

===== БИОИНФОРМАТИКА =====

УДК: 577.214.5

## Короткие стемы в псевдоузловых структурах РНК

Баулин Е.Ф.<sup>\*1,2</sup>, Кориневская А.В.<sup>3</sup>, Тихонова П.О.<sup>3</sup>, Ройтберг М.А.<sup>1,2,3</sup>

<sup>1</sup>Лаборатория прикладной математики, Институт математических проблем биологии РАН – филиал Института прикладной математики им. М.В. Келдыша РАН,  
Пущино, Московская область, 142290, Россия

<sup>2</sup>Факультет инноваций и высоких технологий, Московский физико-технический  
институт (Государственный университет), Долгопрудный, Московская область,  
141700, Россия

<sup>3</sup>Факультет компьютерных наук, Национальный исследовательский университет  
«Высшая школа экономики», Москва, 101000, Россия

**Аннотация.** Данная работа посвящена анализу роли так называемых "коротких" стемов, т.е. стемов длиной в два или три спаривания оснований, в псевдоузловых структурах РНК. Как известно, задача предсказания вторичной структуры РНК, включая псевдоузлы произвольного типа, является NP-сложной задачей, что определяет актуальность поиска путей такого ограничения рассматриваемых типов псевдоузлов, которое позволит решать задачу за полиномиальное время. Псевдоузлы в экспериментально разрешенных пространственных структурах РНК более чем на половину состоят из коротких стемов, энергетическая нестабильность которых подтверждается как различными математическими моделями, так и косвенно экспериментальными данными. Так, например, согласно представленным данным, коротким стемам чаще соответствуют внутренние петли и реже – шпильки, чем в среднем по всем стемам. В работе показано, что при исключении из рассмотрения коротких стемов, разнообразие псевдоузлов сводится к двум базовым типам – H-узлы и Kissing Loops. Единственным исключением является псевдоузел, образованный инtronом группы II бактерии *Oceanobacillus iheyensis*. Экземпляры данного псевдоузла состоят из 12–13 стемов, причем исключение коротких стемов не приводит к изменению сигнатуры псевдоузла. В рассмотренных структурах интроны группы II представлены тремя организмами, однако в двух других организмах подобные псевдоузлы не сохраняют свой тип при исключении коротких стемов. Интересно, что данный псевдоузел образуется только при наличии фрагмента экзона Intron Binding Site 1, а при его отсутствии сводится к базовому типу Kissing Loops.

**Ключевые слова:** вторичная структура РНК, спаривание оснований, стем, короткий стем, псевдоузел, сигнатура псевдоузла, инtron группы II.

## ВВЕДЕНИЕ

Молекулы рибонуклеиновых кислот (РНК) являются одним из важнейших классов биополимеров в живых организмах. Наряду с ДНК и белками, молекулы РНК задействованы во множестве клеточных процессов (см. обзор [1]). В последние 15–20 лет было открыто большое количество классов функциональных некодирующих РНК, что привело к переоценке роли РНК в клетке. На данный момент помимо матричных,

\*baulin@lpm.org.ru

транспортных и рибосомальных РНК, выступающих посредниками в процессе экспрессии генов, известны такие классы, как микро-РНК, транспортно-матричные РНК, малые ядерные РНК, длинные некодирующие РНК и др., которые принимают участие в трансляции [2], регуляции генов [3], процессинге РНК [4] и других процессах в клетках живых организмов.

Функция, выполняемая молекулой РНК, напрямую связана с её пространственной структурой, что определяет актуальность исследований структур РНК и поиска различных закономерностей как на уровне последовательности нуклеотидов, так и на уровне вторичной и пространственной структуры молекулы. Особый интерес представляет такой элемент структур РНК, как псевдоузлы. На данный момент не существует общепринятого мнения о том, к какому уровню структурной организации должны быть отнесены псевдоузлы, являются ли они частью вторичной структуры или третичным мотивом [5]. Однако известно, что псевдоузлы являются важным функциональным элементом, задействованным, например, в образовании катализитических ядер рибозимов [6, 7], сплайсинге инtronов [8], функционировании теломераз [9].

Одной из главных задач исследования структур РНК является задача предсказания вторичной структуры РНК. На данный момент существует несколько алгоритмов, решающих данную задачу за полиномиальное время (см., например, обзор [10]), однако такие алгоритмы либо ограничены классическими структурами РНК, не содержащими псевдоузлов [11, 12], либо допускают небольшой набор псевдоузлов простейших типов [13, 14]. Показано, что предсказание произвольной вторичной структуры РНК, содержащей псевдоузлы, является NP-сложной задачей [15].

Нами было замечено, что псевдоузлы в экспериментально разрешенных структурах РНК более чем на половину состоят из так называемых "коротких" стемов, т.е. стемов, образованных двумя или тремя парами оснований. Согласно эмпирическим термодинамическим параметрам образование стема вносит отрицательный вклад в свободную энергию структуры только при длине в четыре спаривания и более [16–18]. Таким образом, короткие стемы могут быть энергетически нестабильны и не должны вносить существенного вклада в свободную энергию структуры.

В данной работе осуществлен анализ роли коротких стемов в образовании псевдоузловых структур РНК. Показано, что при исключении коротких стемов из рассмотрения, псевдоузлы в экспериментально разрешенных пространственных структурах РНК сводятся к двум простейшим типам. Кроме того, на основе экспериментальных данных были получены косвенные подтверждения энергетической нестабильности коротких стемов.

## МАТЕРИАЛЫ И МЕТОДЫ

### Основные определения

Молекулу РНК будем представлять, как последовательность нуклеотидов, иначе говоря, как символьную последовательность в алфавите  $\{A, C, G, U\}$ . Каждый нуклеотид в молекуле имеет свой номер от 1 до  $L$ , где  $L$  – длина последовательности.

*Спаривание* – это пара нуклеотидов  $(i, j)$ , где  $i < j$ , которые образуют водородные связи. При этом допускаются Уотсон–Криковские связи между комплементарными нуклеотидами и плавающие G–U связи (wobble pairs).

*Стем* – это последовательность пар  $(i, j), (i + 1, j - 1), \dots, (i + k, j - k)$ , такая что:

- 1)  $k \geq 1$ ;
- 2)  $i + k < j - k$ ;
- 3) все пары  $(i + x, j - x)$ , где  $x = 0, \dots, k$ , образуют спаривания.

*Короткий стем* – это стем, для которого  $k < 3$ . Для стема  $(i, j), (i + 1, j - 1), \dots, (i + k, j - k)$  участок  $[i, i + k]$  называется *левым крылом* стема, а участок  $[j - k, j]$  – *правым крылом* стема.

Будем говорить, что спаривания  $(m, n)$  и  $(p, q)$  *конфликтуют*, если  $m < p < n < q$  или  $p < m < q < n$ . *Нить* (или неспаренный участок) – это участок  $[i, j]$ , такой что:

- 1) не существует спаривания  $(k, t)$ , такого что  $i \leq k \leq j$  или  $i \leq t \leq j$ ;
- 2) существуют спаривания, содержащие нуклеотиды  $i - 1$  и  $j + 1$ .

*Элементарный закрытый участок* ( $\mathcal{EZY}$ ) – это минимальный участок  $[i, j]$ , где  $i < j$ , такой что:

- 1) не существует спаривания  $(k, l)$ , такого что  $(i \leq k \leq j; l > j)$  или  $(k < i; i \leq l \leq j)$ ;
- 2) не существует такого нуклеотида  $l$ , что  $i < l < j$  и оба участка  $[i, \dots, l]$  и  $[l + 1, \dots, j]$  удовлетворяют условию 1);
- 3) существуют спаривания  $(i, k)$  и  $(l, j)$ ; при этом возможно  $k = j$  и  $i = l$ .

$\mathcal{EZY}$  называется псевдоузловым  $\mathcal{EZY}$  (или *псевдоузлом*), если спаривания образующих его стемов конфликтуют.  $\mathcal{EZY} [k, l]$  является *суб-ЭZY* для  $\mathcal{EZY} [i, j]$ , если  $i < k < l < j$ , и не существует такого  $\mathcal{EZY} [m, n]$ , что  $i < m < k < l < n < j$ .

В данной работе используются классические определения петель, согласно модели Мэтьюза–Тернера [19]. Петля называется *шпилькой*, если она состоит из одной нити. Петля называется *внутренней*, если состоит из двух нитей. Петля называется *мульти-петлей*, если она состоит из трех и более нитей. Внутренняя петля называется *выпячиванием*, если одна из её нитей имеет нулевую длину. Формальные определения обобщенных петель приведены на сайте базы данных URSDB [20] <http://server3.lpm.org.ru/urs/struct.py?where=3#def>.

## Классификация псевдоузлов

В работе используется топологическая классификация псевдоузлов, основанная на понятии сигнатуры. Классификация схожа с классификацией в [21]; основным отличием является исключение из рассмотрения одиночных спариваний в нашем случае. Определение сигнатуры совпадает с определением в [22]. Описание псевдоузлов с помощью сигнатур в том или ином виде встречается в работах [10, 23–25].

Возьмем  $\mathcal{EZY}$  и обозначим все его стемы латинскими буквами, согласно последовательности их крыльев в направлении от 5'- к 3'-концу РНК. Левым крыльям будут соответствовать прописные буквы (например,  $a$ ), а правым крыльям – заглавные буквы (например,  $A$ ). Таким образом, первому стему будут соответствовать буквы  $aA$ , второму –  $bB$ , и так далее.

*Полной сигнатурой*  $\mathcal{EZY}$  называется последовательность букв, соответствующих крыльям стемов  $\mathcal{EZY}$  в направлении от 5' к 3'-концу цепи РНК (см. рис. 1).

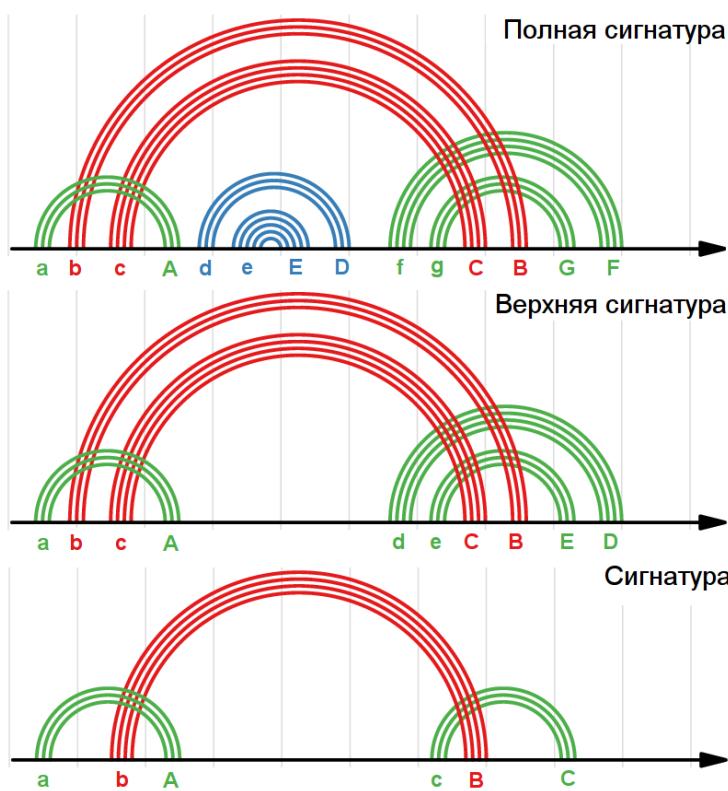
*Верхней сигнатурой*  $\mathcal{EZY}$  называется строка, полученная из полной сигнатуры  $\mathcal{EZY}$  в результате выполнения следующих операций (см. рис. 1):

- 1) удаление букв, соответствующих крыльям всех суб- $\mathcal{EZY}$  данного  $\mathcal{EZY}$ ;
- 2) переименование оставшихся крыльев буквами с начала алфавита с сохранением исходного порядка.

Стемы  $xX$ ,  $yY$  называются *связанными*, если обе строки  $xy$  и  $YX$  являются подстроками верхней сигнатуры.

*Сигнатурой*  $\mathcal{EZY}$  называется строка, полученная из верхней сигнатуры в результате выполнения следующих операций (см. рис. 1):

- 1) удаление всех букв, кроме  $x$  и  $X$  (первой и последней буквы соответствующих подстрок), соответствующих последовательностям связанных стемов;
- 2) переименование оставшихся крыльев буквами с начала алфавита с сохранением исходного порядка следования крыльев.



**Рис. 1.** Пример получения сигнатуры псевдоузла по его полной сигнатуре.

### Описание данных

В качестве источника исходных данных использовалась база данных пространственных структур РНК URSDB [20], т.к. это единственная база данных, в которой размечены сигнатуры всех псевдоузлов, встречающихся в структурах РНК из банка данных Protein Data Bank (PDB, [26]). Среди всех РНК-содержащих документов PDB были отобраны структуры из неизбыточного подмножества структур РНК ([27], версия 2.156\_all от 24 ноября 2017 г.). В подмножество вошло 2300 структур, образованных 3974 молекулами РНК из 1987 документов PDB. Т.к. база данных URSDB учитывает межцепочечные стемы, все сигнатуры псевдоузлов были пересчитаны после исключения таких стемов. Далее сигнатуры были пересчитаны в трех различных вариантах: после удаления стемов длины 2; после удаления стемов длины 3 и после удаления стемов длины 2 и 3. Для проверки наличия в PDB псевдоузлов таких типов, которые по каким-либо причинам не вошли в неизбыточное подмножество структур РНК, все процедуры были также выполнены на всём множестве структур РНК из банка PDB. Полные списки сигнатур приведены в дополнительных материалах к статье.

### РЕЗУЛЬТАТЫ

В рассматриваемое подмножество структур вошло 12793 стема различной длины. Среди всех стемов, входящих в псевдоузлы, коротких стемов оказалось более половины – 1171 из 2303. В процессе работы была проверена гипотеза о том, что короткие стемы предпочитают находиться рядом с внутренними петлями и избегать шпилек. Данная гипотеза была вызвана тем, что внутренние петли в реальных структурах РНК представляют собой набор неканонических спариваний нуклеотидов, а не последовательность неспаренных нуклеотидов, как это следует из графического представления элементов вторичной структуры РНК. Таким образом, в

пространственных структурах последовательные элементы стем – внутренняя петля – стем образуют один неканонический стем, что способствует дополнительной стабилизации структуры. Однако в случаях, когда короткий стем замыкает шпильку, какие-либо дополнительные стабилизирующие факторы отсутствуют. Гипотеза была подтверждена – точный тест Фишера показал  $p\text{-value} < 2.2 \cdot 10^{-16}$  для шпилек и  $p\text{-value} = 6.29 \cdot 10^{-9}$  для внутренних петель (см. табл. 1).

**Таблица 1.** Взаимное расположение стемов и примыкающих петель разных типов

Типы примыкающих петель	Кол-во коротких стемов	Кол-во длинных стемов
Шпильки	2242	3010
Нешпильки: выпячивания, внутренние петли, мультипетли	3895	3646
Внутренние петли	2066	1923
Невнутренние петли: шпильки, выпячивания, мультипетли	4071	4733

**Таблица 2.** Список сигнатур псевдоузлов с короткими стемами и без них (по данным анализа 1987 документов PDB из неизбыточного подмножества структур)

Сигнатура	Кол-во	Кол-во (без коротких стемов)
abAB	289	74
abAcBC	67	43
abcdBCAD	2	0
abcdCADC	2	0
abAcdBDeCE	3	0
abcdCABeDE	1	0
abAcdCeBEfDF	1	0
abAcdeBEfDFC	9	0
abcdCeAEfDFB	1	0
abcdCeBEAfDF	1	0
abAcdCefDFgBGE	1	0
abAcdeDfgFEChBHgI	17	0
abAcdeDfghiHFECjkGKIBLJmIM	2	0
<b>Всего</b>	<b>396</b>	<b>117</b>

Всего в рассматриваемом подмножестве структур РНК было обнаружено 396 псевдоузлов, принадлежащих 13 различным типам, согласно топологической классификации псевдоузлов (см. табл. 2). Наиболее распространёнными типами являются Н-узлы (сигнатура **abAB**, 289 псевдоузлов) и kissing loops (сигнатура **abAcBC**, 67 псевдоузлов). После исключения коротких стемов и пересчета сигнатур было выявлено 117 псевдоузлов (см. табл. 2), принадлежащих к двум простейшим типам (**abAB** и **abAcBC**).

## ОБСУЖДЕНИЕ

Подсчет сигнатур был также произведен на всём множестве структур РНК из банка PDB. Было обнаружено 6986 псевдоузлов. После исключения коротких стемов было выявлено 1276 псевдоузлов, 1270 из которых принадлежат двум простейшим типам, а оставшиеся 6 псевдоузлов имеют сигнатуру **abcdCefAFDEB**.

Все шесть псевдоузлов с сигнатурой **abcdCefAFDEB** были обнаружены в различных экземплярах молекулы интрана группы II бактерии *Oceanobacillus iheyensis*. Всего в PDB представлено 26 структур данной молекулы, которые образуют класс эквивалентности NR\_all\_35054.3. Анализ данных структур показал, что псевдоузел с сигнатурой **abcdCefAFDEB** образуется только в присутствии фрагмента экзона под названием Intron Binding Site 1 (IBS1, последовательность AUAA). Среди 26 структур данного класса шесть структур (идентификаторы 4ds6, 4faq, 4fau, 4y1n и два экземпляра 4y1o) содержат лигированный IBS1 и образуют указанный псевдоузел. Ещё три структуры (4far, Zeoh, Zeog) содержат IBS1 в виде изолированной молекулы и также образуют указанный псевдоузел (однако данные экземпляры содержат межцепочечные стемы, поэтому при подсчетах не учитывались). Таким образом, следует полагать, что данный псевдоузел существует только в момент сплайсинга данного интрана. Представителем данного класса эквивалентности в неизбыточном подмножестве структур РНК является структура с идентификатором 5j01, в которой IBS1 отсутствует, что объясняет отсутствие псевдоузла с сигнатурой **abcdCefAFDEB** в таблице 2.

Интроны группы II представлены в PDB тремя организмами: *Oceanobacillus iheyensis* (26 структур, класс NR\_all\_35054.3), *Lactococcus lactis* (2 структуры – 5g2x, 5g2y, класс NR\_all\_28269.1) и *Pylaiella littoralis* (1 структура – 4r0d, класс NR\_all\_05993.1). Структуры 5g2x и 4r0d содержат IBS1 и также образуют сложные псевдоузлы, однако при исключении коротких стемов их сигнатуры сводятся к сигнатуре **abAcBC**. Таким образом, во всём банке PDB существует всего один уникальный псевдоузел, который не сводится к простейшему типу при исключении коротких стемов.

## ЗАКЛЮЧЕНИЕ

Нами проведён анализ роли коротких стемов в образовании псевдоузловых структур РНК. В ходе работы были получены косвенные подтверждения энергетической нестабильности коротких стемов. Так, было показано, что короткие стемы предпочитают находиться рядом с внутренними петлями, а также избегать шпилек. Анализ сигнатур псевдоузлов показал, что при исключении коротких стемов из рассмотрения и последующего пересчета сигнатур все псевдоузлы из неизбыточного подмножества структур РНК сводятся к двум простейшим типам (**abAB** и **abAcBC**). С учётом этого представляется возможным разработка алгоритма предсказания вторичной структуры РНК, допускающего наличие только псевдоузлов указанных типов. Такой алгоритм может состоять из двух этапов: 1) предсказание вторичной структуры, не допускающей коротких стемов в составе псевдоузлов; 2) последующее наложение на полученную структуру коротких стемов, вносящих отрицательный вклад в свободную энергию структуры.

Анализ всего множества структур РНК из банка PDB выявил единственный псевдоузел, сигнатура которого не сводится к одному из простейших типов при исключении коротких стемов. Этот псевдоузел формируется только в процессе сплайсинга интрана группы II, т.е. не является элементом функциональной молекулы РНК, а именно для таких молекул задача предсказания вторичной структуры представляется наиболее актуальной.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (грант № 16-04-01640).

## СПИСОК ЛИТЕРАТУРЫ

1. Marzluff W.B. Twenty years of RNA: reflections on post-transcriptional regulation. *RNA (New York, NY)*. 2015. V. 21. № 4. P. 687–689 doi: [10.1261/rna.050997.115](https://doi.org/10.1261/rna.050997.115)
2. Eiring A.M., Harb J.G., Neviani P., Garton Ch., Oaks J.J., Spizzo R., Liu Sh., Schwind S., Santhanam R., Hickey Ch.J. et al. miR-328 Functions as an RNA decoy to modulate hnRNP E2 regulation of mRNA translation in leukemic blasts. *Cell*. 2010. V. 140. № 5. P. 652–665. doi: [10.1016/j.cell.2010.01.007](https://doi.org/10.1016/j.cell.2010.01.007)
3. Bartel D.P. MicroRNAs: target recognition and regulatory functions. *Cell*. 2009. V. 136. № 2. P. 215–233. doi: [10.1016/j.cell.2009.01.002](https://doi.org/10.1016/j.cell.2009.01.002)
4. Kapranov P., Cheng J., Dike S., Nix D.A., Duttagupta R., Willingham A.T., Stadler P.F., Hertel J., Hackermüller J., Hofacker I.L. et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science*. 2007. V. 316. № 5830. P. 1484–1488. doi: [10.1126/science.1138341](https://doi.org/10.1126/science.1138341)
5. Shapiro B.A., Yingling Ya.G., Kasprzak W., Bindewald E. Bridging the gap in RNA structure prediction. *Current opinion in structural biology*. 2007. V. 17. № 2. P. 157–165. doi: [10.1016/j.sbi.2007.03.001](https://doi.org/10.1016/j.sbi.2007.03.001)
6. Rastogi T., Beattie T.L., Olive J.E., Collins R.A. A long-range pseudoknot is required for activity of the Neurospora VS ribozyme. *The EMBO journal*. 1996. V. 15. № 11. P. 2820. doi: [10.1002/j.1460-2075.1996.tb00642.x](https://doi.org/10.1002/j.1460-2075.1996.tb00642.x)
7. Ke A., Zhou K., Ding F., Cate J.H., Doudna J.A. A conformational switch controls hepatitis delta virus ribozyme catalysis. *Nature*. 2004. V. 429. P. 201–205. doi: [10.1038/nature02522](https://doi.org/10.1038/nature02522)
8. Adams P.L., Stahley M.R., Kosek A.B., Wang J., Strobel S.A. Crystal structure of a self-splicing group I intron with both exons. *Nature*. 2004. V. 430. P. 45–50. doi: [10.1038/nature02642](https://doi.org/10.1038/nature02642)
9. Theimer C.A., Blois C.A., Feigon J. Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Molecular cell*. 2005. V. 17. № 5. P. 671–682. doi: [10.1016/j.molcel.2005.01.017](https://doi.org/10.1016/j.molcel.2005.01.017)
10. Condon A., Davy B., Rastegari B., Zhao Sh., Tarrant F. Classifying RNA pseudoknotted structures. *Theoretical Computer Science*. 2004. V. 320. № 1. P. 35–50. doi: [10.1016/j.tcs.2004.03.042](https://doi.org/10.1016/j.tcs.2004.03.042)
11. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research*. 2003. V. 31. № 13. P. 3406–3415. doi: [10.1093/nar/gkg595](https://doi.org/10.1093/nar/gkg595)
12. Reeder J., Höchsmann M., Rehmsmeier M., Voss B., Giegerich R. Beyond Mfold: recent advances in RNA bioinformatics. *Journal of biotechnology*. 2006. V. 124. № 1. P. 41–55. doi: [10.1016/j.jbiotec.2006.01.034](https://doi.org/10.1016/j.jbiotec.2006.01.034)
13. Rivas E., Eddy S.R. A dynamic programming algorithm for RNA structure prediction including pseudoknots. *Journal of molecular biology*. 1999. V. 285. № 5. P. 2053–2068. doi: [10.1006/jmbi.1998.2436](https://doi.org/10.1006/jmbi.1998.2436)
14. Lyngsø R.B., Pedersen C.N.S. Pseudoknots in RNA secondary structures. In: *Proceedings of the fourth annual international conference on Computational molecular biology. ACM*. 2000. P. 201–209. doi: [10.1145/332306.332551](https://doi.org/10.1145/332306.332551)
15. Lyngsø R.B., Pedersen C.N.S. RNA pseudoknot prediction in energy-based models. *Journal of computational biology*. 2000. V. 7. № 3–4. P. 409–427. doi: [10.1089/106652700750050862](https://doi.org/10.1089/106652700750050862)
16. Tan Z., Zhang W., Shi Ya., Wang F. RNA folding: structure prediction, folding kinetics and ion electrostatics. In: *Advance in Structural Bioinformatics*. Springer Netherlands, 2015. P. 143–183. doi: [10.1007/978-94-017-9245-5\\_11](https://doi.org/10.1007/978-94-017-9245-5_11)
17. Xia T., SantaLucia J. Jr., Burkard M.E., Kierzek R., Schroeder S.J., Jiao X., Cox Ch., Turner D.H. Thermodynamic parameters for an expanded nearest-neighbor model for

- formation of RNA duplexes with Watson – Crick base pairs. *Biochemistry*. 1998. V. 37. № 42. P. 14719–14735. doi: [10.1021/bi9809425](https://doi.org/10.1021/bi9809425)
18. Mathews D.H., Sabina J., Zuker M., Turner D.H. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *Journal of molecular biology*. 1999. V. 288. № 5. P. 911–940. doi: [10.1006/jmbi.1999.2700](https://doi.org/10.1006/jmbi.1999.2700)
  19. Baulin E., Yacovlev V., Khachko D., Spirin S., Roytberg M. URS DataBase: universe of RNA structures and their motifs. *Database*. 2016. V. 2016. Article No. baw085. doi: [10.1093/database/baw085](https://doi.org/10.1093/database/baw085)
  20. Zuker M., Mathews D.H., Turner D.H. Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide. In: *RNA biochemistry and biotechnology*. Springer Netherlands, 1999. P. 11–43. doi: [10.1007/978-94-011-4485-8\\_2](https://doi.org/10.1007/978-94-011-4485-8_2)
  21. Andersen J.E., Penner R.C., Reidys C.M., Waterman M.S. Topological classification and enumeration of RNA structures by genus. *Journal of mathematical biology*. 2013. V. 67. № 5. P. 1261–1278. doi: [10.1007/s00285-012-0594-x](https://doi.org/10.1007/s00285-012-0594-x)
  22. Bon M., Vernizzi G., Orland H., Zee A. Topological classification of RNA structures. *Journal of molecular biology*. 2008. V. 379. № 4. P. 900–911. doi: [10.1016/j.jmb.2008.04.033](https://doi.org/10.1016/j.jmb.2008.04.033)
  23. Rødland E.A. Pseudoknots in RNA secondary structures: representation, enumeration, and prevalence. *Journal of Computational Biology*. 2006. V. 13. № 6. P. 1197–1213. doi: [10.1089/cmb.2006.13.1197](https://doi.org/10.1089/cmb.2006.13.1197)
  24. Reidys C.M., Huang F.W.D., Andersen J.E., Penner R.C., Stadler P.F., Nebel M.E. Topology and prediction of RNA pseudoknots. *Bioinformatics*. 2011. V. 27. № 8. P. 1076–1085. doi: [10.1093/bioinformatics/btr090](https://doi.org/10.1093/bioinformatics/btr090)
  25. Chiu J.K.H., Chen Y.P.P. Conformational features of topologically classified RNA secondary structures. *PloS one*. 2012. V. 7. № 7. Article No. e39907. doi: [10.1371/journal.pone.0039907](https://doi.org/10.1371/journal.pone.0039907)
  26. Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., Bourne P.E. The protein data bank. *Nucleic acids research*. 2000. V. 28. № 1. P. 235–242. doi: [10.1093/nar/28.1.235](https://doi.org/10.1093/nar/28.1.235)
  27. Leontis N. B., Zirbel C. L. Nonredundant 3D structure datasets for RNA knowledge extraction and benchmarking. *RNA 3D structure analysis and prediction*. Springer Berlin Heidelberg, 2012. P. 281–298. doi: [10.1007/978-3-642-25740-7\\_13](https://doi.org/10.1007/978-3-642-25740-7_13)

Рукопись поступила в редакцию 26.10.2018.

Дата опубликования 18.12.2018.