

УДК: 577.214

## Профили вызванной суперспирализацией дестабилизации дуплекса ДНК (SIDD) для промоторов бактериофага Т7

М. А. Орлов<sup>а</sup>, С. Г. Камзолова, А. А. Рясик,  
Е. А. Зыкова, А. А. Сорокин

Институт биофизики клетки Российской академии наук,  
Россия, 142290, г. Пущино, ул. Институтская, д. 3

E-mail: <sup>а</sup> orlovmikhailanat@gmail.com

Получено 08.01.2018, после доработки — 21.05.2018.

Принято к публикации 21.09.2018.

Для функционирования регуляторных областей ДНК решающее значение имеет не нуклеотидная последовательность (генетический текст), а их физико-химические и структурные свойства. Именно они обеспечивают кодирование ДНК-белковых взаимодействий, лежащих в основе различных процессов регуляции. Среди таких свойств SIDD (Stress-Induced Duplex Destabilization) — характеристика, описывающая склонность участка дуплекса ДНК к плавлению при заданном уровне суперспирализации. Ранее для данного параметра дуплекса показана роль в функционировании областей регуляции различного типа. В данной работе модель SIDD использована для получения профилей вероятности плавления последовательностей промоторов бактериофага Т7. Данный геном характеризуется малым размером (примерно 40 тыс. пар нуклеотидов) и временной организацией экспрессии генов: на первом этапе инфекции ранняя область Т7-ДНК транскрибируется РНК-полимеразой бактерии-хозяина, на более поздних этапах жизненного цикла фагоспецифичная РНК-полимераза последовательно производит транскрипцию областей генов II класса и III класса. При этом механизмы дифференциального узнавания промоторов разных групп ферментом-полимеразой не могут быть основаны исключительно на их нуклеотидной последовательности, в частности в связи с тем, что она очень близка для большинства таких промоторов. В то же время полученные профили SIDD данных промоторов сильно различаются и могут быть разделены на характерные группы, соответствующие функциональным классам промоторов Т7-ДНК. Так, все промоторы ранней области находятся в области влияния одного максимально дестабилизированного участка дуплекса ДНК, соответствующего различным областям конкретных промоторов. Промоторы класса II лишены значительно дестабилизированных областей вблизи точки старта транскрипции. Напротив, промоторы III класса имеют характерные пики профилей вероятности плавления, в каждом случае локализованные в ближней downstream-области. Таким образом, установлены значительные различия профилей для промоторных областей при очень близкой нуклеотидной последовательности (промоторы II и III классов отличаются единичными заменами нуклеотидов), что подтверждает высокую чувствительность рассматриваемого свойства дуплекса к первичной структуре, а также необходимость рассмотрения широкого генетического контекста. Описанные различия профилей вероятности плавления на основе модели SIDD наряду с другими физическими свойствами могут определять дифференциальное узнавание промоторов разных классов РНК-полимеразами.

Ключевые слова: бактериофаг Т7, промотор, РНК-полимераза, физика ДНК, вызванная суперспирализацией дестабилизация дуплекса ДНК

UDC: 577.214

## Stress-induced duplex destabilization (SIDD) profiles for T7 bacteriophage promoters

M. A. Orlov<sup>a</sup>, S. G. Kamzolova, A. A. Ryasik,  
E. A. Zykova, A. A. Sorokin

Institute of Cell Biophysics of Russian Academy of Sciences,  
3 Institutskaya st., Pushchino, 142290, Russia

E-mail: <sup>a</sup> orlovmikhailanat@gmail.com

*Received 08.01.2018, after completion — 21.05.2018.*

*Accepted for publication 21.09.2018.*

The functioning of DNA regulatory regions rely primarily on their physicochemical and structural properties but not on nucleotide sequences, i.e. ‘genetic text’. The formers are responsible for coding of DNA-protein interactions that govern various regulatory events. One of the characteristics is SIDD (Stress-Induced Duplex Destabilization) that quantify DNA duplex region propensity to melt under the imposed superhelical stress. The duplex property has been shown to participate in activity of various regulatory regions. Here we employ the SIDD model to calculate melting probability profiles for T7 bacteriophage promoter sequences. The genome is characterized by small size (approximately 40 thousand nucleotides) and temporal organization of expression: at the first stage of infection early T7 DNA region is transcribed by the host cell RNA polymerase, later on in life cycle phage-specific RNA polymerase performs transcription of class II and class III genes regions. Differential recognition of a particular group of promoters by the enzyme cannot be solely explained by their nucleotide sequences, because of, among other reasons, it is fairly similar among most the promoters. At the same time SIDD profiles obtained vary significantly and are clearly separated into groups corresponding to functional promoter classes of T7 DNA. For example, early promoters are affected by the same maximally destabilized DNA duplex region located at the varying region of a particular promoter. class II promoters lack substantially destabilized regions close to transcription start sites. Class III promoters, in contrast, demonstrate characteristic melting probability maxima located in the near-downstream region in all cases. Therefore, the apparent differences among the promoter groups with exceptional textual similarity (class II and class III differ by only few singular substitutions) were established. This confirms the major impact of DNA primary structure on the duplex parameter as well as a need for a broad genetic context consideration. The differences in melting probability profiles obtained using SIDD model alongside with other DNA physicochemical properties appears to be involved in differential promoter recognition by RNA polymerases.

**Keywords:** T7 bacteriophage, promoter, RNA-polymerase, DNA physics, Stress-induced duplex destabilization

Citation: *Computer Research and Modeling*, 2018, vol. 10, no. 6, pp. 867–878 (Russian).

## Введение

Отличительной особенностью современного этапа исследования механизмов транскрипции (в частности, промоторно-полимеразного узнавания) является рассмотрение в качестве новых промоторных детерминант физико-химических характеристик двойной спирали ДНК. В ходе многочисленных исследований показано, что разные физические свойства дуплекса промоторной ДНК (электростатический профиль, термостабильность и наличие легкоплавких участков, конформационная подвижность, изгибность, наличие изломов, петель и крестообразных структур, динамические характеристики и др.) могут служить детерминирующими элементами прокариотических промоторов, узнаваемыми нативной РНК-полимеразой [Kamzolova, Postnikova, 1981; Margalit et al., 1988; Perez-Martin et al., 1994; Jensen et al., 1999; Yeremian, 2000; Kamzolova et al., 2000; Kamzolova et al., 2005; Wang, Benham, 2006; Sorokin et al., 2006; Osipov et al., 2010; Kamzolova et al., 2014].

Такие новые промоторные детерминанты, являющиеся по своей природе физическими свойствами ДНК, вызывают особый интерес, поскольку существенны на начальных этапах промоторно-полимеразного взаимодействия. Эти этапы, как правило, являются также скоростью-лимитирующими стадиями всего процесса транскрипции. Особенно наглядно это продемонстрировано для РНК-полимеразы бактериофага Т7 и ее многочисленных нативных промоторов [Kamzolova et al., 2014; Сорокин и др., 2016]. Как показано в этом случае, комплементарные электростатические взаимодействия между детерминирующими элементами Т7 промоторов и различными функционально значимыми компонентами промотор-связывающего центра Т7-РНК-полимеразы способствуют наиболее адекватной для каждого индивидуального промотора локализации фермента на промоторной ДНК, что определяет характер и скорость дальнейших конформационных переходов промоторно-полимеразного комплекса, а также его биохимические свойства [Kamzolova et al., 2014]. Следует подчеркнуть, что разные промоторы отличаются по характеру электростатических взаимодействий с активным центром РНК-полимеразы. Таким образом, анализ электростатических профилей промоторной ДНК считается перспективным и необходимым разделом исследования проблемы специфического кодирования промоторно-полимеразного узнавания для конкретных индивидуальных промоторов.

Другим важным разделом данной проблематики являются анализ термодинамических характеристик промоторной ДНК и оценка их вклада в обеспечение функциональной активности промоторов. Под термодинамическими свойствами ДНК понимают характеристики дуплекса, которые используются для описания процесса его плавления (денатурации). При этом плавление может быть вызвано различными физическими (прежде всего нагреванием) или химическими воздействиями, приводящими к расхождению цепей дуплекса ДНК и образованию ее одноцепочечных форм. Соответствующие модели основаны на расчетах энергии Гиббса ( $G^\circ$ ), энтропии ( $S^\circ$ ) и энтальпии ( $H^\circ$ ). Значения этих параметров для олигонуклеотидов (прежде всего димеров) получены в ряде работ и служат основой для исследований более крупных фрагментов ДНК [Breslauer et al., 1986; Sugimoto et al., 1996; SantaLucia, Hicks, 2004]. Конкретные модели рассматривают взаимодействия различного типа (как для отдельных составляющих молекулы ДНК, так и ее окружения) [SantaLucia, Hicks, 2004], при этом в основе многих использован алгоритм «ближайших соседей» [Guerra, 2015].

Примечательно, что исследование роли различных термодинамических свойств геномной ДНК в определении взаимодействующих с РНК-полимеразой участков началось еще до того, как была сформулирована сама концепция промотора [Szybalski et al., 1966; Баев и др., 1972; Адлер и др., 1973; Камзолова и др., 1977; Cherny et al., 1977; Pribnow, 1979]. Основанием для этих исследований послужили данные о том, что ДНК некоторых бактериофагов содержат легкоплавкие области, которые могут использоваться РНК-полимеразой для инициации синтеза РНК [Баев и др., 1972; Камзолова и др., 1977; Cherny et al., 1977]. Кроме того, было известно, что РНК-полимераза обладает большим сродством к денатурированной ДНК и однонитевым

полинуклеотидам по сравнению с нативной ДНК, а поли-d(AT) является более эффективной матрицей, чем естественные ДНК, что также объясняется большей легкостью раскрытия АТ-пар в сравнении с GC-парами [Szybalski et al., 1966; Pribnow, 1979].

В настоящее время идентифицирована первичная структура сотен прокариотических промоторов. Анализ показал, что около 70 % охарактеризованных промоторов имеют повышенное содержание АТ-пар в upstream-области, что предполагает возможность формирования легкоплавких участков. При этом необходимо отметить, что термодинамические характеристики ДНК определяются в большей степени последовательностью оснований, а не их «валовым» составом [Breslauer et al., 1986; Margalit et al., 1988]. Поэтому вопрос о наличии легкоплавких участков в промоторной ДНК и их использовании в качестве сигнальных элементов, узнаваемых РНК-полимеразой в конкретных индивидуальных промоторах, требует специальных исследований.

Примечательно, что все специфически узнаваемые фаговой полимеразой промоторы в геноме рассматриваемого в настоящей работе бактериофага Т7 имеют идентичную нуклеотидную последовательность в области [–7; –3] пар оснований (п. о.) от точки старта транскрипции (ТСТ). Более того, промоторы ряда других бактериофагов (Т3, SP6 и K11), взаимодействующие с соответствующими фагоспецифичными РНК-полимеразами, также содержат эту консенсусную последовательность. Это дает основания предполагать, что данная область промоторов выполняет в них универсальную функцию. С другой стороны, участок [–12; –8] названных промоторов сильно варьирует по нуклеотидному составу, что подтверждает представления о его участии в специфическом узнавании и связывании конкретного промотора с РНК-полимеразой конкретного фага. При этом с помощью различных экспериментальных методик показано, что плавление дуплекса ДНК в ходе инициации транскрипции происходит в downstream-направлении от положения –5 относительно ТСТ. В соответствии с этим в консенсусной области Т7-специфичных промоторов выделяют 2 функциональные области: область связывания РНК-полимеразы (локализована в интервале [–17; –5] пар оснований относительно ТСТ) и область инициации транскрипции (интервал [–4; +6]). Также установлено, что замены нуклеотидов в областях связывания и инициации оказывают наибольшие эффекты на соответствующие процессы [Imburgio et al., 2000].

Разработка новых методов оценки термодинамической стабильности двойной спирали ДНК позволила проанализировать ряд термодинамических характеристик небольших геномов и плазмид [Breslauer et al., 1986; Benham, 1992; 1993; 1996; Yeremian, 2000]. Для некоторых геномов установлено соответствие легкоплавких доменов и функционально значимых участков ДНК (включая ТСТ) [Benham, 1993; 1996]. Для ряда других геномов эта корреляция не обнаружена [Yeremian, 2000]. Все это указывает на актуальность исследования роли термодинамики геномной ДНК в формировании промоторной функции и выяснения участия легкоплавких доменов ДНК как специфических сигнальных элементов для конкретных промоторов.

Многие методы, основанные на оценке термодинамических свойств ДНК, в настоящее время используются для предсказания положения промоторных областей. К их числу относятся расчеты профилей свободной энергии, стэкинг-взаимодействий азотистых оснований, температуры плавления, энтропии [Shahmuradov et al., 2017], а также рассмотренная в данной работе вызванная суперспирализацией дестабилизации дуплекса (SIDD, Stress-induced duplex destabilization) [Wang and Benham, 2006]. Объектом настоящего исследования служит полногеномная ДНК литического бактериофага Т7, а именно ее промоторные области, как взаимодействующие с  $\sigma^{70}$  РНК-полимеразой *E. coli*, так и узнаваемые фагоспецифичной Т7 РНК-полимеразой. Данные исследования представляют дополнительный интерес в связи с тем, что для рассмотренных промоторов ранее были опубликованы данные о различиях электростатических профилей [Kamzolova et al., 2014; Сорокин и др., 2016; Камзолова и др., 2015]. Это делает возможным сравнительный анализ двух различных физических свойств на примере промоторных областей бактериофага Т7.

## Краткое описание модели

Поскольку при физиологических условиях для плавления (открывания) АТ-пары дуплекса ДНК требуется меньше свободной энергии, чем для плавления GC-пары [Breslauer et al., 1986], предсказания положений областей денатурации могут основываться либо на данных о локальном обогащении рассматриваемого участка АТ-парами, либо об энергии, которая необходима для локального плавления дуплекса [Natale et al., 1992]. Данный подход позволяет успешно выявлять наиболее легкоплавкие участки, однако для оценки их размеров необходимо использование более сложного анализа. При этом плавление некоторого участка ДНК зависит не только от его собственных характеристик, но и от его взаимодействий с генетическим окружением (контекстом). Такая зависимость определяет сложное поведение системы при переходе между закрытым и открытым состояниями и, в частности, высокую чувствительность вероятности плавления к малым изменениям нуклеотидного состава.

В основе использованной модели лежит статистико-механический расчет суперспиральной дестабилизации дуплекса ДНК, подробно описанный ранее [Benham, 1992]. При этом соответствующая каждой паре нуклеотидов свободная энергия дуплекса зависит от ряда факторов: от размера рассматриваемого участка, его нуклеотидного состава, уровня взаимной скрученности цепей дуплекса, а также суперспиральной деформации. Исходно выявляют, какое из возможных состояний молекулы ДНК соответствует минимуму свободной энергии. Далее задают пороговое значение энергии и избирают те состояния, свободная энергия которых превышает минимум энергии на величину этого порога или меньше. Вычисляют их статистическую сумму (partition function) и другие статистико-механические параметры. Дополнительно с помощью расчета плотности состояний производят учет тех из них, энергия которых превышает пороговую. Данный этап расчета носит приблизительный характер, однако позволяет добиться высокой точности с отклонениями от истинного значения менее 0.01 % [Benham, 1990].

Описываемая модель позволяет охарактеризовать стабильность участка ДНК двумя величинами. Первая из них — средняя вероятность  $p(x)$  плавления данной пары нуклеотидов в положении  $x$  для рассмотренного ансамбля молекул:

$$p(x) = Z(x) / Z,$$

где  $Z$  — статистическая сумма (partition function), рассчитанная как

$$Z(x) = \sum_{i_x} \exp[-G(i_x) / RT].$$

Суммирование производится по всем состояниям  $i_x$ , расплавленным в положении  $x$ . График зависимости  $p(x)$  от  $x$  описывает вероятность перехода и позволяет выявить участки ДНК, плавление которых наиболее вероятно. Вторая, более чувствительная мера дестабилизации, — дополнительная энергия  $G(x)$ , необходимая для плавления дуплекса по заданной паре оснований  $x$  [Benham, 1993]:

$$G(x) = {}^{\circ}G(x) - {}^{\circ}G,$$

где  ${}^{\circ}G$  — средняя свободная энергия рассматриваемого ансамбля и  ${}^{\circ}G(x)$  — средняя свободная энергия для всех расплавленных по положению  $x$  состояний  $i_x$ .

$${}^{\circ}G = \frac{\sum_{i_x} \exp\left[-\frac{G(i_x)}{RT}\right]}{Z(x)}.$$

В данном случае  $G(i_x)$  соответствует свободной энергии состояния  $i$ , которая также используется для построения профилей, описывающих стабильность дуплекса ДНК. В соответствии с этим возможны два варианта представления результатов расчета в виде профилей:

с использованием значений вероятности перехода  $p$  (профили вероятности плавления) или дополнительной энергии  $-G$ , необходимой для поддержания данной пары нуклеотидов в открытой форме (профили дестабилизации) [Benham, 1993]. В данной работе рассмотрены профили вероятности перехода, поскольку они позволяют более наглядно выделить только сильно дестабилизированные участки ДНК с меньшим вниманием к слабо дестабилизированным областям.

Следует отметить, что авторы метода указывают на сложную взаимосвязь между локальным обогащением участка ДНК АТ-парами и его склонностью к плавлению при воздействии суперспирального стресса. Подчеркивается, что SIDD определяется не только первичной структурой (нуклеотидной последовательностью) рассматриваемого участка, но и тем, насколько эффективно процесс перехода в расплавленное состояние в нем конкурирует с таковым в окружающих участках (генетическом контексте). Авторы приводят пример последовательности ДНК полиомы, содержащей исключительно богатую парами АТ область (15 из 16 нуклеотидов А/Т, в том числе 10 подряд). При этом данная область не входит в число 10 наиболее дестабилизированных участков рассмотренной молекулы ДНК [Benham, 1993]. Подчеркивается также, что SIDD, как свойство дуплекса ДНК, не сводимо к его термостабильности [Wang, Benham, 2006]. При рассмотрении различных физических и текстовых свойств ДНК установлена максимальная эффективность SIDD в качестве промоторной детерминанты, используемой для предсказания локализации промоторов с помощью методов машинного обучения. В той же работе проведен корреляционный анализ таких характеристик, показавший умеренную положительную корреляцию энергии дестабилизации (SIDD) со способностью ДНК к деформации, слабую положительную — с термостабильностью и умеренную отрицательную корреляцию с баллом мотива области  $-10$  [Wang, Benham, 2006]. Все это указывает на важность рассмотрения данного свойства ДНК как при аннотации регуляторных областей, так и для исследований механизмов функционирования промоторов.

## Материалы и методы

Нуклеотидная последовательность и аннотация (включая расположение промоторных областей) генома бактериофага Т7 получены из базы данных NCBI RefSeq (версия генома NC\_001604.1) [<http://www.ncbi.nlm.nih.gov/refseq/>]. Профили вызванной суперспирализацией дестабилизации дуплекса (SIDD) были получены с применением различных методик: на основе статистико-механического подхода [Bi, Benham, 2004] и, для сравнения, альтернативного подхода с рассмотрением непрерывных степеней свободы [Michoel, Van de Peer, 2006]. При этом первый был реализован при помощи двух методик — дистанционного расчета [Bi, Benham, 2004] и локально с применением алгоритма, предложенного теми же авторами в виде программы на языке Perl [Zhabinskaya et al., 2015]. Второй подход реализован с использованием языка Octave [Michoel, Van de Peer, 2006]. Применены стандартные условия расчета: температура — 310 К, плотность суперспирализации —  $-0.055$  и ионная сила —  $0.01$  М. При этом для каждого из промоторов рассмотрены фрагменты ДНК в интервале  $[-500; 250]$  пар оснований относительно ТСТ. Результаты представлены в виде профилей, описывающих вероятность  $P$  пребывания дуплекса в открытом (расплавленном) состоянии по заданной паре оснований.

## Результаты и обсуждение

Карта генома бактериофага Т7, иллюстрирующая положение различных функциональных областей, приведена на рис. 1.

Профили вероятности плавления для  $\sigma^{70}$ -специфичных (ранних) промоторов приведены на рис. 2. Для них характерно наличие дестабилизированной области размером около 80 пар оснований в различных областях вблизи ТСТ. При этом дестабилизированные участки промоторов в каждом случае совпадают с одним и тем же участком генома, находящимся вблизи

5'-конца генома рядом с несколькими сближенными ранними промоторами. Таким образом, данный дестабилизированный участок входит в генетический контекст всех  $\sigma^{70}$ -специфичных промоторов. При этом его локализация относительно ТСТ рассматриваемых промоторов различается: для *phil.3*, *A3*, *A2* и *A1* пик вероятности открывания дуплекса расположен в upstream-области (расстояние до него максимально в случае *phil.3* и минимально — в случае *A1*). У промоторов *1.1B* и *1.1A* данный дестабилизированный участок дуплекса приходится на downstream-область.

Поздние (T7-специфичные) промоторы двух классов демонстрируют характеристические различия в профилях SID. У представителей класса II (рис. 3) заметно отсутствие выраженных элементов профилей (то есть существенно дестабилизированных участков) вблизи ТСТ. Более того, на рассмотренном отрезке [−750; 250] пар оснований относительно ТСТ значимые пики вероятности плавления отмечаются лишь у единичных промоторов. Напротив, промоторы III класса (рис. 4) содержат области с высокой дестабилизацией в ближней downstream-области (примерно соответствующей интервалу [20; 90] пар оснований относительно ТСТ, выделена на рисунке). Необходимо отметить, что T7-специфичные промоторы (*phiOL*, *phiOR* и расположенный на обратной цепи *A0*) не могут быть отнесены ни к одному из классов и не участвуют в продуктивной транскрипции [Сорокин и др., 2016].

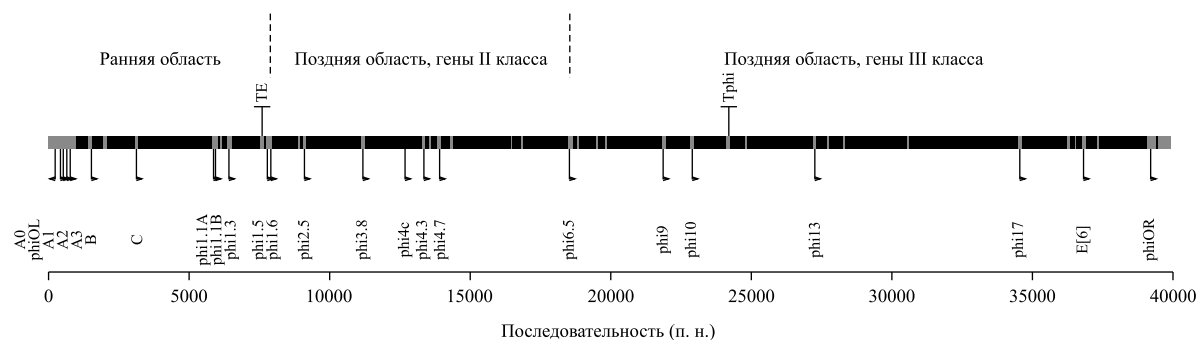


Рис. 1. Генетическая карта генома бактериофага Т7 (адаптировано из [Сорокин и др., 2016]). Вверху обозначены границы областей генома, соответствующих генам и промоторам разных классов. Промоторы обозначены стрелками, указывающими направление транскрипции. Терминаторы транскрипции обозначены T-образными линиями. Кодированные последовательности выделены черным, некодирующие — серым

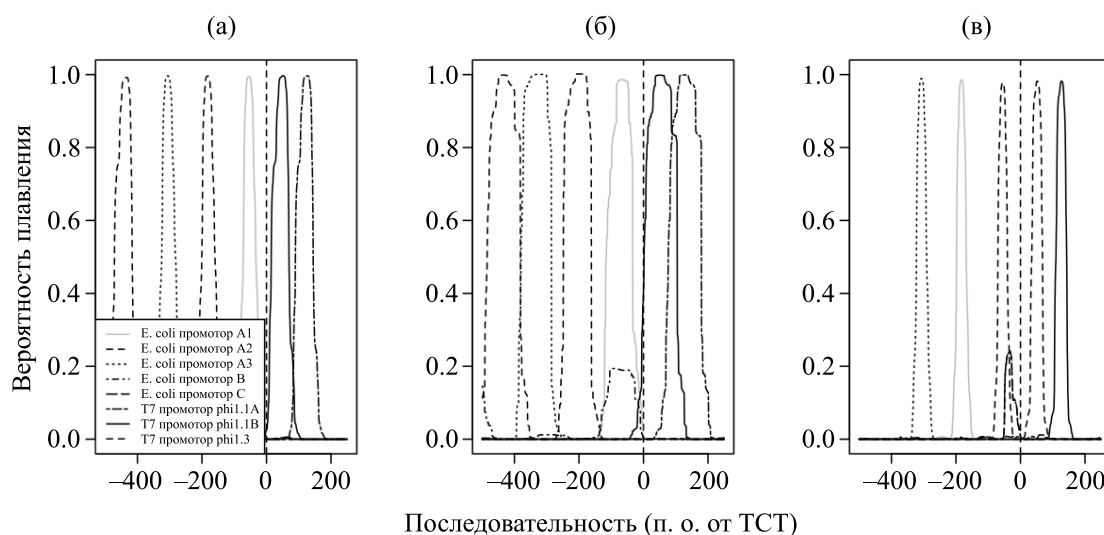


Рис. 2. Профили вероятности плавления для ранних ( $\sigma^{70}$ -специфичных) промоторов, рассчитанных при помощи различных методик: а) [Bi, Benham, 2004]; б) [Michoel, Van de Peer, 2006]; в) [Zhabinskaya et al., 2015]

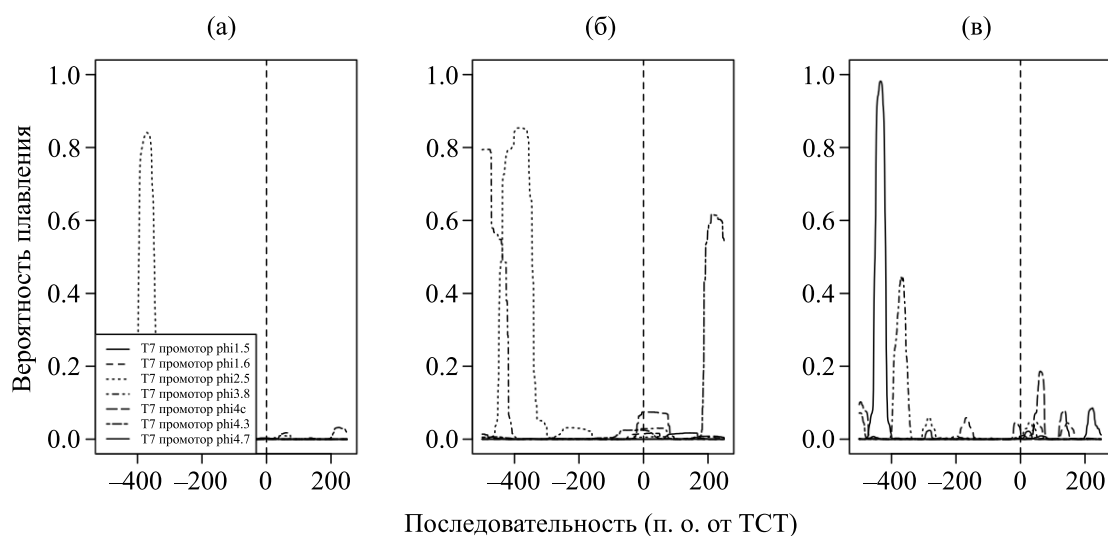


Рис. 3. Профили вероятности плавления для поздних (фагоспецифичных) промоторов II класса. Условные обозначения — те же, что на рис. 2

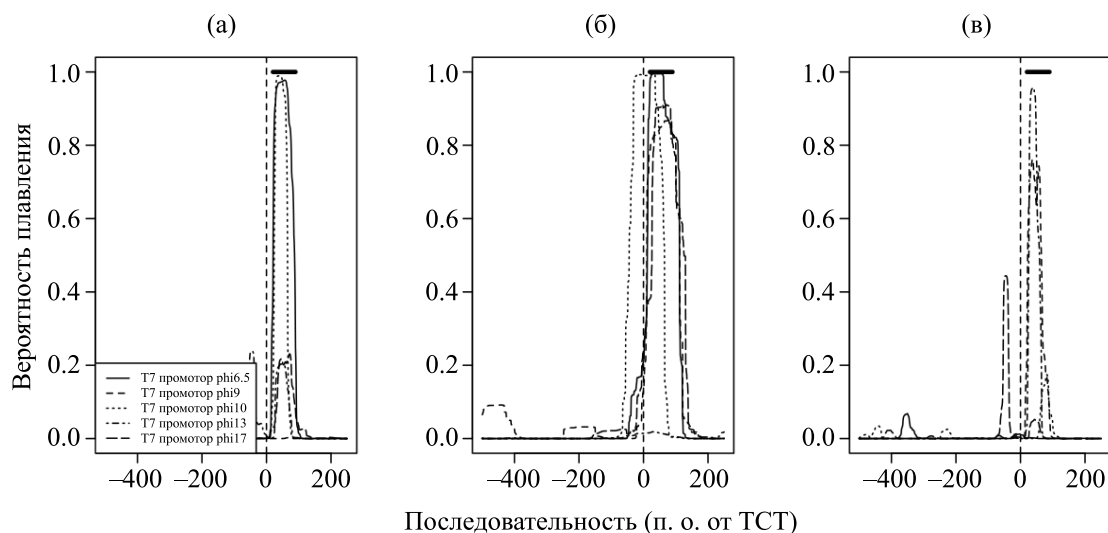


Рис. 4. Профили вероятности плавления для поздних (фагоспецифичных) промоторов III класса. Интервал [20; 90] относительно ТСТ выделен горизонтальной чертой. Условные обозначения — те же, что на рис. 2

Таким образом, два класса узнаваемых специфичной РНК-полимеразой фага промоторов сильно различаются по профилям вероятности плавления (SIDD). Описанные различия позволяют предположить, что для функционирования промоторов II класса вызванная суперспирализацией дестабилизация дуплекса не имеет существенного значения, в то время как наличие характерных для III класса промоторов дестабилизированных участков может свидетельствовать об участии этого параметра дуплекса в их функционировании. Возможно также другое предположение о причинах различия профилей SIDD рассмотренных классов. В работе [McAllister, Carter, 1980] с использованием полученных в результате рестрикции фрагментов Т7-ДНК установлено, что промоторы II класса избирательно ингибируются при способствующих стабилизации дуплекса ДНК условиях. Не исключено, что обе группы промоторов требуют наличия дестабилизированных участков вблизи ТСТ, способствующих переходу комплекса ДНК — РНК-полимераза в открытое состояние. Однако в случае представителей II класса такая дестабилизация может достигаться только в определенных условиях среды. В то же время для от-



крытия близких к точке старта транскрипции участков дуплекса в случае исходно дестабилизированных промоторов III класса особым условием может не требоваться.

Следует отметить, что роль SIDD в дифференцированном узнавании разных классов может быть опосредована изменением силы промоторов (определяемой как относительная нагрузка транскрипта). Ранее установлено, что промоторы III класса характеризуются значительно большей силой в сравнении с промоторами II класса [Kamzolova et al., 2014]. Это различие может объясняться, в частности, различным уровнем дестабилизации. Однако при этом необходимо учитывать, что сила промотора является интегральной характеристикой, которая определяется комбинацией скоростей связывания и открывания промоторов, скоростью формирования исходных межнуклеотидных связей, а также скоростью и эффективностью перехода к стабильному транскрипционному комплексу [McClure, 1985]. В связи с этим, помимо профилей SIDD (которые главным образом влияют на переход комплекса промотор–РНК-полимеразы в открытое состояние на одном из этапов инициации транскрипции), необходимо рассматривать некоторые дополнительные характеристики. Полученные результаты дополняют установленную ранее значимость профилей физических свойств (на примере электростатического потенциала) для дифференцированного узнавания РНК-полимеразой промоторов двух рассмотренных классов поздних промоторов генома T7 (рис. 5, адаптировано из [Kamzolova et al., 2014]). Следует отметить, что в случае электростатических профилей различия между нативными промоторами разных классов состоят в наличии характерных мотивов (долин и пиков профилей), которые делают возможным специфическое связывание промоторов с РНК-полимеразой бактериофага. Как и в случае SIDD, различия в профилях электростатического потенциала промоторов не могут определяться исключительно их первичной структурой и подвержены значительному влиянию контекста. Однако протяженность характеристических элементов профилей и расстояние, на котором генетический контекст может оказывать на них влияние, в случае электростатического потенциала значительно меньше.

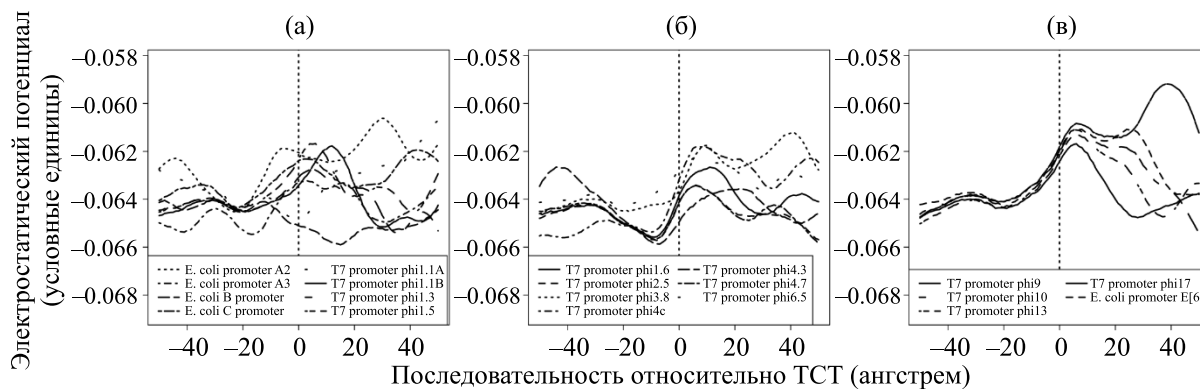


Рис. 5. Электростатические профили ДНК бактериофага T7 (адаптировано из [Kamzolova et al., 2014]): а) ранних  $\sigma^{70}$ -специфичных промоторов; б) поздних фагоспецифичных промоторов II класса; в) поздних фагоспецифичных промоторов III класса

В отношении связи различий в полученных профилях SIDD и нуклеотидной последовательности промоторов важно отметить следующее. Промотор *phi10* (наиболее сильный промотор генома T7, относится к классу III) максимально дестабилизирован, в то время как слабый промотор *phi3.8* (класс II) расположен в области, имеющей гораздо более стабильный дуплекс ДНК. Данные различия хорошо коррелируют с силой этих промоторов и эффективностью инициации транскрипции с них. Однако первичная структура промоторов *phi3.8* и *phi10* различается лишь по 4 из 26 (образующих консенсусную область) пар оснований. Такое небольшое различие не может полностью определять формирование области дестабилизации размером порядка 80 пар оснований, расположенной в upstream-области промотора *phi10* и отсутствующей у *phi3.8*, что также подчеркивает важность учета генетического контекста ДНК.

## Заключение

В настоящей работе с использованием модели вызванной суперспирализацией дестабилизации дуплекса ДНК (SIDD) установлено, что принадлежащие к разным классам, т. е. активные в различных стадиях жизненного цикла, промоторы бактериофага Т7 имеют значительно различающиеся профили вероятности плавления при очень близкой (в ряде случаев — идентичной) первичной структуре. Это подтверждает представления о том, что РНК-полимераза способна к их дифференциальному узнаванию на основе детерминант физической природы (в отличие от «текстовых», напрямую определяемых последовательностью нуклеотидов). Такое узнавание имеет важное регуляторное значение и необходимо для перехода фага из одной стадии жизненного цикла в другую. Необходимо отметить, что отнесение некоторого участка ДНК к промоторам и предсказание его промоторной силы не могут основываться исключительно на анализе свойств самого участка без учета его генетического окружения. В случае SIDD необходимо рассмотрение протяженных (сотни пар оснований) участков геномной ДНК. Также на примере трех различных методик расчета SIDD удалось установить хорошую согласованность результатов, полученных с помощью различных методов и алгоритмов (при наличии небольших расхождений для профилей отдельных промоторов).

## Список литературы (References)

- Адлер В. В., Поверенный А. М., Подгородниченко В. Н., Шапот В. С. Изучение процесса транскрипции с использованием антител к ДНК // Молекуляр. биология. — 1973. — Т. 7, вып. 2. — С. 203–208.
- Adler V. V., Poverennyj A. M., Podgorodnichenko V. N., Shapot V. S. Izuchenie processa transkripcii s ispol'zovaniem antitel k DNK [Transcription studying using antibodies against cDNA] // Molekuljar. biologija. — 1973. — Vol. 7, Iss. 2. — P. 203–208 (in Russian).
- Баев А. С., Любченко Ю. Л., Лазуркин Ю. С., Трифонов Э. Н., Франк-Каменецкий М. Д. Изучение легкоплавких участков ДНК фага Т2 с помощью электронной микроскопии и кинетического формальдегидного метода // Мол. Биол. — 1972. — № 6. — С. 760–766.
- Baev A. S., Ljubchenko Yu. L., Lazurkin Yu. S., Trifonov E. N., Frank-Kameneckij M. D. Izuchenie legkoplavkih uchastkov DNK faga T2 s pomoshh'ju jelektronnoj mikroskopii i kineticheskogo formal'degidnogo metoda [Studying easily melted DNA regions of bacteriophage T2 by means of microscopy and kinetic formaldehyde method] // Mol. Biol. — 1972. — Vol. 6. — P. 760–766 (in Russian).
- Камзолова С. Г., Артюх Р. И., Елфимова Л. И. Изучение матричных свойств Т2-ДНК, модифицированных 2,2',6,6'-тетрометил-4-бромацетокси-пиперидин-1-оксидом, в РНК-полимеразной системе *E. coli* II // Биохимия. — 1977. — Т. 42. — С. 1117–1122.
- Kamzolova S. G., Artjuh R. I., Elfimova L. I. Izuchenie matrichnyh svojstv T2-DNK, modificirovannyh 2,2',6,6'-tetrometil-4-bromacetoksipiperidin-1-oksilom, v RNK-polimeraznoj sisteme *E. coli* II [Studying blueprint properties of T2-DNA modified with 2,2',6,6'-tetromethyl-4-bromacetoxypiperidin-1-oxyl in *E. coli* RNA-polymerase system] // Biohimija. — 1977. — Vol. 42. — P. 1117–1122 (in Russian).
- Камзолова С. Г., Сорокин А. А., Осипов А. А., Бескаравайный П. М. Электростатическая карта генома бактериофага Т7. 1. Сравнительный анализ электростатических свойств сигма70-специфических промоторов Т7 ДНК, взаимодействующих с РНК-полимеразой *E. coli* // Биофизика — 2009. — № 54 (6). — С. 975–983.
- Kamzolova S. G., Sorokin A. A., Osipov A. A., Beskaravajnyj P. M. Jelektrostaticheskaja karta genoma bakteriofaga T7. 1. Sravnitel'nyj analiz jelektrostaticheskikh svojstv sigma70-spezificheskikh promotorov T7 DNK, vzaimodejstvujushchih s RNK-polimerazoj *E. coli* [Electrostatic map of bacteriophage T7 genome. Comparative analysis of electrostatic properties of sigma70-specific T7 DNA promoters recognized by RNA-polymerase of *Escherichia coli*] // Biofizika. — 2009. — Vol. 54 (6). — P. 975–983 (in Russian).
- Сорокин А. А., Дзелядин Т. Р., Орлов М. А., Зыкова Е. А., Камзолова С. Г. Пространственная организация электростатических взаимодействий Т7 РНК-полимеразы с поздними промоторами Т7 ДНК // Вестник биотехнологии и физико-химической биологии им Ю. А. Овчинникова. — 2016. — Т. 12 (4). — С. 64–71.

- Sorokin A. A., Dzhelyadin T. R., Orlov M. A., Zyкова E. A., Kamzolova S. G. Prostranstvennaya organizacija elektrostatičeskikh vzaimodejstvij T7 RNK-polimerazy s pozdnimi promotorami T7 DNK [Spatial organization of electrostatic interactions between T7 RNA-polymerase and late T7 DNA promoters] // Vestnik biotehnologii i fiziko-himicheskoj biologii im Yu. A. Ovchinnikova [Yu. A. Ovchinnikov Bulletin of Biotechnology and Physical and Chemical Biology]. — 2016. — Vol. 12 (4). — P. 64–71 (in Russian).
- Benham C. J. Theoretical Analysis of Heteropolymeric Transitions in Superhelical DNA Molecules of Specified Sequence // Journal of Chemical Physics. — 1990. — Vol. 92. — P. 6294–6305.
- Benham C. J. Energetics of the strand separation transition in superhelical DNA // J. Mol. Biol. — 1992. — Vol. 225. — P. 835–847.
- Benham C. J. Sites of predicted stress-induced DNA duplex destabilization occur preferentially at regulatory loci // Proc. Natl. Acad. Sci. U.S.A. — 1993. — Vol. 90 (7). — P. 2999–3003.
- Benham C. J. Duplex destabilization in superhelical DNA is predicted to occur at specific transcriptional regulatory regions // J. Mol. Biol. — 1996, January. — Vol. 255, No. 3. — P. 425–434.
- Bi C., Benham C. J. WebSIDD: server for predicting stress-induced duplex destabilized (SIDD) sites in superhelical DNA // Bioinformatics (Oxford, England). — 2004. — Vol. 20. — P. 1477–1479.
- Breslauer K. J., Frank R., Blocker H., Marky L. A. Predicting DNA duplex stability from the base sequence // Proc. Natl. Acad. Sci. U.S.A. — 1986. — Vol. 83. — P. 3746–3750.
- Cherny D. I., Alexandrov A. A., Zarudnaya M. I. et al. Investigation of the binding of *Escherichia coli* RNA polymerase to DNA from bacteriophages T2 and T7 by kinetic formaldehyde method and electron microscopy // Eur. J. Biochem. — 1977. — Vol. 79. — P. 309.
- Guerra J. C. O. Thermodynamics of Denaturation Transition of DNA Duplex Oligomers in the Context of Nearest Neighbor Models: A Short Review // Physics Procedia. — 2015. — Vol. 68. — P. 37–42.
- Imburgio D., Rong M., Ma K., McAllister W. T. Studies of promoter recognition and start site selection by T7 RNA polymerase using a comprehensive collection of promoter variants // Biochemistry. — 2000. — Vol. 39 (34). — P. 10419–10430.
- Kamzolova S. G., Postnikova G. B. Spin-labeled nucleic acids // Quart. Rev. Biophys. — 1981. — Vol. 14. — P. 223–288.
- Kamzolova S. G., Sivozhelezov V. S., Sorokin A. A. et al. RNA polymerase–promoter recognition. Specific features of electrostatic potential of early T4 phage DNA promoters // J Biomol Struct Dyn. — 2000, November. — Vol. 18, No. 3. — P. 325–334.
- Kamzolova S. G., Sorokin A. A., Dzhelyadin T. D. et al. Electrostatic potentials of *E. coli* genome DNA // J Biomol Struct Dyn. — 2005, November. — Vol. 23, No. 3. — P. 341–345.
- Kamzolova S. G., Beskaravainy R. M., Osypov A. A. et al. Electrostatic map of T7 DNA: Comparative analysis of functional and electrostatic properties of T7 RNA polymerase-specific promoters // Journal of Biomolecular Structure and Dynamics. — 2014. — Vol. 32. — P. 1184–1192.
- Margalit H., Shapiro B. A., Nussinov R. et al. Helix stability in prokaryotic promoter regions // Biochemistry. — 1988. — Vol. 27. No. 14. — P. 5179–5188.
- McAllister W. T., Carter A. D. Regulation of promoter selection by the bacteriophage T7 RNA polymerase in vitro // Nucleic Acids Res. — 1980. — Vol. 8 (20). — P. 4821–4837.
- McClure W. R. Mechanism and control of transcription initiation in prokaryotes // Annu. Rev. Biochem. — 1985. — Vol. 4. — P. 171–204.
- Michoel T., Van de Peer Y. Helicoidal transfer matrix model for inhomogeneous DNA melting // Phys Rev E Stat Nonlin Soft Matter Phys. — 2006. — Vol. 73:011908.
- Natale D. A., Schubert A. E., Kowalski D. DNA helical stability accounts for mutational defects in a yeast replication origin // Proc. Natl Acad. Sci. USA. — 1992. — Vol. 89. — P. 2654–2658.
- Osypov A. A., Krutinin G. G., Kamzolova S. G. DEPPDB — DNA electrostatic potential properties database: electrostatic properties of genome DNA // Journal of Bioinformatics and Computational Biology. — 2010. — Vol. 08, No. 03. — P. 413.

- Pérez-Martín J., Rojo F., de Lorenzo V.* Promoters responsive to DNA bending: a common theme in prokaryotic gene expression // *Microbiol. Rev.* — 1994. — Vol. 58. — P. 268–290.
- Pribnow D.* Biological Regulation and Development (R. Goldberger, ed.) // Plenum, New York, 1979. — Vol. 1. — P. 219–277.
- [RefSeq database]: <https://www.ncbi.nlm.nih.gov/refseq/> (accessed: 22.12.2017).
- SantaLucia J., Hicks D.* The thermodynamics of DNA structural motifs // *Annu Rev Biophys Biomol Struct.* — 2004. — Vol. 33. — P. 415–440.
- Shahmuradov I. A., Razali R. M., Bougouffa S., Radovanovic A., Bajic V. B.* bTSSfinder: a novel tool for the prediction of promoters in cyanobacteria and *Escherichia coli* // *Bioinformatics.* — 2017. — Vol. 33 (3). — P. 334–340.
- Sorokin A. A., Osypov A. A., Dzhelyadin T. R. et al.* Electrostatic properties of promoter recognized by *E. coli* RNA polymerase Esigma70 // *Journal of Bioinformatics and Computational Biology.* — 2006, April. — Vol. 4, No. 2. — P. 455–467.
- Sugimoto N., Nakano S., Yoneyama M., Honda K.* Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes // *Nucleic Acids Res.* — 1996. — Vol. 24. — P. 4501–4505.
- Szybalski W., Kubinski H., Sheldrick P.* Pyrimidine clusters on the transcribing strand of DNA and their possible role in the initiation of RNA synthesis // *Cold Spring Harb Symp Quant Biol.* — 1966. — Vol. 14. — P. 123–127.
- Wang H., Benham C. J.* Promoter prediction and annotation of microbial genomes based on DNA sequence and structural responses to superhelical stress // *BMC Bioinformatics.* — 2006. — Vol. 7. — P. 248.
- Yeremian E.* Genes and the physics of the DNA double-helix // *Gene.* — 2000. — September. — Vol. 255, No. 2. — P. 139–150.
- Zhabinskaya D., Madden S., Benham C. J.* SIST: stress-induced structural transitions in superhelical DNA // *Bioinformatics (Oxford, England).* — 2015. — Vol. 31 (3). — P. 421–2.