



НАНОСТРУКТУРЫ

математическая физика и моделирование



Nanostructures. Mathematical physics & modelling

НАНОСТРУКТУРЫ

математическая физика и моделирование

Nanostructures.
Mathematical Physics & Modelling

2015, volume 12(2)

Наноструктуры. Математическая физика и моделирование

Объединенная редколлегия:

Д.В. Быков, И.В. Волович, В.Л. Голо, В.В. Гусаров, П.Н. Дьячков, М.В. Карасев (*зам. главного редактора*), Ю.А. Косевич, Ю.Е. Лозовик, М.А. Мазо, В.П. Маслов (*главный редактор*), А.В. Махиборода (*ответств. секретарь*), А.Ю. Морозов, С.А. Никитов, Ю.И. Сезонов, А.К. Скуратов, И.С. Смирнов, Р.А. Сурис, В.А. Тулин, В.Е. Фортон, А.Р. Хохлов, А.В. Чаплик, Л.А. Чернозатонский, К.В. Шайтан; *иностранные члены*: А.С. Бакай, А. Шимизу; *информационное обеспечение наноиндустрии*: В.Н. Азаров, А.Д. Иванников, Ю.Л. Ижванов, В.П. Кулагин, А.Н. Тихонов

Электронная версия журнала размещается на сайте
<http://nano-journal.ru>

Адрес редакции:

Москва 109028, Б. Трехсвятительский пер., 3/12,
НИУ ВШЭ, Московский институт электроники и математики,
nano@miem.edu.ru

Москва

Содержание

В.Н. Блинов, В.Л. Голо, Ю.Ф. Крупянский Торические конформации молекул ДНК	5
О. Вальба, М. Тамм, С. Нечаев Топологические свойства РНК-подобных молекул со случайной первичной структурой	21
В.Б. Ошурко О механизме селективного биологического действия микроволнового и терагерцевого излучения	65
С.М. Першин ВЧ и СВЧ резонансы смешанных квантовых состояний орто-пара спин-изомеров H_2O как фактор их конверсии и воздействия на биообъекты ..	77
Информация и правила для авторов	89

Contents

V.N. Blinov, V.L. Golo, Yu.F. Krupyanskiy, A.A. Tuzhilin Coordinate and momentum tunneling in one-dimensional quantum systems with discrete spectrum	5
O. Valba, M. Tamm, S. Nechaev Topological properties of RNA-like molecules with a random primary structure ...	21
V.B. Oshurko On the mechanism of biological action of microwave and terahertz radiation	65
S.M. Pershin RF&MW Resonances of H₂O ortho-para spin-isomers mixing quantum states as a factor of interconversion and impact on bio-objects	77
The information and rules for authors	89

ТОРИЧЕСКИЕ КОНФОРМАЦИИ МОЛЕКУЛ ДНК

В.Н. Блинов¹, В.Л. Голо^{2, 3, 4}, Ю.Ф. Крупянский²

¹ *Научный Центр РАН в Черноголовке*

² *Институт химической физики РАН им. Н.Н. Семенова*

³ *Механико-математический факультет Московского Государственного
Университета им. М.В. Ломоносова*

⁴ *National Research University, Higher School of Economics*

voislav.golo@gmail.com

Поступила 07.06.2015

Мы рассматриваем тороидальные конфигурации как один из вариантов плотной упаковки ДНК в клетках и вирусах. Экспериментальные исследования плотных упаковок такого типа выявили наличие нетривиальной внутренней структуры. В реальной клетке, однако, такие образования должны сочетать компактность с возможностью осуществления биофизических процессов (репликация и т.п.). Возникает вопрос, как изменяются такие плотные образования во время тех или иных процессов жизнедеятельности клеток. Современная экспериментальная техника не может дать ответ на этот вопрос, поэтому возникает необходимость применения новых инструментов, таких как XFEL.

УДК 517.15

1 Образование компактных, глобулярных форм ДНК

Образование глобул является характерным признаком длинных мягких молекул полимерного типа, что может быть объяснено энтропийными соображениями, [1]. Казалось бы, соответствующая теория может быть применена и к молекуле ДНК, рассматриваемой на масштабе много большем, чем её длина персистентности, благодаря чему можно пренебречь упругими свойствами ДНК. Но оказалось, что В-форма ДНК способна обра-

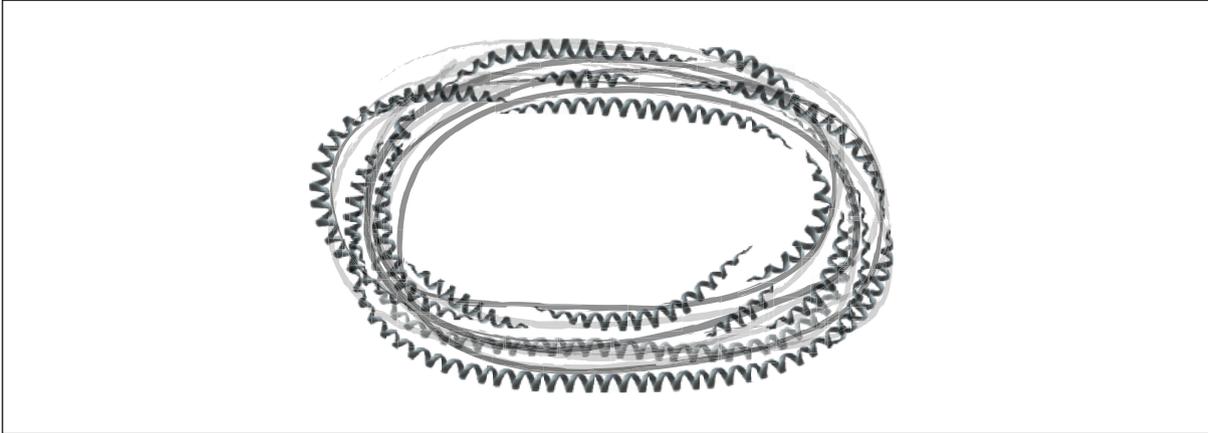


Рис. 1. Тороидная конформация молекулы ДНК

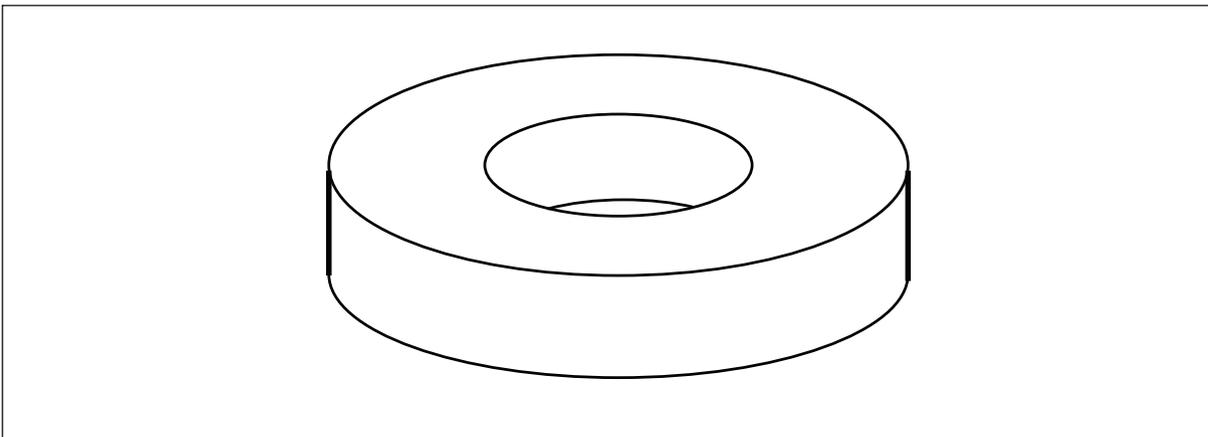


Рис. 2. Та же конформация ДНК что на предыдущем рисунке, 1, но представленная схематически

зовывать более сложные компактные структуры, чем на это способны обычные полимеры, [2], [3].

При соответствующих условиях глобулы ДНК могут иметь тороидальную форму, а также сферическую или цилиндрическую.

Исследование ДНК тороидов значительно стимулируется изучением конформации молекулы ДНК в вирусах и клетках. Тороиды ДНК состоят из отдельных петель, [4], [5], и образуются в результате выпадания, *коллапса*, из первоначально однородного раствора. Коллапс может быть вызван добавлением нейтрального полимера, обычно полиэтиленгликоля (PEG), при наличии достаточной концентрации солевого раствора. Это явление получило название ψ -конденсации. Оно является обратимым переходом и может осуществлено в ту или другую сторону, компактная форма ДНК \rightleftharpoons изотропный раствор, посредством изменения концентрации солевого раствора, или PEG. Таким образом, речь идёт об обратимом конформационном переходе. ДНК тороиды часто имеют внешний диаметр порядка 100 *nm* и внутренний диаметр, дырку, порядка 30 *nm*. При теоретическом рассмотрении можно убедиться, что за счёт изменения упругих свойств ДНК или характеристик противоионов раствора, содержащего молекулу, можно получить сферические образования вместо торических. G. S. Manning, [7], предложил трактовку этих явлений, исходя из предположения, что на пространственных масштабах больше длины персистентности молекула ДНК может быть описана в рамках классической механики деформируемых сред, при учёте электростатических явлений, связанных с зарядами фосфатных групп ДНК.

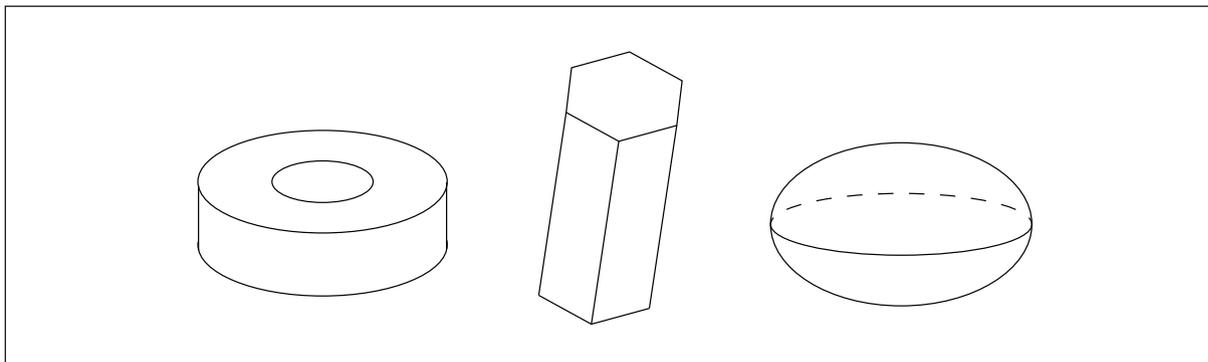


Рис. 3. Схематическое изображение типов компактных образований молекул ДНК

При этом очень существенно, что модули Юнга молекулы и распределение зарядов соответствуют её поперечной анизотропии. Идеи G. S. Manning'a, в той или иной форме, лежат в основе теоретических работ по компактным формам молекулы ДНК. Существенно, что молекула ДНК характеризуется тремя различными пространственными масштабами:

1. микромасштаб порядка 3.4 \AA , соответствующий расстоянию между соседними парами нуклеотидов, образующих двойную спираль диаметра порядка 20 \AA ;
2. мезомасштаб порядка 1000 \AA , нескольких длин персистентности;
3. макромасштаб, порядка длины молекулы ДНК, который может достигать до нескольких мм. Поскольку молекула ДНК на мезо-масштабе нескольких длин персистентности содержит несколько тысяч атомов, представляется, что её можно рассматривать в этой отношении в рамках механики деформируемых сред и применять классическую теорию упругости.

В том, что касается упругих свойств ДНК, на указанном выше мезомасштабе её можно представить себе как классический упругий стержень. При этом возникает проблема расположения молекулы ДНК в ограниченном объёме, например, в клетке, которую приходится решать в рамках теории упругости. Это нетривиальная « контактная » задача теории упругости.

Тороидальные образования молекул ДНК были также обнаружены при условиях, когда цепь молекулы ДНК подвергалась сильному пространственному ограничению, так что система была близка к режиму наиболее плотной упаковки. Таковые были обнаружены в экспериментах с ДНК в наноканалах, жидких каплях, липосомах и вирусных капсидах, включая бактериофаги. Это сравнительно простые системы, если учесть, что образование компактных форм ДНК может иметь место и в живых клетках, в которых исследование этого явления представляет значительные трудности, [8]. Представляет значительный интерес исследование характера расположения молекул ДНК в липосомах, пузырьках-мембранах, самопроизвольно образующихся в смесях фосфолипидов. В частности, это диктуется важными медицинскими применениями.

Следует отметить, что одной из наиболее типичных форм самоорганизации ДНК является её способность образовывать жидкокристаллические (ЖК) холестерические фазы, как *in vivo*, так и *in vitro*. В частности, высказывалось утверждение, что митохондриальная ДНК находится в холестерической ЖК-фазе. Правда, следует иметь ввиду, что утверждение о холестерической ЖК структуре ДНК в хроматине основывается всего лишь на наличии у соответствующих образцов кругового дихроизма, [9].

Возвращаясь к тороидальным формам ДНК заметим, что они, возможно, играют важную роль в динамике жизнедеятельности клетки, [10]. В этом круге вопросов, по-видимому,

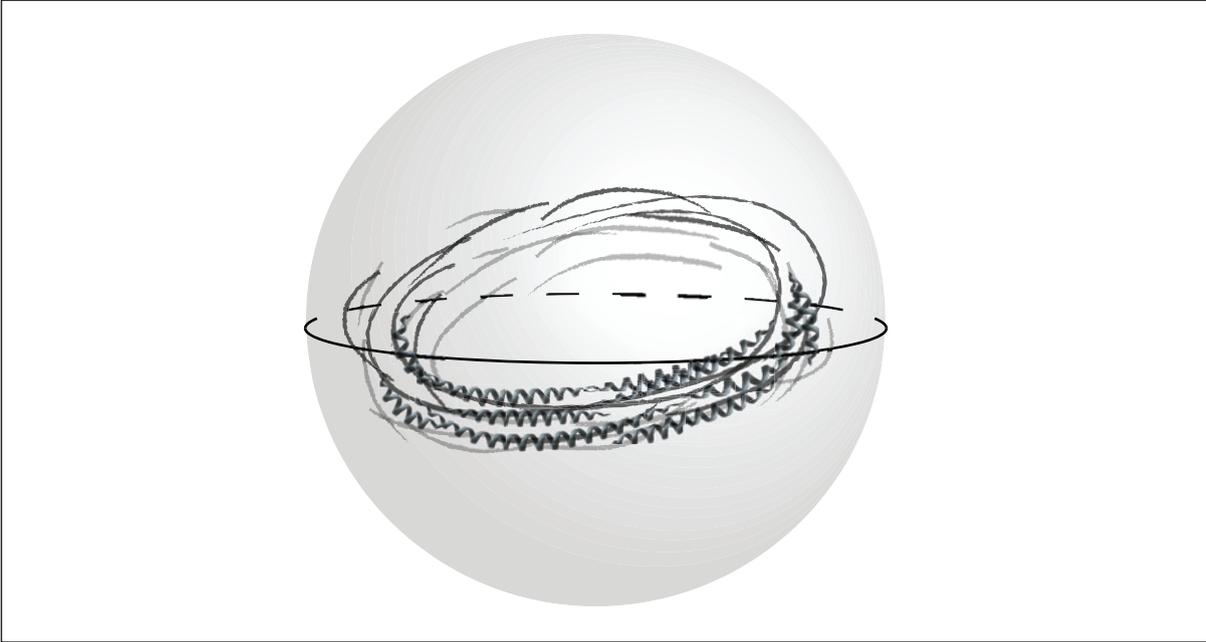


Рис. 4. Схематическое изображение расположения молекулы ДНК в липосоме

также играют существенную роль задачи оптимальной упаковки или укладки молекулы ДНК в клетке, и вытекающие отсюда проблемы нарушения предполагаемой симметрии получающихся структур. По-видимому, роль симметрии в строении тороидальных структур ДНК велика и требует внимательного рассмотрения, что до настоящего момента особо внимательно не исследовалось, хотя уже при изучении ψ -конденсации использовались симметричные соображения с целью упрощения решения теоретических задач, в основном, вычислительного характера.

Тороидальные формы ДНК могут иметь важные медицинские приложения для новых методов генной терапии, с помощью которых лекарственный препарат в виде гена будет доставляться непосредственно в клетку. В этой связи исследование образования тороидальных форм ДНК представляет значительный интерес.

Особую роль играет явление образования торических и нанокристаллических структур при взаимодействии молекул ДНК с белками, что важно для понимания внутриклеточных процессов. В экспериментах *in vitro* было показано, что при взаимодействии стресс-индуцированного белка *DPS* и ДНК образуются стабильные нанокристаллы, в которых ДНК защищена от внешних воздействий. Данные криоэлектронной микроскопии показывают, что параметры решетки комплексов *in vitro* и *in vivo* близки друг к другу. Нанокристаллы в нуклеоиде *E. coli* имеют размер порядка нескольких сотен нанометров, а кристаллы, получаемые *in vitro* – порядка 1 мкм. Данные кристаллы изучались методами криоэлектронной микроскопии и томографии, [10], а также диффузного рентгеновского рассеяния [?]. В результате было охарактеризовано расположение додекамеров *Dps*, однако, точная укладка цепи ДНК в нанокристаллах остается неизвестной.

2 Экспериментальные данные

Известно, [6], что ДНК в живых клетках находится в сильно конденсированном состоянии, и что такого рода конденсация ДНК весьма существенна для жизненного цикла клетки. Можно считать, что тороидальная форма является оптимальной упаковкой молекулы

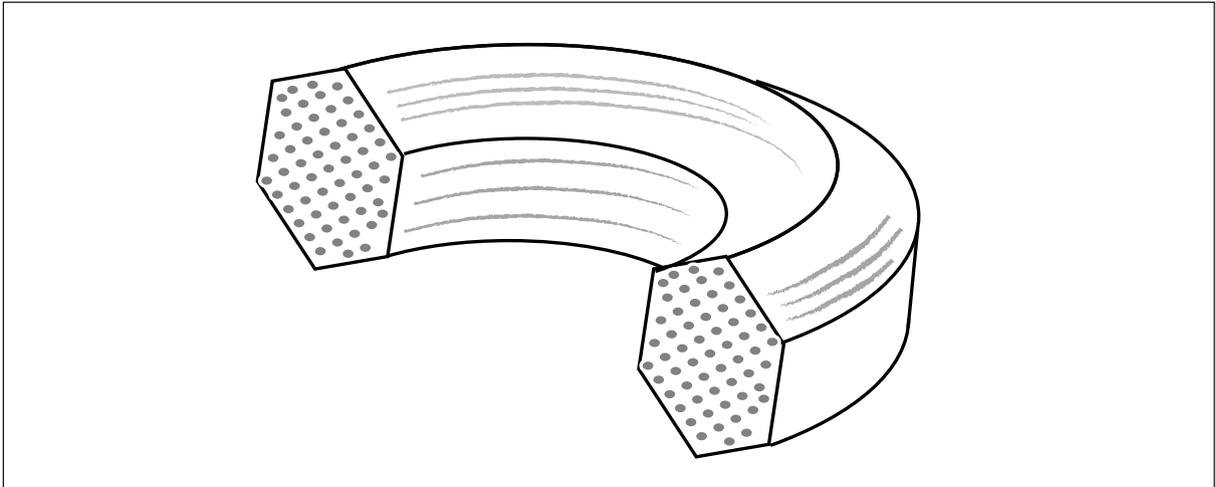


Рис. 5. Поперечное сечение тороида с гексагональной симметрией петель ДНК, выявленное в экспериментах по рентгенографии

ДНК в клетках. Стоит отметить, что явление конденсации ДНК представляет значительный интерес для генной терапии, поскольку может быть использовано для разработки средств доставки искусственных генов.

В работе [6] применялась криоэлектронная микроскопия, что позволило получить разрешение порядка одного нанометра для ДНК тороидов, замороженных в лёд. При этом скорость заморозки такова, что *микрофотографию образца можно рассматривать как моментальный снимок молекулы ДНК в жидкости до заморозки*. Были получены образы, имеющие беспрецедентную точность и четкость изображения, [6].

Укладки ДНК в виде тороидов носят регулярный характер, имеется ряд симметрий и дефектов. Экспериментально установлено, что часть тороидов соответствует регулярной / правильной упаковке молекулы ДНК почти по всей окружности тороида. Однако отмечается, что полная регулярность укладки по всем 360° не имеет места. Часть тороидов обладает упаковкой по гексагональной решетке, но при этом видно, что имеются нарушения правильной укладки. Таким образом, если рассматривать тороид ДНК как своего рода кристаллическую организацию, то она соответствует кристаллу с дефектами. Происхождение и роль этих дефектов не ясны.

В работе [10] было показано, что образование торических форм ДНК играет существенную роль в жизнедеятельности клеток, например, при переходе "изголодавшихся" бактериальных клеток в устойчивое состояние, которое характеризуется образованием пространственно упорядоченных тороидальных структур в хроматине (авторы даже употребляют термин - кристаллических). При этом, по-видимому, существенную роль играет белок *DPS*. Авторы отмечают, что это имеет место только для бурно (экспоненциально) растущих клеток, в то время как споры и бактерии, находящиеся в стационарном состоянии, характеризуются аморфным и однородным состоянием цитоплазмы. Механизм, ответственный Gerald'ом Manning'ом в 60-х годах прошлого века. В рассматриваемом случае он предоставляет возможность качественно описать влияние электростатических сил посредством эффективного механического момента, действующего на молекулу ДНК, рассматриваемую как упругую анизотропную нить. В результате, следуя идеям G. S. Manning'a, конформацию ДНК удастся описать в рамках классической теории Kirchhoff'a одномерной элаستيци, [17].

Будем рассматривать осевую линию молекулы ДНК как упругую нить или стержень. Как было отмечено выше, на промежуточных (мезо)масштабах конформация этой нити опре-

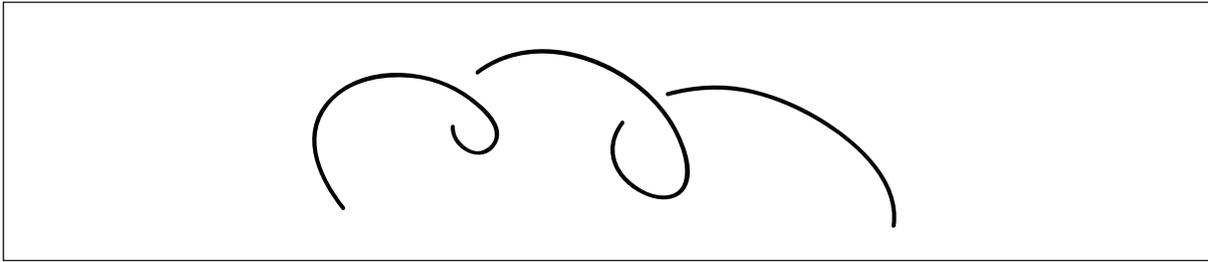


Рис. 6. Кривая, иллюстрирующая решение уравнения Кирхгофа

деляется, в основном, ее упругой энергией. Последняя обычным образом может быть разложена по тензору деформации. Для интересующих нас вопросов существенно наличие внутренне-анизотропной упругости молекулы, т. е. тот факт, что упругие модули не изотропны. Эта анизотропия упругих модулей появляется естественным образом, так как, заменяя двойную спираль одной упругой линией (осью спирали), мы обязательно должны учесть анизотропную структуру ДНК. Во многих работах указанную анизотропию не учитывают, что приводит к существенному загромулению, а то и вовсе потере результатов. На механическом языке полинуклеотидные нити, образующие двойную спираль, могут быть представлены как винтовые бороздки на рассмотренном выше упругом стержне, причем одна из этих бороздок узкая, а другая широкая. Это же обстоятельство приводит к очень важному для ψ -конденсации следствию. Напомним, что молекула ДНК является заряженной и несет элементарный заряд $-2e$ на пару фосфатных остатков. Далее, за счет «конденсации» противоионов из раствора происходит частичная компенсация отрицательного заряда ДНК. Такая задача о заряженной нити (полиионе) в растворе, содержащем противоионы, рассматривалась многими авторами. Было показано, что при заряде полииона, превышающем некоторое критическое значение, противоионы «конденсируются» на полиион, снижая плотность заряда до критического значения. Отличительная черта такой модели, как уже отмечалось выше, состоит в представлении о внутренней анизотропии макромолекулы, в том числе и по распределению плотности заряда в узкой и широкой бороздках полииона. Это означает, в частности, что компенсация заряда за счет противоионов происходит неоднородно и анизотропно по длине молекулы, например, в силу неоднородности поля, создаваемого отрицательными зарядами фосфатных групп в узкой или в широкой бороздке. В результате удастся получить качественно верное описание тороидальных структур, [18]. А именно, полученные таким способом результаты дают описание тороидальных структур и находятся в качественном согласии с экспериментом, в том, что касается численных данных. Во всяком случае, размеры тороидов получаются вполне разумные. Недостатком данного подхода является то, что он не учитывает стерическое отталкивание отдельных сегментов молекулы и оставляет в стороне вопрос о гексагональной симметрии. Во всяком случае здесь необходимо дополнительное исследование. Но если учесть, что гексагональная симметрия обычно вводилась «руками», то вполне возможно, что этот недостаток является в действительности достоинством. Во всяком случае, можно ожидать, что уравнения Kirhhoff'a дадут решения, удовлетворяющие некоторой симметрии. Какой именно и как она будет согласовываться с экспериментом, «remains to be seen».

Здесь следует отметить, что тороидальные формы ДНК могут образовываться при условиях, когда электростатическим взаимодействием можно пренебречь. По-видимому, не существует единого механизма возникновения тороидов, и рассмотренный выше электростатический подход является только одним из возможных.

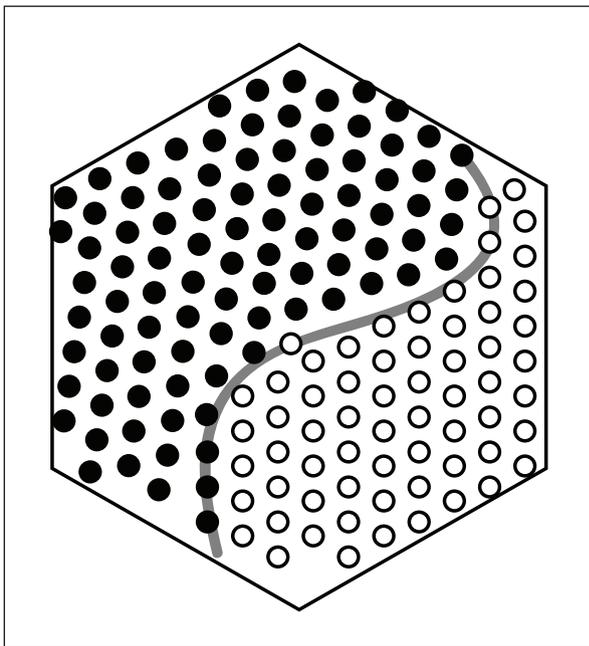


Рис. 7. Нарушение гексагональной симметрии в тороиде

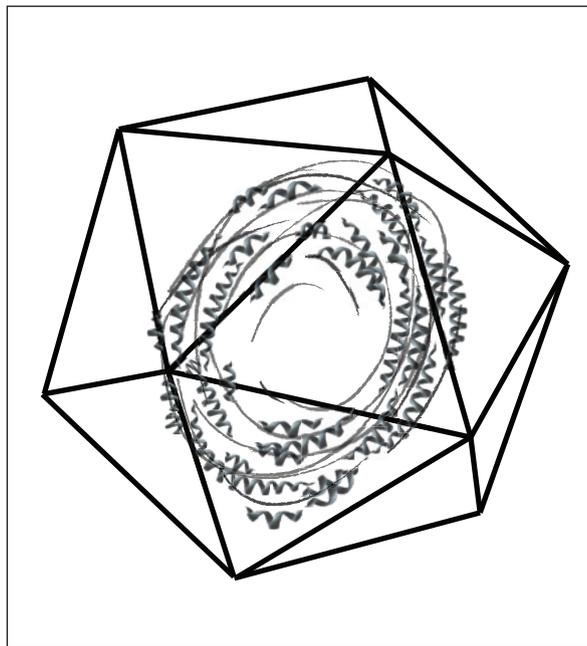


Рис. 8. Компактное образование из молекулы ДНК внутри икосаэдрической капсулы

4 Проблема особенностей укладок и намоток.

Соображения о гексагональной симметрии торических укладок ДНК уже давно используются в теоретических работах, [19]. В основе лежат соображения математической целесообразности, поскольку для того чтобы хоть что-то просчитать авторам нужны упрощающие соображения. Экспериментаторы, в свою очередь, охотно видят то, что им советуют увидеть теоретики (например, работа Hud'a & Downing'a, [6], которая, надо отдать им должное выглядит как вполне честная работа). По-видимому, эти симметричные соображения вполне адекватны для тороидов, образовавшихся в свободном растворе, т. е. когда нет влияния стенок (клеточных мембран, капсидов вирусов и т. п.), но в случае, когда такие ограничения имеются, нужно решать соответствующую задачу теории электро-упругости, и далеко не ясно, сохраняться ли «общепринятые» условия симметрии.

4.1 *Analysis situs* укладок ДНК.

Соотношения гексагональной симметрии, по-видимому, действительно имеют место, но нужно учитывать, что они могут быть приближенного характера. В сущности, проблема похожа на задачу «смотреть садовый шланг таким образом, чтобы его можно было запихнуть в небольшой ящик». Можно ожидать, что тут должны проявиться явления, основанные на требованиях согласования метрических и симметричных свойств системы – мотка ДНК. Задача согласования метрики и симметрии для намоток не изучалась. Она актуальна, и её следует исследовать. Она может быть полезна при анализе экспериментальных данных, которая как показывает работа Hud'a & Downing'a, не тривиальна. Среди прочего, она представляет значительный интерес для исследования икосаэдрических ДНК-вирусов, расположение молекулы ДНК в которых как-то не особенно обсуждалось в литературе (наука безмолвствует, что можно рассматривать как биомолекулярный вариант финальной сцены «Бориса Годунова»).

4.2 Особенности и преимущества вирусной проблемы.

Исследование укладок ДНК в вирусах интересно с одной стороны, потому что это актуальная биологическая задача, результаты исследования которой могут иметь важные практические последствия, значимость которых может превзойти ожидания, и, с другой стороны, она выглядит как более трактательная нежели аналогичная задача расположения тороида ДНК в клетке общего типа. Во всяком случае, её понятно как ставить в рамках теории электро-упругости, а значит что-то можно и разрешить. Постановка задачи конформации тороида ДНК в клетке общего вида - *не ясна*.

4.3 Укладки ДНК в хроматине.

ДНК в клетке хранится в специально упакованном виде. Такая упаковка – *хроматин* – имеет одинаковую структуру в клетках различных организмов и определяется несколькими уровнями. Так, на микроскопическом уровне он представляет собой совокупность *нуклеосом* [20]. Нуклеосома состоит из ядра – белковой глобулы из восьми субъединиц, *гистонов*, – обёрнутого в молекулу ДНК, обычно содержащую 147 пар оснований. В нуклеосомное ядро входит по две субъединицы каждого из четырёх различных гистонов (Н2А, Н2В, Н3 и Н4). Нуклеосомы соединяются между собой посредством гистона Н1, образуя следующий уровень упаковки в структуры больших масштабов, в конечном счёте образуя хроматин. Данная упаковка схематично изображена на рисунке 9; хромосомы представляют собой плотные клубки хроматиновых нитей, каждая из которых есть, в свою очередь комплекс нуклеосом.

Такая упаковка позволяет значительным образом сократить объём, занимаемый молекулой ДНК (в 30-40 раз). Столь сильное сжатие обуславливает значительную деформацию молекулы ДНК: длина персистентности молекулы составляет порядка 500 Å, как раз эту величину имеет петля из молекулы нуклеиновой кислоты в нуклеосоме. При моделировании молекулы ДНК как однородного стержня сжатие было бы равномерным, в результате чего образовывалась бы петля. Реальная ДНК имеет более сложную структуру: борозд-

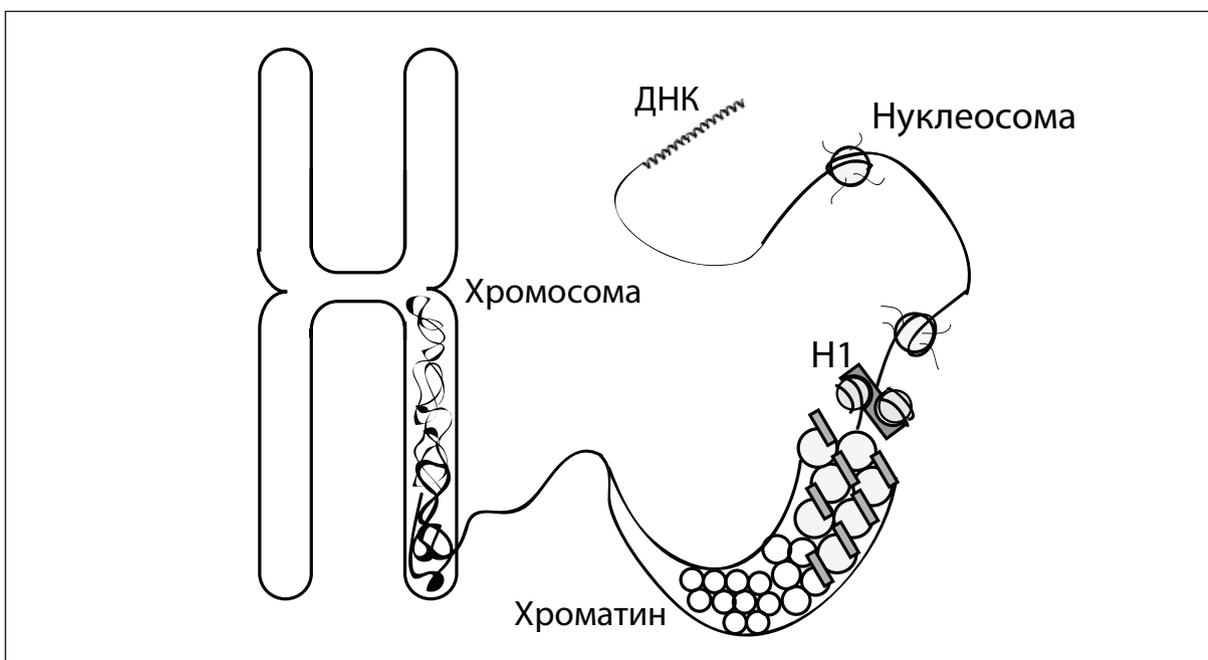


Рис. 9. Иллюстрация структуры хроматина в клетке

ки и наличие двух нитей, что существенным образом влияет на её упругие свойства при намотке на гистоновое ядро. Во-первых, ДНК начинает сильнее изгибаться по бороздкам, то есть, в тех местах, в которых она прилегает одной из бороздок к октамеру. С другой стороны, поскольку отрицательный заряд несёт сахаро-фосфатный остов, в тех областях поверхности октамера, где имеется положительный заряд, энергетически выгоднее оказывается конформация, когда ДНК прилегает к гистонам своим остовом. Таким образом, на молекуле ДНК появляются два типа характерных областей, определяющих её деформацию: это точки электростатического прикрепления ДНК и ядра, а также точки изгиба молекулы по бороздкам. Эти особенности видны как экспериментально (на микроскопической структуре, полученной в рентгеноструктурном анализе), так и в полноатомных моделях в молекулярно-динамических вычислениях. В частности, было показано, что ДНК находится в сверхспирализованном по отношению к обычной В-форме состоянии: расположение двух положительно заряженных областей на гистоновом ядре оказывается ближе, чем шаг спирали, в результате чего ДНК вынуждена "докручиваться", чтобы обе эти области соприкасались с остовом молекулы нуклеиновой кислоты. Интересно, что сам факт наличия изломов (*kinks*) известен достаточно давно, [24].

Важным обстоятельством здесь является то, что роль подобных намоток с биофизической точки зрения исследована плохо. Имеется множество гипотез о том, как структура хроматина связана с действием различных белков, влияющих на регуляцию связанных с ДНК процессов транскрипции и репликации, [27]. С этой точки зрения интересны исследования структуры хроматина и белков, участвующих в его образовании, [26].

Хроматин является крайне сложной структурой, на которую могут значительным образом влиять параметры среды. По этой причине значительный интерес представляют типы деформации двойной спирали и условия, при которых они могут возникать. Так, например, способ намотки молекулы ДНК зависит от нуклеотидной последовательности и виток может содержать 146 пар оснований [25]. Структура нуклеосомы значительным образом зависит от условий, в которых находится комплекс (ионный состав, ионная сила раствора и др., [21]). Это, в частности, обусловлено и свойствами молекулы ДНК как полииона. Можно предположить, что увеличение ионной силы раствора приведёт к усилению экранировки зарядов фосфатного остова молекулы, что, в свою очередь, уменьшит длину персистентности ДНК. Это будет означать, что молекула будет деформироваться значительно легче.

С точки зрения процессов в клетке значительный интерес представляет также механизм прикрепления молекулы ДНК к гистоновому октамеру; нуклеотидная последовательность ДНК влияет на её механические свойства, а потому энергия связи нуклеиновой кислоты с ядром оказывается больше или меньше для различных последовательностей. Известна т.н. 601-я последовательность пар оснований, которая очень хорошо связывается с гистоновым октамером, и часто используется в экспериментальных исследованиях. Можно предположить, что свойства таких важнейших процессов как репликация или транскрипция будут зависеть от нуклеотидной последовательности ДНК. Кроме того, подобные исследования могут выявить корреляцию между последовательностью пар оснований и регуляцией генов.

Наглядно структуру нуклеосомы можно сравнить с толстой нитью, виток которой намотан на боковую поверхность цилиндра. В роли цилиндра при этом выступает гистоновый октамер, имеющий нетривиальную структуру. Как было отмечено ранее, гистоновый октамер состоит из четырёх пар одинаковых белков: H2A, H2B, H3 и H4, которые, объединяясь, образуют нуклеосомное ядро, рис. 12. Внешняя поверхность такого ядра имеет множество полярных аминокислот, обуславливающих значительный заряд на поверхности. Боковая поверхность, при этом, оказывается заряженной положительно, поэтому отрицательно заряженная поверхность молекулы ДНК прилипает к ней, образуя устойчивый нуклеосом-

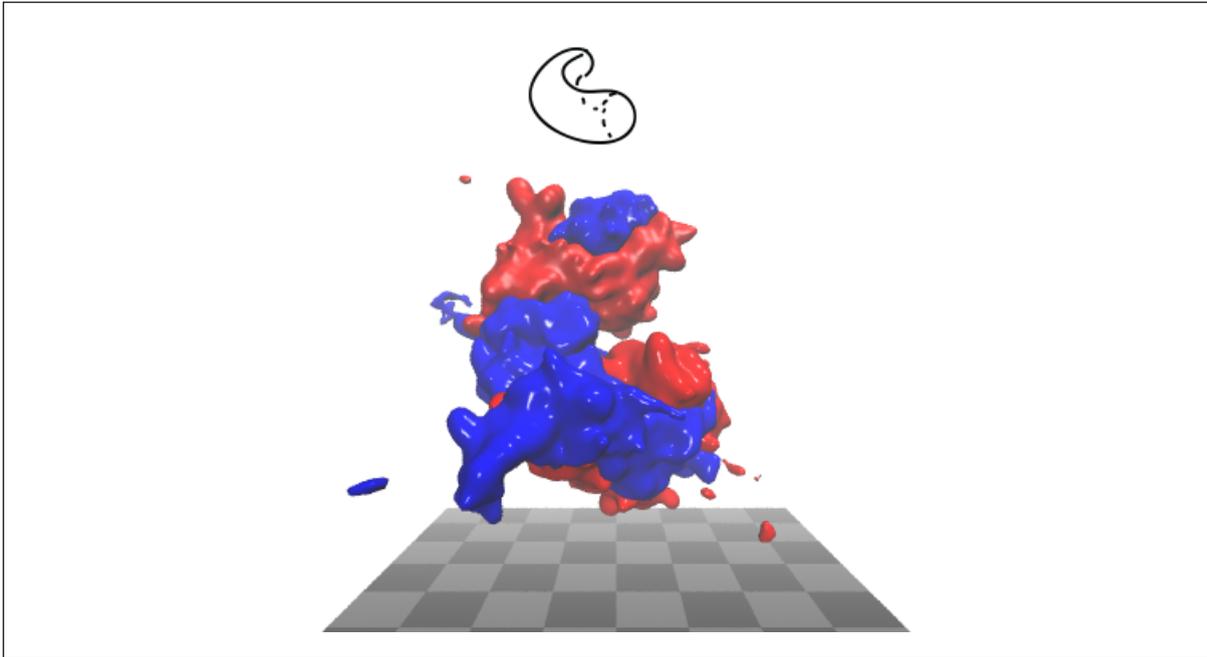


Рис. 10. Четыре субъединицы гистонового октамера: два H2A и два H2B образуют фигуру в форме подковы в горизонтальной плоскости.

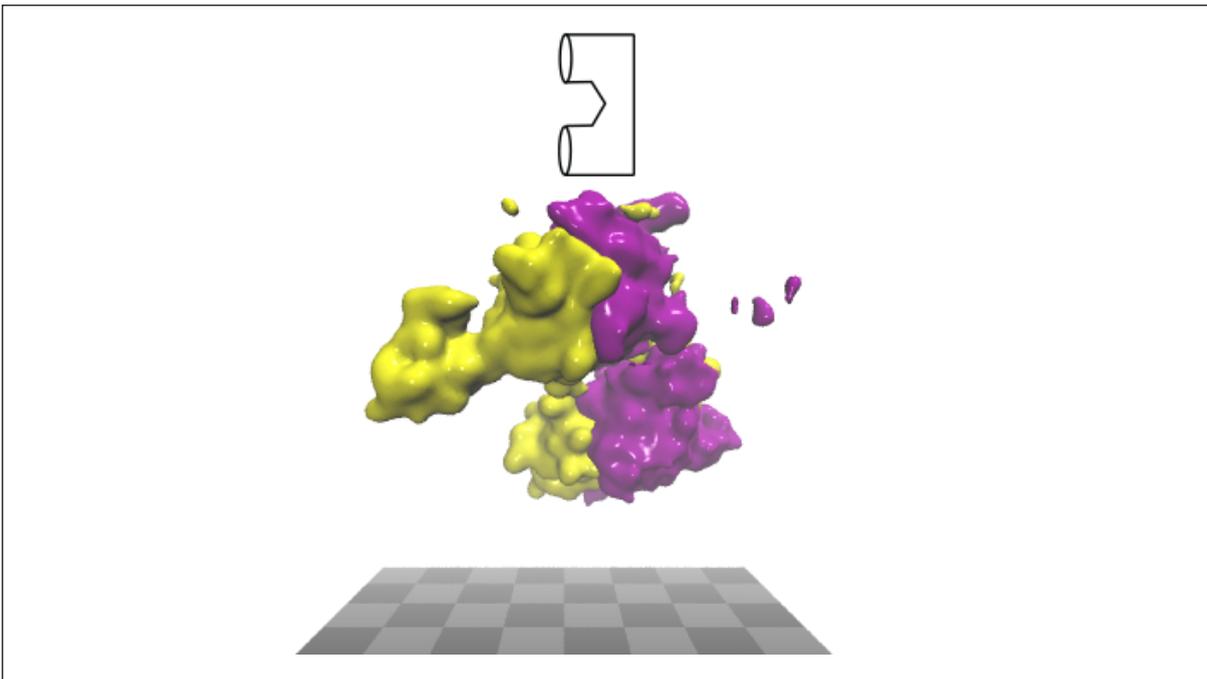


Рис. 11. Четыре субъединицы гистонового октамера: два H3 и два H4 образуют фигуру в форме подковы в вертикальной плоскости

ный комплекс. Схема сборки нуклеосомы из составных частей показана на рисунках 10 - 13.

Кроме положительно заряженной боковой поверхности, внешняя часть октамера содержит фрагменты с отрицательным зарядом, у которых также есть некоторый смысл. Помимо указанных типов гистонов имеются также и другие, которые имеют собственное предназначение – так, например, гистоны H1 служат для скрепления нуклеосом между собой. Намотка молекулы ДНК на гистоновый октамер представляет интерес в связи с двумя

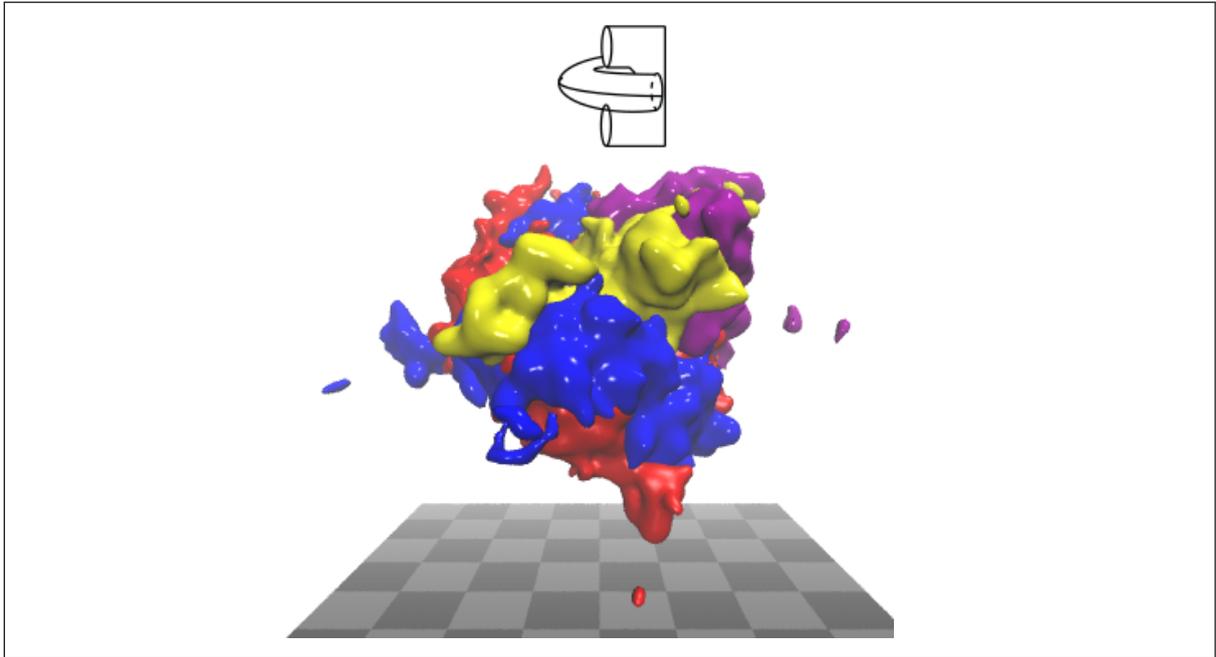


Рис. 12. Соединение двух подков с образованием гистонового октамера катушки для витка ДНК

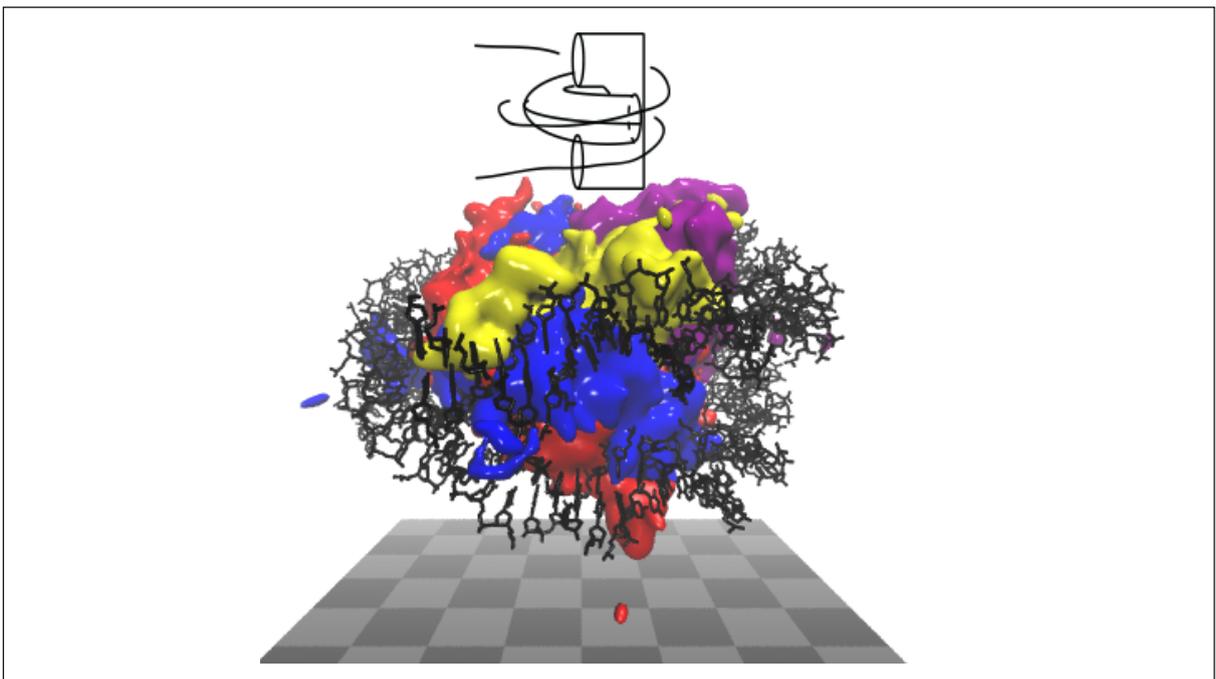


Рис. 13. Нуклеосома: гистоновый октамер с намотанным витком ДНК.

обстоятельствами. Во-первых, последовательность нуклеотидов фрагмента ДНК влияет на то, насколько хорошо связывается ДНК с октамером. Как было отмечено выше, нуклеотидный состав может даже изменить количество пар оснований, связанных с белковым ядром. Во-вторых, молекула ДНК имеет бороздки, расположенные несимметрично, ввиду чего намотка оказывается не равномерной, а с изломами. Таким образом, есть почти прямолинейные отрезки молекулы, а есть с изломами. Наличие изломов достаточно интересно и исследование их положения на молекуле может влиять на энергию деформации молекулы и полную энергию системы.

Интересен также и тот факт, что нуклеосома стабильна (удерживает молекулу ДНК) даже в том случае, когда гистонов в ядре не восемь, а шесть или даже четыре. Так, с боковой поверхности нуклеосомы (со стороны обоих оснований цилиндра) может отсоединяться димер Н3-Н4. В результате, может оказаться, что виток ДНК удерживается двумя димерами Н2А-Н2В, что подтверждается проведённым нами компьютерным моделированием.

Считается, что такой именно тетрасома является промежуточным комплексом при сборке и распаде нуклеосом, [28]. При этом сборка и разборка *in vivo* осуществляется, по-видимому, при помощи вспомогательных белков, схема этого процесса изображена на рисунке 14.

Структуры нуклеосом представляют значительный экспериментальный интерес и активно изучаются в настоящее время (микроскопия и т.п.). В частности, интересен вопрос о том, как ДНК разматывается с гистонного ядра, например, при прохождении через нуклеосому ДНК-полимеразы. Существуют различные модели этого процесса, [29], [30], которые пока опираются на косвенные данные. Следует отметить, что зернистая структура хроматина может объяснить действия различных белков и молекул на характерных расстояниях, кратных длине витка. Подобные гипотезы высказываются и имеют право на существование, однако экспериментально проверить подобные предположения крайне затруднительно. Единственный метод, который может дать некоторую информацию о подобного рода процессах, заключается в компьютерном моделировании таких систем. Молекулярно-динамические вычисления, однако, зачастую носят характер построения одной или нескольких траекторий, особенно в тех случаях, когда речь идёт о больших структурах и комплексах (в случае полноатомной модели к таковым можно отнести одну нуклеосому, а рассмотреть на современных компьютерах хроматин практически невозможно). Такой подход совершенно не учитывает статистические поправки, которые имеют место в реальной системе. Кроме того, имеется множество других проблем, связанных с моделированием молекулярных систем, которые сводятся к тому, что следует относиться к результатам моделирования с большой осторожностью, и проводить численное моделирование одних и тех же систем несколько раз различными методами.

Таким образом, молекулярное моделирование - крайне мощный инструмент для изучения макромолекулярных систем, которым не следует пренебрегать, однако использовать его с большой осторожностью.

Экспериментально основная информация о структуре молекулы получена из рентгеноструктурного анализа (pdb 1kx5 и 3lz0). В этом случае удавалось получить кристалл из нуклеосомных комплексов, после чего извлекать из кристалла структуру ячейки классическими методами. При получении структур таким образом следует помнить о двух моментах. Во-первых, структура нуклеосомы чувствительна ко внешней среде, поэтому полученная таким образом структура, вообще говоря, может отличаться от структуры *in vivo* (как, впрочем, и *in vitro*). С этой точки зрения информация о структуре, интересующая исследователей, не исчерпывается полученной таким образом. Вторым моментом является то, что в таком эксперименте координаты подвижных частей кристаллизуемого комплекса будут найдены неточно. Это, впрочем, часто оказывается не важным, поскольку про подвижность таких фрагментов часто известно *a priori*.

Ответ на вопрос, как выглядит структура нуклеосомы в условиях живой клетки обычно решается путём компьютерного моделирования, поскольку других инструментов на данный момент нет. Разработка методики, которая позволила бы исследовать нуклеосомы на XFEL представляет несомненный интерес, и одновременно является крайне сложной задачей. Моделирование динамики белковых структур и их комплексов проводится преимущественно методом молекулярной динамики, [22], [23]. При этом для анализа получаемых структур могут применяться и другие средства, как графические (например, VMD), так и аналитические (например, X3DNA для анализа деформаций молекулы ДНК). В

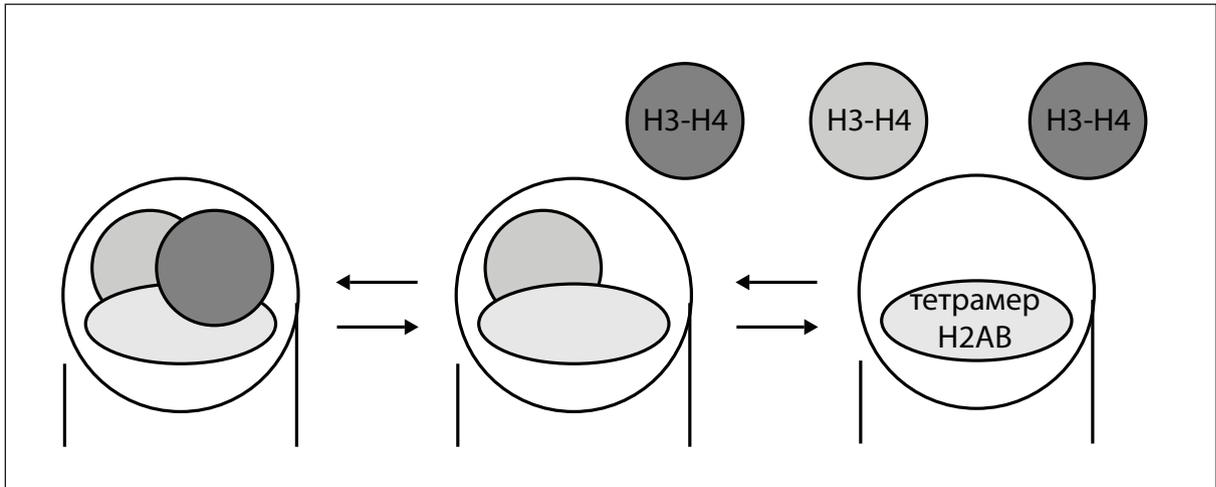


Рис. 14. Подвижность димеров H3-H4 с образованием тетрасомы

молекулярно-динамическом приближении каждый атом (или несколько атомов в случае крупнозернистых моделей) представляется шаром с массой и зарядом, а ковалентные связи – голономными или гармоническими связями конструируемой механической системы. При этом, каждому шару приписываются некоторый набор силовых констант, определяющихся используемым силовым полем и типом обозначаемого им объекта. Моделируемая система собирается из таких элементарных шариков и пружинок, после чего производится численное интегрирование уравнений Ньютона для построенной механической системы. Выше уже отмечалось, что теоретический анализ задач компактной конформации молекул ДНК представляет очень сложную задачу и требует использования различных аппроксимаций, т.е. модельных представлений. При этом необходимо учитывать, что основными силами, определяющими характер этих задач являются упругие и электростатические силы. Таким образом, необходимо последовательно работать в рамках теории электро-упругости.

5 Проблема особенностей укладок и намоток.

К настоящему времени достигнуты значительные успехи в применении электронной микроскопии и рентгеноструктурного анализа к изучению компактных форм ДНК (особо отметим работу Hud'a & Downing'a). Вместе с тем нужно использовать новые инструменты (XFEL), которые позволят проверить старые результаты и получить новые, недоступные для старых методик. Во всяком случае, обещания, которые дают строители XFEL'ей, позволяют надеяться увидеть что-то новое. В этом отношении первостепенной является задача исследования симметричных дефектов тороидальных укладок ДНК, выяснения их структурных особенностей и роли в биологических процессах. Возможно, удастся посмотреть динамику образования укладок в реальном времени. Как утверждают создатели XFEL-аппаратуры, технические возможности для этого будут. Сюда же относится задача исследования расположения тороидов в клетке, по-видимому, в настоящий момент не поддающаяся теоретическому анализу, и исследование конформации ДНК в нуклеосоме и её биологическая значимость (имеющиеся данные о кристаллической структуре получены для нуклеосом в специальных условиях, при которой возможна кристаллизация; при этом, вообще говоря, неясно, так ли будет организована нуклеосома в клетке). Исследование конформации ДНК в вирусах – трудная теоретическая задача. Она представляет значительный интерес для понимания устойчивости вирусной частицы по отношению к внешним воздействиям, например, микроволновому излучению. Здесь необходимо применение мощных новых приборных средств. По-видимому, XFEL тут будет очень кстати.

Опять же, если дать волю воображению, было бы интересно посмотреть в реальном времени («кино») разрушение капсида вирусной частицы, или выбрасывание бактериофагом нуклеиново-кислотной начинки бактериофага. Никакой мало-мальски разумной теории таких процессов в настоящее время нет. Для создания таковой требуется основательная подсказка со стороны экспериментаторов. Значимость этой проблемы не требует пояснений.

В.Л.Г. благодарит Программу Прогресса Фундаментальных Исследований Высшей Школы Экономики за финансовую поддержку.

Список литературы

- [1] *Lifshitz I.M.* Some Problems of the Statistical Theory of Biopolymers // ZhETF, 1969, **28**, p.1280.
Френкель Я.И. Кинетическая теория жидкости // Л.: Наука, 1975.
- [2] *Yevdokimov Yu.M. et al.* A compact form of double-stranded DNA in solution // FEBS Lett. 1972, **23**, p.180.
- [3] *Lerman L.S.* Physico-Chemical Properties of the Nucleic Acids // J. Duchesne, ed., Academic Press, London, pp. 59-76, 1973.
- [4] *Petrov A. S., Harvey S. C.* Structural and thermodynamic principles of viral packaging // Structure, 2007, **15**(1), pp.21–27.
- [5] *Marenduzzo D. et al.* DNA–DNA interactions in bacteriophage capsids are responsible for the observed DNA knotting // Proc. Natl. Acad. Sci. U.S.A., 2009, **106**, pp.22269–74.
- [6] *Hud N. V., Downing K.H.* Cryoelectron microscopy of λ -phage DNA condensates in vitreous ice: The fine structure of DNA toroids // PNAS, 2001, **98**, pp.14925–14930.
- [7] *Manning G. S.* A procedure for extracting persistence lengths from light-scattering data on intermediate molecular weight DNA // Biopolymers, 1981, **20**, pp. 1751–1755.
- [8] *Leforestier A. et al.* Protein-DNA Interactions Determine the Shapes of DNA Toroids Condensed in Virus Capsids // Biophysical Journal, 2011, **100**, pp.2209–2216.
- [9] *Livolant F.* Ordered phases of DNA in vivo and in vitro // Physica A, 1991, **176**, pp. 117-137.
- [10] *Frenkiel-Krispin D. et al.* Nucleoid restructuring in stationary-state bacteria // Molecular Microbiology, 2004, **51**(2), pp.395–405.
- [11] *Vasilevskaya V. V. et. al.* Structure of collapsed persistent macromolecule: Toroid vs. spherical globule // Biopolymers, 1997, **41**(1), p. 51.
- [12] *Fenley M. O., Manning G. S. and Olson W. K.* Approach to the limit of counterion condensation // Biopolymers, 1990, **30** (13-14), pp.1191-1203.
- [13] *Manning G. S.* // Comments Mol. Cell. Biophys. 1982, **5**, 311 (1982).
- [14] *Marky N.L. and Manning G.S.* The elastic resilience of DNA can induce all-or-none structural transitions in the nucleosome core particle // Biopolymers, 1991, **31**(13), pp.1543-57.
- [15] *Manning G. S.* Packaged DNA. An Elastic Model // Cell Biophys. 1985, **7**(1), pp. 57-58.

- [16] *Manning G. S.* Electric and Elastic Instabilities of DNA //in Theoretical Biochemistry and Molecular Biophysics, ed. by O. L. Beveridge and R. Lavery. Adenine Press, New York, p. 191, 1990.
- [17] *Kirchhoff G.* Mechanik // Teubnir, Berlin, 1897.
- [18] *Golo V.L., Kats E. I., Yevdokimov Yu. M.* Toroidal Structures Due to Anisotropy of DNA-like Molecules // J. Biomolecular Structure & Dynamics 1998, **15**, pp.757-764.
- [19] *Ubbink J., Odijk T.* Polymer-and salt-induced toroids of hexagonal DNA // Biophys. J. 1995, **68**, p.54.
- [20] *Kornberg R.D.* Structure of chromatin // Annu. Rev. Biochem. 1977, **46**, pp.931–954.
- [21] *van Holde R.E.* Chromatin // Springer, New York, 1988.
- [22] *Roccatano D., Barthel A., Zacharias M.* Structural Flexibility of the Nucleosome Core Particle at Atomic Resolution Studied by Molecular Dynamics Simulation // Biopolymers, 2007,**85**(5-6), pp. 407-421.
- [23] *Materese C., Savelyev A., Papoian G.* Counterion Atmosphere and Hydration Patterns near a Nucleosome Core Particle // JACS 2009, **131**, pp. 15005-15013.
- [24] *Hogan M.E., Rooney T.F., Austin R.H.* Evidence for kinks in DNA folding in the nucleosome // Nature, 1987, **328**, pp. 554-557.
- [25] *Richmond T.J., Davey C.A.* The structure of DNA in the nucleosome core // Nature, 2003, **423**, pp. 145-150.
- [26] *Rea S. et al.* Regulation of chromatin structure by site-specific histone H3 methyltransferases // Nature, 2000, **406**, pp. 593-599.
- [27] *Narlikar G.J., Fan H.-Y., Kingston R.E.* Cooperation between Complexes that Regulate Chromatin Structure and Transcription // Cell, 2002, **108**(4), pp. 475-487.
- [28] *Donham D.C., Scorgie J.K., Churchill M.E.A.* The activity of the histone chaperone yeast Asf1 in the assembly and disassembly of histone H3/H4–DNA complexes // Nucleic Acids Research, 2011, **39**(13), pp.5449-5458.
- [29] *Luse D.S., Studitsky V.M.* The mechanism of nucleosome traversal by RNA polymerase II. Roles for template uncoiling and transcript elongation factors // RNA Biology, 2011, **8**(4), pp.581-585.
- [30] *Chang H.-W. et al.* Analysis of the mechanism of nucleosome survival during transcription // Nucleic Acids Research, 2013, **42**(3), pp.1619-1627.

TOROIDAL CONFORMATIONS OF DNA MOLECULES

V.N. Blinov¹, V.L. Golo^{2, 3, 4}, Yu.F. Krupyanskiy², A.A. Tuzhilin³

¹*Science Center in Chernogolovka RAS,*

²*Semenov Institute of Chemical Physics RAS*

³*Dept. of Mechanics and Mathematics Lomonosov Moscow State University*

⁴*National Research University, Higher School of Economics*

voislav.golo@gmail.com

Received 07.06.2015

We consider dense toroidal structures of DNA molecules in cells and viruses. The problem needs special studying. For one thing the space scale of the system being small one has to admit that such structures must be compact enough to accommodate a molecule of the DNA, for another the molecule should be easily accessible for various biochemical processes, for example replication. Theoretical treatment of the problem runs across substantial difficulties, and therefore there is a need for the experimental investigation of the DNA conformation in cells and viruses. The XFEL could be instrumental to that end.

ТОПОЛОГИЧЕСКИЕ СВОЙСТВА РНК-ПОДОБНЫХ МОЛЕКУЛ СО СЛУЧАЙНОЙ ПЕРВИЧНОЙ СТРУКТУРОЙ

О. Вальба^{1,2}, М. Тамм^{1,3}, С. Нечаев^{4,5}

¹*Национальный исследовательский университет Высшая школа экономики, Москва*

²*Институт химической физики им. Н.Н. Семенова РАН, Москва*

³*Московский государственный университет им. М.В. Ломоносова, Москва*

⁴*Universitè Paris-Sud/CNRS, Orsay, France*

⁵*Физический институт им. П.Н. Лебедева РАН, Москва*

ovalba@hse.ru

Поступила 20.04.2015

Работа посвящена применению методов статистической физики и теории случайных процессов для исследования топологических свойств РНК-подобных гетерополимеров со случайной первичной структурой. В частности, описывается критическое изменение топологии РНК-подобных структур в зависимости от алфавита, используемого в случайной последовательности, приводится аналитическая оценка точки перехода в рамках комбинаторного и матричного описания.

УДК 538.9

I. ВВЕДЕНИЕ

Структура важнейших биологических макромолекул, таких как дезоксирибонуклеиновые кислоты (ДНК), рибонуклеиновые кислоты (РНК) и белки, играет ключевую роль в их правильном функционировании в клетке. Различают три уровня структурной

упорядоченности биомакромолекул. Одна из основных их особенностей состоит в гетерополимерности. Последовательность звеньев в ДНК, РНК и белках индивидуального организма, она называется первичной структурой, строго зафиксирована. Далее, биополимерные цепи могут формировать спиралеобразные и складчатые участки небольшого масштаба, как в белках, или комплементарно спаренные и петлевые участки, как в РНК. Такие фрагменты называются элементами вторичной структуры. Различают также третичную и четвертичную пространственные структуры биополимеров.

Данная работа посвящена исследованию топологических свойств вторичной структуры молекул РНК-типа. Известно, что биомакромолекулы являются «слабо отредактированными случайными гетерополимерами» [1, 2]. Более того, для ряда свойств распределение мономерных звеньев в первичной структуре, например, функциональных РНК можно считать случайным [3, 4]. В этом случае, модель случайной первичной структуры является базовой моделью, описывающей основной (нулевой) вклад в наблюдаемые физические явления. Основное внимание при этом сфокусировано на нетривиальной вторичной структуре РНК-подобных полимеров, для описания которой привлекаются разнообразные техники, в том числе, техники квантовой теории поля и моделей Изинга [5].

Структура работа такова. В разделе II приводятся алгоритмы описания РНК-подобной структуры и вычисления свободной энергии ее основного состояния. Формулируется вспомогательная статистическая модель, описывающая взаимодействия мономерных звеньев в РНК-подобной структуре с петлевыми участками. Далее, в предложенной модели учитывается вклад внутриветлевого взаимодействия мономеров и приводится соответствующий алгоритм динамического программирования для вычисления энергии такой иерархической структуры.

Раздел III посвящен определению свойств распределения свободной энергии ансамбля случайных последовательностей РНК. Обсуждаются такие характеристики, как среднее значение свободной энергии в ансамбле, флуктуация средней энергии, распределение по длинам петель в пространственных структурах.

Анализ топологических свойств в модели случайной первичной структуры РНК-подобной молекулы в зависимости от используемой в последовательности алфавита описан в отдельный раздел IV. Показывается, что в зависимости от алфавита РНК-подобная структура характеризуется либо максимально связанной вторичной структурой без пропусков (неспаренных мономеров), либо структурой с конечной долей несвязанных мономеров. Для определения точки такого топологического перехода фор-

мулируется модель Бернулли. В рамках предложенной модели приводятся численные и аналитические оценки критической точки перехода.

II. АЛГОРИТМЫ ВЫЧИСЛЕНИЯ СВОБОДНОЙ ЭНЕРГИИ РНК-ПОДОБНЫХ СТРУКТУР

А. Выравнивание последовательностей

Задача о выравнивании двух последовательностей – это задача нахождения эффективного алгоритма поиска наибольшей общей подпоследовательности (НОП) двух произвольных линейных последовательностей. Данная проблема является одной из ключевых задач вычислительной эволюционной биологии. В частности, она позволяет судить о том насколько далеко (в эволюционном смысле) разошлись друг от друга два рассматриваемых гена и какие гены могут являться их общими предками. Задача об НОП широко исследовалась в биологии, компьютерных науках, теории вероятности и позже в статистической физике.

Задача о поиске НОП двух последовательностей формулируется следующим образом. Рассмотрим две произвольные последовательности (в качестве примера рассматриваются последовательности РНК, составленные из 4-х буквенного алфавита A, C, G, U): $S_1 = \{A, C, G, C, U, A, C\}$ длины $m = 7$ и $S_2 = \{C, U, G, A, C\}$ длины $n = 5$. Далее, везде под алфавитом подразумевается количество различных мономерных хвеньев в первичной структуре. Общая подпоследовательность – это подпоследовательность, содержащая буквы (нуклеотиды) как первой, так и второй последовательности, причем подпоследовательность необязательно содержит буквы, идущие непосредственно друг за другом. Так, например, для двух последовательностей S_1 и S_2 можно выделить несколько различных общих подпоследовательностей, например, $\{C, U, A, C\}$ или $\{G, A, C\}$ – обе эти подпоследовательности содержатся в S_1 и S_2 , и являются для них общими. Число возможных общих подпоследовательностей с ростом длин m и n полимеров растет экспоненциально. Алгоритм для определения оптимального выравнивания двух последовательностей впервые был сформулирован в [6]. В наиболее общем смысле каждое выравнивание двух последовательностей характеризуется числом совпадающих и несовпадающих букв и числом пропусков (делеций) в выравненных последовательностях. Для каждого выравнивания можно ввести весовую функцию (cost function), имеющую значение энергии [6]:

$$F = N_{\text{match}} + \mu N_{\text{mis}} + \delta N_{\text{gap}}. \quad (1)$$

В формуле (1) N_{match} , N_{mis} и N_{gap} – число пар совпадающих букв, число пар несовпадающих букв и число делеций в рассматриваемом выравнивании, соответственно. Величины μ и δ – это вклады в весовую функцию от пары несовпадающих букв и делеции; вклад от пары совпадающих нуклеотидов, без потери общности, можно считать равным 1. В таком представлении функция F удовлетворяет очевидному закону сохранения:

$$n + m = 2N_{\text{match}} + 2N_{\text{mis}} + N_{\text{gap}}. \quad (2)$$

Используя (2), формулу (1) можно переписать в виде:

$$\tilde{F} = N_{\text{match}} + \gamma N_{\text{mis}}, \quad (3)$$

где

$$\gamma = \frac{\mu - 2\delta}{1 - 2\delta}. \quad (4)$$

Здесь интерес представляет область $0 \leq \gamma \leq 1$, так как, случай $\gamma < 0$ неотличим от $\gamma = 0$, а случай $\gamma > 1$ соответствует тому, что «несовпадения» более выгодны, чем «совпадения» и может быть учтен простым переопределением этих понятий. Заметим, что, хотя предлагаемая теория применима ко всему доступному интервалу значений γ , все численные результаты настоящей работы получены для случая $\gamma = 0$, который представляется наиболее физически осмысленным. Задача поиска НОП заключается в определении выравнивания с максимальным значением весовой функции F .

Оказывается, что для нахождения весовой функции F удобнее всего использовать рекурсивный алгоритм, известный как метод динамического программирования:

$$\tilde{F}_{i,j}^{\max} = \max \left[\tilde{F}_{i-1,j}^{\max}, \tilde{F}_{i,j-1}^{\max}, \tilde{F}_{i-1,j-1}^{\max} + \zeta_{i,j} \right], \quad (5)$$

где

$$\zeta_{i,j} = \begin{cases} 1, & \text{для } S_1(i) = S_2(j) \\ \gamma, & \text{для } S_1(i) \neq S_2(j) \end{cases} \quad (6)$$

Выражения (5)–(6) имеют следующий смысл. Начиная с левых концов последовательностей, на каждом шаге выбирается такое положение букв в выравнивании, которое вносит наибольший вклад в функцию F . Члены в (5) соответствуют трем возможным ситуациям: пропуску буквы в первой последовательности, пропуску во второй

последовательности и случаю, когда i -ая буква первой последовательности выравнена с j -ой буквой второй последовательности.

В. Комплементарное связывание биополимеров

Цель работы заключается в разработке статистического алгоритма вычисления весовой функции, которая бы характеризовала «похожесть» двух заданных последовательностей со сложной вторичной структурой типа РНК. Эта функция должна включать как энергетический вклад от непосредственного взаимодействия мономеров друг с другом, так и энтропийный вклад, обусловленный наличием ансамбля пространственных конформаций макромолекул. При этом постараемся, по возможности, остаться в рамках статистической физики и избежать неконтролируемых эвристических соображений, апеллирующих к опыту, полученному в результате анализа экспериментальных данных.

Прежде всего покажем, что рекуррентное соотношение (5) имеет прозрачный физический смысл в терминах статистической физики и формулы (5), (6) можно рассматривать как свободную энергию статистической модели, описывающей комплексообразование двух взаимодействующих линейных полимеров в пределе нулевой температуры. Затем, учитывая возможность того, что каждый из полимеров может, помимо собственно комплексообразования, образовывать сложную иерархическую структуру, обобщим выражение для статистической суммы (соответствующей ненулевой температуре) на комплексы с внутренней иерархической структурой. Переходя в конечном выражении снова к пределу $T \rightarrow 0$, найдем искомую весовую функцию.

Рассмотрим вспомогательную статистическую модель, описывающую взаимодействие двух линейных полимеров с произвольными первичными последовательностями. Пусть длины этих последовательностей, измеренные в единицах мономерных звеньев, равны m и n , соответственно. Каждый мономер может быть выбран из c различных мономеров A, B, C, D, \dots (Для последовательностей РНК $c = 4$). Мономеры первой последовательности могут образовывать связи с мономерами второй последовательности. В молекулах РНК такие связи образуются согласно комплементарности азотистых оснований. Будем считать энергию связи между комплементарными нуклеотидами равной $-u$, а энергию между некомплементарными равной $-v$, где u и v — некоторые положительные величины ($|v| > |u|$). Предположим также, что некоторые части полимеров могут образовывать петли. На Рис. 1 схематически представлено взаимодействие

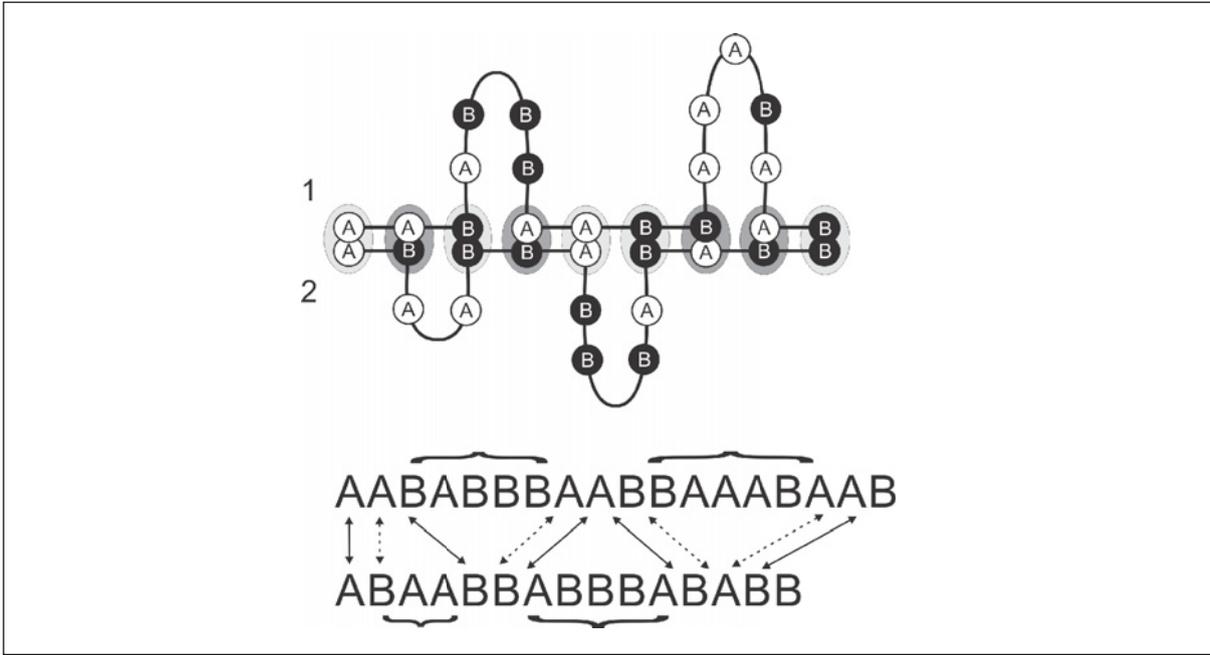


Рис. 1. Взаимодействие мономерных звеньев в РНК-подобной структуре с петлевыми участками как выравнивание соответствующих последовательностей (первичных структур).

двухбуквенных полимеров. Очевидно, что петли соответствуют делециям в задаче о выравнивании двух последовательностей.

Задача заключается в вычислении свободной энергии описанной модели при достаточно низких температурах, при которых энтропийным вкладом можно пренебречь по сравнению с энергетическим. Пусть $G_{m,n}$ – статистическая сумма рассматриваемого комплекса. По смыслу $G_{m,n}$ – это сумма по всем возможным конфигурациям связей. При низких температурах $G_{m,n}$ можно представить как:

$$\begin{cases} G_{m,n} = 1 + \sum_{i,j=1}^{m,n} \beta_{i,j} G_{i-1,j-1} \\ G_{m,0} = 1; G_{0,n} = 1; G_{0,0} = 1. \end{cases} \quad (7)$$

Смысл данной формулы очевиден: начиная с левого конца последовательностей (Рис. 1), находим первый существующий контакт между i -м мономером первой цепи и j -м мономером второй, а далее суммируем по всем возможным расположениям этого контакта. Статистические веса связей $\beta_{i,j}$ определяются энергией контакта между i -ым и j -ым мономерами:

$$\beta_{i,j} = \begin{cases} \beta^+ \equiv e^{u/T}, & S_1(i) \text{ и } S_2(j) \text{ комплементарны} \\ \beta^- \equiv e^{v/T}, & S_1(i) \text{ и } S_2(j) \text{ не комплементарны.} \end{cases} \quad (8)$$

Здесь и далее, $T \equiv k_B T$. Легко проверить, что статистическая сумма вида (7) удовлетворяет рекуррентному соотношению:

$$G_{m,n} = G_{m-1,n} + G_{m,n-1} + (\beta_{m,n} - 1) G_{m-1,n-1}. \quad (9)$$

В свою очередь, статистическая сумма связана со свободной энергией комплекса $F_{m,n}$ и температурой T известным соотношением $G_{m,n} = \exp\{-F_{m,n}/T\}$. Будем интересоваться значением свободной энергии с точностью до знака, тогда для величины $\tilde{F}_{m,n} = -F_{m,n}$, переходя в уравнении (9) к пределу $T \rightarrow 0$, получим:

$$\tilde{F}_{m,n} = \lim_{T \rightarrow 0} T \ln \left(e^{\tilde{F}_{m-1,n}/T} + e^{\tilde{F}_{m,n-1}/T} + (\beta_{m,n} - 1) e^{\tilde{F}_{m-1,n-1}/T} \right). \quad (10)$$

Формулу (10) можно переписать в виде:

$$\tilde{F}_{m,n} = \max \left[\tilde{F}_{m-1,n}, \tilde{F}_{m,n-1}, \tilde{F}_{m-1,n-1} + \eta_{m,n} \right], \quad (11)$$

где введено обозначение:

$$\begin{aligned} \eta_{m,n} &= T \ln(\beta_{m,n} - 1) = \\ &= \begin{cases} \eta^+ = T \ln(e^{u/T} - 1), & \text{если } S_1(i) \text{ и } S_2(j) \text{ комплементарны} \\ \eta^- = T \ln(e^{v/T} - 1), & \text{если } S_1(i) \text{ и } S_2(j) \text{ не комплементарны.} \end{cases} \end{aligned} \quad (12)$$

Принимая η^+ за единицу энергии, перепишем формулу (11) в виде:

$$\tilde{F}_{m,n} = \max \left[\tilde{F}_{m-1,n}, \tilde{F}_{m,n-1}, \tilde{F}_{m-1,n-1} + \tilde{\eta}_{m,n} \right], \quad (13)$$

с

$$\tilde{\eta}_{m,n} = \begin{cases} 1, & \text{если } S_1(i) \text{ и } S_2(j) \text{ комплементарны} \\ a = \frac{\eta^-}{\eta^+} = \frac{\ln(e^{v/T} - 1)}{\ln(e^{u/T} - 1)} \Big|_{T \rightarrow 0} = \frac{v}{u}, & \text{если } S_1(i) \text{ и } S_2(j) \text{ не комплементарны.} \end{cases} \quad (14)$$

И функция $\tilde{F}_{m,n}$ удовлетворяет начальным условиям: $\tilde{F}_{0,n} = \tilde{F}_{n,0} = \tilde{F}_{0,0} = 0$. Видно, что выражение свободной энергии связывания двух полимеров без петлевых взаимодействий имеет вид, совпадающий с (5). Далее, все результаты численного моделирования представлены для $\frac{v}{u} = 0$.

Таким образом, рекурсия, используемая в методе динамического программирования является ничем иным, как рекуррентным соотношением на свободную энергию взаимодействия гетерополимеров в пределе нулевой температуры. В природе существует множество примеров образования подобных гетерополимерных комплексов, например, образование двойной спирали ДНК.

Отметим, что предложенная выше модель является лишь первым приближением к описанию комплексообразования биополимеров. Известно (см., например, [2]), что

для точного количественного описания такого связывания, например, двойной спирали ДНК необходимо учесть еще ряд факторов. Во-первых, не учтены так называемые «петлевые факторы»: при образовании петли возможные конформации полимера ограничены условием, что ее концы обязаны сойтись в одной точке пространства, поэтому образование каждой петли приводит к снижению энтропии комплекса. Во-вторых, в реальной ДНК имеется выраженная кооперативность образования связей: вероятность образования связи выше, если соседние мономеры также образуют связь. В-третьих, не учтено, что гибкость полимера конечна и, тем самым, существует ограничение на минимальную длину петли. И наконец, не было принято во внимание то обстоятельство, что комплементарные пары $A - T$ и $C - G$ имеют различную энергию связи и, что помимо комплементарных пар, возможно образование неканонических пар.

Обобщение выражений (7)–(14) с учетом кооперативности образования связи, минимальной длины петли и различной энергией комплементарных связей — задача вычислительно сложная, но не требующая качественного изменения предложенного формализма, т.к. эти факторы влияют только на локальные свойства полимерных цепей. С другой стороны, петлевой фактор — характеристика нелокальная, зависящая от расстояния между мономерами, образующими связь и в этом случае нельзя описать состояние комплекса уравнениями динамического программирования, вида (7). Однако, поскольку петлевые факторы имеют энтропийную природу, в пределе низких температур ($T \rightarrow 0$) их вклад в свободную энергию гетерополимерного комплекса становится пренебрежимо мал. Ситуация усложняется, если сами петли могут образовывать вторичную структуру (т.е. если внутри петли имеет место взаимодействие между мономерами), а именно такая ситуация типична для последовательностей РНК. В этом случае энергетический вклад от вторичной структуры петли сохраняется и в пределе нулевой температуры, и его учет становится необходим.

С. Связывание РНК с внутриветлевым взаимодействием

В этом разделе обобщается модель взаимодействия двух сополимеров на случай, когда возможно комплементарное связывание внутри петель комплекса. Будем рассматривать иерархические структуры петель типа клеверного листа (Рис. 2(а)), структуры типа псевдоузлов (Рис. 2(б)) в данной работе не рассматриваются.

Как и в предыдущем параграфе для простоты не будем учитывать кооперативность образования связей и различие в энергиях комплементарных пар. Однако, как уже

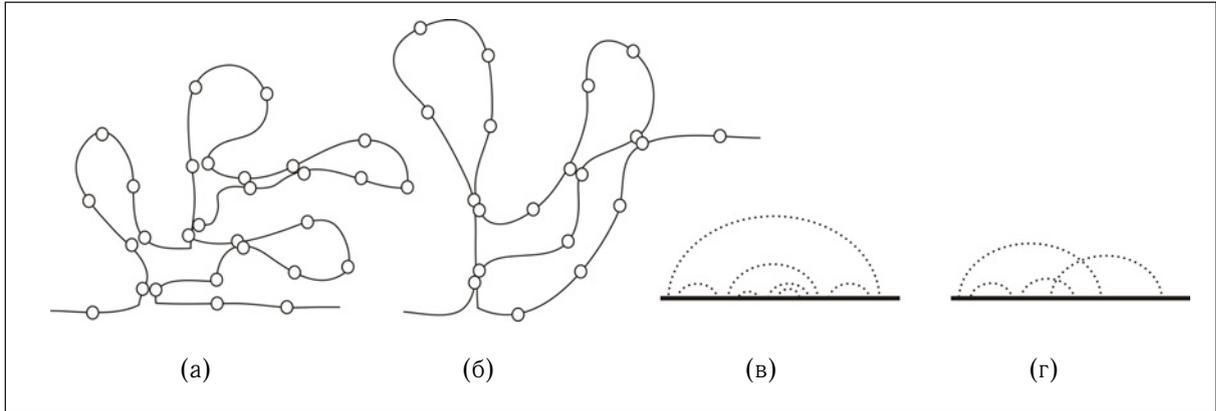


Рис. 2. Клеверная структура РНК (а) и псевдоузел (б); (в) и (г) — арочное представление (а) и (б), соответственно



Рис. 3. Диаграмма для вычисления статистического веса g последовательности

указывалось, модель может быть обобщена с учетом этих факторов. Выражение (7) для статистической суммы $G_{m,n}$ двух взаимодействующих сополимеров можно переписать в виде:

$$\begin{cases} G_{m,n} = g_{1,m}^{(1)} g_{1,n}^{(2)} + \sum_{i,j=1}^{m,n} \beta_{i,j} G_{i-1,j-1} g_{i+1,m}^{(1)} g_{j+1,n}^{(2)} \\ G_{m,0} = g_{1,m}^{(1)}; \quad G_{0,n} = g_{1,n}^{(2)}; \quad G_{0,0} = 1, \end{cases} \quad (15)$$

где $g_{i,j}^{(1)}$ и $g_{i,j}^{(2)}$ обозначены статистические веса участков (с i -го нуклеотида до j -го) первой и второй последовательности, соответственно, удовлетворяющие уравнениям:

$$\begin{cases} g_{i,j}^{(a)} = 1 + \sum_{k=i}^{j-1} \sum_{l=i+1+l}^j \beta'_{k,l} g_{k+1,l-1}^{(a)} g_{l+1,j}^{(a)}; \\ g_{i,i}^{(a)} = 1, \quad a = 1, 2. \end{cases} \quad (16)$$

Эти уравнения отвечают за топологию кактусообразной структуры, свойственной молекулам РНК, диаграмма, описывающая такие структуры представлена на Рис.3. Коэффициенты $\beta'_{i,j}$ – это константы, описывающие взаимодействие внутри последовательности, аналогичные $\beta_{m,n}$. Суммирование по j ведется от $i + 1 + \ell$ до n для того чтобы исключить петли длиной меньше ℓ мономеров. В последующих вычислениях, как правило, предполагается, что $\ell=0$, также обсуждается случай $\ell = 3$. Напомним еще раз, что так

как интерес представляет низкие температуры, можно пренебречь вкладом, связанным с потерей энтропии при образовании петель.

Сложную систему уравнений на статистические веса петлевых участках $g_{i,j}^{(a)}$, $a = 1, 2$ (16) можно решить следующим образом. Для каждой из последовательности РНК можно построить матрицу g , (i, j) -й элемент которой определяет статистический вес участка, начинающейся с i -го нуклеотида и заканчивающейся j -м. Таким образом, статистические веса всех возможных петель описываются матрицами размера $m \times m$ для первой последовательности и $n \times n$ для второй. Из граничных условий (16) можно однозначно определить элементы $g_{i,i+1}^{(a)}$. Из (16) следует, что элементы последующих субдиагоналей $g_{i,i+k}^{(a)}$ зависят только от элементов предыдущих субдиагоналей $g_{i,i+k-1}^{(a)}$ матрицы:

$$g_{i,i+k}^{(a)} = g_{i+1,i+k}^{(a)} + \sum_{s=i+1}^{i+k} \beta'_{i,s} g_{i+1,s-1}^{(a)} g_{s+1,i+k}^{(a)}. \quad (17)$$

Определенные таким образом матрицы статистических весов $g^{(a)}$ всех возможных петель позволяют вычислить статистическую сумму взаимодействия двух РНК с внутри-петлевым взаимодействием (15).

Как и в случае связывания последовательностей без петлевых участков, можно выполнить переход к пределу нулевой температуры – см. выражения (7)-(14). Элементы матрицы свободной энергии при этом можно представить в виде:

$$F_{m,n} = \max_{\substack{i=1,\dots,m \\ j=1,\dots,n}} [f_{1,m}^{(1)} + f_{1,n}^{(2)}, Q_{i,j}^{m,n}] \quad (18)$$

где $f_{i,j}^{(a)} = \lim_{T \rightarrow 0} [T \ln g_{i,j}^{(a)}]$ ($a = 1, 2$) имеют смысл с точностью до знака свободных энергий петлевых участков последовательностей с i -го нуклеотида по j -й, $Q_{i,j}^{m,n}$ – (i, j) -ый элемент суммы (15), который в пределе нулевой температуры есть:

$$Q_{i,j}^{m,n} = F_{i-1,j-1} + f_{i+1,m}^{(1)} + f_{j+1,n}^{(2)} + \tilde{\eta}_{i,j}. \quad (19)$$

Элемент $Q_{i,j}$ описывает энергию комплекса взаимодействующих РНК, не имеющих контакта правее пары (i, j) . Из (17) следует, что функции $f_{i,j}^{(a)}$ удовлетворяют:

$$f_{i,i+k}^{(a)} = \max \left[f_{i+1,i+k}^{(a)}, \max_s \left(f_{i+1,s-1}^{(a)} + f_{s+1,i+k}^{(a)} + \tilde{\eta}_{i,s}^{(a)} \right) \right], \quad (20)$$

здесь величина $\tilde{\eta}_{i,j}$ – величина, как в (14), $\tilde{\eta}_{i,s}^{(a)}$ – аналогичная величина, описывающая взаимодействие внутри петель. На свободную энергию накладываются граничные условия, как это следует из (15):

$$\begin{cases} F_{0,0} = 0; \\ F_{i,0} = f_{1,i}^{(1)}; \quad 1 \leq i \leq m \\ F_{0,j} = f_{1,j}^{(2)}; \quad 1 \leq j \leq n. \end{cases} \quad (21)$$

Таким образом, для того, чтобы вычислить энергию основного состояния комплекса двух взаимодействующих РНК, необходимо построить матрицы $f^{(1)}$ и $f^{(2)}$ и, далее, применяя (18)-(19), определить элементы матрицы F .

Отметим, что выражения (17), (20) можно использовать для непосредственного вычисления свободной энергии основного состояния одноцепочечной РНК.

III. СВОЙСТВА РНК СТРУКТУР СО СЛУЧАЙНОЙ ПОСЛЕДОВАТЕЛЬНОСТЬЮ ЗВЕНЬЕВ

В данном разделе обсуждаются свойства распределения свободной энергии основного состояния в ансамбле РНК-подобных молекул со случайной первичной структурой. Также, приводятся результаты для распределения длин петель в РНК-подобных структурах и обсуждаются аналитические модели их описания.

А. Свободная энергия основного состояния

1. Связывание двух РНК с петлевыми участками

Задача поиска оптимальной конфигурации линейного выравнивания случайных последовательностей неоднократно рассматривалась в литературе (см., например, [7]) в рамках так называемой модели «бернуллиевого сравнения», т.е. в предположении о том, что матричные элементы $\eta_{m,n}$ (14) являются независимыми случайными величинами, принимающими значения 1 с вероятностью $p = c^{-1}$ и 0 с вероятностью $q = 1 - p$, где c — алфавит, используемый в случайной первичной структуре полимера. В работе [7] было показано что для длин последовательностей $n, m \gg 1$ распределение энергии основного состояния имеет вид:

$$\langle F_{m,n} \rangle = \frac{2\sqrt{pmn} - p(m+n)}{q} + \frac{(pmn)^{1/6}}{q} \left[(1+p) - \sqrt{\frac{p}{mn}}(m+n) \right]^{2/3} \chi \quad (22)$$

где χ — случайная величина с распределением Трейси–Видома ($\langle \chi \rangle = -1.7711\dots$ и $\langle \chi^2 \rangle - \langle \chi \rangle^2 = 0.8132\dots$) (более подробное описание этого распределения можно найти, например, в обзоре [8]). При $m = n$, оптимальная конфигурация характеризуется:

$$\langle F_{n,n} \rangle \approx \frac{2}{1 + \sqrt{c}} n + f(c) \langle \chi \rangle n^{1/3}, \quad (23)$$

где

$$f(c) = \frac{c^{1/6} (\sqrt{c} - 1)^{1/3}}{\sqrt{c} + 1}.$$

Флуктуации свободной энергии подчиняются [7]:

$$\sigma \equiv \sqrt{\langle F_{n,n}^2 \rangle - \langle F_{n,n} \rangle^2} \approx \sqrt{\langle \chi^2 \rangle - \langle \chi \rangle^2} f(c) n^{1/3}. \quad (24)$$

Показатель $1/3$ является типичным для стохастической динамики сильно коррелированных систем и относится к классу универсальности Кардара-Паризи-Занга (Kardar-Parisi-Zhang (KPZ)) [9].

Результаты численного моделирования распределения свободной энергии основного состояния для ансамбля случайных первичных структур РНК представлены на Рис. 4. Угловой коэффициент прямой $k \approx 0.65$ (Рис. 4(а)), что хорошо согласуется с величиной $k = \lim_{n \rightarrow \infty} \frac{\langle F_{n,n} \rangle}{n} \rightarrow \frac{2}{3}$, вычисленной по формуле (23). Для флуктуации энергии полученный наклон 0.34 (Рис.4(б)) также близок к значению $\frac{1}{3}$. Таким образом, уравнение (23), полученное в приближении бернуллиевского сравнения, удовлетворительно описывает численно наблюдаемую зависимость энергии основного состояния при связывании сополимеров с петлевыми участками от длины случайных цепей.

2. Связывание двух РНК с внутриветлевым взаимодействием

Аналогичный анализ был проведен и для двух последовательностей, образующих структуру с внутриветлевым взаимодействием и минимальной длиной петли $\ell = 0$. Соответствующие графики зависимости свободной энергии и флуктуации энергии представлены на Рис. 5. Как и для взаимодействия с петлевыми участками, $\langle F_{n,n} \rangle (n) = kn$ при $n \gg 1$ (Рис. 5), но угловой коэффициент прямой $k \approx 0.92$ гораздо выше, что обусловлено взаимодействием нуклеотидов внутри петель. Зависимость флуктуации энергии основного состояния остается такой же (см. Рис. 5(б)).

Оценим аналитически величину коэффициента k в зависимости свободной энергии от длины цепи для внутриветлевого взаимодействия (Рис. 5). Будем рассматривать комплекс, который образуют две случайные последовательности РНК, как структуру, состоящую из петель различных иерархических уровней, занумерованных индексом i (см. Рис. 6).

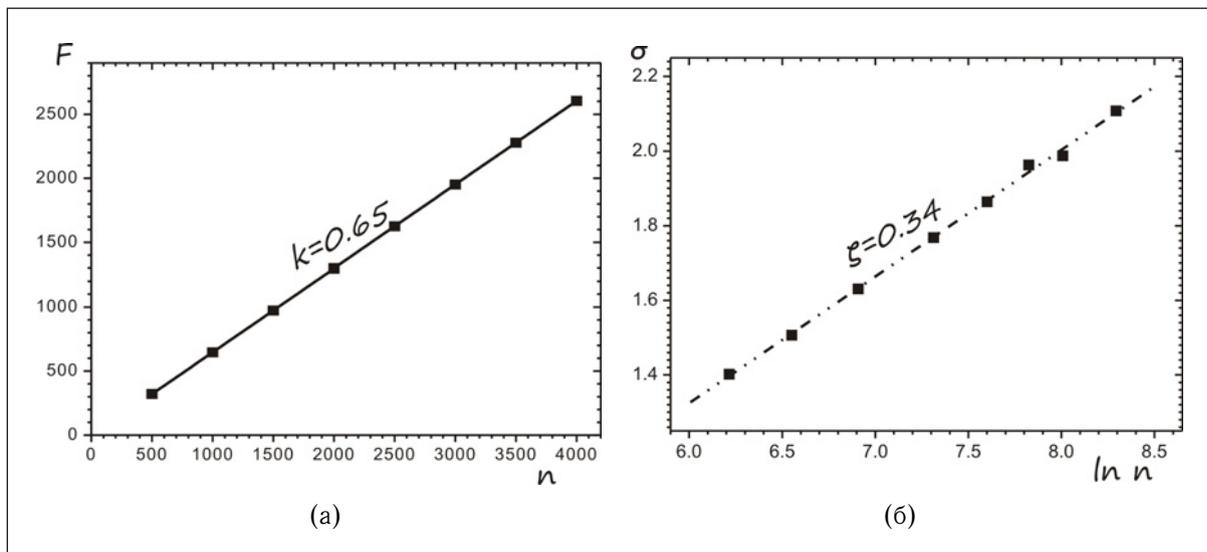


Рис. 4. Взаимодействие РНК с петлевыми участками: зависимость среднего значения свободной энергии основного состояния $F_{n;n}$ (а) и флуктуации энергии σ (б) от длины случайной последовательности n . Усреднение проводилось по ансамблю из 10^5 случайных пар последовательностей для каждого значения длины.

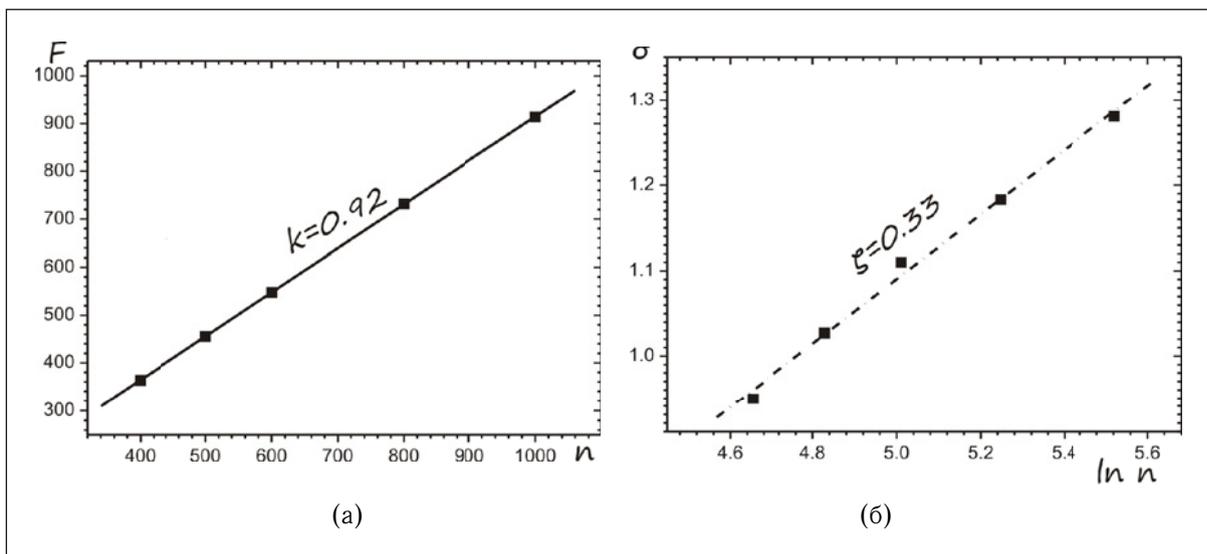


Рис. 5. Связывание РНК с внутриветлевым взаимодействием: зависимость энергии основного состояния $F_{n;n}$ (а) и флуктуации свободной энергии σ (б) от длины случайной последовательности n . Усреднение проводилось по ансамблю из 10^5 случайных пар последовательностей для каждого значения длины

Каждую петлю i -ого иерархического уровня можно рассматривать как комплекс двух взаимодействующих подпоследовательностей из которых она состоит. Из выражения (22) следует, что наибольший вклад в свободную энергию наблюдается для комплекса, состоящего из двух последовательностей равной длины, $m = n$. Это позволяет оценить сверху свободную энергию петли как свободную энергию двух взаимодействующих половинок этой петли. Представление комплекса двух молекул РНК в виде иерархи-

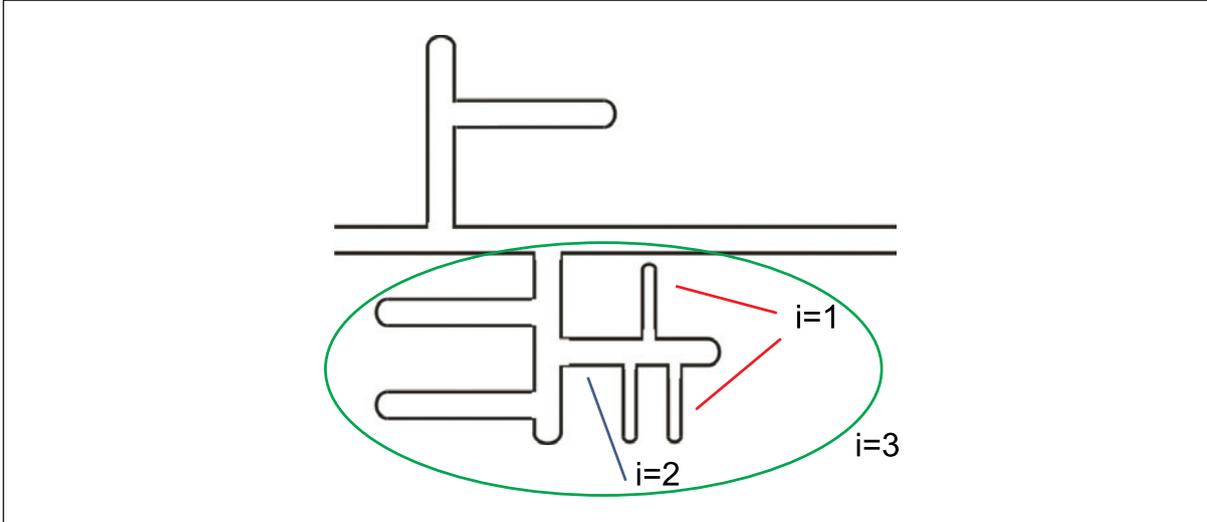


Рис. 6. Иерархическая модель связывания двух полимеров с внутривитлевым взаимодействием. Петли первого ($i = 1$), второго ($i = 2$) и третьего ($i = 3$) иерархических уровней

ской структуры позволяет использовать идеи ренормализационной группы. А именно, комплексы i -ого иерархического уровня содержат петли, которые будем считать комплексами $(i + 1)$ -ого уровня (Рис.6) ($i = 1, 2, \dots$).

Формализуя эту идею, будем полагать, что комплекс двух молекул РНК иерархического уровня i – это комплекс двух последовательностей с петлевыми участками, в которых энергия взаимодействующих мономеров перенормирована энергией петель иерархического уровня $(i + 1)$. Пользуясь тем, что энергия петель в первом приближении пропорциональна длине (23), представим ее в виде: $F_s^{(i)} \approx k_r^{(i)} s$, где s – длина петли, а k_r – соответствующий i -ому уровню коэффициент связывания. Подставляя в формулу (15) статистические веса петель $g_{i,i+s} = e^{-k_r s/T}$, получим выражение для определения свободной энергии комплекса двух случайных РНК-последовательностей ¹:

$$F_{m,n}^{(i+1)} = \max \left[F_{m-1,n} + k_r^{(i)}, F_{m,n-1} + k_r^{(i)}, (F_{m-1,n-1} + u) \mathcal{P}(m, n) \right]. \quad (25)$$

Выражение (25) нужно понимать следующим образом. Прежде всего, определим свободную энергию комплекса $F_{m,n}^{(2)}$, в котором могут образовываться петли только *первого* иерархического уровня. Далее определим энергию связывания на один мономер в петлях *второго* уровня как

$$k_r^{(2)} = \frac{F_{m,n}^{(2)}}{m + n}. \quad (26)$$

¹ Здесь, как и ранее, F имеет смысл свободной энергии с обратным знаком

Подставляя полученный коэффициент связывания снова в формулу (25), получим значения энергии для петель *третьего* иерархического уровня, $k_r^{(3)}$, и т.д. Величина $\mathcal{P}(m, n)$ учитывает ограничение на минимальное количество мономеров, которые могут образовать петлю *i*-ого иерархического уровня:

$$\mathcal{P}(m, n) = \begin{cases} 1 & \text{мономеры } m \text{ и } n \text{ могут образовать связь} \\ 0 & \text{в противном случае} \end{cases} \quad (27)$$

Будем считать, что *m*-й и *n*-й мономеры могут образовать связь, если:

- а) участок $[m - l_{min}^i, m - 1]$ последовательности S_1 *не имеет* связей с участком $[n - l_{min}^i, n - 1]$ подпоследовательности S_2 , где l_{min}^i – минимальное количество нуклеотидов, необходимых для формирования петли определенного уровня (если $m < l_{min}^i$ и/или $n < l_{min}^i$, то рассматриваются соответственно участки последовательностей $[1, m]$ и/или по $[1, n]$);
- б) *m* – 1-й мономер первой последовательности взаимодействует с *n* – 1-м мономером второй последовательности, и при замене $(m - 1) \rightarrow m$, $(n - 1) \rightarrow n$ выполняется а) (или б)).

В таблице I приведены значения для коэффициента связывания и минимальное количество нуклеотидов в петлях *i*-ого уровня; вычисления проводились для случайных последовательностей равной длины $m = n = 10^4$. Длины последовательностей слабо влияют на средний коэффициент связывания, однако рассмотрение больших длин позволяет провести оценку для большего количества иерархических уровней. Отметим, что коэффициент связывания, определяемый по данной иерархической процедуре, медленно (логарифмически) стремится к 1 с ростом количества иерархических уровней (т.е. при $n \rightarrow \infty$). Логарифмическая зависимость обусловлена экспоненциальным ростом минимального числа мономеров, которые могут образовать петлю, $l_{min}^i = 3l_{min}^{i-1} + 6$ ($i > 2$) с увеличением номера иерархического уровня *i* (см. Табл. I).

Таблица 1

Вероятность связывания мономеров в зависимости от числа уровней в иерархической модели взаимодействия двух полимеров

Уровень, <i>i</i>	2	3	4	5	6	7
Минимальная длина петли	2	6	24	78	240	726
Коэффициент связывания	0.851	0.912	0.931	0.937	0.94	0.941

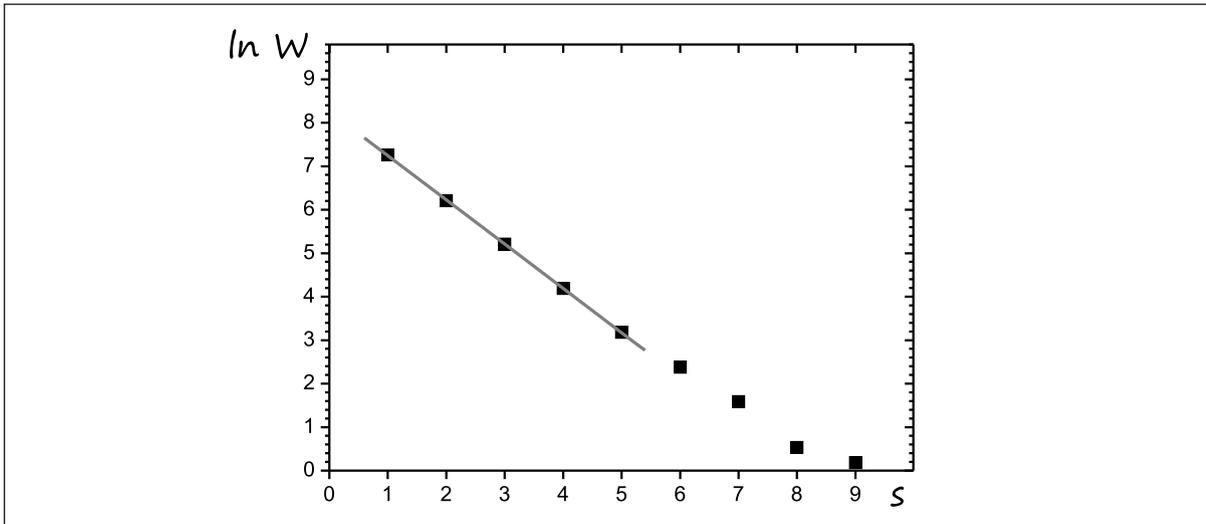


Рис. 7. Распределение длин петель в структуре комплекса с петлевыми участками. Вычисления были выполнены для случайных последовательностей длины $N = 10^4$, результаты усреднялись по набору из 10^5 сополимеров.

Таким образом, численно наблюдаемый коэффициент связывания k (Рис. 5(a)) в действительности зависит от длин рассматриваемых последовательностей и полученное нами значение $k \approx 0.92$ лишь указывает на то, что последовательности длиной $400 \div 1000$ мономеров образуют структуру *всего с двумя–тремя иерархическими уровнями*.

В. Распределение длин петель в РНК-подобных структурах

1. Связывание двух РНК с петлевыми участками

Было проанализировано распределение длин петель в структуре комплекса с петлевыми участками и внутривитлевым взаимодействием. На Рис. 7 представлена зависимость $W(s)$ числа петель различной длины s для структуры с петлевыми участками. Видно, что зависимость с хорошей точностью является экспоненциальной. Такое распределение характерно для системы, в которой связывание различных мономеров в цепи происходит независимо (т.е. вероятность того, что следующий по цепи мономер образует связь, никак не зависит от того, образует ли связь предыдущий мономер). Действительно, величину $k = \frac{\langle F_{n,n} \rangle}{n}$ при $n \gg 1$ можно рассматривать, как вероятность связывания мономера в структуре. Считая, что взаимодействие мономеров независимым, число петель длиной s в структуре двух взаимодействующих сополимеров длиной n можно оценить, как:

$$W(s) = nk^2(1 - k)^s. \quad (28)$$

Такое распределение длин петель при $n \gg 1$ удовлетворяет очевидному соотношению $\sum_{s=1}^n sW(s) = (1 - k)n$. Из Рис. 7 видно, что численные результаты хорошо аппроксимируются в логарифмическом масштабе прямой $y(s) = a - bs$, где с хорошей точностью $a \approx \ln(nk^2)$ и $b \approx \ln(1 - k)$ (см. (28)). Таким образом, в связывании сополимеров с петлевыми участками статистика петель выглядит в точности так, как происходит при *независимом* связывании мономеров. Однако стоит отметить, что модель независимого связывания дает хорошие результаты для последовательностей, в которых количество различных сортов мономеров $c \geq 4$. Для двухбуквенных и трехбуквенных алфавитов, взаимодействие сополимеров оказывается коррелированным, и формула (22) плохо описывает энергию оптимальной конфигурации.

2. Связывание двух РНК с внутриветлевым взаимодействием

Существенно иное поведение имеет статистика петель в комплексах с внутриветлевым взаимодействием. На Рис.8(а) представлена зависимость числа петель с длиной s по набору из 10^3 пар случайных последовательностей. Отметим особенности наблюдаемого распределения. Во-первых, для данной зависимости характерно степенное поведение. Показатель степенной зависимости для РНК разной длины меняется в интервале $[1.38, 1.5]$. Во-вторых, распределения для РНК с различной длиной n совпадают, что позволяет проводить вычисления для набора коротких последовательностей. В-третьих, при малых n ($n \leq 5$) характерно небольшое число петель с нечетной длиной и большое число петель с четной длиной. Последнее обстоятельство связано с тем, что для структуры комплекса с внутриветлевым взаимодействием и $\ell = 0$ характерно высокое значение средней энергии на один нуклеотид ($k \approx 0.92$), обусловленное связыванием внутри петель, а образование петли малой длины с нечетным числом нуклеотидов приводит к потере, по крайней мере, одной возможной связи внутри петли. Таким образом, образование петель с нечетным числом мономеров энергетически невыгодно. Наконец, для распределения характерно наличие плато при больших s , что обусловлено эффектом конечного размера (см., например, [10], где построена теория аналогичного эффекта).

Полученные численные распределения можно интерпретировать следующим образом. Поставим каждой вторичной структуре полимера в соответствие одномерное случайное блуждание на (1+1)-мерной решетке, построенное следующим образом (см. Рис. 9). Каждому мономерному звену соответствует один шаг блуждания. Этот шаг

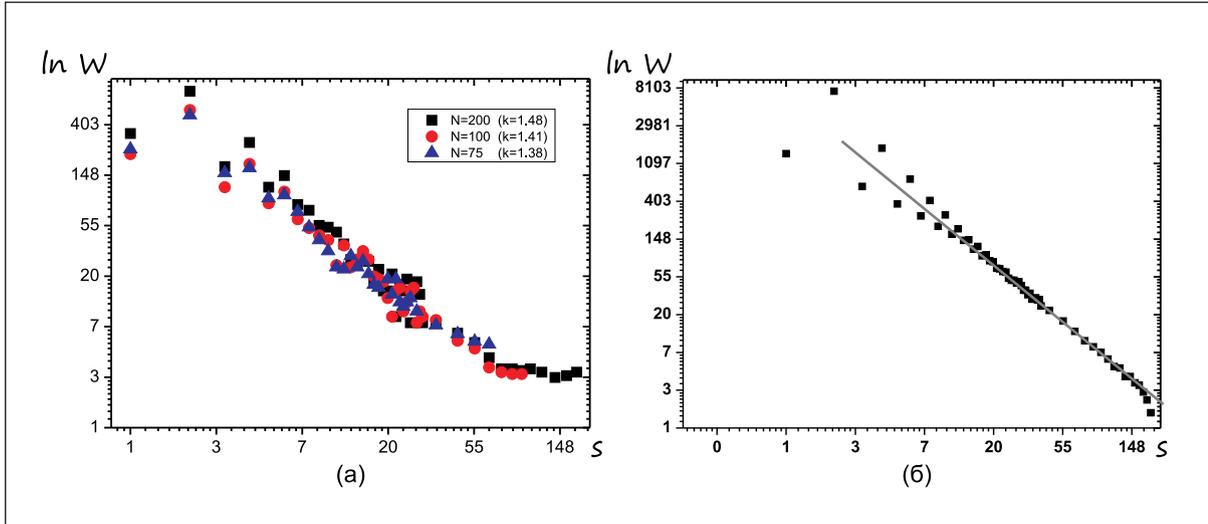


Рис. 8. (а) Распределение длин петель в структуре комплекса с внутриветлевым взаимодействием. Вычисления проводились для последовательностей с длинами $n = m = 75; 100$ и 200 , для каждого n было выполнено 10^3 накоплений, для $n \geq 30$ функция распределения сглаживалась по 10 соседним значениям); (б) Распределение путей Моцкина по длинам (длина пути случайного блуждания 200 шагов, количество накоплений $\sim 10^4$, для $n \geq 30$ функция распределения сглаживалась по 10 соседним значениям).

направлен направо вверх, если мономер является «началом петли» (т.е. связан с мономером, расположенным *после* него по цепи), направо вниз, если он является «концом петли» (т.е. связан с мономером, расположенным *до* него по цепи) или горизонтально, если мономер не образует связи. Легко видеть, что такое построение задает соответствие между РНК-подобными вторичными структурами и так называемыми путями Моцкина [11] — состоящими из горизонтальных и диагональных участков дискретными случайными блужданиями в верхней полуплоскости, концы которых закреплены на оси абсцисс. Возвращение на ось абсцисс соответствует образованию одной петли в структуре комплекса. Как известно, [12], количество различных путей Моцкина $W_M(s, t)$ длины s с заданным количеством горизонтальных шагов t определяется числами Каталана:

$$W_M(s, t) = \binom{s}{t} C_{(s-t)/2} = \binom{s}{t} \frac{1}{\frac{s-t}{2} + 1} \binom{s-t}{\frac{s-t}{2}}, \quad (29)$$

где $\binom{s}{t}$ — биномиальные коэффициенты, $C_{(s-t)/2}$ — числа Каталана. При $s \gg 1$ (29) имеет асимптотическую зависимость $W_M(s, t) \sim s^{-3/2}$ от длины пути. Было построено распределение длин петель для случайных путей Моцкина с вероятностью диагонального шага вверх или вниз равной $p_M \approx \frac{k}{2} = 0.46$, где $k = 0.92$ — наблюдаемое в численном моделировании значение вероятности образования связи, а вероятность

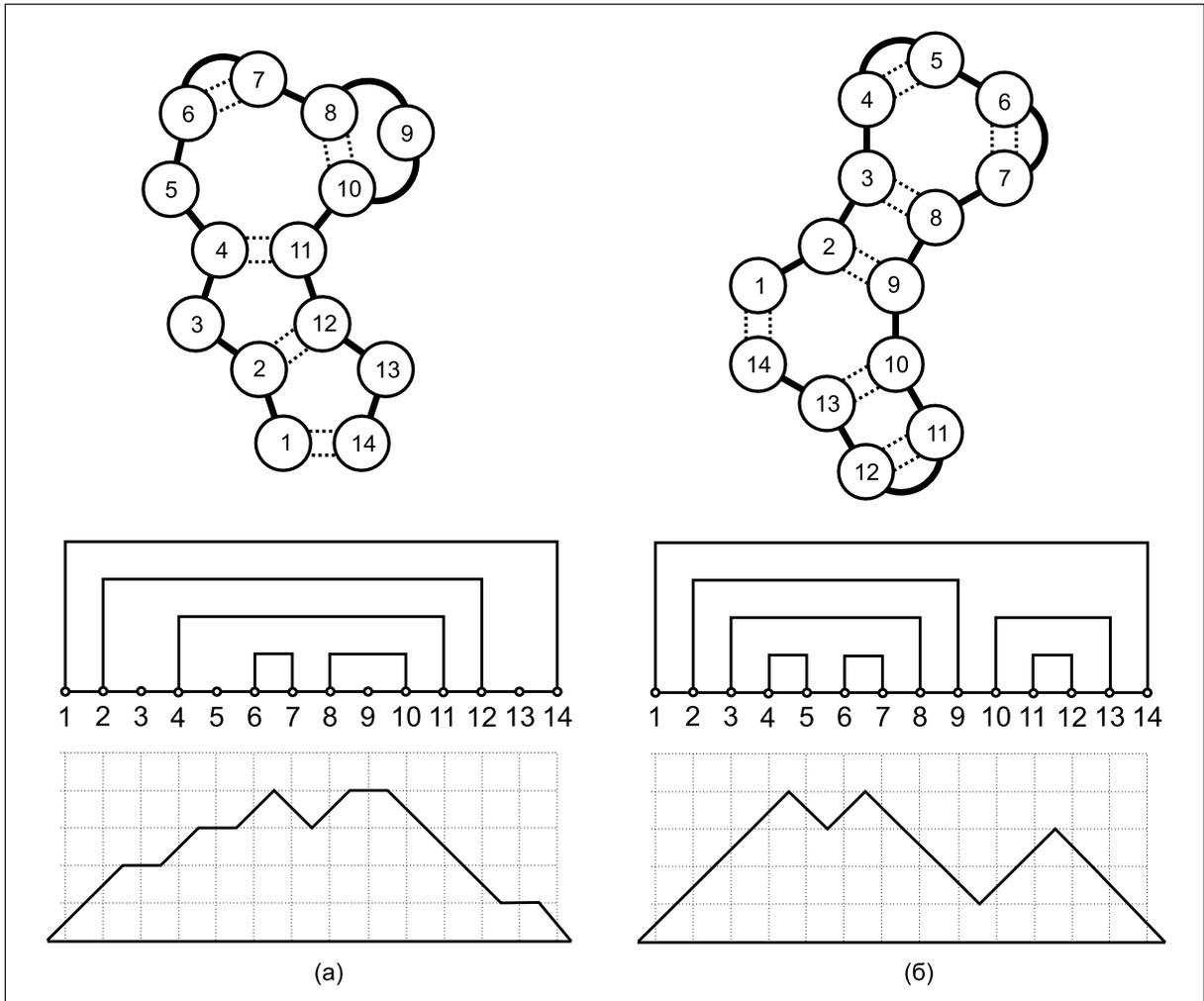


Рис. 8. (а) Распределение длин петель в структуре комплекса с внутриветлевым взаимодействием. Вычисления проводились для последовательностей с длинами $n = m = 75; 100$ и 200 , для каждого n было выполнено 10^3 накоплений, для $n \geq 30$ функция распределения сглаживалась по 10 соседним значениям); (б) Распределение путей Моцкина по длинам (длина пути случайного блуждания 200 шагов, количество накоплений -10^4 , для $n \geq 30$ функция распределения сглаживалась по 10 соседним значениям).

горизонтального шага $1 - 2p_M$. Результат приведен на рисунке 8(б). Видно, что зависимость обладает всеми характерными свойствами, наблюдаемыми для распределения длин петель в структуре с внутриветлевыми взаимодействиями.

Представление структур РНК в виде путей Моцкина, статистика которых известна, позволяет сделать интересное наблюдение. А именно, показать, что для РНК-подобных структур характерно критическая зависимость структуры основного состояния в зависимости от числа различных сортов мономеров, используемых в последовательности.

IV. ТОПОЛОГИЯ РНК-ПОДОБНЫХ МОЛЕКУЛ В ЗАВИСИМОСТИ ОТ АЛФАВИТА СЛУЧАЙНОЙ ПЕРВИЧНОЙ СТРУКТУРЫ

Раздел посвящен исследованию топологии пространственной структуры РНК-подобной молекулы в основном состоянии и ее изменений в зависимости от алфавита, используемого в случайной первичной структуре. А именно, показывается, что существует некоторая критическая точка (критический алфавит) в которой происходит изменение топологии основного состояния РНК-подобной молекулы. В разделе приводятся аналитические и численные оценки критической точки топологического перехода и обсуждается связь данного топологического перехода с температурным фазовым переходом в замороженное состояние.

А. Зависимость свободной энергии РНК-подобных структур от алфавита

Рассмотрим случайную последовательность длиной L и алфавита c , образующую вторичную структуру типа РНК (Рис. 2(а)). Зададимся вопросом о том, к какому пределу стремится доля комплементарных пар в основном состоянии длинной ($L \rightarrow \infty$) цепи РНК. Другими словами, интерес представляет удельная (в расчете на одно звено) энергия основного состояния длинной РНК. Вначале, приведем доводы, подтверждающие наличие критического изменения удельной энергии в зависимости от алфавита. Предположим, что существует критическое значение алфавита $c = c_c$ такое, что при $c < c_c$ доля связанных мономерных звеньев стремится к 1, тогда как при $c > c_c$ предельная доля связанных звеньев меньше 1. Убедиться в этом можно следующим образом. Для того чтобы доля связанных звеньев в РНК-подобной структуре, образуемой случайной последовательностью, была равна 1, каждой последовательности из c^L возможных должен соответствовать так называемый путь Дика (т.е. путь Моцкина, в котором нет горизонтальных шагов) (Рис. 9(б)). Количество путей Дика $G(L)$ длины L определяется формулой (29) ($G(L) = W_M(L, 0)$). При $t = 0$ и при $L \gg 1$ $G(L)$ имеет асимптотическое выражение

$$G(L) \sim \frac{4^{L/2}}{L^{3/2}}. \quad (30)$$

Заметим, что один и тот же путь Дика может описывать несколько РНК-подобных структур. Действительно, каждая пара подъем/спуск в пути Дика может быть, независимо от остальных, реализована c разными способами (в случае РНК возможные

варианты — это А-У, У-А, С-Г и Г-С). Таким образом, число различных первичных структур, для которых существуют полностью связанные вторичные структуры, не превышает

$$W(c, L) = G(L)c^{L/2} \sim \frac{(4c)^{L/2}}{L^{3/2}}. \quad (31)$$

Это оценка сверху, т.к., вообще говоря, одной и той же последовательности может соответствовать несколько различных РНК-подобных структур и, таким образом, несколько путей Дика. Тем не менее, естественно ожидать, что при $c \rightarrow c_c$ число таких последовательностей с двумя и более полностью связанными вторичными структурами становится малым. В таком случае, сравнивая (31) с полным числом возможных первичных структур $W_0(c, L) = c^L$, можно записать ($L \gg 1$):

$$\begin{cases} \lim_{L \rightarrow \infty} \frac{1}{L} \ln W(c, L) > \lim_{L \rightarrow \infty} \frac{1}{L} \ln W_0(c, L), \text{ для } 2 \leq c < c_c \\ \lim_{L \rightarrow \infty} \frac{1}{L} \ln W(c, L) < \lim_{L \rightarrow \infty} \frac{1}{L} \ln W_0(c, L), \text{ для } c > c_c. \end{cases} \quad (32)$$

Откуда, $c_c = 4$. Подчеркнем, что несмотря на то, что эта оценка является грубой оценкой сверху, она демонстрирует характерное изменение свойств РНК-структур со случайной первичной структурой.

Таким образом, при $c < c_c$ в пределе $L \rightarrow \infty$ практически любой последовательности соответствует полностью связанная вторичная структура, и энергия оптимальной конфигурации на одну пару нуклеотидов стремится к 1, в то время как для случайных цепочек с $c > c_c$ доля последовательностей, допускающих совершенную вторичную структуру, экспоненциально мала. Добавление горизонтальных шагов в пути случайных блужданий приводит к увеличению возможных РНК-подобных конфигураций (29), что позволяет сопоставить каждой случайной последовательности из ансамбля c^L путь в случайном блуждании, соответствующий ее оптимальной вторичной структуре. Однако в этом случае доля связанных звеньев в оптимальной вторичной структуре остается в пределе $L \rightarrow \infty$ меньше единицы. Путь Моцкина длиной L , включающий t горизонтальных шагов определяется (согласно (29)) как

$$W_M(L, t) = \frac{L!}{t!(L-t)!} C_{(L-t)}. \quad (33)$$

Для нечетных $(L-t)$ функция $W_M(L, t)$ равна 0. Для четных $(L-t)$ воспользуемся (30) и формулой Стирлинга для оценки асимптотического поведения:

$$\begin{aligned} \frac{1}{L} \ln W_M(L, a) &= -a \ln a - (1-a) \ln \frac{1-a}{4} + o\left(\frac{\ln L}{L}\right) \\ \frac{\partial}{\partial a} \frac{1}{L} \ln W_M(L, a) &\rightarrow +\infty, \text{ при } a \rightarrow +0, \end{aligned} \quad (34)$$

где введено обозначение $a = \frac{t}{L}$ ($a > 0$). Последнее выражение (34) показывает рост W_M для малых, но конечных a .

Как много различных структур могут иметь один и тот же путь Моцкина (Рис. 9(a))? Как и в случае полностью связанных структур, каждая связанная пара имеет вырожденность c , тогда как каждый несвязанный мономер также может быть выбран c разными способами. Суммарная вырожденность Z имеет вид

$$Z(c, L, a) = c^{(L-aL)/2} c^{aL} = c^{L(1+a)/2}, \quad (35)$$

и является возрастающей функцией a .

А теперь оценим минимальное количество несвязанных мономеров (горизонтальных шагов в пути Моцкина), $a(c) = 1 - f(c)$, в основном состоянии при $c > 4$. Наибольшее количество структур, имеющих в основном состоянии долю несвязанных мономеров меньше или равной a определяется выражением:

$$W(c, L, a) = \sum_{j=0}^{aL} Z(c, L, j/L) W_M(j, L). \quad (36)$$

Для $c > 4$ и $a = 0$ эта сумма меньше $W_0(c, L) = c^L$, и растет с увеличением a так, что при некотором \bar{a} величины $W(c, L, a)$ и $W_0(c, n)$ сравниваются. Для $L \gg 1$ сумму (36) можно оценить методом перевала. Введем обозначение

$$\Delta w(a, c) = \lim_{L \rightarrow \infty} \frac{1}{L} \ln \frac{W(c, L, a)}{W_0(c, L)}.$$

Тогда

$$\Delta w(a, c) = \begin{cases} -(1-a) \ln \frac{\sqrt{c}(1-a)}{2} - a \ln a; & a < a_m \\ \ln \left(1 + \frac{\sqrt{c}}{2} \right) > 0; & a > a_m, \end{cases} \quad (37)$$

где $a_m = \frac{\sqrt{c}}{2+\sqrt{c}}$. Для $a < a_m$ сумма в (37) определяется вкладом от верхней границы, тогда как для $a > a_m$ максимум достигается в точке a_m и, таким образом, не зависит от верхнего предела суммирования. Величина $\bar{a}(c)$ определяется из уравнения $\Delta w(a, c) = 0$. На Рис. 10 представлена функция $f(c) = 1 - a(c)$. Напомним, что данная оценка является верхней границей, так как не учитывает корреляции между оптимальными конфигурациями.

Оценка (37) сделана в предположении так называемого среднего поля: связывание на каждой паре подъем/спуск Рис. 9 происходит независимо с вероятностью $1/c$ и все пути случайных блужданий считаются статистически независимыми. В разделе IV С

приводится более точная оценка критического алфавита, учитывающая корреляции между конфигурациями.

Результаты численного моделирования для РНК-подобных структур со случайной последовательностью звеньев различного алфавита представлены на Рис. 10. Для простоты предполагается, что комплементарные связи образуются согласно правилу А–А, т.е., только одинаковые мономеры могут комплементарно связываться. Напомним, что в реальных молекулах РНК действуют перекрестные правила комплементарности. Однако, анализ случайных последовательностей показал, что правила комплементарности незначительно влияют на свойства РНК-подобных структур. Тогда как, комплементарное связывания типа А–А позволяет исследовать цепочки не только с четным алфавитом как в случае перекрестного связывания, но и с нечетным. Соответственно, для каждого значения $c = 3, 4, \dots, 7$ были построены зависимости удельной энергии $f = \langle F \rangle / L$ от длины случайной первичной структуры. Как видно (Рис. 10(а)), удельная энергия при $L \rightarrow \infty$ действительно стремится к некоторому усредненному значению f_∞ , которое является только функцией от c (Рис. 10(б)). Результаты численного моделирования существенно расходятся с аналитической оценкой (Рис. 10(б)). Зависимость, полученная в численном моделировании имеет критическую точку топологического перехода $c_c = 2$, которая является очевидной оценкой снизу. Действительно, рассмотрим произвольную двухбуквенную последовательность, например, *АВААВВВВААВВБАВАААВ* и будем последовательно находить комплементарные пары (в предположении А–А связывания) следующим образом. Ближайшие соседи по цепи одного сорта образуют комплементарную пару, и далее, вычеркиваются из последовательности. Легко видеть, что такая процедура приводит к формированию РНК-подобной структуры. Рассматриваемая цепочка после первой итерации будет выглядеть: *АВВААВАВ*, последующее вычеркивание приведет к *АВАВ*. Понятно, что данная процедура для любой случайной двухбуквенной последовательности приведет к тому, что, либо в остатке будет *АВАВ*, либо последовательность будет полностью вычеркнута. Второй вариант означает, что все мономеры участвуют в формировании вторичной структуры, удельная энергия которой $f = 1$. В случае остатка *АВАВ*, данный участок цепочки образует конфигурацию с двумя пропусками, но, в термодинамическом пределе, $f_\infty = 1$. Если структура образуется согласно перекрестным правилам комплементарности, то остатком будет $|L_1 - L_2|$ букв одного сорта, где L_1 и L_2 — количество мономеров А и В соответственно. В случайной последовательности $|L_1 - L_2| \sim 1/\sqrt{L}$. Таким образом, для случайных последовательностей с алфавитом $c = 2$ можно записать:

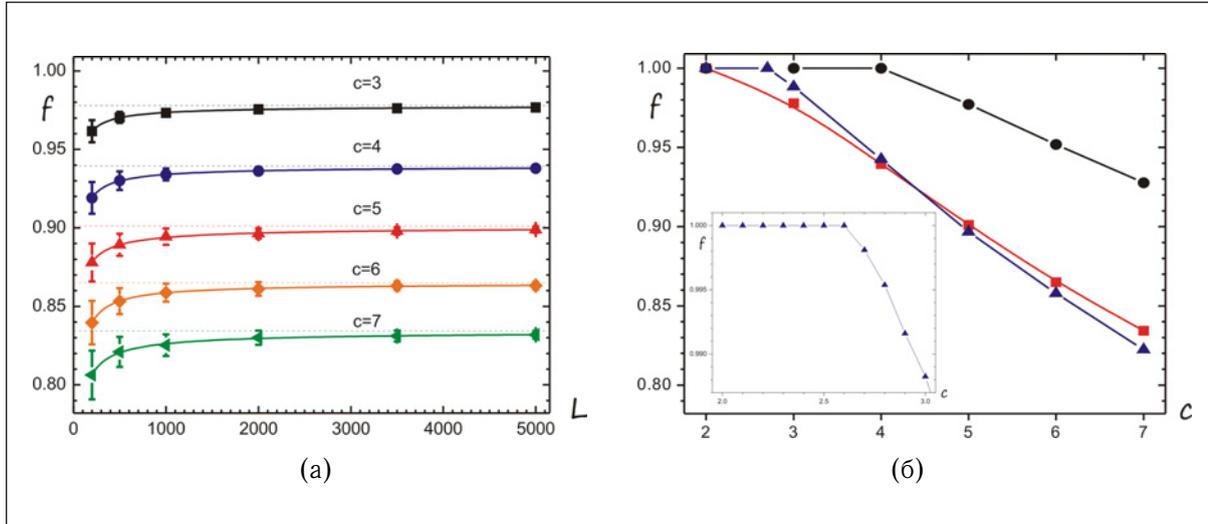


Рис. 10. (а) Зависимость удельной энергии f от длины случайной последовательности с заданным алфавитом c ; (б) зависимость предельного значения энергии f_∞ от алфавита для последовательностей с дискретным алфавитом (красным), в модели Бернулли (синим), и верхняя оценка энергии (черным) в модели независимого связывания. Дополнительный график: зависимость предельного значения энергии от алфавита в модели Бернулли демонстрирует, что критический алфавит является нецелым и принадлежит интервалу $2 < c_c < 3$.

$$f(L) = \begin{cases} 1 - \frac{const}{L}; & \text{для А-А связывания} \\ 1 - \frac{const}{\sqrt{L}}; & \text{для А-В связывания.} \end{cases} \quad (38)$$

Таким образом, аналитическое рассмотрение буквенных последовательностей позволило говорить о том, что критическое значение алфавита $2 \leq c_c \leq 4$. Более строгое рассмотрение [13] показало, что критическое значение алфавита лежит в интервале: $2 < c_c < 3$.

Подводя итог, подчеркнем еще раз, что при изменении алфавита, используемого в первичной структуре случайной последовательности, существует переход от полностью связанной РНК-подобной структуры до структуры с конечной долей несвязанных мономеров. Такой переход в работе называется топологическим. Критическая точка топологического перехода принадлежит интервалу ($2 < c_c < 3$), т.е. эффективно является *нецелым*.

Как можно трактовать нецелый алфавит в случайных последовательностях типа РНК? Далее, попробуем ответить на этот вопрос.

В. Топологический переход в модели Бернулли

Модель случайной последовательности с эффективно нецелым алфавитом может быть построена следующим образом. Будем считать, что матрица контактов η' в урав-

нении (20) является случайной: вероятность того, что $\eta'_{i,j} = 1$, равна p , а вероятность $\eta'_{i,j} = 0$ равна соответственно $1-p$. То есть теперь случайная последовательность характеризуется не первичной структурой — последовательностью мономеров из c различных типов, как это было раньше, а некой матрицей контактов, (i, j) -элемент которой разрешает или запрещает образование комплементарной пары между i и j мономером цепи. Мономеры цепи в данной модели не различаются по сортам и, в целом, любой мономер может образовать связь с любым другим в цепи, однако, в среднем, вероятность такого события равна p . Каждой последовательности в рассматриваемой модели можно сопоставить граф Эрдёша–Реньи, изображающего все возможные контакты между L мономерами. Основное отличие данной модели от дискретных буквенных последовательностей — нарушение свойства транзитивности. Если 1-й мономер может образовать связь со 2-м, а 2-й с 3-м, отсюда, вообще говоря, не следует (как это было для последовательностей с дискретным алфавитом), что 1-й мономер может связаться с 3-м. Однако, как, например, уже упоминалось, подобная модель бернуллиевского сравнения в задачах выравнивания случайных последовательностей является хорошей аппроксимацией. Вероятности p случайной матрицы контактов соответствует алфавит, равный:

$$c_{eff} = \frac{1}{p}. \quad (39)$$

Таким образом, оказывается возможным генерировать случайную последовательность с *любым* нецелым значением алфавита c . На Рис. 10(б) приведена зависимость удельной энергии f_∞ в термодинамическом пределе от алфавита c (39), полученная в численном моделировании. Во-первых, отметим, что значения f_∞ для бернуллиевского алфавита не более, чем на 1% отличается от соответствующих величин для случайных последовательностей с дискретным алфавитом, что оправдывает применимость данной модели. Случайный бернуллиевский полимер характеризуется критической вероятностью p_c . Для $p > p_c$, в термодинамическом пределе, $f_\infty = 1$ (так называемая «полочка» на зависимости удельной энергии (см. дополнительный график на Рис. 10(б)), что соответствует полностью связанной вторичной структуре, тогда как для $p < p_c$, даже в пределе бесконечной длины, основное состояние характеризуется $O(L)$ количеством несвязанных мономеров. Критическое значение вероятности согласно (39) соответствует критическому значению алфавита $c_c = 2.6$. Таким образом, модель Бернулли позволяет численно получить точку перехода.

Для более точной оценки критической точки топологического перехода были проведены следующие численные эксперименты. Рассмотрим ансамбль, состоящий из

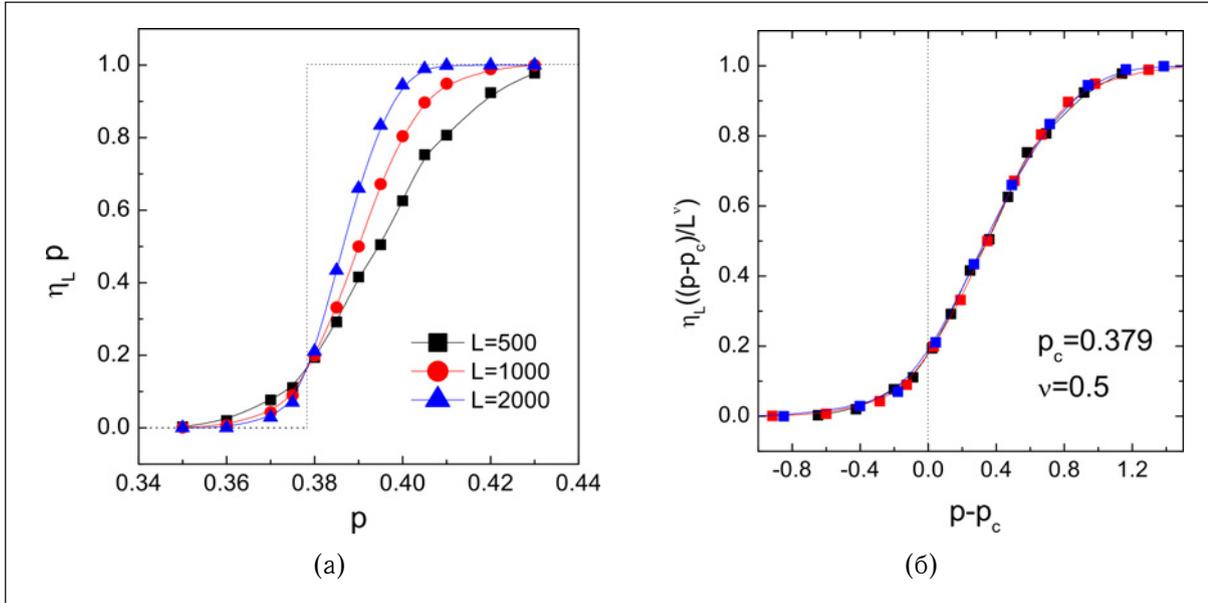


Рис. 11. Зависимость доли полностью связанных РНК-подобных структур в ансамбле случайных первичных структур различной длины (а) от параметра p модели Бернулли; скейлинг-анализ полученных зависимостей (б). Для каждого значения p и L было выполнено 10^5 накоплений.

N ($N = 10^5$) случайных бернуллиевских полимеров длиной L и подсчитаем количество последовательностей с полностью связанной вторичной структурой N_c . Доля полностью связанных структур в таком ансамбле $\eta_L = N_c/N$ есть функция p (см. Рис. 11). Естественно ожидать, что в пределе $L \rightarrow \infty$ (Рис. 11(а)), функция $\eta_L(p)$ вырождается в ступенчатую функцию. Скейлинг-анализ полученных зависимостей $\eta_L(p)$ обеспечивает критическое значение $p_c = 0.37$, что соответствует алфавиту:

$$c_c \approx 2.67.$$

Можно провести аналогию между данным топологическим переходом и переходом, наблюдаемым в теории перколяции [14]. В перколяционной теории задача формулируется следующим образом (одна из возможных формулировок). Рассмотрим протекание жидкости через пористую среду, причем пористую среду будем моделировать дискретной решеткой (сетью) — набором сайтов, между которыми есть связи — каналы. Жидкость протекает по этим каналам, которые могут быть открыты или закрыты с вероятностью p и $1-p$ соответственно. Существует пороговое значение вероятности p_{th} выше которой, протекание через данную среду возможно, т.е. существует связанный кластер на решетке, а ниже которой, построить связанный кластер невозможно. Переход между этими двумя состояниями в теории перколяции называют геометрическим фазовым переходом и относят к переходам второго рода [14].

Таким образом, можно предполагать, что топологический переход между полностью связанной РНК-подобной структурой и структурой с пропусками является фазовым переходом второго рода. В пользу этого предположения также свидетельствует непрерывное изменение вырожденности основного состояния (числа полностью связанных РНК-подобных структур) от вероятности p .

Был также выполнен анализ областей алфавита, лежащего выше и ниже критической точки топологического перехода в модели Бернулли. Во-первых, области характеризуются различной зависимостью от длины случайной последовательности: L :

$$\begin{cases} f(L) \sim 1 - C_1 e^{-L/\ell} & \text{для } p > p_c \\ f(L) \sim f_\infty - C_2 L^{-\alpha} & \text{для } p < p_c, \end{cases} \quad (40)$$

где C_1 и C_2 — некоторые константы. Для допереходной фазы ($p > p_c$) характерно экспоненциальное приближение к предельному значению удельной энергии ($f_\infty = 1$), тогда как в области больших алфавитов ($p < p_c$) энергия приближается к своему предельному значению степенным образом (Рис. 12). Показатель степени α в (40) находится в пределах $[0.75, 1]$ (сравните с (Рис. 10(а))). В допереходной области случайная последовательность из алфавита p может быть охарактеризована некоторой релаксационной длиной ℓ , указывающей на характерный масштаб длин, на котором энергия основного состояния сходится к своему предельному значению $f_\infty = 1$. Ясно, что зависимость релаксационной длины ℓ от вероятности p имеет вертикальную асимптоту в точке $p = p_c$. Естественно ожидать, что асимптотическое поведение $f(L)$ зависит от выбранной модели случайного полимера, в частности от правил комплементарности — см. (38).

Области отличаются также зависимостями флуктуаций свободной энергии от длины случайной последовательности L . Допереходная область характеризуется быстрым (экспоненциальным) падением флуктуаций с ростом L . Тогда как для $p < p_c$, характерен степенной рост флуктуаций с увеличением длины последовательности L (см. Рис. 5(б)).

С. Аналитическая оценка критической точки топологического перехода в модели Бернулли

1. Метод среднего поля

Для простоты переформулируем задачу в терминах планарных диаграмм. Рассмотрим граф, вершины которого (мономеры вдоль цепочки) перенумерованы, а матрица

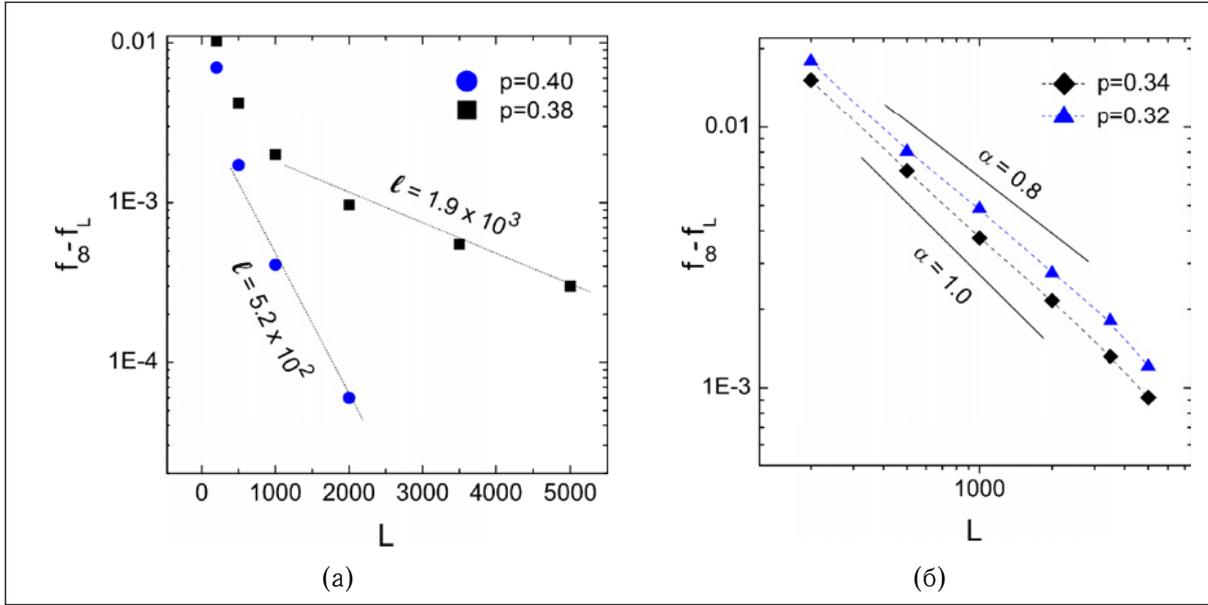


Рис. 12. Асимптотическое поведение удельной энергии $f(L)$ до (а) и после (б) топологического перехода. Зависимость $(f_\infty - f(L))$ в логарифмическом масштабе (а) и двойном логарифмическом масштабе (б) (см. (40)).

контактов V — матрица инцидентности графа. Задача о полностью связанной РНК-подобной структуре на данном графе сводится к вопросу о том, как выбрать среди разрешенных контактов $L/2$ связей, которые обеспечивают планарную структуру на заданном случайном графе, т.е. все вершины входят в конфигурацию ровно один раз и любые пары связей (i_1, j_1) и (i_2, j_2) удовлетворяют соотношению [15]:

$$(j_1 - i_1)(j_2 - i_1)(j_1 - i_2)(j_2 - i_2) > 0. \quad (41)$$

Другими словами, как разместить $L/2$ непересекающихся арок, принимая во внимание ограничения, накладываемые матрицей V . В модели Бернулли каждый элемент V_{ij} равен 1 либо 0 с соответствующими вероятностями p и $1 - p$, кроме того, матрица контактов — симметричная с нулевыми диагональными элементами:

$$P(V_{ij}) = ([p\delta(V_{ij} - 1) + (1 - p)\delta(V_{ij})])\theta(i - j) + \delta(V_{ji} - V_{ij})\theta(j - i) (\delta_{ij} - 1). \quad (42)$$

Здесь $\delta(x)$ и $\theta(x)$ — дельта-функция Дирака и функция Хевисайда, соответственно. Для $p = 1$ (когда все элементы V_{ij} равны 1), количество всех возможных арочных структур, удовлетворяющих (41) определяется числами Каталана (см. (33))

$$\# = C_{L/2} = \frac{L!}{(\frac{L}{2})!(\frac{L}{2} - 1)!}. \quad (43)$$

Когда $p \neq 1$, некоторые из конфигураций $\#$ запрещены матрицей контактов V . Введем обозначение p_1 — вероятность, того, что одна выбранная из $\#$ конфигурация разрешена.

Очевидно, что

$$p_1 = p^{L/2}. \quad (44)$$

Аналогично, определим p_k как вероятность, что k диаграмм из $\#$ разрешены, для $k = 2$, например

$$p_2 = p^{L/2} p^{L/2} p^{-n_{1\cap 2}} = p^L p^{-\kappa_2 L}, \quad (45)$$

где $n_{1\cap 2} \equiv \kappa_2 L$ равно количеству общих арок для двух случайно выбранных планарных диаграмм, усредненному по ансамблю $\#$. Для p_3 можно записать:

$$p_3 = (p^{L/2})^3 p^{-n_{1\cap 2\cap 3}} = p^{3L/2} p^{-C_3^2 \kappa_2 L} p^{\kappa_3 L}. \quad (46)$$

Величины κ_k могут быть вычислены с любой точностью. К примеру, κ_2 лежит строго в интервале $[1/15, 1/14.8]$. Вероятность иметь по крайней мере одну планарную конфигурацию для данной заполненности p матрицы V (42) определяется как:

$$\mathcal{P} = \# p_1 - \frac{\#(\# - 1)}{2} p_2 + C_{\#}^3 p_3 + \dots \quad (47)$$

Предполагая, что все диаграммы в ансамбле $\#$ независимы, т.е., $p_k = p_1^k$, для \mathcal{P} из (47) можно записать:

$$\mathcal{P} = 1 - (1 - p_1)^{\#} = 1 - \exp(-p_1 \#). \quad (48)$$

В пределе больших L , величина \mathcal{P} равна либо нулю, либо единице, в зависимости от соотношения между $\#$ и p_1 . Используя (44), для критического значения вероятности можно записать уравнение:

$$\lim_{L \rightarrow \infty} p_c [\#]^{2/L} = 1. \quad (49)$$

Условие (49) можно интерпретировать как то, что переход наблюдается в точке, при которой плотность единиц в матрице контактов V такая, что *в среднем* разрешена только одна планарная конфигурация. Вспоминая, асимптотику чисел Каталана (45), для критического значения вероятности получим $p_c = 1/4$, что совпадает с верхней оценкой $c_c = 4$ из (IV B).

2. Комбинаторная оценка

Предположение о независимости планарных конфигураций соответствует так называемому приближению среднего поля. Естественным следующим шагом является введение ненулевых корреляций между конфигурациями: $\kappa_k \neq 0$. Чтобы учесть корреляции между различными планарными диаграммами, поступим следующим образом. Перепишем (49) как:

$$\lim_{L \rightarrow \infty} \xi(p_c) [\#]^{2/L} = 1, \quad (50)$$

где $\xi(p)$ — некоторая функция, учитывающая корреляции между планарными диаграммами. Основная идея дальнейшего рассмотрения следующая: арки разной длины встречаются в оптимальной планарной конфигурации с различной вероятностью. Рассмотрим полностью связанную планарную конфигурацию, состоящую из $\mathcal{N} = \frac{L}{2}$ арок, соединяющих L точек. Возвращаясь к представлению планарных диаграмм через пути Дика (см. Рис. 9), можно увидеть, что арка между i -ой и j -ой точками возможна, только если i -й и j -й шаг имеют одну и ту же пространственную координату y . Тогда можно определить вероятность арки между i -ой и j -ой точками как:

$$P(i, j) = \frac{1 \times C_{(j-i-1)/2} \times 1}{2^{j-i+1}}. \quad (51)$$

В знаменателе правой части (51) стоит суммарное число возможных шагов вверх/вниз на длине $(j - i + 1)$, в числителе — “1” соответствуют выбору шага вверх и вниз на позициях i и j соответственно; число Каталана $C_{(j-i-1)/2}$ описывает все возможные конфигурации петли между парой (i, j) (так как i -й и j -й шаги находятся на одной высоте, петля между ними должна быть тоже путем Дика). Вероятности $P(i, j)$ зависят только от длины арки $(j - i)$ и не равны нулю только для арок нечетной длины, т.е., $P(i, i+1) = \frac{1}{4}$, $P(i, i+3) = \frac{1}{16}$, $P(i, i+5) = \frac{2}{32}$, т.д.. Если просуммировать $P(i, j)$ по всем возможным арочным длинам, то результатом будет $\sum_{k=1}^{L-1} P(i, i+k) = \frac{1}{2}$ — вероятность того, что в i -ой позиции находится левая граница арки (шаг вверх).

Отметим, что доля коротких арок чрезвычайно высока. Действительно, вероятность, в типичной арочной конфигурации иметь арку длиной $\ell = 1$ равна $\frac{1}{4}$, арку длиной $\ell = 3$, уже $\frac{1}{16}$, и т.д.. С другой стороны, количество всех возможных кратчайших арок — $(L - 1)$. Поэтому, в типичной конфигурации $\frac{1}{4}$ среди них должны быть «разрешены». Естественно, что веса таких коротких арок в бернулливской модели (элементы $V_{i, i+1}$ матрицы контактов) выше, чем длинных арок.

Принимая во внимание эту выделенность коротких арок, оценим функцию $\xi(p)$ в (50). Вместо независимого выбора набора арок, теперь предположим, что построение типичной арочной конфигурации происходит следующим образом:

1. выбор $\frac{L}{4}$ непересекающихся коротких арок ($\ell = 1$) из $(L - 1)$ возможных
2. выбор остальных $\mathcal{N} - \frac{L}{4} = \frac{L}{4}$ из длинных ($\ell > 2$) арок

Так как общее число длинных арок порядка $L^2 \gg \frac{L}{4}$, будем считать, что длинные арки

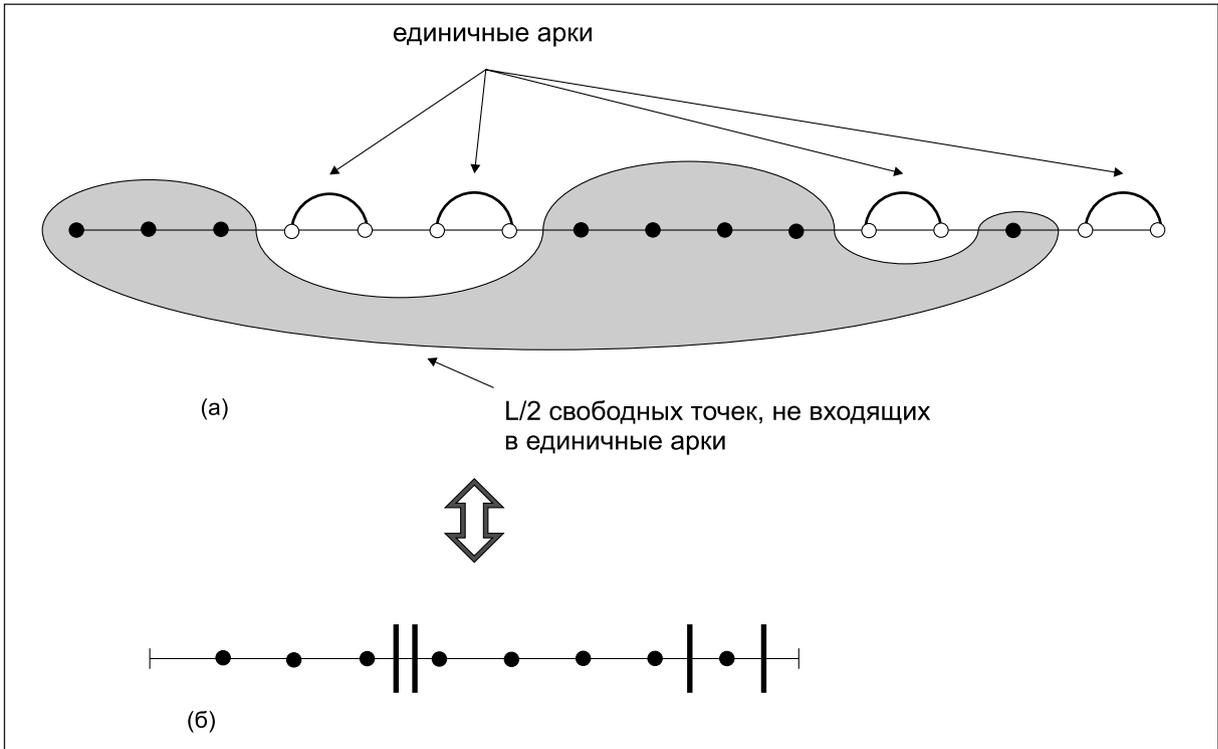


Рис. 13 Пояснение к вычислению $\mathcal{Z}(p)$: (а) Выбор $\frac{L}{4}$ единичных арок на L вершинах случайного графа ($\frac{L}{2}$ свободных вершин) аналогичен комбинаторной задаче о расположении $\frac{L}{2}$ точек по $\frac{L}{4} - 1$ ящикам (б).

выбираются независимо друг от друга с вероятностью p . И, таким образом, вклад от длинных арок в функцию $\xi(p)$ равен $p^{L/4}$.

Иная ситуация при выборе кратчайших арок длиной “1”. Для бернуллиевского полимера с матрицей контактов V только pL единичных арок разрешены. Таким образом, выбор коротких арок для оптимальной конфигурации без пропусков оказывается сильно ограниченным. Вероятность выбрать $L/4$ непересекающихся арок из pL разрешенных можно оценить следующим образом. Определим сначала число способов \mathcal{Z} выбора $\frac{L}{4}$ непересекающихся единичных арок из всех $(L-1)$ возможных (Рис. 13). Единичные арки можно рассматривать как стенки ящиков, тогда задачу можно переформулировать следующим образом. Будем интересоваться количеством способов, которыми можно заполнить $(\frac{L}{4} - 1)$ ящика $L/2$ свободными точками (шарами). Результат известен из комбинаторики и $\mathcal{Z} = C_{L/4}^{3L/4-1}$, где C_m^n — число сочетаний m по n .

Можно считать, что среди них $p(3\frac{L}{4} - 1)$ арок разрешены первичной структурой полимера (матрицей контактов V) и величина $C_{L/4}^{p(3L/4-1)}$ описывает вес коротких арок в полностью связанной РНК-подобной структуре случайного полимера. Учет корреляций

между планарными конфигурациями на уровне единичных дуг приводит к следующими выражению для $\xi(p)$ (50):

$$\xi(p)^{L/2} = p^{L/4} C_{L/4}^{p(3L/4-1)} \left[C_{L/4}^{3L/4-1} \right]^{-1}. \quad (52)$$

В пределе $L \rightarrow \infty$, после упрощений, получим:

$$\ln \xi(p) = \frac{1}{2} \ln p + \frac{3p}{2} \ln \frac{3p}{2} - \frac{3p-1}{2} \ln \frac{3p-1}{2} - \frac{3}{2} \ln \frac{3}{2}. \quad (53)$$

Подставляя этот результат в (50):

$$\begin{aligned} \ln \xi(p_c) &= -\ln 4; \\ p_c &\approx 0.35 \quad (c_c = 2.87). \end{aligned} \quad (54)$$

3. Матричный подход

Еще один подход оценки критического алфавита основан на матричном описании вторичной структуры РНК. Статистическую сумму $Z_L(N, V)$ случайного полимера можно представить через случайные эрмитовы матрицы ϕ как [16]:

$$Z_L(N, V) = \frac{\int d\phi_1 \dots d\phi_L e^{-S_0} \frac{1}{N} \text{tr}(\phi_1 \dots \phi_L)}{\int d\phi_1 \dots d\phi_L e^{-S_0}} \equiv \langle \phi_1 \dots \phi_L \rangle_{S_0}, \quad (55)$$

где

$$S_0 \equiv S_0\{V, \phi_1, \dots, \phi_L\} = \frac{N}{2} \sum_{i,j} (V^{-1})_{ij} \text{tr}(\phi_i \phi_j). \quad (56)$$

В отсутствие замороженного беспорядка, т.е., если $V_{ij} \equiv 1$, задача (55) может быть решена точно. В частности, множитель $a_{L/2,0}$ перед $v^{L/2}$, описывающий планарные конфигурации с $L/2$ арками, т.е. полностью связанные структуры, вносит наибольший вклад в общую статистическую сумму полимера и определяется числами Каталана:

$$\lim_{N \rightarrow \infty} Z_L(N; V) = C_{L/2} \sim \frac{4^{L/2}}{(L/2)^{3/2} \sqrt{\pi}}. \quad (57)$$

Как и ранее будем вычислять функцию $\xi(p)$ в (50) усредняя статистическую сумму $Z_L(N, V)$ по распределению (42). Для этого выполним стандартное преобразование Хаббарда-Стратоновича и будем интегрировать по V с весом (42):

$$\begin{aligned} \int dV P(V) Z_L(N, V) = \\ \text{const} \int \prod_{k=1}^L d\phi_k \frac{1}{N} \text{tr}(\phi_1 \dots \phi_L) \int \prod_{m=1}^L dh_m e^{iN \sum_i \text{tr}(h_i \phi_i)} e^{\mathcal{S}}, \end{aligned} \quad (58)$$

где $\mathcal{S} = \mathcal{S}_0 + U$, и

$$\mathcal{S}_0 = -\frac{pN}{2} \sum_{ij} \text{tr}(h_i h_j), \quad (59)$$

$$U = \frac{p(1-p)N^2}{8} \sum_{ij} [\text{tr}(h_i h_j)]^2 - \frac{p(1-p)(1-2p)N^3}{48} \sum_{ij} [\text{tr}(h_i h_j)]^3 + \dots \quad (60)$$

Величина \mathcal{S}_0 соответствует единичной матрице контактов с дополнительным фактором p . Учет только этого слагаемого, после обратного преобразования Хаббарда-Стратоновича, приводит к $\xi(p) = p$, и оценке $p_c = \frac{1}{4}$, совпадающей с оценкой в предположении среднего поля. Действие U (60) сдвигает значение p_c в сторону меньших значений. Но, так как U содержит бесконечное число слагаемых (60), теория возмущений в данном случае неприменима. В этой связи было предложено следующее приближение: все поля $\{h_i\}_{i=1,\dots,L}$ в (60) эквивалентны, поэтому можно считать, что в среднем, $\langle N \text{tr}(h_i h_j) \rangle_{\mathcal{S}_0} \equiv T$ не зависит от (i, j) . В рамках данного средне-полевого приближения можно сделать замену $e^{\mathcal{S}} = e^{\mathcal{S}_0} e^{(U)}$, где:

$$\langle U \rangle = \frac{p(1-p)N}{8} T \sum_{ij} \text{tr}(h_i h_j) - \frac{p(1-p)(1-2p)N}{48} T^2 \sum_{ij} \text{tr}(h_i h_j) + \dots \quad (61)$$

Упрощение выражения (61) приводит к следующему уравнению на пропагатор T :

$$\frac{1}{T} = \frac{-2}{T} \log \left[1 - p + p \exp \left(-\frac{T}{2} \right) \right]. \quad (62)$$

Выражение (62) дает $T = -2 \log \left[1 - \frac{1-1/\sqrt{e}}{p} \right]$, и окончательно можно написать:

$$\mathcal{S} = -\frac{\xi(p)N}{2} \sum_{ij} \text{tr}(h_i h_j), \quad (63)$$

где

$$\xi(p) = \frac{1}{T} = \left(-2 \log \left[1 - \frac{1-1/\sqrt{e}}{p} \right] \right)^{-1}. \quad (64)$$

Подстановка (64) в (50) приводит к оценке критического алфавита $p_c^* = 0.4551$. Большая расходимость полученного результата с численным $p_c = 0.37$ означает, что предложенного приближения недостаточно для описания топологического перехода.

D. Переход случайной РНК в замороженное состояние, ограниченный топологическим переходом

Рассмотрим как данный топологический переход ограничивает фазовый переход в замороженное состояние. С пионерских работ Бундшу и Хва [17, 18] принято считать, что в этой системе имеет место фазовый переход в «замороженное» состояние при низких температурах. Основываясь на репличном анализе, Лассиг и Визе, [19] и Давид и Визе [20] сформулировали задачу о переходе в терминах теории поля. Ниже приводятся доводы Бундшу и Хва, доказывающие существование фазового перехода и обсуждаются характерные свойства разных фаз.

В зависимости от температуры, случайная РНК находится в одной из фаз: i) «растаявшая» высокотемпературная фаза (molten phase) или ii) «замороженная» низкотемпературная фаза (glass phase). В высокотемпературной фазе большую роль играет энтропия цепочки, нежели порядок мономеров в первичной структуре. Данная фаза хорошо описывается в модели гомополимера, комплементарное связывание не играет роли, и эффективно можно заменить все мономеры мономерами одного типа A . Низкотемпературная фаза, наоборот, определяется, в первую очередь, первичной структурой цепочки, то есть, основной вклад в свободную энергию обусловлен комплементарным связыванием мономеров. Такую фазу принято характеризовать замороженным беспорядком [17, 18]. Температура, при которой РНК переходит из одной фазы в другую, называется температурой фазового перехода и в литературе обозначается T_g .

Был предложен следующий подход к определению температуры фазового перехода. Рассмотрим пару мономеров, чье взаимодействие приводит к образованию петли наибольшего размера, т.е. нуклеотидов с номером 1 и $L/2$ по цепи для последовательности длиной L . Определим энергию выигрыша данного контакта, которая определяется как $\Delta F(L) = k_B T \ln P_{1,L/2}$, где $P_{1,L/2}$ — вероятность связывания 1 и $L/2$ мономера цепи. Данную энергию называют энергией пинча, и из выражения для статистической суммы цепочки следует, что:

$$\Delta F(L) = F_{1,L} - (F_{1,L/2} + F_{L/2+1,L}). \quad (65)$$

Отметим, что аналогичный вопрос исследуется и в теории перколяции, где тоже предполагается взаимосвязь перколяционного перехода и температурного фазового перехода, наблюдаемого, например, в модели Изинга [21].

Были проанализированы температурные зависимости свободной энергии пинча (2) случайной последовательности в модели Бернулли разной вероятности p . Как уже об-

суждалось, температура перехода в замороженное состояние T_g непосредственно связано со средним числом пропусков в структуре основного состояния. В [18] было показано, что температура перехода не превосходит T^*

$$T^* = \lambda^{-1}\sigma, \quad (66)$$

где σ — среднее число пропусков на пару мономеров, а λ определяется из зависимости наибольшего общего непрерывного сегмента ℓ двух половинок последовательности РНК: $\ell = \lambda^{-1} \ln L$ (см. Рис.1.4). Известно, что для цепочек РНК $\lambda = \ln 2$. Для случайного бернуллиевого процесса λ определяется как $\lambda = \ln(1/p)$ [6]. Таким образом, выражение (66) можно переписать в виде

$$T^* = \frac{\sigma}{\ln(1/p)}. \quad (67)$$

Доля несвязанных мономеров σ растет с ростом алфавита $1/p$ сильнее, чем логарифм (см. Рис. 10(б)) и из (67) непосредственно следует, что в допереходной области ($p > p_c$) фазовый переход в замороженное состояние наблюдаться не будет. Температура перехода T_g эффективно равна нулю, т.е., случайный полимер во всем температурном диапазоне находится в расплавленной фазе. Данное предположение дополнительно подтверждается наблюдением того, что для случайных последовательностей с алфавитом $c = 2$ переход имеет место только при накладывании ограничений на структуру, а именно, введением минимального размера петли [22].

Результаты численного моделирования представлены на Рис. 14. Был проанализирован температурный коэффициент $a(T)$ (4) для последовательностей с разной вероятностью p . Температура перехода определяется точкой, в которой нарушается линейная зависимость $a(T) = \frac{3}{2}T$, характерная для расплавленной фазы. Из полученных данных видно, что температура перехода уменьшается с ростом вероятности p и в допереходной области становится равной нулю ($p = 0.5$ на Рис. 14). Вблизи критического значения p_c численный эксперимент усложняется тем, что корректный анализ требует рассмотрения достаточно длинных случайных цепочек (с длиной, превышающей соответствующую релаксационную длину $\ell(p)$, см. (40)), что приводит к существенному увеличению времени численного моделирования. Также стоит отметить, что в связи с наблюдаемой степенной зависимостью свободной энергии основного состояния от длины последовательности ((40)), аппроксимация уравнением (2) вблизи точки $T = 0$, вообще говоря, неверна.

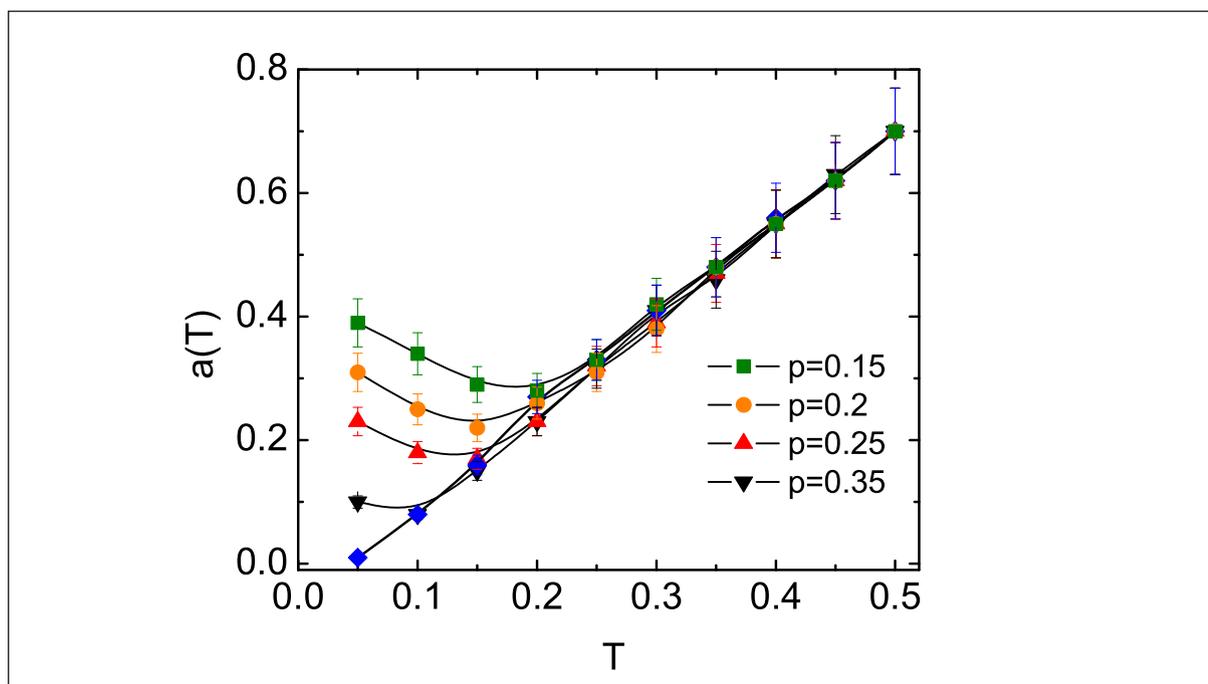


Рис. 14. Зависимость коэффициента $a(T)$ для случайной последовательности разной вероятности p в модели Бернулли.

Предполагается, что критическая точка топологического перехода между полностью связанной РНК-подобной структурой и структурой с пропусками является пороговым значением для термодинамического перехода. В области последовательностей $p > p_c$ возможна только расплавленная фаза вне зависимости от температуры. Рис. 15 показывает фазовую диаграмму на (T, p) плоскости. Это предположение подтверждается исследованием энергии пинча от длины случайной последовательности в точке $T = 0$. Точка пересечения зависимостей для разных длин (см. дополнительный график на Рис. 15) разделяет два топологических режима и близка к наблюдаемому критическому алфавиту.

Е. Другие модели нецелого алфавита

Основной недостаток бернуллиевской модели полимера заключается в отсутствии ясного соответствия матрицы контактов V для произвольного p и первичной структуры полимера. Как уже указывалось, в модели Бернулли нет деления на сорта мономеров, все мономеры, рассматриваются однотипными. В этом разделе, речь пойдет о некоторых подходах генерации полимера с нецелым алфавитом и разными сортами мономеров.

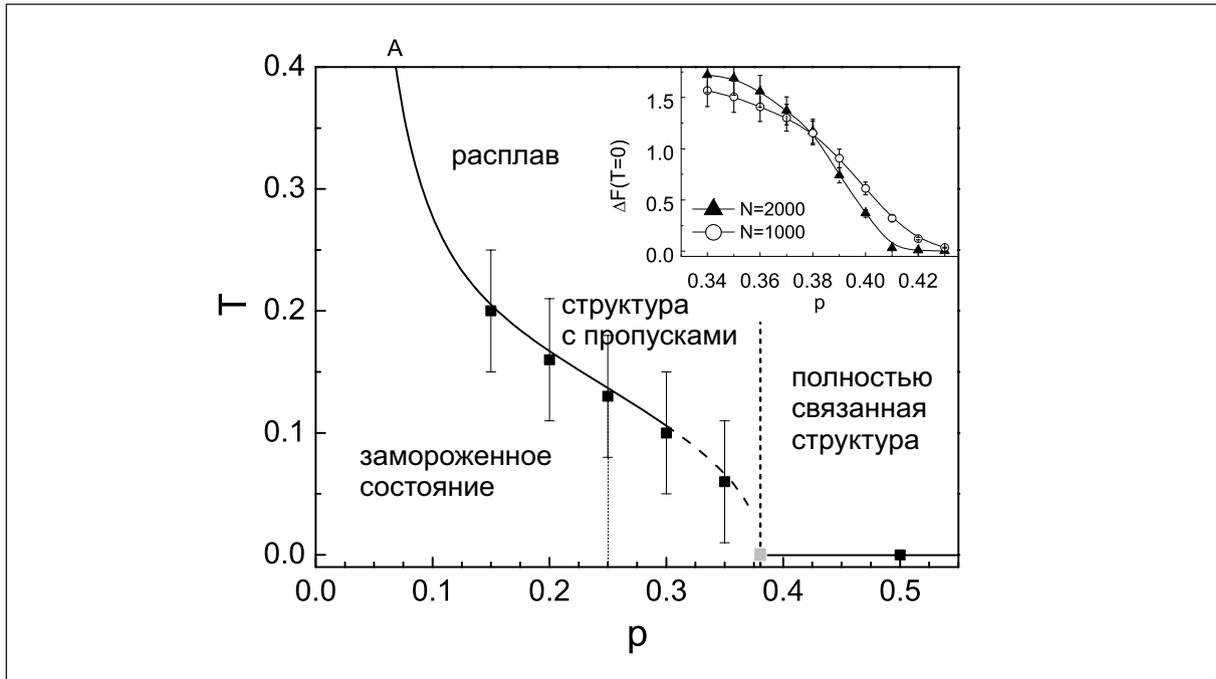


Рис. 15. Фазовый переход в замороженное состояние, ограниченный топологическим переходом в модели Бернулли. Дополнительный график: зависимость энергии пинча в пределе $T \rightarrow 0$ от вероятности p .

1. Метод концентраций

Одна из самых простых моделей нецелого алфавита, сохраняющего свойство транзитивности — модель концентраций. В такой модели предполагается, что случайный полимер состоит из трех типов мономеров, А, В и С, но мономеры распределены в цепочке не случайно, а коррелировано. В модели так называемых «локальных концентраций» предполагается, что концентрация, количество мономеров третьего типа $[C]$ не равна концентрации мономеров $[A] = [B]$, а зависит от алфавита p . В частности, концентрацию $[C]$ можно определить по заполненности матрицы контактов V . Изменение концентрации $[C]$ от 0 до $\frac{1}{3}$ описывает последовательности с нецелым алфавитом от $c = 2$ до $c = 3$. Однако, для алфавитов, немного превышающих $c = 2$ ($c = 2 + \epsilon$), данная модель приводит к случайной двухбуквенной последовательности, слабо разбавленной третьим типом мономеров C . Из-за малого количества $[C]$, эти мономеры появляются редко в цепочке, и из-за специфического комплементарного взаимодействия $C-C$, это приводит к сильным ограничениям на конфигурации основного состояния. Как уже упоминалось, основное состояние характеризуется большим количеством коротких арок, т.е., взаимодействием ближайших соседей в цепочке. Таким образом, важнее оказывается не количество различных типов мономеров в первичной структуре, а их

распределение. Модель концентраций может быть улучшена, если распределять $[C]$ мономеров третьего типа не случайно, а согласно некому распределению, характерному для случайных трехбуквенных (двухбуквенных) последовательностей. Грубо говоря, это приведет к тому, что мономеры третьего типа C будут появляться в первичной структуре блоками. Но даже такая модель обладает существенным недостатком — выделенной ролью мономеров типа C , по сравнению с мономерами типа A и B .

Г. Коррелированная случайная последовательность

Модель, которая устраняет этот недостаток — модель так называемой коррелированной последовательности, в которой распределение трех типов мономеров является не случайным, а сильно коррелированным. Различный алфавит p определяется в таких последовательностях не количеством (концентрацией) мономеров, а тем, насколько сильно скоррелировано появление мономеров различного типа в цепочке. Распределение мономеров в первичной структуре определяется согласно Марковскому процессу [12]:

	A	B	C
A	$1 - 2\epsilon$	ϵ	ϵ
B	ϵ	$1 - 2\epsilon$	ϵ
C	ϵ	ϵ	$1 - 2\epsilon$

т.е., вероятность встретить, например, мономер типа A за мономером A не равна вероятности появления A после B (или C). Изменение ϵ в диапазоне $[0, \frac{1}{3}]$ обеспечивает диапазон алфавитов $[1, 3]$. Взаимосвязь между параметром модели ϵ и алфавитом c можно установить, используя определение информационной энтропии по Шеннону [23]:

$$c = \left(\frac{1}{\epsilon} - 2\right)^{2\epsilon} \frac{1}{1 - 2\epsilon}. \quad (68)$$

Результатом данной модели является полимер с блочной первичной структурой, причем размер блока зависит от параметра модели ϵ . На Рис. 16 представлены результаты численного моделирования в модели такой коррелированной последовательности. Скачкообразного изменения удельной энергии от алфавита c в численном моделировании не наблюдается. Отметим, что даже для алфавитов, эффективное значение которых меньше 2, идеальные полностью связанные структуры не образуются. Объяснить такую зависимость можно следующим образом. После процедуры вычеркивания коррелированный полимер с трехбуквенным алфавитом сводится к последовательности со

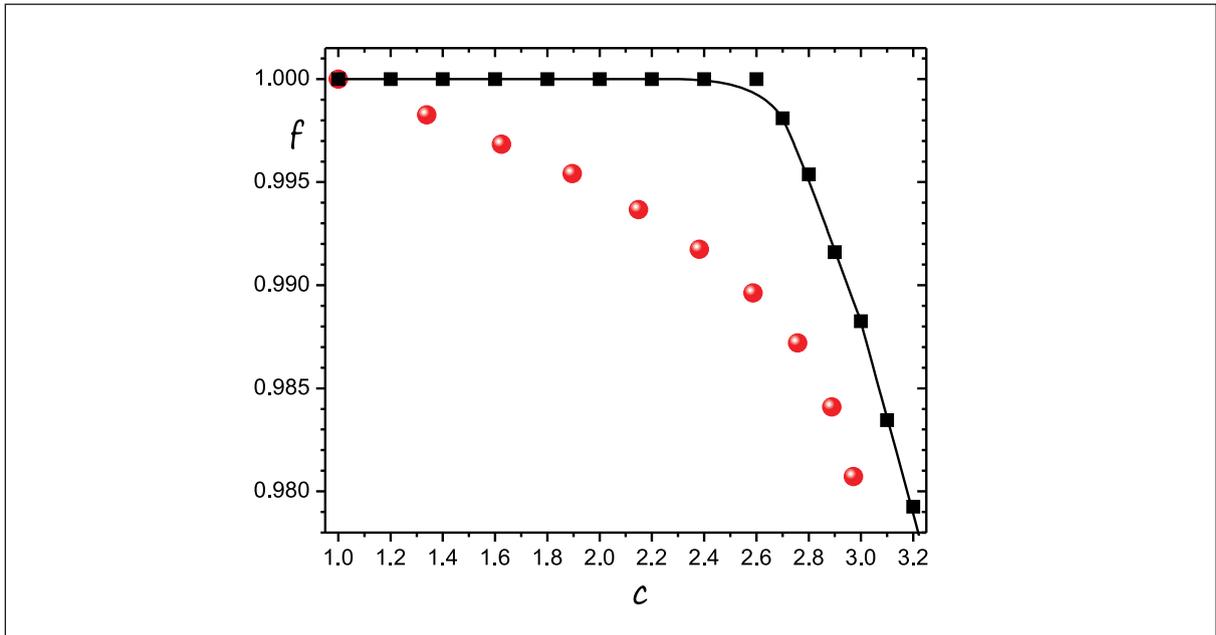


Рис. 16. Модель коррелированной последовательности: зависимость удельной энергии f_∞ от алфавита (красным); для сравнения приведена зависимость предельной энергии в модели Бернулли (черным).

случайно распределенными мономерами трех сортов, для которой, как было показано (Рис. 10) свойственно образование $O(L)$ пропусков. Длина остаточного полимера (после вычеркивания) зависит от параметра ϵ , и для достаточно длинных последовательностей пропорциональна L . Таким образом, любая модель с тремя сортами мономеров (коррелированная последовательность или модель концентраций) всегда сводится к модели случайной последовательности с алфавитом $c = 3$, который лежит в послепереходной области. Тем не менее, стоит отметить, что для зависимости удельной свободной энергии (см. Рис. 16) в модели коррелированной последовательности характерен резкий спад удельной энергии основного состояния при $c > c_c$.

1. Рациональный алфавит

Другая модель, частично сохраняющая свойство транзитивности, — модель рационального алфавита — заключается в следующем. Последовательность с алфавитом $\frac{P}{Q}$ можно представить, как полимер, состоящий из P сортов мономеров, правила комплементарности для которых разрешают Q связей для каждого мономера. Например, алфавит $c = \frac{5}{2}$ означает пятибуквенный алфавит в первичной структуре и с правилами комплементарности, организованными, к примеру, по пятиугольнику (Рис. 17). Такие правила можно построить для рационального алфавита любой величины. Численные

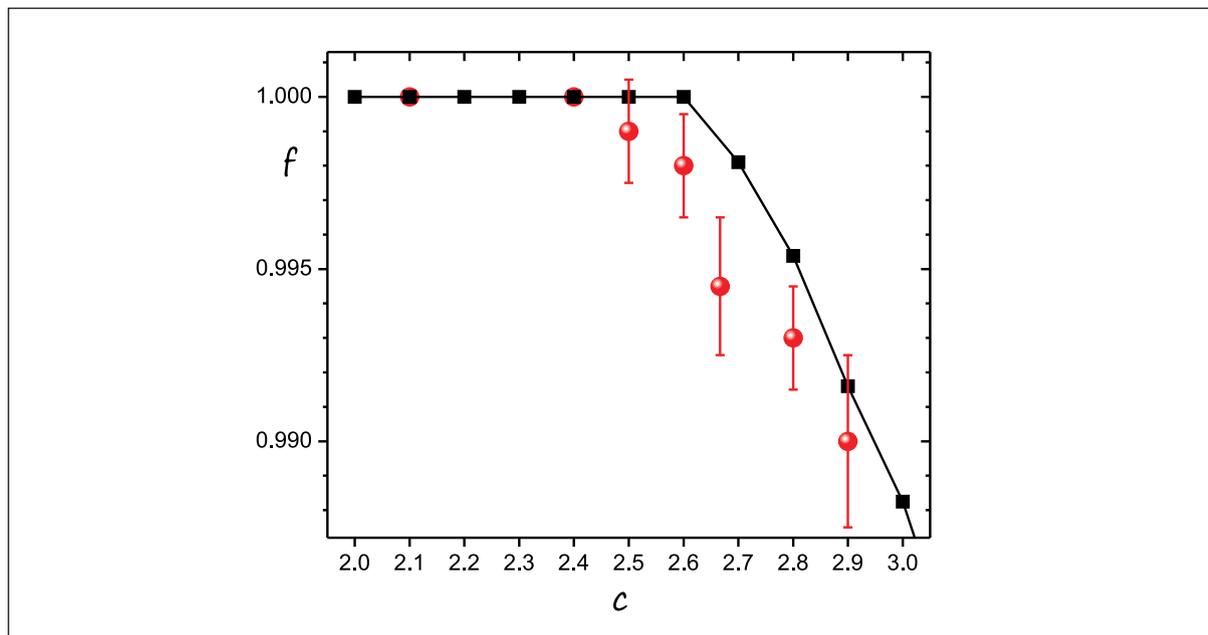


Рис. 17. Модель рационального алфавита: зависимость удельной энергии основного состояния f_∞ от алфавита (красным). Для сравнения приведена зависимость, полученная в модели Бернулли (черным).

результаты для данной модели приведены на Рис. 17 и показывают критическое изменение топологии вторичной структуры РНК.

Отметим, что модель чувствительна к выбору P и Q , так например, один и тот же алфавит $c = 2.8$, представленный как $\frac{14}{5}$ и $\frac{28}{10}$ дает разные результаты для удельной энергии основного состояния. В пределе $P \rightarrow L$ рассматриваемая модель сводится к модели Бернулли. Модель рационального алфавита, интуитивно, кажется ближе к алфавиту, используемого природой в молекулах РНК. Как указывалось, в молекулах РНК, помимо комплементарных пар, образуются неканонические пары, т.е. правила образования связей, во-первых, не транзитивны, а во-вторых, система правил похожа на систему связей в модели рационального алфавита (Рис. 17).

Каков алфавит в реальных молекулах РНК? Понятно, что учет неканонических пар эффективно приводит к уменьшению алфавита. С другой стороны, учет, к примеру, минимальной длины петли увеличивает алфавит в последовательностях РНК. Образование псевдоузлов и стэкинг взаимодействия приводит к сдвигу алфавита к меньшим значениям. Таким образом, фактический алфавит в молекулах РНК определяется многими факторами. Однако, логично предполагать, что алфавит в РНК находится вблизи критического. Почему выгодно реальным молекулам РНК иметь алфавит вблизи критического? Для того чтобы РНК выполняла свою биологическую функцию, она должна удовлетворять следующим критериям: i) ее фолдинг должен быть достаточно уникален

и ii) структура должна быть устойчива к тепловому шуму. Короткие алфавиты $c < c_c$ не обеспечивают первый критерий, так как для допереходной фазы характерна сильная вырожденность основного состояния. С другой стороны, длинным алфавитам $c \gg c_c$ свойственны структуры с длинными петлями — несвязанными мономерами, которые неустойчивы к тепловым флуктуациям. Таким образом, биологический алфавит, по-видимому, находится вблизи критического.

V. ЗАКЛЮЧЕНИЕ

В работе представлен анализ топологических свойств РНК-подобных молекул со случайной первичной структурой методами статистической физики и теории случайных процессов. Получено выражение для статистической суммы, описывающие взаимодействие двух сополимеров, учитывающий способность каждого из сополимеров образовывать РНК-подобную структуру с иерархией петлевых участков. Разработан соответствующий алгоритм динамического программирования вычисления свободной энергии основного состояния таких РНК-подобных молекул. Численно и аналитически показано критическое поведение РНК-подобной структуры в зависимости от используемого в первичной структуре алфавита. Существует две области: для алфавитов $c < c_c$ свойственна максимально связанная вторичная структура без пропусков, тогда как для $c > c_c$ вторичная структура содержит конечную долю несвязанных мономеров. Аналитическая оценка точки топологического перехода $c_c = 2.87$ близка к наблюдаемой в численном моделировании $c_c = 2.67$.

Литература

- [1] *Птицын Б.О., Финкельштейн А.* Физика белка: Курс лекций // Москва: Университет, 2002, 376 С.
- [2] *Гросберг Ю.А., Хохлов Р.А.* Статистическая физика макромолекул // Москва: Наука, 1989, 344 С.
- [3] *Workman C., Krogh A.* No evidence that mRNAs have lower folding free energies than random sequences with the same dinucleotide distribution // *Nucleic Acids Research*, 1999, **27** (24), P. 4816-4822.
- [4] *Clote P., Ferre F., Kranakis E., Krizanc D.* Structural RNA has lower folding energy than randomRNA of the same dinucleotide frequency // *RNA*, 2005, **11**(5), P. 578-591.

- [5] *Brezin E.E., Itzykson C., Parisi G., Zuber J.B.* Planar diagrams // *Communications in Mathematical Physics*, 1978, **59**, P. 5-51.
- [6] *Waterman M.S., Vingron M.* Sequence comparison significance and poisson approximation // *Statistical Science*, 1994, **9**, P. 367-381.
- [7] *Majumdar S.T., Nechaev S.K.* Exact asymptotic results for the bernoulli matching model of sequence alignment // *Physical Review E*, 2005, **72** (2), P. 020901.
- [8] *Kriecherbauer T., Krug J.* A pedestrian's view on interacting particle systems: KPZ universality and random matrices // *Journal of Physics A: Mathematical and Theoretical*, 2010, **43**(40), P. 403001.
- [9] *Kardar M., Parisi G., Zhang Y.C.* Dynamic scaling of growing interfaces // *Physical Review Letters*, 1986, **56** (9), P. 889-892.
- [10] *Тамм М.В., Лисаченко Н.Г., Ерухимович И.Я., Иванов В.А.* Эффекты конечного объема в системе равновесных циклических полимеров: теория и компьютерное моделирование // *Высокомолекулярные соединения*, 2005, **47** (7), С. 348-352.
- [11] *Ландо К.* Лекции о производящих функциях // Москва: Московский центр непрерывного математического образования, 2007, 144 С.
- [12] *Feller W.* An introduction to probability theory and its applications // New York: Wiley, 1968, 509 p.
- [13] *Владимиров А.А.* Паросочетания без пересечений // *Проблемы передачи информации*, 2013, **49**(1), С. 61-65.
- [14] *Grimmett G.* What is Percolation? // New York: Springer, 1999, 444 P.
- [15] *Tamm M., Nechaev S.* Necklace-cloverleaf transition in associating RNA-like diblock copolymers // *Physical Review E*, 2007, **75** (3), P. 031904.
- [16] *Zee A.* Random matrix theory and RNA folding // *Acta Physica Polonica B*, 2005, **36** (9), P. 2829-36.
- [17] *Bundschuh R., Hwa T.* RNA secondary structure formation: A solvable model of heteropolymer folding // *Physical Review Letters*, 1999, **83** (7), P. 1479-1482.
- [18] *Bundschuh R., Hwa T.* Statistical mechanics of secondary structures formed by random RNA sequences // *Physical Review E*, 2002, **65** (3), P. 031903.
- [19] *Lässig M., Wiese K.J.* Freezing of random RNAs // *Physical Review Letters*, 2006, **96** (22), P. 228101.
- [20] *David F., Wiese K.J.* Systematic field theory of the RNA glass transition // *Physical Review Letters*, 2007, **98** (12), P. 128102.

- [21] *Toninelli C., Biroli G., Fisher D.* Jamming percolation and glass transitions in lattice models // *Physical Review Letters*, 2006, **96** (3), P. 035702.
- [22] *Pagnani A., Parisi G., Ricci-Tersenghi F.* Glassy transition in a disordered model for the RNA secondary structure // *Physical Review Letters*, 2000, **84** (9), P. 2026-2030.
- [23] *Shannon C.E., Weaver W.* A mathematical theory of communication // *The Bell System Technical Journal*, 1948, **27**, P. 379-423.

TOPOLOGICAL PROPERTIES OF RNA-LIKE MOLECULES WITH A RANDOM PRIMARY STRUCTURE

O.Valba^{1, 2}, M. Tamm^{1, 3}, S. Nechaev^{4, 5}

¹*National Research University HSE*

²*Semenov Institute of Chemical Physics*

³*Moscow State University*

⁴*Universitet Paris-Sud / CNRS*

⁵*Lebedev Physical Institute*

ovalba@hse.ru

Received 20.04.2015

The paper focuses on the application of methods of statistical physics and stochastic processes for the study of the topological properties of RNA-like heteropolymers with a random primary structure. In particular, it describes the critical behavior of RNA-like secondary structure topology in dependence on the alphabet used in a random sequence. The analytical evaluation of the critical transition point based on the combinatorial and matrix description is given.

О МЕХАНИЗМЕ СЕЛЕКТИВНОГО БИОЛОГИЧЕСКОГО ДЕЙСТВИЯ МИКРОВОЛНОВОГО И ТЕРАГЕРЦЕВОГО ИЗЛУЧЕНИЯ

В.Б. Ошурко

Институт общей физики РАН им. А.М. Прохорова

vbo08@mail.ru

Поступила 24.03.2015

Кратко анализируются известные физические механизмы поглощения микроволнового и терагерцового излучения. Показано, что основным механизмом биологического действия может быть только простой равновесный нагрев среды. Однако, в случае неоднородных сред не исключен также эффект "пространственного селективности" (различный нагрев различных частей системы). Показано, что нелинейная (возрастающая) зависимость диэлектрических потерь в воде от температуры может обеспечить положительную обратную связь между температурой и коэффициентом поглощения излучения. Это положительная обратная связь может значительно усиливаться в областях, имеющих хоть немного больший коэффициент поглощения. Так, разница в 0,01% может приводить к различию температуры порядка десятка градусов. Такой эффект, в принципе, позволяет избирательно воздействовать отдельные клетки заданного сорта или внутриклеточные органеллы и, тем самым, обеспечивает возможность селективного контроля биологических систем.

УДК 539.938, 577.3

В настоящее время хорошо известно, что практически единственным возможным механизмом воздействия электромагнитного излучения таких частот (1 ГГц – 100 ТГц) чисто тепловой, иначе говоря – простой нагрев облучаемой области. Тепловые эффекты микроволнового излучения даже используются в медицине (например, в онкологии) как способ создания локальной «гипертермии», т.е. области повышенного нагрева. Вместе с тем, непрерывно появляются сообщения о некотором более «таинственном» действии такого высокочастотного излучения на человека. Даже если отбросить наивные журналистские измышления о якобы «страшном» вреде мобильных телефонов

или использовании излучения спецслужбами, остается большое количество вполне обоснованных научных сообщений со стороны биологов.

Так, например, в ряде случаев наблюдается усиленная пролиферация (размножение) клеток [1], усиленное производство ферментов или других белков [2], или же, наоборот, конформационные переходы в белках [3], либо при незначительном нагреве, либо при практически полном отсутствии заметного теплового эффекта. Здесь обычно очень важна постановка эксперимента. Как правило, выводы о нетепловом механизме действия излучения обосновываются следующим образом. При воздействии например, непрерывного микроволнового излучения просто регистрируется температура образца (суспензии клеток, пленки и т.п.) и наблюдается специфический выход какого-либо агента (белка) или какой-то специфический биологический эффект. Затем, в контрольном эксперименте образец нагревают с помощью внешнего источника до той же температуры и обнаруживают, что эффекта нет или же он очень мал. На основании такого сравнения делаются выводы о нетепловом механизме воздействия.

Но каким же может быть такой механизм? Достаточно ли обоснованы приведенные выводы о нетепловом механизме? Рассмотрим это подробнее. Собственно целью настоящей работы и является анализ таких возможностей.

Есть и еще одно странное обстоятельство. Как известно, в настоящее время (2012-2015) на Западе чрезвычайно популярны исследования в терагерцевом диапазоне частот – число работ растет лавинообразно. Однако, число работ, посвященных действию терагерцевого излучения на биологические объекты поразительно мало! Даже работ в области микроволн во много раз больше.

Молекулярные механизмы поглощения. Итак, рассмотрим возможные механизмы воздействия. Биологический объект – организм, клетка и т.п. – прежде всего молекулярная система, подчиняющаяся всем законам физики. Поэтому сначала рассмотрим все возможные физические механизмы поглощения излучения молекулами в таком диапазоне частот. Молекула – объект существенно квантовый и подчиняется законам квантовой механики. В соответствии с квантовой механикой, для того, чтобы произошло поглощение, энергия кванта излучения должна быть равна разнице между энергетическими уровнями молекулы. Иначе говоря, излучение должно быть резонансно какому-то типу внутримолекулярного движения со своей лестницей энергетических уровней. Какие виды движения известны в молекуле? Их три: вращения молекул, внутримолекулярные колебания и электронные движения, иначе называемые электронными переходами. Для этих движений давно установлен так называемый принцип Борна-Оппенгеймера. В соответствии с этим принципом, энергии соответствующих движений (а, значит, и разница между соседними уровнями) выстроены так: энергии электронных переходов много больше энергий колебательных движений, а энергии колебательных, в свою очередь, много больше энергий вращений. В шкале энергий это (грубо) выглядит следующим образом: электронные переходы: $1,0 \dots 10 \text{ эВ}$, колебания: $0,01 \dots 1,0 \text{ эВ}$, вращения: $0,01 \dots 10^{-6} \text{ эВ}$.

Теперь сравним: энергия кванта микроволнового излучения с частотой около 1 ГГц составляет несколько микроэлектронвольт (т.е. 10^{-6} эВ). Это важная информация для наших целей. Это означает, что микроволновое излучение в принципе не способно возбуждать непосредственно электронные переходы или молекулярные колебания. Химические реакции, практически всегда связаны с электронными переходами и/или колебаниями. Иначе говоря, никакие химические реакции (вопреки бытующему мнению!) таким излучением напрямую вызываться не могут: просто не хватит энергии.

Единственный вид движения, которому может быть резонансно такое излучение – молекулярные вращения. В самом деле, в газах, где молекулы практически не связаны друг с другом, наблюдают хорошо разрешенные линии вращательного поглощения. Но в конденсированных средах – к которым относятся биообъекты – ситуация

существенно иная. Сильное взаимодействие между близко расположенными молекулами приводит к тому, что энергии состояния зависят от ближайшего окружения и линии переходов сильно «размываются» – это так называемое неоднородное уширение. Для нас это означает лишь то, что получив квант возбуждения, молекула меньше, чем за период вращения передает часть энергии возбуждения соседям, они – дальше и т.п. В итоге, за короткое время (меньше периода вращения и даже периода колебания) вся энергия возбуждения перераспределяется по всем доступным степеням свободы системы в пропорциях, определяемых (грубо) Больцмановским распределением. Это и есть просто нагрев! Итак, резонансное поглощение, например, вращательными переходами может приводить только к нагреву системы.

Но может существовать какой-то более простой классический механизм поглощения такого излучения? Действительно, это – хорошо известное поглощение ионами. (Правда, также известно, что свободный заряд не может поглощать электромагнитное излучение в отсутствие третьего тела – не выполняются законы сохранения импульса и энергии одновременно. Однако, в конденсированных средах этой проблемы нет – третье тело всегда найдется). Несколько огрубляя, можно представить этот механизм следующим образом. Заряд, например, ион в растворе под действием излучения испытывает периодическое действие кулоновских сил со стороны электрической компоненты волны. В итоге, ион начинает совершать «колебания», приводящие к столкновению с окружающими молекулами. В столкновениях кинетическая энергия иона передается окружению, что означает, как и прежде, просто нагрев. Заметим, что молекула воды, обладающая большим дипольным моментом, в рамках этого механизма тоже может рассматриваться просто как два связанных иона, что тоже приводит к простому равновесному нагреву.

Таким образом, мы приходим к неизбежному заключению: единственным возможным механизмом воздействия на биологические объекты является чисто тепловой механизм.

Пространственная селективность. Но тогда получается, что должна быть какая-то методическая ошибка в биологических работах, констатирующих нетепловой характер воздействия. Так ли это? Сходная ситуация, как оказывается, уже была в другой области – в области т.н. «лазерной химии», популярной в семидесятых-восьмидесятых. Тогда также пытались выявить нетепловые химические процессы, вызываемые лазерным излучением. Традиционная постановка эксперимента была очень сходной – в эксперименте регистрировали температуру образца (чаще всего – термпарой), наблюдали получающиеся при лазерном воздействии химические продукты, а затем сравнивали с продуктами простого нагрева (пиролиза) при той же температуре. Продукты почти всегда оказывались разными из чего делался вывод о нетепловом характере лазерного воздействия. Однако, довольно быстро выяснилась причина разницы в продуктах химических реакций: как оказалось продукты чаще всего соответствовали пиролизу, но при значительно более высоких температурах. Тривиальная причина этого – формирование областей с повышенной температурой вследствие «неоднородностей» лазерного излучения во времени и пространстве (т. н. «горячие точки»). Сразу заметим, что никаких горячих точек в микроволновом излучении нет и с этой стороны подход биологов совершенно корректен. Тем не менее, эффект «горячих точек» позволил развить совсем нетривиальный подход в лазерной химии.

Дело в том, что «неоднородности» теплового поля могут быть вызваны не только горячими точками излучения, но и «пространственными неоднородностями» самого образца. Поставим теперь вопрос так: а зачем нужно (или чем интересно) именно нетепловое воздействие излучения? Ответ прост: для возможности управления протекающими процессами. Нетепловой процесс в самом общем случае означает

нетепловое, т. е. неравновесное («небольцмановское») распределение энергии по степеням свободы системы. Это значит, что у нас появляется возможность «селективного», т.е. избирательного воздействия на систему. В частности, становится возможным избирательно воздействовать на один объект не, затрагивая другой. Именно такие процессы были обнаружены группой, в которую входил автор этих строк, на примере инфракрасного лазерного стимулирования неравновесных химических реакций в нитрате целлюлозы в зависимости от надмолекулярной структуры (конформации) полимера. Суть явления состоит в том, что если есть какая-либо неоднородность структуры (кристаллический участок в клубковом полимере, микропримесь с иным сечением поглощения и т.п.), то нагрев такой структуры оказывается неравновесным – отличающимся от нагрева окружения.

Если предположить, что нечто подобное принципиально возможно и при микроволновом или терагерцевом воздействии на биообъекты, то открывается совершенно новый взгляд на возможность нетепловых *селективных* процессов в биообъектах, которые можно вызывать микроволновым излучением.

Надмолекулярное поглощение. Что же может быть такой «неоднородностью» в случае биологического объекта? Собственно, любой элемент – клетка, органелла клетки, большая макромолекула и т. п., имеющая коэффициент (сечение) поглощения, отличающийся от окружающей среды. Как показывают экспериментальные работы, в микроволновом диапазоне есть довольно сильные линии резонансного поглощения колебаниями крупных макромолекул или клетки в целом.

На рис.1 показан спектр микроволнового поглощения вируса *influenza A* в буферном растворе из работы [4]. Как можно видеть, при исключении спектра раствора, имеет место пик поглощения с частотой около 12 ГГц. Для примера сравним со спектром *энтеровируса 71* (рис. 2). Оба вируса отличаются прежде всего размером и формой. Как видно из рисунков, спектр энтеровируса обнаруживает широкий пик около 45 ГГц. Как видно из рис. 3, не только в области десятков гигагерц, но и в области единиц гигагерц обнаруживаются пики поглощения вирусов. На рисунке приведен спектр Мандельштам-Бриллюэновского рассеяния света на кристалле вируса *табачной мозаики*. Обращает на себя внимание тот факт, что и в этом случае наблюдается пик около 12 ГГц, хотя размеры вирусов существенно разные: вирус *influenza A* имеет грубо сферическую форму с диаметром 180-200 нм, тогда как вирус табачной мозаики ближе к цилиндрическому образованию длиной около 300 нм и диаметром 18-20 нм.

Тем не менее, природа поглощения здесь отличается от рассмотренных выше молекулярных механизмов. Из работ [5-6] ясно, что это – поглощение крупномасштабными «акустическими» колебаниями вирусов или их подсистем (макромолекул РНК или белковых оболочек) как целого. Из экспериментов ясно (рис. 1 - рис. 3), что крупномасштабное образование – например, вирус – может представлять собой «неоднородность» в водной среде, имеющую более высокий коэффициент (сечение) поглощения, нежели вода. Тем самым, такая неоднородность будет приводить к пространственно-неравновесному распределению температуры.

Однако, как велика такая разница температур? Нетрудно оценить, что она определяется отношением коэффициентов поглощения неоднородности и среды. (Из приведенных рисунков нельзя вычислить эти коэффициенты, т.к. спектры нормированы на единицу и вычислены из спектров буферного раствора). В случае водной среды, такая разница не может быть большой. Дело в том, что в любом биообъекте, не только вирусе, но даже просто в белке имеется большое количество молекул воды. Количество молекул воды на единицу объема в белке мало отличается от окружающего раствора (~90%). Если основной поглотитель микроволнового излучения – ионы и дипольные молекулы воды (см. ионный механизм выше), то по оценкам, разница в нагреве должна составлять лишь единицы кельвинов.

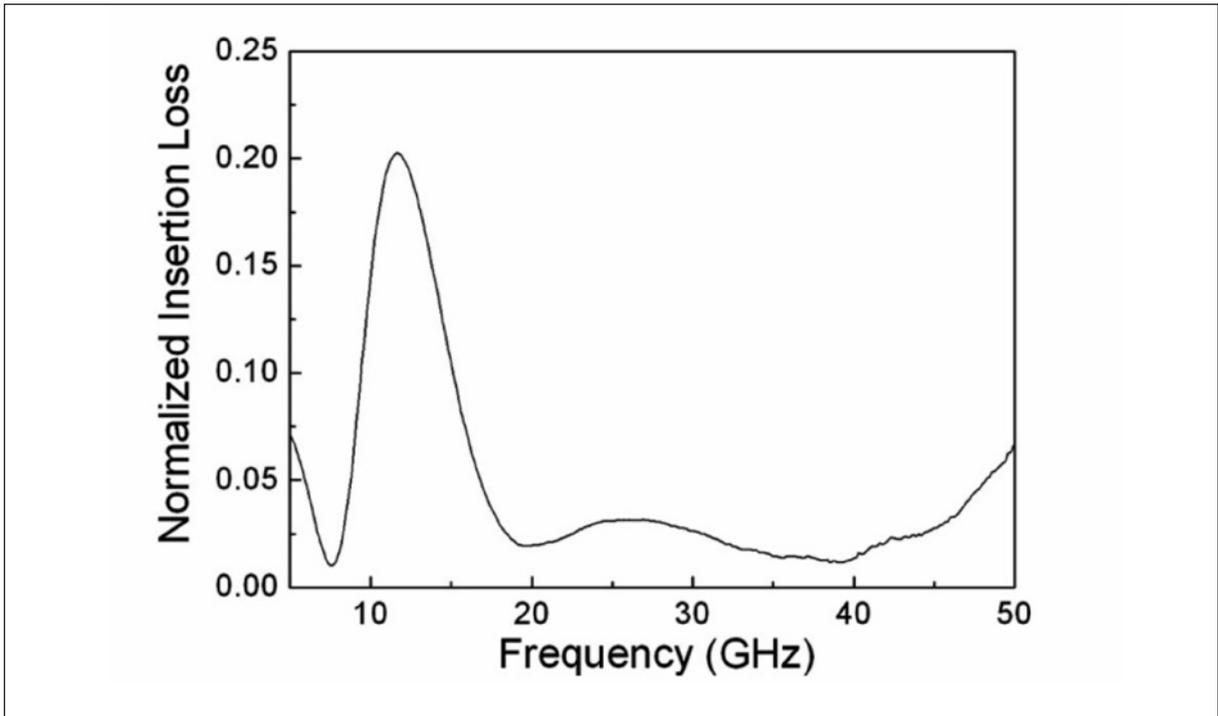


Рис. 1. Спектр микроволнового поглощения вируса influenza A в буферном растворе (спектр буферного раствора учтен) [4].

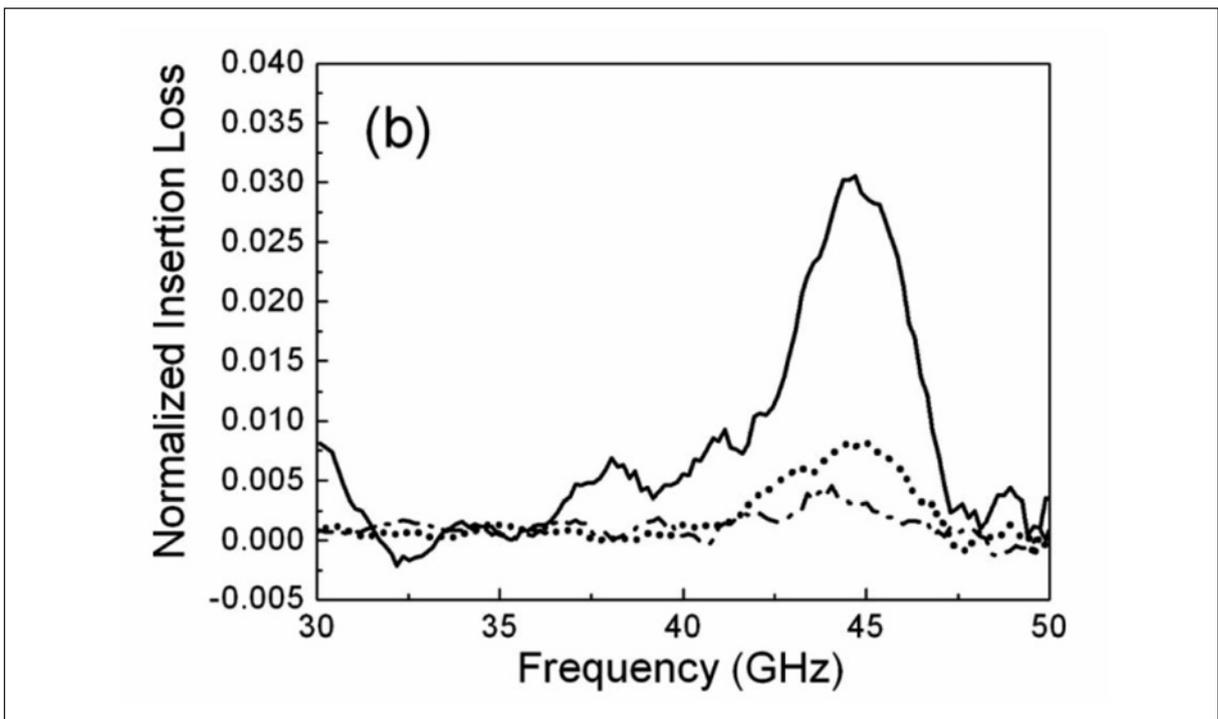


Рис. 2. Спектр микроволнового поглощения энтеровируса (Enterovirus 71) в буферном растворе (спектр буферного раствора учтен) [4].

И снова, очевидно, такой механизм тоже еще не подходит для объяснения биологических явлений. Но может быть все-таки существует какой-то сходный механизм, дающий относительно большую неравновесность?

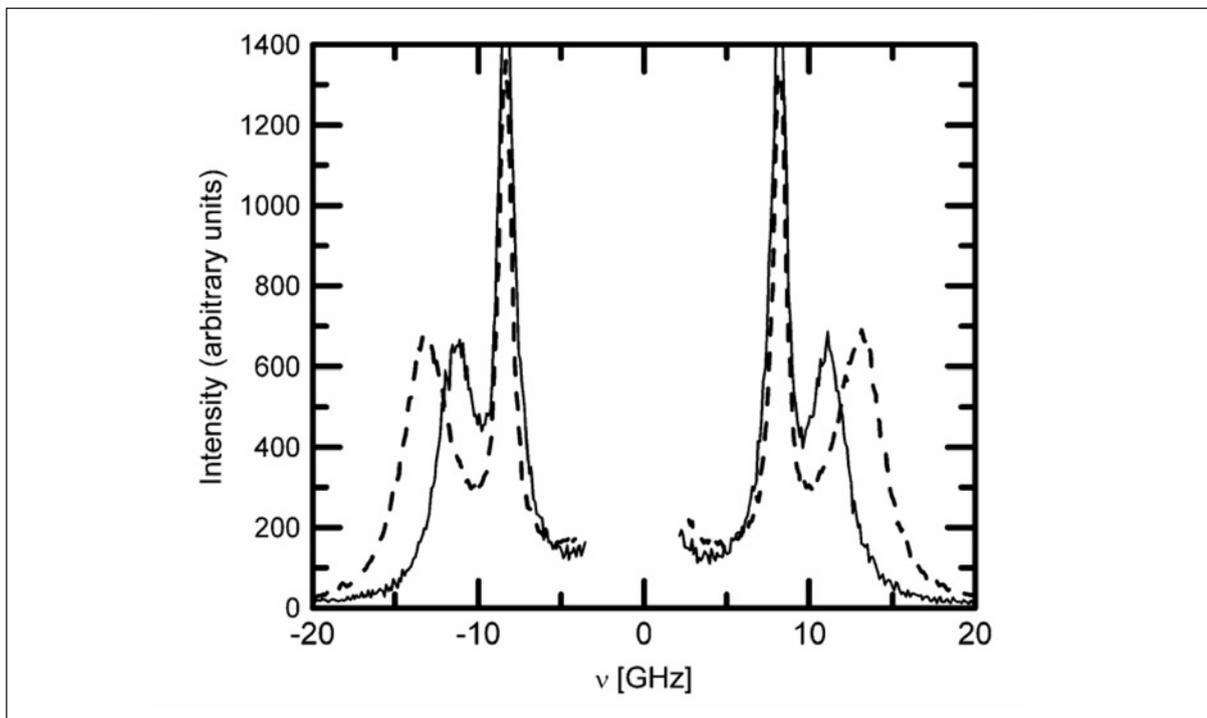


Рис. 3. Спектр Бриллюэновского рассеяния вируса табачной мозаики в виде кристалла [5].

Механизм неравновесности. Любопытно, но основу такого неравновесного механизма подсказывают особенности цитированных работ, отмеченные самими авторами. Так, часто упоминается о «нестабильности» наблюдаемых эффектов (при этом эффекты совсем не малы!); в ряде случаев результаты не подтверждались другими группами [3]; иногда знак эффекта менялся на обратный; относительно низкочастотная модуляция СВЧ излучения могла усилить или затормозить эффект и т. п. С точки зрения физики все это – черты резко нелинейной системы, такой, что малое изменение условий эксперимента приводит к кардинальному изменению фазовой траектории системы.

Такое поведение – «сверхчувствительность» к малым изменениям параметров особенно характерны для систем с положительной обратной связью. Какая же положительная обратная связь возможна при воздействии микроволнового излучения на биологический объект? Такая связь, как оказывается, не только возможна, более того – она очевидна!

На рис. 1 приведены зависимости диэлектрических потерь в воде при различных температурах. (Данные взяты с сайта www.lsbu.ac.uk, хотя эти данные есть в любом соответствующем справочнике). Как известно, диэлектрические потери напрямую связаны с коэффициентом поглощения простой формулой

$$\alpha = \frac{2\pi}{\lambda\sqrt{\varepsilon}} L \quad (1)$$

где L – коэффициент диэлектрических потерь, приведенный на рис. 4-6, λ – длина волны излучения и ε – диэлектрическая проницаемость. (Здесь коэффициент поглощения есть величина, обратная глубине проникновения излучения). Как можно видеть, с ростом температуры широкая полоса диэлектрических потерь в воде – полоса поглощения смещается в низкочастотную сторону с ростом температуры. Это значит, что если выбрать частоту излучения, так, что с ростом температуры поглощение

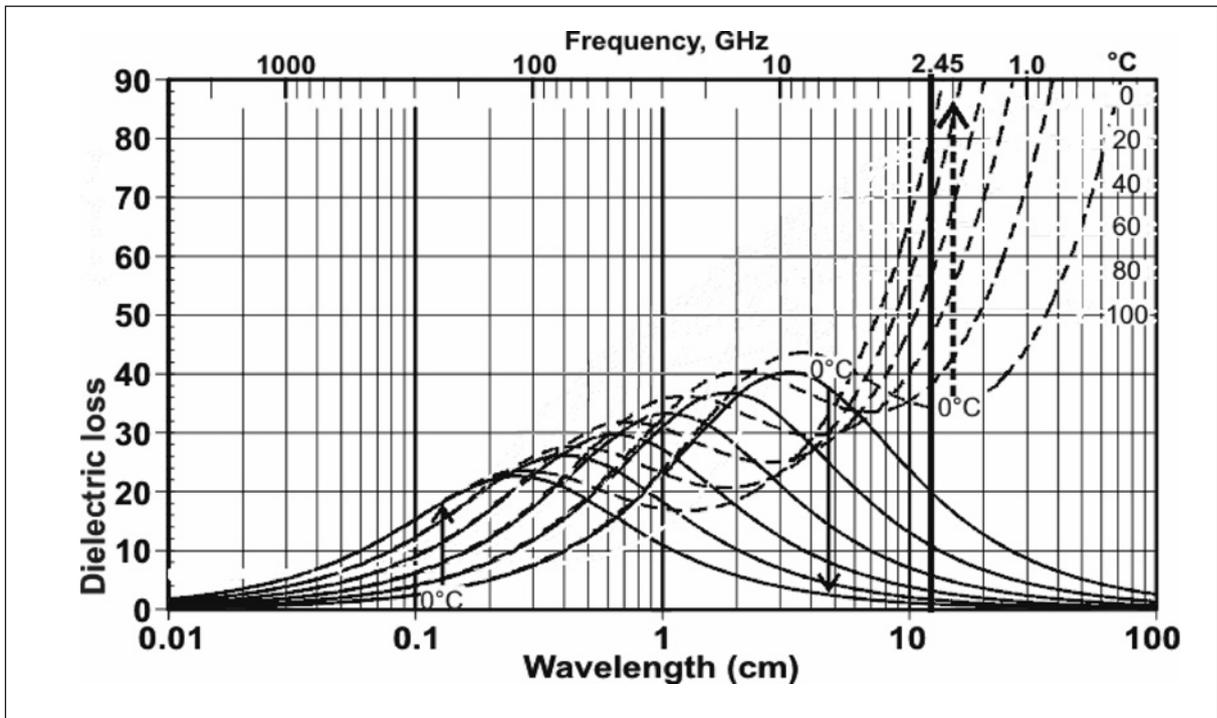


Рис. 4. Зависимость диэлектрических потерь (поглощения) в дистиллированной воде (сплошные кривые) и растворе соли (20 ppt, пунктир) от частоты излучения при различных температурах. Взято из [7].

усиливается (см. рис. 1), то увеличение поглощения вызывает увеличение температуры, которое, в свою очередь ведет к росту нагрева. Таким образом, получаем систему с положительной обратной связью – «поглощение \rightarrow температура \rightarrow поглощение \rightarrow ...». Как видно из рис. 2, этот эффект положительной обратной связи многократно усиливается в присутствии ионов в растворе – а ионы – обязательный элемент любой биосистемы. Более того, видно, что в присутствии ионов почти нет нужды в выборе частоты – излучение с практически любой частотой в области 1- 5 ГГц (см. рис. 4) будет вызывать эту обратную связь. Наконец, на рис. 5 явно приведена кривая зависимости диэлектрических потерь в растворе с концентрацией $NaCl$ 10 ppt от температуры. В области температур 20 C – 100 C, кривая поглощения, как видим, нелинейно, почти экспоненциально растет с ростом температуры.

Итак, неравновесный механизм таков. И в воде, и в биообъекте (например вирусе, клетке и т.п.) возникает нагрев, нелинейно усиливающийся во времени за счет положительной обратной связи *поглощение \rightarrow температура \rightarrow поглощение \rightarrow ...*. Но если сечение (и коэффициент) поглощения биообъекта хоть немного выше, чем окружения, то *скорость* роста его температуры будет выше, чем у раствора и по прошествии времени температура объекта может очень сильно отличаться от температуры среды.

Тогда возникает вопрос: как сильно должны отличаться коэффициенты поглощения для обеспечения разницы хотя бы в десятки кельвинов при известном поглощении воды (рис. 5)? Достаточно ли известной (показанной на рис. 6) нелинейности для такого эффекта?

Для ответа на эти вопросы построим простую математическую модель на основе уравнения теплопроводности:

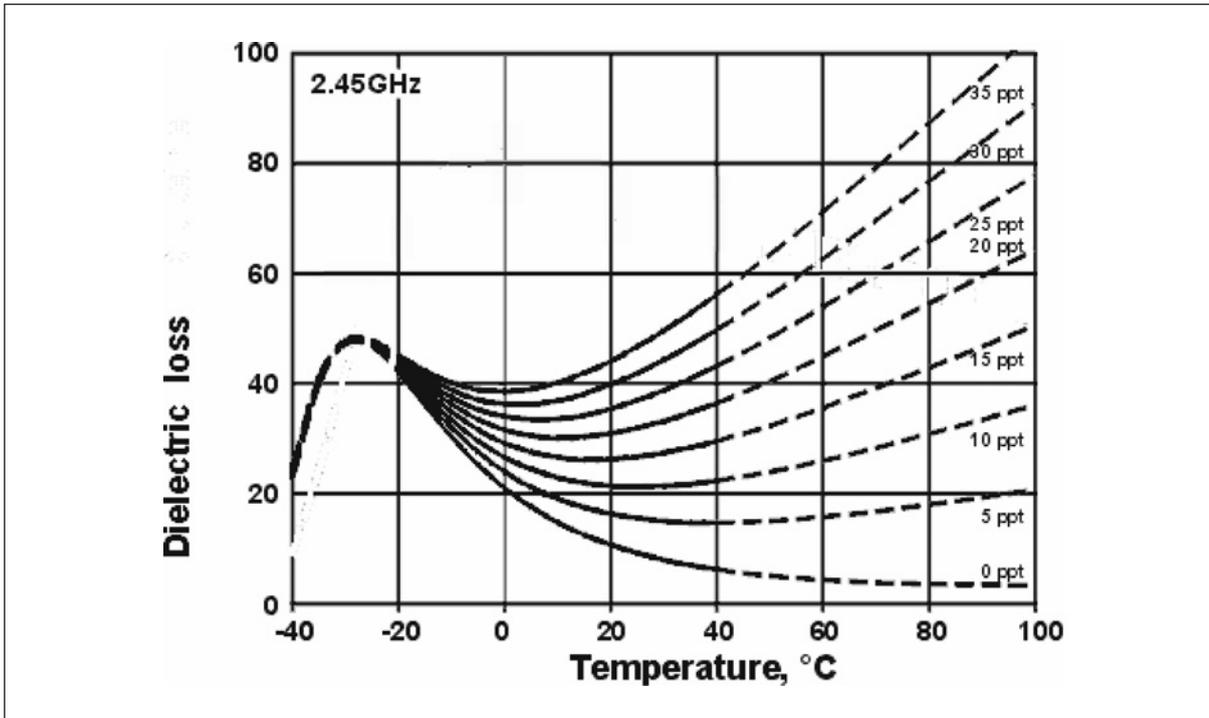


Рис. 5. Зависимость диэлектрических потерь (поглощения) в растворе соли от температуры на частоте 2,45 ГГц при различных концентрациях. Взято из [8].

$$\frac{\partial T(x,t)}{\partial t} = \chi \frac{\partial^2 T(x,t)}{\partial x^2} + Q - \beta(T - T_0) \quad (2)$$

Здесь $T(x,t)$ – распределение температуры во времени и пространстве, χ – теплопроводность, Q – источники тепла, T_0 – температура холодильника, β – коэффициент холодильника. Последний член описывает отток тепла к холодильнику (термостату). Для простоты мы рассматриваем одномерное уравнение теплопроводности. В это уравнение в качестве источников тепла Q входит член, описывающий поглощение излучения

$$Q = \frac{dT}{dt} = \frac{I_0 k(T, x)}{C_p \rho} \quad (3)$$

где в левой части уравнения стоит скорость «продукции» температуры за счет поглощения излучения, I_0 – интенсивность излучения, C_p – теплоемкость, ρ – плотность. (Это – обычное выражение для источников тепла в лазерной физике).

Заметим, что здесь коэффициент поглощения является функцией температуры, определяемой кривой диэлектрических потерь (рис. 6). Это значит, что важно правильно аппроксимировать эту кривую какой-либо подходящей функцией. Возьмем, например, в качестве аппроксимации функцию

$$k(T) = k_0 \exp(aT) \quad (4)$$

где k_0 и a – подбираемые параметры.

Расположим теперь в районе нуля $x = 0$ «объект», т.е. область от $-R$ до R , где коэффициент поглощения k_l хотя и близок к k_0 , но несколько отличается от него, $k_l > k_0$. Иначе говоря,

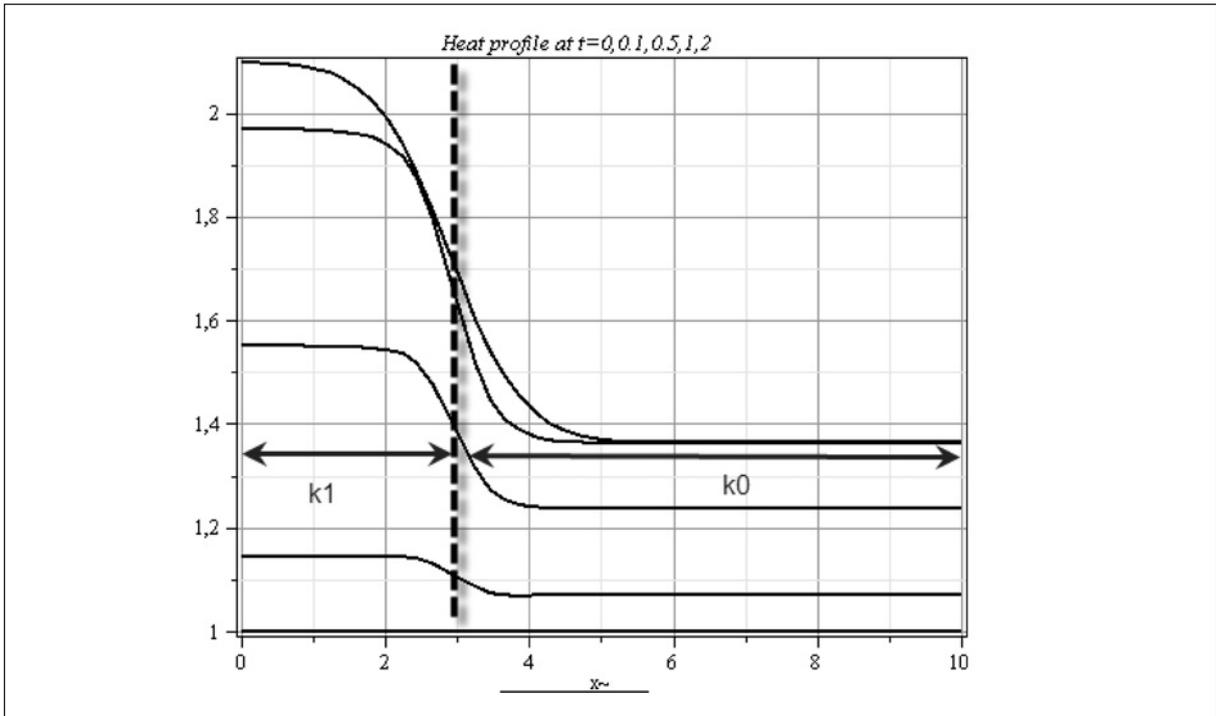


Рис. 6. Относительные профили пространственного распределения температуры в различные моменты времени согласно модели (см. текст). Область от 0 до 3,0 принадлежит «объекту» с коэффициентом поглощения k_1 , от 3,0 до 10,0 - окружающей среде с коэффициентом поглощения k_0 . Разница между k_1 и k_0 - всего 0,01% ($k_1 = 1.0001 k_0$).

$$k(T) = k_1 \exp(aT) \cdot \theta(R - x) + k_0 \exp(aT) \cdot \theta(x - R) \quad (5)$$

где θ - ступенчатая функция Хевисайда. В начальный момент времени считаем температуру везде одинаковой ($T_0 = 300 K$).

Рассмотрим теперь динамику распределения температуры во времени в такой модели. Пусть коэффициент поглощения k_1 объекта («вируса»), находящегося в резонансе с излучением отличается от коэффициента поглощения окружающей среды («воды») k_0 на очень малую величину $\sim 0.01\%$ (т.е. $k_1 / k_0 = 1.0001$). Будем считать, что скорость отвода тепла β много меньше, чем скорость нагрева излучением - это практически всегда так для биообъектов). Результаты численного решения такой задачи (4)-(7) с функцией «обратной связи» вида (4) и $k_1 / k_0 = 1.0001$ приведены на рис. 9-10.

(Заметим, что если кривую обратной связи $k(T)$ аппроксимировать квадратичной функцией, а не экспонентой, то задача может быть решена аналитически. Однако, для дальнейших выводов это несущественно).

Как можно видеть, несмотря на ничтожно малую разницу в поглощении, по прошествии некоторого времени температура резонансного объекта может быть в несколько раз больше, чем температура среды! (На рис. 6. температура объекта с k_1 - область 0..3 в 1.7 раза выше, чем среды).

На рис. 7 хорошо видна динамика такого неравновесного нагрева: скорость нагрева объекта значительно выше, чем среды, а при выключении излучения ($t = 3$) отток тепла начинает снижать и выравнивать температуру.

В данной модели эта динамика, очевидно, такова, что если просто продолжать нагрев, то формально можно довести разницу температур до любого значения! Ясны

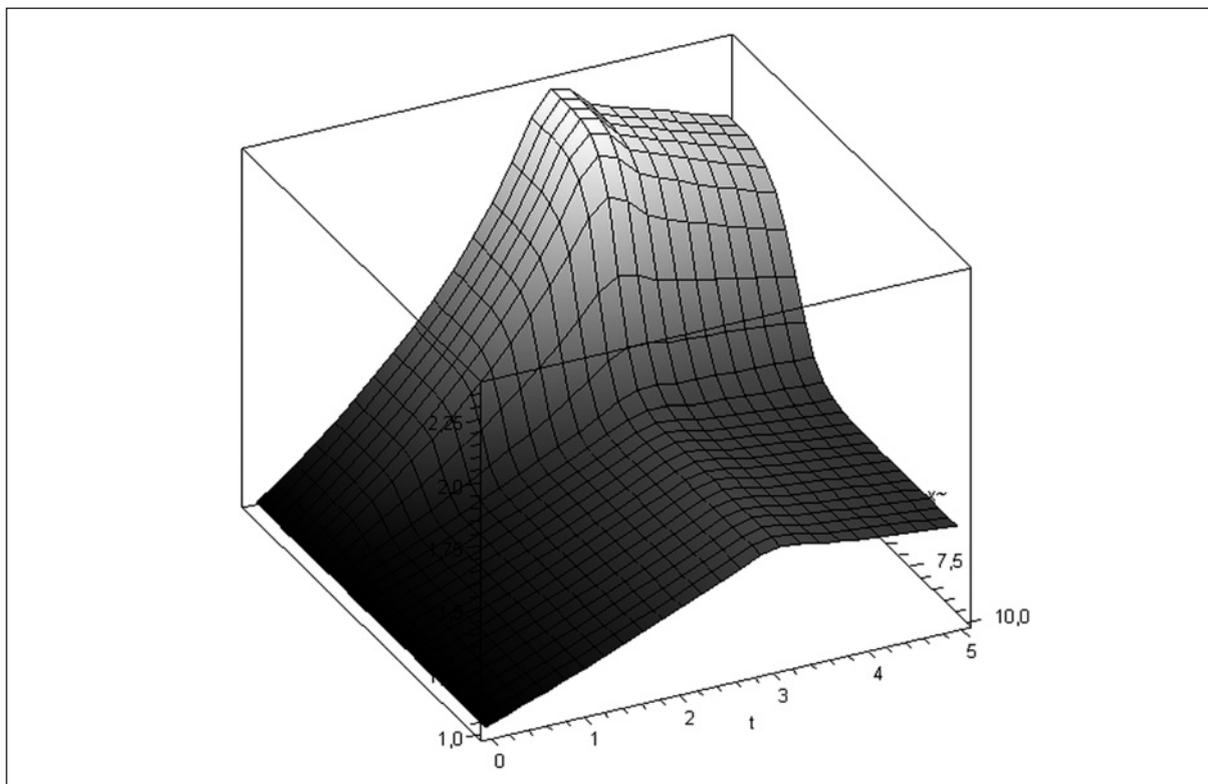


Рис. 7. Пространственно-временное распределение температуры вокруг «резонансного» биообъекта с учетом обратной связи в поглощении. Область по X от 0 до 3,0 принадлежит «объекту» с коэффициентом поглощения k_1 , от 3,0 до 10,0 - окружающей среде с коэффициентом поглощения k_0 . (Разница между k_1 и k_0 – всего 0,01% ($k_1 = 1.0001 k_0$)). Излучение выключается при $t = 3$.

управляющие параметры: полученный сценарий всегда реализуется, если константа α в (4), характеризующая экспоненту положительной обратной связи (рис. 5) превосходит скорости отвода тепла β и теплопроводность χ (релаксацию к полному равновесию).

Итак, из приведенной простой модели ясно, что даже ничтожно малая разница в поглощении, может вызывать большую разницу в результирующей температуре, т. е. сильно неравновесный нагрев.

Селективное воздействие на биообъекты. Как же можно это использовать для селективного (избирательного) воздействия на биообъекты? Ответ достаточно очевиден. Предположим, что у нас имеется в растворе два биологических объекта, качественно отличающиеся, например, только размером и формой. Вирусы или бактерии – хороший пример этого: какие-либо макроскопические химические различия между различными вирусами, клетками и т. п. фактически отсутствуют (химически биообъекты созданы из одних и тех веществ). Это означает, что невозможно различить такие объекты просто спектральными методами (спектры поглощения и испускания в инфракрасной, видимой и ультрафиолетовой областях будут практически неотличимы (если только не применять специфические красители и т. п., что не всегда возможно). Однако, если размер и форма объектов такова, что один объект имеет пик поглощения на частоте излучения (т. е. в резонансе с излучением), а другой -нет, то, очевидно, выбирая длину волны излучения мы можем избирательно «перегреть» либо один, либо другой вирус.

Другой вариант – белки, например белки-прионы или ферменты. Аналогично, выбор длины волны и интенсивности излучения, как видим, может приводить к избирательной денатурации, например, прионов заданного типа, не трогая других.

Еще более тонко такой механизм может позволить управлять ферментативными реакциями. Известно, белки-ферменты активны только в довольно узком диапазоне температур. Например, нагрев до 30-40 °C приводит к увеличению числа оборотов (скорости) ферментативной реакции, а более 60 °C – к денатурации фермента. Это означает, что некоторой интенсивности можно стимулировать активность только (резонансного) заданного фермента, а при более высокой – подавлять ее.

Заметим, что микроволновая гипертермия уже применяется для лечения раковых опухолей. Из приведенной модели ясно, что тонкий подбор частоты и, главное, интенсивности излучения, в принципе, может приводить к избирательному уничтожению только раковых клеток.

Таким образом, как видим принципиально все-таки существует связанный с водой нелинейный механизм биологического действия микроволнового/терагерцевого излучения, который может приводить к неравновесному распределению возбуждения и быть основой для селективного (избирательного) управления процессами в биологических объектах.

Литература

1. C. Pérez-Castejón, R.N. Pérez-Bruzón, M. Llorente, N. Pes, C. Lacasa, T. Figols, M. Lahoz, C. Maestú, A. Vera-Gill, A. del Moral and M.J. Azanza, “Exposure to ELF-pulse modulated X band microwaves increases in vitro human astrocytoma cell proliferation” // *Histology and Histopathology*, 2009, **24**, 1551-1561.
2. Birendra N., Pramanik, Urooj A. Mirza, Yao Hain Ing, Yan-Hui Liu, Peter L. Bartner, Patricia C. Weber and Ajay K. Bose, “Microwave-enhanced enzyme reaction for protein mapping by mass spectrometry: A new approach to protein digestion in minutes”// *Protein Science*, 2002, **11**, 2676–2687.
3. David I. de Pomerai, Brette Smitha, Adam Dawea, Kate Northa, Tim Smitha, David B. Archera, Ian R. Ducea, Donald Jonesb, E. Peter M. Candido, “Microwave radiation can alter protein conformation without bulk heating” // *European Biochemical Societies Letters*, 2003, **543**, 93-97.
4. Tzu-Ming Liu, Hung-Ping Chen, Li-Tzu Wang, Jen-Ren Wang, Tang-Nian Luo, Yi-Jan Chen, Shen-Iuan Liu, and Chi-Kuang Sun, “Microwave resonant absorption of viruses through dipolar coupling with confined acoustic vibrations” // 2009, *Applied Physics Letters*, **94**, 043902.
5. B. Stephanidis, S.Adichtchev, P.Gouet, A. McPherson, A.Mermet, “Elastic Properties of Viruses” // *Biophysical Journal*, 2007, **93**, 1354–1359.
6. V. N. Blinov and V. L. Golo, “Acoustic spectroscopy of DNA in the gigahertz range” // *Physical Review E*, 2011, **83**, 021904.
7. <http://www1.lsbu.ac.uk/water>
8. P. O. Risman and B. Wäppling-Raaholt, “Retro-modelling of a dual resonant applicator and accurate dielectric properties of liquid water from –20 °C to +100 °C” // *Meas. Sci. Technol.*, 2007, **18**, 959-966.

ON THE MECHANISM OF BIOLOGICAL ACTION OF MICROWAVE AND TERAHERTZ RADIATION

V.B. Oshurko

Prokhorov General Physics institute of Russian Academy of Sciences

vbo08@mail.ru

Received 24.03.2015

All known physical mechanisms of microwave and terahertz radiation absorption are briefly analyzed. It is shown that basic mechanism of biological action is just an equilibrium heating of media. However, in case of non-homogeneous media the effect of "spatial selectivity" (e.g. different heating of different parts) is also possible. It was shown that nonlinear (increasing) dependence of dielectric loss in water on temperature can provide positive feedback between temperature and absorption of radiation. This positive feedback can greatly amplify the effect of spatial selectivity and provides a possibility of selective control of biological systems.

ВЧ И СВЧ РЕЗОНАНСЫ СМЕШАННЫХ КВАНТОВЫХ СОСТОЯНИЙ ОРТО-ПАРА СПИН-ИЗОМЕРОВ H₂O КАК ФАКТОР ИХ КОНВЕРСИИ И ВОЗДЕЙСТВИЯ НА БИООБЪЕКТЫ

С.М. Першин

Научный центр волновых исследований ИОФ им. А.М. Прохорова РАН

pershin@kapella.gpi.ru

Поступила 24.03.2015

Недавно А.И. Коновалов с сотрудниками обнаружил увеличение электропроводности одновременно с увеличением диаметра нанообъектов (замедлении диффузии) при некоторых концентрациях водных растворов. Установлено, что эти особенности не проявляются при экранировании образцов от внешних электромагнитных полей (ЭМП) и восстанавливаются при их облучении ЭМП. Для объяснения наблюдаемого парадокса предложена новая концепция воды с учетом квантовых отличий орто-пара спиновых изомеров H₂O: спиновая селективность при гидратации и спонтанном формировании льдо-подобных структур, квантовые биения и спин-конверсия, индуцированная резонансной накачкой микроволновыми полями (слабыми ЭМП), которые ранее не учитывались. С учетом квантовых отличий изомеров H₂O рассмотрен их транспорт через мембраны клеток. Предложено считать, что дипольный ключ в водных каналах мембран выполняет роль вентиля, пропускающего только вращающиеся молекулы. Показано, что выявленная селективность более предпочтительна для орто-H₂O, которые не имеют уровня с нулевой вращательной энергией. Анализ спектра вращательных переходов спиновых изомеров H₂O показал, что существуют близкорасположенные пары уровней орто-пара H₂O, которые образуют смешанные квантовые состояния и допускают резонансную накачку в широком диапазоне частот ЭМП, включая микроволны СВЧ.

УДК УДК 577.352; УДК 541.11

1. Введение

Предметом настоящей работы является анализ и физическое обоснование экспериментально установленных ряда аномалий водных растворов низкой концентрации, а также влияние ВЧ и СВЧ резонансов смешанных квантовых состояний орто-пара H₂O при электромагнитном воздействии на биообъекты. Предложен новый подход к физике воды, который учитывает квантовые отличия орто/пара спиновых изомеров H₂O в воде.

Этот подход позволяет физически непротиворечиво интерпретировать наблюдаемые особенности, что и является целью данной работы.

Известно, что физический механизм повышения биологической активности [1], изменения свойств мембран и увеличения флуктуаций ИК поглощения [2] водных растворов аномально низкой концентрации ($\sim 10^{-12}$ - 10^{-18} М) остается пока неустановленным. Позднее [3-6], А.И. Коновалов с сотрудниками выявил парадоксальное поведение параметров водных растворов низких концентраций - увеличение электропроводности при замедлении диффузии (увеличение эквивалентного диаметра нанообъектов, измеренного методом динамического рассеяния света), которые коррелировали с повышением биоактивности [1]. При этом авторы, впервые, насколько нам известно, в окрестности этих же концентраций обнаружили проявление оптической активности (вращение плоскости поляризации) и сделали предположение, что нанообъекты гидратируются упорядоченными структурами воды в процессе её самоорганизации. Кроме того, было обнаружено, что наблюдаемый парадокс и экстремумы других параметров в области низких концентраций не проявляются в условиях экранирования образцов от внешних ЭМП в пермалловом контейнере [6]. Эти факты указывают на неизвестные ранее фундаментальные закономерности, которые не удовлетворяют существующим моделям воды [7] и противоречат кинетической теории жидкости [8]: одновременное повышение подвижности и размера структурных форм гидратных оболочек молекул.

2. Орто-пара спиновые изомеры H_2O

Для снятия обнаруженного противоречия мы предложили учесть малоизвестные степени свободы H_2O – квантовые отличия спиновых изомеров H_2O [9,10] и сверхтонкие взаимодействия в виде конверсии изомеров, а также их проявление в свойствах воды. Этот подход позволил нам показать ранее [11], что значения температур особых точек T_c воды, тяжелой воды и льда не являются случайными. Они отражают резонансное совпадение энергии броуновского движения kT_c (экстремум максвелловского распределения по скоростям) и энергии квантов вращательных переходов $h\Omega_{mn}$ орто или пара спин-изомеров H_2O (или одновременно обоих, например, в окрестности температур 4, 19 и 60 °С)

$$kT_c = h\Omega_{mn} \quad (1)$$

Выполнение условия резонанса (1) и увеличение вероятности орто-пара конверсии в смешанных квантовых состояниях [12] при возбуждении или релаксации молекул H_2O позволили нам объяснить некоторые биофизические эксперименты [13,14], которые не получили пока удовлетворительной интерпретации.

Для анализа роли спин-изомеров H_2O в свойствах воды и водных растворов малых концентраций [1-6] обоснуем их присутствие на основе экспериментальных данных, которые не упоминались нами ранее [14]. Наличие мономеров H_2O в воде не является очевидным фактом. При этом регистрация их вращения даёт основание для введения в рассмотрение орто и пара спин-изомеров H_2O , поскольку отличия их волновых функций проявляется в спектрах вращения H_2O , которые обеспечивают отнесение линий и идентификацию изомеров. Заметим, что сигнал магнитно-резонансного томографа в единицах протонной плотности пропорционален, как показано нами ранее [15], содержанию орто-изомера H_2O , магнитный момент которого равен единице, а пара-изомера – нулю [9-14].

Декаду назад авторам работы [16] удалось обнаружить (при повышении спектрального разрешения до $\sim 2 \text{ см}^{-1}$) модуляцию коэффициента поглощения воды при увеличении частоты в области крыла Релея (10 - 100 см^{-1}), которая показана (круги) на Рис. 1 (правая ось ординат). Мы предположили, что наблюдаемая модуляция является прояв-

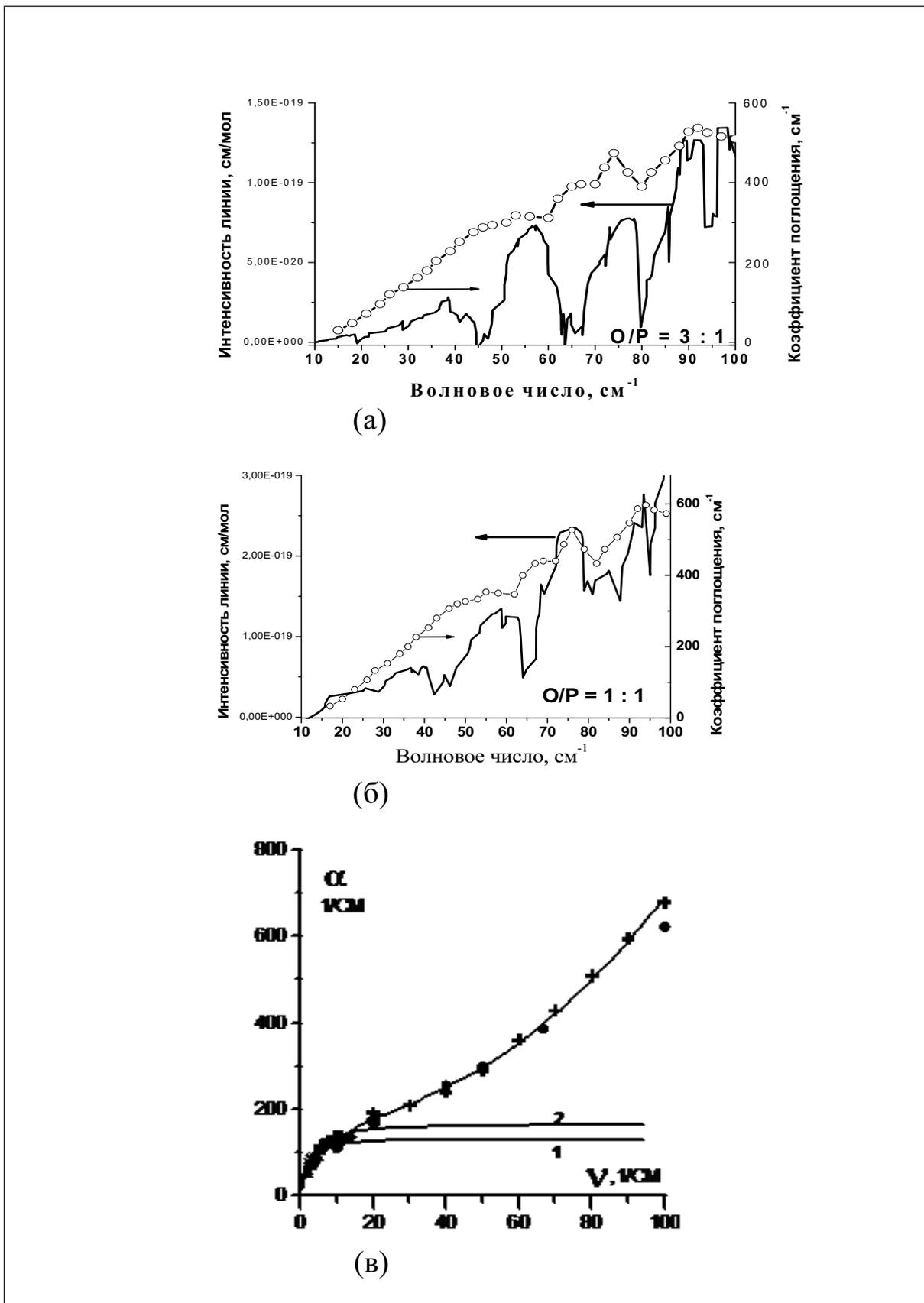


Рис. 1. Спектр поглощения синхротронного излучения в воде (круги) [16] с разрешением 2 см⁻¹ и сглаженный с таким же шагом спектр вращения молекул H₂O (линия) из базы HITRAN при орто/пара отношении 3:1 (а) и 1:1 (б), а также спектр с монотонной зависимостью (расчёт – линия, эксперимент – точки) поглощения воды (в) с разрешением 10 см⁻¹ из работы [18].

лением вклада вращательного спектра мономеров (спин-изомеров [9,10]) H_2O , присутствующих в воде. Затем взяли вращательный спектр H_2O с равновесным отношением орто/пара = 3:1 из базы данных HITRAN и провели его сглаживание с шагом 2 см^{-1} (для удобства сравнения с экспериментом). После нормирования полученной огибающей на максимальное значение коэффициента поглощения [16] в окрестности частоты 100 см^{-1} была получена немонотонная зависимость (рис. 1а, линия), которая удовлетворительно отражала модуляцию экспериментальных данных (круги на рис. 1а). Далее, принимая во внимание наши данные о неравновесном отношении орто/пара=1:1 в воде комнатной температуре [17], мы уменьшили амплитуду и вклад линий орто- H_2O в суммарный спектр в три раза. После этой процедуры соответствие огибающей (линия на рис. 1б) экспериментальным данным (круги) заметно улучшилось. Полученное совпадение с учетом вклада полосы межмолекулярных колебаний в окрестности частоты 50 см^{-1} указывает на существование (вращение) орто-пара спиновых изомеров H_2O в воде с неравновесным соотношением 1:1. Здесь следует заметить, что общепринятым результатом в настоящий момент считается монотонная зависимость (линия расчета и точки эксперимента на рис. 1в из работы [18]) коэффициента поглощения (диэлектрической проницаемости) воды, полученной при низком спектральном разрешении по сравнению с работой [16].

В работе [19] была также обнаружена модуляция кинетики флюоресценции реакционных центров, которую авторы (в отличие от работы [16]) отнесли к вращению мономеров H_2O . Заметим, что реакционные центры имеют замкнутые полости с мономерами H_2O (данные рентгеноструктурного анализа), в которых они могут свободно вращаться. Такие факты дают основание полагать, что и в объемной воде существуют полости для вращения H_2O . Оставалось неясным, почему авторы работы [19] выбрали частоты только некоторых мод H_2O и их обертоны для сравнения с данными эксперимента, пренебрегая десятками других мод, частоты которых близки к выбранным. Принимая во внимание квантовые отличия орто и пара спиновых изомеров H_2O (D_2O), нам удалось обосновать подобное «прореживание» вращательного спектра частот H_2O [20] (и D_2O тоже [21]) тем, что переходы с частотами, наблюдаемыми в эксперименте [19], образуют смешанные квантовые состояния [12], в которых наблюдается спин-конверсия изомеров. Селекция [20,21] именно таких вращательных переходов обеспечила хорошее совпадение с модуляцией кинетики [19].

Недавно [22] была выявлена корреляция увеличения электропроводности воды с увеличением ее плотности в окрестности известной особой температурной точки $4 \text{ }^\circ\text{C}$. Отметим, что этот факт согласуется с обнаруженным ранее кратным увеличением «подвижности» воды в окрестности $4 \text{ }^\circ\text{C}$ [23]. Следует отметить, что подобный парадокс «выше плотность раствора - выше подвижность», но уже с водным раствором-суспензией гемоглобина, был обнаружен ранее [24] в окрестности другой особой температурной точки воды $36.6 \text{ }^\circ\text{C}$. Поскольку наблюдаемые особенности [22-24] не получили авторской интерпретации, нами было дано их физическое обоснование с привлечением механизма орто-пара конверсии H_2O в окрестности температур особых точек [13,14], которая индуцирована резонансными (1) столкновениями молекул с переходом в смешанное квантовое состояние [12]. В такие состояния молекула(ы) H_2O реакционных центров попадает при возбуждении (синхронно с экситонами [20,21]) и релаксации в экспериментах накачка-зондирование [19].

Наконец, сравнительно недавно [25] удалось зарегистрировать колебательно-вращательные линии орто и пара спин-изомеров в воде методом Фурье-спектроскопии. Эти данные подтверждают наши результаты о наблюдении вращения молекул H_2O , полученные ранее [26,27] при четырехволновом смешении.

Приведенный набор экспериментальных данных, с учётом обсуждавшихся ранее [14], даёт основание полагать, что спиновые изомеры существуют и вращаются в воде и

водных растворах. Конверсия изомеров H₂O изменяет их отношение, подвижность (вязкость) растворов [13,14,24]. Физически ясно, что снижение вязкости воды [23] и растворов [24] обеспечивает увеличение электропроводности [22] также как и в экспериментах Коновалова [3-6] (даже без учёта возможности повышения подвижности ионов).

Рассмотрим далее, может ли повышение электропроводности коррелировать с увеличением толщины слоя гидратации (диаметра наноккомплексов [3-6]) с определенной структурой водородосвязанных молекул H₂O.

3. Льдоподобные структуры в воде

В серии экспериментальных результатов опубликованных работ, перечисленных в работе [14], было установлено, что предпочтительной структурой водородосвязанных комплексов воды и гидратных оболочек биомолекул в водных растворах является динамическая структура гексагонального льда. Дополним этот список [14] новыми данными для обоснования парадокса Коновалова [6].

Применение методов статистики дробных моментов к анализу ОН полосы комбинационного рассеяния (КР) в воде позволило нам [28] выделить в ней компоненту 3200 см⁻¹, характеристическую для структуры гексагонального льда I_h [7], вклад которой должен уменьшаться с повышением температуры воды в соответствии с физикой теплового разрушения водородных связей. Нам удалось зарегистрировать ожидаемое уменьшение методом КР [29] (см. рис.2). Из рисунка видно, что при температуре воды 99 °С вклад компоненты 3200 см⁻¹ заметно меньше, чем при 13 °С, но остаётся значимым, что указывает на существование в воде динамических структур подобных льду I_h .

В другой серии измерений нам удалось зарегистрировать аномально большие деформации огибающей ОН полосы. На рис. 3 показаны спектры ОН полосы КР импульсов второй гармоники Nd:YAG лазера в воде комнатной температуры, которые были записаны последовательно через 10 секунд. Из рисунка видно, что первый спектр (толстая линия) был подобен спектру КР во льдах с доминирующей компонентой 3200 см⁻¹ (сравни с рис.2) что указывает на флуктуационную сборку льдоподобного каркаса из молекул H₂O в объёме перетяжки (который содержит порядка ~10¹⁹ молекул) пучка лазера за время усреднения по 8 импульсам длительностью 10 наносекунд. Напротив, второй спектр (тонкая линия) с огибающей ОН полосы характерной для нагретой воды.

Серия работ [30-32] о структурных превращениях льдов [30], выращенных из свободных молекул H₂O на поверхности сверхтекучего гелия, а также эволюция структур в кластерах молекулярных пучков H₂O [31,32], указывает, что структура гексагонального льда наиболее вероятна при спонтанном переключении водородных связей из-за тепловых флуктуаций.

Особый интерес для нас представляет работа [32] с пучками молекул H₂O, в которой обнаружен размер-зависимый фазовый переход «беспорядок-порядок». Этот спонтанный переход от аморфной к упорядоченной фазе (подобной кристаллу льда I_h) был обнаружен по появлению в спектре характеристической компоненты 3200 см⁻¹ при достижении порогового числа молекул (275) в кластере. Отсюда следует, что в воде [29] и в пленке воды, в которой регистрировали структуру льда [33] при комнатной температуре атомно-силовым микроскопом, а также в слоях гидратирования макромолекул в водных растворах [34] размеры «динамических» льдоподобных структур нанообъектов могут достигать сотен нанометров, что наблюдалось Коноваловым с сотрудниками [4-6]. Заметим, что на границе раздела, например, с воздухом или с гидрофильной поверхностью толщина пограничного слоя (из которого вытесняются полистирольные микрочастицы) в воде может достигать нескольких микрон [35].

На основе приведенных данных можно заключить, что динамическая структура гексагонального льда водородосвязанных каркасов в воде более вероятна, чем другие. То-

гда внутри шестигранного (вдоль оси «с») канала-каркаса с поперечником до 6 Å [18] гексагонального льда всегда могут локализоваться молекулы H_2O [36] и ионы H_3O^+ и OH^- [37], что согласуется с известной клатратной моделью О.Я. Самойлова [18,38]. Впервые, насколько нам известно, предположение о наличии льдоподобных кластеров в воде было сформулировано в работе Röntgen W.C. [39]. Подобная локализация мономеров H_2O , скорее всего, имеет место не только при формировании структур льда I_h в воде, но и в кластерах пучков [32]. Об этом свидетельствует спектральная компонента заметной амплитуды колебательно-вращательных переходов на частоте мономеров в окрестности $\sim 3700 \text{ см}^{-1}$ (см. Рис.2 в работе [32]).

Следует заметить, что в отличие от компактной полости в модели Самойлова [38] мы подчеркиваем осесимметричность и протяженность льдоподобных каналов, которые обеспечивают кратное увеличение коэффициента самодиффузии мономеров и ионов (электропроводность) [37] из-за большой длины свободного пробега по каналу без столкновений. Ранее [40] подобный механизм был принят для обоснования аномально большого (10-кратного) увеличения диффузии альфа частиц вдоль оси каналов кристаллических решёток мишени.

Таким образом, приведенные выше данные допускают, что в воде и водных растворах сосуществуют молекулы H_2O и льдоподобные структуры-каркасы. Кроме того, локализация свободных мономеров H_2O в каналах [36] позволяет нам ввести в рассмотрение их квантовые отличия по спиновому состоянию (орто и пара спин-изомеры). Как будет показано ниже, все эти факторы взаимосвязаны и необходимы для обоснования физики парадокса Коновалова, а также влияние электромагнитного облучения на био-объекты.

4. Модель

Физически ясно, что повышение биологической активности [1,3] и электропроводности [4-6] растворов в области низких концентраций обусловлены увеличением подвижности (диффузии) молекул воды и ионов, а также снижением гидродинамической вязкости раствора. Напротив, замедление диффузии, D , обусловлено увеличением диаметра, $2r$, нанобъектов или вязкости, η , которые связаны соотношением Стокса-Эйнштейна

$$D = kT(6\pi\eta r)^{-1} \quad (2)$$

Отсюда следует, что в гомогенных растворах эти параметры антикоррелируют: $D \sim (2r\eta)^{-1}$. Напротив, корреляция этих параметров в водных растворах составляет основу парадокса Коновалова [3-6].

Обнаруженное противоречие снимается, если расширить существующие модели воды и растворов на основе новых данных следующими допущениями:

1. динамические каркасы водородосвязанных комплексов и слоев гидратации имеют структуру гексагонального льда с каналами вдоль оси «с», длина которых превышает среднюю длину свободного пробега молекул вне каналов, а поперечный размер каналов, $\sim 6 \text{ Å}$ [18,38,41], допускает локализацию (как в контейнере) и движение в них мономеров H_2O и ионов;
2. спиновые изомеры H_2O проявляют квантовые отличия: орто- H_2O имеют магнитный момент и взаимодействуют (вода омагничивается [42]) с магнитным полем (см. магниторезонансную томографию [15]), напротив, пара- H_2O не имеет магнитного момента; термы их вращательных переходов разные, изомеры орто- H_2O вращаются всегда, напротив, часть пара- H_2O не вращается и проявляет спиновую селективность связывания при гидратации неорганических [10] и органических соединений [43], а также биомолекул [27,44,45] и при образовании водородосвя-

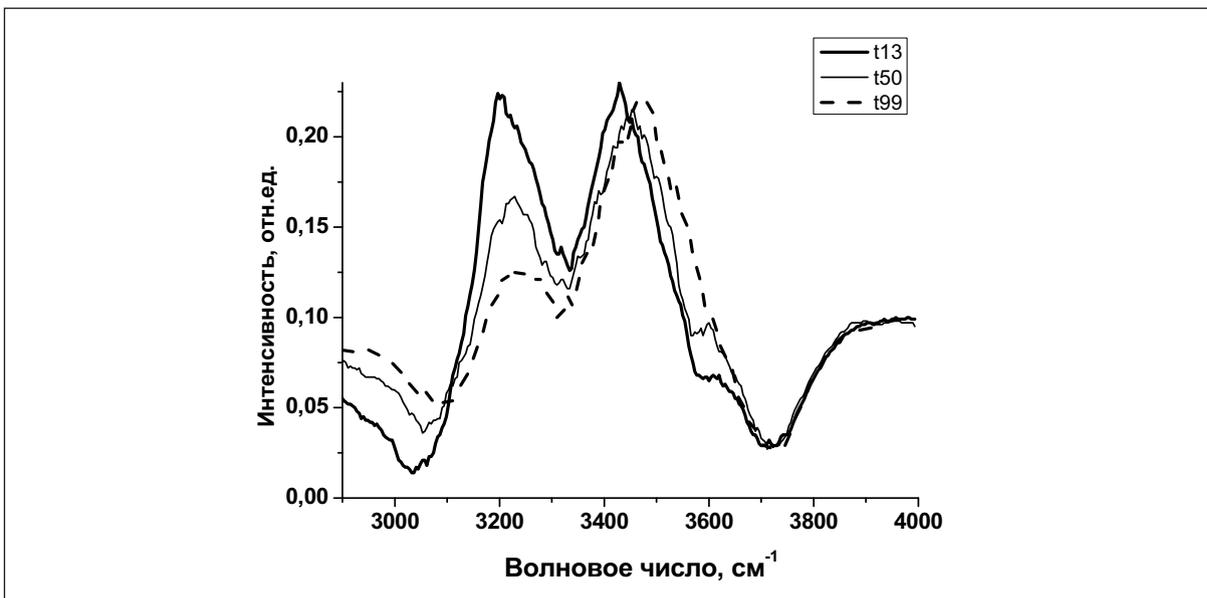


Рис. 2. Уменьшение вклада спектральной компоненты льда 3200 см⁻¹ в ОН полосу комбинационно-го рассеяния в воде с повышением температуры от 13 °C (толстая линия) до 50 °C (тонкая линия) и 99 °C (пунктир).

занных комплексов [46]; при этом близкорасположенные энергетические уровни орто и пара изомеров H₂O образуют смешанные квантовые состояния, вероятность спиновой конверсии в которых возрастает, например, при резонансной накачке электромагнитным излучением [12];

3. вода при комнатной температуре – неравновесная жидкость с отношением орто/пара = 1:1 [17], что соответствует спиновой температуре ~30 К и указывает на её перегрев (на 260-270 К), обеспечивая неустойчивость, изменчивость и нестабильность ее параметров, особенно при спиновой конверсии в неоднородном магнитном поле в присутствии, например, триплетного кислорода [47].

Тогда большой (до сотен нанометров) эффективный диаметр нанообъектов, обнаруженных Коноваловым с сотр. [3-6] в растворах низких концентраций энергетически выгоднее формировать в виде льдоподобных, водородосвязанных каркасов пара-изомеров H₂O [46]. Об этом свидетельствует уменьшение интенсивности линий пара-H₂O в спектре при гидратации белков [27] по мере их выбывания в связанное состояние с одновременным повышением числа орто-изомеров пропорционально интенсивности спектральной линии [27], которые могут локализоваться в каналах-конейнерах. Известно, что аномально большая электропроводность гексагонального льда обусловлена высокой подвижностью ионов H₃O⁺ в каналах кристаллической структуры льда I_h [37]. Физика явления обоснована квантовыми различиями иона H₃O⁺ (по «конфигурации максимальной плотности электронов») от иона OH⁻ [37]. Степень этих отличий оставляют иону H₃O⁺ единственную траекторию «движения по пустотам структур гексагонального льда, в то время как ион OH⁻ движется по каркасу водородных связей» [37] с меньшей скоростью. При этом двукратная разность подвижности ионов H₃O⁺ и OH⁻ может приводить к локальному разделению зарядов и формированию ζ-потенциала в окрестности концентраций с экстремум электропроводности [3-6]. Такой же механизм повышения электропроводности в растворах Коновалова возможен с увеличением диаметра наноассоциатов и протяженности каналов с льдоподобной структурой. При этом вращение орто-изомеров H₂O и увеличение их числа [27] способствует снижению гидродинамической вязкости раствора [24] и поверхностного натяжения при тех же концентрациях (см. рис. 12 из обзора Коновалова [48]).



Рис. 3. Спектры комбинационного рассеяния в воде при комнатной температуре, записанные с интервалом 10 секунд

Кроме того, увеличение числа орто-изомеров H_2O при гидратации белков, измеренное нами ранее [27] по росту интенсивности линий, может обеспечить повышение биологической активности [1,3], которая коррелирует с увеличением диаметра наноассоциатов Коновалова [48], поскольку только вращающиеся мономеры H_2O проходят ключ-замок аквапоринового канала мембран [14,49] и, следовательно, в большей степени определяют биоактивность клетки, чем пара- H_2O , часть из которых не вращается и может быть связана в комплексы с другими молекулами.

Таким образом, учёт отличий спиновых изомеров воды, позволяет обосновать наблюдаемую Коноваловым корреляцию электропроводности и диаметра нанообъектов в растворах. Более того, физически ясно, что квантовые отличия двух спиновых изомеров H_2O и их соотношение должны проявляться в свойствах воды (растворов) после конденсации орто- и пара- H_2O по мере их спиновой конверсии при охлаждении воды и льда, поскольку в окрестности абсолютного нуля должны оставаться только пара-изомеры H_2O [47]. Известно, что спиновая конверсия относится к сверхтонким взаимодействиям с энергией $\sim 10^{-8}$ эВ ($\sim 10^{-5}$ cm^{-1} или $\sim 10^{-27}$ Дж) [10-14], что по порядку величины совпадает с зеемановским расщеплением в магнитном поле Земли. Это даёт основание ожидать проявление восприимчивости воды и водных растворов к слабым (сравнимым с магнитным полем Земли) электромагнитным воздействиям, например, к микроволновой (СВЧ) и ВЧ накачке смешанных состояний [12] (см. ниже) или конверсии спина в магнитном поле триплетного кислорода [47].

Отсюда следует, что экранировка ЭМП, например, в пермалловом контейнере [3-6,48] также будет влиять на свойства воды, поскольку снимается резонансная накачка смешанных квантовых состояний орто-пара изомеров H_2O [49] в широком (от кГц до ТГц) диапазоне частот. Такая экранировка от действия «слабых» ЭМП естественной (магнитные бури, колебания ионосферы и пр.) и антропогенной природы снижает вероятность орто-пара конверсии и формирование льдоподобных каркасов гидратации, что проявляется в подавлении эффекта Коновалова [48]. Это весьма значимый результат, который показывает, что амплитуда спонтанных квантовых флуктуаций спиновых состояний изомеров H_2O совместно с их термоиндуцированными столкновениями недо-

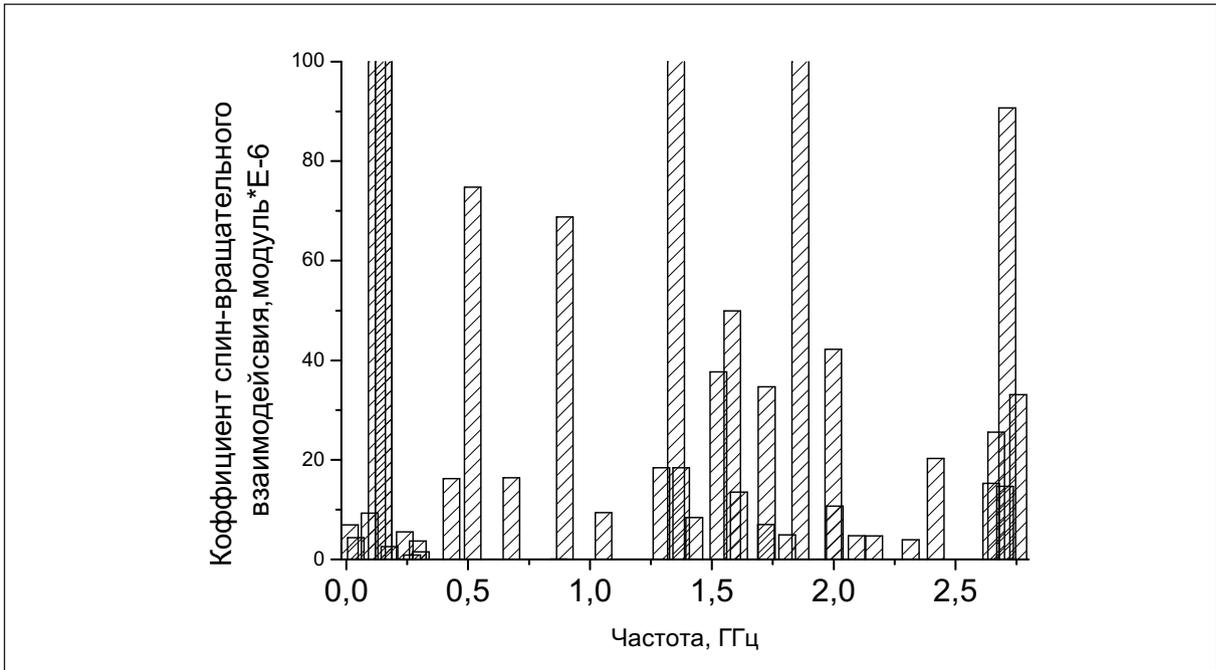


Рис. 4. Зависимость коэффициента спин-вращательного взаимодействия от частоты резонансов смешанных квантовых состояний в диапазоне от килогерц до гигагерц

статочны для обеспечения необходимой скорости конверсии в условиях экранирования (отсутствии «накачки» смешанных квантовых состояний «слабыми, шумовыми» ЭМП).

5. ВЧ и СВЧ резонансы смешанных квантовых состояний H_2O

Как было показано ранее [11-14, 50] энергетические уровни вращательных переходов спин-изомеров H_2O расположены весьма близко и образуют смешанные квантовые состояния с возможностью орто-пара конверсии. Так авторам работы [12] удалось управлять скоростью спиновой конверсии посредством сближения этих уровней в модулированном электрическом поле, используя эффект Штарка. В работе [51] провели анализ вращательных спектров изомеров H_2O и получили таблицы коэффициента спин-вращательного взаимодействия на частотах, которые определяются энергетическим зазором между близкими уровнями в смешанных квантовых состояниях орто и пара H_2O .

На рис.4 представлена зависимость этого коэффициента от частоты в ВЧ - СВЧ диапазоне (от кГц до нескольких ГГц). Из рисунка видно, что резонансную накачку [12] смешанных квантовых состояний, индуцирующую орто-пара конверсию, можно осуществлять во всем этом диапазоне. Здесь необходимо учесть, что эффект Штарка [12] будет смещать эти уровни и обеспечивать более полное заполнение некоторых «окон прозрачности». Следует отметить, что рабочие частоты средств мобильной связи 0.9 и 1.8 ГГц, а также бытовых СВЧ печей (2.45 ГГц) попадают в полосы резонансной накачки смешанных квантовых состояний орто и пара H_2O . Достаточно высокая спектральная плотность распределения (рис.4) коэффициента и малая энергия ($\sim 10^{-27}$ Дж) спин-вращательного сверхтонкого взаимодействия [10-14] даёт основание предполагать, что отклик на облучение растворов [52] слабыми низкочастотными ЭМП, которые помещались в соленоид в пермалловом контейнере, скорее всего, был обусловлен резонансной накачкой смешанных квантовых состояний. Несомненно, результаты этого эксперимента в условиях глубокой экранировки от действия внешних ЭМП на водный раствор в пермалловом контейнере указывают на присутствие в воде большого количества «резонансных приёмни-

ков», чтобы суммировать излучение соленоида. Из вышеизложенного следует, что такими «приёмниками» могут быть только молекулы-мономеры [53], а именно, орто-пара спин-изомеры H_2O , с узкими резонансами, поскольку сверхнизкая концентрация растворов варьировалась в пределах $10^{-12} - 10^{-18}$ М [52].

Здесь следует отметить, что многолетняя дискуссия [7, 8, 11-15, 18, 38, 41 и ссылки в них] о наличии мономеров H_2O в воде не отрицала их присутствия, а касалась только их количества (от 1% [54,55] до 10.5 % [56]) и места локализации [18, 26, 27, 36, 38]. Наша модель захвата и локализации мономеров в полостях льдоподобных каркасов хорошо согласуется с интенсивно обсуждавшейся ранее [38,57] моделью гидрофобных мономеров H_2O как самостоятельных химических единиц внутри льдоподобных каналов, в которых *все* водородные связи (ковалентные по своему характеру со свойством насыщения [57]) заняты. Недавно [58], на основе анализа данных, полученных современными методами терагерцовой спектроскопии воды и тяжелой воды (D_2O), была сделана оценка количества свободных мономеров от 10% (при температуре 5 °С) до 18% (при 50 °С).

6. Выводы

Проведенный анализ показывает, что свободные мономеры воды совместно с молекулами, связанными в динамические структуры, определяют её физические свойства. При этом метаболизм (транспорт воды через водные каналы аквапориновых белков мембран) в биологических объектах обусловлен преимущественно мономерами H_2O [59,60]. Однако заметим, что в работах [58-60], как и, в ранее упоминавшихся [18, 38, 52, 54-57], не были приняты во внимание квантовые различия спиновых изомеров H_2O и D_2O [11-15,], а также температурная эволюция их орто/пара отношения [17], которые позволяют обосновать ВЧ и СВЧ накачку смешанных квантовых состояний слабыми ЭМП.

Несомненно, предложенный подход требует разработки новых неинвазивных методов изучения роли сверхтонких взаимодействий, спиновой селективности при образовании комплексов и эволюции орто/пара отношения спиновых изомеров H_2O в физике водных растворов низких концентраций и взаимодействия с ЭМП. Из проведенного рассмотрения следует, что орто/пара отношение спиновых изомеров H_2O является ключевым фактором свойств воды, что согласуется с первым, насколько нам известно, теоретическим анализом роли магнитных взаимодействий и спиновых частиц (изомеров орто- H_2O) в формировании комплексов большого диаметра [61-62].

Работа выполнялась при частичной поддержке гранта РФФИ 11-02-00034а и гранта НШ № 214.2012.2, программы РАН «Спектроскопия и стандарты частоты», а также программы Президиума РАН №28 «Физика, химия и биология воды»

Литература

1. Бурлакова Е.Б., Конрадов А.А., Худяков И.В. // Изв. АН СССР. Сер. Биол. 1992. №2. С. 184-193; Белов В.В., Мальцева Е.Л., Пальмина Н.П., Бурлакова Е.Б. // ДАН. 2004. Т. 399. №4. С. 1-4.
2. Зубарева Г.М., Каргаполов А.В., Ягузинский Л.С. // Биофизика. Флуктуации коэффициентов пропускания воды и водных растворов солей в инфракрасной области спектра. 2003. Т.48. №2. С. 197-200.
3. Пальмина Н.П., Часовская Т.Е., Рыжкина И.С., Муртазина Л.И., академик Коновалов А.И. Водные растворы фенозана калия: влияние на структуру биологических мембран и электропроводность. // ДАН. 2009. Т. 429. №1. С. 128-131.
4. Коновалов А.И., Рыжкина И.С., Муртазина Л.И. и др. // Изв. АН. Сер. хим. 2008. № 6. С. 1207-1214.
5. Рыжкина И.С., Муртазина Л.И., Киселева Ю.В., Коновалов А.И. // ДАН. 2009. Т. 428. № 4. С. 487-491.

6. Рыжжина И.С., Киселева Ю.В., Тимошева А.П., Сафиуллин Р.А., Кадилов М.К., Валитова Ю.Н., академик Коновалов А.И. // ДАН. 2012. Т. 447. №1. С. 1-7. Труды международного конгресса «Слабые поля в биологии и медицине», Санкт-Петербург, июль 2012г.
7. Chaplin M., Water Structure and Science, www.btinternet.com/martin.chaplin/phase.html.
8. Френкель Я.И. Кинетическая теория жидкости, Л.: Из-во АН СССР. 1945.
9. Landau L.D., Lifshitz E.M. // Course of Theoretical Physics. vol. 3. Quantum Mechanics [third edition]. Oxford. Pergamon Press, 1980.
10. Tikhonov V.I., and Volkov A.A. // Separation of Water into Its Ortho and Para Isomers. Science. 2002. V. **296**. PP. 2363-2366.
11. Pershin S.M. // Coincidence of rotational energy of ortho-para molecules and translation energy near specific temperatures in water and ice. Phys. of Wave Phenomena, 2008, V. 16. # 1. PP. 15-25.
12. Nagels B., Calas N., Roozmond D.A., Hermans L.J.F., and Chapovsky P.L. // Level-Crossing Resonances in Nuclear Spin Conversion of Molecules, Phys. Rev. Lett., **77**(23), 4732-4735, (1996).
13. Pershin S. M. // Ortho/Para H₂O Conversion in Water and a Jump in Fluidity of Erythrocytes through a Microcapillary at the Temperature 36.6 +/- 0.3⁰C. Phys. of Wave Phenomena. 2009. V. 17. #4. PP. 241-250
14. Першин С.М. // Влияние квантовых отличий орто и пара спин-изомеров H₂O на свойства воды: биофизический аспект, Биофизика, **58**(5), 910-918, (2013).
15. S.M. Pershin, A.F. Bunkin, N.V. Anisimov, and Yu. A. Pirogov, Water Enrichment by H₂O *ortho*-Isomer: Four-Photon and NMR Spectroscopy, *Laser Physics*, **19**(3), 410–413 (2009)

RF&MW RESONANCES OF H₂O ORTHO-PARA SPIN-ISOMERS MIXING QUANTUM STATES AS A FACTOR OF INTERCONVERSION AND IMPACT ON BIO-OBJECTS

S.M. Pershin

PWave Research Center of Prokhorov general physics institute, Russian Academy of Sciences

pershin@kapella.gpi.ru

Received 24.03.2015

An increasing of electroconductivity simultaneously with increasing of nanoobjects diameter (decreasing of diffusion coefficient) in highly delution of aqueous solutions have been observed by A. Konovalov with colleagues recently. It was established that aqueous solutions which were subjected to shielding from electromagnetic fields (EMF) by permalloy container do not manifest these anomalous properties. The novel concept of water properties was proposed to understand these phenomena and observing paradox- electroconductivity growing vs diffusion coefficient decreasing. This concept includes the H₂O ortho-para spin-isomers quantum differences: spin-selectivity in formation of hydration layers of nanoobject and spontaneous transition to ice-like structure in water clusters, quantum beating and spin-conversion which can be induced by week EMF as a resonance pump in RF and MW range – which did not any time before. This concept is used to explain the transportation features of H₂O monomers through the cell membrane aquaporine channels. It is shown that the dipole key inside channel operates like valve which permeates only rotating H₂O. This mode is more convenient to H₂O ortho-isomer because it has no zero rotating quantum state. Moreover there are close locating quantum levels of ortho-para H₂O that are named as quantum mixing states to induce spin conversion process which can be pump by resonance EMF in RF and MW frequency range also.

Информация и правила для авторов

Общие положения

Журнал «Наноструктуры. Математическая физика и моделирование» (сокращенно: НМФМ) публикуется с 2009 года и является периодическим научным изданием. Электронная версия журнала размещается на сайте <http://www.nano-journal.ru>. Основная цель издания: представление новых теоретических и вычислительных методов моделирования наноструктур и мягкой материи, общих подходов в исследовании мезосистем, а также ключевых экспериментальных результатов в данной области и связанных с этим проблем математической физики.

Журнал НМФМ имеет междисциплинарный характер и в силу этого несет определенную образовательную направленность, а не только узко научную. Работы, представляемые в журнал, должны содержать вводные сведения, которые обеспечат понимание постановок задач и восприятие результатов не только прямыми специалистами. Определения понятий, объяснение обозначений и терминов, оценки характерных параметров, теоретические предпосылки и идеи, используемые методы, и т.п., должны быть кратко объяснены в тексте статьи, имея в виду читателей, специализирующихся в иных направлениях. Должны быть описаны базовые математические модели и уравнения. Во Введении и в последующих разделах очерчивается стратегия и основные трудности, это увязывается с используемыми моделями. Структура статьи ориентируется на прояснение общей логики и методики исследования, содержит резюмирующие выводы. В тексте должны быть рассмотрены характерные примеры (хотя бы, методические), ясно иллюстрирующие предлагаемые алгоритмы.

Журнал публикует научные обзоры, исследовательские статьи и краткие научные сообщения, а также избранные аналитические и информационно-образовательные материалы, тексты докладов и циклов лекций, прочитанных в университетах, научных центрах, на школах-семинарах, конференциях, нигде ранее не публиковавшиеся и не принятые к публикации в других изданиях. Язык публикации в журнале НМФМ, как правило, русский. Работы, представляемые в журнал, не могут иметь научно-популярный или компилятивный характер. Все статьи рецензируются и могут быть отклонены редколлегией журнала. В случае принятия работы к печати ее авторы передают издателью журнала НМФМ право на разовую безвозмездную публикацию текста и его размещение в электронной версии на сайте журнала. Перевод опубликованных в журнале статей на другие языки может осуществляться только с разрешения и при участии авторов.

Порядок представления статей

- В редакцию изначально представляются:
 - файл статьи, файлы с иллюстрациями;
 - сопроводительное письмо, можно в электронной форме, содержащее сведения об объеме статьи и обо всех авторах (фамилии, имена, отчества, полные названия мест работы, почтовый адрес с индексом, номер контактного телефона с кодом города, электронный адрес автора, ответственного за переписку с редакцией); предпочтительно, чтобы это письмо было выполнено на бланке учреждения, в котором работает кто-то из авторов, было заверено печатью и содержало утверждение о возможности открытого опубликования статьи;
 - файл с переводом на английский язык названия статьи, фамилий и инициалов авторов, аннотации, ключевых слов.
- Авторские файлы могут быть присланы на электронный адрес: papers@nano-journal.ru; (резервный адрес в случаях затруднений с пересылкой: nano@miem.edu.ru) или переданы в редакцию на любом электронном носителе. Авторы получают из редакции подтверждение о получении их материалов.
- Телефон (факс) редакции: +7 (495) 916-8876. Адрес редакции: Москва 109028, Б. Трехсвятительский пер., 3/12, Московский институт электроники и математики (МИЭМ), комн. 449.

Общие требования к представляемым файлам

- Допускается использование текстовых редакторов WORD и LATEX. К рабочим файлам должна быть приложена их pdf-копия. В названии файлов используется латинский алфавит, пробелы заменяются знаком `_`. Шапка статьи содержит название, инициалы и фамилии авторов, место работы, электронный адрес, краткую аннотацию, ключевые слова. В аннотации не следует использовать формулы и ссылки на текст работы или список литературы; в конце она должна содержать индекс УДК (к английской версии аннотации можно добавить индексы зарубежных рубрикаторов).
- Объем кратких сообщений 4-8 страниц, исследовательских статей, как правило, до 20 страниц, а обзоров – более 20 страниц. Верхняя граница согласуется с редколлегией. При подсчете объема нужно ориентироваться на страницы формата А4, шрифт 12, знаков в строке 80, интервалов между строками 1.
- Авторы не должны злоупотреблять сокращениями, составленными из заглавных начальных букв терминов. Предпочтительней каждый раз использовать полное наименование объекта. Возможно использование только устоявшихся аббревиатур.

Требования к файлам Word

- Рекомендуемый шрифт – Times New Roman.
- Строки в пределах абзаца не должны разделяться символом возврата каретки (Enter).
- Нельзя использовать автоматическое создание сносок, автоматический перенос или автоматический запрет переносов, создание списков, автоматический отступ и т.п.
- Ссылки на список литературы даются цифрами в квадратных скобках: [1], [5,6,7], [1-9].
- Все без исключения формулы и обозначения размерности, даже состоящие из одной латинской буквы, и в тексте и вынесенные в отдельную строку, всегда набираются в формульном редакторе и никогда – в обычном текстовом редакторе.

- При создании таблицы рекомендуется использовать возможности Word или MS Excel. Таблицы, набранные вручную (с помощью большого числа пробелов), не принимаются.

Требования к иллюстрациям

- Иллюстрации представляются в отдельных файлах, черно-белыми. Они должны иметь разрешение не менее 600 dpi.
- Форматы файлов – TIFF, EPS, PSD, JPEG.

Требования к списку литературы

- Ф.И.О. авторов или редакторов выделяются курсивом.
- Для статей приводится название. Названия отделяются от выходных данных знаком //. Расположение выходных данных указано на образце ниже. Номер тома выделяется жирным шрифтом, номер выпуска дается в скобках. Указываются номера первой и последней страниц статьи, либо уникальный номер статьи и ее объем. Для книг желательно указывать их объем. Если известна ссылка на электронный архив или сайт, то ее желательно указать.

Фамилия И.О. Название статьи // Назв. журн., 2000, **1** (1), 1-6.

Family F.M. and Family F. Title of the paper // Name of the Journal, 2006, **73**, 165313, 9 pp.

Фамилия И.О., Фамилия И.О. Название книги // Наука, С.-П., 1999, 176 стр.

Family F.M. Title of the paper // In book: Family F.M. (et al. eds), Title of the collection, Publisher, Boston, 2005, 9-24.

Family F.M. (ed.), Title of the collection // Publisher, N.Y., 2005, 324 pp.

Фамилия И.О. Название доклада // Доклад на конференции «Название конференции (место и дата проведения)»; ссылка на электронный ресурс.

Наноструктуры. Математическая физика и моделирование

Журнал зарегистрирован

в Министерстве РФ по делам печати,
телерадиовещания и средств массовых коммуникаций.

Свидетельство о регистрации

ПИ № ФС77-34934 от 29 декабря 2008 г.

Учредители

Московский государственный институт электроники и математики (МИЭМ),

ФГУ ГНИИ ИТТ «Информика»,

Европейский центр по качеству

Председатель Совета Учредителей

Тихонов А.Н.

Издатель

Европейский центр по качеству

ПОДПИСКА НА ЖУРНАЛ НМФМ

На первое полугодие 2015 г. подписаться на журнал можно в любом отделении связи по каталогу
Агентства Роспечать «Журналы России»,
рубрика «Физико-математические науки», подписной индекс 70017.

Редакция предлагает подписчикам возможность безвозмездно получить подборку прошлых
выпусков журнала. Пришлите на электронный адрес nano@miem.edu.ru (или на почтовый адрес:
109028, Москва, Б. Трехсвятительский пер., 3, к. 452, редакция журнала НМФМ)
копию подписной квитанции, а также адрес для отсылки выпусков.