# Московский государственный университет имени М.В. Ломоносова физический факультет

На правах рукописи

Нечипуренко Дмитрий Юрьевич

Анализ специфичности расщепления ДНК ультразвуком

03.01.02 - Биофизика

Диссертация на соискание ученой степени кандидата физико-математических наук

Научный руководитель доктор физ.-мат. наук, профессор Твердислов Всеволод Александрович

Москва - 2010

# Оглавление

Введение	3
Глава 1. Литературный обзор	8
1.1 Подходы к изучению контекстно-зависимых физико-химических свойств ЛНК	8
1 2 Физика упругих растяжений ЛНК	17
1 3 Расшепление полимеров пол лействием ультразвука	23
Глава 2. Ультразвуковое расшепление ЛНК: обработка и анализ экспериментальн	ых
ланных	
2.1 Описание эксперимента и процедуры обработки экспериментальных	
данных	27
2.2 Характерные свойства ультразвукового расщепления ДНК	30
2.3 Статистический анализ специфичности расшепления ДНК	
VЛЬТDАЗВVКОМ	34
Глава 3. Моделирование расшепления фрагментов ЛНК под действием акустическ	кой
кавитации	45
3.1 Моделирование динамики кавитационного пузырька	45
3.2 Моделирование взаимодействия фрагмента ДНК с кавитационным	
течением	51
3.3 Моделирование кинетики расшепления фрагментов ДНК	
под действием акустической кавитации	55
3.4 Обсуждение	61
Глава 4. Подходы к интерпретации специфичности расщепления ДНК	
ультразвуком	64
4.1 Особенности конформационной динамики В- формы ДНК	64
4.2 Влияние конформационной подвижности дезоксирибозы на эффективн	ость
ультразвукового расщепления ДНК	67
4.3 Обсуждение	72
Глава 5. Исследование особенностей ультразвукового расщепления функционалы	ных
областей ДНК	75
5.1. Особенности ультразвукового расщепления ДНК λ-фага	75
5.2 Применение данных по специфичности ультразвукового	
расщепления ДНК для анализа промоторных областей генома человека	79
Основные результаты и выводы	84
Литература	85
Публикации по теме диссертации	89
Благодарности	90

# Введение

#### Актуальность проблемы

В процессах функционирования ДНК в живой клетке существенную роль играют взаимодействия ДНК с белковыми комплексами и другими лигандами. Эти взаимодействия, как правило, реализуются по принципу «узнавания» молекулами - лигандами определенных сайтов молекулы ДНК.

Среди множества видов ДНК - белкового узнавания выделяют два принципиально различных типа – это так называемое «прямое» и «непрямое» узнавание. В случае «прямого» узнавания белок распознает определенную последовательность пар оснований ДНК, образуя с их функциональными группами сеть контактов, присущую только данной последовательности нуклеотидов и соответствующую геометрии белковой молекулы. В случае же «непрямого» узнавания избирательность связывания белковой молекулы с ДНК определяется локальными и зависящими от нуклеотидной последовательности конформационно-динамическими характеристиками ДНК – такими как гибкость, термодинамическая стабильность двойной спирали, её геометрия, подвижность определенных молекулярных групп и т.п. Таким образом, зависимость локальных конформационно-динамических свойств от последовательности пар оснований в ДНК играет важнейшую роль при функционировании молекулы ДНК в клетке. Поэтому, изучение контекстно-зависимых конформационных и динамических свойств молекулы ДНК является одной из важнейших задач молекулярной биофизики.

Существует ряд экспериментальных подходов к изучению структурных свойств молекулы ДНК. Для исследования влияния последовательности нуклеотидов на конформацию ДНК применяются методы рентгеноструктурного анализа, ЯМР, и ИК - спектроскопии. Экспериментальные данные о гибкости молекулы ДНК получают при помощи анализа её подвижности в геле и исследуя расщепление ДНК неспецифичными эндонуклеазами, а изменения геометрии малой бороздки двойной спирали вдоль ДНК изучают методами химического расщепления молекулы гидроксильными радикалами и другими химическими агентами.

3

При помощи перечисленных экспериментальных методик достаточно сложно извлечь информацию о динамических характеристиках изучаемых фрагментов ДНК. Здесь на помощь приходят методы молекулярного моделирования: молекулярная динамика, метод Монте-Карло, а также квантово-химические расчеты.

В Институте молекулярной биологии РАН в настоящее время развивается новый экспериментальный метод, позволяющий изучать конформационно-динамические свойства двойной спирали ДНК. Метод основан на анализе картин расщепления фрагментов ДНК под действием ультразвука высокой интенсивности. Контекстная специфичность расщепления, то есть зависимость профилей ультразвукового расщепления фрагментов ДНК от их нуклеотидной последовательности, позволяет изучать влияние последовательности пар оснований в ДНК на ее структурные свойства в масштабах от нескольких десятков до сотен нуклеотидов.

Явление контекстной специфичности разрывов ДНК под действием ультразвука представляет несомненный научный интерес как дополнительный источник информации о контекстно-зависимых характеристиках ДНК. Тем не менее, физика этого явления практически не изучена, что вызывает серьезные трудности при попытке интерпретации полученных результатов.

Анализ и моделирование расщепления ДНК ультразвуком, которым посвящена данная работа, необходимы для более глубокого исследования физики этого процесса с целью дальнейшего применения и развития основанного на этом явлении методе изучения конформационно-динамических свойств ДНК.

# Цели и задачи диссертационной работы

Целью данного исследования являлось выявление основных закономерностей процесса расщепления ДНК под действием ультразвука, разработка физических моделей, адекватно описывающих характерные особенности этого явления и применение полученных данных по расщеплению ДНК для анализа функциональных участков ДНК человека. Для достижения этих целей решались следующие основные задачи:

- установить характерные особенности расщепления ДНК ультразвуком;
- провести анализ контекстной специфичности расщепления;
- разработать модель, описывающую процесс расщепления фрагментов ДНК;
- сравнить полученные теоретические результаты с экспериментальными;

- разработать модель, качественно описывающую явление контекстной специфичности расщепления ДНК ультразвуком.

- исследовать особенности теоретических профилей ультразвукового расщепления, построенных для промоторных последовательностей ДНК человека.

#### На защиту выносятся следующие положения и результаты:

- получены относительные частоты ультразвукового расщепления фрагментов ДНК в ди- и тетрануклеотидном приближении;

- выявлено увеличение степени ультразвукового расщепления фосфодиэфирной связи, следующей за дезоксицитидином (в направлении от 5' к 3' концу фрагмента ДНК);

- предложен подход к моделированию процесса расщепления ДНК под действием кавитационных эффектов, который позволяет описать характерные особенности наблюдаемого расщепления;

- разработана модель, позволяющая качественно описать явление контекстной специфичности ультразвукового расщепления;

- продемонстрирована возможность применения полученных относительных частот ультразвукового расщепления для анализа промоторных участков ДНК человека.

## Апробация результатов диссертационной работы

По материалам диссертации опубликовано 7 научных работ, из них 5 статей в рецензируемых научных журналах ВАК России:

«Биофизика», «Журнал Структурной Химии», «Journal of Biomolecular Structure & Dynamics», «In collection: NATO Science for Peace and Security Series B: Physics and Biophysics, Nanomaterials for Application in Medicine and Biology», «Biophysical Journal».

Основные результаты исследований, представленные в диссертационной работе, докладывались на следующих международных и российских конференциях:

13-ой международной конференции «Математика, компьютер, образование. (Дубна 2006), 15-ой международной конференции «Математика, компьютер, образование» (Дубна, 2008), 17 -ой международной конференции «Математика, компьютер, образование» (Дубна, 2010) и 15-ом Симпозиуме по межмолекулярному

взаимодействию и конформациям молекул (Петрозаводск, 2010), The second Saint-Petersburg International Conference on NanoBio Technologies, (NanoBio' 08, Санкт Петербург, Россия, 2008), 7<sup>th</sup> EBSA European Biophysics Congress, (EBSA, Генуя, Италия, 2009), Solvation and Ionic Effects in Biomolecules: Theory to Experiment, (Цахкадзор, Армения, 2010).

Список опубликованных статей по теме диссертации приведен в конце настоящего автореферата.

#### Научная новизна и практическая значимость работы

Все представленные выше результаты получены впервые. Предложенный подход к моделированию расщепления ДНК под действием ультразвука высокой интенсивности позволяет описать характерные особенности ультразвукового расщепления ДНК. Выявленная корреляция относительных частот ультразвукового расщепления с имеющимися в литературе данными о конформационной подвижности дезоксирибозы позволяет построить модель, качественно описывающую явление контекстной специфичности расщепления ДНК. В соответствии с предложенной моделью, различие в относительных степенях ультразвукового расщепления является следствием отличия конформационной динамики дезоксирибозных групп сахарофосфатного остова ДНК. Таким образом, относительные частоты ультразвукового расщепления, по всей видимости, позволяют описывать влияние нуклеотидной последовательности ДНК на подвижность определенных участков сахарофосфатного остова. Полученные результаты могут быть использованы для выявления функциональных сайтов ДНК при анализе геномных последовательностей.

# Личный вклад автора

Все результаты оригинальных теоретических исследований получены лично автором, либо при его непосредственном участии. Экспериментальные результаты были получены в лаборатории ДНК-белковых взаимодействий Института молекулярной биологии им. В.А. Энгельгардта РАН к.х.н. С.Л. Гроховским.

6

# Структура и объем диссертации

Диссертационная работа состоит из введения, пяти глав и заключения. Каждая глава снабжена краткой аннотацией, состоит из нескольких разделов и заключения. В конце работы приведен библиографический список используемой литературы и список публикаций автора по теме диссертации. Полный объем диссертационной работы составляет 90 страниц, включая 30 рисунков.

# Глава 1. Литературный обзор

### 1.1. Подходы к изучению контекстно-зависимых физико-химических свойств ДНК

В настоящее время для компьютерного анализа стали доступны нуклеотидные последовательности геномов множества организмов. Большое значение приобретает развитие эффективных методов обнаружения функциональных сайтов расшифрованных последовательностей [Liolios et al, 2006; Wang et al, 2006]. Несмотря на значительные успехи методов биоинформатики, задача точного определения регуляторных участков в геноме остается одной из важнейших в молекулярной биологии. Определение промоторных областей в ДНК, то есть участков, отвечающих за регуляцию работы генов, важно не только для обнаружения новых генов, но и для локализации областей генома, в которых следует проводить экспериментальный и теоретический поиск сайтов, наиболее важных для регуляции транскрипции.

Известно, что свойства промоторных участков ДНК, как правило, сильно отличаются от свойств других областей генома [Florquin et al, 2005]. Большинство существующих алгоритмов поиска промоторных участков в геномах опираются на свойства нуклеотидных последовательностей уже известных промоторов и используют такие методы, как дискриминантный анализ, цепи Маркова и искусственные нейронные сети [Sonnenburg et al, 2006]. Программы, реализующие алгоритмы такого рода, требуют большие объемы данных для «обучения», и являются зачастую ориентированными на определенный вид организмов, а также далеко не всегда способны быстро работать при анализе на уровне целого генома [Bajic et al, 2004; Sonnenburg et al, 2006].

В последние годы успешно развиваются новые подходы к поиску регуляторных участков в геномах, основанные на использовании данных о контекстной зависимости различных физико-химических характеристик ДНК [Florquin et al, 2005; Abelev et al, 2008; Dineen et al, 2009]. Основной идеей таких подходов является теоретическое построение и последующий анализ профиля, характеризующего изменение вдоль ДНК определенного физико-химического параметра. При этом используются самые разные локальные характеристики ДНК – гибкость, энергия стэкинга, температура плавления и

многие другие. Зависимость используемых физических характеристик ДНК от нуклеотидной последовательности, как правило, описывается в рамках ди-, три-, и тетрануклеотидных приближений. Существенную роль играет вырожденность таких «физических кодов» по нуклеотидной последовательности ДНК: разные последовательности могут приводить к схожим профилям изменения локального структурного параметра ДНК. Вырожденность такого рода позволяет предложить новые критерии для анализа функциональных участков генома [Parker et al, 2009]. Описанные подходы представляют собой важное дополнение к общепринятым методам биоинформатики.

Стоит отметить, что при описании характерных свойств регуляторных последовательностей ДНК различные физико-химические характеристики молекулы в определенном смысле могут дополнять друг друга [Florquin et al, 2005]. В процессах взаимодействия ДНК с регуляторными и структурными белками важную роль играют локальные конформационно-динамические характеристики молекулы ДНК: в конечном итоге белок «узнает» не определенную последовательность пар оснований в ДНК, а структуру молекулы, как правило, меняя её локальную геометрию. Описано множество вариантов ДНК-белкового узнавания [Becker et al, 2006; Rohs et al, 2009], и для разных классов белков ключевыми факторами, определяющими «прочность» связывания, могут быть самые разные локальные характеристики ДНК: гибкость, геометрия спирали, стабильность дуплекса, подвижность определенных молекулярных групп и Как известно. все эти параметры в ДНК от нуклеотидной Т.Л. зависят последовательности, причем влияние последовательности локальные на характеристики ДНК может выходить далеко за рамки ди- и тетрануклеотидного описания [Faiger et al, 2006]. Таким образом, сложность и многообразие процессов регуляции генетической экспрессии требуют учёта неоднородности разных физических параметров молекулы ДНК при описании характерных свойств её функциональных сайтов.

Особенности ДНК-белковых взаимодействий, как правило, изучаются на модельных системах. Примером такой модельной системы, анализу которой посвящено множество работ, является система регуляции экспрессии бактериофага лямбда – (далее –  $\lambda$ –фага) [Becker et al, 2006].

На рисунке 1.1. приведена структура комплекса белка-репрессора системы регуляции экспрессии λ-фага с операторным участком ДНК.



**Рис. 1.1**. Структура комплекса белка-репрессора с операторным участком ДНК λфага, полученная методом рентгеноструктурного анализа. Разрешение структуры 1.8 A, ее код в базе данных NDB (http://ndbserver.rutgers.edu/): PDR010.

Как видно из рисунка, данный комплекс характеризуется наличием двух участков ДНК, в которых имеются специфические контакты с молекулой белка. Между ними находится центральный участок, не вступающий в прямое взаимодействие с белковым комплексом. Было показано, что константа образования данного комплекса существенным образом зависит от нуклеотидной последовательности центральной части операторного участка: при внесении точечных мутаций в данную область константа связывания может изменяться в десятки раз. Особенности образования данного комплекса являются характерными для так называемого «непрямого» типа ДНК-белкового узнавания. Было показано, что существенное влияние на величину константы образования данного комплекса оказывают определенные конформационнодинамические свойства центральной части операторного участка [Becker et al, 2006].

Константа реакции образования ДНК-белкового комплекса может быть описана в рамках равновесной термодинамики:

$$K \sim e^{-\frac{\Delta F}{kT}} \tag{1.1}$$

где К – константа равновесия реакции образования ДНК-белкового комплекса,

 $\Delta F$  – изменение свободной энергии системы при образовании комплекса, k – постоянная Больцмана, T – абсолютная температура. При расчете изменения свободной энергии, как правило, пренебрегают изменениями в структуре белковой молекулы. Общепринятым является предположение о том, что на стадии «узнавания», изменения конформации белковой молекулы, как правило, значительно меньше, чем изменения, происходящие в структуре ДНК, что может выражено следующим образом:

$$\Delta F = \Delta E_{DNA} - T \Delta S_{DNA} + V \tag{1.2},$$

где  $\Delta E_{DNA}$  – изменение энергии, связанное с изменением конформации ДНК,  $\Delta S_{DNA}$  – изменение энтропии ДНК, а V – потенциальная энергия взаимодействия ДНК с белковой молекулой. Таким образом, изменение свободной энергии системы в данном приближении определяется тремя основными факторами. При исследовании специфичности образования ДНК – белкового комплекса (например – при исследовании влияния последовательности центральной части операторного участка на эффективность образовании комплекса репрессор-оператор), часто используется следующее предположение, которое устанавливает иерархию перечисленных выше факторов. Взаимодействие белка с молекулой ДНК приводит к изменению конформации ДНК. Величина константы образования комплекса для данной нуклеотидной последовательности операторного участка зависит от энергетического и энтропийного барьера, которые нужно «преодолеть» для перехода ДНК в данную конформацию. То есть потенциальная энергия взаимодействия определяет оптимальную конформацию ДНК в комплексе, а величины  $\Delta E_{DNA}$  и  $\Delta S_{DNA}$ определяются исходя из особенностей данной конформации, а также конформационнодинамических свойств ДНК для данной последовательности.

Для определения величин  $\Delta E_{DNA}$  и  $\Delta S_{DNA}$ , как правило, используются данные о конформационной подвижности ДНК, полученные из анализа кристаллографических данных для свободных и связанных олигонуклеотидов ДНК, данных ЯМР и результатов компьютерного моделирования. При описании структуры ДНК достаточно часто применяется приближение жестких пар оснований – в таком случае, локальная конформация ДНК определяется шестью параметрами – тремя угловыми и тремя трансляционными, которые характеризуют относительное расположение двух подряд идущих пар оснований (см. Рис. 1.2).



**Рис.1.2**. Конформационные параметры ДНК, используемые при анализе структурных свойств ДНК в приближении жестких пар оснований.

Элементарным звеном для анализа в таком приближении является динуклеотидный блок. Учитывая симметрию ДНК, возможны 10 различных типов динуклеотидных блоков:

AA (TT), AC (GT), AG (CT), AT, CA (TG), CC (GG), CG, GA (TC), GC, TA,

В скобках указаны комплементарные динуклеотиды в противоположной цепи ДНК, то есть в данном случае блок АА и блок ТТ в виду симметрии представляют один и тот же динуклеотидный блок двойной спирали ДНК.

Анализируя данные о конформации различных олигонуклеотидов, для каждого динуклеотидного блока находят средние значения структурных параметров, а также значения констант жесткости, которые в гармоническом приближении описывают энергетические аспекты деформации динуклеотидного блока [Fujii et al, 2007; Olson et al, 1998]. Полученные значения используют для вычисления энергетических и энтропийных барьеров деформации ДНК в комплексах с различными белками с целью

описания наблюдаемой специфичности связывания [Becker et al, 2006]. В работе [Koudelka, 1998] показано, например, что константа связывания белка-репрессора с операторным участком значимо коррелирует с величиной торсионной подвижности центральных динуклеотидных блоков (то есть с величиной, обратно пропорциональной жесткости по отношению к изменению угла относительного поворота пар оснований (см. Рис. 1.2.)) а также со значениями углов поворота (twist) в данном сайте для свободной ДНК.

Исследование зависимости конформационно-динамических параметров ДНК (то есть параметров, характеризующих как равновесную локальную конформацию ДНК, так и способность к ее изменению) от последовательностей пар оснований, как правило, проводят в рамках ди- и тетрануклеотидных приближений. Показано, что тип ближайших по цепи нуклеотидов оказывает существенное влияние на конформационную динамику динуклеотидных блоков ДНК [Dixit et al, 2005; Fujii et al, 2007]. Стоит отметить, что результаты, полученные различными методами исследования структурных свойств, далеко не всегда согласуются количественно, что приводит к дополнительным сложностям. Тем не менее, существует общая закономерность, выявленная при анализе структурных свойств ДНК различными методами на динуклеотидном уровне описания, которая может быть описана следующим образом: D(PyPu) > D(PuPu) > D(PuPy),

где под *D* понимается общая подвижность динуклеотидного блока (величина, пропорциональная объему конформационного пространства), Ру – пиримидин, то есть нуклеотид С или Т, а Ри – пурин – то есть нуклеотид А или G [Fujii et al, 2007]. Последовательность РуРи отвечает динуклеотидному блоку ДНК, в котором в направлении от 5' к 3' концу цепи за пиримидином следует пурин.

Важно отметить, что существующих данных рентгеноструктурного анализа и ЯМР не достаточно для исследования эффектов последовательности ДНК на конформационную динамику на тетрануклеотидном уровне, поэтому для этой цели, как правило, используются результаты компьютерного моделирования [Dixit et al, 2005; Fujii et al, 2007; Packer et al, 2000].

Во многих случаях связывания белка с ДНК имеет место сильный изгиб молекулы ДНК. На рисунке 1.3 приведена структура комплекса ДНК с ТАТА – связывающим белком ( TATA-binding protein, TBP ). ТВР специфически «узнает»

ТАТА- участок ДНК в промоторе и входит в состав белкового комплекса, инициирующего транскрипцию.



**Рис.1.3.** Структура комплекса ДНК с ТАТА – связывающим белком, полученная методами рентгеноструктурного анализа. Фиолетовым цветом изображен белок, а остальными цветами – молекула ДНК (PDB code: 1cdw).

Связывание ТВР с ДНК приводит к изгибу оси двойной спирали на угол, составляющий около 80°. Считается, что конформационные изменения, происходящие в ДНК в результате образования данного комплекса, в конечном итоге приводят к локальному плавлению ДНК – то есть расхождению цепей сахарофосфатного остова – образованию так называемого транскрипционного пузыря, необходимого для работы РНК- полимеразы, отвечающей за процесс транскрипции [Schramm et al, 2002].

Важно отметить, что так называемые ТАТА – содержащие промоторы присутствуют только в 10-20 % процентов генов эукариот, а для промоторов, не содержащих ТАТА – последовательность, особенности предшествующих транскрипции взаимодействий ДНК с белками остаются плохо изученными [Schramm et al, 2002].

Одной из наиболее широко известных экспериментальных моделей для изучения изгибной жесткости ДНК является взаимодействие ДНК с ферментом ДНКзой I, осуществляющим расщепление сахарофосфатного остова ДНК. Показано, что в зависимости от нуклеотидного состава, скорости реакции расщепления ДНК

данным ферментом могут изменяться в десятки раз [Brukner et al, 1990]. Так как для расщепления ДНК ферменту необходимо изогнуть ДНК в сторону широкой «бороздки», результаты специфичности расщепления ДНК используются для описания зависимости локальной гибкости молекулы, контекстной которая считается пропорциональной логарифму скорости реакции расщепления. На основании результатов по расшеплению ДНК данным ферментом были получены эффективные параметры, характеризующие специфичность расщепления к нуклеотидной последовательности ДНК на уровне ди- и тринуклеотидного приближения [Brukner et al, 1990; Brukner et al 1995].

Известно, что для большинства полученных структур ДНК-белковых комплексов имеют место плотные контакты белка с широкой или узкой «бороздкой» ДНК. Так, например, в случае взаимодействия одного из белков семейства НОХ (семейство белков, имеющих ДНК-связывающий домен, и участвующих в регуляции транскрипции на эмбриональной стадии развития организма) с ДНК имеет место связывание неструктурированной части полипептидной цепи с узкой бороздкой ДНК за счет эффекта фокусировки электростатического потенциала, имеющего место при сужении бороздки [Rohs et al, 2009].

Изменение параметров узкой бороздки ДНК в растворе исследуют при помощи анализа картин химического расщепления ДНК различными агентами – в том числе -ОН – радикалами. Считается, что эффективность расщепления ДНК определяется параметрами узкой бороздки ДНК – чем уже бороздка, тем более «труднодоступным» является сахарный цикл по отношению к взаимодействию с ОН-радикалами [Greenbaum et al, 2007]. На основе данных по специфичности расщепления ДНК ОНрадикалами были получены величины, характеризующие зависимость расщепления от нуклеотидной последовательности ДНК на уровне ди, три и тетрануклеотидного приближений [Greenbaum et al, 2007]. При помощи полученных данных было продемонстрировано, что функциональные области ДНК [Parker et al, 2009].

При анализе особенностей промоторных областей генома человека также использовались упомянутые данные по специфичности расщепления ДНК ДНКзой I [Pedersen et al, 1998]. На Рис. 1.4 приведен профиль, характеризующий среднее

изменение локальной гибкости ДНК вдоль промоторных областей в геноме человека. В этой работе были отобраны 624 нуклеотидные последовательности из промоторных областей генов, обладающие низким уровнем сходства. Последовательности были выравнены, после чего для каждой из них строился теоретический профиль скорости расщепления ДНК ДНКазой I [Brukner et al, 1995]. Зависимость, представленная на Рис.1.4 624 была получена усреднением профилей. log(k/k)(a) -0.012 -0.014 -0.016 -0.018 -0.020 -0.022

**Рис.1.4.** Изменение вдоль ДНК логарифма отношения теоретической величины скорости расщепления ДНК под действием ДНКзы I - k - к максимально возможному значению скорости расщепления k<sub>o</sub>. График получен усреднением 624 профилей и характеризует среднее теоретическое изменение гибкости вдоль промоторных областей ДНК человека. Нумерация ведется относительно точки старта транскрипции i=+1.

-200

-150

-100

-50

+1

50

100

150

200

l

Как видно из рисунка, область до точки старта транкрипции характеризуется более низкими значениями локальной гибкости ДНК, в то время как область после старта транскрипции – более высокими значениями гибкости. Пик, наблюдаемый в районе -30 нуклеотида отвечает резкому увеличению гибкости, имеющему место в ТАТА – содержащих промоторах. Периодическое изменение гибкости в области после точки старта транскрипции связывают с возможностью образования нуклеосомы в данной области ДНК, в то время как низкое значение гибкости в области до начала транскрипции интерпретируют таким образом, что в таком случае уменьшается вероятность образования комплекса с ДНК с гистонами – процесса, нежелательного с точки зрения возможности регуляции экспрессии гена, так как данная область ДНК содержит сайты связывания с регуляторными белками [Сао et al, 2008].

Данный пример приведен с целью демонстрации использования данных по контекстной зависимости структурных свойств ДНК для анализа функциональных участков и выявления критериев поиска таких областей в геномах. В настоящее время для решения этой задачи используются самые разные характеристики ДНК: энергия стэкинг-взаимодействия [Abeel et al, 2008], температура плавления ДНК [Dineen et al, 2009], различные конформационно-динамические параметры [Cao et al, 2008; Goni et al 2007].

На основе явления специфичности ультразвукового расщепления ДНК к нуклеотидной последовательности [Гроховский, 2006] в настоящее время развивается новый подход к изучению контекстно-зависимых свойств ДНК [Grokhovsky et al, 2011]. Как будет показано в данной работе, наблюдаемая специфичность, по всей видимости, связана с особенностями конформационно-динамических свойств сахарофосфатного остова ДНК, проявляющимися в процессе растяжения фрагментов при их движении в высокоградиентном потоке жидкости вблизи кавитационных пузырьков.

Используемые в диссертации модели описывают процесс растяжения молекулы ДНК под действием внешних сил, поэтому, в качестве введения к используемым подходам в следующем разделе приводятся некоторые результаты экспериментальных и теоретических исследований механических свойств отдельных молекул ДНК.

# 1.2. Физика упругих растяжений ДНК

#### Методы исследования упругих свойств молекулы ДНК

Упругие свойства двунитевой молекулы ДНК исследуются различными методами, например, с помощью магнитных бусин [Smith et al, 1992], стеклянных игл [Cluzel et al, 1996], оптических ловушек [Smith et al, 1996; Wang et al, 1997] и атомно - силовой микроскопии [Binnig et al, 1986; Hansma, 1997]. Магнитные бусины прикрепляют к концам молекул ДНК. Внешнее магнитное поле действует на бусины, вызывая натяжение молекул. Использование такого «магнитного пинцета» позволяет достигать растягивающих напряжений в диапазоне 0.01 - 10 пкН (пиконьютонов). Для

получения нагрузок в интервале 0.1 - 100 пкН можно использовать методику «оптического пинцета» [Ashkin et al, 1986; Svoboda et al, 1994]. Принцип работы пинцета заключается в том, что к исследуемому образцу прикрепляется бусинка, на которую фокусируется лазерный луч. Изменяя мощность лазерного пучка, можно регулировать силу светового давления на бусину, а, следовательно, и натяжение исследуемой молекулы. Использование атомно-силовой микроскопии позволяет достичь натяжения молекул в диапазоне 10 - 10000 пкН.

Эксперименты по исследованию упругости двунитевой молекулы ДНК показали, что каждому диапазону сил соответствует своя природа и свой коэффициент упругих растяжений. Различают, по крайней мере, четыре различных типа поведения двунитевой молекулы ДНК в зависимости от величины силы натяжения.

#### Режимы упругих растяжений молекулы ДНК

#### Энтропийная жесткость

Вследствие тепловых флуктуаций двунитевая молекула ДНК в растворе постоянно находится в «искривленном» состоянии. Это приводит к тому, что расстояние между концами молекулы ДНК меньше ее общей длины. Будем говорить, что макросостояние молекулы ДНК характеризуется расстоянием между концами молекулы, а микросостояние - тем, как «искривлена» молекула ДНК. Растяжение молекулы ДНК приводит к увеличению расстояния между концами молекулы, что уменьшает количество микросостояний, соответствующих данному макросостоянию.

В этом процессе происходит уменьшение энтропии системы за счет работы внешней растягивающей силы. Таким образом, молекула ДНК оказывает сопротивление растяжению, что в терминах механики можно описать как проявление жесткости молекулы. Такую жесткость молекулы ДНК называют энтропийной жесткостью [Bustamante et al, 1994].

Для теоретического описания энтропийной жесткости ДНК используются две модели. В модели «свободно сочлененной цепи» полагается, что молекула ДНК состоит из независимо ориентированных сегментов, длина *b* которых определяется жесткостью молекулы. В модели «червеобразной цепи» молекула ДНК считается сгибаемым стержнем длины *L*, слабо искривленным вследствие тепловых флуктуаций. Для такого стержня можно ввести понятие персистентной длины *P*, т.е. расстояния, на

котором наблюдается корреляция между начальным и конечным сегментами стержня. Определение персистентной длины выглядит следующим образом:

$$\langle (\vec{n}, \vec{n'}) \rangle = e^{-s/P}$$
(1.3),

где:  $\vec{n}$  и  $\vec{n'}$  - векторы нормали к двум элементам стержня, находящимся на расстоянии *s* друг от друга, а  $\langle (\vec{n}, \vec{n'}) \rangle$  среднее значение их скалярного произведения. Чем жестче молекула, тем больше величина *P*. Персистентная длина двунитевой молекулы ДНК в водно-солевом растворе составляет примерно 50 нм, т.е. около 150 пар оснований.

В работе [Bustamante et al, 1994] приведена приближенная формула расчета удлинения х в зависимости от длины молекулы L и действующей на нее силы F, полученная на основе червеобразной модели ДНК:

$$\frac{FP}{k_B T} = \frac{1}{4\left(1 - \frac{x}{L}\right)^2} + \frac{x}{L} - \frac{1}{4}$$
(1.4)

где:  $k_B$  – константа Больцмана, T – температура. При малой величине отношения x/L формула (1.4) переходит в закон Гука:

$$F = \frac{3k_B T}{2P} \frac{x}{L}$$
(1.5),

с коэффициентом жесткости  $k_{ДHK} = 3k_BT/2PL$ , обратно пропорциональным общей длине и персистентной длине молекулы. Для двунитевой молекулы ДНК длиной 10 мкм коэффициент жесткости, рассчитанный в соответствии с (1.5), равен приблизительно

 $10^{-5}$  пкН/нм. Такое же выражение для коэффициента жесткости можно получить и в модели свободно сочлененной цепи [Grosberg et al, 1994], приняв размер сегмента равным b = 2P.

Эксперименты по исследованию упругости молекулы ДНК показали границы применимости этих двух моделей. Оказалось, что модель свободно сочлененной цепи хорошо описывает поведение молекулы при нагрузках до 0.1 пкН. Важно отметить, что

модель червеобразной цепи может быть применена в более широком диапазоне сил 0.01 -10 пкН. Сравнение экспериментальных данных с теоретическими кривыми приводится на Рис. 1.5.

#### Собственная эластичность

Выше упоминалось, что предложенные модели не могут описать поведение молекулы ДНК при нагрузках выше 10 пкН. Действительно, при увеличении нагрузок расстояние между концами молекулы ДНК данной длины становится больше расстояния между концами молекулы ДНК в В – форме. При таких нагрузках структура молекулы ДНК меняется, и основной вклад в жесткость молекулы уже не является энтропийным. Эксперименты, проделанные оптическим пинцетом, демонстрируют наличие упругих растяжений в интервале нагрузок 5 - 50 пкН.

Если в модели «червеобразной цепи» учитывать внутренние упругие свойства стержня, то в первом приближении, полагая, что общая длина молекулы ДНК возрастает линейно с ростом приложенной силы, получаем:

$$x = L\frac{F}{S} \tag{1.6}$$

где: *S* – модуль растяжения стержня, связанный с его персистентной длиной *P* соотношением:  $P = \frac{Sr^2}{4k_{_R}T}$  где *r* - радиус стержня.

Оптимальное согласие формулы (1.6) с опытами, проделанными с помощью оптического пинцета [Smith et al, 1996; Wang et al, 1997] в растворе 150 мМ Na+ получается при *S* равным примерно 1000 пкH.

#### В-Ѕ переход

Экспериментальные данные, представленные на Рис. 1.6, показали, что когда молекула ДНК испытывает нагрузки около 65 пкН или больше, она резко меняет свою структуру, растягиваясь на 70% по сравнению с канонической В-формой [Cluzel et al, 1996; Smith et al, 1996], то есть происходит переход в так называемую S-форму,

различные модели которой ждут своего экспериментального подтверждения. Переход из В в S форму происходит кооперативно и в очень узком диапазоне сил. При дальнейшем увеличении нагрузок поведение S–ДНК становится таким же, как у двух однонитевых молекул ДНК. Таким образом, S–ДНК разделяется на две нити, т.е. происходит ее плавление.



**Рис. 1.5.** Экспериментальные и теоретические зависимости растяжения x/L двунитевой молекулы ДНК от приложенной силы F [Bustamante et al, 1994].

(xx) - экспериментальные значения; зеленая и синяя кривые «—» и «—» отвечают теоретически рассчитанным зависимостям по модели червеобразной цепи при <math>P = 53 нм (синяя кривая соответствует точно рассчитанной зависимости, а зеленая – экстраполяции в соответствии с формулой (1.4)); черная кривая «—» соответствует зависимости, полученной в рамках модели свободно сочлененной цепи при b = 2P = 106 нм, а коричневая кривая «—» соответствует закону Гука (1.5).



**Рис.** 1.6. Экспериментальная зависимость растяжения x/L двунитевой и однонитевой молекул ДНК от приложенной внешней силы F [Bustamante et al, 1994].

«—•—» - растяжение двунитевой молекулы ДНК, «– –» - расчетная кривая по уравнению (1.6), «▲». «■», «•»– растяжение однонитевой молекулы ДНК в 2 мМ Na+, 5 мМ Mg2+ и 150 мМ Na+ соответственно.

## Разрыв ковалентных связей в молекуле ДНК

Какое натяжение необходимо, чтобы разорвать ковалентные связи в молекуле ДНК? Теоретические оценки показывают, что данная сила должна превышать 5000 пкН. Однако в экспериментах в движущемся потоке разрыв молекул ДНК происходил уже при 100-300 пкН. Единичные двунитевые молекулы ДНК, растянутые водным мениском, разрывались при 960 пкН [Bensimon et al, 1995]. Короткие двунитевые молекулы ДНК в атомно-силовой микроскопии [Lee et al, 1994] выдерживали натяжение более 1700 пкН. Весьма сложно установить истинное значение силы, необходимой для разрыва двунитевой молекулы ДНК, т.к. она оказывается зависящей от длины молекулы и свойств растворителя. У полисахаридных молекул в водном растворе сила, необходимая для их разрыва, была определена при помощи атомносиловой микроскопии и составляет около 1000 пкН [Grandbois et al, 1999].

#### 1.3. Расщепление полимеров под действием ультразвука.

Явление ультразвукового расщепления полимерных соединений в растворе активно используется в самых разных технологических операциях. Важной особенностью данного процесса является то, что крупные молекулы расщепляются намного эффективнее, чем короткие, позволяя существенным образом изменять распределение компонентов раствора по степени полимеризации.

Характерные особенности процесса расщепления полимеров под действием ультразвука высокой интенсивности свидетельствуют о том, что данное явление относится к классу так называемых механохимических реакций – то есть реакций, катализируемых действием механических сил [Basedow et al, 1977]. Механохимическая реакция является сложным многостадийным процессом, который включает механическую деформацию молекулы, предшествующую химической реакции. В случае ультразвукового расщепления полимеров в растворе, деформация молекулы возникает в результате действия гидродинамических сил. Силы, приводящие к расщеплению полимеров под действием ультразвука высокой интенсивности, связаны с кавитационными процессами, которые происходят в озвучиваемом растворе [Basedow et al, 1977].

В соответствии с теорией механохимических реакции [Бутягин, 1971], существует два прямых механизма перехода механической энергии в энергию, необходимую для протекания химической реакции. Первый механизм связан с механической деформацией химической связи – изменением межатомных расстояний и деформацией электронных облаков под действием внешней силы, а второй – с возбуждением колебательных степеней свободы связи при локальном разогреве молекулы, связанным с переходом механической энергии деформации в тепло.

Первый механизм можно считать равновесным процессом, если характерное время действия внешней силы намного превышает время релаксации к распределению Максвелла-Больцмана. Второй механизм связан с существенно неравновесным процессом и применим к описанию реакций в случае действия импульсных сил – например, при действии ударных волн. Кинетика механохимической реакции в первом случае может быть описана в рамках теории активированного комплекса – или на основе модифицированного уравнения Аррениуса [Schmidt et al, 2008], что будет использовано в третьей главе диссертации.

#### Акустическая кавитация

Под явлением акустической кавитацией понимается совокупность процессов, связанных с образованием в озвучиваемом растворе так называемых кавитационных пузырьков. Образование пузырьков является результатом разрыва жидкости под действием отрицательных давлений ультразвука: в фазе разряжения звуковой волны на имеющихся в жидкости микропузырьках образуется разрыв в виде полости, которая заполняется насыщенным паром и диффундирующим в нее растворенным газом. В фазе сжатия пар конденсируется, а имеющийся в полости газ подвергается сильному сжатию. В момент «схлопывания» давление и температура газа достигают высоких значений, что порождает в близкой окрестности пузырька импульс высокого давления. Акустическая кавитация представляет собой эффективный механизм концентрации энергии. При кавитации относительно низкая средняя плотность энергии звукового поля трансформируется в высокую плотность энергии в малом объеме внутри и вблизи схлопывающегося пузырька. Полная энергия захлопывающегося пузырька невелика, однако сферическая сходимость пузырька приводит к образованию очень больших локальных плотностей энергии, а, следовательно, высоких температур и давлений [Brennen, 1995].

#### Модель расщепления полимеров под действием акустической кавитации.

Экспериментальные данные, а также теоретические расчеты свидетельствуют о том, что при определенных условиях схлопывание кавитационного пузырька может вызывать высокоградиентные течения в окружающей жидкости, достаточные для разрыва полимерных молекул [Эльпинер, 1973; Basedow et al, 1977].

В настоящее время общепринятой является модель расщепления, предложенная Томасом [Thomas, 1959]: при схлопывании кавитационного пузырька возникает сильный радиальный градиент скорости течения жидкости, который приводит к растяжению полимера и, в конечном итоге, к его расщеплению (см. Рис. 1.7).



**Рис.1.7.** Иллюстрация к модели расщепления полимеров под действием акустической кавитации. При схлопывании кавитационного пузырька радиуса R жидкость устремляется к центру пузырька. Скорость движения жидкости имеет максимальное значение у поверхности пузырька U=dR/dt и уменьшается с ростом расстояния от поверхности пузырька. Наличие градиента скорости течения жидкости приводит к тому, что скорости течения на противоположных концах полимера различны. Это приводит к возникновению растягивающего усилия F в полимере, имеющего максимальное значение в центре молекулы. Таким образом, полимер, расположенный на расстоянии r от центра пузырька и ориентированный параллельно скорости течения жидкости, подвержен действию растягивающего усилия, которое может привести к его расщеплению

В третьей главе диссертации на основе данной модели предложен подход к моделированию расщепления ДНК под действием акустической кавитации. Важно отметить, что данная модель позволяет описывать расщепление молекулы ДНК в рамках основного механизма механохимического расщепления – то есть на основе предположения о понижении активационного барьера реакции за счет действия растягивающего усилия. Такой подход позволяет описать позиционный эффект расщепления ДНК ультразвуком (см. Главу 2), то есть зависимость скорости расщепления фосфодиэфирной связи от ее положения в цепи молекулы [II'icheva et al,

1999]. Экспериментальные профили расщепления ДНК под действием ультразвука демонстрируют значительное увеличение скорости расщепления фрагмента ДНК в центральной части молекулы, что свидетельствует в пользу выбранной модели расщепления, согласно которой сила растяжения имеет максимальное значение в центре фрагмента.

Следует отметить, что наблюдаемое расщепление ДНК под действием ультразвука высокой интенсивности также может быть связано с действием ударных волн, образующихся при схлопывании кавитационных пузырьков. В случае действия ударной волны импульсная механическая деформация молекулы может происходить из-за резкого изменения давления растворителя. Была предложена модель расщепления полимеров под действием ударных волн кавитации [Gooberman, 1960], позволяющая описать позиционный эффект расщепления, наблюдаемый в экспериментах по ультразвуковой деградации макромолекул полистирола. Однако, возможность применения данной модели для описания расщепления коротких фрагментов ДНК – длиной от нескольких десятков до нескольких сотен нуклеотидных пар – требует дополнительного исследования.

В данной работе рассмотрена традиционная модель расщепления, описывающая разрыв ДНК как результат взаимодействия молекулы с высокоградиентным течением жидкости, возникающим вблизи схлопывающегося кавитационного пузырька. Возможность расщепления ДНК под действием ударных волн кавитации требует дополнительных как экспериментальных, так и теоретических исследований.

26

# Глава 2. Ультразвуковое расщепление ДНК: обработка и анализ экспериментальных данных

В данной главе описана экспериментальная процедура расщепления ДНК ультразвуком, методика получения и обработки данных. Приведены характерные свойства наблюдаемого расщепления, а также результаты статистического анализа специфичности ультразвукового расщепления ДНК.

#### 2.1. Описание эксперимента и процедуры обработки экспериментальных данных.

Фрагменты ДНК получали расщеплением  $\lambda$  - фага и плазмид pBR322, pUC18 и pGEM7(f+) (Promega), содержавших в полилинкерах различные вставки, соответствующими рестриктазами. Для введения радиоактивной метки в 3'-конец фрагментов использовали [ $\alpha$ -<sup>33</sup>P]dATP или [ $\alpha$  -<sup>33</sup>P]dCTP (" $\Phi$ ГУП" Институт реакторных материалов, Заречный, Свердловская обл.), остальные немеченые dNTP и фрагмент Кленова ДНК-полимеразы I Escherichia coli (Boehringer Mannheim, Германия). Выделение фрагментов ДНК проводили в 5%-ном полиакриламидном геле толщиной 1мм, с последующей элюцией и осаждением. Последовательности фрагментов и другие материалы, касающиеся эксперимента, приведены на сайте <u>http://grok.imb.ac.ru\en</u>

## Облучение растворов фрагментов ДНК ультразвуком

Для приготовления образцов 10 мкл раствора фрагмента ДНК (примерно 10<sup>4</sup> Бк) в воде смешивались с 10 мкл 0.2 М NaOAc, pH 6.0 в тонкостенных полипропиленовых пробирках на 0.2 мл (N801-0540, Perkin-Elmer, USA). Конечная концентрация фрагмента составляла 5 - 10 мкг/мл или ~10 мкМ п.н.

Облучение проводилось на ультразвуковом диспергаторе УЗДН-2Т (Украина) при частоте 22 кГц. Фотография и схема установки приведена на Рис.2.1. Пробирки помещали в тефлоновое кольцо с центральным отверстием 15 мм и с радиальными отверстиями для пробирок так, чтобы концы пробирок, в которых находились тестируемые растворы, находились на расстоянии около 0.5 см от поверхности торца излучателя, диаметр которого составлял 12 мм. Кольцо и излучатель помещались в баню с водой и мелкоразмолотым льдом, которая вращалась со скоростью два оборота в минуту. Мощность ультразвука определялась калориметрически и составила более 70 Ватт.



Рис.2.1. Схема установки для облучения фрагментов ДНК ультразвуком

Для контроля кавитации использовалась тестовая пробирка, содержащая 20 мкл 0.05М КЈ в 0.025%-ном растворе крахмала. После облучения в течении 8 мин степень окрашивания раствора была эквивалентна добавлению примерно 10<sup>-4</sup> М перекиси водорода. Для контроля меньших доз экспозиции использовался аналогичный тест, содержащий раствор крахмала, насыщенный CCl<sub>4</sub> при 20°C. Следует отметить, что наблюдаемый выход перекиси водорода соответствует известным из литературы результатам, полученным при нормальной температуре и интенсивности ультразвука 2 Ватт/см<sup>2</sup> [Маргулис, 1984]. Известно, что низкая частота ультразвука и низкая температура раствора приводят к усилению кавитационных эффектов [Basedow et al, 1977]

# Разделение фрагментов в денатурирующем геле

После облучения к смеси добавляли 90 мкл раствора 0.15 M NaCl, 50 мМ Трис-HCl (pH 7.5), 10 мМ EDTA, 10 мкг/мл тРНК. Смесь экстрагировали фенолом, ДНК осаждали этанолом, промывали 70%-ным этанолом, высушивали, растворяли в 1 мкл 95%-ного формамида, содержащего 15 мМ EDTA (pH 8.0), 0.05% бромфенолового синего и 0.05% ксиленцианола FF, нагревали 1 мин при 90°C, быстро охлаждали до 0°C и наносили на денатурирующий ПААГ длиной 40 см с градиентной толщиной 0.15-0.45 мм. Электрофорез проводили 55 мин при 100 Вт (2.5 кВ) при температуре 60-70°C. Перед экспонированием гель фиксировали в 10%-ной уксусной кислоте и высушивали на стекле, предварительно обработанном гамма-метакрилпропилоксисиланом (LKB, Швеция) и экспонировали с люминесцентным экраном с последующим сканированием на приборе "Cyclone Storage Phosphor System" (Packard BioScience Company, USA).

#### Анализ гелей

Для анализа гелей, типичная картина которых показана на Рис.2.1, использовалась компьютерная программа SAFA, разработанная группой из Стендфорского университета [Das et al, 2005]. Эта программа позволяет после корректировки треков дорожек геля вычислять интенсивности всех полос на всех дорожках, и соотносить эти полосы с заданной последовательностью нуклеотидов при помощи дорожки, содержащей результаты химического расщепления известной последовательности по пуринам. Анализ ряда полученных гелей, проведённый с помощью программы OptiQuant (Packard Instrument Company), показывал сходные результаты.

Значение интенсивности каждой полосы на геле пропорционально концентрации фрагмента ДНК определенной длины и, поэтому, соответствует скорости расщепления определенной фосфодиэфирной связи начального фрагмента. Степень расщепления ДНК – то есть доля расщепленных фрагментов - повышается при увеличении времени облучения фрагмента ультразвуком. Также следует отметить наблюдаемый эффект увеличения скорости расщепления фрагментов ДНК в их центральной части – так называемый позиционный эффект.

На абсолютную величину интенсивности полосы на геле влияет множество факторов, которые могут меняться от эксперимента к эксперименту, поэтому, при анализе экспериментальных данных следует оперировать относительными значениями интенсивностей полос, а не абсолютными значениями.

В дальнейшем, абсолютная величина интенсивности полосы на геле будет называться интенсивностью расщепления *I*, а относительная велична – относительной частотой расщепления *R*.

Для вычисления относительных частот расщепления, значение интенсивности каждой полосы делилось на среднее значение интенсивностей определенного количества соседних полос:

$$R_n = (2m+1)\frac{I_n}{\sum_{k=n-m}^{n+m}I_k},$$

где n, k – номера полос, m – параметр, характеризующий величину «окна», по которому проводится усреднение. Варьирование величины такого «окна» показало, что оптимальным является величина в 31 нуклеотид (т.е. в 31 полосу) - уменьшение этой величины приводит к увеличению разброса данных, а дальнейшее увеличение практически не сказывается на соотношении полученных значений, но приводит к уменьшению числа анализируемых значений на величину «окна» для каждого геля. Эта операция в дальнейшем будет называться процедурой нормализацией данных расщепления. Описанная процедура нивелирует эффекты уменьшения интенсивности расщепления к концам фрагмента и различия в интенсивностях дорожек на разных гелях поэтому, используется для статистического анализа влияния И, последовательности пар оснований на скорость ультразвукового расщепления ДНК.

#### 2.2. Характерные свойства ультразвукового расщепления ДНК

Для того чтобы продемонстрировать характерные картины ультразвукового расщепления, полученные при помощи электрофореза в полиакриламидном геле, далее приведен Рис.2.2, представляющий изображение геля, полученного в результате ультразвукового облучения фрагментов ДНК различной длины. Гель разбит на три части, в каждой из которых дорожки соответствуют расщеплению фрагментов ДНК определенной длины. Слева приведены дорожки, отвечающие расщеплению самого длинного фрагмента ДНК – длиной 311 пар оснований (дорожки 1-6), в центре – дорожки 7-12 – расщепление фрагмента средней длины (251 п.о.) и справа – данные по расщеплению короткого фрагмента (216 пар оснований) – дорожки 13-17. Последовательность нуклеотидов короткого фрагмента также содержится в среднем и длинном фрагментах, а последовательность среднего фрагмента – в длинном фрагменте. Видно, что при увеличении длительности облучения ультразвуком от двух до 16 минут для всех фрагментов наблюдается значительное увеличение общего уровня расщепления ДНК.

1 2 3 4 5 6



 Рис.2.2
 Изображение
 геля,

 полученного
 в
 результате

 электрофореза
 облученных

 ультразвуком
 меченых
 фрагментов

 ДНК различной начальной длины.
 Соблика
 Соблика

Дорожки 2 - 6 соответствуют облучению двунитевых фрагментов ДНК длиной 311 пар оснований при различных временах облучения. 7-12 Дорожки соответствуют облучению фрагментов длиной в 251 пару оснований, а дорожки 13-17 облучению фрагментов длиной 216 нуклеотидных пар. Дорожки 1 и 18 соответствуют химическому расщеплению фрагментов ДНК по пуринам.

Каждая полоса на геле соответствует фрагментам ДНК определенной длины, образованным в результате разрыва первоначального фрагмента по соответствующей фосфодиэфирной связи. Интенсивность засветки изображения полосы на экране, то есть степень ее «почернения», считается пропорциональной концентрации соответствующих фрагментов.

Из рисунка видно, что добавление тиомочевины не оказывает значимого влияния на картину расщепления ( см. Дорожки 11 и 12). Также не оказывает видимого влияния добавление других скевенджеров активных радикалов (то есть соединений, блокирующих действие активных радикалов) - аскорбата натрия и дитиотреитола (данные не приведены). Напротив, добавление 50% глицерина, приводящее к увеличению вязкости раствора, приводит к резкому увеличению уровня расщепления (дорожки 3,8 и 14). Дорожки, соответствующие расщеплению ДНК при добавлении глицерина, выглядят также, как дорожки без глицерина, соответствующие более длительному времени облучения. Таким образом, увеличение вязкости раствора приводит к увеличению общей степени расщепления ДНК, оставляя неизменными относительные частоты расщепления. Этот эффект является одним из характерных признаков механохимических реакций.

Рисунок 2.2 также демонстрирует позиционный эффект расщепления – то есть ослабление степени расщепления ДНК к концам фрагментов. Эта важная особенность ультразвукового расщепления ДНК также является характерным признаком механохимической природы наблюдаемого расщепления. Следует отметить, что существенная величина позиционного эффекта, а также малое влияние добавления скевенжеров активных радикалов на наблюдаемые картины расщепления, позволяют заключить, что расщепление ДНК, связанное с действием активных радикалов, в описанных экспериментах пренебрежимо мало.

На Рисунке 2.3 представлены результаты компьютерной обработки геля, изображенного на Рис.2.2. Здесь приводятся профили интенсивностей расщепления *I*, полученные для нескольких полос геля. Профили относительных частот расщепления *R*, расчитанные для двух полос, демонстрируют результат нормализации профилей интенсивности. Величины относительных частот расщепления, а также последовательность пар оснований фрагмента представляют входные данные для статистического анализа специфичности ультразвукового расщепления ДНК к нуклеотидной последовательности.



**Рис.2.3.** Профили ультразвукового расщепления, полученные в результате обработки указанных дорожек на геле, приведенном на Рис 2.2. Высота столбца пропорциональна концентрации соответствующих ему фрагментов ДНК, а его цвет – типу нуклеотида, за которым происходит разрыв (в направлении от 5' к 3' концу цепи, содержащей метку). Нижние диаграммы получены путем нормализации начальных данных с целью устранения позиционного эффекта, наблюдаемого на первых трех диаграммах.

# 2.3. Статистический анализ специфичности расщепления ДНК ультразвуком

Были проанализированы картины расщепления 48 различных радио-меченных фрагментов ДНК, имеющих длины от 150 до 500 пар оснований. Для статистического анализа использовались центральные части гелей, в которых полосы были четко разделены. Так как эксперименты по облучению фрагментов одной и той же последовательности демонстрировали некоторый разброс данных, они проводились несколько раз. Следует отметить, что максимальные и минимальные значения интенсивностей ультразвукового расщепления (то есть точки, соответствующие границам распределения полос по интенсивностям) не учитывались при статистическом анализе. Общее число таких выброшенных значений составило около 4% от полного числа точек. Группа максимальных значений, отброшенная при анализе, является следствием дефектов на геле или присутствием чужеродных фрагментов (пример такого дефекта виден в левом нижнем углу Рис.2.2). Вторая группа, - группа с минимальными значениями интенсивности расщепления является следствем неверной апроксимации суммарной интенсивности полосы из-за ее искривления или перекрывания с соседней полосой (такие перекрывания также можно увидеть в верхней части дорожек на Рис.2.2).

Общее число проанализированных полос составило около 20500. Для анализа зависимости относительной частоты разрыва фосфодиэфирной связи от локальной нуклеотидной последовательности использовались ди- и тетрануклеотидное приближения. Для исследования влияния типа ди- и тетрануклеотида, в котором происходит расщепление, на величину относительной частоты расщепления применялся однофакторный дисперсионный анализ, непараметрические критерии (Kruskal-Wallis test, Brown-Mood test). Результаты анализа позволяют сделать вывод о том, что влияние типа динуклеотида на относительную частоту его ультразвукового расщепления статистически значимо на уровне  $p << \alpha = 0.05$ . В Таблице 1 и на Рис.2.4 приведены значения выборочных характеристик и 95% доверительные интервалы для средних значений относительных частот расщепления.

Как видно из Таблицы 1 и Рисунка 2.4, значения средних относительных частот ультразвукового расщепления комплементарных динуклеотидов различаются значимо. Это указывают на независимость расщепления комплементарных цепей. Обозначения:

N – объём выборки;

 $\overline{R}$  – средняя величина;

*S* – стандартное отклонение;

	Ν	$\overline{R}$	S
AA	1636	0.919	0.129
AC	1076	0.913	0.128
AG	1028	0.900	0.124
AT	1374	0.904	0.119
СА	1265	1.160	0.209
CC	1141	1.007	0.144
CG	1230	1.444	0.334
СТ	1077	1.130	0.198
GA	1153	0.970	0.133
GC	1317	0.954	0.146
GG	1168	0.922	0.145
GT	1101	0.952	0.126
ТА	1065	0.973	0.120
TC	1173	0.912	0.131
TG	1305	0.979	0.126
TT	1672	0.932	0.127

Для того, чтобы выяснить для каких динуклеотидов их интенсивности ультразвукового расщепления значимо не равны, и для какого динуклеотида относительную частоту расщепления можно считать максимальной (минимальной), использовались параметрические (Tukey-Kramer, GT2-, T'-) и непараметрический (Kruskal-Wallis test) методы множественного сравнения. Результаты проведенного множественного сравнения показывают, что средние значения ультразвукового расщепления в динуклеотидах СС, СТ, СА, СG статистически значимо отличаются друг от друга и от средних значений для всех других динуклеотидов. Выборочное среднее для CG является максимальным значением относительной частоты ультразвукового расщепления динуклеотидов.

Исследование влияния типа динуклеотида во втором положении (3'-концевого) на относительную частоту ультразвукового расщепления в каждой из четырех групп динуклеотидов, различающихся по первому (5'-концевому) нуклеотиду, позволяет сделать вывод, что в каждой из четырех групп такое влияние значимо на уровне  $p << \alpha = 0.05$ . Применение методов множественного сравнения для каждой из четырех групп динуклеотидов дает следующие результаты:

a) в группе динуклеотидов AN (AA, AC, AG, AT) нельзя выделить динуклеотид, среднее значение ультразвукового расщепления которого значимо больше или меньше, чем у всех других;

b) в группе динуклеотидов CN (CA, CC, CG, CT) все средние значения различаются значимо, причем среднее в группе CC минимально, а среднее в CG – максимально;

с) в группе динуклеотидов GN (GA, GC, GG, GT) среднее в GG минимально, а среднее в GA – максимально;

d) в группе динуклеотидов TN (TA, TC, TG, TT) среднее в TC минимально.

Необходимо отметить, что результаты, полученные при помощи параметрических методов анализа, подтверждаются и непараметрическими методами.

36


**Рис.2.4.** Выборочные средние относительных частот расщепления и их 95% доверительные интервалы. Обозначения: ∘, ▲, ■, ◊ - среднее, I - 95% доверительный интервал.

Анализ влияния фланкирующих нуклеотидов N<sub>1</sub> (5'-концевого) и N<sub>4</sub> (3'-концевого) (где N – любой из нуклеотидов A, C, G, T) для всех тетрануклеотидов N<sub>1</sub>N<sub>2</sub>'N<sub>3</sub>N<sub>4</sub> на относительную частоту ультразвукового расщепления в их центре показал (положение разрыва условно отмечено штрихом), что это влияние статистически значимо ( $p << \alpha = 0.05$ ) в каждой из 16 групп тетрануклеотидов. Для четырех групп тетрануклеотидов – NCAN, NCGN, NTAN, NTTN – можно выделить статистически достоверные максимальные средние значения относительной частоты разрыва. Они соответствуют тетрануклеотидам GCAG, GCGA, GTAG и GTTA.

На Рис.2.5 и в Таблице 2 приведены значения выборочных характеристик относительных частот разрыва в центральных фосфодиэфирных связях 256 тетрануклеотидов.





Рис.2.5. (a-d) Выборочные средние относительных частот расщепления центральной фосфодиэфирной связи в 256 тетрануклеотидов N<sub>1</sub>N<sub>2</sub>'N<sub>3</sub>N<sub>4</sub> (положение разрыва условно отмечено штрихом) и их 95% доверительные интервалы. Обозначения: о, ▲, ■, ◊ - среднее, I - 95% доверительный интервал.

**Таблица 2.** Выборочные характеристики относительных частот ультразвукового расщепления центральной фосфодиэфирной связи в 256 тетрануклеотидах.

Обозначения:

N – объём выборки;

- $\overline{R}$  средняя величина;
- S стандартное отклонение;
- $S_{\overline{R}}$  стандартное отклонение среднего.

	Ν	$\overline{R}$	S	$S_{\overline{R}}$
AAAA	232	0.894	0.132	0.009
AAAC	151	0.883	0.139	0.011
AAAG	93	0.915	0.107	0.013
AAAT	96	0.951	0.142	0.013
CAAA	114	0.921	0.169	0.012
CAAC	80	0.877	0.118	0.015
CAAG	77	0.938	0.142	0.015
CAAT	77	0.935	0.126	0.015
GAAA	119	0.921	0.113	0.012
GAAC	49	0.887	0.117	0.019
GAAG	88	0.936	0.113	0.014
GAAT	132	0.917	0.127	0.011
TAAA	108	0.952	0.125	0.013
TAAC	73	0.916	0.097	0.015
TAAG	55	0.997	0.155	0.018
TAAT	95	0.929	0.124	0.013

	N	$\overline{R}$	S	$S_{\overline{R}}$
AATA	110	0.925	0.121	0.012
AATC	65	0.913	0.117	0.014
AATG	98	0.913	0.127	0.013
AATT	129	0.875	0.115	0.010
CATA	75	0.910	0.137	0.016
CATC	67	0.801	0.189	0.023
CATG	70	0.879	0.103	0.012
CATT	88	0.863	0.108	0.011
GATA	91	0.907	0.083	0.009
GATC	92	0.950	0.120	0.012
GATG	77	0.929	0.113	0.013
GATT	76	0.913	0.107	0.012
TATA	67	0.913	0.129	0.016
TATC	89	0.858	0.114	0.012
TATG	69	0.946	0.106	0.013
TATT	115	0.914	0.132	0.012

	Ν	$\overline{R}$	S	$S_{\overline{R}}$
AACA	96	0.938	0.112	0.011
AACC	77	0.881	0.131	0.015
AACG	100	0.923	0.128	0.013
AACT	78	0.919	0.127	0.014
CACA	58	0.943	0.151	0.020
CACC	85	0.820	0.130	0.014
CACG	46	0.871	0.174	0.026
CACT	73	0.900	0.116	0.014
GACA	59	0.935	0.166	0.022
GACC	34	0.921	0.117	0.020
GACG	83	0.913	0.156	0.017
GACT	61	0.914	0.129	0.016
TACA	38	0.986	0.094	0.015
TACC	79	0.887	0.113	0.013
TACG	71	0.921	0.101	0.012
TACT	43	0.973	0.114	0.017

	Ν	$\overline{R}$	S	$S_{\overline{R}}$
AAGA	70	0.905	0.098	0.012
AAGC	91	0.845	0.142	0.015
AAGG	72	0.924	0.133	0.016
AAGT	79	0.899	0.108	0.012
CAGA	81	0.935	0.146	0.016
CAGC	100	0.846	0.136	0.014
CAGG	82	0.932	0.128	0.014
CAGT	92	0.897	0.137	0.014
GAGA	50	0.880	0.118	0.017
GAGC	59	0.889	0.109	0.014
GAGG	48	0.914	0.117	0.017
GAGT	50	0.874	0.107	0.015
TAGA	41	0.935	0.082	0.013
TAGC	50	0.907	0.114	0.016
TAGG	40	0.886	0.145	0.023
TAGT	26	0.975	0.197	0.039

**Таблица 2** (продолжение). Выборочные характеристики относительных частот ультразвукового расщепления центральной фосфодиэфирной связи в 256 тетрануклеотидах.

Обозначения:

N – объём выборки;

- $\overline{R}$  средняя величина;
- S стандарнтое отклонение;
- $S_{\overline{R}}$  стандартное отклонение среднего.

	Ν	R	S	$S_{\overline{R}}$
ACAA	57	1.103	0.176	0.023
ACAC	53	1.108	0.177	0.024
ACAG	86	1.183	0.213	0.023
ACAT	67	1.116	0.181	0.022
CCAA	73	1.025	0.178	0.021
CCAC	45	0.989	0.149	0.022
CCAG	96	1.190	0.213	0.022
CCAT	81	1.073	0.168	0.019
GCAA	126	1.223	0.204	0.018
GCAC	71	1.174	0.215	0.026
GCAG	89	1.356	0.263	0.028
GCAT	73	1.186	0.203	0.024
TCAA	93	1.127	0.171	0.018
TCAC	94	1.112	0.144	0.015
TCAG	84	1.231	0.266	0.029
TCAT	80	1.129	0.200	0.022
	N	$\overline{R}$	S	$S_{\overline{R}}$
ACCA	N 78	<u>R</u> 0.975	S 0.112	<i>S</i> <sub><i>R</i></sub> 0.013
ACCA ACCC	N 78 72	R           0.975           0.986	S 0.112 0.147	$S_{\bar{R}}$ 0.013 0.017
ACCA ACCC ACCG	N 78 72 79	R           0.975           0.986           1.064	S 0.112 0.147 0.143	$     S_{\bar{R}} \\     0.013 \\     0.017 \\     0.016 $
ACCA ACCC ACCG ACCT	N 78 72 79 45	$ \bar{R} $ 0.975 0.986 1.064 1.012	S 0.112 0.147 0.143 0.101	$\begin{array}{c} S_{\bar{R}} \\ 0.013 \\ 0.017 \\ 0.016 \\ 0.015 \end{array}$
ACCA ACCC ACCG ACCT CCCA	N 78 72 79 45 45	$     \overline{R}     0.975     0.986     1.064     1.012     0.930 $	S 0.112 0.147 0.143 0.101 0.093	$\begin{array}{c} S_{\bar{R}} \\ 0.013 \\ 0.017 \\ 0.016 \\ 0.015 \\ 0.014 \end{array}$
ACCA ACCC ACCG ACCT CCCA CCCC	N 78 72 79 45 45 107	$     \bar{R}     0.975     0.986     1.064     1.012     0.930     0.912 $	S 0.112 0.147 0.143 0.101 0.093 0.115	$\begin{array}{c} S_{\bar{R}} \\ 0.013 \\ 0.017 \\ 0.016 \\ 0.015 \\ 0.014 \\ 0.011 \end{array}$
ACCA ACCC ACCG ACCT CCCA CCCC CCCG	N 78 72 79 45 45 107 103	$\begin{array}{c} \overline{R} \\ 0.975 \\ 0.986 \\ 1.064 \\ 1.012 \\ 0.930 \\ 0.912 \\ 0.992 \end{array}$	S 0.112 0.147 0.143 0.101 0.093 0.115 0.147	$\begin{array}{c} S_{\bar{R}} \\ 0.013 \\ 0.017 \\ 0.016 \\ 0.015 \\ 0.014 \\ 0.011 \\ 0.014 \end{array}$
ACCA ACCC ACCG ACCT CCCA CCCC CCCG CCCG	N 78 72 79 45 45 107 103 59	$ \overline{R}           0.975           0.986           1.064           1.012           0.930           0.912           0.992           0.941  $	S 0.112 0.147 0.143 0.101 0.093 0.115 0.147 0.138	$\begin{array}{c} S_{\bar{R}} \\ 0.013 \\ 0.017 \\ 0.016 \\ 0.015 \\ 0.014 \\ 0.011 \\ 0.014 \\ 0.018 \end{array}$
ACCA ACCC ACCG ACCT CCCA CCCC CCCG CCCT GCCA	N 78 72 79 45 45 107 103 59 118	$\begin{array}{c} \overline{R} \\ 0.975 \\ 0.986 \\ 1.064 \\ 1.012 \\ 0.930 \\ 0.912 \\ 0.992 \\ 0.941 \\ 1.080 \end{array}$	S 0.112 0.147 0.143 0.101 0.093 0.115 0.147 0.138 0.153	$\begin{array}{c} S_{\bar{R}} \\ 0.013 \\ 0.017 \\ 0.016 \\ 0.015 \\ 0.014 \\ 0.011 \\ 0.014 \\ 0.018 \\ 0.014 \end{array}$
ACCA ACCC ACCG ACCT CCCA CCCC CCCG CCCT GCCA GCCC	N 78 72 79 45 45 107 103 59 118 50	$\begin{array}{c} \overline{R} \\ 0.975 \\ 0.986 \\ 1.064 \\ 1.012 \\ 0.930 \\ 0.912 \\ 0.992 \\ 0.941 \\ 1.080 \\ 1.085 \end{array}$	S           0.112           0.147           0.143           0.101           0.093           0.115           0.147           0.138           0.153           0.223	$\begin{array}{c} S_{\bar{R}} \\ 0.013 \\ 0.017 \\ 0.016 \\ 0.015 \\ 0.014 \\ 0.011 \\ 0.014 \\ 0.018 \\ 0.014 \\ 0.012 \end{array}$
ACCA ACCC ACCG ACCT CCCA CCCC CCCG CCCT GCCA GCCC GCCC	N 78 72 79 45 45 107 103 59 118 50 48	$\begin{array}{c} \overline{R} \\ 0.975 \\ 0.986 \\ 1.064 \\ 1.012 \\ 0.930 \\ 0.912 \\ 0.992 \\ 0.941 \\ 1.080 \\ 1.085 \\ 1.130 \\ \end{array}$	S0.1120.1470.1430.1010.0930.1150.1470.1380.1530.2230.130	$\begin{array}{c} S_{\bar{R}} \\ 0.013 \\ 0.017 \\ 0.016 \\ 0.015 \\ 0.014 \\ 0.011 \\ 0.014 \\ 0.018 \\ 0.014 \\ 0.032 \\ 0.019 \\ \end{array}$
ACCA ACCC ACCG ACCT CCCA CCCC CCCG CCCT GCCA GCCC GCCC	N 78 72 79 45 45 107 103 59 118 50 48 61	$\begin{array}{c} \overline{R} \\ 0.975 \\ 0.986 \\ 1.064 \\ 1.012 \\ 0.930 \\ 0.912 \\ 0.992 \\ 0.941 \\ 1.080 \\ 1.085 \\ 1.130 \\ 1.086 \\ \end{array}$	S         0.112         0.147         0.143         0.101         0.093         0.115         0.147         0.138         0.153         0.223         0.130         0.150	$\begin{array}{c} S_{\bar{R}} \\ 0.013 \\ 0.017 \\ 0.016 \\ 0.015 \\ 0.014 \\ 0.011 \\ 0.014 \\ 0.018 \\ 0.014 \\ 0.012 \\ 0.019 \\ 0.019 \\ 0.019 \end{array}$
ACCA ACCC ACCG ACCT CCCA CCCC CCCG CCCT GCCA GCCC GCCC	N 78 72 79 45 45 107 103 59 118 50 48 61 53	$\begin{array}{c} \overline{R} \\ 0.975 \\ 0.986 \\ 1.064 \\ 1.012 \\ 0.930 \\ 0.912 \\ 0.992 \\ 0.941 \\ 1.080 \\ 1.085 \\ 1.130 \\ 1.086 \\ 0.999 \\ \end{array}$	S0.1120.1470.1430.1010.0930.1150.1470.1380.1530.2230.1300.1500.127	$\begin{array}{c} S_{\bar{R}} \\ 0.013 \\ 0.017 \\ 0.016 \\ 0.015 \\ 0.014 \\ 0.011 \\ 0.014 \\ 0.018 \\ 0.014 \\ 0.032 \\ 0.019 \\ 0.019 \\ 0.018 \\ 0.018 \end{array}$
ACCA ACCC ACCG ACCT CCCA CCCC CCCG CCCT GCCA GCCC GCCC	N 78 72 79 45 45 107 103 59 118 50 48 61 53 88	$\begin{array}{c} \overline{R} \\ 0.975 \\ 0.986 \\ 1.064 \\ 1.012 \\ 0.930 \\ 0.912 \\ 0.992 \\ 0.941 \\ 1.080 \\ 1.085 \\ 1.130 \\ 1.086 \\ 0.999 \\ 0.943 \\ \end{array}$	S         0.112         0.147         0.143         0.101         0.093         0.115         0.147         0.138         0.153         0.223         0.130         0.127         0.113	$\begin{array}{c} S_{\bar{R}} \\ 0.013 \\ 0.017 \\ 0.016 \\ 0.015 \\ 0.014 \\ 0.011 \\ 0.014 \\ 0.018 \\ 0.014 \\ 0.032 \\ 0.019 \\ 0.019 \\ 0.019 \\ 0.018 \\ 0.012 \end{array}$
ACCA ACCC ACCG ACCT CCCA CCCC CCCG CCCT GCCA GCCC GCCG GCCC GCCT TCCA TCCC	N 78 72 79 45 45 107 103 59 118 50 48 61 53 88 70	$\begin{array}{c} \overline{R} \\ 0.975 \\ 0.986 \\ 1.064 \\ 1.012 \\ 0.930 \\ 0.912 \\ 0.992 \\ 0.941 \\ 1.080 \\ 1.085 \\ 1.130 \\ 1.086 \\ 0.999 \\ 0.943 \\ 1.035 \\ \end{array}$	S0.1120.1470.1430.1010.0930.1150.1470.1380.1530.2230.1300.1500.1270.1130.149	$\begin{array}{c} S_{\bar{R}} \\ 0.013 \\ 0.017 \\ 0.016 \\ 0.015 \\ 0.014 \\ 0.011 \\ 0.014 \\ 0.018 \\ 0.014 \\ 0.012 \\ 0.019 \\ 0.019 \\ 0.019 \\ 0.012 \\ 0.018 \\ 0.012 \\ 0.018 \end{array}$

	Ν	$\overline{R}$	S	$S_{\overline{R}}$
ACGA	86	1.537	0.331	0.036
ACGC	96	1.362	0.309	0.032
ACGG	61	1.483	0.283	0.036
ACGT	53	1.381	0.369	0.051
CCGA	63	1.432	0.309	0.039
CCGC	91	1.257	0.220	0.023
CCGG	91	1.417	0.311	0.033
CCGT	55	1.328	0.246	0.033
GCGA	100	1.783	0.577	0.058
GCGC	87	1.462	0.358	0.038
GCGG	102	1.549	0.334	0.033
GCGT	72	1.427	0.324	0.038
TCGA	47	1.543	0.371	0.054
TCGC	104	1.300	0.282	0.028
TCGG	65	1.543	0.265	0.033
TCGT	58	1.263	0.227	0.030

	Ν	$\overline{R}$	S	$S_{\overline{R}}$
ACTA	44	1.160	0.210	0.032
ACTC	74	1.064	0.134	0.016
ACTG	58	1.157	0.243	0.032
ACTT	69	1.060	0.158	0.019
CCTA	23	1.031	0.145	0.030
CCTC	73	0.980	0.128	0.015
CCTG	73	1.118	0.157	0.018
CCTT	67	1.029	0.120	0.015
GCTA	40	1.227	0.166	0.026
GCTC	90	1.256	0.227	0.024
GCTG	112	1.303	0.216	0.020
GCTT	88	1.210	0.225	0.024
TCTA	31	1.098	0.068	0.012
TCTC	56	0.992	0.131	0.018
TCTG	101	1.149	0.195	0.019
TCTT	81	1.042	0.188	0.021

**Таблица 2** (продолжение). Выборочные характеристики относительных частот ультразвукового расщепления центральной фосфодиэфирной связи в 256 тетрануклеотидах.

Обозначения:

N – объём выборки;

- $\overline{R}$  средняя величина;
- *S* стандарнтое отклонение;
- $S_{\overline{R}}$  стандартное отклонение среднего.

	N	$\overline{R}$	S	$S_{\overline{R}}$
AGAA	96	0.986	0.117	0.012
AGAC	30	0.979	0.137	0.025
AGAG	48	0.956	0.144	0.021
AGAT	69	1.003	0.127	0.015
CGAA	100	0.982	0.205	0.020
CGAC	84	0.914	0.150	0.016
CGAG	52	0.962	0.102	0.014
CGAT	61	1.008	0.223	0.029
GGAA	102	0.968	0.123	0.012
GGAC	26	0.979	0.136	0.027
GGAG	31	0.898	0.175	0.031
GGAT	98	0.998	0.155	0.016
TGAA	88	0.936	0.124	0.013
TGAC	87	0.964	0.117	0.013
TGAG	77	0.997	0.106	0.012
TGAT	105	0.987	0.123	0.012

	Ν	$\overline{R}$	S	$S_{\overline{R}}$
AGGA	37	0.943	0.100	0.016
AGGC	70	0.905	0.150	0.018
AGGG	89	0.989	0.153	0.016
AGGT	46	0.921	0.155	0.023
CGGA	89	0.962	0.112	0.012
CGGC	69	0.831	0.133	0.016
CGGG	63	0.902	0.159	0.020
CGGT	100	0.925	0.148	0.015
GGGA	76	0.939	0.165	0.019
GGGC	71	0.881	0.155	0.018
GGGG	84	0.958	0.160	0.018
GGGT	71	0.948	0.121	0.014
TGGA	54	0.944	0.152	0.021
TGGC	93	0.859	0.142	0.015
TGGG	65	0.885	0.134	0.017
TGGT	93	0.922	0.130	0.013
	N	=	C	C

	Ν	$\overline{R}$	S	$S_{\overline{R}}$
AGCA	95	0.975	0.136	0.014
AGCC	46	0.922	0.136	0.020
AGCG	73	1.029	0.136	0.016
AGCT	84	0.974	0.145	0.016
CGCA	99	0.969	0.115	0.012
CGCC	106	0.886	0.173	0.017
CGCG	92	0.974	0.171	0.018
CGCT	82	0.929	0.146	0.016
GGCA	75	0.913	0.128	0.015
GGCC	68	0.868	0.121	0.015
GGCG	104	0.967	0.169	0.017
GGCT	58	0.982	0.144	0.019
TGCA	88	0.931	0.134	0.014
TGCC	55	0.958	0.143	0.019
TGCG	94	1.016	0.174	0.018
TGCT	87	1.008	0.155	0.017

	Ν	$\overline{R}$	S	$S_{\overline{R}}$
AGTA	36	0.972	0.083	0.014
AGTC	38	0.909	0.170	0.028
AGTG	59	0.972	0.099	0.013
AGTT	112	0.965	0.125	0.012
CGTA	60	0.987	0.081	0.010
CGTC	57	0.908	0.165	0.022
CGTG	32	0.981	0.178	0.031
CGTT	91	0.940	0.138	0.014
GGTA	87	0.959	0.123	0.013
GGTC	30	0.904	0.139	0.025
GGTG	88	0.966	0.134	0.014
GGTT	103	0.938	0.118	0.012
TGTA	76	0.967	0.126	0.014
TGTC	62	0.921	0.127	0.016
TGTG	83	0.960	0.135	0.015
TGTT	89	0.925	0.138	0.015

**Таблица 2** (продолжение). Выборочные характеристики относительных частот ультразвукового расщепления центральной фосфодиэфирной связи в 256 тетрануклеотидах.

Обозначения:

- N объём выборки;
- $\overline{R}$  средняя величина;
- *S* стандарнтое отклонение;
- $S_{\overline{R}}$  стандартное отклонение среднего.

	Ν	$\overline{R}$	S	$S_{\overline{R}}$
ATAA	103	0.974	0.109	0.011
ATAC	94	0.963	0.148	0.015
ATAG	49	1.020	0.164	0.023
ATAT	98	0.945	0.136	0.014
CTAA	30	1.005	0.077	0.014
CTAC	19	0.895	0.073	0.017
CTAG	24	0.998	0.065	0.013
CTAT	64	0.948	0.128	0.016
GTAA	93	1.001	0.127	0.013
GTAC	43	0.986	0.130	0.020
GTAG	28	1.068	0.157	0.030
GTAT	92	1.002	0.100	0.010
TTAA	103	0.925	0.111	0.011
TTAC	84	0.915	0.106	0.012
TTAG	53	1.012	0.113	0.016
TTAT	86	0.960	0.104	0.011
	N	$\overline{R}$	S	$S_{\overline{R}}$
ATCA	N 104	<u>R</u> 0.967	S 0.165	$S_{\overline{R}}$ 0.016
ATCA ATCC	N 104 79	R           0.967           0.877	S 0.165 0.096	$S_{\bar{R}}$ 0.016 0.011
ATCA ATCC ATCG	N 104 79 69	R           0.967           0.877           0.942	S 0.165 0.096 0.101	$S_{\bar{R}}$ 0.016 0.011 0.012
ATCA ATCC ATCG ATCT	N 104 79 69 64	$ $	S 0.165 0.096 0.101 0.110	$S_{\bar{R}}$ 0.016 0.011 0.012 0.014
ATCA ATCC ATCG ATCT CTCA	N 104 79 69 64 95	$     \overline{R}     0.967     0.877     0.942     0.941     0.892 $	S 0.165 0.096 0.101 0.110 0.109	$S_{\bar{R}}$ 0.016 0.011 0.012 0.014 0.011
ATCA ATCC ATCG ATCT CTCA CTCC	N 104 79 69 64 95 73	$     \bar{R}     0.967     0.877     0.942     0.941     0.892     0.842 $	S 0.165 0.096 0.101 0.110 0.109 0.166	$\begin{array}{c} S_{\bar{R}} \\ 0.016 \\ 0.011 \\ 0.012 \\ 0.014 \\ 0.011 \\ 0.019 \end{array}$
ATCA ATCC ATCG ATCT CTCA CTCC CTCC	N 104 79 69 64 95 73 51	$\begin{array}{c} \overline{R} \\ 0.967 \\ 0.877 \\ 0.942 \\ 0.941 \\ 0.892 \\ 0.842 \\ 0.937 \\ \end{array}$	S 0.165 0.096 0.101 0.110 0.109 0.166 0.107	$\begin{array}{c} S_{\bar{R}} \\ \hline 0.016 \\ 0.011 \\ 0.012 \\ 0.014 \\ 0.011 \\ 0.019 \\ 0.015 \end{array}$
ATCA ATCC ATCG ATCT CTCA CTCC CTCG CTCG	N 104 79 69 64 95 73 51 71	$\begin{array}{c} \overline{R} \\ 0.967 \\ 0.877 \\ 0.942 \\ 0.941 \\ 0.892 \\ 0.842 \\ 0.937 \\ 0.864 \end{array}$	S 0.165 0.096 0.101 0.100 0.109 0.166 0.107 0.123	$\begin{array}{c} S_{\bar{R}} \\ \hline 0.016 \\ 0.011 \\ \hline 0.012 \\ 0.014 \\ \hline 0.011 \\ 0.019 \\ \hline 0.015 \\ 0.015 \end{array}$
ATCA ATCC ATCG ATCT CTCA CTCC CTCC CTCG CTCT GTCA	N 104 79 69 64 95 73 51 71 70	$\begin{array}{c} \overline{R} \\ \hline 0.967 \\ 0.877 \\ 0.942 \\ 0.941 \\ 0.892 \\ 0.842 \\ 0.937 \\ 0.864 \\ 0.983 \\ \end{array}$	S0.1650.0960.1010.1100.1090.1660.1070.1230.129	$\begin{array}{c} S_{\bar{R}} \\ \hline 0.016 \\ \hline 0.011 \\ \hline 0.012 \\ \hline 0.014 \\ \hline 0.011 \\ \hline 0.019 \\ \hline 0.015 \\ \hline 0.015 \\ \hline 0.015 \\ \hline 0.015 \\ \hline \end{array}$
ATCA ATCC ATCG ATCT CTCA CTCC CTCG CTCT GTCA GTCC	N 104 79 69 64 95 73 51 71 70 36	$\begin{array}{c} \overline{R} \\ 0.967 \\ 0.877 \\ 0.942 \\ 0.941 \\ 0.892 \\ 0.842 \\ 0.937 \\ 0.864 \\ 0.983 \\ 0.932 \\ \end{array}$	S         0.165         0.096         0.101         0.110         0.109         0.166         0.107         0.123         0.129         0.080	$\begin{array}{c} S_{\bar{R}} \\ \hline 0.016 \\ 0.011 \\ 0.012 \\ 0.014 \\ 0.011 \\ 0.019 \\ 0.015 \\ 0.015 \\ 0.015 \\ 0.013 \\ \end{array}$
ATCA ATCC ATCG ATCT CTCA CTCC CTCC CTCG CTCT GTCA GTCC GTCC	N 104 79 69 64 95 73 51 71 70 36 48	$\begin{array}{c} \overline{R} \\ 0.967 \\ 0.877 \\ 0.942 \\ 0.941 \\ 0.892 \\ 0.842 \\ 0.937 \\ 0.864 \\ 0.983 \\ 0.932 \\ 0.970 \\ \end{array}$	S0.1650.0960.1010.1100.1090.1660.1070.1230.1290.0800.132	$\begin{array}{c} S_{\bar{R}} \\ \hline 0.016 \\ \hline 0.011 \\ \hline 0.012 \\ \hline 0.014 \\ \hline 0.011 \\ \hline 0.019 \\ \hline 0.015 \\ \hline 0.015 \\ \hline 0.015 \\ \hline 0.015 \\ \hline 0.013 \\ \hline 0.019 \\ \end{array}$
ATCA ATCC ATCG ATCT CTCA CTCC CTCG CTCT GTCA GTCC GTCG GTCG	N 104 79 69 64 95 73 51 71 70 36 48 30	$\begin{array}{c} \overline{R} \\ 0.967 \\ 0.877 \\ 0.942 \\ 0.941 \\ 0.892 \\ 0.842 \\ 0.937 \\ 0.864 \\ 0.983 \\ 0.932 \\ 0.970 \\ 0.962 \\ \end{array}$	S         0.165         0.096         0.101         0.101         0.109         0.166         0.107         0.123         0.129         0.080         0.132         0.144	$\begin{array}{c} S_{\bar{R}} \\ \hline 0.016 \\ 0.011 \\ 0.012 \\ 0.014 \\ 0.011 \\ 0.019 \\ 0.015 \\ 0.015 \\ 0.015 \\ 0.015 \\ 0.013 \\ 0.019 \\ 0.026 \end{array}$
ATCA ATCC ATCG ATCT CTCA CTCC CTCG CTCT GTCA GTCC GTCC	N 104 79 69 64 95 73 51 71 70 36 48 30 81	$\begin{array}{c} \overline{R} \\ 0.967 \\ 0.877 \\ 0.942 \\ 0.941 \\ 0.892 \\ 0.842 \\ 0.937 \\ 0.864 \\ 0.983 \\ 0.932 \\ 0.970 \\ 0.962 \\ 0.917 \\ \end{array}$	S0.1650.0960.1010.1100.1090.1660.1070.1230.1290.0800.1320.1440.115	$\begin{array}{c} S_{\bar{R}} \\ \hline 0.016 \\ \hline 0.011 \\ \hline 0.012 \\ \hline 0.014 \\ \hline 0.011 \\ \hline 0.019 \\ \hline 0.015 \\ \hline 0.015 \\ \hline 0.015 \\ \hline 0.015 \\ \hline 0.013 \\ \hline 0.026 \\ \hline 0.013 \\ \hline \end{array}$
ATCA ATCC ATCG ATCT CTCA CTCC CTCG CTCT GTCA GTCC GTCG GTCG	N           104           79           69           64           95           73           51           71           70           36           48           30           81           97	$\begin{array}{c} \overline{R} \\ 0.967 \\ 0.877 \\ 0.942 \\ 0.941 \\ 0.892 \\ 0.842 \\ 0.937 \\ 0.864 \\ 0.983 \\ 0.932 \\ 0.970 \\ 0.962 \\ 0.917 \\ 0.845 \\ \end{array}$	S0.1650.0960.1010.1100.1090.1660.1070.1230.1290.0800.1320.1440.1150.131	$\begin{array}{c} S_{\bar{R}} \\ \hline 0.016 \\ \hline 0.011 \\ \hline 0.012 \\ \hline 0.014 \\ \hline 0.011 \\ \hline 0.019 \\ \hline 0.015 \\ \hline 0.015 \\ \hline 0.015 \\ \hline 0.015 \\ \hline 0.013 $
ATCA ATCC ATCG ATCT CTCA CTCC CTCG CTCT GTCA GTCC GTCG GTCG	N           104           79           69           64           95           73           51           71           70           36           48           30           81           97           100	$\begin{array}{c} \overline{R} \\ 0.967 \\ 0.877 \\ 0.942 \\ 0.941 \\ 0.892 \\ 0.842 \\ 0.937 \\ 0.864 \\ 0.983 \\ 0.932 \\ 0.970 \\ 0.962 \\ 0.917 \\ 0.845 \\ 0.920 \\ \end{array}$	S0.1650.0960.1010.1090.1660.1070.1230.1290.0800.1320.1440.1150.1310.136	$\begin{array}{c} S_{\bar{R}} \\ \hline 0.016 \\ 0.011 \\ 0.012 \\ 0.014 \\ 0.011 \\ 0.019 \\ 0.015 \\ 0.015 \\ 0.015 \\ 0.015 \\ 0.013 \\ 0.013 \\ 0.013 \\ 0.013 \\ 0.014 \\ \end{array}$

	Ν	$\overline{R}$	S	$S_{\overline{R}}$
ATGA	85	1.017	0.105	0.011
ATGC	82	0.979	0.116	0.013
ATGG	66	1.006	0.123	0.015
ATGT	80	0.951	0.137	0.015
CTGA	89	1.008	0.141	0.015
CTGC	101	0.973	0.139	0.014
CTGG	82	1.008	0.109	0.012
CTGT	71	0.993	0.135	0.016
GTGA	86	1.026	0.125	0.013
GTGC	73	0.929	0.139	0.016
GTGG	55	1.011	0.121	0.016
GTGT	49	0.968	0.124	0.018
TTGA	97	0.983	0.125	0.013
TTGC	82	0.917	0.118	0.013
TTGG	102	0.960	0.132	0.013
TTGT	110	0.932	0.118	0.011
	•			
	N	$\overline{R}$	S	$S_{\overline{R}}$
ATTA	107	0.908	0.105	0.010
ATTC	95	0.908	0.122	0.013
ATTC	05	1.004	0.116	0.012

	- • •			
ATTC	95	0.908	0.122	0.013
ATTG	95	1.004	0.116	0.012
ATTT	105	0.938	0.123	0.012
CTTA	48	0.910	0.108	0.016
CTTC	90	0.895	0.098	0.010
CTTG	67	0.954	0.119	0.014
CTTT	97	0.926	0.152	0.015
GTTA	81	1.041	0.145	0.016
GTTC	66	0.954	0.133	0.016
GTTG	111	0.966	0.127	0.012
GTTT	138	0.945	0.125	0.011
TTTA	93	0.937	0.117	0.012
TTTC	127	0.890	0.133	0.012
TTTG	117	0.932	0.119	0.011
TTTT	236	0.897	0.140	0.009

Статистический анализ показал, что влияние фланкирующих (то есть ближайших по цепи) нуклеотидов на величину относительной частоты ультразвукового расщепления ДНК является значимым для всех динуклеотидов, причем как видно из рисунков 2.5, это влияние модулируется типами фланкирующих нуклеотидов схожим образом для различных динуклеотидов.

Полученные в данной главе результаты используются в следующих главах диссертации при построении модели расщепления ДНК по действием ультразвука (Глава 3), для интерпертации специфичности расщепления (Глава 4), а также при исследовании теоретических профилей расщепления промоторных последовательностей ДНК человека (Глава 5 диссертации).

Исследование характерных свойств ультразвукового расщепления ДНК, основные результаты которого приведены в разделе 2.2 этой главы, позволяет сделать определённые выводы о механизме расщепления фраментов ДНК под действием ультразвука. Ниже приведены характерные свойства наблюдаемого расщепления:

1.Увеличение степени расщепления фосфодиэфирных связей, находящихся в центральной части фрагмента ДНК;

2.Увеличение общего уровня расщепления при увеличении длины облучаемых фрагментов;

3.Увеличение общего уровня расщепления при увеличении вязкости раствора;4.Добавление соединений, способствующих нейтрализации свободных радикалов, не оказывает влияния на наблюдаемые картины расщепления ДНК.

Указанные свойства свидетельствуют о механохимической природе наблюдаемого расщепления фрагментов ДНК [Basedow et al, 1977]. Данный вывод используется в следующей главе диссертации при выборе модели расщепления ДНК под действием ультразвука.

44

### Глава 3. Моделирование расщепления фрагментов ДНК под действием акустической кавитации

Как уже было отмечено в Главе 1, в настоящее время расщепление полимеров под действием ультразвука высокой интенсивности связывают с кавитационными эффектами в облучаемом растворе. Общепринятой является модель расщепления, предложенная Томасом [Thomas, 1959], в соответствии с которой расщепление полимеров происходит в результате их взаимодействия с высокоградиентными течениями жидкости вблизи схлопывающихся кавитационных пузырьков (см. раздел 1.3). Описанная модель, как правило, используется для оценки растягивающих усилий, действующих на длинные полимерные цепи - длиной от нескольких сотен полимерных звеньев.

В Главе 3 на основе модели Томаса разрабатывается подход к описанию процесса расщепления молекул ДНК под действием акустической кавитации, который объединяет гидродинамическую модель кавитационных процессов, механическую модель ДНК, а также статистическую модель расщепления фосфодиэфирных связей под действием растягивающего усилия на основе теории механохимических реакций [Basedow et al, 1977].

#### 3.1. Моделирование динамики кавитационного пузырька.

Для того, чтобы получить оценку величин градиентов скорости течения жидкости, возникающих при схлопывании кавитационного пузырька, было проведено моделирование динамики его движения. Основной целью данного моделирования было получение оценки возникающих градиентов «снизу» – то есть расчет минимально возможных значений градиентов скорости течения, имеющих место при схлопывании кавитационных пузырьков в описанном эксперименте по расщеплению ДНК ультразвуком.

45

Для моделирования динамики кавитационного пузыря использовалось уравнение Рэлея-Плессета:

$$R\frac{d^2R}{dt^2} + \frac{3}{2}\left(\frac{dR}{dt}\right)^2 = \frac{1}{\rho}\left[P_i - P_\infty - \frac{2\sigma}{R} - \frac{4\mu}{R}\left(\frac{dR}{dt}\right)\right]$$
(3.1),

где R – радиус кавитационного пузырька, t – время,  $\rho$  – плотность жидкости,  $\sigma$  –её поверхностное натяжение,  $\mu$  – вязкость жидкости,  $P_i$  – давление газа внутри пузырька, а  $P_{\infty}$  - давление окружающей пузырек жидкости.

Данная модель основана на следующих приближениях: форма пузырька в процессе его движения не изменяется и описывается сферой определенного радиуса. Жидкость, окружающая пузырек, считается несжимаемой, а термодинамическое состояние газа внутри пузырька характеризуется наличием определенного давления, имеющего одно и то же значение для любой области внутри пузырька.

Параметры жидкости, входящие в уравнение (3.1) были взяты из экспериментальных данных, полученных при температуре воды 0-5 <sup>0</sup>C и приведенных в справочниках:

$$\sigma$$
= 75 10<sup>-3</sup> H/м, μ=1,7 мПа с,  $\rho$ =10<sup>3</sup> кг/м<sup>3</sup>.

Изменение давления окружающей пузырек жидкости описывалось гармоническим законом:

$$P_{\infty} = P_0 - P_A \cos(2\pi f t)$$

где  $P_0 = 10^5$  Па - атмосферное давление,  $P_A = 2.4 \ 10^5$  Па – амплитуда изменения ультразвукового давления, f = 22 кГц – частота ультразвука. Величина  $P_A$  связана с интенсивностью ультразвуковой волны уравнением:

$$P_A = \sqrt{2I\rho c}$$

в котором I – интенсивность ультразвуковой волны, принятая равной 2 Ватт/ см<sup>2</sup> (см. раздел 1.1), а c = 1480 м/с – скорость звука в воде.

Для описания сжатия газа в пузырьке использовалось адиабатическое приближение:

$$P_i = P_{i0} \left(\frac{R_0}{R}\right)^{3\gamma},$$

где *P<sub>i0</sub>* – давление насыщенного пара в пузырьке в момент, когда его радиус принимает максимальное значение  $R_m \approx 80$  мкм, а  $\gamma$  – экспериментальный показатель адиабаты для водяного пара, равный 1.3. Давление насыщенного пара воды достаточно сильно зависит от ее температуры: так при 0  $^{0}$ С оно составляет 0.61 кПа, а при 7  $^{0}$ С – уже 1 кПа. Тем не менее, как будет показано ниже, давление газа внутри пузырька является определяющим фактором его динамики лишь на конечной стадии схлопывания, которая по указанным ниже причинам в работе не рассматривалась. Адиабатическое приближение, используемое при расчетах, позволяет получить оценку «сверху» для величины внутреннего давления газа в пузырьке. Следует отметить, что внутреннее давление пара В пузырьке является единственным активным фактором, противодействующим сжатию пузырька. Как показано в работе [Flynn, 1964], учет явлений тепло- и массопереноса при схлопывании пузырька приводит к более высоким значениям скорости схлопывания, а значит, и большим величинам радиального градиента течения.

В зависимости от значения начального радиуса кавитационного пузырька  $R_0$ возможны два варианта его динамического поведения: в первом случае, если радиус пузырька меньше критического значения  $R_c = 2\sigma/(P_A - P_0 + P_i) \approx 1$  мкм,

пузырек схлопнется (в результате доминирования сил поверхностного натяжения), а во втором случае, - если радиус пузырька больше  $R_c$ , то пузырек начнет раздуваться под действием отрицательного внешнего давления. Важно отметить, что при  $R_0 > R_c$ , максимальный радиус пузырька  $R_m$  слабо зависит от начального значения  $R_0$ . Уравнение (3.1) решалось методом Рунге-Кутты 4 ого порядка с шагом по времени  $\Delta t$ =10<sup>-10</sup> с. На рисунке 3.2 представлен график зависимости радиуса кавитационного пузырька от времени.



**Рис.3.2.** Зависимость радиуса кавитационного пузырька R от времени t для различных значений начального радиуса  $R_0$ .

По достижении максимального значения  $R_m$  начинается процесс схлопывания, который можно условно разделить на три этапа: на первом этапе происходит медленное уменьшение радиуса пузырька в связи с сжимающим действием внешнего давления. Затем скорость схлопывания начинает резко увеличиваться, что связано со вторым членом уравнения (3.1), квадратичным по скорости схлопывания пузырька. Скорость схлопывания за короткий промежуток времени увеличивается до сотен метров в секунду. При достижении скорости схлопывания величины, равной  $10^3$  м/с расчет прекращался.

На рисунках 3.3 и 3.4 приведены результаты расчетов скорости схлопывания, соответствующих динамики кавитационного пузырька с минимальным начальным радиусом  $R_0=1.1$  мкм. На Рис.3.3 приведена зависимость скорости схлопывания от времени, условно соответствующая первому этапу схлопывания. Временная шкала в данном случае соответствует шкале на рисунке 3.2.



**Рис.3.3.** Зависимость скорости схлопывания  $\frac{dR}{dt}$  кавитационного пузырька от времени t. Приведены результаты моделирования для  $R_0=1.1$  мкм, что соответствует нижней кривой на Рис.3.2.

Рисунок 3.4 демонстрирует дальнейшее изменение скорости схлопывания, но уже на других временных масштабах. Таким образом, Рис.3.4 является продолжением Рис.3.3 и отражает зависимость скорости схлопывания от времени для второй стадии схлопывания кавитационного пузырька.



*Рис. 3.4.* Зависимость скорости схлопывания кавитационного пузырька от времени на второй стадии схлопывания.

Из Рис. 3.4 видно, что изменение скорости схлопывания на несколько сотен метров в секунду достигается на временном интервале порядка десятка наносекунд. Анализ временного поведения различных членов уравнения (3.1) показал, что для второй стадии схлопывания влияние давления газа внутри пузырька на динамику схлопывания является пренебрежимо малым.

Явления, происходящие на конечной, - третьей, стадии схлопывания, которые в данной работе не рассматривались, имеют характерные времена порядка наносекунд и связаны с катастрофическим ростом температуры и давления внутри пузырька. На этой стадии схлопывания важную роль играет упругость пара, заключенного в пузырьке и явления теплопереноса. С конечной стадией схлопывания кавитационного пузырька связаны химические и оптические эффекты кавитации, образование ударных волн, а также микроструй, вызванных отклонением формы пузырька от сферической [Brennen, 1995].

Результаты численного интегрирования уравнения (1), приведенные выше, далее использовались для вычисления радиального градиента скорости течения жидкости в окружающем пузырек пространстве. Из условия несжимаемости жидкости следует:

$$G \equiv \left| \frac{\partial V}{\partial r} \right| = 2 \frac{R^2}{r^3} \left| \frac{dR}{dt} \right|$$
(3.2),

где G – радиальный градиент скорости течения жидкости, V – радиальная скорость жидкости на расстоянии r от центра пузырька (r > R).

На рисунке 3.5 приведены результаты расчета градиента скорости течения жидкости для второй стадии схлопывания кавитационного пузырька. Видно, что градиент скорости течения уменьшается с ростом расстояния до центра пузырька и растет по мере его схлопывания.



**Рис.3.5.** Зависимость радиального градиента скорости течения жидкости G вблизи кавитационного пузырька на второй стадии схлопывания от расстояния до поверхности пузырька d и времени t.

#### 3.2. Моделирование взаимодействия фрагмента ДНК с кавитационным течением.

Для исследования динамики фрагмента ДНК при его движении в жидкости с высоким градиентом скорости течения использовалась следующая модель [Doi et al, 1986]: фрагмент представлялся системой шариков определенного гидродинамического радиуса  $R_D$ =1 нм, соединенных пружинами с константой жесткости *k*=6000 пН.



**Рис.3.6**. Иллюстрация к механической модели фрагмента молекулы ДНК. *R*-гидродинамический радиус шарика, т – его масса, k – жесткость пружины.

Сила взаимодействия каждого звена такой модельной системы с жидкостью определяется законом Стокса:

$$F = 6\pi\mu R_D V \tag{3.3},$$

где F – модуль силы взаимодействия,  $R_D$  - гидродинамический радиус шарика, V – модуль скорости шарика относительно жидкости,  $\mu$  – динамическая вязкость жидкости. Уравнение движения Ньютона для i - ого звена модельной системы будет иметь следующий вид:

$$m\ddot{x}_{i} = \frac{k(x_{i+1} - x_{i} - l_{0})}{l_{0}} - \frac{k(x_{i} - x_{i-1} - l_{0})}{l_{0}} + \alpha(V_{a} - \dot{x})$$
(3.4),

где *m* – масса *i*-ого шарика,  $x_i$  – его координата, k – жесткость пружин,  $l_0$  – длина нерастянутой пружины,  $\alpha$  – константа, характеризующая силу вязкого трения между шариком и жидкостью ( $\alpha = 6\pi\mu R_D$ ),  $V_a$  – проекция абсолютной скорости течения жидкости на направление оси *x*. Ось *x* направлена вдоль системы звеньев и соответствует направлению их условной нумерации (т.е.  $x_{i+1} > x_i$ ).

Собственная длина  $l_0$  элементарного упругого сегмента – то есть пружины была принята равной  $2R_D$ , что как видно из Рис. 3.6. соответствует блоку длиной около 6 нуклеотидных пар. В таком случае значение константы  $\alpha$  соответсвует величине, полученной из анализа диффузии олигонуклеотидов ДНК [Fernandes et al, 2002].

Система «помещалась» в область с одномерным течением вдоль ось х и имеющим постоянный и однородный градиент скорости течения вдоль оси. Динамика фрагмента ДНК рассчитывалась численным интегрированием системы уравнений типа (3.4), составленных для всех звеньев, при помощи модифицированного алгоритма Верле, после чего находился профиль, характеризующий изменение силы растяжения вдоль системы, то есть зависимость величины  $F=k(x_{i+1}-x_i - l_0)$  от номера звена *i*. На рисунке 3.7 приведены результаты расчетов для системы, состоящей из 52 звеньев, то есть имеющей эффективную длину вдоль оси *x* порядка 300 пар оснований.



**Рис.3.7.** Зависимость силы растяжения F звена модельной системы от его положения вдоль цепи і . Приведены данные расчетов для трех значений градиента скорости течения жидкости G.

Полученные результаты показывают, что при значении градиента скорости течения более 10<sup>8</sup> с<sup>-1</sup> на фрагмент ДНК, движущийся вместе с жидкостью действуют растягивающие усилия величиной несколько тысяч пиконьютонов, что по порядку величины соответствует пределу прочности молекулы ДНК на разрыв (см. раздел 1.2. Главы 1).

Стоит отметить, что процесс растяжения модельной системы имеет нелинейный характер: при удлинении цепи сила растяжения увеличивается, что приводит к дополнительному растяжению. Таким образом, наличие положительной обратной связи при растяжении полимера при его движении в области и сильным градиентом скорости течения, приводит к нелинейной зависимости силы растяжения от величины градиента скорости течения. Эта зависимость приведена на Рис. 3.8.



**Рис.** 3.8. Зависимость силы растяжения **F** в центре фрагмента от величины продольного градиента скорости течения жидкости G.

Величина силы растяжения полимера, возникающая при его движении в жидкости, в которой существует градиент скорости течения, зависит от ориентации полимера по отношению к направлению вектора этого градиента. Результаты, полученные выше, относятся к случаю параллельного расположения фрагмента ДНК и вектора градиента скорости течения, при котором величина растягивающего усилия имеет максимум, так как градиент скорости течения в данной модели является продольным, то есть сонаправлен с вектором скорости течения в любой точке жидкости. Стоит отметить, что в данной модели цепь полимера изначально является прямолинейной. Так как персистентная длина молекулы ДНК составляет величину порядка 150 пар оснований (см. раздел 1.2 Главы 1.), то для фрагмента длиной около 300 пар оснований такое приближение не совсем оправдано. Тем не менее, в случае высоких градиентов течения, имеющих место вблизи схлопывающихся кавитационных пузырьков, растяжение полимера, очевидно, приводит к его «линеаризации». Структура течения также способствует «выравниванию» фрагментов вдоль вектора скорости течения жидкости.

Сила растяжения модельной системы также существенным образом зависит от таких параметров модели, как вязкость жидкости, жесткость сегментов и количество звеньев в цепи. Вязкость жидкости в модели соответствовала вязкости воды при температуре 0 °C:  $\mu$ =1,7 мПа, а константа жесткости, k=6000 пкН – получена при помощи моделирования растяжения коротких фрагментов ДНК в рамках молекулярнодинамического подхода (неопубликованные результаты), и принимает промежуточное значение по отношению к экспериментальным данным по растяжению единичных молекул ДНК (см. раздел 1.2 Главы 1) (*k*≈1000 пкН ) [Bustamante et al, 1994] и экспериментальным данным по определению скорости звука в ДНК (*k*≈8000 пкН) [Hakim et al, 1984]. Важно отметить, что полученные значения растягивающего усилия, действующего в центральной части фрагмента и составляющего при  $G > 10^8 \text{ c}^{-1}$ величину в несколько тысяч пиконьютонов могут приводить к значительному изменению конформации полимера. В экспериментах по растяжению единичных молекул ДНК [Smith et al, 1996] конформационный переход ДНК из В в S – форму происходил уже при 70 пкН. Однако, в описанных экспериментах (см. раздел 1.2) процесс растяжения характеризовался временами порядка десятков секунд, в то время как растяжение фрагментов ДНК под действием акустической кавитации, в соответствии с полученными результатами, происходит на временах порядка десятка наносекунд. Переход ДНК в S-форму, сопровождающийся резким изменением длины молекулы более чем в полтора раза в таком случае может быть затруднен из-за наличия множества связанных с ДНК молекул растворителя. Модель, предлагаемая в четвертой главе диссертации строится на предположении, что фрагмент ДНК все же не претерпевает серьезных конформационных переходов в процессе взаимодействия с высокоградиентным течением жидкости.

## 3.3. Моделирование кинетики расщепления фрагментов ДНК под действием акустической кавитации.

Как известно, наличие механического напряжения в молекуле может привести к Химическая молекулы, ee разрыву. реакция диссоциации катализируемая воздействием, механическим является частным случаем так называемых механохимических реакций. Кинетика механохимических реакций, как правило, описывается в рамках теории активированного комплекса или же эмпирическим уравнением, схожим с уравнением Аррениуса.

В соответствии с модифицированной теорией активированного комплекса, скорость механохимической реакции в случае наличия «катализирующего» реакцию усилия f может быть выражена следующим образом [Schmidt et al, 2008]:

$$k_{p} = \frac{kT}{2\pi \hbar} e^{\frac{\Delta S}{k}} e^{-\frac{\Delta E_{f}}{kT}}$$
(3.5)

где  $k_p$  – скорость реакции, k- постоянная Больцмана, T – абсолютная температура,  $\hbar$  – постоянная Планка,  $\Delta S$  – изменение энтропии при образовании активированного комплекса,  $\Delta E_f$  - энергетический барьер образования активированного комплекса при наличии усилия f (энергия активации):

$$\Delta E_f = E^* - E_0 - f \Delta x \tag{3.6},$$

где  $E^*$  - энергия активированного комплекса,  $E_0$  – энергия основного состояния (начальная энергия реагентов), f – проекция внешней силы на направление, соответствующее условной координате реакции ( в случае реакции диссоциации или гидролиза это направление совпадает с направлением ковалентной связи),  $\Delta x$  – изменение условной координаты реакции при образовании активированного комплекса.

В случае реакции гидролиза фосфодиэфирной связи сахарофосфатного остова ДНК величина активационного барьера  $\Delta E^*$  в отсутствии внешней силы составляет величину около 26 ккал/моль, а энтропийный барьер – около 34 кал/( К<sup>·</sup> моль). Таким образом, скорость гидролиза ДНК по C3<sup>°</sup>-O3<sup>°</sup> связи при нормальных условиях составляет величину порядка  $k_0=2 \cdot 10^{-13} \text{ c}^{-1}$ .

При наличии растягивающего усилия *F*, скорость реакции, в соответствии с уравнениями (3.5) и (3.6) может быть записана в виде:

$$k_f = k_0 e^{\frac{f\Delta x}{kT}}$$
(3.7),

где f – величина силы, действующей на растяжение ковалентной связи C3'-O3', которая зависит от распределения «нагрузки» F по степеням свободы молекулы, - то есть в том числе от ориентации связи по отношению к направлению действия растягивающей силы F. Величина  $\Delta x$  представляет изменение длины ковалентной связи, необходимой для гидролиза связи. Предполагая, что основную часть активационного барьера реакции гидролиза фосфодиэфирной связи составляет энергия, необходимая для достаточной «деформации» ковалентной связи C3'-O3', величину  $\Delta x$  можно определить из условия:

$$V_m(\Delta x) = \Delta E^* \tag{3.8}$$

где  $V_m(\Delta x)$  – потенциал Морзе, описывающий деформацию ковалентной связи C3'-O3' [Basedow et al, 1977]:  $V_m(\Delta x)=D(1-e^{-\beta\Delta x})^2$ ,

для которого D=83 ккал/моль,  $\beta=2.58$  Å<sup>-1</sup>. Из уравнения (3.8) следует, что величина  $\Delta x$  составляет около 0.3 Å.

При помощи полученных соотношений можно оценить долю разорванных фрагментов ДНК за один цикл схлопывания кавитационного пузырька: результаты моделирования, полученные в разделах 3.1 и 3.2, могут быть объединены с целью расчета кинетики расщепления ДНК на второй стадии схлопывания кавитационного пузырька.

Пусть в некоторый момент времени t стадии схлопывания градиент скорости течения жидкости в окружающем пузырек пространстве описывается функцией G(r,t). Если концентрацию фрагментов ДНК n(r,t) считать постоянной величиной  $n(r,t)=n_0$ , то в сферическом слое радиуса r и толщины dr будет находится  $4\pi r^2 dr n_0$  фрагментов ДНК. Считая, что некоторая их часть  $\beta$  ориентирована по скорости течения, можно определить действующую на эти фрагменты со стороны жидкости растягивающую силу F, которая зависит от величины градиента скорости скорости течения G(r,t), то есть F(G(r,t)). Здесь под F понимается сила растяжения, действующая в центре фрагмента ДНК. Общая сила растяжения F распределяется по степеням свободы молекулы и частично передается на каждую фосфодиэфирную связь, увеличивая вероятность ее гидролиза. Таким образом, за время dt гидролиз определенной фосфодиэфирной связи в центральной части фрагмента произойдет в среднем в

$$dN = \beta 4\pi r^2 dr n_0 k_f dt$$

фрагментах ДНК, где под  $k_f$  понимается скорость механохимической реакции гидролиза, которая определяется в соответствии с уравнением (3.7). Существенным пунктом в данной модели является переход от общей силы растяжения F к ее составляющей f, действующей на уровне фосфодиэфирной связи C3'-O3':

$$f = \gamma F \tag{3.9},$$

где  $\gamma$  – эффективный множитель, далее называемый трансмиссионным коэффициентом. Как видно из (3.9), данный множитель характеризует долю растягивающего усилия *f*, приходящегося на фосфодиэфирную связь, по отношению к величине суммарного натяжения *F*. Так как сахарофосфатный остов ДНК состоит из двух цепей, то  $\gamma \leq 0.5$ . Суммируя, доля разорванных фрагментов ДНК на определенной стадии схлопывания может быть выражена следующим образом:

$$k_{om\mu} = \frac{\iint_{t,r} 4\pi r^2 n_0 \beta k_0 e^{\frac{\gamma F(G(r,t))}{kT}} dt dr}{\int_{r} 4\pi r^2 n_0 dr}$$
(3.10),

где интегрирование ведется по соответствующему промежутку времени, а для каждого момента времени *t* интегрирование ведется по диапазону  $R(t) \le r \le R_{max}$ , где

R(t) – радиус кавитационного пузырька в момент времени *t*, а  $R_{max} = R(t) + \Delta R$ . Параметр  $\Delta R$  был принят равным 50 микрон и имеет смысл ограничения диапазона интегрирования в силу быстрого уменьшения величины градиента скорости течения с ростом расстояния *r* от центра пузырька (см. уравнение 3.2).

Для расчета величины  $k_{omh}$  использовались результаты моделирования динамики кавитационного пузырька. Рассматривался временной промежуток продолжительностью 80 нс, соответствующий второй стадии схлопывания (см. Рис. 3.4). Как уже было отмечено, значения G(r,t) вычислялись при помощи соотношения (3.2). Зависимость F(G), полученная в разделе 3.2 и приведенная на Рис. 3.4. была апроксимирована полиномом четвертой степени методом наименьших квадратов. Интеграл (3.10) вычислялся прямым суммированием с шагом по времени  $\Delta t=10^{-10}$  с и шагом  $\Delta r=10^{-7}$  м по параметру r.

Ниже приведены результаты расчетов величины  $k_{omh}$  для различных значений  $\gamma$ . График, представленный на Рис.3.9, получен для  $\beta=1$ . Приведенные величины для  $k_{omh}$ 

соответствуют вероятности разрыва фосфодиэфирной связи в центральной части фрагмента ДНК за секунду, то есть за 22000 цикла схлопывания.



**Рис. 3.9.** Зависимость доли расщепленных фрагментов ДНК  $k_{omh}$  в пересчете на 1 секунду (то есть 22000 эффективных циклов схлопывания кавитационного пузырька) от значения трансмиссионного коэффициента у.

Экспериментальное значение усредненной вероятности ультразвукового расщепления фрагмента ДНК длиной 312 пар оснований за секунду было получено при помощи анализа данных электрофореза и составляет величину порядка  $10^{-5}$ . Таким образом, теоретические результаты для  $\gamma=0.4$  соответствуют экспериментальным данным, если предположить, что вероятность образования кавитационного пузырька за цикл разряжения ультразвуковой волны составляет величину порядка  $10^{-2}$  – то есть концентрация зародышей кавитации составляет величину порядка  $10^3$  мл<sup>-1</sup>.

Данные расчетов, приведенные на Рис.3.9., можно также сравнить с экспериментальными данным по ультразвуковому расщеплению фрагментов ДНК, содержащих изначальный разрыв по одной из цепей. Показано, что частота расщепления фосфодиэфирных связей противоположной цепи ДНК, находящихся напротив разрыва, увеличивается в 10-30 раз, в зависимости от места расположения ника (то есть начального разрыва) в фрагменте [Il'icheva et al, 2009]. Наличие разрыва в одной из цепей ДНК приводит к тому, что большая часть нагрузки, связанной с

действием силы растяжения, переносится неразорванную цепь сахарофосфатного остова, что можно приближенно описать следующим образом:

$$\gamma_n = 2 \gamma_0, \tag{3.11}$$

где  $\gamma_n$  и  $\gamma_0$  – трансмиссионные коэффициенты в случае наличия разрыва в противоположной цепи и при его отсутствии, соответственно.

Отношение теоретических констант скоростей разрыва k, рассчитанных на основании уравнения (3.10), для значений  $\gamma_n$  и  $\gamma_0$ , и полученных при  $\gamma_0=0.2$ ,  $\gamma_0=0.3$  и  $\gamma_0=0.4$ , равны соответственно: 16; 7 и 6, то есть находятся в качественном согласии с экспериментальными данными, приведенными выше.

Уравнение (3.10) может быть модифицировано с целью нахождения теоретического профиля ультразвукового расщепления, то есть зависимости вероятности расщепления фрагмента ДНК по данной фосфодиэфирной связи от положения этой связи в цепи полимера:

$$k_{omh}^{i} = \frac{\iint_{r} 4\pi r^{2} n_{0} \beta k_{0} e^{\frac{\gamma F_{i}(G(r,t))}{kT}} dt dr}{\int_{r} 4\pi r^{2} n_{0} dr}$$
(3.12)

где *i* –номер звена модельной системы, в котором происходит расщепление фосфодиэфирной связи, а  $F_i$  – сила растяжения этого звена. Показано, что отношение силы растяжения центрального звена модельной системы к силе растяжения *i*-ого звена практически не зависит от величины градиента скорости течения жидкости (данные не приводятся), поэтому для описания позиционной зависимости силы растяжения достаточно ввести функцию  $\alpha(i)$ :

$$F_i = F_{\max} \cdot \alpha(i), \tag{3.13}$$

где  $F_{\text{max}}$  – сила растяжения центрального звена модельной системы, то есть  $\alpha(i) \leq 1$ . На рисунке 3.10 представлены результаты расчета относительных скоростей реакции гидролиза на основании уравнения (3.11), в котором ориентационный фактор  $\beta$  также принят равным единицы, а  $\gamma = 0.4$ .



**Рис.3.10.** Сравнение теоретического профиля относительных скоростей разрыва звеньев модельной системы с экспериментальным профилем расщепления ДНК ультразвуком.

Как видно из рисунка, полученные теоретические результаты позволяют качественно описать позиционный эффект, наблюдаемый при анализе картин расщепления ДНК ультразвуком.

#### 3.4. Обсуждение

Полученные результаты свидетельствуют о применимости модели Томаса к описанию ультразвукового расщепления фрагментов ДНК длиной от нескольких десятков до сотен нуклеотидных пар. Предложенный подход к описанию процесса расщепления основывается на этой модели и объединяет:

1. гидродинамическую модель динамики кавитационного пузырька;

2. механическую модель, описывающую взаимодействие фрагмента ДНК с высокоградиентным течением жидкости, возникающим в процессе схлопывания кавитационного пузырька;

3. статистическую модель кинетики процесса расщепления фосфодиэфирных связей фрагмента ДНК под действием возникающих сил растяжения.

Полученные результаты демонстрируют возможность расщепления ДНК вблизи кавитационного пузырька уже на промежуточной стадии схлопывания, при которой величина градиента скорости течения жидкости может составлять величину порядка

10<sup>8</sup> с<sup>-1</sup>, а растягивающие молекулу усилия, возникающие при ее взаимодействии с окружающей жидкостью, - порядка нескольких тысяч пиконьютонов. Особенность распределения растягивающего усилия вдоль полимера также позволяет описать наблюдаемый эффект позиционной зависимости вероятности расщепления фосфодиэфирных связей ДНК от их положения вдоль цепи.

Полученные результаты носят качественный характер в силу ряда факторов, усложняющих более точное описание наблюдаемого процесса расщепления ДНК под действием ультразвука. Предложенные модели опираются на приближение сильно разбавленного раствора ДНК, которое не вполне правомерно: концентрация ДНК, имеющая место в эксперименте, составляет величину 5 - 10 мкг/мл, а наличие в буферном растворе различных компонентов (например, Na<sup>+</sup>) может приводить к изменению большинства параметров, используемых при расчетах.

Принятое при моделировании значение интенсивности ультразвуковой волны I=2 Ватт/см<sup>2</sup> может быть завышенным, так как было определено косвенным образом – при помощи сравнения с экспериментальными данными по выходу перекиси водорода, полученным при более высокой температуре. Тем не менее, представленные результаты моделирования предлагают скорее нижний предел значений градиентов скорости течения жидкости, так как не учитывают ряд факторов, приводящих к увеличению скорости схлопывания пузырька, не говоря уже о наличии третьей стадии схлопывания, которая в предложенной модели не рассматривалась. Также остается «открытым» вопрос о конформации полимера при наличии значительной силы растяжения, действующей на временных масштабах порядка нескольких наносекунд.

Значение константы жесткости k=6000 пкН, принятое в механическое модели постоянным, зависит от силы растяжения ДНК, что в свою очередь влияет на динамику ее растяжения. Стоит также отметить, что полученные времена релаксации полимера к стационарному состоянию растяжения при его движении в жидкости с высоким значением продольного градиента скорости течения, составляли величину порядка нескольких наносекунд (данные не приведены). Таким образом, использование равновесных значений сил растяжения F(G) в уравнении (3.10) как функций градиента скорости течения G, также является приближением, приводящим к завышению значений растягивающих усилий. Растягивающие усилия F также существенным образом зависят от вязкости окружающей полимер жидкости и гидродинамического радиуса звена модельной системы – параметров, значения которых известны лишь приближено. Все эти величины, включая трансмиссионный множитель у, влияют на значение силы растяжения *f*, действующей на уровне ковалентной связи и являющейся определяющей для кинетики процесса расщепления в силу экспоненциальной зависимости скорости расщепления от величины f. Таким образом, изменяя величину у, можно моделировать влияние на теоретическую скорость расщепления ДНК таких параметров как вязкость жидкости, гидродинамический радиус сегмента цепи, длина фрагмента и его жесткость. Также стоит отметить, что предположение о гидролизе ДНК в качестве основного механизма разрыва цепи полимера по действием растягивающего усилия может быть дополнено рассмотрением реакции прямой диссоциации ковалентной связи под действием экстремально больших значений растягивающих усилий, по всей видимости, имеющих место на конечной стадии схлопывания кавитационного пузырька, а также в случае расщепления под действием ударных волн кавитации.

Следует отметить, что сравнение результатов моделирования с экспериментальными данными по скорости расщепления ДНК, приведенное выше носит условный характер, так как концентрация зародышей кавитации в описанных экспериментах неизвестна. Этот параметр может принимать значения от единиц до сотен тысяч зародышей на миллилитр раствора [Arora et al, 2007], и существенным образом зависит от его физико-химических свойств. Полученная оценка концентрации зародышей - 10<sup>3</sup> мл<sup>-1</sup>, при которой теоретические значения скорости расщепления ДНК (у=0.4) качественно соответствуют экспериментальным данным, – также существенным образом зависит от величины транмиссионного коэффициента у. При у=0.2 оценка концентрации зародышей кавитации приводит к величине порядка 10<sup>4</sup> мл<sup>-1</sup>. Таким полученные оценки соответствуют образом, достаточно высоким значениям концентрации зародышей кавитации. Важно отметить, что имеются данные, свидетельствующие о том, что присутствие молекул ДНК может приводить к увеличению концентрации зародышей кавитации в растворе [Lentz et al, 2005].

## Глава 4. Подходы к интерпретации специфичности расщепления ДНК ультразвуком

Среди особенностей наблюдаемого расщепления ДНК ультразвуком можно выделить две основные черты – это наличие позиционной зависимости частоты расщепления и специфичность расщепления к нуклеотидной последовательности облучаемых фрагментов. Данная глава посвящена исследованию обнаруженной специфичности ультразвукового расщепления ДНК, а также построению модели, предлагающей интерпретацию этого явления.

#### 4.1. Особенности конформационной динамики В- формы ДНК

Как уже было отмечено ранее, существует множество физико-химических характеристик ДНК, зависящих от последовательности нуклеотидов в молекуле. Основная задача данного раздела – выявление таких структурных свойств ДНК, которые в наибольшей степени согласуются со статистическими данными по специфичности расщепления ДНК ультразвуком.

Одной из наиболее успешных и активно используемых параметризаций контекстно-зависимых свойств ДНК является описание термодинамических параметров ДНК в рамках модели ближайших соседей – то есть на уровне динуклеотидного приближения. Стоит отметить, что параметры свободной энергии образования дуплекса ДНК, характеризующие термодинамику плавления двойной спирали молекулы, не согласуются с полученным данными по ультразвуковому расщеплению ДНК – так, например, динуклеотиды d(CpG) и d(GpC) характеризуются практически одинаковыми значениями свободных энергий образования дуплекса, в то время как относительные частоты их ультразвукового расщепления отличаются более чем в 1.5 раза.

Также известно, что образование необычных структур в ДНК, таких как кинки (изломы) и пузыри (локальное раскрытие дуплекса), наиболее часто происходит в А-Т богатых участках молекулы. Тем не менее, частота ультразвукового расщепления ДНК в таких областях значимо не отличается от среднего уровня расщепления.

Как было показано в Главе 3, по всей видимости, существенную роль в процессе расщепления ДНК под действием акустической кавитации играет распределение растягивающего усилия по степеням свободы молекулы: в зависимости от конформации и динамических особенностей сахарофосфатного остова, растягивающее усилие может по-разному передаваться на фосфодиэфирную связь, приводя к увеличению, или наоборот – уменьшению вероятности ее расщепления. Таким образом, в процессе расщепления ДНК под действием ультразвука могут проявляться конформационно-динамические особенности сахарофосфатного остова ДНК.

В В-форме ДНК существует ряд локальных конформационных движений, динамика которых определяется нуклеотидным составом молекулы, а также физикохимическими параметрами растворителя. Наиболее изученными среди них является S $\leftrightarrow N$  интерконверсия дезоксирибозы [Isaacs et al, 2001], а также ВІ $\leftrightarrow$ ВІІ переходы сахарофосфатного остова [Heddi et al, 2006].

Интерконверсия сахарного цикла связана с транс↔гош<sup>+</sup> вариациями угла δ (см. Рис.4.1), влияющего на ориентацию фосфодиэфирной связи СЗ'-ОЗ', а также с вращением цикла по отношению к гликозидной связи (изменение угла χ).



**Рис.4.1.** Фрагмент сахарофосфатного остова В-формы ДНК с обозначениями углов внутреннего вращения.

Как уже было отмечено, одним из основных результатов статистического анализа специфичности ультразвукового расщепления ДНК является значимое увеличение относительной частоты расщепления фосфодиэфирной связи, следующей за дезоксицитидином, то есть связи, соединяющей сахарный цикл, примыкающий к цитозину, с фосфатной группой следующего по цепи нуклеотида (см. Рис.2.4 и таблицу 1). Так как на Рис.4.1. приведены два подряд идущих дезоксицитидина, то сказанное выше относится к обеим связям C3'-O3', изображенным на рисунке.

Таким образом, можно предположить, что некоторое структурное свойство участка, соединяющего дезоксицитидин со следующим по цепи нуклеотидом, должно иметь особенность, отличающую это сочленение от похожих сайтов, образованных другими типами нуклеотидов. Это предположение подтверждается как теоретическими, так и экспериментальными данными.

Квантово-химические расчеты *Ab initio* демонстрируют уникальные свойства дезоксицитидина [Foloppe et al, 1999]: область минимальной потенциальной энергии как при *S*, так и при *N* конформации дезоксирибозы соответствует углам  $\chi$ , относящимся к A-форме ДНК. Более того, конформация, соответствующая *N*-типу дезоксирибозы в дезоксицитидине обладают меньшей энергией, чем для *S*-типа сахарного цикла, который свойственен B-форме ДНК [Foloppe et al, 1999]. Таким образом, конформационные предпочтения дезоксицитидина не вполне соответствуют B-форме ДНК, что может приводить к конформационным переходам его дезоксирибозы. В действительности, эти переходы были обнаружены при помощи ЯМР-методов и молекулярно-динамических расчетов: показано, что дезоксирибоза, примыкающая к цитозину, подвержена S $\leftrightarrow$ N интерконверсии в значительно большей степени, чем дезоксирибозы других нуклеотидов [Duchardt et al, 2008; Wu et al, 2003].

Другой тип конформационных движений в В-ДНК – так называемый ВІ $\leftrightarrow$ ВІІ переход – определяется согласованным изменением углов є и  $\zeta$  (см. Рис. 4.1) и подразумевает переход от канонического состояния ВІ (соответственно транс и гош<sup>-</sup> конформация углов є и  $\zeta$ ) к состоянию ВІІ (значения углов - гош<sup>-</sup> и транс). Вероятность переходов, а также заселенность конформационных состояний также зависит от локального нуклеотидного состава ДНК. Существующая статистика по

конформеров **BI/BII** равновесным долям для различных нуклеотидных последовательностей [Heddi et al, 2010] не находит значимой корреляции с величинами относительных частот ультразвукового расщепления ДНК. Высокий процент конформеров BII типа в динуклеотидах d(CpG) и d(CpA), имеющих высокий уровень ультразвукового расщепления, также имеет место в динуклеотидах d(TpG), d(GpG), d(GpC) и d(GpA), со значительно меньшими значениями относительных частот ультразвукового расщепления. Более того, относительная частота расщепления в динуклеотиде d(CpT) высока, в то время как для этого динуклеотида BII состояний не было обнаружено. Таким образом, склонность к ВІ↔ВІІ переходам для определенного сайта ДНК, по всей видимости, не оказывает значимого влияния на вероятность ультразвукового расщепления фосфодиэфирной связи в этой области.

Подводя итог данному разделу, можно заключить, что единственной особенностью конформационной динамики В-ДНК, которая согласуется с полученными данными по специфичности ультразвукового расщепления, является S↔N интерконверсия дезоксирибозы, примыкающей к цитозину. В следующем разделе на основе данного заключения построена модель, при помощи которой дается интерпретация специфичности расщепления ДНК под действием ультразвука высокой интенсивности.

# 4.2. Влияние конформационной подвижности дезоксирибозы на эффективность ультразвукового расщепления ДНК.

Как уже было отмечено, при  $S \leftrightarrow N$  интерконверсии дезоксирибозы происходит переориентация C3'-O3' связи. Для более детального анализа особенностей *N*-конформации дезоксирибозы были исследованы кристаллографические структуры свободных олигонуклеотидов ДНК, содержащих C3'- эндо конформацию дезоксирибозы (то есть минорную – *N*-конформацию). Были рассмотрены структуры с кодовыми именами 1fq2, 1jgr, 2fih и 355d, приведенные в структурной базе данных

нуклеиновых кислот <u>http://www.ndbserver.rutgers.edu/</u>, причем в каждой из структур содержалось несколько C3'-ендо конформаций дизоксирибозы.

Анализ кристаллографических структур олигонуклеотидов ДНК, содержащих *N*–конформацию дезоксирибозы показал, что в СЗ' эндо конформации (то есть в минорной, *N*-конформации), уменьшается угол между осью спирали и направлением СЗ'-ОЗ' связи, как показано на Рис. 4.2.



**Рис.4.2.** Расположение фосфодиэфирной связи по отношению к оси спирали ДНК в случае C2' эндо и C3' эндо конформации дезоксирибозы. C3' эндо конформация изображена менее ярко. Ось z – ось двойной спирали ДНК.

Для N конформации получено следующее значение среднего угла между осью спирали *z* и направлением C3'-O3' связи:  $\alpha_N = 64^\circ \pm 3^\circ$ , а для S-конформации (то есть C2' – эндо конформации сахарного цикла, свойственной B-форме ДНК):  $\alpha_S = 81^\circ \pm 5^\circ$ .

Как было отмечено в Главе 3, значение трансмиссионного коэффициента  $\gamma$ , характеризующего долю растягивающего усилия f, приходящегося на фосфодиэфирную связь, по отношению к величине суммарного натяжения F, зависит от распределения последнего по степеням свободы молекулы. Предположим, что

данная цепь сахарофосфатного остова имеет некоторое натяжение F<sub>o</sub>, сонаправленное с осью спирали (см Рис.4.2.), то есть направленное вертикально. Рассмотрим распределение сил, действующих на атом ОЗ' в рамках молекулярно-динамической модели, включающей вибрационную энергетическую компоненту ( то есть энергию, связанную с деформацией валентной связи) и энергию, связанную с деформацией валентных углов. Энергетическая составляющая, связанная с деформацией углов внутреннего вращения, а также с невалентными взаимодействиями, не рассматривается в силу малости значений сил, вызываемых такого рода взаимодействиями по сравнению с силами растяжения, приводящими к разрыву молекулы. Проекция силы растяжения остова на направление связи, действующая со стороны нижней части молекулы на атом ОЗ'(см. Рис.4.2) может быть связана только с деформацией валентной связи СЗ'-ОЗ', так как силы, действующие на атом ОЗ', связанные с деформацией валентных углов, образованных атомами Х - СЗ'- ОЗ' (Х-любой атом). очевидно, направлены перпендикулярно к направлению валентной связи СЗ'-ОЗ'. Поэтому, для силы растяжения f, действующей на уровне фосфодиэфирной связи C3'-ОЗ' можно записать простое соотношение:

$$f = F_0 \cos(\alpha) \tag{4.1}$$

где  $\alpha$  – угол между направлением действия силы  $F_o$  и направлением связи. Чем меньше угол между связью C3'-O3' и осью спирали, тем большую величину составляет сила f, и, следовательно, тем выше вероятность гидролиза этой связи. Таким образом, для трансмиссионных коэффициентов в случае различных конформационных состояний N и S дезоксирибозы , можно записать:

$$\gamma_N = \gamma_0 \cos(\alpha_N),$$
  

$$\gamma_S = \gamma_0 \cos(\alpha_S)$$
(4.2),

где коэффициент у<sub>0</sub> описывает долю нагрузки, приходящейся на данную цепь остова.

Указанные в литературе процентные доли N – конформации дезоксирибозы дезоксицитидина в динуклеотиде d(CpG) составляют величину около 30 % [Wu et al, 2003].

Предположим, что при наличии растягивающего усилия, дезоксирибоза, примыкающая к цитозину, остается в том же состоянии, что и до него. В таком случае для полной скорости механохимической реакции гидролиза фосфодиэфирной связи в динуклеотиде d(CpG) можно записать:

$$k_{CG} = P_S k_S + P_N k_N \tag{4.3},$$

где  $k_{CG}$  – полная скорость механохимической реакции гидролиза по фосфодиэфирной связи, примыкающей к дезоксицитидину,  $P_S$  - вероятность пребывания дезоксирибозы дезоксицитидина в S – состоянии,  $k_S$  – скорость реакции гидролиза фосфодиэфирной связи, примыкающей к дезоксирибозе в S конформации, а  $P_N$  и  $k_N$  – аналогичные вероятность и скорость гидролиза для N состояния, соответственно. В соответствии с указанным выше предположением,  $P_S = 0.7$ ;  $P_N = 0.3$ .

Для скорости гидролиза центральной фосфодиэфирной связи в динуклеотиде d(ApA) можем по аналогии записать:

$$k_{AA} = k_S \tag{4.4},$$

так как для дезоксирибозы аденина *N*-конформации не было обнаружено. Используя экспериментальное соотношение  $k_{CG} \approx 1.6 k_{AA}$ , полученное из Таблицы 1 Главы 2, а также (4.3) и (4.4), получаем:  $k_N \approx 3 k_S$ 

С другой стороны, в соответствии с уравнением (3.10) и (4.2) для величин  $k_N$  и  $k_S$  можно записать:

$$k_{N} = \frac{\iint_{t,r} 4\pi r^{2} n_{0} \beta k_{0} e^{\frac{\gamma_{0} \cos(\alpha_{N})F(G(r,t))}{kT}} dt dr}{\int_{r} 4\pi r^{2} n_{0} dr}$$

$$k_{S} = \frac{\iint_{t,r} 4\pi r^{2} n_{0} \beta k_{0} e^{\frac{\gamma_{0} \cos(\alpha_{S})F(G(r,t))}{kT}} dt dr}{\int_{r} 4\pi r^{2} n_{0} dr}$$

$$(4.5)$$

Вычисление величин  $k_N$  и  $k_S$  на основе уравнения (4.5) при различных значениях параметра  $\gamma_0$  и значений углов, соответствующих кристаллографическим данным, дает величину отношения  $k_N / k_S > 10$ . Естественно предположить, что при высоких значениях растягивающего усилия, угол между осью спирали и направлением связи будет уменьшаться – в том числе и для S-конформации дезоксирибозы, - что может уменьшить теоретическое значение отношения  $k_N / k_S$ .

Также неясен вопрос о возможности конформационных переходов сахарного цикла, индуцированных действием растягивающего усилия – то есть правомерность использования для величин  $P_N$  и  $P_S$  значений, соответствующих экспериментальным данным, полученным при отсутствии действия внешней силы. Тем не менее, полученные результаты позволяют заключить, что предложенная модель качественно описывает эффект увеличения вероятности расщепления фосфодиэфирной связи при увеличении конформационной подвижности примыкающей к ней дезоксирибозы. Далее кратко подведены итоги раздела.

В данном разделе описана модель, предлагающая интерпретацию специфичности ультразвукового расщепления ДНК и основанная на результатах Главы 3 и раздела 4.1. Основная идея предложенной модели заключается в том, что при наличии растягивающего усилия, направленного вдоль фрагмента ДНК, уменьшение угла между направлением химической связи и осью спирали приводит к увеличению энергии деформации этой связи. В свою очередь, это приводит к уменьшению барьера а значит, к увеличению энергетического гидролиза, скорости механохимической реакции, приводящей к расщеплению фрагмента ДНК. Таким образом, при увеличении конформационной подвижности дезоксирибозы увеличивается время пребывания сахарного цикла в N конформации, при которой вероятность осуществления механохимической реакции выше, а значит относительная частота расщепления увеличивается.

71

#### 4.3. Обсуждение

Отправным пунктом предложенной модели является предположение о влиянии угла между осью спирали и направлением ковалентной связи, подверженной разрыву, на вероятность расщепления ДНК. Важно отметить, что данная гипотеза позволяет описать специфичность расщепления без дополнительных предположений об особенностях конформационной динамики дезоксирибозы. Угол между направлением фосфодиэфирной связи и вектором растягивающего усилия является динамической переменной, зависящей не только от состояния сахарного цикла, но и от других особенностей локальной конформационной динамики молекулы, определяемых последовательностью нуклеотидов. Важно отметить, что *S*↔*N* интерконверсия дезоксирибозы, по всей видимости, также имеет место и в дезокситимидине: в работе [Wu et al, 2003] доля *N*-конформациий для первой дезоксирибозы в d(TpT) динуклеотиде составила более 20%, в то время как значение средней относительной частоты ультразвукового расщепления для данного динуклеотида составляет небольшую величину - 0.932. По всей видимости, существенным является не только состояние дезоксирибозы, но и подвижность сахарофосфатного остова в целом. Это предположение подкрепляется тем, что рассматривая относительные частоты ультразвукового расщепления в динуклеотидах типа d(CpN) и d(TpN) (см. Рис.2.4) можно увидеть хорошо известную закономерность, связывающую локальную конформационную подвижность ДНК с пурин-пиримидиновой типом последовательности: если в направлении от 5' к 3' концу цепи вслед за пиримидином ( С или Т) следует пурин (А или G), то конформационная подвижность будет выше, чем если за пиримидином следовал бы пиримидин. В общем случае эта закономерность, как было отмечено в Главе 1, записывается следующим образом:

D(PyPu) > D(PuPu) > D(PuPy),

где под *D* понимается общая подвижность динуклеотидного блока - величина, пропорциональная объему конформационного пространства, Ру – пиримидин, а Ри – пурин. Последовательность РуРи отвечает динуклеотидному блоку ДНК, в котором в направлении от 5' к 3' концу цепи за пиримидином следует пурин. Если
просуммировать величины относительных частот ультразвукового расщепления  $\overline{R}$ , приведенные в Таблице 1 Главы 2, рассчитывая для каждого динуклеотидного блока сумму относительных частот расщепления комплементарных динуклеотидов, входящих в блок, то можно получить следующий результат:

$$\overline{R}$$
 (PyPu)= 2.32;

- $\overline{R}$  (PuPu)= 1.92;
- $\overline{R}$  (PuPy)= 1.86,

согласующийся с рядом D (PyPu) > D (PuPu) > D (PuPy).

Полученные величины относительных частот ультразвукового расщепления ДНК для 256 тетрануклеотидов также позволяют провести их сравнение с некоторыми экспериментальными данными, полученными при помощи ЯМР спектроскопии и Например, было молекулярной динамики. показано, что конформационная подвижность центрального блока d(CpG) выше в контексте тетрануклеотида d(TpCpGpA), чем в d(ApCpGpT) [Lefebvre et al, 1995]. Величины относительных частот ультразвукового расщепления центральной фосфодиэфирной связи В этих тетрануклеотидах составляют 1.543 и 1.381, соответственно (см. Таблицу 2 в Главе 2).

Как известно, зависимость локальных особенностей конформационной динамики ДНК от нуклеотидной последовательности может выходить за рамки тетрануклеотидного уровня описания. В качестве примера можно рассмотреть подвижность центрального d(CpG) динуклеотидного блока в гексануклеотидах d(GpApCpGpTpC) и d(ApApCpGpTpT), которая по данным ЯМР-спектроскопии совместно с молекулярно-динамическими расчетами в первом гексануклеотиде значительно выше, чем во втором [Cordier et al, 1999]. В проанализированной выборке по ультразвуковому расщеплению, данные гексануклеотиды встречаются 11 и 9 раз, соответственно, а средние величины относительной частоты расщепления центрального динуклеотида d(CpG) для данных контекстов составляют 1.61 и 1.32 соответственно, что также качественно согласуется с экспериментальными данными.

Таким образом, специфичность ультразвукового расщепления может отражать не только особенности конформационной динамики дезоксирибозы, но и общую подвижность сахарофосфатного остова ДНК, так как и то, и другое может существенным образом влиять на распределение растягивающей нагрузки по степеням свободы молекулы. Следует отметить, что с точки зрения основной гипотезы, обозначенной в начале данного раздела, также не важен механизм разрыва ковалентной связи (это может быть не только гидролиз, но и прямая диссоциация связи), и ее тип – то есть это может быть как C3'-O3', так и O3'-P связь.

Следует подчеркнуть, что в случае реакции гидролиза существенным фактором, определяющим скорость реакции, является доступность фосфодиэфирной связи для молекул растворителя. Величина доступной для растворителя эффективной поверхности фосфодиэфирной группы атомов влияет на энтропийный множитель в выражении для скорости химической реакции (3.5) и может зависеть от локальных конформационно-динамических свойств сахарофосфатного остова. Таким образом, при описании наблюдаемой специфичности расщепления важную роль также может играть вклад энтропийного фактора гидролиза фосфодиэфирной связи.

## Глава 5. Исследование особенностей ультразвукового расщепления функциональных областей ДНК

В данной главе рассмотрены экспериментальные данные ультразвукового расщепления регуляторных областей ДНК *λ*-фага, а также проведено исследование профилей ультразвукового расщепления, построенных для промоторов МҮС–кластера генома человека с помощью статистических данных по специфичности ультразвукового расщепления ДНК.

#### 5.1 Особенности ультразвукового расщепления ДНК λ-фага.

Как было показано в предыдущей главе, данные по ультразвуковому ДНК, расщеплению ПО всей видимости, отражают зависимость локальных конформационно-динамических свойств сахарофосфатного ДНК остова ОТ последовательности нуклеотидов, причем в некоторых случаях эта зависимость не может быть описана на уровне тетрануклеотидного приближения. Ниже приведен ультразвукового расщепления фрагмента ДНК, содержащего профиль блок регуляторной последовательности ДНК λ-фага.



**Рис.5.1.** Нормализованный профиль ультразвукового расщепления фрагмента ДНК, содержащего регуляторный участок последовательности ДНК λ-фага. Три операторных участка – OR1, OR2 и OR3, с которыми связываются белки – регуляторы транскрипции, отмечены серым цветом.

Наблюдаемая периодичность приведенного профиля ультразвукового расшепления, составляющая величину около 20 нуклеотидов, не может быть описана в рамках тетрануклеотидного приближения. Ниже приведен график, характеризующий отношение экспериментальных степеней расщепления (приведенных на Рис.5.1) к средним значениям относительных частот ультразвукового расщепления, полученным в рамках тетрануклеотидного уровня описания и приведенным в Таблице 2 Главы 2. Экспериментальное значение относительной частоты расщепления ДНК (массив этих значений представлен Рис.5.1) делилось на относительную величину на ультразвукового расщепления соответствующего тетрануклеотида:

$$R_{omh}(i) = \frac{R_{exp}(i)}{\overline{R}_{N_{i-1}N_iN_{i+1}N_{i+2}}}$$
(5.1)

где  $R_{exp}$  (*i*) – экспериментальная величина относительной частоты расщепления,  $\overline{R}_{N_{i-1}N_iN_{i+1}N_{i+2}}$  - средняя относительная частота ультразвукового расщепления центральной фосфодиэфирной связи тетрануклеотида  $N_{i-1}N_iN_{i+1}N_{i+2}$ . После вычисления массива значений по формуле (5.1), соответсвующий профиль был сглажен методом линейной фильтрации с окном в пять нуклеотидов. На Рис.5.2. представлены результаты расчета.



**Рис. 5.2.** Профиль, характеризующий отношение экспериментальной относительной частоты расщепления  $R_{exp}$  к ожидаемому значению относительной частоты ультразвукового расщепления  $\overline{R}_{N_{i-1}N_iN_{i+1}N_{i+2}}$ , полученному в рамках тетрануклеотидного приближения, в зависимости от номера фосфодиэфирной связи *i*. Приведенная нумерация соответствует порядку следования полос на Рис. 5.1.

Как видно из Рис.5.2, наличие периодичности в экспериментальном профиле расщепления не может быть описано при помощи статистических данных по расщеплению, полученных в рамках тетрануклеотидного приближения.

Можно предположить, что наблюдаемый профиль ультразвукового расщепления отражает некую структурную особенность операторного блока. В таком случае обнаруженная периодичность степени ультразвукового расщепления около 20 нуклеотидов может отражать специфику изменения определенных конформационнодинамических свойств вдоль данного участка ДНК. Стоит отметить, что данная периодичность не соответствует особенности расположения операторных участков: Операторный участок OR2 расположен на «расстоянии» в 22 нуклеотида от участка OR1, а OR3 – на «расстоянии» 23 нуклеотидов от OR2. Это хорошо видно из рисунка 5.1. – в котором первый минимум расщепления приходится на среднюю часть OR1, а последний – на начало OR3.

Альтернативное объяснение выявленной периодичности относительных частот ультразвукового расщепления связано с особенностями процесса расщепления ДНК. Согласно предложенной в Главе 4 гипотезе, основным фактором, влияющим на вероятность расщепления фосфодиэфирной связи, является ее ориентация по отношению к растягивающему усилию. Важно отметить, что для свободной ДНК в растворе угол между фосфодиэфирной связью и локальной осью двойной спирали является динамической переменной, а направление локальной оси ДНК может не совпадать с направлением глобальной оси двойной спирали, определяемой в масштабах десятков нуклеотидных пар. В качестве иллюстрации к сказанному выше далее приведена структура фрагмента ДНК λ - фага, построенная на основе известных контекстно-зависимых конформационных параметров ДНК, полученных при помощи анализа данных ЯМР (сервис представлен на сайте http://hydra.icgeb.trieste.it/dna/model it.html). Данная структура приведена в качестве иллюстрации, так как отражает некоторые особенности изменения структурных параметров вдоль ДНК, хорошо известные как из данных кристаллографии и ЯМР, так и молекулярно-динамических расчетов.



**Рис. 5.3.** Фрагмент структуры ДНК, построенной для последовательности λ-фага в рамках параметризации структурных параметров, полученных при помощи данных ЯМР. Приведенная структура демонстрирует изменение угла между локальной осью двойной спирали (направление которой можно считать перпендикулярным к плоскости, образованной парами оснований) и глобальной осью ДНК.

Ниже для одной из цепей полученной структуры приведен график зависимости угла *α* между направлением фосфодиэфирной связи C3'-O3' и главной осью двойной спирали ДНК от номера данной связи вдоль цепи.



**Рис.5.4**. Зависимость угла а между направлением фосфодиэфирной связи C3'-O3' и осью спирали от номера связи вдоль данной цепи сахарофосфатного остова.

Как видно из графика, данный угол изменяется с периодичностью в 10 нуклеотидов, что и следовало ожидать от структуры свободной ДНК.

При растяжении молекулы угол между фосфодиэфирной связью и силой растяжения влияет на степень деформации связи и, следовательно, вероятность её

разрыва. Поэтому, выявленная при анализе картины ультразвукового расщепления ДНК λ-фага периодичность величиной около 20 нуклеотидов может быть связана с изменением общей периодичности структуры ДНК – так как известно, что при наличии растягивающего усилия ДНК может частично расплетаться [Gore et al, 2006]. В таком случае период в 20 нуклеотидов может быть связан с изменением параметров двойной спирали ДНК при действии растягивающего усилия.

# 5.2. Применение данных по специфичности ультразвукового расщепления ДНК для анализа промоторных областей генома человека.

Как уже отмечалось в Главе 1, в настоящее время для решения задачи локализации функциональных областей в геномах все чаще используются данные о контекстной зависимости различных физико-химических свойств ДНК. Данный раздел посвящен исследованию возможности применения данных по специфичности расщепления ДНК ультразвуком для выявления особенностей промоторных областей генома человека, которые затем могут быть использованы в качестве дополнительного критерия при поиске функциональных сайтов в геноме человека.

В данном теоретического разделе для анализа промоторных последовательностей генома человека используются приведенные в Главе 2 данные о специфичности ультразвукового расщепления ДНК. Полученные величины относительных частот расщепления, как было показано в Главе 4, по всей видимости, отражают локальную конформационную подвижность сахарофосфатного остова ДНК: чем больше относительная частота ультразвукового расщепления данной фосфодиэфирной связи, тем более подвижным является сахарофосфатный остов в соответствующем участке молекулы.

Следует отметить, что полученные величины относительных частот ультразвукового расщепления позволяют различать особенности структурной динамики сахарофосфатного остова в каждой из цепей ДНК: в Главе 2 было показано, что величины относительной частоты расщепления фосфодиэфирных связей в комплементарных динуклеотидах могут значимо различаться (см. Раздел 2.3 Главы 2).

Поэтому, при исследовании геномных последовательностей появляется возможность анализировать конформационно-динамические особенности обеих цепей двойной спирали ДНК.

Используя значения, приведенные в таблице 2 Главы 2, были рассчитаны профили ультразвукового расщепления для последовательностей 648 промоторов генома человека (MYC кластер, взятый из базы данных Transcriptional Regulatory Element Database: http://rulai.cshl.edu/cgi-bin/TRED/tred.cgi?process=home).

Все последовательности изначально были выровнены: рассматривался блок длиной в 1000 нуклеотидов, содержащий 700 нуклеотидов до точки старта транскрипции и 300 – после.

Для каждой последовательности строился теоретический профиль ультразвукового расщепления на основании данных по относительным частотам расщепления тетрануклеотидов (Таблица 2 Главы 2), после чего полученные профили сглаживались методом линейной фильтрации с окном в пять нуклеотидов. Для построения профилей использовались обе цепи: вычислялась суммарная относительная частота расщепления по обеим цепям ДНК.

На Рис. 5.5 приведен профиль ультразвукового расщепления, полученный усреднением 648 профилей.



**Рис. 5.5.** Зависимость относительной частоты ультразвукового расщепления ( $\overline{R}$ ) от позиции соответствующей фософодиэфирной связи (n) по отношению к точке начала транскрипции (0). Приведен средний для 648 последовательностей профиль ультразвукового расщепления.

Из Рис.5.5 видно, что рассматриваемая характеристика имеет явную аномалию в районе старта транскрипции. Полученный профиль соответствует обнаруженным ранее результатам - возрастанием к месту сайта старта транскрипции частоты встречаемости цитозинов и гуанинов в последовательности [Abeel et al, 2008; Cao et al, 2008]) и является следствием увеличения относительной частоты ультразвукового расщепления фосфодиэфирных связей, следующих за дезоксицитидином.

Для того, чтобы нивелировать вклад этого фактора, была использована величина, которая характеризует влияние фланкирующих (то есть ближайших по цепи) нуклеотидов на ультразвуковое расщепление фосфодиэфирной связи в данном динуклеотиде:

$$S(n) = \frac{\overline{R}(N_{n-1}N_n * N_{n+1}N_{n+2}) - \overline{R}(N_n * N_{n+1})}{\overline{R}(N_{n-1}N_n * N_{n+1}N_{n+2})}$$
(5.2)

где  $\overline{R}(N_{n-1}N_n*N_{n+1}N_{n+2})$  – средняя относительная частота расщепления центральной фосфодиэфирной связи тетрануклеотида  $N_{n-1}N_nN_{n+1}N_{n+2}$ , а  $\overline{R}(N_n*N_{n+1})$  - средняя относительная частота расщепления динуклеотида  $N_nN_{n+1}$ . Звездочка обозначает условное положение разрыва.

При S(n) < 0 фланкирующие нуклеотиды приводят к уменьшению ультразвукового расщепления по сравнению со средним для данного динуклеотида значением и, следовательно, к уменьшению локальной подвижности сахарофосфатного остова, а при S(n) > 0 – наоборот, - к увеличению ультразвукового расщепления по сравнению со средним значением для данного динуклеотида, а значит, к увеличению локальной подвижности остова

На Рис.5.6 представлен усреднённый профиль, полученный аналогично первому рисунку, но для величины *S*(*n*), соответствующей кодирующей нити.



**Рис. 5.6.** Зависимость относительного изменения частоты ультразвукового расщепления (S) при переходе от динуклеотидного к тетрануклеотидному приближению от положения соответствующей фосфодиэфирной связи (n) по отношению к точке начала транскрипции (0). Приведен средний для 648 последовательностей профиль величины S(n).

Из Рис. 5.6 видно, что область до сайта старта транскрипции в среднем характеризуется величинами S(n) < 0, причем переход к более высоким значениям S(n) достигается на сравнительно небольшом участке – в области от -150 до 1 нуклеотида, после чего величины S(n) принимают значения, близкие к нулевому.

Полученный результат позволяет предположить, что в промоторных участках ДНК человека нуклеотиды эволюционно отобраны таким образом, что имеет место тенденция к уменьшению подвижности сахарофосфатного остова. Уменьшение конформационной подвижности промоторных участков ДНК может иметь своей целью увеличение доступности регуляторных сайтов для факторов, регулирующих экспрессию генов, так как по мере уменьшения подвижности сахарофосфатного остова ДНК, как известно, происходит уменьшение вероятности образования комплексов ДНК с гистонами [Pedersen et al, 1998].

Уменьшение гибкости и стабильности ДНК в промоторных участках было неоднократно подчеркнуто различными авторами и используется в пакетах программ по предсказанию сайтов старта транскрипции в геномных последовательностях [Florquin et al, 2005; Abeel et al, 2008; Dineen et al, 2009; Cao et al, 2008]. Как уже отмечалось, ключевым звеном многих алгоритмов является построение профилей различных физико-химических характеристик ДНК. Как правило, такие алгоритмы используют параметризацию, полученную в рамках динуклеотидного приближения.

Проведенный анализ свидетельствует о наличии «эффектов последовательности» в промоторных участках генома человека, проявляющихся на уровне тетрануклеотидного приближения, а также о возможности выявления этих эффектов при помощи полученных данных по расщеплению фрагментов ДНК ультразвуком.

#### Основные результаты и выводы

1. Получены средние относительные частоты ультразвукового расщепления ДНК, описывающие зависимость вероятности расщепления сахарофосфатного остова от последовательности нуклеотидных пар в рамках ди- и тетрануклеотидного приближений.

2. Характер изменения частоты расщепления фосфодиэфирных связей от их положения в цепи фрагмента, а также специфика зависимости общего уровня расщепления ДНК от физико-химических параметров облучаемого раствора свидетельствуют о механохимической природе наблюдаемых разрывов в фрагментах ДНК под действием ультразвука.

3. Развит подход к моделированию расщепления фрагментов ДНК под действием кавитационных эффектов в облучаемом растворе, на основе которого показано, что возникающие вблизи кавитационных пузырьков высокие градиенты скорости течения жидкости (порядка  $10^8 \text{ c}^{-1}$ ) способны приводить к разрыву фрагментов ДНК длиной в несколько сотен нуклеотидных пар.

4. Предложена модель, качественно описывающая явление контекстной специфичности расщепления ДНК ультразвуком. В соответствии с разработанной моделью, наблюдаемое увеличение частоты расщепления фосфодиэфирной связи, примыкающей к дезоксицитидину, связано с особенностью конформационной динамики дизоксирибозы этого нуклеотида.

5. При помощи полученных статистических даных по ультразвуковому расщеплению ДНК показано, что в промоторных областях генов МҮС- кластера человека, имеет место тенденция к уменьшению ультразвукового расщепления по сравнению с кодирующими областями.

#### Литература

- Abeel T., Saeys Y., Bonnet E., Rouzé P., Van de Peer Y. Generic eukaryotic core promoter prediction using structural features of DNA // Genome Research. 2008. Vol. 18. P. 310-23.
- Arora M., Ohl C.D., Lohse D. Effect of nuclei concentration on cavitation cluster dynamics. // J Acoust Soc Am. 2007. V.121(6). P.3432-6.
- Ashkin A., Dziedzic J.M., Bjorkholm J.E., Chu S. Observation of a single-beam gradient force optical trap for dielectric particles. // Optics Lett. 1986. V.11.P. 288-290.
- Bajic V.B., Tan S.L., Suzuki Y., Sugano S., Tan S.L., Suzuki Y., Sugano S., Suzuki Y., Sugano S., Sugano S. Promoter prediction analysis on the whole human genome // Nat. Biotechnol. 2004. V. 22. P.1467–1473.
- Basedow A. M. and Ebert. E.B. Ultrasonic degradation of polymers in solution. // Advances in Polymers Science. A. Abe, A.-C. Albertsson, ., J. Genzer, editors. Springer, Berlin/Heidelberg. 1977. V.22. P. 83–148.
- Becker N.B., Wolff L., Everaers R. Indirect readout: detection of optimized subsequences and calculation of relative binding affinities using different DNA elastic potentials // Nucleic Acids Res. 2006. Vol. 34. P. 5638-5649.
- Bensimon D., Simon A.J., Croquette V., Bensimon A. Stretching DNA with a receding meniscus: experiments and models. Phys Rev Lett. 1995. V.74. P. 4754-4757.
- Binnig G, Quate CG, Gerber C. Atomic force microscope. // Phys Rev Lett. 1986. V. 56. P.930-933.
- Brennen C.E. Cavitation and bubble dynamics. // Oxford University Press.1995.
- Brukner I., Jurukovski V., Savic A. Sequence-dependent structural variations of DNA revealed by DNase I.// Nucleic Acids Research. 1990.V. 18(4). P. 891–894.
- Brukner I., Sánchez R., Suck D., Pongor S. Sequence-dependent DNA bending as revealed by DNase I : Bending Parameters for trinucleotides.// EMBO Journal.1995.V. 14. P. 1812-1818.
- Bustamante C., Marko J.F., Siggia E.D., Smith S. Entropic elasticity of ambda-phage DNA. // Science. 1994. V. 265. P. 1599-1601.
- *Cao XQ, Zeng J, Yan H.* Structural property of regulatory elements in human promoters // Phys Rev E Stat Nonlin Soft Matter Phys. 2008. Vol. 77(4 Pt 1). P. 041908.
- Cluzel P, Lebrun A, Heller C, Lavery R, Viovy J-L, Chatenay D, Caron F. DNA: an extensible molecule.// Science.1996. V. 271. P.792-794.
- *Cordier C., Marcourt L., Dodin G.* Conformational variation of the central CG site in d(ATGACGTCAT)<sub>2</sub> and d(GAAAACGTTTTC)<sub>2</sub>. An NMR, molecular modelling and 3D-homology investigation. // Eur. J. Biochem. 1999.V. 261.P. 722–733.

- *Das R., Laederach A., Altman R.B.* SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. RNA. 2005.V.11. P.344–354.
- Dineen D.G., Wilm A., Cunningham P., Higgins D.G. High DNA melting temperature predicts transcription start site location in human and mouse // Nucleic Acids Res. 2009. Vol. 37(22). P.7360-7.
- *Doi M. and Edwards S.F.* The Theory of Polymer Dynamics // Clarendon Press, Oxford. 1986.
- *Duchardt E., Nilsson L. and Schleucher J.* Cytosine ribose flexibility in DNA: a combined NMR <sup>13</sup>C spin relaxation and molecular dynamics simulation study.// Nucleic Acids Res. 2008. V. 36. P. 4211–4219.
- Faiger H., Ivanchenko M., Cohen I., Haran T.E., Ivanchenko M., Cohen I., Haran T.E., Cohen I., Haran T.E., Haran T.E. TBP flanking sequences: Asymmetry of binding, long-range effects and consensus sequences // Nucleic Acids Res. 2006. Vol. 34. P. 104–119.
- *Fernandes M.X., Ortega A., López Martínez M.C., García de la Torre J.* Calculation of hydrodynamic properties of small nucleic acids from their atomic structure. //Nucleic Acids Res. 2002. V.30(8).P.1782-8
- Florquin K, Saeys Y, Degroeve S, Rouzé P, Van de Peer Y. Large-scale structural analysis of the core promoter in mammalian and plant genomes // Nucleic Acids Res. 2005. Vol. 33(13). P. 4255-64.
- Flynn H.G. Physics of acoustic cavitation in liquids.// In: Physical acoustics.1964. V.1-B.p.51.
- Foloppe N. and MacKerell A.D. Jr. Contribution of the phosphodiester backbone and glicosil linkage intrinsic torsional energetic to DNA structure and dynamics. // J. Phys. Chem. 1999. V.103. P. 10955–10964.
- Foloppe N. and MacKerell A. D. Jr. Intrinsic conformational properties of deoxyribonucleosides: implicated role for cytosine in the equilibrium among the A, B, and Z forms of DNA.// Biophys. J. 1999. V.76. P. 3206–3218.
- Gooberman G. Ultrasonic degradation of polystyrene. Part.1. A proposed mechanism for degradation. // J Polymer Sci. V.42. P.25
- Goñi JR, Pérez A, Torrents D, Orozco M. Determining promoter location based on DNA structure first-principles calculations.// Genome Biol. 2007. V.8(12). R263.
- Gore J., Bryant Z., Nöllmann M., Le M.U., Cozzarelli N.R., Bustamante C. DNA overwinds when stretched. // Nature. 2006. V.442(7104). P. 836-9.
- Grandbois M., Beyer M., Rief M., Clausen-Schaumann H., Gaub H.E. How strong is a covalent bond? // Science. 1999. V. 283. P. 1727-1730.
- Greenbaum J. A., Pang Bo, and Tullius T. D. Construction of a genome-scale structural map at single-nucleotide resolution // Genome Res. 2007. Vol. 17., P. 947–953

- Grokhovsky S.L., Il'icheva I.A., Nechipurenko D.Y., Golovkin M.V., Panchenko L.A., Polozov R.V. and Nechipurenko D.Y.. Sequence-Specific Ultrasonic Cleavage of DNA. // Biophysical Journal. 2011.V.100.P.117-125.
- *Grosberg A.Y., Khokhlov A.R.* Statistical Physics of Macromolecules.// Woodbury, New York: American Institute of Physics. 1994.
- Hakim H. B., Lindsay S. M. and Powell J. The speed of sound in DNA.// Biopolymers. 1984.V.23.P.1185-1192.
- Hansma H.G. Properties of biomolecules measured from atomicforce microscope images: a review. // J Struct Biol. 1997. V.119. P.99-108.
- Heddi B., Foloppe N.,..., Hartmann B. Quantification of DNA BI/BII backbone states in solution. Implications for DNA overall structure and recognition. // J. Am. Chem. Soc. 2006. V.128. P. 9170–9177.
- *Heddi B., Oguey C., ., Hartmann B.* Intrinsic flexibility of B-DNA: the experimental TRX scale. // Nucleic Acids Res. 2010. V.38. P. 1034–1047.
- Il'icheva I.A., Nechipurenko D.Y., Grokhovsky S.L. Ultrasonic cleavage of nicked DNA .// J Biomol Struct Dyn. 2009. Vol. 27(3). P. 391-8.
- *Isaacs R. J. and H. P. Spielmann.* NMR evidence for mechanical coupling of phosphate B(I)-B(II) transitions with deoxyribose conformational exchange in DNA.// J. Mol. Biol. 2001. V.311. P.149–160.
- *Koudelka G.* Recognition of DNA structure by 434 repressor.// Nucleic Acids Res. 1998. 26. P. 669–675.
- *Lee G.U., Chrisey L.A., Colton R.J.* Direct measurement of the forces between complementary strands of DNA. // Science. 1994. V.266. P.771-773.
- *Lefebvre A., Mauffret O., ..., Fermandjian S.* Structural behavior of the CpG step in two related oligonucleotides reflects its malleability in solution. // Biochemistry. 1995. V.34. P. 12019–12028.
- Lentz Y.K., Anchordoquy T.J., Lengsfeld C.S. DNA acts as a nucleation site for transient cavitation in the ultrasonic nebulizer. // J Pharm Sci. 2006. V. 95(3). P. 607-19.
- Liolios K., Tavernarakis N., Hugenholtz P., Kyrpides N.C., Tavernarakis N., Hugenholtz P., Kyrpides N.C., Hugenholtz P., Kyrpides N.C., Kyrpides N.C. The Genomes On Line Database (GOLD) v.2: A monitor of genome projects worldwide // Nucleic Acids Res. 2006. Vol. 34. P. D332–D334.
- Olson, W. K., A. A. Gorin, ., V. B. Zhurkin. DNA sequence dependent deformability deduced from protein-DNA crystal complexes. // Proc. Natl. Acad. Sci. USA. 1998. 95:11163–11168.
- *Packer M.J., Dauncey M.P., Hunter C.A.* Sequence-dependent DNA Structure: Tetranucleotide Conformational Maps.// Journal of Molecular Biology. V.295 (1). 2000. P. 85-103.
- Parker S.C., Hansen L., Abaan H.O., Tullius T.D., Margulies E.H. Local DNA topography correlates with functional noncoding regions of the human genome // Science. 2009. Vol. 324(5925). P. 389-92.
- Pedersen A.G., Baldi P., Chauvin Y., Brunak S. DNA Structure in Human RNA Polymerase II Promoters // J. Mol. Biol. 1998. Vol. 281. P. 663-673.

- Rohs R., West S.M., Liu P., Honig B. Nuance in the double-helix and its role in protein–DNA recognition // Curr. Opin. in Struct. Biol. 2009. Vol. 19. P. 171-177.
- Schmidt S. W., Beyer M.K. and Clausen-Schaumann H. Dynamic strength of the silicon-carbon bond observed over three decades of force-loading rates. // J. Am. Chem. Soc. 2008. V.130. P.3664–3668.
- Schramm, L., Hernandez, N. Recruitment of RNA polymerase III to its target promoters. Genes Dev. 2002. V. 16. P. 2593–2620.
- Smith S.B., Finzi L., Bustamante C.. Direct mechanical measurement of the elasticity of single DNA molecules by using magnetic beads. //Science. 1992.V. 258.P.1122-1126.
- Smith S.B., Cui Y., Bustamante C. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules.// Science. 1996.V. 271.P.795-799.
- Solovyev V., Kosarev P., Seledsov I., Vorobyev D., Kosarev P., Seledsov I., Vorobyev D., Seledsov I., Vorobyev D., Vorobyev D. Automatic annotation of eukaryotic genes, pseudogenes and promoters // Genome Biol. 2006. Vol. 7. P. 11–12.
- Sonnenburg S.O., Zien A., Rätsch G., Zien A., Rätsch G., Rätsch G. ARTS: Accurate recognition of transcription starts in human // Bioinformatics. 2006. Vol. 22. P. 472–480.
- Svoboda K., Block S.M. Biological applications of optical forces.//Annu Rev Biophys Biomol Struct. 1994. V. 23. P. 247-285.
- Thomas J. R. Sonic degradation of high polymers in solution. //J. Phys. Chem. 1959. V. 63. P.1725.
- Wang G., Zhang W., Zhang W. A steganalysis-based approach to comprehensive identification and characterization of functional regulatory elements // Genome Biol. 2006. Vol. 7. R49. doi: 10.1186/gb-2006-7-6-r49.

*Wang M.D., Yin H., Landick R., Gelles J., Block S.M.* Stretching DNA with optical tweezers.// Biophys. J. 1997. V. 72. P.1335-1346.

*Wu Z., Delaglio F., Tjandra N., Zhurkin V.B., Bax A.* Overall structure and sugar dynamics of a DNA dodecamer from homo- and heteronuclear dipolar couplings and <sup>31</sup>P chemical shift anisotropy.// J Biomol NMR. 2003. V. 26(4). P. 297-315.

*Бутягин П.Ю.*, Кинетика и природа механохимических реакций. // Успехи химии. 1971. т.40. с.1935-1959.

*Гроховский С.Л.* Специфичность расщепления ДНК ультразвуком // Мол. биология. 2006. т. 40, с. 317–325..

*Маргулис М.А.* Основы звукохимии. Химические реакцим в акустических полях. // Высшая школа. Москва. 1984.

Эльпинер И.Е. Биофизика ультразвука.// Серия: Физика жизненных процессов. М. Наука 197

#### Публикации по теме диссертации

- Нечипуренко Ю.Д., Полозов Р.В., Нечипуренко Д.Ю., Ильичева И.А., Воробьев Е.А., Гроховский С.Л. и Гурский Г.В. Математические модели регуляции экспрессии генов: механические возмущения структуры ДНК. // Математика. Компьютер. Образование: Сб. научных трудов. Том.2 / Под ред. Г.Ю.Ризниченко. -М.-Ижевск: НИЦ "Регулярная и хаотическая динамика". 2006, С. 392-402.
- Yury Nechipurenko, Sergey Grokhovsky, Georgy Gursky, Dmitry Nechipurenko and Robert Polozov. DNA-Based Nanostructures: Changes of Mechanical Properties of DNA upon Ligand Binding. *In collection: NATO Science for Peace and Security Series B: Physics and Biophysics, Nanomaterials for Application in Medicine and Biology*, Springer. 2008, P. 59-67.
- Гроховский С.Л., Ильичева И.А., Нечипуренко Д.Ю., Панченко Л.А., Полозов Р.В., Нечипуренко Ю.Д. Локальные неоднородности структуры и динамики двухспиральной ДНК: исследование при помощи ультразвука. Биофизика. 2008, T.53, C. 417-425.
- Ю.Д. Нечипуренко, М.В. Головкин, Д.Ю. Нечипуренко, И.А. Ильичева, Л.А. Панченко, Полозов Р.В., Гроховский С.Л. Характерные особенности расщепления ДНК ультразвуком. Ж. Структурной химии. 2009, Т. 50, № 5, С. 1040 – 1047.
- 5. A. Il'icheva, D. Yu. Nechipurenko, S. L. Grokhovsky Ultrasonic Cleavage of Nicked DNA J. of Biomolecular Structure & Dynamics. 2009, 27, № 3, P. 391-397.
- Головкин М.В., Нечипуренко Д.Ю., Ильичева И.А., Панченко Л.А., Полозов Р.В., Гроховский С.Л. Нечипуренко Ю.Д. Математические методы анализа электрофоретических картин расщепления ДНК. Компьютерные исследования и моделирование. 2009, Т.1, №3, С. 287–295.
- Sergei L. Grokhovsky, Irina A. Il'icheva, Dmitry Yu. Nechipurenko, Michail V. Golovkin, Larisa A. Panchenko, Robert V. Polozov and Yury D. Nechipurenko. Sequence-Specific Ultrasonic Cleavage of DNA. *Biophysical Journal*. 2011,V.100, P.117-125.

### Благодарности

Автор выражает благодарность к.х.н. Гроховскому С.Л., к.ф.-м.н. Ильичевой И.А., д.ф.-м.н. Полозову Р.В. и д.ф.-м.н. Нечипуренко Ю.Д. за внимание и постоянную помощь в работе, а также к.т.н. Панченко Л.А. за участие в проведении статистического анализа, д.ф.-м.н., профессору Твердислову В.А. и к.ф.-м.н. Есиповой Н.Г. за ценное обсуждение результатов работы.

Автор также благодарен своим родителям, бабушке и деду за доброту, понимание и заботу.