синтезированной РНК, что накладывает значительное количество ограничений на эту концепцию.

РНК-катализируемая репликация. Концепция РНК-мира подразумевает существование рибозима, обладающего свойствами РНК-зависимой РНК-полимеразы, которая синтезировала бы комплементарные РНК для воспроизводства дополнительных копий самой себя. Образование способной к такой активности молекулы РНК могло происходить по следующему сценарии

рию: благодаря неферментативной репликации образовывались популяции реплицирующихся РНК. Среди них некоторые РНК получили бы большее распространение в случае, если бы они обладали более высокой скоростью воспроизводства. Таким образом, мог произойти отбор наиболее подходящих матричных последовательностей и РНК, способствующий увеличению эффективности процесса репликации. При продолжающемся селективном давлении этот процесс химической репликации будет вытеснен биохимическим процессом репликации [38].

Возможность существования рибозимной РНК-полимеразы была показана в работе Bartell и Szostak [42], которые *in vitro* воспроизвели эволюцию большой популяции РНК со случайными последовательностями и в результате отобрали рибозим РНК-лигазы размером ~100 нуклеотидов, катализирующий соединение двух связанных с матрицей олигонуклеотидов [38]. Далее была проведена серия экспериментов по эволюции *in vitro* для преобразования этой лигазы в настоящую РНК-полимеразу, которая действует на отдельной матрице РНК. В результате дальнейшей оптимизации рибозимной полимеразы была увеличена длина синтезируемого фрагмента, скорость полимеризации, эффективность работы на сложных матрицах и точность рибозима [38]. Полученный в итоге полимеразный рибозим был способен достаточно быстро синтезировать множество сложноструктурированных РНК [43]. Таким образом, эта оптимизированная РНК-полимераза могла катализировать реципрокный синтез как РНК, так и её комплементарной цепи, обеспечивая экспоненциальную амплификацию коротких матриц РНК в форме полимеразной цепной реакции, и кроме того, была способна синтезировать полимеры ДНК на матрице РНК [44]. Несмотря на то что наиболее оптимизированная форма такой полимеразы не может реплицировать длинные последовательности РНК, исследователи данного направления эволюции репликации Joyce и Szostak полагают, что в конечном итоге в лаборатории будет разработана процессивная РНК-репликаза (молекулярная машина, которая продвигается по полинуклеотидной матрице и включает мономеры в растущую комплементарную цепь) [38].

Происхождение генетического кода в рамках гипотезы мира РНК. Генетический код представляет собой систему, определяющую трансляцию нуклеотидной последовательности нуклеиновых кислот в аминокислотную последовательность полипептидов, и является практически универсальной системой живых организмов [12, 45]. За расшифровку генетического ко-

да Khorana и Nirenberg вместе с Holley в 1968 г. получили Нобелевскую премию по физиологии и медицине. С момента завершения процесса расшифровки стандартного генетического кода, в котором приняли участие несколько исследовательских групп [46–49], множество исследователей были вовлечены в изучение его структуры и свойств. Существует несколько теорий происхождения этой системы строгого соответствия кодируемых аминокислот триплетам нуклеотидов, которые не являются взаимоисключающими, а, вероятно, описывают разные этапы формирования структуры генетического кода.

Стереохимическая гипотеза. Согласно стереохимической теории, генетический код должен был возникнуть благодаря физико-химическим взаимодействиям между нуклеотидами антикодонов или кодонов и соответствующими аминокислотами.

Для экспериментальной проверки гипотезы стереохимического взаимодействия была разработана специальная шкала (англ. Polar Requirement, PR), отражающая растворимость аминокислот в пиридине и его производных (аналогах пиримидина) [50]. Используя эту шкалу, Woese [51] показал, что аминокислоты со схожими значениями PR, как правило, кодируются родственными кодонами, обычно – кодонами с одинаковым вторым основанием [50, 51]. С развитием технологий получения аптамеров возникли новые аргументы в поддержку стереохимической теории. В экспериментах SELEX было показано, что фракции случайных аптамеров РНК, отобранных на основе их способности связываться с определенными аминокислотными остатками, значительно обогащены соответствующими триплетами. Критика экспериментов SELEX обычно связана с тем, что специфические статистически значимые корреляции были обнаружены для незначительного количества аминокислот, при этом в некоторых случаях последовательности были обогащены кодонами, а в других – антикодонами [51–55]. Кроме того, средство было показано для аминокислот, большинство из которых имело сложное строение и которые могли быть недоступны для формирования структуры генетического кода на ранних этапах его эволюции [12, 45, 56].

В серии недавно опубликованных протеомных исследований было установлено, что профили распределения плотности пиримидина в последовательностях мРНК коррелируют с профилями распределения значений PR, кодируемых белков. Иными словами, участки мРНК, богатые пиримидинами, кодируют участки белка, которые проявляют высокую склонность к взаимодействию с миметиками пиримидина, и

наоборот [51, 57–60]. Неясно, однако, могли ли предполагаемые стереохимические взаимодействия между нуклеиновыми кислотами и пептидами, а не свободными аминокислотами, играть значимую роль в происхождении кода [12].

Для объяснения наблюдаемых корреляций в рамках стереохимической гипотезы различными авторами было предложено множество моделей стереохимического взаимодействия между аминокислотами и нуклеотидами нуклеиновых кислот [61–65]. Следует отметить, что только некоторые из предложенных механизмов (например, предложенный Woese et al. [50]) потенциально способны объяснить процесс прямого формирования полипептидной цепи на матрице нуклеотидной последовательности, так как для возникновения пептидной связи было необходимо в первую очередь обеспечить сближение аминокислот в пространстве. Существенная разница в размерах между нуклеотидными триплетами и аминокислотами затрудняет пространственное сближение аминокислот в процессе прямого матричного синтеза. Тем не менее, несмотря на свои недостатки, стереохимическая гипотеза может значительно облегчить понимание того, как происходило формирование и ранняя эволюция генетического кода [56].

Гипотеза расширения кода. Сценарий эволюции кода, предложенный ещё Crick [14], включал примитивную легко реплицирующуюся нуклеиновую кислоту, несколько наиболее доступных аминокислот и первичный код, базирующийся на двух основаниях. Crick полагал, что двух оснований могло быть достаточно для кодирования ограниченного репертуара аминокислот в составе примитивных белков. При этом кодон должен был оставаться триплетным на протяжении всей эволюции генетического кода, так как нарушение принципа непрерывности кода с высокой вероятностью оказалось бы летальным даже для примитивного организма. Woese et al. [50], основываясь на паттернах кластеризации аминокислот, наблюдаемых в структуре стандартного генетического кода, выдвинул гипотезу о том, что примитивный код мог обладать низкой специфичностью: отдельные триплеты кодировали группы химически родственных аминокислот, при этом система распознавания базировалась на одном основании. Таким образом, примитивный код, вероятно, обладал гораздо более высокой степенью вырожденности и характеризовался относительно низким потенциалом кодирования. Данные представления были развиты в модели «2-1-3» [66], которая предлагала сценарий расширения генетического кода и увеличения его разрешаю-

щей способности путём последовательного назначения специфичности второму, затем первому и, в некоторых случаях, третьему нуклеотиду кодона.

Метаболическая теория. Теория коэволюции (метаболическая теория) Wong [67, 68] утверждает, что структура стандартного кода отражает пути биосинтеза аминокислот. Он разделил все протеиногенные аминокислоты на аминокислоты пребиотического происхождения – аминокислоты Фазы 1 (Gly, Ala, Ser, Asp, Glu, Val, Leu, Ile, Pro и Thr) и аминокислоты, возникшие в результате биосинтеза из пребиотических – аминокислоты Фазы 2 (Phe, Tyr, Arg, His, Trp, Asn, Gln, Lys, Cys и Met) [67, 68]. Основываясь на существующих представлениях о биосинтезе аминокислот, Wong предложил следующую схему эволюции биохимических путей синтеза аминокислот Фазы 2 из путей синтеза их предшественников, т.е. аминокислот Фазы 1: Ser – Trp, Ser – Cys, Val – Leu, Thr – Ile, Gln – His, Phe – Tyr, Glu – Gln и Asp – Asn [68]. По его предположению, возникающая в системе генетического кода аминокислота наследовала кодон своего предшественника. Критика теории коэволюции «код-путь» Wong отражена в работах Amirnovin [69] и Ronneberget et al. [70].

Теория «замороженной случайности». Согласно современным представлениям, расширение системы генетического кода по описанному выше сценарию было возможно только в случае кодирования незначительного количества белков, характеризующихся относительно простой структурой. По мере эволюционного развития организмов усложнялись как структура и специфичность функций кодируемых белков, так и их репертуар. В конечном итоге должен был наступить момент, когда любое незначительное изменение в существующей системе кодирования и введение новой аминокислоты приводило бы к критическому нарушению первичной структуры значимого количества кодируемых белков [14]. Таким образом, наблюдаемая высокая консервативность структуры генетического кода обусловлена давлением отрицательного отбора, сила которого увеличивалась по мере усложнения системы [14, 45]. Соответственно, в какой-то момент код перестал накапливать изменения, и, обсуждая эту проблему, Crick [14] предложил термин «замороженной случайности» (англ. «frozen accident»).

Теория минимизации ошибок. Crick [14] и Woese [71] полагали, что формирование наблюдаемой структуры стандартного генетического кода происходило случайным образом и не включало какие-либо механизмы оптимизации. Однако это плохо согласуется с тем фактом, что

стандартный генетический код характеризуется высокой степенью устойчивости к ошибкам (в частности, синонимичные замены, как правило, не приводят к нарушению структуры кодируемого белка; кроме того, замена первого нуклеотида в кодоне часто приводит к включению аминокислоты со схожими физико-химическими свойствами). Устойчивость кода к ошибкам была подтверждена в многочисленных экспериментах *in silico* [56, 72, 73]. Было показано, например, что вероятность достичь аналогичного уровня минимизации ошибок путём случайных перестановок кодонов стандартного генетического кода составляет менее 10^{-6} [56]. Высокая устойчивость стандартного генетического кода к трансляционным ошибкам стала главным аргументом в пользу теории формирования кода под давлением естественного отбора. Теория получила название теории минимизации ошибок [12, 45]. В исследовании адаптивного ландшафта генетического кода авторы изучили возможную эволюцию устойчивости генетического кода к трансляционным ошибкам и подтвердили гипотезу о том, что эволюция стандартного генетического кода могла включать в себя стадию его частичной оптимизации с целью минимизировать эффект мутаций, возникающих в процессе трансляции [73].

Таким образом, сценарий возникновения современной системы кодирования, называемой стандартным генетическим кодом, мог включать следующие стадии: на этапе, когда пул протеиногенных аминокислот характеризовался невысоким разнообразием, назначение кодонов могло определяться одним или двумя нуклеотидами, согласно стереохимическим взаимодействиям; затем по мере увеличения доступности аминокислот эволюция происходила по пути расширения кода и увеличения специфичности кодирования, при этом биохимически родственные аминокислоты кодировались родственными кодонами; по мере усложнения кодируемых белков код был оптимизирован таким образом, чтобы минимизировать возможные ошибки трансляции; и, в конце концов, когда любые изменения в системе устоявшегося соответствия аминокислот кодонам могли стать летальными для организма, код был «заморожен» в его локальном состоянии, а его дальнейшая эволюция стала невозможной за редкими исключениями, что обеспечило наблюдаемую универсальность генетического кода. На сегодняшний день быстроразвивающаяся синтетическая биология уже обладает инструментами и технологиями, которые позволяют модифицировать стандартный генетический код [74]. Дальнейшие исследования в этой области помо-

гут углубить наше понимание механизмов эволюции канонического кода.

Возникновение разнообразия молекул РНК. Молекулы РНК являются центральным звеном процесса передачи информации от генов к белкам (мРНК), функционирования самого механизма трансляции (рРНК и тРНК) и осуществления различных регуляторных функций (различные классы некодирующих РНК). Для формирования современной трансляционной системы в процессе эволюции должны были возникнуть предшественники этих молекул. Так, в ряде гипотез предлагаются модели последовательных преобразований молекул РНК, в ходе которых сформировались структуры, выполняющие функции мРНК, тРНК и рибосом.

Теория рекомбинирующихся РНК. В рамках РНК-мира существует гипотеза, выдвинутая Спириным [75], в которой предполагается, что первичный трансляционный аппарат возник до появления аппарата энзиматической репликации генетического материала и базировался на нескольких видах специализированных РНК. Согласно этой гипотезе, многообразие молекул РНК, в том числе молекул с рибозимной активностью, достигалось благодаря рекомбинации вследствие спонтанной неэнзиматической трансэтерификации абиогенно синтезированных олигорибонуклеотидов. Благодаря такой рекомбинации коротких молекул РНК, комплементарно связывающихся с полинуклеотидной матрицей, происходило объединение фрагментов и первичное размножение молекул РНК. А с появлением рибозимов с полимеразной активностью эффективность репликации должна была значительно возрасть [75, 76]. Согласно данной гипотезе, рекомбинация молекул РНК приводила к образованию первичного трансляционного аппарата: прото-мРНК, набора прото-тРНК, каталитически активной проторибосомной РНК с пептидилтрансферазной активностью и проторибосомной РНК, способной взаимодействовать одновременно с каталитической проторибосомной РНК, прото-мРНК и прото-тРНК. Эта система уже могла осуществлять первичную трансляцию [75]. В дальнейшем эта гипотеза развилась в предположение, что только рекомбинаций полирибонуклеотидов недостаточно для появления и эволюции каталитических РНК, и сначала должны возникнуть нуклеозидтрифосфатзависимые хеликазы на основе РНК [76]. Такие рибозимные хеликазы могли быть способны раскручивать стабильные двухспиральные РНК, неизбежно образующиеся в процессе синтеза РНК на комплементарных матрицах. Благодаря этому мог возникнуть механизм репликации РНК с использова-

нием двухспиральных РНК, что положило начало быстрой эволюции древнего мира РНК [76]. Можно предположить, что образующиеся пептиды, не так уж и необходимые для такого мира рекомбинирующихся РНК, могли стабилизировать РНК, защищая от гидролиза и потери правильной пространственной структуры [77].

Теория прямого кодирования РНК-матрицей.

Самый простой способ кодирования пептидов по шаблону РНК заключается в непосредственном последовательном связывании активированных аминокислот на матрице РНК. Такой механизм предложен Yarus et al. [52, 53] в теории прямого кодирования РНК (англ. Direct RNA Template, DRT) и основан на стереохимической гипотезе возникновения генетического кода. Согласно этой модели, активированные аминокислоты непосредственно связывались непосредственно с РНК-матрицей (рис. 2, а). Взаимодействие аминокислоты при этом могло быть как с последовательностью кодона, так и с последовательностью антикодона. Матрица РНК представляла собой аптамероподобные структуры, распознающие активированные аминокислоты (рис. 2, а; б, часть 1). Стереохимическое взаимодействие определённых аминокислот с матрицей РНК сопровождалось образованием пептидной связи либо химическим, либо рибозим-опосредованным катализом, который осуществлялся или за счёт активности самой матрицы DRT, или реализовывался сторонним рибозимом [58]. Так, активированные аминокислоты полимеризовались, и формировался кодируемый DRT-пептид. На следующем этапе система DRT могла эволюционировать в сторону современного непрямого кодирования, использующего промежуточное звено — тРНК. Такая промежуточная аминоацил-РНК образовывалась через присоединение антикодонного фрагмента DRT к активированной рибозой группе аминокислоты (рис. 2, б, часть 2). Таким образом, происходило преобразование исходной функции связывания аминокислоты в функцию кодирования нуклеиновой кислоты [55]. На более позднем этапе произошёл переход к единообразной версии непрямого кодирования (рис. 2, б, часть 3), включающего отдельную мРНК, активированные аминокислоты, связанные с их антикодонами (аминоацил-тРНК), а также РНК с пептидилтрансферазной активностью — риборибосому [52, 78].

Основная проблема гипотезы прямого кодирования аминокислот с РНК-матрицы — это выраженное пространственное несоответствие размеров аминокислот и нуклеотидных триплетов (рис. 2, в). Такое положение дел оставляет только один возможный вариант последова-

тельного прямого кодирования — спиралевидное закручивание нуклеотидной нити вокруг аминокислотной. Однако в этом случае могли бы возникать различные межнуклеотидные взаимодействия, которые многократно повышали бы вероятность образования РНКовых шпилек и подобных структур. Большим вопросом остаётся и процесс извлечения пептидного стержня из РНКовой спирали. Единственное решение такой проблемы — непоследовательное кодирование аминокислот, т.е. расположение кодирующих триплетов не подряд (рис. 2, а). Переход же к сплошному кодированию мог произойти только после подключения тРНК к процессу трансляции.

Идея напрямую кодирующей РНК была развита также в работе Ma [80] в гипотезу возникновения механизма трансляции DRT с учётом принципа репликативной экономии (англ. Direct RNA Template with replication parsimony). Под «репликативной экономией» в данной гипотезе подразумевается тот факт, что под давлением отбора из-за сложности репликации существовала тенденция к «скупому», «экономному» использованию генетического материала, поэтому на ранних этапах не могли возникнуть сложные, но «бессмысленные» в данный момент структуры, которые потом обрели бы функциональность. Данная гипотеза DRT с учётом принципа репликативной экономии пытается последовательно объяснить эволюционные события появления таких промежуточных звеньев, как тРНК, мРНК или рРНК [80].

Гипотеза подразумевает, что из-за нестабильности РНК некоторые молекулы DRT могли деградировать или быть недореплицированными из-за низкой эффективности рибозимных РНК-полимераз. Эти события могли привести к образованию фрагментов, содержащих основной домен аминокислотного аптамера, окружённого последовательностями спейсеров. При этом полная репликация DRT производила цепи, комплементарные DRT — C-DRT (рис 3, а). Фрагменты DRT «распознавали» C-DRT согласно принципу комплементарности. Аптамерная область фрагментов трансформировалась в конформацию с большой «петлёй распознавания», и они могли превратиться в адаптеры, использующие C-DRT в качестве матрицы (рис. 3, б). При этом самоаминоацилирование могло происходить благодаря рибозимной активности при условии, что фрагмент РНК содержал сайт связывания аминокислоты и один из концов цепи находился вблизи сайта связывания аминокислоты. Это могло быть возможным при формировании L-образной тРНК-подобной структуры, т.е. прото-тРНК, имеющей

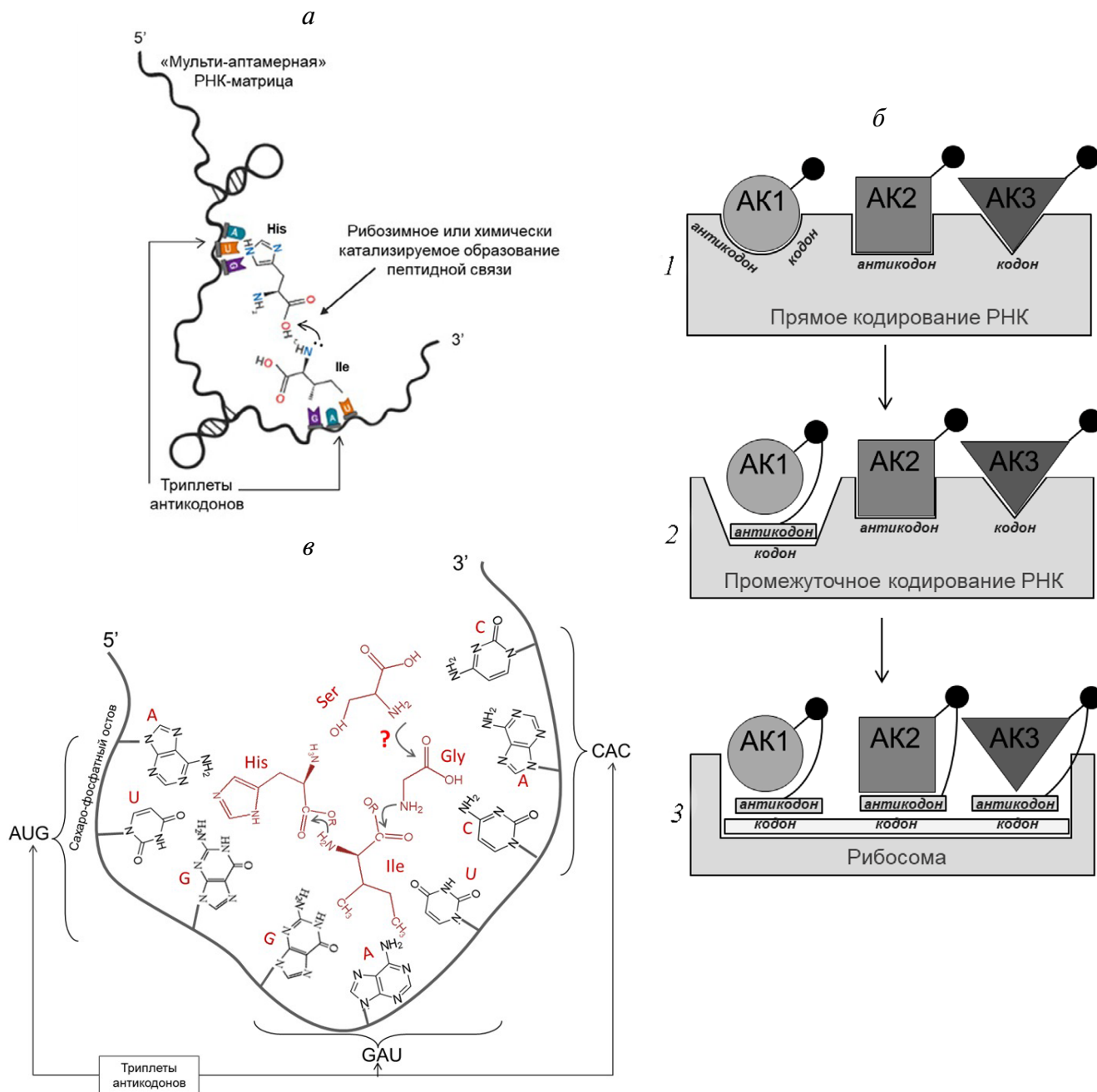


Рис. 2. Модель прямого кодирования на матрице РНК (англ. Direct RNA Template, DRT). *а* – Схематическое изображение матрицы DRT на примере фрагмента РНК, непоследовательно связывающего гистидин и изолейцин. Модифицировано по [79]. *б*, часть 1 – Первичная трансляция, при которой активированные аминокислоты связываются для полимеризации с сайтами (кодонами или антикодонами) на матрице РНК (DRT). Активирующие карбоксильные группы (*) участвуют в полимеризации аминокислот с получением закодированного пептида. *б*, часть 2 – Промежуточное кодирование РНК. Образуется аминоксил-РНК (подобие тРНК) путём объединения антикодона кодирующей РНК с аминокислотой. *б*, часть 3 – Переход к единой версии современного непрямого кодирования: появление отдельной прото-мРНК, примитивных аминоксил-тРНК, состоящих из активированных форм аминокислот, связанных с их антикодонами, и рибосомы. Модифицировано по [52, 53]. *в* – Пространственное несоответствие триплетов и аминокислот. При прямом сплошном кодировании аминокислоты либо «не дотягиваются» друг до друга, либо РНКовая нить закручивается, формируя кольца или спирали

«ногу», вертикально связывающуюся с С-DRT, и «руку», горизонтально доставляющую аминокислоты. Прото-мРНК могла бы возникнуть в результате утраты спейсерных последователь-

ностей С-DRT, находящихся между прото-тРНК-связывающими сайтами (рис. 3, *в*) [80]. Возникшая прото-тРНК распознавала прото-мРНК за счёт связывания довольно большого

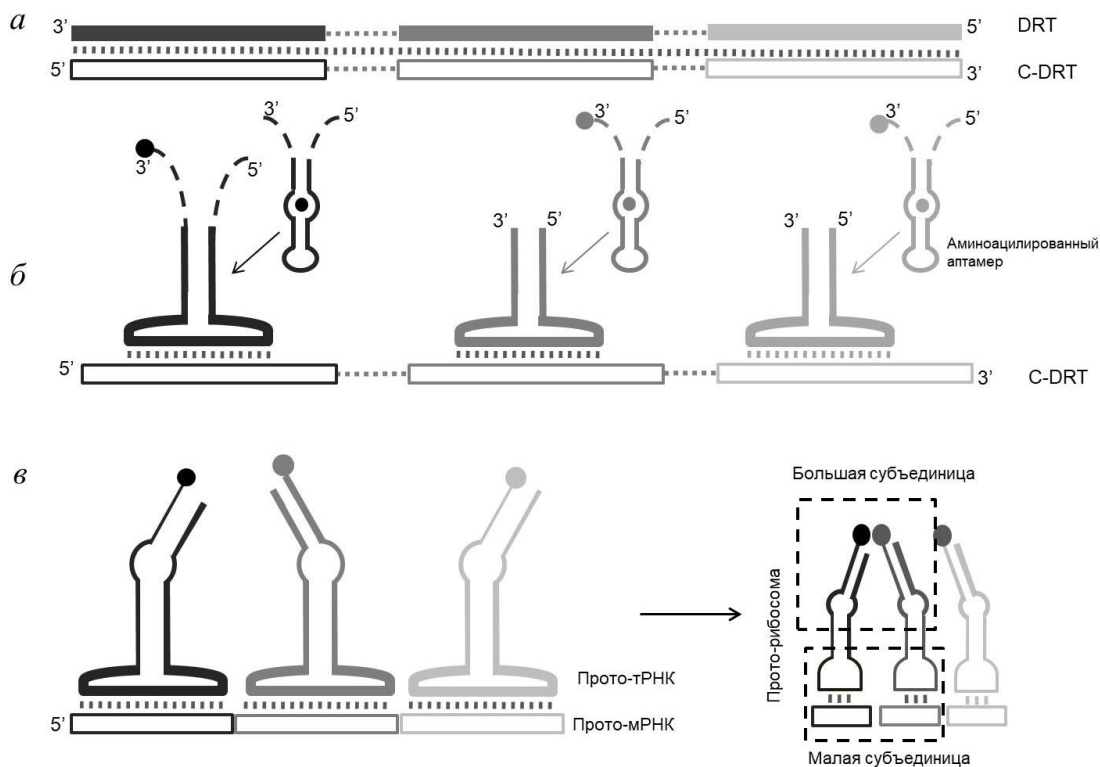


Рис. 3. Модель прямого кодирования на матрице РНК с учётом принципа репликативной экономии (англ. Direct RNA Template with replication parsimony). *a* – Репликация DRT приводит к образованию комплементарной цепи – C-DRT. *б* – В результате деградации или неполной репликации DRT могут образоваться сегменты с аминоксил-аптамерным доменом, фланкированным некоторыми спейсерными последовательностями. Они могут трансформироваться в конформацию с большой «петлём распознавания», и, таким образом, возникают адаптеры, использующие C-DRT в качестве матрицы. *в* – Возникновение мРНК: C-DRT теряет последовательности между сайтами связывания аминоксил-аптамера, появление малой субъединицы прото-рРНК способствует стабилизации кодон-антикодонного взаимодействия триплетных пар прото-тРНК и прото-мРНК. Появление большой субъединицы прото-рРНК способствует сближению аминоксилот, что приводит к образованию пептидных связей. Модифицировано по [80]

количества нуклеотидных остатков, в результате чего образовывался стабильный комплекс. При сокращении количества «антикодонных» оснований до трёх ослаблялось связывание между заряженной аминокислотой прото-тРНК и прото-мРНК, при этом появление малой субъединицы прото-рРНК могло обеспечить большую стабильность комплекса. Большая субъединица прото-рибосомы возникала в рамках этой гипотезы для обеспечения эффективности реакции пептидилпереноса и удерживала заряженные аминокислоты рядом друг с другом, работая как рибозим, таким же образом, как и большая субъединица рРНК в современной рибосоме [80].

Возникновение тРНК. Современные транспортные РНК состоят в среднем из 76 нуклеотидов (73–93), обладают универсальной последовательностью ССА на 3'-конце и вторичной структурой клеверного листа. Вторичная структура клеверного листа молекулы тРНК складывается в L-образную трёхмерную конформацию

посредством сложных третичных взаимодействий, в том числе между D- и T-плечами, образованными акцепторным стержнем плюс T-стержнем с T-петлём и D-стержнем с D-петлём вместе с антикодонным стержнем и петлём. Небольшой размер молекулы тРНК делает её привлекательным кандидатом на роль первого возникшего компонента генетической системы. Это предположение основывается на том, что до возникновения гиперциклической самовоспроизводящейся системы (см. гиперцикл Eigen–Schuster [7]) длина нуклеотидных цепей, которые могли быть точно реплицированы при дарвиновском отборе, составляла не более 100 нуклеотидов, что сопоставимо с размером молекулы тРНК.

В процессе выявления возможного механизма возникновения тРНК, проанализировав более тысячи последовательностей тРНК, Rodin et al. [81] реконструировали консенсусную структуру акцепторных доменов тРНК. Консенсусная последовательность представляла

собой двухцепочечный палиндром длиной 11 нуклеотидов с комплементарными триплетами в центре, каждый из которых был окружен мотивами 3'-ACCD и NGGU-5' на каждой нити, где D – это дискриминаторное основание. Такой палиндром мог реплицироваться в аналогичный палиндром длиной 25 пар нуклеотидов с комплементарным триплетом в середине, и далее самопроизвольная элонгация обоих палиндромов должна была привести к образованию двух спиралей с внутренними комплементарными триплетами. Наконец, удаление 5'-UGGN с 5'-конца привело к образованию современного тРНК-подобного клеверного листа (рис. 4, а). В исследовании Rodin et al. [81] также была обнаружена следующая закономерность: в парах тРНК с комплементарными антикодонами их основания во второй позиции акцепторного ствола также были комплементарными. За единственным исключением (пара тРНК-Phe/тРНК-Glu), такой параллелизм особенно впечатляет для пар тРНК, распознаваемых аминокил-тРНК-синтетазами (aaRS) из противоположных классов.

Возможным доказательством того, что тРНК возникла путём лигирования двух шпилечных молекул РНК, является следующее наблюдение. Число возможных последовательностей РНК, состоящих из четырёх нуклеотидных оснований, равно 4^n (где n – число нуклеотидов). По математической оценке Tamura [82], не учитывая функциональную частоту, общая масса, необходимая для образования нуклеотидной цепи длиной, аналогичной современной тРНК (~76 нуклеотидов), составила бы ~1/25 от общей массы Земли. Однако для полу-тРНК (~35 нуклеотидов) требуемая масса составляла бы не более 100 г, что говорит о вероятности образования современных тРНК путём, похожим на шпильки [82]. Анализ, проведённый в ряде работ, ука-

зывает на то, что большинство последовательностей тРНК возникли в результате двойной дубликации шпильки [83, 84]. Комплементарное спаривание между нуклеотидами акцепторного ствола и основаниями антикодонных ствола/петли в 5'-половине тРНК и 3'-половине тРНК соответствует двойному складыванию шпильки (рис. 4, б). Таким образом, это может свидетельствовать о том, что образование двойной шпильки в древнем пребиотическом мире лежит в основе эволюции современной структуры тРНК [82].

Возникновение рРНК. Существует предположение, также основанное на моделях объединения фрагментов молекул тРНК, которое описывает эволюцию рибосом путём слияния (аккреции) тРНК-подобных РНК [85–88]. Димерная центральная часть рибосомной РНК вокруг пептидилтрансферазного центра сходна с димерами тРНК, связанными комплементарными антикодонами. Отметим, что в современных митохондриальных рибосомах позвоночных регулярные митохондриальные тРНК конститутивно выполняют функции 5S рРНК [89], что может свидетельствовать в пользу такого сценария эволюции рРНК. Кластеризация вторичных структур РНК [90] обнаружила две основные группы вторичных структур РНК, одна из которых характеризуется небольшими, предположительно, древними тРНК-подобными вторичными структурами, и предполагаемая более поздняя группа, состоящая из более крупных рРНК-подобных вторичных структур.

Существует так называемая аккреционная модель эволюции рибосомы, основанная на исследовании тРНК-рРНК-эволюционной оси [87]. Модель предполагает возникновение предковых сегментов РНК и их объединение в структуры большой и малой прото-рРНК. Авторы аккреционной модели делят эволюцию системы

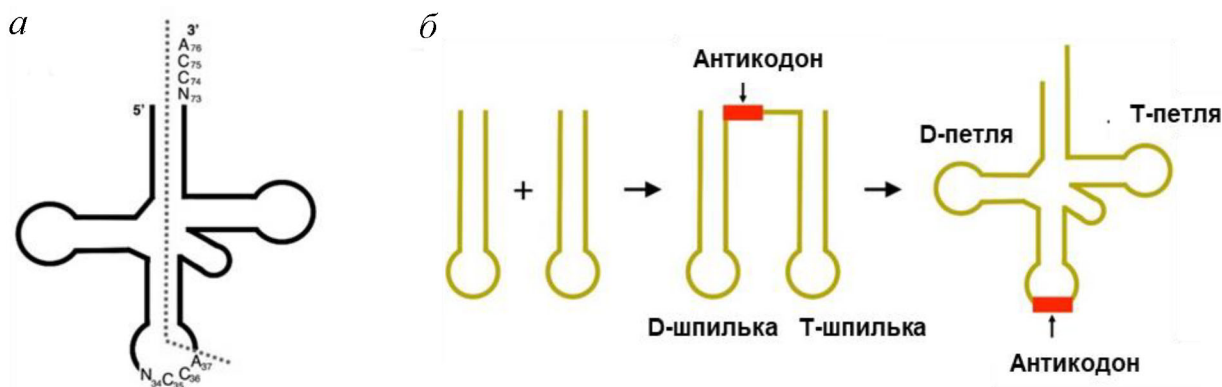


Рис. 4. Образование молекулы тРНК путём лигирования двух олигомерных молекул на примере глициновой тРНК (а). Модель образования тРНК из двух шпилечных структур (б). Модифицировано по [82]

трансляции на шесть фаз (рис. 5). В первой фазе формировались шпилечные структуры РНК, содержащие ССА-последовательность на 3'-конце и стабилизовавшиеся взаимодействиями с катионами металлов. Они служили мономерами как для будущих прото-тРНК, так и прото-рРНК. Во второй фазе шпилечные фрагменты неспецифически объединялись в олигомеры будущей большой прото-рРНК и малой прото-рРНК. Большая прото-рРНК могла представлять собой рибозим с пептидилтрансферазным центром, катализирующим неспецифическую полимеризацию аминокислот. Шпилечные структуры с заряженными аминокислотами ССА-хвостами могли доставлять аминокислоты в пептидилтрансферазный центр прото-рРНК. Функция малой прото-рРНК могла заключаться в формировании ассоциации с одноцепочечной РНК (прото-мРНК). В третьей фазе происходило объединение шпилечных структур прото-тРНК в современную L-структуру тРНК. Зарождающаяся ассоциация между субъединицами прото-рРНК опосредовалась прото-мРНК и тРНК. Малая прото-рРНК и прото-мРНК рекрутировались сформировавшейся тРНК: тРНК на одном конце связывалась с пептидилтрансферазным центром большой прото-рРНК, а на другом – с прото-мРНК, которая, в свою очередь, связывалась с одноцепочечной областью малой прото-рРНК. В четвертой фазе происходило дальнейшее объединение шпилечных структур субъединиц прото-рРНК, сформировались А- и Р-сайты, большая прото-рРНК неcodируемо полимеризовала аминокислоты, а прото-мРНК и малая прото-рРНК действовали как кофакторы, позиционирующие тРНК в А- и

Р-сайты рибозима. В пятой фазе рибосома приобрела примитивную способность к декодированию, стала энергозависимой и транслоцирующей, возник механизм храповика. Шестая фаза знаменует завершение формирования общего ядра рРНК, взаимодействий с рибосомными белками и оптимизацию генетического кода. Переход от синтеза неcodированных гетерогенных аминокислотных олигомеров к белкам с помощью рибосомы давал преимущества, поскольку некоторые продукты реакции связывались с рибосомой. Таким образом, рибосома последовательно приобретала способность к структурированию РНК, катализу, ассоциации субъединиц, коэволюции субъединиц, декодированию и передаче энергии и породила существующие симбиотические отношения белка и нуклеиновой кислоты [87].

Гипотеза de Farias et al. [91] подчёркивает центральную роль молекул тРНК в возникновении и развитии матричных процессов. Согласно этой гипотезе, тРНК дали начало первым генам (мРНК) и пептидилтрансферазному центру рибосомы (рРНК). Эволюционные превращения генетической системы в рамках этой гипотезы отражены на рис. 5. Первые гены образовались из тРНК путём структурных изменений (структура тРНК-подобной мРНК) и связывали другие тРНК (каноническая структура тРНК клеверного листа) с помощью петли прото-антикодона. Аминокислоты и небольшие пептиды при этом работали подобно кофакторам, стабилизируя альтернативные конформации прото-тРНК. Связывание двух или более тРНК повышало стабильность этих молекул. Связывающиеся с аминокислотами тРНК могли взаимодей-

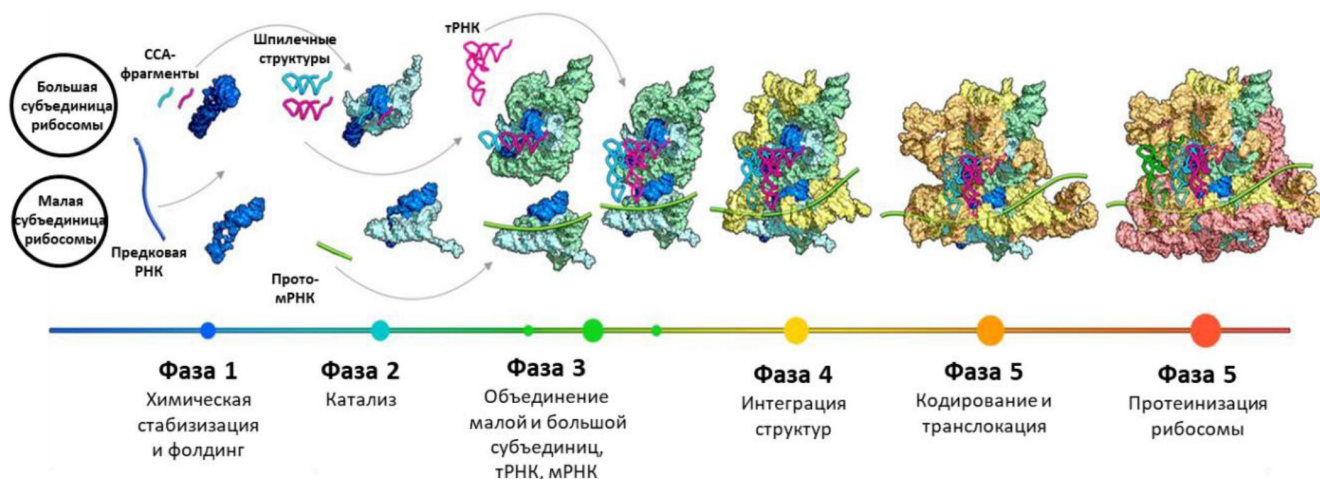


Рис. 5. Шесть фаз аккреционной модели эволюции рибосом. По работе Petrov, A. S., et al. (2015) History of the ribosome and the origin of translation, *Proc. Natl. Acad. Sci. USA*, **112**, 15396-15401, doi: 10.1073/pnas.1509761112 [87]. (PNAS не несёт ответственность за точность перевода.)

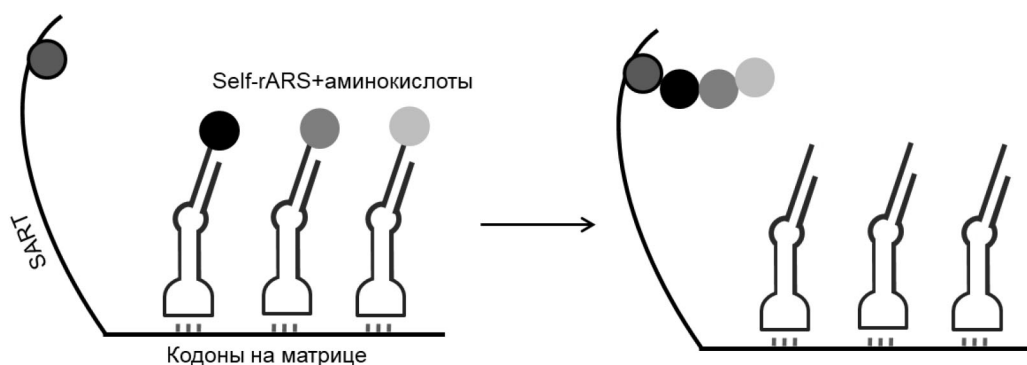


Рис. 6. Модель самоаминоацилирующихся рибозимных матриц (англ. Self-rARS Template, SART). Синтез боковой цепи из трёх аминокислот на нуклеотиде-мишени X (темно-серый круг) из конъюгатов аминокислота-Self-rARS, связанных со специфическими аминокислотами кодонными сайтами на матрице SART. Модифицировано по [94]

ствовать с другими тРНК в открытой конформации (тРНК-подобная структура мРНК). Это взаимодействие стабилизировало сложную клеверную тРНК-каноническую структуру. Кроме того, образование комплекса прото-тРНК/мРНК стабилизировало взаимодействие с прото-рРНК (прото-пептидилтрансферазный центр), образуя тройную структуру прото-тРНК/мРНК/рРНК [91]. Предполагается, что первые пептиды, которые были зафиксированы в биологической системе, обладали способностью связывать и стабилизировать тРНК в различных структурных конформациях, что, в свою очередь, увеличивало стабильность и эффективность процесса, тем самым устанавливая первую положительную обратную связь между пептидами и нуклеиновыми кислотами [92]. С увеличением стабильности системы синтеза примитивного белка её эффективность возросла, а скорость включения аминокислот в образующиеся прото-белки увеличилась [91].

Возникновение аминоацилирования РНК. Гипотеза самоаминоацилирующихся рибозимных матриц. Для образования системы кодирования, трансляции и репликации необходимо, чтобы сформировался процесс, избирательно вызывающий обогащение только функциональными РНК (фРНК). Для этого необходим механизм, позволяющий разделять молекулы на фРНК, способные связывать лиганды метаболитов, для которых селективно перезапускается репликация, и нефункциональные РНК, не связывающие лиганды и разрушающиеся [93]. Для иллюстрации такого решения Wong et al. [94] была предложена модель самоаминоацилирующихся рибозимных матриц (англ. Self-rARS Template, SART) (рис. 6).

Изучение рибозимов, обладающих аминоацилирующей активностью, привело к обнаруже-

нию самоаминоацилирующих рибозимов (англ. Self-rARS) с различной структурой и аминокислотной специфичностью [95–98]. Это открытие позволяет предположить, что такие Self-rARS были отнюдь не редким явлением в гипотетическом мире РНК. Эти Self-rARS осуществляли бы различные типы реакций. Помимо катализа *цис*-аминоацилирования их собственных 2'(3')-ОН, 5'-ОН или внутреннего 2'-ОН, некоторые из них могли быть способны трансааминоацилировать субстратную РНК. Конъюгаты аминокислот и rARS, образованные Self-rARS посредством *цис*-аминоацилирования, могли связываться с кодонными сайтами на матрице функциональной РНК и включать их аминоацильные фрагменты в пептид, связанный с ней. При этом могла использоваться либо собственная трансааминоацилирующая активность, либо принимал участие специализированный трансааминоацилирующий рибозим, например, пяти-нуклеотидный rARS, который работает с различными аминокислотами [97]. В любом случае, порядок кодонов сайтов для различных собственных rARS на РНК-матрице определял аминокислотную последовательность получающегося в результате пептида [94].

Надо отметить, что существование в рамках РНК-мира рибозимных aaRS, функционирующих столь же эффективно, что и современные белковые аминоацилирующие ферменты, привело бы к ситуации, когда отсутствовала бы необходимость в этих компонентах системы трансляции именно белковой природы.

Гипотеза РНК-коферментов – кодирующих коферментных «ручек». Существует другая гипотеза, связывающая РНК-мир и аминокислоты, – The Coding Coenzyme Handle Hypothesis (CCH) [99], в которой аминокислоты действуют как коферменты для рибозимов. Данная теория

базируется на том, что многие современные коферменты обладают нуклеотидоподобными фрагментами, что может свидетельствовать об их возможном происхождении от мира РНК. В рибо-организмах нуклеотидоподобные части коферментов могли быть использованы в качестве «ручек» (возможно, несущих другие мелкие молекулы), благодаря чему рибозимы могли легко захватывать их и управлять ими.

В рамках данной гипотезы ранние адаптеры, предки современных антикодонов, были заряжены аминокислотами, действующими как коферменты рибозимов в метаболически сложном мире РНК. Предковые аминоксил-адаптерные синтетазы могли быть подобны современным самосплайсинговым интронам тРНК. Кодон/антикодон-дискриминаторный базовый комплекс, встроенный в эти синтетазы, мог бы сыграть важную роль в распознавании аминокислот.

СН-гипотеза также пытается объяснить, каково было адаптивное преимущество генетического кода. Специфичность однозначного связывания аминокислот с «ручками» определяла то, какая аминокислота должна была быть присоединена, а рибозимы были первыми катализаторами этого процесса установления соответствия между триплетами и аминокислотами. Позже в ходе эволюции «ручки» превратились в адаптеры (тРНК), рибозимы были заменены белковыми аминоксил-тРНК-синтетазами, и многие рибозимы стали молекулами мРНК, потеряв свою первоначальную ферментативную активность. Таким образом, этот механизм мог являться преадаптацией к кодированию и трансляции [100].

Изначально предполагалось, что «ручки» могли быть моно- или динуклеотидными [99], однако только тринуклеотидные соединения могли обеспечить стабильное связывание [100]. В работе Kazakov и Altman [101] показано, что тринуклеотид может катализировать металл-ион-зависимое расщепление специфического мотива последовательности РНК GAAA.

Согласно теории коферментных «ручек» [100], предполагается, что сначала возникла антикодоновая шпилька тРНК, а акцепторный ствол тРНК был эволюционным следствием. Свидетельством этого являются данные [102], благодаря которым была обнаружена эволюционная связь между антикодоном и первым-третьим триплетом (и иногда семидесятым-семьдесят вторым) в акцепторном стволе. Далее шпилька антикодона могла стать более длинной, несущей повтор антикодона на 3'-конце молекулы благодаря появлению адаптеров, которые связывают аминокислоту с по-

мощью более реакционноспособных сложноэфирных связей.

НУКЛЕОПЕПТИДНЫЙ МИР

Некоторые исследователи [103–105] предположили, что полноценного мира РНК на самом деле никогда не существовало: вместо этого они предполагают очень раннюю коэволюцию РНК и пептидов в «рибонуклеопептидном» мире. Свидетельством в пользу такого сценария является предполагаемое постепенное добавление рибосомных белков к рибосомным РНК [106]. Таким образом, возможно, образовавшиеся рибосомные РНК начали взаимодействовать с аминокислотами или пептидами, что повысило стабильность, эффективность и специфичность примитивной рибосомы [104]. На более поздней стадии эволюции генетически кодируемые белки могли заменить эти пептиды. Следовательно, строгое сохранение структуры рибосомных белков и их повсеместное присутствие во всех трёх доменах жизни отражают их важный вклад в поддержании структуры и функции рибосомы [104]. Кроме того, рибосомы не могут функционировать без своих белковых частей, что может свидетельствовать об одновременном возникновении белковой части и РНК [107]. Соответственно, другие рибонуклеопротеиды, такие как сплайсосома и РНКазы Р, могли столкнуться с той же эволюционной судьбой. И хотя сплайсосома, как предполагается, возникла из самосплайсирующихся интронов группы II, которые не нуждаются в белках для собственного сплайсинга, белки им всё же необходимы для фолдинга и стабилизации РНК, а также обратной транскрипции [108].

В целом большинство рибозимов имеют ограниченные каталитические ресурсы и не демонстрируют каких-либо доказательств значительной сложности [29]. Взаимодействие между пептидами и рибозимами могло существовать на ранних этапах эволюции и эффективно способствовать увеличению каталитического разнообразия рибозимов, придавая новые функции рибозимам и ускоряя эволюционный процесс [104].

Система из двух полимераз. К коэволюционным теориям развития ранних событий эволюции можно отнести теорию о двух полимеразах Kunin [109]. Он предложил теорию зарождения репликационно-трансляционной системы, включающуюся в возникновении tandemно работающих молекул: РНК, катализирующей образование пептидных связей, и олигопептида, катализирующего репликацию РНК. Благодаря

способности молекулы РНК катализировать образование пептидных связей синтезировались олигопептиды из аминокислот, доступных в «первичном бульоне». По такому механизму могло происходить присоединение аминокислот как к коротким пептидам, так и непосредственно к аминокислотам. Короткие пептиды могли также образовывать пептидные связи друг с другом вследствие неспецифического пептидилтрансферазного катализа молекулой РНК. Рост цепи прекращался из-за пространственных помех, поскольку пептидная цепь становилась длиннее и приобретала трёхмерную структуру. Среди таких синтезированных пептидов могли быть и пептиды, обладающие активностью РНК-зависимой РНК-полимеразы. Эта активность олигопептидов позволила бы реплицировать молекулы РНК. Реакция полимеризации использовала РНК в качестве матрицы и рибонуклеотиды, доступные в окружающей среде, и могла протекать в две температурно-зависимые стадии. При повышенных температурах водородные связи между участками молекул РНК разрушались, делая молекулы доступными для РНК-зависимой РНК-полимеразы, катализировавшей полимеризацию РНК. Когда температура снижалась, молекулы РНК укладывались во вторичную структуру и снова становились способны катализировать образование пептидных связей. Смена дня и ночи могла быть причиной таких температурных колебаний [109].

Последовательность молекул РНК в рамках данной гипотезы зависела от процесса естественного отбора. Наследственность в системе поддерживалась через сохранение нуклеотидной последовательности молекул РНК. Неточная репликация и, возможно, другие типы мутаций могли давать материал для отбора. Улучшение каталитической активности РНК привело бы к увеличению концентрации каталитических пептидов в ближайшем микроокружении, а каталитические пептиды, в свою очередь, реплицировали бы ближайшую доступную РНК, т.е. в большинстве случаев ту, которая послужила матрицей для их синтеза. В результате лучший РНК-катализатор имел больше шансов быть воспроизведённым. Таким образом, такая двухполимеразная система может рассматриваться как примитивный дарвиновский организм [109].

При кажущейся простоте теории сам автор теории указывает на её проблемы. 1. Достаточно ли специфичности системы для успешного синтеза РНК-полимераз? 2. Будут ли неизбежно образованы случайные ферменты, расходующие «впустую» активированные мономеры? 3. Является ли синтез пептидов возможным при нали-

чии только одной молекулы РНК? [109]. Описанная двухполимеразная система может быть инициирована появлением одного компонента — одна каталитическая РНК может синтезировать пептид с РНК-полимеразной активностью, тем самым создавая всю систему, а пептидная РНК-полимераза, не имеющая матрицы, могла бы способствовать случайному синтезу молекул РНК с периодическим образованием функциональных рибозимов. Это означает, что появление любого компонента системы могло запустить автокаталитический каскад [109]. При этом данная гипотеза не описывает появление системы кодирования, и обе полимеразы формируются случайным образом среди пула нефункциональных РНК и олигопептидов. Такая гипотетическая пептидная полимеразы не кодирована, а значит, не сможет воспроизводиться, а полимеразы из РНК не смогут вести процессивный синтез пептида, обладающего полимеразной активностью.

Гипотеза прогенов. Существует альтернативная РНК-миру гипотеза о возникновении генетической системы клетки — гипотеза прогенов, предложенная Альтштейном [107, 110, 111]. Данная гипотеза объяснила возникновение одновременного синтеза полипептидов и полинуклеотидов, которые эволюционируют не последовательно, а взаимосвязанно. Согласно этой гипотезе, нуклеотиды сначала объединялись в аминокислотированные тринуклеотиды (прогены), которые являлись единственным субстратом для одновременного синтеза полинуклеотида и полипептида, при этом последовательность аминокислот кодировалась в полинуклеотиде, а образующийся полипептид обладал свойствами процессивной полимеразы (прогенлигазы). Общая формула прогена $NpNpNr \sim pX \sim aa$, где N — нуклеозид, p — фосфат, X — бифункциональный агент, например рибоза или другой углевод, aa — аминокислота, \sim — макроэргическая связь. При этом аминокислота специфична по отношению к своему триплету.

В рамках данной гипотезы прогены образуются путём соединения динуклеотида $NpNp$ и аминокислотированного нуклеотида $Np \sim pX \sim aa$, специфически обусловленного взаимодействием динуклеотида и аминокислоты третьего нуклеотида. Механизм образования прогенов представлен на рис. 7. Этот механизм позволяет объяснить первичное возникновение генетического кода и отбор веществ для возникающей генетической системы, включая гомохиральность [95, 96]. Фосфат в нуклеотидах прогена находится на 3'-, а не на 5'-конце (это повышает вероятность образования фосфодиэфирной связи), и третий нуклеотид должен обязательно

быть дезоксирибонуклеотидом, второй – чаще дезоксирибонуклеотидом, а первый – чаще рибонуклеотидом, т.е. полинуклеотид является полирибодезоксирибонуклеотидом.

Такие прогены являются мономерами для одновременного синтеза полинуклеотида и полипептида (рис. 7). Два прогена за счёт слабого стэкинг-взаимодействия сближаются друг с другом, и второй проген удерживается рядом с первым взаимодействием с аминокислотой первого. Стабилизируется комплекс комплементарным олигонуклеотидом, состоящим из 6 нуклеотидов. *N*-Конец аминокислоты второго прогена сближается с активированным *C*-концом аминокислоты первого прогена; образуется дипептид, связанный со вторым прогеном. *N*-Концевая аминокислота дипептида, Glu или Asp (из-за их кислотных боковых групп), влияет на образование фосфодиэфирной связи между первым и вторым прогенами. В результате образуется гексануклеотид, связанный с дипептидом. Дипептид обладает повышенной (по сравнению с одной аминокислотой) способностью удерживать следующий проген, и весь процесс повторяется, при этом дипептид переносится на третий проген (транспептидация) и помогает последнему присоединяться к гексануклеотиду. Образуется нонануклеотид с трипептидом. Таким образом, при присоединении прогенов каждый триплет запоминает кодируемую аминокислоту и устанавливается связь «генотип (порядок триплетов в полинуклеотиде)–фенотип (порядок аминокислот в полипептиде)» [107].

Растущий пептид постоянно участвует в формировании системы: если последняя аминокислота, присоединяющаяся к пептиду, увеличивает вероятность включения следующего прогена, рост всей системы продолжается; если уменьшает – рост системы останавливается. Таким образом, возникающий полипептид-фермент растёт в тесном взаимодействии с его будущим субстратом (прогеном) и, как очень редкое

событие, формируется связанная пара «ген–процессивная полимераз (прогенлигаза)», причём полимераз имеет тропность к 3'-концу полинуклеотида.

Полимераза, расположенная на 3'-конце гена, перемещается к 5'-концу, соединяя прогены, комплементарные матрице. Одновременно синтез комплементарного полинуклеотида происходил посредством репликации и транскрипции, поскольку исходный ген был представлен (–)-цепью, и сопровождается синтезом «неправильного» белка, кодируемого (–)-цепью. Полимераза перемещается к 3'-концу вновь синтезированной (+)-цепи и реплицирует её, одновременно формируя новую молекулу полимеразы и новую (–)-цепь. Две молекулы полимеразы дополнительно создают две новые (+)-цепи. Затем синтезируются две новые (–)-цепи и две дополнительные молекулы полимеразы (всего 4 молекулы полимеразы).

Таким образом, гипотеза прогенов предлагает единый механизм, который позволяет объяснить: 1) возникновение генетического кода; 2) механизм отбора компонентов для образования генетической системы, в том числе возникающей гомохиральности; и 3) возникновение и размножение бимолекулярной генетической системы, состоящей из гена полинуклеотида и белковой процессивной полимеразы, кодируемой этим геном.

Согласно данной гипотезе, возникшая бимолекулярная система – это первый вирусоподобный организм (*Protoviroidum primum*), который размножается и эволюционирует путём единого процесса репликации, транскрипции и трансляции и является прародителем жизни на Земле [107]. Экспериментальная химическая проверка механизма образования прогенов принципиально возможна, но пока не осуществлена.

Гипотеза урзимов. Согласно другой гипотезе, описанной Li et al. [112], катализ и кодирование

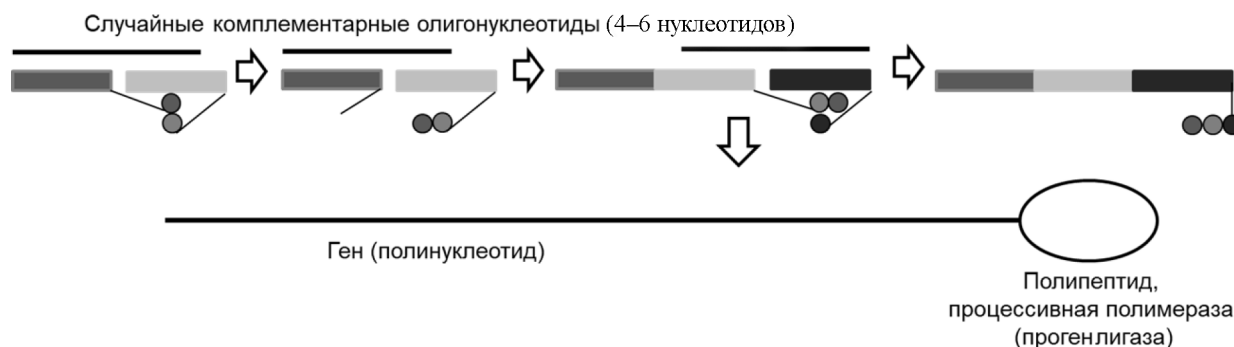


Рис. 7. Формирование бимолекулярной генетической системы из прогенов. Модифицировано по [107]

совместно развились благодаря одновременному образованию пептидов и РНК.

В центре этой гипотезы лежат «урзимы», представляющие собой небольшие (120–130 аминокислот), высококонсервативные фрагменты доменов двух суперсемейств аминоацил-тРНК-синтетаз (aaRS). Два этих суперсемейства – aaRS I и aaRS II – были когда-то закодированы противоположными цепями одного и того же гена [113]. В отличие от рибозимов, которые лежат в основе гипотезы о мире РНК, пары урзимов I- и II-классов представляют собой консенсусные предковые формы, достаточные для кодон-направленного синтеза неслучайных пептидов. Наличие механизма «активации» аминокислот экспоненциально увеличивает эффективность их полимеризации, а возникновение урзимов решает эту проблему. Урзимы значительно снижают структурную сложность, необходимую для осуществления трансляции рудиментарного генетического кода. Таким образом, урзимы проявляют каталитические свойства задолго до полного формирования генетического кода.

Урзимная модель коэволюции предполагает стереохимическое «комплементарное» взаимодействие между РНК и пептидами, основывающееся на примитивном генетическом коде, в котором каждый нуклеотид в дуплексе РНК соответствует двум аминокислотам, а каждый дипептид определяет соответствующее основание. Подобное взаимодействие описано в работе Carter и Kraut [114]. Стереохимическое кодирование могло генерировать пептиды и соответствующие им РНК-гены. При этом такие гены будут иметь длину гена тРНК (~72 основания). Таким образом, мир, основанный на коэволюции РНК и белков, состоял из коротких РНК-«генов» и кодируемых ими небольших белков, специфически аминоацелирующих молекулы РНК, благодаря которым осуществлялась рибозимная трансляция.

Автор такой гипотезы о коэволюции РНК и белков, Carter [115], предлагает следующий сценарий пребиотической эволюции: в течение длительного периода химической эволюции мономеры аминокислот и нуклеиновых кислот начали собираться в ковалентные комплексы, включающие структурно «комплементарные» олигонуклеотиды и дипептиды; далее могли происходить синтез и лигирование пептидов и олигонуклеотидов, специфичность которых была бы ограничена спариванием нуклеотидных оснований и грубым стереохимическим кодированием между двумя типами полимеров, присоединяющих новые мономеры таким образом, чтобы стабилизировать двойной спиральный

комплекс пептид–РНК; комплексы росли в длину, возможно, до размера предполагаемого гена, кодирующего 23-аминокислотные смысловые/антисмысловые продукты. 23-аминокислотные смысловые/антисмысловые пептиды стали родоначальниками aaRS ферментов класса I и II; 23-аминокислотные пептиды были лигированы в 46-аминокислотный белок, обладавший большей эффективностью аминоацелирования. Некоторые из образовавшихся пептидов всё более специализировались как полимеразы нуклеиновых кислот, другие аминоацелировали двухцепочечную РНК, двухцепочечная РНК сохранила свою роль шаблона репликации и развила пептидилтрансферазную активность, став большой рибосомальной субъединицей. Одноцепочечная РНК стала всё чаще брать на себя доминирующую роль матрицы и эволюционировала в современную мРНК [115].

Авторы гипотезы об урзимах, Li et al. [112], указывают на два недостатка РНК-мира: 1) отсутствие естественных рибозимов, реплицирующих РНК, активирующих аминокислоты и аминоацелирующих тРНК, и 2) зависимость рибозимного пути от раннего появления точной репликации. Такой путь не предлагает механизма для преодоления «обрыва Эйгена» [116]. Оба этих слабых места, согласно Li et al. [112], объясняются совместной эволюцией пептидных катализаторов и РНК.

Данная гипотеза интересна для понимания эволюции трансляции, но не даёт чёткого проверяемого ответа на вопрос о кодировании других пептидов. Непонятно, как может эволюционировать такая система, в которой белки для репликации не кодируются РНК.

БЕЛКОВЫЙ МИР

Гипотеза GADV-белкового мира. Гипотеза GADV-белкового мира основана на наблюдении, что белки, содержащие примерно одинаковое количество четырёх аминокислот (глицина – G (gly), аланина – A (ala), аспарагиновой кислоты – D (asp) и валина – V (val)), могут образовывать основные вторичные структуры белка – α -спирали и β -слои [117, 118]. В рамках этой теории предполагается, что основные GADV-белки могли полимеризоваться случайно из этих четырёх аминокислот, формируя глобулы. Согласно этой гипотезе, пребиотический мир до появления генов включал стадию, на которой GADV-белки были способны к псевдо-репликации. Возникновение примитивных биологических систем в таком случае должно было проходить следующие стадии: 1) пребиотичес-

кий синтез и накопление gly, ala, asp и val; 2) случайное формирование белков из этих четырёх аминокислот; 3) формирование первичной метаболической системы за счёт активности GADV-белков, синтез нуклеотидов и олигонуклеотидов РНК; 4) формирование первичного генетического кода и образование генов; 5) образование двойной спирали РНК, возникновение наследственности и передачи генетического материала поколениям — начало биологической эволюции; 6) эволюция генетического кода и генетических систем.

Согласно этой гипотезе, примитивный генетический код состоял из предварительного GNC-кода, включавшего четыре кодона — GGC, GCC, GAC и GUC, кодирующих 4 основные аминокислоты GADV-мира: gly, ala, asp и val. Затем произошло расширение генетического кода до 16 кодонов (GGC, GGG, GCC, GCG, GAC, GAG, GUC, GUG, CUC, GUG, CCC, CCG, CAC, CAG, CGC и CGG), кодировавших 10 аминокислот (gly, ala, asp, val, glu, leu, pro, his, gln и arg) [119]. Порядок появления аминокислот в пребиотическом GADV-мире примерно совпадает с установленным Trifonov [120] эволюционным порядком образования аминокислот, основанном на 60 различных критериях.

Примитивные одноцепочечные нуклеиновые кислоты состояли из GNC-повторяющихся последовательностей и были получены путём образования фосфодиэфирной связи между кодонами GNC. Двухцепочечные молекулы образовывались при комплементарном спаривании молекул (GNC)_n или кодонов GNC, и, таким образом, была сформирована первичная система репликации [121].

Проблемой гипотезы «белкового мира» является то, что стабильная и функциональная структура возможна только для коротких пребиотических пептидов, а случайное формирование правильной пространственной укладки сложных пептидов/белков в таких условиях практически невозможно. Кроме того, пробелом данной теории является отсутствие чёткого пути для репликации и передачи информации на ранних этапах эволюции, основанных на пептидах.

ЛИПИДНЫЙ МИР

Гипотезы липидного мира [122–124] предполагают, что простые гидрофобные молекулы липидов могут сохранять и распространять информацию, таким образом, претерпевая эволюцию.

Модель GARD. Модель липидного мира Lancet et al. [124] основывается на предположе-

нии, что воспроизведение состава компонентов пребиотического мира предшествует репликации непосредственно информационных последовательностей. Согласно этому взгляду на протобиологию, первые репликаторы были скоплениями спонтанно объединяющихся, гетерогенных и в основном неканонических (не образующих бислойные структуры) амфифильных веществ. Это предположение подтверждается построением химической кинетики модели репликации мицеллярного автокатализа (англ. Graded autocatalysis replication domain, GARD) [122]. Согласно GARD, в пребиотическом мире существовал отбор неравновесных агрегаций веществ (композомов), которые могли быть способны к автокаталическому гомеостатическому росту [124]. Такой процесс, наряду со случайным делением композома, воплощал в себе примитивное клеточное размножение. При наличии в среде амфифильных веществ из них самоформируются мицеллы, этот факт может служить предпосылкой для возникновения самоподдерживающихся автокаталитических липидных систем. В таком случае GARD-эволюция мира-до-РНК проявлялась в отборе различных композомов в пределах разреженного ландшафта пригодности в ответ на химические изменения окружающей среды. По модели GARD «генотипом» являлся сам композом, т.е. состав автокаталитических мицелл, и он же определял способность к репликации, т.е. «фенотип» системы. «Геном» композома в такой системе оказывает прямое кинетическое влияние на эффективность и точность гомеостатического роста, следовательно, на его репликацию. Таким образом, авторы гипотезы указывают на корреляцию между мутациями в системе и её репродукцией без необходимости в посреднике, таком как кодируемый белок или функциональная РНК [124]. Гипотеза предполагает раннее объединение ключевых компонентов биологии в липидной протоклетке: амфифильные композомы могли включать аминокислоты и пептиды, азотистые основания РНК и олигонуклеотиды, способные к комплементарному спариванию, кофакторы, хелаторы металлов, тиолы, олигофосфаты и многие другие соединения, между которыми формировались метаболические связи. Данная модель не даёт ответов на вопросы о возникновении именно генетических систем, но связывает при этом между собой разные классы молекул. Кроме того, моделирование GARD демонстрирует в образующихся системах те же термодинамические и кинетические особенности, которые соответствуют предполагаемому предшественнику современных живых клеток. GARD композомы характеризуются ав-

торами как диссипативные системы, что также является отличительными признаками жизни [124]. При этом сами авторы гипотезы указывают на скудость экспериментальных данных для взаимно каталитических сетей.

РНК-липидный мир. Идея возникновения жизни из амфифильных композитов [122] была развита в гипотезе РНК-липидного мира Mallik и Kundu [123]. Гипотеза о коэволюции липидов и РНК основана на предположении, что современные стабильные взаимодействия липид-РНК — это молекулярные «окаменелости» древней стадии эволюции, когда мир РНК возник из мира липидов. Предполагается, что древние рибозимы были «защищены» липидами, что обеспечило их выживание в жёстких условиях пребиотического океана. Присутствующие в окружающей среде ионы металлов могли катализировать репликацию РНК внутри защищённых липидных мицелл. Авторами представлен возможный в рамках этой модели механизм трансляции и перехода от РНК-липидного мира к РНК-белковому: РНК-прото-рибосомы, прикрепленные к внутреннему слою липидного пузырька, случайно полимеризовали аминокислоты. Синтез случайных полипептидов в большом пуле везикул мог привести к образованию функциональных белков, которые могли сыграть важную роль в ранней эволюции [123].

ЗАКЛЮЧЕНИЕ

Наши знания и представления о современных организмах и молекулах, которые они содержат, обеспечивают формирование предпосылок для предположения о том, что развитие биологических систем началось с эволюции семейств молекул, которые могли бы катализировать их собственную репликацию. Существует большое количество гипотез, которые описывают подобные сценарии, наиболее широко известные из них мы постарались представить в данном обзоре. Однако практически ни одна из этих гипотез не описывает без пробелов и допущений все этапы ранней эволюции генетических систем.

Гипотеза РНК-мира сегодня является главенствующей научной идеей, и существует немало доводов в её пользу, которые мы описали и которым посвящено множество работ, в том числе рассмотренных в настоящем обзоре. В рамках этой теории разработано много вероятных механизмов эволюции компонентов трансляции. При этом главная проблема мира РНК до сих пор не решена — не достигнуто полное самовоспроизведение РНК. Открытие процессивной рибозимной полимеразы (репликазы), способ-

ной реплицировать длинные полинуклеотидные последовательности, решительно укрепило бы позиции гипотезы мира РНК, но такого достижения ещё нет [38]. В настоящее время не известно ни одной построенной только из нуклеотидов процессивной полимеразы. Яркими примерами таких молекулярных машин являются разнообразные современные ДНК- и РНК-полимеразы, которые имеют чисто белковую или нуклеопротеидную природу [125]. Очень выразительна ситуация с рибосомой, которая, по сути, является молекулярной машиной, полимеризующей аминокислоты при движении по полинуклеотидной матрице (мРНК). Известно, что рибосомная РНК является рибозимом, способным соединять аминокислоты пептидной связью [125]. Однако рибосома содержит 30–40% белка, без которого не может выполнять свою функцию. После 1990 г. были искусственно получены рибозимы со свойствами РНК-полимеразы [22, 23], однако это полимеразы дистрибутивного типа, не способные двигаться по матрице. Таким образом, рибозимы могут образовывать и фосфодиэфирные связи между нуклеотидами, и пептидные связи между аминокислотами, но, вероятно, не способны без белка выполнять функцию молекулярных машин. Если это действительно так, гипотеза мира РНК является таким же тупиковым направлением для решения проблемы происхождения жизни, как и гипотеза Опарина [126], поскольку не может предложить конкретного механизма возникновения генетической системы.

Кроме того, несмотря на то что РНК обладает способностью выполнять многие функции белковых ферментов, нет данных об участии рибозимов в функционировании клеточных систем, которые должны были возникнуть на самых ранних этапах биологической эволюции. Так, рибозимы не принимают участия в синтезе нуклеотидов, в работе энергетических систем, в аминокислотировании тРНК. При этом искусственно получены рибозимы, способные эффективно и специфично аминокислотировать тРНК [95–98], но все естественные aaRS являются белками. Эти факты трудно совместимы с гипотезой мира РНК и позволяют думать о том, что кодированный синтез белка имел место на самых ранних этапах эволюционного развития живого мира.

Представляется разумной идея о коэволюции двух типов макромолекул — нуклеиновых кислот и белков. Такой сценарий не столь экономичен в эволюционном отношении, как живой мир на основе одной только РНК, однако, вероятно, что пептиды (и полипептиды) играли важную роль в раннем биологическом мире. Ги-

потезой, которая хорошо дополняет мир РНК вероятными ранними белковыми молекулами, требовавшимися для осуществления трансляции, является гипотеза урзимов. Эволюционно консервативные aaRS двух классов могли способствовать более точному установлению соответствия тРНК и аминокислот в трансляции. Комплементарные цепи РНК кодировали каждая свой пептид, невысокая точность репликации могла привести к образованию новой молекулы, при этом её функциональность могла не нарушиться при сохранении пространственной структуры. И таким образом, гипотеза урзимов избегает «обрыва Эйгена», потому что наиболее важным её ограничением является не точность последовательности, а сохранение структурных элементов и каталитических активностей коротких полипептидных шпилек. При этом маловероятным совпадением кажется сохранение комплементарности цепочек РНК, кодирующих такие пептиды, при низкой точности репликации.

Идея о том, что генетические системы сформировались на основе двух коэволюционирующих макромолекул разных классов, реализована в гипотезе двух полимераз, а также в гипотезе прогенов. В последней на основе одного принципа (образование аминокислотированных нуклеотидных триплетов) теоретически решаются несколько важнейших проблем теории возникновения жизни: отбор веществ, включая гомохиральность, механизм возникновения первич-

ного генетического кода и первой бимолекулярной генетической системы, реплицирующейся на основе единого процесса репликации-транскрипции-трансляции. Автором предлагается уникальная концепция возможного первого живого неклеточного организма – протовироида и описывается принципиальный механизм его биологической (дарвиновской) эволюции в направлении к клетке. Главный недостаток гипотезы заключается в отсутствии её экспериментальной проверки, которая, однако, принципиально возможна.

Критика рассмотренных представлений о возникновении и эволюции трансляции не означает, что важные экспериментальные данные, полученные на основе гипотезы РНК-мира, и идеи, выдвинутые в этой и других гипотезах, не могут быть использованы для понимания важных аспектов эволюции живого мира и не могут быть положены в основу будущих экспериментов.

Финансирование. Работа выполнена при поддержке Программы стратегического академического лидерства РУДН.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Соблюдение этических норм. Настоящая статья не содержит описания каких-либо исследований с участием людей или животных в качестве объектов.

СПИСОК ЛИТЕРАТУРЫ

- Dodd, M. S., Papineau, D., Grenne, T., Slack, J. F., Rittner, M., et al. (2017) Evidence for early life in Earth's oldest hydrothermal vent precipitates, *Nature*, **543**, 60-64, doi: 10.1038/nature21377.
- Benner, S. A. (2010) Defining life, *Astrobiology*, **10**, 1021-1030, doi: 10.1089/ast.2010.0524.
- Wolf, Y. I., and Koonin, E. V. (2007) On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization, *Biol. Direct*, **2**, 14, doi: 10.1186/1745-6150-2-14.
- Nikolaeva, D. D., Gelfand, M. S., and Garushyants, S. K. (2021) Simplification of ribosomes in bacteria with tiny genomes, *Mol. Biol. Evol.*, **38**, 58-66, doi: 10.1093/molbev/msaa184.
- Eigen, M. (1971) Selforganization of matter and the evolution of biological macromolecules, *Naturwissenschaften*, **58**, 465-523, doi: 10.1007/bf00623322.
- Penny, D. (2005) An interpretive review of the origin of life research, *Biol. Philos.*, **20**, 633-671, doi: 10.1007/s10539-004-7342-6.
- Eigen, M., and Schuster, P. (1977) The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle, *Naturwissenschaften*, **64**, 541-565, doi: 10.1007/BF00450633.
- Szathmáry, E. (1986) Some remarks on hypercycles and the stochastic corrector model, *Endocyt. Cell Res.*, **3**, 337-339.
- Szathmáry, E., and Demeter, L. (1987) Group selection of early replicators and the origin of life, *J. Theor. Biol.*, **128**, 463-486.
- Zintzaras, E., Santos, M., and Szathmáry, E. (2002) "Living" under the challenge of information decay: the stochastic corrector model vs. hypercycles, *J. Theor. Biol.*, **217**, 167-181.
- Gago, S., Elena, S. F., Flores, R., and Sanjuan, R. (2009) Extremely high mutation rate of a hammerhead viroid, *Science*, **323**, 1308, doi: 10.1126/science.1169202.
- Koonin, E. V., and Novozhilov, A. S. (2017) Origin and evolution of the universal genetic code, *Annu. Rev. Genet.*, **51**, 45-62, doi: 10.1146/annurev-genet-120116-024713.
- Rich, A. (1962) *Horizons in Biochemistry* (Kasha, M., and Pullman, B., eds.) Academic Press, New York.
- Crick, F. H. (1968) The origin of the genetic code, *J. Mol. Biol.*, **38**, 367-379, doi: 10.1016/0022-2836(68)90392-6.
- Orgel, L. E. (1968) Evolution of the genetic apparatus, *J. Mol. Biol.*, **38**, 381-393, doi: 10.1016/0022-2836(68)90393-8.
- Gilbert, W. (1986) Origin of life: The RNA world, *Nature*, **319**, 618.
- Le Vay, K., and Mutschler, H. (2019) The difficult case of an RNA-only origin of life, *Emerg. Top. Life Sci.*, **3**, 469-475, doi: 10.1042/etls20190024.
- Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., et al. (1982) Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena, *Cell*, **31**, 147-157, doi: 10.1016/0092-8674(82)90414-7.
- Robertson, H. D., Altman, S., and Smith, J. D. (1972) Purification and properties of a specific *Escherichia coli*

- ribonuclease which cleaves a tyrosine transfer ribonucleic acid precursor, *J. Biol. Chem.*, **247**, 5243-5251.
20. Klemm, B. P., Wu, N., Chen, Y., Liu, X., Kaitany, K. J., et al. (2016) The diversity of ribonuclease P: Protein and RNA catalysts with analogous biological functions, *Biomolecules*, **6**, doi: 10.3390/biom6020027.
 21. Walter, N. G., and Engelke, D. R. (2002) Ribozymes: catalytic RNAs that cut things, make things, and do odd and useful jobs, *Biologist*, **49**, 199-203.
 22. Johnston, W. K., Unrau, P. J., Lawrence, M. S., Glasner, M. E., and Bartel, D. P. (2001) RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension, *Science*, **292**, 1319-1325, doi: 10.1126/science.1060786.
 23. Wöchner, A., Attwater, J., Coulson, A., and Holliger, P. (2011) Ribozyme-catalyzed transcription of an active ribozyme, *Science*, **332**, 209-212, doi: 10.1126/science.1200752.
 24. Bowman, J. C., Hud, N. V., and Williams, L. D. (2015) The ribosome challenge to the RNA world, *J. Mol. Evol.*, **80**, 143-161, doi: 10.1007/s00239-015-9669-9.
 25. Ouzounis, C., and Kyriades, N. (1996) The emergence of major cellular processes in evolution, *FEBS Lett.*, **390**, 119-123, doi: 10.1016/0014-5793(96)00631-x.
 26. Ganoza, M. C., Kiel, M. C., and Aoki, H. (2002) Evolutionary conservation of reactions in translation, *Microbiol. Mol. Biol. Rev.*, **66**, 460-485, doi: 10.1128/mmr.66.3.460-485.2002.
 27. Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G., et al. (2012) One core, two shells: bacterial and eukaryotic ribosomes, *Nat. Struct. Mol. Biol.*, **19**, 560-567, doi: 10.1038/nsmb.2313.
 28. Hsiao, C., Lenz, T. K., Peters, J. K., Fang, P. Y., Schneider, D. M., et al. (2013) Molecular paleontology: a biochemical model of the ancestral ribosome, *Nucleic Acids Res.*, **41**, 3373-3385, doi: 10.1093/nar/gkt023.
 29. Bernhardt, H. S. (2012) The RNA world hypothesis: the worst theory of the early evolution of life (except for all the others)(a), *Biol. Direct*, **7**, 23, doi: 10.1186/1745-6150-7-23.
 30. Bregestovski, P. D. (2015) "RNA World", a highly improbable scenario of the origin and early evolution of life on earth, *J. Evol. Biochem. Physiol.*, **51**, 72-84, doi: 10.1134/S0022093015010111.
 31. Wills, P. R., and Carter, C. W., Jr. (2018) Insuperable problems of the genetic code initially emerging in an RNA world, *Biosystems*, **164**, 155-166, doi: 10.1016/j.biosystems.2017.09.006.
 32. Orgel, L. E. (2004) Prebiotic chemistry and the origin of the RNA world, *Crit. Rev. Biochem. Mol. Biol.*, **39**, 99-123, doi: 10.1080/10409230490460765.
 33. Kim, H. J., Ricardo, A., Illangkoon, H. I., Kim, M. J., Carrigan, M. A., et al. (2011) Synthesis of carbohydrates in mineral-guided prebiotic cycles, *J. Am. Chem. Soc.*, **133**, 9457-9468, doi: 10.1021/ja201769f.
 34. Спирин А. С. (2007) Когда, где и в каких условиях мог возникнуть и эволюционировать мир РНК? *Палеонтологический журнал*, **41**, 11-19.
 35. Tupper, A. S., Shi, K., and Higgs, P. G. (2017) The role of templating in the emergence of RNA from the prebiotic chemical mixture, *Life*, **7**, 41, doi: 10.3390/life7040041.
 36. Szostak, J. W. (2012) The eightfold path to non-enzymatic RNA replication, *J. Syst. Chem.*, **3**, 2, doi: 10.1186/1759-2208-3-2.
 37. Joyce, G. F., and Orgel, L. E. (1986) Non-enzymic template-directed synthesis on RNA random copolymers. Poly(C, G) templates, *J. Mol. Biol.*, **188**, 433-441, doi: 10.1016/0022-2836(86)90166-x.
 38. Joyce, G. F., and Szostak, J. W. (2018) protocells and RNA Self-Replication, *Cold Spring Harb. Perspect. Biol.*, **10**, doi: 10.1101/cshperspect.a034801.
 39. James, K. D., and Ellington, A. D. (1999) The fidelity of template-directed oligonucleotide ligation and the inevitability of polymerase function, *Orig. Life Evol. Biosph.*, **29**, 375-390, doi: 10.1023/a:1006544611320.
 40. Prywes, N., Blain, J. C., Del Frate, F., and Szostak, J. W. (2016) Nonenzymatic copying of RNA templates containing all four letters is catalyzed by activated oligonucleotides, *eLife*, **5**, doi: 10.7554/eLife.17756.
 41. Szostak, J. W. (2011) An optimal degree of physical and chemical heterogeneity for the origin of life? *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **366**, 2894-2901, doi: 10.1098/rstb.2011.0140.
 42. Bartel, D. P., and Szostak, J. W. (1993) Isolation of new ribozymes from a large pool of random sequences [see comment], *Science*, **261**, 1411-1418, doi: 10.1126/science.7690155.
 43. Horning, D. P., and Joyce, G. F. (2016) Amplification of RNA by an RNA polymerase ribozyme, *Proc. Natl. Acad. Sci. USA*, **113**, 9786-9791, doi: 10.1073/pnas.1610103113.
 44. Samanta, B., and Joyce, G. F. (2017) A reverse transcriptase ribozyme, *eLife*, **6**, doi: 10.7554/eLife.31153.
 45. Koonin, E. V. (2017) Frozen accident pushing 50: Stereochemistry, expansion, and chance in the evolution of the genetic code, *Life (Basel)*, **7**, doi: 10.3390/life7020022.
 46. Nirenberg, M. W., and Matthaei, J. H. (1961) The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides, *Proc. Nat. Acad. Sci. USA*, **47**, 1588-1602.
 47. Gardner, R. S., Wahba, A. J., Basilio, C., Miller, R. S., et al. (1962) Synthetic polynucleotides and the amino acid code, VII, *Proc. Natl. Acad. Sci. USA*, **48**, 2087.
 48. Wahba, A. J., Gardner, R. S., Basilio, C., Miller, R. S., Speyer, J. F., et al. (1963) Synthetic polynucleotides and the amino acid code, VIII, *Proc. Natl. Acad. Sci. USA*, **49**, 116.
 49. Söll, D., Ohtsuka, E., Jones, D., Lohrmann, R., Hayatsu, H., et al. (1965) Studies on polynucleotides, XLIX. Stimulation of the binding of aminoacyl-sRNA's to ribosomes by ribotrinucleotides and a survey of codon assignments for 20 amino acids, *Proc. Natl. Acad. Sci. USA*, **54**, 1378.
 50. Woese, C. R., Dugre, D. H., Dugre, S. A., Kondo, M., and Saxinger, W. C. (1966) On the fundamental nature and evolution of the genetic code, *Cold Spring Harb. Symp. Quant. Biol.*, **31**, 723-736, doi: 10.1101/sqb.1966.031.01.093.
 51. Woese, C. R. (1968) The fundamental nature of the genetic code: prebiotic interactions between polynucleotides and polyamino acids or their derivatives, *Proc. Natl. Acad. Sci. USA*, **59**, 110-117, doi: 10.1073/pnas.59.1.110.
 52. Yarus, M., Widmann, J. J., and Knight, R. (2009) RNA-amino acid binding: a stereochemical era for the genetic code, *J. Mol. Evol.*, **69**, 406-429, doi: 10.1007/s00239-009-9270-1.
 53. Yarus, M. (1998) Amino acids as RNA ligands: A direct-RNA-template theory for the code's origin, *J. Mol. Evol.*, **47**, 109-117, doi: 10.1007/pl00006357.
 54. Yarus, M. (2000) RNA-ligand chemistry: a testable source for the genetic code, *RNA*, **6**, 475-484, doi: 10.1017/s1355838200002569.
 55. Yarus, M., Caporaso, J. G., and Knight, R. (2005) Origins of the genetic code: the escaped triplet theory, *Annu. Rev. Biochem.*, **74**, 179-198, doi: 10.1146/annurev.biochem.74.082803.133119.
 56. Koonin, E. V., and Novozhilov, A. S. (2009) Origin and evolution of the genetic code: the universal enigma, *IUBMB Life*, **61**, 99-111, doi: 10.1002/iub.146.
 57. Hajnic, M., Osorio, J. I., and Zagrovic, B. (2014) Computational analysis of amino acids and their sidechain analogs in crowded solutions of RNA nucleobases with implications for the mRNA-protein complementarity hypothesis, *Nucleic Acids Res.*, **42**, 12984-12994.
 58. Hlevnjak, M., Polyansky, A. A., and Zagrovic, B. (2012) Sequence signatures of direct complementarity between mRNAs and cognate proteins on multiple levels, *Nucleic Acids Res.*, **40**, 8874-8882, doi: 10.1093/nar/gks679.

59. Polyansky, A. A., Hlevnjak, M., and Zagrovic, B. (2013) Analogue encoding of physicochemical properties of proteins in their cognate messenger RNAs, *Nat. Commun.*, **4**, 2784, doi: 10.1038/ncomms3784.
60. Polyansky, A. A., Hlevnjak, M., and Zagrovic, B. (2013) Proteome-wide analysis reveals clues of complementary interactions between mRNAs and their cognate proteins as the physicochemical foundation of the genetic code, *RNA Biol.*, **10**, 1248-1254, doi: 10.4161/rna.25977.
61. Gamow, G. (1954) Possible relation between deoxyribonucleic acid and protein structures, *Nature*, **173**, 318.
62. Melcher, G. (1974) Stereospecificity of the genetic code, *J. Mol. Evol.*, **3**, 121-140, doi: 10.1007/bf01796558.
63. Hendry, L., Bransome, E., Hutson, M., and Campbell, L. (1981) First approximation of a stereochemical rationale for the genetic code based on the topography and physicochemical properties of "cavities" constructed from models of DNA, *Proc. Natl. Acad. Sci. USA*, **78**, 7440-7444, doi: 10.1073/pnas.78.12.7440.
64. Balasubramanian, R., Seetharamulu, P., and Raghunathan, G. (1980) A conformational rationale for the origin of the mechanism of nucleic acid-directed protein synthesis of "living" organisms, *Orig. Life*, **10**, 15-30, doi: 10.1007/bf00928940.
65. Shimizu, M. (1982) Molecular basis for the genetic code, *J. Mol. Evol.*, **18**, 297-303, doi: 10.1007/bf01733895.
66. Massey, S. E. (2006) A sequential "2-1-3" model of genetic code evolution that explains codon constraints, *J. Mol. Evol.*, **62**, 809-810, doi: 10.1007/s00239-005-0222-0.
67. Wong, J. T. (1975) A co-evolution theory of the genetic code, *Proc. Natl. Acad. Sci. USA*, **72**, 1909-1912, doi: 10.1073/pnas.72.5.1909.
68. Wong, J. T. (2005) Coevolution theory of the genetic code at age thirty, *Bioessays*, **27**, 416-425, doi: 10.1002/bies.20208.
69. Amirnovin, R. (1997) An analysis of the metabolic theory of the origin of the genetic code, *J. Mol. Evol.*, **44**, 473-476, doi: 10.1007/pl00006170.
70. Ronneberg, T. A., Landweber, L. F., and Freeland, S. J. (2000) Testing a biosynthetic theory of the genetic code: fact or artifact? *Proc. Natl. Acad. Sci. USA*, **97**, 13690-13695, doi: 10.1073/pnas.250403097.
71. Woese, C. R. (1965) On the evolution of the genetic code, *Proc. Natl. Acad. Sci. USA*, **54**, 1546-1552, doi: 10.1073/pnas.54.6.1546.
72. Haig, D., and Hurst, L. D. (1991) A quantitative measure of error minimization in the genetic code, *J. Mol. Evol.*, **33**, 412-417, doi: 10.1007/BF02103132.
73. Novozhilov, A. S., Wolf, Y. I., and Koonin, E. V. (2007) Evolution of the genetic code: partial optimization of a random code for robustness to translation error in a rugged fitness landscape, *Biol. Direct*, **2**, 24, doi: 10.1186/1745-6150-2-24.
74. Mukai, T., Lajoie, M. J., Englert, M., and Söll, D. (2017) Rewriting the genetic code, *Annu. Rev. Microbiol.*, **71**, 557-577.
75. Спирин А. (2001) Биосинтез белков, мир РНК и происхождение жизни, *Вестник Российской академии наук*, **71**, 320-328.
76. Spirin, A. (2013) The emergence of molecular machines as a prerequisite of the ancient RNA world evolution, *Paleontol. J.*, **47**, 1016-1029.
77. Schimmel, P. (2011) The RNP bridge between two worlds, *Nat. Rev. Mol. Cell. Biol.*, **12**, 135, doi: 10.1038/nrm3061.
78. Nissen, P., Hansen, J., Ban, N., Moore, P. B., and Steitz, T. A. (2000) The structural basis of ribosome activity in peptide bond synthesis, *Science*, **289**, 920-930, doi: 10.1126/science.289.5481.920.
79. Dunn, I. S. (2011) RNA templating of molecular assembly and covalent modification patterning in early molecular evolution and modern biosystems, *J. Theor. Biol.*, **284**, 32-41, doi: 10.1016/j.jtbi.2011.06.009.
80. Ma, W. (2010) The scenario on the origin of translation in the RNA world: in principle of replication parsimony, *Biol. Direct*, **5**, 65, doi: 10.1186/1745-6150-5-65.
81. Rodin, S., Rodin, A., and Ohno, S. (1996) The presence of codon-anticodon pairs in the acceptor stem of tRNAs, *Proc. Natl. Acad. Sci. USA*, **93**, 4537-4542, doi: 10.1073/pnas.93.10.4537.
82. Tamura, K. (2015) Origins and early evolution of the tRNA molecule, *Life (Basel)*, **5**, 1687-1699, doi: 10.3390/life5041687.
83. Di Giulio, M. (1992) On the origin of the transfer RNA molecule, *J. Theor. Biol.*, **159**, 199-214, doi: 10.1016/s0022-5193(05)80702-7.
84. Widmann, J., Di Giulio, M., Yarus, M., and Knight, R. (2005) tRNA creation by hairpin duplication, *J. Mol. Evol.*, **61**, 524-530, doi: 10.1007/s00239-004-0315-1.
85. Caetano-Anolles, D., and Caetano-Anolles, G. (2016) Piecemeal buildup of the genetic code, ribosomes, and genomes from primordial tRNA building blocks, *Life (Basel)*, **6**, doi: 10.3390/life6040043.
86. De Farias, S. T., Rêgo, T. G., and José, M. V. (2021) Origin of the 16S ribosomal molecule from ancestor tRNAs, *J. Mol. Evol.*, **89**, 249-256, doi: 10.1007/s00239-021-10002-8.
87. Petrov, A. S., Gulen, B., Norris, A. M., Kovacs, N. A., Bernier, C. R., et al. (2015) History of the ribosome and the origin of translation, *Proc. Natl. Acad. Sci. USA*, **112**, 15396-15401, doi: 10.1073/pnas.1509761112.
88. Demongeot, J., and Seligmann, H. (2020) Comparisons between small ribosomal RNA and theoretical minimal RNA ring secondary structures confirm phylogenetic and structural accretion histories, *Sci. Rep.*, **10**, 1-14.
89. Amunts, A., Brown, A., Toots, J., Scheres, S. H., and Ramakrishnan, V. (2015) The structure of the human mitochondrial ribosome, *Science*, **348**, 95-98.
90. Seligmann, H., and Raoult, D. (2018) Stem-loop RNA hairpins in giant viruses: invading rRNA-like repeats and a template free RNA, *Front. Microbiol.*, **9**, 101.
91. De Farias, S. T., Rego, T. G., and Jose, M. V. (2016) tRNA core hypothesis for the transition from the RNA world to the ribonucleoprotein world, *Life (Basel)*, **6**, doi: 10.3390/life6020015.
92. Noller, H. F. (2012) Evolution of protein synthesis from an RNA world, *Cold Spring Harb. Perspect. Biol.*, **4**, a003681, doi: 10.1101/cshperspect.a003681.
93. Wong, J. T. (2014) Emergence of life: From functional RNA selection to natural selection and beyond, *Front. Biosci. (Landmark Ed)*, **19**, 1117-1150, doi: 10.2741/4271.
94. Wong, J. T., Ng, S. K., Mat, W. K., Hu, T., and Xue, H. (2016) Coevolution theory of the genetic code at age forty: Pathway to translation and synthetic life, *Life (Basel)*, **6**, doi: 10.3390/life6010012.
95. Illangasekare, M., Sanchez, G., Nickles, T., and Yarus, M. (1995) Aminoacyl-RNA synthesis catalyzed by an RNA, *Science*, **267**, 643-647, doi: 10.1126/science.7530860.
96. Lee, N., Bessho, Y., Wei, K., Szostak, J. W., and Suga, H. (2000) Ribozyme-catalyzed tRNA aminoacylation, *Nat. Struct. Biol.*, **7**, 28-33, doi: 10.1038/71225.
97. Suga, H., Hayashi, G., and Terasaka, N. (2011) The RNA origin of transfer RNA aminoacylation and beyond, *Philos. Trans. R. Soc. B*, **366**, 2959-2964, doi: 10.1098/rstb.2011.0137.
98. Turk, R. M., Chumachenko, N. V., and Yarus, M. (2010) Multiple translational products from a five-nucleotide ribozyme, *Proc. Natl. Acad. Sci. USA*, **107**, 4585-4589, doi: 10.1073/pnas.0912895107.
99. Szathmari, E. (1993) Coding coenzyme handles: a hypothesis for the origin of the genetic code, *Proc. Natl. Acad. Sci. USA*, **90**, 9916-9920, doi: 10.1073/pnas.90.21.9916.
100. Szathmari, E. (1999) The origin of the genetic code: Amino acids as cofactors in an RNA world, *Trends Genet.*, **15**, 223-229, doi: 10.1016/s0168-9525(99)01730-8.
101. Kazakov, S., and Altman, S. (1992) A trinucleotide can promote metal ion-dependent specific cleavage of RNA, *Proc. Natl. Acad. Sci. USA*, **89**, 7939-7943, doi: 10.1073/pnas.89.17.7939.

102. Rodin, S. N., and Ohno, S. (1997) Four primordial modes of tRNA-synthetase recognition, determined by the (G,C) operational code, *Proc. Natl. Acad. Sci. USA*, **94**, 5183-5188, doi: 10.1073/pnas.94.10.5183.
103. Di Giulio, M. (2008) An extension of the coevolution theory of the origin of the genetic code, *Biol. Direct*, **3**, 37, doi: 10.1186/1745-6150-3-37.
104. Saad, N. Y. (2018) A ribonucleopeptide world at the origin of life, *J. Syst. Evol.*, **56**, 1-13, doi: 10.1111/jse.12287.
105. Altstein, A. D., and Efimov, A. V. (1988) Physico-chemical basis of the genetic code origin: stereochemical analysis of interactions of amino acids and nucleotides based on the progene hypothesis, *Mol. Biol. (Mosk)*, **22**, 1411-1429.
106. Fox, G. E. (2010) Origin and evolution of the ribosome, *Cold Spring Harb. Perspect. Biol.*, **2**, a003483, doi: 10.1101/cshperspect.a003483.
107. Altstein, A. D. (2015) The progene hypothesis: the nucleoprotein world and how life began, *Biol. Direct*, **10**, 67, doi: 10.1186/s13062-015-0096-z.
108. Lambowitz, A. M., and Zimmerly, S. (2011) Group II introns: mobile ribozymes that invade DNA, *Cold Spring Harb. Perspect. Biol.*, **3**, a003616, doi: 10.1101/cshperspect.a003616.
109. Kunin, V. (2000) A system of two polymerases – a model for the origin of life, *Orig. Life Evol. Biosph.*, **30**, 459-466, doi: 10.1023/a:1006672126867.
110. Альштейн А. Д., Каверин Н. Н. (1980) О происхождении вирусных генетических систем, *Журн. Всесоюз. хим. общ.им. Д.И. Менделеева*, **25**, 383-390.
111. Альштейн А. Д. (1987) Происхождение генетической системы: гипотеза прогенов, *Мол. Биол.*, **21**, 309-321.
112. Li, L., Francklyn, C., and Carter, C. W., Jr. (2013) Aminoacylating urzymes challenge the RNA world hypothesis, *J. Biol. Chem.*, **288**, 26856-26863, doi: 10.1074/jbc.M113.496125.
113. Rodin, A. S., Rodin, S. N., and Carter, C. W., Jr. (2009) On primordial sense-antisense coding, *J. Mol. Evol.*, **69**, 555-567, doi: 10.1007/s00239-009-9288-4.
114. Carter, C. W., Jr., and Kraut, J. (1974) A proposed model for interaction of polypeptides with RNA, *Proc. Natl. Acad. Sci. USA*, **71**, 283-287, doi: 10.1073/pnas.71.2.283.
115. Carter, C. W. (2015) What RNA world? Why a peptide/RNA partnership merits renewed experimental attention, *Life (Basel)*, **5**, 294-320, doi: 10.3390/life5010294.
116. Koonin, E. V. (2011) *The Logic of Chance: The Nature and Origin of Biological Evolution*, FT Press.
117. Ikehara, K. (2002) Origins of gene, genetic code, protein and life: comprehensive view of life systems from a GNC-SNS primitive genetic code hypothesis, *J. Biosci.*, **27**, 165-186, doi: 10.1007/BF02703773.
118. Ikehara, K. (2005) Possible steps to the emergence of life: the [GADV]-protein world hypothesis, *Chem. Record*, **5**, 107-118, doi: 10.1002/tcr.20037.
119. Ikehara, K., Omori, Y., Arai, R., and Hirose, A. (2002) A novel theory on the origin of the genetic code: a GNC-SNS hypothesis, *J. Mol. Evol.*, **54**, 530-538, doi: 10.1007/s00239-001-0053-6.
120. Trifonov, E. N. (2004) The triplet code from first principles, *J. Biomol. Struct. Dyn.*, **22**, 1-11, doi: 10.1080/07391102.2004.10506975.
121. Ikehara, K. (2014) [GADV]-protein world hypothesis on the origin of life, *Orig. Life Evol.*, **44**, 299-302, doi: 10.1007/s11084-014-9383-4.
122. Segre, D., Ben-Eli, D., Deamer, D. W., and Lancet, D. (2001) The lipid world, *Orig. Life Evol.*, **31**, 119-145, doi: 10.1023/a:1006746807104.
123. Mallik, S., and Kundu, S. (2012) The lipid-RNA world, arXiv preprint arXiv:1211.0413.
124. Lancet, D., Zidovetzki, R., and Markovitch, O. (2018) Systems protobiology: origin of life in lipid catalytic networks, *J. R. Soc. Interface*, **15**, doi: 10.1098/rsif.2018.0159.
125. Krebs, J. E., Goldstein, E. S., and Kilpatrick, S. T. (2017) *Lewin's genes XII*, Jones & Bartlett Learning.
126. Miller, S. L., Schopf, J. W., and Lazcano, A. (1997) Oparin's "Origin of Life": sixty years later, *J. Mol. Evol.*, **44**, 351-353.

THE ORIGIN OF THE GENETIC CODE AND TRANSLATION WITHIN CURRENT CONCEPTS OF THE ORIGIN OF LIFE

Review

L. G. Kondratyeva¹, M. S. Dyachkova², and A. V. Galchenko^{3*}

¹ Shemyakin–Ovchinnikov Institute of bioorganic chemistry, Russian Academy of Sciences, 117997 Moscow, Russia

² Vavilov Institute of General Genetics, 119991 Moscow, Russia

³ Peoples' Friendship University of Russia (RUDN University), 117198 Moscow, Russia; e-mail: gav.jina@gmail.com

The origin of the genetic code and translation system is perhaps the central and the most difficult problem in studying the origin of life and one of the most difficult in evolutionary biology at all. There are plenty of hypotheses for the emergence and development of present genetic systems, considering the origin and early evolution of the genetic code, as well as the emergence of replication and translation. The most famous hypotheses are discussed in this review. However, none of these hypotheses describes without gaps and assumptions all stages of the early evolution of genetic systems. The RNA-world hypothesis is the dominant scientific idea of the early evolution of biological and pre biological objects for now. Its main advantage is that it offers RNA as the first living systems to be self-sufficient. Namely, RNA are capable of functioning as a catalytic component of the system, and at the same time as a template. However, there are also significant limitations. In particular, ribozyme processive polymerase has not yet been discovered or synthesized. Taking into account the proteins and nucleic acids mutual need in each other in the present world, many authors propose scenarios of early evolution based on the co-evolution of these two classes of organic molecules. Hypotheses of this kind postulate that the emergence of translation was necessary for the replication of nucleic acids, in contrast to the world of RNA, where the emergence of translation was preceded by the era of self-replicating RNA. And although such scenarios are less economical from the point of view of the evolution, since they require the simultaneous emergence and evolution of two classes of organic molecules, as well as synchronization of replication and translation emergence, their great advantage is that they offer the development of a processive and much more accurate protein-dependent replication.

Keywords: evolution, origin of life, translation, genetic code, RNA world, progene, protein world, lipid world