

УДК 577.322

ФОРМИРОВАНИЕ ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ МОЛЕКУЛ РНК

© 2012 г. Е. И. Леонова^{1,2}, М. В. Баранов^{1,3}, О. В. Галзитская^{1*}

¹Институт белка Российской академии наук, Пущино, Московская обл., 142290

²Филиал Института биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова Российской академии наук, Пущино, Московская обл., 142290

³Биологический факультет Московского государственного университета им. М.В. Ломоносова, Москва 119991

Поступила в редакцию 03.03.2011 г.

Принята к печати 21.06.2011 г.

Рассмотрены экспериментальные и теоретические подходы для изучения самоорганизации молекул РНК и выявления нуклеотидов, играющих важную роль при сворачивании, на примере таких молекул как тРНК и рибозимы. Так, показано, что нуклеотиды в области петель D- и T-шпилек вовлекаются в структуру тРНК последними или, точнее, не входят в ядро сворачивания тРНК. Модель иерархического сворачивания тРНК, которой придерживались длительное время, оказалась неточна, как было показано с помощью специально разработанного метода SHAPE. В обзоре также рассмотрены алгоритмы и программы, используемые для предсказания вторичных структур и моделирования сворачивания пространственных структур РНК.

Ключевые слова: сворачивание РНК, ядро сворачивания, водородная связь, стэкинг- и гидрофобные взаимодействия, неиерархическая модель сворачивания.

FORMATION OF SPATIAL STRUCTURE OF RNA MOLECULES, by E. I. Leonova¹, M. V. Baranov^{1, 2}, O. V. Galzitskaya* (¹Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region, 142290 Russia, *e-mail: ogalzit@vega.protres.ru; ²Branch of Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Pushchino, Moscow Region, 142290 Russia; ³Biological Department, Moscow State University, Moscow 119991 Russia). In this review we consider several experimental and theoretical approaches for investigation of RNA folding and determination of nucleotides that play an important role upon folding of such molecules as tRNA and several classes of ribozymes. It has been shown that nucleotides in the D- and T-loop regions are the last to be involved in tRNA structure or they are not included in the folding nucleus of tRNA. Using the specially elaborated method SHAPE it has been demonstrated that the model of hierarchical folding which was recognized for a long time is not correct for tRNA folding. In the second part of the given review the algorithms and programs used for the prediction of secondary structures of RNA as well as for modeling of RNA folding are considered.

Keywords: RNA folding, folding nucleus, H-bonds, base stacking, hydrophobic interactions, non-hierarchical model of folding.

Огромное количество разнообразных биохимических процессов, в которых РНК играет одну из ключевых ролей, поддерживает работоспособность живой клетки. И это неудивительно, так как для того, чтобы синтезировать белок, нужно задействовать целый комплекс рибосомных, матричных, транспортных, а также регуляторных РНК. Несмотря на незначительные отличия в строении ДНК и РНК, а именно: вместо дезоксирибозы в состав нуклеотида входит рибоза, а вместо тимина – урацил, функции их заметно отличаются. Это связано с тем, что ДНК в клетке име-

ет двухцепочечную форму, что делает ее очень жесткой, а РНК – одноцепочечную, что делает ее гибкой и дает ей возможность сворачиваться в большое разнообразие структурных форм. От конечной конформации РНК зависят и ее функции, которые намного шире, чем быть просто посредником между ДНК и белком, только два процента генома кодирует мРНК, тогда как большая часть транскрибируется в “некодирующие РНК” [1]. “Некодирующие РНК” выполняют огромное количество функций: от регуляторных на уровне транскрипции и трансляции до каталитических.

Принятые сокращения: FRET – флуоресцентный резонансный перенос энергии; SAXS – метод малоуглового рентгеновского рассеяния; SHAPE – избирательное ацилирование по гидроксильной группе, анализируемое методом достройки праймера (selective 2-hydroxyl acylation analysed by primer extension); NMIA – N-метилизатоиевый ангидрид (N-methylisatoic anhydride).

* Эл. почта: ogalzit@vega.protres.ru

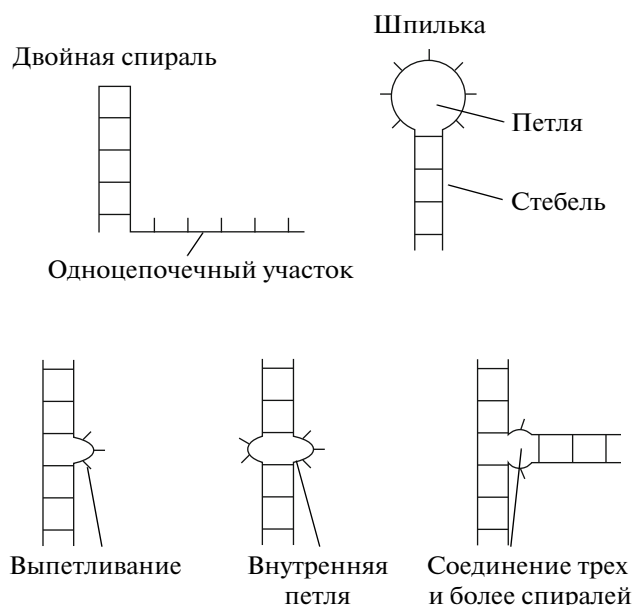


Рис. 1. Элементы вторичной структуры РНК.

Помимо рибосомной и тРНК к ним относятся: теломеразная РНК, входящая в состав фермента теломеразы и используемая в качестве матрицы для синтеза повторяющихся сегментов теломерной ДНК [2, 3]; малые ядерные РНК, регулирующие процессы транскрипции мРНК и процессы посттранскрипционных модификаций предшественников мРНК и тРНК [4, 5]; рибозимы, выполняющие каталитическую функцию [6, 7]; малые интерферирующие РНК, которые вовлечены в защиту клетки от вирусов, репрессию транскрипции, регуляцию некоторых генов и формирование центрального гетерохроматина, более того, система РНК-интерференции определяет эмбриональное развитие и клеточную дифференцировку [8, 9]; РНК-переключатели, которые могут включать или выключать как дальнейшую транскрипцию, так и трансляцию мРНК [10, 11], и многие другие [1]. И все же, функции многих “некодирующих РНК” пока остаются неизвестными [1].

Главное отличие процесса сворачивания РНК от белков состоит в том, что для РНК, прежде чем она приобретает пространственную структуру, необходимо установить нативо-подобные контакты, связанные со вторичной структурой [12–15]. В то время как в белковых структурах часть контактов, связанных с третичной структурой, может возникать еще до полного образования вторичной структуры.

Проводя аналогию между РНК и белками, можно отметить, что для тех и других справедлив парадокс Левинтала, согласно которому они находят нативную структуру без перебора всех возможных конформаций полипептидной и нуклео-

тидной цепей [16]. В процессе сворачивания цепь РНК, как и белковая глобула, проходит через множество промежуточных состояний, которые могут играть ключевую роль в кинетике процесса. Энергетический ландшафт молекул белков и РНК, отражающий свободную энергию как для переходных состояний, так и для интермедиатов сворачивания, содержит в себе много схожих по энергиям переходов. Вследствие того, что молекула РНК состоит из комбинации только 4 нуклеотидов, а белок — из 20 аминокислотных остатков, можно предположить, что разнообразие переходных состояний у белков существенно больше, чем у РНК. Однако число возможных путей сворачивания для молекул РНК сравнимо с аналогичным числом путей для белков сравнимого размера.

ОСОБЕННОСТИ ОРГАНИЗАЦИИ ПЕРВИЧНОЙ, ВТОРИЧНОЙ И ТРЕТИЧНОЙ СТРУКТУР РНК

Как и у белка, у РНК можно выделить три фундаментальных уровня организации: первичный, вторичный и третичный.

Первичная структура биологически активной РНК образуется после многочисленных посттранскрипционных модификаций и часто отличается от первичного транскрипта. Эти модификации включают в себя метилирование нуклеотидов по 2'-гидроксильной группе в рибозе, образование модифицированных оснований, таких как псевдоуридил и дигидроуридил, вырезание или вставки нуклеотидов. Поэтому, чтобы правильно определить первичную структуру РНК, необходимо опираться на данные секвенирования и масс-спектрометрии биологически активной РНК [17].

Вторичная структура РНК включает такие элементы как: шпильки, спирали, симметричные и асимметричные выпетливания, стыки или соединения трех и более спиралей, внутренние петли (рис. 1). Несмотря на то, что молекулы РНК построены из одной нуклеотидной цепи в ней имеются комплементарные участки друг другу, которые могут взаимодействовать, образуя двойные спирали.

Третичная структура образуется за счет взаимодействия между собой элементов вторичной структуры. До недавнего времени было мало известно об этом уровне организации РНК, пока не научились получать кристаллы для дальнейшего изучения методом рентгеноструктурного анализа и проводить измерения довольно крупных макромолекул методом ядерного магнитного резонанса (ЯМР). Так получили структуры с высоким разрешением отдельных молекул РНК, таких как тРНК, домена Р4–Р6 интрона группы I рибозима из *Tetrahymena thermophila* и др. Особое внимание

стоит обратить на структуры псевдоузлов, которые нередко появляются в молекулах РНК [18] (рис. 2). Эти структуры важны для выполнения РНК своих функций, однако почти во всех компьютерных алгоритмах предсказания структуры РНК псевдоузлы во внимание не принимают, так как это заметно усложняет задачу.

Замечено, что при малых концентрациях Mg^{2+} молекулы тРНК принимают некомпактную структуру, которая содержит элементы вторичной структуры, отличные от “клеверного листа” [13]. Следовательно, для формирования и стабилизации третичной структуры необходимы двухвалентные ионы металлов, особенно ионы магния: “РНК в природе существует, в основном, в виде Mg^{2+} -соли” [19].

Тема взаимодействия ионов металла с РНК представляет собой большой научный интерес. Много статей посвящено различным экспериментальным и теоретическим подходам, которые иногда противоречат друг другу. Например, существует точка зрения, согласно которой, “в отличие от двухцепочечных ДНК, конденсация противоионов на тРНК практически не происходит”, а отрицательно заряженные фосфатные группы на тРНК индуцируют заряды на аминокислот-тРНК-синтетазе, благодаря чему фермент узнает свою тРНК [20]. Молекулы одноцепочечной РНК также использовали для деконденсации ДНК, связанной с такими лигандами как спермидин, принимая во внимание тот факт, что РНК имеет большую константу связывания с лигандами [21].

ЭКСПЕРИМЕНТАЛЬНОЕ ОПРЕДЕЛЕНИЕ ПУТЕЙ СВРАЧИВАНИЯ РИБОЗИМОВ

В связи с тем, что функции РНК зависят от ее конформации, как и у белков, для изучения процесса ее сворачивания стали успешно применять методы, разработанные для белков, такие как Ф-анализ [22], флуоресцентный резонансный перенос энергии (FRET) [23], метод малоуглового рентгеновского рассеяния (SAXS) [24]. Специально для РНК разработали метод SHAPE (Selective 2'-Hydroxyl Acylation analyzed by Primer Extension), позволяющий обнаружить подвижные нуклеотиды в РНК любого размера [25].

Все перечисленные методы стали успешно применять для исследования формирования сложной пространственной структуры РНК. Чтобы изучить образование переходного состояния на пути сворачивания РНК, Ф-анализ используют совместно с методом FRET. Подобных исследований пока проведено немного, сложность состоит в том, что для РНК, как и для белков, характерен множественный путь сворачивания и, соответственно, большое количество переходных состояний. Ниже мы опишем работы по исследо-

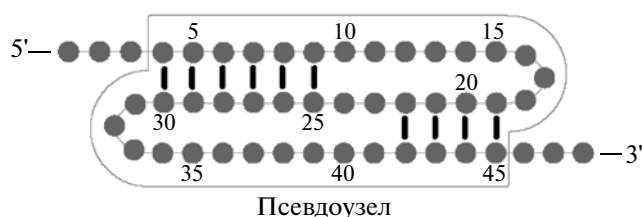


Рис. 2. Упрощенная структура псевдоузла.

ванию сворачивания различных РНК на основании перечисленных выше методов. Мы остановимся чуть более подробно на рассмотрении нескольких тРНК и домена Р4–Р6 интрона группы I рибозима из *Tetrahymena thermophila*. Особенность данных молекул – они либо сами выполняют важнейшие функции в клетке, как тРНК, либо они являются частью более крупного рибозима, как в случае домена Р4–Р6.

Домен Р4–Р6 интрона группы I рибозима из *Tetrahymena thermophila* – это самосплайсирующий интрон предшественника рибосомной РНК инфузории *Tetrahymena thermophila*. Данный рибозим отнесен к группе I интронов на основании наличия нескольких коротких консервативных элементов последовательности, определяющих вторичную структуру [26]. Рибозим состоит из девяти спаренных спиральных участков Р1–Р9 [26].

Изучен вклад участка Р1 в образование третичной структуры целого рибозима интрона группы I РНК [27]. Структура Р1 представляет собой дуплекс, построенный из шести пар оснований. Считают, что присоединение Р1 к остальной, уже сформированной, молекуле рибозима происходит в самую последнюю очередь, и поэтому является лимитирующей стадией для образования функционирующей молекулы. Эксперимент проводили следующим образом: последовательно делали восемь искусственных модификаций химических групп в дуплексе Р1, далее методом FRET рассматривали влияние этих мутаций на правильное пристраивание Р1 к коровой части рибозима. Модификации в дуплексе Р1 осуществляли, в основном, в виде замен 2'-гидроксила рибоз на водород (при этом получали 2'-дезоксирибозу) или на метильную группу – OCH_3 (продукт такой замены – 2'-метоксирибоза). В случае неканонической пары GU (расположена в области сайта щепления) производили замену аминокислоты G на водород с целью преобразования G в I (инозин). В другом эксперименте, с этой же парой GU, отдельно делали замену U на C с целью образования канонической пары GC в этом месте. Присоединение дуплекса Р1 к коровой части молекулы рибозима характеризовало значение сигнала $FRET \geq 0.75$, для неприсоединенного дуплекса Р1 это значение составляло ≤ 0.5 . Пользуясь

этими данными, можно вычислить константу равновесия при связывании K_{dock} , которая равна отношению времен, проведенных молекулой в полностью свернутом и свободном состояниях. Как утверждают авторы, все вышеперечисленные модификации приводили к уменьшению констант равновесия в 500 раз, а констант присоединения меньше чем в два раза. Чтобы рассчитать Φ -величины для замененных остатков в Р1, используют константы скорости присоединения и равновесия k_{dock} и K_{dock} соответственно. При этом считают, что в начальном состоянии Р1 не присоединен к остальной молекуле рибозима и находится как бы в свободном состоянии. Формулу для расчета Φ -величин в этом случае записывают следующим образом:

$$\Phi = \frac{\Delta\Delta G(k_{dock,modified} \rightarrow unmodified)}{\Delta\Delta G(K_{dock,modified} \rightarrow unmodified)} \quad (1)$$

$\Phi \sim 1$ означает, что данный остаток сформировал нативные контакты со своим окружением до переходного состояния, а $\Phi \sim 0$, говорит о том, что остаток еще не сформировал нативных контактов. С помощью Φ -анализа определили, что Р1-дуплекс присоединяется к каталитическому кору в последнюю очередь ($\Phi = 0$). Главным выводом статьи было предположение о том, что у больших РНК в нативном состоянии возможно наличие локально развернутых участков.

Теперь чуть более подробно остановимся на результатах, полученных для домена Р4–Р6 рибозима интрона группы I из *Tetrahymena thermophila*. В соответствии с экспериментальными данными, доступными для домена Р4–Р6, выяснили, что его сворачивание является высококооперативным, также показано, что контакты, соответствующие третичной структуре, в отдельно взятом домене Р4–Р6 образуются в два раза быстрее, чем в том же самом домене, но в составе целого рибозима [28]. Авторы этой работы исследовали сворачивание 160 н. домена Р4–Р6 рибозима интрона группы I из *Tetrahymena thermophila*. Интересно объяснение причины выбора авторами такого объекта для исследования, как домен Р4–Р6. Во-первых, в то время самым хорошо изученным объектом была тРНК. По сравнению с ней домен Р4–Р6 существенно больше (160 н.), поэтому исследование сворачивания крупных молекул является следующим шагом для изучения самоорганизации рибонуклеиновых кислот. Во-вторых, домен Р4–Р6 можно изучать отдельно от целой молекулы, так как этот домен не нуждается в контроле своего сворачивания со стороны остального мультидомена. Цель рассматриваемой работы – определение свободной энергии активационного барьера и нахождение положения переходного состояния при сворачивании домена Р4–Р6. Чтобы проследить за формированием равновесных

образований третичной структуры, участки Р4–Р6 поместили пиреновой меткой и далее, с помощью метода остановленного потока (*stop-flow*), наблюдали за кинетикой сворачивания при различных концентрациях Mg^{2+} . По изменению интенсивности флуоресценции после внесения различных концентраций Mg^{2+} измеряли наблюдаемую константу скорости для сворачивания k_{obs} . Далее рассчитали свободную энергию активации ΔG^\ddagger для Mg^{2+} -индуцированного сворачивания домена Р4–Р6 по формуле:

$$k_{obs} = (k_B T/h) \exp(-\Delta G^\ddagger/k_B T). \quad (2)$$

Так, при концентрациях $Mg^{2+} \sim 10$ мМ участки домена Р4–Р6 сворачиваются за миллисекунды с наблюдаемой константой скорости $k_{obs} = 15\text{--}31$ с⁻¹. При данной константе скорости выяснили, что изменение свободной энергии барьера находится в диапазоне $\Delta G^\ddagger \approx 8\text{--}16$ ккал/моль. Такой достаточно большой диапазон значений свободной энергии можно объяснить неопределенностью предэкспоненциального фактора, который используют в формуле для расчета отношений k_{obs} и ΔG^\ddagger . В данном случае применили уравнение Эйринга, в котором предэкспоненциальный множитель изначально предложен для реакций малых молекул, происходящих в газовой фазе. Однако молекула домена Р4–Р6 достаточно крупная, что ставит под сомнение использование множителя $(k_B T/h)$. Несмотря на это, все энергии активации у авторов были на первый взгляд достаточно достоверными, так как они по величине соответствуют величинам свободноэнергетического барьера при сворачивании в третичную структуру молекул белков и РНК (обычно это значение всего несколько ккал/моль).

Чтобы определить, какие нуклеотиды входят в ядро сворачивания и влияют на ход сворачивания домена Р4–Р6, специально для эксперимента ввели несколько мутаций в домен Р4–Р6 (рис. 3). По соотношению изменений свободных энергий в процессе сворачивания и стабильности определили значения Φ -величин. Несмотря на термодинамические отличия, значения скорости сворачивания домена Р4–Р6 с введенными в него мутациями и дикого типа оказались похожими. А значение $\Phi = \Delta\Delta G^\ddagger/\Delta\Delta G^o \sim 0$ указывает на то, что нативные контакты для данных нуклеотидов в переходном состоянии еще не сформированы.

Оказалось, что наблюдаемая константа скорости для сворачивания k_{obs} зависит от концентрации Mg^{2+} и значительно увеличивается при 1 мМ Mg^{2+} . При повышении концентрации ионов магния идет постепенное увеличение константы скорости сворачивания, которая выходит на плато при концентрации 30 мМ Mg^{2+} . Это можно объяснить тем, что в переходном состоянии молекула

по большей части остается развернутой, соответственно, для образования переходного состояния необходимо не больше одного иона магния на молекулу, а на стадии формирования дальних взаимодействий – уже большее количество ионов магния. Но дальнейшее увеличение концентрации магния не оказывает влияния на скорость сворачивания молекулы, так как происходит конформационная перестройка участков структуры, в которых уже связаны ионы магния.

В целом, авторы считают, что их одностадийное описание сворачивания рибозима не отображает полной картины, так как она не учитывает интермедиатов сворачивания, и объясняют это тем, что пока они не обладают достаточными данными для визуализации всех стадий процесса. Пока неизвестно, как устроена молекула в развернутом состоянии, какого рода взаимодействия необходимо разрушить в ней для начала сворачивания и достижения переходного состояния. Скорее всего, такие взаимодействия могут быть ненативными в отсутствие ионов Mg^{2+} . Большая величина активационного барьера не согласуется с практически развернутым переходным состоянием. Это скорее означает, что существуют денатурированные молекулы с остаточной структурой, не соответствующей нативной, которые нужно развернуть при сворачивании в уникальную структуру.

Шпилечный рибозим вируса табака с кольцевыми точками (tobacco ringspot virus) катализирует реакцию разрезания и сшивки продуктов репликации РНК вируса по типу катящегося кольца. Рибозим, сворачивание которого изучено в работе [29], состоит из двух спиральных доменов (спираль–петля–спираль) А и В, которые могут находиться в вытянутой (развернутой) конформации или компактной (свернутой), ферментативно-активной форме (рис. 4). Его каталитический центр образован за счет сближения двух доменов и взаимодействия между имеющимися в них петлями А и В, что показано методом FRET. Этим методом измерены скорости сворачивания и разворачивания фрагмента. Оказалось, что константа скорости сворачивания составляла 0.018 c^{-1} и была одинакова для всех 760 траекторий сворачивания. При денатурации значения констант скоростей разворачивания разделили на 4 группы: $k_1 = 0.001\text{ c}^{-1}$, $k_2 = 0.1\text{ c}^{-1}$, $k_3 = 0.8\text{ c}^{-1}$ и $k_4 = 6\text{ c}^{-1}$. В некоторых случаях молекула разворачивалась по нескольку часов, а в некоторых совсем не развернулась.

Чтобы выделить контакты, которые образовались в переходном состоянии шпилечного рибозима, авторы работы использовали Ф-анализ [22] и рассмотрели основания, участвовавшие в образовании рибозного замка.

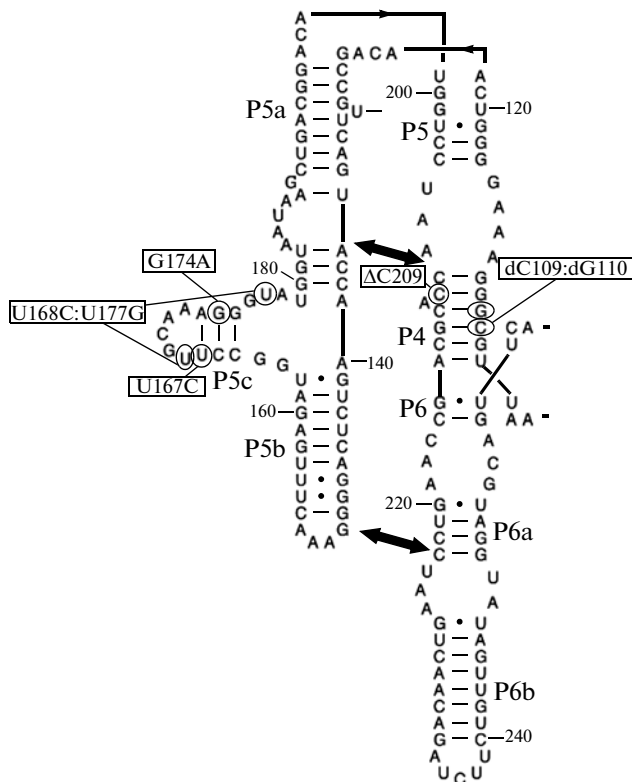


Рис. 3. Схематичное изображение вторичной структуры Р4–Р6 домена с указанием мест проведенных мутаций.

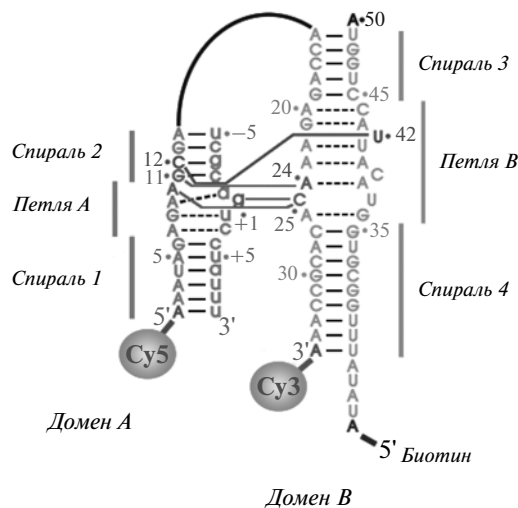


Рис. 4. Схематичное изображение свернутого фрагмента шпилечного рибозима, помеченного флуорофорами. Сплошными линиями показаны контакты, соответствующие третичной структуре, между парами оснований: g+1-C25; G11-A и A11-C25 (рибозный замок); C12-U42.

Мутация dG11 при стандартных условиях увеличила значения константы разворачивания k_U в 12.3 раз, а константы сворачивания k_N – только в

2.9 раз. Значения Φ -величин оказались невысокими. Это означает, что рибозный замок между основаниями G11, A10 и A24, C25 в доменах А и В соответственно сформирован частично (рис. 4). Мутация U42 (dC12) увеличила k_U в 4.4 раза, а k_N – только в 1.7 раз. Аналогично значения Φ -величин были невысокими, а, следовательно, контакты между U42 и C12 доменов В и А также сформированы частично. Последняя мутация пары g + 1 : C25, которую заменили на более слабую a + 1 : U25, увеличила k_U в 34 раза, при этом не изменила показания k_N . Φ -величины при этом равны нулю, это означает, что контакты на момент переходного состояния не сформированы совсем. Такая серия экспериментов позволила выделить нативные контакты, уже образовавшиеся в ядре сворачивания шпилечного рибозима. И хотя уотсон-криковская пара g_1 : C25 еще не сформировалась в переходном состоянии, контакты, образующие рибозный замок, уже возникают. Следует отметить, что образование контактов, соответствующих третичной структуре, в переходном состоянии не наблюдали при исследовании сворачивания домена P4–P6 рибозима интрона группы I из *Tetrahymena thermophila* [28]. Сворачивание данного фрагмента шпилечного рибозима изучалось при различных концентрациях магния. В результате обнаружено, что повышение концентрации магния значительно понижает энергию активации переходного состояния. Сделан вывод, что электростатические взаимодействия делают молекулу в переходном состоянии более компактной и близкой к нативной структуре. Концентрация магния в стандартных условиях приравнивалась к 12 мМ Mg^{2+} . Оказалось, что при увеличении концентрации магния значения Φ -величин практически не менялись.

Рибонуклеаза Р РНК (RNase P RNA) – это молекула РНК с каталитическими свойствами, отвечающая за процессинг тРНК. Авторы работы [14] изучали образование интермедиатов у S-доменов рибонуклеаз Р РНК *Bacillus subtilis* и *Escherichia coli*, которые принадлежат к разным классам рибонуклеаз Р и отличаются последовательностями на концах (рис. 5). Исследование проводили с помощью методов SAXS и молекулярной динамики Ланжевена (LD). Анализировали две модели интермедиатов. Вытянутость структуры моделей достигали за счет электростатических взаимодействий (первая) и дополнительных связей между нуклеотидными остатками (вторая). На первой модели изучали влияние электростатических взаимодействий на процесс сворачивания, тогда как на второй – влияние специальных контактов в структуре РНК на процесс сворачивания. Ранее авторы обнаружили, что, в основном, равновесные интермедиаты 154 остатков S-домена *Bacillus subtilis* имеют более вытянутую форму, чем

нативная структура [14]. В последней работе авторы попытались выяснить, что влияет на вытянутость формы интермедиатов. С помощью метода SAXS авторы описали структуру интермедиата, которая состоит из трех спиральных доменов, соединяющихся в центре структуры. Первый и второй домены стыкуются друг с другом под углом $\sim 180^\circ$, тогда как в нативной структуре – под углом $\sim 60^\circ$. Третий домен расположен перпендикулярно по отношению к первым двум как у интермедиатов, так и у нативной структуры. Исследование проводили при различных концентрациях моновалентных катионов. При этом конформационные изменения были незначительны, а первый и второй домены – вытянуты. Кроме того, авторы проанализировали изменение конформаций при добавлении Mg^{2+} с учетом того факта, что бивалентные ионы взаимодействуют гораздо сильнее, чем моновалентные [14]. Тем не менее, сильного изменения в конформации интермедиатов не произошло. Только при добавлении Mg^{2+} больше 10 мМ молекулы свернулись в нативные структуры. Подобное исследование провели для домена P4–P6 рибозима интрона группы I из *Tetrahymena thermophila*, результат оказался похожим – при концентрациях Mg^{2+} меньше 10 мМ интермедиаты оставались вытянутыми. Сделан вывод о том, что вытянутую форму интермедиатов нельзя объяснить только с точки зрения электростатики. При рассмотрении пяти LD-траекторий большая часть взаимодействий между остатками менялась, но при этом все три спиральных домена не изменяли своего положения в пространстве. Определено, что одна из пар оснований в структуре представлена ненативной парой оснований между C134 и G176, которая определяет вытянутость формы интермедиата. В нативной конформации C134 находится далеко от G176. Затем пару C134–G176 заменили на C–U. Измеряя радиус инерции R_g и распределение расстояний P_r , определили, что в отсутствие моновалентных ионов характер сворачивания мутанта и дикого типа очень похож. Но при добавлении 0.2 М NaCl R_g мутанта значительно уменьшился, и это привело к более компактной форме интермедиата. Наблюдения за S-доменом рибонуклеазы Р *E.coli* дали те же результаты. Все это привело к выводу, что вытянутая форма интермедиатов S-домена зависит от некоторых контактов в структуре, а не от влияния электростатического отталкивания [14].

ЭКСПЕРИМЕНТАЛЬНОЕ ОПРЕДЕЛЕНИЕ ПУТЕЙ СВРАЧИВАНИЯ тРНК В НАТИВНУЮ СТРУКТУРУ

Молекулы тРНК не очень различаются по нуклеотидной последовательности, их длина изменяется от 74 до 95 н., а для митохондриальных тРНК – от 50 до 60 [19]. Анализ последовательностей тРНК из

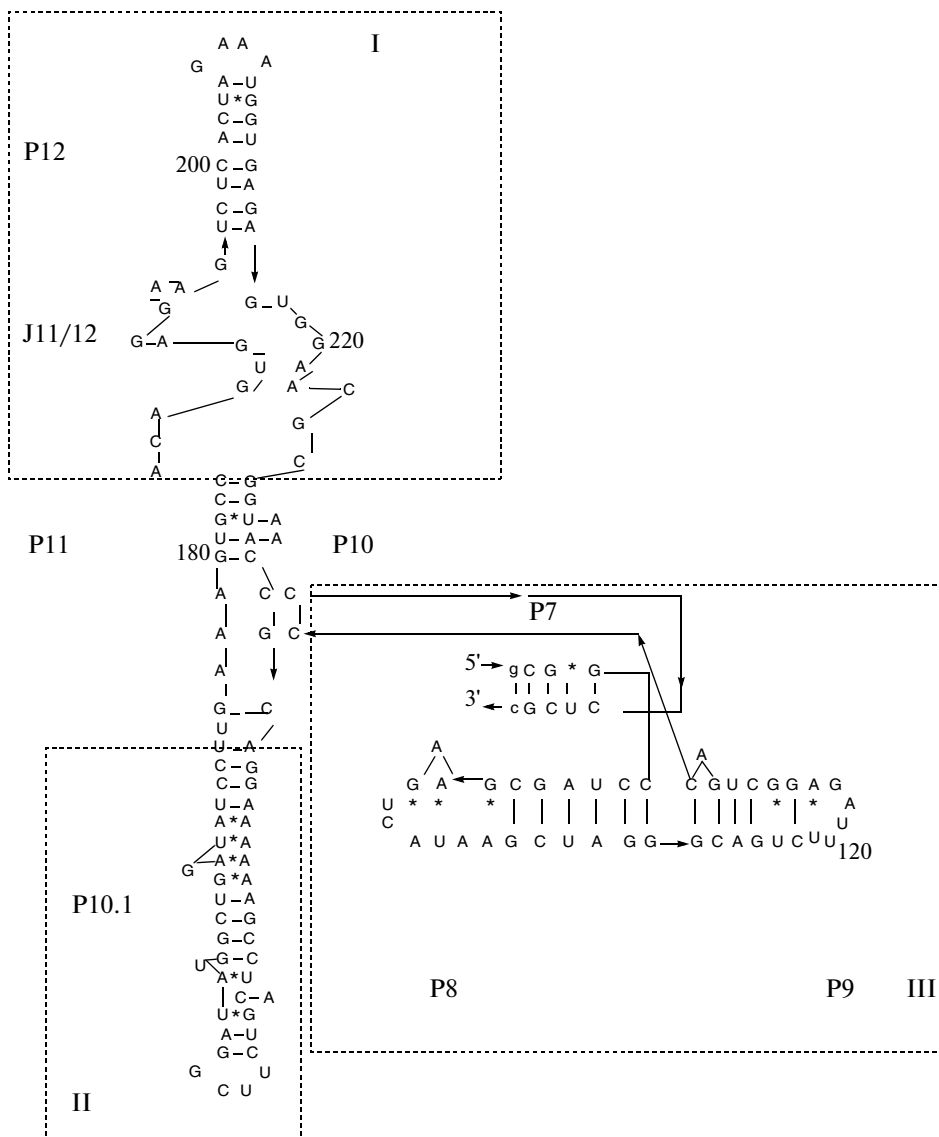


Рис. 5. Вторичная структура S-домена рибонуклеазы Р РНК *B. subtilis*.

организмов, принадлежащих ко всем царствам природы (или, как сейчас принято писать в литературе, доменам жизни), выявил консервативные участки в молекулах тРНК. К ним относятся: U8, A14, мотив GG в петле D, U33, несколько остатков в петле T и последовательность CCA на 3'-конце молекулы. Таким образом, консервативные участки в молекулах тРНК различных организмов сохраняются, чтобы принять участие во взаимодействии с другими белками и рРНК. Есть места, где расположены только пуриновые, или только пиримидиновые основания, их называют полуконсервативными остатками. Еще одной из ярких особенностей тРНК является присутствие нескольких модифицированных (минорных) оснований. Они образуются в определенных позициях тРНК в результате посттранскрипционных

модификаций с помощью специальных ферментов. Наиболее часто встречаются: риботимидин (T), псевдоуридин (Ψ), 5,6-дигидроуридин (D), инозин (I) и другие.

Вторичная структура у всех тРНК похожа на “клеверный лист” и образована [19]:

1. Дигидроуридиловой (D) шпилькой, содержащей дигидроуридиловые остатки, которые различны по длине у разных тРНК;
2. Тимидил-псевдоуридиловой (ТΨ) шпилькой, имеющей постоянную последовательность в петле GTΨCGA или GTΨCAA;
3. Антикодоновой (AC) шпилькой, где для всех тРНК характерно наличие ровно 7 нуклеотидов в петле;

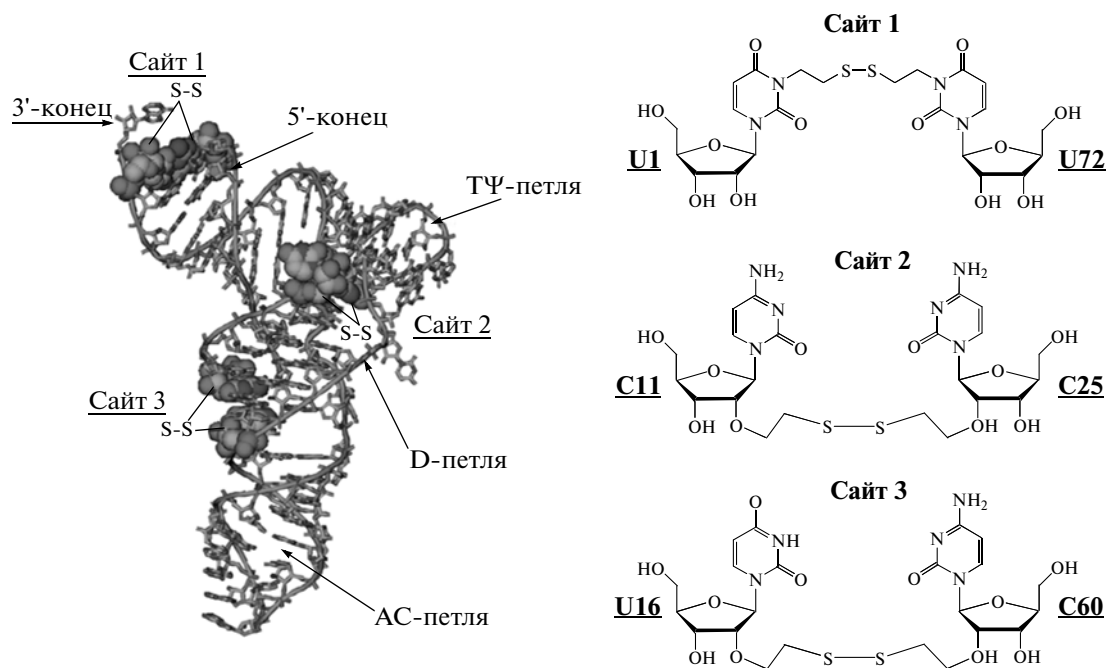


Рис. 6. *a* – Молекула тРНК^{Phe} с обозначенными на ней дисульфидными связями. Сайт 1 – дисульфидная связь между U1 и U72 нуклеотидами; сайт 2 – между C11 и C25; сайт 3 – между U16 и C60. *б* – Детальный рисунок дисульфидных связей.

4. Акцепторным (AA) стеблем, на 3'-конце которого имеется универсальная последовательность CCA;

5. Вариабельной (V) петлей, сильно отличающейся по длине у разных тРНК.

Все спаренные участки шпильек образуют двойную спираль А-формы. Формирование третичной структуры происходит за счет взаимодействия элементов вторичной структуры, образуя L-форму (рис. 6*a*).

Одна из первых экспериментальных работ по определению ядер сворачивания для молекул РНК посвящена изучению переходного состояния в процессе разворачивания тРНК с использованием Ф-анализа [30]. Искусственно синтезированную тРНК^{Phe} с немодифицированными основаниями выбрали для эксперимента. В основу лег метод, предложенный Кларке и Фёрштом, изучения белков с помощью внесения дополнительных дисульфидных связей [22]. В зависимости от места внесения дисульфидных связей можно как дестабилизировать переходное состояние, уменьшив при этом скорость разворачивания, так и стабилизировать переходное и нативное состояния, не изменяя при этом скорость разворачивания. Соответственно, если дисульфидным мостиком будет связана пара, которая рвется до или во время переходного состояния в процессе разворачивания, то скорость разворачивания уменьшится. Дисульфидная связь между парой, образованной по-

сле переходного состояния в процессе разворачивания, не изменит скорость разворачивания. Поэтому, чтобы выявить, какие нативные контакты образуются в последний момент сворачивания, авторы работы провели несколько денатураций тРНК^{Phe} [30]. В первый раз дисульфидная связь была образована между нуклеотидами U1 и U72 в AA-черешке (рис. 6*б*, сайт 1). Во второй раз – между нуклеотидами C11 и C25 D-шпильки, которая удерживала вторгшуюся V-петлю (рис. 6*б*, сайт 2). И, наконец, в последний раз дисульфидным мостиком были связаны петли U16-D и C60-TΨC. Каждый раз вычисляли разницу свободной энергии $\Delta G_{F-\#}$ между свернутой молекулой тРНК и переходным состоянием. Затем сравнивали значения $\Delta G_{F-\#}$ молекулы с дисульфидной связью с величиной $\Delta G_{F-\#}$ молекулы без этой связи, т.е. дикого типа. Разворачивание проводили при температуре 20–35°C (ниже температуры плавления тРНК). Результаты эксперимента показали, что при введении дисульфидной связи в AA-черешок и D-шпильку значения $\Delta G_{F-\#}$ не меняются. При внесении дисульфидной связи между D- и TΨC-петлями разница между $\Delta G_{F-\#}$ составляла около 0.80 ккал/моль. При этом разворачивание тРНК с дисульфидной связью проводили в две фазы, первая из которых зависела от температуры и соответствовала плавлению третичной структуры, а вторая не зависела от температуры и отображала небольшие конфигурационные изменения после плавления третичной

структуры. На основании полученных данных авторы сделали вывод, что нативные контакты между D- и ТΨ-петлями образуются в последнюю очередь при сворачивании тРНК^{Phe}. Возможно, это связано с сильным отталкиванием между отрицательно заряженными фосфатными группами. Более того, по мнению авторов, подобный порядок образования третичной структуры может быть свойствен и другим молекулам тРНК.

Метод FRET использовали при изучении инициации обратной транскрипции генома ВИЧ [31]. Затравкой для синтеза ДНК служит одна из акцепторных цепей тРНК^{Lys}. Соответственно, для гибридизации этой нити с участком праймера вирусной РНК AA-черешок необходимо развернуть. С помощью флуоресцентных красителей на 3'- и 5'-концах тРНК^{Lys} показано, что нуклеокапсидный белок вируса значительно ускоряет денатурацию акцепторных цепей. Но при этом обязательно присутствие самого генетического материала вируса (рис. 7), так как в отсутствие его нуклеокапсидный белок не работает. Следовательно, денатурация AA-черешка происходит одновременно с гибридизацией вирусной РНК.

До недавнего времени считали, что РНК сворачивается строго иерархично. Предполагали, что вначале формируются энергетически стабильные элементы вторичной структуры, которые включают образование пар нуклеиновых оснований. Затем происходит формирование элементов третичной структуры РНК [32, 33]. Авторы работы [15] считают, что эта модель сворачивания РНК неточна, так как им удалось с помощью специально разработанного нового метода SHAPE показать, что часто потеря третичных контактов происходит одновременно с разрывом пар оснований элементов вторичной структуры для искусственно синтезированной тРНК^{Asp} без модифицированных оснований.

Метод SHAPE позволяет получить информацию о локальном окружении конкретного нуклеотида другими компонентами молекулы тРНК, как функциональную зависимость от температуры. С этой целью используют способный ацилировать 2'-гидроксил группу нуклеотида РНК структурно-чувствительный реагент NMIA (N-methylisatoic anhydride), который не участвует в образовании вторичной и третичной структур. Реакционная способность 2'-гидроксил-группы может снизиться и под влиянием соседнего аниона 3'-фосфодиэфира (рис. 8). Так, свободные нуклеотиды вступают в данную реакцию более охотно, чем те, конформационная подвижность которых затруднена образованием пары с другими нуклеотидами. С увеличением температуры флуктуационная подвижность молекулы тРНК выросла, некоторые из ее нуклеотидов освободились из пары и реагировали с NMIA.

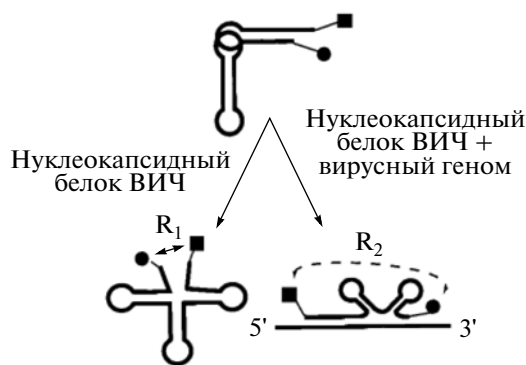


Рис. 7. Схематичное изображение тРНК, помеченной флуорофорами. R₁ и R₂ – расстояния между флуорофорами. Разворачивание AA-черешка не происходит (слева). Разворачивание AA-черешка происходит одновременно с гибридизацией с вирусной РНК (справа).

Основная задача метода SHAPE – обеспечить взаимодействие исследуемой РНК с NMIA, после чего с помощью обратной транскриптазы запустить синтез кДНК (рис. 9). Транскрипция будет обрываться именно на тех нуклеотидах цепочки РНК, где произошла модификация рибоз реагентом NMIA по 2'-положениям.

Исследования проводили при различных температурах от 35–75°C с интервалом ≤3.5°C. При каждой температуре тРНК^{Asp} обрабатывали NMIA и затем с помощью удлиняющегося праймера (Primer Extension) определяли подвижные нуклеотиды. При этом обнаружили пять групп:

1. Первая группа представляет 10 позиций, включающих в себя все остатки ТΨ-шпильки, которые остаются несвязанными с NMIA при всех температурах.
2. Вторая группа – это 10 нуклеотидов тРНК^{Asp} (G18, U19, C20, U33, G34, U35, G37, C38, C75 и A76), которые, наоборот, реагируют с NMIA при всех температурах. Эти нуклеотиды расположены в основном в D-петле и 3'-конце акцепторного стебля.
3. В третьей группе реактивность нуклеотидов постепенно увеличивалась с ростом температуры.
4. В четвертой группе некоторые нуклеотиды в D- и ТΨ-петлях реагировали при температуре 35°C с NMIA и резко увеличивали свою активность с ростом температуры.
5. И, наконец, пятая группа – это 3 позиции, которые реагировали с NMIA с самого начала, но потом изменяли свою реакционную способность. Например, G73 сначала увеличивал свою реакционную способность с ростом температуры, но потом уменьшал.

Чтобы интерпретировать полученные результаты и применить их для описания процесса сво-

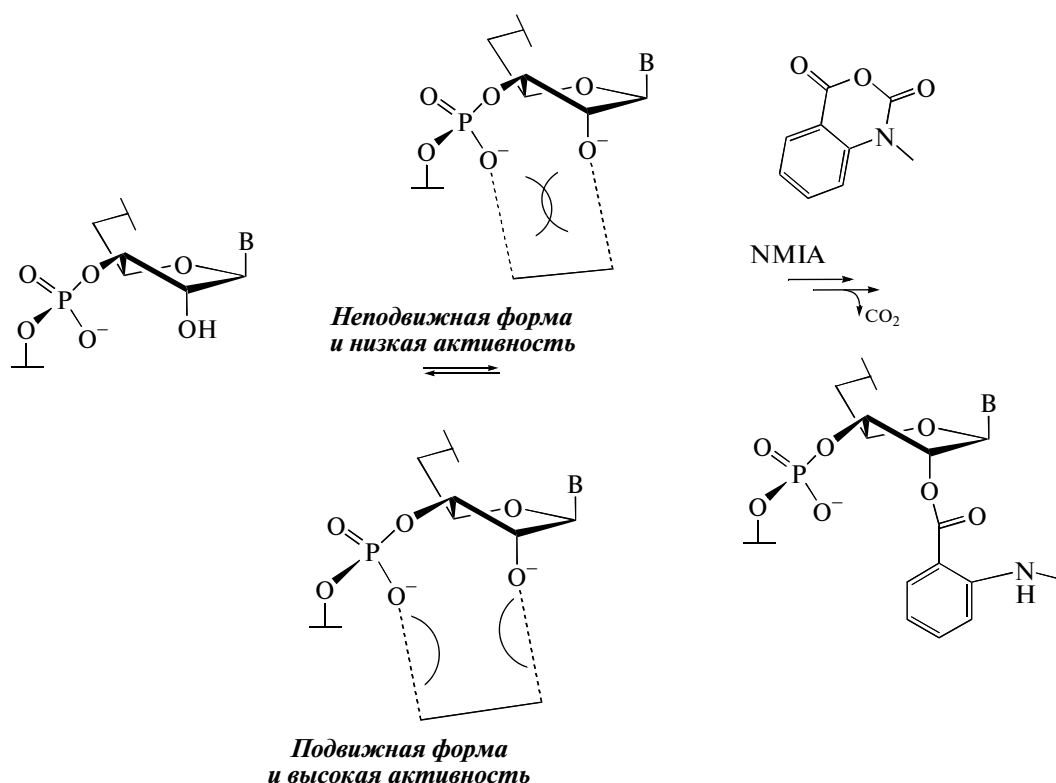


Рис. 8. Реакция 2'-гидроксила с NMIA [15]. Так как рибоза может находиться в различных конформациях, то при конформации, когда 3'-фосфодиэфир находится близко к 2'-гидроксила, он не вступает в химическую реакцию с NMIA, а при конформации, когда 3'-фосфодиэфир находится дальше, — вступает.

рачивания интермедиатов для тРНК^{Asp} реакционную способность при разворачивании нуклеотидов разделили на 8 групп по температуре плавления, при этом нуклеотиды каждой группы оказались в соседних участках тРНК^{Asp}.

При температуре 35°C данные, полученные с помощью метода SHAPE, соответствовали параметрам канонической структуры тРНК^{Asp}, при которой нуклеотиды во всех петлях и на 3'-конце акцепторного черешка и одна пара (G–U) в антикодонной шпильке реагировали с NMIA, в то время как нуклеотиды, вовлеченные в образование вторичной и третичной структур, не реагировали с NMIA.

На втором этапе при температуре 51°C сначала разрушались третичные взаимодействия между D- и ТЧ-петлями и связь между A9 и U13, что соответствует вторичной структуре и взаимодействию между пиримидин-богатými участками цепи G10–U13.

На третьем этапе при температуре 53°C полностью развернулся стебель D и разрушились стэкинг-взаимодействия между G6, A7 и C49, принадлежавшими ТЧ и акцепторному черешку.

Так же исчезли третичные взаимодействия между пурином-богатými участками (G22–U25).

При каждом повышении температуры данные, полученные методом SHAPE, показали сложность взаимодействий между нуклеотидами в равновесных конформационных состояниях тРНК^{Asp}, которые не строго соответствуют модели иерархического сворачивания, как это считалось, а всего лишь частично. В работе показано, что шпильки могут плавиться отдельными участками, асимметрично, одновременно с этим одноцепочечные участки могут сдвигаться и образовывать новые пары. Иными словами, вся молекула тРНК может сворачиваться в альтернативные структуры при высоких температурах. На основании проделанного эксперимента авторы работы сделали вывод, что равновесные конформационные состояния тРНК^{Asp} сложно предсказать с помощью только иерархической модели сворачивания тРНК.

Можно надеяться, что при дальнейшем исследовании с помощью метода SHAPE-анализа удастся накопить необходимую информацию для определения взаимосвязи между нуклеотидами при сворачивании и для других молекул РНК.

ПРЕДСКАЗАНИЕ СТРУКТУР РНК

В теоретических исследованиях основное внимание уделяют как предсказанию вторичной и третичной структуры РНК, так и изучению кинетики сворачивания РНК, которая может быть описана в виде свободно-энергетического ландшафта. В данном параграфе мы рассмотрим алгоритмы для решения первой задачи более подробно.

Сворачивание молекул РНК и образование вторичной структуры имеют очень важное значение для выполнения РНК своих функций в клетке. Изучение вторичной структуры РНК началось сразу же после важного открытия уотсон-криковского спаривания между основаниями в 1953 г. В последние 25 лет важность компьютерных методов для решения проблем молекулярной биологии, включая сворачивание молекул РНК, возрастает.

Количество информации о структурах РНК стремительно увеличивается с каждым годом. Однако и число возможных вторичных структур РНК значительно возрастает при увеличении длины последовательности той или иной РНК. Например, для 16S рРНК, состоящей из около 1500 н., существует приблизительно 15000 возможных спиралей и только меньше 100 из них войдут в конечную структуру. Так была сделана оценка максимального количества комбинаторных перестановок всех возможных спиралей, которые в конечном итоге приведут к образованию совершенно различных структурных моделей для 16S рРНК, количеством 4.3×10^{393} [34]. Подобное утверждение раскрывает сущность парадокса Левинталя [16], который состоит в следующем: развернутый полимер (белок или РНК) из огромного числа возможных конформаций выбирает свою нативную структуру за доли секунды, в то время как по теоретическим подсчетам такой поиск должен был бы длиться больше времени жизни всей вселенной. Левинталь предположил, что нативная структура белка определяется не стабильностью, не термодинамикой, а кинетикой, т.е. она соответствует не глобальному, а быстро достижимому минимуму свободной энергии цепи. На самом деле, структура стремится не просто к глобальному минимуму, а к быстро достижимому минимуму свободной энергии [35].

Проблема пути сворачивания РНК, как и в случае с белками, до сих пор вызывает множество споров. Экспериментальные исследования крупных молекул полимеров проводятся по большей части в условиях *in vitro*, которые, как известно, не могут полностью моделировать среду в клетке, т.е. *in vivo*. До сих пор до конца неясно, что происходит в клетке. В компьютерных экспериментах пытаются воспроизвести все условия, при которых происходит сворачивание РНК в растворе. Однако очень часто условия, при которых про-

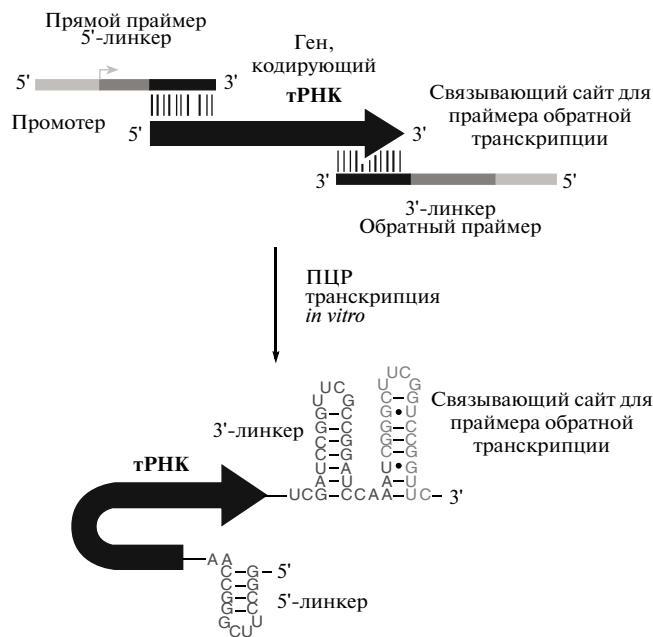


Рис. 9. РНК-транскрипт и “структурная кассета”, которая пришивается к РНК посредством транскрипции с матрицы, образованной с помощью ПЦР.

дится компьютерный эксперимент, сильно упрощены, и вводятся лишь некоторые термодинамические параметры. Подобные упрощения часто не влияют на чистоту эксперимента и используются для оптимизации и ускорения расчетов.

Все существующие на сегодняшний день алгоритмы предсказывают структуру либо одной молекулы РНК, либо структуру комплексов, например РНК–ДНК или рибозим–рибозим.

Вторичную структуру единственной нуклеиновой кислоты без псевдоузлов предсказывают следующие алгоритмы.

Алгоритм минимизации свободной энергии Цукера и Стиглера [36]. В качестве входного источника данных этот алгоритм использует последовательность РНК и с помощью метода динамического программирования производит поиск структуры с минимумом свободной энергии. Основа данного метода – модель ближайшего соседа. Усовершенствования скорости выполнения и оптимизации внесли в данный алгоритм Лингсё и соавт. [37]. Мэтьюс предложил для увеличения точности рассматривать до 750 структур в области энергетического минимума [38]. Программа *mfold* [39] Цукера и Стиглера доступна по адресу: <http://mfold.bioinfo.rpi.edu>.

Алгоритм расчета статистической суммы. Маккаскилл [40] предложил другой алгоритм динамического программирования для молекул РНК без псевдоузлов. В данном случае проводят расчет

множества структур и, далее, их статистической суммы:

$$Q = \sum_S e^{-\Delta G(S)/RT}, \quad (3)$$

где суммирование осуществляют по всем возможным структурам, в которые РНК может свернуться. Когда статистическая сумма Q будет рассчитана, можно вычислить вероятность той или иной структуры S :

$$P(S) = \frac{1}{Q} e^{-\Delta G(S)/RT}. \quad (4)$$

Однако для РНК в биологическом плане более характерен набор похожих структур (или их еще называют кинетической группой объектов [41]), между которыми постоянно с той или иной скоростью происходит обмен [40]. Поэтому в алгоритме упор делается на расчет вероятностей структур РНК, входящих в одну группу или класс похожих структур. Эти подструктурные классы важны, так как они позволяют отобразить наиболее характерные черты ансамбля структур. В конечном счете, может быть рассчитана равновесная вероятность появления каждой возможной пары оснований и для лучшей наглядности может быть составлено соотношение вероятности оснований и оптимальной структуры. Маккаскилл оценил этот метод на четырех природных РНК последовательностях с известными структурами. Он показал, что реальные пары оснований предсказаны с высокой, но не с максимально возможной, вероятностью. Алгоритм расчета статистической суммы включили в пакет программ Vienna [42].

Таким образом, можно сделать вывод о том, что у обоих алгоритмов минимизации свободной энергии и расчета статистической суммы есть два главных недостатка: 1) они не могут предсказать структуры с псевдоузлами; 2) они в основном используют упрощенные термодинамические модели.

В дополнение можно добавить, что существуют еще несколько подходов для предсказания структур РНК. Например, целый класс *алгоритмов сравнительного анализа*, разработанный в лаборатории Гутелла. Алгоритм сравнительного анализа основан на двух простых, но эффективных принципах [34]: 1) “различные РНК последовательности могут сворачиваться в одни и те же вторичные структуры”; 2) “уникальная структура и функция РНК созданы в ходе эволюционного процесса путем мутаций и отбора”. Результат предсказания обычно сравнивают с доступными закристиализованными структурами. Класс сравнительного анализа включает несколько методов, один из которых основан на выявлении структурных мотивов (предсказывает 70% п.н.).

Также известны случаи объединения термодинамической модели и сравнительного анализа. Хофакер [43] представил метод компьютерного расчета вторичной структуры для набора выровненных РНК последовательностей, используя одновременно термодинамическую стабильность и корреляцию последовательностей. Данная программа, известная под названием *RNAalifold*, доступна пользователям интернета в пакете программ Vienna.

МОДЕЛИРОВАНИЕ СВРАЧИВАНИЯ МОЛЕКУЛ РНК

Несмотря на простоту первичной структуры молекул РНК, они участвуют в большом количестве биологических процессов. Пространственная структура и сворачивание больших молекул РНК – предмет исследования многих ученых в последнее время [44].

Во вторичной структуре РНК контактов, образованных дальними участками по цепи, меньше, чем в белках [45]. Однако именно их поведение определяет характер перехода спираль–клубок для больших молекул РНК: спирали, образованные спариванием далеких по цепи участков, плавятся одновременно по механизму фазового перехода второго рода, в то же время спирали, образованные близкими по цепи участками, плавятся практически независимо [46].

Исследование кинетики самоорганизации гетерополимерных цепей показало, что скорость их сворачивания сильно зависит, как от внешних условий [47–50], так и от самой последовательности [51]. В случае белковых моделей показано [51–53], что геометрические свойства нативного состояния влияют на скорость сворачивания в это состояние и его стабильность. Так установлено, что нелокальные контакты важны для восстановления нативной структуры [53], а структуры белковых моделей, имеющие большее число нелокальных контактов, сворачиваются в стабильную структуру на два порядка быстрее, чем структуры с большим числом локальных контактов [51]. Естественно предположить, что для сворачивания больших молекул РНК, которым нужно преодолеть огромный перебор возможных структур, существуют геометрические факторы, ускоряющие это процесс.

В работах [54, 55] показано, что таким фактором в модели вторичной структуры РНК является наличие в нативной структуре сильных энергетических контактов, образованных удаленными участками по цепи. Показано, что “геометрически” отредактированные последовательности, в которых сделано соответствие между энергиями нативных контактов: самые сильные контакты образованы между самыми дальними участками

по цепи, находят свое нативное состояние в оптимальной области внешних параметров (температуры и эффективной силы взаимодействия между звеньями) на порядок быстрее, чем случайные последовательности.

Алгоритм с использованием дискретной молекулярной динамики, ранее применяемый для расчетов с белками, опробовали на РНК, он хорошо предсказывает как 3D-структуры, так и саму кинетику сворачивания [56]. Чтобы провести данным методом расчеты, структуру РНК упростили и рассмотрели ее как модель “бусины на нитке”. Суть модели в том, что полноатомная модель, полученная экспериментально, переводится в грубую модель (coarse-grained) таким образом, что каждому нуклеотиду соответствует 3 точки: P — центр масс фосфатной группы, практически совпадает с фосфором в PO₄ группе; S — центр масс рибозы; B — точка, соответствующая основанию, которая описывается центром масс шестичленного кольца в пуринах и пиримидинах. Модель протестировали на 153 последовательностях РНК, доступных из банка нуклеиновых кислот (NDB). При этом отбирали молекулы длиной от 10 до 100 н. (так как более короткие молекулы РНК не обладают достаточной третичной структурой). При этом в отобранных молекулах представлены основные виды третичной структуры: L-подобная тРНК, шпильки, псевдоузлы. Для каждой отдельной РНК сначала моделировали развернутое состояние, после чего запускали алгоритм обмена репликами при различных температурах.

Основной алгоритм расчета кинетики (пути) сворачивания РНК основан на методе дискретной молекулярной динамики [56]. Чтобы представить энергетический ландшафт пути сворачивания, применяли алгоритм обмена репликами. Энергетический ландшафт сильно изрезан: он состоит из “щероховатостей”, “клонов”, ведущих к энергетическим ямам, которые соответствуют стабильным состояниям молекулы или ее минимумам свободной энергии. Достижение глобального минимума ускоряется при увеличении температуры. А чтобы преодолеть энергетические барьеры на ландшафте, используют алгоритм обмена репликами. Суть его состоит в том, что одновременно запускается симуляция сворачивания множества одинаковых моделей исследуемой молекулы при разных температурах. Информацию, полученную от индивидуального моделирования сворачивания каждой молекулы, объединяют при помощи алгоритма Монте-Карло. Функция последнего состоит в том, что через определенные промежутки времени он запускает обмен температурами T_i и T_j между репликами i и j (каждая из которых обладает энергиями E_i и E_j). Обмен температурами осуществляется в рамках

алгоритма Монте-Карло по схеме Метрополиса с вероятностью обмена $p = 1$ если $\Delta = (1/k_B T_i - 1/k_B T_j)(E_j - E_i) \leq 0$, и $p = \exp(-\Delta)$, если $\Delta > 0$ [57]. Алгоритм обмена репликами используют для того, чтобы быстро заполнить конформационное пространство, доступные для РНК. В алгоритме для простоты используют 8 температур: 0.200, 0.208, 0.214, 0.220, 0.225, 0.230, 0.235, и 0.240 (температура дана в абстрактных единицах).

Рассмотренную выше модель успешно использовали для предсказания ядер сворачивания для тРНК молекул. Алгоритм для изучения процессов, включающих сворачивание/разворачивание белков [58], успешно применен для расчетов на тРНК [59]. Для четырех молекул тРНК, структуры которых получены в свободном состоянии (тРНК^{Phe}, тРНК^{Asp}, тРНК^{Met}, тРНК^{Lys}), рассчитаны профили Ф-величины. Эти профили согласуются с экспериментальными данными, свидетельствующими, что нуклеотидные остатки в области петель D- и T-шпилек вовлекаются в структуру тРНК последними или, точнее, не входят в ядро сворачивания тРНК. Высокие значения рассчитанных Ф-величин в области антикодоновой шпильки показывают, что там локализовано ядро сворачивания для тРНК.

Ввиду большого разнообразия биологических функций РНК понятно, что разные молекулы могут использовать различные пути сворачивания. Вопросы о том, какие элементы структуры лимитируют сворачивание разных типов РНК, и каким образом макромолекула выбирает кинетический путь сворачивания, остаются пока открытыми. Наша задача состоит в разработке как теоретических, так и экспериментальных методов определения таких структур и описания кинетических путей сворачивания. Рассматриваемые и обсуждаемые в этом обзоре данные необходимо принимать во внимание при рассмотрении процесса моделирования сворачивания молекул РНК.

Авторы благодарны рецензенту, А.К. Сурину и О.М. Селивановой за критические замечания и ценные советы при прочтении данной статьи.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (№ 11-04-00763), Российской академии наук (программы “Молекулярная и клеточная биология” (01200959110) и “Фундаментальные науки – медицине”), Федерального агентства по науке и инновациям (№ 16.512.11.2204).

СПИСОК ЛИТЕРАТУРЫ

1. Макарова Ю.А., Крамеров Д.А. 2007. Некодирующие РНК. *Биохимия*. **72**, 1427–1448.
2. Koziel J.E., Fox M.J., Steding C.E., Sprouse A.A., Herbert B.S. 2011. Medical genetics and epigenetics of telomerase. *J. Cell Mol. Med.* **15**, 457–567.

3. Caslini C. 2010. Transcriptional regulation of telomeric non-coding RNA: implications on telomere biology, replicative senescence and cancer. *RNA Biol.* **7**, 18–22.
4. Макарова Ю.А., Крамеров Д.А. 2007. Малые ядерные РНК. *Молекуляр. биология.* **41**, 246–259.
5. Taft R.J., Pang K.C., Mercer T.R., Dinger M., Mattick J.S. 2010. Non-coding RNAs: regulators of disease. *J. Pathol.* **220**, 126–139.
6. Zhang J., Lau M.W., Ferré-D'Amaré A.R. 2010. Ribozymes and riboswitches: modulation of RNA function by small molecules. *Biochemistry.* **49**, 9123–9131.
7. Ferré-D'Amaré A.R., Scott W.G. 2010. Small self-cleaving ribozymes. *Cold Spring Harb. Perspect. Biol.* **2**, a003574.
8. Peralta-Zaragoza O., Bermúdez-Morales V.H., Madrid-Marina V. 2010. RNA interference: biogenesis molecular mechanisms and its applications in cervical cancer. *Rev. Invest. Clin.* **62**, 63–80.
9. Ketting R.F. 2011. The many faces of RNAi. *Dev. Cell.* **20**, 148–161.
10. Mulhbach J., St-Pierre P., Lafontaine D.A. 2010. Therapeutic applications of ribozymes and riboswitches. *Curr. Opin. Pharmacol.* **10**, 551–556.
11. Baird N.J., Kulshina N., Ferré-D'Amaré A.R. 2010. Riboswitch function: flipping the switch or tuning the dimmer? *RNA Biol.* **7**, 328–332.
12. Thirumalai D., Hyeon C. 2005. RNA and protein folding: common and variations. *Biochemistry.* **44**, 13, 4957–4970.
13. Draper D. 1996. Parallel worlds. *Nature Struct. Biol.* **3**, 397–400.
14. Baird N.J., Gong H., Zaheer S.S., Freed K.F., Pan T., Sosnick T.R. 2010. Extended structures in RNA folding intermediates are due to nonnative interactions rather than electrostatic repulsion. *J. Mol. Biol.* **397**, 1298–1306.
15. Wilkinson K.A., Merino E.J., Weeks K.M. 2005. RNA SHAPE chemistry reveals nonhierarchical interactions dominate equilibrium structural transitions in tRNA^{Asp} transcripts. *J. Amer. Chem. Soc.* **127**, 4659–4667.
16. Levinthal C. 1968. Are there pathways for protein folding? *J. Chem. Phys.* **65**, 44–45.
17. Batey R.T., Rambo R.P., Doudna J.A. 1999. Tertiary motifs in RNA structure and folding. *Angew. Chem. Int.* **38**, 2326–2343.
18. Dost B., Han B., Zhang S., Bafna V. 2006. Structural alignment of pseudoknotted RNA. *Recomb. LNBI. Springer-Verlag Berlin Heidelberg.* **3909**, 143–158.
19. Спиринов А.С. 2011. *Молекулярная биология. Рибосомы и биосинтез белка.* М.: Издательский центр Академия, 496 с.
20. Ivanov V.V., Stepanenko V.A., Zrel'ov P.V., Zheltukhin E.I., Chirgadze Yu.N., Sivozhelezov V.S., Polozov R.V. 2009. Toward protein-DNA stereochemical recognition codes. In: *Mathematical Modeling and Computational Physics (MMCP'2009): Book of Abstracts of the International Conference.* Dubna, July 7–11.
21. Tsumoto K., Yoshikawa K. 1999. RNA switches the higher-order structure of DNA. *Biophys. Chem.* **82**, 1–8.
22. Fersht A. 1999. *Structure and mechanism in protein science.* N.Y.: W.H. Freeman and Company.
23. Förster T. 1948. Zwischenmolekulare energiewanderung und fluoreszenz. *Ann. Physik.* **437**, 55–75.
24. Свергун Д.И., Фейгин Л.А. 1985. Рентгеновское и нейтронное малоугловое рассеяние. М.: Наука.
25. Merino E.J., Wilkinson K.A., Coughlan J.L., Weeks K.M. 2005. RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). *J. Amer. Chem. Soc.* **127**, 4223–4231.
26. Федоренко Е.С., Иродов Д.М., Кордюм В.А. 1995. Свойства рибозима из tetrahymena Thermophila. *Биополимеры и клетка.* **11**, 20–29.
27. Bartley L.E., Zhuang X., Das R., Chu S., Herschlag D. 2003. Exploration of the transition state for tertiary structure formation between an RNA helix and a large structured RNA. *J. Mol. Biol.* **328**, 1011–1026.
28. Silverman S.K., Cech T.R. 2001. An early transition state for folding of the RNA P4-P6 RNA domain. *RNA Society.* **7**, 161–166.
29. Bokinsky G., Rueda D., Misra K., Rhodes M.M., Gordus A., Babcock H.P., Walter N.G., Zhuang X. 2003. Single-molecule transition-state analysis of RNA folding. *Proc. Natl. Acad. Sci. USA.* **100**, 9302–9307.
30. Maglott E.J., Goodwin J.T., Glick G.D. 1999. Probing the structure of an RNA tertiary unfolding transition state. *J. Amer. Chem. Soc.* **121**, 7461–7462.
31. Chan B., Weidemaier K., Yip W.T., Barbara P.F., Musier-Forsyth K.M. 1999. Intra-tRNA distance measurements for nucleocapsid protein-independent tRNA unwinding during priming of HIV reverse transcription. *Proc. Natl. Acad. Sci. USA.* **96**, 459–464.
32. Tinoco I.J., Bustamante C. 1999. How RNA folds. *J. Mol. Biol.* **293**, 271–281.
33. Privalov P.L., Filimonov V.V. 1978. Thermodynamic analysis of transfer RNA unfolding. *J. Mol. Biol.* **122**, 447–464.
34. Gutell R.R., Lee J.C., Cannone J.J. 2002. The accuracy of ribosomal RNA comparative structure models. *Curr. Opin. Struct. Biol.* **12**, 301–310.
35. Финкельштейн А.В., Птицин О.Б. 2005. *Физика белка: Курс лекций с цветными и стереоскопическими иллюстрациями и задачами.* М.: Книжный дом “Университет”.
36. Zuker M., Stiegler P. 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucl. Acids Res.* **9**, 133–148.
37. Lyngsø R.B., Zuker M., Pedersen C.N. 1999. Fast evaluation of internal loops in RNA secondary structure prediction. *Bioinformatics.* **15**, 440–445.
38. Mathews D.H., Sabina J., Zuker M., Turner D.H. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911–940.
39. Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucl. Acids Res.* **31**, 34063415.
40. McCaskill J.S. 1990. The equilibrium partition function and base pair binding probabilities for RNA secondary structure. *Biopolymers.* **29**, 1105–1119.
41. Andronescu M.S. 2000. *Algorithms for predicting the secondary structure of pairs and combinatorial sets of nu-*

- cleic acid strands. Master Science Thesis. Bucharest, Romania: Academy of Economic Studies. 115 p.*
42. Schuster P., Fontana W., Stadler P.F., Hofacker I.L. 1994. From sequences to shapes and back: a case study in RNA secondary structures. *Proc. Biol. Sci.* **255**, 279–284.
 43. Hofacker I.L., Fekete M., Stadler P.F. 2002. Secondary structure prediction for aligned RNA sequences. *J. Mol. Biol.* **319**, 1059–1066.
 44. Woodson S.A. 2010. Compact intermediates in RNA folding. *Annu. Rev. Biophys.* **39**, 61–77.
 45. De Gennes P.G. 1968. Statistics of branching and hairpin helices for the dAT copolymer. *Biopolymers.* **6**, 715–729.
 46. Гутин А.М., Галзитская О.В. 1993. Переход спираль в клубок в простейшей модели больших природных РНК: I. Учет только нативных спиралей. *Биофизика.* **38**, 84–92.
 47. Socci N.D., Onuchic J.N. 1994. Folding kinetics of proteinlike heteropolymers. *J. Chem. Phys.* **101**, 1519–1528.
 48. Galzitskaya O.V., Finkelstein A.V. 1995. Folding of chains with random and edited sequences: similarities and differences. *Protein Eng.* **8**, 883–892.
 49. Galzitskaya O.V., Finkelstein A.V. 1996. Computer simulation of secondary structure folding of random and “edited” RNA chains. *J. Chem. Phys.* **105**, 319–325.
 50. Sali A., Shakhnovich E., Karplus M. 1994. Kinetics of protein folding. A lattice model study of the requirements for folding to the native state. *J. Mol. Biol.* **235**, 1614–1636.
 51. Abkevich V.I., Gutin A.M., Shakhnovich E.I. 1995. Impact of local and non-local interactions on thermodynamics and kinetics of protein folding. *J. Mol. Biol.* **252**, 460–471.
 52. Leopold P.E., Montal M., Onuchic J.N. 1992. Protein folding funnels: a kinetic approach to the sequence-structure relationship. *Proc. Natl. Acad. Sci. USA.* **89**, 8721–8725.
 53. Go N., Taketomi H. 1978. Respective roles of short- and long-range interactions in protein folding. *Proc. Natl. Acad. Sci. USA.* **75**, 559–563.
 54. Galzitskaya O.V. 1997. Geometrical factor and physical reasons for its influence on the kinetic and thermodynamic properties of RNA-like heteropolymers. *Fold Des.* **2**(3), 193–201.
 55. Галзитская О.В. 1997. Влияние энергии дальних контактов на время поиска нативной структуры для РНК-подобных гетерополимеров. *Молекуляр. биология.* **31**, 488–491.
 56. Ding F., Sharma S., Chalasani P., Demidov V.V., Broude N.E., Dokholyan N.V. 2008. *Ab initio* RNA folding by discrete molecular dynamics: From structure prediction to folding mechanisms. *RNA Society.* **14**, 1164–1173.
 57. Metropolis N., Rosenbluth A.W., Rosenbluth M.N., Teller A.H., Teller E. 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087–1092.
 58. Galzitskaya O.V., Finkelstein A.V. 1999. A theoretical search for folding/unfolding nuclei in 3D protein structures. *Proc. Natl. Acad. Sci. USA.* **96**, 11299–11304.
 59. Переяславец Л.Б., Баранов М.В., Леонова Е.И., Галзитская О.В. 2011. Предсказание ядер сворачивания в молекулах тРНК. *Биохимия.* **76**, 299–308.