

## ПЕТЛИ И ПОВТОРЫ В БЕЛКАХ КАК ОТПЕЧАТКИ МОЛЕКУЛЯРНОЙ ЭВОЛЮЦИИ

©2012 г. Е. И. ДЕРЮШЕВА<sup>1</sup>, О. М. СЕЛИВАНОВА<sup>2</sup>,  
И. Н. СЕРДЮК<sup>2</sup>

<sup>1</sup> Тульский государственный университет, Тула

<sup>2</sup> Институт белка РАН, Пуцино, Московская область

I. Введение. II. Петли и повторы в белках. III. Роль неструктурированных участков в семействе элонгационных факторов. IV. Роль повторов в семействе рибосомных белков S1. V. Заключение.

### I. ВВЕДЕНИЕ

Идея о том, что уникальная третичная структура белка необходима для его функционирования, берет свое начало из экспериментов Э.Фишера, осуществленных более ста лет назад на экстрактах пивных дрожжей. Он высказал гипотезу, согласно которой в основе специфического взаимодействия молекул при химических реакциях лежит их пространственное соответствие типа «ключ–замок» [1].

Однако последующие исследования ряда белков доказали, что этот принцип не всегда выполняется. Первое сомнение в универсальности гипотезы «ключ–замок» появилось в 1950-е, когда Ф.Каруш обратил внимание на то, что обычный сывороточный альбумин проявляет способность связываться со многими малыми гидрофобными молекулами и высказал предположение о том, что гипотеза Фишера должна быть дополнена представлением о конфигурационной приспособляемости белков [2]. Однако идея Каруша о конфигурационных изменениях в белках в процессе их взаимодействия в лигандами, лежащих в основе множественности функций белка долгое время игнорировалась.

---

*Принятые сокращения:* G-семейство – семейство ГТФ-связывающих белков (GTP); EF-1 – фактор элонгации 1 (под EF-1 в данной работе понимается фактор элонгации EF-Tu и EF-Ts для эубактерий, aEF-1 для архей и EEF-1 для эукариот); EF-2 – фактор элонгации 2 (под EF-2 в данной работе понимается фактор элонгации EF-G для эубактерий, aEF-2 для архей и EEF-2 для эукариот); PNPase – полинуклеотидфосфорилаза.

*Адрес для корреспонденции:* janed1986@ya.ru.

Тем не менее, первые сравнения структуры лизоцима в двух функциональных состояниях (с лигандом и без него) показали, что в белках действительно имеют место пространственные смещения отдельных атомов [3]. Эти изменения носили локальный характер и не превышали 1 Å. Однако в ряде случаев изменения такого рода могли принимать более масштабный характер, как это было показано для дрожжевой гексокиназы при ее связывании с глюкозой [4], или для эубактериального элонгационного фактора EF-1 при замене ГДФ на ГТФ [5].

Намного позже был открыт ряд белков, которые при высокой гомологии последовательностей образовывали разные структуры, а для ряда амилоидогенных белков данные электронной и атомно-силовой микроскопии показывали, что часто один и тот же белок образует в растворе полиморфные структуры [6, 7].

Кроме того, были открыты белки, которые не имеют стабильной третичной структуры и находятся в растворе в состоянии, традиционно считавшимся денатурированным. Они получили название белков с природной или внутренней неупорядоченностью [8]. «Приобретаемая» структура для таких белков часто зависит от вида лиганда, и при этом может сильно изменяться. Разнообразие образуемых структур и своеобразная «структурная гибкость» (пластичность) определяет полифункциональность таких белков [9]. Таким образом, центральная догма молекулярной биологии для этих специфических белков преобразуется в следующую: один ген кодирует один белок, трехмерную структуру которого определяет партнер по взаимодействию или особенности окружающей его среды. Сформированная структура и определяет функцию или функции данного белка. Несмотря на то, что порядка 30% белков частично или полностью неструктурированы, до сих пор не было описано ни одного случая установленной связи между количеством и спецификой неупорядоченных участков и функциональной активностью этих белков.

Отметим тот факт, что полифункциональность не всегда является следствием наличия в белках дополнительных гибких участков структуры. Для ряда белков характерно множественное копирование собственной доменной структуры [10, 11]. При этом, в таких белках, каждый домен, несмотря на одну и ту же структуру выполняет различную функцию. Объяснение такого копирования и, как следствие, полифункциональности белков до сих пор не найдено.

В данной работе мы предлагаем новый взгляд на роль гибких участков (петель) в семействе элонгационных факторов, а также на число структурных доменов (повторов) в семействе рибосом-

ных белков S1, в процессе их эволюционного развития и, как следствие, в их функционировании. Эти семейства являются типичными представителями полифункциональных природно неструктурированных белков. Отметим, что рибосомный природно неструктурированный белок S1 к тому же содержит различное количество структурных повторяющихся доменов в разных организмах. Обнаруженная в этих белках корреляция между числом петель и повторов и сложностью организма позволила нам рассматривать их число как стабильный признак принадлежности организма к определенной систематической группе. Кроме того, выявленные закономерности позволили судить о «направлении» эволюции как этих белков, так и содержащих их организмов, что говорит об уникальности выявленных признаков с точки зрения молекулярной эволюции.

## **II. ПЕТЛИ И ПОВТОРЫ В БЕЛКАХ**

### **БЕЛКИ С ОДНОЙ ФУНКЦИЕЙ**

Принято считать, что каждый белок как вещество с определенным химическим строением выполняет одну узкоспециализированную функцию и лишь в нескольких отдельных случаях – несколько взаимосвязанных функций. В основном этот принцип верен для глобулярных белков – белков, в молекулах которых полипептидные цепи плотно свёрнуты в компактные квазисферические структуры – глобулы [12]. Этот тип белков представляет одну из трех главных групп белков: две другие группы – это фибриллярные белки и мембранные белки, которые обычно обнаруживаются в мембранах клетки и ее органелл [13, 14]. Функции глобулярных белков чрезвычайно многообразны. Они могут быть энзимами (например, амилаза и лактоза), биологическими мессенджерами (пептидные гормоны – пролактин и инсулин), транспортными белками (глобулин и альбумин), а также структурными белками внутри клеток (тубулин). Глобулярные белки функционируют как антитела (глобулин) в иммунной системе организма, участвуют в процессах репликации и репарации ДНК (полимераза).

Глобулярная структура белков обусловлена в основном гидрофобно-гидрофильными взаимодействиями. В глобулярных белках аминокислотная последовательность укладывается таким образом, чтобы увеличить растворимость белка в воде полярные группы атомов размещаются на поверхности белка (где они участвуют во взаимодействии с молекулами воды). Сворачивание и укладка цепей под

действием внутримолекулярных сил, которые существуют между разными группами атомов внутри молекулы, а также сил взаимодействия между группами атомов белка и молекул, непосредственно окружающих белок, определяют форму молекулы белка – его третичную структуру. При этом, элементы вторичной структуры (спирали, листы, повороты и петли) играют роль строительных блоков, из которых строится структура высшего порядка, а вся их «функциональная нагрузка» сводится, как правило, к сближению в пространстве аминокислот, образующих сайты связывания. В глобулярных белках основная роль участков, связывающих элементы вторичной структуры (петель) – обеспечить их «структурную гибкость» относительно друг друга [15].

Однако, как будет показано ниже, существует ряд белков, для которых именно размер и положение петель является определяющим для функционирования белковой молекулы.

#### БЕЛКИ СО МНОГИМИ ФУНКЦИЯМИ

##### *Природно неструктурированные белки*

С помощью ряда физических методов, таких как ЯМР (близость спектра неупорядоченного белка к суммарному спектру отдельных боковых групп свободных аминокислот); кругового дихроизма (низкая интенсивность сигнала в области 210–240 нм); малоуглового рассеяния (нетипично большие размеры и специфический ход кривой рассеяния в области, отвечающей за форму молекулы); гидродинамики (большие объемы молекул в растворе), было обнаружено, что многие белки в изолированном состоянии не обладают уникальной третичной структурой, хотя и имеют при этом четкую функцию при физиологических условиях. Кроме того, по мере появления все большего и большего числа рентгеновских структур белков, расшифрованных с атомарным разрешением, стало очевидным отсутствие на картах электронной плотности у некоторых из них относительно небольших участков полипептидной цепи. Это дало основание считать такие участки гибкими [16]. Белки, содержащие такие участки, получили название белков с природной или внутренней неупорядоченностью [8]. Доля неструктурированных областей в таких белках может быть разной, начиная от последовательности из нескольких аминокислот и заканчивая полностью неупорядоченной последовательностью длиной в десятки, а иногда и в сотни аминокислот [9].

Главное отличие этих белков от структурированных (глобулярных) белков состоит в том, что они не имеют уникальной третичной структуры в изолированном виде, а приобретают ее после взаимо-

действия с партнерами. Их конформация в комплексе определяется партнером по взаимодействию, а не только собственной аминокислотной последовательностью, как это характерно для структурированных (глобулярных) белков.

Открытие природно неструктурированных белков позволило расширить представление о том, какую конформацию белка можно считать нативной [17]. В дополнение к аксиоматическому нативному состоянию – жесткой уникальной структуре – предлагается рассматривать еще три: 1) расплавленная глобула (компактная структура при несколько большем гидродинамическом объеме, с хорошо развитой вторичной структурой, при отсутствии или следами кооперативно плавящейся третичной структуры), 2) полностью разупорядоченная цепь (набор быстро флуктуирующих конформаций, ассоциирующихся со статистическим гауссовым клубком) [18, 19], 3) доменная структура, соединенная гибкими и достаточно длинными перетяжками и, как правило, содержащая на концах достаточно длинные неупорядоченные области. Эти четыре состояния охватывают все конформации, которые известны сегодня из физических представлений о структуре белка.

Кроме того, все эти белки, как правило, полифункциональны – помимо основной функции они имеют дополнительные функции. По этой причине их часто называют белками «по совместительству». Функция белка «по совместительству» может изменяться в зависимости от многих факторов, таких как клеточная локализация, тип клеток, олигомерное состояние, концентрация клеточного лиганда, субстрата и тому подобное [20]. При этом достаточно часто оказывается, что функционально важные белковые участки находятся именно вне глобулярных доменов.

Всё это противоречит классическому представлению о том, что функционально активный белок обязательно должен быть глобулярным. Кроме того, масштаб конформационных изменений в природно неструктурированных белках при их взаимодействии с партнером гораздо больше, чем в структурированных белках.

#### *Белки со структурными повторами*

Взаимное расположение нескольких полипептидных цепей в составе единого белкового комплекса принято называть четвертичной структурой. В состав белка с четвертичной структурой могут входить как идентичные, так и различающиеся полипептидные цепочки. В стабилизации четвертичной структуры принимают участие те же типы взаимодействий, что и в стабилизации третичной. Отметим, что белки длиной более 200 а.о. – мультидоменные, причем каждый отдельный

домен такого белка может выполнять разные функции. При этом, одними из наиболее интересных с функциональной точки зрения белков, являются белки, содержащие множественное копирование одного и того же домена. Несмотря на то, что такие повторы были найдены в составе большого количества биологически важных белков, до сих пор не было описано ни одного случая установленной связи между количеством этих повторов и специфической активностью белка [10, 11]. Однако можно уверенно утверждать тот факт, что, именно данные повторы, обеспечивая множественные межбелковые взаимодействия, приводят к полифункциональности этих белков в различных организмах.

Рассмотренные в данной работе белки (факторы элонгации, семейство рибосомных белков S1) являются представителями рассмотренной выше группы белков – белков с различной степенью природной неструктурированности. Кроме того, рибосомный белок S1 содержит различное количество структурных повторяющихся доменов в разных организмах. Именно этот факт и объясняет интерес к этим белкам с точки зрения их эволюционного развития. По этой причине, начнем с общей характеристики особенностей выбранных объектов исследования.

#### *Семейство элонгационных факторов*

Выбор элонгационных факторов в качестве объекта исследования связан с несколькими обстоятельствами. Во-первых, они обладают набором признаков, характерных для белков с внутренней неупорядоченностью – специфичностью аминокислотного состава, большими размерами молекул в растворе [9, 21]. Во-вторых, в дополнение к основной канонической функции, эукариотические факторы образуют комплексы с разными лигандами, начиная от актина и заканчивая РНК [9]. Столь большое количество лигандов (партнеров) является характерным признаком белков с внутренней неупорядоченностью и, поэтому, представляет особый интерес сравнение числа и положения петель в факторах элонгации из разных организмов. В-третьих, они присутствуют в клетках любых организмов, так как принимают непосредственное участие в процессе биосинтеза белка [21]. Кроме того, элонгационные факторы – это белки G-семейства, которые были нами подробно рассмотрены ранее [22, 23].

Элонгационный фактор EF-1 в бактериальных клетках образует комплекс с ГТФ и тРНК в процессе биосинтеза белка, а в клетках эукариот образует комплексы с различными лигандами, такими как актин, тубулин и регуляторными белками [21]. Рентгеноструктурные

исследования факторов элонгации в изолированном состоянии и в комплексе с различными лигандами позволили составить полное представление об их структуре и принципах функционирования для эубактерий, архей и эукариот (рис. 1). Фактор состоит из трех доменов, разделенных двумя перетяжками. Домен I (нуклеотид-связывающий домен) соединяется с доменом II достаточно длинной перетяжкой (около 16 Å в *T. thermophilus*). Домен III связан с доменом I нековалентными взаимодействиями.

Элонгационный фактор EF-2 катализирует процесс транслокации [24]. Он состоит из пяти структурных доменов, первый из которых образует структурное ядро, свойственное всем другим белкам G-семейства. Внутри ядра находится субдомен G', который отсутствует в других G-белках. Структуры эубактериального и эукариотического белка EF-2 приведены на рис. 2.

Как видно из представленных структур, белок имеет вытянутую в пространстве конформацию. Кроме того, как у эубактериального, так и у эукариотического факторов присутствуют участки, которые не наблюдаются на картах электронной плотности

#### *Семейство рибосомных белков S1*

Согласно литературным данным, рибосомный белок S1 идентифицирован только в эубактериях и не идентифицирован в архебактериях и в эукариотах. Этот белок полифункциональный. Наряду с рядом рибосомных функций, таких как взаимодействие с 30S-рибосомной субчастицей, взаимодействие с мРНК и тмРНК [25], он выполняет и ряд внерибосомных функций [26]. Например, в составе одной из четырех субъединиц фаговой Qbeta-репликазы он необходим для репликации (+) цепи Qbeta-РНК, а при взаимодействии с рибонуклеазой regV бактериофага T4 повышает ее активность [27].

Изначально, считалось, что в рибосомный белок S1 из *E. coli* содержит четыре уникальных повтора [28]. Этот повтор был назван S1 РНК-связывающим мотивом или S1-доменом. Позже, с помощью выравнивания белковой последовательности самой на себя, в белке S1 из *E. coli* было идентифицировано шесть гомологичных S1 повторов [29]. Стоит отметить, что число S1-доменов в эубактериальных организмах меняется в строго ограниченном диапазоне: от одного до шести [30].

Известно, что белок полинуклеотидфосфорилаза (PNPase) из *E. coli* на С-конце содержит один S1-домен, имеющий высокую гомологию с первоначально выделенными четырьмя S1-повторами. Полученная методом ЯМР трехмерная структура S1 РНК-связывающего домена



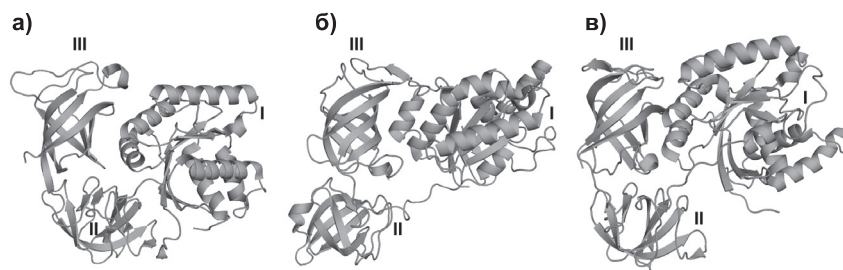


Рис. 1. Структура элонгационного фактора EF-1:

- а) эубактериального из *T. thermophilus* (PDB код 1HA3),
- б) архейного из *S. solfataricus* (PDB код 1JNY),
- в) эукариотического из *S. cerevisiae* (PDB код 1F60).

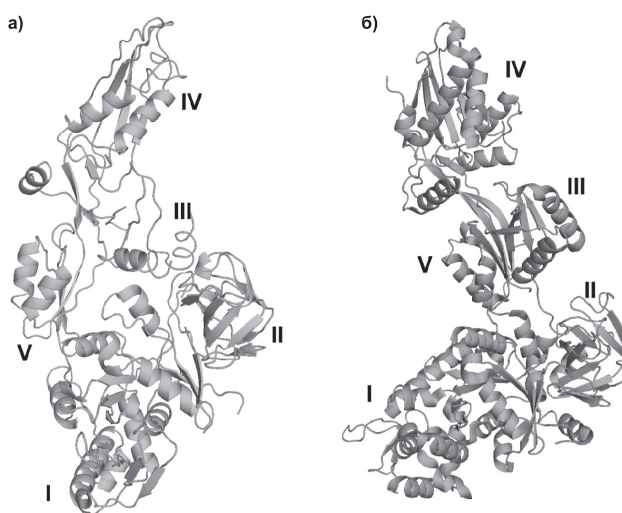


Рис. 2. Структура элонгационного фактора EF-2:

- а) эубактериального из *T. thermophilus* (PDB код 1ELO),
- б) эукариотического из *S. cerevisiae* (PDB код 1N0U).

из PNPase *E. coli* представляет  $\beta$ -баррель с дополнительной  $\alpha$ -спиралью между третьим и четвертым  $\beta$ -листами [31]. Именно такую ОВ-укладку (oligonucleotide binding fold) вторичной структуры принято рассматривать как основной структурный элемент семейства рибосомных белков S1 [10] (рис. 3а).

В S1-домене располагается ряд консервативных остатков (Phe-19, Phe-22, His-34, Asp-64 и Arg-68), собранных в кластер на его поверхности. Они расположены в петле между первым и вторым  $\beta$ -листами



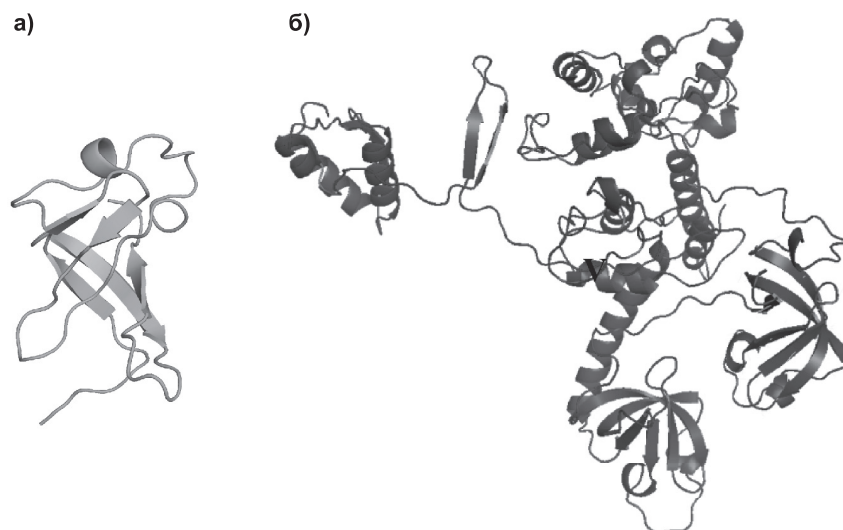


Рис. 3. а) Структура S1 домена из полинуклеотидфосфорилазы *E. coli* (PDB код 1SR0). б) Предсказанная структура белка S1 для *E. coli* программой Robbeta.

(Phe-19), в середине второго  $\beta$ -листа (Phe-22), в конце третьего  $\beta$ -листа (His-34) и в петле между последними двумя  $\beta$ -листами (Asp-64, Arg-68). Эти остатки пространственно сближены, располагаются с одной стороны бочонка и формируют участок, который рассматривается как РНК-связывающий сайт всего семейства рибосомных белков [10].

При выравнивании последовательностей эубактериальных белков S1, содержащих разное количество доменов, с S1 РНК-связывающим доменом PNPase из *E. coli*, нами ранее было обнаружено, что в семействе выявляется один домен, имеющий максимальную гомологию с доменом S1 из PNPase. Этот консервативный домен мигрирует по цепи, локализуется в белках с разным количеством доменов с определенной закономерностью, и, на наш взгляд, может рассматриваться как центральный РНК-связывающий домен в семействе рибосомных белков [30].

Несмотря на то, что белок S1 является самым большим из рибосомных белков, его трехмерная структура до сих пор не расшифрована. Однако, спомощью программы предсказания трехмерной структуры белков по аминокислотной последовательности Robetta [32] нам удалось оценить возможную структуру этого белка (рис. 3б).

Из предсказанной модели видно, что домены белка сильно разнесены в пространстве. Причем, кроме третьего домена (который мы

рассматриваем как консервативный [30]), все остальные домены не компактны и сильно разупорядочены. Отметим также, что линейные параметры данной модели находятся в хорошем соответствии с параметрами структуры реального белка S1, полученной методом криоэлектронной микроскопии с низким разрешением 11.5 Å [33].

## II. РОЛЬ НЕСТРУКТУРИРОВАННЫХ УЧАСТКОВ В СЕМЕЙСТВЕ ЭЛОНГАЦИОННЫХ ФАКТОРОВ

### ПРЕДСКАЗАНИЕ НЕСТРУКТУРИРОВАННЫХ УЧАСТКОВ В БЕЛКАХ

Поскольку развернутые участки белковой цепи играют важную роль в процессе функционирования белка, определению их точных границ уделяется большое внимание. Так, например, неструктурированные участки белковой цепи на N- и C-концах часто приводят к сложностям при экспрессии, выделении, очистке и кристаллизации таких белков.

Умение предсказывать существование неупорядоченных участков белковой структуры важно как для понимания функции белка, так и путей его сворачивания. Эти же данные необходимы и для дизайна белков, поскольку необходимо знать, какие особенности первичной структуры будут определять уникальность трехмерной структуры белка. Сведения о положении неструктурированных участков в белковой цепи помогают предварительно оценить компактность трехмерной структуры белков. Кроме того, умение правильно предсказывать неструктурированные участки необходимо для правильной интерпретации данных ЯМР [34, 35].

Так как белки с внутренней неупорядоченностью имеют специфичный аминокислотный состав, то возможно разделение всех аминокислот на две группы, в одну из которых входят аминокислоты, способствующие структурированности (Trp, Phe, Ile, Tyr, Val, Leu) а в другую – неструктурированности (Gly, Ser, Pro). Это означает, что аминокислотная последовательность кодирует беспорядок в белках [36]. На этой основе был предложен ряд программ для предсказания неупорядоченных областей в белке по его аминокислотной последовательности, таких как PONDR [37], RONN [38], DisEMBL [39], PreLINK [40], IUPred [41], GlobPlot 2 [42], FoldIndex [43]. Однако, как показывает практика, ни одна из этих программ не дает четких и точных предсказаний [23].

Для характеристики выбранных объектов исследования мы использовали Программу FoldUnfold, разработанную в Институте белка РАН [44, 45]. Принцип работы данной программы не раз освещался в литературе, поэтому мы лишь кратко опишем ее основные

особенности. Программа использует такое свойство остатков, как ожидаемое среднее число сближенных остатков (средняя плотность окружения), вычисленное по аминокислотной последовательности. Этот параметр определяет статус остатка, а именно: будет ли он свернут или развернут в нативном состоянии.

Для предсказания разупорядоченных участков белковой цепи по аминокислотной последовательности рассчитывается ожидаемое число сближенных остатков и строится профиль ожидаемого числа с выбранным окном усреднения. На профиле ожидаемых контактов проводится граница, разделяющая структурированные и неструктурированные участки. Если среднее ожидаемое число сближенных остатков для целого белка меньше чем 20.4, то такой белок следует рассматривать как неструктурированный полностью. Если это число больше 20.4, то белок рассматривается как структурированный, но может иметь неструктурированные участки [45]. По алгоритму программы размер неструктурированных участков в таком белке должен быть равен или больше размера усредняемого окна. Поэтому количество предсказанных участков зависит от ширины окна. Как было показано, ширина окна в 11 аминокислотных остатков оптимальна для поиска относительно коротких неструктурированных участков в полипептидной цепи [45]. В случае же поиска длинных неструктурированных участков в частично неструктурированных белках необходимо увеличить ширину окна до нескольких десятков аминокислот. Для поиска же коротких петель, соединяющих элементы вторичной структуры, необходимо использовать ширину усредняемого окна в 3 аминокислотных остатка, являющуюся оптимальной для такой задачи [23].

#### ЧИСЛО И ПОЛОЖЕНИЕ ПЕТЕЛЬ КАК ПРИЗНАК ДЛЯ КЛАССИФИКАЦИИ ОРГАНИЗМОВ

Для ответа на вопрос: «Существует ли взаимосвязь между количеством гибких участков (петель) в семействе элонгационных факторов и принадлежностью организма к тому или иному отделу?» были проанализированы более 300 аминокислотных последовательностей этих белков.

Нами было найдено, что суммарное число аминокислотных остатков в элонгационных факторах EF-1 отражает разделение живого мира на три Надцарства. Длины полипептидной цепи элонгационных факторов растут при усложнении организмов. Надцарство Эубактерий занимают интервал 393–406 а.о., Надцарство Архей – интервал 422–444 а.о., а Надцарство Эукариот – интервал 458–464 а.о.

Затем с помощью программы FoldUnfold при ширине усредняемого окна в 11 а.о. нами были проанализированы все гибкие предсказанные участки. Максимальное число предсказанных петель оказалось равным шести. Из шести разных петель три петли (А, В, С) выявляются в первом домене, одна петля (D) – во втором, и две петли (Е и F) – в третьем домене фактора, причем наличие всех шести петель никогда не обнаруживается в одном факторе. Положение дополнительных предсказанных петель на структуре петель для EF-1 из *S. cerevisiae* показано на рис. 4.

Факторы бактерий Надцарства Эубактерий содержат минимальное число петель. У *E. coli* нет ни одной дополнительной петли кроме эффекторной, тогда как у термофильных бактерий (*T. aquaticus* и *T. thermophilus*) выявляется одна термофильная петля. Сине-зеленые водоросли содержат одну или две дополнительные петли. Одна выявляется на месте термофильной петли (петля С), а другая выявляется в третьем домене фактора (петля Е как у *S. platensis*). Для факторов Надцарства Эукариот предсказывается существование дополнительных петель, число которых равно трем или четырем. Так, факторы животных содержат две петли (А и В) в первом домене фактора и две петли (Е и F) в третьем домене. Несколько особняком располагается отдел членистоногих: у них выявляется петля D. В Царстве растений не выявляются петли В и D. У всех эукариотических факторов на С-конце всегда присутствует полностью неупорядоченная область F длиной около 20 аминокислот.

Длины перетяжек, соединяющих первый и второй домены, в среднем одинаковы как в эубактериальных, так и в эукариотических факторах. Однако длины петель между вторым и третьим доменами в среднем на 8–10 аминокислот больше в эукариотических факторах. Этот факт мы рассматриваем как аргумент в пользу того, что эукариотические факторы обладают большей междоменной подвижностью, чем эубактериальные.

Принципиально новым результатом для понимания структуры эукариотических факторов является предсказание наличия у них на С-конце неструктурированной области длиной около 20 аминокислот (петля F). Длина этой области не зависит от источника выделения. Дополнительные петли внутри каждого из доменов, большая неструктурированная петля на С-конце и высокая междоменная подвижность в эукариотических факторах может служить объяснением того факта, что до сих пор ни один из факторов Надцарства Эукариот не закристаллизован в изолированной форме [46].

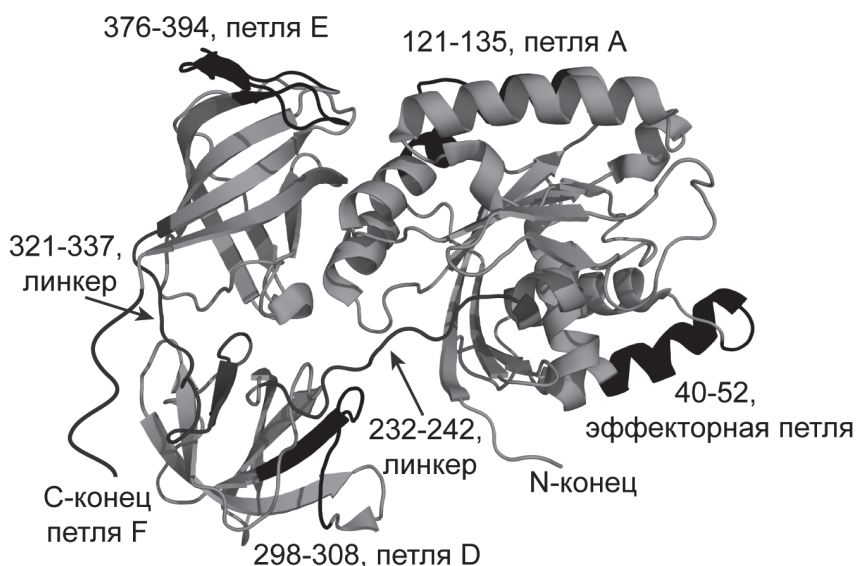


Рис. 4. Рентгеновская структура элонгационного фактора EF-1 *S. cerevisiae* (PDB код 1F60). Черным выделены предсказанные петли и перетяжки между доменами.

Факторы Надцарства Архей выглядят как специализированные белковые структуры, число петель в которых (две или три) и их расположение по доменам фактора определяются их функцией. Так, исключительным признаком серных бактерий является наличие петли А. Кроме того, у них предсказываются еще две петли (D и E). Гипертермофилы содержат петлю В, которой нет у остальных архей. Галлофилы не содержат петель А, В и С в первом домене, но всегда содержат петлю D во втором домене и петлю E в третьем домене. Общей чертой, объединяющей архейные факторы с зубактериальными факторами, является отсутствие петли F. Таким образом, данные, полученные при сопоставлении свойств белковых факторов, служат еще одним подтверждением того, что Археи представляют собой самостоятельное Надцарство, свойства которого существенно отличаются от таковых Надцарства Эубактерий.

Предсказываемые для архейного фактора две дополнительные петли длиной 14 и 18 а. о. в районе 283–296 и 364–381 а. о. имеют суммарную длину, близкую к разности между средней длиной архейных и зубактериальных факторов. Этот факт означает, что дополнительное удлинение полипептидной цепи архейных факторов по сравнению с зубактериальными идет на образование этих петель (рис. 5).

	Петли						кол-во петель
	A	B	C	D	E	F	
Эубактерии							
Бактерии							
Протеобактерии	-	-	-	-	-	-	0
Гипертермофильные бактерии	-	-	+	-	-	-	1
Цианобактерии	-	-	+	-	+	-	2
Эукариоты							
Протисты							
Хоанозоа	+	-	-	-	+	-	2
Зигмикота	+	-	-	-	+	-	2
Хитридномикота	+	-	-	-	+	-	2
Животные							
Черви	+	±	-	-	+	+	3-4
Членистоногие	+	±	-	±	+	+	4
Хордовые	+	+	-	-	+	+	4
Грибы	+	-	-	+	+	+	4
Растения	+	-	-	-	+	+	3
Кренархеоты	+	-	-	+	+	-	3
Археи							
Эвриархеоты							
Гипертермофилы	-	+	-	+	+	-	3
Галобактерии	-	-	-	+	+	-	2
Метаногены	-	-	-	+	+	-	2
домены фактора	I			II	III		

схема расположения доменов фактора EF-1

Рис. 5. Предсказываемые петли в факторах элонгации EF-1 для трех Надцарств Живого мира.

Аминокислотный состав некоторых петель весьма консервативен и удивительно сходен у факторов из различных источников. Так, структурный мотив петли А для всех факторов внутри Надцарства Эукариот почти не отличается (GEFEAGISKNGQTR – у животных, GGFERGISKDGQTR – у растений и GEFEAGISKDGQTR – у грибов), тогда как та же петля А у типа кренархеоты Надцарства Археи имеет совершенно другой мотив (KGEYEAGMSAEG). Также консервативна петля Е у эукариот (DRR(T)SGK). Интересно, что структурный мотив петли В у гипертермофилов архей (MDATEPPFSEK) напоминает такой же петли В у животных (MDSTEPPYS), а мотив петли D у серных бактерий и гипертермофилов (PGDNIGF) близок к мотиву такой же петли для некоторых представителей эукариот (PGDNVGF).

Протисты содержат петли А и Е, типичные для Царства Животных и Грибов. Это подтверждает возможность существования такой супер-группы как Opisthokonta, выделенной рядом авторов в отдельную эволюционную группу организмов [47, 48].

В элонгационных факторах EF-2 суммарное число аминокислотных остатков также как для EF-1 отражает разделение живого мира на три Надцарства. Длины полипептидной цепи элонгационных факторов растут при усложнении организмов. Надцарство Эубактерий занимает интервал 691–704 а.о., Надцарство Архей – интервал 727–737 а.о., а Надцарство Эукариот – интервал 842–858 а.о.

Анализ дополнительных петель, предсказанных программой FoldUnfold для представителей Трех Надцарств Живого мира белка EF-2 показал, что их число равно 10, четыре из которых предсказываются в первом домене, по одной во втором и третьем доменах, и четыре – в четвертом домене. Как и для белка EF-1, элонгационные факторы бактерий содержат минимальное число петель. Так, например у протеобактерий и гипертермофилов не предсказывается не одной дополнительной петли, а у цианобактерий и фирмикутов предсказывается только одна дополнительная петля. Для эукариот число дополнительных петель колеблется от пяти (растения) до восьми (членистоногие и грибы). Тип хордовых Царства Животных стоит особняком, так как содержит петлю В. Кроме того, все эукариотические факторы, имеют петли С, Н, и I. Элонгационные факторы Надцарства Археи имеют специфический набор предсказываемых петель. Так например, все археи имеют петлю А, которой нет у других организмов, а петля G есть у всех архей кроме серных бактерий. Стоит отметить, что аминокислотный состав предсказываемых петель для факторов элонгации довольно консервативен для представителей разных Царств (рис. 6).

Таким образом, в рассмотренных белковых последовательностях факторов элонгации отличие одного организма от другого сводится к присутствию той или иной петли, а также специфики их аминокислотного состава. У простейших организмов их число оказалось минимальным, тогда как у высших – максимальным.

К сожалению, имеющиеся в литературе данные пока недостаточны для установления четкой связи между всеми петлями и функцией белка, за исключением эффекторной и термофильной петель [49, 50]. Кроме того, подробное исследование G-доменов семейства ГТФ-связывающих белков [24, 26] позволило нам установить функцию петли В в факторах элонгации EF-1. Петля В в факторах элонгации EF-1, предсказанная при ширине окна в 11 а.о. соответствует одной из



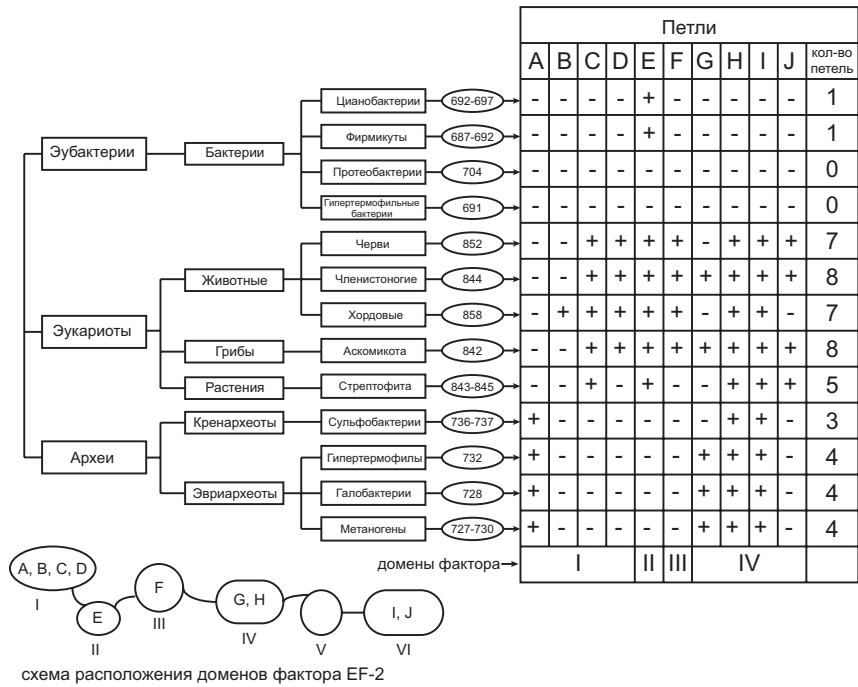


Рис. 6. Предсказываемые петли в факторах элонгации EF-2 для трех Надцарств Живого мира.

пяти петель, формирующих нуклеотид-связывающий сайт гуанина в G-домене. Эта петля имеет консервативный состав NKXD (X – любая аминокислота) и частично участвует в распознавании гетероцикла гуанина.

Также отметим, что во всех дополнительных петлях присутствует лизин – аминокислота, ответственная за особенности строения факторов элонгации [9].

### III. РОЛЬ ПОВТОРОВ В СЕМЕЙСТВЕ РИБОСОМНЫХ БЕЛКОВ S1

#### ПРЕДСКАЗАНИЕ ПОВТОРОВ ПО АМИНОКИСЛОТНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ

Информация о белковых доменах, их количестве и границах в различных белковых семействах можно найти в таких базах данных как SMART [51], Pfam [52], PROSITE [53]. В основе каждой базы данных, дающей информацию о структуре белков, их укладке, доменной организации лежат определенные алгоритмы работы. Поскольку эти алгоритмы в ряде случаев различны, а в базах данных, имеющих один и тот же алгоритм, обычно вводятся дополнительные условия или ограничения, выходные данные для одного и того же объекта могут существенно различаться [30].

По этой причине для выделения структурных доменов в семействе рибосомных белков S1 мы использовали собственный алгоритм [30]. Опишем его кратко. Для определения вторичной структуры, первичные последовательности семейства рибосомных белков S1, имеющиеся в базе данных SwissProt [54], были проанализированы с помощью программы PsiPred [55]. Далее, основываясь только на предсказанной вторичной структуре для S1 РНК-связывающего домена PNPase из *E. coli* и без привлечения баз данных SMART, Pfam и PROSITE, с помощью множественного выравнивания мы выделили в белковых последовательностях структурные домены, число которых изменялось от единицы до шести. Отметим, что для большинства рассмотренных бактерий, количество доменов S1 совпадает для трех баз данных (SMART, PROSITE, Pfam). Например, все бактерии из отряда *Bacillales* содержат четыре S1 повтора, а у отряда *Rhodobacterales* (Alpha proteobacteria) – шесть повторов. Однако для ряда бактерий были обнаружены сильные расхождения [30].

#### ЧИСЛО СТРУКТУРНЫХ ПОВТОРОВ КАК СТАБИЛЬНЫЙ ПРИЗНАК ПРИНАДЛЕЖНОСТИ МИКРООРГАНИЗМА К ОПРЕДЕЛЕННОМУ ОТДЕЛУ

В настоящее время все бактерии принято делить на 25 основных групп или отделов (согласно выравниванию последовательностей 16S рРНК [56]). Некоторые отделы включают в себя только нескольких видов бактерий (например, Thermomicrobia, Chrysiogenetes, Fibrobacteres, Deferribacteres), в то время как к другим отделам, таким как Proteobacteria, Bacteroidetes, Cyanobacteria, Actinobacteria и Firmicutes относятся около 90–95% всех известных бактерий.

В данной работе было рассмотрено более 400 аминокислотных последовательностей белка S1 следующих отделов бактерий: Tenericutes, Cyanobacteria, Firmicutes, Actinobacteria, Nitrospira, Deinococcus-Thermus, Bacteroidetes, Planctomycetes, Spirochaetes, Chlorobi, Chlamydiae и Proteobacteria.

Как было отмечено выше, число структурных S1-доменов в эубактериальных организмах варьируется в строго ограниченном диапазоне: от одного до шести. Длина одного S1 домена около 70 а.о. Учитывая длину перетяжек между структурными доменами, легко вычислить число структурных доменов бактериальных белков различной длины. Однако, белок S1 из *T. pallidum* длиной 863 а.о. содержит также как и белок из *E. coli* длиной 557 а.о. только шесть структурных доменов. Увеличение белковой последовательности свыше 600 а.о. происходит либо за счет удлинения его концов, либо за счет появления других структурных доменов.

На рис. 7 представлены все исследуемые отделы бактерий и количество обнаруженных в них S1 доменов.

Наименьшая длина белка S1 наблюдается у представителей семейства Mycoplasmataceae (*M. mobilis* 116 а.о) и Spiroplasmataceae (*Sp. kunkelii* 111 а.о). Рибосомный белок этих бактерий содержит один S1 домен. Цианобактерии (Cyanobacteria), с длиной белка S1 в среднем около 350 а.о., содержат три S1 домена, а представители отдела Firmicutes с длиной белка 390 а.о. содержат в зависимости от класса три (для Clostridia) или четыре S1 (для Bacillaceae) домена. Все рассмотренные бактерии, относящиеся к отделу Actinobacteria, имеют четыре S1 домена. Бактерии малочисленного монотипического (состоит из одного класса Deinococci) отдела Deinococcus-Thermus имеют длину S1 белка в среднем около 530 а.о. и всегда содержат пять S1 доменов. Также пять S1 доменов обнаруживаются в бактериях отдела Nitrospira. К граммотрицательным бактериям, содержащим шесть S1 доменов, относятся Proteobacteria ( $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ), Spirochaetes, Bacteroidetes и Chlamydiae. Длина рибосомного белка S1 у всех этих бактерий составляет в среднем около 570 а.о. Также шестидоменными являются и рибосомные белки S1 бактерий отделов Chlorobi (зеленые серные бактерии) и Planctomycetes. Интересным является тот факт, что во всех исследуемых отделах было обнаружено всего несколько бактерий, содержащих два S1 домена, например *S. epidermidis* относящаяся к отделу Firmicutes. Все остальные представители различных отделов и классов четко сохраняют фиксированное число структурных S1 доменов. Таким образом, число структурных доменов является стабильным признаком принадлежности микроорганизма к тому или иному отделу.

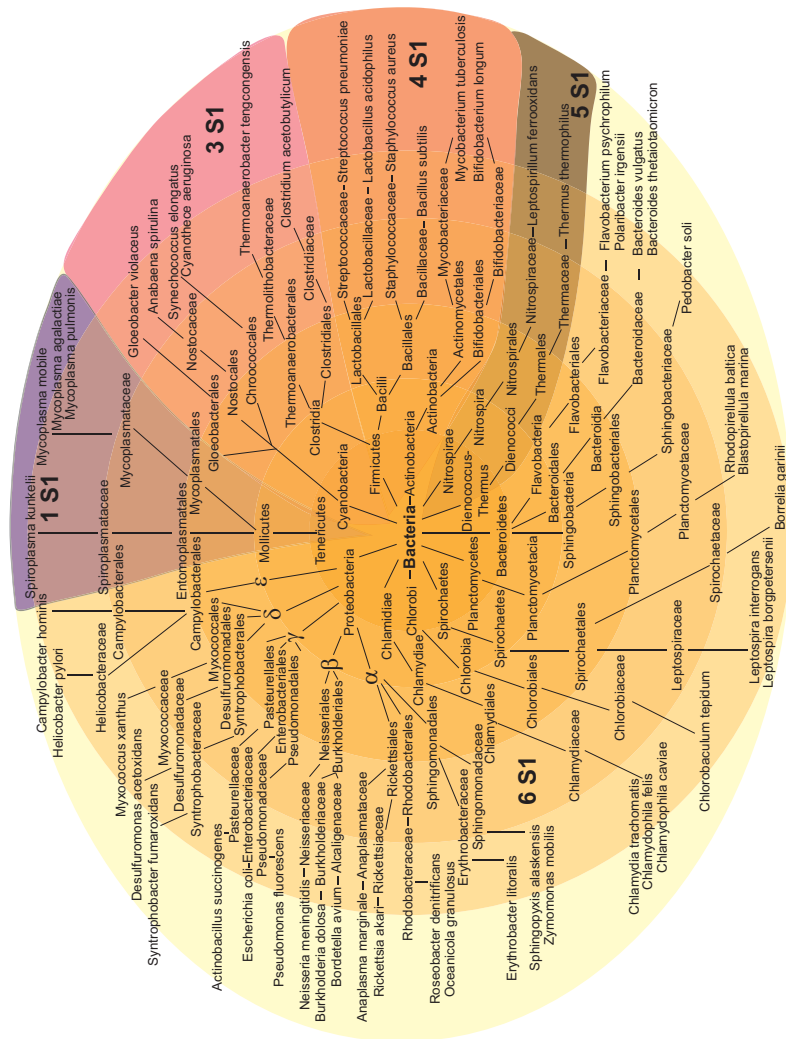


Рис. 7. Число доменов рибосомного белка S1 для Надцарства Эубактерий. На схеме представлено деление исследуемых бактерий на отделы, классы, отряды, семейства (от центра к периферии).

#### УНИКАЛЬНЫЙ ПОВТОР (ДОМЕН) КАК КОНСЕРВАТИВНЫЙ РНК-СВЯЗЫВАЮЩИЙ ЦЕНТР СЕМЕЙСТВА

Так как S1 домен является одним из «древнейших» белковых доменов, его наличие у бактерий в различных сочетаниях, скорее всего, является прямым следствием эволюционного развития. Возможны два пути эволюционного развития белковых молекул организмов – «вверх» и «вниз».

Принцип прогрессивной эволюции базируется на представлениях о Дарвиновской дивергенции: ветви филогенетического дерева расходятся, чтобы никогда не сойтись. Построенные описанным выше способом деревья отражают только вертикальную, направленную «вверх», эволюцию, при которой наследуется базовый геном, часто подверженный различным рекомбинациям и мутациям, приводящим в итоге к фенотипическим изменениям организмов [57]. Однако известны и случаи влияния на размер и функциональность генов и белков так называемой редукционной эволюции. При таком влиянии отсекается все «ненужное», т.е. со временем происходит сокращение числа генов, и, как следствие, белковых продуктов, а также утрата различных функций и даже целых клеточных органелл. Такая «закономерность» особенно свойственна патогенным бактериям, так как они используют для своей жизни метаболические системы клетки «хозяина» и не нуждаются в собственных аналогичных системах [58]. Еще одним важным ключевым направлением эволюции является «способность» горизонтального, или, как его принято называть, латерального переноса генов между организмами. Горизонтальный перенос генов – один из главных механизмов видообразования в мире прокариот. На сегодняшний день уже доказано, что не менее 80% генов в каждом зубактериальном геноме на том или ином этапе эволюции участвовали в процессе горизонтального переноса, а многочисленные данные геномики подтверждают факт переноса массивных генов между организмами не только внутри различных Царств, но также и между ними. Также доказана возможность обмена целыми геномами, что приводит к мгновенному превращению одного вида бактерий в другой. О «чужеродном» происхождении гена говорит высокая степень его сходства с гомологичным геном из отдаленного таксона при отсутствии подобного гена у близких «родственников» [59].

В нашем случае «вверх» означает что более «древними» являются бактерии отдела Tenericutes рода *Mycoplasma* (микоплазмы) с наименьшим количеством S1 доменов, а развитие «вниз» подразумевает эволюционное развитие от шестидоменных протеобактерий.

В литературе по данному вопросу существуют две точки зрения. Согласно одной из них микоплазмы являются выжившей ветвью примитивных микроорганизмов, из которых впоследствии произошли зубактерии и эукариоты. Они появились как продукт прогрессивной эволюции еще до образования присущей бактериям клеточной стенки [60]. Другая точка зрения состоит в том, что микоплазмы являются регрессивной ветвью эволюции некоторых или грамположительных бактерий, или кластридий (отдел Firmicutes). Последняя гипотеза находит экспериментальное подтверждение и рассматривается в двух возможных вариантах: все микоплазмы происходят либо от предка, общего с грамположительными бактериями, либо от разных бактерий [61]. На основании проведенного сравнения последовательностей олигонуклеотидов 16S рРНК нескольких видов микоплазм и грамположительных бактерий из родов *Clostridium*, *Bacillus*, *Lactobacillus*, *Streptococcus* было высказано обоснованное предположение об их эволюционном родстве с отделом Firmicutes [60, 62]. Более детальный анализ последовательностей 16S РНК показал, что филогенетически микоплазмы ближе всего к кластридиям [63]. В свою очередь наиболее вероятными предками кластридий считаются грамположительные бактерии с низким содержанием G–C пар в ДНК.

Продоланный нами анализ структурных особенностей семейства рибосомных белков S1, принадлежащих к разным отделам, свидетельствует в пользу второго («вниз») эволюционного пути развития зубактерий. Это заключение подтверждают несколько фактов. Во-первых, микоплазмы – самые маленькие, самостоятельно воспроизводящиеся живые бактерии, выполняющие все необходимые им функции. Во-вторых, биохимические эксперименты по изучению фрагментов шестидоменного белка S1 из *E. coli* (обрезание доменов на концах последовательности) показали, что функциональность белка при этом сохраняется, а снижается лишь эффективность выполняемых белком функций [64, 65]. Кроме того, S1 домен «переходит» в другие белки (как зубактериальные, так и архейные) чаще всего одной «копией» (фактор инициации, рибонуклеаза).

В составе эукариотических клеток белок S1 не был найден, хотя в некоторых эукариотических белках было обнаружено множественное «копирование» S1 домена (от девяти до двенадцати копий). Например, дрожжевой белок Rrp5p, участвующий в процессинге пре-рРНК и созревании 40S и 60S субчастицы, содержит двенадцать tandemно повторяющихся S1 доменов [66].

Такое большое количество S1 доменов может объясняться латеральным переносом в эукариотический химерный геном «чужих»,

кодирующих белок генов, приобретенных сравнительно недавно и еще не подвергшихся амелиорации в составе генома. При этом партнерами для латерального переноса генов часто являются филогенетически отдаленные организмы. Такие случаи в большей степени свойственны белкам, ответственным за белок-белковые и ДНК-белковые взаимодействия, т.е. участвующим в сборке сложных комплексов или в регуляции действия генов. Напомним, что S1 домен является РНК-связывающим.

Как уже было отмечено ранее, проведенные исследования семейства рибосомных белков S1 эубактерий позволили выявить в нем уникальный консервативный домен, обладающий РНК-связывающими свойствами, который мигрирует по цепи, локализуется в белках с разным количеством доменов с определенной закономерностью [30]. Суммируя все вышесказанное можно утверждать, что именно этот домен и являлся конечным продуктом действия эволюции: при минимальном размере белка обеспечивается выполнение им всех необходимых функций.

## **V. ЗАКЛЮЧЕНИЕ**

Мы обнаружили, что в качестве нового признака и отпечатка молекулярной эволюции для семейства элонгационных факторов может выступать количество гибких петель. Как было показано выше, в простейших организмах их число оказалось минимальным, тогда как в высших – максимальным. Это дало основание высказать гипотезу об эволюционном развитии факторов элонгации: «Природа, придерживаясь принципа бережливой изобретательности, оперирует разными универсальными вставками (в данной работе петли), комбинируя их число и положение в факторах таким образом, чтобы белок мог выполнять ряд специализированных функций: одну – в простейших микроорганизмах и несколько – в высших».

Длина белка увеличивается с увеличением сложности организма, что является характерной особенностью прогрессивной эволюции. Однако при этом вторичная структура и доменная организация эубактериальных, архейных и эукариотических факторов оставалась весьма консервативной. Различие же в длине факторов обусловлено появлением дополнительных гибких участков в трехмерной структуре белка, а также увеличением длины междоменных перетяжек, приводящих к большей подвижности доменов относительно друг



друга в пространстве. Кроме того, наличие дополнительных петель со специфической последовательностью в эукариотических факторах элонгации может служить возможным объяснением полифункциональности этих белков. Как уже упоминалось в обзоре, в дополнении к своей основной функции, взаимодействие с тРНК, эукариотические факторы образуют комплексы с различными лигандами, начиная от кальмодулина и заканчивая РНК. Столь многообразная природа лигандов требует наличия у этих белков разных функциональных участков, в роли которых могут выступать обнаруженные нами петли.

При исследовании семейства рибосомных белков S1 было обнаружено, что число структурных доменов (повторов) может рассматриваться как стабильный признак принадлежности организма к тому или иному отделу.

Обнаруженные закономерности, по нашему мнению, позволяют судить о «направлении» эволюции этих белков. Для семейства элонгационных факторов это, без сомнения, действие прогрессивной эволюции, приведшей к появлению дополнительных петель в эукариотических организмах и, как следствие, их полифункциональности. Для семейства рибосомных белков S1 – это изменение числа структурных доменов, являющееся результатом латерального переноса гена белка S1 в ходе редуцированной эволюции между основными отделами Надцарства Эубактерий с сохранением в качестве «конечного» продукта эволюции белка минимального размера. Это позволило паразитическим организмам класса Mollicutes, несмотря на маленькие размеры, выполнять все необходимые им функции.

Предложенный новый взгляд на роль гибких участков (петель) в семействе элонгационных факторов, а также на число структурных доменов (повторов) в семействе рибосомных белков S1, отличается от всех принятых подходов к выбору признаков систематики и, как нам представляется, открывает новые возможности для исследования белковых последовательностей с точки зрения молекулярной эволюции и систематики.

## ЛИТЕРАТУРА

1. Lemieux, U., Spohr, U. (1994) Adv. Carbohydrate Chem. Biochem., **50**, 1–20.
2. Karush, F. (1950) J. Am. Chem. Soc., **72**, 2705–2713.
3. Blake, C., Koenig, D., Mair, G., North, A., Phillips, D., Sarma, V. (1965) Nature, **206**, 757–761.
4. McDonald, R., Steitz, T., Engelman, D. (1979) Biochemistry, **18**, 338–342.
5. Kjeldgaard, M., Nissen, P., Thirup, S., Nyborg, J. (1993) Structure, **1**, 35–50.
6. Dobson, C. (1999) Trends Biochem. Sci., **24**, 329–332.
7. Pedersen, J., Andersen, C., Otzen, D. (2010) FEBS J., **277**, 4591–4601.
8. Uversky, A., Gillspie, J., Fink, A. (2000) Proteins, **41**, 415–427.
9. Сердюк И.Н. (2007) Молекулярная биология, **42**, 287–313.
10. Murzin, A. (1993). The EMBO J., **12**, 861–867.
11. Mosavi, L., Cammett, T., Desrosiers, D., Peng, Z. (2004) Protein Sci., **13**, 1435–1448.
12. Betz, S. (1993) Protein Sci., **2**, 1551–1558.
13. Fraser, R., MacRae, T., Suzuki, E. (1979) J. Mol. Biol., **129**, 463–448.
14. Daley, D. (2008) Curr. Opin. Struct. Biol., **18**, 420–424.
15. Timasheff, S., Gorbunoff, M. (1967) Annu. Rev. Biochem., **36**, 13–54.
16. Dunker, A., Lawson, J. (2001) J. Mol. Graphics Modell., **19**, 26–59.
17. Dunker, A., Brown, C., Lawson, J. (2002) Biochemistry, **41**, 6573–6582.
18. Ptitsyn, O. (1995) Adv. Protein Chem., **47**, 83–229
19. Receveur-Brechot, V., Bourish, J., Uversky, V. (2006) Proteins, **62**, 24–45.
20. Jeffrey, C. (1999) TIBS, **24**, 8–11.
21. Budkevich, T., Timchenko, A., Tiktopulo, E. (2002) Biochemistry, **41**, 15342–15349.
22. Дерюшева Е.И., Галзитская О.В., Сердюк И.Н. (2008) Молекулярная биология, **42**, 1067–1078.
23. Дерюшева Е.И., Левин Д.М., Сердюк И.Н. (2009) Известия ТулГУ. Естественные науки, **3**, 229–237.
24. Evarsson, A., Brazhnikov, E., Garber, M. (1994) The EMBO J., **13**, 3669–3677.
25. McGinness, K., Sauer, R. (2004) PNAS, **101**, 13454–13459.
26. Dahlberg, A., Dahlberg, J. (1975) Proc. Natl. Acad. Sci. USA, **72**, 2940–2944.
27. Blumenthal, T., Young, R., Brown, S. (1976) J. Biol. Chem., **251**, 2740–2743.
28. Regnier, P., Grunberg-Manago, M., Portier, C. (1987) J. Biol. Chem., **262**, 63–68.
29. Subramanian A. (1983) Prog. Nucleic Acid Res. Mol. Biol., **28**, 101–142.
30. Дерюшева Е.И., Мачулин А.В., Селиванова О.М., Сердюк И.Н. (2010) Молекулярная биология, **44**, 728–734.
31. Bycroft, M., Hubbard, T., Proctor, M., Freud, S., Murzin, G. (1997) Cell, **24**, 235–242.
32. Kim, D., Chivian, D., Baker, D. (2004) Nucleic Acids Res., **32**, W526–W531.
33. Frank, J. (2001) Proc. Natl. Acad. Sci. USA, **98**, 11991–11996.
34. Fink, A. (2005) Curr. Opin. Struct. Biol., **15**, 35–41.

35. Dyson, H., Wright, P. (2005) *Nat. Rev. Mol. Cell Biol.*, **6**, 197–208.
36. Tompa, P. (2002) *Trends Biochem. Sci.*, **27**, 527–533.
37. Romero, P., Obradovic, Z., Li, X., Garner, E., Brown, C., Dunker, A. (2001) *Proteins*, **42**, 38–48.
38. Yang, Z., Thomson, R., McNeil, P., Esnouf, R. (2005) *Bioinformatics*, **21**, 3369–3376.
39. Linding, R., Jensen, L., Diella, F., Bork, P., Gibson, T., Russell, R. (2003) *Structure*, **11**, 1453–1459.
40. Coeytaux, K., Poupon, A. (2005) *Bioinformatics*, **21**, 1891–1900.
41. Dosztanyi, Z., Csizmok, V., Tompa, P., Simon, I. (2005) *Bioinformatics*, **21**, 3433–3434.
42. Linding, R., Russell, R., Neduva, V., Gibson, T. (2003) *Nucleic Acids Res.*, **31**, 3701–3708.
43. Prilusky, J., Felder, C., Zeev-Ben-Mordehai, T., Rydberg, E., Man, O., Beckmann, J., Silman, I., Sussman, J. (2005) *Bioinformatics*, **21**, 3435–3438.
44. Галзитская О.В., Гарбузинский С.А., Лобанов М.Ю. (2006) *Молекулярная биология*, **40**, 341–348.
45. Galzitskaya, O., Garbuzynskiy, S., Lobanov, M. (2006) *Bioinformatics*, **22**, 2948–2949.
46. Andersen, G., Pedersen, L., Valente, L. (2000) *Mol. Cell*, **6**, 1261–1266.
47. Steenkamp, E., Wright, J., Baldauf, S. (2005) *Mol. Biol. Evol.*, **23**, 93–106.
48. Galzitskaya, O., Deryusheva, E., Serdyuk, I. (2008) *J. Comput. Sci. Syst. Biol.*, **1**, 73–80.
49. Kjeldgard, M., Nissen, P., Thirup, S., Nyborg, J. (1993) *Structure*, **1**, 35–50.
50. Kjeldgard, M., Nyborg, J. (1992) *J. Mol. Biol.*, **223**, 721–742.
51. Schultz, J., Milpetz, F., Bork, P., Ponting, C. (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 5857–5864.
52. Sonnhammer, E., Eddy, S., Birney, E., Bateman, A., Durbin, R. (1998) *Nucleic Acids Res.*, **26**, 320–322.
53. Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., Cuche, B., Castro, E., Lachazel, C., Langendijk-Genevaux, P., Sigrist, C. (2008) *Nucleic Acids Res.*, **36**, 245–249.
54. Bairoch, A., Apweiler, R. (2000) *Nucl. Acids Res.*, **28**, 45–48.
55. Bryson, K., McGuffin, L., Marsden, R., Ward, J., Sodhi, J., Jones, D. (2005) *Nucl. Acids Res.*, **33**, 36–38.
56. Guttel, R., Larsen, N., Woese, C., (1994) *Microbiol. Rev.*, **58**, 10–26
57. Doolittle, W. (1999) *Science*, **284**, 2124–2128.
58. Lartigue, C., Glass, J., Alperovich, N., Pieper, R., Parmar, P., Hutchinson, C., Smith, H., Venter J. (2007) *Science*, **317**, 632–638.
59. Dagan, T., Artzy-Randrup, Y., Martin, W. (2008) *Proc. Natl. Acad. Sci. USA*, **105**, 10039–10044.
60. Bhugra, B. (1992) *Mol. Microbiol.*, **6**, 1149–1154.
61. Neimark, H.C. (1979) Phylogenetic relations between Mycoplasmas and other prokaryotes. The mycoplasmas. / Eds. M.F.Barile and S.Razin. New York, **1**, 43–61.
62. Борхсениус С.Н. Микоплазмы. Молекулярная и клеточная биология, взаимодействие с иммунной системой млекопитающих, патогенность, диагностика (2000) М: Наука, 123 с.
63. Blanchard, A. (1990) *Mol. Microbiol.*, **4**, 669–676.
64. Amblar, M., Barbas, A., Gomez-Puertas, P., Arraiano, C. (2007) *RNA*, **13**, 317–327.

65. *Boni, I., Artamonova, V., Dreyfus, M.* (2000) *J. Bacteriol.*, **182**, 5872–5879.
66. *De Boer, P., Vos, H., Faber, A., Vos, J., Raue, H.* (2006) *RNA*, **12**, 263–271.