

ПИСЬМА В РЕДАКЦИЮ

Жизнь возникла как генетический код

(к дискуссии по статье Г.Р. Иваницкого

"XXI век: что такое жизнь с точки зрения физики"

[УФН 180 337 (2010); УФН 182 1235 (2012); УФН 182 1238 (2012)])

Ю.Н. Корыстов

Цель этого письма — конкретизация сценария возникновения жизни. Дано определение жизни как самовоспроизводящейся информационной системы. Кратко рассмотрено современное состояние проблемы происхождения жизни. Предложен первый бимолекулярный сценарий происхождения жизни, основанный на взаимодействии аминокислот с триплетами нуклеотидов, соответствующими их кодонам. По этому сценарию воспроизводство информации осуществлялось благодаря образованию водородных связей с комплементарными нуклеотидами и, следовательно, образованию комплементарных кодонов, которые воспроизводили исходный кодон. Принципиальным следствием бимолекулярного сценария происхождения жизни является то, что для кодонов всех аминокислот, включённых в жизнь, должны существовать комплементарные кодоны. Показано полное соответствие кодонов с комплементарными кодонами для 21 аминокислоты. Этот факт является существенным аргументом в пользу предлагаемой гипотезы, поскольку при случайном выборе аминокислотами кодонов такое совпадение невозможно.

Ключевые слова: аминокислоты, происхождение жизни, генетический код

PACS numbers: 87.14. –g, 87.18. –h, 87.23.kg

DOI: <https://doi.org/10.3367/UFNr.2016.11.037996>

Содержание

1. Введение (235).
 2. Определение жизни (235).
 3. Современное состояние проблемы происхождения жизни (236).
 4. Бимолекулярный сценарий возникновения жизни (236).
 5. Некоторые следствия бимолекулярного сценария возникновения жизни (237).
 6. Заключение (240).
- Список литературы (240).

1. Введение

В статье Г.Р. Иваницкого [1] постулировано возникновение жизни в результате процесса игры взаимодействий материи, в которой одна её часть приобрела свойство запоминать вероятности появления удач и неудач в предыдущих раундах этой игры, повышая тем самым свой шанс на существование в последующих раундах. В комментарии В.И. Кляцкина [2] к этой статье обосновывается несогласие с тем, что жизнь возникла в процессе игры, и автор заключает, что это событие произошло с вероятностью единица в существовавших на Земле условиях. В ответе на этот комментарий Г.Р. Иваницкий отмечает: «Главная загадка возникновения живого заключается не в самом факте структуризации, а в интервале времени от его возникновения до исчезновения, т.е. времени его жизни *t*. Возникший кластер, для того чтобы превратиться в

"живую структуру", должен запомнить условия своего возникновения и передать результат запоминания потомству» [3]. В настоящем сообщении предлагается конкретный сценарий возникновения жизни, согласно которому при наличии четырёх канонических нуклеотидов и аминокислот с вероятностью единица возникают все возможные триплеты нуклеотидов, а запоминаются (воспроизводятся) только те, с которыми взаимодействуют аминокислоты.

2. Определение жизни

Проблема происхождения жизни неразрывно связана с её определением. Невозможно обсуждать происхождение чего-либо без понимания того, что именно произошло. Поэтому перед рассмотрением гипотезы происхождения жизни попытаемся дать её определение. К настоящему времени предложено множество определений жизни, апеллирующих для этого к различным подходам: физиологическому, метаболическому, биохимическому, генетическому и термодинамическому. Каждый из этих подходов либо не охватывает всех проявлений жизни, либо допускает включение в жизнь заведомо неживых систем [4]. Помоему, наиболее близкое к верному пониманию сущности жизни было дано в работах Эйгена: "все реакции в живой системе контролируются программой, действующей из информационного центра, чьей целью является самовоспроизводство самой программы" [5, 6]. В это утверждение необходимо внести одну поправку: программа — алгоритм (инструкция) структур, а реакции она контролирует косвенно, через структуры. Выделим принципиальные моменты сущности жизни по Эйгену: 1) воспроизводство информации, 2) необходимость структур и реакций для воспроизводства информации, 3) информация является алгоритмом (инструкцией) для воспроизводящих её структур и реакций. Обозначим совокупность структур и реакций по воспроизводству информации термином "оператор". Учитывая приведённые выше сущностные

Ю.Н. Корыстов. Институт теоретической и экспериментальной биофизики РАН, ул. Институтская 3, 142290 Пущино, Московская обл., Российская Федерация
E-mail: ykorystov@rambler.ru

Статья поступила 23 августа 2016 г.,
после доработки 17 ноября 2016 г.

характеристики жизни, можно сформулировать её определение: жизнь — самовоспроизводящаяся информационная система (СИС), необходимыми элементами которой являются информация (алгоритм структуры оператора) и оператор, воспроизводящий информацию. Поскольку в настоящее время существуют и другие СИС (культура, материальным носителем информации в которой являются нейроны мозга человека), то для полного определения жизни необходимо указать материальный носитель информации и оператор, специфичные для жизни. Таким образом, полное определение жизни формулируется следующим образом: *жизнь — самовоспроизводящаяся информационная система (СИС), необходимыми элементами которой является информация (алгоритм структуры оператора), материализованная последовательностью азотистых оснований, и состоящий из аминокислот оператор, воспроизводящий информацию.*

3. Современное состояние проблемы происхождения жизни

В настоящее время большинство авторов согласны с тем, что воспроизводство информации является ключевой проблемой возникновения жизни. Гипотезы о происхождении жизни разделяются на две группы лишь по вопросу того, что возникло первым: метаболизм ("metabolism first") или информация ("genetics-first") [7–9]. После открытия каталитической активности РНК была предложена гипотеза о том, что жизнь возникла в виде самовоспроизводящихся молекул РНК, способных также катализировать химические реакции, "РНК-мир" [10–12]. Однако поскольку в первичном бульоне существовали не только канонические нуклеотиды, но и большое число других нуклеотидов, то спонтанное возникновение самовоспроизводящейся РНК (не менее 40 исключительно канонических нуклеотидов) невероятно [3, 13]. Кроме того, при таком подходе к происхождению жизни появляются неразрешимые проблемы с происхождением генетического кода, трансляции и включением аминокислот и белков в генетические программы. Понимание того, что гипотезы происхождения жизни разделены с существующей жизнью логической пропастью, пришло лишь в последнее время со стороны специалистов по информатике [14]. Приведём некоторые цитаты из этой работы: "отсутствие разделения между алгоритмом и его реализацией указывает на то, что мономолекулярные системы отделены от известной жизни логической и организационной пропастью, которая не может быть преодолена усложнением пассивных аппаратных средств", и далее: "жизнь должна была быть бимолекулярной на старте". Авторы также отмечают, что все перечисленные выше гипотезы происхождения жизни упускают из рассмотрения существенное свойство биологической информации — то, что это алгоритм, инструкция. От себя добавим, что это инструкция оператора. Впервые в литературе о происхождении жизни указано, что информация должна быть не только воспроизводящейся (как в РНК-мире), но и содержательной. Об алгоритме первичной информации авторы ничего не пишут. В настоящей работе предлагается бимолекулярный вариант возникновения жизни, согласно которому содержанием начальной информации были аминокислоты, обсуждается также механизм воспроизводства информации.

4. Бимолекулярный сценарий возникновения жизни

Предполагается, что L-аминокислоты взаимодействовали со спонтанно образующимися триплетами канонических нуклеотидов, соответствующими их кодоном. Взаимодействие аминокислот с кодонами увеличивало время существования последних на срок, достаточный для образования водородных связей с комплементарными каноническими нуклеотидами. При стэкинге оснований в связанных

водородными связями комплементарных нуклеотидах они полимеризовались [15] в комплементарный кодон. Воспроизводство информации исходного кодона происходило, когда другая аминокислота, для которой возникший комплементарный кодон является своим, стабилизировала комплементарный кодон, и он воспроизводил исходный кодон за счёт образования водородных связей с комплементарными нуклеотидами. Таким образом, эта система воспроизводила информацию (не прямо, а через посредника), материализованную последовательностью азотистых оснований, с помощью операторов — аминокислот. Содержанием информации являлась структура операторов — аминокислот. Как следует из приведённого выше определения жизни, эта система уже была живой.

Предлагаемая в настоящей работе гипотеза происхождения жизни базируется на трёх положениях:

1. В пребиотических условиях образуются азотистые основания и триплеты канонических нуклеотидов.

2. Азотистые основания образуют стопки за счёт стэкинга и водородные связи с комплементарными основаниями.

3. L-аминокислоты взаимодействуют со своими кодонами так, что время существования кодона увеличивается и становится достаточным для образования водородных связей с тремя комплементарными каноническими нуклеотидами, — воспроизводство информации через комплементарный кодон.

Давно было показано, что при ультрафиолетовом облучении раствора аденина, рибозы и фосфатов образуется АТФ [16]. В 2009 году были получены пиримидиновые нуклеотиды из смеси цианоацетилена, цианамиды, глицеральдегида, гликоальдегида и фосфата [17]. При ультрафиолетовом облучении раствора часть цитозина превращалась в урацил. В работах [18, 19] было показано, что добавление к смеси, использованной в работе [17], аминокислот с небольшим избытком (1 %) L-стереоизомеров приводит к образованию хирально чистых, содержащих только D-рибозу, рибонуклеотидов. Азотистые основания и аминокислоты образуются также при нагревании формамида, а в присутствии источника фосфата — циклические нуклеотиды, которые могут полимеризоваться до олигонуклеотидов разной длины при стэкинге их оснований [15].

Справедливость второго положения общеизвестна. В одной из последних работ на эту тему показано, что при нейтральных значениях pH среды происходит "стэкинг" канонических оснований и образование водородных связей с комплементарными основаниями [20]. Это происходит потому, что при значениях pH среды около нейтральных основания не заряжены и взаимодействуют друг с другом за счёт стэкинга и образования водородных связей. По мнению автора работы, именно поэтому канонические азотистые основания были выбраны в качестве информационного носителя жизни. Такие взаимодействия характерны также для азотистых оснований, находящихся в составе нуклеотидов [21].

Предположение о том, что генетический код обусловлен взаимодействием аминокислот с нуклеотидными последовательностями, было выдвинуто давно [22, 23]. Рассмотрим в общем виде, каким должно быть это специфическое взаимодействие и какие эксперименты могут его проверить. Если генетический код образовался в результате взаимодействия аминокислот со стопками оснований в кодоне, то его триплетность для всех аминокислот указывает на то, что это свойство кода определяется общими для всех аминокислот группами. Единственными общими для всех аминокислот группами являются аминок- и карбоксильные группы. Следовательно, факт триплетности кода свидетельствует о том, что аминокислоты взаимодействовали с триплетами оснований именно этими

группами. Специфичность же кода должна определяться R-группами аминокислот, потому что только ими они и отличаются. Однако ни с какими химическими свойствами R-групп специфичность кода не коррелирует. Известное свойство генетического кода — "блочность" — объединяет аминокислоты с совершенно разными структурами и химическими свойствами R-групп, кроме того, кодоны одной аминокислоты могут попадать в разные блоки. Ясно, что это искусственное объединение кодонов аминокислот по принципу отличия лишь по одному основанию из трёх.

Приведённые выше аргументы указывают на то, что специфичность кода обусловлена не химическими, а физическими (ван-дер-ваальсовыми: диполь-дипольное, индукционное и дисперсионное) взаимодействиями. Существенной общей физической характеристикой аминокислот является то, что при нейтральных значениях pH это диполи. Величина дипольных моментов и частота осцилляций диполей зависят от R-групп аминокислот и, кроме того, эти параметры различны в конформерах одной и той же аминокислоты [24–26]. Можно предположить, что специфическое взаимодействие аминокислоты (в форме какого-то её конформера) с определённой последовательностью оснований в кодоне происходит лишь тогда, когда частота осцилляций аминокислотного диполя совпадает с частотой осцилляций индуцированного диполя в триплете нуклеотидов. Таким образом, для проверки гипотезы связывания аминокислот с их кодонами необходимо исследовать взаимодействие свободных L-аминокислот с триплетами канонических нуклеотидов, соответствующими их кодонам не только по составу, но и по последовательности нуклеотидов. Насколько мне известно, таких исследований проведено ещё не было.

Так, например, в работе [27] исследовали взаимодействие иммобилизованных на резине аминокислот с нуклеотидами и олигонуклеотидами. Было обнаружено связывание нуклеотидов и олигонуклеотидов с иммобилизованным глицином и триптофаном, возрастающее с увеличением числа нуклеотидов в олигонуклеотидах. Однако ни один из трёх исследованных триплетов нуклеотидов не совпадал ни с одним из кодонов этих двух аминокислот. Так что специфичность взаимодействия аминокислот с их кодонами в этой работе не изучалась. Иммобилизованные аминокислоты существенно отличаются по конформерам и дипольным характеристикам от свободных аминокислот, и эта модель, по-видимому, не соответствует взаимодействию свободных аминокислот с кодонами.

Другое направление исследований этой проблемы представлено в работах Yarus с соавторами [28–30]. В первой работе этого цикла было исследовано взаимодействие аргинина с GTP-связывающим участком рибосомальной РНК [28]. Было показано, что аргинин связывается с этим участком за счёт R-группы, а амино- и карбоксильные группы в связывании не участвуют. В последующих работах в РНК были обнаружены связывающие участки для ещё семи аминокислот [30]. Хотя авторы указывают, что связывающие аминокислоту участки обогащены кодонами и антикодонами для данной аминокислоты, в связывании участвуют и другие нуклеотиды и, следовательно, в связывании задействован огромный по сравнению с кодоном участок РНК. Кроме того, это связывание определяется не амино- и карбоксильными, а R-группами. Возможно, обнаруженное в работах Yarus с соавторами связывание аминокислот с РНК имеет какое-то регуляторное значение, но отношение этого типа связывания к происхождению генетического кода сомнительно.

Таким образом, в ряде работ показано, что аминокислоты взаимодействуют с нуклеотидами, олигонуклеотидами и РНК, однако взаимодействие аминокислот с их кодонами пока не исследовано.

Одним из основных факторов деструкции органических молекул в пребиотических условиях было ультрафиолетовое излучение, интенсивность которого превышала современный уровень на два порядка из-за отсутствия озонового экрана [31, 32]. Азотистые основания, рибоза и аминокислоты поглощают в ультрафиолетовом диапазоне. При поглощении кванта молекула переходит в возбуждённое состояние, которое может привести к разрыву связи (разрушение молекулы), или поглощённая энергия рассеется в тепло через колебания и вращение молекулы, обмен атомами водорода и другие механизмы и молекула останется неповреждённой. При образовании комплексов молекул: нуклеотиды, стэкинг нуклеотидов, комплементарные пары нуклеотидов — возрастает число взаимодействий в комплексах и, соответственно, возможности рассеяния энергии кванта в тепло. В результате устойчивость комплексов к ультрафиолету на несколько порядков выше, чем отдельных молекул [33]. Связывание аминокислот с кодонами могло снизить вероятность их разрушения (кодонов и аминокислот) и увеличивать время существования кодонов.

5. Некоторые следствия бимолекулярного сценария возникновения жизни

Рассмотрим одно из принципиальных следствий гипотезы. Согласно предлагаемой гипотезе, воспроизводство информации происходило не прямо, а через комплементарный кодон. Из этого следует, что для кодонов всех аминокислот, включённых в жизнь, должны существовать комплементарные кодоны. В таблице приведены пары аминокислот, имеющие комплементарные кодоны.

Приведённая таблица показывает полное соответствие кодонов с комплементарными кодонами для 21 аминокислоты: нет ни одного кодона без комплементарного ему кодона. Этот факт является серьёзным аргументом в пользу предлагаемой гипотезы, поскольку при случайном выборе аминокислотами кодонов такое совпадение невероятно. Кроме того, полное совпадение кодонов и комплементарных кодонов для 21 аминокислоты свидетельствует о том, что изначально были вовлечены в жизнь 21 аминокислота. Начальная кодировка сохранилась на протяжении всей эволюции жизни для 20 канонических (наиболее распространённых) аминокислот.

С появлением в процессе эволюции аппарата, синтезирующего полипептиды, возникла необходимость в остановке трансляции, и эта функция перешла к редко употребляемым кодонам: кодоны селеноцистеина — UGA и глицина — UAG (оставшийся только у зелёных водорослей) становятся стоп-кодонами. Селеноцистеин же перекодируется с помощью дополнительной последовательности нуклеотидов (selenocysteine insertion sequence), следующих за UGA. В таком сочетании аппарат синтеза белка интерпретирует кодон UGA как кодирующий селеноцистеин.

Одна аминокислота (22-я) — пирролизин [34] — была включена в жизнь, по-видимому, позднее при эволюции одной из её ветвей — метанобразующих архей. Учитывая сложность синтеза пирролизина, требующего нескольких ферментов [35], кажется сомнительной возможность синтеза этой аминокислоты в пребиотической среде. В этом случае стоп-кодон UAG был перекодирован в кодон для пирролизина также с помощью следующей за ним дополнительной последовательности нуклеотидов, и в этом сочетании UAG узнаётся аппаратом синтеза белка архей как кодон пирролизина. Таким образом, при включении в жизнь новых аминокислот или их перекодировке в условиях уже существующего синтеза белка триплета оснований не достаточно, для этого необходима дополнительная информация (последовательность нуклеотидов).

В настоящее время считается, что аминокислота кодируется преимущественно двумя первыми азотистыми

Таблица. Пары аминокислот, имеющих комплементарные кодоны

Аминокислота	Кодоны	Комплементарные кодоны	Аминокислота
Ala/A — аланин (неполярная) ¹	GCU, GCC, GCA, GCG ²	CGA, CGG, CGU, CGC, AGA, AGG	Arg/R — аргинин (основная)
Ser/S — серин (полярная)	UCU, UCC, UCA, UCG, AGU, AGC ³	CGA, CGG, CGU, CGC, AGA, AGG	Arg/R — аргинин (основная)
Asn/N — аспарагин (полярная)	AAU, AAC	UUA, UUG, CUU, CUC, CUA, CUG	Leu/L — лейцин (неполярная)
Asp/D — аспарагиновая кислота (кислотная)	GAU, GAC	UUA, UUG, CUU, CUC, CUA, CUG	Leu/L — лейцин (неполярная)
Glu/E — глутаминовая кислота (кислотная)	GAA, GAG	UUA, UUG, CUU, CUC, CUA, CUG	Leu/L — лейцин (неполярная)
Cys/C — цистеин (полярная)	UGU, UGC	ACU, ACC, ACA, ACG	Thr/T — треонин (полярная)
Sec/U — селеноцистеин (полярная)	UGA	ACU, ACC, ACA, ACG	Thr/T — треонин (полярная)
Trp/W — триптофан (неполярная)	UGG	ACU, ACC, ACA, ACG	Thr/T — треонин (полярная)
Gln/Q — глутамин (полярная)	CAA, CAG	GUU, GUC, GUA, GUG	Val/V — валин (неполярная)
His/H — гистидин (основная)	CAU, CAC	GUU, GUC, GUA, GUG	Val/V — валин (неполярная)
Gly/G — глицин (неполярная)	GGU, GGC, GGA, GGG, (UAA, UAG — зел. вод.)	CCA, CCG, CCU, CCC	Pro/P — пролин (неполярная)
Gly/G — глицин (неполярная)	GGU, GGC, GGA, GGG, (UAA, UAG — зел. вод.)	AUA, AUU, AUC	Ile/I — изолейцин (неполярная)
Ile/I — изолейцин (неполярная)	AUA, AUU, AUC	UAU, UAC	Tyr/Y — тирозин (полярная)
Met/M — метионин (неполярная)	AUG	UAU, UAC	Tyr/Y — тирозин (полярная)
Lys/K — лизин (основная)	AAA, AAG	UUU, UUC	Phe/F — фенилаланин (неполярная)

¹ R-группы.
² Жирным шрифтом отмечены комплементарные кодоны.
³ Серин стабилизирует не только комплементарные кодоны для аргинина, но также 4 попарно комплементарных кодона (выделены курсивом). Это означает, что серин — единственная аминокислота, способная воспроизводить свою информацию без посредников (других аминокислот).

основаниями, а третье не столь существенно, поэтому оно может заменяться мутациями, не влияющими на код. Этим объясняется вырожденность кода. Это правило, однако, справедливо только для кодонов четырёх аминокислот: аланина, треонина, валина и пролина. Для всех остальных кодонов оно не выполняется. Так, в кодонах аргинина и лейцина заменяется не только третье, но и первое основание, а третьи основания других аминокислот заменяются в меньшей степени, вплоть до полного отсутствия замен (кодоны метионина, селеноцистеина, триптофана), несмотря на то что эволюционный путь для всех кодонов был одинаков. Следовательно, причина разного числа кодонов на аминокислоту не в том, что третьи основания в кодонах могут заменяться, а первые и вторые — нет.

Положим, что все основания во всех кодонах могут заменяться с одинаковой частотой в результате ошибок ДНК и РНК полимераз, и рассмотрим наиболее простой случай — одна замена (две замены одновременно исключим, как очень маловероятные) в любом из трёх оснований кодона триптофана: UGG. В результате всех замен в трёх основаниях образуется следующий спектр кодонов: AGG (аргинин), CGG (аргинин), GGG (глицин), UAG (глицин), UCG (серин), UUG (лейцин), UGU (цистеин), UGA (селеноцистеин), UGC (цистеин). Как видно, все вновь образованные в результате замен кодоны заняты другими аминокислотами, т.е. результатом мутации является замена аминокислоты, а не образование нового (дополнительного для данной аминокислоты) кодона.

Возникает вопрос: когда кодоны были заняты аминокислотами и почему на одну аминокислоту приходится разное количество кодонов? Предположим, что вырож-

денность генетического кода (аминокислота может кодироваться не только одним, но и от 2 до 6 кодонов) появилась в самом начале (при возникновении жизни). Это свойство генетического кода может быть обусловлено двумя причинами: 1) разные аминокислоты могут стабилизировать кодоны, взаимодействуя не только с тремя основаниями кодона, но и с двумя или с тремя в несколько другом сочетании и 2) разные конформеры аминокислоты могут стабилизировать разные кодоны. Трём аминокислотам (метионин, триптофан, селеноцистеин) соответствует только один кодон (код не вырожден). Следовательно, для стабилизации триплетов оснований этими аминокислотами необходимо их взаимодействие со всеми тремя основаниями, расположенными в определённом порядке (или из всех конформеров данных аминокислот стабилизирует триплет оснований только один). Девять аминокислот кодируются двумя кодонами, в которых первые два основания одинаковы, а третье основание — одно из двух пар U–C или A–G (ни в одном из девяти кодонов это правило не нарушается). Очевидно, для стабилизации кодонов этими аминокислотами двух оснований не достаточно, необходимо третье основание, любое из двух пар (или по два конформера данных аминокислот способны стабилизировать триплеты оснований). Изолейцин — единственная аминокислота, которой принадлежат три кодона: первые два основания одинаковы для всех трёх, в двух кодонах третье основание — пара U–C, а в третьем A. Для этой аминокислоты третье основание может быть любым из трёх (или три конформера изолейцина стабилизируют триплеты оснований). И, наконец, если для стабилизации триплета аминокислоте достаточно взаимодействие только с двумя основаниями, тогда

третье основание может быть любым. При этом условия на каждую аминокислоту должно приходиться по четыре кодона: первые два основания фиксированы, а третье — любое из четырёх (четыре конформера данных аминокислот стабилизируют триплеты оснований). По четыре кодона имеют аланин, валин, треонин, пролин. Четыре аминокислоты кодируются шестью кодонами. Для лейцина и аргинина к четырём кодонам с одинаковыми первыми основаниями добавляются по два кодона с заменой первого основания в лейциновых кодонах U на C, в аргининовых C на A, при этом вторые и третьи основания остаются такими же, как в двух из четырёх кодонов. Другие две аминокислоты с шестью кодонами, серин и глицин, представляют особые случаи. Серин помимо стабилизации четырёх кодонов по двум первым основаниям (UC), два из которых комплементарны двум кодонам аргинина, стабилизирует также два кодона, комплементарные к своим же двум кодонам. Это единственная аминокислота, способная стабилизировать комплементарные кодоны, что даёт возможность репликации кодонов без посредников. Глицину, если исключить зелёные водоросли, принадлежит 4 кодона: два первых основания одинаковы, а третье — любое. Однако два кодона глицина в зелёных водорослях совсем другие: у зелёных водорослей первые два основания UA вместо обычных GG. Это единственный случай отличия кодонов одной аминокислоты сразу по двум первым основаниям. Причём нельзя списать это на результат мутаций после расхождения зелёных водорослей с другими ветвями эволюции жизни, поскольку для всех шести кодонов глицина существуют комплементарные кодоны, а это означает, что глицин стабилизировал все шесть кодонов изначально.

Если начальная репликация осуществлялась через образование комплементарных кодонов, то из этого следует:

1. Аргинин стабилизирует 6 кодонов и воспроизводит 6 комплементарных кодонов, из которых 4 стабилизируются аланином и 2 — серином. Кодоны этих аминокислот воспроизводят кодоны аргинина (Arg, Ala, Ser: основная, неполярная, полярная).

2. Лейцин стабилизирует 6 кодонов, воспроизводит 6 комплементарных кодонов, из которых 2 стабилизируются аспарагином, 2 — аспарагиновой кислотой и 2 — глутаминовой кислотой. Кодоны этих аминокислот воспроизводят кодоны лейцина (Leu, Asn, Asp, Glu: неполярная, полярная, кислотная, полярная).

3. Треонин стабилизирует 4 кодона и воспроизводит 4 комплементарных кодона, из которых 2 стабилизируются цистеином, 1 — селеноцистеином и 1 — триптофаном (Thr, Cys, Sec, Trp: полярная, полярная, полярная, неполярная).

4. Валин стабилизирует 4 кодона и воспроизводит 4 комплементарных кодона, из которых 2 стабилизируются глутамином и 2 — гистидином. Кодоны этих аминокислот воспроизводят кодоны валина (Val, Gln, His: неполярная, полярная, основная).

5. Глицин стабилизирует 6 кодонов (два из них остались только у зелёных водорослей) и воспроизводит 6 комплементарных кодонов, из которых 4 стабилизируются пролином и 2 стабилизируются изолейцином. В свою очередь изолейцин стабилизирует 3 кодона и воспроизводит ещё один комплементарный кодон, стабилизируемый тирозином. Комплементарный к тирозиновому кодон стабилизируется метионином. Пролиновые и изолейциновые кодоны воспроизводят кодоны глицина, тирозиновый — изолейциновый, метиониновый — тирозиновый (Gly, Pro, Ile, Tug, Met: неполярная, неполярная, неполярная, полярная, неполярная).

6. Лизин стабилизирует 2 кодона и воспроизводит 2 комплементарных кодона, которые стабилизируются

фенилаланином. Фенилаланиновые кодоны воспроизводят лизиновые (Lys, Phe: основная, неполярная).

7. Серин стабилизирует 4 попарно комплементарных кодона и поэтому самовоспроизводится. (Ser: полярная).

Таким образом, при наличии азотистых оснований и аминокислот жизнь возникла спонтанно как семь самовоспроизводящихся информационных систем (СИС). Матрица (триплет азотистых оснований) содержала информацию о структуре оператора (аминокислоте), воспроизведшего эту информацию. В шести из семи систем информация воспроизводилась оператором опосредованно через комплементарные кодоны и их аминокислоты. С появлением этих СИС начала работать дарвинская эволюция: конкуренция СИС за азотистые основания и аминокислоты, отбор на стабильность информации (выживание) и скорость её воспроизводства (размножение). Поскольку воспроизводство информации через посредников гораздо менее вероятно, чем без них, то эволюция первичных СИС шла в направлении объединения комплементарных кодонов в единый комплекс. Другим существенным фактором, ведущим к объединению кодонов, было ультрафиолетовое излучение, поскольку комплексы менее чувствительны к его разрушающему действию [30]. Это привело к появлению РНК, содержащих от 2 до 5 кодонов (или от 6 до 15 нуклеотидов) и соответствующих им новых операторов, стабилизирующих комплексы кодонов, — пептидов. На этой стадии самовоспроизводство серина утратило эволюционное значение, поскольку он входил также в состав трипептида (система 1), а будущее было за комплексами кодонов и пептидами. Можно предположить два возможных варианта образования пептидов:

1. Пептидные связи образовывались во время взаимодействия аминокислот с комплексом кодонов, поскольку при этом аминокислоты фиксировались и сближались, причём так, что аминогруппа одной аминокислоты сближалась с карбоксильной — другой (комплексы с обратной ориентацией кодонов, при взаимодействии с которыми аминокислоты сближались амино- или карбоксильными группами, не воспроизводились). Так была задана ориентация кодонов в матрице, сохранившаяся до сих пор. Расстояние, на которое сближаются азотистые основания при стэкинге (0,45 нм) [21] существенно больше длины пептидной связи (0,132 нм), но, возможно, кулоновское взаимодействие противоположно заряженных аминокислотных групп сближало соседние аминокислоты на дистанцию, достаточную для образования пептидной связи.

2. Часть РНК, образовавшихся в результате объединения кодонов, были способны катализировать образование пептидной связи. Как показано в работах [36, 37], короткоцепочечные РНК с такой каталитической активностью действительно существуют.

В результате эволюции СИС к воспроизводству без посредников возникли молекулы РНК и пептиды. Оба типа молекул обладают как структурными, так и каталитическими функциями. Однако возникновение эффективного метаболического ансамбля невозможно без его изоляции, поскольку в неограниченной водной среде как сами каталитические молекулы, так и их продукты разносятся конвекционными потоками и течениями. Поэтому следующим эволюционным этапом жизни должно было быть её отделение от окружающей среды. Основными источниками энергии для абиогенного синтеза органических соединений были атмосферное электричество и ультрафиолетовое излучение [38]. Следовательно, эти реакции происходили либо в атмосфере, с последующим выпадением продуктов на поверхность водоёмов с осадками, либо вблизи от поверхности водоёмов, поскольку ультрафиолет глубоко не проникает. Известно, что жирные кислоты, которые также синтезируются в абиогенных

реакциях, образуют плёнки на поверхности воды. Часть РНК и пептидов могла попадать в возникающие при волнении или выпадении осадков пузырьки, ограниченные такой плёнкой. Это замкнутое пространство ещё не являлось оператором, поскольку информации о плёнке не было в матрице, это пока была только экологическая ниша. Однако уже на этом этапе плёнка из жирных кислот могла модифицироваться пептидами. Обратим внимание на то, что все системы 1–6 генерировали амфифильные пептиды. Амфифильные соединения способны встраиваться в плёнку из жирных кислот, возможно, стабилизируя её и создавая транспортные каналы для аминокислот и нуклеотидов.

6. Заключение

Представленная в работе гипотеза является первым вариантом бимолекулярного сценария возникновения жизни, который, по мнению авторов работы [11], не отделён от существующей жизни логической пропастью. По этому сценарию короткоцепочечные РНК и пептиды возникли одновременно. Последующее объединение первичных самовоспроизводящихся систем создаст огромный информационный потенциал для эволюции жизни к всё более сложным операторам: рибосомам, клеткам, растениям, животным и человеку. Предваряя критику гипотезы, отметим, что наименее обоснованным положением гипотезы является предположение о том, что L-аминокислоты специфически связываются со своими кодонами (триплетами только канонических нуклеотидов: канонические основания с D-рибозой) и это взаимодействие увеличивает время существования кодона. К сожалению, этот вопрос не изучен ни теоретически, ни экспериментально. Возможно, настоящая гипотеза послужит стимулом для исследований в этом направлении.

Список литературы

- Иваницкий Г Р *УФН* **180** 337 (2010); Ivanitskii G R *Phys. Usp.* **53** 327 (2010)
- Кляцкин В И *УФН* **182** 1235 (2012); Klyatskin V I *Phys. Usp.* **55** 1152 (2012)
- Иваницкий Г Р *УФН* **182** 1238 (2012); Ivanitskii G R *Phys. Usp.* **55** 1155 (2012)
- Jortner J *Phil. Trans. R. Soc. Lond. B* **361** 1877 (2006)
- Eigen M *Naturwissenschaften* **58** 465 (1971)
- Eigen M *Steps Towards Life. A Perspective on Evolution* (Oxford: Oxford Univ. Press, 1996)
- Orgel L E *Trends Biochem. Sci.* **23** 491 (1998)
- Lazcano A, Miller S L *Cell* **85** 793 (1996)
- Иваницкий Г Р *Выражи закономерностей: правило БИО — стержень науки* (М.: Наука, 2011)
- Gilbert W *Nature* **319** 618 (1986)
- Cech T R *Gene* **135** 33 (1993)
- Joyce G F *Nature* **418** 214 (2002)
- Joyce G F, Orgel L E *The RNA World. The Nature of Modern RNA Suggests a Prebiotic RNA* (Eds R F Gesteland, T R Cech, J F Atkins) 2nd ed. (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, 1999) p. 49
- Walker S I, Davies P C W *J. R. Soc. Interface* **10** 20120869 (2013)
- Saladino R et al. *Biochimie* **94** 1451 (2012)
- Ponnampuruma C, Sagan C, Mariner R *Nature* **199** 222 (1963)
- Powner M W, Gerland B, Sutherland J D *Nature* **459** 239 (2009)
- Hein J E, Tse E, Blackmond D G *Nature Chem.* **3** 704 (2011)
- Hein J E, Blackmond D G *Acc. Chem. Res.* **45** 2045 (2012)
- Krishnamurthy R *Acc. Chem. Res.* **45** 2035 (2012)
- Norberg J, Nilsson L *J. Am. Chem. Soc.* **117** 10832 (1995)
- Woese C R *Proc. Natl. Acad. Sci. USA* **54** 1546 (1965)
- Orgel L E *J. Mol. Biol.* **38** 381 (1968)
- Gronert S, O'Hair R A J *J. Am. Chem. Soc.* **117** 2071 (1995)
- Cao X, Fischer G *J. Phys. Chem. A* **103** 9995 (1999)
- Kaczor A et al. *J. Phys. Chem. A* **110** 2360 (2006)
- Saxinger C, Ponnampuruma C *Origins Life* **5** 189 (1974)
- Yarus M *Science* **240** 1751 (1988)
- Yarus M, Caporaso J G, Knight R *Annu. Rev. Biochem.* **74** 179 (2005)
- Yarus M, Widmann J J, Knight R *J. Mol. Evol.* **69** 406 (2009)
- Garcia-Pichel F *Orig. Life Evol. Biosphere* **28** 321 (1998)
- Cockell C S, Horneck G *Photochem. Photobiol.* **73** 447 (2001)
- Mulkidjanian A Y, Galperin M Y *Chem. Biodiversity* **4** 2003 (2007)
- Atkins J F, Gesteland R *Science* **296** 1409 (2002)
- Krzycki J *Curr. Opin. Microbiol.* **8** 706 (2005)
- Turk R M, Chumachenko N V, Yarus M *Proc. Natl. Acad. Sci. USA* **107** 4585 (2010)
- Yarus M *Phil. Trans. R. Soc. B* **366** 2902 (2011)
- Chyba C, Sagan C *Nature* **355** 125 (1992)

Life originated as a genetic code (contribution to the discussion of the paper by G.R. Ivanitskii

"21st century: what is life from the perspective of physics" [*Usp. Fiz. Nauk.* **180** 337 (2010); *Phys. Usp.* **53** 327 (2010); *Usp. Fiz. Nauk.* **182** 1235 (2012); *Phys. Usp.* **55** 1152 (2012); *Usp. Fiz. Nauk.* **182** 1238 (2012); *Phys. Usp.* **55** 1155 (2012)])

Yu.N. Korystov

*Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences,
ul. Institutskaya 3, 142290 Pushchino, Moscow region, Russian Federation
E-mail: ykorystov@rambler.ru*

This letter is aimed at concretizing the origin-of-life scenario. The letter defines life as a self-reproducing information system and briefly discusses the current status of the origin-of-life problem. The first bimolecular origin-of-life scenario is proposed which is based on the interaction of amino acids with triplets of nucleotides corresponding to their codons. In this scenario, the reproduction of information occurs by hydrogen bonding with complementary nucleotides and hence is due to the formation of complementary codons that reproduce the primary codons. The bimolecular scenario fundamentally necessitates the existence of complementary codons for all amino acids associated with life. For 21 amino acids, a complete correspondence is shown to exist between codons and complementary codons, a fact which greatly favors the proposed hypothesis because this correspondence is impossible if amino acids choose codons at random.

Keywords: amino acids, life origin, genetic code

PACS numbers: **87.14.-g**, **87.18.-h**, 87.23.kg

Bibliography — 38 references

Uspekhi Fizicheskikh Nauk **187** (2) 235–240 (2017)

DOI: <https://doi.org/10.3367/UFNr.2016.11.037996>

Received 23 August 2016, revised 17 November 2016

Physics–Uspekhi **60** (2) (2017)

DOI: <https://doi.org/10.3367/UFNe.2016.11.037996>