

ОБЗОРЫ

УДК 577.212.8

**ФАКТОРЫ ОТБОРА ПРОТОБИОМОНОМЕРОВ
ПРИ ВОЗНИКНОВЕНИИ ГЕНЕТИЧЕСКОГО КОДА**

Саралов А.И.

*Институт экологии и генетики микроорганизмов Уральского отделения
Российской академии наук, Пермь, e-mail: saralov@iegm.ru*

Для объяснения происхождения универсального генетического кода выдвинут ряд взаимодополняющих гипотез от «замерзшей случайности» до представлений о коэволюции кодонов с путями синтеза аминокислот. Слабее изучены факторы отбора необходимых компонентов для создания прообраза генетического кода от абиогенеза до начальных биогенных стадий. Первичный отбор минимального числа биомономеров (достаточных для поэтапной самоорганизации генетического кода) из хаотичного многообразия органических соединений мог происходить в зависимости от их молекулярной массы (M), температуры плавления (t_m), растворимости в воде, структуры и элементного состава – основы любого соединения. На диаграммах в координатах M и t_m протеиногенные аминокислоты вместе с кодирующими их кодонами располагаются согласованно относительно четко обособленных семейств пуриновых и пиримидиновых оснований. Подтверждены коэволюционные связи кодонов с путями синтеза аминокислот и с хронологией их появления в коде. Эволюционный отбор термостойких протобиомономеров сопровождался замещением атомов кислорода атомами азота от простых питательных веществ к аминокислотам и азотистым основаниям. Показано, что элементному составу M и t_m при первичном дифференцированном отборе протобиомономеров и их пространственных изомеров принадлежала исключительно важная роль. Причем в данной работе впервые рассмотрена t_m протобиомономеров в качестве детерминирующего фактора отбора специфических компонентов при возникновении генетического кода.

Ключевые слова: генетический код, происхождение, эволюция, новая гипотеза

**SELECTION FACTORS FOR THE PROTOBIOMONOMERS
AT THE ORIGIN OF GENETIC CODE**

Saralov A.I.

*Institute of Ecology and Genetics of Microorganisms, Ural Branch of Russian Academy of Sciences,
Perm, e-mail: saralov@iegm.ru*

To explain the origin of the universal genetic code, a number of complementary hypotheses are put forward from a «frozen accident» to ideas about the co-evolution of codons with amino acid synthesis pathways. Less studied the factors of selection of the necessary components to create a prototype of the genetic code from abiogenesis to the initial biogenic stages. The initial selection of the minimum number of biomonomers (sufficient for phased self-organization of the genetic code) from the chaotic diversity of organic compounds could occur depending on their molecular weight (M), melting point (t_m), water solubility, structure and elemental composition – the basis of any compound. On the charts in the coordinates M and t_m , the proteinogenic amino acids, together with the coding codons, are arranged consistently with respect to clearly distinct families of purine and pyrimidine bases. The co-evolutionary links of codons with the pathways of the synthesis of amino acids and the chronology of their appearance in the code were confirmed. The evolutionary selection of heat-resistant protobiomonomers was accompanied by the replacement of oxygen atoms by nitrogen from simple nutrients to amino acids and nitrogenous bases. It was shown that the elemental composition, M and t_m in the primary differentiated selection of protobiomonomers and their spatial isomers belonged to an extremely important role. Moreover, in this work, for the first time, t_m protobiomonomers are considered as the determining factor for the selection of specific components when a genetic code occurs.

Keywords: genetic code, origin, evolution, novel hypothesis

Генетический код определяет соответствие последовательности в молекулах РНК и ДНК с последовательностью аминокислот в молекулах белка [1]. «Каноническую» версию генетического кода обычно представляют в виде таблицы из 64 триплетов для 20 протеиногенных аминокислот и трех стоп-кодонов остановки трансляции, допускающая ее «универсальность» почти для всех живущих прокариот и эукариот.

Впоследствии, однако, было установлено, что иногда через стоп-кодоны могут синтезироваться дополнительные неканонические аминокислоты: селеноцистеин

(SeC) из серина через UGA [2] и пирролизин (Pyl) из лизина и орнитина через UAG [3]. Обнаруженная способность ряда кодонов выполнять двойственную функцию позволила предположить существование альтернативных версий генетического кода, в частности, в митохондриях, хлоропластах и бактериальных симбионтах [4].

При изучении проблемы происхождения генетического кода был выдвинут ряд взаимодополняющих гипотез от «замерзшей случайности» до теории коэволюции кодонов и путей синтеза аминокислот при создании триплетного кода по мере появления

новых аминокислот в примордиальных биохимических системах [5–7]. Причем первое основание кодона они неизменно связывали с синтетическими путями аминокислот при поэтапной самоорганизации генетического кода. Затем при обсуждении селекции биологически кодируемых аминокислот отмечена важность особенностей пребиотического синтеза, энергетических биосинтетических издержек при анаэробии, заряда, полярности, растворимости в воде, формы функциональных групп, хиральности, связи со стабильностью протеиновых структур и РНК-миром [8, 9]. Аминоацилирование тРНК и гомохиральность «одноруких» сахаров и аминокислот, вероятно, произошли при переходе от химической эволюции к рибозимам и РНК-миру [10, 11]. Несмотря на большое число работ по проблеме происхождения генетического кода, можно констатировать, что практически не изучены физико-химические закономерности отбора протобиомономеров из множества первозданных органических соединений с разнообразными структурами и свойствами.

Цель работы: определение детерминирующих факторов отбора компонентов нуклеиновых кислот, протеиногенных аминокислот и их предшественников при поэтапном формировании генетического кода от абиогенеза до начальных биогенных стадий.

Абиогенный синтез протобиомолекул

В примитивной атмосфере на первобытной Земле, возможно, содержались малые молекулы, преимущественно стабильные CO_2 , N_2 и H_2O , а также CH_4 и NH_3 , состоящие из атомов водорода, углерода, кислорода и азота. В зависимости от видов потоков энергии (электрические разряды, ультрафиолетовое излучение, бета-частицы или теплота), проходящих через одну из таких газобразных смесей, в условиях лабораторных экспериментов удалось получить альдегиды, сахара, карбоновые кислоты, аминокислоты, пуриновые и пиримидиновые азотистые основания [12]. В хронологии появления аминокислот в генетическом коде установлено совпадение десяти первых протеиногенных аминокислот с набором выявленных аминокислот в метеоритах и синтезированных в имитирующих экспериментах Миллера [12, 13]. Все эти мономеры – основные строительные блоки живых систем и бесчисленное множество иных соединений могли образоваться из малых молекул в качестве вероятных продуктов «химического реактора» при абиогенных геофизических процессах на древней Земле и других планетах [14].

В 2000 г. вне Срединно-Атлантического хребта было обнаружено обширное ги-

дротермальное поле, названное «Lost City» (Потерянный город) [15]. В отличие от вулканогенных «черных курильщиков», эти гидротермальные флюиды менее горячие ($150\text{--}200^\circ\text{C}$), сильнощелочные (pH 9–10) и богаты водородом. Их подводные постройки формируются в процессе серпентинизации при взаимодействии морской воды с железом-магниевыми силикатными минералами типа оливина. При гидроксилации оливина высвобождаются серпентин, магнетит и H_2 . Именно в подобных щелочных гидротермальных вентях с температурой до 200°C могла зародиться жизнь.

В древних гидротермальных системах с CO , CO_2 , N_2 , H_2 и H_2O могли синтезироваться обычные пребиотические молекулы формальдегида (HCHO) и циановодорода (HCN), затем молекулы моносахаров из HCHO , пуринов из HCN , а из смеси HCN и HCHO – цитозин, урацил и тимин [16]. В условиях щелочных гидротерм типа «Lost City» из неорганических прекурсоров (CO_2 , H_2 , NH_3 и H_2S), предполагается, осуществлялся и абиогенный синтез аминокислот [17]. Показана возможность пребиотического синтеза Pro и Arg в сети цианосульфидных реакций, катализируемых Cu (I) [18].

Возможности биогенных элементов определять структуру и физико-химические свойства протобиомономеров

Около 99% массы всего живого построено в основном из четырех легких атомов (H, C, O, N), способных образовывать прочные ковалентные связи и активно реагировать друг с другом [19]. Структуры с гидрофобными атомами и радикалами (S, S-S, Fe-S, Se, I, CH_3) более упорядочены и стабильны, но менее растворимы в воде.

Разнообразие органических веществ резко возрастает, если в молекулы кроме углерода и водорода дополнительно входят атомы кислорода: углеводы, карбоновые кислоты, спирты, альдегиды, кетоны, эфиры, жиры и соединения со смешанными функциями (таблица). Монокарбоновая пировиноградная кислота имеет низкую $t_{\text{пл}}$ и низкое отношение $t_{\text{пл}}, ^\circ\text{C} : M$, но неограниченную растворимость в воде. При переходе к дикарбоновым кислотам резко повышаются $t_{\text{пл}}$ и отношение $t_{\text{пл}} : M$, но утрачивается неограниченная растворимость в воде. Следует отметить, что могут существенно различаться физико-химические свойства ($t_{\text{пл}}$ и растворимость в H_2O) у пространственных изомеров органических кислот, углеводов, аминокислот и азотистых оснований. Например, у *цис-транс*-изомеров дикарбоновых малеиновой и фумаровой

кислот повышение $t_{пл}$ в 2,2 раза у фумарового *транс*-изомера сопровождается падением растворимости в воде в сто с лишним раз. Существенно могут различаться $t_{пл}$ и растворимость в воде у структурных изомеров α , β , γ -аминокислот в зависимости от взаимного расположения amino- и гидроксильных групп. Напротив, у структурных изомеров α -аминокислот, различающихся порядком расположения углеводородных радикалов $-CH_2$ и $-CH_3$, отличия растворимости в воде и $t_{пл}$ выражены значительно слабее. В свете этих данных представляется вполне оправданным предположение, что норлейцин и норвалин могли быть довольно обычными компонентами протеинов на ранних стадиях клеточной эволюции [20].

Полимерные цепи нуклеиновых кислот построены из четырех гетероциклических азотистых оснований А, С, G, U или Т [1]. Четыре основания С, G, U, Т состоят из че-

тырех атомов Н, С, N и О, но атом кислорода отсутствует в бициклическом аденине $C_5H_5N_5$. Они характеризуются термостабильностью, близкими значениями отношения $t_{пл}$ к М, низкой растворимостью в воде. Гидрофобность азотистых оснований тесно связана с явлением резонанса, когда каждая молекула плоского гетероцикла описывается наложением (гибридизацией) сразу нескольких валентных структур. Наибольшей энергией резонанса среди оснований обладает сравнительно устойчивая молекула аденина. Столь высокой стабильности, как у аденина, нет у его структурного изомера 2-аминопурина, обладающего способностью к интенсивной флуоресценции. Эволюционный отбор на фотостабильность, обусловленную коротким временем жизни их электронного возбуждения светом, вероятно, прошли при абиогенезе и другие природные азотистые основания нуклеиновых кислот [21].

Физико-химические свойства главных и минорных протобиомономеров

Соединение и формула	М	$t_{пл}$, °C	$\frac{t_{пл}}{M}$	Растворимость в H_2O (25 °C)
Пировиноградная кислота $C_3H_4O_3$	88,1	13,6	0,15	∞
Малеиновая кислота $C_4H_4O_4$	116,1	140	1,20	79
Фумаровая кислота $C_4H_4O_4$	116,1	302	2,60	0,7
β -D-Фруктоза $C_6H_{12}O_6$	180,2	105	0,58	375
β -D-Галактоза $C_6H_{12}O_6$	180,2	167	0,93	68
β -Аланин $H_2N(CH_2)_2COOH$	89,1	207	2,32	54,5
α -Аланин $CH_3CH(NH_2)COOH$	89,1	316	3,55	16,7
L-Норлейцин $CH_3(CH_2)_3CH(NH_2)COOH$	131,2	301	2,29	1,5
L-Лейцин $\begin{matrix} CH_3 \\ \diagdown \\ CH \\ \diagup \\ CH_3 \end{matrix} CHCH_2CH(NH_2)COOH$	131,2	295	2,25	2,4
L-Изолейцин $CH_3CH_2CH \begin{matrix} \diagup \\ CH_3 \\ \diagdown \\ CH(NH_2)COOH \end{matrix}$	131,2	285	2,17	4,1
Цитозин $C_4H_5N_3O$	111,1	224	2,02	0,77
Урацил $C_4H_4N_2O_2$	112,1	283	2,52	0,36
5-Метилцитозин $C_5H_7N_3O$	125,1	221	1,77	0,45
Тимин (метилурацил) $C_5H_6N_2O_2$	126,1	282	2,24	0,40
2-Аминопурин $C_5H_5N_5$	135,1	282	2,09	0,12
Аденин $C_5H_5N_5$	135,1	305	2,28	0,09
N6-Метиладенин $C_6H_7N_5$	149,2	300	2,02	1,18
Гипоксантин $C_5H_4N_4O$	136,1	310	2,28	0,09
Гуанин $C_5H_5N_5O$	151,1	346	2,29	0,004
7-Метилгуанин $C_6H_7N_5O$	165,2	370	2,24	$\leq 0,004$

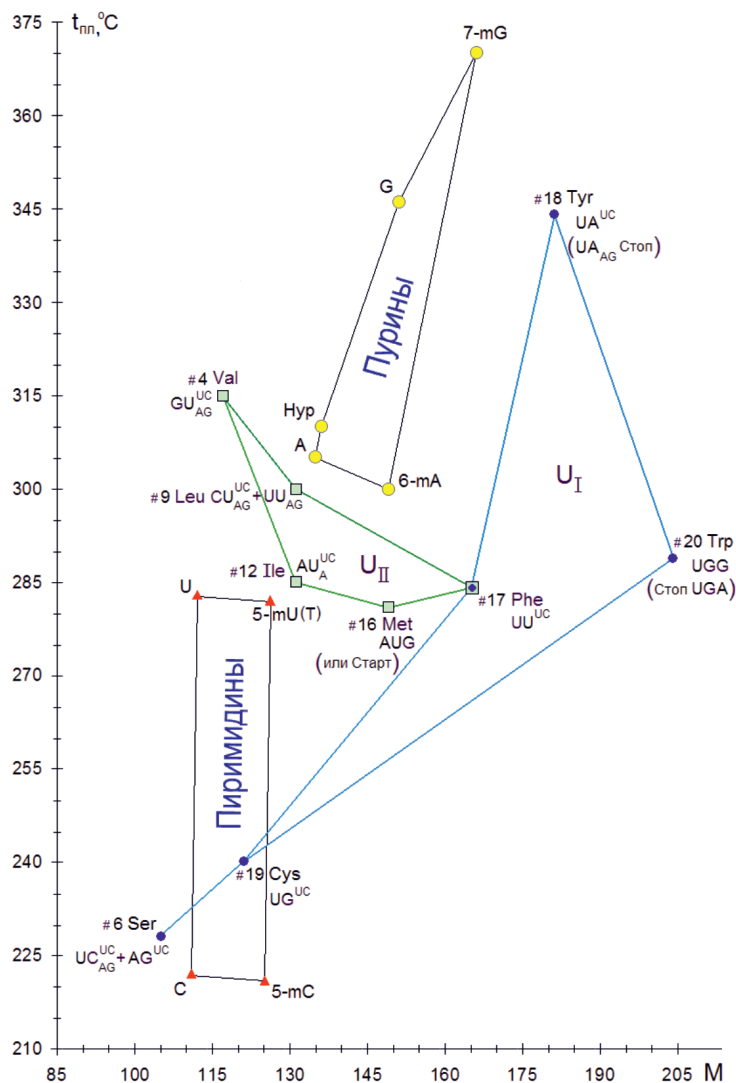


Рис. 1. Диаграмма, отражающая соотношение M и $t_{пл}$ у пуриновых и пиримидиновых оснований, протеиногенных аминокислот, кодируемых кодонами, содержащими урацил в первой (U) или второй (U_{II}) позициях. Сокращенное обозначение оснований в кодонах: Y – пиримидины, R – пурины, N – любые, $NN^Y = NN^{UC}$; $NN_R = NN_{AG}$; $NN_R^Y = NN_{AG}^{UC}$. # – номер появления аминокислот в формируемом коде

Попытки модифицировать главные и минорные основания, например, путем замещения в гетероцикле одного из атомов углерода на атомы азота или дополнительное введение в боковые группы атомов фтора или серы сопровождаются резкими негативными изменениями $t_{пл}$ и особенно растворимостью в воде (таблица).

Влияние молекулярной массы и температуры плавления на взаимосвязи протеиногенных аминокислот и кодонов

Молекулярная масса протеиногенных аминокислот изменяется в более широком

интервале (от 75,1 до 204,2), чем у главных азотистых оснований (от 111,1 до 151,1). Однако температуры плавления, при которых их кристаллы и жидкости находятся в равновесии, одинаково высоки и имеют близкие диапазоны изменчивости: у аминокислот от 186 до 344 °C, у оснований от 222 до 346 °C. На диаграмме в координатах M и $t_{пл}$ главные и минорные азотистые основания группируются в два четко обособленных семейства: пуриновые и пиримидиновые (рис. 1). Впервые установлено, что в системе прямоугольных координат $t_{пл}$ и M (рис. 1, 2), при постоянных значениях

остальных параметров, канонические аминокислоты вместе с кодирующими их кодонами располагаются довольно закономерно относительно обособленных семейств пуриновых и пиримидиновых оснований. При незначительном повышении M у главных C_4 -пиримидиновых и C_5 -пуриновых оснований наблюдается существенный рост их t_{nn} . Однако увеличение M за счет метилирования главных оснований, которые являются важными минорными основаниями, обычно не сопровождается повышением t_{nn} . Причем на диаграммах в координатах M и t_{nn} канонические аминокислоты вместе с кодирующими их кодонами располагаются довольно согласованно относительно семейств пуриновых и пиримидиновых оснований.

Пять аминокислот, кодируемых кодонами с урацилом во второй позиции (U_{II}), располагаются в узком диапазоне M и t_{nn} (между M и t_{nn} семейств пуриновых и пиримидиновых оснований) (рис. 1). Здесь кодоны с U_{II} различаются по степени вырожденности: 1-2-3-4-6. При этом каждая аминокислота кодируется разным количеством триплетов в зависимости от хронологии появления в генетическом коде: #4-9-12-16-17 [13]. Причем однократный кодон AUG может

кодировать серусодержащий метионин, но он же является стартовым кодоном, после которого начинается синтез полипептида на рибосоме.

Двукратно вырожденные кодоны с урацилом в первой позиции и пиримидинами UC в третьей позиции (UN^{UC}) кодируют четыре аминокислоты, у которых при сравнительно небольшом повышении M наблюдается резкое повышение t_{nn} (от низших значений у семейств пиримидинов до максимальных у семейств пуринов). В пределе после двукратного кодона UA^{UC} для наиболее термостойкого Tyr следует двукратный стоп-кодон UA_{AG} с пуринами в третьей позиции. Также после однократного кодона UGG с урацилом в первой позиции и с гуанином в третьей позиции для наиболее массивного Trp следует однократный стоп-кодон UGA . В итоге стоп-кодона, прерывающие биосинтез белка при трансляции на рибосоме в первой позиции, содержат пиримидин урацил, а в третьей позиции лишь бициклические пуриновые основания (UA_{AG} и UGA). Кроме того, урацил в первой и второй позициях послужил для разграничения стоп-кодонов (U) и старт-кодонов (U_{II}).

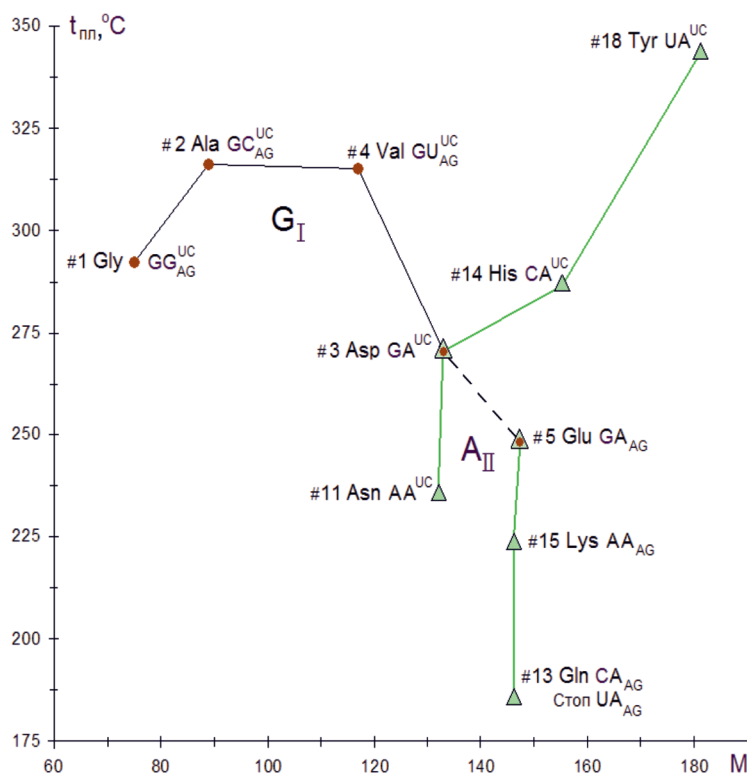


Рис. 2. Влияние M и t_{nn} на последовательность расположения протеиногенных аминокислот и соответствующих им кодонов, содержащих гуанин в первой (G_I) или аденин во второй (A_{II}) позициях

На следующей диаграмме (рис. 2) в системе прямоугольных координат $t_{пл}$ и M представлены канонические аминокислоты, кодируемые триплетами оснований с гуанином в первой позиции (G_1) и аденином во второй позиции (A_2). Там группа кодонов с G_1 кодирует пять наиболее древних C_2 – C_5 аминокислот (#1–5). Первые три наиболее термостойкие неполярные углеводородные аминокислоты (#1Gly, #2Ala, #4Val с $t_{пл}, ^\circ C: M$ 3,89–3,59–2,69) кодируются синонимическими квартетами кодонов. Менее термостойкие дикарбоновые полярные аминокислоты (#3Asp и #5Glu с $t_{пл}, ^\circ C: M$ 2.04 и 1.69) кодируются двукратно вырожденными кодонами (GA^{UC} и GA_{AG}).

Семь аминокислот, кодируемых двукратными триплетами с A_2 , характеризуются наиболее широким диапазоном $t_{пл}$ от $186^\circ C$ у #13Gln до $344^\circ C$ у #18Tyr с $t_{пл}, ^\circ C: M$ от 1,27 до 1,90. Эти аминокислоты в зависимости от содержания в третьей позиции кодонов пиримидинов (NA^{UC}) или пуринов (NA_{AG}) принадлежат двум разным ветвям. Одна ветвь из сравнительно более термостойких аминокислот с $t_{пл}$ 236–344 $^\circ C$ кодируется дублетами кодонов с пиримидинами в третьей позиции NA^{UC} , где N – A, G, C, U отражает закономерную последовательность оснований в первой позиции. Вторая ветвь для двух стоп-кодонов остановки трансляции на рибосоме и трех менее термостойких аминокислот с разной $t_{пл}$ 249–186 $^\circ C$, но с одинаковой M 146.2–147.1, кодируются кодонами с пуринами в третьей позиции NA_{AG} , где N – G, A, C, U последовательность оснований в первой позиции, завершаемая стоп-кодонами UA_{AG} . Итак, есть основание полагать, что двукратно и однократно вырожденные кодоны с U_1 фактически завершали формирование генетического кода на ранних стадиях протоклеточной эволюции, включая кодирование #17–22 аминокислот и терминирующих кодонов с пуринами в третьей позиции (UN_{AG} и UGA).

Связь кодонов с путями синтеза аминокислот

В свете теории коэволюции кодонов и путей синтеза протеиногенных аминокислот [5–7] модифицирована схема, отображающая связи последовательных метаболических превращений простых питательных веществ и аминокислот с кодирующими их кодонами. Отмечено два блока синтеза аминокислот. Первый блок связан с фосфорилированием и путями распада глюкозы, которые ведут к C_3 -соединениям, преимущественно пировиноградной кислоте. Второй блок синтеза аминокислот связан с циклом трикарбоно-

вых кислот с ветвями от щавелевоуксусной и α -кетоглутаровой кислот.

Исключительно важен тот факт, что первое основание кодона неизменно связано с путями синтеза аминокислот и с хронологией их появления в коде. Наиболее древние пять аминокислот кодируются кодонами с гуанином в первом положении (G_1). На очередном эволюционном этапе ряд аминокислот образуется в результате превращения других аминокислот, кодируемых кодонами с цитозином в первой позиции. Процесс синтеза аминокислот в первом блоке завершается кодированием кодонами с урацилом в первой позиции (U_1), а во втором блоке – с аденином (A_1). Учитывая закономерные последовательности синтеза аминокислот в схеме, по-видимому, следует внести коррективы в хронологию их появления в коде по Трифонову [13]. Будет правильнее переставить номера у пар #7Glu/#5Pro и #19Met/#16Cys на #5Glu/#7Pro и #16Met/#19Cys.

Отмечено существенное повышение $t_{пл}$ при прямом аминировании кетокислот при замене атома кислорода на NH_3 . При этом M синтезируемых углеводородных аминокислот увеличивается лишь на 1. Так, превращение пирувата в Ala сопровождается скачкообразным увеличением $t_{пл}$ с 14 до $316^\circ C$, «кетовалина» в Val – с 32 до $315^\circ C$, «кетолейцина» в Leu – с 10 до $295^\circ C$. Для синтеза #2Ala, #4Val и #9Leu характерны низкие энергетические затраты (1–2 АТФ) [9]. Тогда как для биосинтеза #12Ile (изомера Leu) требуется 11 АТФ, для #16Met – 23 АТФ. Они принадлежат разным ветвям биосинтеза и, естественно, отличаются составом кодонов с урацилом или аденином в первой позиции (U_1 и A_1).

Заключение

В проблеме возникновения генетического кода слабо изучены факторы дифференцированного отбора протобиомономеров при абиогенных и начальных биогенных стадиях. Первичный отбор минимального числа биомономеров мог происходить в зависимости от их молекулярной массы, температуры плавления, растворимости в воде, структуры и элементного состава. Отмечено, что очень существенно могут различаться физико-химические свойства изомеров органических кислот, углеводов и азотистых оснований (особенно $t_{пл}$ и растворимость в воде). Эволюционный отбор термостойких протобиомономеров в протоклетках мог происходить путем поэтапного замещения атомов кислорода атомами азота, в частности, от важнейших промежуточных продуктов обмена веществ к аминокислотам.

Впервые показано, что на диаграммах в системе прямоугольных координат t_{III} и M протеиногенные аминокислоты вместе с кодирующими их кодонами располагаются согласованно с четко обособленными семействами пуриновых и пиримидиновых оснований. Двукратно и однократно вырожденные кодоны, содержащие урацил в первой позиции (U_I), фактически завершили формирование генетического кода на ранних стадиях протоклеточной эволюции, включая кодирование #17-22 аминокислот и терминирующих кодонов с пуринами в третьей позиции.

Подтверждена коэволюционная связь кодонов разной степени вырожденности с путями синтеза аминокислот и с хронологией их появления в формируемом генетическом коде. Часть однократно вырожденных триплетов, содержащих урацил в первой (U_I) или второй (U_{II}) позициях и лишь один из пуринов в третьей позиции, фактически завершили формирование генетического кода, включая кодирование #16 Met (AUG), #20 Trp (UGG), #21 Sec (UGA) и #22 Pyl (UAG), ряда стоп- (U_{AG} , UGA) и старт-кодонов (обычно AUG, реже GUG или UUG). Тем самым из состава однократных триплетов был полностью исключен сравнительно легкоплавкий и наиболее растворимый в воде моноциклический цитозин.

Комплекс физико-химических процессов позволил отобрать из хаотического многообразия органических соединений «первичного бульона» минимальное количество протобиономеров с определенными значениями M и t_{III} и отношением $t_{III}:M$ (5 главных азотистых оснований, 20 протеиногенных аминокислот, несколько моносахаридов и жирных кислот). В данной работе впервые указано на t_{III} в числе детерминирующих факторов отбора специфических протобиономеров при возникновении и поэтапной самоорганизации генетического кода.

Работа выполнена в рамках государственного задания номер госрегистрации темы: 01201353247.

Список литературы

1. Нуклеиновые кислоты: от А до Я. Под ред. Мюллера С. М.: БИНОМ. Лаборатория знаний, 2013. 413 с.
2. Hohn M.J., Palioura S., Su D., Yuan J., Söll D. Genetic analysis of selenocysteine biosynthesis in the archaeon *Methanococcus marisplacidus*. *Molecular Microbiology*. 2011. Vol. 81. No. 1. P. 249–258. DOI: 10.1111/j.1365-2958.2011.07690.x.
3. Gaston M.A., Zhang L., Green-Church K.B., Krzycki J.A. The complete biosynthesis of the genetically encoded amino acid pyrrolysine from lysine. *Nature*. 2011. Vol. No. 7340. 471. P. 647–650. DOI: 10.1038/nature09918.
4. Turanov A.A., Lobanov A.V., Fomenko D.E., Morrison H.G., Sogin M.L., Klobutcher L.A., Hatfield D.L., Glad-

shev V.N. Genetic code supports targeted insertion of two amino acid one codon. *Science*. 2009. Vol. 323. No. 5911. P. 259–261. DOI: 10.1126/science.1164748.

5. Wong J.T. Question 6: coevolution theory of the genetic code: a proven theory. *Origins of Life and Evolution of Biospheres*. 2007. Vol. 37. No. 4–5. P. 403–408. DOI: 10.1007/s11084-007-9094-1.

6. Taylor F.J.R., Coates D. The code within the codons. *Biosystems*. 1989. Vol. 22. No. 3. P. 177–187. DOI: 10.1016/0303-2647(89)90059-2.

7. Di Giulio M. The lack of foundation in the mechanism on which are based the physico-chemical theories for the origin of the genetic code is counterposed to the credible and natural mechanism suggested by the coevolution theory. *Journal of Theoretical Biology*. 2016. Vol. 399. P. 134–140. DOI: 10.1016/j.jtbi.2016.04.005.

8. Cleaves H.J. The origin of the biologically codes amino acids. *Journal of Theoretical Biology*. 2010. Vol. 263. No. 4. P. 490–498. DOI: 10.1016/j.jtbi.2009.12.014.

9. Doig A.J. Frozen, but no accident – why the 20 standard amino acids were selected. *FEBS Journal*. 2017. Vol. 284. No. 9. P. 1296–1305. DOI: 10.1111/febs.13982.

10. Carter C.W.Jr. What RNA World? Why a peptide/RNA partnership merits renewed experimental attention. *Life*. 2015. Vol. 5. No. 1. P. 294–320. DOI: 10.3390/life5010294.

11. Wagner A.J., Zubarev D.Yu., Aspuru-Guzik A., Blackmond D.G. Chiral sugars drive enantioenrichment in prebiotic amino acid. *ACS Central Science*. 2017. Vol. 3. No. 4. P. 322–328. DOI: 10.1021/acscentsci.7b00085.

12. Miller S.L. Production of some organic compounds under possible primitive earth conditions. *Journal of the American Chemical Society*. 1955. Vol. 77. No. 9. P. 2351–2361. DOI: 10.1021/ja01614a001.

13. Trifonov E.N. The triplet code from first principles. *Journal of Biomolecular Structure & Dynamics*. 2004. Vol. 22. No. 1. P. 1–11. DOI: 10.1080/07391102.2004.10506975.

14. Никитин М. Происхождение жизни. От туманности до клетки. М.: Альпина Диджитал, 2016. 363 с.

15. Kelley D.S., Karson J.A., Blackman D.K., Früh-Green G.L., Butterfield D.A., Lilley M.D., Olson E.J., Schrenk M.O., Roe K.K., Lebon G.T., Rivizzigno P. An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30°N. *Nature*. 2001. Vol. 412. No. 6843. P. 145–149. DOI: 10.1038/35084000.

16. LaRowe D.E., Regnier P. Thermodynamic potential for the abiotic synthesis of adenine, cytosine, guanine, thymine, uracil, ribose, and deoxyribose in hydrothermal systems. *Origins of Life and Evolution of Biospheres*. 2008. Vol. 38. No. 5. P. 383–397.

17. Kitadai N. Energetics of amino acid synthesis in alkaline hydrothermal environments. *Origins of Life and Evolution of Biospheres*. 2015. Vol. 45. No. 4. P. 377–409. DOI: 10.1007/s11084-015-9428-3.

18. Patel B.H., Percivalle C., Ritson D.J., Duffy C.D., Sutherland J.D. Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nature Chemistry*. 2015. Vol. 7. No. 4. P. 301–307. DOI: 10.1038/nchem.2202.

19. National Center for Biotechnology Information. PubChem compound database. 2019. [Electronic resource]. URL: <https://pubchem.ncbi.nlm.nih.gov> (date of access: 15.09.2019).

20. Alvarez-Carreno C., Bacerra A., Lazcano A. Norvaline and norleucine may have been more abundant protein components during early stages of cell evolution. *Origins of Life and Evolution of Biospheres*. 2013. Vol. 43. No. 4–5. P. 363–375. DOI: 10.1007/s11084-013-9344-3.

21. Powner M.W., Gerland B., Sutherland J.D. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature*. 2009. Vol. 459. No. 7244. P. 239–242. DOI: 10.1038/nature08013.