

**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
имени М.В. Ломоносова  
Механико-математический факультет**

На правах рукописи

Сабитов Денис Иджадович

**Механические модели пространственных структур  
молекул РНК.**

Специальность: 01.02.01 – Теоретическая механика

**АВТОРЕФЕРАТ**  
диссертации на соискание ученой степени  
кандидата физико-математических наук

Москва 2010

Работа выполнена на кафедре теоретической механики и мехатроники  
механико-математического факультета МГУ им. М.В. Ломоносова

Научный руководитель: доктор физико-математических  
наук Е.И. Кугушев

Официальные оппоненты: доктор физико-математических  
наук, профессор И.И. Косенко

кандидат физико-  
математических наук, доцент  
А.В. Родников

Ведущая организация: Вычислительный центр  
им. А.А. Дородницына  
Российской академии наук

Защита состоится 15 октября 2010 года в 16:30 часов на  
заседании специализированного совета Д 501.001.22 по механике при  
Московском государственном университете им. М.В. Ломоносова по  
адресу: 119992, Москва, Ленинские горы, МГУ, механико-  
математический факультет, аудитория 16-10.

С диссертацией можно ознакомиться в библиотеке механико-  
математического факультета МГУ (Главное здание, 14 этаж).

Автореферат разослан

2010 года.

Ученый секретарь диссертационного  
совета Д 501.001.22  
доцент

В.А. Прошкин

## Общая характеристика работы

**Актуальность работы.** В работе рассматриваются вопросы, связанные с развитием, изучением и применением тонкостержневой механической модели пространственных структур молекул рибонуклеиновых кислот (РНК). Исследование таких структур является одной из важных задач молекулярной биологии. Применение прямых (физико-химических) методов изучения таких структур является сложной и дорогостоящей задачей. Исследование пространственных структур кольцевых молекул дезоксирибонуклеиновых кислот (ДНК) показало, что использование механических моделей при изучении таких структур может не только упростить решение поставленных задач, но и позволить изучать структуры молекул в тех случаях, когда это невозможно сделать прямыми методами. Используемая для этого механическая модель (тонкий прямой однородный стержень, замкнутый в кольцо) не может быть использована при изучении пространственных форм молекул РНК, пространственная структура которых может содержать систему кинематических циклов с определенными топологическими характеристиками. Описание такой структуры возможно в рамках тонкостержневой механической модели как система тонких упругих стержней, соединенных по заданным правилам и принимающих равновесные пространственные формы. Применение таких новых механических моделей при изучении свойств сложных молекулярных структур является важной и актуальной задачей, решаемой в рамках теоретической механики.

**Цель работы.** Применение классической теории тонких упругих стержней для развития адекватной механической модели пространственных структур молекул РНК; изучение свойств тонкостержневой механической модели; применение этой модели для построения пространственных структур реальных молекул РНК.

**Научная новизна.** Все основные результаты, полученные в работе, являются новыми, ранее неизвестными. Они базируются на классических методах теоретической и прикладной механики. Среди новых результатов следует отметить развитие тонкостержневой механической модели пространственных структур РНК. В отличие от известных подходов, модель основана на использовании тонких упругих стержней общего вида. Разработан вариационный метод решения краевых задач для поиска равновесных состояний систем тонких упругих стержней с учетом топологических характеристик кинематических циклов. Аналитически показана корректность

тонкостержневой механической модели плоских форм кольцевых молекул ДНК. Решена задача идентификации геометрических параметров тонкостержневой механической модели по данным рентгеноструктурного анализа. Новым также является построение пространственных структур некоторых реальных молекул РНК.

**Теоретическая и практическая ценность.** Работа носит теоретико-прикладной характер, полученные в ней результаты дают возможность изучать свойства систем, состоящих из длинных упругих нитей с поперечными соединениями.

**Апробация работы и публикации.** Результаты, представленные в диссертации, докладывались автором и обсуждались на следующих научных семинарах и конференциях:

- Конференция, посвященная 90-летию со дня рождения Алексея Андреевича Ляпунова, Новосибирск, 2001 г.;
- Научная конференция Ломоносовские чтения МГУ им. М.В. Ломоносова, апрель 2002 г.;
- Семинар по динамике относительного движения кафедры теоретической механики и мехатроники МГУ под руководством чл.-корр. РАН В.В. Белецкого, проф. Ю.Ф. Голубева, доц. К.Е. Якимовой, доц. Е.В. Мелкумовой, 2002 г.;
- XII Международная конференция по вычислительной механике и современным прикладным программным системам. Владимир, 2003 г.;
- Конференция молодых ученых Института механики МГУ им. М.В. Ломоносова, Москва, 2004 г.;
- Пятый международный симпозиум по классической и небесной механике, Великие Луки, 23-28 августа 2004 г.;
- Международная школа-семинар по геометрии и анализу памяти Н.В. Ефимова, Абрау-Дюрсо, Ростов-на-Дону, 2004 г.;
- Семинар по аналитической механике и устойчивости движения кафедры теоретической механики и мехатроники МГУ под руководством академика В.В. Румянцева, чл.-корр. РАН В.В. Белецкого, проф. А.В. Карапетяна, 2006 г.;
- Семинар им. В.В. Румянцева по аналитической механике и устойчивости движения кафедры теоретической механики и мехатроники МГУ под руководством чл.-корр. РАН В.В. Белецкого, проф. А.В. Карапетяна, проф. Я.В. Татарина, 2009 г.;
- Семинар отдела механики ВЦ РАН под рук. проф. С.Я. Степанова, 2010 г.;

**Публикации.** Основные результаты диссертационной работы изложены в печатных работах, список которых приведен в конце автореферата.

**Структура работы.** Диссертационная работа состоит из введения, двух частей, в которые входят семь глав, заключения и списка литературы из 55 наименований. Общий объем диссертации – 127 страниц.

## Содержание работы

Во **введении** описана предметная область и цель диссертации – создание и применение механической модели пространственных структур рибонуклеиновых кислот (РНК), дан краткий обзор работ, относящихся к указанной тематике. Приведено краткое описание содержания диссертации. В конце введения перечислены основные результаты работы.

Диссертационная работа разбита на две части. Первая часть (главы 1-4) посвящена описанию механической модели пространственных структур молекул нуклеиновых кислот и аналитическому исследованию некоторых её свойств. Во второй части (главы 5-7) рассматриваются вопросы, связанные с применением описанной выше механической модели РНК для построения пространственных структур молекул. Важными вопросами здесь являются методы решения краевых задач при определении равновесных форм упругих стержней и методы определения параметров модели.

**Первая часть** диссертации состоит из четырех глав (главы 1-4).

В **первой главе** описываются основные понятия, относящиеся к изучению процессов структурообразования биополимерных макромолекул. Дается краткое описание методов определения вторичных структур РНК. Пространственная структура молекулы в тонкостержневой модели формируется на основе известной вторичной структуры. В завершении этой главы приводится общее описание тонкостержневой модели, с помощью которой можно строить и описывать пространственные структуры РНК и других биополимеров.

Молекула РНК представляет собой полинуклеотид – цепочку соединенных в линейной последовательности мономеров-нуклеотидов. В РНК нуклеотиды бывают 4х типов: аденин (А), гуанин (G), цитозин (С) и урацил (U). Нуклеотидная последовательность молекулы называется ее *первичной структурой*. Некоторые нуклеотиды в молекулярной цепочке связаны попарно т.н. Уотсон-

Криковскими связями. Структура этих связей называется *вторичной структурой* РНК. В силу конечности множества возможных Уотсон-Криковских связей число вторичных структур, которые может принимать данная молекула РНК, конечно, но весьма значительно. Определение реальной вторичной структуры РНК по ее известной первичной структуре является важной задачей, но в диссертации она не рассматривается. Под *третичной структурой* молекулы РНК (ДНК) понимается пространственная форма, которую принимает ее молекулярная цепочка в пространстве под воздействием Уотсон-Криковских и других более слабых потенциалов. Первичная структура у молекулы РНК одна, а возможных вторичных и третичных структур много.

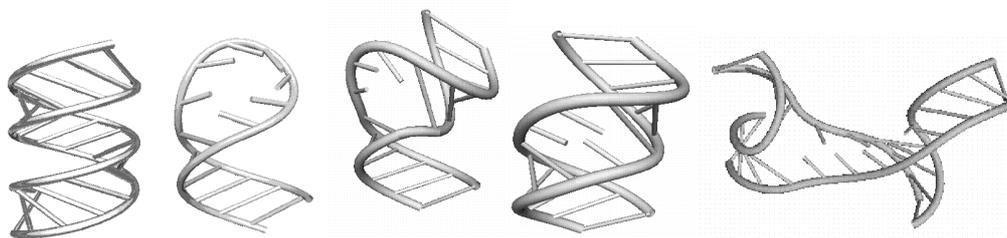


Рис 1.1. Элементы третичной структуры РНК (слева направо): стебель, шпильчатая петля, боковая петля, внутренняя петля, многозвенная петля.

С точки зрения вторичной структуры все нуклеотиды в молекуле можно разбить на два класса – спаренные (т.е. образующие Уотсон-Криковскую связь с каким-либо другим нуклеотидом) и свободные (неспаренные). Формально, вторичная структура РНК – это описание всех спаренных и свободных оснований в молекулярной цепи. Отрезок молекулярной цепи, состоящий из неспаренных оснований, называется *однонитевым*. Отрезки  $[n_1, n_2]$  и  $[n_3, n_4]$  спаренных оснований так, что  $n_1$  спарен с  $n_4$ ,  $n_1 + 1$  с  $n_4 - 1, \dots, n_2$  с  $n_3$ , образуют *двухнитевый* или *двуспиральный* участок во вторичной структуре РНК. Однонитевые участки (и семейства таких участков) называются *петлями*, а двухнитевые – *стеблями*. Стебель можно представить себе, как участок винтовой лестницы, где ступеньки – это поперечные Уотсон-Криковские связи (см. Рис. 1.1). Длиной стебля называется число пар оснований в нём:  $n_2 - n_1 + 1 = n_4 - n_3 + 1$ . Таким образом, вторичная структура РНК – это совокупность стеблей и петель.

**Вторая глава** посвящена изучению разнообразия третичных структур РНК. Пространственная структура молекулы РНК определяется ее вторичной структурой, причем разным вторичным

структурам соответствуют существенно отличающиеся пространственные структуры. В главе дается оценка различных вторичных структур РНК, которая также является оценкой снизу числа различных пространственных структур. Две молекулы, имеющие одну вторичную структуру, могут иметь различные пространственные структуры. В рамках тонкостержневой модели это объясняется тем, что в структуре могут присутствовать кинематически замкнутые циклы, состоящие из одного или нескольких стержней, и пространственные формы таких циклов в равновесном состоянии зависят от значения их топологических параметров (т.н. коэффициента зацепления цикла).

Будем называть длиной структуры количество Уотсон-Криковских связей в ней. Обозначим за  $V(N, n)$  среднее число структур длины  $n$  при случайном выборе молекулярной цепи, состоящей из  $N$  нуклеотидов. Рассмотрим структуру длины  $n$ . В ней какие-то  $2n$  нуклеотидов попарно связаны Уотсон-Криковскими связями. Всего возможно  $C_N^{2n}$  различных выборов  $2n$  нуклеотидов. После того, как нуклеотиды выбраны, вторичная структура определяется системой связей между ними. Перенумеруем выбранные нуклеотиды последовательно от 1 до  $2n$  и расположим их в вершинах правильного  $2n$ -угольника. Соединим связанные нуклеотиды диагоналями. Каждой вторичной структуре взаимно однозначно соответствует набор из  $n$  непересекающихся диагоналей. Обозначим за  $s(n)$  число различных расположений  $n$  непересекающихся диагоналей в правильном  $2n$ -угольнике, что отвечает рассмотрению вторичной структуры в виде хорд на окружности с выполнением стерического условия. Тогда (с учетом того, что вероятность комплементарности выбранных нуклеотидов равна  $\frac{1}{4^n}$ ) имеем

$$V(N, n) = \frac{1}{4^n} C_N^{2n} s(n) = \frac{1}{4^n} \frac{N!}{(2n)!(N-2n)!} s(n) \quad (2.1)$$

Величина  $s(n-1) = \frac{1}{2n-1} C_{2n-1}^{n-1}$  называется *числом Каталана*.

Показывается, что верна следующая оценка:

$$s(n) < ns^2(n-1) < n(n-1)^2 s^4(n-2) < \dots < \prod_{k=0}^{n-2} (n-k)^{2^k} < n^{\sum_{k=0}^{n-2} 2^k} = n^{2^{n-1}-1}$$

Поскольку  $s(n) > 2^{n-1}$ , то

$$V(N, n) = \frac{1}{4^n} C_N^{2n} s(n) > \frac{1}{2^{n+1}} \frac{N!}{(N-2n)!}, \quad (2.2)$$

причём степень полинома от  $N$ , стоящего в правой части, равна  $2n$ . Отсюда можно сделать два вывода: 1) с ростом молекулярной цепи количество возможных вторичных структур, состоящих из  $n$  Уотсон-Криковских связей, растет как  $N^{2n}$ , 2) с ростом молекулярной цепи количество всех возможных вторичных структур растет быстрее, чем любой полином (иначе говоря, скорость роста количества вторичных структур больше полиномиальной).

В завершение главы проводится оценка разнообразия структур двух классов малых РНК. Для малых ядерных РНК, длина которых  $\approx 30$  нуклеотидов, а  $n$  около 5, число возможных структур составляет не менее  $1.23232 \cdot 10^6$ . Для транспортных РНК средние значения длины молекулы  $N=75$  и количество Уотсон-Криковских связей составляет  $n=20$ . Из (2.2) получим:  $V(75,20) > 1.75679 \cdot 10^{19}$ .

В третьей главе дается формальное описание тонкостержневой механической модели третичной структуры молекул рибонуклеиновых кислот. Аналитически эта модель основана на использовании уравнений равновесия тонких упругих стержней. В полном виде уравнения выглядят следующим образом:

$$\begin{aligned} \frac{d}{ds} (B_1(\omega_1 - \omega_1^0)) + B_3\omega_2(\omega_3 - \omega_3^0) - B_2\omega_3(\omega_2 - \omega_2^0) &= 0 \\ \frac{d}{ds} (B_2(\omega_2 - \omega_2^0)) + B_1\omega_3(\omega_1 - \omega_1^0) - B_3\omega_1(\omega_3 - \omega_3^0) - F_3 &= 0 \\ \frac{d}{ds} (B_3(\omega_3 - \omega_3^0)) + B_2\omega_1(\omega_2 - \omega_2^0) - B_1\omega_2(\omega_1 - \omega_1^0) + F_2 &= 0 \end{aligned} \quad (3.1)$$

$$\begin{aligned} \frac{d}{ds} F_1 + \omega_2 F_3 - \omega_3 F_2 &= 0 \\ \frac{d}{ds} F_2 + \omega_3 F_1 - \omega_1 F_3 &= 0 \\ \frac{d}{ds} F_3 + \omega_1 F_2 - \omega_2 F_1 &= 0 \end{aligned} \quad (3.2)$$

$$\begin{aligned} \frac{d\vec{r}}{ds} &= \vec{e}_1 \\ \frac{d\vec{e}_1}{ds} &= \omega_3 \vec{e}_3 \times \vec{e}_1 - \omega_2 \vec{e}_3 \\ \frac{d\vec{e}_3}{ds} &= \omega_1 \vec{e}_1 \times \vec{e}_3 + \omega_2 \vec{e}_1 \end{aligned} \quad (3.3)$$

Здесь  $\vec{e}_1$  - касательный вектор к осевой линии стержня,  $\vec{e}_3 \perp \vec{e}_1$ ,  $\vec{e}_2 = \vec{e}_3 \times \vec{e}_1$  - образуют главный сопутствующий репер осевой линии стержня, совпадающий с репером Френе при повороте на угол  $\theta$  вокруг  $\vec{e}_1$ ;  $\vec{r}$  - радиус-вектор осевой линии стержня;  $B_1$  - крутильная жесткость,  $B_{2,3}$  - изгибные жесткости стержня;  $F_i$  - компоненты силы, действующей на левое поперечное сечение стержня в точке  $s$ ;  $\omega_i$  - вектор Дарбу осевой линии стержня в изогнутом состоянии;  $\omega_i^0$  - вектор Дарбу осевой линии стержня в свободном состоянии. В соответствии с аналогией Кирхгофа уравнения (3.1, 3.2) совпадают с уравнениями движения тяжелого твердого тела с неподвижной точкой (возможно, с добавлением гиростатов). Уравнения (3.3) описывают геометрическую форму осевой линии.

Третичные структуры РНК состояются из простых элементов (стебли и петли). Для нахождения формы этих элементов необходимо решение соответствующих краевых задач уравнений равновесия. Предполагается, что вторичная структура молекулы известна. Вторичная структура разбивается на основные элементы: стебли, петли различных типов, однонитевые незамкнутые участки. В результате такого разбиения у молекулярных нитей, входящих в состав каждого элемента, образуются незамкнутые концы. Для каждого элемента строится трехмерная модель. Стебель представляет собой двойную спираль, параметры которой совпадают с параметрами двойной спирали А-формы РНК. Форма петель определяется путем решения соответствующих краевых задач. При нахождении формы однонитевых незамкнутых участков считалось, что их форма соответствует свободному состоянию тонкого упругого стержня, моделирующего молекулу. Все незамкнутые концы молекулярных нитей указанных выше элементов структуры РНК имеют относительную ориентацию, соответствующую двойной спирали А-формы РНК. На завершающем этапе построения пространственной модели молекулы РНК найденные элементы ее пространственной структуры соединяются в соответствующих незамкнутых концах своих молекулярных нитей. В результате этого процесса все концы молекулярной цепочки, которые были разомкнуты в результате разбиения вторичной структуры на элементы, оказываются соединенными, и нити элементов образуют единую молекулярную цепь.

Показывается, что однородный, нейтральный биополимер в пространстве принимает форму винтовой линии (Рис. 3.1). Это является обоснованием того, что в тонкостержневой модели стержень

в свободном состоянии имеет форму винтовой линии. Приводятся уравнения равновесия тонкого упругого стержня общего вида, описываются краевые условия для системы уравнений равновесия, необходимые для получения элементов структур РНК.

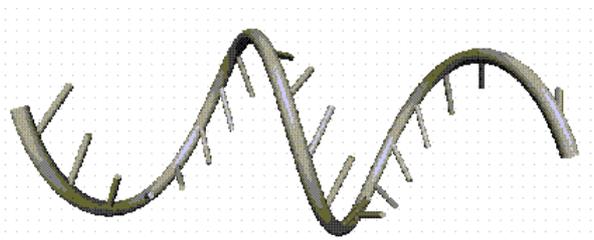


Рис. 3.1. Тонкий упругий стержень в свободном состоянии.

При заданных краевых условиях задача определения равновесной формы стержня может иметь более одного решения (даже счетное число решений). Геометрия этих форм может быть существенно различной, и, следовательно, биохимические свойства молекул будут также различны. Этот эффект возникает тогда, когда в структуре присутствуют кинематически замкнутые циклы, состоящие из одного или нескольких стержней, и пространственные формы таких циклов в равновесном состоянии зависят от значения их топологических параметров. Основными характеристиками, которые позволяют различать и классифицировать различные равновесные формы упругого стержня, моделирующего молекулу, являются числа вращения соответствующих элементов третичной структуры.

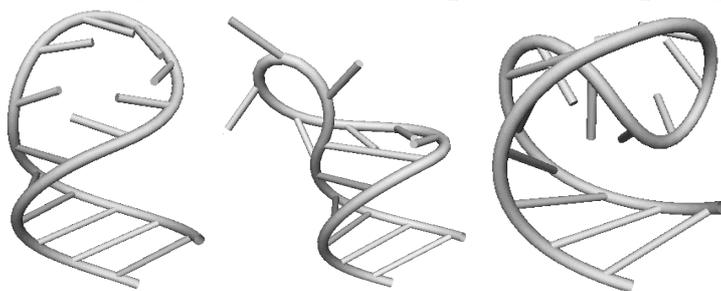


Рис. 3.2. Шпилечная петля: нейтральная форма, изменение формы петли при полном повороте краевой ориентации главного трехгранника по часовой стрелке и против часовой стрелки.

Дается краткое описание этих характеристик – различные решения краевой задачи отличаются *числом вращения*, т.е. числом оборотов правого краевого условия, переводящим найденное решение в некоторое выделенное нами нейтральное решение. Заметим, что в строгой формулировке надо говорить об относительных

коэффициентах зацепления найденных решений. Пример изменения пространственной формы шпилечной петли в зависимости от числа вращения представлен на Рис. 3.2.

В четвертой главе рассматривается вопрос о возможности моделирования пространственных форм ДНК с помощью прямолинейного тонкого упругого стержня. Как известно, молекула ДНК на сравнительно коротких своих участках имеет форму двух винтовых линий, закрученных относительно общей оси. При этом толщина молекулы и шаг витков малы по сравнению с общей её длиной.

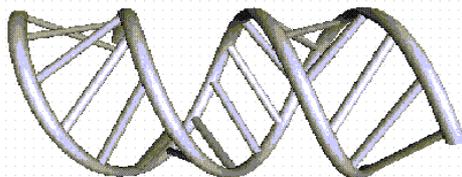


Рис. 4.1. Двухспиральный участок молекулы ДНК.

При моделировании пространственных форм длинных участков ДНК возникает вопрос: можем ли мы заменить сложную модель молекулы из двух винтовых линий гораздо более простой моделью, в которой длинная молекула представлена в виде прямолинейного стержня? В диссертации вопрос о применимости модели прямолинейного стержня рассматривается только в плоской постановке, когда молекула ДНК принимает плоские формы. ДНК представляет собой длинное нитевидное образование, или, точнее, двухнитевую структуру, в форме двойной спирали (см. Рис. 4.1). Иногда концы этой спирали замыкаются, и нить ДНК склеивается в кольцо. При математическом анализе возможных пространственных форм ДНК наиболее распространенной является модель, в которой нить ДНК представляются однородным тонким упругим стержнем, изгибающимся под действием внешних сил. В таких работах принято считать, что упругий стержень в свободном состоянии является прямолинейным. Стержень считается однородным (его упругие характеристики не изменяются вдоль длины), и изгибные жесткости одинаковы во всех направлениях. В то же время реальная молекула ДНК больше напоминает нить, которая в свободном состоянии принимает форму винтовой линии, ее упругие характеристики (в первом приближении) меняются периодически, с частотой, соответствующей длине витка спирали. Однако, исследование подобной модели намного сложнее, поэтому использование упрощенной модели более предпочтительно.

По сравнению с длиной молекулы (до нескольких сантиметров)

шаг витка ( $3 - 4 \text{ \AA}$ ) и толщина ДНК ( $20 \text{ \AA}$ ) незначительны. Поэтому использование упрощенной прямолинейной модели считается допустимым. В диссертации рассматривается вопрос о ее применимости. Проблема ставится следующим образом. Допустим, у нас есть тонкий упругий стержень общего вида, характеристики которого очень быстро периодически меняются вдоль его оси. Рассмотрим тонкий прямолинейный стержень с осредненными характеристиками. Будут ли его равновесные формы близки к равновесным формам исходного упругого стержня? В терминах пространственных структур ДНК постановка такова. Мы берем в качестве модели ДНК спиралевидный стержень с малой толщиной спирали. При этом стержень находится под действием каких-то сил и моментов. Устремляем толщину спирали к нулю, но сохраняем при этом геометрическую форму спирали и действующие нагрузки, и пытаемся показать, что форма такого спиралевидного стержня будет стремиться к форме прямолинейного стержня при тех же нагрузках. То, что в пределе форма одной модели стремится к форме другой, будет означать, что использование прямолинейного стержня при изучении пространственных форм ДНК является корректным.

Указанная проблема рассматривается в простейшей, плоской постановке, когда ДНК принимает только плоские формы. В этом случае структура молекулы описывается как равновесная форма плоского тонкого упругого стержня, к концам которого приложена нагрузка. Геометрия стержня в свободном (ненапряженном) состоянии периодически повторяется. Показано, что с ростом частоты повторения геометрии свободного стержня до бесконечности равновесные формы, которые он принимает, стремятся к равновесным формам тонкого прямолинейного стержня при некоторой нагрузке. Для полноты изложения приводятся общие сведения из теории плоских упругих стержней, а также необходимые утверждения из теории слабой сходимости функций. Доказывается, что использование модели тонкого прямолинейного стержня для плоского случая корректно, поскольку форма криволинейного стержня под нагрузкой стремится к форме прямолинейного стержня при устремлении частоты повторения геометрии свободного стержня к бесконечности. При этом жесткость предельного стержня  $A^*$  связана с исходной жесткостью  $A$  соотношением  $A^* = \frac{Ax_1(T)}{T}$ , где  $T$  – длина исходного стержня на периоде формы,  $x_1(T)$  – длина проекции периода формы на ось  $Ox$ , вдоль которой располагается стержень в свободном состоянии. В качестве примера возьмем стержень, имеющий форму

синусоиды:  $x(t) = t$ ,  $y(t) = \sin(t)$ . В этом случае

$$A^* = \frac{2\pi \cdot A}{\int_0^{2\pi} \sqrt{1 + \cos^2(s)} ds} \approx 0.82236 \cdot A.$$

В завершающем разделе на основе аналогии Кирхгофа полученные результаты интерпретируются в терминах движения математического маятника, к которому приложено быстро меняющийся периодический момент. Результат этого раздела можно трактовать следующим образом. На конечном интервале времени влияние высокочастотного возмущающего момента на движение математического маятника эквивалентно изменению начальной скорости.

**Вторая часть** диссертации состоит из трёх глав (главы 5 – 7). В ней рассматриваются вопросы применения тонкостержневой модели к построению пространственных структур реальных молекул РНК.

В **пятой главе** рассматриваются краевые задачи, которые появляются при построении пространственных структур РНК, и описывается метод их решения. При решении краевой задачи нужно учитывать два фактора. Во-первых, решение должно удовлетворять краевым условиям, и, во-вторых, полученная форма должна иметь заданные топологические характеристики. В простой постановке учет этих условий может происходить поэтапно: первый этап – нахождение приближенного решения, удовлетворяющего поставленным краевым условиям, второй этап – коррекция полученного решения для достижения нужных значений топологических характеристик равновесной формы, третий этап – уточнение приближенного решения для достижения необходимой точности. В комплексной и более эффективной вариационной постановке учет всех этих условий можно производить сразу при построении пространственной формы молекулы. Данный раздел посвящен описанию задач, связанных с указанными вопросами.

В качестве примера постановки краевой задачи рассматривается шпилечная петля (см. Рис. 5.1). Тонкий упругий стержень, моделирующий шпилечную петлю, является гладким продолжением одной из нитей двуспирального участка и гладко переходит другим концом в другую его нить. Поэтому начальное положение и ориентация стержня совпадают с конечным положением и ориентацией первой нити двуспирального участка, а конечное положение и ориентация с начальным положением и ориентацией второй нити. Параметры двуспирального участка нам известны – это параметры А-формы двойной спирали.

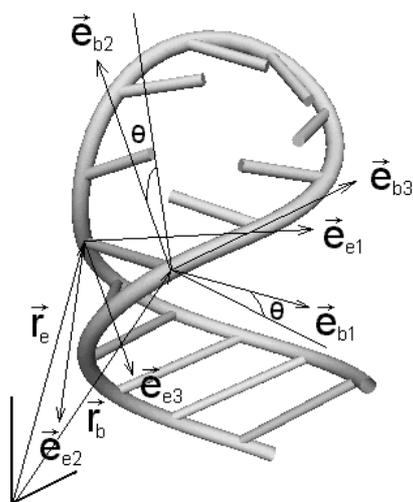


Рис. 5.1. Краевая задача для шпилечной петли.

Поясним обозначения:  $\vec{r}$  – радиус-вектор осевой линии стержня,  $\vec{r}(0) = \vec{r}_b$ ,  $\vec{r}(L) = \vec{r}_e$ ;  $\vec{e}_i$  – векторы главных сопутствующих осей,  $\vec{e}_1(0) = \vec{e}_{b1}$ ,  $\vec{e}_1(L) = \vec{e}_{e1}$ ,  $\vec{e}_3(0) = \vec{e}_{b3}$ ,  $\vec{e}_3(L) = \vec{e}_{e3}$ . Угол  $\theta$  – это угол поворота главных сопутствующих осей относительно осей естественного трехгранника Френе. Индексы  $b$  и  $e$  соответствуют начальной и конечной точкам шпилечной петли.

В **шестой главе** рассматриваются вопросы, связанные с идентификацией параметров модели третичных структур РНК. При этом решается обратная задача: по имеющимся данным о пространственных структурах некоторых молекул РНК, полученных физическими методами, подбираются значения параметров модели для более точного построения пространственных форм молекул.

Качество предсказания пространственной структуры молекулы РНК напрямую зависит от правильного определения параметров механической модели. В этой главе рассматриваются вопросы, связанные с идентификацией параметров механической модели третичных структур РНК. Для этого решается обратная задача: по имеющимся данным о пространственных структурах некоторых молекул РНК, полученных физическими методами, определяются значения параметров модели для более точного построения пространственных форм молекул.

Перечислим параметры тонкостержневой механической модели.

Упругие параметры стержня: жесткости  $B_i$  и угол  $\theta$  поворота главного сопутствующего репера по отношению к трехграннику Френе;

Геометрические параметры стержня для однонитевых участков, определяющие его форму в свободном состоянии: вектор Дарбу  $(\omega_1^0, \omega_2^0, \omega_3^0)$ ;

Три координаты вектора  $\vec{u}$  Уотсон-Криковской связи в трехграннике Френе и параметры  $\phi$  и  $\Delta Z$ , которые описывают положение второй винтовой нити двухспирального участка по отношению к первой. Вторая нить получается из первой поворотом вокруг оси винтовой линии на угол  $\phi$  и сдвигом по этой же оси на расстояние  $\Delta Z$ ;

Геометрические параметры стержня для двухспиральных участков: вектор Дарбу  $(\omega_1^d, \omega_2^d, \omega_3^d)$ .

Отметим, что параметры  $\phi$  и  $\Delta Z$  можно вычислить, используя вектор Уотсон-Криковской связи  $\vec{u}$ , вектор направления оси двухспирального участка  $\vec{e}$  и радиус  $r$  винтовой линии двухспирального участка. Величина  $\Delta Z$  есть проекция вектора  $\vec{u}$  на вектор  $\vec{e}$ . Для определения угла  $\phi$  найдем проекцию вектора  $\vec{u}$  на плоскость с нормалью  $\vec{e}$ :

$$l = \sqrt{|\vec{u}|^2 - \Delta Z^2} \quad (6.1)$$

и воспользуемся теоремой косинусов:

$$\cos \phi = \frac{2r^2 - l^2}{2r^2}. \quad (6.2)$$

Параметры модели разбиваются на несколько групп. Это геометрические параметры стержня, моделирующего молекулярную цепь, его упругие характеристики, геометрические параметры жестких перемычек, моделирующих Уотсон-Криковскую связь, и геометрические параметры двухспиральных участков. Задача идентификации параметров модели естественным образом разделяется на два этапа. Первый этап – это идентификация геометрических параметров двухспиральных участков и жестких перемычек, моделирующих Уотсон-Криковскую связь. Вторым этапом – это идентификация геометрических и упругих параметров стержня, моделирующего молекулярную цепь, с использованием результатов, полученных на первом этапе. В диссертации представляются результаты, в основном относящиеся к первому этапу задачи идентификации. Основная идея идентификации параметров состоит в использовании опубликованной информации о третичных структурах некоторых реальных молекул транспортных РНК, полученной методом рентгеноструктурного анализа. Эти данные включают, в частности, координаты атомов рибозофосфатных остовов обеих

нитей, составляющих двуспиральные участки стеблей тРНК. Идентифицируемые параметры подбираются таким образом, чтобы нити модельного двуспирального участка как можно лучше приближались к нитям рибозофосфатных остовов двуспирального участка реальной молекулы.

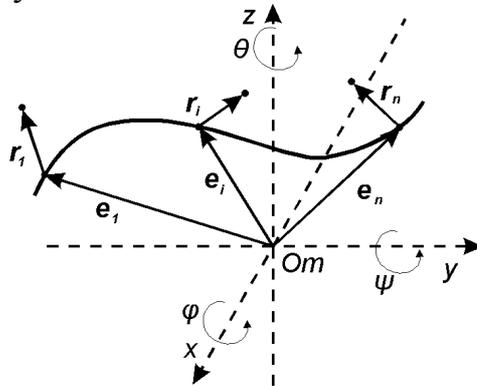


Рис. 6.1. Предварительное совмещение кривых.

Изложим теперь кратко содержание этой главы. Материал, относящийся к задаче идентификации параметров, разбит на три раздела. В первом разделе дается краткое описание используемой математической модели и ее параметров, подлежащих идентификации. Рассказывается о постановке задачи идентификации, дается описание данных рентгеноструктурного анализа молекул тРНК, называются основные шаги процесса идентификации. Далее описывается генетический алгоритм поиска экстремумов и способ построения хромосомы для задачи идентификации. Также рассматривается задача сравнения двух ломаных с различным числом звеньев и три способа ее решения. В третьем разделе описываются результаты идентификации геометрических параметров стержня. В задаче сопоставления геометрических форм пространственных кривых используется квазимеханический алгоритм сравнения. Для приведения двух ломаных в наилучшее относительное расположение (см. Рис. 6.1) решается система неявных разностных уравнений:

$$\begin{cases} \bar{x}^{i+1} = \bar{x}^i + \tau \cdot \bar{N}(\bar{x}^{i+1}, \bar{\phi}^{i+1}) \\ \bar{\phi}^{i+1} = \bar{\phi}^i + \tau \cdot \bar{M}(\bar{x}^{i+1}, \bar{\phi}^{i+1}) \end{cases}$$

где  $\bar{N}(\bar{x}, \bar{\phi}) = k \sum_{i=1}^n \frac{\bar{r}_i(\bar{x}, \bar{\phi})}{|\bar{r}_i(\bar{x}, \bar{\phi})|}$ ,  $\bar{M}(\bar{x}, \bar{\phi}) = l \sum_{i=1}^n \bar{e}_i(\bar{x}, \bar{\phi}) \times \bar{r}_i(\bar{x}, \bar{\phi})$ ,  $\bar{x} = (cx, cy, cz)$

- координаты центра масс кривой,  $\bar{\phi} = (\phi, \varphi, \theta)$  - углы поворота кривой вокруг осей Кенига,  $k, l$  - безразмерные коэффициенты, влияющие на скорость и характер сходимости процесса. Процесс минимизирует

функционал  $k \sum_{i=1}^n d_i$ , где  $d_i$  – расстояние от  $i$ -й ближайшей точки до соответствующей ей точки набора.

Функционал может иметь не одну точку экстремума. Следует ожидать, что их может быть достаточно много. Поэтому в строгом математическом смысле задача не является корректной (т. е. ее решение не единственно). Однако, наличие минимума у оценочного функционала можно гарантировать, т. к. он ограничен снизу на компактном пространстве параметров. Поэтому осмысленной является задача изменения геометрических параметров модели таким образом, чтобы оценочный функционал стал меньше при разумной форме и взаимном расположении сопоставляемых кривых в точке экстремума. В силу сказанного результаты процесса идентификации необходимо контролировать для того, чтобы отбрасывать заведомо неподходящие варианты. В разделе приводятся результаты процессов идентификации параметров по рентгеноструктурным моделям РНК (пример приведен на Рис. 6.2).



Рис. 6.2. Упругий стержень, полученный в результате идентификации параметров, аппроксимирует участки из 5 и 9 нуклеотидов. Атомы фосфора рентгеноструктурной модели обозначены сферами.

В качестве параметров двухнитевого участка молекулы РНК, помимо вектора Дарбу  $(\omega_1^d, \omega_2^d, \omega_3^d)$ , необходимо иметь также угол  $\phi$  и смещение  $\Delta Z$ .

При построении моделей однонитевых участков молекулы РНК, каковыми являются петли различных типов, возникает необходимость в решении краевых задач для системы уравнений (3.1-3.3). Знание параметров Уотсон-Криковской связи необходимо для вычисления граничных условий этих краевых задач. Однако, на практике в этих двух случаях удобнее использовать непосредственно параметры  $\Delta Z$  и  $\phi$ . С помощью несложных геометрических построений можно

получить из них и координаты вектора Уотсон-Криковской связи в главном сопутствующем репере стержня.

Из рентгеноструктурных данных о третичной структуре молекулы РНК нам известны, в частности, координаты атомов фосфора, находящихся в двуспиральном участке молекулы. Модель двуспирального участка представляет собой две винтовые линии с общей осью. Для поиска этой оси на заданном участке строится цилиндр, такой, чтобы сумма квадратов расстояний от атомов фосфора до поверхности цилиндра была минимальна. Осевая линия этого цилиндра принимается за осевую линию двуспирального участка.

При поиске параметров  $\Delta Z$  и  $\phi$  подсчитывается среднее арифметическое проекций векторов Уотсон-Криковских связей в двуспиральном участке на осевую линию участка и определяется смещение  $\Delta Z$ . Для вычисления угла  $\phi$  требуется среднее арифметическое длин Уотсон-Криковских связей в стебле  $\bar{l}$ . Подставив в формулу (6.1) в качестве  $|\vec{u}|$  среднюю длину  $\bar{l}$ , получим среднюю длину проекции связи на плоскость, перпендикулярную оси цилиндра. Зная радиус цилиндра, вычислим значение  $\phi$ , используя формулу (6.2).

Параметры механической модели, полученные в результате описанного в этой главе метода, были использованы для построения пространственных структур реальных молекул тРНК.

В **седьмой главе** описывается алгоритм построения пространственных структур молекул РНК по известной вторичной структуре.

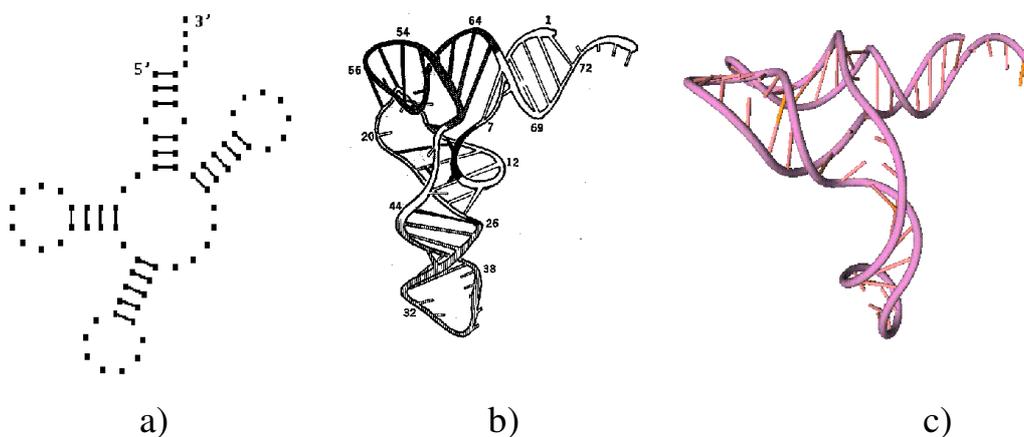


Рис. 7.1. Пространственная структура молекулы тРНК Phenilalanine Yeast. а) Вторичная структура, б) Рентгеноструктурный анализ. с) Компьютерные вычисления на основе тонкостержневой модели.

Приводится краткое описание процесса построения пространственной структуры РНК на основе механической модели, описанной в предыдущих главах. Предполагается, что вторичная структура молекулы известна. Вторичная структура разбивается на основные элементы: стебли, петли различных типов, однонитевые незамкнутые участки. В результате такого разбиения у молекулярных нитей, входящих в состав каждого элемента, образуются незамкнутые концы. Для каждого элемента при помощи алгоритма, описанного в предыдущих главах, строится трехмерная модель. Стебель представляет собой двойную спираль, параметры которой совпадают с параметрами двойной спирали А-формы РНК. Форма петель определяется путем решения соответствующих краевых задач. Из физических соображений при решении этих краевых задач задавались нулевые числа вращения. При нахождении формы однонитевых незамкнутых участков считалось, что их форма соответствует свободному состоянию тонкого упругого стержня, моделирующего молекулу. Все незамкнутые концы молекулярных нитей указанных выше элементов структуры РНК имеют относительную ориентацию, соответствующую двойной спирали А-формы РНК. На завершающем этапе построения пространственной модели молекулы РНК найденные элементы ее пространственной структуры соединяются в соответствующих незамкнутых концах своих молекулярных нитей. В результате этого процесса все концы молекулярной цепочки, которые были разомкнуты в результате разбиения вторичной структуры на элементы, оказываются соединенными, и нити элементов образуют единую молекулярную цепь.

В конце раздела приводятся примеры структур реальных молекул РНК, полученные на основе тонкостержневой механической модели. Несколько примеров построенных структур даны на Рис. 7.1 и 7.2.

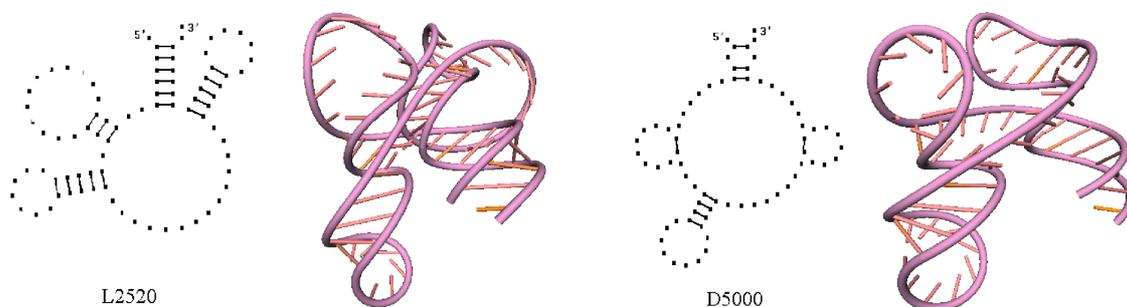


Рис. 7.2. Вторичные и третичные структуры тРНК: Аспарагиновая кислота *Asterina Pectini* и Лейцин *Euglena Gracilis*.

**В заключении** приведены основные результаты и выводы:

- Тонкостержневая механическая модель модифицирована для учета топологических характеристик элементов пространственных структур РНК, содержащих кинематически замкнутые циклы. Разработаны вариационный (прямой) и трехэтапный методы решения краевых задач для поиска равновесных состояний тонких упругих стержней при построении пространственных структур молекул РНК с учетом топологических характеристик их элементов.
- Аналитически решен вопрос о корректности применения модели прямо-линейного тонкого упругого стержня для изучения плоских форм кольцевых молекул ДНК. Получена оценка разнообразия потенциально возможных пространственных структур РНК.
- Решена задача идентификации геометрических параметров тонкостержневой механической модели по данным о пространственных структурах реальных молекул РНК, полученных физическими методами.
- На основе тонкостержневой механической модели найдены пространственные структуры реальных молекул тРНК. Для тех молекул, реальные структуры которых были получены ранее физическими методами, тонкостержневые модели оказались близкими по форме и повторяют все их основные топологические свойства.

**По теме диссертации опубликованы следующие работы:**

1. Козлов Н.Н, Кугушев Е.И., Сабитов Д.И., Энеев Т.М. Компьютерный анализ процессов структурообразования нуклеиновых кислот. Препринт ИПМ им. М.В. Келдыша РАН, 2002, N 42.
2. Сабитов Д. И. Об идентификации параметров математической модели пространственной структуры РНК. Препринт ИПМ им. М.В. Келдыша РАН, 2004, N 19.
3. Сабитов Д.И. Коэффициент зацепления как характеристика элементов пространственных структур молекул РНК. Труды Международной школы-семинара по геометрии и анализу памяти Н.В. Ефимова. Абрау-Дюрсо, Ростов-на-Дону, 2004, с.222-223.
4. Кугушев Е.И., Сабитов Д.И. О тонких упругих стержнях с быстро меняющимися периодическими характеристиками. Препринт Института прикладной математики им. М.В. Келдыша РАН, N 33, 2005, с. 28.
5. Кугушев Е.И., Сабитов Д.И. О плоских тонких упругих стержнях с быстроменяющимися периодическими характеристиками. Вестник МГУ, сер. мат.-мех., 2009, N 4, с. 42-48