

**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
имени М.В. Ломоносова**

---

**ФАКУЛЬТЕТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И КИБЕРНЕТИКИ**

**ISSN 2411-1473**

**Современные  
информационные технологии  
И  
ИТ-образование**

**Научный журнал**

**Том 2 (№ 11)**

**Москва  
2015**

УДК [004:377/378](063)  
ББК 74.5(0)я431+74.6(0)я431+32.81(0)я431  
С 56

**Современные информационные технологии и ИТ-образование. Т. 2 (№ 11),  
2015. - 614 с. (ISSN 2411-1473)**

В данном выпуске журнала представлены доклады X Юбилейной международной научно-практической конференции «Современные информационные технологии и ИТ-образование», прошедшей в Московском государственном университете имени М.В. Ломоносова 20-22 ноября 2015 года.

Журнал «Современные информационные технологии и ИТ-образование» включен в наукометрическую базу «Российский индекс научного цитирования» с размещением полнотекстовых версий в научной электронной библиотеке eLIBRARY.RU. URL: [http://elibrary.ru/title\\_about.asp?id=52785](http://elibrary.ru/title_about.asp?id=52785)



*Издание осуществлено при финансовой поддержке  
Российского фонда фундаментальных исследований  
(Грант РФФИ № 15-07-20760\_з)*

**Учредитель:**

Фонд содействия развитию интернет-медиа, ИТ-образования, человеческого потенциала «Лига интернет-медиа»

**Издатель:**

Фонд содействия развитию интернет-медиа, ИТ-образования, человеческого потенциала «Лига интернет-медиа»

**Адрес редакции:**

119991, г. Москва, ГСП-1, Ленинские горы, д. 1, стр. 52, факультет ВМК МГУ имени М.В. Ломоносова, каб. 375. E-mail: [sukhomlin@mail.ru](mailto:sukhomlin@mail.ru), тел./факс (495) 939-46-26.

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор).

Свидетельство о регистрации средства массовой информации ПИ № ФС77-61433 от 10 апреля 2015 г.

Издается с 2005 года. Выходит 1 раз в год.

**Редакционная коллегия журнала:**

**Главный редактор:**

**Сухомлин В.А.** - доктор технических наук, профессор, заведующий лабораторией ОИТ факультета ВМК МГУ имени М.В. Ломоносова, Президент Фонда «Лига интернет-медиа»;

**Члены редакционной коллегии:**

Веремей Е.И. - доктор физ.-мат. наук, профессор, СПбГУ;

Гергель В.П. - доктор физ.-мат. наук, профессор, ННГУ им. Н.И. Лобачевского;

Самуйлов К.Е. - доктор физ.-мат. наук, профессор, РУДН;

Калиниченко Л.А. - доктор физ.-мат. наук, профессор, вед. н.с. ИПИ РАН ФИЦ ИУ РАН;

Лугачев М.И. - доктор экономических наук, профессор, МГУ имени М.В. Ломоносова;

Любецкий В.А. - доктор физ.-мат. наук, профессор, ИППИ РАН им. А.А. Харкевича;

Нечаев В. В. - доктор технических наук, профессор, МИРЭА;

Посыпкин М.А. - доктор физ.-мат. наук, вед. н. с. ИППИ РАН им. А.А. Харкевича;

Язенин А.В. - доктор физ.-мат. наук, декан факультета ПМиК, профессор, ТвГУ;

Намиот Д.Е. - кандидат физ.-мат. наук, с.н.с. факультета ВМК МГУ имени М.В. Ломоносова;

Зубарева Е.В. - кандидат пед. наук, доцент, н.с. факультета ВМК МГУ имени М.В. Ломоносова;

Сотникова М.В. - кандидат физ.-мат. наук, доцент СПбГУ.

Статьи, поступающие в редакцию, рецензируются. За достоверность сведений, изложенных в статьях, ответственность несут авторы публикаций. Мнение редакции может не совпадать с мнением авторов материалов. При перепечатке ссылка на журнал обязательна.

Материалы публикуются в авторской редакции. При перепечатке и цитировании материалов ссылка на журнал «Современные информационные технологии и ИТ-образование» обязательна.

# Королев С.А.<sup>1</sup>, Селиверстов А.В.<sup>2</sup>, Зверков О.А.<sup>3</sup>, Любецкий В.А.<sup>4</sup>

<sup>1</sup>ИППИ РАН, г. Москва, стажер-исследователь, [korolev@iitp.ru](mailto:korolev@iitp.ru)

<sup>2</sup>ИППИ РАН, г. Москва, к.ф.-м.н., в.н.с., [slvstv@iitp.ru](mailto:slvstv@iitp.ru)

<sup>3</sup>ИППИ РАН, г. Москва, к.ф.-м.н., н.с., [zverkov@iitp.ru](mailto:zverkov@iitp.ru)

<sup>4</sup>ИППИ РАН, г. Москва, д.ф.-м.н., зав.лаб., [lyubetsk@iitp.ru](mailto:lyubetsk@iitp.ru)

## КЛАССИЧЕСКАЯ АТТЕНУАТОРНАЯ РЕГУЛЯЦИЯ, ЗАВИСИМАЯ ОТ КОНЦЕНТРАЦИИ ТРИПТОФАНА, У АКТИНОБАКТЕРИЙ

### КЛЮЧЕВЫЕ СЛОВА

*Актинобактерии, классическая аттенуаторная регуляция, лидерный пептид, моделирование транскрипции.*

### АННОТАЦИЯ

*Предсказаны и подтверждены моделированием новые случаи классической аттенуаторной регуляции, зависимой от концентрации триптофанил-тРНК, у актинобактерий. Показано, что экспрессия генов, кодирующих ферменты пути синтеза триптофана, регулируется аттенуаторной регуляцией только у видов двух родов *Corynebacterium* и *Streptomyces*. Обсуждается совершенствование методов моделирования аттенуаторной регуляции, вовлекающей псевдоузлы и триплексы на РНК.*

Лидерные гены играют ключевую роль в аттенуаторной регуляции, основанной на сопряжении транскрипции и трансляции у прокариот. Впервые такая регуляция была описана для *Escherichia coli* [1], но также известна и для актинобактерий. В частности, экспериментальное подтверждение такой регуляции, зависимой от концентрации триптофана, получено для *Corynebacterium glutamicum* [2] и *Streptomyces venezuelae* [3]. Широкомасштабный биоинформатический поиск проведен в работах [4–6].

Аттенуаторная регуляция определяется структурой, которая включает ген лидерного пептида с регуляторными кодонами в нём и связанные с ним альтернативные вторичные структуры мРНК, одни из которых приводят к преждевременной терминации транскрипции структурного гена, а другие позволяют РНК-полимеразе продолжать транскрипцию структурного гена. Эта альтернатива зависит от скорости, с которой рибосома выполняет трансляцию лидерного пептида. В свою очередь скорость зависит от концентрации аминоацил-тРНК, которая в свою очередь зависит от концентрации соответствующей аминокислоты и аминоацил-тРНК синтетазы.

Нами проведён поиск лидерных генов перед всеми размеченными в аннотациях генами 196 геномов актинобактерий, доступных в базе данных NCBI. При этом лидерные пептиды, кодируемые лидерными генами, содержат три подряд или две близкие пары остатков триптофана. Далее гены, отмеченные в аннотации, называются структурными. Предсказанные нами лидерные гены не перекрывают структурные гены и обычно имеют небольшую длину.

Моделирование классической аттенуаторной регуляции выполнено нашей программой, описанной в [7].

На рисунке 1 показана зависимость числа предсказанных лидерных генов с кодонами триптофана от расстояния между лидерным и структурным генами у актинобактерий.

Классическая аттенуаторная регуляция, зависимая от концентрации триптофана или триптофанил-тРНК, предсказана только у некоторых видов двух родов *Corynebacterium* и *Streptomyces*. При этом расстояние между лидерным и структурным генами колеблется в пределах от 50 до 296 п.н. У многих штаммов коринебактерий *C. pseudotuberculosis* и *C. ulcerans* лидерные гены предсказаны перед обоими генами *trpB* и *trpE*. У штаммов *C. diphtheriae* – только перед опероном, включающим *trpB1* и другие гены пути синтеза триптофана. У *C. aurimucosum* лидерные гены предсказаны перед геном *trpS*, кодирующим триптофанил-тРНК синтетазу, и большим опероном, содержащим ген *trpE* и другие гены пути синтеза триптофана, включая *trpB* и *trpA*. У

*Streptomyces* гены антранилат синтазы, предположительно регулируемые классической аттенуацией, не входят в единый оперон с другими генами пути синтеза триптофана. Ранее нами исследована потенциальная аттенуаторная регуляция гена *phzE* (называемого также *trpE1*), кодирующего антранилат синтазу [4], перед которым нами предсказан лидерный ген, содержащий три подряд кодона триптофана. У близкого вида *S. cattleya* перед геном *phzB*, кодирующим антранилат синтазу, также предсказан лидерный ген. Аналогичная картина наблюдается у *S. scabiei* и *S. coelicolor* для гена *trpE*.

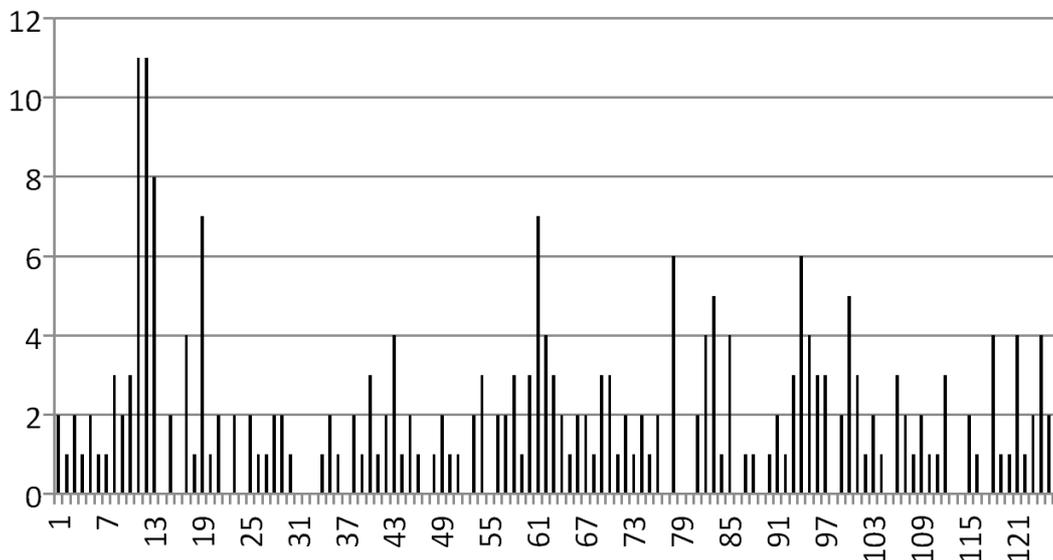


Рис.1. Зависимость числа предсказанных лидерных генов с кодонами триптофана от расстояния между лидерным и структурным генами у актинобактерий

Выполненное нашими методами моделирование подтверждает эффективность классической аттенуации для генов *trpE* и *trpB* у *C. pseudotuberculosis* и гена *trpB1* у *C. diphtheria* 31A. Результат моделирования показан на рисунке 2. Однако такая регуляция гена *trpB1* не подтверждается для штамма *C. diphtheriae* INCA 402.

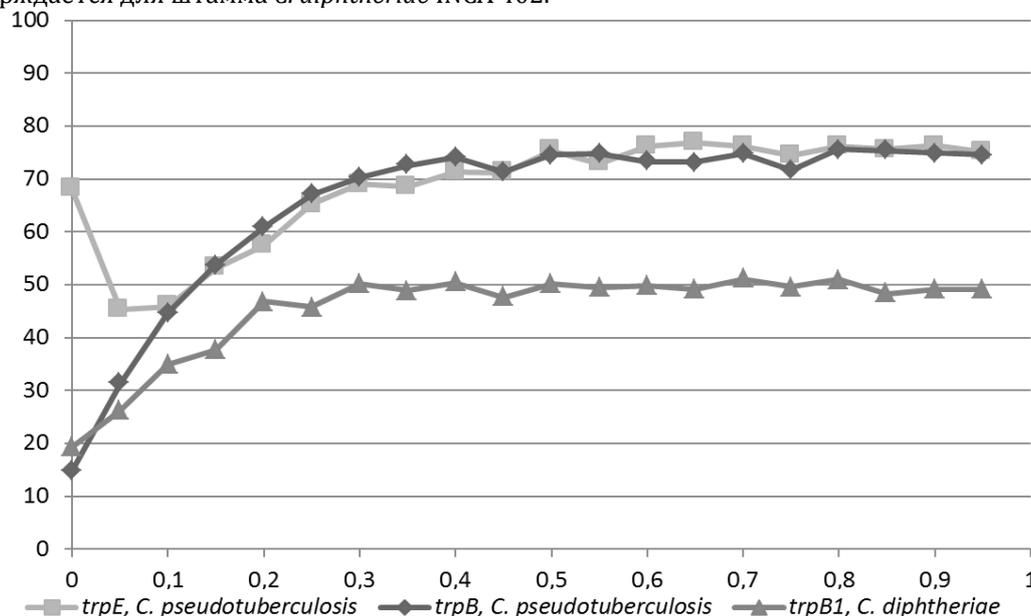


Рис.2. Зависимость частоты терминации транскрипции до начала транскрипции структурного гена от концентрации триптофана в модели для генов *trpE* и *trpB* у *C. pseudotuberculosis*, гена *trpB1* у *C. diphtheria* 31A.

Классическая аттенуаторная регуляция, зависящая от концентрации триптофанил-тРНК предсказана для видов двух родов *Corynebacterium* и *Streptomyces*. Эти предсказания хорошо согласуются с экспериментами [1–2]. Отсутствие такой регуляции у других видов может быть

связано с регуляцией на другом уровне. Некоторые из ферментов пути синтеза триптофана содержат на N-конце домен PF04715, вовлечённый в ингибирование фермента триптофаном [8].

Например, в геноме *Streptomyces avermitilis* MA-4680 (GenBank: NC\_003155) есть два гена, предположительно вовлечённых в синтез антранилата из хоризмата, которые имеют низкое сходство между собой. Первый ген *trpE* (с координатами 7417784..7419265) кодирует короткий белок длиной 493, который содержит N-концевой домен PF04715 и домен PF00425, связывающий хоризмат. Второй ген *phzE* (или *trpE1* с координатами complement(7320283..7322268)) кодирует длинный белок, содержащий домен PF00425, связывающий хоризмат, и домен PF00117, характерный для глутаминамидотрансфераз, но не содержит домена, участвующего в ингибировании триптофаном. Третий ген *pabB* (с координатами 8122893..8124014) и четвёртый *pabAB* (с координатами complement(1468399..1470597)). Последний ген также кодирует белок с N-концевым доменом PF04715. Домены определены по базе данных Pfam [9]. Координаты указаны по данным из GenBank, доступным по адресу <http://www.ncbi.nlm.nih.gov/nuccore/148878541>. Ранее нами исследована потенциальная аттенуаторная регуляция гена *phzE* (называемого также *trpE1*) [4], перед которым нами предсказан лидерный ген, содержащий три подряд кодона триптофана. Почему такая регуляция не найдена перед генами *trpE* и *pabAB*? Одно из возможных объяснений в том, что эти гены кодируют ферменты, которые ингибируются триптофаном. Тогда как фермент, кодируемый геном *phzE*, не ингибируется, поскольку не содержит PF04715 на N-конце. У близкого вида *Streptomyces cattleya* антранилат синтаза, кодируемая геном *phzB*, содержит домены PF00425 и PF00117, но не содержит домен PF04715. Перед этим геном также предсказан лидерный ген. Аналогичная картина наблюдается у *Streptomyces scabiei* (ген *trpE*). У *Streptomyces* гены антранилат синтазы, предположительно регулируемые классической аттенуацией, не входят в единый оперон с другими генами пути синтеза триптофана.

Отметим, что у актинобактерий известны и другие типы аттенуаторной регуляции. Один из них, вовлекающий белок Rho, описан в работе [10]. Но и случаи аттенуации транскрипции, не связанные с дополнительными факторами, могут значительно отличаться от классической аттенуации. Структура РНК может включать псевдоузлы и триплексы [5], существенно затрудняющие моделирование регуляции. Разработка новых алгоритмов, учитывающих эти эффекты, позволяет уточнить список регулируемых генов и эффективность предсказанной регуляции.

Напомним, что триплексы – это три участка РНК, один из которых состоит полностью из пуриновых нуклеотидов, а два других участка имеют относительно пуринового участка параллельную и антипараллельную ориентации. Пара антипараллельных участков образует спираль РНК, в которой комплементарными считаются уотсон-криковские пары и GU пара. Таким образом, два участка триплекса образуют обычную спираль, а еще один, параллельный, участок связан с ней хугстиновскими водородными связями. Последний участок называется третьим плечом триплекса, которое, как и составляющие его нуклеотиды, в записи отделяется знаком \* от самой спирали (от ее нуклеотидов). В триплексе возможны триады нуклеотидов: C\*GU, G\*GC, G\*GU, U\*AU, A\*AU, A\*GC, C\*GC и некоторые другие. Триады характеризуются рядом особенностей, например, триада C\*GC прочная лишь в слабокислой среде, когда происходит протонирование по атому N3 цитозина. Имеется много публикаций о РНК-триплексах, среди которых отметим [10–16]. Если третье плечо триплекса находится на 5'-конце его спирали, то необходимо, чтобы оно располагалось на определенном расстоянии от соседнего к нему плеча спирали, достаточном для сворачивания триплекса. Мы принимали: если третье плечо содержит  $k$  нуклеотидов, то оно отделено не менее чем на  $k+6$  нуклеотидами от соседнего плеча. Если третье плечо находится на 3'-конце спирали, то оно может быть и вблизи нее, как показано в [15–16].

Разрешено образование псевдоузлов у всех вторичных структур в каждом текущем окне на РНК между рибосомой, транслирующей лидерный пептид, и РНК-полимеразой. При этом энергия вторичной структуры вычисляется на основе усовершенствования метода, представленного в [17], и учитывается температура типичная для среды обитания бактерии.

При моделировании аттенуаторной регуляции у некоторых видов актинобактерий эффективная зависимость от концентрации триптофана частоты преждевременной терминации транскрипции наблюдается только при учёте РНК-триплексов и псевдоузлов. Однако полученные результаты носят предварительный характер, поскольку точные значения энергий этих структур не известны. В частности, важную роль играет правильное вычисление изменения энтропии при формировании триплексов.

С другой стороны, анализ траекторий моделирования в нашей модели позволяет

приписать некоторым шпилькам роль терминатора, а другим – антитерминатора, и построить адекватное регуляции (если она имеется) локальное множественное выравнивание потенциальных регуляторных областей. Мы ожидаем, что существенные элементы регуляторных структур консервативны, то есть хорошо выравниваются друг с другом. И действительно, такие выравнивания обычно удаётся построить [4–6] Это даёт возможность независимой проверки правильности предсказания, в частности, правильности выбора формул для вычисления энергии и констант для вычисления скоростей образования или распада элементов структуры РНК: спиралей и триплексов.

Описанные методы могут применяться и для предсказания эффективности регуляции инициации трансляции в зависимости от скорости трансляции лидерного пептида. Типичным примером такой регуляции служит LEU-регуляция гена *leuA*, которая обнаружена у большинства актинобактерий из порядка Actinomycetales [4–5]. Здесь предсказание регуляции основано на консервативности структуры РНК, включающей псевдоузел, и позиционную связь этой структуры с участком регуляторных кодонов и сайтом связывания рибосомы перед кодирующей областью *leuA*.

*Работа выполнена за счёт гранта Российского научного фонда (проект № 14–50–00150).*

## Литература

1. Das A., Crawford I.P., Yanofsky C. Regulation of tryptophan operon expression by attenuation in cell-free extracts of *Escherichia coli* // *Journal of Biological Chemistry*, 1982, vol. 257, no. 15, pp. 8795–8798.
2. Heery D.M., Dunican L.K. Cloning of the *trp* gene cluster from a tryptophan hyperproducing strain of *Corynebacterium glutamicum*: Identification of a mutation in the *trp* leader sequence // *Applied and Environmental Microbiology*, 1993, vol. 59, no. 3, pp. 791–799.
3. Lin C., Paradkar A.S., Vining L.C. Regulation of an anthranilate synthase gene in *Streptomyces venezuelae* by a *trp* attenuator // *Microbiology*, 1998, vol. 144, no. 7, pp. 1971–1980.
4. Seliverstov A.V., Putzer H., Gelfand M.S., Lyubetsky V.A. Comparative analysis of RNA regulatory elements of amino acid metabolism genes in Actinobacteria // *BMC Microbiology*, 2005, vol. 5, no. 54.
5. Лопатовская К.В., Селиверстов А.В., Любецкий В.А. Аттеноаторная регуляция оперонов биосинтеза аминокислот и аминоацил-тРНК у бактерий: сравнительный геномный анализ // *Молекулярная биология*, 2010, том 44, № 1, стр. 140–151.
6. Lyubetsky V.A., Korolev S.A., Seliverstov A.V., Zverkov O.A., Rubanov L.I. Gene expression regulation of the PF00480 or PF14340 domain proteins suggests their involvement in sulfur metabolism // *Computational Biology and Chemistry*, 2014, vol. 49, pp. 7–13.
7. Lyubetsky V.A., Pirogov S.A., Rubanov L.I., Seliverstov A.V. Modeling classic attenuation regulation of gene expression in bacteria, *Journal of Bioinformatics and Computational Biology*, 2007, vol. 5, no. 1, pp. 155–180.
8. Spraggon G., Kim C., Nguyen-Huu X., Yee M.C., Yanofsky C., Mills S.E. The structures of anthranilate synthase of *Serratia marcescens* crystallized in the presence of (i) its substrates, chorismate and glutamine, and a product, glutamate, and (ii) its end-product inhibitor, L-tryptophan // *Proc Natl Acad Sci U S A*, 2001, vol. 98, pp. 6021–6026.
9. Finn R.D., Bateman A., Clements J., Coggill P., Eberhardt R.Y., Eddy S.R., Heger A., Hetherington K., Holm L., Mistry J., Sonnhammer E.L.L., Tate J., Punta M. The Pfam protein families database // *Nucleic Acids Research*, 2014, Database Issue 42, pp. D222–D230.
10. Lyubetsky V.A., Korolev S.A., Seliverstov A.V., Zverkov O.A., Rubanov L.I. Gene expression regulation of the PF00480 or PF14340 domain proteins suggests their involvement in sulfur metabolism // *Computational Biology and Chemistry*, 2014, vol. 49, pp. 7–13.
11. Chastain M., Tinoco I.Jr. Poly(rA) binds poly(rG) poly(rC) to form a triple helix // *Nucleic Acids Res.*, 1992, vol. 20, no. 2, pp. 315–318.
12. Кнорре Д.Г., Мызина С.Д. *Биологическая химия*. М.: Наука. 2000.
13. Semerad C.L., Maher L.J. Exclusion of RNA strands from a purine motif triple helix // *Nucleic Acids Res.*, 1994, vol. 22, no. 24, pp. 5321–5325.
14. Carmona P., Molina M. Binding of oligonucleotides to a viral hairpin forming RNA triplexes with parallel G\*GC triplets // *Nucleic Acids Res.*, 2002, vol. 30, no. 6, pp. 1333–1337.
15. Klinck R., Guitteta E., Liquier J., Taillandier E., Gouyetteb C., Huynh-Dinhby T. Spectroscopic evidence for an intramolecular RNA triple helix // *FEBS Lett.*, 1994, vol. 355, pp. 297–300.
16. Holland J.A., Hoffman D.W. Structural features and stability of an RNA triple helix in solution // *Nucleic Acids Res.*, 1996, vol. 24, no. 14, pp. 2841–2848.
17. Isambert H., Siggia E.D. Modeling RNA folding paths with pseudoknots: application to hepatitis delta virus ribozyme // *Proc. Nat. Acad. Sci. USA.*, 2000, vol. 97, no. 12, pp. 6515–6520.