

СВЯЗЫВАНИЕ ЛИГАНДОВ С НУКЛЕИНОВЫМИ КИСЛОТАМИ В РАСТВОРЕ И НА МИКРОЧИПАХ

© 2022 г. Ю.Д. Нечипуренко*, **, А.С. Бучельников**

*Институт молекулярной биологии им. В.А. Энгельгардта РАН, Москва, ул. Вавилова, 32

E-mail: nech99@mail.ru

**Севастопольский государственный университет, 299053, Севастополь, ул. Университетская, 33

Поступила в редакцию 02.04.2022 г.

После доработки 02.04.2022 г.

Принята к публикации 04.04.2022 г.

Многие процессы в живой клетке включают в себя равновесное связывание лигандов (белков, биологически активных соединений и олигонуклеотидов) с нуклеиновыми кислотами. Модели и методы теории физической адсорбции позволяют количественно описать такое связывание в случае, когда лиганд занимает протяженный участок на ДНК и когда существуют кооперативные взаимодействия между адсорбированными молекулами лиганда. Применение модели Изинга позволяет построить решение уравнений адсорбции в самом общем виде, для широкого класса систем. Физико-химические эксперименты в модельных системах *in vitro*, где присутствуют, с одной стороны, такие биологические макромолекулы как белки, сложные белковые комплексы, полимеры РНК и ДНК, а с другой — малые молекулы лигандов, позволили пролить свет на множество механизмов, лежащих в основании клеточных процессов. Диаграммы футпринтинга или данные, полученные на микрочипах, позволяют получить детальную информацию о системе, где происходит связывание лигандов с ДНК. Краткий обзор теоретических исследований в этой области содержится в данной работе. Описано, как можно использовать расчеты энтропии адсорбционной системы, чтобы представить происходящие в ней процессы кооперативного связывания. В приложении дана методика анализа связывания на микрочипах.

Ключевые слова: связывание белков с ДНК, кооперативные взаимодействия, изотермы адсорбции, модели адсорбции, модель Изинга, микрочип.

DOI: 10.31857/S000630292203005X, EDN: ANDOHM

Развитие экспериментальных методов анализа связывания лигандов с нуклеиновыми кислотами после открытия структуры ДНК привело к появлению сотен теоретических работ, в которых рассматривалось связывание лигандов с линейными полимерами. Подходы статистической термодинамики к данным системам были применены впервые в работах Д. Брэдли и С. Лифсона [1] и С. Латта и Х. Собера [2]. Уравнения адсорбции, описывающие связывание лигандов, получили широкое распространение после работы Д. Крозерса [3] (следует заметить, что с точки зрения статистической термодинамики задача адсорбции протяженного лиганда на линейном полимере принадлежит классу решеточных моделей, и здесь впервые решение получено Л. Тонксом еще в 1936 г. [4]).

Важный этап развития теории адсорбции связан с учетом кооперативных взаимодействий между адсорбированными лигандами. Кооперативность может быть обусловлена, например, контактными взаимодействиями между лиганда-

ми, которые занимают на полимере соседние центры связывания. Лиганды при связывании на макромолекуле могут как способствовать связыванию, так и препятствовать ему, в первом случае говорят о положительной кооперативности, во втором — об отрицательной кооперативности или антикооперативности взаимодействия. Описание систем, в которых наблюдаются кооперативные взаимодействия, было проведено впервые в работах [5, 6]. Г.В. Гурский, А.С. Заседателей и М.В. Волькенштейн решили ряд теоретических задач, которые позволили описывать связывание биологически активных соединений с нуклеиновыми кислотами при помощи систем алгебраических уравнений и рекуррентных соотношений. Описанию кооперативного связывания лигандов с ДНК была посвящена также работа Дж. МакГи и П. фон Хиппеля [7]. В этой работе были наглядно истолкованы полученные Заседательным с соавторами уравнения для простейшего случая контактных кооперативных взаимодействий между адсорбированными лигандами.

В конце 70-х годов XX века двумя группами исследователей был обнаружен аллостерический эффект ДНК — возможность распространения вдоль ДНК кооперативных взаимодействий между адсорбированными лигандами. В работах отечественных исследователей этот эффект называли дальнедействием и описывали при помощи потенциала кооперативных взаимодействий [8, 9]. Аллостерические взаимодействия лежат в основании механизмов регуляции генов (см., например, работы [10–13]).

Дальнейшее развитие теории адсорбции в применении к связыванию лигандов с нуклеиновыми кислотами позволило учесть ситуации, когда лиганды образуют различные типы комплекса или связываются на ДНК в различных ориентациях [14, 15]. В русле этого подхода для широкого класса адсорбционных задач применялась одномерная модель Изинга [16]. Заметим, что эта модель для описания ферромагнетизма была предложена в 1920 г. В. Ленцем [17]. Э. Изинг смог найти решение уравнений, которые возникают в соответствующей этой модели математической задаче (поэтому в ряде источников используется название «модель Ленца–Изинга») [18]. Использование математического решения задачи адсорбции для этой модели позволило получить систему уравнений адсорбции в самом общем виде [16, 19, 20].

Заседателев, Гурский и Волькенштейн получили уравнения, описывающие кооперативное связывание протяженного лиганда на полимере, однако они не рассматривали дисперсию адсорбции (так как воспользовались приближением «бесконечного полимера»). Позже И. Эпштейн показал, что приближение бесконечного полимера хорошо описывает ситуацию, когда длина полимера уже в десяток раз превышает длину лиганда [21]. В ряде работ было показано, что дисперсия адсорбции может быть весьма информативной характеристикой системы [22–25].

При связывании на матрицах нуклеиновых кислот может быть задействовано одновременно несколько систем реакционных центров: в живой клетке с однотяжевыми и двутяжевыми участками молекул ДНК и РНК связывается большое число различных соединений. В монографии [26] рассматривается как конкуренция лигандов за одни и те же связывающие места, так и координация при связывании лигандов на одной или двух системах реакционных центров. Эта работа в некотором смысле содержит итог развития теории адсорбции в применении к описанию связывания лигандов с нуклеиновыми кислотами с 70-х годов прошлого века, в нее вошли как оригинальные работы, так и ряд обзоров [27, 28].

СВЯЗЫВАНИЕ ПРОТЯЖЕННЫХ ЛИГАНДОВ НА ПОЛИМЕРЕ

Пусть в растворе при постоянной температуре и постоянном давлении находятся молекулы лиганда и молекулы полимера с адсорбированными лигандами. С точки зрения физики процесс обратимого связывания лигандов на поверхности макромолекулы представляет собой явление физической адсорбции и относится к классу решеточных моделей: решетка-полимер состоит из реакционных центров (сайтов), которые могут связывать лиганд [26, 29].

Обозначим число лигандов, адсорбированных на полимере в равновесии, как q , а длину полимера — N . Предположим, что лиганд протяженный, то есть он при связывании закрывает L звеньев полимера, делая их недоступными для связывания других молекул лиганда.

В случае $L \ll N$ заполнение полимера лигандом r (которое определяется как q/N) и концентрация свободного лиганда в растворе x связаны следующим соотношением [5]:

$$\frac{r}{Kx} = \left(\frac{1-rL}{1-rL+r} \right)^L (1-rL+r), \quad (1)$$

где K — равновесная константа связывания лиганда с одним связывающим местом на полимере.

Модель, описывающая связывание протяженного лиганда на полимере, рассматривалась и до работ Заседателева с соавторами, соответствующие уравнения были получены ранее в работах [1–3], однако эти уравнения имели другую, параметрическую форму записи. Уравнение (1) можно назвать центральным уравнением в теории адсорбции протяженных лигандов. В англоязычной литературе за ним закреплено имя «уравнение МакГи–фон Хиппеля», так как эти авторы объяснили смысл данного уравнения при помощи условных вероятностей. Однако первыми в таком виде уравнения записали отечественные исследователи (как и уравнения, описывающие кооперативные взаимодействия [5]). Позже было показано, что уравнение (1) может быть использовано для построения асимптотического метода анализа изотерм адсорбции и для классификации моделей адсорбции [26, 30].

Интерпретировать уравнения адсорбции в рамках вероятностного подхода можно следующим образом (см., например, работу [26]). Рассмотрим полимер с адсорбированными лигандами как результат проведения схемы испытаний Бернулли с двумя исходами. Введем в рассмотрение вероятность встретить в последовательности из свободных звеньев полимера и адсорбированных лигандов свободное звено (не закрытое лигандом):

$$P = \frac{1 - rL}{1 - rL + r}. \quad (2)$$

Тогда уравнение (1) можно записать в следующем виде:

$$\frac{1}{Kx} = \frac{P^L}{1 - P}. \quad (3)$$

Параметр P изменяется от 0 до 1, и если задать значение этого параметра, то можно из уравнений (2) и (3) легко найти значения x и r и построить изотерму адсорбции.

Вероятностная интерпретация уравнения (1) заключается в том, что правая часть уравнения соответствует относительному числу свободных связывающих мест на полимере с адсорбированными лигандами (одно место составляют L расположенных подряд не закрытых лигандом звеньев полимера). Величина r дает относительное число адсорбированных лигандов. В таком случае уравнение (1) можно рассматривать как уравнение химического равновесия для лиганда длины L .

В реальной ситуации полимер имеет не бесконечную длину, благодаря этому в системе наблюдается дисперсия: разные полимеры в растворе несут разное число адсорбированных лигандов. Ранее было показано, как построить схему испытаний Бернулли и оценить сходимость величины q/N к значению r . Эту величину можно интерпретировать как вероятность события, что определенное звено полимера занято лигандом [26].

Заметим, что уравнения (1)–(3) являются довольно простыми алгебраическими уравнениями, и при усложнении моделей адсорбции — учете образования разных типов комплекса и кооперативных взаимодействий — вместо уравнений (1)–(3) появляется более сложная система уравнений (см., например, работы [19, 26, 31, 32]).

Отдельно следует упомянуть об избирательном связывании лигандов с ДНК, когда разным связывающим местам соответствуют разные энергии связывания. В этом случае принято говорить о связывании на гетерополимере. Для описания такого связывания вместо одной константы химического равновесия K , описывающей связывание лиганда с единичным связывающим местом на ДНК, вводят в рассмотрение набор констант $\{K_i\}$ и набор соответствующих изменений свободной энергии Гиббса $\{\Delta G_i\}$, которые связаны соотношением $K_i = \exp(-\Delta G_i/RT)$. Анализ связывания и расчет уравнений адсорбции в этом случае удобно проводить при помощи построения большой статистической суммы адсорбционной системы или с использованием метода рекуррентных соотношений [6]. Пример применения модели адсорбции лигандов на гетерополимере приведен в «Приложении».

Наиболее полную информацию об избирательном связывании лигандов на разных местах ДНК дают диаграммы футпринтинга, на которых представлены данные о том, как связанный лиганд закрывает отдельные межуэотидные связи в ДНК от расщепления разного рода агентами. Ранее в работах [33, 34] было показано, как можно рассчитывать такие диаграммы и сравнивать результаты с экспериментальными данными в работах (см. также монографию [26], где данному анализу посвящена отдельная глава).

ПРИМЕНЕНИЕ ТЕОРИИ АДсорбЦИИ ДЛЯ ОПИСАНИЯ БИОЛОГИЧЕСКИХ СИСТЕМ

Математические модели позволяют количественно описать процессы, происходящие в живой клетке и в модельных системах при связывании лигандов с ДНК. Кооперативные взаимодействия между лигандами были обнаружены и описаны в работах [9, 10, 26, 35]. Позже были получены данные, подтверждающие участие таких взаимодействий в процессах генетической регуляции, и Д. Крозерс уделил этому особое внимание [36] (см. также специальный номер журнала «Biopolymers», посвященный его памяти [37]).

Методы статистической термодинамики применялись для анализа широкого класса систем: рассматривалось связывание белков и антибиотиков с ДНК [38], проводился анализ расположения нуклеосом на сателлитной ДНК [39], было описано связывание матричной РНК с рибосомальными РНК в составе рибосомы [40]. Также эти методы были использованы для анализа кооперативных взаимодействий лигандов с рибосомой [41, 42], для описания шума экспрессии генов [25, 43], для анализа вхождения малых РНК в комплекс РНК-индуцированного сайленсинга генов RISC [44] и анализа данных по связыванию олигонуклеотидов на микрочипах [29, 45–49].

Множество работ в области применения теории адсорбции к описанию связывания лигандов с нуклеиновыми кислотами контрастирует с отсутствием обобщающих монографий в этой области. Хотелось бы видеть современные труды, где были бы описаны разные аспекты такого связывания, подобно тому, как описано связывание лигандов с белками в монографии Дж. Ваймана и С. Гилла [29]. В одной из глав монографии Г. Стормо рассматриваются термодинамические аспекты связывания белков с ДНК [50], однако в ней нет систематического изложения методов анализа такого связывания.

Заметим, что относительно недавно вышел ряд работ, в которых содержится описание систем, где связывание лигандов с матрицами нуклеиновых кислот происходит по новым схемам

(см., например, работы [51, 52]). Развита новая модель адсорбции, получены уравнения связывания. В монографии [26] такие модели впервые изложены системно: с единой точки зрения впервые описана совокупность методов и подходов статистической термодинамики, позволяющих анализировать системы, в которых изучается связывание лигандов с нуклеиновыми кислотами *in vitro*. На ряде примеров показано, как такие подходы работают, какие параметры связывания они позволяют оценить.

Методы анализа данных в области теории адсорбции дают в руки исследователя разнообразный теоретический «инструментарий», позволяющий анализировать как реальные биологические системы, так и модельные системы, в которых изучается связывание лигандов с нуклеиновыми кислотами. Примером такого инструментария является модель Изинга, которая хорошо известна в самых разных приложениях статистической физики. Любопытно, что развитие теории связывания лигандов с двумерными платформами микрочипов закономерно привело к применению двумерной модели Э. Изинга [29, 45–49].

Детальный анализ связывания биологически активных соединений с нуклеиновыми кислотами содержится в работах М.П. Евстигнеева с соавторами [53–55], отдельно следует упомянуть оригинальную концепцию интерцепторно-протекторного действия ДНК-связывающих лигандов [53], развиваемую этими авторами.

Обширный класс фундаментальных и прикладных задач молекулярной биофизики можно решать при помощи теории адсорбции. Любопытны новые подходы, которые позволяют связать адсорбцию с конденсацией ДНК [56–60] и с плавлением ДНК (см., например, работу [61]), исследовать противовирусную активность лигандов (см., например, работу [62]), связывание лигандов с квадруплексами [63–65]. Особый интерес представляют работы, в которых изучается кооперативное связывание белков с ДНК (см., например, работу [66]).

Статистическая термодинамика, описывающая связывания лигандов с ДНК и РНК, находится на переднем крае науки, в области развития новых технологий постгеномной эры. Модели адсорбции оказываются необходимыми и востребованными в анализе экспериментов по футпринтингу [33, 34], в объяснении экспериментов по гибридизации геномных ДНК и РНК на микрочипах [67], в описании регуляции экспрессии генов при помощи малых РНК (см., например, работу [44]), при конструировании новых лекарственных соединений и создании биосенсоров [47, 51] — т. е. во многих областях молекулярной биологии и биотехнологии. В связи с развитием

новых методов и технологий в молекулярной биологии появилась надежда решить и фундаментальные задачи: разобраться в том, как работают факторы транскрипции, малые РНК и лекарственные соединения, которые регулируют процессы экспрессии генов. Теоретические и экспериментальные методы, разработанные за последние годы, дают возможность вплотную подойти к решению этих задач. Заметим, что концептуально молекулярная биология уже созрела для количественного подхода к описанию регуляции генов (см., например, работу [68]) и, как мы отмечали во «Введении», аллостерические взаимодействия между лигандами могут играть важнейшую роль в такой регуляции [69].

Для строгого термодинамического описания кооперативного связывания необходимо понимание того, как изменяются функции состояния адсорбционной системы, среди которых ключевую роль, по нашему мнению, играет статистическая, или конфигурационная, энтропия. Рассмотрим далее, как описанный формализм можно применить для расчетов энтропии на примере простейшей модельной системы, в которой существуют кооперативные взаимодействия между адсорбированными лигандами.

СТАТИСТИЧЕСКАЯ ЭНТРОПИЯ ПРИ ОПИСАНИИ КООПЕРАТИВНОГО СВЯЗЫВАНИЯ

Ниже мы постараемся изучить поведение статистической энтропии на простом примере адсорбции малых молекул на линейном полимере и показать, что расчеты энтропии позволяют пролить свет на происходящие в системе процессы. Заметим, что данный обзор концептуально связан с работой [70], вышедшей пять лет назад. В настоящей публикации мы расширили и обновили список литературы, а также включили расчеты конфигурационной энтропии [71].

В дополнение к величинам K и x , введенным выше при рассмотрении модели адсорбции лигандов на гомополимере, обозначим через ω параметр кооперативности. Значения $\omega < 1$ соответствуют отрицательной кооперативности, или антикооперативности; при $\omega = 1$ адсорбция некооперативна, параметр $\omega > 1$ соответствует положительной кооперативности [72].

Все термодинамические параметры системы, в том числе статистическую энтропию, можно рассчитать, зная большую статистическую сумму системы Ξ_N . Существует несколько методов нахождения большой статистической суммы системы, в которой происходит адсорбция лигандов на полимерной решетке. Среди них матричный [18] и комбинаторный [73] методы являются наиболее известными.

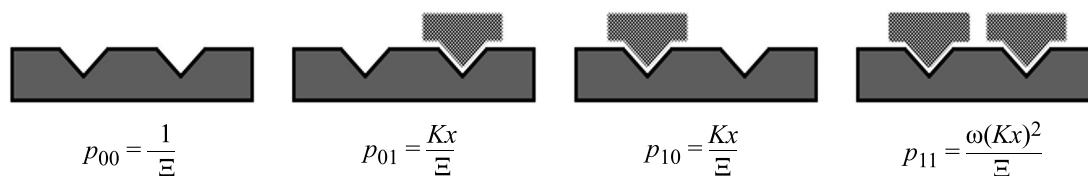


Рис. 1. Показаны разные состояния простейшей решетки и соответствующие им вероятности.

Рассмотрим простейшую систему с полимером, имеющим только два центра связывания ($N = 2$). Большая статистическая сумма такой системы может быть записана следующим образом:

$$\Xi = 1 + 2Kx + \omega(Kx)^2.$$

Определим статистическую, или конфигурационную энтропию исследуемой системы как меру неопределенности, возникающую при адсорбции лигандов на решетке. Действительно, если известно, что на одной решетке связан один лиганд, существует неопределенность: какой именно из двух реакционных центров занят лигандом, а какой свободен? Из статистической термодинамики известно, что такая энтропия S в расчете на моль решеточного полимера равна:

$$S = -R \sum_{i=0}^1 \sum_{j=0}^1 p_{ij} \ln p_{ij}, \quad (4)$$

где R – универсальная газовая постоянная, p_{ij} – вероятность пребывания системы в одном из четырех возможных состояний. Всего в рассматриваемой системе возможны четыре состояния, укажем для каждого из них свою вероятность (см. рис. 1).

Найдем энтропию системы в явном виде на основании формулы (4):

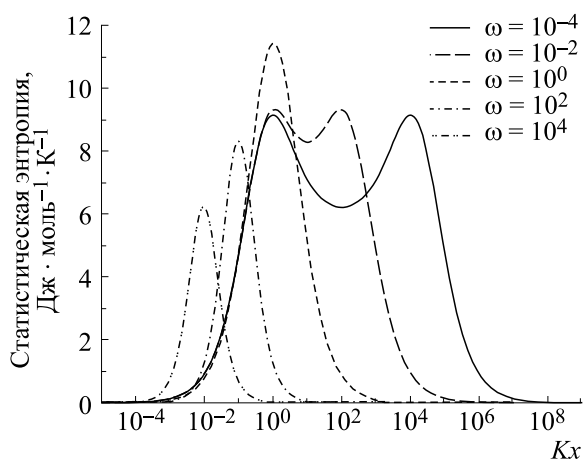


Рис. 2. Зависимость статистической энтропии от произведения Kx при различных значениях параметра кооперативности ω .

$$S = -R(p_{00} \ln p_{00} + p_{01} \ln p_{01} + p_{10} \ln p_{10} + p_{11} \ln p_{11})S,$$

или

$$S = -R \left[\frac{Kx}{\Xi} (2(1 + \omega Kx) \ln Kx + \omega Kx \ln \omega) - \ln \Xi \right]. \quad (5)$$

Формула (5) позволяет представить статистическую энтропию как функцию двух независимых переменных – $S(Kx, \omega)$. Зависимость энтропии от величины Kx при разных значениях ω показана на рис. 2.

Мы видим, что зависимость энтропии от концентрации лиганда в растворе и от заполнения матрицы лигандом оказывается нетривиальной: при значениях параметра кооперативности $\omega < 1$ график имеет два максимума.

На рис. 3 показаны зависимости плотности вероятности состояний решетки от Kx . При малом параметре кооперативности ($\omega \ll 1$) три состояния последовательно сменяют друг друга. По мере увеличения концентрации лиганда начинает снижаться вероятность того, что решетка полностью пуста, и увеличиваться вероятность того, что хотя бы один центр связывания решетки занят. С дальнейшим ростом концентрации лиганда увеличивается вероятность полностью занятой решетки. В пределе, при $\omega \rightarrow 0$, можно ожидать, что последнее событие наступит при бесконечно большой концентрации лиганда, что невозможно, и это вполне отвечает нашим представлениям. Действительно, при бесконечно антикооперативном процессе существует запрет на существование двух соседних занятых сайтов – модель «исключенного соседа» [74]. Эта модель может служить переходом к решеточной модели «газа твердых стержней», в которой один лиганд при связывании занимает сразу несколько реакционных центров [2–4].

При увеличении параметра кооперативности вероятность полного заполнения решетки лигандами наступает уже при меньших концентрациях. При этом состояние решетки с одним занятым сайтом становится все менее вероятным. При дальнейшем увеличении параметра кооперативности вероятность встретить в растворе наполовину заполненную решетку падает практически до нуля. В пределе, при $\omega \rightarrow \infty$ состояние системы определяется двумя возможными конфигурациями: полностью пустая либо полностью заполнен-

ная решетка, причем смена конфигураций, т. е. кооперативный переход, происходит при бесконечно малой концентрации лиганда.

Из сопоставления рис. 2 и 3 следует важный вывод: энтропия системы максимальна в месте пересечения плотностей вероятности, т.е. когда соответствующие события равновероятны. Равная вероятность наступления событий означает отсутствие преимущественного состояния решетки и приводит к неопределенности, количественной мерой которой является статистическая энтропия. При отрицательной или положительной кооперативности процессы связывания лигандов с решеткой зависят друг от друга, в то время как при некооперативном связывании эти процессы протекают независимо. По этой причине наибольшее значение энтропии в максимуме (см. рис. 2 и 3) наблюдается именно для некооперативного процесса ($\omega = 1$).

Интересно рассмотреть зависимость максимальных значений энтропии S_{\max} , т.е. значений энтропии в пике каждого графика на рис. 2, от параметра кооперативности ω . Можно показать, что

$$S_{\max}(Kx, \omega) = \begin{cases} S\left(\frac{1}{\omega}, \omega\right) & \text{при } \omega \leq 1, Kx = 1 \\ S\left(\frac{1}{\sqrt{\omega}}, \omega\right) & \text{при } \omega \geq 1 \end{cases}$$

График зависимости максимальных значений энтропии с точностью до универсальной газовой постоянной от параметра кооперативности изображен на рис. 4.

Из рисунка видно, что максимальная статистическая энтропия соответствует системе с некооперативными взаимодействиями ($\omega = 1$), в которой равновероятны все четыре состояния системы. По мере уменьшения параметра кооперативности ($\omega \ll 1$) состояние с полностью заполненной решеткой становится все менее вероятным, а при увеличении параметра кооперативности ($\omega \gg 1$) все менее вероятными становятся состояния с наполовину заполненной решеткой.

Как видно из рис. 4, максимальная энтропия системы с одним молекул решеточных полимеров с точностью до постоянной Авогадро находится в полном согласии с уравнением Больцмана:

$$S = R \ln \Omega = N_A k \ln \Omega,$$

где Ω — термодинамическая вероятность, т. е. число микроскопических состояний системы, которым можно реализовать данное макроскопическое состояние, в нашем случае — число конфигураций адсорбционной системы при данном значении параметра кооперативности.

Заметим, что в реальных системах, т.е. при связывании лигандов с рецепторами, субстратов с рибосомами и др., связывание лигандов с пер-

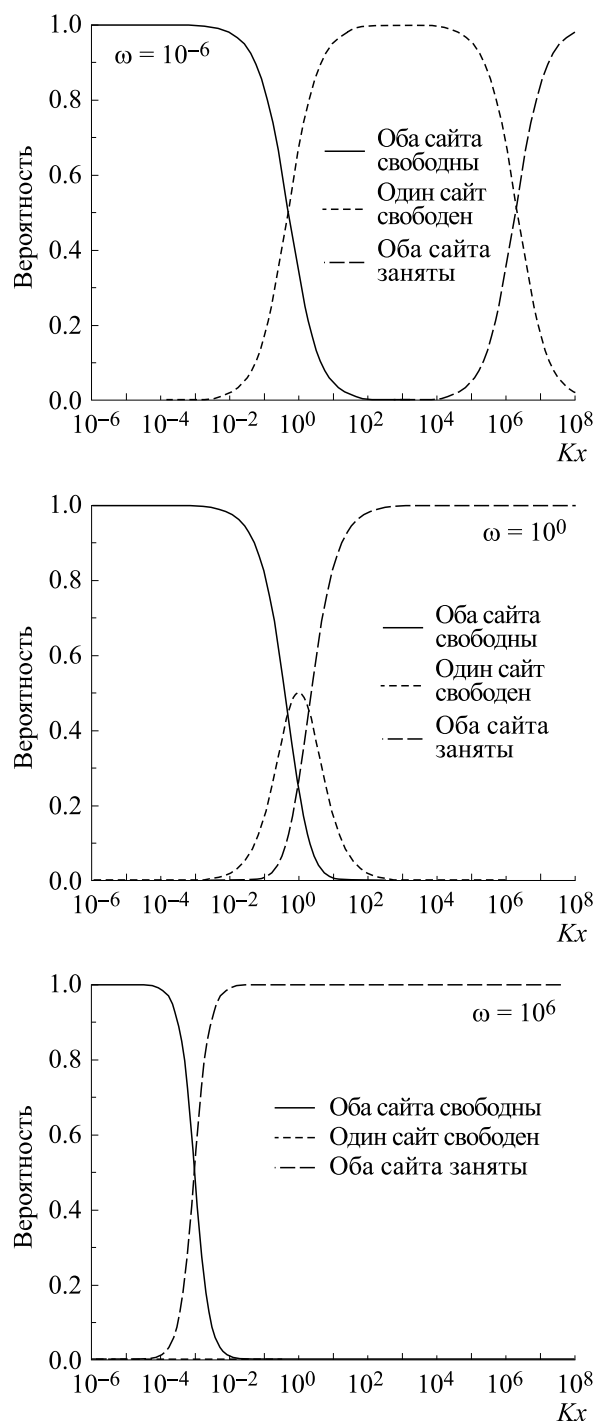


Рис. 3. Зависимости плотностей вероятности состояний решетки от концентрации лиганда.

вым и вторым центрами может происходить с разными константами. Пример такого случая мы рассмотрели в работах [41, 42].

При анализе экспериментальных данных исследователю необходимо решать так называемую обратную задачу: восстановить параметры моде-

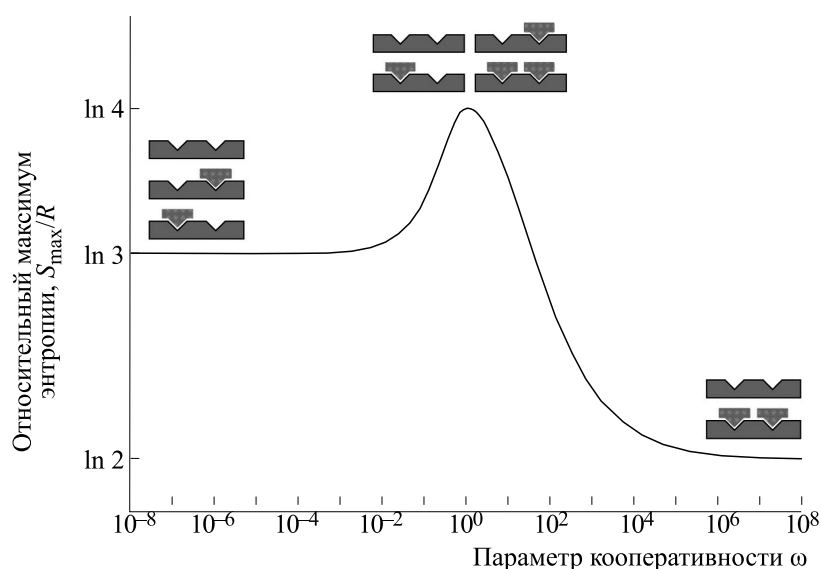


Рис. 4. Зависимость максимума энтропии от параметра кооперативности.

ли на основании кривых связывания. В этом случае возникает неопределенность, и параметр кооперативности можно определить только при определенных допущениях. Существует также известная проблема, связанная с тем, что результаты расчета кривых в модели, в которой существуют антикооперативные взаимодействия между лигандами, совпадают с результатами расчетов в случае, если есть гетерогенное связывание, но нет кооперативности. В этом случае на помощь исследователю могут прийти эксперименты по кинетике связывания [75].

Ниже, в «Приложении», мы продемонстрируем применение методов статистической термодинамики к анализу связывания на платформах микрочипов молекул матричных РНК в ситуации, когда необходимо провести различие между высокоспецифичным связыванием (сигналом) и низкоспецифичным (шумом). Заметим, что данное «Приложение» может иметь большое практическое значение для анализа данных по связыванию мишеней на биочипах. Кроме известных зарубежных работ в этой области и работ, которые мы используем для развития методов анализа данных [67, 76–78], следует упомянуть примеры успешного применения биочипов в Институте молекулярной биологии им. В.А. Энгельгардта РАН (см., например, работы [79–83]).

ЗАКЛЮЧЕНИЕ

Применение теории адсорбции для описания биологических систем обусловлено участием специалистов в области статистической физики в решении задач, связанных с биополимерами. В настоящее время необходимо привлечь внимание

молекулярных биологов и биофизиков к теории адсорбции: по сути, она помогает решать множество актуальных задач, о чем мы писали выше. В отечественной и мировой литературе почти не существует крупных обзорных работ, в которых с единой точки зрения рассматривались бы сотни примеров успешного применения теории адсорбции к описанию взаимодействия лигандов с нуклеиновыми кислотами. Наша работа продолжает серию обзоров и глав в монографиях, в которых мы ставим перед собой цель заполнить лакуны в описании применения теории адсорбции к связыванию лигандов с нуклеиновыми кислотами [26, 28, 70].

Теория адсорбции применялась к биологическим системам на протяжении всего XX века и накопила уже достаточно опыта, чтобы выступить в этой области в роли научной дисциплины. Эта дисциплина выработала свой язык, который позволил сформулировать взгляд на такую фундаментальную проблему биологии, как регуляция генетической экспрессии, в рамках представлений статистической физики (см., например, работы [25, 36, 37, 43]). С другой стороны, применение теории адсорбции к анализу экспериментальных данных позволяет связать наблюдаемые величины с модельными свойствами молекулярно-биологических систем: определять энергии взаимодействия лигандов с нуклеиновыми кислотами, энергии взаимодействия между адсорбированными лигандами и т.п. Описание реальных систем при помощи теории адсорбции дает возможность как предсказывать их поведение при изменении внешних условий, так и уточнять черты моделей связывания лигандов с ДНК. Без применения теории адсорбции невозможно сейчас

представить эффективное решение и таких прикладных задач, как анализ связывания биологически активных соединений на микрочипах и создание биосенсоров.

ПРИЛОЖЕНИЕ

Сравнение кросс-гибридизации со специфической гибрилизацией. Рассмотрим экспериментальную систему, в которой происходит гибрилизация — связывание молекул-мишеней из раствора с олигонуклеотидами-зондами, расположенными на платформе микрочипа [67] (см. также работу [76]). Заметим, что под терминами микрочип и биочип мы понимаем здесь одно и то же: матрицы ячеек на плоской поверхности, в каждой из которых иммобилизованы молекулярные зонды определенного типа.

Для оценки специфичности гибрилизации здесь мы будем использовать две базы данных, полученных в двух эквивалентных экспериментальных системах, представляющих микрочипы с одними и теми же пробами [67]. В первой системе к микрочипу добавляются мишени, которые имеют специфические места связывания для проб, во второй — те же самые мишени, только без специфических мест связывания (что можно считать контролем). Соответственно, в первой базе представлены данные для проб, каждая из которых может находиться в одном из трех состояний: быть свободной (не связанной с мишенью) (I), быть связанной специфично (II) и связанной неспецифично (III). Во второй базе представлены данные для проб, для которых нет специфических мишеней — и для них возможны только состояния (I) и (III).

Рассмотрим связывание одной молекулы зонда с мишенью для первой экспериментальной системы. Будем считать, что это связывание проходит в равновесных условиях. Для того, чтобы оценить вероятность связывания зонда на мишени по специфическому механизму, с образованием совершенного дуплекса, по сравнению с вероятностями связывания на всех неспецифических местах (приводящего к кросс-гибридизации), применим следующий прием.

Предположим, что каждая длинная молекула мишени связывается только с одним зондом. В этом случае мы можем представить, что молекула зонда равновесно связывается с набором молекул-мишеней, каждая из которых имеет свой номер $i = 1, \dots, N$. Таким образом, мы заменяем реальную молекулу мишени на набор коротких молекул, каждая из которых соответствует одному месту связывания с молекулой зонда. Подобную схему связывания рассматривали ранее авторы работы [77]. Однако в отличие от этой работы, где оценивалось связывание только зондов с мише-

нями, которые не содержали мест специфического связывания, в нашей системе будет присутствовать мишень со специфичным местом $i = s$.

Молекула зонда может пребывать в $N + 1$ состоянии: либо она будет свободна, либо связана с одной мишенью из набора N типов (на который мы разбили реальную мишень). Методы статистической термодинамики позволяют оценить вероятность того, что зонд будет связан с i -й мишенью. Связывание зонда с i -й мишенью характеризуется равновесной константой K_i .

Воспользуемся формализмом статистических сумм. Представим большую статистическую сумму системы в виде:

$$\Xi_N = 1 + \sum_i K_i x,$$

где x — концентрация свободных молекул мишеней в растворе, статистический вес 1 соответствует ситуации, когда мишень не связана с зондом. Тогда вероятность обнаружить зонд связанным с i -й мишенью равна

$$P_i = \frac{d \ln \Xi_N}{d \ln K_i} = \frac{K_i x}{1 + \sum_i K_i x}.$$

Если рассматривается связывание со специфичной матрицей $i = s$, то

$$P_s = \frac{K_s x}{1 + \sum_i K_i x}.$$

Здесь в знаменателе в общей сумме есть член, соответствующий $j = s$, то есть $K_s x$.

Вероятность кросс-гибридизации равна:

$$P_{\text{cross}} = \frac{\sum_{j \neq s} K_j x}{1 + \sum_i K_i x}. \quad (\text{A1})$$

Если мы сигналом будем называть вероятность связывания зонда на специфическом месте, а шумом — вероятность кросс-гибридизации, то оценку величины отношения «сигнал—шум» дает отношение соответствующих вероятностей:

$$R_{SN} = \frac{P_s}{P_{\text{cross}}} = \frac{K_s x}{\sum_{j \neq s} K_j x}.$$

Для интерпретации экспериментальных данных, полученных в двух описанных выше системах, то есть на двух микрочипах, применим наш подход.

Вероятность кросс-гибридизации P_{cross} мы можем оценить на основании данных из эксперимента во второй системе, где зонды связываются с мишенями, в которых нет специфического места. Если предположить, что в этой системе экс-

перимент проводится при аналогичных условиях, то весь сигнал гибридизации в этом эксперименте будет равен сигналу кросс-гибридизации:

$$P'_{\text{cross}} = \frac{\sum_{j \neq s} K_j x'}{1 + \sum_{i \neq s} K_i x'}$$

где x' — концентрация свободных мишеней во второй системе. Если предположить, что концентрации мишеней в этих системах мало отличаются, т. е. что $x' \sim x$, то, используя соотношение (A1), можно показать, что отношение сигнала кросс-гибридизации в первой системе к сигналу кросс-гибридизации во второй системе равно:

$$\frac{P_{\text{cross}}}{P'_{\text{cross}}} = \frac{1 + \sum_{j \neq s} K_j x}{1 + \sum_i K_i x} \quad (\text{A2})$$

Перейдем к экспериментальным оценкам. В первой системе зонд может находиться в трех состояниях: он может быть не связан с мишенью (обозначим вероятность этого состояния как P_0), быть связанным со специфичным местом (соответствующая вероятность равна P_s) и быть связанным неспецифично (соответствующая вероятность равна P_{cross}):

$$P_0 + P_s + P_{\text{cross}} = 1.$$

Сигнал от общей гибридизации в первой системе пропорционален величине $P_s + P_{\text{cross}}$. Вычитая из этого сигнала P'_{cross} , мы получаем некоторую оценку для P_s , которая в соответствии с соотношением (A2) оказывается тем точнее, чем меньше отношение сигнала к шуму R_{SN} . По сути, тогда будет сделано следующее приближение: мы считаем полный набор состояний в первой системе для эксперимента с микрочипом, когда в мишени есть специфические места. Потом сигнал кросс-гибридизации на этом чипе (его оценивает формула (A1)) мы заменяем на сигнал кросс-гибридизации, полученный в эксперименте, где нет специфического места [67]. Уточнить оценки параметров мы сможем, если учтем детали эксперимента. Следует оценить отличие x' и x в этих системах, которое обусловлено двумя факторами: во-первых, зонды, специфично связанные с мишенями, выключают из реакции протяженные участки ДНК, которые прилегают к специфическому месту связывания и тем самым уменьшают эффективное значение x , во-вторых, в системе, где нет участков ДНК со специфичными местами связывания, вообще уменьшается общая концентрация мишеней в растворе, которая равна сумме концентраций мишеней, связанных на зондах и свободных мишеней в растворе. Учет двух этих

факторов может позволить уточнить оценки отношения «сигнал—шум» в рассматриваемой системе.

Заметим, что оценка специфического и неспецифического связывания была проведена в работе [78], авторы которой развивали модель, сходную с нашей. Однако в этой работе сумму констант, соответствующих неспецифическому связыванию, заменяли на одну константу с усредненной энергией.

Мы описали здесь модель, в которой молекула зонда может находиться в одном из трех состояний: быть свободной (I), связываться со специфическими мишенями (II), либо с неспецифическими (III), и эта модель применялась для анализа данных в работе [67].

БЛАГОДАРНОСТИ

Авторы считают своим приятным долгом поблагодарить А.С. Заседателя, Н.Г. Есипову и В.В. Аристову за полезные дискуссии, А. Анашкину и О. Матвееву — за ценные замечания и обсуждение.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена при частичной финансовой поддержке Программы фундаментальных научных исследований в Российской Федерации на долгосрочный период (2021–2030 годы) (тема 121052600299-1).

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Настоящая статья не содержит каких-либо исследований с участием людей или животных в качестве объектов исследований.

СПИСОК ЛИТЕРАТУРЫ

1. D. F. Bradley and S. Lifson, in *Molecular Associations in Biology*, Ed. by B. Pullman (Acad. Press, London, New York, 1968), pp. 261–270.
2. S. A. Latt and H. A. Sober, *Biochemistry* **6** (10), 3293 (1967).
3. D. M. Crothers, *Biopolymers* **6** (4), 575 (1968).
4. L. Tonks, *Phys. Rev.* **50** (10), 955 (1936).
5. А. С. Заседатель, Г. В. Гурский и М. В. Волькенштейн, *Молекуляр. биология* **5** (20), 245 (1971).
6. Г. В. Гурский, А. С. Заседатель и М. В. Волькенштейн, *Молекуляр. биология* **6**, 479 (1972).
7. J. D. McGhee and P. H. von Hippel, *J. Mol. Biol.* **86** (2), 469 (1974).

8. A. S. Krylov, S. L. Grokhovsky, A. S. Zasedatelev et al., *Nucl. Acids Res.* **6** (1), 289 (1979).
9. M. Hogan, N. Dattagupta, and D. M. Crothers, *Nature* **278**, 521 (1979).
10. X. Qu, J. O. Trent, I. Fokt et al., *Proc. Natl. Acad. Sci. USA* **97**, 12032 (2000).
11. D. M. Chenoweth and P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **106**, 13175 (2009).
12. A. Balaceanu, A. Pérez, P. D. Dans et al. *Nucl. Acids Res.* **46** (15), 7554 (2018).
13. J. Dale, C. P. Howe, H. Toncova, et al., *Phys. Chem. Chem. Phys.* **23** (28), 15352 (2021).
14. Г. В. Гурский и А. С. Заседателев, *Биофизика* **23**, 932 (1978).
15. G. V. Gursky and A. S. Zasedatelev, *Sov. Sci. Rev. D, Physicochem. Biol.* **5**, 53 (1984).
16. Ю. Д. Нечипуренко, А. С. Заседателев и Г. В. Гурский, *Биофизика* **24**, 351 (1979).
17. W. Lenz, *Z. Phys.* **21**, 613 (1920).
18. E. Ising, *Z. Phys.* **31**, 253 (1925).
19. Ю. Д. Нечипуренко и др., *Молекуляр. биология* **18**, 332 (1984).
20. Yu. D. Nechipurenko and G. V. Gursky, *Biophys. Chem.* **24** (3), 195 (1986).
21. I. R. Epstein, *Biophys. Chem.* **8** (4), 327 (1978).
22. Б. Йованович и Ю. Д. Нечипуренко, *Молекуляр. биология* **24**, 478 (1990).
23. V. Arakelyan, Yu. Babayan, and G. Potikyan, *J. Biomol. Struct. Dyn.* **18**, 231 (2000).
24. Ю. Д. Нечипуренко, А. М. Вольф и Ю. М. Евдокимов, *Биофизика* **48** (5), 802 (2003).
25. Ю. Д. Нечипуренко, А. М. Вольф и Г. В. Гурский, *Биофизика* **48** (6), 986 (2003).
26. Ю. Д. Нечипуренко, *Анализ связывания биологически активных соединений с нуклеиновыми кислотами* (Ин-т компл. исслед., М.—Ижевск, 2015).
27. Ю. Д. Нечипуренко, *Биофизика* **27**, 391 (1982).
28. Ю. Д. Нечипуренко и Г. В. Гурский, *Биофизика* **48** (5), 773 (2003).
29. J. Wyman and S. J. Gill, *Binding and Linkage: Functional Chemistry of Biological Macromolecules* (University Science Books, Mill Valley CA, 1990).
30. Ю. Д. Нечипуренко, *Молекуляр. биология* **18**, 1066 (1984).
31. A. R. Wolfe and T. Meehan, *J. Mol. Biol.* **223** (4), 1063 (1992).
32. D. Yu. Lando and Yu. D. Nechipurenko, *J. Biomol. Struct. Dyn.* **26**, 187 (2008).
33. В. Ф. Рябоконт, Ю. Д. Нечипуренко и Г. В. Гурский, *Докл. РАН* **398**, 832 (2004).
34. Yu. D. Nechipurenko, B. Jovanovic, V. F. Riabokon, et al., *Ann. N. Y. Acad. Sci.* **15**, 660 (2005).
35. Ю. Д. Нечипуренко, А. С. Крылов, А. С. Заседателев и др., *Молекуляр. биология* **18**, 332 (1984).
36. D. M. Crothers, *Science* **339** (6121), 766 (2013).
37. J. V. Chaires and S. D. Levene, *Biopolymers* **103**, 469 (2015).
38. Ю. Д. Нечипуренко и Г. В. Гурский, *ДАН СССР* **281**, 213 (1985).
39. Ю. Д. Нечипуренко и М. В. Волькенштейн, *ДАН СССР* **286**, 216 (1986).
40. Ю. Д. Нечипуренко, Н. В. Попов, М. А. Исаев и др., *Биофизика* **40** (6), 1208 (1995).
41. S. B. Bourd, M. K. Kukhanova, B. P. Gottikh, et al., *Eur. J. Biochem.* **135** (3), 465 (1983).
42. М. К. Куханова, С. Б. Бурд, Л. С. Викторова и др., *Молекуляр. биология* **18** (3), 691 (1984).
43. М. В. Головкин, Ю. Д. Нечипуренко и Г. В. Гурский, *Биофизика* **54** (4), 581 (2009).
44. O. Matveeva, Yu. Nechipurenko, L. Rossi, et al., *Nucl. Acids Res.* **35** (8), e63 (2007).
45. М. В. Головкин, О. В. Матвеева и Ю. Д. Нечипуренко, *Биофизика* **54** (5), 820 (2009).
46. М. В. Ходыков, А. А. Анашкина, М. В. Головкин и др., *Биофизика* **56** (6), 1053 (2011).
47. Я. В. Стирманов, М. В. Ходыков, О. В. Матвеева и др., *Биофизика* **58** (6), 981 (2013).
48. Ya. V. Stirmanov, O. V. Matveeva, and Yu. D. Nechipurenko, *J. Biomol. Struct. Dyn.* **37** (12), 3103 (2019).
49. Yu. D. Nechipurenko, in *Advances in Medicine and Biology*, Ed. by L. V. Berhardt (Nova Science Publishers, Hauppauge, New York, 2019), vol. 136, pp. 193–206.
50. G. D. Stormo, *Introduction to protein-DNA interactions: Structure, thermodynamics, and bioinformatics* (Cold Spring Harbor Laboratory Press, N. Y., 2013).
51. Yu. M. Yevdokimov, S. G. Skuridin, Yu. D. Nechipurenko, et al., *Int. J. Biol. Macromol.* **36**, 103 (2005).
52. V. B. Teif, N. Kepper, K. Yserentant, et al., *J. Phys.: Condens. Matter* **27**, 064110 (2015).
53. М. Р. Евстигнеев, А. С. Buchelnikov, and V. P. Evstigneev, *Prog. Biophys. Mol. Biol.* **149**, 131 (2019).
54. С. Ф. Барановский, Д. Н. Чернышев, А. С. Бучельников и др., *Биофизика* **56** (2), 235 (2011).
55. М. Р. Евстигнеев and A. V. Shestopalova, in *Application of Computational Techniques in Pharmacy and Medicine. Challenges and Advances in Computational Chemistry and Physics*, Ed. by L. Gorb, V. Kuz'min, and E. Muratov (Springer, Dordrecht, 2014), vol. 17, pp. 21–57.
56. D. Yu. Lando and V. B. Teif, *J. Biomol. Struct. Dyn.* **17**, 903 (2000).
57. В. Б. Тейф и Д. Ю. Ландо, *Молекуляр. биология* **35**, 117 (2001).
58. Ю. Д. Нечипуренко, А. М. Вольф, В. И. Саянов и др., *Журн. эксперим. и теорет. физики* **125** (1), 103 (2004).
59. Е. А. Воробьев, Ю. Д. Нечипуренко, В. И. Саянов и др., *Биофизика* **52** (4), 636 (2007).
60. V. B. Teif and K. Bohinc, *Prog. Biophys. Mol. Biol.* **105**, 208 (2011).
61. E. Bereznyak, N. Gladkovskaya, E. Dukhopelnykov, et al., *AIMS Biophysics* **2**, 423 (2015).
62. А. Н. Суровая, Н. П. Бажулина, С. Ю. Лепехина и др., *Биофизика* **61** (2), 270 (2016).
63. А. Г. Kudrev, *Talanta* **116**, 541 (2013).
64. Н. С. Ильинский, А. М. Варижук, А. Д. Вениаминов и др., *Молекуляр. биология* **48** (6), 891 (2014).
65. A. Beniaminov, G. Chashchina, A. Shchyolkina, et al., *Biochimie* **191**, 33 (2021).
66. M. Melikishvili and M. G. Fried, *Biopolymers* **103** (9), 509 (2015).
67. O. Matveeva, Yu. Nechipurenko, N. Nazipova, et al., *Bioinformatics* **32** (17), i552 (2016).
68. E. Segal and J. Widom, *Nat. Rev. Genet.* **10** (7), 443 (2009).

69. S. Kim, E. Broströmer, D. Xing, et al., *Science* **339** (6121), 816 (2013).
70. Ю. Д. Нечипуренко, в кн. *Физико-химические механизмы и регуляция процессов трансформации энергии в биологических структурах*, под ред. Г. Ю. Ризниченко и А. Б. Рубина (Ин-т компл. исслед., М.–Ижевск, 2017), сс. 102–125.
71. V. V. Aristov, A. S. Buchelnikov, and Yu. D. Nechipurenko, *Entropy* **24** (2), 172 (2022).
72. S. H. Na and J. E. Ferrell, *Science* **352** (6288), 990 (2016).
73. G. Scatchard, *Ann. N. Y. Acad. Sci.* **51** (4), 660 (1949).
74. M. S. Rocha, *Biopolymers* **93** (1), 1 (2010).
75. F. Sevlever, J. P. Di Bella, and A. C. Ventura, *PLoS Comput. Biol.* **16** (6), e1007929 (2020).
76. J. Hooyberghs, P. Van Hummelen, and E. Carlon, *Nucl. Acids Res.* **37**, 53 (2009).
77. C. Furusawa, N. Оно, S. Suzuki, et al., *Bioinformatics* **25** (1), 36 (2009).
78. L. Zhang, M. F. Miles, and K. D. Aldape, *Nat. Biotechnol.* **21**, 818 (2003).
79. Д. А. Грядунов, Б. Л. Шаскольский, Т. В. Наседкина и др., *Acta Naturae* **10** (4), 4 (2018).
80. С. А. Лапа, Р. А. Мифтахов, Е. С. Клочихина и др., *Молекуляр. биология* **55** (6), 944 (2021).
81. Д. О. Фесенко, И. С. Абрамов, В. Е. Шершов и др., *Молекуляр. биология* **52** (6), 997 (2018).
82. Д. О. Фесенко, Т. О. Гусейнов, С. А. Лапа и др., *Молекуляр. биология* **52** (3), 533 (2018).
83. D. O. Fesenko, A. V. Chudinov, S. A. Surzhikov, et al., *Genet. Test. Mol. Biomarkers* **20** (4), 208 (2016).

Ligand Binding to Nucleic Acids in Solution and on Microchips

Yu.D. Nechipurenko*, ** and A.S. Buchelnikov**

**Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, ul. Vavilova 32, Moscow, 119991 Russia*

***Sevastopol State University, Universitetskaya ul. 33, Sevastopol, 299053 Russia*

The equilibrium binding of ligands (proteins, biologically active compounds, and oligonucleotides) to nucleic acids underlies many processes in the living cell. Model study and physical adsorption theory methods can be used to give quantitative description of this binding in cases when the ligand occupies an extended region on DNA and when cooperative interactions occur between adsorbed ligand molecules. With the Ising model it is possible to find an analytical solution to the equations of adsorption in the most general form for a wide range of models. Physico-chemical experiments in *in vitro* model systems containing biological macromolecules such as proteins, the complex web of protein interactions, RNA and DNA polymers, on the one hand, and small molecule ligands on the other hand, shed light on many mechanisms underlying cellular processes. Footprinting diagrams or microchip data sheet provide the detailed information on the system where ligands bind to DNA. This article presents a brief overview of theoretical research in this area. How to use entropy formula for the adsorption system is shown to learn more about the cooperative binding processes occurring in this system. An approach for the analysis of binding on microchips is described in the Appendix.

Keywords: DNA-protein binding, cooperative interactions, binding curve, adsorption model, Ising model, microchip