

ЕРЕВАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Цатурян Айк Левонович

ГЕТЕРОГЕННОСТЬ НУКЛЕОТИДНОЙ
ПОСЛЕДОВАТЕЛЬНОСТИ В
КОНФОРМАЦИОННЫХ ПЕРЕХОДАХ
НУКЛЕИНОВЫХ КИСЛОТ

03.00.02 Биофизика и биоинформатика

ДИССЕРТАЦИЯ

на соискание ученой степени

кандидата физико-математических наук

Научный руководитель

д. ф.-м. н., доцент

Е.Ш. Мамасакхлисов

Ереван 2017

Содержание

Введение	4
Глава 1. Литературный обзор	10
1.1. Структура и биологические функции нуклеиновых кислот	10
1.2. Типы РНК	11
1.3. Вторичная структура РНК	12
1.4. Третичная структура нуклеиновых кислот	15
1.5. Термодинамика нуклеиновых кислот	17
1.6. Алгоритмы динамического программирования	24
1.7. Существующие теории переходов во вторичной структуре ДНК ...	28
Глава 2. Конформационные переходы в оц-РНК с двухбуквенным и четырёхбуквенным алфавитом в последовательности	38
2.1. Двухлитерная модель	40
2.2. Четырёхлитерная модель	50
2.3. Метод ограниченного отжига	57
2.4. Низкотемпературная поведение	61
Глава 3. Коллапс и гибридизация оц-РНК: неравновесное стационарное состояние	69
3.1. Модель оц-РНК, как случайно разветвленного полимера	71
3.2. Свободная энергия и ненулевое количество реплик	74
3.3. Параметризация матрицы Паризи для $n \neq 0$	76
3.4. Фазовый переход и удельная теплоемкость	78
Глава 4. Влияние интеркалирующих лигандов на изотерму гибридации ДНК – чипа	87
4.1. Неконкурентная гибридизация	89
4.2. Изотермы адсорбции и гибридизации	91
4.3. Конкурентная гибридизация на поверхности	92

Выводы	99
Литература	100

Введение

Актуальность работы. Одной из основных задач молекулярной биофизики является определение физических закономерностей строения и биологической функции одно- и двунитевых нуклеиновых кислот. Общеизвестно, что биологическая функция биополимеров определяется их пространственной структурой. В связи с этим важным является определение основных факторов и закономерностей, влияющих на конформации и конформационные переходы. Одним из таких факторов является влияние нуклеотидной последовательности на образование вторичной структуры нуклеиновых кислот и взаимное влияние вторичной и третичной структур нуклеиновых кислот. Кроме того, остается ряд открытых вопросов, связанных с влиянием различных низкомолекулярных соединений (лигандов) на структуру и конформационные переходы в нуклеиновых кислотах.

Исследование данных вопросов представляет интерес также и с точки зрения биоинформатики. Алгоритмы расчета термодинамических параметров и предсказания оптимальной вторичной структуры однонитевой РНК в настоящее время широко используются в биологических исследованиях. Корректный учет эффекта гетерогенности последовательности существенно улучшит существующие алгоритмы и будет способствовать разработке новых подходов к данной проблеме.

Кроме того, исследование термодинамики и кинетики гибридизации нуклеиновых кислот представляет существенный интерес для создания биологических сенсоров и способствует определению оптимальных условий работы подобного рода устройств.

В свете изложенного, актуальность настоящей работы определяется разработкой новых подходов к исследованию эффектов

нуклеотидной последовательности при плавлении и гибридизации нуклеиновых кислот в растворе и на границах раздел раствор – твердое тело.

Целью работы является:

1. Построение теории, описывающей термодинамические свойства однонитчатой РНК со случайной последовательностью нуклеотидов.
2. Расчет термодинамических характеристик однонитчатой РНК.
3. Построение теории, описывающей взаимное влияние пространственной и вторичной структур однонитчатой РНК.
4. Изучение неравновесных эффектов, возникающих при коллапсе и гибридизации однонитчатой РНК.
5. Расчет равновесных характеристик гибридизации ДНК – зондов иммобилизованных на границе раздела твердое тело – раствор с ДНК – мишенями в растворе в присутствии лигандов.
6. Изучение эффекта интеркалирующих лигандов на чувствительность и избирательность ДНК – сенсора.

Научная новизна заключается в построении модели, описывающей образование вторичной структуры однонитчатой РНК, основанная на методе отжига с ограничениями для случаев двух и четырехбуквенного алфавита. Проведено сопоставление и показано количественное согласие с результатами численного моделирования. Получена температурная зависимость таких характеристик денатурации однонитчатой РНК, как свободная энергия, степень спиральности, энтропия и теплоемкость. Для случайной последовательности нуклеотидов как двух, так и четырех типов показана возможность наличия двух, или трех структурных переходов. Показана возможная связь низкотемпературного структурного перехода с холодной денатурацией РНК, наблюдаемой на эксперименте. Показано, что данный эффект имеет место как для двух, так и для четырехбуквенного алфавита нуклеотидов при соответствующем выборе параметров взаимодействия.

Также в работе изучено влияние различных характерных времен релаксации пространственной и вторичной структур однонитевой РНК на ее структуру и термодинамические свойства. Предложен общий формализм, позволяющий описывать стационарное неравновесное состояние таких систем. В частности, показано, что при коллапсе однонитевой РНК имеет место неравновесный фазовый переход второго рода. Определены и рассчитаны энтропия и теплоемкость системы в стационарном неравновесном состоянии.

Основные положения, выносимые на защиту.

1. Однонитевая РНК со случайной двухбуквенной последовательностью нуклеотидов при определенном выборе параметров взаимодействия демонстрирует два пика на температурной зависимости удельной теплоемкости системы. Подобное поведение указывает на наличие двух структурных переходов.
2. В узком диапазоне параметров энергии образования пар оснований однонитевая РНК со случайной четырехбуквенной последовательностью демонстрирует три пика на температурной кривой для удельной теплоемкости.
3. Как для двух, так и для четырехбуквенного алфавита детально проанализирована холодовая денатурация РНК и показано, что это явление может быть объяснено в терминах конкуренции между энергиями образования Уотсон – Криковских пар и пар одноименных нуклеотидов
4. При коллапсе однонитевой РНК она претерпевает неравновесный фазовый переход второго порядка, обусловленный различными характерными временами релаксации пространственных степеней свободы и элементов вторичной структуры.
5. Интеркалирующие лиганды повышают чувствительность и избирательность ДНК – сенсора за счет связывания с нативными

участками гибризованных ДНК на поверхности раздела твердое тело – раствор.

Научная и практическая ценность работы обусловлена существенным значением, которое играют термодинамические и кинетические эффекты при функционировании биологических макромолекул и их комплексов. В этой связи, теоретическое исследование эффектов взаимодействия между нуклеотидами различных типов и других характеристик биологических макромолекул имеет важное значение для интерпретации экспериментальных результатов и их предсказания. При этом, понимание основных принципов, лежащих в основе самоорганизации и конформационных перестроек в биологических макромолекулах имеет большое практическое значение для решения проблем биологии и медицинских приложений. В частности, полученные результаты несомненно будут полезны при оценке оптимальных режимов работы ДНК – сенсоров. Кроме того, рассчитанные в работе термодинамические характеристики оц-РНК, несомненно, позволяет усовершенствовать существующие биоинформационные алгоритмы, используемые при расчете стабильности вторичной структуры РНК.

Аппробация работы

Материалы диссертации были представлены и докладывались на

- International Young Scientists Conference “Perspectives for Development of Molecular and Cellular Biology-3”, The Institute of Molecular Biology NAS RA, Yerevan, Armenia, September 26-29, 2012.
- Disordered and Ordered Materials Analysis and Characterization - DOM2015, International Symposium and Young Scientist School 24-30 September 2015, Yerevan, Armenia.

Материалы диссертации также докладывались на научных семинарах

кафедры молекулярной физики Ереванского Государственного Университета и на семинарах Учебно-научный центр повышения квалификации ЕГУ по теме «Алгебраические и геометрические исследования по физике конденсированных сред».

Публикации.

По теме диссертации опубликовано 4 работы (4 статьи в реферируемых журналах).

Структура диссертации.

Диссертация состоит из Введения, четырех глав и Выводов (109 страниц текста). Она содержит 39 рисунка и список литературы из 148 наименований. Во введении сформулирована цель работы, показана научная новизна и практическая значимость результатов, перечисляются основные положения представленные к защите. Первая глава посвящена обзору структуры, термодинамики и биологических свойств нуклеиновых кислот. Представлены основные свойства ставших классическими моделей ДНК. Также, литературный обзор содержит описание основных моделей вторичной структуры РНК и динамических алгоритмов расчета ее термодинамических параметров. Вторая глава посвящена модели однонитевой РНК, учитывающую гетерогенность нуклеотидной последовательности. В рамках этой модели рассчитаны основные термодинамические параметры системы. В основу расчетов термодинамических функций взят метод отжига с ограничением. Третья глава диссертационной работы посвящена взаимному влиянию вторичной и пространственной влиянию структур, возникающих в однонитевой РНК. Изучено стационарное неравновесное состояние, возникающее на временах наблюдения, промежуточных для характерных времен релаксации вторичной и третичной структур РНК. Четвертая глава посвящена гибридизации мишеней ДНК с ДНК – зондами, иммобилизованными на границе раздела твердое тело – раствор. В ней представлены результаты,

описывающие влияние интеркалирующих лигандов на чувствительность и избирательность ДНК – сенсоров. Работа завершается Выводами.

Гл.1. Литературный обзор

1.1. Структура и биологические функции нуклеиновых кислот

Структура нуклеиновых кислот

Нуклеиновые кислоты являются линейными полимерами, мономеры которых - нуклеотиды. Нуклеотид состоит из сахарного кольца, фосфатной группы и азотистого основания. Основная цепь нуклеиновой кислоты состоит из сахарных колец связанных между собой фосфатными группами и каждое сахарное кольцо связано с одним из четырех видов азотистых оснований. Фосфатные группы соединяют 5' углерод одного сахарного кольца с 3' углеродом другого, образуя сахарофосфатный остов. Сахарные кольца бывают двух видов: рибоза и дезоксирибоза. И поэтому существует два типа нуклеиновых кислот: дезоксирибонуклеиновая кислота (ДНК) и рибонуклеиновая кислота (РНК). Между ними три главных отличия. Во-первых, химически ДНК отличается от РНК тем, что ОН группа сахара РНК в ДНК заменена на протон Н. Во-вторых, в РНК вместо основания тимин, который есть в ДНК, появляется урацил. То есть, ДНК состоит из азотистых оснований гуанина, цитозина, аденина и тимина (G, C, A, T), а РНК – из гуанина, цитозина, аденина и урацила (G, C, A, U). В-третьих, в природе, как правило, встречается одноцепочечная РНК, а ДНК состоит из двух цепей. Это приводит к тому, что структуры молекул сильно отличаются. ДНК образует двойную спираль из двух комплементарных ниток, связанных водородными связями, а РНК в основном – это одна нить, и лишь некоторые участки могут образовывать спирали (на этом основана вторичная структура РНК). Поэтому, структуры РНК больше похожи на глобулярные структуры белков, чем на двойную спираль ДНК и имеют третичную структуру. Некоторые третичные структуры РНК хорошо изучены. Главная роль ДНК состоит в сохранении генетической информации. А роль белков состоит в обеспечении биохимического катализа. Таким образом, получается, что РНК

как структурно, так и функционально находится между ДНК и протеинами. Поэтому мы с уверенностью можем сказать, что РНК очень важная и своеобразная молекула.

1.2. Типы РНК

Транспортная РНК (тРНК)

Главная цель этого типа рибонуклеиновых кислот состоит в доставке аминокислот в рибосому в процессе синтеза белка. Общее число нуклеотидов транспортной РНК равно 76 [1, 2]. Для каждой аминокислоты существует «своя» тРНК. В средней части центральной петли находится антикодон, который состоит из трех нуклеотидов. Кодон мРНК и антикодон соответствующей тРНК одинаковы. Вторичную структуру тРНК называют «клеверным листом» и она хорошо изучена (Рис. 1.1А). Третичная структура тРНК похожа на букву «Г» (Рис. 1.1Б).

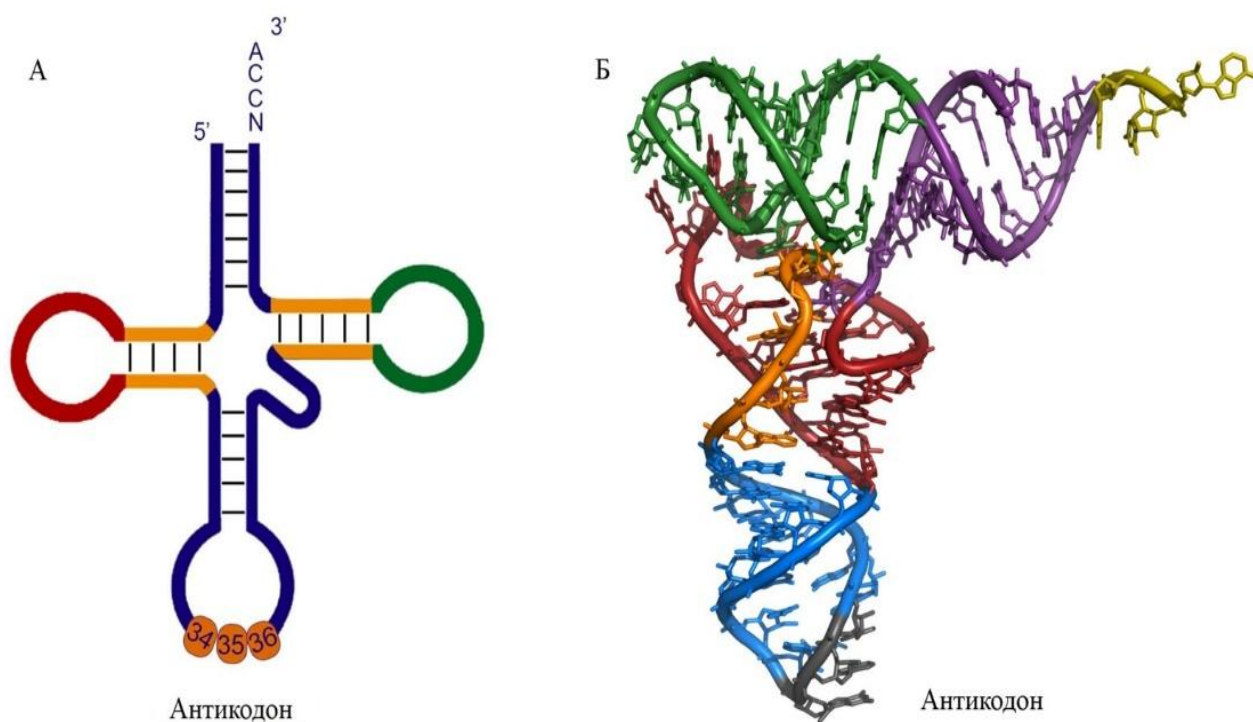


Рис. 1.1. Вторичная и третичная структуры транспортной РНК

Матричная РНК (мРНК)

Матричная РНК представляет собой точную копию некоторого участка одной из нитей ДНК и содержит информацию о белке, который будет синтезироваться в рибосоме. Ее центральная часть кодирует белок. Матричная РНК состоит из нескольких тысяч нуклеотидов.

Рибосомная РНК (рРНК)

Рибосомные РНК составляют каркас рибосомы, в которой проходит синтез белка. Рибосома содержит центры связывания для мРНК и тРНК. Это и является ее главной ролью. Диаметр рибосомы равен 250 Å и она состоит из двух субъединиц, каждая из которых в свою очередь состоит из трех рРНК и около 56 различных белков [3-5]. Главная цель рибосомы заключается в выполнении синтеза белка - одного из самых важных процессов в клетке. В процессе синтеза она считывает информацию из мРНК и следуя этой информации связывает соответствующие аминокислоты, которые доставляются в рибосому с помощью тРНК. Следовательно, мы можем сказать, что тРНК и мРНК играют очень важную роль в правильном функционировании рибосомы и, как следствие, невозможно синтезировать белок без этих молекул. Рибосомные РНК представляют собой длинные последовательности. Существует много баз данных, которые позволяют нам понять их структурные модели [6-8].

1.3. Вторичная структура РНК

Если в молекуле РНК есть два комплементарных участка, то они могут образовать спиральную структуру. Спиральные участки появляются вследствие возникновения водородных связей между азотистыми основаниями. Возможны три вида водородных связей: между азотистыми основаниями С-G, А-U и G-U. Пара G-U менее стабильна, чем пара С-G или

пара A-U. Как правило, спиральные участки состоят, по крайней мере, из двух пар, так как одна изолированная пара довольно нестабильна. Максимальное число пар в спиральных участках не превышает десяти. Между парами существуют стэкинг взаимодействия, которые имеют большой вклад в стабилизации спиральной структуры. При стэкинге, соседние азотистые основания примерно параллельны друг другу. Для вычисления свободной энергии спиральной структуры часто используется модель ближайших соседей. Это вклад в свободную энергию от каждой из соседних пар оснований. Если у нас короткие цепи, то используя разные методы мы можем вычислить изменения энергии и энтропии спиральных участков.

На Рис. 1.2 показаны несколько структур, которые могут появиться между нитями в однонитевой РНК. Шпилька (hairpin loop) соединяет две комплементарные последовательности друг с другом, образуя на конце неспаренный участок — петлю. Петля, возникшая между двумя спиральными участками, называется внутренней (internal loop). Множественными петлями (multi-branched loop) называются те петли, которые встречаются между тремя или больше спиральными участками. Вспучивание (bulge loop), двунитевой участок (stem) и псевдоузел (pseudoknot) тоже встречаются в однонитевой РНК. О псевдоузлах мы поговорим позже. Свободная энергия нескольких видов петель вычислена экспериментально, но, как правило, параметры спиральных структур известны намного точнее, чем параметры петель [10]. Например, у нас нет никаких термодинамических данных о множественных петлях. Мы предполагаем, что свободная энергия петли не зависит от нуклеотидной последовательности. Это зависит от числа непарных нуклеотидов в петле. Исключением являются тетрапетли (tetraloops). Тетрапетли это специальные последовательности, которые состоят из четырех однонитевых участков. Благодаря им термодинамическая стабильность растет за счет

взаимодействия между непарными основаниями в петле шпильки, длиной в 4 нуклеотида, где они часто встречаются. В алгоритме прогнозирования вторичной структуры мы должны посчитать свободные энергии для каждой возможной структуры. После этого мы должны сравнить свободные энергии для всех этих структур. Вместо тех термодинамических параметров, которые не были напрямую измерены, мы можем взять их разумные и приблизительные значения (энтропия образования петли, свободная энергия образования пары оснований). Свободная энергия вторичной структуры всей цепи будет определена с помощью свободных энергии отдельных участков цепи.

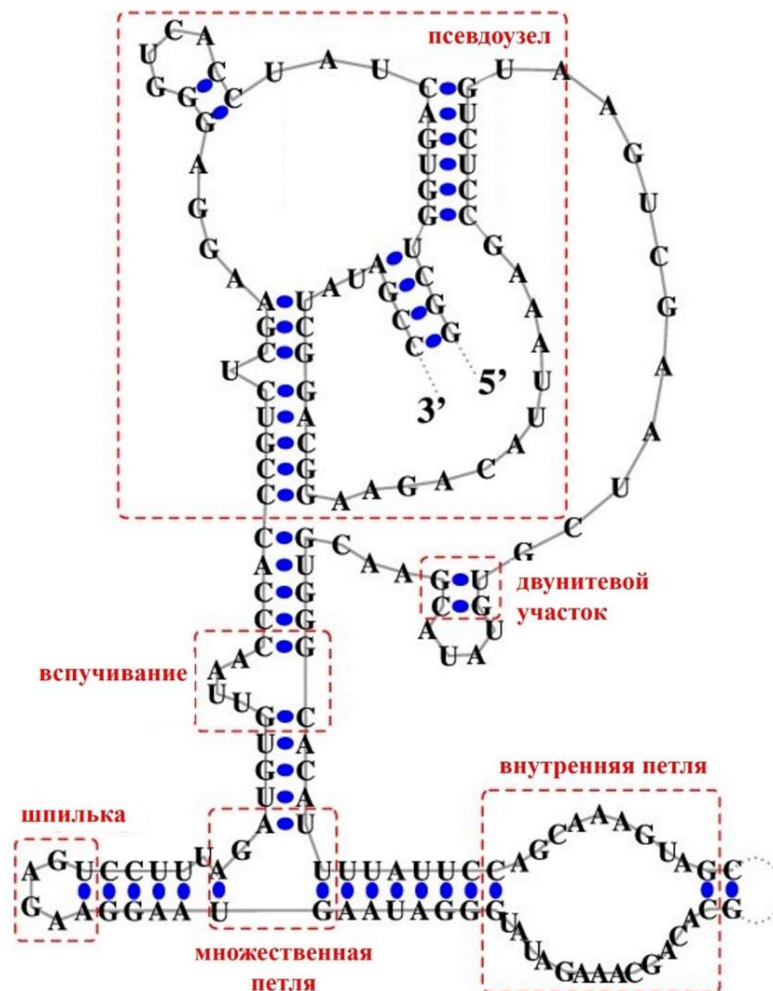


Рис. 1.2. Некоторые структуры, которые могут образоваться в одноцепочечной ДНК

1.4. Третичная структура нуклеиновых кислот

Методы и алгоритмы определения вторичной структуры нуклеиновых кислот развиваются намного быстрее, чем методы и алгоритмы определения третичной структуры. До недавнего времени имелось очень мало экспериментальной информации о третичной структуре. В этом обзоре мы больше будем говорить о вторичной структуре. Несмотря на то, что в последнее время данных о третичной структуре стало намного больше, мы можем утверждать, что информация о вторичной структуре нуклеиновых кислот тоже очень важна. Вторичная структура это рисунок, показывающий нам связанные и несвязанные нуклеотиды в этой, конкретной структуре. Во вторичных структурах пары оснований должны удовлетворять некоторым ограничениям. Предположим у нас есть цепь состоящая из оснований пронумерованных от 1 до N . Предположим, что основания i и j комплементарны. Они смогут образовать пару, если между ними находится по крайней мере 3 нуклеотида. Иными словами, петля должна содержать 3 или более непарных нуклеотидов. Предположим, что образовались пары между нуклеотидами (i, j) и (k, l) . Эти пары смогут будут совместимы, если они смогут находится в цепи одновременно. Для этого они не должны перекрываться ($i < j < k < l$) или одна из них должна содержать другую ($i < k < l < j$). Структуры, где они перекрываются ($i < k < j < l$) называются псевдоузлами (Рис. 1.3). Большинство методов динамического программирования не учитывают существование псевдоузлов. Вторичная структура последовательности содержит информацию о связанных и несвязанных основаниях, но она не содержит информацию о третичной структуре цепи. Но если у нас есть информация о третичной структуре данной цепи, мы можем немного изменить диаграмму вторичной структуры. Например, те части цепи, которые находятся близко друг друга в третичной структуре, мы можем нарисовать возле друг друга во вторичной структуре.

Таким образом, наша диаграмма вторичной структуры последовательности будет содержать некоторую информацию о ее третичной структуре. Как правило, составляя диаграммы вторичной структуры цепи особо не учитывается третичная структура. Несмотря на это, в случае с РНК, из ее вторичной структуры мы можем получить довольно много информации о ее третичной структуре. Мы можем получить информацию о доменной структуре молекулы и взаимных расположениях ее важных частей. Следовательно, вторичная структура РНК содержит намного больше информации о форме этой молекулы, чем соответствующие диаграммы вторичных структур белков, которые являются линейными полимерами, содержащими α -спирали и β -листы.

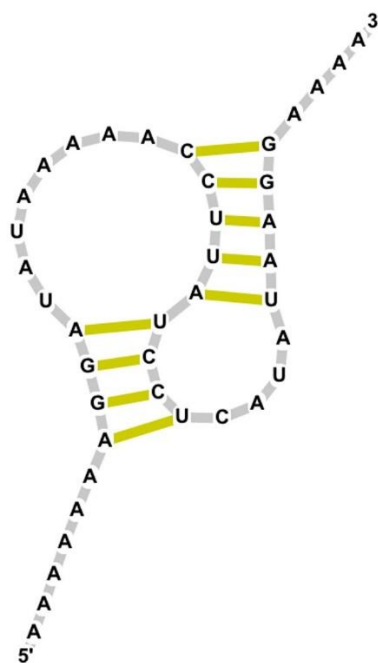


Рис. 1.3. Схематическое изображение псевдоузла

Главное преимущество вторичной структуры РНК состоит в том, что с термодинамической точки зрения спирали очень сильно связаны. Во время фолдинга РНК вначале образуется стабильная вторичная структура [11-13], а потом образуется третичная структура, которая незначительно меняет

вторичную структуру. После этого образуется третичная структура, так как молекула может согнуться вокруг некоторых частей. Взаимодействия в третичной структуре могут изменить только слабые, с точки зрения стабильности, элементы вторичной структуры. Это обусловлено тем, что их сила довольно слаба чтобы изменить вторичную структуру. Эти взаимодействия могут изменить расположение оснований в самых нестабильных спиральных. В отличие от РНК, в белках вторичная структура сама по себе довольно нестабильна. Поэтому очень сложно отделить их вторичные и третичные структуры. Как правило, мы пренебрегаем существованием псевдоузлов при подсчете параметров описывающих вторичную структуру. Существует очень много причин для этого. Одна из них состоит в том, что алгоритм, позволяющий нам предсказать вторичную структуру не может учитывать псевдоузлы. Например, в малой субъединице рРНК число неперекрывающихся спиралей намного больше числа псевдоузлов. Таким образом, мы можем получить достаточно точные результаты не учитывая вклад псевдоузлов. Но очевидно, что некоторые типы псевдоузлов довольно часто встречаются в РНК и они имеют функциональное значение. Теперь мы имеем много данных о вторичных и третичных структурах псевдоузлов [14-17]. Как результат, новые алгоритмы динамического программирования могут учитывать существование псевдоузлов [18]. Главная проблема этих алгоритмов заключается в отсутствии тех термодинамических параметров псевдоузлов, которые нам нужны.

1.5. Термодинамика нуклеиновых кислот

В этом разделе мы поговорим об общих механизмах плавления ДНК и экспериментальных результатах плавления, которые представлены в обзоре [19]. Как было сказано в предыдущем разделе, дезоксирибонуклеиновая кислота (ДНК) состоит из двух полинуклеотидных нитей. Они скручены в

двойную спираль как показано на Рис. 1.4. Эти две нити полностью комплементарны. В ДНК нуклеотиды аденин и тимин связываются между собой двумя водородными связями, а нуклеотиды гуанин и цитозин – тремя водородными связями. Диаметр ДНК примерно равен 20\AA . Расстояние между двумя соседними повторяющимися единицами приблизительно равно 3.4\AA . Число нуклеотидов, составляющих один виток спирали ДНК, колеблется от десяти до двенадцати повторяющихся единиц и зависит от формы ДНК (A, B, Z). Разделив 360° на число нуклеотидов в одном витке, мы получим угол поворота из расчета на одну повторяющуюся единицу.

Один из самых фундаментальных термодинамических процессов происходящих в ДНК это плавление. Это явление также называют переходом спираль – клубок. Схема плавления ДНК представлена на Рис. 1.5. Во время этого процесса водородные связи между азотистыми основаниями начинают разрушаться и, в конце концов, мы получаем две отдельные цепочки ДНК, которые могут быть рассмотрены как Гауссовы клубки.

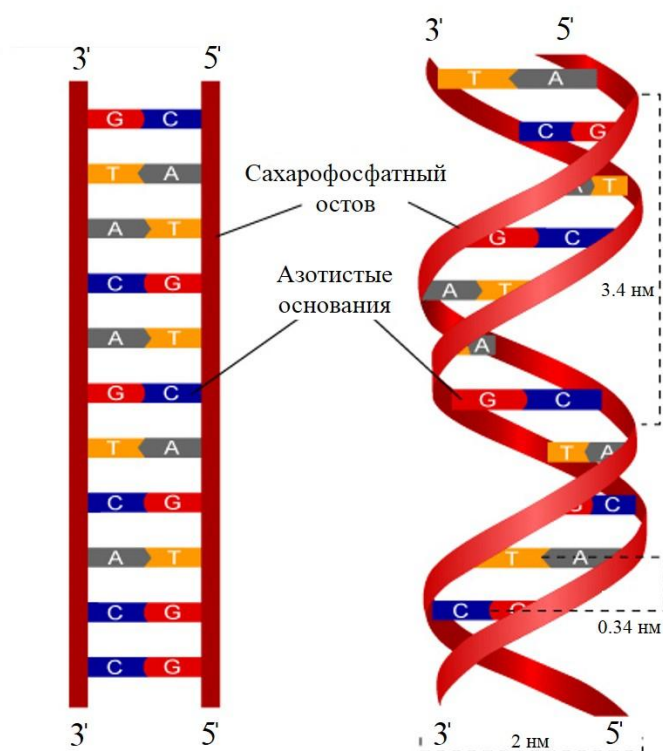


Рис. 1.4. Двойная спираль ДНК

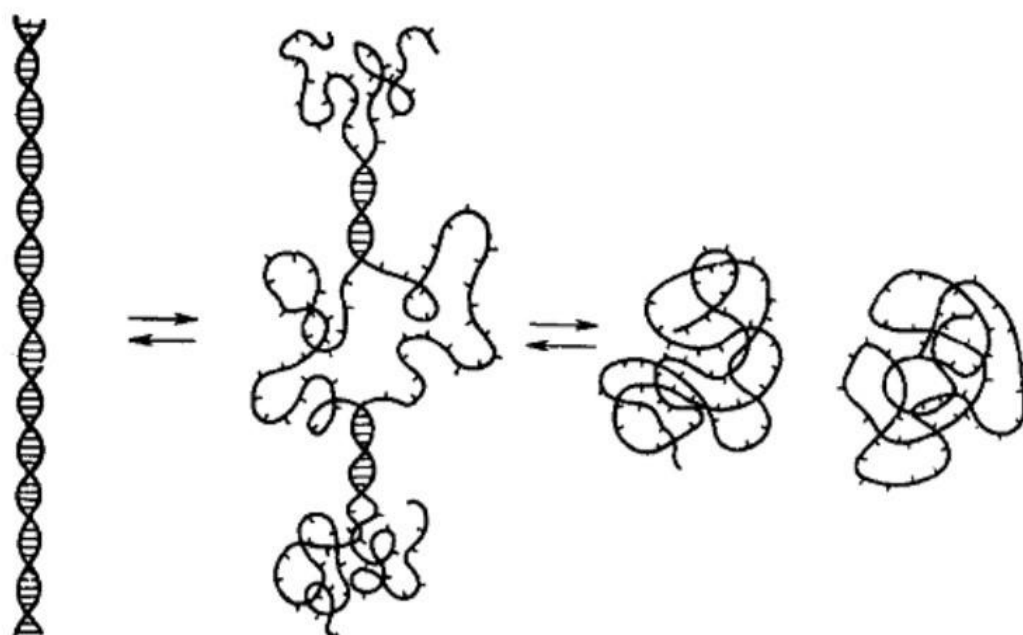


Рис. 1.5. Схематическое изображение перехода спираль – клубок в ДНК [19]

Переход спираль – клубок является обратимым процессом, то есть снижение температуры ведет к ренатурации ДНК. Но если ДНК полностью расплавлена, то вероятность восстановления спиральной структуры практически нулевая. Это является следствием очень большого влияния кинетических факторов. Теперь поговорим об экспериментальных данных касающихся плавления ДНК.

Существует целый ряд методов позволяющих экспериментально исследовать переходы спираль – клубок в ДНК. Один из них основан на поглощении ультрафиолетового излучения раствором ДНК). Метод основан на зависимости степени поглощения от структуры ДНК. То есть, поглощения для нуклеотидов в спиральных и нуклеотидов в клубках отличаются [20]. Причиной является отсутствие стэкинг взаимодействия в клубке в отличие от спирали. Степень денатурации описывается величиной $(D - D_{min}) / (D_{max} - D_{min})$, где D – оптическая плотность раствора, а D_{min} и

D_{max} – оптические плотности спиральной и клубковой структур соответственно. На Рис. 1.6 показана температурная зависимость оптической плотности для двухнитевого гомополинуклеотида (кривая плавления). Кривая плавления двухнитевого гомополинуклеотида хорошо изучена в [21]. Кривую плавления можно охарактеризовать двумя параметрами: температурой плавления (T_m) и интервалом плавления (ΔT). Ширина интервала плавления вычисляется с помощью формулы

$$\Delta T = 1 / \left| \frac{\partial \theta}{\partial T} \right|_{max}. \quad (1.1)$$

На Рис. 1.7 показана связь температуры плавления T_m и интервала плавления ΔT для ДНК тимуса теленка, которая была изучена в [23,24].

Одной из главных характеристик кривой плавления ДНК является ГЦ-состав. Зависимость температуры плавления от ГЦ-состава показана на Рис. 1.8. ГЦ-состав определяется по формуле

$$x_0 = (N_G + N_C) / (N_G + N_C + N_A + N_T), \quad (1.2)$$

где величины N_G , N_C , N_A и N_T – число азотистых оснований гуанина, цитозина, аденина и тимина соответственно. Как показано на Рис. 1.8 зависимость T_m от x_0 линейная.

Температура плавления ДНК существенным образом зависит от состава растворителя. Наличие двойной спирали ДНК возможно только в среде с достаточной концентрацией положительных ионов, таких как ионы натрия или калия. Для нейтрального водородного показателя рН мы можем использовать следующую эмпирическую формулу:

$$T_m = 176 - (2,6 - x_0)(36 - 7,04 \cdot \lg[Na^+]), \quad (1.3)$$

где величина $[Na^+]$ – молекулярная концентрация ионов натрия. Зависимость температуры плавления от логарифма концентрации ионов натрия показана на Рис. 1.9 [26]. Эта зависимость была получена для четырех разных ДНК с помощью формулы (1.3). Температура плавления довольно низкая при $pH < 5$ и $pH > 9$. Для этих ДНК ширина интервала

плавления приблизительно равна 3° . Для гомополинуклеотида этот параметр равен $0,5^\circ$. Главная часть исследования сделана для стандартных условий ($\text{pH} = 7$, $[\text{Na}^+] = 0,196 \text{ M}$).

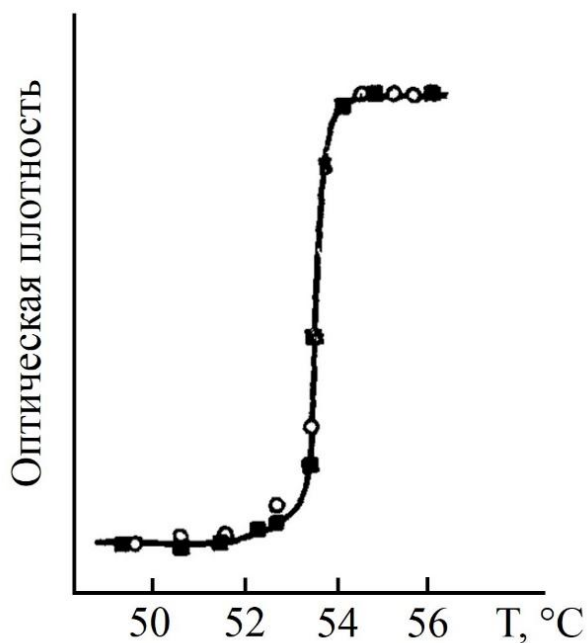


Рис. 1.6. Кривая плавления гомополинуклеотида [22].

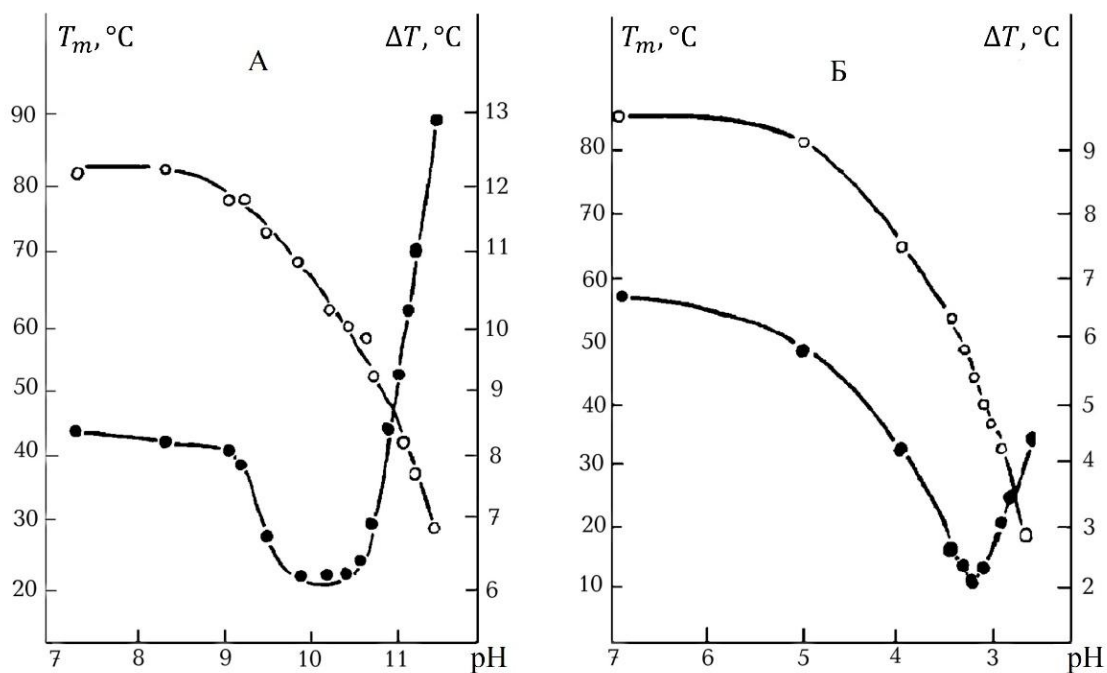


Рис. 1.7. Связь температуры плавления T_m (○) и интервала плавления ΔT (●) для ДНК тимуса теленка [23, 24].

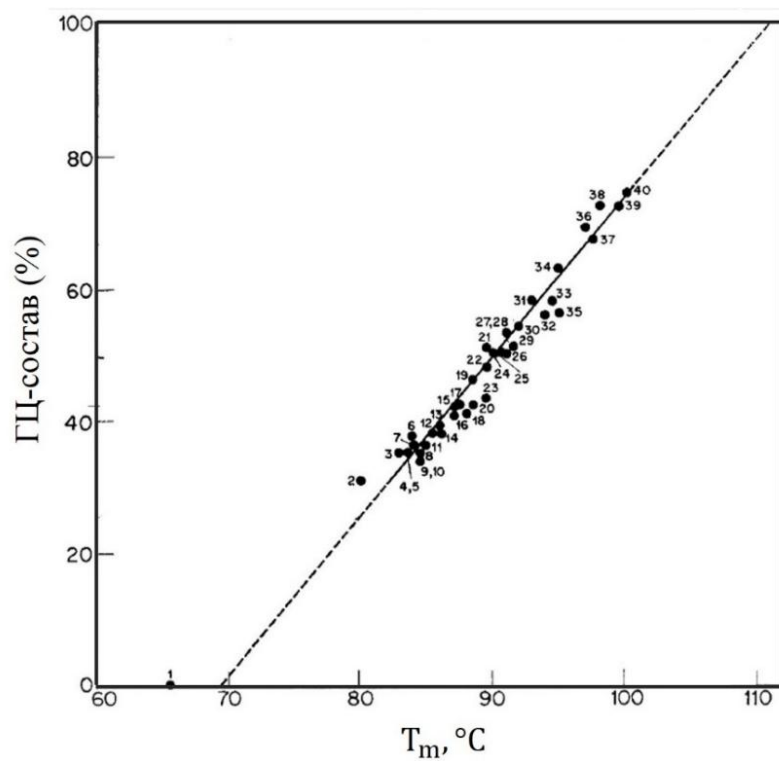


Рис. 1.8. Зависимость температуры плавления от ГЦ-состава [25].

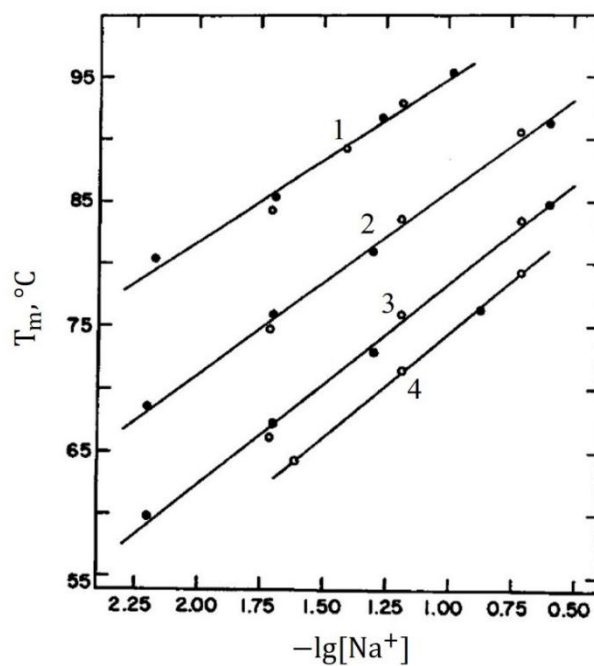


Рис. 1.9. Связь между температурой плавления и концентрацией ионов натрия в растворе. Линия 1 получена для *M. lysodeikticus* ($x_0 = 0.72$), линия 2 – *E. coli* ($x_0 = 0.5$), линия 3 – *S. saprophyticus* ($x_0 = 0.33$), линия 4 – *M. mycoides* var. *capri* ($x_0 = 0.24$) [26].

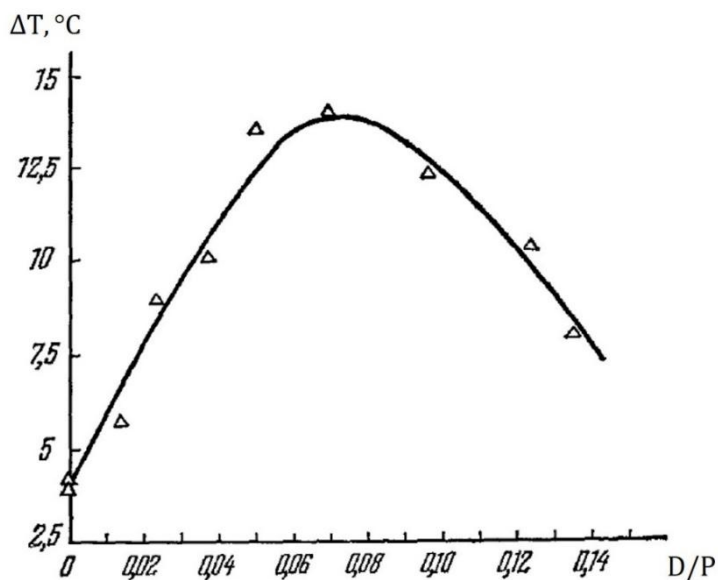


Рис. 1.10. Зависимость интервала плавления ДНК от концентрации рибонуклеазы в растворе [27].

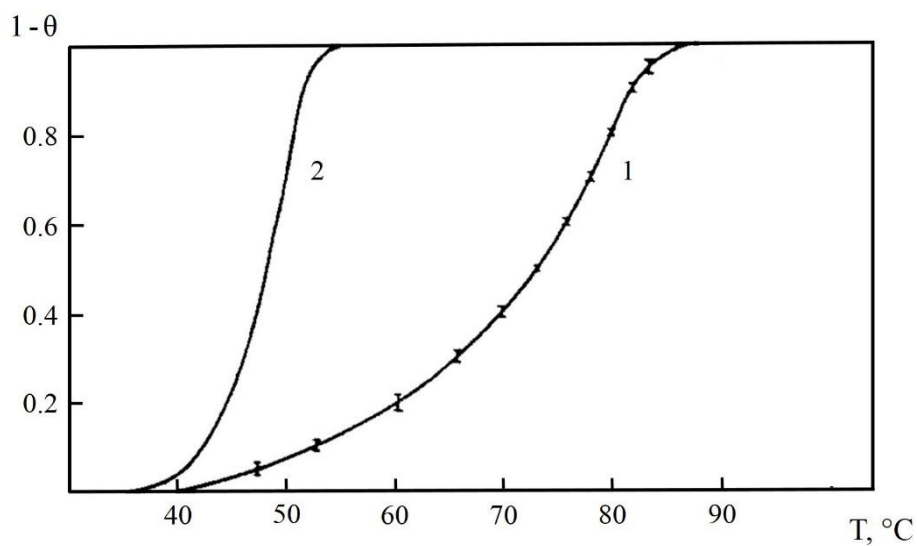


Рис. 1.11. Кривые плавления кольцевой, замкнутой ДНК полиомы (1) и для той же ДНК с разрывом в одной из нитей (2) в 7.2 М-ном растворе $NaClO_4$.

Вещества, которые могут связываться с ДНК, известные как лиганды, имеют очень большой вклад в кривые плавления. Например, лигандами являются ионы тяжелых металлов (Cu, Fe,...). Влияние органических лигандов на кривую плавления можно увидеть на Рис. 1.10, где показана зависимость интервала плавления от концентрации белков рибонуклеазы (величина D – молярная концентрация белков рибонуклеазы в растворе, величина P – молярная концентрация повторяющихся единиц в ДНК). Обычно $D \ll P$. Надо отметить, что в ходе эксперимента эти лиганды перераспределяются на ДНК. При данной температуре они принимают термодинамически наиболее выгодное положение. Эксперименты проводились для линейных, незамкнутых, двухцепочных ДНК. В случае с кольцевыми, замкнутыми ДНК, экспериментальные результаты другие. Характеристики кривых плавления для таких ДНК изучены в статье [28]. Температура плавления кольцевых ДНК на 20° больше, чем температура плавления линейных ДНК (Рис. 1.11). Расплавленные цепи в кольцевых ДНК остаются скручены друг вокруг друга, что и становится причиной более высокой температуры плавления. Как следствие, энтропия плавленного состояния более низкая, чем для плавленного состояния линейных ДНК. Кроме того, ширина интервала плавления для кольцевых ДНК в 2-3 раза больше, чем для линейных ДНК.

1.6. Алгоритмы динамического программирования

Самая стабильная вторичная структура молекулы РНК характеризуется минимальной свободной энергией. Мы можем получить такие структуры учитывая все возможные пары оснований и считая свободную энергию для каждой из этих структур [29]. Это возможно для

очень коротких последовательностей, так как число возможных конформации молекулы растет экспоненциально с ростом числа нуклеотидов в РНК. Однако существуют алгоритмы динамического программирования, которые позволяют нам вычислить свободные энергии для более длинных последовательностей. Эти алгоритмы основаны на рекуррентных соотношениях, которые позволяют получить термодинамические величины для более длинных последовательностей основываясь на уже полученных величинах для более коротких цепей. Теперь мы обсудим алгоритм программирования для очень простого выбора параметров взаимодействия. В рамках этой модели будем предполагать, что каждая пара оснований добавляет -1 к свободной энергии всей цепи и пренебрегаем энтропийными потерями при образовании петли. Следовательно, структурой с наименьшей свободной энергией будет та, в которой мы имеем наибольшее количество пар оснований. Поэтому, эта модель называется «модель максимальных совпадений» [30]. Предположим, что энергия связывания ($\epsilon_{i,j}$) между основаниями i и j равна -1 , если эти основания комплементарны, и она равна ∞ , если они не комплементарны. Следовательно, мы должны найти минимальную энергию $E_{i,j}$ цепи (i,j) . Если последний нуклеотид j образует пару с нуклеотидом k , то наша цепь разделится на две субцепи: от i до $k-1$ и от $k+1$ до $j-1$. Мы исключаем существование псевдоузлов. Иными словами, нуклеотиды с разных субцепей не могут образовать пары. Если j и k образовали пару, то энергия участка (i,j) будет равна $E_{i,k-1} + E_{k+1,j-1} + \epsilon_{k,j}$. Если они не образовали пару, то энергия будет равна $E_{i,j-1}$. Следовательно, минимальная энергия этого участка будет определяться как

$$E_{i,j} = \min(E_{i,j-1}, \min_{i \leq k \leq j-4} (E_{i,k-1} + E_{k+1,j-1} + \epsilon_{k,j})). \quad (1.4)$$

Мы предполагаем, что $E_{i,j} = 0$, если $j - i \leq 4$. После этого мы должны найти $E_{i,j+1}$, $E_{i,j+2}$ и так далее. В конце концов мы получим энергию $E_{1,N}$ для всей цепи с длиной N . Этот алгоритм оценивает вклад отдельных пар оснований в энергию вторичной структуры РНК. Предположим, мы имеем последовательность нуклеотидов от B_1 до B_n и эта последовательность расположена на окружности как показано на Рис. 1.12. Пусть основания B_x и B_y образовали пару. Наша задача заключается в том, чтобы понять будут ли основания B_x и B_y образовать пару в той вторичной структуре, которую мы ищем. Дуга $B_x B_y$ делит нашу окружность на две части: верхнюю и нижнюю секции (Рис. 1.12). Пренебрегаем псевдоузлами, то есть нуклеотиды из разных секции не могут образовать пару друг с другом. Иными словами, если два нуклеотида образовали пару, то они оба находятся или в верхней секции, или в нижней секции. Следовательно, энергия вторичной структуры будет равна сумме энергии верхней и нижней секции и энергии связи $B_x B_y$.

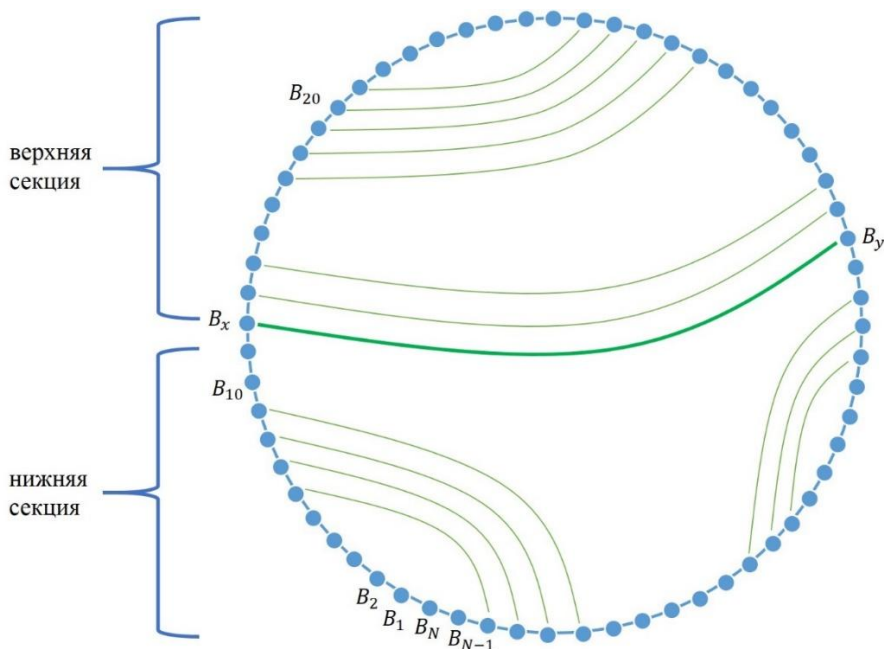


Рис. 1.12. Схематическое представление вторичной структуры РНК без псевдоузлов.

Если у нас есть реальная биологическая последовательность, то мы должны учитывать все возможные взаимодействия. Для таких последовательностей рекуррентные соотношения будут иметь более сложный вид.

Предположим, что наша последовательность состоит только из азотистых оснований А и U. Пусть вероятность азотистого основания А в нашей цепи равна P . Тогда вероятность U будет равна $(1 - P)$. Давайте посчитаем статистическую сумму, где суммирование проводится по всем возможным конформациям, исключая псевдоузлы. Для этого выделим участок (i, j) . Предположим, что $j + 1$ образовал пару с k . В этом случае мы получаем две субпоследовательности: от i до $k - 1$ и от $k + 1$ до j . В этом случае статистическая сумма для субцепи однонитевой РНК будет рекуррентно определяться [31, 32] как

$$Z_{i,j} = Z_{i,j-1} + \sum_{k=1}^{j-1} Z_{i,k-1} q_{ij} Z_{k+1,j-1}, \quad (1.5)$$

где $Z_{i,j}$ статсумма участка (i, j) , а величина $q_{ij} = \exp(-\beta \epsilon_{ij})$ является статистическим весом пары оснований между i и j .

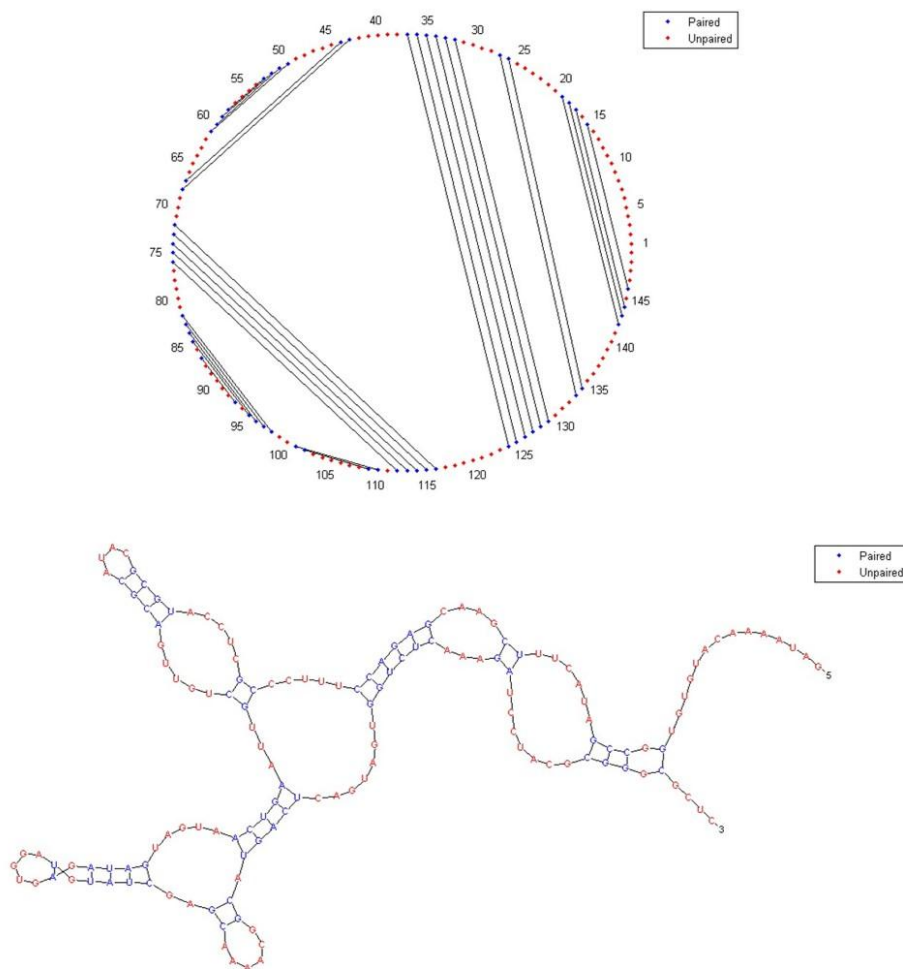


Рис. 1.13. Схематическое представление вторичной структуры РНК для цепи состоящей из 150 нуклеотидов.

На Рис. 1.13 показана вторичная структура РНК состоящей из 150 нуклеотидов полученное с помощью рекуррентного соотношения (1.5).

1.7. Существующие теории переходов во вторичной структуре ДНК

Модель Зимма – Брэгга

Эта модель [33] является первой последовательной и наиболее изученной статистической теорией перехода спираль – клубок. Она основана на одномерной модели Изинга. Предположим, что число нуклеотидов в цепи равно N . В этой теории состояние повторяющейся единицы определяется состоянием атома кислорода в карбоксильной

группе. Если эта повторяющаяся единица образовала водородную связь, то это состояние будет описываться числом **1**. Все другие состояния будем обозначать числом **0**. В итоге, мы будем иметь последовательность из нулей и единиц для каждой конфигурации. Параметр s определяется изменением свободной энергии, когда длина нашей спирали увеличивается на один мономер.

$$s = \exp\left(-\frac{\Delta F}{RT}\right), \quad (1.6)$$

где $\Delta F = F_h - F_c$. Если образовал пару мономер, который следует за тремя или более «свободными» мономерами, то свободная энергия растет. Параметр кооперативности σ описывает прирост свободной энергии:

$$\sigma = \exp\left(-\frac{F_s}{RT}\right), \quad (1.7)$$

где величина F_s – это прирост свободной энергии. При $N \gg 1$ статсумма имеет следующий вид:

$$Z = \text{Tr} P^N, \quad (1.8)$$

где N – число повторяющихся единиц, а P – матрица статистических весов. В этом случае

$$P = \begin{pmatrix} 1 & \sigma s \\ 1 & s \end{pmatrix}. \quad (1.9)$$

Отсюда мы получим секулярное уравнение:

$$(1 - \lambda)(s - \lambda) = \sigma s. \quad (1.10)$$

Это простое уравнение и мы можем найти точное решение. В [33] показано, что переход спираль – клубок происходит в интервале

$$1 - \sqrt{\sigma} \ll s \ll 1 + \sqrt{\sigma}. \quad (1.11)$$

Учитывая, что $\sigma \ll 1$, мы получим

$$\Delta T = 2\sqrt{\sigma} \frac{KT_m^2}{\Delta H}. \quad (1.12)$$

Таким образом, термодинамический анализ перехода спираль – клубок становится возможным с помощью модели Зимма – Брэгга.

Энтропия петли в модели Поланда – Шераги

Эта модель помогает описать существование петель в ДНК и получить разумные результаты [34]. Существуют два основные взаимодействия в ДНК: водородные связи и стэкинг. Водородные связи образуются между двумя основаниями, которые находятся в разных цепях. Стэкинг происходит между двумя соседними азотистыми основаниями. Предположим, что статвес водородных связей равен t , а статвес стэкинг взаимодействий – τ . Если у нас есть упорядоченная последовательность состоящая из j повторяющихся единиц, то статвес последовательности можно записать как

$$v_j = t(t\tau)^j \equiv \sigma\omega^j, \quad (1.13)$$

где $\sigma \equiv t$ и $\omega \equiv t\tau$. Производящая функция [35, 36]

$$V(x) = \frac{\sigma\omega}{(x-\omega)}. \quad (1.14)$$

Пусть $\sigma = 1$. В этом случае мы можем игнорировать неоднородности в t и τ зависящие от нашей последовательности. Отсюда

$$V(x) = \frac{\omega}{(x-\omega)}. \quad (1.15)$$

Если взять

$$U(x) = \frac{1}{V(x)}, \quad (1.16)$$

то получим

$$U(x) = \frac{x}{\omega} - 1. \quad (1.17)$$

Как было показано в [37], если наша петля состоит из N оснований, то ее энтропия будет иметь следующий вид

$$S(N) = R(N\ln\Omega - [A + \frac{3}{2}\ln N]). \quad (1.18)$$

Точное значение константы A невозможно найти. Член $RN\ln\Omega$ является конформационной энтропией свободной цепи.

Предположим, что наша цепь расположена на двумерной решетке. Если в цепи есть петля, то это означает, что концы петли должны совпадать. Как следствие, число возможных шагов влево должно быть равным числу шагов вправо, и число шагов вверх должно быть равным числу шагов вниз. Следовательно, если число оснований равно N , то число шагов влево – вправо будет равно числу шагов вверх – вниз и будет равно $\frac{N}{2}$. Следовательно, число конформации петли

$$Q = \frac{\left[\left(\frac{N}{2}\right)!\right]^2}{\left[\left(\frac{N}{4}\right)!\right]^4}. \quad (1.19)$$

Если используем формулу Стирлинга

$$n! = e^{-n} n^n (2\pi n)^{\frac{1}{2}}, \quad (1.20)$$

то получим

$$Q = N \ln 2 - \left[\ln \left(\frac{\pi}{4} \right) + \ln N \right]. \quad (1.21)$$

Для трехмерного случая мы будем иметь, что возможное число шагов в направлениях $\pm x$, $\pm y$ и $\pm z$ будет равно $\frac{N}{3}$, при больших N . Следовательно, число конформации петли будет

$$Q = \frac{\left[\left(\frac{N}{3}\right)!\right]^3}{\left[\left(\frac{N}{6}\right)!\right]^6}. \quad (1.22)$$

Используя формулу Стирлинга

$$\ln n! = n \ln n - n,$$

то получим

$$\ln Q = N \ln 2 - \left[\ln \left(\frac{\pi}{6} \right)^{\frac{3}{2}} + \frac{3}{2} \ln N \right]. \quad (1.23)$$

Очевидно, что формулы (1.18) и (1.23) похожи. Поэтому, мы можем написать

$$\ln Q = Na' - (b' + c \ln N), \quad (1.24)$$

где $c = 1$ для двумерной решетки и $c = \frac{3}{2}$ для трехмерной решетки. Следовательно, мы можем сказать, что энтропия петли будет определяться с помощью следующей формулы

$$S_{loop}(N) = R[Na - (b + c \ln N)], \quad (1.25)$$

где $c = d/2$, если мы пренебрегаем взаимодействия исключенного объема. Величина d это размерность пространства. Статвес свободной цепи

$$u_N = (e^a)^N, \quad (1.26)$$

и отсюда

$$u_N = e^{\frac{[S_{loop}(N)]}{R}} = (e^a)^N e^{-bN - c}. \quad (1.27)$$

Для нуклеиновых кислот $N = 2(i + 1) \approx 2i$. Если $d = 3$, то

$$u_i = (constant)u^i i^{-3/2}. \quad (1.28)$$

В статье [38] показано, что в случае учета взаимодействий дальнего порядка,

$$c \sim 1,75, \quad (1.29)$$

если у нас трехмерное пространство и

$$c \sim 1,46 \quad (1.30)$$

для двумерного пространства.

В частности Мукамелом в статье [39] была рассмотрена модель типа Поланда и Шераги. В рамках этой модели авторы рассмотрели эффекты взаимодействия исключенного объема. Несмотря на то, что они рассматривали эти взаимодействия приближенно, полученные результаты позволили понять связь между механизмами образования пар оснований и природой фазового перехода.

Согласно этой модели, мономеры в ДНК могут находиться только в двух состояниях: связанных и несвязанных. Следовательно, цепь выглядит, как последовательность таких участков. Энергия связывания для всех мономеров одна и та же. Статвес связанной пары

$$\omega = \exp\left(-\frac{E_0}{T}\right), \quad (1.31)$$

где $E_0 < 0$ является энергией связывания, а величина T – температурой. Если сегмент цепи состоит из k связанных единиц, то статвес этого сегмента будет определяться как

$$\omega^k = \exp\left(-\frac{kE_0}{T}\right). \quad (1.32)$$

Статвес несвязанной цепи с длиной k можно определить с помощью изменения энтропии. Для больших k оно равняется $\frac{As^k}{k^c}$, где величина s является неуниверсальной константой. Экспонента c описывает свойства петли. Авторы рассматривают случай, когда $A = 1$. Большая каноническая статсумма будет определена как

$$Z = \sum_{M=0}^{\infty} G(M) z^M = \frac{V_0(z)U_L(z)}{1-U(z)V(z)}, \quad (1.33)$$

где $G(M)$ – каноническая статсумма цепи с длиной M , а величина z – фугативность,

$$U(z) = \sum_{k=1}^{\infty} \frac{s^k}{k^c} z^k, \quad (1.34)$$

$$V(z) = \sum_{k=1}^{\infty} \omega^k z^k, \quad (1.35)$$

$$U_L(z) = 1 + U(z),$$

$$V_0(z) = 1 + V(z).$$

Средняя длина цепи может быть получена с помощью статсуммы:

$$\langle L \rangle = \frac{\partial \ln Z}{\partial \ln z}. \quad (1.36)$$

При $\langle L \rangle \rightarrow \infty$, степень спиральности θ будет функцией от температуры.

Среднее число связанных пар оснований будет определяться как

$$\langle m \rangle = \frac{\partial \ln Z}{\partial \ln \omega}. \quad (1.37)$$

Следовательно,

$$\theta = \lim_{L \rightarrow \infty} \frac{\langle m \rangle}{\langle L \rangle} = \frac{\partial \ln z^*}{\partial \ln \omega}, \quad (1.38)$$

где z^* – фугативность, когда $\langle L \rangle \rightarrow \infty$. Используя

$$V(z) = \frac{\omega z}{(1-\omega z)}, \quad (1.39)$$

мы получим, что

$$U(z^*) = \frac{1}{(\omega z^*)} - 1. \quad (1.40)$$

Природа перехода будет определена благодаря зависимости z^* от ω . В [39, 40] показано, что существуют 3 случая:

1. при $c \leq 1$, фазового перехода не наблюдается,
2. при $1 < c \leq 2$, переход является непрерывным,
3. при $c > 2$, мы имеем фазовый переход первого порядка.

Экспонента $c = dv$, где величина d – размерность пространства. Если цепь является идеальной, то $c = \frac{d}{2}$. Следовательно, при $d \leq 2$, перехода не наблюдается, при $2 < d \leq 4$, переход непрерывный и, при $d > 4$, мы имеем переход первого порядка.

Модель Пейрара – Бишопа

Метод интеграла перехода использовался для анализа моделей Пейрара – Бишопа [41 - 45]. В этой модели авторы использовали тот факт, что существует аналогия между изучением конформационных свойств с помощью статистической физики и уравнением диффузии. Проблема денатурации ДНК смоделирована как частица в потенциале Морзе, которая описывает образование водородных связей. Введем пару переменных для каждой повторяющейся единицы. Эта пара описывает отклонение сегментов цепи в рамках каждой повторяющейся единицы в направлении параллельном оси спирали ДНК. Некоторое отклонение дозволено, но если отклонение больше дозволенного, то водородные связи разрушаются. Также был изучен эффект гармонического потенциала, который имитирует стэкинг между соседними повторяющимися единицами.

Гамильтониан этой модели будет иметь следующий вид:

$$H = \sum_n \frac{p_n^2}{2m} + W(y_n, y_{n-1}) + V(y_n), \quad (1.41)$$

где $p_n = m \frac{dy_n}{dt}$, величина m – приведенная масса оснований. Потенциал $V(y_n)$ описывает взаимодействие между двумя повторяющимися единицами. Другими словами, она описывает образование водородных связей. Потенциал $W(y_n, y_{n-1})$ описывает взаимодействия между двумя повторяющимися единицами вдоль одной цепи ДНК, то есть стэкинг взаимодействия. Удобно использовать потенциал Морзе, так как он является стандартным для описания взаимодействий приводящих к образованию химических связей и он имеет подходящую форму. Мы имеем сильно отталкивание на коротких расстояниях, состояние равновесия отвечающее минимальной энергии затухание с увеличением расстояния. Благодаря Гамильтониану, мы можем найти зависимость среднего отклонения от равновесного состояния от энергии образования пары оснований. Среднее отклонение описывает степень денатурации.

Предварительная модель перехода спираль – клубок в двухцепочной ДНК

Рассмотрим двухцепочный гомополинуклеотид с комплементарным связыванием. Это будет возможным, если рассмотрим реальную ДНК с приближением, что энергии связей АТ и GC равны. Чтобы рассмотреть порядок фазового перехода в двухцепочной ДНК мы должны рассмотреть гомополимерную ДНК с комплементарным образованием пар оснований. Экспериментально это может быть получено используя инверсию стабильности предложенный в [46]. При подходящей концентрации соединения алкиламмония, стабильность пар АТ и GC не только становится равной, а даже может наблюдаться инверсия стабильности. В случае равных стабильностей пар АТ и GC, двухцепочная ДНК ведет себя как гомополимер с симметричными петлями. В другом случае мы можем рассмотреть

случайный гетерополимер состоящий только из комплементарных пар АТ. Одна цепь является случайной последовательностью нуклеотидов А и Т, а другая цепь полностью комплементарна первой. Мы можем сказать то же самое для GC пар. В этом случае энергия образования водородной пары будет одинакова вдоль всей цепи. Мы можем предположить, что водородные связи образуются только между нуклеотидами имеющими одинаковый номер. Следовательно, все петли симметричны. Схема макромолекулы показана на Рис. 1.14.

Мы изучаем образование водородных связей между комплементарными повторяющимися единицами в двух цепях. Предположим, что первые повторяющиеся единицы связаны. Гамильтониан макромолекулы [47]

$$-\beta H = J \sum_{i=1}^N \delta_1^{(i)}, \quad (1.42)$$

где $\beta = T^{-1}$, $J = \frac{U}{T}$, величина U – энергия образования водородной связи между одной, комплементарной парой. Величина $\delta_1^{(i)}$ равна 1, если образовалась связь между комплементарными нуклеотидами под номером i , и она равна 0, если связь не образовалась. Так как первая пара связана, то величина $\delta_1^{(i)}$ ненулевая, если две цепи образовали замкнутое кольцо между первой и i -ой повторяющимися единицами. Существование других колец в интервале $[1, i]$ возможны. Фактически, эта модель является моделью Поланда – Шераги [34]. Статсумма для Гамильтониана будет иметь следующий вид

$$A = \sum_{\{\vec{\gamma}_i\}} \exp(-\beta H) = \sum_{\{\vec{\gamma}_i\}} \prod_{i=1}^N (1 + v \delta_1^{(i)}), \quad (1.43)$$

где $v = e^J - 1$. Пусть $\vec{\gamma}_i$ будет набором всех возможных значений $1, 2, \dots, Q$, которые являются номерами конформации цепи. Статсумма может быть представлена в виде ряда по степеням v . Используя соотношение

$$\delta_1^{(k)} \delta_1^{(m)} = \delta_1^{(k)} \delta_{k+1}^{(m-k)}$$

мы можем записать член относящийся к v^f как

$$v^f \delta_1^{(k_1)} \delta_{k_1+1}^{(k_2-k_1)} \delta_{k_2+1}^{(k_3-k_2)} \dots \delta_{k_{f-1}+1}^{(k_f-k_{f-1})}. \quad (1.44)$$

Используя циклические граничные условия и определяя $m_i = k_i - k_{i-1}$, мы получим

$$\Lambda = Q^N \sum_f v^f \sum_{m_1} \varphi(m_1) \dots \sum_f \varphi(m_f), \quad (1.45)$$

где

$$\varphi(m) = Q^{-m} \sum_{\gamma_1} \sum_{\gamma_2} \dots \sum_{\gamma_m} \delta_1^{(m)}. \quad (1.46)$$

Согласно (1.46), величина $\varphi(m)$ является соотношением числа состояний соответствующих образованию петли с длиной m и всех состояний цепи с длиной m . Следовательно, функция $\varphi(m)$ может быть интерпретирована как вероятность образования петли с длиной m . Используя условие

$$\sum_{i=1}^f m_i = N$$

и умножая (1.45) на коэффициент

$$\delta(\sum_{k=1}^f m_k - N),$$

мы получим

$$\Lambda = \frac{1}{2\pi i} \oint z^{-N-1} \sum_{f=1}^N (v \sum_{m=1}^{\infty} \varphi(m) z^m)^f dz. \quad (1.47)$$

В [47] функция $\varphi(m)$ была выбрана как

$$\varphi(m) = \begin{cases} Q^{-m}, & m \leq \Delta \\ Q^{-\Delta}, & m > \Delta \end{cases}. \quad (1.48)$$

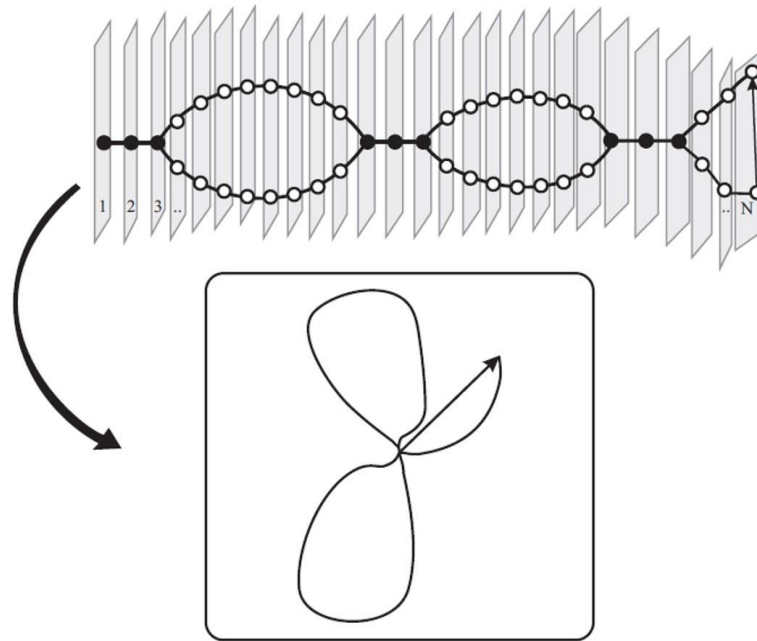


Рис. 1.14. Схематическое представление модели.

Используя метод перевала, мы можем показать, что характеристическое уравнение для свободной энергии в термодинамическом пределе такое же, как для модели ОМПЦ, которая является одномерной моделью типа Поттс. Это представление функции $\varphi(m)$ является эмпирическим и пренебрегает возможным образованием петель короче Δ . Величина Δ характеризует жесткость отдельной цепи.

Гл.2. Конформационные переходы в оц-РНК с двухбуквенным и четырехбуквенным алфавитом в последовательности

РНК играет ключевую роль во всех живых организмах. Кроме передачи генетической информации, РНК активно участвует в разнообразных процессах в клетке [48]. Цепи РНК состоят из четырех различных оснований А, С, G и U, которые могут образовывать двойные спиральные структуры, состоящей из Уотсона-Криковских пар оснований А-U или G-C. Таким образом, последовательность оснований контролирует трехмерную структуру РНК. Следовательно, последовательность РНК, как структура имеет важное значение для надлежащего функционирования организма [49]. Экспериментально стабильность вторичной структуры РНК изучалась в исследованиях денатурации, где цепи РНК разворачиваются путем изменения температуры или концентрации денатурата. В стандартном сценарии, вторичные структуры РНК расплавляются при высокой температуре, чтобы максимизировать конформационную энтропию цепи [50]. Но плавления РНК наблюдается и при низких температурах: явления, которое много обсуждалось и до сих пор лишь частично объяснено [51,52]. Считается, что холодовая денатурация белков является следствием немонотонной зависимости гидрофобного эффекта от температуры [53-55], но основан ли механизм холодовой денатурации РНК на том же эффекте остается неясным [52]. Замороженный беспорядок последовательности может только ослабить стабильность вторичной структуры РНК при низких температурах, что предлагает альтернативный механизм холодовой денатурации РНК. Из-за сложной теоретической природы фолдинга РНК, он широко исследуется с учетом влияния псевдоузлов [56], петель в цепи [57], внешних растягивающих сил [58], а также наиболее значимого в данном контексте, беспорядок в последовательности [59-63]. Были применены многочисленные методы к моделям РНК с замороженным беспорядком.

Ранее, для исследования моделей РНК с замороженным беспорядком в последовательности применялись репличные и численные методы [56,57,59-61]. Здесь нами использовался метод ограниченного отжига [64-67], результаты которого мы сначала сравнили с результатами, полученными с помощью численных методов [68,69] для цепи РНК с конечной длиной. Показано, что подход ограниченного отжига удивительно точное описание фолдинга РНК с замороженным беспорядком. Мы проанализировали термодинамические свойства РНК с неупорядоченной последовательности в пространстве параметров натянутом на энергии образования пар оснований и параметров функции распределения последовательности. Мы показали, что температурно-зависимая удельная теплоемкость системы демонстрирует два пика, в то время как степень спиральности понижается как при низких, так и при высоких температурах. Эффект имеет место только для конкурентного спаривания оснований (выгодного для разных и невыгодного для одинаковых оснований).

Мы изучали последовательности РНК используя 2 алфавита: двухлитерный, состоящий только из оснований А и U, и четырехлитерный, состоящий из оснований А, U, С и G. Гамильтоньян для цепи РНК состоящей из N оснований будет иметь следующий вид:

$$H(M, \{h_N\}) = \sum_{1 \leq i < j \leq N} m_{ij} \epsilon_{ij}, \quad (2.1)$$

где величина M является матрицей пар оснований размером $N \times N$ с элементами $m_{ij} = 1$, если основания i и j образовали пару, и $m_{ij} = 0$, если i и j не образовали пару. В двухлитерном случае энергия взаимодействия

$$\epsilon_{ij} = \epsilon_0 + \epsilon h_i h_j \quad (2.2)$$

будет состоять из константы ϵ_0 и члена, зависящего от переменных последовательности $h_i = \pm 1$, описывающих тип нуклеотида: если $h_i = +1$, то i -ый нуклеотид будет А, а если $h_i = -1$, то i -ый нуклеотид – U. В четырехлитерном случае

$$\epsilon_{ij} = \epsilon_0 + \epsilon h_i h_j + \epsilon' g_i g_j \quad (2.3)$$

появится дополнительная степень свободы $g_i = \pm 1$. Статистическая сумма будет иметь следующий вид:

$$Z_N(\{h_N, g_N\}) = \sum'_M \exp[-\beta H(M, \{h_N, g_N\})], \quad (2.4)$$

где $\beta = 1/k_B T$ и сумма производится по всем тем реализациям M , при которых в матрице есть максимум один ненулевой элемент в каждой строке и в каждом столбце и нет псевдоузлов. Смысл первой части состоит в том, что каждое основание может образовать пару только с одним другим основанием, а отсутствие псевдоузлов вводит следующее ограничение: если у нас есть две пары (i, j) и (k, l) с $i < j, k < l$, то либо одна пара находится в другой ($i < k < l < j$), либо они отдельные ($i < j < k < l$), то есть они не перекрываются [59]. Функция распределения нуклеотидов в последовательности в двухлитерном случае будет иметь следующий вид:

$$P_2(\{h_n\}) = \prod_{i=1}^N \rho(h_i, q), \quad (2.5)$$

где $\rho(h_i) = q\delta(h_i - 1) + (1 - q)\delta(h_i + 1)$. Величины q и $1 - q$ являются вероятностями, что $h_i = +1$ и $h_i = -1$ соответственно. Для четырехлитерного случая, уравнение (2.5) будет иметь следующий вид:

$$P_4(\{h_n\}, \{g_n\}) = \prod_{i=1}^N \rho(h_i, q) \rho(g_i, p).$$

В случае без псевдоузлов статистическая сумма субцепи РНК $Z_{i,j}$ между нуклеотидами i и j может быть рекурсивно определена как [68, 69]

$$Z_{i,j} = Z_{i,j-1} + \sum_{k=i}^{j-1} Z_{i,k-1} Q_{k,j} Z_{k+1,j-1}, \quad (2.6)$$

где величина $Q_{k,j} = \exp(-\beta \epsilon_{kj})$ является статистическим весом пары между нуклеотидами k и j и $Z_N = Z_{1,N}$.

2.1. Двухлитерная модель

Рассмотрим двухлитерный случай. Для данной последовательности $\{h_N\}$ свободная энергия, приходящаяся на один нуклеотид определяется как

$$\beta f(\{h_n\}) = -N^{-1} \ln Z_N(\{h_n\}).$$

Смысл самоусредняемости состоит в том, что термодинамический предел $N \rightarrow \infty$ функции $f(\{h_n\})$ становится независимым от $\{h_n\}$ и равен свободной энергии усредненной по последовательностям с соответствующей функцией распределения [70]:

$$\beta \bar{f} = -\frac{1}{N} \overline{\ln Z_N(\{h_n\})} \quad (2.7)$$

где знаком \bar{A} обозначено среднее по функции распределения $P_2(\{h_n\})$ в формуле (2.5). Этот факт используется при оценке свободной энергии для ограниченного отжига:

$$\beta f_{ca}(\mu) = -\frac{1}{N} \ln Z_N^{ca} = -\frac{1}{N} \overline{\ln Z_N(\{h_n\}) e^{-N\mu\alpha(\{h_n\})}}, \quad (2.8)$$

где величина $\alpha(\{h_n\})$ является соответствующим образом определенной интенсивной величиной зависящей от последовательности. Предполагается, что ее средняя равна нулю. Согласно [66], f_{ca} удовлетворяет следующему неравенству:

$$\bar{f} \geq f_{ca} \geq f_a, \quad (2.9)$$

где величина $\beta f_a = -\frac{1}{N} \overline{\ln Z_N(\{h_n\})}$ приведенная свободная энергия в расплавленном случае. Таким образом, $f_{ca}(\mu)$ улучшает оценку полученную в расплавленном приближении и представляет собой нижнюю границу замороженной свободной энергии \bar{f} . Наилучшая оценка получается при максимизации f_{ca} по неопределенному множителю Лагранжа μ :

$$f_{ca}^* = f_{ca}(\mu^*) = \max_{\mu} f_{ca}(\mu). \quad (2.10)$$

Простейшее ограничение, которое можно наложить на замороженные переменные $\{h_n\}$ это линейная зависимость:

$$\alpha(\{h_n\}) = \frac{1}{N} \sum_{i=1}^N [h_i - (2q - 1)].$$

После некоторых арифметических упрощений [71] из формулы (2.8) получаем:

$$Z_N^{ca} = e^{N\mu(2q-1)} \Omega^N Z_N^{hom}(\epsilon_{ca}), \quad (2.11)$$

где величина $Z_N^{hom}(\epsilon_{ca})$ является статистической суммой гомополимерной цепи РНК, полученной из формулы (2.4) при $\epsilon_{ij} = \epsilon_{ca}$. Здесь

$$\epsilon_{ca} = \epsilon_0 - (1/\beta) \ln(W/\Omega^2),$$

$$\Omega = qe^{-\mu} + (1 - q)e^{\mu},$$

$$W = e^{-\beta\epsilon} [q^2 e^{-2\mu} + (1 - q)^2 e^{2\mu}] + 2q(1 - q)e^{\beta\epsilon}.$$

Поскольку Z_N^{hom} при $N \gg 1$ оценивается как [61]

$$Z_N^{hom} \sim N^{-3/2} (1 + 2\sqrt{Q_{ca}})^N,$$

где

$$Q_{ca} = e^{-\beta\epsilon_{ca}},$$

то для вариационной свободной энергии получим следующее выражение:

$$\frac{f_{ca}(\mu)}{k_B T} = -\mu(2q - 1) - \ln \Omega(\mu) - \ln(1 + 2\sqrt{Q_{ca}}). \quad (2.12)$$

Максимизируя это выражение мы получим свободную энергию ограниченного отжига f_{ca}^* .

На Рис. 2.1(а) мы сравнили f_{ca}^* (синяя линия) для $q = 0.75$ с точными результатами полученными благодаря рекурсивному алгоритму в уравнении (2.6) для $N = 50$. Красные линии получены для 30 отдельных случайных последовательностей, а пунктирная линия является соответствующим замороженным средним \bar{f} полученным из уравнения (2.7). Как мы видим, зависимость получена с использованием метода ограниченного отжига хорошо согласуется с зависимостью, полученной с использованием численных методов. Энергетические параметры $\epsilon_0 < 0$ и $\epsilon = 1.5|\epsilon_0|$

выбраны таким образом, чтобы энергия пары AU $\epsilon_0 - \epsilon = -2.5|\epsilon_0|$ была более выгодна, чем энергия пар AA или UU $\epsilon_0 + \epsilon = 0.5|\epsilon_0|$.

Очень хорошее соответствие с численными методами получается и для удельной теплоемкости (Рис. 2.1(б)):

$$C_V = -T \frac{\partial^2 f}{\partial T^2}.$$

Как мы видим на Рис. 2.1(б), при помощи метода ограниченного отжига мы получили двойной пик на графике удельной теплоемкости (синяя линия). Тот же двойной пик получается и замороженного среднего (пунктирная линия). Разброс отдельных реализаций последовательности (красные линии) является существенным, показывая, что некоторые цепи ($N = 50$) далеки от предела самоусредняемости. Структурное значение двух пиков для удельной теплоемкости может быть оценено с помощью степени спиральности θ :

$$\theta = \frac{2}{N} \langle \sum_{i < j} m_{ij} \rangle, \quad (2.13)$$

где символ $\langle A \rangle$ означает тепловое усреднение по всем разрешенным парам оснований матрицы M , как в уравнении (2.4). $\theta = 1$ соответствует структуре, при которой все основания образовали пару, а $\theta = 0$ – структуре, при которой ни одно основание не образовало пару. В подходе ограниченного отжига, θ задается ранее полученным выражением для гомополимерной РНК [61]:

$$\theta = \frac{2\sqrt{Q_{ca}}}{1+2\sqrt{Q_{ca}}}.$$

В численных методах θ получается исходя из вероятности образования пары между нуклеотидами i и j . В статье [61]:

$$\langle m_{ij} \rangle = Q_{ij} Z_{i+1, j-1} Z_{j+1, N+i-1} / Z_{1, N},$$

где статистическая сумма $Z_{j+1, N+i-1}$ вычисляется из уравнения (2.6) для периодически повторяющейся последовательности. На Рис. 2.1(в) мы

сравнили величину θ полученную с помощью ограниченного отжига и численных методов. Как видим, согласие довольно хорошее. На графике видно, что степень спиральности демонстрирует максимум при промежуточных температурах, что указывает на присутствие плавления при высоких и низких температурах.

Для понятия механизма холодовой денатурации рассмотрим 2 фракции: η^+ , которая энергетически невыгодна (например пары AA и UU) и η^- , которая энергетически выгодна (например пара AU). Эти величины будут определяться с помощью формулы

$$\eta^{\pm} = \frac{2}{N} \overline{\langle \sum_{ij} \delta(h_i h_j \mp 1) m_{ij} \rangle}, \quad (2.14)$$

которая прямо получается с помощью соответствующих производных свободной энергии [71]. По определению

$$\theta = \eta^+ + \eta^-.$$

На Рис. 2.1(г) мы сравнили величину η^+ полученную с помощью метода ограниченного отжига и численных алгоритмов. Видно, что доля невыгодных контактов довольно резко падает при низких температурах. Во вставке

к Рис. 2.1(г) результаты ограниченного отжига для η^+ и η^- показывают, что число энергетически выгодных пар η^- монотонно растет при понижении температуры в широком интервале приблизительно в той же области, в которой наблюдается высокотемпературный пик в C_V , в то время как η^+ резко падает в интервале, где наблюдается низкотемпературный пик для C_V . Подобное температурное поведение величин η^+ и η^- , таким образом, не только объясняет немонотонность θ на Рис. 2.1(в), но и двойной пик C_V на Рис. 2.1(б).

То, что структура с двойным пиком C_V действительно является следствием замороженного беспорядка, показано на вставке в Рис. 2.1(б),

где мы сравнили результаты полученные для ограниченного отжига (синяя линия), результаты для расплавленного беспорядка (зеленая линия) и результаты для однородной РНК с усредненной энергией пар оснований равной $\bar{\epsilon} = \overline{\epsilon_0 + \epsilon h_i h_j} = \epsilon_0 + \epsilon(2q - 1)^2 = -5/8|\epsilon_0|$ (коричневая линия). Вычисления в расплавленном случае дают лишь один широкий высокотемпературный пик, в то время как при однородном случае получается только низкотемпературный пик. Причина заключается в отсутствии конкуренций между выгодными и невыгодными парами оснований, что приводит к монотонному увеличению степени спиральности θ , как показано на Рис. 2.1(в).

После того, как мы проверили эффективность метода ограниченного отжига для одного значения параметра, мы можем использовать этот подход для изучения фазового поведения всей цепи РНК. На Рис. 2.2 показаны температурные зависимости (а) степени спиральности θ и (б) удельной теплоемкости C_V для ранее зафиксированных энергетических параметров $\epsilon_0 < 0$, $\epsilon = 3|\epsilon_0|/2$ и для разных вероятностей $1/2 \leq q \leq 1$ (отметим, что система симметрична по отношению к точке $q = 1/2$, так что результаты для области $q < 1/2$ будут идентичны). Кроме гомополимерного случая $q = 1$ и точки симметрии $q = 1/2$, во всех остальных случаях получается двойной пик для удельной теплоемкости, а степень спиральности понижается как при низких, так и при высоких температурах. Особый интерес представляет случай $q = 1/2$. Минимизируя уравнение (2.12) для $q = 1/2$ получаем $\mu^* = 0$. Это соответствует отождествленному среднему в уравнении (2.8), что, как мы уже показали на Рис. 2.1, приводит к одиночному пику C_V и монотонному изменению θ . Заметим, что для всех $q \neq 1/2$, число возможных невыгодных пар AA и UU равно $q^2 + (1 - q)^2$ и оно больше числа возможных выгодных пар AU и UA, которое равно

$2q(1 - q)$. Таким образом, при высоких температурах пары AA и UU становятся статистически более выгодны. Этот факт объясняет почему точка симметрии $q = 1/2$ является особой: только в этой точке выгодные и невыгодные пары имеют одинаковый статистический вес. Мы приходим к выводу, что двойной пик C_V и немонотонность θ получаются для всех значений q , кроме особых случаев ($q = 0; 1/2; 1$).

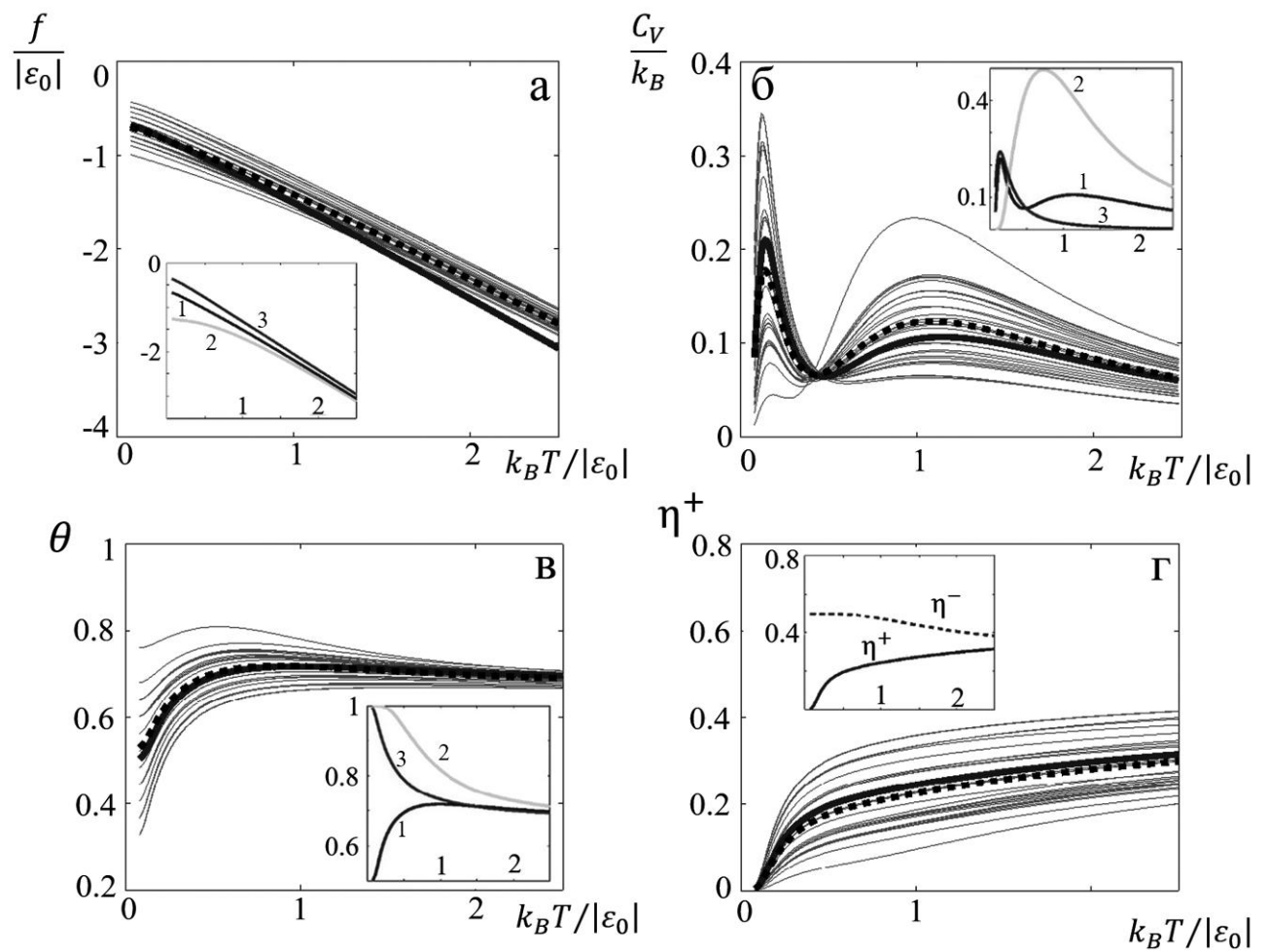


Рис. 2.1. Температурные зависимости (а) свободной энергии, (б) удельной теплоемкости, (в) степени спиральности и (г) энергетически невыгодной фракции для параметров $\epsilon_0 < 0$, $\epsilon = 1.5|\epsilon_0|$ и $q = 0.75$. Тонкие линии – численные результаты для цепи с длиной 50 нуклеотидов, пунктирная линия

– замороженное среднее, толстая линия – метод ограниченного отжига при $N \rightarrow \infty$. Во вставках (а), (б) и (в) линия 1 - метод ограниченного отжига, линия 2 – расплавленный случай, а линия 3 – случай гомополимера. Во вставке (г) сравнены фракции энергетически выгодных и невыгодных пар оснований.

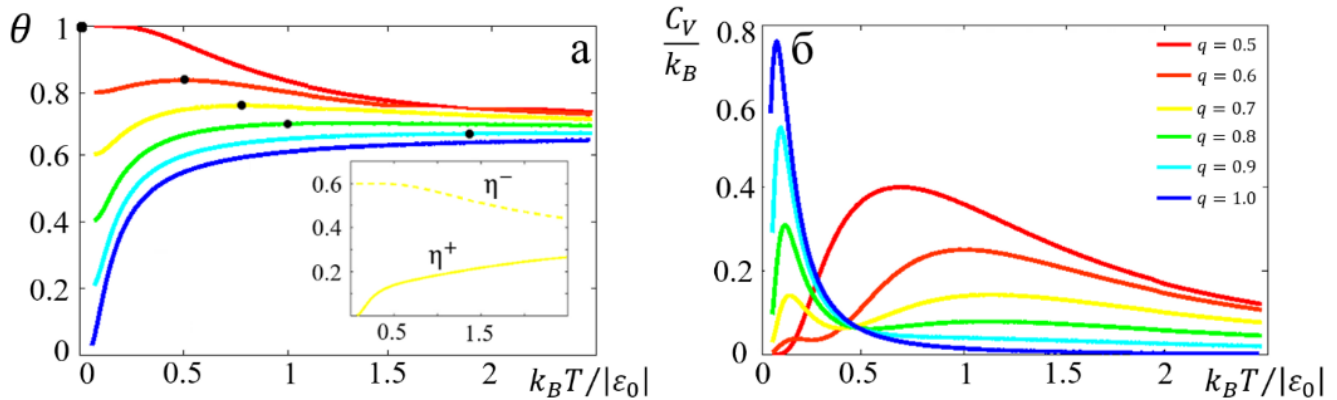


Рис. 2.2. Температурные зависимости (а) степени спиральности и (б) удельной теплоемкости для конкурентного случая $\epsilon_0 + \epsilon > 0$ ($\epsilon_0 < 0$, $\epsilon = 3|\epsilon_0|/2$) для разных вероятностей q . Черные точки в (а) показывают максимальное значение. Во вставке сравнены фракции энергетически выгодных и невыгодных пар оснований для $q = 0.7$.

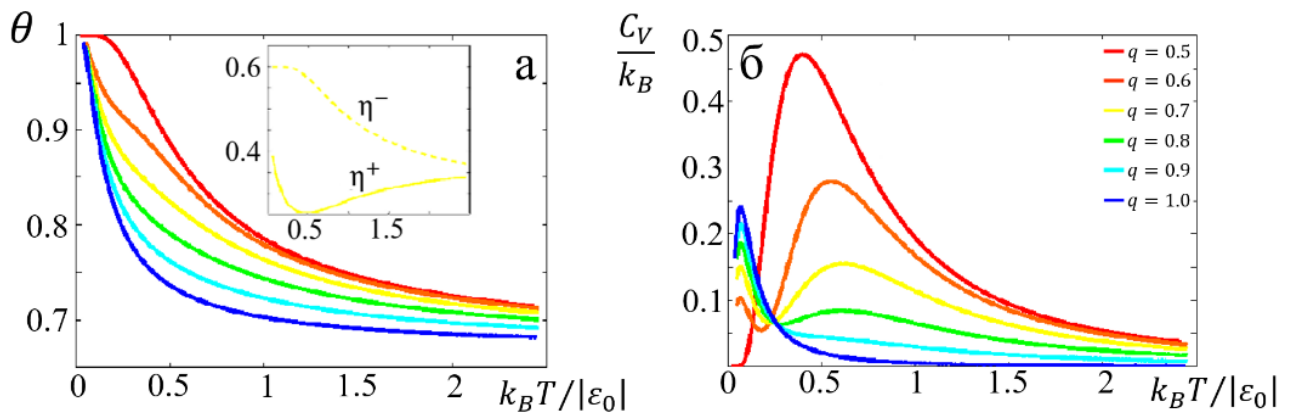


Рис. 2.3. Температурные зависимости (а) степени спиральности и (б) удельной теплоемкости для неконкурентного случая $\epsilon_0 + \epsilon < 0$ ($\epsilon_0 < 0$, $\epsilon = 2|\epsilon_0|/3$) для разных вероятностей q . Во вставке сравнены фракции энергетически выгодных и невыгодных пар оснований для $q = 0.7$.

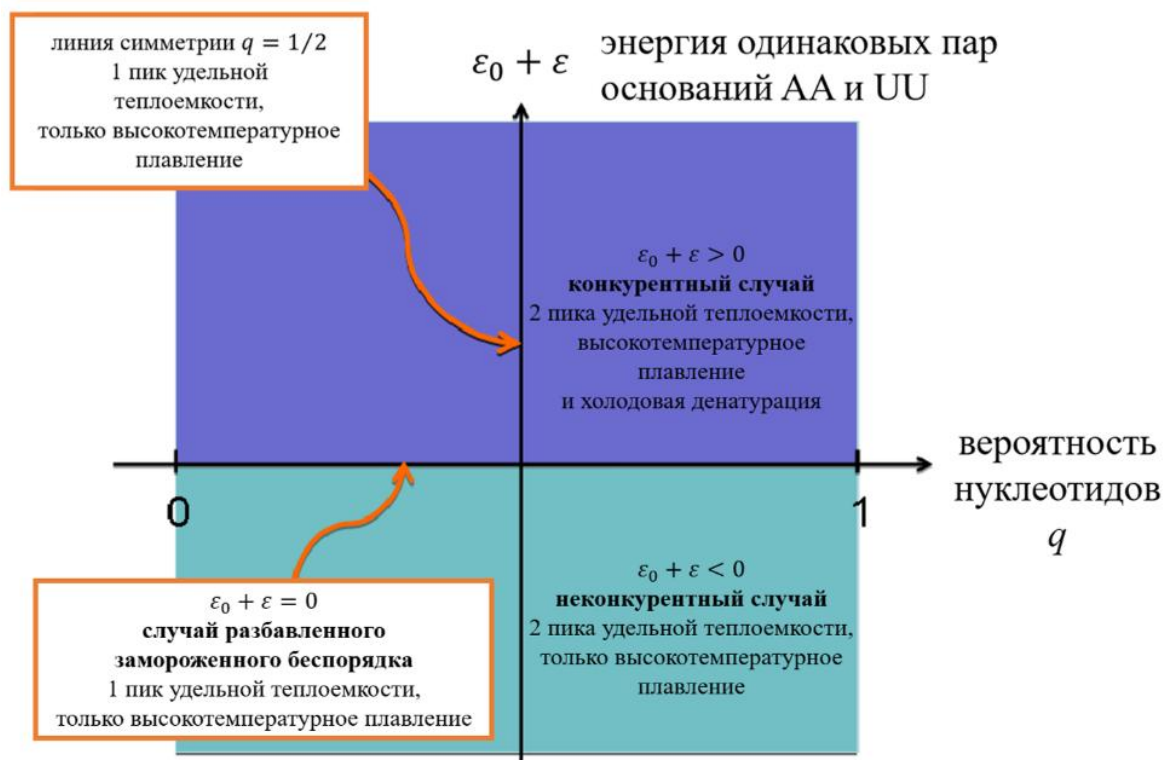


Рис. 2.4. Фазовая диаграмма поведения фолдинга РНК как функции от вероятности нуклеотидов q и энергии одинаковых пар оснований $\epsilon_0 + \epsilon$.

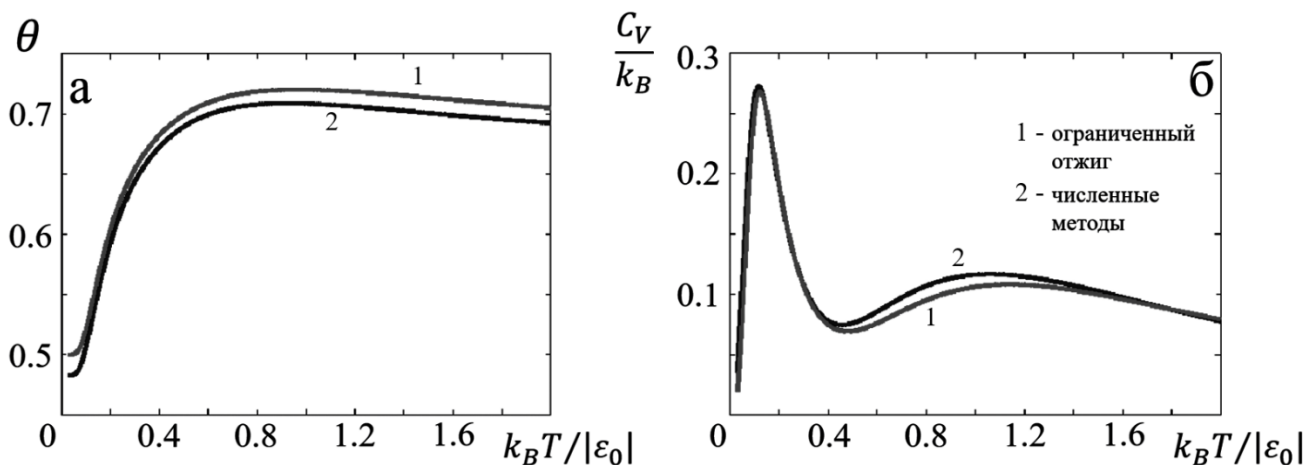


Рис. 2.5. Температурные зависимости (а) степени спиральности и (б) удельной теплоемкости для четырехлитерного случая с энергетическими параметрами $\epsilon = 3|\epsilon_0|/2$ и $\epsilon' = -|\epsilon_0|/10$ с вероятностями $p = q = 0.75$.

До сих пор мы рассматривали такие энергетические параметры ($\epsilon_0 < 0$, $\epsilon = 3|\epsilon_0|/2$), для которых получалась конкуренция между выгодными AU парами ($\epsilon_0 - \epsilon < 0$) и невыгодными AA и UU парами ($\epsilon_0 + \epsilon > 0$). Теперь рассмотрим неконкурентный случай, при котором $\epsilon_0 - \epsilon < 0$ и $\epsilon_0 + \epsilon < 0$. На Рис. 2.3 показано, что при $\epsilon_0 < 0$, $\epsilon = 2|\epsilon_0|/3$ для разных значений q теплоемкость по-прежнему имеет два пика, но холодная денатурация исчезает, а степень спиральности монотонно растет с понижением температуры.

В конкурентном случае низкотемпературный пик на графике удельной теплоемкости C_V был следствием того, что при понижении температуры число невыгодных пар оснований η^+ постепенно уменьшалось (вставка на Рис. 2.2(а)). В неконкурентном случае низкотемпературный пик является следствием обратного эффекта: число невыгодных пар оснований увеличивается при понижении температуры (вставка в Рис. 2.3(а)).

Рис. 2.4 описывает фолдинг РНК на фазовой диаграмме как функцию вероятности последовательности q и энергии невыгодной пары оснований $\epsilon_0 + \epsilon$ (энергия между парами AU и UA является более выгодной: $\epsilon_0 - \epsilon < 0$). Для конкурентного случая $\epsilon_0 + \epsilon > 0$ (верхняя половина) два пика удельной теплоемкости C_V являются следствием холодной и тепловой денатурации, а степень спиральности принимает свое максимальное значение при промежуточных температурах. Для неконкурентного случая $\epsilon_0 + \epsilon < 0$ (нижняя половина) опять получаются два пика удельной теплоемкости, но только высокотемпературный пик является следствием

плавления, а степень спиральности монотонно стремится к $\theta = 1$ при повышении температуры. Линии $q = 0$ и $q = 1$ соответствуют гомополимерному случаю, а линия симметрии $q = 1/2$ отоженному сценарию, которая приводит к поведению аналогичному случаю гомополимера. Линия $\epsilon_0 + \epsilon = 0$ соответствует случаю разбавленного замороженного беспорядка, когда происходит только тепловое плавление и удельная теплоемкость имеет только один пик.

Похожее поведение наблюдается и для четырехлитерного случая. На Рис. 2.5(а) и Рис. 2.5(б) мы сравнили степень спиральности и удельную теплоемкость полученную с помощью подхода ограниченного отжига и численных методов для (красная и синяя линии) $q = p = 0.75$. Мы видим, что двойной пик удельной теплоемкости получается и в четырехлитерном случае, так что наши результаты о холодной денатурации РНК являются надежными. Механизм экспериментально наблюдаемого феномена холодной денатурации РНК [51,52] широко обсуждался в контексте положительной разницы удельной теплоемкости между расплавленной и нативной структурами в результате гидрофобного эффекта [53,54], электростатики [51,52] и конкуренции между внутримолекулярными и межмолекулярными водородными связями [73]. Наши результаты показывают, что гетерогенность системы само по себе ослабляет тенденцию к образованию вторичной структуры при низких температурах для асимметричной функции распределения последовательности. Данный эффект имеет место только при конкурентных энергиях образования пар оснований и приводит к холодной денатурации, что позволяет говорить о дополнительных факторах, ответственных за явление холодной денатурации.

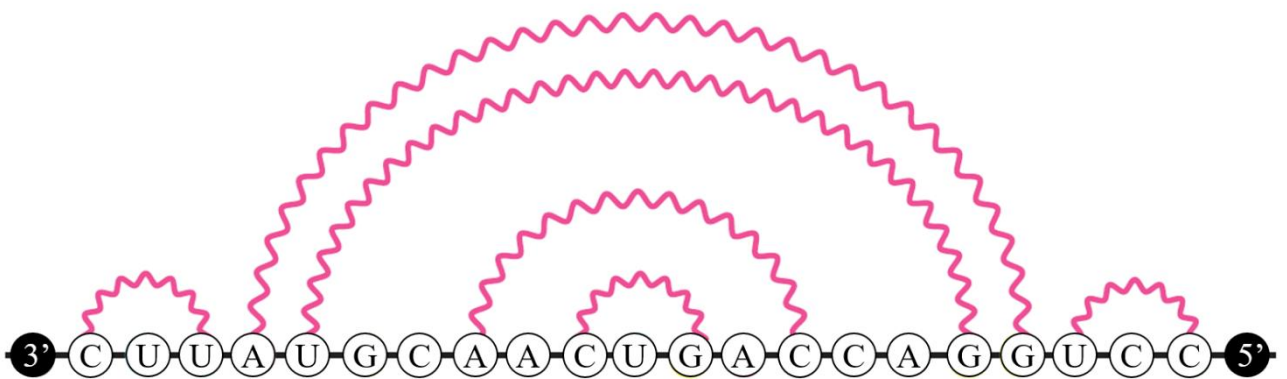
2.2. Четырехлитерная модель

Здесь мы рассмотрим так называемую модель «энергии связывания», т.е. мы пренебрегаем стэкинг взаимодействием и энтропийными потерями при образовании петель, и рассматриваем РНК состоящую из четырех видов оснований {U, A, C, G}. Возможны 10 видов водородных связей:

	U	A	C	G
U	UU	UA	UC	UG
A		AA	AC	AG
C			CC	CG
G				GG

Для однозначного описания каждого нуклеотида мы совмещаем два поля Изинга h_i и g_i , каждая для длины N . Следовательно каждый нуклеотид может быть описан следующим образом:

	$h_i = +1$	$h_i = -1$
$g_i = +1$	C	G
$g_i = -1$	U	A



h	+	+	+	-	+	-	+	-	-	+	+	-	-	+	+	-	-	-	+	+	+
g	+	-	-	-	-	+	+	-	-	+	-	+	-	+	+	-	+	+	-	+	+

Рис. 2.6. Последовательность вложенных друг в друга и независимых пар оснований для вторичной структуры РНК. Здесь молекула изображена в виде диаграммы непересекающихся дуг (фиолетовые волнистые линии), представляющие собой водородные связи. Последовательности генерируется с помощью двух независимых случайных величин h и g . Для этой конкретной последовательности мы использовали две равные вероятности $p_h = 0.5$ и $p_g = 0.5$.

Обозначим каждую последовательность РНК, как (h, g) , (Рис. 2.6), и кроме того, мы определим матрицу пар оснований S , которая полностью определяет вторичную структуру РНК, как $N \times N$ симметричную матрицу с компонентами $s_{i,j}$ равными единице, если (i, j) образовали связь и равной нулю в противном случае. Гамильтониан для конфигурации последовательности (h, g) записывается в виде суммы по неповторяющимся парам оснований

$$H(S, h, g) = \sum_{(i,j) \in S} \epsilon_{i,j} = \sum_{1 \leq i < j \leq N} s_{i,j} \epsilon_{i,j}, \quad (2.15)$$

где

$$\epsilon_{i,j} = \epsilon_0 + \epsilon_1 h_i h_j + \epsilon_2 g_i g_j. \quad (2.16)$$

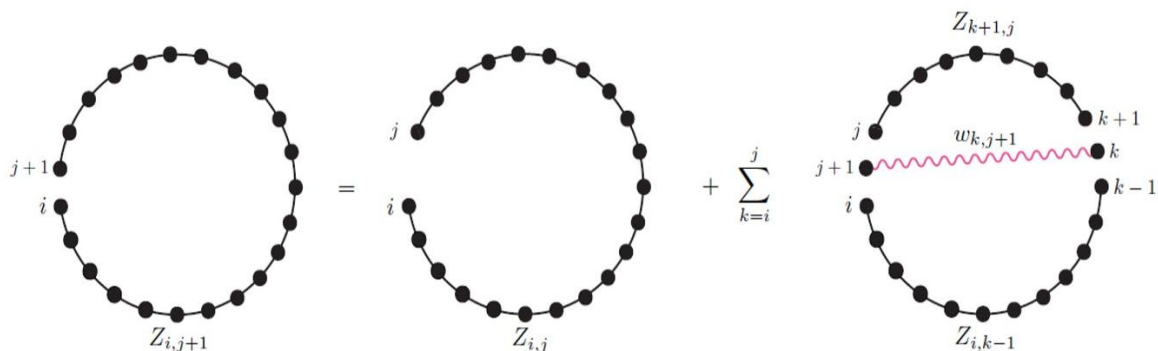


Рис. 2.7. Иерархическая рекурсивная схема
для функции распределения (2.28).

В отличие от элементов матрицы пар оснований $s_{i,j}$, переменные h_i и g_i заморожены и могут принимать только значения $+1$ или -1 и не могут измениться чтобы уменьшить суммарную энергию. Константа $\epsilon_0 < 0$ определяет природу фонового взаимодействия между нуклеотидами, которое рассматривается нами как притягивающее. Это гарантирует, что если $\epsilon_1 = \epsilon_2 = 0$, то степень спиральности вынуждена достигать своего максимального значения θ_{MAX} , при $T \rightarrow 0$, как в стандартном сценарии. Дополнительные константы $\epsilon_1 > 0$ и $\epsilon_2 < 0$ позволяют нам различать энергии связывания не Уотсон – Криковских пар, в отличие от ранее изученных спиновых решеточных моделей [74,75]. В частности, при этом выборе мы можем определить четыре фундаментальные силы

$$\epsilon_{GC} = \epsilon_0 - \epsilon_1 + \epsilon_2, \quad \text{для GC и UA,} \quad (2.17)$$

$$\epsilon_{GA} = \epsilon_0 + \epsilon_1 - \epsilon_2, \quad \text{для GA и CU,} \quad (2.18)$$

$$\epsilon_{GU} = \epsilon_0 - \epsilon_1 - \epsilon_2, \quad \text{для GU и CA,} \quad (2.19)$$

и

$$\epsilon_{GG} = \epsilon_0 + \epsilon_1 + \epsilon_2 \quad \text{для GG, CC, AA и UU.} \quad (2.20)$$

У нас есть дополнительное ограничение $\epsilon_1 > |\epsilon_2|$, так что все канонические пары оснований, включая пару GU имеют меньшую энергию, чем ϵ_{GG} .

Для каждой последовательности (h, g) возьмем в качестве h_i и g_i замороженные статистически – независимые случайные переменные, имеющие одинаковые функции распределения вида

$$P(h) = \prod_{i=1}^N \rho(h_i), \quad P(g) = \prod_{i=1}^N \rho(g_i). \quad (2.21)$$

где

$$\rho(h_i) = p_h \delta(h_i - 1) + (1 - p_h) \delta(h_i + 1), \quad (2.22)$$

и аналогично

$$\rho(g_i) = p_g \delta(g_i - 1) + (1 - p_g) \delta(g_i + 1). \quad (2.23)$$

Так как h и g взаимно независимы, вероятности раздельно нормализованы:

$$\sum_{\{h\}} P(h) = 1, \sum_{\{g\}} P(g) = 1. \quad (2.24)$$

Б. Функция распределения и степень спиральности

Функция распределения является суммой по всем допустимым реализациям матрицы пар оснований

$$Z_N(h, g) = \sum_{\{S\}} e^{-\beta H(S, h, g)}, \quad (2.25)$$

где $\beta = (-k_B T)^{-1}$. Индекс суммирования $\{S\}$ обозначает совокупность всех вторичных структур без псевдоузлов для данной последовательности (h, g) , то есть всех симметричных матриц $N \times N$ со строками и столбцами, каждый из которых содержит максимум один элемент, равный единице, а все остальные равны нулю. Свободная энергия

$$\beta f(h, g) = -\frac{1}{N} \ln Z_N(h, g) \quad (2.26)$$

явно зависит от беспорядка и не обеспечивает необходимую термодинамическую информацию для описания общего поведения системы. Величина, которую мы хотим получить это свободная энергия, полученная путем усреднения по всем неупорядоченным реализациям (h, g) , которая обозначается через $\overline{(\dots)}$. Так как беспорядок в нашей модели кодируется в виде двух отдельных переменных, то свободная энергия вычисляется как

$$\beta \overline{f(h, g)} = -\frac{1}{N} \overline{\ln Z_N(h, g)} = \beta \sum_{\{h, g\}} P(h) P(g) f(h, g), \quad (2.27)$$

где сумма берется по всем независимым реализациям $h_i = \pm 1$ и $g_i = \pm 1$.

Рекурсивный алгоритм вычисления функции распределения был представлен в [76] и определяется как

$$Z_{i, j+1} = Z_{i, j} + \sum_{k=i}^j \omega_{k, j+1} Z_{i, k-1} Z_{k+1, j}, \quad (2.28)$$

где величина $Z_{i,j}$ является статсуммой отрезка (i,j) , показанного на Рис. 2.7. Первый член в правой части уравнения соответствует вероятности того, что основание $j + 1$ не образовало пару. Второй член описывает все возможные конфигурации, если нуклеотид k внутри отрезка (i,j) образовал пару с нуклеотидом $j + 1$. Статвес этой пары

$$\omega_{k,j+1} = \exp(-\beta\epsilon_{k,j+1}). \quad (2.29)$$

Используя рекурсивное уравнение (2.28), мы можем определить функцию распределения всей цепи $Z_N = Z_{1,N}$ без псевдоузлов, с граничными условиями $Z_{i,i} = Z_{i,i-1} = 1$ для каждой i .

Конформация молекулы, описываемая ее фолдингом, может быть количественно определена с помощью степени спиральности как статвес оснований образовавших пару:

$$\theta = \frac{2}{N} \overline{|S|}, \quad (2.30)$$

где $\overline{\{\dots\}}$ - среднее по каноническому ансамблю. Число оснований образовавших пару в структуре S

$$|S| \equiv \sum_{i<j}^N s_{i,j}. \quad (2.31)$$

Так как $\overline{|S|} \in [0, N/2]$, то степень спиральности является функцией от температуры и находится в интервале $[0, 1]$ с минимальными и максимальными значениями при отсутствии пар оснований или если все основания образовали пары соответственно. Отметим, что

$$\begin{aligned} \overline{|S|} &= \frac{1}{Z_N(h, g)} \sum_{\{S\}} |S| e^{-\beta\epsilon_0 |S|} e^{-\beta\epsilon_1 \sum_{i<j} s_{i,j} h_i h_j} e^{-\beta\epsilon_2 \sum_{i<j} s_{i,j} g_i g_j} \\ &= \left(-\frac{\partial}{\partial(\beta\epsilon_0)}\right) \ln Z_N(h, g), \end{aligned} \quad (2.32)$$

степень спиральности может быть выражена для замороженного беспорядка как

$$\theta = 2 \frac{\partial}{\partial(\beta \epsilon_0)} \overline{\beta f(h, g)}. \quad (2.33)$$

Численно она может быть вычислена с помощью вероятности образования пары между основаниями i и j , которая определяется как [74]

$$P_{i,j} = \langle s_{i,j} \rangle = \omega_{i,j} \frac{Z_{i+1,j-1} Z_{j+1,N+i-1}}{Z_{1,N}}, \quad (2.34)$$

откуда получим

$$\theta = \frac{2}{N} \overline{\sum_{i < j} P_{i,j}}. \quad (2.35)$$

Здесь $Z_{i+1,j-1}$ является статсуммой внутренней последовательности $(i+1, \dots, j-1)$, а $Z_{j+1,N+i-1}$ – статсуммой внешней последовательности $(1, \dots, i-1, j+1, \dots, N)$, см. Рис. 2.8, где последняя величина может быть получена с помощью расширения рекуррентного соотношения на сдублированную последовательность $(1, \dots, N, N+1, \dots, 2N)$.

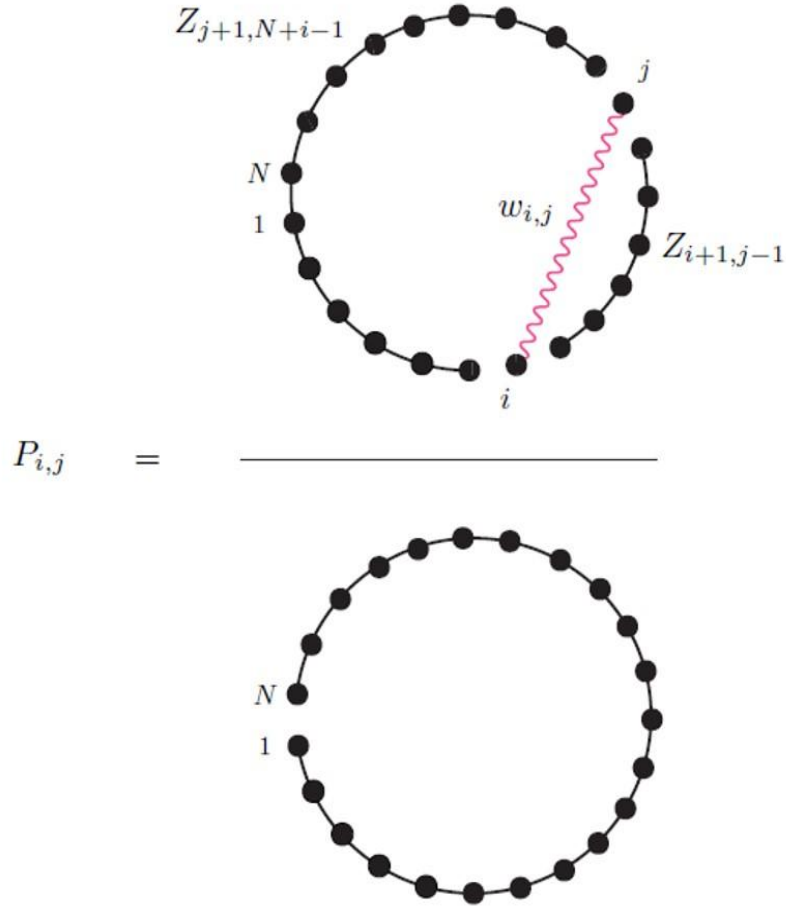


Рис. 2.8. Схематическое представление вероятности образования пар оснований $P_{i,j}$ в уравнении (2.34).

Если $\epsilon_1 = \epsilon_2 = 0$, то из уравнения (2.15) мы получим, что энергия связывания независима: $\epsilon_{i,j} = \epsilon_0$. В этом случае энергия структуры S будет зависеть только от числа пар оснований и статистическая сумма будет оцениваться как [74]

$$Z_N = \sum_{\{S\}} \omega^{|S|} \sim \xi(\omega) N^{-3/2} (1 + 2\sqrt{\omega})^N, \quad (2.36)$$

где $\xi(\omega)$ является функцией масштабирования для статистического веса гомополимера

$$\omega = \exp(-\beta\epsilon_0). \quad (2.37)$$

В этом случае выражение степени спиральности непосредственно получается из статсуммы для гомополимера

$$\theta^{hom} = \frac{2}{N} \langle |S| \rangle = \frac{2}{N} \frac{\partial \ln Z_N}{\partial \ln \omega}. \quad (2.38)$$

Используя $\omega \partial / \partial \omega = \partial / \partial \ln \omega$, получаем

$$\theta^{hom} = \frac{2}{N} \omega \frac{\partial \ln Z_N}{\partial \omega} \approx \frac{2\sqrt{\omega}}{1+2\sqrt{\omega}}. \quad (2.39)$$

При высоких температурах

$$\theta_{\infty} \equiv \lim_{T \rightarrow \infty} \theta^{hom} = \frac{2}{3}, \quad (2.40)$$

в то время как значение для нулевой температуры $\theta_0 \equiv \lim_{T \rightarrow 0} \theta^{hom}$ равно единице для притягивающего фонового взаимодействия, где $\epsilon_0 < 0$, и равно нулю в противном случае.

2.3. Метод ограниченного отжига

Самым известным подходом для изучения систем с замороженным беспорядком, таких как спиновые стекла является метод реплик. Основная идея заключается в том, чтобы вместо усреднения логарифма функции распределения (2.27) получить гораздо легче вычисляемое среднее для случая расплавленного беспорядка, которое определяется как

$$\beta f^a = -\frac{1}{N} \ln \overline{Z_n(h, g)} = -\frac{1}{N} \ln Z_N^a. \quad (2.41)$$

Другой подход был введен в работе [77] и был успешно применен для случайного поля модели Изинга, так называемый подход ограниченного отжига. Метод работает следующим образом. Объединяя случайную величину h с множителем Лагранжа μ_h , и самоусредняемой функцией α_h [78], среднее от логарифма можно аппроксимировать логарифмом от статистической суммы, усредненной с весом, зависящим от вариационного параметра μ_h . Таким образом, мы можем сделать более простое усреднение с наложением соответствующих ограничений. С помощью этих ограничений можем свести к минимуму разницу между точным значением замороженной

средней и величиной, полученной этим метод. Таким образом получим довольно хорошую нижнюю оценку f^a .

Формализм для полного алфавита РНК

Согласно [75] строим две независимые функции переменных h и g , вычитая среднее значение из их первого момента функции распределения

$$\alpha_{[h,g]} = \frac{1}{N} \sum_{i=1}^N [[h,g]_i - (2p_{[h,g]} - 1)]. \quad (2.42)$$

Обе эти функции являются интенсивными и самоусредняемыми к нулю, что следует непосредственно из

$$\overline{[h,g]} = \sum_{\{[h,g]\}} P([h,g])[h,g] = \sum_{[h,g]=\pm 1} \rho([h,g])[h,g] = 2p_{[h,g]} - 1. \quad (2.43)$$

Записав функцию статистическую сумму в виде

$$Z_N(h,g) = \sum_{\{S\}} e^{-\beta \epsilon_0 |S|} \prod_{k < l}^N \exp(-\beta \epsilon_1 s_{k,l} h_k h_l) \times \\ \times \prod_{m < n}^N \exp(-\beta \epsilon_2 s_{m,n} g_m g_n), \quad (2.44)$$

получаем:

$$Z_N^{ca}(\mu_h, \mu_g) = \exp[N\mu_h(2p_h - 1)] \exp[N\mu_g(2p_g - 1)] \times \\ \times \sum_{\{S\}} e^{-\beta \epsilon_0 |S|} \prod(\mu_h, \mu_g), \quad (2.45)$$

где

$$\prod(\mu_h, \mu_g) = \sum_{\{h,g\}} \prod_{i=1}^N \rho(h_i) e^{-\mu_h h_i} \rho(g_j) e^{-\mu_g g_j} \times \\ \times \prod_{k < l}^N \exp(-\beta \epsilon_1 s_{k,l} h_k h_l) \prod_{m < n}^N \exp(-\beta \epsilon_2 s_{m,n} g_m g_n).$$

Ключевым элементом в данном случае является тот факт, что произведение дает вклад равный единице только если $s_{k,l} = 1$, или $s_{m,n} = 1$ и что каждое основание может участвовать в образовании только одной пары. При усреднении по беспорядку $|S|$ раз возникают множители вида $e^{-\beta \epsilon_1 h_k h_l}$ и

$e^{-\beta\epsilon_2 g_m g_n}$. Т.о., каждый член усредняется независимо, как это следует из взаимной независимости переменных h и g . Получаем

$$\begin{aligned} \prod(\mu_h, \mu_g) &= \left(\sum_{h=\pm 1} \rho(h) e^{-\mu_h h} \sum_{g=\pm 1} \rho(g) e^{-\mu_g g} \right)^{N-2|S|} \times \\ &\times \left(\sum_{h, h'=\pm 1} \rho(h)\rho(h') e^{-\mu_h h} e^{-\mu_{h'} h'} e^{-\beta\epsilon_1 h h'} \right)^{|S|} \times \\ &\times \left(\sum_{g, g'=\pm 1} \rho(g)\rho(g') e^{-\mu_g g} e^{-\mu_{g'} g'} e^{-\beta\epsilon_1 g g'} \right)^{|S|} = \\ &= \Omega_h^N(\mu_h) \Omega_g^N(\mu_g) \left(\frac{\gamma_h(\mu_h)}{\Omega_h^2(\mu_h)} \frac{\gamma_g(\mu_g)}{\Omega_g^2(\mu_g)} \right)^{|S|}, \end{aligned}$$

где используются следующие сопряженные величины

$$\begin{aligned} Y_{[h,g]}(\mu_{[h,g]}) &= e^{-\beta\epsilon_{[1,2]}} p_{[h,g]}^2 e^{-2\mu_{[h,g]}} + \\ &+ e^{-\beta\epsilon_{[1,2]}} (1 - p_{[h,g]})^2 e^{2\mu_{[h,g]}} + \\ &+ e^{-\beta\epsilon_{[1,2]}} 2p_{[h,g]}(1 - p_{[h,g]}), \end{aligned} \quad (2.46)$$

и

$$\Omega_{[h,g]}(\mu_{[h,g]}) = p_{[h,g]} e^{-\mu_{[h,g]}} + (1 - p_{[h,g]}) e^{\mu_{[h,g]}} \quad (2.47)$$

для h и g соответственно. Мы можем написать это в более простой форме, если введем две новые константы энергии взаимодействия

$$\overline{\epsilon}_{[h,g]}(\mu_{[h,g]}) = -\frac{1}{\beta} \ln \frac{\gamma_{[h,g]}(\mu_{[h,g]})}{\Omega_{[h,g]}^2(\mu_{[h,g]})}, \quad (2.48)$$

где вся информация об усреднении по беспорядку теперь включена в вариационные параметры μ_h и μ_g . Тогда выражение (45) сводится к

$$Z_N^{ca}(\mu_h, \mu_g) = \Xi(\mu_h, \mu_g) Z_N^{hom}(\mu_h, \mu_g), \quad (2.49)$$

где

$$\Xi(\mu_h, \mu_g) = e^{N\mu_h(2p_h-1)} e^{N\mu_g(2p_g-1)} \Omega_h^N(\mu_h) \Omega_g^N(\mu_g).$$

Здесь величина $Z_N^{hom} = \sum_{\{S\}} [\omega^{ca}]^{|S|}$ является статсуммой гомополимера при ограниченном отжиге с статвесом $\omega^{ca} = e^{-\beta\epsilon^{ca}}$. Энергия взаимодействия пары

$$\epsilon^{ca}(\mu_h, \mu_g) = \epsilon_0 + \bar{\epsilon}_h(\mu_h) + \bar{\epsilon}_g(\mu_g). \quad (2.50)$$

При максимизация свободной энергии $f^{ca}(\mu_h, \mu_g)$ по параметрам μ_h и μ_g получаем

$$f^{ca}(\mu_h, \mu_g) = -\mu_h(2p_h - 1) - \mu_g(2p_g - 1) - \ln \Omega_h(\mu_h) - \ln \Omega_g(\mu_g) - \ln \left(1 + 2\sqrt{\omega^{ca}(\mu_h, \mu_g)} \right), \quad (2.51)$$

Здесь решаются уравнения вида

$$1 - 2p_{[h,g]} = \frac{\partial \ln \Omega_{[h,g]}}{\partial \mu_{[h,g]}} - \frac{2\sqrt{\omega^{ca}}}{1+2\sqrt{\omega^{ca}}} \left(\frac{\partial \ln \Omega_{[h,g]}}{\partial \mu_{[h,g]}} - \frac{1}{2} \frac{\partial \ln Y_{[h,g]}}{\partial \mu_{[h,g]}} \right), \quad (2.52)$$

для h и g соответственно. Тогда, для степени спиральности в приближении ограниченного отжига, получаем

$$\theta^{ca}(\mu_h, \mu_g) = \frac{2\sqrt{\omega^{ca}(\mu_h, \mu_g)}}{1+2\sqrt{\omega^{ca}(\mu_h, \mu_g)}} \quad (2.53)$$

Расплавленный и гомополимерный случаи

Нас интересуют два предельных случая, которые получаются при $p_{[h,g]} = 0.5$ и $p_{[h,g]} = 1$, которые описывают расплавленный беспорядок и случай гомополимера соответственно. Действительно, если $p_h = p_g = 0.5$, то

$$\Omega_{[h,g]}(\mu_{[h,g]}) = \cosh \mu_{[h,g]} \quad (2.54)$$

и

$$Y_{[h,g]}(\mu_{[h,g]}) = \frac{1}{2} (e^{-\beta\epsilon_1} \cosh 2\mu_{[h,g]} + e^{\beta\epsilon_1}). \quad (2.55)$$

Из уравнения (2.52) получаем, что

$$0 = \tanh \mu_{[h,g]} \left(1 + \frac{\cosh 2\mu_{[h,g]}}{\cosh 2\mu_{[h,g]} + e^{2\beta\epsilon_1}} \frac{4\sqrt{\omega^{ca}}}{1 + \tanh^2 \mu_{[h,g]}} \right),$$

единым решением которого является $\mu_h = \mu_g = 0$. В этом случае статсумма, полученная методом ограниченного отжига сводится к статсумме для расплавленного случая:

$$Z_N^{ca}(0,0) = \overline{Z_N(h,g)} = Z_N^a. \quad (2.56)$$

Следовательно, свободная энергия ограниченного отжига не улучшает нижнюю оценку в расплавленном случае.

В случае гомополимера ($p_{[h,g]} = 1$), мы получаем

$$\Omega_{[h,g]}(\mu_{[h,g]}) = e^{-\mu_{[h,g]}} \quad (2.57)$$

и

$$Y_{[h,g]}(\mu_{[h,g]}) = e^{-\beta\epsilon_{[1,2]}} e^{-2\mu_{[h,g]}}. \quad (2.58)$$

Таким образом получаем, что $\epsilon^{ca} = \epsilon_{GG} = \epsilon_0 + \epsilon_1 + \epsilon_2$ и уравнение (2.52) имеет место для всех $\mu_{[h,g]} \in \mathbb{R}$.

2.4. Низкотемпературная поведение

Холодовая денатурация

Для изучения общего поведения четырехлитерной модели мы выбираем энергию ϵ_{GG} и исследуем разные точки на плоскости $\{p_h, p_g\} \in [0,1]$, показанной на Рис. 2.9. В каждой области, помеченной римскими цифрами каждая точка является зеркально симметричной относительно обеих осей, соединяющих точку $(0.5,0.5)$ с точками $(0.0,0.5)$ и $(0.5,0.0)$ соответственно. Следовательно, для любой точки $(p_h, p_g) \in [0,1]$, точки

$$(1 - p_h, p_g) \Delta (1 - p_h, 1 - p_g) \Delta (p_h, 1 - p_g) \quad (2.59)$$

симметричны и внутри этих двух наборов областей {I, IV, V, VIII} и {II, III, VI, VII} дают одинаковые результаты. В дальнейшем мы представим результаты из областей I и II для каждого набора энергетических степеней свободы помня о лежащей в основе этого симметрии.

Для простоты возьмем $\epsilon_0 = -1$, составляя привлекательный фон взаимодействия, а также рассмотрим возможные значения отношения

$$\Lambda \equiv \left| \frac{\epsilon_1}{\epsilon_0 + \epsilon_2} \right| \quad (2.60)$$

для разных точек $\{p_n, p_g\}$. Другое условие, накладываемое на энергетические параметры состоит в том, что энергии как Уотсон – Криковские пар, так и пар между одноименными основаниями (2.19) должны быть отрицательными

$$\Lambda > \frac{\epsilon_2 - \epsilon_0}{\epsilon_2 + \epsilon_0}. \quad (2.61)$$

На Рис. 2.10. (а) (с) мы сравнили удельную теплоемкость

$$C_V^{ca}(\tilde{\mu}_n, \tilde{\mu}_g) = -T \frac{\partial^2 f^{ca}(\tilde{\mu}_n, \tilde{\mu}_g)}{\partial T^2} \quad (2.62)$$

и степень спиральности $\theta^{ca}(\tilde{\mu}_n, \tilde{\mu}_g)$ для ограниченного отжига при $\Lambda = 1/2$ с точным замороженным средним полученным для 50 случайных последовательностей с длиной $N = 100$. Выбранная точка на плоскости вероятностей $(p_n, p_g) = (0.75, 0.75)$ на границе между областями I и II не влияет на поведение спиральности при низких температурах и теплоемкость имеет основной пик при $T \approx 0.7$, а также менее выраженный пик при $T \approx 0.2$.

Для той же точки на плоскости вероятности $(p_n, p_g) = (0.75, 0.75)$, при увеличении значения $\Lambda = 2$ так, чтобы $\epsilon_{GG} > 0$, ситуация резко меняется. Теперь степень спиральности не достигает своего максимума, когда температура понижается, но стабилизирует саму себя при зафиксированном

значении $\theta_0 \approx 0.5$, см. Рис. 2.10 (d). Удельная теплоемкость ведет себя так же, как перед двойным пиком в предыдущем случае, но теперь низкотемпературный пик более выраженный, чем высокотемпературный. Высокотемпературный пик почти исчезает по сравнению к случаю, когда $\epsilon_{GG} < 0$.

При $\Lambda > 1$, проявляется эффективная конкуренция между GC и парами с одинаковой буквой и наблюдается холодовая денатурация. В этом случае вторичная структура РНК при низких температурах ослабевает за счет замороженного беспорядка в первичной структуре. При неконкурентном режиме ($\Lambda < 1$) холодовая денатурация не наблюдается и мы имеем только высокотемпературное плавление, что видно на Рис. 2.11, где мы сравнили удельную теплоемкость и степень спиральности на границе между областями I и II при разных вероятностях (p_n, p_g) . Из Рис. 2.11 видно, что в этом случае удельная теплоемкость всегда имеет два пика. Критическое значение $\Lambda = 1$ является точкой, при которой низкотемпературный пик становится доминирующим и для которого $\theta_0 \neq 1$. При $\Lambda > 1$ пик растет, а степень спиральности принимает свое минимальное значение $\theta_0 = 0$.

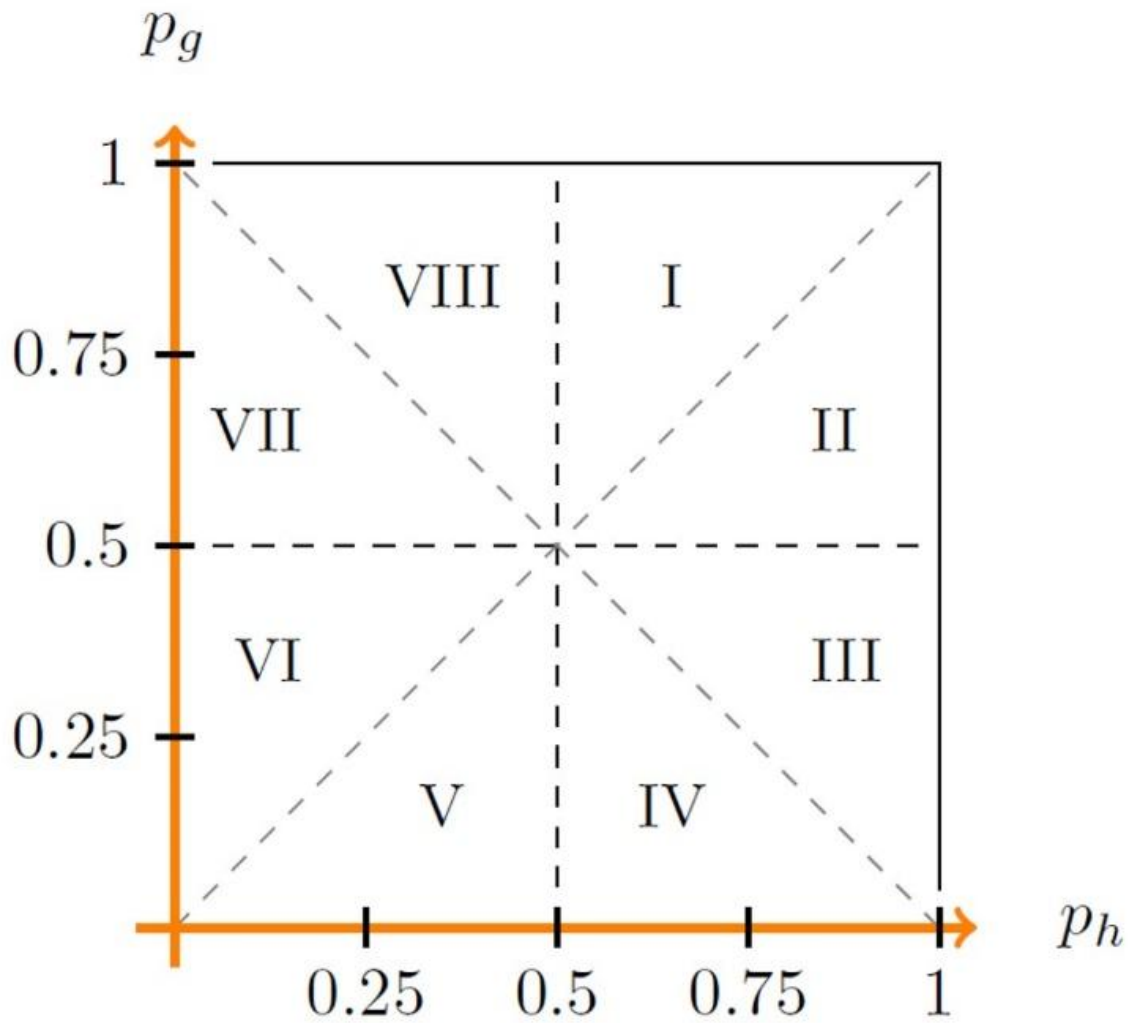


Рис. 2.9. Плоскость сечения $\{p_h, p_g\} \in [0,1]$ для зафиксированной энергии ϵ_{GG} . Результаты для точек на диагональных линиях симметричны относительно осей p_h и p_g и центра симметрии $(0.5,0.5)$.

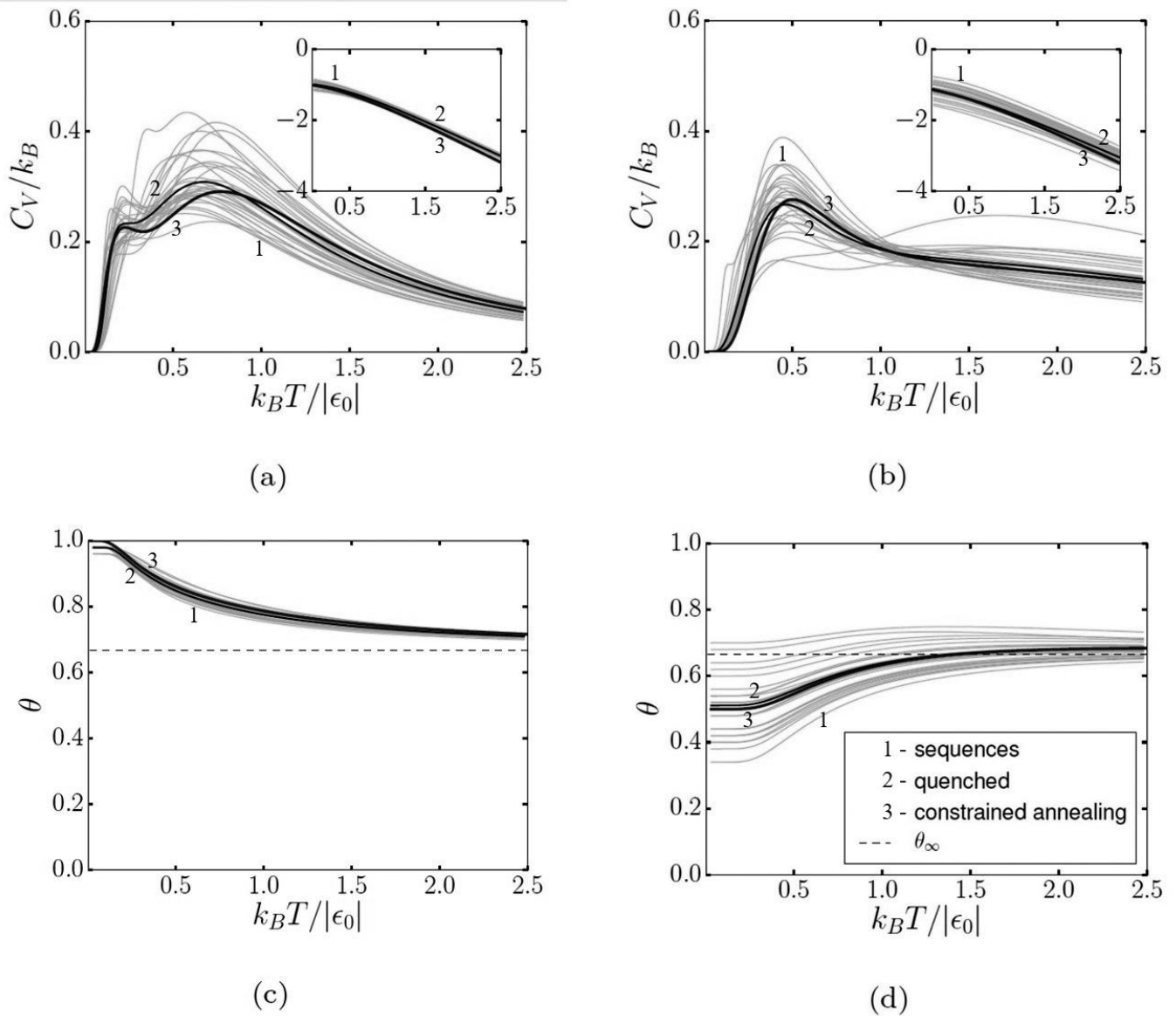


Рис. 2.10. Удельная теплоемкость (a) (b) и степень спиральности (c) (d) для неконкурентных и конкурентных сценариев соответственно. Замороженная средняя (черная сплошная линия) получается из 50 случайных последовательностей (красные линии) с $N = 100$ и $(p_h, p_g) = (0.75, 0.75)$. Параметры энергии $\epsilon_0 = -1$, $\epsilon_1 = |\epsilon_0|$ и $\epsilon_2 = -|\epsilon_0|$ для (a) (c) и $\epsilon_0 = -1$, $\epsilon_1 = 3.0|\epsilon_0|$ и $\epsilon_2 = -0.5|\epsilon_0|$ для (b) (d). Верхняя вкладка на рисунке с теплоемкостью показывает поведение свободной энергии.

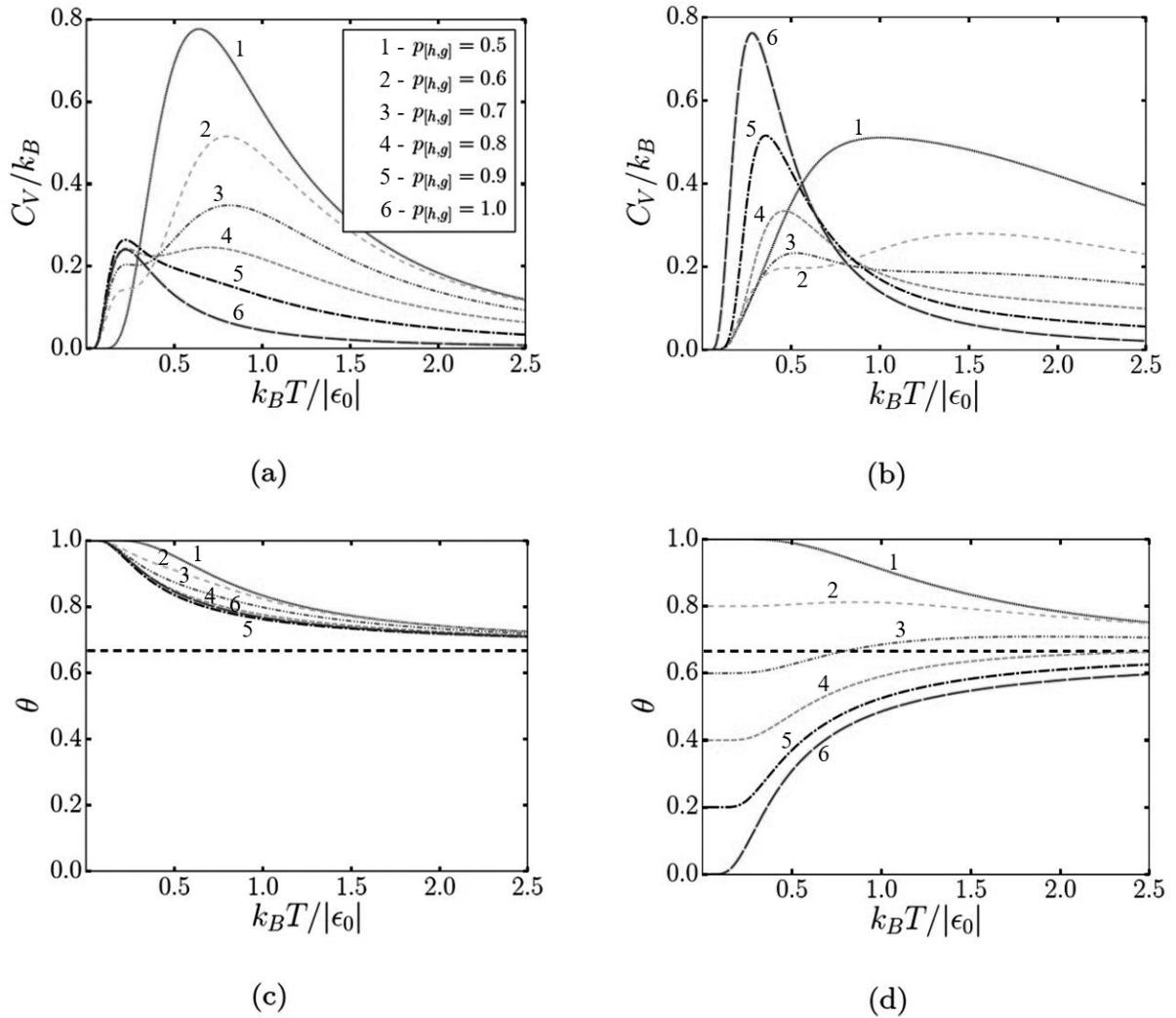


Рис. 2.11. Удельная теплоемкость (a) (b) и степень спиральности (c) (d) для разных вероятностей $0.5 \leq (p_h, p_g) \leq 1.0$ в случае ограниченного отжига. Графики (a) и (c) получены для $\Lambda = 1/2$ и $\epsilon_{GG} = -|\epsilon_0|$, а графики (b) и (d) - для $\Lambda = 2$ и $\epsilon_{GG} = 1.5|\epsilon_0|$. Пунктирная черная линия определяет асимптотическое значение ур. (26) в высокотемпературной расплавленной фазе.

Конечные скачки в графике удельной теплоемкости

Введем более высокий уровень сложности для функции энергии, чем гомополимер с двумя дополнительными степенями свободы ϵ_1 и ϵ_2 . Этому соответствует новый тип поведения в изменении внутренней энергии относительно температуры. Как будет показано, три пика на графике

удельной теплоемкости появляются как при конкурентном, так и при неконкурентном режиме до тех пор, пока выполняются определенные условия для Λ . В частности, третий пик появляется в режиме низких температур если

$$\frac{\epsilon_0}{\epsilon_0 + \epsilon_2} < \Lambda < \frac{\epsilon_0 + 2\epsilon_2}{\epsilon_0 + \epsilon_2}, \quad \Lambda \neq 1. \quad (2.63)$$

Эквивалентно это будет иметь место только при условии, что $|\epsilon_{GG}| < \epsilon_2$ и $\epsilon_{GG} \neq 0$, в то время как для критической энергии $\epsilon_{GG} = 0$ график с двумя пиками восстановится. Это видно на Рис. 2.12 и Таб. 1, где мы сравниваем удельную теплоемкость и поведение спиральности при $(p_h, p_g) = (0.75, 0.75)$.

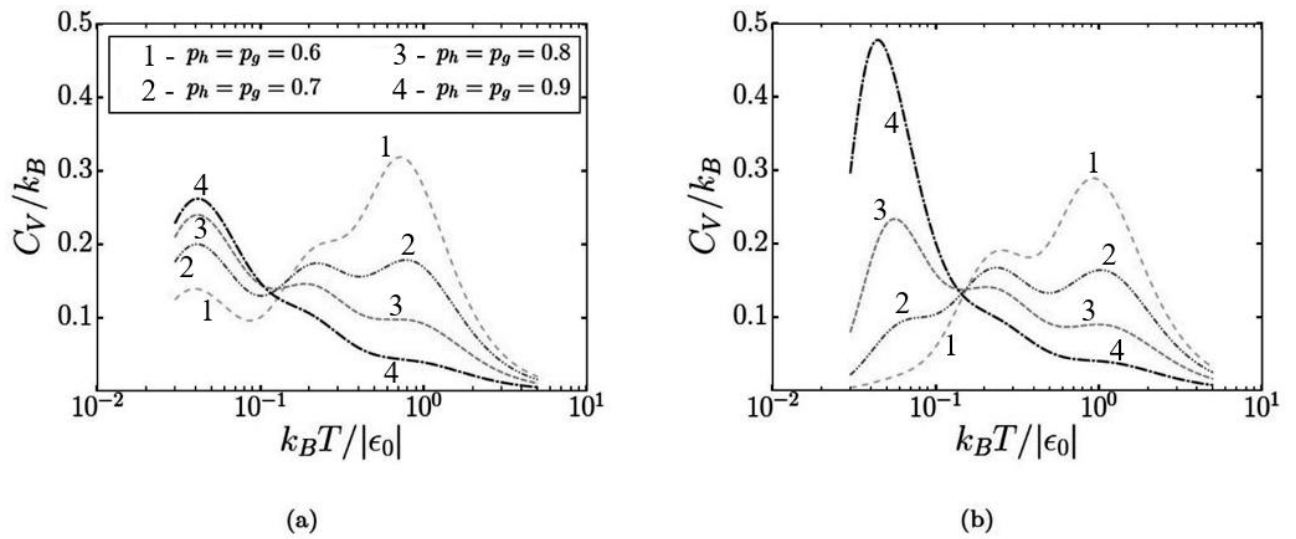


Рис. 2.12. Удельная теплоемкость для разных вероятностей $0.5 < (p_h, p_g) < 1.0$ в случае ограниченного отжига для $\epsilon_0 = -1$ и $\epsilon_2 = -0.3|\epsilon_0|$. Графики (a) получен для $\epsilon_1 = 1.1|\epsilon_0|$ и $\epsilon_{GG} = -0.2|\epsilon_0|$, а график (b) – для $\epsilon_1 = 1.5|\epsilon_0|$ и $\epsilon_{GG} = 0.2|\epsilon_0|$.

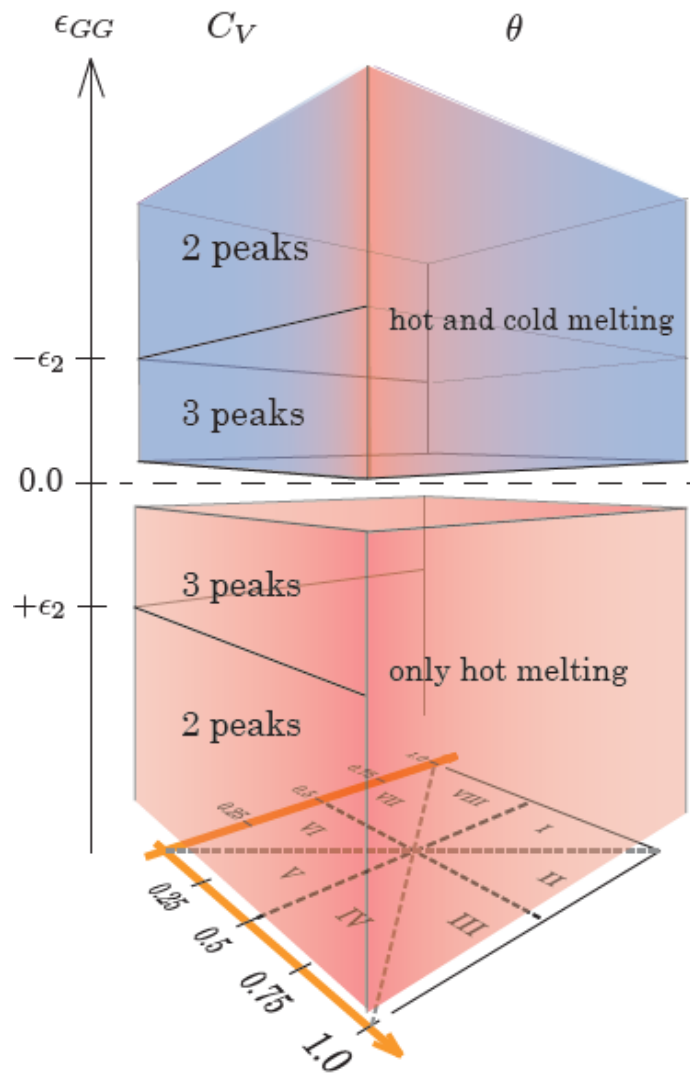


Рис. 2.13. Глобальная фазовая диаграмма, описывающая температурное поведение теплоемкости и степени спиральности в трехмерном пространстве $\{\epsilon_{GG}, p_h, p_g\}$. По вертикали через точку $(p_h, p_g) = (0.5, 0.5)$ проходит линия симметрии, соответствующая замороженной модели с разбавлением. На четырех ребрах параллелепипеда теплоемкость имеет всего один пик для любых значений энергий взаимодействия.

Физическая интерпретация такого поведения может быть найдена в микроскопическом описании конечных скачков удельной теплоемкости. На самом деле каждый пик связан с наличием микроскопической степени свободы в гамильтониане. Так же, как для двухатомных молекул, где C_V зависит от поступательных, вращательных и колебательных квантовых

степеней свободы, в нашей модели три пика отражают факт наличия трех степеней свободы ϵ_0 , ϵ_1 и ϵ_2 . Если мы возьмем $\epsilon_1 = \epsilon_2 = 0$, то восстановим случаи гомополимера РНК с одним пиком. Когда мы добавляем $\epsilon_1 \neq 0$ или $\epsilon_2 \neq 0$, то появляется второй пик [75]. Если все три энергии ненулевые, то у нас появляется третий пик, когда выполняется условие (2.63). Проявление холодной денатурации и поведение теплоемкости в зависимости от ϵ_{GG} и концентраций (p_h, p_g) иллюстрируется на Рис. 2.13.

Гл.3. Коллапс и гибридизация оц-РНК: неравновесное стационарное состояние

Гибридизация одноцепочечных нуклеиновых кислот является ключевым процессом в любой активности ДНК-сенсора или РНК-сенсора [79] и во многих методах молекулярной биологии, таких как полимеразная цепная реакция (ПЦР) [80], *in situ* гибридизация [81], и т.д.

Силы, которые стабилизируют вторичную структуру РНК, сильнее, чем взаимодействия, ответственные за третичную структуру, и, следовательно, эти две структуры характеризуются двумя различными энергетическими масштабами. Согласно одной из принятых в настоящее время концепций фолдинга РНК, сначала формируются элементы вторичной структуры, такие как спирали, петли, шпильки, а затем складываются вместе, образуя трехмерную третичную структуру [82]. Это так называемый иерархический механизм фолдинга.

Однако некоторые эксперименты показывают, что скорость фолдинга больших РНК ниже, чем прогнозируется иерархическим механизмом [83]. Это может означать, что два последовательных этапа фолдинга не полностью независимы. Ландшафт энергетической функции большой РНК чрезвычайно прочен, содержит несколько глубоких минимумов, которые действуют как кинетические ловушки в путях фолдинга [83]. Молекула может оставаться в ловушке в состояниях, отличных от нативной структуры для периодов времени, даже превышающих среднее время жизни РНК в живой клетке [84,85]. Этот эффект был назван эффектом памяти при самоорганизации РНК. В отличие от проблем фолдинга белка и расщепления ДНК, которые теоретически интенсивно изучались [86–96], до настоящего времени не существует полностью надежной теоретической модели образования третичной структуры РНК. Следует отметить, что были разработаны отличные модели для описания образования вторичной

структуры РНК [97–102]. Однако полный путь к формированию третичной структуры остается относительно мало изученным. В частности, взаимное взаимодействие между вторичной и третичной структурами требует более четкого описания [103–107].

Вторичная структура РНК определяется парами образованными между нуклеотидами. В [106–108] показано, что характерная вторичная структура РНК может быть представлена в виде дерева, ветви которого образуются относительно короткими двухцепочечными спиралями. Это описание выглядит очень подходящим, и модель отражает многие аспекты вторичной структуры. Однако здесь вопрос об ансамбле последовательностей, совместимых с этой схемой, остается без ответа. Иерархический сценарий фолдинга был изучен, например, в [106], где фолдинг РНК с фиксированной вторичной структурой описывается моделью разветвленного полимера с замороженными случайными ветвлениями. В [107] для изучения равновесных характеристик РНК была использована модель отождествленного полимера со случайными ветвлениями. Полностью отождествленные разветвленные структуры описывают ансамбль вторичных структур, в котором третичная структура формируется в результате существенных перестроек вторичных структурных элементов. Этот сценарий типичен для больших РНК, к которым, скорее всего, не применим механизм иерархического фолдинга. Хотя модель полимера со случайными отождествленными разветвлениями может быть успешно применена для изучения равновесных особенностей фолдинга РНК, она не может описать кинетику фолдинга так же удачно, как и равновесные особенности. В то же время кинетические эффекты в процессе фолдинга считаются очень важными из-за существования долгоживущих промежуточных состояний, как упомянуто выше.

3.1. Модель оц-РНК, как случайно разветвленного полимера

Рассмотрим следующий мысленный эксперимент. Молекула одноцепочечной РНК сначала растворяется в растворителе при температуре T' , которая удовлетворяет неравенствам $\theta < T' < T_m$, где T_m – является температурой плавления, а θ – температура Флори [110]. В этих условиях молекула РНК представляет собой статистический клубок с хорошо определенной вторичной структурой. Затем перенесем очень небольшое количество нашего раствора, содержащего РНК, в такой же растворитель, но с температурой T так, чтобы $T < \theta$. Теперь содержание вторичной и третичной структур не соответствует температуре термостата, и система оказывается в неравновесном состоянии. Вначале вторичная и пространственная структуры в этом состоянии все еще соответствуют температуре T' но они начинают релаксировать к состоянию соответствующему новой температуры T . В конце процесса РНК придет в компактное глобулярное состояние с некоторой вторичной структурой.

Третичная структура РНК стабилизируется взаимодействием между различными элементами вторичной структуры: спиральными участками, шпильками, внутренними петлями, несоответствиями и т.д. (Рис. 3.1). Взаимодействия между спиральными стеблями можно считать однородными, так как азотистые основания расположены внутри двойной спирали.

Чтобы огрубленно описать конформацию РНК, рассмотрим каждый непарный участок РНК как эффективный мономер. Мономеры пронумерованы $i = 1, \dots, N$. Центр масс мономера i помещается в точке с координатами x_i в d -мерном пространстве. В терминах случайно разветвленного полимера вторичная структура описывается матрицей $\hat{B} = \|\|B_{ij}\|\|$, где $B_{ij} = 1$, если i -ый и j -ый мономеры связаны спиральными участками и $B_{ij} = 0$, в противном случае. На

Рис. 3.2 показан фрагмент цепочки с описанием матрицы \hat{B} . Как и в [106,108], здесь мы применяем простейшее приближение, в котором молекула моделируется как корневое дерево с координационным числом, равным трем. Спирали между мономерами моделируются как пружины. Введем следующий гамильтониан для модели:

$$H = H_{II} + \sum_{i < j} v_{ij} \delta(x_i - x_j) + V_{conf}, \quad (3.1)$$

где член $H_{II}(\{x_i\}; \hat{B}) = \frac{dT}{2l^2} \sum_{i < j} B_{ij} (x_i - x_j)^2$ описывает упругость спиральной пружины между i -м и j -м эффективными мономерами, а l является равновесным расстоянием между соседними мономерами. Здесь оно совпадает со средней длиной спирального участка. Предполагается, что постоянная упругости спиральной пружины равна $\frac{dT}{l^2}$. Здесь v_{ij} – второй вириальный коэффициент взаимодействия между i -м и j -м эффективными мономерами, а $V_{conf}(\{x_i\})$ – сжимающий потенциал, описывающий однородное притяжение между спиральными участками.

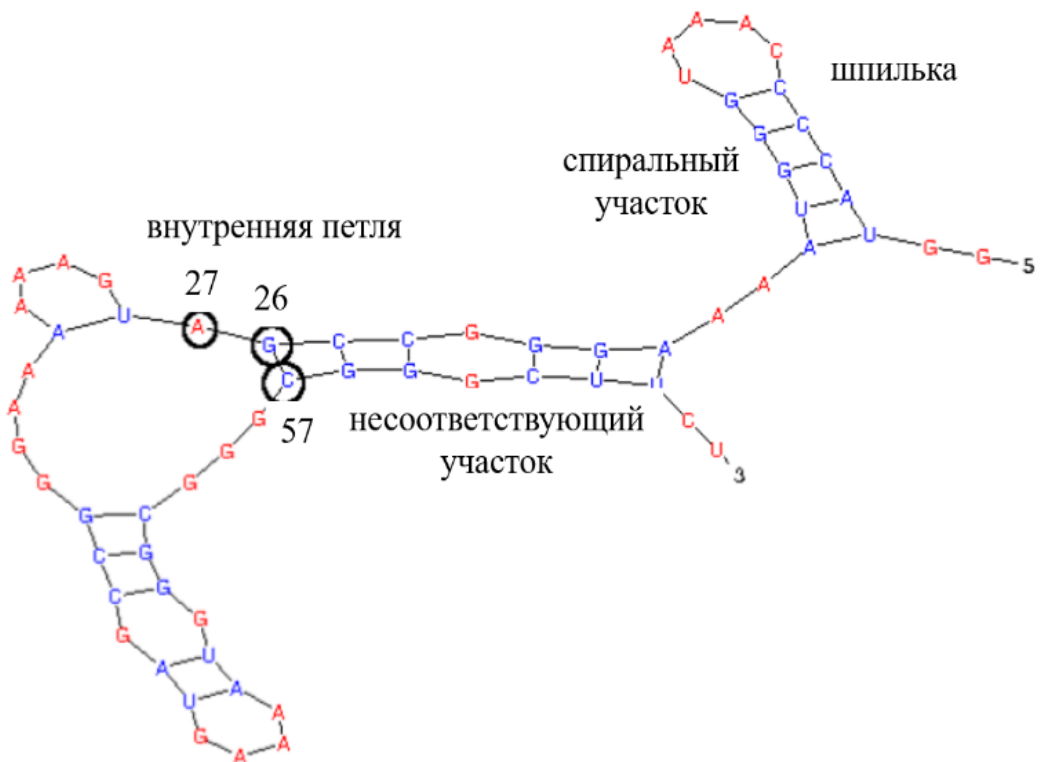


Рис. 3.1. Фрагмент цепи РНК с типичными фрагментами вторичной структуры

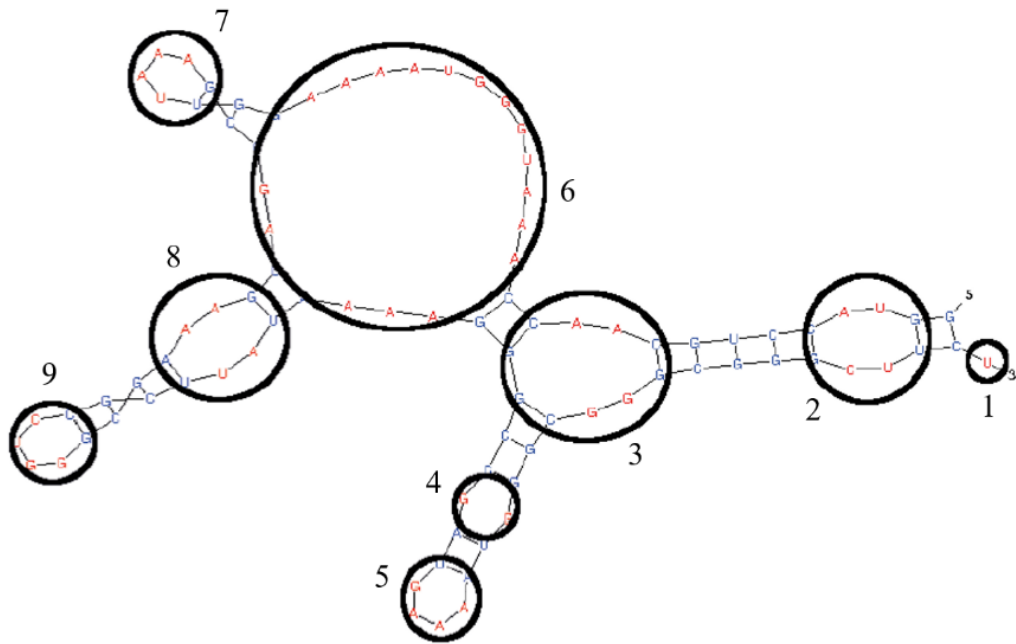


Рис. 3.2. Фрагмент цепи РНК объясняющий модель. Некоторые мономеры в соответствии с нашей моделью заключены в кружки. Для этой конфигурации ненулевыми элементами симметричной матрицы \hat{B} являются

$$B_{1,2} = B_{2,1} = B_{2,3} = B_{3,2} = B_{3,4} = B_{4,3} = B_{3,6} = B_{6,3} = B_{4,5} = B_{5,4} = B_{6,7} = B_{7,6} = B_{6,8} = B_{8,6} = B_{8,9} = B_{9,8} = 1$$

Взаимодействия между неспаренными областями (петлями) i и j определяются их размером и нуклеотидными последовательностями. Так как многие нуклеотиды вносят вклад во взаимодействие этих эффективных мономеров i и j , разумно рассматривать коэффициенты v_{ij} как статистически независимые гауссовские случайные переменные с распределением

$$p(v_{ij}) \sim \exp(-v_{ij}^2/(2\Lambda^2)), \quad (3.2)$$

где Λ является дисперсией распределения. Коллапс молекулы одноцепочечной РНК обусловлен в основном электростатическими

взаимодействиями и исследован экспериментально [111,112] и теоретически [103]. В отличие от [103], здесь мы явно не учитываем наличие противоионов. Их влияние присутствует неявно в сжимающем потенциале

$$\frac{V_{conf}(\{x_i\})}{T} = \frac{1}{2}v_0 \sum_{i<j} \delta(x_i - x_j) + \frac{1}{3!}g_0 \sum_{i<j<k} \delta(x_i - x_j)\delta(x_j - x_k), \quad (3.3)$$

которое обеспечивает существование глобулярного состояния молекулы РНК. Здесь величины $v_0 < 0$ и $g_0 > 0$ – соответственно второй и третий вириальные коэффициенты взаимодействий между спиральными участками.

Возможны два типа конформационных перестроек в РНК: перегруппировка вторичной структуры с характерным временным масштабом τ_2 и флуктуации третичной структуры с характерным временным масштабом τ_3 . Разница между масштабами времени τ_2 и τ_3 является существенной. Как показано в работе [112], коллапс РНК с длиной 400 нуклеотидов занимает около 3–4 мс, в то время как на два порядка короче РНК (21 нуклеотид) образует шпильку примерно за 10 мс [113]. Поэтому в дальнейших расчетах предполагается, что $\tau_3 \ll \tau_2$. Причиной этого является не только более высокая стабильность пар оснований по сравнению с третичными контактами. Важную роль также играют кинетические факторы. Образование пар оснований происходит путем разматывания одиночных цепей с дальнейшим образованием дуплекса, и процесс разматывания одиночных цепей кинетически затруднен.

3.2. Свободная энергия и ненулевое количество реплик

На временной шкале τ , такой, что $\tau_3 \ll \tau \ll \tau_2$, вторичная структура и пространственное расположение эффективных мономеров (непарных областей) не находятся в тепловом равновесии, и это неравновесное стационарное состояние можно описать в терминах эффективной статистической суммы [114]:

$$Z = \langle \langle (Z_{\hat{B}, \hat{v}})^n \rangle_{\hat{B}} \rangle_{\hat{v}}, \quad (3.4)$$

где $Z_{\hat{B}, \hat{v}}$ – статистическая сумма разветвленной молекулы с заданным характерным ветвлением \hat{B} и матрицей взаимодействия $\hat{v} = \|\|v_{ij}\|\|$. Символом $\langle \dots \rangle_{\hat{B}}$ обозначено среднее по всем возможным ветвящимся структурам, где каждый мономер принадлежит к одному из двух типов: а) концевой мономер, который связан только с одним другим мономером, и б) мономер на участке разветвления, который соединен точно с тремя другими. $\langle \dots \rangle_{\hat{v}}$ – среднее по межмономерным взаимодействиям, а $n = \frac{T}{T'}$, где T и T' являются эффективными температурами огрубленных пространственных и вторичных структур соответственно. Эффективная статистическая сумма (3.4) рассчитывается с использованием метода реплик, разработанного для систем с замороженным беспорядком (например в [114,115,91]). В нашем случае предел $n \rightarrow 0$ соответствует замороженному беспорядку, а $n = 1$ описывает полностью расплавленный беспорядок, и $0 < n < 1$ для частично расплавленного беспорядка.

Следуя методу, описанному в [106], мы можем написать статистическую сумму (3.4) в виде

$$Z = \int D\rho e^{-\beta F\{\rho\}}, \quad (3.5)$$

где $\beta = 1/T$, $\rho(X) = \sum_{i=1}^N \prod_{a=1}^n \delta(x_i^a - x^a)$, $X = (x^1, \dots, x^n)$.

$$F\{\rho\} = E\{\rho\} - TS\{\rho\}$$

является n -репличной свободной энергией, где

$$E\{\rho\} = \sum_a V_{conf}(c_a) - \frac{\beta \Lambda^2}{4} \sum_{a \neq b} \int dx dy q_{ab}^2(3, x, y),$$

$$S\{\rho\} \cong -\frac{l^4}{2} \int dX \frac{(\Delta \rho(X))^2}{\rho(X)}. \quad (3.6)$$

Здесь величина $c_a(x) = \int dX \rho(3, X) \delta(x^a - x)$ является однорепличная плотность мономеров, $q_{ab}(x, y) = \int dX \rho(3, X) \delta(x^a - x) \delta(x^b - y)$ – двурепличный параметр перекрывания, а Δ – оператор Лапласа в $(n \times d)$ -мерном пространстве. Далее во всех уравнениях принимаем, что $k_B = 1$.

Энергетический член в уравнении (3.6) получается усреднением n -той степени статистической суммы по переменным v_{ij} , а энтропийный член включает усреднение по всем возможным вариантам ветвления, соответствующим корневому дереву с координационным числом, равным трем. Для функции $\rho(X)$ используется следующая форма:

$$\rho(X) = \rho_0 \exp \left[-\frac{1}{2} \sum_{a,b} k_{ab} x^a x^b \right]. \quad (3.7)$$

Сжимающий потенциал записывается как вириальное расширение $V_{conf}(c) = NT \left(\frac{1}{2} v_0 c + \frac{1}{3!} g_0 c^2 \right)$. В компактно упакованной цепочке плотность мономеров ведет себя как $c_0 \approx l^{-d}$, сжимающий потенциал $V_{conf}(c) \approx NT \frac{v_0}{4l^d}$.

3.3. Параметризация матрицы Паризи для $n \neq 0$

Для частично расплавленной системы $0 < n = \frac{T}{T'} < 1$. В этом случае мы параметризуем недиагональные элементы матрицы \hat{k} с помощью функций $k(u)$, где $u \in [n, 1]$, а диагональные элементы как $k_{aa} = \bar{k}$ [116]. Обратная матрица \hat{m} параметризуется с помощью $m(u)$, $u \in [n, 1]$ и $m_{aa} = \tilde{m}$, где [116,117]

$$\tilde{m} = \frac{1}{\bar{k} - \langle k \rangle - [k](n)} \left[1 - \frac{1}{n} \frac{[k](n)}{\bar{k} - \langle k \rangle} - \int_n^1 \frac{du}{u^2} \frac{[k](u) - [k](n)}{\bar{k} - \langle k \rangle - [k](u)} \right], \quad (3.8)$$

$$m(u) = - \frac{1}{\bar{k} - \langle k \rangle - [k](n)} \left[\frac{1}{n} \frac{[k](n)}{\bar{k} - \langle k \rangle} + \frac{1}{u} \frac{[k](u) - [k](n)}{\bar{k} - \langle k \rangle - [k](u)} + \int_n^u \frac{dv}{v^2} \frac{[k](v) - [k](n)}{\bar{k} - \langle k \rangle - [k](v)} \right], \quad (3.9)$$

где $[k](u) = uk(u) - \int_n^u dvk(v)$.

Используя пробная функция (3.7), мы можем записать выражение для энтропии (3.6) в виде

$$S\{\rho\} \cong \frac{l^d dNn}{2} \left[\int_n^1 duk(u)^2 - \bar{k}^2 \right]. \quad (3.10)$$

Двурепличный параметр перекрывания $q_{ab}(x_1, x_2)$ в уравнений (3.6) описывается следующим уравнением:

$$q_{ab}(x_1, x_2) = \frac{N}{(2\pi)^d} (\det \hat{M})^{-d/2} \exp\left(-\frac{1}{2} \sum_{\alpha, \beta=1}^2 M_{\alpha\beta}^{-1} x_\alpha x_\beta\right), \quad (3.11)$$

где

$$\hat{M} = \begin{pmatrix} k_{aa}^{-1} & k_{ab}^{-1} \\ k_{ba}^{-1} & k_{bb}^{-1} \end{pmatrix}. \quad (3.12)$$

Подставляя (3.11) в уравнение (3.6), получим энергию в непрерывном пределе ($0 < n < 1$):

$$E\{\rho\} = \frac{\beta\Lambda^4}{4} nN^2 \pi^d (2\pi)^{-2d} \int_n^1 du (\phi - m(u)^2)^{-d/2}, \quad (3.13)$$

где величина

$$\phi = \tilde{m}^2 \cong l^2 N^{4/d} \quad (3.14)$$

является константой. Следовательно, функционал свободной энергии $F\{\rho\}$ будет иметь следующий вид:

$$F\{\rho\} = nV_{conf}(c_0) + \frac{\beta\Lambda^4}{4} nN^2 \pi^d (2\pi)^{-2d} \int_n^1 du (\tilde{m}^2 - m(u)^2)^{-d/2} - T \frac{l^4 dNn}{2} \left[\langle k \rangle^2 + \int_n^1 \frac{du}{u^2} [k](u)^2 - \tilde{k}^2 \right]. \quad (3.15)$$

Учитывая тождества

$$\begin{aligned} \langle k \rangle &= \int_n^1 \frac{du}{u^2} [k](u), \\ \int_n^1 duk(u)^2 &= \langle k \rangle^2 + \int_n^1 \frac{du}{u^2} [k](u)^2 \end{aligned} \quad (3.16)$$

и уравнения (3.8), (3.9) функционал свободной энергии $F\{\rho\}$ может быть прямо выражена через \tilde{k} и $[k](u)$. Вариация свободной энергии $F\{\rho\}$ по \tilde{k} и $[k](u)$ в качестве независимых «переменных» дает два решения, удовлетворяющих условиям $[k](x) = const$ и

$$[\tilde{k} - \langle k \rangle - [k](3.x)]^4 = \frac{\alpha}{3} (\phi - m(x)^2)^{\frac{d}{2}-2} [\phi + (d+1)m(x)^2], \quad (3.17)$$

где $\alpha = \frac{\beta\Lambda^2}{4} \frac{N\pi^d}{2(2\pi)^{2d} T l^4}$. Легко найти, что в пределе длинной цепочки $N \gg 1$ решением уравнения (3.17) являются $m(x) = const$ и

$$m(x)^2 \cong \phi - \phi \left(\frac{\chi\epsilon}{x} \right)^{1/\gamma}, \quad (3.18)$$

где $\frac{1}{\epsilon} = 2(\alpha/3)^{1/4} \phi^{\frac{2-d}{8}}$, $\chi = \frac{3\left(\frac{d}{2}+1\right)+(d+1)\left(\frac{d}{2}-1\right)}{(d+2)^{5/4}}$ и $\gamma = \frac{1}{8}(4-d)$. Точка u_0 , в которой совпадают решения $m(x) = const$ и уравнения (3.18), оценивается как

$$u_0 = \frac{3\chi}{2(d+2)^{3/4}}. \quad (3.19)$$

С помощью уравнений (3.18), (3.19) получен следующий анзац для функции $[k](u)$:

$$[k](u) \cong \begin{cases} -z_0 u^\delta, & \text{если } n \leq u \leq u_0 \\ -z_0 u_0^\delta, & \text{если } u_0 \leq u \leq 1 \end{cases}, \quad (3.20)$$

где $z_0^2(d, \beta, \Lambda, l) \propto (\beta \Lambda l^{-d})^{1/\gamma(3,d)}$, $\gamma(d) = \frac{(4-d)}{8}$, $\delta(d) = \frac{d+4}{4-d}$ и $u_0(d) = \frac{3}{2} \frac{3\left(\frac{d}{2}+1\right)+(d+1)\left(\frac{d}{2}-1\right)}{(d+2)^2} \approx 0.6$ если $d = 3$.

3.4. Фазовый переход и удельная теплоемкость

Подставляя выражение (3.20) в функционал свободной энергии (3.15) видно, что свободная энергия приходящаяся на один мономер имеет две ветви [109]:

$$f = \frac{F\{\rho\}}{Nn} = T \frac{v_0}{4l^d} + \begin{cases} f_{<}, & \text{если } n \leq u_0 \\ f_{>}, & \text{если } n > u_0 \end{cases}. \quad (3.21)$$

Это указывает на существование фазового перехода при температуре

$$T_c' = \frac{T}{u_0}. \quad (3.22)$$

Неупорядоченная свободная энергия

$$f_{<} = f_{<}^0 - \frac{n^\zeta}{\zeta} (a_1 - Tl^4 dz_0^2), \quad (3.23)$$

где $n \leq u_0$, $\zeta = \frac{3d+4}{4-d}$, $a_1(d, \beta, \Lambda, l) \propto \beta^{-1}(\beta \Lambda l^{-d})^{1/\gamma(d)}$. В то же время при $n \geq u_0$ неупорядоченная свободная энергия может быть записана в виде

$$f_{>} = f_{>}^0 - n \left(a_1 u_0^{\zeta-1} + Tl^4 dz_0^2 u_0^{\zeta+1} \frac{1}{n^2} \right), \quad (3.24)$$

где $f_{>}^0 = f_{>}^0(T, d, \Lambda, l)$. Полная энтропия двухтемпературной системы [40] определяется как

$$S_{tot} = S_{II} + S, \quad (3.25)$$

где $S_{II} = -\frac{\partial}{\partial T'} F \Big|_{T=const}$ и $S = -\frac{\partial}{\partial T} F \Big|_{T'=const}$. Суммарная удельная теплоемкость, которая недавно была изучена в работе [109] имеет следующий вид:

$$C_{tot} = C_{II} + C, \quad (3.26)$$

где величина $C_{II} = T' \frac{\partial S_{II}}{\partial T'}$ является удельной теплоемкостью вторичной структуры и $C = T \frac{\partial S}{\partial T}$.

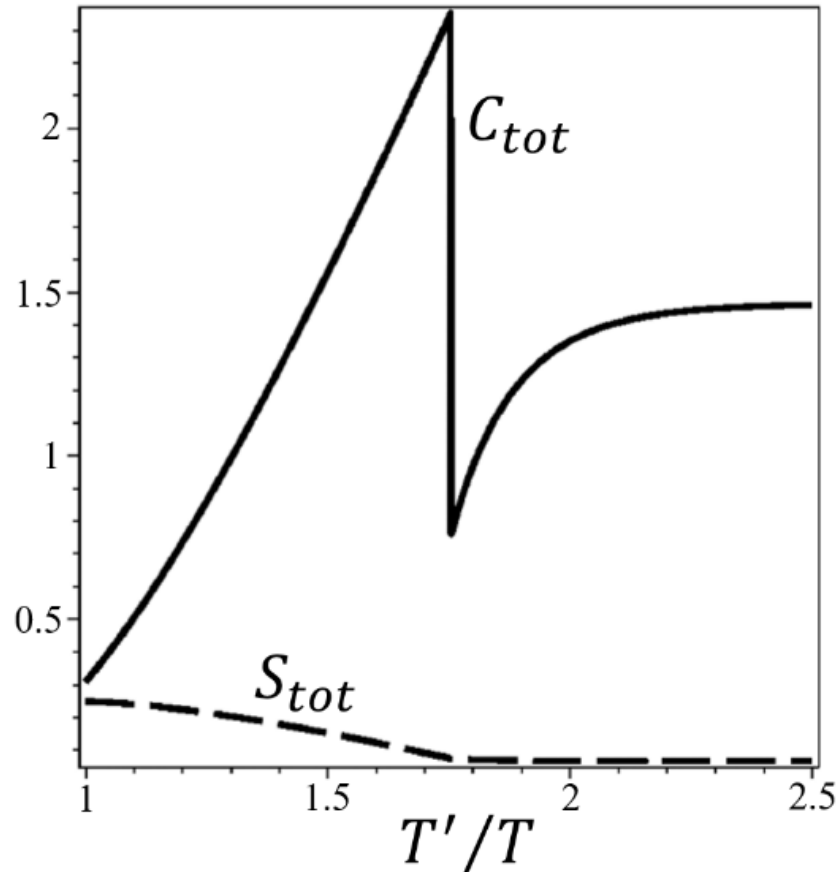


Рис. 3.3. Температурные поведения суммарной удельной теплоемкости (сплошная линия) и суммарной энтропии (3.пунктирная линия) для $d = 3$ и $\frac{\beta\Lambda}{l^d} = 140$.

Как показано на Рис. 3.3, график суммарной удельной теплоемкости c_{tot} имеет резкий пик в окрестностях критической точки T'_c , что указывает на наличие фазового перехода. Наблюдаемый разрыв свидетельствует о существовании неравновесного фазового перехода второго порядка. При уменьшении температуры T' суммарная удельная теплоемкость испытывает скачок $\Delta c_{tot}|_{T'=T'_c} > 0$ в критической точке $T' = T'_c$, но не показывает расходимости, подобной

$$c_{tot} \sim \left(\frac{|T' - T'_c|}{T'_c} \right)^{-\alpha},$$

которая типична для фазового перехода второго порядка. Возможной причиной может быть среднеполевой характер теории, основанной на репликах. В то же время полная энтропия S_{tot} сама по себе непрерывна в окрестности критической точки (Рис. 3.3), а скрытая теплота перехода равна нулю.

Результаты и обсуждение

Основной характеристикой стеклообразной фазы является существование большого числа минимумов свободной энергии, разделенных барьерами, которые становятся бесконечно большими в термодинамическом пределе. Две фазы, разделенные критической точкой (3.22), характеризуются параметром порядка, описывающим структурное перекрытие между минимумами a и b [106,115]:

$$Q_{ab} = \frac{1}{N} \sum_j \delta(x_j^a - x_j^b) = \frac{1}{N} \int dx q_{ab}(3, x, x). \quad (3.27)$$

Принимая во внимание уравнение (3.11), в непрерывном пределе ($0 < n < 1$) параметр перекрытия $q(u; x_1; x_2)$ можно записать в виде

$$q(u; x_1; x_2) = N \frac{\exp[-(x_1^2 - x_1 r)(\sqrt{\phi} - m(u))/(\phi - m(u)^2)] \exp[-r^2 \sqrt{\phi}/2(\phi - m(u)^2)]}{(2\pi)^{2d(1-n/4)} (\phi - m(u)^2)^{d/2}}, \quad (3.28)$$

где $r = x_2 - x_1$ и $r = |r|$. С помощью последнего уравнения можно оценить длину корреляции между репликами:

$$R(u)^2 \cong \frac{\phi - m(u)^2}{\sqrt{\phi}}. \quad (3.29)$$

Используя уравнения (3.18) и (3.19), из уравнения (3.29) можно получить, что

$$R(u) \cong l \begin{cases} u^{-\frac{4}{4-d}} & \text{если } n \leq u \leq u_0 \\ u_0^{-\frac{4}{4-d}} & \text{если } u_0 \leq u \leq 1 \end{cases}. \quad (3.30)$$

Таким образом, в частично расплавленном случае масштаб перекрытия структур соответствующих разным минимумам является таким же, как и в случае молекулы РНК с полностью замороженными ветвлениями [106]. В непрерывном пределе ($0 < n < 1$) структурный параметр перекрытия (3.27) масштабируется как

$$Q(u) \sim R(u)^{-d}. \quad (3.31)$$

Если функция $Q(u)$ постоянна, то мы имеем только чистые состояния, макроскопически не отличающиеся друг от друга. Если функция $Q(u)$ не является константой, то должны существовать макроскопически различные чистые состояния: $Q(1)$ идентифицируется с параметром Эдвардса-Андерсона [119]. Поведение функции $Q(u)$ показано на Рис. 3.4. Как обычно, функция $Q(u)$ имеет плато при $Q = Q(1)$.

Для описания ландшафта свободной энергии нам нужно оценить больцмановские веса минимумов свободной энергии $\{p_a\}$. Согласно [42], эти веса должны удовлетворять условию

$$\sum_a p_a^2 = \text{длине плато}. \quad (3.32)$$

Таким образом $\sum_a p_a^2 = 1 - u_0(d) \approx 0.4$ если $d = 3$. Так как $p_a < 1$, то в этой сумме должны доминировать несколько состояний, что приводит к стекловидному ландшафту свободной энергии, показанному на Рис. 3.6а.

Для описания статистических свойств перекрывания двух минимумов свободной энергии можно ввести следующую функцию распределения вероятностей [115–117]:

$$P(Q) = \sum_{a,b} p_a p_b \delta(Q_{ab} - Q), \quad (3.33)$$

где величина p_a является больцмановским весом a -ого минимума свободной энергии. С учетом соотношения $P(Q) = du/dQ$, получаем

$$P(Q) \sim Q^{\frac{4-5d}{4d}} \quad (3.34)$$

для $Q < Q(1)$. Из профиля функции $P(Q)$, показанной на Рис. 3.5 видно, что возможно (и вероятность этого достаточно велика) найти низкоэнергетические минимумы с некоторым структурным подобием (величина $P(Q)$ отлична от нуля до $\frac{Q}{Q(1)} \sim 1$). Кроме того, непрерывность функции $Q(u)$ указывает на то, что конфигурационное пространство системы является ультраметрическим.

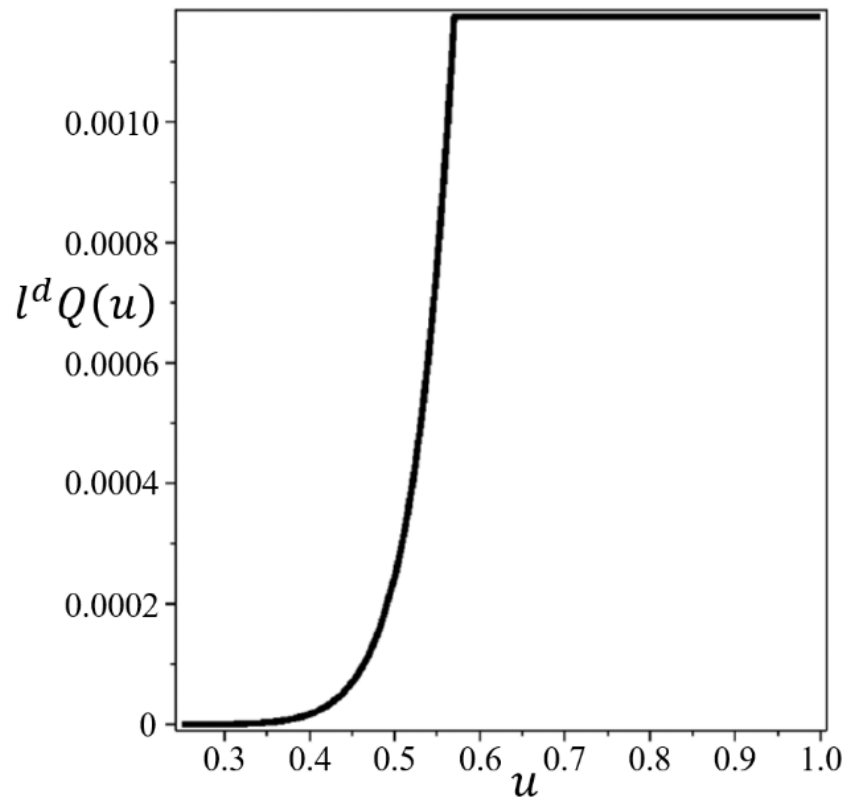


Рис. 3.4. Поведение функционала $l^d Q(u)$ для $d = 3$ и $n = 0.25$.

Таким образом, при температурах $T' > T'_c$, что соответствует состоянию, достаточно далекому от равновесия, ландшафт свободной энергии имеет ту же структуру, что и в случае расплавленной структуры ветвлений. Однако для температур $T' < T'_c$ у нас есть реплично симметричная фаза с отличным от нуля спин-стекольным параметром порядка (так называемый репличный магнит) [121].

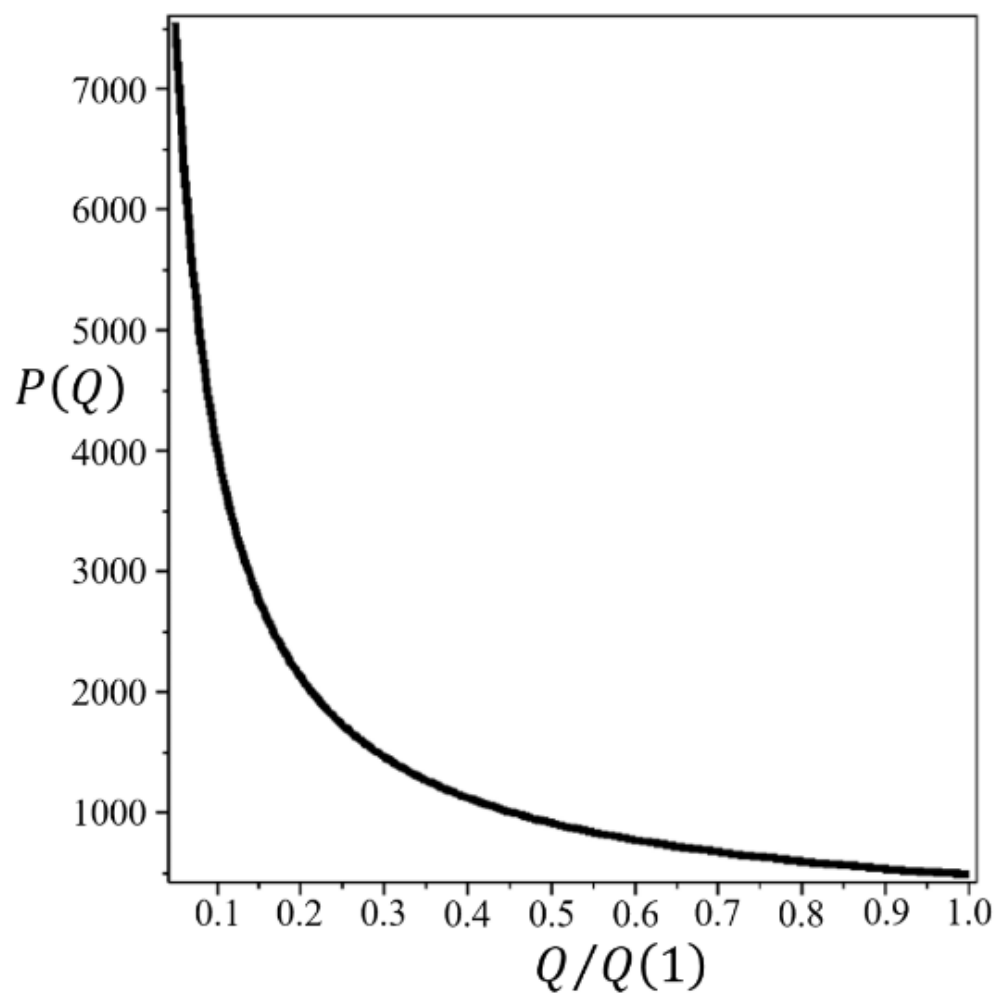


Рис. 3.5. Поведение функции $P(Q)$ для $d = 3$ и $n = 0.25$.

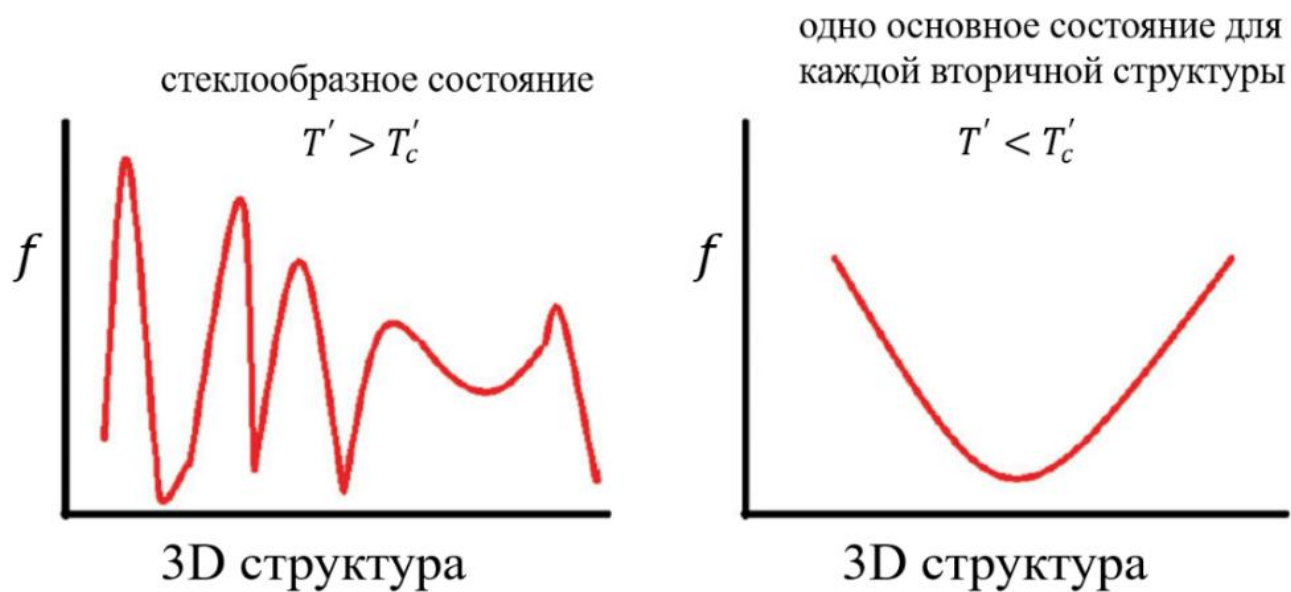


Рис. 3.6. Эскиз ландшафта свободной энергии.

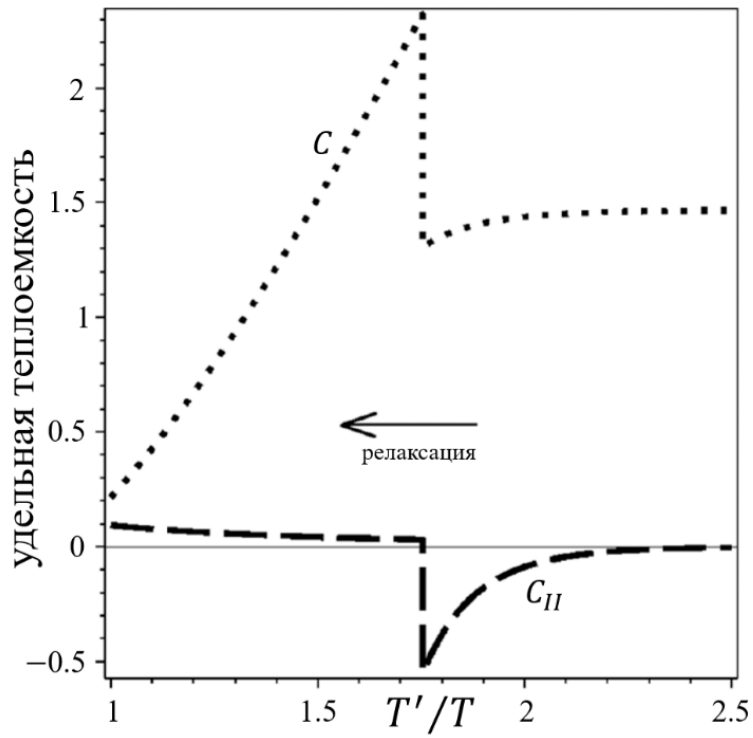


Рис. 3.7. Температурные поведения удельной теплоемкости вторичной структуры C_{II} и условной удельной теплоемкости C для $d = 3$ и $\frac{\beta\Lambda}{l^d} = 140$.

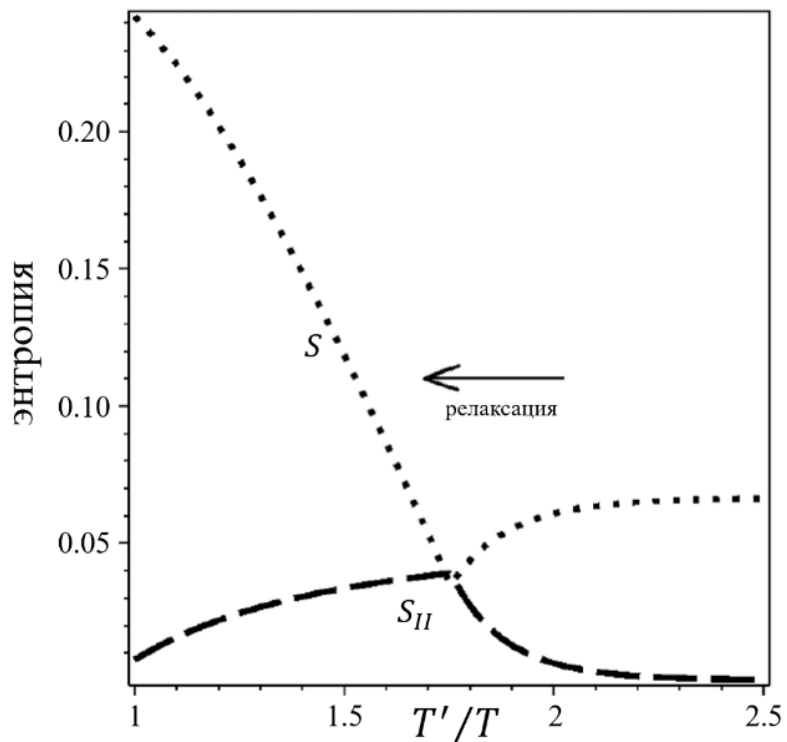


Рис. 3.8. Температурные поведения энтропии вторичной структуры S_{II} и условной энтропии S для $d = 3$ и $\frac{\beta\Delta}{l^d} = 140$.

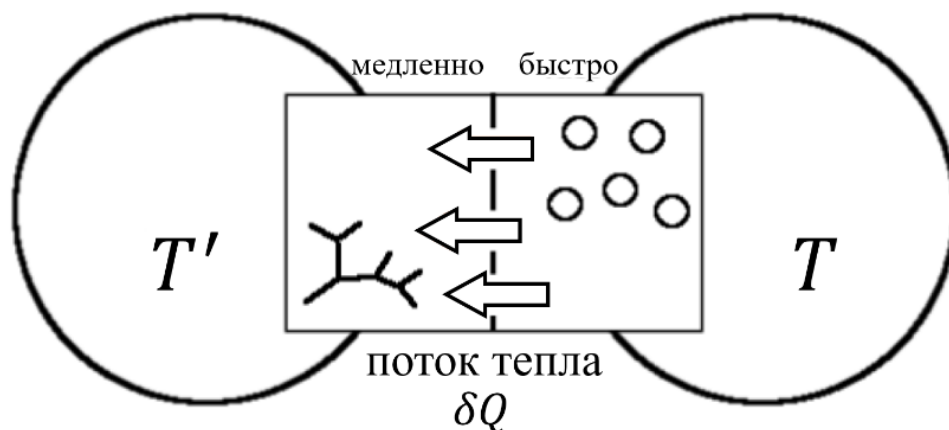


Рис. 3.9. Вторичная структура холодильника.

Понижение температуры T' в сторону температуры T соответствует приближению одноцепочечной РНК к равновесию в процессе термодинамической релаксации и обеспечивает условия для гибридизации РНК и перестройки вторичной структуры.

Таким образом, в процессе релаксации наблюдаются два режима, характеризующиеся различными ландшафтами свободной энергии, (Рис. 3.6). Если эффективная температура РНК T' удовлетворяет условию $T' > T'_c$, то получается стеклообразное состояние (Рис. 3.6а), где у нас есть несколько третичных структур, соответствующих минимумам свободной энергии для каждой вторичной структуры. В то же время, если эффективная температура T' становится меньше критической температуры T'_c , для каждой вторичной структуры имеется только одно основное состояние (Рис. 3.6б). Таким образом, усиление гибридизации в процессе термодинамической релаксации сопровождается настройкой третичной структуры в соответствии с установленной вторичной структурой.

В процессе релаксации удельная теплоемкость вторичной структуры становится отрицательной ($C_{II} < 0$) при $T' > T'_c$, в то время как общая удельная теплоемкость C_{tot} остается положительной (Рис. 3.3 и Рис. 3.7). Отрицательная теплоемкость C_{II} получается в неравновесном стационарном состоянии на временных масштабах τ , так чтобы $\tau_3 \ll \tau \ll \tau_2$. Отрицательная теплоемкость для неравновесных систем обсуждалась в [122–124].

В стеклообразной фазе ($T' > T'_c$) охлаждение вторичной структуры сопровождается уменьшением условной энтропии S , а энтропия вторичной структуры S_{II} увеличивается (Рис. 3.8). Фактически, в стеклообразной фазе вторичная структура работает подобно холодильнику, который отбирает тепло $\delta Q \leq TdS < 0$ от пространственных степеней свободы, используя сценарий «получаешь больше при меньших усилиях», описанный в [122] (Рис. 3.9). Взяв тепло из быстрых (пространственных) степеней свободы, мы, вероятно, уменьшаем число минимумов в ландшафте свободной энергии (Рис. 3.6а), вызывая переход к ландшафту, показанному на Рис. 3.6б.

Гл.4. Влияние интеркалирующих лигандов на изотерму гибридизации ДНК - чипа

ДНК-чипы являются перспективными инструментами с большим разнообразием областей применения, таких как медицинская диагностика, мониторинг окружающей среды, защиты от биологического оружия и т. д. [125, 126]. Важным направлением разработки ДНК-чипов является повышение селективности и чувствительности за счет усиления электрического сигнала и стабильности целевого зонда гибридизации. Эффективность таких устройств, как ДНК-сенсоры и ДНК-чипы зависит от точности определения экспериментальных параметров, отвечающих за термостабильность дуплексов нуклеиновых кислот и время формирования дуплексов ДНК [127]. Некоторые из факторов влияют на термодинамику гибридизации, в частности: поверхностная плотность одноцепочечных ДНК (длина 25–49 нуклеотидов), иммобилизованных на поверхности, и наличие конкурирующей гибридизации. Стабильность дуплексов ДНК–ДНК и ДНК–РНК определяется двумя ключевыми факторами: последовательностью и внешними факторами, такими, как pH , ионная сила, концентрации низкомолекулярных соединений (лигандов), наличие межфазных границ, геометрических ограничений и т.д.). Лучшее понимание физико-химических процессов, лежащих в основе гибридизации ДНК и РНК на поверхности электрического преобразователя имеет важное значение для улучшения эффективности ДНК-чипов и их изготовления [128].

Требованиями, предъявляемыми к ДНК-сенсорам является высокая чувствительность и селективность, которые, в свою очередь, требуют максимальной эффективности гибридизации и минимальной неспецифической адсорбции на границе раздела твердой и жидкой фаз. Гибридизация нуклеиновых кислот в значительной степени зависит от

температуры, концентрации солей, вязкости, GC-состава и других физико-химических характеристик.

Повышение селективности и чувствительности ДНК-сенсоров может быть достигнуто с помощью электрохимически активных соединений с более высоким сродством к двухцепочечной ДНК, чем к одноцепочечной ДНК. Этот вид соединений может существенно повысить стабильность двухцепочечных участков и в то же время, амплитуду генерируемого сигнала, что, в свою очередь, повысит чувствительность ДНК-сенсора. Такими лигандами являются, например, интеркаляторы, молекулы с плоской гетероциклической структурой, которые помещаются между азотистыми основаниями и меняют локальную структуру двухцепочечной ДНК [129–131].

Термодинамика и кинетика гибридизации как в объеме [132,133], так и на поверхности [128,134–140] была тщательно изучена в последние годы. Спектр рассматриваемых проблем включает в себя, например, кинетику гибридизации на поверхности [136, 138], влияние солей на гибридизацию ДНК в объеме [133], изотермы гибридизации на поверхности [128] и т. д. В то же время взаимодействия ДНК-лиганд также были рассмотрены в огромном количестве работ, посвященных интеркаляции [129–131] и связыванию лигандов в малой бороздке [141,142], их кросс-докингу [143] и т. д. Однако влияние взаимодействия ДНК-лиганд на термодинамику и кинетику гибридизации никогда не было рассмотрено ранее. В контексте развития ДНК-сенсоров, теоретический анализ влияния интеркаляции лигандов на гибридизацию ДНК на поверхности становится необходимым.

Настоящая работа посвящена изучению изотермы гибридизации ДНК на поверхности в присутствии лигандов, которые связываются с двухцепочечными участками ДНК. На практике ДНК-чипы погружаются в целевой раствор на относительно короткий промежуток времени и кинетика гибридизации играет решающую роль. Однако понимание равновесных

свойств также необходимо для сравнительной оценки важности кинетических и термодинамических факторов для производительности ДНК-чипов.

4.1. Неконкурентная гибридизация

Свободная энергия

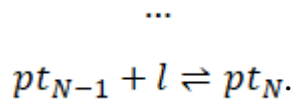
Рассмотрим равновесные изотермы гибридизации для двух идеализированных, но экспериментально доступных ситуаций, когда ДНК-чип погружаем в раствор, содержащий интеркалированные лиганды. В результате имеем только один тип одноцепочечной мишени (Рис. 4.1a) или раствор содержит мишени двух разных типов, которые не гибридизуются в объеме, но оба способны гибридизоваться с тем же зондом на поверхности (Рис. 4.1b).

Рассмотрим совокупность N_0 однонитевых пробных p молекул ДНК, где N_0 из них гибридизованы с мишенью t . Гибридизация p и t создает на поверхности двухцепочечный олигонуклеотид pt . В простейшем случае для одного вида мишени, состоящей из одноцепочечной ДНК, поверхность будет покрыта только свободными зондами p и гибридизированными олигонуклеотидами pt . В этом случае мы имеем одну реакцию



В этом случае реакции конкурентной гибридизации отсутствуют (Рис. 4.1a). Зависимость степени гибридизации $x = N_{pt}/N_0$ от концентрации мишеней c_t описывается с помощью изотермы гибридизации. Для интеркалирующих лигандов l реакции связывания будут иметь следующий вид:





где pt является свободным дуплексом, а pt_j является дуплексом мишень–зонд связанным с лигандом l .

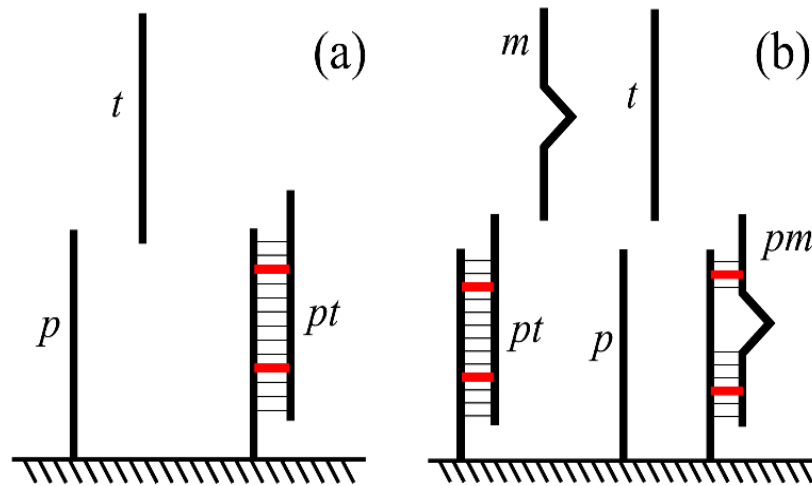


Рис. 4.1. Схемы (a) неконкурентной и (b) конкурентной гибридизации на поверхности в присутствие лигандов.

В отсутствие лигандов свободная энергия слоя с зондами будет иметь следующий вид [135]:

$$G = G_0 + N_{pt}\mu_{pt}^0 + (N_0 - N_{pt})\mu_p^0 + N_0 \sum \gamma_{el} + k_B T \left[N_{pt} \ln \left(\frac{N_{pt}}{N_0} \right) + (N_0 - N_{pt}) \ln \left(\frac{N_0 - N_{pt}}{N_0} \right) \right] \quad (4.3)$$

где Σ – площадь, приходящаяся на один зонд, G_0 – плотность свободной энергии голой поверхности, μ_{pt}^0 и μ_p^0 – химические потенциалы зондов pt и p в исходном состоянии и γ_{el} – электростатическая плотность свободной энергии зондирующего слоя.

Если интеркаляция является единственным механизмом связывания лигандов, то образование комплекса ДНК–лиганд будет ограничено только двуцепочечными областями и свободная энергия слоя с зондами составит:

$$G_L = G + N_{pt} \left\{ m\mu_b^0 + k_B T \left[m \ln \left(\frac{m}{N} \right) + (N - m) \ln \left(\frac{N - m}{N} \right) \right] \right\}, \quad (4.4)$$

где m это число связанных лигандов из расчета на одну гибридизованную пробу pt , а μ_b^0 – химический потенциал связанного лиганда в исходном состоянии. Предполагается, что доступное количество мест связывания на дуплексе pt совпадает с длиной N . Таким образом, свободная энергия слоя с зондами запишется в виде функции от независимых величин: числа гибридизованных зондов N_{pt} и числа связанных лигандов $N_b = mN_{pt}$. Свободная энергия будет

$$G_L(N_{pt}, N_b) = G(N_{pt}) + N_b \mu_b^0 + k_B T \left[N_b \ln \left(\frac{N_b}{NN_{pt}} \right) + (NN_{pt} - N_b) \ln \left(\frac{NN_{pt} - N_b}{NN_{pt}} \right) \right]. \quad (4.5)$$

4.2. Изотермы адсорбции и гибридизации

Состояние равновесия для реакций (4.1) и (4.2) будет определяться с помощью условий:

$$\mu_{pt} = \mu_p + \mu_t \quad (4.6)$$

и

$$\mu_b = \mu_l, \quad (4.7)$$

где μ_{pt} – химический потенциал гибридизованного зонда pt , μ_t – химический потенциал мишени, μ_p – химический потенциал зонда, а величины μ_b и μ_l – химические потенциалы связанных и несвязанных лигандов, соответственно [144].

Обменный химический потенциал гибридизированного зонда ($\Delta\mu_{pt} = \mu_p - \mu_t$) напишется в виде

$$\Delta\mu_{pt} = \frac{\partial G_L}{\partial N_{pt}} = \Delta\mu_{pt}^0 + N \frac{\partial \gamma_{el}}{\partial \sigma} + k_B T \ln \frac{x}{1-x} + k_B T \ln(1 - r), \quad (4.8)$$

где $r = \frac{N_b}{NN_{pt}}$ описывает степень адсорбции лигандов l в двухцепочечной ДНК. Плотность электростатической свободной энергии γ_{el} рассматривается как функция от плотности зарядов на поверхности σ . В то же время, химический потенциал связанных лигандов [144] составляет

$$\mu_b = \frac{\partial G_L}{\partial N_b} = \mu_b^0 + k_B T \ln \frac{r}{1-r}. \quad (4.9)$$

В приближении слабого раствора химический потенциал мишени будет иметь следующий вид

$$\mu_t = \mu_t^0 + k_B T \ln c_t, \quad (4.10)$$

а химический потенциал свободных лигандов в растворе

$$\mu_l = \mu_l^0 + k_B T \ln c_l, \quad (4.11)$$

где величины c_t и c_l являются объемными концентрациями мишеней и лигандов соответственно. Учитывая уравнения (4.6) – (4.11), получим изотерму гибридизации

$$\frac{x(1-r)^N}{c_t(1-x)} = K_t \exp\left(-\frac{N}{k_B T} \frac{\partial \gamma_{el}}{\partial \sigma}\right), \quad (4.12)$$

где $K_t = \exp\left(-\frac{\Delta G^0}{k_B T}\right)$ и $\Delta G^0 = \mu_{pt}^0 - \mu_p^0 - \mu_t^0$. Равновесное распределение l между связанными и свободными состояниями будет описываться изотермой адсорбции

$$\frac{r}{c_l(1-r)} = K_l, \quad (4.13)$$

где $K_l = \exp\left(-\frac{\Delta g^0}{k_B T}\right)$ и $\Delta g^0 = \mu_b^0 - \mu_l^0$.

4.3. Конкурентная гибридизация на поверхности

Рассмотрим второй сценарий, в котором растворитель содержит мишени двух различных типов t и m , которые не гибридизуются в объеме, но оба способны к гибридизации с одним и тем же зондом p на поверхности (Рис. 4.1b). Здесь t является последовательностью, полностью

комплементарной к зонду p , а t является несогласованной последовательностью только частично комплементарной к зонду p . Предполагается, что доступное количество мест связывания для интеркалированных лигандов на дуплексе pt равно M , где $M < N$.

Согласно подходу разработанному в [135], свободную энергию слоя зондов, свободных от лигандов, можно записать в следующем виде:

$$G_C = G_0 + N_{pt}\mu_{pt}^0 + N_{pm}\mu_{pm}^0 + (N_0 - N_{pt} - N_{pm})\mu_p^0 + N_0 \sum \gamma_{el} + k_B T \left[N_{pt} \ln \left(\frac{N_{pt}}{N_0} \right) + N_{pm} \ln \left(\frac{N_{pm}}{N_0} \right) + (N_0 - N_{pt} - N_{pm}) \ln \left(\frac{N_0 - N_{pt} - N_{pm}}{N_0} \right) \right], \quad (4.14)$$

где величина N_{pm} – количество дуплексов pm и μ_{pm}^0 – химический потенциал этих дуплексов в исходном состоянии.

Если интеркаляция является единственным механизмом связывания, то образование комплексов ДНК–лиганд будет ограничено только двухцепочечными участками, и тогда свободная энергия слоя с зондами будет:

$$G_{CL} = G_C + N_{pt} \left\{ n\mu_b^0 + k_B T \left[n \ln \left(\frac{n}{N} \right) + (N - n) \ln \left(\frac{N - n}{N} \right) \right] \right\} + N_{pm} \left\{ m\mu_b^0 + k_B T \left[m \ln \left(\frac{m}{M} \right) + (M - m) \ln \left(\frac{M - m}{M} \right) \right] \right\}, \quad (4.15)$$

где n – число лигандов, связанных с зондом pt и m – число лигандов, связанных с зондом pm . Таким образом, свободная энергия слоя с зондами зависит от числа полностью комплементарно гибридизированных зондов N_{pt} , числа не полностью комплементарно гибридизированных зондов N_{pm} и количеств лигандов, связанных с дуплексами pt и pm (для N_1 и N_2 соответственно). Здесь $N_1 = nN_{pt}$ и $N_2 = mN_{pm}$.

$$\begin{aligned}
G_{CL}(N_{pt}, N_{pm}, N_1, N_2) &= G_C(N_{pt}, N_{pm}) + (N_1 + N_2)\mu_b^0 + \\
&k_B T \left[N_1 \ln \left(\frac{N_1}{NN_{pt}} \right) + (NN_{pt} - N_1) \ln \left(\frac{NN_{pt} - N_1}{NN_{pt}} \right) \right] + \\
&k_B T \left[N_2 \ln \left(\frac{N_2}{MN_{pm}} \right) + (MN_{pm} - N_2) \ln \left(\frac{MN_{pm} - N_2}{MN_{pm}} \right) \right]. \quad (4.16)
\end{aligned}$$

Химический потенциал лигандов, связанных с комплементарным дуплексом pt , будет иметь следующий вид:

$$\mu_b^1 = \frac{\partial G_{CL}}{\partial N_1} = \mu_b^0 + k_B T \ln \frac{r_1}{1-r_1}, \quad (4.17)$$

где величина $r_1 = \frac{N_1}{NN_{pt}}$ характеризует степень адсорбции лигандов на дуплексе pt . Химический потенциал лигандов, связанных с частично комплементарным дуплексом pm , будет:

$$\mu_b^2 = \frac{\partial G_{CL}}{\partial N_2} = \mu_b^0 + k_B T \ln \frac{r_2}{1-r_2}, \quad (4.18)$$

где $r_2 = \frac{N_2}{MN_{pm}}$.

В то же время химические потенциалы обмена гибридизованных зондов pt и pm будут иметь следующий вид:

$$\begin{aligned}
\Delta\mu_{pt} &= \Delta\mu_{pt}^0 + N \frac{\partial \gamma_{el}}{\partial \sigma} + k_B T \ln \frac{x}{1-x-y} + N k_B T \ln(1 - r_1), \\
\Delta\mu_{pm} &= \Delta\mu_{pm}^0 + N \frac{\partial \gamma_{el}}{\partial \sigma} + k_B T \ln \frac{y}{1-x-y} + N k_B T \ln(1 - r_2), \quad (4.19)
\end{aligned}$$

где $y = \frac{N_{pm}}{N_0}$.

Равновесное состояние между лигандами, мишенями, несовместимыми последовательностями в растворе и гибридизованными зондами будет описываться следующими условиями:

$$\begin{aligned}
\mu_b^1 &= \mu_l, \\
\mu_b^2 &= \mu_l, \\
\Delta\mu_{pt} &= \mu_t, \\
\Delta\mu_{pm} &= \mu_m, \quad (4.20)
\end{aligned}$$

где μ_m является потенциалом свободных несовместимых последовательностей m . В приближении слабого раствора величину μ_m можно оценить как $\mu_m = \mu_m^0 + k_B T \ln c_m$, где величина c_m является концентрацией последовательностей m .

Результаты и обсуждение

Изотерма гибридизации мишеней в отсутствие лигандов $x_0(c_t)$ [135] воспроизводится путем подстановки значения $r = 0$ в уравнение (4.12). Рассмотрим сдвиг изотермы гибридизации вызванное лигандами:

$$\delta x = x - x_0.$$

Равновесная степень адсорбции r^* составляет

$$r^* = \frac{c_l K_l}{1 - c_l K_l}. \quad (4.21)$$

Таким образом, эффект адсорбции интеркалирующих лигандов сводится к перенормировке константы связывания K_t :

$$\tilde{K}_t = K_t \exp(-N \ln(1 - r^*)). \quad (4.22)$$

Плотность электростатической свободной энергии слоя с зондами γ_{el} была оценена в работе [135] в приближении двухкомпонентного ящика [145–148]. В этом приближении ступенчатый профиль распределения мономеров позволяет рассматривать полиэлектролиты на поверхности как непрерывную область с равномерным распределением заряда. При высоком содержании солей экранирования в заряженном слое приводит к следующему выражению для плотности электростатической свободной энергии

$$\frac{\gamma_{el}}{k_B T} = 4\pi\sigma^2 l_B \frac{r_D^2}{H}, \quad (4.23)$$

где величина $l_B = \frac{e^2}{\epsilon k_B T}$ – длина Бьоррума, ϵ – диэлектрическая проницаемость, r_D – Дебаевская длина экранирования. Здесь H – это толщина слоя с зондами. При этом предполагается, что заряды распределяются в этом слое равномерно. Так как каждая цепь содержит заряд $-eN$, плотность заряда σ зависит от степени гибридизации x как

$$\sigma = \sigma_0(1 + x), \quad (4.24)$$

где $\sigma_0 = \frac{NN_0}{A}$ и величина A – площадь слоя с зондами.

Изотермы гибридизации, полученные с помощью уравнения (4.12), (4.21), (4.22) и (4.24) показаны на Рис. 4.2.

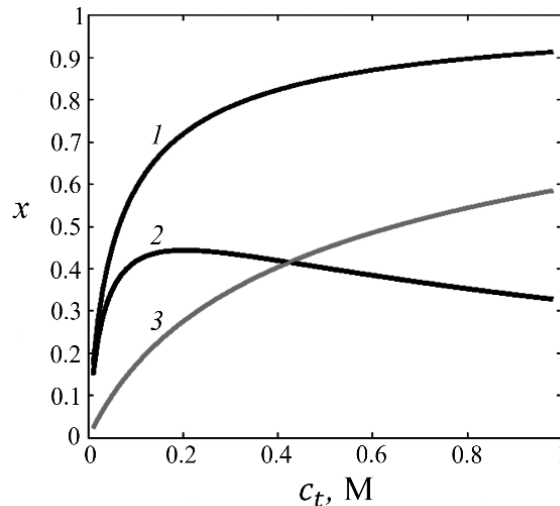


Рис. 4.2. Зависимость степени гибридизации от концентрации мишеней.

1 – в присутствие лигандов, 2 – сдвиг изотермы гибридизации и кривая 3 – без лигандов. Результаты получены для следующих значений параметров: $l_B \approx 7 \text{ \AA}$, $r_D \approx 3 \text{ \AA}$, $K_t = 10^9 \text{ M}^{-1}$, $\Gamma \approx 2.57$.

Таким образом, интеркалирующие лиганды вызывают существенное увеличение степени гибридизации. Одним из важных параметров, ответственных за чувствительность ДНК–чипа является концентрация мишеней, которая приводит к половинной гибридизации, c_{50}^t [128–135].

Сдвиг концентрации c_{50}^t по сравнению со случаем без лигандов δc_{50}^t :

$$\delta c_{50}^t = \frac{e^{3\Gamma/2}}{K_t} \left[\left(\frac{c_l K_l}{c_l K_l + 1} \right)^N - 1 \right], \quad (4.25)$$

где $\Gamma = 8\pi N \sigma_0 l_B \frac{r_D^2}{H}$. Сдвиг концентраций для половинной гибридизации δc_{50}^t зависит от концентрации лигандов l (Рис. 4.3.).

Таким образом, интеркалирующие лиганды, связанные с гибридизированными зондами pt , существенно снижают концентрацию половинной гибридизации, а затем влияют на чувствительность ДНК-чипа в случае неконкурентной гибридизации.

Из системы уравнений (4.20) получаем изотерму гибридизации в случае конкурентной гибридизации на поверхности:

$$\begin{aligned} \frac{x}{1-x-y} &= c_t \tilde{K}_t \exp\left(-\frac{N}{k_B T} \frac{\partial \gamma_{el}}{\partial \sigma}\right) \\ \frac{y}{1-x-y} &= c_m \tilde{K}_m \exp\left(-\frac{N}{k_B T} \frac{\partial \gamma_{el}}{\partial \sigma}\right) \end{aligned} \quad (4.26)$$

где $\tilde{K}_t = K_t (1 - r_1^*)^{-N}$ и $\tilde{K}_m = K_m (1 - r_2^*)^{-M}$ являются перенормированными константами связывания, аналогичные результатам, полученным для случая без лигандов [135]:

$$\begin{aligned} K_t &= \exp\left(-\frac{\mu_{pt}^0 - \mu_p^0 - \mu_t^0}{k_B T}\right) \\ K_m &= \exp\left(-\frac{\mu_{pm}^0 - \mu_p^0 - \mu_m^0}{k_B T}\right) \end{aligned} \quad (4.27)$$

Если химический потенциал связанных лигандов μ_b^0 в исходном состоянии один и тот же для дуплексов pt и pm , то равновесная степень спиральности будет $r_1^* = r_2^*$ и получим, что

$$r^* = \frac{c_l K_l}{c_l K_l + 1}, \quad (4.28)$$

где константа связывания лигандов

$$K_l = \exp\left(-\frac{\mu_b^0 - \mu_l^0}{k_B T}\right). \quad (4.29)$$

Таким образом, в случае конкурентной гибридизации на поверхности интеркаляция лигандов значительно увеличивает константы связывания

мишеней и несовместимых последовательностей, вследствие чего повышается чувствительность ДНК-чипа.

В то же время, влияние лигандов является более выраженным для целевых последовательностей t из-за экспоненциальной зависимости от количества мест связывания N и M для констант связывания

$$\begin{aligned} \tilde{K}_t &= K_t (1 - r^*)^{-N} \\ \tilde{K}_m &= K_m (1 - r^*)^{-M} \end{aligned} \quad (4.30)$$

Данный эффект особенно важен при высоких значениях параметра $K_l c_l$, что соответствует высокой концентрации лигандов и/или большим значениям константы связывания K_l . Если количество мест связывания M для несовместимых дуплексов pm существенно отличается от количества мест связывания комплементарных дуплексов pt , то можно получить значительное увеличение селективности.

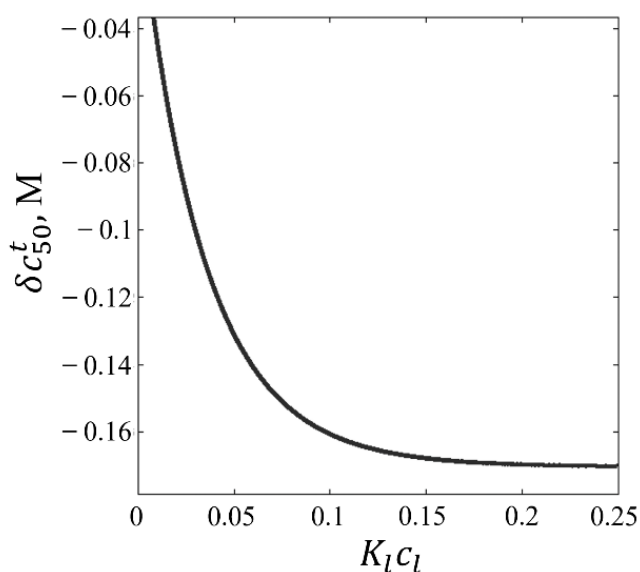


Рис. 4.3. Зависимость сдвига концентраций для половинной гибридизаций от $K_l c_l$.

Заключение

Исследованы термодинамические свойства поверхности ДНК-чипа с закрепленными к нему ДНК-зондами, взаимодействующие с ДНК-

мишенями и лигандами в растворе. Проанализированы некоторые факторы, влияющие как на термодинамику гибридизации ДНК на границе раздела твердое тело–раствор. Для случаев неконкурентной и конкурентной гибридизации ДНК на поверхности исследованы такие термодинамические характеристики системы, как изотермы гибридизации и концентрация ДНК–мишеней, соответствующие половинной гибридизации ДНК–мишеней с пробами. Проведен сравнительный анализ изотерм гибридизации ДНК–зонд с ДНК–мишенями и ДНК с последовательностями, лишь частично комплементарными ДНК–зондам, который показывает, что связывание с интеркалирующими лигандами приводит к увеличению избирательности и чувствительности ДНК–чипов.

Выводы

1. Однонитевая РНК со случайной последовательностью нуклеотидов демонстрирует два, или три пика на температурной зависимости удельной теплоемкости системы как при притяжении, так и при отталкивании между одинаковыми нуклеотидами и различным содержании различных типов нуклеотидов. Подобное поведение указывает на наличие двух, или трех структурных переходов.
2. Низкотемпературный пик теплоемкости соответствует холодной денатурации РНК, когда степень спиральности существенно падает при понижении температуры. Данный эффект обусловлен большим количеством термодинамически невыгодных контактов и подтвержден экспериментально.
3. Коллапс однонитевой РНК сопровождается существенной перестройкой, как вторичной, так и пространственной структур и является неравновесным фазовым переходом второго рода.
4. Фазовый переход имеет место между стеклообразной фазой, где реализуется несколько вариантов третичной структуры для каждой вторичной и фазой, где одной вторичной структуре соответствует одна третичная. Т.о. гибридизации оц-РНК сопровождается настройкой третичной структуры в соответствие со вторичной.
5. В стеклообразной фазе вторичная структура работает как холодильник и забирает энергию из пространственных степеней свободы.
6. Связывание интеркалирующих лигандов с ДНК в растворе и на поверхности ДНК – сенсора повышает чувствительность и избирательность работы ДНК – сенсора.

Литература:

1. D. Soll, *The RNA World*, chapter Transfer RNA: an RNA for all seasons, pages 157-183, Cold Spring Harbor Laboratory Press, 1993.
2. M. Sprinzl, C. Steegborn, F. Hubel, S. Steinberg, The transfer RNA database. *Nucl. Acids Res.*, 24:68-72, 1996.
3. P.B. Moore, *The RNA World*, chapter Ribosomes and the RNA world, pages 119-135, Cold Spring Harbor Laboratory Press, 1993.
4. H.F. Noller, *The RNA World*, chapter On the origin of the ribosome: coevolution of subdomains of RNA and RNA, pages 137-156, Cold Spring Harbor Laboratory Press, 1993.
5. R.A. Zimmermann, A.E. Dahlberg, *Protein Biosynthesis*, chapter Ribosomal RNA: Structure, Evolution, Processing, and Function. CRC Press, 1996.
6. Y. Van de Peer, A. Caers, P. De Rijk, R. De Wachter, Database on the structure of small ribosomal subunit RNA, *Nucl. Acids Res.*, 26:179-182, 1998.
7. P. De Rijk, A. Caers, Y. Van de Peer, R. De Wachter, Database on the structure of small ribosomal subunit RNA, *Nucl. Acids Res.*, 26:183-186, 1998.
8. B.L. Maidak, J.R. Cole, C.T. Jr Parker, G.M. Garrity, N. Larsen, B. Li, T.G. Lilburn, M.J. McCaughey, G.J. Olsen, R. Overbeek, S. Pramanik, T.M. Schmidt, J.M. Tiedje, C.R. Woese, A new version of the rdp (ribosomal database project), *Nucl. Acids Res.*, 27:171-173, 1999.
9. S.M. Freier, R. Kierzek, J.A. Jaeger, N. Sugimoto, M.H. Caruthers, T. Nielson, D.H. Turner, Improved free energy parameters for prediction of RNA duplex stability, *PNAS*, 83:9373-9377, 1986.
10. J. Jr SantaLucia, D.H. Turner, Measuring the thermodynamics of RNA secondary structure formation, *Biopolymers*, 44:309-319, 1997.
11. A.M. Pyle, J.B. Green, RNA folding, *Curr. Opin. Struct. Biol.*, 5:303-310, 1995.
12. P. Brion, E. Westhof, Hierarchy and dynamics of RNA folding, *A. Rev. Biophys. Biomol. Struct.*, 26:113-137, 1997.
13. I. Jr. Tinoco, C. Bustamante, How RNA folds, *J. Mol. Biol.*, 293(2):271-281, 1999.
14. F.H.D. van Batenburg, A.P. Gulyaev, C.W.A. Pleij, J. Ng, J. Oliehoek, Pseudobase: a database with RNA pseudoknots, *Nucl. Acids Res.*, 28:201-204, 2000.
15. E. Westhof, L. Jaeger, RNA pseudoknots, *Cur. Opin. Struct. Biol.*, 2:327-333, 1992.
16. C.W. Hilbers, P.J.A. Michiels, H.A. Heus, New developments in structure determination of pseudoknots, *Biopolymers*, 48:137-153, 1998.

17. T. Hermann, D. J. Patel, Stitching together RNA tertiary architectures, *J. Mol. Biol.*, 294:829-849, 1999.
18. E. Rivas, S.R. Eddy, A dynamic programming algorithm for RNA structure prediction including pseudoknots, *J. Mol. Biol.*, 285:2053-2067, 1999.
19. A.A. Vedenov, A.M. Dykhne, M.D. Frank-Kamenetskii, The helix-coil transition in DNA, *Sov. Phys. Usp.*, 14:715-736, 1972.
20. Yu.S. Lazurkin, *Physical Methods of Studying Proteins and Nucleic Acids*, Nauka, 1967.
21. M.J. Chamberlin, Comparative properties of DNA, RNA, and hybrid homopolymer pairs, *Fed. Proc.*, 24:1446-1457, 1965.
22. R.B. Inman, R.L. Baldwin, Helix-random coil transitions in DNA homopolymer pairs, *J. Mol. Biol.*, 8:452-469, 1964.
23. S.G. Haroutiunian, Y.B. Dalyan, V.M. Aslanian, D.Yu. Lando, A.A. Akhrem, A new method for determining the relative effect of ligand on at- and gc- base pairs in DNA: application to metal ions, protons and two amino acids, *Nucl. Acids Res.*, 18(21):6413-6417, 1990.
24. D.Yu. Lando, S.G. Haroutiunian, A.M. Kulba, Y.B. Dalyan, P. Orioli, R. Mangani, A.A. Akhrem, Theoretical and experimental study of DNA helix-coil transition in acidic and alkaline medium, *J. Biomol. Struct. Dyn.*, 12:355-366, 1994.
25. J. Marmur, P. Doty, Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature, *J. Mol. Biol.*, 5:109-118, 1962.
26. R.J. Owen, L.R. Hill, S.P. Lapage, Determination of DNA base compositions from melting profiles in dilute buffers, *Biopolymers*, 7:503-516, 1969.
27. Yu.N. Kosaganov, Yu.S. Lazurkin, N.V. Sidorenko, *Molekulyarnaya Biologiya*, 1:352, 1967.
28. J. Vinograd, J. Lebowitz, R. Watson, Early and late helix-coil transitions in closed circular DNA. the number of superhelical turns in polyoma DNA, *J. Mol. Biol.*, 33(1):173-197, 1968.
29. P.G. Higgs, RNA secondary structure: physical and computational aspects, *Q. Rev. Biophys.*, 33:199-253, 2000.
30. R. Nussinov, A.B. Jacobson, Fast algorithm for predicting the secondary structure of single-stranded RNA, *Proc. Natl. Acad. Sci. USA*, 77(11):6309-6313, 1980.
31. M. Zuker, P. Stiegler, Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information, *Nucleic Acids Res.*, 9(1):133-148, 1981.
32. J.S. McCaskill, The equilibrium partition function and base pair binding probabilities for RNA secondary structure, *Biopolymers*, 29:1105-1119, 1990.
33. B.H. Zimm, Theory of "melting" of the helical form in double chains of the DNA, *J. Chem. Phys.*, 33(5):1349-1356, 1960.

34. D. Poland, H.A. Scheraga, Occurrence of a phase transition in nucleic acid models, *J. Chem. Phys.*, 45:1464, 1966.
35. S. Lifson, Partition functions of linear chain molecules, *J. Chem. Phys.*, 40:3705, 1964.
36. A. Litan, S. Lifson, Statistical mechanical treatment of multistranded polynucleotide molecules, *J. Chem. Phys.*, 42:2528, 1965.
37. P.J. Flory, Theory of elastic mechanisms in fibrous proteins, *J. Am. Chem. Soc.*, 78:5222-5235, 1956.
38. M.E. Fisher, Effect of excluded volume on phase transition in biopolymers, *J. Chem. Phys.*, 45:1469, 1966.
39. Y. Kafri, D. Mukamel, L. Peliti, Melting and unzipping of DNA, *Eur. Phys. J. B*, 27(1):135-146, 2002.
40. Y. Kafri, D. Mukamel, L. Peliti, Why is the DNA denaturation transition first order? *Phys. Rev. Lett.*, 85:4988-4991, 2000.
41. M. Peyrard, A.R. Bishop, Statistical mechanics of a nonlinear model for DNA denaturation, *Phys. Rev. Lett.*, 62(23):2755-2758, 1989.
42. T. Dauxois, M. Peyrard, A.R. Bishop, Dynamics and thermodynamics of a nonlinear model for DNA denaturation, *Phys. Rev. E*, 47(1):684-695, 1993.
43. T. Dauxois, M. Peyrard, A.R. Bishop, Entropy-driven DNA denaturation, *Rap. Comm. Phys. Rev. E*, 47(1):R44-R47, 1993.
44. N. Theodorakopulos, T. Dauxois, M. Peyrard. Order of the phase transition in models of DNA thermal denaturation, *Phys. Rev. Lett.*, 85(1):6-9, 2000.
45. M. Peyrard, Nonlinear dynamics and statistical physics of DNA, *Nonlinearity*, 17:R1-R40, 2004.
46. W.B. Melchior Jr., P.H. Von Hippel, Alteration of the relative stability of da-dt and dg-dc base pairs in DNA, *Proc. Nat. Acad. Sci. USA*, 70(2):298-302, 1973.
47. V.F. Morozov, E.Sh. Mamasakhlisov, Sh.A. Hayryan, Chin-Kun Hu, Microscopical approach to the helix-coil transition in DNA, *Physica A*, 281:51-59, 2000.
48. R.F. Gesteland, J. F. Atkins, *The RNA World*, Cold Spring Harbor Monograph Series 2, Cold Spring Harbor Laboratory Press, New York, 1993.
49. P. G. Higgs, RNA secondary structure: a comparison of real and random sequences, *J. Phys. I France*, 3:43-59, 1993.
50. D. Lambert, D.E. Draper, Denaturation of RNA secondary and tertiary structure by urea: simple unfolded state models and free energy parameters account for measured m-values, *Biochemistry*, 51(44):9014-9026, 2012.
51. P.J. Mikulecky, A.L. Feig, Cold denaturation of the hammerhead ribozyme, *J. Am. Chem. Soc.*, 124(6):890-891, 2002.

52. P.J. Mikulecky, A.L. Feig, Heat capacity changes in RNA folding: application of perturbation theory to hammerhead ribozyme cold denaturation, *Nucleic Acids Res.*, 32(13):3967-3976, 2004.
53. P.L. Privalov, Cold denaturation of proteins, *Crit. Rev. Biochem. Mol. Biol.*, 25(4):281-305, 1990.
54. A.V. Finkelstein, *Protein Physics*, Academic Press, New York, 2002.
55. F. Sedlmeier, D. Horinek, R.R. Netz, Entropy and enthalpy convergence of hydrophobic solvation beyond the hard-sphere limit, *J. Chem. Phys.*, 134:055105, 2011.
56. G. Vernizzi, H. Orland, A. Zee, Enumeration of RNA Structures by Matrix Models, *Phys. Rev. Lett.*, 94:168103, 2005.
57. T.R. Einert, P. Nager, H. Orland, R.R. Netz, Impact of Loop Statistics on the Thermodynamics of RNA Folding, *Phys. Rev. Lett.*, 101:048103, 2008.
58. M. Manosas, F. Ritort, Thermodynamic and kinetic aspects of RNA pulling experiments, *Biophys. J.*, 88:3224-3242, 2005.
59. P.G. Higgs, Overlaps between RNA Secondary Structures, *Phys. Rev. Lett.*, 76:704, 1996.
60. A. Pagnani, G. Parisi, F. Ricci-Tersenghi, Glassy Transition in a Disordered Model for the RNA Secondary Structure, *Phys. Rev. Lett.*, 84:2026, 2000.
61. R. Bundschuh, T. Hwa, Statistical mechanics of secondary structures formed by random RNA sequences, *Phys. Rev. E*, 65:031903, 2002.
62. F. Krzakala, M. Mezard, M. Mueller, Nature of the glassy phase of RNA secondary structure, *Europhys. Lett.*, 57:752-758, 2002.
63. M. Lassig, K.J. Wiese, The freezing of random RNA, *Phys. Rev. Lett.*, 96:228101, 2006.
64. T. Morita, *Statistical Mechanics of Quenched Solid Solutions with Application to Magnetically Dilute Alloys*, *J. Math. Phys. (N.Y.)*, 5:1401, 1964.
65. R. Kühn, Gibbs vs. Non-Gibbs in the Equilibrium Ensemble Approach to Disordered Systems, *Markov Proc. Relat. Fields*, 10:523-545, 2004.
66. M. Serva, G. Paladin, Gibbs thermodynamic potentials for disordered systems, *Phys. Rev. Lett.*, 70:105-108, 1993.
67. T. Liu, R. Bundschuh, Quantification of the differences between quenched and annealed averaging for RNA secondary structures, *Phys. Rev. E*, 72:061905/1-061905/9, 2005.
68. M. Zuker, P. Stiegler, Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information, *Nucleic Acids Res.*, 9:133-148, 1981.
69. R. Lorenz, S.H. Bernhart, C. Höner zu Siederdisen, H. Tafer, C. Flamm, P.F. Stadler, I.L. Hofacker, *ViennaRNA Package 2.0*, *I. Algorithms for Molecular Biology*, 6:26-39, 2011.
70. M. Mezard, G. Parisi M. Virasoro, *Spin Glass Theory and Beyond*, World Scientific, Singapore, 1987.
71. See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.113.068101>.

72. S.R. Morgan, P.G. Higgs, Barrier heights between ground states in a model of RNA secondary structure, *J. Phys. A*, 31:3153–3170, 1998.
73. A.V. Badasyan, Sh.A. Tonoyan, Y.Sh. Mamasakhlishov, A. Giacometti, A.S. Benight, V.F. Morozov, Competition for hydrogen-bond formation in the helix-coil transition and protein folding, *Phys. Rev. E*, 83:051903, 2011.
74. S. Wolfsheimer, B. Burghardt, A. Mann, A.K. Hartmann, RNA secondary structures: complex statics and glassy dynamics, *J. Stat. Mech.*, 3:P03005(1-28), 2008.
75. G.N. Hayrapetyan, F. Iannelli, J. Lekscha, V.F. Morozov, R.R. Netz, Y.S. Mamasakhlishov, Reentrant Melting of RNA with Quenched Sequence Randomness, *Phys. Rev. Lett.*, 113:068101, 2014.
76. P. Bhadola, N. Deo, Genus distribution and thermodynamics of a random matrix model of RNA with Penner interaction, *Phys. Rev. E*, 88:032706(1-14), 2013.
77. G. Paladin, M. Pasquini, M. Serva, Constrained Annealing for Spin Glasses, *J. Phys. I (France)*, 5:337-354, 1995.
78. M.V. Dotsenko, *An Introduction to the Theory of Spin Glasses and Neural Networks*, World Scientific, 1994.
79. T.G. Drummond, M.G. Hill, J.K. Barton, Electrochemical DNA sensors, *Nat. Biotechnol.*, 21:1192-1199, 2003.
80. J.M.S. Bartlett, *D. Stirling in PCR Protocols* edited by J.M.S. Bartlett, D. Stirling, Humana Press, Totowa, 545, 2003.
81. L. Jin, R.V. Lloyd, *J. Clin. In situ hybridization: Methods and applications*, *J. Clinical Lab. Anal.*, 11:2-9, 1997.
82. I. Tinoco Jr., C. Bustamante, How RNA folds, *J. Mol. Biol.*, 293:271-281, 1999.
83. D.K. Treiber, J.R. Williamson, Exposing the kinetic traps in RNA folding, *Curr. Opin. Struct. Biol.*, 9:339-345, 1999.
84. X. Zhuang, H. Kim, M.J. Pereira, H.P. Babcock, N.G. Walter, S. Chu, Correlating structural dynamics and function in single ribozyme molecules, *Science*, 296:1473-1476, 2002.
85. S.Wood, A.R. Ferre-D'Amare, D. Rueda, Allosteric Tertiary Interactions Preorganize the c-di-GMP Riboswitch and Accelerate Ligand Binding, *ACS Chem. Biol.*, 7:920-927, 2012.
86. E.I. Shakhnovich, A.M. Gutin, Implications of thermodynamics of protein folding for evolution of primary sequences, *Nature*, 346:773-775, 1990.
87. T. Garel, H. Orland, Mean-Field Model for Protein Folding, *Europhys. Lett.*, 6:597, 1988.
88. P.G. Wolynes, J.N. Onuchik, D. Thirumalai, Navigating the folding routes, *Science*, 267:1619-1620, 1995.

89. E.I. Shakhnovich, A.M. Gutin, Formation of unique structure in polypeptide chains: Theoretical investigation with the aid of a replica approach, *Biophys. Chem.*, 34:187-199, 1989.
90. V.S. Pande, A.Yu. Grosberg, T. Tanaka, Heteropolymer freezing and design: Towards physical models of protein folding, *Rev. Mod. Phys.*, 72:259, 2000.
91. T. Garel, H. Orland, E. Pitard, in *Spin Glasses and Random Fields*, edited by A.P. Young, World Scientific, Singapore, 443, 1998.
92. Y.Sh. Mamasakhlishov, Sh. Hayryan, C.-K. Hu, Random Sequences with Power-law Correlations Exhibit Protein-like Behavior, *J. Chem. Phys.*, 126:145103, 2007.
93. D. Poland, H.A. Scheraga, *J. Chem. Phys.*, 45:1456, 1966.
94. D. Cule, T. Hwa, Denaturation of Heterogeneous DNA, *Phys. Rev. Lett.*, 79:2375-2378, 1997.
95. C. Richard, A.J. Guttmann, Poland–Scheraga Models and the DNA Denaturation Transition, *J. Stat. Phys.*, 115(3/4):925-947, 2004.
96. G.N. Hayrapetyan, Y.Sh. Mamasakhlishov, V.F. Morozov, V.V.Papoyan, V.B. Priezzhev, Two-dimensional random walk and critical behavior of double-strand DNA, *J. Phys. A*, 46:035001, 2013.
97. R. Bundschuh, T. Hwa, Phases of the Secondary Structures of RNA Sequences, *Europhys. Lett.*, 59:903-909, 2002.
98. M. Müller, F. Krzakala, M. Mezard, The secondary structure of RNA under tension, *Eur. Phys. J. E*, 9(1):67-77, 2002.
99. F. David, C. Hagendorf, K.J. Wiese, A growth model for RNA secondary structures, *J. Stat. Mech.*, 04:P04008(1-43), 2008.
100. M. Bon, G. Vernizzi, H. Orland, A. Zee, Topological classification of RNA structures, *J. Mol. Biol.*, 379:900-911, 2008.
101. T.R. Einert, R.R. Netz, Theory for RNA Folding, Stretching, and Melting Including Loops and Salt, *Biophys J.*, 100(11):2745–2753, 2011.
102. M. Bon, C. Micheletti, H. Orland, A Monte Carlo algorithm to predict RNA secondary structures with pseudoknots, *Nucl. Acids Res.*, 41:1895-1900, 2013.
103. Y.Sh. Mamasakhlishov, Sh. Hayryan, V.F. Morozov, C.-K. Hu, RNA folding in the presence of counterions, *Phys. Rev. E*, 75:061907, 2007.
104. A. Siber, R. Podgornik, Nonspecific interactions in spontaneous assembly of empty versus functional single-stranded RNA viruses, *Phys. Rev. E*, 78:051915, 2008.
105. C. Hajdin, F. Ding, N.V. Dokholyan, K.M. Weeks, How good is that RNA tertiary structure prediction? *RNA*, 16:1340-1349, 2010.
106. A.M. Gutin, A.Yu. Grosberg, E.I. Shakhnovich, Globular state of branched random heteropolymers, *J. Phys. A: Math. Gen.*, 26:1037, 1993.

107. D. Schwab, R.F. Bruinsma, R. F. Flory Theory of the Folding of Designed RNA Molecules, *J. Phys. Chem. B*, 113:3880-3893, 2009.
108. P.G. de Gennes, Statistics of branching and hairpin helices for the dAT copolymer, *Biopolymers*, 6:715-729, 1968.
109. Y.Sh. Mamasakhlisov, Sh. Hayryan, V.F. Morozov, C.-K. Hu, Kinetics of the long ssRNA: Steady state, *EPL*, 106:48007, 2014.
110. A.R. Shultz, P.J. Flory, Phase Equilibria in Polymer-Solvent Systems, *J. Am. Chem. Soc.*, 74:4760-4767, 1952.
111. R. Shiman, D.E. Draper, Stabilization of RNA tertiary structure by monovalent cations, *J. Mol. Biol.*, 302:79-91, 2000.
112. R. Das, L.W. Kwok, I.S. Millett, Y. Bai, T.T. Mills, J. Jacob, G.S. Maskel, S. Seifert, S.G. Mochrie, P. Thiyagarajan, S. Doniach, L. Pollack, D. Hershlag, The fastest global events in RNA folding: electrostatic relaxation and tertiary collapse of the Tetrahymena ribozyme, *J. Mol. Biol.*, 332:311-319, 2003.
113. W. Zhang, S.-J. Chen, RNA hairpin-folding kinetics, *Proc. Natl. Acad. Sci. U.S.A.*, 99:1931-1936, 2002.
114. V. Dotsenko, S. Franz, M. Mezard, Partial annealing and overfrustration in spin-glasses and neural networks, *J. Phys. A: Math. Gen.*, 27:2351-2365 1994.
115. K. Binder, A.P. Young, Spin glasses: Experimental facts, theoretical concepts, and open questions, *Rev. Mod. Phys.*, 58:801-976, 1986.
116. M. Mezard, G. Parisi, Replica field theory for random manifolds, *J. Phys. (Paris)*, 1:809-836, 1991.
117. V. Dotsenko, Introduction to the Replica Theory of Disordered Statistical Systems, Cambridge University Press, Cambridge, 220, 2001.
118. A.E. Allahverdyan, K.G. Petrosyan, Anomalous Latent Heat in Nonequilibrium Phase Transitions, *Phys. Rev. Lett.*, 96:065701, 2006.
119. G. Parisi, Order Parameter for Spin-Glasses, *Phys. Rev. Lett.*, 50:192, 1983.
120. M. Mezard, G. Parisi, N. Sourlas, G. Toulouse, M. Virasoro, Nature of the Spin-Glass Phase, *Phys. Rev. Lett.*, 52:1156-1159, 1984.
121. T. Temesvari, C. De Dominicis, I.R. Pimentel, Generic replica symmetric field-theory for short range Ising spin glasses, *Eur. Phys. J. B*, 25:361-372, 2002.
122. R.K.P. Zia, E.L. Praestgaard, O.G. Mouritsen, Getting more from pushing less: Negative specific heat and conductivity in nonequilibrium steady states, *Am. J. Phys.*, 70:384-392, 2002.
123. T. Tao, T. Odagaki, A. Yoshimori, Cooling rate dependence of specific heat in systems out of equilibrium, *J. Chem. Phys.*, 122:044505, 2005.

124. J. Bisquert, Master equation approach to the non-equilibrium negative specific heat at the glass transition, *Am. J. Phys.*, 73:735-741, 2005.
125. D. Ivniński, I. Abdel-Hamid, P. Atanasov, E. Wilkins. *Biosensors and Bioelectronics*, 14:599, 1999.
126. J. Labuda, A.M.O. Brett, G. Evtugyn, M. Fojta, M. Mascini, M. Ozsoz, I. Palchetti, E. Paleček, J. Wang, *Electrochemical Nucleic Acid-Based Biosensors: Concepts*, *Pure Appl. Chem.*, 82:1161-1187, 2010.
127. J.H. Watterson, P.A.E. Piunno, U.J. Krull, Towards the optimization of an optical DNA sensor: control of selectivity coefficients and relative surface affinities, *Anal. Chem. Acta.*, 457:29-38, 2002.
128. A. Halperin, A. Buhot, E. B. Zhulina, On the hybridization isotherms of DNA microarrays: The Langmuir model and its extensions, *J. Phys. Condens. Matter*, 18:S463-S490, 2006.
129. G. Ananyan, A. Avetisyan, L. Aloyan, Y. Dalyan, The stability of DNA–porphyrin complexes in the presence of Mn(II) ions, *Biophys. Chem.*, 156:96-101, 2011.
130. A.A. Ghazaryan, Y.B. Dalyan, S.G. Haroutiunian, A. Tikhomirova, N. Taulier, J.W. Wells, T.V. Chalikian, Thermodynamics of Interactions of Water-Soluble Porphyrins with RNA Duplexes, *J. Amer. Chem. Soc.*, 128:1914-1921, 2006.
131. R.F. Pasternack, J.I. Goldsmith, S. Szep, E.J. Gibbs, A spectroscopic and thermodynamic study of porphyrin/DNA supramolecular assemblies, *Biophys. J.*, 75:1024-1031, 1998.
132. D.M. Hinckley, G.S. Freeman, J.K. Whitmer, J.J. de Pablo, An experimentally-informed coarse-grained 3-site-per-nucleotide model of DNA: Structure, thermodynamics, and dynamics of hybridization, *J. Chem. Phys.*, 139:144903, 2013.
133. D.M. Hinckley, J.P. Lequeieu, J.J. de Pablo, Coarse-grained modeling of DNA oligomer hybridization: Length, sequence, and salt effects, *J. Chem. Phys.*, 141:035102, 2014.
134. A.W. Peterson, R.J. Heaton, R.M. Georgiadis, The effect of surface probe density on DNA hybridization, *Nucl. Acids Res.*, 29:5163-5168, 2001.
135. A. Halperin, A. Buhot, E.B. Zhulina, Sensitivity, Specificity and the Hybridization Isotherms of DNA Chips, *Biophys. J.*, 86:718-730, 2004.
136. M.F. Hagan, A.K. Chakraborty, Hybridization Dynamics of Surface Immobilized DNA, *J. Chem. Phys.*, 120:4958-4968, 2004.
137. M.M.A. Seckar, W. Bloch, P.M.S. John, Comparative study of sequence-dependent hybridization kinetics in solution and on microspheres, *Nucleic Acids Res.*, 33:366-375, 2005.
138. N.V. Sorokin, V.R. Chechetkin, S.V. Pankov, O.G. Somova, M.A. Livshits, M.Y. Donnikov, A.Y. Turygin, V.E. Barsky, A.S. Zasedatelev, Kinetics of hybridization on surface oligonucleotide microchips: Theory, experiment, and comparison with hybridization on gel-based microchips, *J. Biomol. Struct. Dyn.*, 24:57-66, 2006.

139. D. Irving, P. Gong, R. Levicky, DNA Surface Hybridization: Comparison of Theory and Experiment *J. Phys. Chem. B*, 114:7631-7640, 2010.
140. T.J. Schmitt, T.A. Knotts IV, Thermodynamics of DNA hybridization on surfaces, *J. Chem. Phys.*, 134:205105, 2011.
141. S.M. Nelson, L.R. Ferguson, W.A. Denny, Non-covalent ligand/DNA interactions: Minor groove binding agents, *Mutat. Res.*, 623:24-40, 2007.
142. V.V. Kostjukov, A.A.H. Santiago, F.R. Rodriguez, S.R. Castilla, J.A. Parkinson, M.P. Evstigneev, Energetics of ligand binding to the DNA minor groove, *Phys. Chem. Chem. Phys.*, 14:5588, 2012.
143. C.G. Ricci, P.A. Netz, Docking studies on DNA-ligand interactions: building and application of a protocol to identify the binding mode, *J. Chem. Inf. Model.*, 49:1925-1935, 2009.
144. C. Tanford, Chemical potential of bound ligand, an important parameter for free energy transduction, *Proc. Natl. Acad. Sci. USA*, 78:270-273, 1981.
145. P. Pincus, Colloid stabilization with grafted polyelectrolytes *Macromolecules*, 24:2912-2919, 1991.
146. J. Wittmer, J.F. Joanny, Charged diblock copolymers at interfaces, *Macromolecules*, 26:2691-2697, 1993.
147. O.V. Borisov, E.B. Zhulina, T.M. Birshtein, Diagram of the States of a Grafted Polyelectrolyte Layer, *Macromolecules*, 27:4795-4803, 1994.
148. I.Y. Wong, N.A. Melosh, An Electrostatic Model for DNA Surface Hybridization, *Biophys. J.*, 98:2954-2963, 2010.