

**НАЦИОНАЛЬНАЯ АКАДЕМИЯ НАУК УКРАИНЫ
ИНСТИТУТ МОЛЕКУЛЯРНОЙ БИОЛОГИИ И ГЕНЕТИКИ**

На правах рукописи

ЗАРУДНАЯ МАРГАРИТА ИВАНОВНА

УДК 577.21

**СТРУКТУРНЫЕ ПЕРЕХОДЫ В ГОМОПОЛИРИБОНУКЛЕОТИДАХ И
ИХ БИОЛОГИЧЕСКОЕ ЗНАЧЕНИЕ**

03.00.03 - молекулярная биология

ДИССЕРТАЦИЯ

на соискание ученой степени

доктора биологических наук

Научный консультант:

Говорун Дмитрий Николаевич,

доктор биологических наук, профессор,

член-корреспондент НАН Украины

КИЕВ - 2015

СОДЕРЖАНИЕ

ПЕРЕЧЕНЬ УСЛОВНЫХ СОКРАЩЕНИЙ	6
ВВЕДЕНИЕ	8
РАЗДЕЛ 1. ОБЗОР ЛИТЕРАТУРЫ	16
1.1. Структурные переходы в гомополирибонуклеотидах	16
1.1.1. Поли(А)	16
1.1.2. Поли(С)	20
1.1.3. Поли(Г), поли(У)	24
1.2. Гомополи(олиго)рибонуклеотиды в природных РНК	27
1.2.1. Полиаденилирование про-мРНК высших эукариотов	27
1.2.2. Полиаденилирование про-мРНК дрожжей, растений и вирусов	33
1.2.2.1. Дрожжи	33
1.2.2.2. Растения	35
1.2.2.3. Вирусы высших эукариотов	38
1.2.3. Тракты поли/олиго(С) и олиго(У)	44
1.2.4. Тракты олиго(Г)	47
1.2.5. G-квадруплексы РНК	49
1.3. Структура области димеризации геномной РНК ВИЧ-1	53
1.4. Выводы к разделу 1	58
РАЗДЕЛ 2. МАТЕРИАЛЫ И МЕТОДЫ	60
2.1. Измерение осмотического давления ДНК и определение эффективного диаметра двойной спирали	60
2.1.1. Препараты	60
2.1.2. Осмометр	61
2.1.3. Вычисление величины эффективного диаметра ДНК	64
2.2. Электрофорез гомополирибонуклеотидов	66
2.2.1. Проведение электрофореза	66

2.2.2. Определение электрофоретической подвижности и коэффициента задержки	68
2.3. Потенциометрическое титрование гомополирибонуклеотидов и расчет кривых протонной буферной емкости	70
2.3.1. Потенциометрическое титрование	70
2.3.2. Метод протонной буферной емкости	70
2.4. Предсказание вторичной структуры РНК	72
2.4.1. Последовательности РНК и используемые программы	72
2.4.2. Определение вторичной структуры фрагментов РНК	74
РАЗДЕЛ 3. ВЛИЯНИЕ ИОННОЙ СИЛЫ РАСТВОРА НА ЭФФЕКТИВНЫЕ РАЗМЕРЫ ПОЛИНУКЛЕОТИДОВ	77
3.1. Зависимость осмотического давления растворов ДНК от ионной силы раствора	78
3.2. Определение эффективного диаметра ДНК	83
3.3. Выводы к разделу 3	87
РАЗДЕЛ 4. ИССЛЕДОВАНИЕ СТРУКТУРНЫХ ПЕРЕХОДОВ В ГОМОПОЛИМЕРАХ МЕТОДАМИ ЭЛЕКТРОФОРЕЗА И ПРОТОННОЙ БУФЕРНОЙ ЕМКОСТИ	88
4.1. Зависимость электрофоретической подвижности гомополирибонуклеотидов от условий среды	88
4.2. Структурные переходы в поли(А)	100
4.2.1. Электрофоретическое исследование двухспиральных форм поли(А)	100
4.2.2. Электрофоретическое исследование «замороженной» формы поли(А)	110
4.2.3. Исследование структурных переходов в поли(А) методом протонной буферной емкости	117
4.2.3.1. Кривая буферной емкости поли(А) и ее разложение по базисным функциям	118

4.2.3.2. Диаграмма распределения отдельных форм поли(А) в процессе протонирования	124
4.2.3.3. Зависимость параметров структурных переходов в поли(А) от условий среды	127
4.2.4. Гипотеза о стабилизации поли(А) вертикальными водородными связями	129
4.2.4.1. <i>DS</i> -форма поли(А)	129
4.2.4.2. Гипотеза о структуре <i>DS</i> -формы поли(А)	132
4.2.5. Специфические особенности структурных переходов в поли(А)	136
4.3. Структурные переходы в поли(С) и поли(dС)	140
4.3.1. Зависимость протонной буферной емкости поли(С) от рН	140
4.3.2. Зависимость протонной буферной емкости поли(dС) от рН	152
4.3.3. Обнаружение методом электрофореза двух различных форм протонированного поли(С)	155
4.3.3.1. Электрофорез поли(С) в трис-ацетатном буфере	156
4.3.3.2. Электрофорез поли(С) в трис-фосфатном буфере	160
4.3.3.3. Модель образования <i>i</i> -мотива поли(С)	162
4.4. Структурные переходы в поли(Г)	166
4.4.1. Электрофорез поли(Г) при разных ионных силах	166
4.4.2. Электронные спектры поглощения толуидинового голубого, связанного с гомополирибонуклеотидами в агарозных гелях	172
4.5. Выводы к разделу 4	177
РАЗДЕЛ 5. СТРУКТУРНЫЕ МОДЕЛИ УЧАСТИЯ САМОАССОЦИАТОВ ГОМОПОЛИМЕРОВ В БИОЛОГИЧЕСКИХ ПРОЦЕССАХ	179
5.1. Возможное вовлечение внутримолекулярных двойных спиралей олиго(А) в биологические процессы	179
5.1.1. Стабилизация ARE-содержащих мРНК белками семейства Hu/ELAV	181

5.1.2. Саморегуляция синтеза цитоплазматического поли(А)- связывающего белка	184
5.1.3. Контроль длины поли(А)-хвостов мРНК эукариотов	186
5.2. Молекулярные модели участия G-квадруплексов РНК в биологических процессах	191
5.2.1. Структурные модели области димеризации геномной РНК ВИЧ-1	191
5.2.1.1. Модели DLS геномных РНК ВИЧ-1 _{Mal} и ВИЧ-1 _{Lai}	191
5.2.1.2. Возможное участие шпильки TAR в димеризации геномной РНК ВИЧ-1	203
5.2.2. Структурная модель поли(А)-сайта про-мРНК вируса SV40 L	206
5.2.3 Структурная модель комплексного поли(А)-сайта про-мРНК ВИЧ-1	211
5.2.3.1. Вторичная структура основного поли(А)-сайта про-мРНК ВИЧ-1	211
5.2.3.2. Вторичная структура основного вспомогательного элемента комплексного поли(А)-сайта про-мРНК ВИЧ-1	215
5.2.3.3. Вторичная структура комплексного поли(А)-сайта про- мРНК ВИЧ-1	223
5.3. Выводы к разделу 5	231
РАЗДЕЛ 6. БАЗА ПОЛИ(А)-САЙТОВ ПРО-МРНК ЧЕЛОВЕКА	233
6.1. Организация основных сайтов полиаденилирования про-мРНК высших эукариотов	234
6.2. Вспомогательные элементы сайтов полиаденилирования про-мРНК высших эукариотов	245
6.3. Выводы к разделу 6	253
ЗАКЛЮЧЕНИЕ	254
ВЫВОДЫ	269
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ	273

ПЕРЕЧЕНЬ УСЛОВНЫХ СОКРАЩЕНИЙ

I	ионная сила
K _R	коэффициент задержки
R	относительная электрофоретическая подвижность
T _m	температура плавления
u	электрофоретическая подвижность
ВИЧ-1	вирус иммунодефицита человека
гРНК	геномная РНК
гяРНП	гетерогенный ядерный рибонуклеопротеиновый (комплекс)
ДОВ	дисперсия оптического вращения
КД	круговой дихроизм
НТО	нетранслируемая область
ТГ	толуидиновый голубой О
ЯМР	ядерный магнитный резонанс
ARE	AU-богатый элемент
ARS	аденин-богатая последовательность
AuxDSE	вспомогательный «нижний» элемент
CF	фактор расщепления
CPSF	фактор специфичности расщепления и полиаденилирования
CstF	фактор стимуляции расщепления
DIS	сайт инициации димеризации
DLS	структура связи между димерами
DSE	«нижний» элемент
GRS	G-богатая последовательность
Nab2	поли(A)-связывающий белок, присутствующий в ядре в избытке
NC	нуклеокапсидный белок
PABPN1	ядерный поли(A)-связывающий белок 1
PABPC	цитоплазматический поли(A)-связывающий белок

PAP	поли(А)-полимераза
PBS	сайт связывания праймера
PTB	белок, связывающий полипиримидиновый тракт
RNAP II	РНК-полимераза II
SD	донор сплайсинга
SV40 L	обезьяний вирус 40 (поздние гены)
URE	U-богатый элемент
USE	«верхний» элемент

ВВЕДЕНИЕ

Актуальность темы. Вторичная структура и структуры более высоких уровней про-мРНК и мРНК, играют важную роль в их функционировании [1, 2], однако исследование структурных особенностей этих РНК, в отличие от транспортных, рибосомных и ряда малых РНК, находится в настоящее время на начальном этапе. Структуру про-мРНК и мРНК нельзя полностью охарактеризовать без учета присутствующих в них трактов гомополинуклеотидов и гомоолигонуклеотидов, которые имеют уникальные структурные особенности [3], и вносят особый вклад в общую архитектуру этих биополимеров.

Гомополимерные тракты выполняют функции регуляторных элементов во многих важных процессах жизненного цикла про-мРНК и мРНК. В клетке содержатся многочисленные белки, специфически связывающиеся с гомополимерами. Например, широко распространены такие белки как PABPCs (cytoplasmic poly(A)-binding proteins) [4], связывающиеся с поли(А), α CPs (α -complex proteins) [5], взаимодействующие с олиго(С)-трактами, NuR, специфически связывающийся с длинными поли(А) последовательностями [6], гяРНП Н, взаимодействующий с короткими олиго(Г)-трактами [7] и G-квадруплексами [8], и др. Однако молекулярные механизмы многих процессов, происходящих с участием гомополимерных трактов, выявлены не полностью или вообще неизвестны. Например, недостаточно хорошо изучены механизмы специфической стабилизации поли(А)-хвостов мРНК эукариотов различными белковыми комплексами [9], терминации полиаденилирования про-мРНК [10], саморегуляции синтеза PABPC [11] и других процессов.

При исследовании взаимодействия клеточных белков с поли(А), поли(С) и поли(У) структурное многообразие гомополимеров обычно не учитывается. Так, например, не учитывается, что одноцепочечный поли(А) существует не менее, чем в двух различных конформациях [12-15], для одноцепочечной спирали поли(С) в литературе предложены две различные модели [16, 17], а

ряд уридиновых остатков принимает уникальную конформацию [18]. Тем более, не учитывается, что эти гомополимеры и олигомеры, структурные характеристики которых во многом подобны, могут образовывать структуры высших порядков. Несмотря на то, что такие структуры в поли(А) и поли(С) индуцируются протонированием оснований при слабокислых рН, они могут образовываться и в клетке за счет переноса протонов от кислых аминокислотных групп белков на цитозиновые или адениновые основания [19, 20].

Кроме того, гомополирибонуклеотиды используются в молекулярной биологии как инструменты для первичного выявления специфичности РНК-связывающих белков, если точная последовательность места связывания неизвестна (например [21]). Отдельные гомополинуклеотиды рассматриваются как потенциальные терапевтические агенты или мишени [22-24], а также обсуждаются перспективы их применения в области РНК-нанотехнологий [25], в частности, в конструировании современных биомолекулярных устройств, функционирование которых основано на структурных переходах.

Несмотря на интенсивное исследование структурных особенностей поли(А), поли(С), поли(Г) и поли(У) на протяжении нескольких десятков лет, осталось немало спорных и невыясненных вопросов. В частности, неизвестно, какую структуру имеет двухцепочечная форма поли(А), основания которой не связаны водородными связями с фосфатами [26-29], способен ли поли(С) образовывать *i*-мотив [30]. Нет единого мнения о том, при каких концентрациях моновалентных катионов происходит переход поли(Г) из одноцепочечной спирали в квадруплекс ([31-33] и ссылки в этих работах) и др.

Таким образом, углубленное комплексное исследование физико-химических характеристик гомополимеров с помощью новых методов и подходов для выяснения механизмов важных биологических процессов, которые протекают с участием клеточных белков, специфически взаимодействующих с этими макромолекулами, является насущной и актуальной задачей.

Связь работы с научными программами, планами, темами.

Диссертационная работа соответствует основному плану научных исследований отдела молекулярной и квантовой биофизики Института молекулярной биологии и генетики НАН Украины. Она выполнена в рамках бюджетных тем «Изучение специфичности белково-нуклеиновых взаимодействий в модельных системах разного уровня сложности» (№ гос. регистрации 01.91.0002584, 1991-1995 г.г.); «Изучение физико-химической природы протекания элементарных процессов белково-нуклеинового узнавания в модельных системах типа «мономер-мономер» и «мономер-полимер» (№ гос. регистрации 0198U005247, 1996-2000 г.г.); «Исследование глубинных физико-химических механизмов элементарных процессов белково-нуклеинового узнавания» (№ гос. регистрации 0101U000603, 2001-2005 г.г.); целевых программ Отделения биохимии, физиологии и молекулярной биологии НАН Украины: «Молекулярные основы функционирования генома и его регуляция» (Физиолого-биохимические и молекулярно-генетические основы функционирования живых систем и разработка принципов управления ими), номер гос. регистрации 0104U000436, 2002-2006 г.г.); «Изучение фундаментальных основ организации и экспрессии генов и структурно-функциональные исследования протеомов организмов в норме и в условиях патологии, разработка основ новейших биотехнологий для диагностики и лечения тяжелых патологий» («Фундаментальные основы геномики и протеомики», номер гос. регистрации 0107U000337, 2007-2011 г.г.); Структурная и функциональная геномика для изучения ряда проблем функционирования вирусов, бактерий и высших эукариотов» (Функциональная геномика, протеомика и метаболомика в системной биологии», номер гос. регистрации 0112U002108, 2012-2016 р.р.), а также полученных на конкурсных основах отечественных и международных проектов – грант международного научного фонда “Україна”, шифр K1F100, 1994-1995 г.г.; «Создание биоинформационного ресурса для диагностики СПИДа (база данных вторичных структур контрольных элементов генома

ВИЧ-1)» (программа Государственного фонда фундаментальных исследований Украины “Фундаментальные основы новейших генных и клеточных технологий”, договор 18/037, 2006-2007 г.г.); “Филогенетическое исследование структурных вариантов сигнальных элементов, расположенных в 5' UTR генома ВИЧ-1” (программа “Новейшие медико-биологические проблемы и окружающая среда человека”, договор № 10, 2007-2010 г.г.); «Роль вторичной структуры РНК в репликации ВИЧ-1: исследование молекулярно-генетическими, филогенетическими и биоинформационными методами» (проект Украинского научно-технологического центра, договор № 4302, 2008-2010 г.г.). Часть диссертационной работы (1983-1984 г.г.) была выполнена в лаборатории физики биополимеров Института молекулярной генетики АН СССР (г. Москва, Россия).

Цель и задачи исследования.

Цель работы – экспериментально изучить физико-химические характеристики гомополирибонуклеотидов, их взаимодействие с производными аминокислот и структурные переходы в зависимости от условий среды с помощью новых физико-химических методов, а также разработать и обосновать на основе этих данных молекулярные модели биологических процессов, которые могут происходить с участием самоассоциатов гомополимеров.

Для достижения цели были поставлены следующие взаимосвязанные задачи:

1. Исследовать на примере ДНК влияние ионной силы раствора на эффективные размеры полинуклеотидов методом осмометрии.
2. Изучить структурные переходы в поли(А), поли(С), поли(Г) и поли(У) и их взаимодействие с производными основных аминокислот методом аналитического электрофореза в зависимости от условий среды.
3. Изучить структурные переходы в поли(А) и поли(С) методом протонной буферной емкости в зависимости от условий среды.

4. Разработать и обосновать молекулярные модели биологических процессов, которые могут проходить с участием самоассоциатов гомополирибонуклеотидов.
5. Создать базу комплексных сигналов полиаденилирования про-мРНК человека и проанализировать встречаемость различных вспомогательных элементов поли(А)-сайтов, а также последовательностей, способных образовывать четырехцепочечные структуры.

Объект исследования: природа структурных переходов в гомополимерах в зависимости от условий среды; молекулярные механизмы функционирования трактов гомополи(олиго)рибонуклеотидов в природных РНК.

Предмет исследования: структурные переходы в синтетических гомополирибонуклеотидах поли(А), поли(С), поли(U) и поли(Г) и модели функционирования.

Методы исследования: для исследования объемных эффектов в полинуклеотидах на примере ДНК использован метод осмометрии, структурные переходы в гомополирибонуклеотидах изучены методами электронной спектроскопии, аналитического электрофореза и протонной буферной емкости. При создании моделей сайтов полиаденилирования вирусных про-мРНК компьютерное предсказание вторичной структуры РНК осуществляли с помощью программ «mfold» и «UNAFold».

Научная новизна полученных результатов. Впервые определен эффективный диаметр двойной спирали ДНК в широком интервале ионных сил (0,01-0,1 М). Для изучения структурных характеристик гомополирибонуклеотидов в зависимости от условий среды впервые применены методы аналитического электрофореза и протонной буферной емкости. С их помощью получены новые данные о физико-химической природе структурных переходов в поли(А), поли(С) и поли(Г). Впервые предложены структурные механизмы образования двух различных

одноцепочечных и двух различных двухцепочечных форм поли(А), «замороженной» формы поли(А) и четырехцепочечной структуры поли(С). Разработаны и обоснованы молекулярные модели участия самоассоциатов поли(А) и поли(Г) в ряде важных биологических процессов.

Практическое значение полученных результатов. Поскольку гомополирибонуклеотиды рассматриваются как потенциальные терапевтические агенты или мишени, а также как претенденты на использование в области РНК-нанотехнологий, важно иметь детальные сведения о структурных характеристиках этих полимеров при создании лекарственных препаратов и конструировании соответствующих биомолекулярных наноустройств. Новые данные, полученные в диссертационной работе, могут внести вклад в эти области исследований. Проведенное в работе широкомасштабное филогенетическое исследование структурных элементов сайта полиаденилирования про-мРНК ВИЧ-1 может оказаться полезным при конструировании терапевтических антиретровирусных агентов, направленных на подавление функций сигнальных элементов, существенных для репликации вируса.

Личный вклад соискателя. Личный вклад диссертанта является определяющим на всех этапах исследования. Основные идеи, разработанные в диссертации, принадлежат автору. В большинстве работ, опубликованных совместно с соавторами, личный вклад соискателя заключается в выполнении значительного объема экспериментальной части и непосредственном участии в анализе и интерпретации полученных результатов. Автор выражает глубокую благодарность всем своим соавторам публикаций за сотрудничество, прежде всего, к.б.н. с.н.с. Потягайло А. Л. и к.б.н. с.н.с. Коломиец И. Н. за вклад в работу по созданию базы поли(А)-сайтов человека, а также н.с. Степанюгину А. В. за участие в проведении исследований гомополимеров методом электрофореза. Особую благодарность автор выражает научному консультанту – члену-корреспонденту НАН Украины, профессору Д. М.

Говоруну за внимание к работе, плодотворные дискуссии и моральную поддержку.

Апробация результатов диссертации. Результаты исследований докладывались на таких научных конференциях: Симпозиум с участием стран-членов СЭВ и СФРЮ «Физико-химические свойства биополимеров в растворе и клетках» (Пушино, СССР, 1985); I-й съезд Украинского биофизического общества (Киев, Украина, 1994); II съезд Украинского биофизического общества (Харьков, Украина, 1998); Конференция по физике биологических систем (с международными участниками, Пуша-Водица (Киев), Украина, 1998); XIV международная школа-семинар «Спектроскопия молекул и кристаллов» (Одесса, Украина, 1999); II съезд биофизиков России (Москва, Россия, 1999); XIV международный круглый стол «Нуклеозиды, нуклеотиды и их биологическое применение» (Сан-Франциско, США, 2000); VIII Украинский биохимический съезд (Черновцы, Украина, 2002); Установочный съезд украинского общества клеточной биологии (Львов, Украина, 2004); I Украинская научная конференция «Проблемы биологической и медицинской физики» (Харьков, Украина, 2004); IV международная конференция «Биоресурсы и вирусы» (Киев, Украина, 2004); 5-я Всероссийская научно-практическая конференция «Генодиагностика инфекционных болезней» (Москва, Россия, 2004); IV съезд Украинского биофизического общества (Донецк, Украина, 2006); 5-ая Европейская конференция по клиническим и социологическим исследованиям по СПИДу и лекарственным препаратам (Вильнюс, Литва, 2009); 34 конгресс FEBS (Прага, Чехия, 2009); X Украинский биохимический съезд (Одесса, Украина, 2010); 7-ая ежегодная научная встреча «Мосты в науках о жизни» (Будапешт, Венгрия, 2012); XI Украинский биохимический конгресс (Киев, Украина, 2014), а также на научных семинарах и научных конференциях ИМБиГ НАН Украины и отдела молекулярной и квантовой биофизики.

Публикации. По теме диссертации опубликовано 49 научных работ, из них – 1 глава в международной научной монографии, 29 статей в

отечественных и международных научных изданиях и 19 тезисов сообщений в сборниках международных и всеукраинских научных конференций.

Структура и объем работы. Диссертация состоит из введения, обзора литературы, материалов и методов исследования, двух экспериментальных разделов с обсуждением результатов, раздела, посвященного молекулярным моделям, раздела, посвященного созданию базы поли(А)-сайтов про-мРНК человека, заключения, выводов и списка использованных литературных источников, насчитывающего 355 наименований. Работа изложена на 312 страницах машинописного текста, содержит 96 рисунков и 16 таблиц.

РАЗДЕЛ 1

ОБЗОР ЛИТЕРАТУРЫ

1.1. Структурные переходы в гомополирибонуклеотидах

1.1.1. Поли(А). При слабощелочных и нейтральных значениях pH поли(А) состоит из спиральных и разупорядоченных областей ([34] и ссылки в этой работе), причем спираль существует не менее чем в двух конформационных состояниях [12-15]. Снижение pH индуцирует кооперативное протонирование аденинов и переход одноцепочечной формы поли(А) в двухцепочечную спираль, который регистрируется методом потенциометрического и спектрофотометрического титрования [35]. Модель полностью протонированной двухспиральной формы поли(А) была предложена в 1961 г. Rich et al. [36] на основании данных рентгеноструктурного анализа (рис. 1.1).

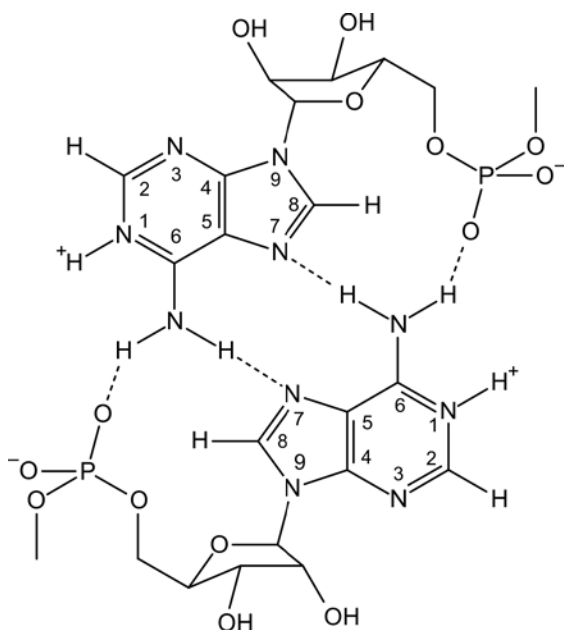


Рис. 1.1. Схема спаривания оснований в высокопротонированной форме поли(А) [36]. Пунктиром изображены межмолекулярные водородные (Н) связи

Цепи в двойной спирали поли(А) ориентированы параллельно. Протонирование происходит по атому N1 аденина, не участвующего в

образовании водородных связей. Водородные связи образуются между атомами азота N7 каждого основания одной цепи и одним из атомов водорода аминогруппы N6H2 противоположной цепи. Второй атом водорода аминогруппы образует водородную связь с фосфатной группой противоположной цепи.

В 1969 г. несколько групп исследователей обнаружили, что при кислых рН поли(А) может существовать не в одной, а в двух двухспиральных формах, которые различаются по спектрам кругового дихроизма (КД) [37], дисперсии оптического вращения (ДОВ) [27] и рентгеновским диаграммам [26]. Adler et al. [27] показали методом ДОВ, что в узком интервале рН, сразу после перехода из одноцепочечного в двухцепочечное состояние, поли(А) существует только в одной конформации, которую они назвали *B*-формой (более щелочной (basic) из двух двухцепочечных форм). При дальнейшем подкислении *B*-форма постепенно превращается в *A*-форму (acidic), и при низких рН поли(А) существует только в *A*-форме. Сопоставляя результаты исследования поли(А) при концентрации КСl, равной 0,1 М, методом ДОВ и методом потенциометрического титрования, эти авторы предположили, что *B*-форма является, по-видимому, полупротонированной структурой, поскольку только эта форма регистрируется методом ДОВ в области полупротонирования поли(А) при рН 5,8-5,9. При промежуточных рН обе формы находятся в динамическом равновесии.

Аналогичные результаты получены Finch & Klug [26] методом рентгеноструктурного анализа волокон поли(А), вытянутых из кислых растворов. Переходы между формами регистрировались путем выдерживания волокон или в парах CH_3COOH (для образования *A*-формы), или в парах NH_3 (для образования *B*-формы). Авторы определили, что спираль *A*-формы более вытянута, чем спираль *B*-формы. *B*-форма имеет 10 пар нуклеотидов (п.н.) на виток спирали и расстояние между нуклеотидными остатками вдоль оси спирали составляет 0,36 нм, а у *A*-формы эти параметры равны соответственно 8,4 п.н. и 0,38 нм. Однако модель структуры *B*-формы поли(А) авторы не

представили. Следует отметить, что такая модель не представлена в литературе и до сих пор.

Позднее наличие двух двухспиральных форм поли(А) было подтверждено Janik et al. методом полярографии [28], Scovell – методом комбинационного рассеяния (КР) [38], Lerner & Kearns – методом ЯМР спектроскопии [29] и Кузнецовым и Воронцовой – методом потенциометрического титрования изоионных растворов поли(А) [39]. Однако эти формы не регистрируются такими методами, как потенциометрическое титрование солевых растворов поли(А) [35, 40] и УФ-спектроскопия [27].

Предположение Adler et al. [27] о том, что сразу после перехода поли(А) из одноцепочечной формы в двухцепочечную этот полимер существует только в одной из двух двухспиральных форм, поддерживается данными рентгеноструктурного анализа [26], а также результатами Janik et al. [28], исследовавшими переходы в поли(А) методом полярографии и реакцией с формальдегидом. Данную форму, в которой только примерно 50 % аденинов взаимодействуют с формальдегидом, авторы называют «промежуточной» формой, а высокопротонированную форму – «плотно упакованной» (tightly packed).

Lerner & Kearns [29], исследуя промежуточную полупротонированную структуру поли(А) методом ЯМР на ядрах ^1H и ^{31}P в условиях, близких к условиям работы Adler et al. [27], пришли к выводу, что в этом состоянии поли(А) представляет собой промежуточную структуру, стэкинг-взаимодействия оснований в которой подобны взаимодействиям в высокопротонированной форме [36], однако сахарофосфатный остов промежуточной формы существует в двух конформациях, медленно переходящих одна в другую. В одной из этих конформаций фосфаты образуют водородные связи с аминогруппами оснований. При дальнейшем протонировании происходит постепенный переход в кислую форму, в которой все фосфаты образуют водородные связи и «приближены» к оси спирали, но стэкинг оснований остается почти неизменным.

Как видно из вышеизложенного, различные двухспиральные формы поли(А) не имеют в литературе фиксированного названия. В частности, термины «А- и В-форма» не употребляются в работах Janik et al. [28], Lerner & Kearns [29] и Scovell [38]. Кроме того термины «А- и В-форма» могут вызывать ассоциацию с классическим названием форм ДНК и РНК, а не с терминами «acidic» и «basic». Исходя из этих соображений, мы назвали В-форму – *DS*-формой (double stranded) и А-форму – *DS_P*-формой. Нижний индекс «Р» (phosphate) относится к дуплексу, в котором основания участвуют в образовании водородной связи с фосфатами. Далее в тексте мы будем употреблять эти новые названия.

В ряде работ отмечено, что в кислых растворах поли(А) наряду с *DS*- и *DS_P*-формами присутствует также третья форма [26, 28, 39, 41]. Finch & Klug [26] назвали эту форму фазой *D*. Эта фаза напоминает по своим параметрам щелочную форму поли(А), описанную Morgan & Byrne [42]. Кузнецов и соавт. [39] предполагают, что наличие определенного количества одноцепочечной формы поли(А) в кислых растворах является свойством самой системы. По мнению Bobst [41] третья форма поли(А) может состоять из больших агрегатов двухспиральных цепей и/или «сетей» агрегированных одноцепочечных макромолекул. Образование агрегатов в кислых растворах поли(А) наблюдали еще в самых ранних работах, посвященных исследованию физико-химических характеристик этого полинуклеотида ([35] и ссылки в этой работе); при этом было показано, что агрегация увеличивается при повышении ионной силы и концентрации макромолекул, причем агрегаты не диссоциируют при разбавлении растворов.

Janik et al. [28] провели наиболее детальное исследование третьей кислой формы поли(А). Они показали, что эта форма образуется из нейтральной формы в том же интервале рН, что и *DS_P*-форма. Авторы предположили, что в условиях, способствующих плохой подгонке цепей, например при быстром подкислении, молекулы поли(А) «замерзают» в одноцепочечной конформации при наличии беспорядочно распределенных коротких двухспиральных

областей, мешающих подгонке остающихся одноцепочечных участков. Авторы допускают, что двухспиральные области являются местами объединения отдельных одноцепочечных молекул в объемистые агрегаты (сети).

1.1.2. Поли(С). Конформационные переходы в поли(С) наиболее близки по характеру к переходам в поли(А) [3]. При щелочных и нейтральных рН оба полимера представляют собой одноцепочечные структуры, состоящие из спиральных и разупорядоченных участков. Для одноцепочечной спирали поли(С) в литературе предложены две модели. Согласно рентгеноструктурным данным, это правая спираль, у которой виток состоит из 6 нуклеотидов (нт) и расстояние между нуклеотидными остатками вдоль оси спирали составляет 3,1 Å [16]. По данным ЯМР спектроскопии нейтральный поли(С), вероятно, существует в водном растворе в виде левой спирали, параметры которой равны соответственно 8 нт и 2,9 Å [17]. Эта спираль стабилизирована водородными связями между амино- и карбонильной группами соседних оснований.

При подкислении раствора оба полимера – поли(А) и поли(С) – образуют двойные спирали ($pK \sim 5,8$ в 0,1 М NaCl [43]). В обоих случаях конформационные переходы имеют высококооперативный характер. В отличие от классической антипараллельной ориентации цепей в двойных спиральях РНК и ДНК, цепи в двойных спиральях поли(С) и поли(А) параллельны [3].

Основные различия между структурными переходами в поли(С) и поли(А) заключаются в следующем. В поли(А) присоединяющиеся к нуклеотидным основаниям протоны не участвуют в образовании водородных связей между ними, а в поли(С) – участвуют. Температура плавления (T_m) двухспирального поли(А) непрерывно возрастает с увеличением степени протонирования полимера [40], а зависимость T_m двухспиральной формы поли(С) от рН имеет колоколообразный характер: при степени протонирования полимера меньше 50 % она возрастает, а затем при

дальнейшем уменьшении pH падает [44-47]. При высокой степени протонирования (~ 70 %) двойная спираль поли(С), в отличие от поли(А), разрушается [43, 45, 47, 48].

Особенности структурных переходов в поли(С) связаны с характером протонирования его нуклеотидных оснований. На рис. 1.2 показана схема спаривания оснований в двойной полупротонированной спирали поли(С), впервые предложенная Akinrimisi [44] и Langridge & Rich [49].

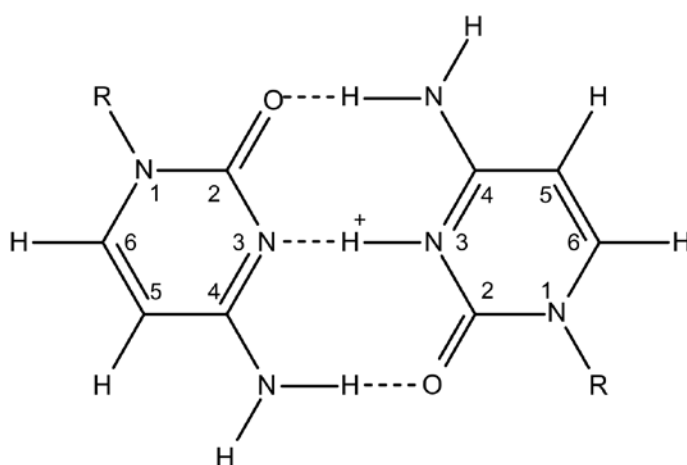


Рис. 1.2. Схема спаривания оснований в двухспиральной форме поли(С) [44, 49]. Пунктиром изображены межмолекулярные водородные (H) связи

Следует отметить, что эта схема основывается на результатах исследования кристаллической структуры цитозин-5-уксусной кислоты и других низкомолекулярных модельных соединений, а также на некоторых физико-химических свойствах самого полимера. Структурная модель двойной спирали поли(С) была предложена Langridge & Rich [49] на основании данных рентгеноструктурного анализа, но позднее Arnott et al. [16] установили, что такие же дифрактограммы соответствуют не дуплексу, а одноцепочечному поли(С). В настоящее время структурные параметры для двойных спиралей поли(С) и поли(dС) отсутствуют.

Как видно из рис. 1.2, на пару комплементарных оснований в двойной спирали поли(С) приходится один протон, т. е. при степени протонирования полимера 50 % все водородные связи между основаниями уже образованы. При этом между электронейтральным Cyt и Cyt(N3H)⁺, несущим на себе избыточный положительный заряд, образуются, как и в случае классической Уотсон-Криковской пары Gua:Cyt, три водородные связи. Малеев и соавт. [50] показали также, что стабильность как односпиральной, так и двухспиральной поли(С) в большой степени обусловлена специфическим гидратным окружением цепей.

Экспериментальные данные свидетельствуют о том, что поли(С) может быть протонирован полностью. Однако единого устоявшегося мнения о механизме протонирования полупротонированной двойной спирали поли(С) в настоящее время нет. Авторы ряда работ [45, 47, 51, 52], в которых структурные переходы поли(С) исследовали методом потенциометрического титрования, а также некоторыми другими физико-химическими методами, пришли к довольно противоречивым выводам.

Так, Hartman & Rich [45] полагают, что дополнительные протоны могут присоединяться к электронейтральным цитозиновым остаткам только посредством раскрытия пары. Этот процесс снижает термодинамическую стабильность двойной спирали поли(С) и регистрируется, как понижение T_m . В то же время Klump [47] считает маловероятным присоединение протонов к образовавшейся двойной спирали поли(С) в ее центре по положению атома азота N3 непротонированного Cyt и склоняется к мнению о внешнем механизме протонирования. В этой связи он предлагает обратить внимание на возможность существования равновесия между протонированными иминными и аминными прототропными таутомерами цитозина и представляет схему спаривания цитозиновых остатков, подобную той, что показана на рис. 1.2, но с локализацией избыточного положительного заряда не на имино- (как это традиционно представляют [3]), а на аминогруппе, считая, что электрическое отталкивание протонов в этом случае будет меньше.

Авторы вышеупомянутых работ [45] и [47], так же как и других, например [43, 48], считают, что при высокой степени протонирования оснований (более 70 %) двойные спирали поли(С) разрушаются. В то же время Сулейманова, Аполонник и Кузнецов [51, 52], напротив, утверждают, что полностью протонированные двойные спирали поли(С) существуют как стабильные конформации, не выдвигая, однако, при этом каких-либо предположений об их структуре.

В двойной спирали поли(dС), как показано методом КД-спектроскопии [43], основания спарены таким же образом, как и в поли(С) (рис. 1.2). Но этот полимер кооперативно переходит из одноцепочечного состояния в двухспиральное при значительно более щелочных рН. Авторы работы [43] связывают это с тем, что одноцепочечный поли(dС) является наименее структурированной макромолекулой среди всех гомополинуклеотидов, содержащих адениновые и цитозиновые основания, поэтому его переход в двойную спираль с большим стэкингом спаренных оснований является энергетически выгодным при значениях рН, намного превышающих рК протонирования мономера. Следует отметить, что двойная полупротонированная спираль поли(dС) значительно стабильнее аналогичной спирали поли(С) – ее T_m не менее чем на 18 °С выше T_m поли(С) [44, 53].

В 1993 г. при исследовании методом ЯМР олигонуклеотида dTC₅ был открыт i-мотив ДНК [54], который представляет собой четырехцепочечную структуру, содержащую ту же самую базовую единицу – пару Cyt·Cyt⁺ (рис. 1.2), что и в моделях двухспиральных поли(С) и поли(dС). Для примера на рис. 1.3. показан i-мотив dTC₂. В i-мотиве две двухцепочечные структуры (дуплексы), состоящие из параллельно ориентированных цепей со спаренными основаниями, связаны между собой таким образом, что пары оснований этих дуплексов полностью интеркалированы; при этом сами дуплексы ориентированы антипараллельно. В дальнейшем образование i-мотива было подтверждено для различных полупротонированных dС-богатых олигонуклеотидов [30], включая dC₈ [55] и dC₁₂ [56], а также было показано,

что такой мотив, но с гораздо меньшей стабильностью, могут образовывать и С-богатые олигорибонуклеотиды [57].

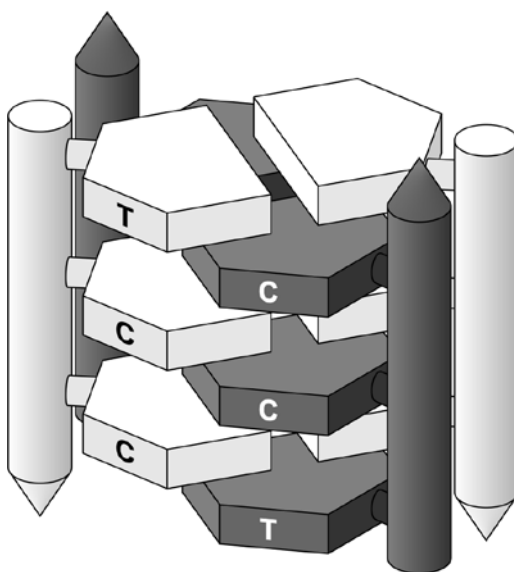


Рис. 1.3. Схема i-мотива для dTC₂. Адаптировано из [54]

1.1.3. Поли(G), поли(U). Поли(G) способен образовывать четырехцепочечные структуры (квадруплексы), содержащие водородсвязанные тетрады гуаниновых оснований [3]. Схема спаривания оснований в такой структуре показана на рис. 1.4.

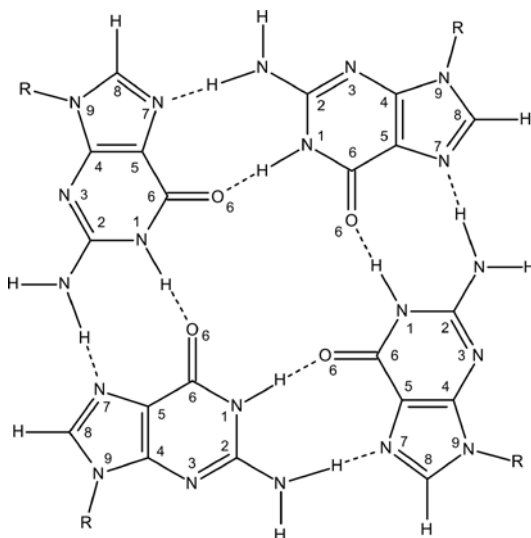


Рис. 1.4. Схема спаривания оснований в G-квадруплексах [3]. R – сахарофосфатный остаток. Пунктиром изображены межмолекулярные водородные (H) связи

Четырехцепочечные спирали стабилизируются в присутствии соли. Литературные данные о зависимости структуры поли(G) от концентрации моновалентных катионов противоречивы. С одной стороны, исследования, проведенные методами электронной спектроскопии [58], электронной микроскопии [33], кинетики тритиевого обмена [59] и другими методами, свидетельствуют в пользу того, что при очень низких ионных силах (H_2O , 2 mM NaCl или 1 mM фосфатный буфер, содержащий 1 mM ЭДТА) поли(G) находится преимущественно в четырехцепочечной форме. С другой стороны иммунологическое исследование [31] показало, что при очень низкой концентрации соли (в H_2O или 5 mM трис) этот полимер находится преимущественно в одноцепочечной форме. Аналогичный результат получен при изучении поли(G) в H_2O методом комбинационного рассеяния света (КРС) [32]. Протонирование атома азота N7 гуанина поли(G) ведет к образованию двухцепочечной спирали [3, 33], но ее детальная структура до сих пор неизвестна.

Поли(U) в отличие от поли(A), поли(C) и поли(G) не образует одноцепочечные спирали. Выше 15 °C в полимере и динуклеотиде UpU стэкинг-взаимодействия незначительны. Ряд уридиновых остатков в одноцепочечной РНК обладает уникальными конформационными характеристиками [18], которые, как предполагают авторы, могут узнаваться определенными белками.

Полиуридиловая кислота образует самоассоциаты (внутримолекулярные шпильки) лишь при температуре ниже комнатной (рис. 1.5 а). Предполагается [60], что при прогревании сначала происходит образование множества коротких шпилек, а затем их кооперативное превращение в одноцепочечную разупорядоченную структуру. Переход из одноцепочечного клубка в двухцепочечную структуру происходит при нейтральном рН при ~ 5° C в 1 M Na^+ или в 0,15 M Na^+ и 0,01 M MgCl_2 [61]. Схема спаривания оснований в двухспиральных поли(U) показана на рис. 1.5 б [62].

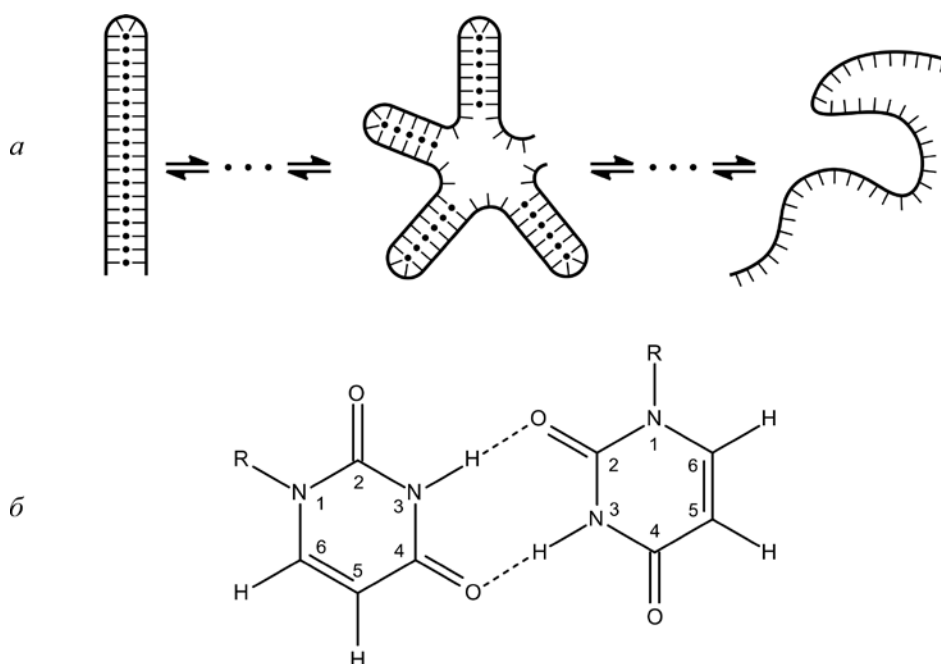


Рис. 1.5. Структурные переходы в поли(U) (адаптировано из [60]) (а). Схема спаривания оснований в двойной спирали поли(U) [62] (б). R – сахарофосфатный остаток. Пунктиром изображены межмолекулярные водородные (H) связи

Вторичная структура может поддерживаться и при более высоких температурах – так, например, спермин может повышать температуру плавления поли(U) до 28° С [63], еще большим стабилизирующим действием обладает полилизин, в его присутствии температура плавления поли(U) повышается до 35 °С [64].

Таким образом, все гомополирибонуклеотиды могут существовать в разных формах. Структурные переходы в поли(A) и поли(C) вызываются протонированием оснований, в поли(G) – взаимодействием с катионами или протонированием оснований, а в поли(U) – взаимодействием с катионами или поликатионами.

Как следует из приведенного краткого обзора, в ряде случаев литературные данные о конформациях гомополирибонуклеотидов отсутствуют или противоречивы. В дополнение к невыясненным вопросам,

упомянутым во вступлении, можно добавить, например, следующие вопросы: какие отличия имеются в структурах двух разных форм одноцепочечного поли(А), какова структура «замороженной» формы этого гомополинуклеотида, каков механизм распада двойной спирали поли(С) и др. Применение в диссертационной работе методов, ранее не использованных в литературе для исследования гомополирибонуклеотидов, позволило дать ответы на ряд вопросов.

1.2. Гомополи(олиго)рибонуклеотиды в природных РНК

1.2.1. Полиаденилирование про-мРНК высших эукариотов.

Полиадениловая кислота является самым известным природным гомополирибонуклеотидом. Поли(А)-последовательности, ковалентно присоединенные к 3'-концам мРНК, участвуют в таких внутриклеточных процессах, как транспорт мРНК из ядра в цитоплазму, контроль их времени жизни и инициация трансляции [4]. Полиаденилированию подвергаются предшественники мРНК (про-мРНК) всех живых организмов от бактерий до человека [65, 66].

Поли(А)-хвосты мРНК высших эукариотов образуются в результате двухстадийной реакции: сначала происходит специфическое эндонуклеолитическое расщепление про-мРНК, а затем – присоединение поли(А)-хвоста к вновь образовавшемуся 3'-концу [67]. Реакция осуществляется большим комплексом белков. Для первой стадии реакции необходимы: фактор специфичности расщепления и полиаденилирования CPSF, фактор стимулирования расщепления CstF, факторы расщепления CF I_m и CF II_m (m – mammalian), поли(А)-полимераза PAP и РНК полимеразы II (RNAP II). Во второй стадии – синтезе поли(А)-хвоста – участвуют PAP, CPSF и поли(А)-связывающий белок PABPN1. Эндонуклеаза, расщепляющая про-мРНК, была найдена значительно позднее других факторов полиаденилирования, ей оказалась субъединица фактора CPSF (CPSF-73) [68].

Необходимым условием осуществления реакции полиаденилирования является наличие в про-мРНК как минимум двух сигнальных последовательностей (рис. 1.6) – поли(А)-сигнала PAS, расположенного в пределах 40 нуклеотидов до места расщепления, и U/GU-богатого «нижнего» сигнального элемента DSE, расположенного примерно на таком же расстоянии за этим местом [69].

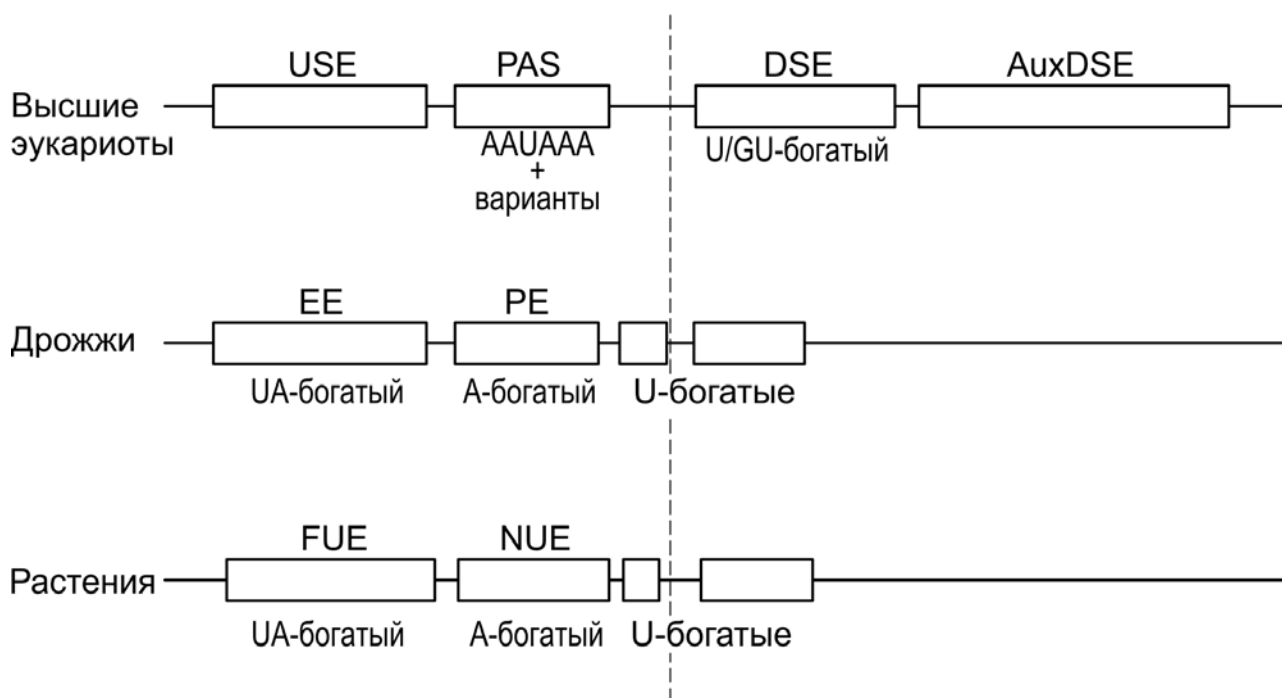


Рис. 1.6. Организация сайтов полиаденилирования в про-мРНК высших эукариотов, дрожжей и растений. Адаптировано из [69]. Пунктир разделяет области транскриптов, расположенные до места расщепления и за этим местом

Поли(А)-сигнал (гексамер AAUAAA и его производные) и DSE образуют основной сайт полиаденилирования про-мРНК высших эукариотов. Сборка комплекса расщепления осуществляется кооперативно. Фактор CPSF посредством субъединицы CPSF-160 узнает гексамер AAUAAA, а субъединица CstF-64 фактора CstF связывается с U/GU-богатым элементом. Эти белки взаимодействуют между собой. Взаимодействие CPSF и CF I_m с

про-мРНК является одним из наиболее ранних событий в сборке комплекса полиаденилирования.

Характерной чертой сигнала полиаденилирования AAUAAA является то, что большинство единичных замен оснований в этом элементе значительно уменьшают эффективность реакций расщепления и присоединения поли(А)-последовательности. Исключением является гексамер AUUAAA, который направляет реакцию расщепления с эффективностью примерно 66 % по сравнению с каноническим гексамером. Кроме того полиаденилирование расщепленных про-мРНК происходит с эффективностью более 10 % в присутствии 8 из 18 возможных вариантов гексамера с единичными заменами [70].

Точная консенсусная последовательность «нижнего» U/GU-богатого элемента поли(А)-области про-мРНК высших эукариотов в литературе не определена. На основании литературных данных и исследования созданной нами базы сайтов полиаденилирования про-мРНК человека (см. раздел 6) мы предположили, что общий консенсус для «нижнего» элемента не существует. Этот сигнал состоит из различных простых элементов (пентамеров), расположенных на разных расстояниях друг от друга [65, 71, 72]. Консенсус для U-богатых пентамеров (UREs) определен экспериментально Chen et al. [73] как последовательность, состоящая из четырех или пяти U и одного A или C, или G, расположенных в любом месте пентамера («четыре из пяти» U-богатый элемент). По результатам нашего анализа GU-богатый элемент представляет собой все возможные варианты пентамера из двух динуклеотидов GU и одного U (GUUGU, UGUGU и GUGUU).

In vivo процесс полиаденилирования сопряжен с процессами сплайсинга и терминации транскрипции [74]. При удалении 3'-концевого интрона узнавание экзона реализуется посредством взаимодействия между белками, связанными с 3'-сайтом сплайсинга и с сайтом полиаденилирования. При этом процессы полиаденилирования и удаления последнего интрона взаимно стимулируют друг друга. Со стороны аппарата сплайсинга главными

участниками являются U2AF65 (субъединица фактора U2AF), U2 мяРНП и SRm160. U2AF65, связанный с пиримидиновым трактом на 3' сайте сплайсинга, стимулирует процесс расщепления и полиаденилирования, способствуя связыванию CF I_m с сайтом полиаденилирования. С другой стороны, связывание U2AF65 с 3' сайтом сплайсинга стимулируется PAF. Сопряжение двух процессов происходит также посредством взаимодействия между субъединицей SF3b U2 мяРНП, связанной с районом ветвления РНК в интроне, и CPSF, связанного с сайтом полиаденилирования. Коактиватор сплайсинга SRm160 способствует привлечению CPSF к сайту полиаденилирования.

РНК-полимераза II выполняет в процессе полиаденилирования две функции – доставляет соответствующие белковые факторы к сигналам про-мРНК, направляющим реакцию полиаденилирования, и непосредственно принимает участие в реакции расщепления транскрипта [75]. На рис 1.7 представлена схема из нашего обзора [76], обобщающая литературные данные о координации процессов полиаденилирования и терминации транскрипции.

Эффективная терминация транскрипции, по-видимому, осуществляется при одновременном взаимодействии RNAP II как с сигнальными элементами в ДНК, так и с сайтом полиаденилирования транскрипта. Контакт RNAP II с про-мРНК осуществляется посредством уникального С-концевого домена (СТД) наибольшей субъединицы фермента (рис. 1.7 а). С этим доменом ассоциированы белковые факторы расщепления. Реакция расщепления сопровождается деградацией 3'-концевой части транскрипта (рис. 1.7 б). Освобождение RNAP II от РНК приводит к дестабилизации связи фермента с двухспиральным участком ДНК, расположенным сразу же за расплетенным транскрибируемым участком, и к отделению RNAP II от матрицы.

Эффективность процесса полиаденилирования про-мРНК высших эукариотов может модулироваться добавочными сигналами, расположенными как до места расщепления (upstream sequence elements, USEs), так и за ним (auxiliary downstream sequence elements, AuxDSEs).

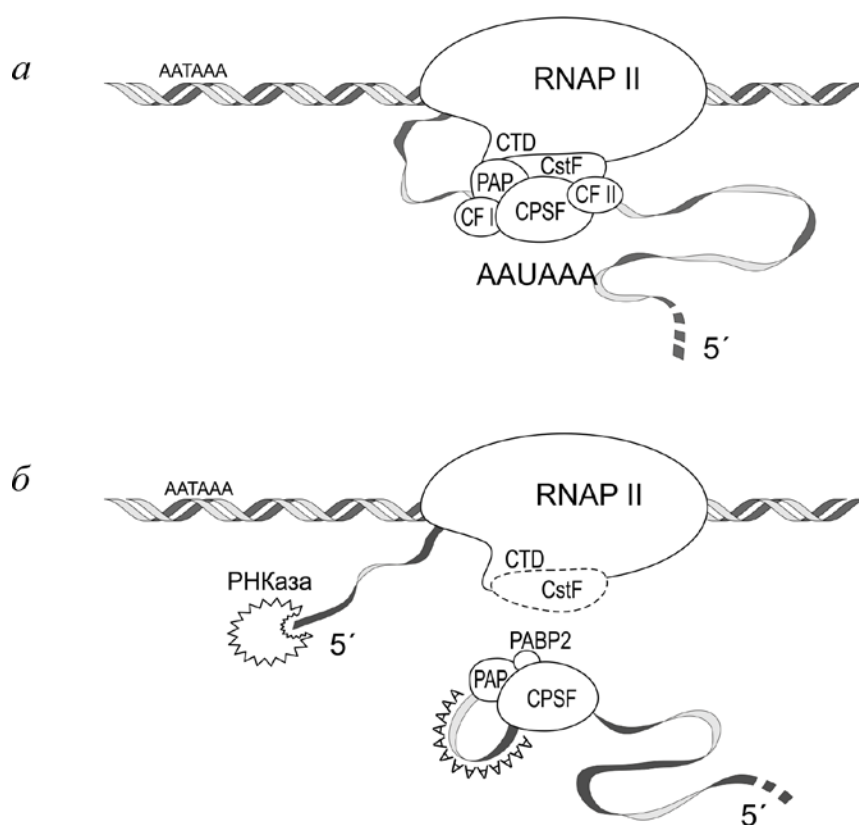


Рис. 1.7. Схема сопряжения процессов терминации транскрипции и полиаденилирования; *а* – участие элонгирующей РНК-полимеразы II в образовании комплекса расщепления транскрипта; *б* – полиаденилирование 5'-концевой части расщепленного транскрипта и деградация его 3'-концевой части [76]

Многие про-мРНК содержат по несколько сайтов полиаденилирования, которые расположены в 3'-нетранслированных областях (НТО) разных экзонов или же в 3'-НТО одного экзона [74]. В первом случае использование альтернативных сайтов приводит к синтезу разных белков. Во втором случае про-мРНК с разными 3'-НТО транслируются в тот же самый белок, но он может быть синтезирован с различной эффективностью, если между тандемно расположенными поли(А)-сайтами имеются элементы, регулирующие стабильность или трансляцию мРНК. Альтернативные поли(А)-сайты могут использоваться в разных тканях, а также на разных стадиях клеточного цикла или развития организма.

По механизму воздействия на процесс полиаденилирования про-мРНК вспомогательные элементы можно разделить на 3 группы. Во-первых, USEs и AuxDSEs могут служить местами связывания для вспомогательных белков, которые оказывают либо положительное, либо отрицательное влияние на процесс полиаденилирования. Например, ближний поли(А)-сайт про-мРНК COX-2 (cyclooxygenase-2) содержит три USEs (AUUUGAA, AUUUCUAA AUUUCUUA [77]), с которыми специфически связываются белки p54^{nrb} (p54 nonO-related binding protein), PTB (polypyrimidine tract binding protein), PSF (PTB associated splicing factor) и U1A [78]. Взаимодействие этих белков с USEs увеличивает эффективность полиаденилирования. С другой стороны, две молекулы U1A, связываясь с трактами AUUGU/CAC, ингибируют поли(А) полимеразу и соответственно процесс полиаденилирования ряда про-мРНК, например [79].

Во-вторых, вспомогательные элементы могут служить местами связывания для основных белковых факторов аппарата полиаденилирования. Некоторые из этих элементов имеют ту же последовательность нуклеотидов, что и основные элементы сигнала полиаденилирования, а другие представляют собой отличающиеся последовательности, с которыми белковые факторы связываются через другие РНК-связывающие мотивы. В первом случае вспомогательный элемент может служить «ловушкой» для основного белкового фактора [80]. Во втором случае вспомогательный элемент поли(А)-сайта может служить первичным и/или временным местом связывания белкового фактора (например CPSF), до того момента, когда он сможет связаться с основным сигналом [81].

В-третьих, USEs и AuxDSEs могут играть структурную роль, в частности, обеспечивая такое сворачивание области поли(А)-сайта, при котором сигнальные элементы не образуют пары оснований с другими участками про-мРНК и полностью доступны для белковых факторов полиаденилирования, например, как это предполагается в случае про-мРНК аденовируса-2 L4 [82].

Вспомогательные элементы могут также образовывать стабильные структуры (например, псевдоузлы), которые будут препятствовать перемещению относительно слабо связанного с про-мРНК фактора CstF вдоль РНК и таким образом «удерживать» его в окрестности поли(А)-сайта [83].

1.2.2. Полиаденилирование про-мРНК дрожжей, растений и вирусов.

1.2.2.1. Дрожжи. Сайт полиаденилирования про-мРНК дрожжей содержит четыре сигнала [69, 74, 84]: PE (positioning element), расположенный на расстоянии 10-30 нуклеотидов до места расщепления, (оптимальный вариант – AAUAAA), EE (efficiency element), локализованный за 25-40 нуклеотидов до места расщепления или еще дальше, (оптимальная последовательность – UAUUAU) и две U-богатые области, непосредственно примыкающие с обеих сторон к месту расщепления (рис. 1.6).

Аппарат полиаденилирования дрожжей состоит из двух комплексов: CF I (cleavage factor I) и CPF (cleavage and polyadenylation factor) [69, 74, 84-87]. Комплекс CF I подразделяется на два подкомплекса: CF IA и CF IB, а CPF объединяет факторы CF II, PF I (polyadenylation factor 1), поли(А)-полимеразу (Pap1) и еще ряд белков (таблица 1.1). CF I и CF II участвуют в реакции расщепления, а CF I, CPF, Pap 1 и поли(А)-связывающий белок (Pab1) необходимы для реакции полиаденилирования. Несмотря на кажущее различие в сборке комплексов расщепления и полиаденилирования, многие индивидуальные субъединицы белков дрожжей в высокой степени гомологичны соответствующим белкам высших эукариотов. Гомологичные полипептиды обозначены в таблице 1.1 одинаковыми цифрами в третьих колонках. Varabino et al. [88] предполагают, что гены субъединиц факторов полиаденилирования про-мРНК дрожжей и животных эволюционировали из общих древних генов и кодируемые ими полипептиды сохранили аналогичные функции, несмотря на то, что они распределились различным образом между белковыми факторами в результате разной эволюции сигналов полиаденилирования про-мРНК позвоночных и дрожжей.

Таблица 1.1

Белки, участвующие в процессе расщепления и полиаденилирования про-мРНК эукариотов [69, 74, 84-87]

Высшие эукариоты			Дрожжи			
Белок	Субъединица	Гомолог	Белок		Гомолог	
	Symplekin	1	CPF	CF II	Pta1	1
CPSF	CPSF160	2			Yhh1	2
	CPSF100	3			Ydh1	3
	CPSF73	4			Ysh1	4
	CPSF30	5		PF I	Yth1	5
	hFip1	6			Fip1	6
CstF	CstF77	7		Pfs2	-	
	CstF64	8		Mpe1	Mpe1	-
	CstF50	-		Ref2	Ref2	-
CF I _m	CF I _m -25	-		Ssu72	Ssu72	-
	CF I _m 68 (59)	-	Glc7	Glc7	-	
CF II _m	hPcf11	9	Pap 1	Pap 1	12	
	hClp1	10	CF I	CF IA	Rna14	7
WDR33	-	Rna15			8	
PP1- α	-	Pcf11			9	
PP1- β	-	Clp1			10	
PAP	PAP	12		CF IB	Hrp1	-
PABPN1	PABPN1	-		Pab 1p	Pab 1p-70	-

CF I связывается с про-мРНК выше места расщепления, а CPF как на месте расщепления, так и на прилегающих к нему участках. Специфическое и прочное связывание белка Hrp1 (единственный полипептид фактора CF IB) с U/A-богатым элементом про-мРНК EE, способствует связыванию Rna15 (полипептид фактора CF IA) с A-богатым PE, хотя без взаимодействия с

другими белками Rna15 предпочтительно связывается с GU-богатыми последовательностями, как и его гомолог у высших эукариотов полипептид CstF64.

Факторы, представленные в последней строке таблицы (поли(А)-связывающие белки), хотя и не гомологичны друг другу, но выполняют одинаковую функцию. Они ответственны за контроль длины синтезируемых поли(А)-последовательностей [67]. Длина поли(А)-хвостов про-мРНК высших эукариотов составляет ~ 250 нуклеотидов, про-мРНК дрожжей – ~ 70-80 нуклеотидов. Механизм, с помощью которого в про-мРНК эукариотов синтезируются поли(А) хвосты определенной длины, в точности не известен. В параграфе 5.1.3. представлен предложенный нами механизм, основанный на структурных переходах в поли(А).

1.2.2.2. Растения. В настоящее время процесс полиаденилирования про-мРНК растений, в отличие от позвоночных и дрожжей, исследован менее детально. Организация сайтов полиаденилирования про-мРНК растений (рис. 1.6) подобна организации сайтов про-мРНК дрожжей [69]. Поли(А)-сайты состоят из U/UG-богатого FUE (far upstream element), А-богатого NUE (near upstream element) и двух U-богатых элементов, непосредственно примыкающих с обеих сторон к месту расщепления. NUE подобен поли(А)-сигналу у высших эукариотов и PE у дрожжей. Сигнал FUE определяет эффективность поли(А)-сайта. Он может располагаться на расстоянии от 30 нт до 150 нт до места расщепления транскрипта.

В растениях найдены белки, гомологичные всем основным факторам полиаденилирования высших эукариотов и дрожжей, кроме CF I (высшие эукариоты) и Hrp1 (дрожжи). Исследуя взаимодействия между различными полипептидами аппарата полиаденилирования про-мРНК *Arabidopsis thaliana*, Hunt et al. [89, 90] распределили эти белки по трем гипотетическим комплексам, два из которых сравнимы с объединениями белков в высших эукариотах и дрожжах, а третий специфичен для растений. Центрами комплексов являются полипептиды AtCPSF100, AtCLPS и AtFIPS. Первый

комплекс включает AtCPSF160, AtCPSF100, AtCPSF73-I, AtCPSF73-II, AtCPSF30, AtPAPS2 и FY (ортолог Pfs2 дрожжей). За исключением FY и AtPAPS2 этот комплекс соответствует классическому фактору CPSF высших эукариотов. Второй комплекс соответствует факторам CF II_m (высшие эукариоты) и CF IA (дрожжи), он включает полипептиды PCFS4 и CLPS3. Третий комплекс включает FIPS5, три изоформы PAPS, три варианта PABN и другие полипептиды. Авторы предложили модель, согласно которой комплекс полиаденилирования растений состоит из относительно постоянного ядра (включающего субъединицы CPSF и CstF) и многочисленных комбинаций периферийных субъединиц. Такая организация может объяснить дегенеративность сигнальных элементов поли(A)-сайтов про-мРНК растений.

Несмотря на определенные различия в организации поли(A)-сайтов и в составе основных белковых комплексов аппарата полиаденилирования, поли(A)-хвосты мРНК высших эукариотов, дрожжей и растений, выполняют одинаковые функции, взаимодействуя с поли(A)-связывающими белками. Во-первых, цитоплазматический поли(A)-связывающий белок (PABPC), связанный с поли(A)-хвостом мРНК, стимулирует инициацию трансляции посредством специфического связывания с фактором инициации трансляции eIF4F (через субъединицу eIF4G). Это взаимодействие значительно увеличивает сродство eIF4G к кэп-связывающему белку eIF4B и способствует образованию замкнутой в кольцо структуры, облегчающей реинициацию трансляции, а также защищающей мРНК от деградации [4]. Образованию кольцевой структуры может также способствовать взаимодействие фактора терминации трансляции eIF3 с PABPC.

Во-вторых, поли(A)-хвосты в комплексе с различными белками принимают участие в различных процессах стабилизации и деградации мРНК [9]. Деградация мРНК представляет собой многостадийный процесс, в котором деаденилирование является обычно первой и чаще всего лимитирующей стадией. У высших эукариотов поли(A)-хвосты мРНК с начальной длиной ~ 250 нт сначала укорачиваются деаденилазным комплексом PAN (poly(A)

nuclease)2-PAN3 до ~ 110 нт, а затем комплексом CCR4 (carbon catabolite repressor protein 4)-CAF1 (CCR4-associated factor) до 10-18 нт. Начальное укорачивание является дистрибутивной реакцией, а деградация до олиго(А)-хвоста осуществляется процессивно. Комплекс LSm (like-Sm)1-7-Pat1, связанный с олиго(А)-хвостом, стимулирует удаление кэп-структуры комплексом белков DCP (decapping protein)1-DCP2, что обеспечивает доступ экзонуклеазе XRN1 (exoribonuclease 1) к 5'-концу мРНК и экзосоме и/или комплексу Lsm1-7-Pat1 к ее 3'-концу.

ПАВРС играет в контроле времени жизни мРНК двойную роль, способствуя как ее стабилизации, так и дестабилизации. Предполагается, что стабилизация обеспечивается ПАВРСs, связанными с поли(А)-хвостом кооперативно, а дестабилизация обусловлена доставкой и/или стимулированием специфических поли(А)-деаденилаз, с которыми ПАВРС (связанный с поли(А)-хвостом некооперативно) взаимодействует или непосредственно или через взаимодействие с другим белком [9].

Стабильность мРНК может также регулироваться специфически при наличии в ней определенных сигнальных элементов, расположенных чаще всего в 3' НТО и реже в кодирующей области. В частности, в кодирующей области мРНК *c-fos* находится G/A-богатый элемент mCRD (major coding-region determinant), который направляет процесс деаденилирования, сопряженный с трансляцией. С mCRD связывается комплекс из пяти белков: UNR (upstream of N-ras), PAIP1 (PABP-interacting protein 1), AUF1 (A/U-rich RNA-binding factor 1), NSAP1 (NS1-associated protein 1) и ПАВРС, связанный с поли(А)-хвостом [9]. Образование этого комплекса предотвращает деаденилирование мРНК *c-fos*. Во время трансляции рибосома смещает белковый комплекс, разрушая связь между mCRD и поли(А)-хвостом, при этом он, вероятно, становится доступным для действия деаденилазы CCR4, привлекаемой UNR. Механизм, предотвращающий деаденилирование, неизвестен. Предложенные нами механизмы стабилизации мРНК *c-fos* и ряда других мРНК описаны в параграфе 5.1.1.

1.2.2.3. Вирусы высших эукариотов. Наличие в эукариотических мРНК специфических концевых структур – кэп-структуры на 5'-конце и поли(А)-хвоста на 3'-конце – является, как известно, отличительной чертой эукариотических мРНК. Хотя эти концевые структуры не являются необходимыми для трансляции, они в значительной степени стимулируют белковый синтез [91], поэтому вирусы либо обеспечивают ими свои мРНК, либо используют иные подходы, чтобы предотвратить преимущественную трансляцию экпированных и полиаденилированных хозяйских мРНК.

Поли(А)-хвосты мРНК ряда вирусов, например пикорнавирусов, кодируются геномом [92]. У этих вирусов мРНК не имеют на 5'-конце кэп-структуры, но они используют для инициации трансляции элементы IRES (internal ribosome entry site). Следует отметить, что поли(А)-хвосты мРНК могут стимулировать инициацию трансляции независимо от кэп-структуры, способствуя доставке рибосом к внутренним местам инициации трансляции [91].

Существуют вирусные мРНК, которые приобретают поли(А)-хвосты путем повторного копирования вирусной полимеразой олиго(У)-участков матричной РНК. Этим способом образуется, например, поли(А)-хвост мРНК вируса гриппа А ([93] и ссылки в этой работе).

Иная стратегия в отношении процесса полиаденилирования выработана вирусом оспы (*vaccinia virus*). Этот вирус кодирует поли(А)-полимеразу, не обладающую особой специфичностью; она связывается с РНК на участке, содержащем мотив $U_2-N_{25}-U$ [94]. Вирусная РАР полиаденилирует небольшие РНК, в частности тРНК, а также фрагменты вирусных и хозяйских мРНК [95]. Предполагается, что полиаденилированные продукты, связывают хозяйский РАВРС, значительно уменьшая его концентрацию в цитоплазме. Это приводит к селективному ингибированию трансляции хозяйских мРНК, поскольку инициация трансляции вирусных мРНК еще возможна при низких концентрациях РАВРС.

Про-мРНК ряда ДНК-содержащих вирусов «кэпируются» и полиаденилируются соответствующими хозяйскими аппаратами. Они имеют в 3' НТО сигнальные последовательности полиаденилирования, аналогичные последовательностям хозяйских мРНК. К таким транскриптам относится, например, мРНК вируса обезьяны SV 40 [96, 97]. Хозяйские аппараты для полиаденилирования своих про-мРНК используют также ретровирусы. Геном ретровирусов представляет собой РНК, которая обратнo транскрибируется в двухцепочечную ДНК и, проникая через ядерную мембрану, интегрируется в хромосомную ДНК клетки. Рассмотрим подробнее процесс полиаденилирования одного из ретровирусов, а именно вируса иммунодефицита человека 1 (ВИЧ-1), поскольку в параграфе 5.2.3 будет обсуждаться созданная нами модель комплексного поли(А)-сайта про-мРНК этого вируса.

В связи с механизмом обратной транскрипции провирусная ДНК ВИЧ-1 содержит на обоих концах длинные повторы LTRs (long terminal repeats), состоящие из трех участков: U3 (unique 3'), R (repeat) и U5 (unique 5'), а вирусный транскрипт содержит концевые повторы, соответствующие участкам R и U5 в ДНК (рис. 1.8 из нашей работы [98]). Поскольку основным поли(А)-сайт про-мРНК ВИЧ-1 кодируется на границе участков R и U5 провирусной ДНК, его идентичные копии присутствуют на обоих концах транскрипта. В связи с этим использование 5'- и 3'-поли(А)-сайтов должно строго регулироваться для того, чтобы предотвратить преждевременное полиаденилирование на 5'-конце вирусной про-мРНК и стимулировать реакцию на 3'-конце. Ингибирование процесса полиаденилирования на 5'-конце осуществляется, в частности, U1 мяРНП (малый ядерный рибонуклеопротеиновый комплекс), связанным с основным донорным местом сплайсинга [99] (рис. 1. 8), которое присутствует только на 5'-конце транскрипта. В то же время использование основного 3'-поли(А) сайта

стимулируется основным и минорными USEs, которые присутствуют исключительно на 3'-конце транскрипта [100, 101].

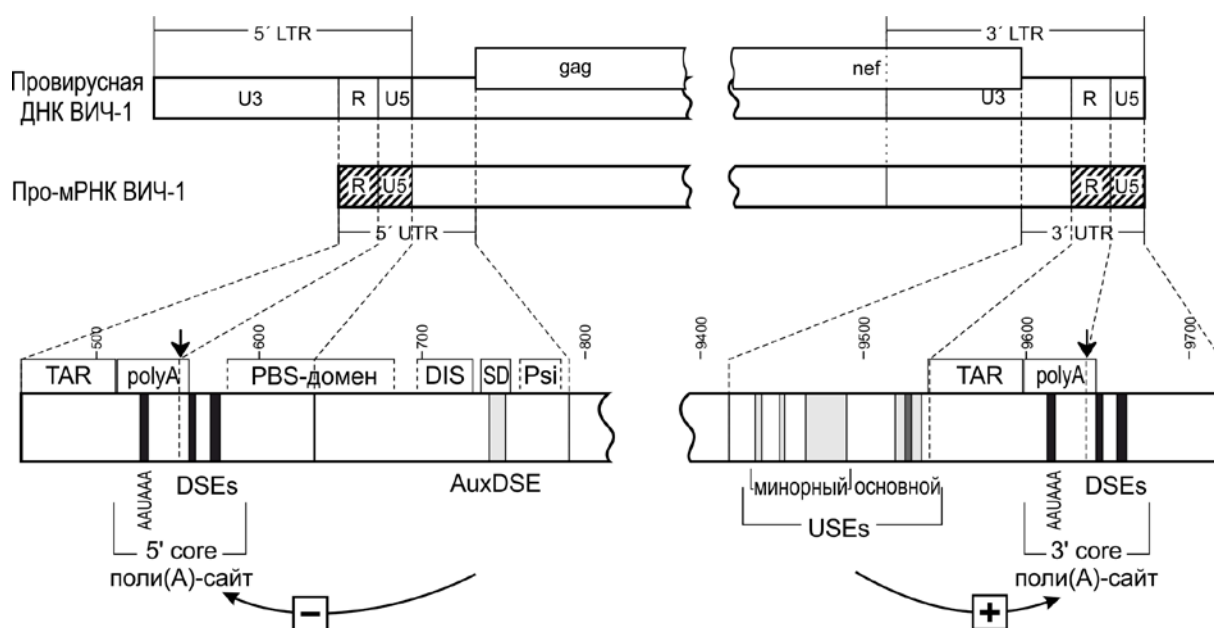


Рис. 1.8. 5' и 3' поли(А)-сайты про-мРНК ВИЧ-1 [98]. Нуклеотиды пронумерованы в соответствии с нумерацией реферативной последовательности (изолят ВИЧ-1 НХВ2, номер K03455 в GenBank). Идентичные области в про-мРНК ВИЧ заштрихованы. Структурные элементы в про-мРНК ВИЧ-1: шпилька TAR (trans-acting responsive element), домен PBS (primer binding site), шпилька DIS (dimerization initiation site), шпилька SD (major splice donor), шпилька Psi (packaging signal). 5'- и 3'-места расщепления транскрипта указаны стрелками. Сигнальные элементы основных 5'- и 3'-поли(А)-сайтов отмечены черным цветом, AuxDSE и USEs – серым, U-богатый тракт основного USE – темно-серым цветом

В литературе опубликован ряд моделей вторичной структуры 5'-концевой области мРНК ВИЧ-1 (например [102-104]) и соответственно 5'-поли(А)-сайта, расположенного в этой области. Очевидно, что все эти модели относятся также и к 5'-концу транскрипта, т.е. к про-мРНК. В то же время,

насколько нам известно, в литературе отсутствуют модели комплексного 3'-поли(А)-сайта про-мРНК ВИЧ-1. На рис. 1.9 показана модель вторичной структуры 5'-конца генома ВИЧ-1, адаптированная из работы Pollom et al. [105], в которой представлена пересмотренная (на основе улучшенного алгоритма сворачивания РНК) структура генома ВИЧ-1 из работы [104].

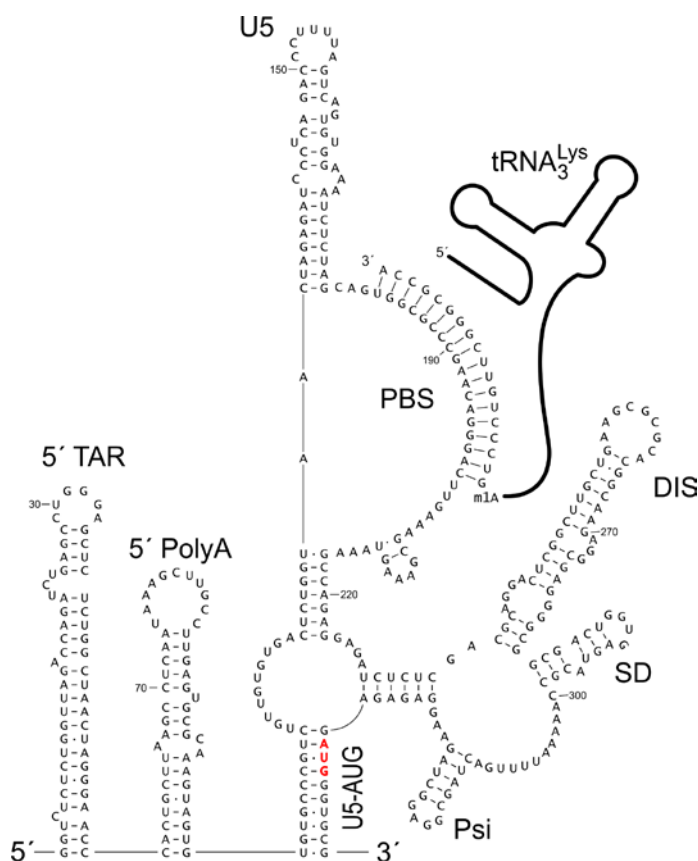


Рис. 1.9. Вторичная структура 5'-концевой области геномной РНК NL4-3 ВИЧ-1 (адаптировано из [105])

5'-концевая область генома ВИЧ-1 сворачивается в несколько шпилек и доменов. Шпилька TAR, связываясь с вирусным белком Tat, стимулирует элонгацию транскрипции провирусной ДНК ВИЧ-1. Домен PBS содержит сайт связывания праймерной тРНК^{Lys}, необходимый для инициации обратной транскрипции. Второй домен, присутствующий в 5'-концевой области транскрипта содержит в верхней части три шпильки: DIS, существенную для

инициации димеризации генома, а также SD и Psi, которые являются элементами комплексного сигнала упаковки геномной РНК.

Шпилька poly(A) содержит основной сигнал полиаденилирования высших эукариотов и их вирусов – гексамер AAUAAA. Гексамер доступен для взаимодействия с CPSF лишь частично, поскольку его первые два нуклеотида вовлечены в образование пар оснований в верхней части шпильки (рис. 1.9). Частичное экранирование гексамера в шпильке polyA играет функциональную роль. Стабильность этой шпильки находится в определенных границах, обеспечивая одновременно и репрессию полиаденилирования на 5'-конце, и стимуляцию на 3'-конце транскрипта [81]. Идентичная шпилька polyA присутствует во всех моделях 5'-конца генома ВИЧ-1, в то время как структура U/GU-богатых «нижних» элементов зависит от модели.

В единственной публикации, посвященной вторичной структуре 3'-концевой области про-мРНК ВИЧ-1, предсказана структура неполного поли(А)-сайта без основного и минорных USEs [106]. В этой работе фрагменты РНК для сворачивания были выбраны, основываясь на определении самодостаточных структурных доменов в полной геномной РНК. USEs не были включены в исследуемые модельные РНК. Изучали последовательности геномов разных субтипов, составленные из двух участков: первый из которых соответствовал примерно 700 нт 5'-концевой области транскрипта, а второй примерно 180 нт 3'-концевой области (от шпильки TAR до конца транскрипта).

Авторы выявили, что при сворачивании таких модельных РНК гексамер AAUAAA в 5'-поли(А)-сайте преимущественно находится в составе шпильки polyA, в то время как в 3'-поли(А)-сайте он встречается и в шпильке, и в составе длинного нерегулярного дуплекса, причем чаще в составе дуплекса. В дуплексе последовательность 3'-концевой шпильки polyA образует пары оснований с областью генома, кодирующей белок GAG, при этом в одноцепочечном состоянии остаются только два нуклеотида гексамера.

Авторы показали, что такой же дуплекс образуется и в модельной РНК, в которой второй участок соответствует концу геномной РНК ВИЧ-1 (от шпильки TAR до места расщепления транскрипта). Образование дуплекса в геномной РНК было экспериментально подтверждено в работе [107].

«Верхние» элементы поли(А)-сайта про-мРНК ВИЧ-1 детально исследованы в работах [100, 101]. Последовательности этих USEs показаны на рис. 1.10 (в координатах про-мРНК НХВ2).

```

9361 ctagcatttc atcacgtggc ccgagagctg catccggagt acttcaagaa ctgctga□cat
9421 cgagcttgct acaagggact ttccgctggg gactttccag ggagggcgtgg cctgggcggg
9481 actggggagt ggcgagccct cagatcctgc atataagcag ctgctttttg cctgtactgg
9541 gtctctctgg ttagaccaga tctgagcctg ggagctctct ggctaactag ggaaccsact
9601 gcttaagcct caataaagct tgccttgagt gcttcaagta gtgtgtgccc gtctgtgtg
9661 tgactctggg aactagagat ccctcagacc cttttagtca gtgtggaaaa tctctagca

```

Рис. 1.10. Последовательность 3'-концевой области про-мРНК ВИЧ-1 НХВ2. Стоп-кодон обведен рамкой. Сигнальные элементы основного поли(А)-сайта выделены жирным шрифтом. Место расщепления транскрипта указано стрелкой. Тракты, соответствующие основному и минорному USEs [100], подчеркнуты двумя и одной линией соответственно. Пунктирные линии и линии, состоящие из точек, указывают местоположение кластерных замен оснований в промоторной и энхансерной областях соответственно [101]

Согласно Valsamakis et al. [100], тракт CAGCUGCUUUUUGCCUGU (9518-9535 нт), расположенный непосредственно перед шпилькой TAR, является первой частью основного USE, которая оказывает наибольшее влияние на процесс полиаденилирования. Его удаление приводит к уменьшению эффективности реакции на 70 %. Поскольку вторая часть основного USE (с меньшим влиянием на процесс) включает первые 14 нт шпильки TAR, удаление которых в первую очередь будет влиять на процесс инициации транскрипции, мы приняли за основной USE его первую часть.

Тракт 9464-9481 нт, делеция которого приводит к уменьшению эффективности полиаденилирования только на 20 %, назван Valsamakis et al. [100] минорным USE. Кластеры точечных замен оснований в энхансерной и промоторной областях провирусной ДНК (рис. 1.10) приводят к снижению эффективности реакции полиаденилирования всего лишь на 13 % и 17 % соответственно [101].

Основной USE является местом связывания белка CPSF [108], который, по-видимому, связывается с U-богатой областью этого элемента посредством субъединицы hFip1 (Factor interacting with PAP) [109]. Это взаимодействие приводит к повышению стабильности комплекса CPSF-про-mRNA и увеличению эффективности процесса полиаденилирования. Klasens et al. [81] предложили следующий механизм функционирования основного USE. Этот элемент служит местом первоначального предварительного связывания CPSF с про-mRNA. Во время «дыхания» шпильки polyA, когда гексамер AAUAAA временно становится полностью доступен, CPSF связывается с ним, или взаимодействуя одновременно и с USE и гексамером, или перемещаясь с дополнительного сигнала на основной.

1.2.3. Тракты поли/олиго(C) и олиго(U). Участки поли(C) длиной от нескольких десятков до нескольких сотен нуклеотидов содержатся в 5'-нетранслируемой области геномов кардиовирусов и афтоввирусов из семейства пикорнавирусов [110]. Роль поли(C)-трактов в репликации этих вирусов в настоящее время неясна [111]. В частности, наличие длинного поли(C)-участка более существенно при инфицировании мышей для вируса Менго, чем для вируса энцефаломиокардита. Неясно также, какой именно гомополимер является в действительности функционально значимым – поли(C)-тракт, комплементарный ему поли(G)-участок в минус цепи РНК вируса, двухцепочечный тракт поли(G)·поли(C), образующийся как промежуточное соединение в процессе синтеза минус цепи, или же все эти участки.

Следует также отметить, что хозяйский поли(C)-связывающий белок PCBP1/PCBP2 [5], связывается с 5' и 3' концами геномной РНК афтоввируса FMDV (foot-and-mouth disease), вероятно, стабилизируя контакт между

концами генома, образованный РНК-РНК взаимодействиями [112]. Неясно, какую роль во взаимодействии между РСВР и геномной РНК FMDV играет поли(С)-тракт, поскольку поли(С) способен прочно связываться с РСВР и конкурентно блокировать, в свою очередь, связывание этого белка с клеточными и вирусными РНК [113].

Поли(С)-связывающие белки подразделяются на две группы: белки hnRNPs К/J и α CPs (α -complex proteins) [5]. Белки α CP-1/ α CP-2 имеют альтернативные названия РСВР1/РСВР2 или hnRNP-E1/hnRNP-E2. РСВРs содержат три домена КН (hnRNP K homology), которые могут одновременно взаимодействовать с тремя С-богатыми трактами. Стабильность и специфичность связывания РСВРs и их КН-доменов с С-богатыми олигонуклеотидами различны. В частности, Thisted et al. [114] определили методом SELEX (systematic evolution of ligands by exponential enrichment) оптимальные мишени для двух широко распространенных РСВРs: hnRNP К и α CP-2KL. В случае hnRNP К мишенью является один короткий олиго(С)-тракт, а в случае α CP-2KL – три таких тракта.

Поли(С)-связывающие белки участвуют в процессах стабилизации и трансляции клеточных и вирусных мРНК, а также репликации вирусных РНК [5, 115-117]. РСВРs связываются с С-богатыми последовательностями, состоящими из коротких олиго(С)-трактов (2-5 нт), разделенных небольшими (1-8 нт) линкерами из разных оснований. Например, белок hnRNP-E1, существенный для трансляции мРНК FR- α (folate receptor) [118], связывается с последовательностью CUCCAUUCCCACUCCCU (С-повторы подчеркнуты). Более длинный функциональный олиго(С)-тракт (14 нт) найден в мРНК pp2AC лягушки [119]. Этот тракт является цитоплазматическим элементом полиаденилирования, с которым специфически связывается белок, гомологичный белку РСВР-2 человека.

Ji et al. [120] показали, что С-богатая последовательность, связанная с α CP, не только стабилизирует мРНК α -глобина человека в цитоплазме, но и

служит в качестве USE в процессах расщепления и полиаденилирования транскрипта. Белок стимулирует эти процессы, взаимодействуя с компонентами аппарата полиаденилирования – CPSF, симплекином и PAP. Недавно Ji et al. [121] выявили в линии кроветворных клеток человека целую группу про-мРНК с C-богатыми USEs, расположенными в основном в пределах 30-40 нт до поли(А)-сигналов. C-богатые последовательности, взаимодействующие с α CP, также участвуют в процессе сплайсинга про-мРНК α -глобина [122].

Определенные биологические функции могут выполнять не только поли(С)-участки в РНК или одиночные олиго(С)-тракты, но и кластеры олиго(С)-трактов способные образовывать i-мотив. I-мотивы РНК значительно менее стабильны, чем i-мотивы ДНК [57] и, насколько нам известно, их исследованию посвящены лишь единичные работы. В частности, Lacroix et al. [123] показали, что температура плавления i-тетраплекса, образованного трактом CCCUCCCUUUUCCCUCCS, равна 25 °С при рН 5,6, а аналогичной структуры ДНК – 54 °С. Brooks et al. [124] подразделили внутримолекулярные i-мотивы ДНК на два класса по размеру петель. Например, i-мотив класса I в промоторном элементе RET содержит пять пар оснований C⁺-C с петлями 2 : 3 : 2 нт, а i-мотив класса II в Vcl-2 имеет семь пар C⁺-C с петлями 8 : 5 : 7 нт. Стабильность i-мотива зависит от числа пар C⁺-C и специфических взаимодействий между основаниями в петлях [124, 125]. В частности, петля длиной 14 нт в i-мотиве, образованном вблизи промоторной области n-Мус, сворачивается в шпильку, стабилизирующую структуру [126]. Dai et al. [127] показали также, что длина каждой из петель в стабильном i-мотиве должна составлять не менее 2 нт.

Функционально значимые олигомерные тракты уридиловой кислоты довольно часто встречаются в составе клеточных РНК как внутри, так и на 3'-конце молекул. Так, например, добавление олиго(U)-хвостов, осуществляемое концевыми уридил-трансферазами (TUTases, terminal uridylyl transferases),

требуется для деградации гистоновых мРНК, про-микроРНК, продуктов расщепления мРНК в процессе РНК-интерференции и др. [128, 129]. Обнаружение рядом исследовательских групп в 2013 г. экзорибонуклеазы Dis3L2, ферментативная активность которой значительно стимулируется наличием 3'-концевого олиго(U)-тракта в РНК, позволило Gallouzi & Wilusz [130] предположить, что присоединение олиго(U)-хвостов к мРНК может оказаться еще одним широко распространенным механизмом их деградации, наряду с механизмами, включающими деаденилирование.

Что касается внутренних олиго(U)-трактов, то они часто встречаются в 3' UTR мРНК, содержащих сигнал ARE (AU-rich element), внутри или вблизи этого элемента. Олиго(U)-тракты служат местами связывания для Hu-белков, гомологичных белку дрозофилы ELAV (embryonic lethal abnormal vision), участвующих в контроле времени жизни мРНК [131]. Сегменты олиго(U) длиной до трех десятков остатков уридиловой кислоты входят в состав элементов цитоплазматического полиаденилирования CPE (cytoplasmic polyadenylation element), расположенных в 3' UTR некоторых матерински наследуемых мРНК. Эти элементы, с которыми связывается белок CPEB, контролируют полиаденилирование и последующую трансляцию мРНК во время раннего эмбриогенеза [132]. Интересно отметить, что в отличие от CPEB, родственный ему белок CPEB4, связывается с элементом UUUUU, только если последний находится во внутренней петле шпильки, т. е. этот белок узнает олиго(U) в определенной вторичной структуре [133]. Белки CPEB2-4 и CPEB выполняют различные функции, в частности CPEB3 регулирует трансляцию.

1.2.4. Тракты олиго(G). Короткие олиго(G)-тракты были обнаружены в составе «нижнего» вспомогательного элемента GGGGGAGGUGUGGG в сайте полиаденилирования про-мРНК SV40 L (simian virus late). С этим элементом связывается белок гетерогенного ядерного рибонуклеопротеинового комплекса (гяРНП) H/H', стимулирующий процесс полиаденилирования [134]. Позднее было показано, что G-богатая последовательность (GRS) является

специфическим вспомогательным элементом поли(А)-сайта не только в случае про-мРНК SV40 L. GRSEs, расположенные дальше основного сайта полиаденилирования, были найдены примерно в 34 % про-мРНК высших эукариотов [135]. Все исследованные G-богатые элементы связывались с белком гяРПП Н/Н' и это взаимодействие приводило к стимулированию процесса полиаденилирования. Однако сродство белка Н/Н' к разным GRSEs значительно различалось, что влияло на их способность позитивно влиять на реакцию. Авторы работы [135] предположили, что GRS являются распространенными вспомогательными элементами поли(А)-сайтов высших эукариотов.

Белок гяРПП Н/Н' принадлежит к подсемейству Н белков гяРПП, в которое входят также гяРПП F и гяРПП 2Н9 ([7 и ссылки в этой работе]). ГяРПП Н и Н', степень идентичности которых составляет 96 %, а также гяРПП F содержат три мотива qRRMs (quasi-RNA recognition motifs,), а гяРПП 2Н9 – два. Все белки группы Н специфически взаимодействуют с последовательностями РНК, содержащими тракт GGGA, а гяРПП Н/Н' узнает также тракты GGGGGC [7] и GGGU [136].

Учитывая эти данные, можно предположить, что другие белки группы Н также могут связываться с G-богатыми AuxDSE, так как большинство вспомогательных «нижних» элементов, найденных в работе [135], содержат тракты GGGA, а остальные содержат тракты GGGU или GGGGGC. Роль связывания гяРПП 2Н9 с GRS в процессе полиаденилирования в настоящее время неясна. Белок гяРПП F способен ингибировать этот процесс [137]. Авторы работы [137] предполагают, что связывание гяРПП F или гетеродимера Н'/F вблизи основного «нижнего» элемента поли(А)-сайта может ингибировать взаимодействие CstF с про-мРНК и/или реакцию расщепления. Samatanga et al. [138] показали, что третий qRRM белка гяРПП F специфически связывается исключительно с одноцепочечными G-трактами. Они предложили модель, согласно которой гяРПП F может связываться в

процессе транскрипции с G-трактами про-мРНК, имеющими потенциал для образования G-квадруплексов, и предотвращать образование этих структур.

Все белки группы Н белков гяРНП принимают участие в процессе сплайсинга, однако механизм их действия в настоящее время неизвестен. В некоторых случаях гяРНП Н может стерически мешать связыванию белков SR с регуляторными элементами сплайсинга, ингибируя их функцию, например [136]. В других случаях, например [139], взаимодействие гяРНП Н со своим местом связывания необходимо для сборки белков на регуляторных элементах сплайсинга и их функционирования. Аналогичная роль была предложена для гяРНП Н в процессе полиаденилирования. Arhin et al. [135] предположили, что этот белок непосредственно взаимодействует с CstF, стимулируя сборку факторов полиаденилирования на основном поли(А)-сайте. В качестве альтернативы, авторы предположили, что белок Н может изменять структуру синтезируемого транскрипта таким образом, чтобы основные элементы поли(А) сайта становились доступными для факторов полиаденилирования. В этой связи интересен тот факт, что места связывания белков гяРНП Н/Н'/F могут образовывать G-квадруплексы. Этот вопрос более подробно будет обсуждаться в параграфе 5.2.2.

Millevoi et al. [140] показали, что гяРНП Н может стимулировать процесс полиаденилирования не только посредством взаимодействия с CstF, но и с PAP. Они также показали, что комплексы гяРНП Н с G-богатыми трактами, расположенными вблизи основного сайта полиаденилирования, могут быть стабилизированы фактором сплайсинга РТВ (polypyrimidine tract-binding protein), связанным с USE.

1.2.5. G-квадруплексы РНК. Биологи начали проявлять интерес к четырехцепочечным структурам гомополинуклеотидов в начале 90-х годов прошлого столетия, когда выяснилось, что G-квадруплексы ДНК могут быть вовлечены в такие важные клеточные процессы как мейоз, регуляция транскрипции и др. [141]. Большинство работ было посвящено исследованию G-квадруплексов ДНК и только в редких работах сообщалось о квадруплексах

РНК. В частности Christiansen et al. [142] показали, что G-квадруплекс образуется в 3'-НТО мРНК IGF-II (insulin-like growth factor). Авторы предположили, что наличие этой четырехцепочечной структуры предотвращает спаривание сайта эндонуклеотического расщепления с другими участками РНК, обеспечивая его доступность для взаимодействия с соответствующими макромолекулами.

Мы первые предположили, что G-квадруплексы РНК могут играть роль в процессе полиаденилирования, обнаружив кластеры G-повторов в области, расположенной непосредственно за основным сайтом полиаденилирования в ряде про-мРНК высших эукариотов и их вирусов [65]. В дальнейшем выяснилось, что кластеры олиго(G)-трактов, способные самоассоциироваться с образованием четырехцепочечных G-квадруплексов, являются важными регуляторами многих биологических процессов, в которых участвуют про-мРНК и мРНК, но механизмы их функционирования остаются еще во многом неясными [24, 143-145].

G-квадруплексы состоят из G-тетрад (рис. 1.11), расположенных друг над другом и стабилизированных стэкинг-взаимодействиями [146]. Стабильность этих структур зависит от присутствия катиона, который располагается внутри или между тетрадами и действует как «координатор» структуры. Конформационные свойства G-квадруплексов определяются природой центрального иона, относительной ориентацией цепей (параллельная, антипараллельная), конформацией гликозидной связи (*syn*, *anti*), последовательностью и длиной петель, соединяющих олиго(G)-тракты, количеством цепей в квадруплексе и количеством G-тетрад. Эти структуры могут быть образованы четырьмя (рис. 1.11 *a, б*), двумя (рис. 1.11 *в, г*) или одной (рис. 1.11 *д, е*) макромолекулой (в случае ее самоассоциации). Цепи в G-квадруплексах ДНК могут быть ориентированы как параллельно, так и антипараллельно (рис. 1.11 *a-e*) а в G-квадруплексах РНК [24, 144] они ориентированы преимущественно параллельно (рис. 1.11 *a, e*).

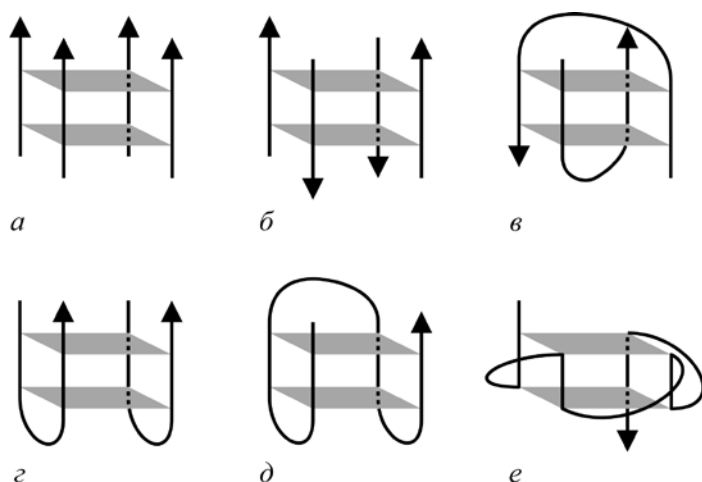


Рис. 1.11. Схемы образования тетрамеров (*a*, *b*), димеров (*c*, *d*) и мономеров (*e*) G-квадруплексов.

Адаптировано из [146]

Наличие 2'-ОН группы в РНК обуславливает преимущественно параллельную топологию G-квадруплексов РНК и их повышенную стабильность относительно G-квадруплексов ДНК [147]. Параллельная ориентация цепей является следствием стерических ограничений, накладываемых гидроксильной группой, а повышенная стабильность G-квадруплексов РНК обусловлена тем, что 2'-ОН группа контактирует с множеством акцепторов водородной связи внутри молекулы, в частности с атомами кислорода фосфатов и остова цепи.

Квадруплексы, образованные из двух цепей, имеют две внутренние петли, соединяющие тетрады, а во внутримолекулярной структуре их три. Петли соединяют основания (расположенные диагонально или соседние) либо в одной и той же тетраде (рис. 1.11 *c-d*), либо в разных тетрадах (рис. 1.11 *e*). На рис. 1.11 показаны лишь некоторые возможные ориентации цепей и способы соединения тетрад петлями.

Квадруплекс, состоящий из G-тетрад, может быть стабилизирован стэкинг-взаимодействиями с парами оснований или триадами [146, 148]. В настоящее время идентифицированы такие триады как A·(A-T) [148], G·(C-A) [149], G·(T-T) [150], T·(A-A) [151] и T·T·T [150]. Более того, в образование квадруплекса могут быть вовлечены тетрады из других оснований: G·C·G·C [148], A·T·A·T [152], A·A·A·A [153], C·C·C·C [154], T·T·T·T [155], U·U·U·U [156] и T·G·T·G [157]. В составе квадруплексов были найдены также

смешанные циклические структуры из пяти, шести и семи оснований, например гексада G·G·A·G·G·A [158].

В настоящее время правила, определяющие стабильность квадруплексов, полностью не изучены. В общем, стабильность G-квадруплексов РНК и ДНК уменьшается с увеличением длины петель и G-квадруплексы РНК более стабильны, чем эквивалентные структуры ДНК ([159, 160] и ссылки в этих работах). В отличие от G-квадруплексов ДНК, стабильность квадруплексов РНК зависит только от общей длины петель, но не от их последовательностей [160]. Этот вывод был сделан авторами для квадруплексов РНК с тремя G-тетрадами и общей длиной петель 4, 5 или 7 нт. Показано также, что структуры с тремя G-тетрадами и общей длиной петель 12 или 15 нт все еще являются достаточно стабильными [160]. Также известно, что стабильные квадруплексы ДНК с тремя G-тетрадами могут быть образованы даже в том случае, когда одна из петель является очень длинной (15 нт и больше), а две другие очень короткие [159]. Более того G-тракты в G-квадруплексах ДНК с тремя тетрадами могут прерываться различным образом другими основаниями, которые образуют балджи между тетрадами, при этом квадруплексы во многих случаях остаются достаточно стабильными [161].

В недавней работе [162] проведено детальное исследование G-квадруплексов РНК, как с тремя, так и с двумя тетрадами. В частности, показано, что мономерные квадруплексы с тремя тетрадами могут быть высокостабильными ($T_m = 60-70$ °C), даже если каждая из трех петель имеет длину 15 нт. Стабильность квадруплексов РНК с двумя тетрадами зависит от длины петель в значительно большей степени. При концентрации $K^+ = 100$ мМ только квадруплексы с общей длиной петель не более 7 нт относительно стабильны ($T_m \geq 40$ °C). Однако квадруплексы с более длинными петлями могут быть стабильными при более высокой (физиологической) концентрации K^+ . Кроме того, стабильность G-квадруплексов зависит от состава петель. В частности, вышеприведенные данные для G-квадруплексов РНК с двумя тетрадами относятся к петлям, состоящих из одних уридиновых остатков. Для

структур, в которых каждая из петель состоит из 1 нт, T_m квадруплексов с петлями, состоящими из А, U и С равняется 70 °С, 60 °С и 43 °С соответственно [162].

В литературе описано несколько стабильных G-квадруплексов РНК с двумя G-тетрадами, например [8 (Supplemental), 163]. Длина петель в этих квадруплексах варьирует от 1 до 6 нт, а общая длина – от 3 нт до 9 нт.

1.3. Структура области димеризации геномной РНК ВИЧ-1

Вирионы вируса иммунодефицита человека, содержат две гомологичные одноцепочечные геномные РНК, которые нековалентно соединены друг с другом [164]. Наиболее стабильно они связаны между собой на участке, расположенном вблизи их 5'-концов. Считается, что диплоидный геном обеспечивает дополнительную матрицу в случае повреждения одной из геномных РНК. Кроме того, в результате генетической рекомбинации во время обратной транскрипции увеличивается скорость эволюции вируса и улучшается его приспособляемость [165].

В начале 90-х годов прошлого столетия рядом авторов было сделано предположение о том, что димеризация геномной РНК вируса ВИЧ-1 может осуществляться посредством образования межмолекулярных четырехцепочечных структур, содержащих пуриновые квартеты. В частности, в работах [166] и [167] показано, что G-богатые последовательности РНК ВИЧ-1 длиной примерно 100 нт, расположенные в 5'-концевой области гена *gag*, способны самоассоциироваться *in vitro* с образованием G-квадруплексов.

Позднее выяснилось, что самоассоциироваться могут также фрагменты РНК ВИЧ-1, включающие шпильку DIS (рис. 1.9), расположенную до GRS, [168, 169]. Преимущество шпильки DIS по сравнению с GRS в процессе инициации димеризации наглядно продемонстрировано в работе [170]. Авторы исследовали в одинаковых условиях димеризацию ряда фрагментов РНК ВИЧ-1_{MAL}, в том числе фрагментов 1-311 нт (со шпилькой DIS), 311-415 нт (с GRS)

и 1-615 нт (с DIS и GRS). Они показали, что параметры димеризации фрагментов 1-311 нт и 1-615 нт близки между собой и заметно отличаются от параметров димеризации фрагмента 311-415. В частности, фрагмент 1-311 нт димеризовался в 10 раз быстрее, чем фрагмент 311-415 нт. Эти экспериментальные результаты свидетельствуют о том, что шпилька DIS играет ведущую роль в инициации димеризации 5'-концевой области РНК HIV-1, но не исключают участия G-квадруплексов в других стадиях процесса димеризации.

Специфическими особенностями шпильки DIS являются палиндром длиной 6 нт и примыкающие к нему пуриновые основания в апикальной области, а также внутренняя петля в стебле шпильки. На первой стадии димеризации палиндромы шпилек DIS двух РНК образуют двойную спираль (рис. 1.12 *а, б*). Такая структура димера получила название «целующиеся петли» (kissing loop dimer) [168, 169]. При определенных условиях (присутствие нуклеокапсидного белка NC или повышенная температура ([171] и ссылки в этой работе)) осуществляется конформационный переход, в результате которого область контакта удлиняется и образуется линейный дуплекс (рис. 1.12 *в*). В этом дуплексе спарены не только палиндромы концевых петель шпилек DIS, но и участки, образующие их стебли. Пурины в основании апикальной петли играют важную роль в димеризации шпилек DIS, в частности, стабилизируя двойную спираль, образованную палиндромами ([172] и ссылки в этой работе).

Полагают, что *in vivo* структурный переход из димера «целующиеся петли» в линейный дуплекс осуществляется во время созревания вирионов [171]. Белок NC и другие белки, образующиеся в результате протеолитического расщепления полипротеинов Gag и Gag-Pro-Pol, необходимы для образования зрелых стабильных димеров [173].

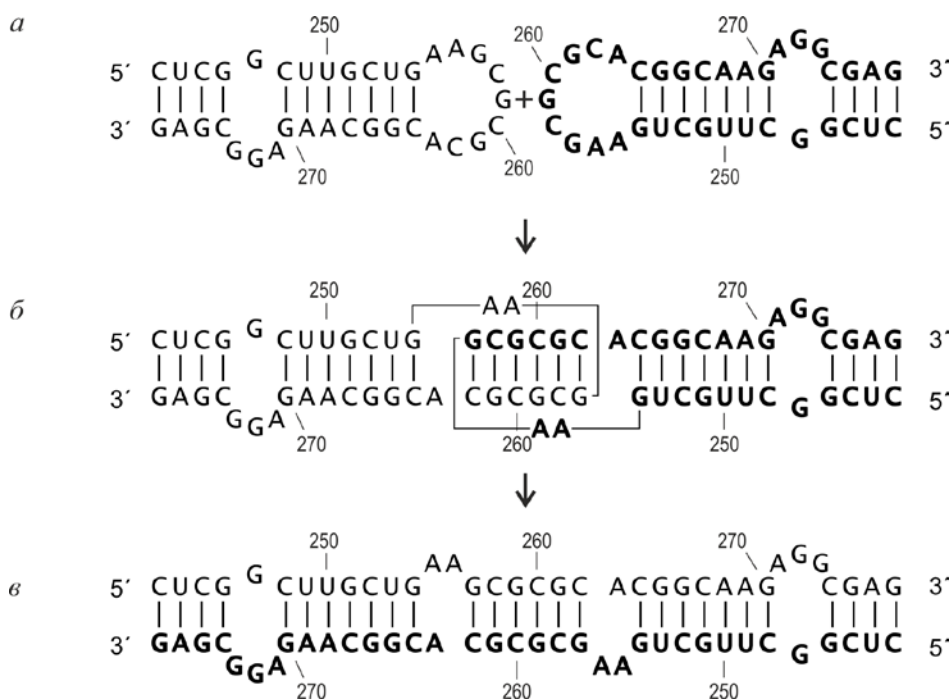


Рис. 1.12. Димеризация шпильки DIS; *а* – шпильки DIS, *б* – димер «целующиеся» петли, *в* – линейный дуплекс

Несмотря на то, что шпилька DIS играет важную роль в репликации вируса ВИЧ-1, тем не менее, она не является абсолютно необходимой для этого процесса [164]. Замены оснований или делеции в шпильке могут приводить к дефектам в упаковке вируса и снижению его инфекционности, однако при этом они могут не влиять на стабильность димеров. Димеры, выделенные из мутантных вирусов, имеют примерно такую же температуру плавления (T_m), как и выделенные из вирионов дикого типа. Авторы ряда работ, например [174, 175], предполагают, что контакты между шпильками DIS существенны лишь для инициации процесса димеризации, а стабильность димеров определяется контактами в других областях РНК. Какие же области РНК ВИЧ-1, кроме шпильки DIS, могут участвовать в процессе димеризации генома?

В качестве возможной DLS (dimer linkage structure) рассматривалась, например, последовательность AAGCUU [174], локализованная в концевой петле шпильки poly(A), однако экспериментальная проверка, проведенная

Расселом и др. [176], не подтвердила это предположение. Авторы [176] нашли другие участки, важные для процесса димеризации, а именно – два GU-богатых тракта, расположенных в области U5. По мнению Рассела и соавт. [176], GU-тракты не играют структурную роль, а могут быть непосредственно вовлечены во взаимодействие между двумя мономерными РНК. Однако, учитывая наличие дуплекса U5-AUG в современных моделях 5'-концевой части геномной РНК ВИЧ-1 (рис. 1.9), представляется более вероятным, что эти тракты все же важны именно для поддержания определенной структуры DLS. Дуплекс U5-AUG способствует сохранению шпилек DIS, SD и Psi и их эффективному функционированию, а его разрушение должно в определенной мере дестабилизировать шпильки, включая шпильку DIS, и ингибировать димеризацию, а также другие процессы, что, вероятно, и наблюдалось в работе [176].

Наиболее вероятным другим участником в димеризации генома ВИЧ-1 является шпилька TAR [177-180]. Верхняя часть этой шпилки содержит палиндромную последовательность GGGAGCUCUC, в которой только первые четыре нуклеотида локализованы в апикальной петле и свободны для взаимодействия с другими последовательностями, а остальные нуклеотиды участвуют в образовании пар оснований в стебле шпилки. В настоящее время роль палиндрома в димеризации геномной РНК ВИЧ-1 с участием шпилки TAR неясна. Jalalirad et al. [180] считают, что палиндром не образует комплекс с палиндромом второй шпилки TAR. По их мнению, основную роль играет бальдж длиной 3 нт, расположенный в верхней части стебля шпилки. Он может взаимодействовать с другим участком в геноме или служить в качестве гибкого соединения между верхней и нижней частями стебля, или связываться с каким-нибудь белком. Мы предполагаем (параграф 5.2.1.2.), что палиндром все же может участвовать в процессе димеризации, а бальдж, связанный с определенным белком, может стимулировать этот процесс.

Мы также разделяем мнение авторов обзора [164] о том, что и G-квадруплексы могут играть роль в димеризации генома ВИЧ-1. Они могут

участвовать в стабилизации димера генома на стадии созревании вируса в присутствии нуклеокапсидного белка. В параграфе 5.2.1.1. представлены наши структурные модели области димеризации генома ВИЧ-1, включающие как двухцепочечные, так и четырехцепочечные участки.

Предположение о наличии более чем одного контакта между двумя геномными РНК ВИЧ-1 согласуется с данными о наличии множественных контактов в димерных геномах других ретровирусов. В частности, согласно модели, предложенной Purzycka et al. [181], в первой стадии димеризации геномной РНК ВИЧ-2 принимают участие TAR и SL 1, отличающиеся по структуре и последовательности от соответствующих элементов TAR и DIS в геноме ВИЧ-1. В геноме ВИЧ-2 TAR представляет собой домен, состоящий из трех шпилек (I, II, III). Палиндромная последовательность UUGGCCGG охватывает апикальную петлю и один нуклеотид стебля шпильки III. SL 1, как и DIS, является отдельной шпилькой с палиндромом GGUACC, расположенном в апикальной петле. Пять нуклеотидов в нижней части 5'-цепи стебля шпильки SL 1 входят в состав палиндромной последовательности GGAGUGCUCC (PAL). Нуклеотиды, образующие пары оснований в SL 1, подчеркнуты. Палиндромные последовательности в апикальных областях шпилек TAR III и SL 1 участвуют в образовании димера генома ВИЧ-2 с относительно низкой стабильностью (loose димер). При созревании вирусной частицы происходит дестабилизация шпильки SL 1 под действием белка NCp8 и высвобождение PAL. Взаимодействие между PAL-последовательностями двух геномов приводит к стабилизации димера.

В димеризацию близкородственных ретровирусов MuLV (Moloney murine leukemia virus) и MuSV (murine sarcoma virus) вовлечены четыре структурных элемента [182]. Две палиндромные последовательности (PAL1 и PAL2) образуют протяженные межмолекулярные дуплексы. Две шпильки (SL1 и SL2) с апикальными петлями из четырех нуклеотидов (GACG), расположенные за PAL2, образуют стабильные комплексы «целующиеся петли», содержащие только две межмолекулярные пары G·C [183].

1.4. Выводы к разделу 1

Все гомополирибонуклеотиды могут принимать различные формы в зависимости от условий среды. Поли(А) при нейтральных и слабощелочных рН существует в двух различных одноцепочечных формах. При протонировании адениновых остатков этот гомополинуклеотид может находиться в двух различных двухспиральных формах, а также в «замороженной» форме.

Поли(С), как и поли(А), переходит при протонировании оснований из одноцепочечной структуры в двухцепочечную, но в отличие от поли(А), при высокой степени протонирования дуплекс поли(С) разрушается.

Поли(У) при физиологических условиях находится в одноцепочечной неспиральной форме. При пониженной температуре он может образовывать шпильчатые структуры при высоких концентрациях монокатионов или в присутствии поликатионов.

Структурные переходы в поли(Г) происходят как при протонировании оснований, так и при изменении ионной силы раствора. При нейтральных рН при повышении ионной силы полимер переходит из одноцепочечной спирали в четырехцепочечную (G-квадруплекс), а при низких рН находится в двухцепочечной форме.

Детальные механизмы структурных переходов различных гомополирибонуклеотидов полностью не выяснены, а для ряда форм отсутствуют какие-либо структурные модели. Например, в случае поли(А) такие модели не представлены для трех из пяти различных форм гомополинуклеотида.

Все гомополи(олиго)рибонуклеотиды встречаются в природных РНК и выполняют определенные регуляторные функции. Наиболее известными из них являются поли(А)-последовательности и олиго(Г)-тракты. Поли(А) в виде поли(А)-хвостов встречаются в про-мРНК всех живых организмов от бактерий до человека. Участие поли(А)-хвостов в контроле времени жизни мРНК

является общей функцией для прокариотов и эукариотов, а в эукариотах они также участвуют в процессах транспорта мРНК из ядра в цитоплазму и трансляции. С поли/олиго(А)-трактами специфически взаимодействуют различные поли(А)-связывающие белки. Механизмы функционирования поли(А)-трактов в ряде биологических процессов в настоящее время не установлены. К таким процессам, в частности, относятся: терминация полиаденилирования про-мРНК, саморегуляция синтеза цитоплазматического РАВ и стабилизация мРНК ELAV-подобными белками.

Последовательности РНК, содержащие олиго(Г)-повторы, могут сворачиваться в мономерные, димерные и тетрамерные G-квадруплексы. В настоящее время признано, что G-квадруплексы РНК являются регуляторами многих биологических процессов. В частности, внутримолекулярные (мономерные) G-квадруплексы участвуют в процессах сплайсинга и полиаденилирования про-мРНК, а также в процессе трансляции мРНК. Межмолекулярные G-квадруплексы, образованные двумя молекулами РНК, могут играть важную роль в димеризации геномов ретровирусов, включая геном вируса иммунодефицита человека.

Механизмы функционирования G-квадруплексов, так же как и поли(А)-трактов, остаются еще во многом неясными и далеко не все биологические функции этих и других гомополирибонуклеотидов выявлены в настоящее время. Например, кажется маловероятным, чтобы в биологических системах использовалась только одна форма поли(А) из пяти различных структур. С другой стороны, трудно представить, что различные поли(А)-связывающие белки, включая их посттрансляционные модификации и комплексы с белками-партнерами, модулирующими их сродство к РНК, взаимодействуют с единственной формой поли(А). Таким образом, необходимо дальнейшее изучение физико-химических характеристик гомополирибонуклеотидов и их структурных переходов, а также выявление возможных биологических процессов, происходящих с их участием.

РАЗДЕЛ 2

МАТЕРИАЛЫ И МЕТОДЫ

2.1. Измерение осмотического давления ДНК и определение эффективного диаметра двойной спирали

2.1.1. Препараты. Использовали ДНК плазмиды Col E1 с молекулярным весом $M = 4,35 \cdot 10^6$ [184], выделенную по методу [185] и любезно предоставленную Ю. Л. Любченко. ДНК обрабатывали панкреатической ДНКазой I (фирмы «Sigma», США) по [186] для получения открытой кольцевой формы. Линейную ДНК получали из открытой кольцевой обработкой рестриктазой EcoR1, выделенной по методу [187] и любезно представленной С. А. Козьявкиным. Реакцию проводили в буфере 50 мМ трис-HCl, pH 8,0, содержащем 7,5 мМ MgCl₂ и 50 мМ NaCl, при концентрации ДНК 100 мкг/мл в течение 1-го часа при 36 °С, после чего проводили фенольную депротеинизацию.

Количественный состав растворов ДНК определяли методом электрофореза. Электрофорез проводили в трубках с 0,8 % агарозным гелем в буфере 40 мМ трис-CH₃COOH, pH 7,9, содержащем 5 мМ NaCH₃COOH и 1 мМ ЭДТА, при напряженности поля 5 В/см. Окрашенные бромистым этидием (1 мкг/мл) гели сканировали на флуориметре, сконструированном А. С. Боровиком. Результаты электрофореза представлены на рис. 2.1. В препарате № 1 содержалось 80 % линейной формы ДНК, ~ 5 % кольцевых молекул и 15 % линейных фрагментов ДНК Col E1 меньшей молекулярной массы (рис. 2.1 *a*), препарат № 2 содержал 75-80 % ДНК в открытой кольцевой форме, незначительное количество сверхспирализованных молекул и димеров и 17-21 % линейных молекул (рис. 2.1 *b*). Растворы ДНК концентрировали следующим образом. Их помещали в диализный мешок, который снаружи обдували воздухом с помощью настольного вентилятора при температуре 5-8 °С.

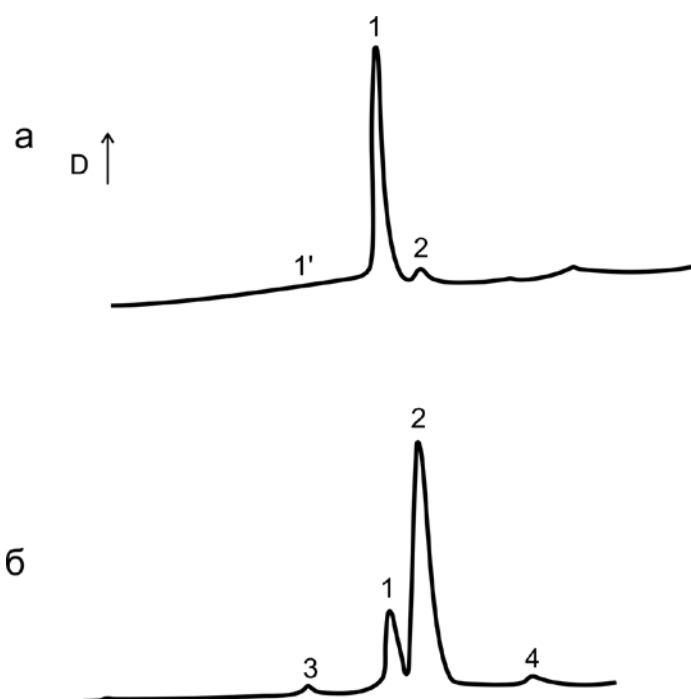


Рис. 2.1. Результаты электрофореза растворов ДНК Col E1; *a* – препарат № 1 (линейная ДНК); *б* – препарат № 2 (открытая кольцевая ДНК). Пики: 1 – линейная форма, 1' – фрагменты линейной формы, 2 – открытая кольцевая форма, 3 – сверхспиральная форма, 4 – димеры. *D* – оптическая плотность. Пики идентифицированы путем сравнения с образцами ДНК, охарактеризованными другими методами

Концентрацию ДНК изменяли диализом в буфере 2 мМ трис-НСl, рН 7,5, содержащем 1мМ ЭДТА и определенное количество NaCl. При достижении концентрации 500 $\mu\text{г}/\text{мл}$ растворы ДНК дополнительно очищали низкоскоростным центрифугированием в течение 20 мин.

2.1.2. Осмометр. Для измерения осмотического давления растворов ДНК нами был специально создан осмометр [188] с измерительной ячейкой малого объема и мембраной, высокопроницаемой для растворителя и малопроницаемой для ДНК (рис. 2.2). Такими мембранами оказались ядерные лавсановые фильтры, изготовленные в лаборатории Г. Н. Флерова в объединенном Институте ядерных исследований в г. Дубне (Россия) и любезно предоставленные нам Б. В. Мчедlishвили. Для измерения

осмотического давления ДНК Col E1 была выбрана мембрана с диаметром пор 830 \AA , пористостью 4,5 %, толщиной 8 мкм.

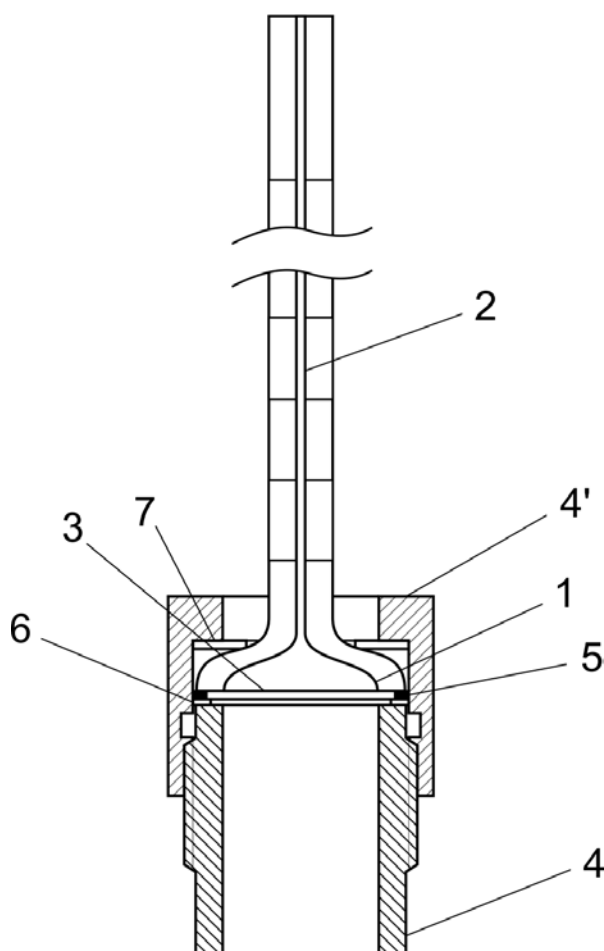


Рис. 2.2. Осмометр (пояснения в тексте)

Осмометр изготовлен из стекла и нержавеющей стали. Он состоит из стеклянной ячейки 1 (внутренний диаметр 1,9 см, объем 0,7 мл), верхняя часть которой выполнена в виде капилляра 2 (внутренний диаметр 0,1 см, длина 14 см). На капилляре вытравлены деления ценой 1 мм на протяжении 12 см (на рисунке обозначены только деления ценой в 1 см). Мембрана 3 прижимается к нижней части путем свинчивания частей металлического корпуса 4 и 4'. Полную герметизацию осмометра обеспечивает кольцевая прокладка круглого сечения из силиконовой резины 5. Для предотвращения деформации прокладки при уплотнении между ней и деталью 4 помещается тефлоновое

кольцо 6. Стеклянная ячейка отделена от верхней части металлического корпуса 4 резиновой прокладкой 7.

Перед сборкой прибора все его части тщательно промывали. Через капилляр многократно пропускали под давлением сначала раствор детергента, затем свежеприготовленный раствор хромовой смеси, после чего обычным способом ополаскивали водой. Собранный прибор помещали в стакан с растворителем, таким образом, чтобы предотвратить появление пузырьков воздуха под мембраной. Для этого осмометр переворачивали капилляром вниз, заливали его нижнюю часть доверху буфером, покрывали ее кружком фильтровальной бумаги и в таком виде опускали прибор уже в рабочем положении в стакан с растворителем, а затем удаляли кружок бумаги. Этот прием полностью обеспечивает отсутствие пузырьков воздуха под мембраной.

Заполнение ячейки производили с помощью пипетки с прикрепленной тонкой полиэтиленовой трубкой, доходящей через капилляр до дна ячейки. После заполнения ячейки осмометра раствором ДНК верхний конец капилляра закрепляли в штативе. стакан с растворителем устанавливали на подъемный столик, для того, чтобы иметь возможность плавно и быстро регулировать глубину погружения капилляра в растворитель. Растворитель содержал 2 мМ трис-НСl, рН 7,5, 1 мМ ЭДТА и NaCl в количестве, требуемом для достижения определенной ионной силы. Объем растворителя составлял 130 мл.

Используемые нами мембраны были малопроницаемые для ДНК. Потеря ДНК за несколько суток не превышала 1 %, т.е. эффекты диффузии ДНК на результаты измерений были незначительны. Причем, как показали эксперименты, через пленку диффундировали все компоненты ДНК, а не какие-нибудь ее низкомолекулярные фракции.

Измерения осмотического давления ДНК проводили в диапазоне концентраций ДНК 1-4 мг/мл, ионов Na^+ от 10^{-3} до 0,1 М в (0,8-1,6) мМ трис-НСl, рН 7,5 с препаратами ДНК № 1 (линейная форма) и № 2 (открытая кольцевая форма).

2.1.2. Вычисление величины эффективного диаметра ДНК. В теории реальные заряженные цепи представляют в виде последовательности свободно-сочлененных незаряженных цилиндров длины l (куновский статистический сегмент) и эффективного диаметра $d_{эфф}$ [189, 190]. Эффективный электростатический диаметр определяется как диаметр жесткого незаряженного цилиндра длины l , имеющего тот же исключенный объем v (объем, из которого данный сегмент вытесняет все другие), что и реальный сегмент молекулы при данных ионных условиях.

Эффективный диаметр может быть определен посредством измерения второго вириального коэффициента раствора сегментов ДНК. Этот коэффициент может быть найден с помощью измерения осмотического давления раствора ДНК при не слишком высоких ионных силах в разбавленных растворах или в таких условиях, когда полимерные клубки настолько переплетаются между собой, что сегменты становятся практически равномерно распределенными в растворе, независимо от их принадлежности, и концентрированный раствор молекул все еще остается разбавленным для сегментов. В этих условиях осмотическое давление определяется вторым вириальным коэффициентом, зависящим от их взаимодействия. Такая область, называемая областью среднего или самосогласованного поля, должна существовать согласно [191, 192] для полужестких молекул с достаточно большим отношением l/d . Чем больше величина $p = l/d$ (жесткость), тем шире область самосогласованного поля.

Определение $d_{эфф}$ посредством измерений в области разбавленных растворов затруднено из-за слабых эффектов и вследствие того, что в уравнение для осмотического давления входит специальный фактор, являющийся функцией исключенного объема. Точная форма этой функции, зависящая от молекулярной модели, используемой в той или иной теории, является проблематичной.

Поэтому целесообразнее выполнять измерения в области самосогласованного поля. Область самосогласованного поля (предельные

значения концентраций ДНК и температуры) для раствора полужестких незаряженных цепей, в зависимости от параметров цепи была определена в работе Бирштейн [192]. Используя полученные в работе Stigter [193] теоретические значения $d_{эфф}$ в качестве предварительных оценок, мы пришли к выводу, что раствор ДНК находится в области самосогласованного поля при концентрации ДНК 1-4 мг/мл и ионных силах выше $5 \cdot 10^{-3}$ М.

Теория Флори-Хаггинса для концентрированных растворов справедлива для области самосогласованного поля. Эта теория дает следующие выражения для осмотического давления P и второго вириального коэффициента раствора сегментов ДНК A_2 [194]:

$$P/RT = (1/M) c + A_2 c^2 \quad (2.1)$$

$$A_2 = N_A v / 2M_s^2, \quad (2.2)$$

где M – молекулярная масса ДНК,

c – концентрация ДНК,

N_A – число Авогадро,

M_s – молекулярная масса статистического сегмента Куна.

$1/M$ много меньше $A_2 c$. Используя выражение Онзагера для исключенного объема сегмента с диаметром $d_{эфф}$ и длиной l [195], получаем:

$$v = P/2l^2 d_{эфф} \quad (2.3)$$

$$d_{эфф} = 4 (M_s/l)^2 A_2 / P N_A \quad (2.4)$$

Отношение молекулярного веса сегмента ДНК (как и любого другого ее участка) к его длине известно с хорошей точностью и составляет $M_s/l = 195$ г/моль·Å. Поэтому измерение осмотического давления ДНК в области

самосогласованного поля позволяет определить по формулам (2.1) и (2.4) эффективный диаметр ДНК, причем в вычисления не входит длина куновского сегмента l .

В данной работе мы измерили осмотическое давление растворов ДНК Col E1 (линейной и кольцевой формы) в области самосогласованного поля и вычислили значения второго вириального коэффициента A_2 и эффективного диаметра ДНК $d_{эфф}$ в зависимости от ионной силы раствора [188, 196].

2.2. Электрофорез гомополирибонуклеотидов

2.2.1. Проведение электрофореза. Использовали препараты поли(А), поли(С), поли(У) и поли(Г) фирм «Serva» (ФРГ), «Reanal» (Венгрия) или «Calbiochem» (США). Фирму обозначали нижним индексом S, R или C соответственно, например, поли(А)_S. В целях предохранения растворов полинуклеотидов от действия нуклеаз их хранили замороженными при -20° С. Использовали дигидрохлорид этилового эфира лизина и гидрохлориды метиловых эфиров метионина и тирозина фирмы «Fluka» (Швейцария), гидрохлориды метиловых эфиров валина и метионина фирмы «Serva» (ФРГ), дигидрохлорид метилового эфира аргинина фирмы «Serva» (ФРГ), агарозу фирм «Sigma», (США), «Serva» (ФРГ) или «BioRad» (США), краситель толуидиновый голубой О (ТГ) фирмы «Serva» (ФРГ).

Для предотвращения образования «замороженной» формы растворы поли(А) готовили при рН ~ 7,5, затем доводили рН до требуемого значения кислым буфером (рН 4,5) и после 10-минутного прогрева при 75-80 °С на водяной бане раствор выдерживали 1 час при комнатной температуре. В отдельных экспериментах растворы поли(А), приготовленные при нейтральных рН, диализовали против кислого буфера по методике Janik et al. [28] или же на кислые гели наносили нейтральные растворы поли(А). В экспериментах, описанных в параграфе 4.4., раствор поли(Г) (2,7 мг/мл)

диализовали против буфера 10 мМ трис-ацетат, рН 7,0, содержащего 1 мМ ЭДТА.

Ультрацентрифугирование полинуклеотидов проводили на аналитической центрифуге «МОМ-3170b» (Венгрия) при 40000 об/мин в буфере 0,01 М трис-НСl, рН 7,7, содержащем 0,1 М NaCl и 2 мМ ЭДТА.

Электрофорез проводили в агарозных гелях при комнатной температуре в трубках диаметром 0,6 см в электрофоретической камере «ОЕ-110» (Венгрия) при напряжении ~ 60 В в течение 80-110 мин. В камеру можно вставить 14 трубок.

Буферные растворы для проведения электрофореза при слабощелочных рН (7,4-7,8) содержали трис-НСl, NaCl и динатриевую соль ЭДТА (от 0,3 до 2 мМ). До 40 мМ ионную силу повышали увеличением концентрации буфера трис-НСl, затем добавлением NaCl. Для проведения электрофореза при кислых рН буферные растворы содержали трис (от 3 до 25 мМ), уксусную кислоту, с помощью которой изменяли рН растворов, и 0,3-2 мМ динатриевой соли ЭДТА. Ионную силу повышали либо увеличением концентрации триса, либо добавлением ацетата натрия. Для электрофореза поли(С) наряду с трис-ацетатным буфером использовали также трис-фосфатный буфер, содержащий трис от 5 до 70 мМ, фосфорную кислоту, с помощью которой изменяли рН растворов, и 1 мМ динатриевой соли ЭДТА.

Эфиры аминокислот и ацетат магния вводили в буфер для электрофореза. В связи с тем, что нагревание растворителей, содержащих гидрохлориды эфиров аминокислот, может приводить к заметному уменьшению рН, производные аминокислот вводили в растворы агарозы сразу же после приготовления последних, интенсивно перемешивая их на магнитной мешалке с подогревом. Значение рН растворителей при этом практически не изменялось.

Электрофоретические эксперименты проводили, используя два варианта. В одном случае агарозу готовили в том же буферном растворе, в котором проводили электрофорез; в другом – в буфер для электрофореза помещали

трубки с агарозными гелями, приготовленными в разных растворителях. На трубку наносили по 7 мкг полинуклеотида в буфере 5 мМ трис-НСl, рН 7,5, содержащем 2 мМ ЭДТА, в присутствии 5 %-ного фикола-400 и 0,02 %-ного бромфенолового синего. Обычно электрофорез проводили в течение 80-120 мин. За время электрофореза изменение рН растворов в катодном и анодном отделениях не превышало 0,35.

Независимо от рН раствора, в котором проводили электрофорез, гели окрашивали на протяжении одного часа 0,01 %-ным раствором толуидинового голубого О (ТГ) при рН 7,4-7,8 [197] в 5 мМ трис-НСl – и отмывали этим же буфером. При этом зоны поли(А) окрашивались в сиреневый цвет, поли(У) и поли(С) – в розовый, а поли(G) – в фиолетовый цвет. Стойкое окрашивание имело место только для поли(А), окрашенные зоны остальных полинуклеотидов со временем обесцвечивались.

Окрашенные и отмытые от красителя гели сканировали на приборе «Хромоскан-200» (Англия). Сканирование каждого геля проводилось дважды.

Электронные спектры поглощения в диапазоне длин волн 400-750 нм в разных участках зоны окрашенного полинуклеотида снимали на спектрофотометре “Shimadzu” (Япония), при этом контролем служил спектр участка геля, не содержащего полинуклеотид.

Ионную силу вычисляли по формуле: $I = \sum c_i \cdot Z_i^2 / 2$, где c_i – молярная концентрация, а Z_i – заряд i -го иона, присутствующего в растворе.

2.2.2. Определение электрофоретической подвижности и коэффициента задержки. Как известно, электрофорезом называется перемещение под действием электрического поля заряженных частиц к катоду или аноду в зависимости от знака их суммарного заряда [197, 198]. Скорость движения частицы (см/сек) при напряженности электрического поля 1 В/см называется электрофоретической подвижностью u :

$$u = (D/t) / (E/l) \text{ см}^2 \cdot \text{с}^{-1} \cdot \text{В}^{-1}, \quad (2.5)$$

где D – расстояние, пройденное молекулой,
 t – время электрофореза,
 E – разность потенциалов на концах рабочего канала,
 l – длина канала.

При наложении электрического поля скорость движения частицы быстро увеличивается до тех пор, пока электрическую силу не уравновесит сила трения. Затем частица движется с постоянной скоростью в соответствии с уравнением:

$$u = Q/f, \quad (2.6)$$

где Q – суммарный заряд частицы,
 f – коэффициент трения.

Гели, используемые при электрофорезе, не только служат поддерживающей средой, но и играют роль молекулярного сита. Между логарифмом электрофоретической подвижности макромолекулы и концентрацией геля имеет место линейное соотношение:

$$\ln u = \ln u_0 - K_R T, \quad (2.7)$$

где u_0 – электрофоретическая подвижность в свободной среде,
 K_R – коэффициент задержки, пропорциональный эффективному размеру макромолекулы,
 T – концентрация геля.

K_R зависит от массы и формы мигрирующей частицы. Соотношение (2.7) выведено теоретически [199] и подтверждено экспериментально для разных типов гелей и молекул. В частности, оно выполняется для денатурированных

рибосомных РНК (мол. масса не больше ~ 2 Мда) в агарозных гелях в интервале T от 0,5 % до 2,5 % [200]; хотя в ряде случаев графики Фергюсона (зависимости $\ln u$ от T) для полимеров в агарозных гелях нелинейны из-за зависимости эффективных параметров агарозного волокна как от концентрации геля, так и от размеров мигрирующих частиц [201].

На графиках, относящихся к электрофоретическим экспериментам, каждая точка (кружок) соответствует, как правило, среднему результату двух измерений, то есть каждый отдельный эксперимент проводили в двух трубках с одинаковой концентрацией агарозы, растворенной в одинаковых условиях.

2.3. Потенциометрическое титрование гомополирибонуклеотидов и расчет кривых протонной буферной емкости

2.3.1. Потенциометрическое титрование. Поли(А)_R (40мг/мл) диализовали против буфера 4мМ трис-НСl, рН 7,8, содержащего 0,05 М NaCl и 0,2 мМ ЭДТА. Потенциометрическое титрование проводили на рН-метре «рН-121» при комнатной температуре. Титруемый раствор поли(А) (20 мл) содержал 2-4 мМ поли(А) (в молях нуклеинового фосфора), 1,5 мМ трис, 0,2 мМ ЭДТА и 0,05 М NaCl. Чтобы предотвратить образование «замороженной» формы раствор поли(А) и контрольный раствор, содержащий вместо поли(А) эквивалентный объем диализного буфера, после титрования до рН 5,2-5,5 прогревали 10 мин при 70-80 °С и охлаждали 1 час при комнатной температуре. При титровании использовали 0,1 М HCl или 0,1 М NaOH.

2.3.2. Метод протонной буферной емкости. Кривая потенциометрического титрования обычно служит для определения рК и количества протон-акцепторных групп исследуемого объекта, для чего ее разбивают на составляющие, соответствующие титрованию групп различного типа. Качество разложения кривой титрования по кривым титрования групп разного типа может быть значительно улучшено при использовании метода протонной буферной емкости [202]. Метод заключается в дифференцировании

кривых потенциометрического титрования с последующим их аддитивным разложением по базисным функциям (пикам), чьи рК соответствуют константам равновесия полимера между двумя состояниями, характеризуемыми определенной конформацией и степенью ионизации. Позиции пиков на шкале рН, их интенсивность и полуширина определяются количеством и качеством гидрофильных групп макромолекулы и набором ее конформационных состояний. Метод буферной емкости был успешно применен ранее для исследования структурных переходов в ДНК [203].

Базисная функция групп i -го типа имеет следующий вид [203, 204]:

$$\beta_i(\text{pH}) = \frac{c_i \gamma_i \ln 10}{10^{\gamma_i(\text{pH} - \text{pK}_i)} + 10^{\gamma_i(\text{pK}_i - \text{pH})} + 2}, \quad (2.8)$$

где c_i , - концентрация способных к ионизации атомных групп i -го типа, определяющих i -й пик на кривой буферной емкости,

pK_i - термодинамический параметр ионизации атомных групп i -го типа,

γ_i - параметр кооперативности, показывающий, во сколько раз сужается рН-диапазон присоединения или отщепления протонов по сравнению с некооперативным процессом для такого же количества протон-акцепторных групп.

Буферную емкость β определяли как $\Delta b / \Delta \text{pH}$ [204], где Δb – изменение концентрации титранта при добавлении его малой порции, ΔpH – малое изменение рН реакционной среды, вызванное Δb . При этом использовали интервалы: $\Delta \text{pH} = 0,01$ на участках с высокой буферной емкостью и $\Delta \text{pH} \leq 0,1$ на остальных участках. Значения параметров пиков рассчитывали методом подбора таким образом, чтобы расчетная суммарная кривая для всех пиков совпадала с экспериментальной кривой буферной емкости. В число базисных функций для подгонки входило количество всех узких пиков на кривой буферной емкости, соответствующих конформационным переходам, и

минимальное количество широких пиков. Количество последних предполагали, исходя из наличия возможных протонакцепторных центров в гомополимерах, способных протонироваться в исследуемом интервале pH. Расчеты и визуализацию данных осуществляли с помощью компьютерной программы Mathcad PLUS 5.0.

В данной работе мы применили метод протонной буферной емкости для изучения структурных переходов в поли(А) и поли(С). В случае поли(С) обрабатывали литературные кривые потенциометрического титрования, а в случае поли(А) – кривые, полученные как в наших экспериментах, так и литературные кривые.

2.4. Предсказание вторичной структуры РНК

2.4.1. Последовательности РНК и используемые программы.

Предсказывали вторичную структуру области про-мРНК SV40 L, охватывающей основной сайт полиаденилирования (2648-2731 нт), а также структуру области комплексного сайта полиаденилирования про-мРНК ВИЧ-1 (9493-9672 нт, в координатах генома НХВ2) и его фрагментов. Исследовали все геномы ВИЧ-1, содержащие последовательность поли(А)-сайта, которые были представлены в GenBank (США) на конец 2010 г. В том случае, если изоляты ВИЧ-1, выделенные из одного и того же пациента, имели геномы с идентичными последовательностями поли(А)-сайтов, для анализа брали только один геном.

Все исследуемые последовательности сворачивали с помощью программы «mfold» [205], а последовательности комплексных поли(А)-сайтов были дополнительно свернуты программой «UNAFold» (Unified Nucleic Acid Foldings) [206]. Программа «mfold» вычисляет структуру РНК по принципу минимизации свободной энергии, используя технику динамического программирования и экспериментально определенные термодинамические параметры и правила, учитывающие стабильность различных мотивов

вторичной структуры, таких как двойные спирали, апикальные и внутренние петли шпилек, бальджи и одноцепочечные участки между шпильками. Этот алгоритм не учитывает третичные взаимодействия и псевдоузлы. Программа «UNAFold», которая в настоящее время приходит на смену «mfold», является интегрированной программой, предсказывающей сворачивание нуклеиновых кислот, гибридизацию и профили плавления.

«Mfold» и «UNAFold» относятся к числу программ, которые определяют не только структуру РНК с минимальной свободной энергией, но также и субоптимальные структуры. *In vivo* субоптимальные структуры могут быть функциональными наряду с оптимальными, в частности, в случае их стабилизации специфически связывающими белками. Для контроля количества свернутых субоптимальных структур и инкремента их свободных энергий в программах «mfold» и «UNAFold» используются следующие параметры: размер окна (Window) и процент субоптимальности (P). Мы начали использовать в своей работе программу «mfold» с 2002 г., учитывая, что во всех известных нам из литературы примерах вторичные структуры коротких фрагментов РНК, предсказанные этой программой, совпадали со структурами, определенными экспериментально с помощью химического и ферментативного зондирования. К таким структурам, относятся, например: шпилька, содержащая «нижний» элемент поли(А)-сайта про-мРНК секреторного иммуноглобулина класса М мыши [207], ряд шпилек в нетранслируемой области геномной РНК ВИЧ-1 [208] и др.

Сворачивания фрагментов РНК в диссертационной работе проводились автоматически с использованием оригинальной программы, созданной к.б.н., с.н.с. Потягайло А. Л. Для предсказания вторичной структуры небольших фрагментов поли(А)-сайтов про-мРНК ВИЧ-1 использовали параметры: window = 0 и P = 10 %, а для сворачивания областей 9537-9672 нт, 9516-9644 нт и 9493-9644 нт использовали параметры: window = 0 и P = 5 %.

Поскольку обе программы, «mfold» и «UNAFold», могут не представлять некоторые структуры из набора структур с близкими свободными энергиями, в

некоторых случаях мы применяли специальные ограничения, чтобы найти структуры, содержащие U-богатый тракт USE и U/GU-богатый DSE в одноцепочечной форме. Мы налагали запрет одновременно на образование пар оснований трактом UUUUU (или соответствующим пентамером с мутациями) и трактом GUUGU (или UGUGU).

Предсказанные структуры поли(A)-сайтов про-мРНК ВИЧ-1 и его фрагментов хранятся в созданной нашей группой базе данных CESSHIV-1 (Control Elements Secondary Structures of HIV-1 genome) [209, 210], доступной на сайте <http://www.cesshiv1.org>.

2.4.2. Определение вторичной структуры фрагментов РНК.

Современные компьютерные программы для предсказания структуры РНК в основном определяют ее вторичную структуру. Что касается третичной структуры, то, насколько нам известно, в настоящее время существуют только программы, предсказывающие образование псевдоузлов, например [211]. Есть программы, с помощью которых можно найти последовательности с потенциалом для образования G-квадруплексов, например [212], но структура квадруплексов при этом не предсказывается, в частности из-за отсутствия строгих правил их образования (см. раздел 1). Очевидно, что наличие в РНК участков, образующих четырехцепочечные структуры, следует учитывать при предсказании вторичной структуры РНК разными программами.

При сворачивании фрагментов РНК очень большой длины (несколько сотен нуклеотидов), например геномных РНК вирусов, резко возрастает количество предсказанных альтернативных структур с близкой свободной энергией. Обычно в таких случаях используют дополнительную информацию для ограничения числа возможных структур. В частности, используют филогенетический подход, сравнивая эволюционно родственные последовательности для идентификации консервативных структурных мотивов, а также используют методы химического и ферментативного зондирования для выявления спаренных и неспаренных оснований.

Эффективным современным методом химического зондирования является метод SHAPE (selective 2'-hydroxyl acylation analysed by primer extension) [104]. Однако, как предполагают Pollom et al. [105], некоторые области природных РНК сворачиваются в уникальные структуры, а другие существуют в многочисленных конформациях. В результате, данные SHAPE-анализа являются усредненными данными для двух или более состояний. Авторы этой работы предсказали вторичную структуру генома вируса обезьян SIVmac239 программой *RNAstructure* с включением в алгоритм сворачивания данных SHAPE. Авторы определили, что в представленной ими структуре с минимальной свободной энергией доля нуклеотидов, чье участие в образовании пар оснований находится в согласии с SHAPE-реактивностями, составляет в среднем 74 %, а для некоторых областей, включающих известные консервативные элементы, этот параметр является более высоким.

В данной работе при определении вторичной структуры поли(А)-сайта про-мРНК ВИЧ-1 мы учли наличие в исследованных РНК участков, способных образовывать четырехцепочечные структуры, а также использовали филогенетический подход для подтверждения предсказанных структур.

В качестве последовательностей с потенциалом для образования внутримолекулярных G-квадруплексов РНК мы, исходя из литературных данных (см. параграф 1.2.5.), выбирали тракты, состоящие из четырех G-повторов длиной ≥ 2 нт, причем G-повторы длиной 2 нт должны были быть разделены линкерами из любых оснований общей длиной не более 9 нт, а повторы из трех и более G-остатков могли быть разделены линкерами длиной до 15 нт. Для облегчения поиска таких G-богатых трактов использовали простую программу, подчеркивающую в исследуемой последовательности РНК все олиго(G). Отбор последовательностей проводили вручную.

Таким образом, для изучения объемных эффектов в полинуклеотидах, обусловленных наличием в них многочисленных отрицательно заряженных фосфатных групп, применен метод осмометрии, что позволило впервые

экспериментально определить значение эффективного диаметра ДНК при промежуточных значениях ионной силы раствора. Для исследования структурных переходов в гомополирибонуклеотидах впервые использованы методы электрофореза и протонной буферной емкости, с помощью которых были выявлены такие характеристики этих макромолекул, которые трудно было бы получить с помощью других физико-химических методов исследования. Учет возможности образования трактами гомоолигонуклеотидов уникальных структур в РНК при ее сворачивании позволил, в частности, впервые создать структурную модель комплексного сайта полиаденилирования про-мРНК ВИЧ-1.

РАЗДЕЛ 3

ВЛИЯНИЕ ИОННОЙ СИЛЫ РАСТВОРА НА ЭФФЕКТИВНЫЕ РАЗМЕРЫ ПОЛИНУКЛЕОТИДОВ

Нуклеиновые кислоты, в том числе и гомополирибонуклеотиды, исследованию структурных характеристик которых посвящена диссертационная работа, являются сильно заряженными полиэлектролитами. Электростатические эффекты, зависящие от ионных условий в растворе, во многом определяют стабильность полинуклеотидов, их структуру, реактивность и способность связывать различные лиганды [213]. Поэтому для детального исследования свойств гомополимеров эксперименты в работе проводились в широком интервале ионных сил, от 3 до 70 мМ. В частности, некоторые формы удалось зарегистрировать только при низких I , при которых приобретают большое значение объемные (электростатические) эффекты. Для того чтобы иметь представление о том, в какой степени размеры гомополимеров зависят от ионной силы, что особенно важно при их изучении гидродинамическими методами, такими как электрофорез, мы исследовали объемные эффекты полинуклеотидов на примере ДНК методом осмометрии [188, 196].

ДНК была выбрана для этих целей, поскольку по своей архитектуре гомополимеры, как линейные структуры, ближе к ДНК, чем к длинным природным РНК. Для описания глобальных конформационных свойств ДНК в растворе используется модель полимерной цепи, состоящей из непроницаемых незаряженных сегментов с эффективным диаметром ($d_{эфф}$), который учитывает как геометрические размеры молекулы, так и электростатическое взаимодействие между сегментами (параграф 2.1). В то же время про-мРНК и мРНК представляют собой разветвленные структуры, состоящие из нерегулярных дуплексов, нерегулярных шпилек и линкеров, которые в целом не описываются подобной моделью.

Эффективный диаметр является одной из важнейших характеристик, используемых для описания конформационных свойств полиэлектролитов. Вместе с персистентной длиной он определяет эффективный исключенный объем полимерного сегмента. Теоретические значения эффективного диаметра ДНК в водном растворе, содержащем ионы Na^+ , были получены Stigter [193] в интервале ионных сил раствора от 0,005 до 1 М. Экспериментально величины $d_{\text{эфф}}$ были измерены Brian et al. [214] методом ультрацентрифугирования при высоких ионных силах (0,2-2,0 М) и $I = 0,005$ М. В области промежуточных ионных сил на период до 1985 г. экспериментальные данные отсутствовали.

3.1. Зависимость осмотического давления растворов ДНК от ионной силы раствора

Установление ионной силы во всем вязком объеме высококонцентрированных растворов ДНК происходит относительно медленно, как это, например, видно на рис. 3.1, где показан результат замены растворителя во внешнем объеме осмометра, ионная сила которого составляла 15,1 мМ, на растворитель с $I = 6,3$ мМ при концентрации ДНК 4,1 мг/мл. Пунктиром на рисунке отмечены начальный и конечный уровни. Начальное уменьшение высоты осмотического столба на рис. 3.1. связано с тем, что при замене буферного раствора (первые 20 минут без измерений на рисунке) уровень растворителя во внешнем объеме был выше уровня, который затем установили для измерений. Поэтому сначала наблюдали понижение мениска со временем, обусловленное достижением прежнего равновесия (пунктир слева), а затем по мере замены растворителя в ячейке осмометра – переход к новому равновесию. Конечный уровень в этом эксперименте определен через двое суток после изменения ионной силы растворителя. Изменение высоты поднятия раствора в капилляре осмометра за три часа составляет 80 % от изменения за двое суток. Через несколько часов после замены растворителя изменение уровня раствора ДНК в капилляре происходит довольно медленно.

Как показали опыты, проведенные при различных ионных силах растворителя, значения высот осмотического столба растворов ДНК, определенные через сутки и через двое суток после изменения ионной силы растворителя, различались лишь в пределах ошибки измерения, поэтому можно полагать, что через сутки измерялись равновесные значения осмотического давления.

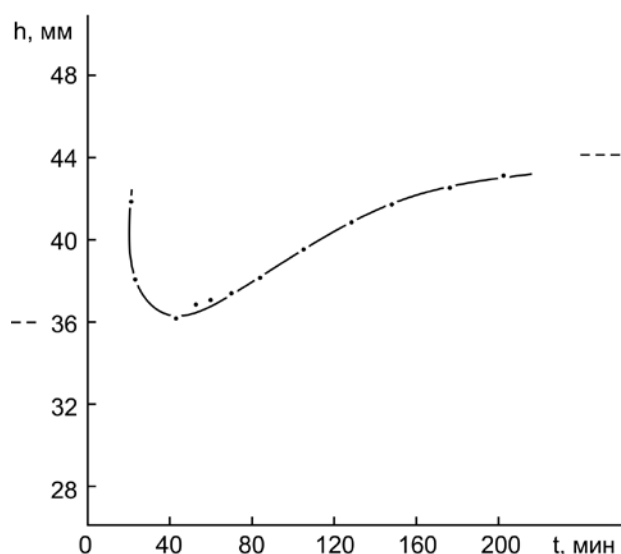


Рис. 3.1. Зависимость высоты поднятия h раствора ДНК в капилляре осмометра от времени при изменении ионной силы от 15,1 мМ до 6,3 мМ; $c = 4,1 \pm 0,4$ мг/мл. Пунктиром отмечены начальный и конечный равновесные уровни

Поскольку малейшее загрязнение капилляра осмометра могло привести к остановке движения столба жидкости в нем и, следовательно, к неправильному определению осмотического давления, для проверки воспроизводимости результатов осмотическое давление измеряли двумя способами: статическим (измерялась высота осмотического столба через сутки после изменения ионной силы растворителя) и динамическим. Динамические измерения проводили следующим образом: изменяя уровень растворителя во внешнем объеме осмометра (изменением высоты столика) смещали

равновесие. Например, сначала устанавливали уровень раствора в капилляре выше равновесного. Затем вернув уровень растворителя в первоначальное положение, наблюдали понижение мениска раствора со временем. Аналогичным образом отмечали повышение мениска во времени. Равновесный уровень определяли графически как общую асимптоту падающей и восходящей кривых, причем определение проводили при двух-трех уровнях растворителя во внешнем объеме, различающихся не менее чем на 1 см.

В качестве примера на рис. 3.2 изображены результаты нескольких измерений осмотического давления. Кривые 1 на рис. 3.2 (*а* и *б*) относятся к чистому растворителю, для которого равновесие устанавливается в первые же минуты. На рис. 3.2 *а* изображены результаты измерений для линейной ДНК, а на рис. 3.2 *б* – для кольцевой.

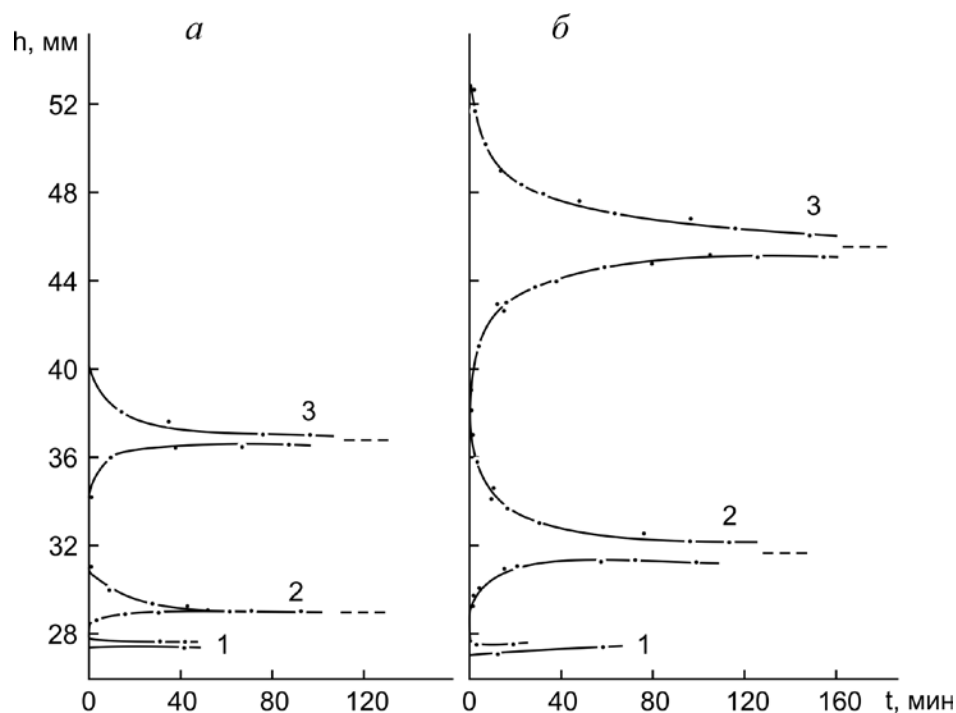


Рис. 3.2. Зависимость высоты поднятия h растворов ДНК и растворителя в капилляре осмометра от времени; *а* – препарат № 1, $c = 1,85 \pm 0,1$ мг/мл, $I = 106,6$ мМ, (кривая 2) и 2,4 мМ (кривая 3); *б* – препарат № 2, $c = 4,1 \pm 0,4$ мг/мл, $I = 102,0$ мМ, (кривая 2) и 5,7 мМ (кривая 3). Кривые 1 (*а* и *б*) получены для растворителя

Как видно на рис. 3.2, чем ниже ионная сила в системе, тем выше значение равновесного уровня и тем дольше устанавливается равновесие после его смещения. Зависимость высоты осмотического столба Δh от ионной силы раствора для препаратов 1 и 2 при нескольких концентрациях ДНК представлена на рис. 3.3 *а* (линейная ДНК) и 3.3 *б* (кольцевая ДНК). Для каждой формы ДНК эксперименты проведены с одним и тем же раствором ДНК, концентрацию которого уменьшали путем извлечения части ДНК из ячейки осмометра после очередной серии измерений.

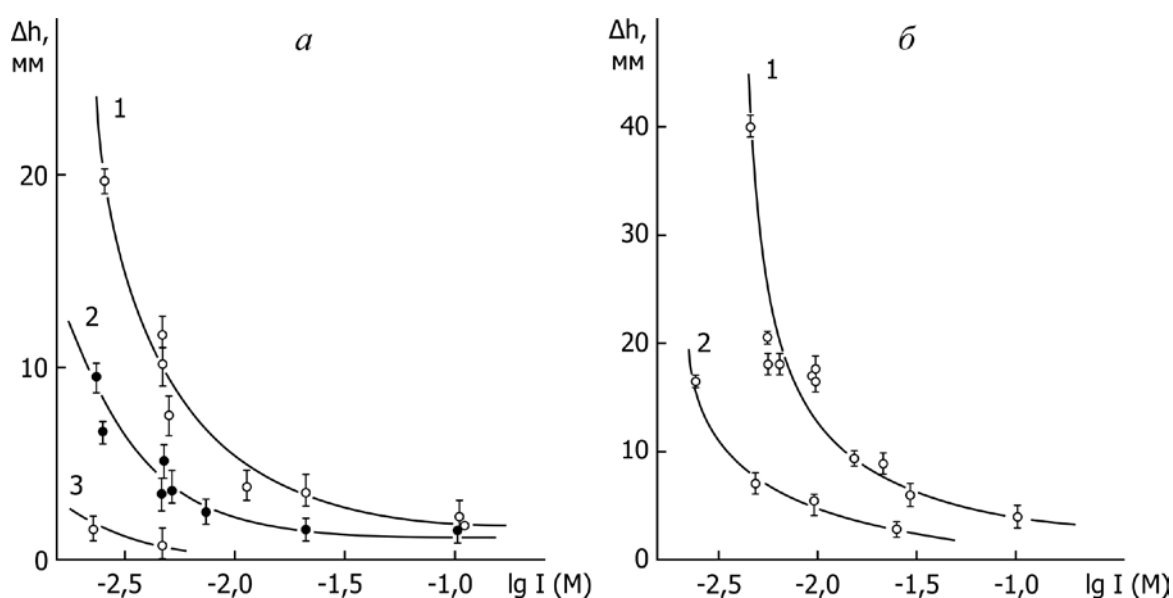
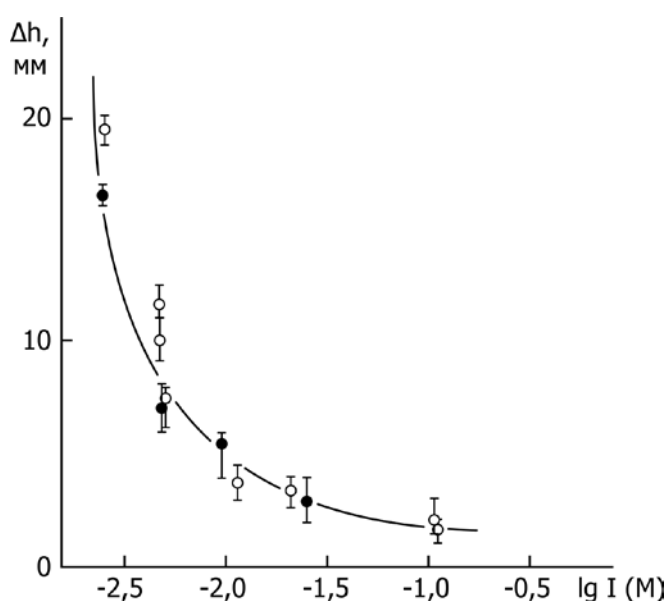


Рис. 3.3. Зависимость высоты осмотического столба Δh раствора линейной ДНК (препарат № 1) от I (*а*). 1 – $c = 2,75 \pm 0,1$ мг/мл, 2 – $c = 1,85 \pm 0,1$ мг/мл, 3 – $c = 1,00 \pm 0,05$ мг/мл. Зависимость высоты осмотического столба Δh раствора кольцевой ДНК (препарат № 2) от I (*б*). 1 – $c = 4,1 \pm 0,4$ мг/мл, 2 – $c = 2,7 \pm 0,1$ мг/мл

На рис. 3.3 видно, что при концентрации ДНК, превышающей 1 мг/мл, в области низких ионных сил значения осмотического давления растворов как линейной, так и кольцевой ДНК, сравнительно велики: они составляют несколько сантиметров водяного столба при погрешности измерений $\pm 0,5$ мм.

При уменьшении ионной силы наблюдается резкое увеличение осмотического давления растворов ДНК, обусловленное электростатическими взаимодействиями. Следует отметить, что значения высот осмотического столба, определенные при какой-либо ионной силе в системе, не зависели в пределах ошибки измерения от того, при какой ионной силе – более низкой, или более высокой – раствор ДНК выдерживался ранее, что свидетельствует об обратимом изменении измеряемого нами осмотического давления растворов ДНК при изменении ионной силы и, следовательно, о том, что нами измерялись действительно равновесные значения высоты осмотического столба.

В области средних и высоких ионных сил, при которых заряды на молекулах ДНК сильно экранированы, осмотическое давление растворов ДНК существенно ниже, но все же при концентрациях $c > 2$ мг/мл превышает ошибку измерения. Кривая 1 на рис. 3.3 а (результаты для линейной ДНК) и кривая 2 на рис. 3.3 б (результаты для кольцевой ДНК) получены для растворов ДНК примерно одинаковой концентрации. Результат сравнения этих кривых показан на рис. 3.4, где обе зависимости изображены вместе.



3.4. Сравнение зависимостей осмотического давления от ионной силы для растворов линейной (препарат № 1) (○) и кольцевой (препарат № 1) (●) ДНК по данным рис. 3.3

Как видно на рис. 3.4, в пределах ошибки измерения значения осмотического давления для кольцевой и линейной ДНК совпадают. По-видимому, чувствительность применяемого нами метода недостаточна для обнаружения разницы в поведении растворов ДНК, обусловленной существованием топологических ограничений.

3.2. Определение эффективного диаметра ДНК

На рис. 3.5 изображена концентрационная зависимость высоты осмотического столба Δh раствора ДНК в координатах $(\Delta h/c, c)$ при $I = 0,01$ и $0,1$ М. Зависимость оказывается линейной при ионных силах раствора выше $0,01$ М.

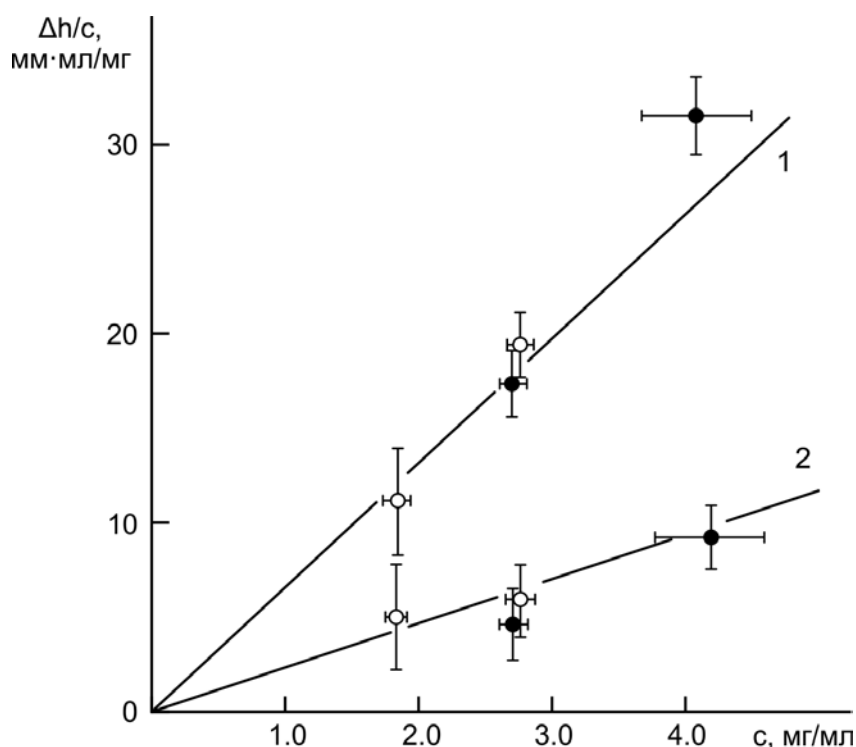


Рис. 3.5. Зависимость высоты осмотического столба Δh от концентрации ДНК c в координатах $(\Delta h/c, c)$ при $I = 0,01$ (1) и $0,1$ М (2). \circ — линейная ДНК (препарат № 1); \bullet — кольцевая ДНК (препарат № 2)

При $I = 0,01$ М наблюдается небольшое, но, видимо, достоверное отклонение от линейности при $c = 4$ мг/мл, а при $I = 0,005$ М зависимость становится нелинейной (не показано). При этой ионной силе, по-видимому, уже большой вклад в осмотическое давление дают тройные взаимодействия сегментов и взаимодействия высших порядков, тогда как при более высоких ионных силах оно определяется парными взаимодействиями. Наклон прямых на рис 3.5 пропорционален второму вириальному коэффициенту A_2 в вириальном разложении осмотического давления P от концентрации ДНК c . Из наклона этих прямых при различных ионных силах раствора были вычислены вторые вириальные коэффициенты сегментов ДНК A_2 , а по формуле (2.4) вычислены величины эффективного диаметра ДНК (см. параграф 2.1.). Значения A_2 и d_{eff} приведены в таблице 3.1.

Таблица 3.1

Зависимость второго вириального коэффициента A_2 и эффективного диаметра ДНК d_{eff} от ионной силы раствора I .

$I, \text{М}$	$A_2 \cdot 10^4, \text{мл} \cdot \text{моль} / \text{г}^2$	$d_{eff}, \text{Å}$
0,01	$26,8 \pm 1,9$	218 ± 16
0,02	$17,3 \pm 1,9$	141 ± 16
0,05	$12,1 \pm 1,8$	98 ± 15
0,10	$9,6 \pm 1,4$	78 ± 11
0,20	$8,0 \pm 1,4$	65 ± 11

Примечание. При $I = 0,2$ М величины A_2 и d_{eff} получены экстраполяцией зависимости Δh от I к этой величине.

После опубликования наших результатов по определению величины эффективного диаметра ДНК при разных ионных силах раствора эта величина была определена еще двумя другими экспериментальными методами: измерением вероятности заузливания ДНК при ее случайной циклизации [215, 216] и измерением светорассеяния фрагментов ДНК [217]. Значения

эффективного диаметра ДНК при различных концентрациях NaCl, полученные в разных работах, собраны вместе в работе [218] и показаны на рис. 3.6. Наши данные отмечены ромбами. Как видно на рисунке, между всеми величинами наблюдается удовлетворительное соответствие.

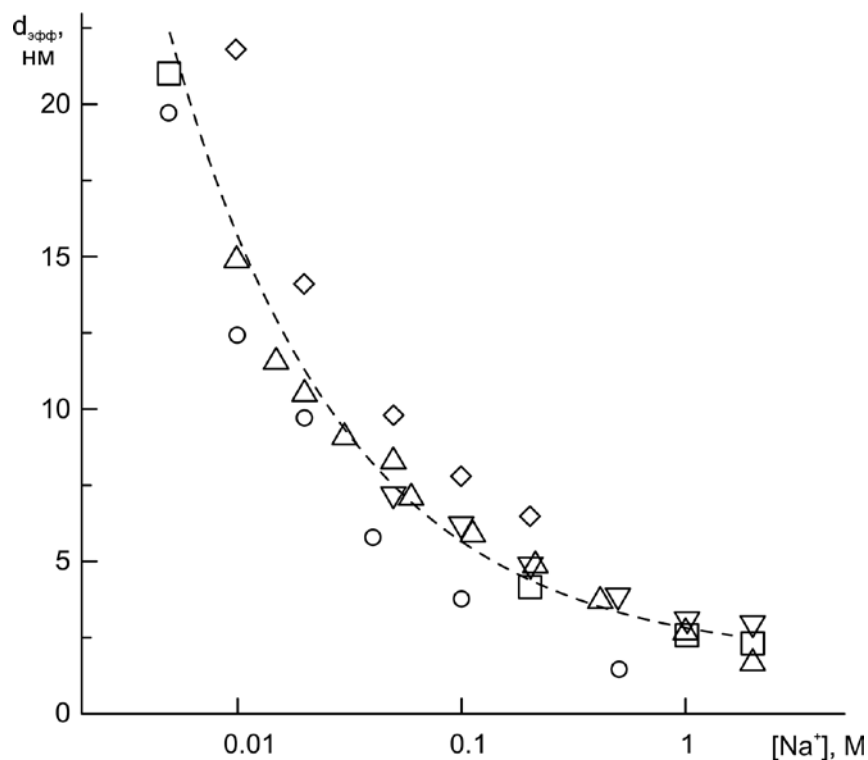


Рис. 3.6. Эффективный диаметр ДНК как функция концентрации NaCl в растворе: \square – экспериментальные данные, основанные на равновесном распределении фрагментов ДНК в ультрацентрифуге [214], \diamond – измерения осмотического давления (наши данные), \triangle , ∇ – измерения вероятности заузливания ДНК при ее случайной циклизации [215, 216], \circ – измерения светорассеяния фрагментов ДНК [217]. Пунктирная линия соответствует теоретическим расчетам, выполненным Stigter [193] в рамках теории полиэлектролитов. Адаптировано из [218]

Другой важнейшей характеристикой, определяющей эффективный исключенный объем полимерного сегмента, является персистентная длина, характеризующая жесткость полинуклеотидной цепи. Суммарные

литературные данные по зависимости персистентной длины от концентрации соли в растворе, представленные в работе [219], приведены на рис. 3. 7.

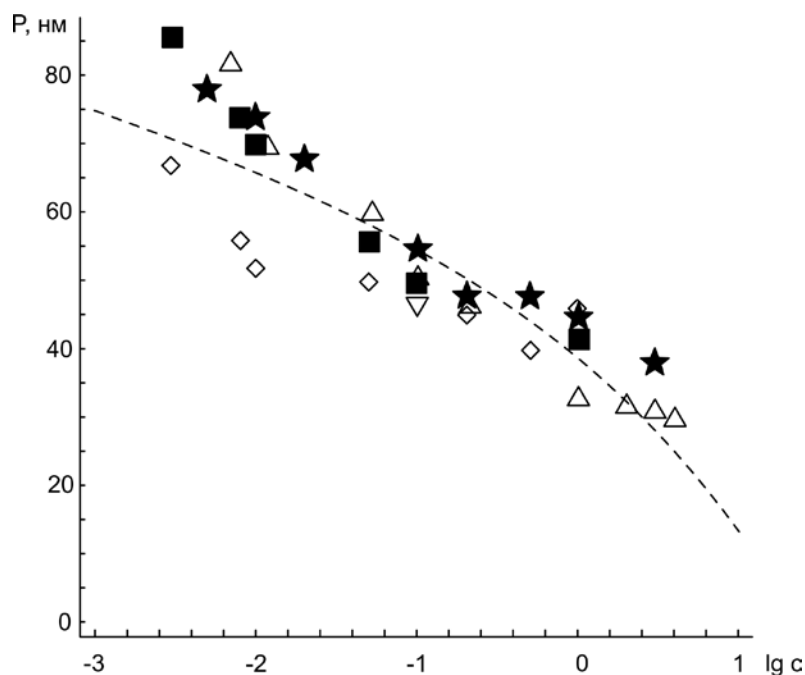


Рис. 3.7. Зависимость персистентной длины ДНК от концентрации соли. ★, ■, ◇ и △ – экспериментальные данные из работ [220-222] и [184] соответственно. Пунктирная кривая соответствует теоретическим расчетам [219]. Адаптировано из [219]

Из сравнения рисунков 3.6 и 3.7 видно, что в области низких ионных сил зависимость эффективного диаметра ДНК от ионной силы раствора значительно более резкая, чем аналогичная зависимость для персистентной длины. В интервале ионных сил 10-100 мМ, в котором в основном проведены наши эксперименты по изучению конформационных характеристик гомополирибонуклеотидов, персистентная длина ДНК изменяется примерно в 1,2 раза, а эффективный диаметр примерно в 2,8 раз. Таким образом, именно изменение величины эффективного термодинамического диаметра полинуклеотидов в этом интервале ионных сил, по-видимому, будет вносить основной вклад в изменение их размеров.

3.3. Выводы к разделу 3

Разработан простой прибор, обладающий достаточно малым временем установления равновесия и позволяющий непосредственно измерять осмотическое давление растворов ДНК достаточно высокого молекулярного веса. Исследованы зависимости осмотического давления ДНК от концентрации полинуклеотида и от ионной силы раствора. В пределах ошибки эксперимента значения осмотического давления для линейной и открытой кольцевой формы ДНК не различаются.

Результаты измерения использованы для вычисления второго вириального коэффициента A_2 раствора сегментов ДНК и определения эффективного диаметра двойной спирали ДНК при промежуточных ионных силах раствора. Величина эффективного диаметра ДНК увеличивается от ~ 78 Å до ~ 218 Å при уменьшении ионной силы от 0,1 до 0,01 М. Результаты исследования достаточно хорошо согласуются с теорией и имеющимися в литературе экспериментальными данными, полученными другими методами.

РАЗДЕЛ 4.

ИССЛЕДОВАНИЕ СТРУКТУРНЫХ ПЕРЕХОДОВ В ГОМОПОЛИМЕРАХ МЕТОДАМИ ЭЛЕКТРОФОРЕЗА И ПРОТОННОЙ БУФЕРНОЙ ЕМКОСТИ

Комплексное исследование синтетических гомополирибонуклеотидов различными физико-химическими методами, такими как потенциометрическое титрование, кристаллография, КД- и ДОВ-спектроскопия, ЯМР и другие методы, выявило, что все гомополимеры канонических оснований претерпевают структурные изменения при изменении рН или/и концентрации соли [3]. Однако, молекулярные механизмы структурных переходов гомополимеров полностью не выяснены и по ряду проблем не существует единого мнения. Поэтому важно использование новых методов и подходов для их исследования. В данной работе мы применили методы аналитического электрофореза в агарозных гелях [223-228] и метод протонной буферной емкости [229-232] для исследования структурных переходов в гомополирибонуклеотидах.

4.1. Зависимость электрофоретической подвижности гомополирибонуклеотидов от условий среды

Метод электрофореза, оказавшийся перспективным для фракционирования и характеристики белков и нуклеиновых кислот, и широко распространенный в настоящее время, не был использован в литературе для изучения структурных переходов в гомополирибонуклеотидах. Насколько нам известно, лишь две работы были посвящены электрофорезу синтетических полирибонуклеотидов. В работе [233] определена зависимость подвижности поли(А) в полиакриламидном геле от коэффициента седиментации при нейтральных значениях рН. Показано также, что полученная калибровочная

кривая может быть использована и для поли(U). В работе [234] полимеризация полиакриламидных гелей была проведена в присутствии электронейтральных аналогов полинуклеотидов: поли(1-винилурацила) (поли(vU)) или поли(9-виниладенина) (поли(vA)). В гелях с добавлением поли(vU) резко уменьшалась подвижность поли(A), но не поли(U) или поли(C), а в гелях с поли(vA) уменьшалась подвижность не только комплементарного ему поли(U), но и, в меньшей степени, поли(A). Авторы сделали вывод о том, что присутствие аналогов оснований в геле может быть использовано для электрофоретического разделения полинуклеотидов, основанного как на образовании пар оснований, так и на стэкинг-взаимодействиях.

Поскольку данные об электрофорезе синтетических гомополирибонуклеотидов в агарозных гелях в литературе отсутствовали, прежде всего, мы подобрали оптимальные условия электрофореза этих макромолекул и краситель для окрашивания гелей (см. параграф 2.2). Концентрация агарозы в интервале 0,8-2,4 % оказалась наиболее подходящей для проведения экспериментов.

На рис. 4.1 показаны результаты электрофореза поли(A)_C и поли(A)_S – средние коэффициенты седиментации ($5,0 \pm 0,1$) S и ($6,2 \pm 0,3$) S соответственно – при разных концентрациях геля в одном эксперименте. Ультрацентрифугирование показало гетерогенность препаратов полинуклеотидов, но мы не подвергали их дополнительной обработке с целью выделения гомогенных фракций. Из рисунка видно, что расстояние, пройденное препаратом в агарозном геле, зависит при прочих равных условиях от длины молекул полирибонуклеотида и от концентрации геля. Оно тем больше, чем короче молекулы полимера (зоны 3 и 4, рис. 4.1) и чем больше размер пор агарозы (зоны 1-3, рис. 4.1).

Относительную электрофоретическую подвижность препарата полинуклеотида R рассчитывали как отношение расстояния, пройденного максимумом зоны полимера, к расстоянию, пройденному максимумом зоны бромфенолового синего, используемого в качестве лидирующего красителя.

Величина R , как показали специально проведенные эксперименты, не зависела от времени электрофореза и от количества полимера нанесенного на гель.

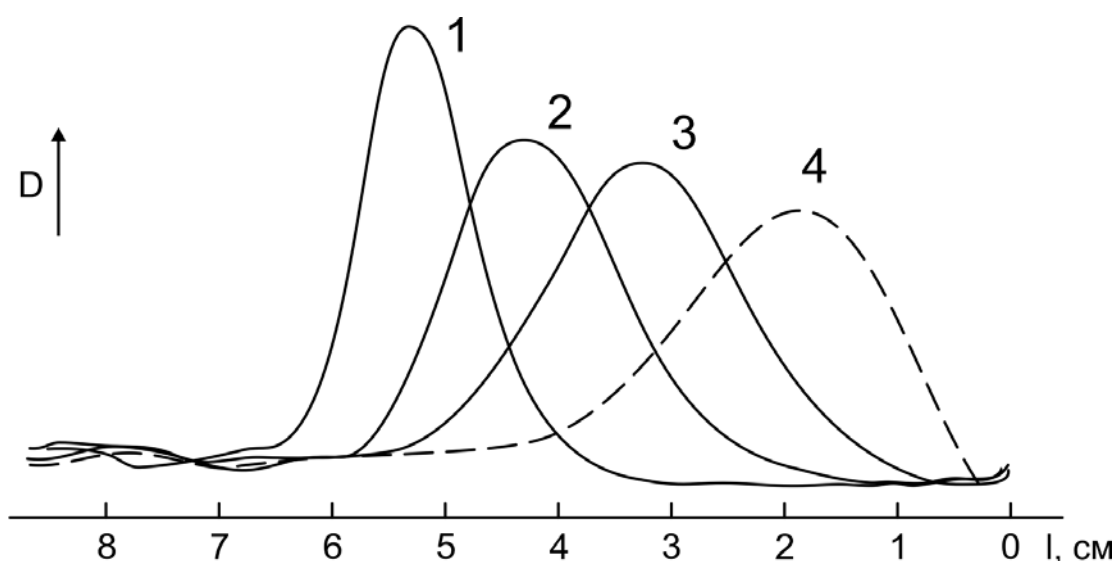


Рис. 4.1. Денситограммы поли(А)_с (зоны 1-3) и поли(А)_с (зона 4). Концентрация агарозы: 1 – 0,8 %; 2 – 1,4 %; 3 и 4 – 2 %; $I = 21$ мМ. Здесь и ниже D – оптическая плотность, l – расстояние от старта

Проверку проводили в интервале времен от 80 до 120 мин и в интервале концентраций полимера от 4 до 10 мкг при разных концентрациях агарозы. Значения R полирибонуклеотидов хорошо воспроизводятся как в одном эксперименте (в разных трубках) с точностью 0,01, так и при повторении экспериментов в одинаковых условиях с точностью $\leq 0,03$. Значения подвижностей u (см. параграф 2.2) воспроизводятся с меньшей точностью, чем значения относительных подвижностей.

Относительная электрофоретическая подвижность R , равна отношению электрофоретических подвижностей u полимера и лидирующего красителя. В нашем случае: $R = D_p/D_b = u_p/u_b$, где символы p и b относятся соответственно к полимеру и к бромфеноловому синему. Для R соотношение (2.7) имеет вид:

$$\ln R = \ln R_0 - K'_R T, \quad (4.1)$$

где R_0 – относительная электрофоретическая подвижность в свободной среде, K'_R – разность коэффициентов запаздывания для полинуклеотида и бромфенолового синего.

В качестве примера на рис. 4.2 приведена зависимость $\ln R$ от T (график Фергюсона) для поли(A)_C при $I = 16$ мМ. Как видно на рисунке, эта зависимость хорошо описывается прямой линией. Линейные соотношения (2.7) и (4.1) выполнялись в пределах ошибки эксперимента в интервале концентраций агарозы, используемых в работе (от 0,8 до 2,4 %), для всех исследованных полирибонуклеотидов и для лидирующего красителя бромфенолового синего при разных ионных силах раствора. В большинстве случаев зависимости $\ln R$ от T определяли при четырех (трех) значениях концентрации агарозы; в отдельных случаях, как на рис. 4.2 – при пяти и больше.

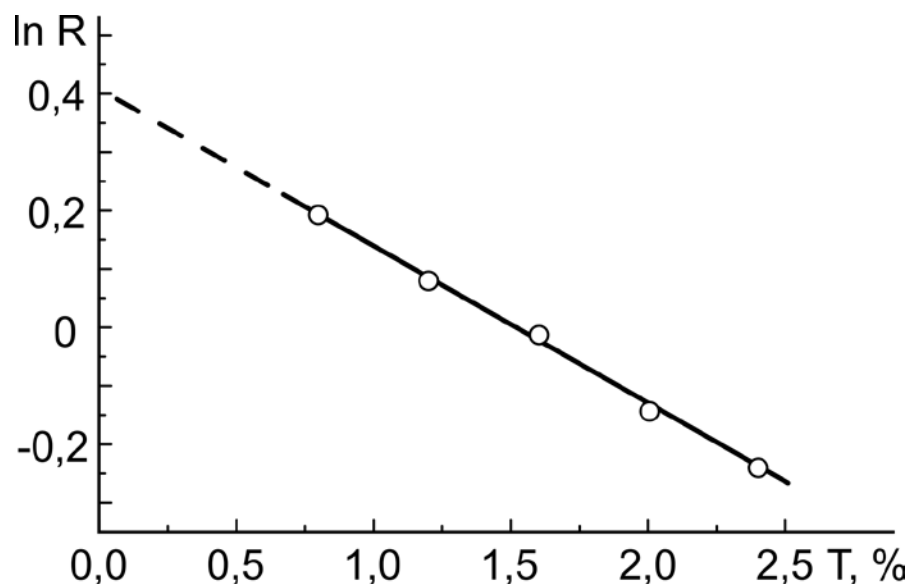


Рис. 4.2. График Фергюсона для поли(A)_C при $I = 16$ мМ, pH = 7,7

На рис. 4.3 (светлые кружки) показаны графики Фергюсона для разных гомополирибонуклеотидов – поли(G)_R (рис. 4.3 а), поли(A)_C (рис. 4.3 б), и поли(U)_S (рис. 4.3 в) при $I = 21$ мМ.

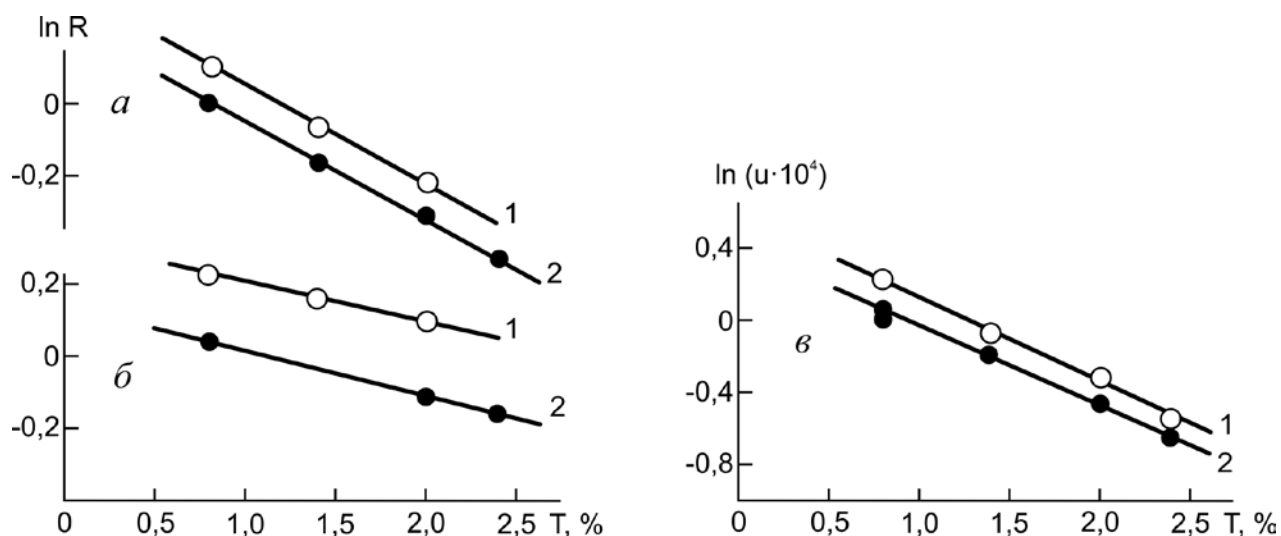


Рис. 4.3. Графики Фергюсона полирибонуклеотидов в отсутствие (1) и в присутствии (2) этилового эфира лизина. $I = 21$ мМ; *a* – поли(A)_C; *б* – поли(G)_R, *в* – поли(U)_S

Как видно на рис. 4.3, наклон графика K'_R для поли(A)_C значительно больше, чем для поли(G)_R. Поскольку коэффициент задержки, определяемый из наклона, пропорционален эффективному молекулярному размеру (параграф 2.2), то эффективный размер макромолекул в препарате поли(G)_R меньше, чем в препарате поли(A)_C.

На рис. 4.4 (светлые кружки) показаны зависимости электрофоретической подвижности поли(A)_C (*a*, *б*) и поли(G)_R (*в*) от ионной силы при разных концентрациях агарозы. На рис. 4.4. видно, что с уменьшением ионной силы подвижность гомополимера может, как уменьшаться, так и увеличиваться. Указанные зависимости соответствуют изменениям характеристик полинуклеотидов как сильно заряженных полиэлектролитов. Как указывалось в параграфе 2.2, электрофоретическая подвижность молекул прямо пропорциональна отношению их суммарного заряда к коэффициенту трения.

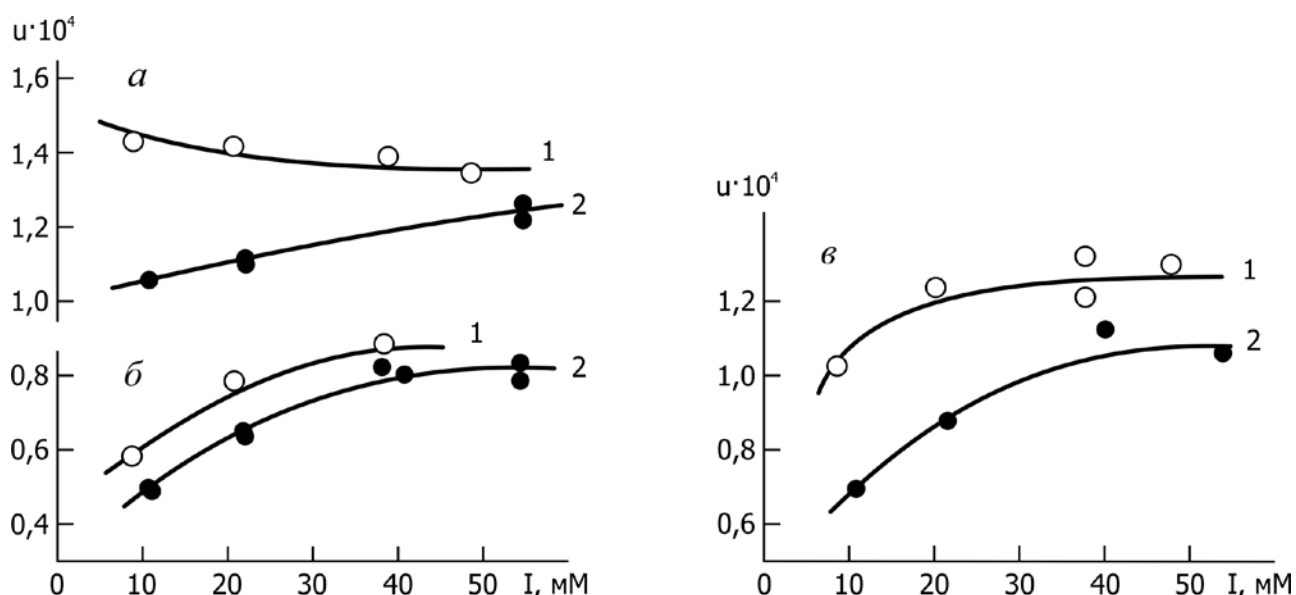


Рис. 4.4. Зависимости электрофоретической подвижности полирибонуклеотидов от ионной силы в отсутствие (1) и в присутствии (2) этилового эфира лизина; *a*, *б* – поли(А)_С, *в* – поли(Г)_Р; *a* – 0,8 %-ная агароза; *б*, *в* – 2,4 %-ная агароза

Уменьшение ионной силы приводит к увеличению заряда полимеров вследствие снижения уровня экранирования отрицательно заряженных фосфатных групп и к увеличению размеров макромолекул из-за возрастания электростатического отталкивания этих групп. Но при больших концентрациях агарозы эффект молекулярного сита геля проявляется в большей степени, чем при малых; увеличение коэффициента трения более существенно, чем изменение заряда, и подвижность уменьшается (кривые 1 на рис. 4.4 *б* и *в*). При концентрации агарозы 0,8 % (крупнопористый гель) для поли(А)_С более существенным оказывается изменение заряда, и подвижность возрастает (кривая 1 на рис. 4.4 *а*).

На рис. 4.5 показаны зависимости коэффициентов задержки K_R от ионной силы для поли(А)_С и поли(Г)_Р. Как видно на рисунке, K_R гомополирибонуклеотидов растут с уменьшением ионной силы и особенно резко при I ниже ~ 20 мМ.

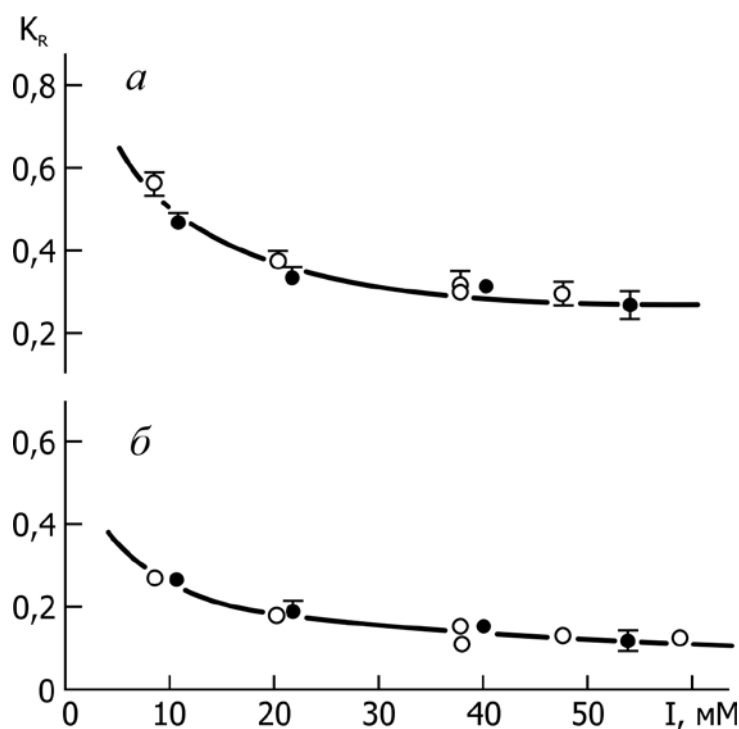


Рис. 4.5. Зависимость коэффициента задержки полирибонуклеотидов от ионной силы в отсутствие (○) и в присутствии (●) этилового эфира лизина; *а* – поли(А)_С, *б* – поли(Г)_Р

Представленные зависимости согласуются с зависимостями эффективного диаметра и персистентной длины ДНК от ионной силы раствора (рисунки 3.6 и 3.7). При этом все три полинуклеотида находятся в разных формах: ДНК в двухспиральной форме, поли(А) в виде одноцепочечных спиралей и неупорядоченных областей [34], а поли(Г), как будет показано ниже, в форме четырехцепочечной спирали. Зависимости R полинуклеотидов от I и K_R от I носят такой же характер, как и зависимости u от I и K_R от I , хотя коэффициент задержки для бромфенолового синего и его подвижность зависят от ионной силы. Эти кривые характеризуются меньшим разбросом точек.

Одной из задач диссертационной работы было исследование методом электрофореза взаимодействия производных аминокислот с гомополимерами в связи с проблемой белково-нуклеинового узнавания. Чувствительность данного метода оказалась достаточной высокой только для регистрации

взаимодействия гомополимеров с производными заряженных аминокислот – эфирами лизина и аргинина.

В разделе 2.2. указано, что увеличение ионной силы в наших экспериментах сначала производили увеличением концентрации буфера трис-НСl, а затем добавлением NaCl. Следует отметить, что соотношение катионов трис и натрия в буфере для электрофореза не влияло в пределах ошибки эксперимента на величину относительной электрофоретической подвижности полинуклеотидов. Например, если изменение ионной силы производили увеличением концентрации NaCl при постоянной концентрации трис (5 мМ), получали такую же зависимость от ионной силы. В отдельных экспериментах при $I = 20$ мМ и 55 мМ вместо трис-НСl использовали трис-СН₃СООН и соответственно NaСН₃СООН вместо NaCl, при этом значения R совпадали. На величину относительной электрофоретической подвижности полинуклеотидов не влияло также присутствие в растворах монокатионов производных аминокислот.

С другой стороны, когда в буфер для электрофореза вводили дигидрохлорид этилового эфира лизина наблюдалось заметное уменьшение электрофоретической подвижности гомополимеров (рис. 4.3. и 4.4, темные кружки), эффект был более значителен при низкой ионной силе и различен для разных полимеров. При этом значения коэффициентов задержки в пределах ошибки измерения не различались (рис. 4.5). Степень уменьшения подвижности при одинаковых условиях эксперимента убывает в следующем ряду: поли(G) > поли(A) ≥ поли(C) > поли(U). Эти результаты свидетельствуют в пользу того, что методом электрофореза регистрируется взаимодействие производного лизина с гомополирибонуклеотидами, и это взаимодействие наиболее сильное в случае поли(G).

Полученные результаты находятся в соответствии с литературными данными. В работе Брускова и Бушуева [235] на основании сравнения степени увеличения ширины линий в спектре ЯМР лизина при взаимодействии с полинуклеотидами получен следующий ряд устойчивости: поли(G) > поли (I)

> поли(С) ≥ поли(А) ≥ поли(У). Подобный ряд получен Porschke [236] для взаимодействия трипептида лизина с полирибонуклеотидами: поли(А) > поли(С) > поли(У).

При слабощелочных рН эфир лизина представляет собой смесь форм с одним и двумя зарядами (по ε- и α-аминогруппам). Для того чтобы изучить взаимодействие гомополирибонуклеотидов с отдельными формами производного лизина, мы провели эксперименты при рН 5,5. В этих условиях обе аминогруппы практически полностью ионизированы. На рис. 4.6 представлена зависимость относительной электрофоретической подвижности поли(Г) от ионной силы в отсутствие и присутствии эфиров аминокислот при рН 5,5.

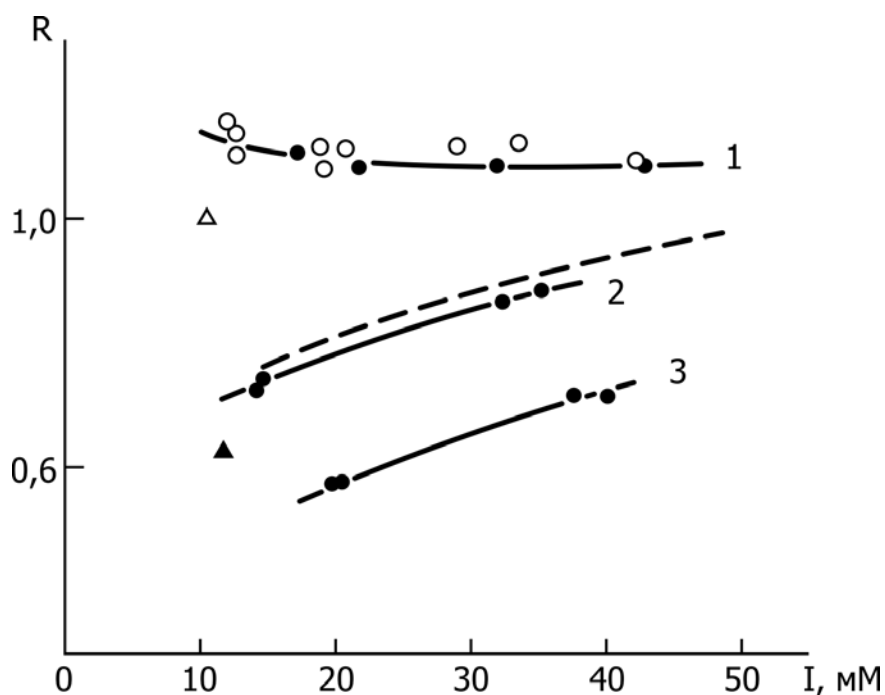


Рис. 4.6. Зависимость относительной электрофоретической подвижности поли(Г)_S от ионной силы в отсутствие (○) и присутствии эфиров аминокислот (●) при рН 5,5. Условия: 2,0 %-ная агароза; 1 – 5 мМ эфир метионина; 2 – 1 мМ эфир лизина; 3 – 4 мМ эфир лизина; пунктирная линия – 3,8 мМ эфир лизина, рН 7,6; Δ – без ацетата магния, ▲ – 0,7 мМ ацетат магния, рН 7,7

Как видно на рисунке 4.6, зависимость при рН 7,6 в присутствии ~ 4 мМ эфира лизина (пунктирная линия), близка к зависимости, полученной при рН 5,5 в присутствии всего лишь 1 мМ эфира лизина (кривая 2), а в присутствии 4 мМ лиганда (кривая 3) электрофоретическая подвижность поли(G) уменьшена при рН 5,5 в значительно большей степени, чем при рН 7,6.

На рис. 4.7 показаны зависимости электрофоретической подвижности поли(G) и поли(U) от рН в присутствии 4 мМ эфира лизина. Значения подвижностей представлены в виде отношения к подвижностям в отсутствие лиганда (u_L/u). Как видно на рис. 4.7, при увеличении рН, т.е. при уменьшении доли дикатиона эфира лизина в его смеси с монокатионом, отношения u_L/u для обоих полирибонуклеотидов приближаются к единице.

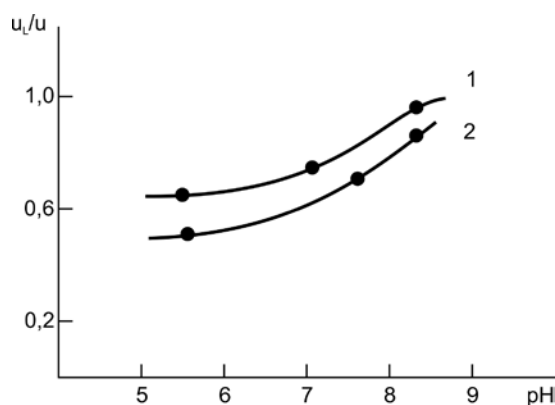


Рис. 4.7. Зависимости отношения подвижностей полинуклеотидов в присутствии 4 мМ эфира лизина (u_L) к подвижностям в отсутствие лиганда (u) от рН. 2 %-ная агароза; 1 – поли(U)_C; 2 – поли(G)_S; I = 20 мМ

Этот результат свидетельствует о том, что изменение подвижности полирибонуклеотидов в присутствии эфира лизина в условиях наших экспериментов обусловлено не обеими формами производного лизина, а практически лишь одной формой – дикатионом. В условиях наших экспериментов не вызывает изменения электрофоретической подвижности при постоянной ионной силе и монокатион производной другой аминокислоты – эфира метионина – (рис. 4.6, темные кружки на кривой 1).

Единичные эксперименты были проведены с ацетатом магния в концентрации 0,7 мМ при рН 7,7. Концентрация ЭДТА в растворах при этом составляла 0,1 мМ. Как видно на рис. 4.6, дикатион магния, так же как и дикатион эфира лизина, значительно уменьшает электрофоретическую

подвижность поли(G), а также поли(U) и поли(A) (не показано). Добавление в растворы ЭДТА до концентрации 2 мМ полностью снимает эффект. Подвижности полинуклеотидов в этом случае такие же, как и в отсутствие катионов магния.

На рис. 4.6 видно, что подвижности поли(G) в контрольных экспериментах (светлые кружки и треугольник) различаются. Подвижность поли(G) в отсутствие дикатиона эфира лизина больше, чем в отсутствие ацетата магния. Аналогичный результат получен и с другими полинуклеотидами. Это может быть связано с различной подвижностью полинуклеотидов в растворах разного состава. В первом случае концентрация динатриевой соли ЭДТА сравнима с концентрацией трис, а во втором – в десятки раз меньше. Но возможно и другое объяснение. Концентрация ЭДТА 0,1 мМ может быть недостаточна для полного осаждения примесей двухвалентных металлов в буферных растворах при проведении электрофореза, и они могут вызывать некоторое уменьшение подвижности полинуклеотидов.

Выше было указано, что при слабощелочных рН в присутствии эфира лизина изменялись только значения подвижностей разных полинуклеотидов, но значения K_R , в пределах ошибки измерения не менялись (рис. 4.5). На рис. 4.8 представлены зависимости коэффициентов задержки гомополирибонуклеотидов от ионной силы в отсутствие и присутствии эфира лизина при рН 5,5. Как видно на рис. 4.8, при рН 5,5 значения K_R , в пределах ошибки измерения тоже не меняются. Близость значений коэффициентов задержки полирибонуклеотидов в присутствии и отсутствие дикатионов производных аминокислот свидетельствует о том, что в условиях наших экспериментов эффективные размеры гомополимеров зависят только от ионной силы растворителя, но не от его состава, а подвижность зависит также и от величины заряда катионов, экранирующих отрицательно заряженные фосфатные группы.

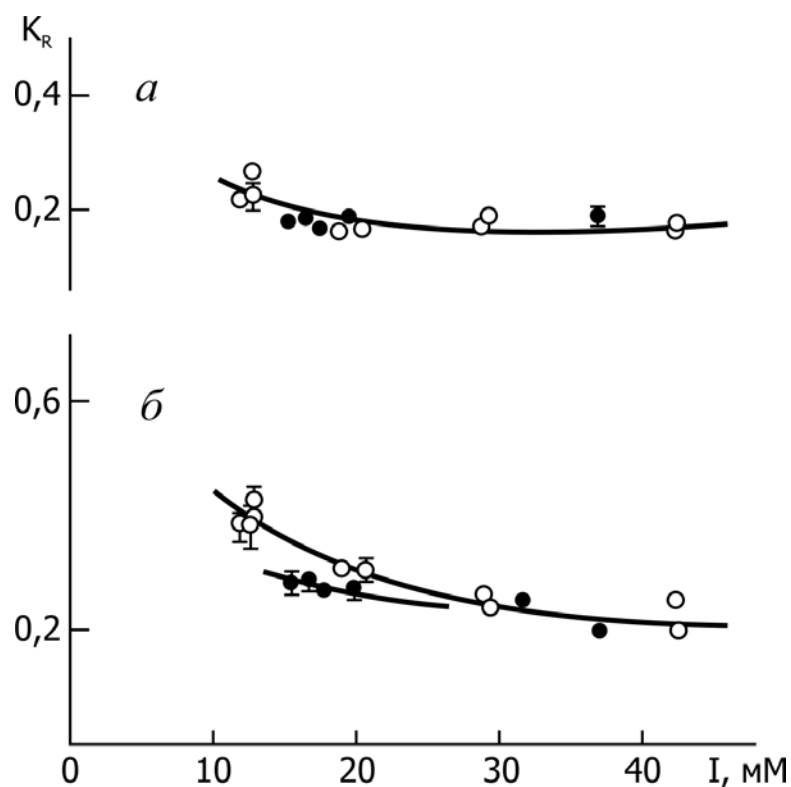


Рис. 4.8. Зависимости коэффициентов задержки гомополирибонуклеотидов от ионной силы в отсутствие (○) и присутствии 2-4 мМ эфира лизина (●) при pH 5,5; *a* – поли(G)_s, *б* – поли(U)_c

Исключение составляет поли(U) при больших концентрациях лизина и низкой ионной силе (рис. 4.8 *б*). Значение K_R в этих условиях ниже, чем в контрольных экспериментах, т. е. форма поли(U) в этом случае более компактна. Как уже упоминалось в параграфе 1.1.3., образованию единой большой шпильки поли(U) предшествует образование множества коротких шпилек (рис. 1.5). Вероятно, именно этот эффект и наблюдается при больших концентрациях дикатиона эфира лизина. В белке эфир лизина может соответствовать двум остаткам этой аминокислоты, расположенным рядом. Не исключено, что в клетке имеются поли(U)-связывающие белки с высоким содержанием лизина, способные индуцировать образование коротких шпилек в поли(U) при физиологических концентрациях соли.

4.2. Структурные переходы в поли(А)

4.2.1. Электрофоретическое исследование двухспиральных форм поли(А). На рис. 4.9 и 4.10 показаны денситограммы поли(А)_S и поли(А)_R, при разных значениях рН. Представленные результаты получены в одном электрофоретическом эксперименте, т.е. составы растворителей в трубках с агарозными гелями были различными (второй вариант проведения электрофореза). Значение рН изменяли, меняя концентрации уксусной кислоты и триса таким образом, чтобы значение ионной силы растворов оставалось постоянным. Как видно на рисунках 4.9 и 4.10, при переходе от нейтрального растворителя к кислым не только меняется форма зоны поли(А), но и происходит ее раздвоение, причем быстрее мигрирующая зона появляется при закислении постепенно, уже после того как произошло резкое изменение формы зоны в области рН 6,5 (рис. 4.9 д, 4.10 д).

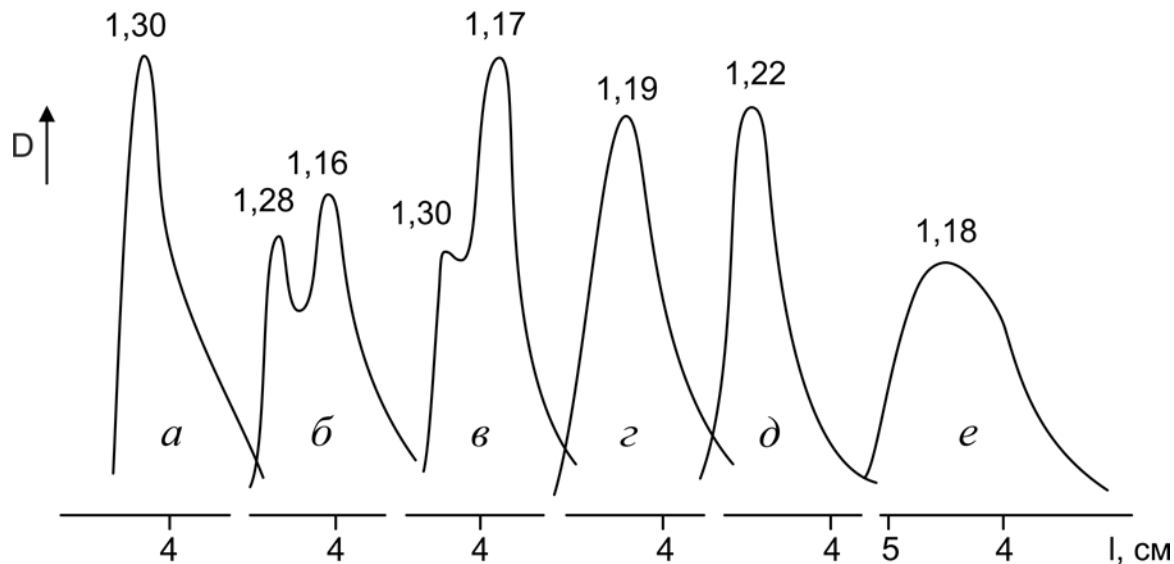


Рис. 4.9. Денситограммы поли(А)_S при разных рН; *a* – 4,9; *б* – 5,4; *в* – 5,8; *г* – 6,0; *д* – 6,5; *e* – 7,3; 1,0 %-ная агароза; $I = 14-15$ мМ; время электрофореза – 80 мин. Цифрами над пиками обозначены величины R

Сопоставление с литературными данными позволяет идентифицировать зоны поли(А) следующим образом. Широкие колоколообразные зоны при рН 7,3 (рисунки 4.9 *e*, 4.10. *e*) соответствуют одноцепочечному поли(А), имеющему клубкообразную форму. Большая ширина зон обусловлена гетерогенностью молекул, которые во время электрофореза делятся по размеру вследствие эффекта молекулярного сита. Такая же форма зон и при рН 6,8 (не показано).

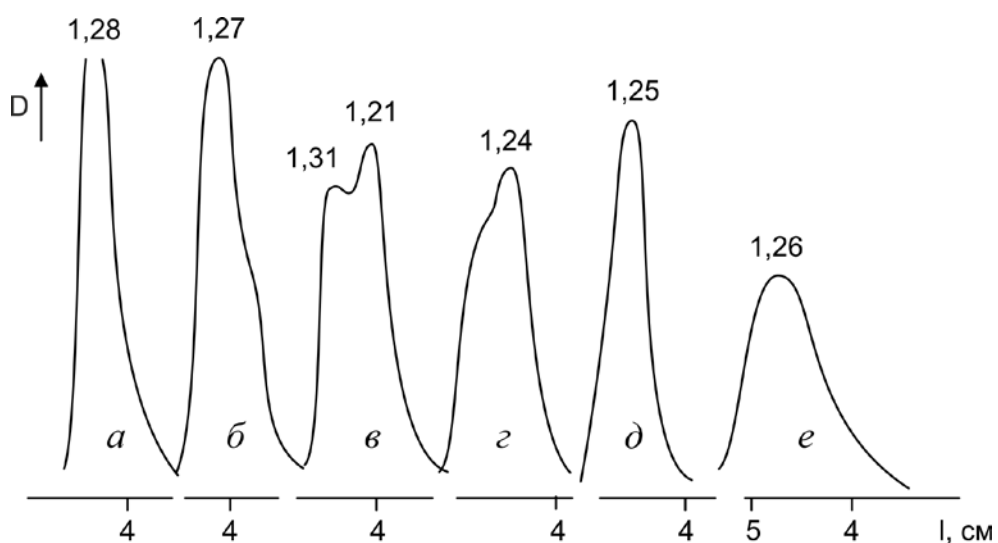


Рис. 4.10. Денситограммы поли(А)_R при разных рН. Условия эксперимента, как на рис. 4.9

Изменение формы зон при рН ~ 6,5 связано с переходом одноцепочечного поли(А) в двухспиральную форму. Молекулярная масса препаратов поли(А), с которыми проводились эксперименты, не превышала 1 МДа; двухспиральные полинуклеотиды такого размера представляют собой стержнеобразные структуры [237]. Очевидно, в данных условиях эксперимента двухцепочечные спирали поли(А) в форме «стержней» легче проходят через поры агарозного геля, чем клубкообразные молекулы. Они хуже разделяются по размеру, что приводит к сужению зон.

По литературным данным [26-28], первой при уменьшении pH образуется полупротонированная DS -форма поли(А), при дальнейшем подкислении она превращается в полностью протонированную DS_p -форму. В определенном интервале pH в зависимости от условий эксперимента обе формы сосуществуют. Поэтому мы считаем, что узкие (остроконечные) зоны при $pH \sim 6,5$ (рис. 4.9 δ , 4.10. δ) соответствуют DS -форме поли(А), двойные зоны – смеси DS - и DS_p -форм, причем DS_p -форме соответствует быстрее мигрирующая часть двойной зоны; зоны при $pH 4,9$ (рис. 4.9 a и 4.10 a) в основном соответствуют DS_p -форме поли(А).

Такая же картина изменения формы зон поли(А) при изменении pH наблюдается и при проведении электрофореза первым, стандартным, способом, т.е. когда гели агарозы готовили в тех же буферах, в которых проводили электрофорез. Изменения формы зон в обоих вариантах проведения электрофореза происходят примерно при одних и тех же значениях pH. По-видимому, во втором варианте проведения электрофореза введенные в гели во время их приготовления катионы водорода и трис и анионы ацетата мигрируют в соответствующие электродные буферы относительно медленно, так что большую часть пути полинуклеотид мигрирует к аноду в среде первоначального состава.

Перекрытие электрофоретических зон в области сосуществования двух форм, возможно, объясняется не только близостью величин электрофоретических характеристик DS - и DS_p -форм поли(А), но и тем, что эти формы находятся в динамическом равновесии [27, 29]. Следует отметить, что при кислых pH обе формы поли(А) регистрируются методом электрофореза только при относительно низких ионных силах (< 30 мМ), при более высоких I наблюдается одна зона полимера. Возможны два объяснения. Электрофоретические подвижности DS - и DS_p -форм поли(А) могут по-разному зависеть от I и при высокой ионной силе, в отличие от низкой, их значения могут сближаться. С другой стороны от ионной силы может зависеть время перехода одной формы в другую. При высокой ионной силе переходы

могут быть настолько быстрыми, что при электрофорезе наблюдается только одна зона с промежуточным значением подвижности. При низкой ионной силе, вследствие увеличения эффективных размеров макромолекул и их жесткости, переходы замедляются и регистрируются две формы.

На рисунках 4.9 и 4.10 представлены денситограммы поли(А) разных фирм. Согласно характеристикам, данным фирмами, поли(А)_S не содержит соли и белка, ее коэффициент седиментации превышает 10 S, коэффициент седиментации поли(А)_R – в пределах от 4,6 до 5,6 S. При хранении препаратов полинуклеотидов их размеры уменьшаются вследствие появления однонитевых разрывов. В данном исследовании длина молекул поли(А) не определялась какими-либо другими методами, но из электрофоретических данных следует, что средний эффективный размер молекул поли(А)_S больше, чем поли(А)_R, поскольку коэффициенты задержки (см. параграф 2.2) для нейтральной формы полимера составляли при $I = 13$ мМ соответственно $0,26 \pm 0,02$ и $0,20 \pm 0,02$.

Как видно на рисунках 4.9 и 4.10, при одинаковых значениях рН количественное соотношение *DS*- и *DS_p*-форм поли(А) в сдвоенных зонах различно для разных препаратов. Однако, возможно, этот эффект является кажущимся. Форма сдвоенной зоны зависит от степени перекрытия составляющих ее зон и от их формы. Но зоны *DS*- и *DS_p*-форм поли(А)_R несколько уже, чем поли(А)_S, поскольку молекулы в первом случае короче и хуже разделяются по размеру. В результате и формы сдвоенных зон в обоих случаях могут быть разными.

На рисунках 4.9 и 4.10 над зонами указаны значения относительных электрофоретических подвижностей. Разумеется, в случае перекрывающихся зон трудно судить об электрофоретических характеристиках полинуклеотидов, но некоторые оценки могут быть проведены. Модельные построения показывают, что при перекрытии зон положения пиков на результирующей кривой сдвинуты по сравнению с их истинными положениями; пики сближены, а их высота завышена, но величина этих эффектов зависит от

степени перекрытия. Приведенные на рисунках 4.9 и 4.10 значения R следует рассматривать с учетом того, что для быстрее мигрирующих зон указанные величины могут быть несколько занижены, а для медленнее мигрирующих – завышены. Как видно на рисунках 4.9 и 4.10, значения R быстрее мигрирующих зон при разных рН примерно совпадают; то же относится и к медленнее мигрирующим зонам. В области рН $\sim 6,5$ уже возможны искажения значения R вследствие перекрытия зон DS - и нейтральной форм, поскольку ширина перехода поли(А) из одноцепочечной спирали в DS -форму (как мы определили методом УФ-спектроскопии), составляет в условиях наших экспериментов $\sim 0,3$ ед. рН. Причем в этом случае перекрываются зоны разной формы (узкая и широкая).

Следует отметить, что в интервале рН 5,0-8,5 возможное изменение свойств агарозного геля мало влияет на электрофоретические характеристики гомополирибонуклеотидов. При электрофорезе поли(Г) и поли(У), которые не претерпевают структурных изменений при уменьшении рН, ни форма зон, ни значения R не изменялись. Для примера на рис. 4.11 показаны денситограммы гомополинуклеотидов при рН 8,4 и 5,5.

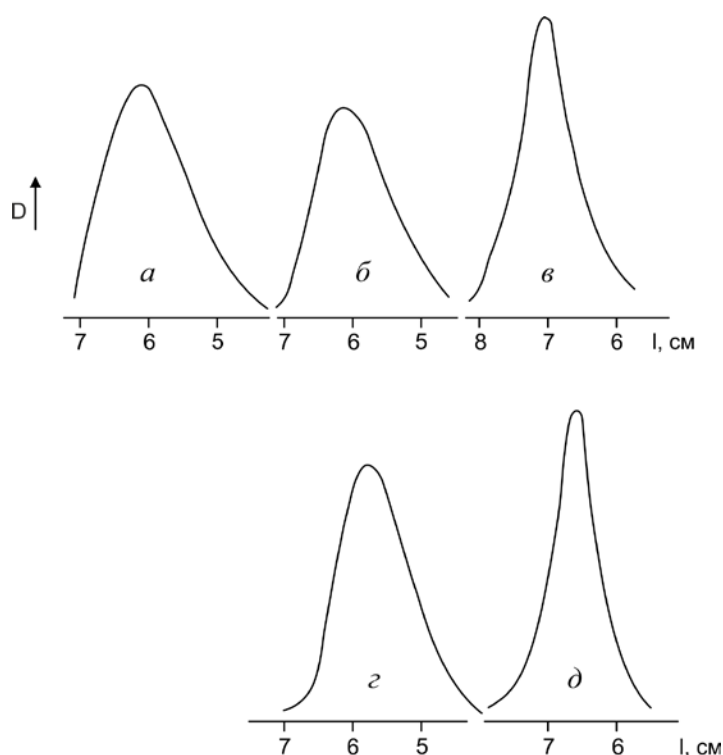


Рис. 4.11.
Денситограммы поли(А)_с, поли(У)_с и поли(Г)_с; а – поли(А)_с; б, г – поли(У)_с; в, д – поли(Г)_с. 0,8 %-ная агароза; а-в – рН 8,4; I = 12,4 мМ; г, д – рН 5,5; I = 15,1 мМ

Как видно на рис. 4.11, в обоих случаях поли(U) и поли(G) мигрируют одной зоной. Зона поли(U) имеет колоколообразную форму, подобную зоне поли(A) при слабощелочных рН, а макромолекулы поли(G) в исследуемых нами препаратах хуже разделяются по размеру в агарозных гелях и мигрируют узкой зоной, как при слабощелочных, так и при кислых рН.

На рис. 4.12 показана зависимость относительной электрофоретической подвижности пиков поли(A) в сдвоенных зонах от ионной силы. Электрофорез проводили стандартным способом. Используются результаты экспериментов при рН в области 5,0-5,8. Темные кружки относятся к быстрее мигрирующим зонам. Как видно на рисунке, во всей использованной области рН значения R , относящиеся к этим зонам, ложатся на одну кривую, так же как и значения, относящиеся к медленнее мигрирующим зонам.

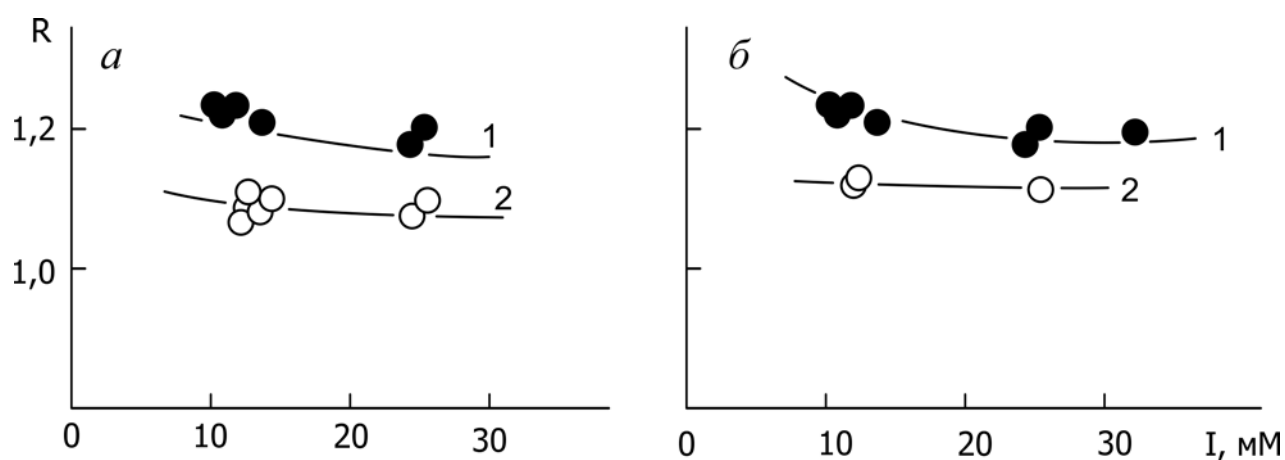


Рис. 4.12. Зависимости относительной электрофоретической подвижности DS_P - (1) и DS -форм (2) поли(A) от ионной силы; а – поли(A)_S; б – поли(A)_R; 1,4 %-ная агароза. Представлены данные экспериментов при рН 5,0-5,8

Таким образом, значения R для быстрее мигрирующих зон в сдвоенных зонах совпадают со значениями R зон, наблюдаемых при самых низких рН, которые соответствуют DS_P -форме поли(A). На рис. 4.12 видно, что значения

электрофоретических подвижностей для DS - и DS_P -форм поли(A)_R (рис. 4.12 б) выше, чем для поли(A)_S (рис. 4.12 а). Это объясняется тем, что, как уже отмечалось выше, молекулы поли(A)_R короче. На рисунках 4.9 и 4.10 различие менее заметно, поскольку приведены результаты при меньшей концентрации агарозы, при которой молекулы хуже разделяются по размеру.

Из представленных данных следует, что DS_P -форма поли(A) имеет большую электрофоретическую подвижность, чем DS . DS_P -форма поли(A) может мигрировать быстрее из-за того, что в соответствии с предположением Finch & Klug [26], молекулы в этой форме могут быть «тоньше», чем в DS -форме, так как определенный из рентгеноструктурных данных размер элементарной ячейки в первом случае меньше.

Как следует из предыдущего параграфа, заметное изменение электрофоретических характеристик полинуклеотидов в присутствии эфиров аминокислот наблюдалось только в случае производных аминокислот с двумя зарядами. При этом происходит уменьшение электрофоретической подвижности. В этом параграфе мы попытались выяснить, оказывают ли производные аминокислот влияние на структурные переходы в поли(A) в области кислых значений pH? В частности, если ароматические аминокислоты могут интеркалировать в двухспиральные формы поли(A), будет ли интеркаляция приводить к изменению электрофоретических характеристик полинуклеотида? С этой целью были проведены эксперименты с эфиром тирозина.

Эфир тирозина смешивали с поли(A) при pH 5,5 в молярном отношении 40:1, выдерживали 2 часа при комнатной температуре и наносили на гель, содержащий 5 мМ эфира тирозина. Значительных изменений электрофоретического поведения поли(A) в таких экспериментах не наблюдалось, хотя отмечалось небольшое, практически в пределах ошибки эксперимента, увеличение количества DS -формы. Не наблюдалось изменений в экспериментах, в которых поли(A), не смешивая предварительно с лигандом, наносили на гель, содержащий эфир тирозина. При этом поли(A) сначала

вводили в течение короткого времени в гель, затем электрофорез останавливали для образования комплекса поли(А) с производным аминокислоты непосредственно в геле и через 2 ч возобновляли электрофорез.

В описанных выше экспериментах с эфиром тирозина производное аминокислоты вводили только в гели. За время электрофореза в наших условиях эфир тирозина, мигрирующий навстречу поли(А), проходит расстояние, равное примерно половине длины геля, и в результате полинуклеотид, проходящий примерно такое же расстояние, 90 % пути мигрирует в среде, содержащей лиганд. Расстояние, пройденное эфиром тирозина, рассчитывали исходя из его концентрации в верхнем электродном буфере после окончания электрофореза, определенной методом УФ-спектроскопии.

В электродные буферные растворы в этих экспериментах вводили HCl в концентрации, равной концентрации гидрохлорида эфира тирозина. Дело в том, что введение HCl (в качестве контрольного вещества по отношению к гидрохлоридам эфиров аминокислот) только в гели, но не в электродные буферы приводит при электрофорезе поли(А) к изменению соотношения количеств DS - и DS_p -форм: увеличивается количество DS -формы, как видно на рис. 4.13 *a-v*. По-видимому, это связано с тем, что анион хлора мигрирует быстрее, чем поли(А), увеличивая pH в области геля, которую он покидает и в которую вступает полинуклеотид. Когда только в гели вводили гидрохлориды эфиров валина и метионина, они оказывали на поли(А) такое же действие, как и HCl в равной концентрации, как видно на рис. 4.13 *d, e*. Однако в случае гидрохлорида эфира тирозина (рис. 4.13 *z*) количество DS_p -формы уменьшается в несколько большей степени, чем в случае других аминокислот или HCl. Возможно, эфир тирозина способен в какой-то степени интеркалировать в двойную спираль поли(А), причем более эффективно в DS -форму, чем в DS_p . В эксперименте, результаты которого показаны на рис. 4.13, поли(А) перед нанесением на гели смешивали с производными аминокислот в молярном отношении 40:1 и выдерживали 2 ч при комнатной температуре.

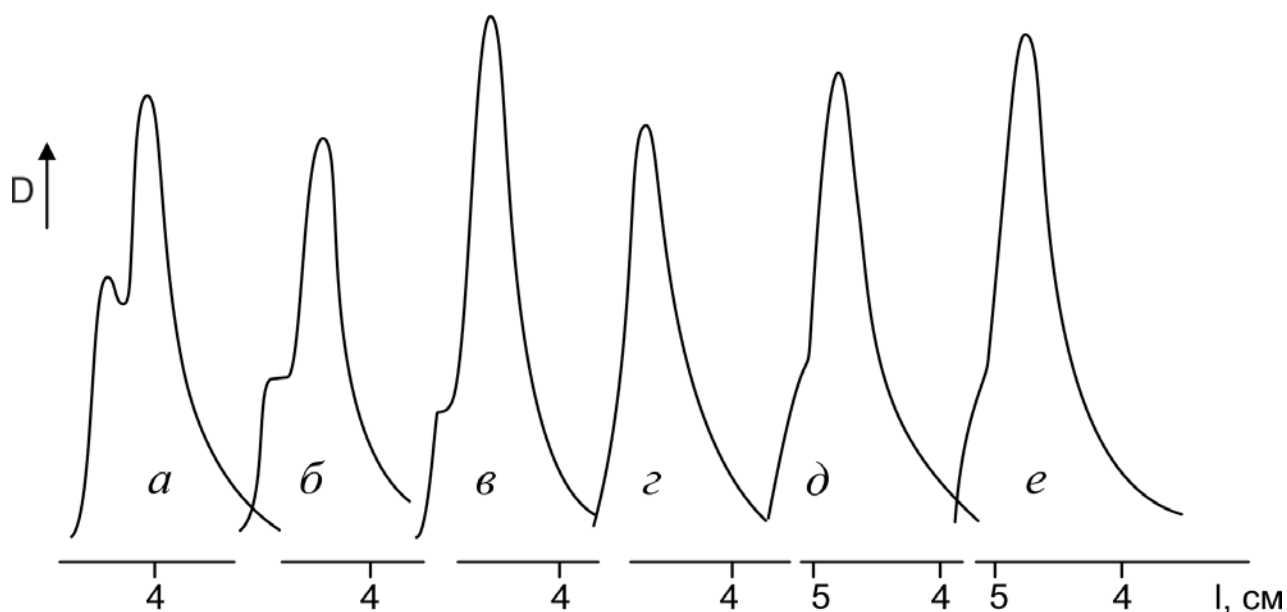


Рис. 4.13. Результаты электрофореза поли(А)_s в присутствии HCl и эфиров аминокислот; *a* – в отсутствии HCl и эфиров аминокислот; *б* – 4 мМ HCl, *в* – 5 мМ HCl; *г*, *д* и *е* – 5 мМ эфира тирозина, метионина и валина соответственно. 1 %-ная агароза; $I = 16$ мМ; pH 5,5; время электрофореза – 80 мин

На рис. 4.14 показаны результаты электрофореза поли(А) в присутствии дигидрохлорида метилового эфира аргинина. Поскольку, как уже отмечалось выше, в присутствии дикатионов электрофоретическая подвижность полинуклеотидов уменьшается, для сравнения результатов при одинаковом расстоянии миграции, т.е. при одинаковой степени разрешения зон, электрофорез проводили в течение разного времени. Например, расстояние, пройденное поли(А) в присутствии эфира аргинина за 80 мин (рис. 4.14 *г*), близко к пройденному полинуклеотидом в отсутствии производного аминокислоты за 60 мин (рис. 4.14 *д*).

Сравнение данных рис. 4.14 *б-д* показывает, что в присутствии эфира аргинина уменьшается количество *DS*-формы поли(А) и увеличивается количество *DS_p*-формы. В этом эксперименте эфир аргинина в концентрации

0,8 мМ вводили только в гели, в электродные буферы вводили соответственно 1,6 мМ HCl. Аналогичный результат наблюдали и в эксперименте, в котором эфир аргинина вводили и в гели, и в электродные буферы.

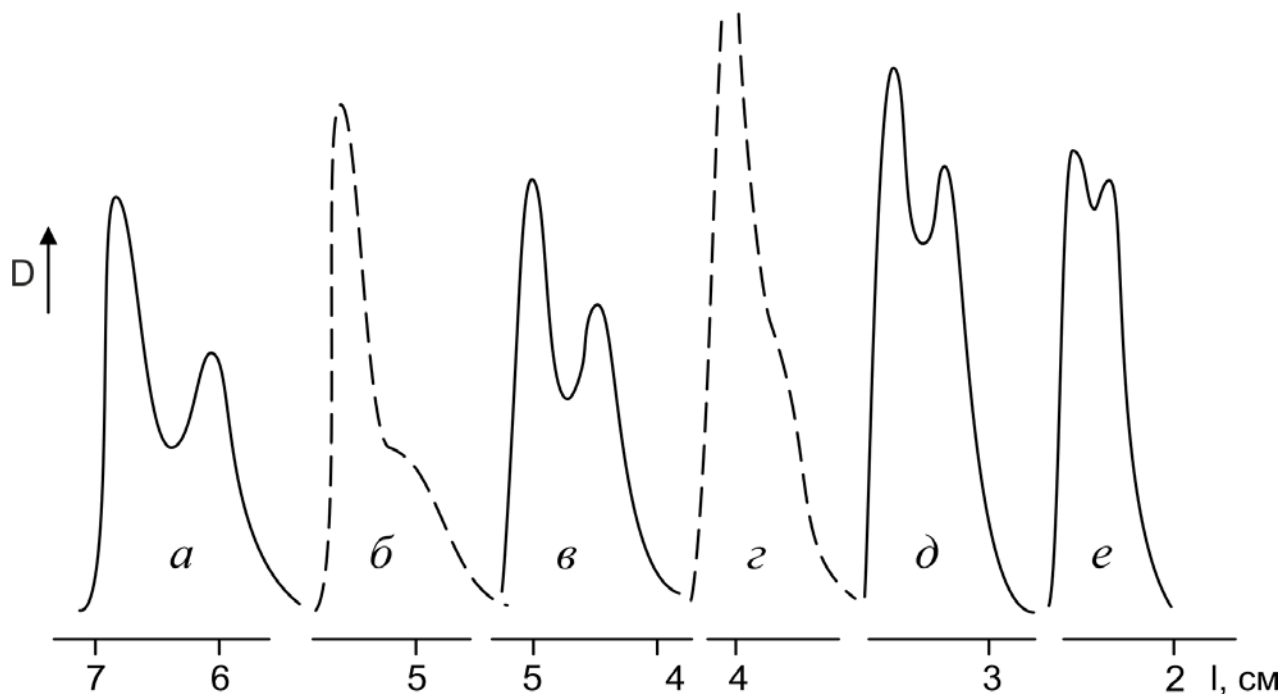


Рис. 4.14. Денситограммы поли(A)₅ при разной длительности электрофореза; *a, б* – 100 мин; *в, г* – 80 мин; *д* – 60 мин; *е* – 45 мин; 1,0 %-ная агароза; $I = 14$ мМ; pH 5,3; *б, г* – 0,8 мМ эфира аргинина; *a, в, д* и *е* – 1,6 мМ HCl

В последнем случае электрофоретическая подвижность поли(A) в присутствии аргинина уменьшалась на 20 %, а в первом случае – на 16 %, т.е. несмотря на то, что дикатионы мигрируют навстречу полинуклеотиду быстрее, чем монокатионы, в экспериментах, в которых эфир аргинина вводили только в гели, вследствие уменьшения подвижности полинуклеотидов в присутствии лигандов с двумя зарядами, поли(A) все же большую часть пути мигрировал в среде лиганда. Увеличение количества DS_p -формы в присутствии эфира аргинина, по-видимому, связано со структурным переходом поли(A) из DS - в DS_p -форму при образовании комплекса с производным аргинина.

Как уже упоминалось в параграфе 1.1.1., структурная модель DS -формы поли(А) до сих пор не представлена в литературе, поэтому важно иметь дополнительную информацию об ее физико-химических характеристиках. Методом электрофореза мы выявили, что при низких ионных силах ($I < 30$ мМ) DS -форма имеет меньшую электрофоретическую подвижность, чем DS_p -форма, а при больших ионных силах значения подвижностей двух форм сближаются. Взаимодействие поли(А) с дикатионом производного аргинина способствует переходу DS -формы полимера в DS_p -форму.

4.2.2. Электрофоретическое исследование «замороженной» формы поли(А). Специфической особенностью поли(А) является образование объемистых структур при кислых рН. На рис. 4.15 показаны результаты электрофореза препарата поли(А)_С, растворенного в трис-ацетатном буфере, рН 5,5 (сплошная линия).

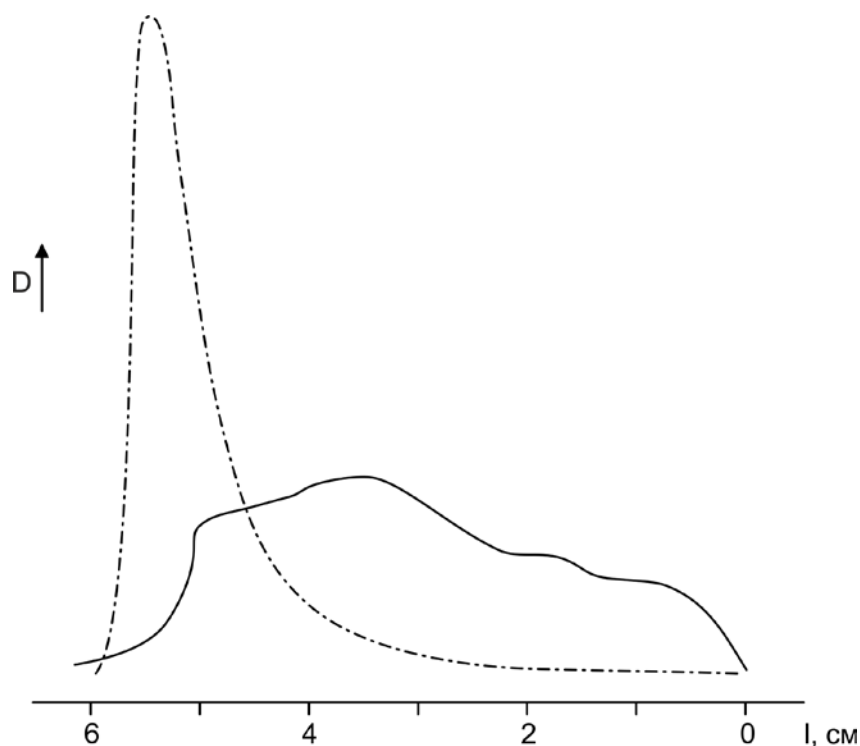


Рис. 4.15. Денситограммы поли(А)_С. 0,8 %-ная агароза; $I = 43$ мМ; рН препарата: 5,5 – сплошная линия, 7,5 – штрих-пунктирная линия

Как видно на рис. 4.15, препарат размыт вдоль всего пути миграции, а в некоторых условиях часть препарата совсем не проникает в гель и остается на старте. Картина электрофореза не менялась, когда образцы выдерживали 20 ч перед нанесением на гель или готовили растворы при разной ионной силе, или предварительно растворяли препараты при нейтральном pH, а затем, интенсивно перемешивая, подкисляли. Однако если кислый раствор поли(А) прогреть несколько минут при ~ 75 °С перед нанесением на гель, препарат мигрирует при кислых pH в виде отдельной зоны; такой же результат достигается и в другом случае, если на кислый гель наносить нейтральный раствор поли(А) (рис. 4.15, штрих-пунктирная линия). Следует отметить, что на рис. 4.15 зона поли(А) при pH 5,5 и $I = 43$ мМ не является двойной. Причины регистрации двух кислых форм поли(А) методом электрофореза только при низких ионных силах (< 30 мМ) обсуждались в параграфе 4.2.1.

По всей видимости, размытые зоны поли(А) в наших экспериментах относятся к «замороженной» форме полимера, описанной Janik et al. [28], т.е. агрегатам всевозможных размеров, образующимся вследствие плохой подгонки цепей в двойных спиралях поли(А) при быстром подкислении. Прогрев, как и в работе [28], облегчает процесс разделения нитей и их последующую правильную подгонку, а нанесение нейтрального раствора препарата на кислый гель, по-видимому, обеспечивает его медленное подкисление, так же как при диализе нейтрального раствора поли(А) против кислого буфера.

Учитывая литературные и наши данные о свойствах «замороженной» формы поли(А), мы предложили следующий механизм ее образования. При подкислении раствора поли(А) двойные спирали могут образовываться не только между разными макромолекулами, но и внутри одной и той же молекулы при ее сворачивании (рис. 4.16). При тепловом движении участок одноцепочечной поли(А) может принять такую конфигурацию, в которой два близких сегмента будут ориентированы параллельно (рис. 4.16 а, б). При протонировании остатков аденина эти сегменты могут образовать

внутримолекулярную двойную спираль (рис. 4.16 *в*). Как было упомянуто в обзоре литературы, цепи в двухспиральном поли(А) ориентированы параллельно [36]. При одновременном участии той же самой молекулы как в образовании внутримолекулярных, так и межмолекулярных двойных спиралей вероятность «распутывания» образовавшейся мультимолекулярной структуры уменьшается и образуется ядро «замороженной» формы (рис. 4.16 *з*).

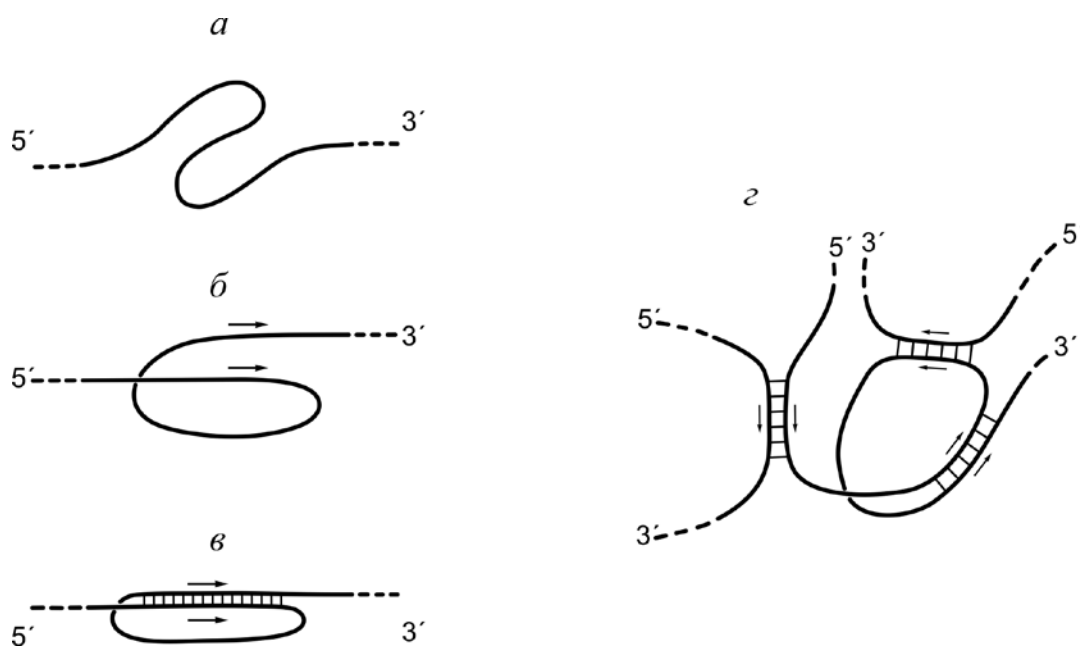


Рис. 4.16. Схемы образования внутримолекулярной двойной спирали в поли(А) (*а-в*) и зародыша «замороженной» формы (*з*). Стрелками отмечено 5' → 3' направление цепей

Предложенная модель согласуется со всеми свойствами «замороженной» формы поли(А), описанными в литературе и обнаруженными в наших исследованиях. Представленная структура образуется при кислых рН в условиях, когда уже полностью закончился переход из односпиральной в двухспиральную форму. Тем не менее, наряду с двухцепочечными участками она содержит также и одноцепочечные области [26, 28, 39] и может принимать большие размеры [35, 41], включая агрегаты, не способные проникать в гель (наши наблюдения). Молекулы поли(А) действительно «замерзают» в

одноцепочечной конформации при наличии беспорядочно распределенных коротких двухспиральных областей, мешающих подгонке остающихся одноцепочечных участков [28]. Агрегаты поли(А) («замороженная» форма) обладают очень низкой скоростью диссоциации [35], поскольку трудно представить, как может структура, показанная на рис. 4.16 з, «распутаться» без разрушения меж- или внутримолекулярных двойных спиралей путем прогрева или подщелачивания раствора.

Следует отметить, что о возможности существования внутримолекулярных двойных спиралей впервые сообщил Steiner в 1958 г. [238] в связи с обсуждением свойств кислой поли(А). Он предположил, что двойная спираль может образовываться или путем межмолекулярной ассоциации двух и более одноцепочечных цепей, или путем межмолекулярной реорганизации, при которой одиночная цепь изгибается таким образом, чтобы образовать двойную цепь, с замкнутой петлей на одном конце (рисунок в работе [238] не приведен).

Позднее Steiner & Beers [35] показали, что степень агрегации поли(А) при кислых рН повышается с увеличением ионной силы и концентрации макромолекул. Авторы предположили, что значительное уменьшение степени агрегации протонированного поли(А) при $I = 0,01$ М и низкой концентрации полимера ($C < 0,1$ г/л) может быть, в частности, связано с конкурентным образованием внутримолекулярных дуплексов.

Гипотеза об образовании внутримолекулярных спиралей в поли(А) вполне согласуется с литературными данными об образовании самоассоциатов других гомополирибонуклеотидов. Экспериментально доказано, что области ДНК и РНК, содержащие кластеры олиго(Г)- или олиго(С)-трактов, могут сворачиваться с образованием четырехцепочечных структур – соответственно G-квадруплексов и i-мотивов [30, 146]. Кроме того, внутримолекулярные спирали в поли(А), возможно, были зарегистрированы в работе Smith et al. [239], исследовавших поли(А) методом силовой атомной микроскопии. Авторы наблюдали линейные цепи этого полимера, содержащие

«шишкообразные» структуры, природу которых они не установили. Учитывая, что эксперименты проводились в очищенной воде (Milli-Q), которая слегка подкислена [240], а также тот факт, что при низкой ионной силе рК перехода поли(А) в двухцепочечную форму увеличивается [40], можно предположить, что в этих экспериментах происходило протонирование аденина и образование внутримолекулярных двухцепочечных спиралей в поли(А) («шишек»).

С целью проверки гипотезы о существовании внутримолекулярных дуплексов поли(А) мы провели более детальное исследование «замороженной» формы. Изучение этой формы представляет интерес также в связи с тем, что в большинстве работ, посвященных исследованию двуспирального поли(А), не принимались специальные меры по предотвращению ее образования и, соответственно, не учитывалось ее влияние на результаты экспериментов.

Если внутримолекулярные дуплексы обуславливают появление «замороженной» формы поли(А), то эффективность ее образования должна зависеть от факторов, влияющих на гибкость полимерной цепи, в частности, от длины макромолекулы и ионной силы раствора. На рис. 4.17 представлены денситограммы поли(А)_s, приготовленного в условиях, предотвращающих образование замороженной формы (параграф 2.2).

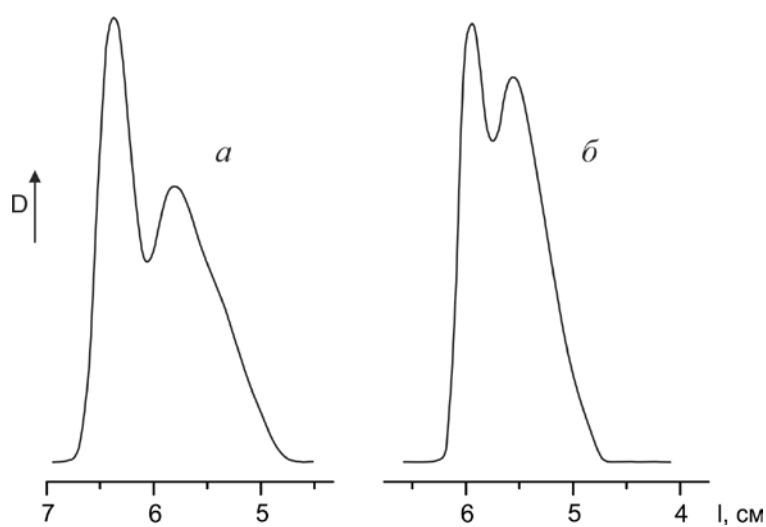


Рис. 4.17. Денситограммы поли(А)_s в трис-СН₃СООН буфере; *a* – I = 11,7 мМ; *б* – I = 24,1 мМ; рН = 5,5; 0,8 %-ная агароза. Препараты приготовлены в условиях, предотвращающих образование «замороженной» формы

Электрофорез проведен при рН 5,5 и ионной силе 11,7 мМ и 24,1 мМ. Как видно на рисунке, при более высокой ионной силе (рис. 4.17 б) зоны поли(А) более узкие и медленнее мигрирующие вследствие уменьшения полного заряда и размера молекул полимера.

На рис. 4.18 б и в представлены денситограммы препарата поли(А)_s, растворенного непосредственно в кислом буфере (рН 5,5) при ионных силах 8,3 и 27,9 мМ соответственно. Как видно на рис. 4.18, при этих значениях I зона полимера значительно размыта, и часть образца остается на старте. В этом случае трудно определить, как изменяется относительное количество «замороженной» формы поли(А) при изменении ионной силы. Прогрев образца поли(А)_s, растворенного непосредственно в кислой среде, не приводит к эффективному удалению «замороженной» формы (4.18 а), хотя значительно уменьшает ее количество.

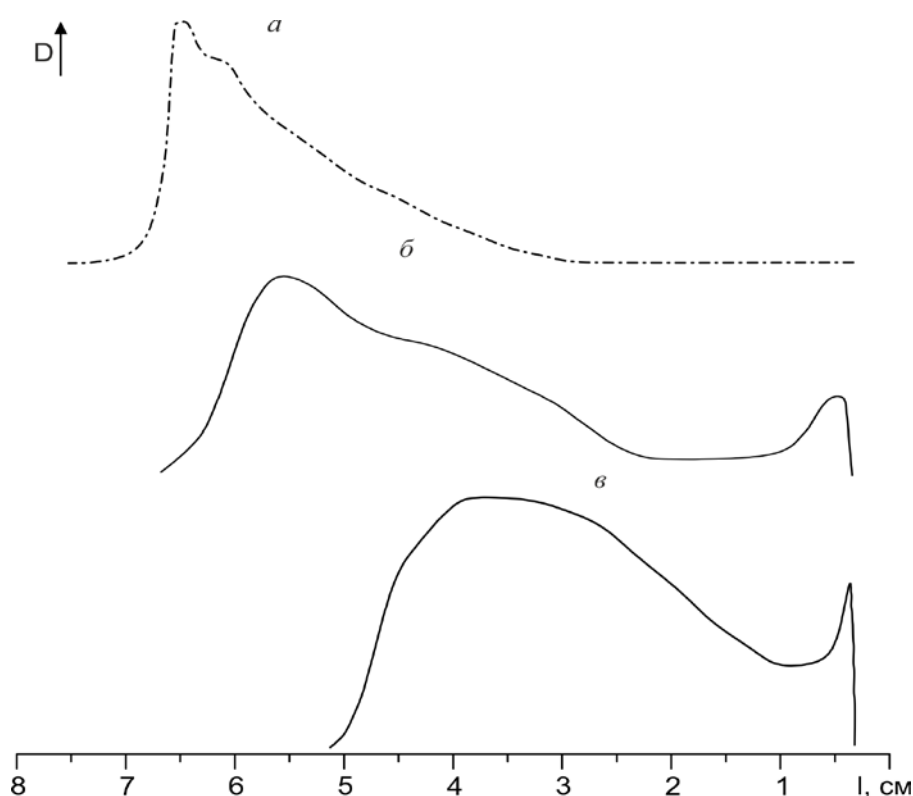


Рис. 4.18. Денситограммы поли(А)_s в трис-СН₃СООН буфере; а, б – I = 8,3 мМ; в – I = 27,9 мМ; рН = 5,5; 0,8 %-ная агароза; а – препарат прогрет

«Замороженная» форма образуется в большом количестве и при растворении в кислом буфере препаратов поли(A)_C, близких по размерам молекул к препаратам поли(A)_S. Однако при приготовлении непосредственно в кислом буфере растворов поли(A)_R (с меньшей молекулярной массой) образуется лишь малое количество «замороженной» формы (рис. 4. 19). Таким образом, увеличение длины поли(A)-цепей и/или ионной силы, способствуя возрастанию гибкости полимера, приводит к увеличению количества «замороженной» формы. Эти результаты свидетельствуют в пользу нашего предположения о том, что поли(A) способен эффективно сворачиваться с образованием внутримолекулярных двойных спиралей, а «замороженная» форма возникает вследствие одновременного образования как межмолекулярных, так и внутримолекулярных дуплексов.

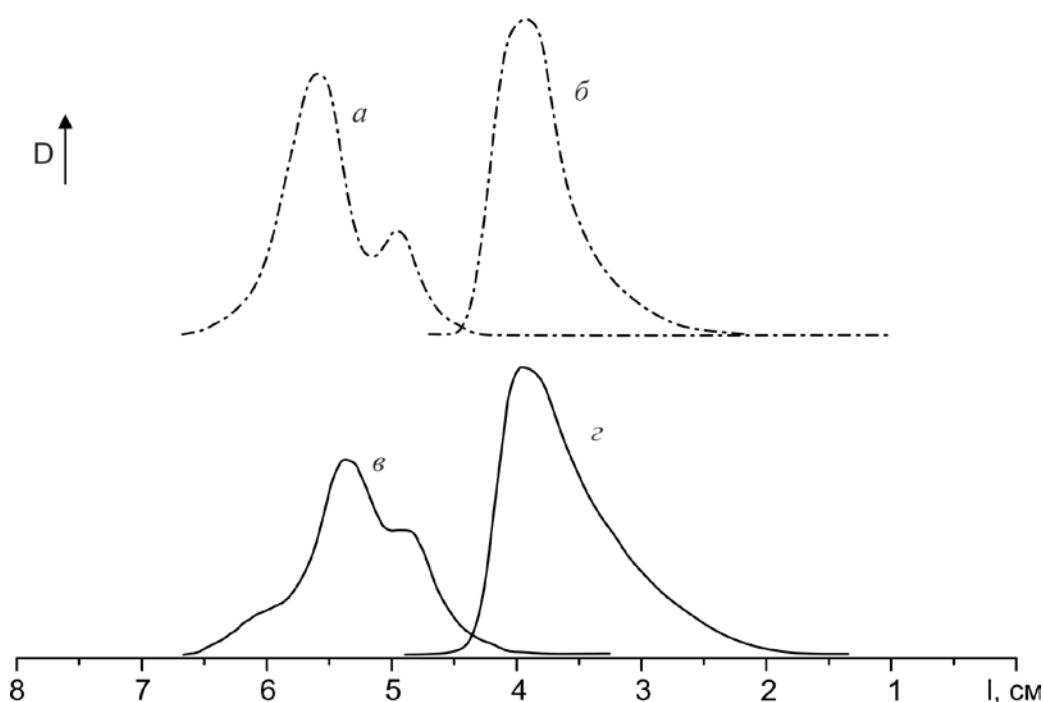


Рис. 4.19. Денситограммы поли(A)_R в трис-CH₃COOH буфере; *a, в* – $I = 6,7$ мМ; *б, г* – $I = 25,9$ мМ; pH = 5,5; 0,8 %-ная агароза; *a, б* – препарат прогрет

В пользу образования внутримолекулярных двойных спиралей в поли(A) при кислых pH свидетельствуют также результаты недавно опубликованной

работы Kim et al. [241], исследовавших структурные переходы в поли(dA) методом FRET (fluorescence resonance energy transfer) и FSC (fluorescence correlation spectroscopy). Они показали, что при кислых pH в сильно разбавленных растворах поли(dA) ($c < 20$ нМ) образуется компактная, предположительно, одноцепочечная форма, содержащая короткие двухцепочечные участки. Авторы назвали эту форму S-формой (shrunken form). Наличие такой формы подтверждено измерением в этой работе гидродинамического радиуса S-формы. S-форма подобна структуре, изображенной на нашем рисунке 4.16 в.

Исследование «замороженной» формы поли(A) методом электрофореза дополнило информацию об ее свойствах, имеющуюся в литературе. Следует отметить, что согласно Janik et al. [28], «замороженная» форма образуется только из нейтральной формы поли(A) при быстром подкислении раствора до достаточно низких pH. Некоторые авторы, например Petrovic & Polavarapu (2005 г.) [242] и Kim et al (2011 г.) [241], ссылаясь, в частности, на эту работу [28], пишут, что «замороженная» форма поли(A) образуется из нейтральной формы только при низких pH ($< 3,8$). Кроме того авторы [242] называют «замороженной» формой гель поли(A). Как следует из наших данных, «замороженная» форма не является наиболее кислой формой поли(A), она появляется при протонировании поли(A) фактически при тех же значениях pH, что и другие кислые формы, вследствие одновременного образования, по нашему мнению, как межмолекулярных, так и внутримолекулярных двойных спиралей. Поэтому возможность ее образования следует учитывать при исследовании протонированных форм поли(A) при всех значениях pH, при которых образуются двухспиральные формы.

4.2.3. Исследование структурных переходов в поли(A) методом протонной буферной емкости. Поскольку структурные переходы в поли(A) и поли(C) вызываются протонированием оснований, наряду с другими физико-химическими методами для исследования этих гомополирибонуклеотидов используется метод потенциометрического титрования. Для изучения

структурных переходов в поли(А) этот метод применяли, в частности в работах [27, 39, 35, 40], при этом наличие двух двухспиральных форм (DS и DS_p) было зарегистрировано только при титровании изоионных растворов этого гомополимера, в которых все катионы, экранирующие фосфатные группы, были замещены на протоны [39]. Однако обе двухспиральные формы могут быть выявлены и при титровании солевых растворов поли(А), если для анализа кривых потенциметрического титрования использовать высокочувствительный метод протонной буферной емкости, разработанный и обоснованный для биохимических объектов Опанасенко и соавт. [202]. Метод заключается в дифференцировании исходной кривой потенциметрического титрования и разложении полученной кривой, т.е. зависимости буферной емкости от рН, на базисные функции, соответствующие протонированию химических групп различного типа (см. параграф 2.3).

4.2.3.1. Кривая буферной емкости поли(А) и ее разложение по базисным функциям. На рис. 4.20 показана кривая потенциметрического титрования поли(А), рассчитанная как разность кривых титрования раствора поли(А) и растворителя. Она построена в координатах (α , рН), где α - степень протонирования поли(А), определяемая как отношение количества протонированных оснований к общему их числу.

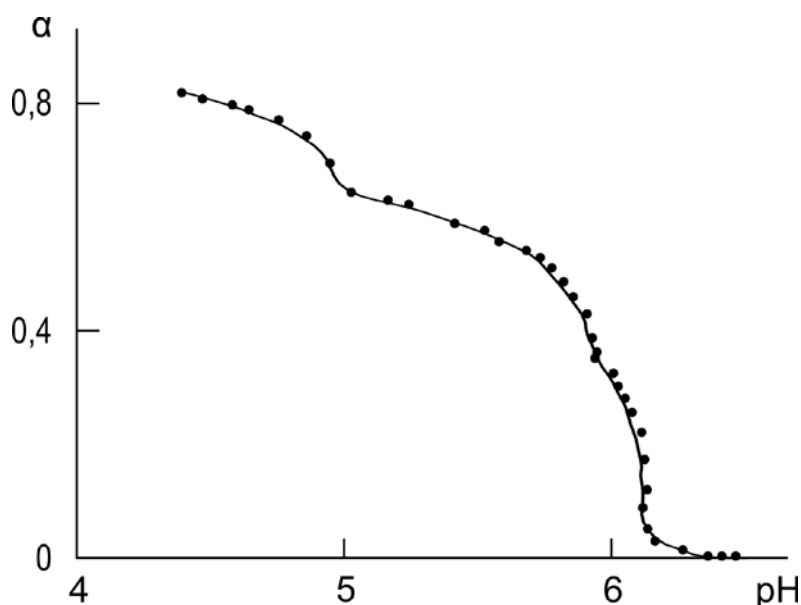


Рис. 4.20. Кривая потенциметрического титрования поли(А)_R. $I = 0,05$ М; $T = 13,5$ °С; 2,8 мМ поли(А)

Растворы титровали в присутствии небольших концентраций трис – это позволяло регистрировать начальный участок кривой потенциметрического титрования поли(А) при титровании кислотой. В отсутствие низкомолекулярного буфера рН раствора поли(А) неконтролируемо снижается до величины, при которой сам полинуклеотид начинает проявлять буферные свойства. Кривая потенциметрического титрования поли(А) имеет в основном такую же форму, как и кривые, приведенные в работах других авторов [27, 39, 35, 40].

Отличительной особенностью кривых титрования поли(А) является наличие большого (до $\alpha \sim 0,45$) начального участка с высокой буферной емкостью, соответствующего переходу поли(А) из одноцепочечной формы в двухцепочечную [35]. Как видно на рис. 4.20, в области $\alpha \sim 0,6$ имеется еще один небольшой участок с высокой буферной емкостью. Такой участок также наблюдается на кривых потенциметрического титрования, приведенных в литературе, но авторами не обсуждается [27, 39, 35, 40].

На рис. 4.21 показаны кривые рН-зависимости буферной емкости раствора поли(А) (темные кружки) и контрольного раствора (светлые кружки) при $I = 0,05$ М. В работе не прилагались специальные усилия для удаления из растворов атмосферного CO_2 – он удалялся в значительной степени при прогреве растворов (параграф 2.3), так как при этом значения β в диапазоне рН 6-7 (pK_1 угольной кислоты равен 6,4) для растворов, используемых длительное время, уменьшались до значений, соответствующих свежеприготовленным растворам.

Как видно на рис. 4.21, в области рН протонирования поли(А) значения β растворителя в основном малы по сравнению со значениями β раствора поли(А). В диапазоне рН, предшествующем области протонирования поли(А) ($pH > 6,4$), значения β раствора поли(А) и контрольного раствора практически совпадали во всех экспериментах.

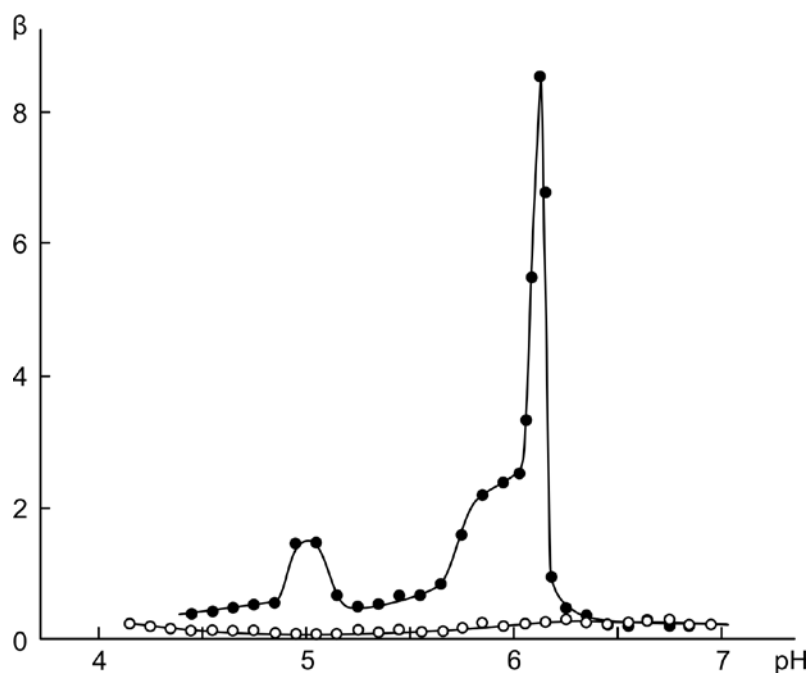


Рис. 4.21. Кривые буферной емкости раствора поли(A)_R (●) и растворителя (○). Условия эксперимента, как на рис. 4.20

На рис. 4.22 показана кривая pH-зависимости буферной емкости поли(A), вычисленная как разность кривых для раствора поли(A) и растворителя, представленных на рис. 4.21. Как видно на рисунке, она имеет три пика. Два участка протонирования поли(A) в области pH 5,7-6,2, которые трудно выделить на кривой титрования, легко различаются на кривой буферной емкости. Подобные кривые буферной емкости, имеющие три пика, получаются также, как будет показано ниже, при обработке кривых потенциометрического титрования, приведенных в литературе.

На рис. 4.22 представлено также разложение кривой pH-зависимости буферной емкости поли(A) по базисным функциям вида (2.8). Поскольку кривая буферной емкости полимера представляет собой сумму кривых, соответствующих однопипным группам ($\gamma \leq 1$) и pH-индуцированным конформационным переходам ($\gamma > 1$) [202], то из рис. 4.22 следует, что при протонировании поли(A) имеют место три конформационных перехода: один

высококооперативный (с pK 6,16) и два менее кооперативных (с pK 5,97 и 5,00). Параметры пиков представлены в таблице 4.1.

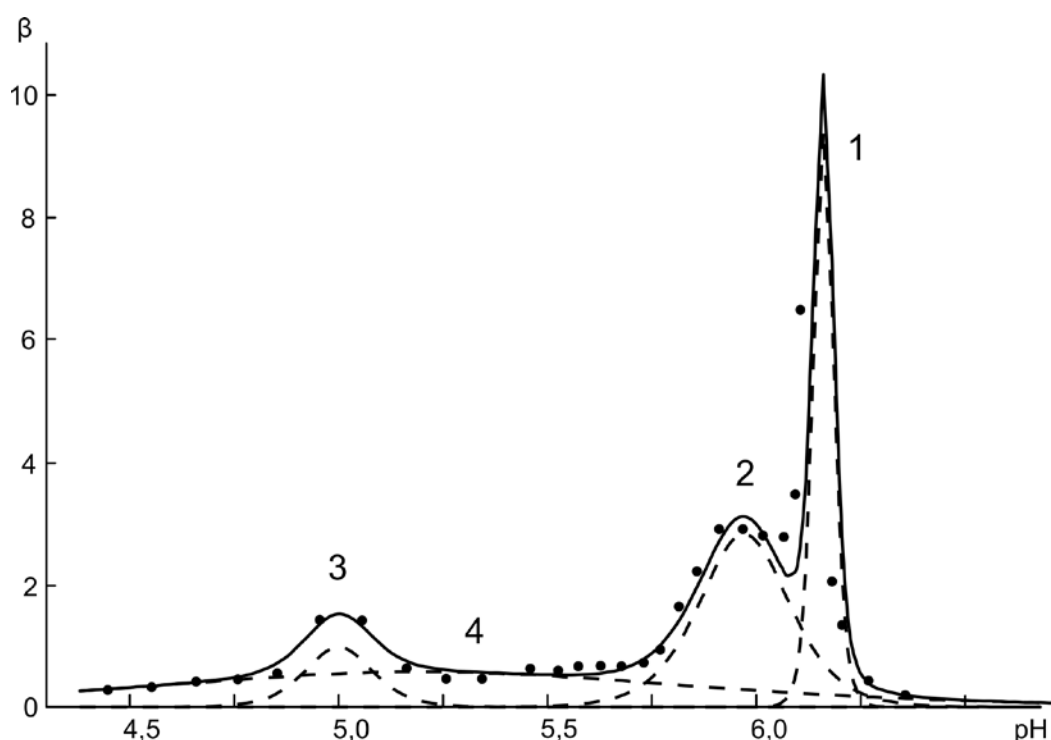


Рис. 4.22. Кривая буферной емкости поли(A)_R и ее разложение на составляющие. Пунктирные линии – базисные функции; сплошная линия – суммарная кривая. Условия эксперимента, как на рис. 4.20

Таблица 4.1

Параметры пиков разложения кривой буферной емкости поли(A) при $I = 0,05$ M на базисные функции β_i .

Номер пика	pK_i	γ_i	c_i/c
1	$6,16 \pm 0,01$	27 ± 3	$0,20 \pm 0,03$
2	$5,97 \pm 0,01$	$6 \pm 0,5$	$0,27 \pm 0,02$
3	$5,00 \pm 0,01$	$8 \pm 0,5$	$0,07 \pm 0,01$
4	$5,20 \pm 0,20$	1	$0,33 \pm 0,10$

Примечание. Определение параметров дано в параграфе 2.3; c – общая концентрация поли(A).

В области полупротонирования поли(А) (рН 5,8) второй конформационный переход, как это видно на рис. 4.22, в основном уже завершен. Исходя из литературных данных о том, что сразу после перехода из одноцепочечной формы в двухцепочечную, поли(А) существует только в *DS*-форме [26-28], мы относим первый (самый узкий) пик на кривой буферной емкости к образованию *DS*-формы, второй пик, перекрывающийся с первым, – к образованию *DS_p*-формы.

Третий пик, наблюдаемый в области рН 4,8-5,2, мы относим к конформационному изменению *DS*-формы поли(А), описанному Lerner & Kearns [29]. Исследуя зависимости химических сдвигов ³¹P в спектрах ЯМР поли(А) от рН, авторы показали, что в интервале рН ниже 5,4 происходит, по видимому, небольшое конформационное изменение двойной спирали по мере ее приближения к состоянию, в котором все основания протонированы. Это конформационное изменение связано с фосфатами, не участвующими в образовании водородных связей (т.е. с *DS*-формой поли(А)).

Кривая буферной емкости поли(А) может быть удовлетворительно представлена в виде суммы четырех базисных функций (рис. 4.22). Три базисных функции соответствуют вышеназванным конформационным переходам. Четвертая базисная функция отражает процесс некооперативного ($\gamma = 1$) присоединения протонов к двухспиральному поли(А). Параметры этого пика вследствие его большой ширины не удастся определить с большой точностью. В работе использованы значения рК для четвертого пика 5,0-5,4 и значения c_4/c в интервале 0,2-0,4.

В эксперименте, результаты которого показаны на рис. 4.20-4.22, поли(А) титровали до рН 5,3, прогревали, как указано в параграфе 2.3, затем продолжали титрование до появления геля. В эксперименте, показанном на рис. 4.23, поли(А) после титрования до рН 5,2 и прогрева титровали обратно щелочью. Как видно из рис. 4.22 и 4.23 а, кривые буферной емкости поли(А) (полученные при той же ионной силе и температуре) в интервале рН кооперативного образования *DS*- и *DS_p*-форм при прямом и обратном

титровании существенно не отличаются. Сравнение кривых прямого и обратного титрования проводилось с целью проверки влияния возможных примесей «замороженной» формы поли(А) на форму кривой буферной емкости.

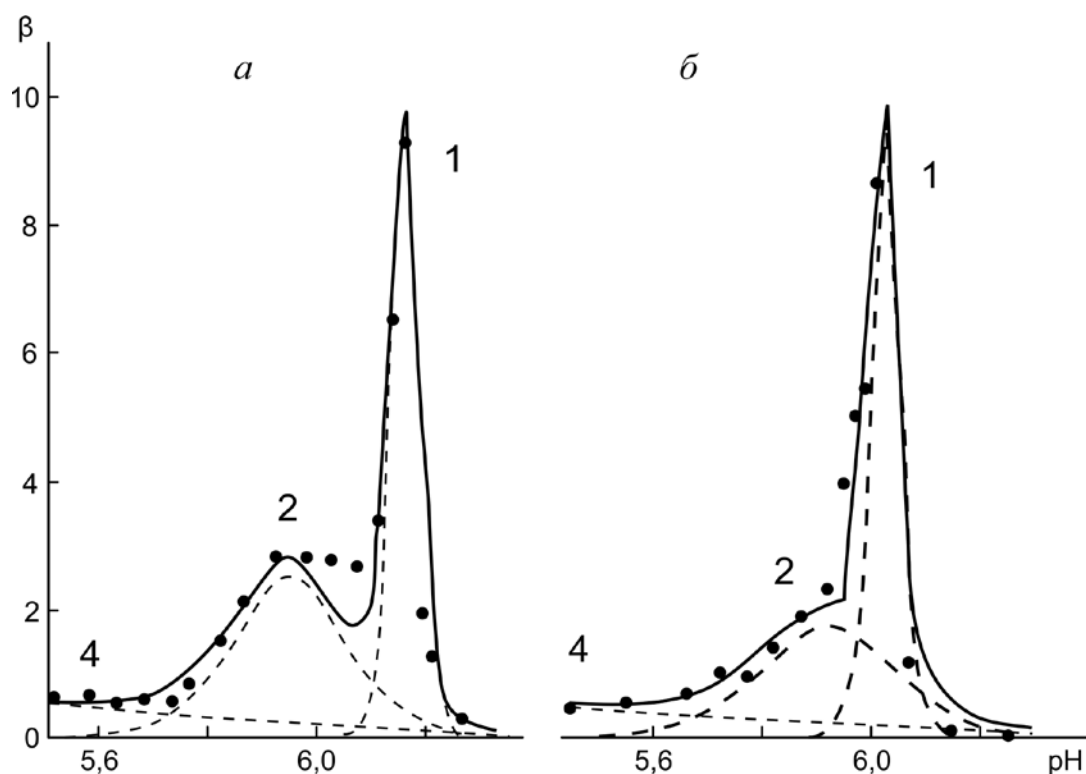


Рис. 4.23. Кривая буферной емкости поли(А)_Р и ее разложение на составляющие. Пунктирные линии – базисные функции, сплошные линии – суммарные кривые. $I = 0,05 \text{ M}$; $2,5 \text{ mM}$ поли(А)_Р; *a* – $T = 13,5 \text{ }^\circ\text{C}$. *б* – $T = 24,5 \text{ }^\circ\text{C}$

Сравнение результатов ряда экспериментов по титрованию прогретых и непрогретых образцов поли(А) показывает, что значения β в интервале рН 5,2-5,6, в котором в основном происходит процесс некооперативного присоединения протонов к поли(А), заметно ниже в случае непрогретого раствора. Этот факт, возможно, свидетельствует об образовании «замороженной» формы поли(А) в процессе титрования непрогретого образца, что приводит к уменьшению концентрации двухспиральных форм, а соответственно и c_4/c .

4.2.3.2. Диаграмма распределения отдельных форм поли(А) в процессе протонирования. Чтобы проиллюстрировать характер распределения отдельных форм поли(А) в процессе протонирования, мы построили модельную диаграмму, показанную на рис. 4.24. При ее построении использованы результаты разложения кривой буферной емкости поли(А), полученные при $I = 0,05$ М и $T = 24,5$ °С (рис. 4.23 б), по базисным функциям, а также литературные данные [27].

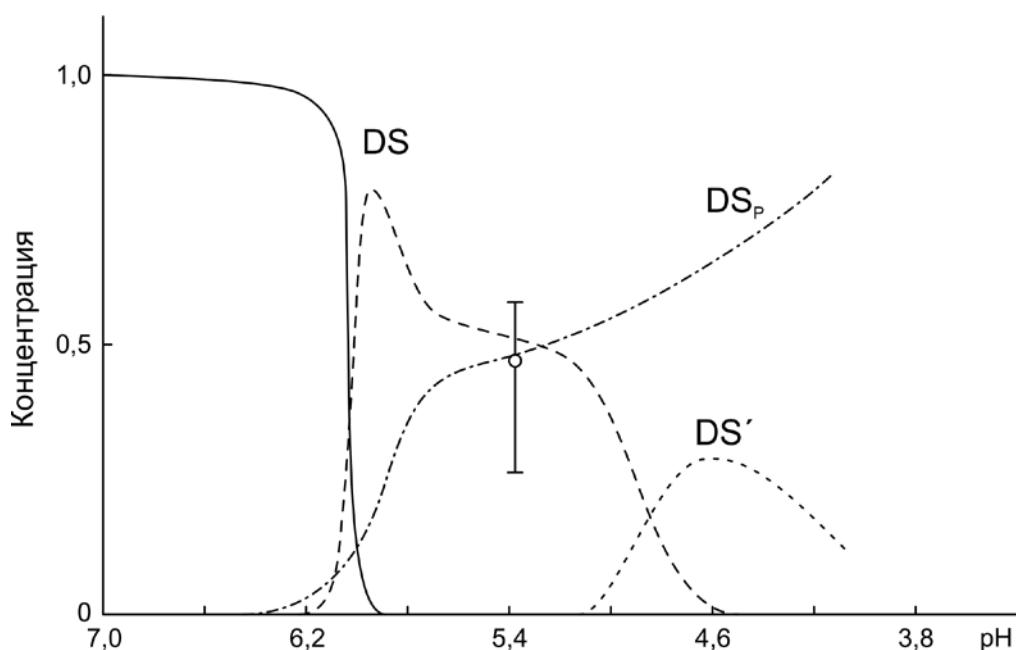


Рис. 4.24. Зависимость концентрации (в относительных единицах) отдельных форм поли(А) от рН при $I = 0,05$ М. Сплошная линия относится к нейтральной форме. DS' обозначает DS -форму, претерпевшую конформационное изменение. \circ – выбранное значение c_2/c

Ширина перехода поли(А) из одноцепочечной формы в двухцепочечную, которую можно определить из кривых спектрофотометрического титрования ([27] и наши данные), меньше 0,3 ед. рН, т.е. при рН ниже 5,90, при котором практически завершился кооперативный процесс образования DS -формы (рис. 4.23 б), одноцепочечная форма поли(А) отсутствует. Степень протонирования поли(А) в интервале рН

= 5,3-5,9, где отсутствуют одноцепочечная и DS' -форма, можно выразить следующим соотношением:

$$\alpha = \alpha_1 c_1 + \alpha_2 c_2, \quad (4.1)$$

где α_1 и α_2 – степени протонирования,

c_1 и c_2 – концентрации (в относительных единицах) соответственно DS - и DS_P -форм; при этом:

$$c_1 + c_2 = 1. \quad (4.2)$$

Количество протонированных остатков, входящих в состав DS -формы (или DS_P -формы) поли(А), можно вычислить как сумму остатков, протонированных в ходе кооперативного и некооперативного процессов. При рН 5,4, при котором завершился кооперативный переход в DS_P -форму (рис. 4.23 б), количество протонов кооперативно присоединенных к DS_P -форме поли(А) (c_2/c), определяется из разложения кривой буферной емкости. Количество некооперативно присоединенных протонов можно определить по площади соответствующего участка под кривой базисной функции четвертого пика. Мы предположили, что количество протонов, некооперативно присоединенных к DS - и DS_P -формам поли(А), пропорционально их концентрациям. Считая, что

$$\alpha_1 < \alpha_2 < 1, \quad (4.3)$$

т.е., что степень протонирования DS -формы не превышает в ходе титрования степень протонирования более кислой DS_P -формы, мы можем определить при рН 5,4 интервал значений c_2 , удовлетворяющий соотношениям (4.1-4.3). Такой интервал показан на кривой зависимости концентрации DS_P -формы от рН (рис. 4.24, штрих-пунктирная линия), как ошибка c_2 при значении рН 5,4. Значение

c_2 на модельной диаграмме при этом рН выбрано в пределах ошибки произвольно, но с учетом того обстоятельства, что по результатам ДОВ [27] при рН 5,0 ($I = 0,1$ М) в DS_p -форме находится 60 % поли(А) и в DS -форме – 40 %.

При построении диаграммы в области рН ниже 5,4 мы предположили, что при рН 4,0, при котором методом ДОВ DS -форма поли(А) уже не регистрируется [27], поли(А) все же не находится полностью в DS_p -форме. Метод ДОВ, по-видимому, не является достаточно чувствительным для регистрации небольших концентраций разных форм поли(А), поскольку не регистрирует наличие небольшого количества DS_p -формы наряду с DS -формой при переходе поли(А) из одноцепочечной формы в двойную спираль. Наличие этой формы при переходе видно на кривых буферной емкости, полученных в наших экспериментах, а также при обработке кривых потенциометрического титрования, представленных в литературе.

Диаграмма построена в интервале рН $> 4,1$, поскольку в данном эксперименте, в котором концентрация поли(А) равна 2,5 мМ, при этом значении рН образуется гель. Менее концентрированные растворы превращались в гель при более низких значениях рН. Образование гелей при низких рН известно из литературы [243].

«Модельность» диаграммы заключается лишь в том, что она построена для произвольно выбранного (из определенного интервала, показанного на рис. 4.24 как ошибка) значения концентрации c_2/c при значении рН, при котором завершился кооперативный переход в DS_p -форму. В случае поли(А) метод протонной буферной емкости не позволяет точно определить концентрации разных двухспиральных форм, а лишь сделать их оценку. Но характер изображенных на диаграмме кривых не зависит от выбранного значения концентрации. То есть, как это видно из рис. 4.24, DS_p -форма поли(А) образуется в два этапа, первый из которых кооперативен, но степень кооперативности образования этой формы значительно меньше, чем DS -формы.

4.2.3.3. Зависимость параметров структурных переходов в поли(А) от условий среды. Для того чтобы сравнить полученные нами результаты с литературными данными, а также выявить зависимость параметров кооперативных переходов поли(А) от температуры и ионной силы, мы провели дифференцирование потенциометрических кривых титрования поли(А), представленных в литературе [35, 40]. На рис. 4.25 показаны результаты разложения по базисным функциям кривых буферной емкости поли(А), полученных при постоянной температуре и разных значениях ионной силы.

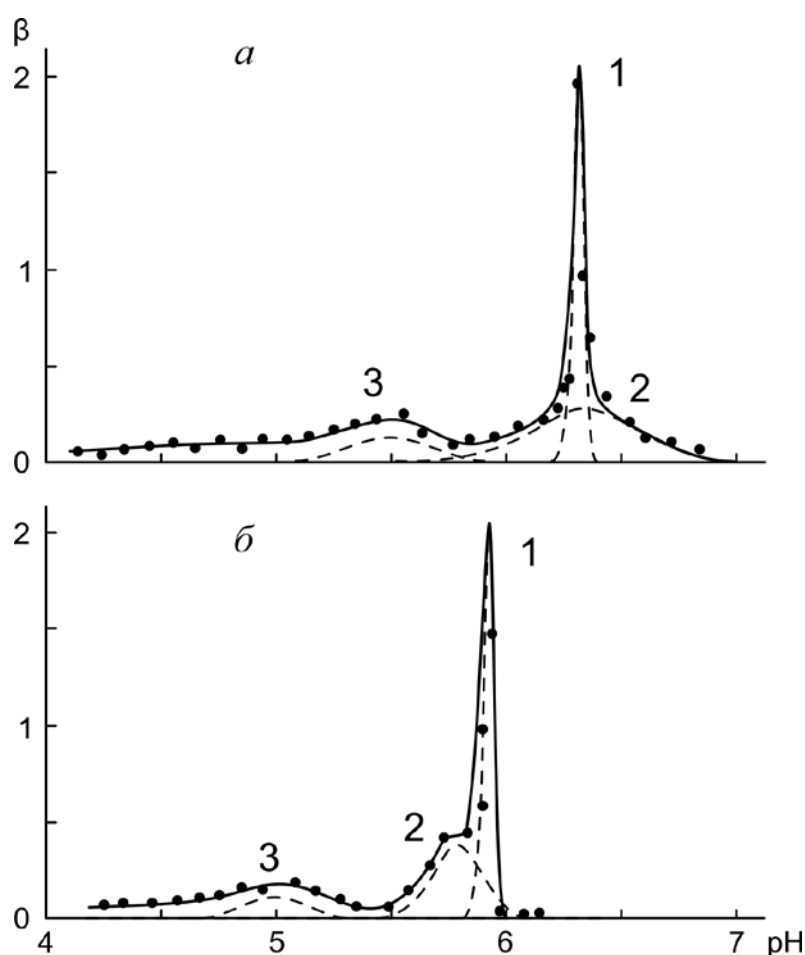


Рис. 4.25. Кривые буферной емкости поли(А), рассчитанные по данным работы [35], и их разложение на составляющие. $T = 26\text{ }^{\circ}\text{C}$; *а* – $I = 0,01\text{ M}$; $0,53\text{ mM}$ поли(А); *б* – $I = 0,1\text{ M}$; $0,46\text{ mM}$ поли(А). Показаны базисные функции кооперативных переходов (пунктирные линии) и суммарные кривые

Из рисунка 4.25 также видно, что при низкой ионной силе процесс образования DS_p -формы (пик 2) становится менее кооперативным и увеличивается количество остатков аденина, протонируемых при образовании этой формы. Кроме того, в этом случае DS -форма (пик 1) появляется при закислении при значительно более низких значениях pH, чем DS_p -форма. Похожие кривые буферной емкости наблюдаются также при обработке кривых потенциометрического титрования, полученных при низкой ионной силе в других работах [39, 40] и наших экспериментах. Кривая pH-зависимости поли(А) при $I = 0,1$ М (рис. 4.25 б) имеет такую же форму, как и кривая, полученная в наших экспериментах при $I = 0,05$ и $T = 24,5$ °С (рис. 4.23 б).

На рис. 4.23 б показана кривая буферной емкости поли(А), полученная при температуре на 11 °С выше, чем в эксперименте, показанном на рис. 4.23 а. Как видно на рисунке, при повышении температуры пики DS - и DS_p -форм смещаются в кислую область, причем смещение больше в случае DS -формы; пики сближаются. Аналогичный характер имеет и зависимость рК конформационных переходов в поли(А) от температуры, рассчитанная на основании литературных данных [40].

На рис. 4.26 а показаны зависимости рК от температуры для всех трех конформационных переходов. Из зависимостей рК конформационных переходов поли(А) от ионной силы и температуры следует, что пики DS - и DS_p -форм полимера на кривой буферной емкости должны в наибольшей степени отстоять друг от друга по шкале pH в условиях высокой ионной силы и низкой температуры.

На рис. 4.26 б показана в качестве примера кривая буферной емкости поли(А) при таких условиях ($I = 0,15$ М и $T = 10,0$ °С), рассчитанная по данным работы [40]. При разложении этой кривой по базисным функциям использованы пять базисных функций. Пятая базисная функция относится к титрованию одноцепочечной поли(А) (рК $\sim 3,8$ [39]), представляющей собой, по-видимому, «замороженную» форму.

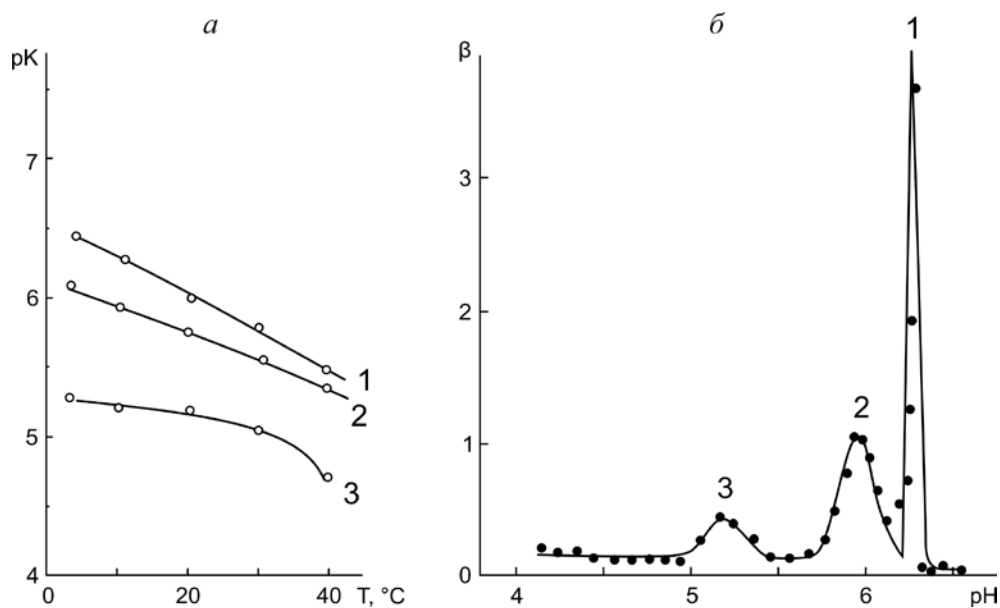


Рис. 4.26. Зависимость pK конформационных переходов в поли(А) от температуры, рассчитанная по данным работы [40] (а). $I = 0,15 \text{ M}$; $1,1 \text{ mM}$ поли(А). Кривая буферной емкости поли(А), рассчитанная по данным работы [40] (б). $I = 0,15 \text{ M}$; $T = 10 \text{ }^\circ\text{C}$; $1,1 \text{ mM}$ поли(А). Показана суммарная кривая

Как видно на рисунке, пики DS - и DS_p -форм в этом случае перекрываются лишь незначительно.

4.2.4. Гипотеза о стабилизации поли(А) вертикальными водородными связями.

4.2.4.1. DS -форма поли(А). Способность поли(А) к переходу в две двухспиральные формы, основания которых участвуют в образовании отличающихся водородных связей, является уникальной среди нуклеиновых кислот. Ее физико-химическая природа до сих пор не установлена. Структурная модель DS -формы поли(А) в литературе отсутствует. Известно только, что в обеих формах основания связаны между собой водородными связями одинаковым образом (модель, предложенная Rich et al. для DS_p -формы, рис. 1.1) и стэкинг-взаимодействия оснований примерно одинаковы [29].

Возникает вопрос, действительно ли стабилизация DS -формы поли(А) обеспечивается только стэкинг-взаимодействиями и двумя водородными связями между основаниями или существуют еще какие-то дополнительные стабилизирующие силы? Здесь интересно отметить, что аналог поли(А), в котором один из атомов водорода аминогруппы аденина замещен на метильную группу, не образует двойной спирали [244]. В результате такого замещения поли(6-МеА) не способен принимать DS_P -форму, в которой каждое основание участвует в формировании трех водородных связей, но форма с двумя водородными связями, DS -форма, в принципе, могла бы быть этим аналогом образована. Хотя отсутствие такой формы может быть связано со стерическими ограничениями, вносимыми метильной группой, не исключена и другая, более вероятная интерпретация, а именно – что второй атом водорода аминогруппы играет определенную роль в стабилизации DS -формы.

Стерические ограничения в поли(6-МеА) в действительности не являются критическими при образовании двойной спирали, поскольку при дальнейшей модификации этого полимера, т. е. замене атома водорода при атоме углерода С2 на аминогруппу, двойная спираль при соответствующих условиях образуется [245]. Причем, как показали авторы этой работы, она образуется с теми же N6Н...N7 Н-связями, что и в двойных спиральях поли(А). Из этих данных следует, что в данном случае какую-то стабилизирующую роль играет аминогруппа в положении 2.

Рассмотрим вопрос о структуре DS -формы поли(А) с другой стороны. Не обладает ли эта форма какими-нибудь специфическими особенностями, характерными для одной из форм одноцепочечного поли(А). Дело в том, что одноцепочечный поли(А), подобно двухцепочечному, существует в двух разных формах [12, 13]. Изучая динамику конформационных переходов «спираль-клубок» в одноцепочечном поли(А) с помощью так называемой кабельной техники температурного скачка, Porschke [12] впервые обнаружил, что временные константы релаксации полимера, регистрируемые в интервале длин волн 240-270 нм (τ_α) отличаются от констант, наблюдаемых при 285 нм

(τ_β), в 2-3 раза: так, например, в 0,1 М NaCl $\tau_\alpha = 0,27$, а $\tau_\beta = 0,53$ мкс. Porschke сделал вывод о том, что упорядоченный поли(A) существует не менее чем в двух конформационных состояниях.

Два времени релаксации отмечены не только для поли(A), но и для олигомеров – вплоть до динуклеотида, а также для поли(dA) [13], причем отношение τ_β/τ_α для поли(dA) превышает в несколько раз отношение для поли(A). Показано также, что τ_α , измеренные в D₂O, увеличиваются как в случае поли(A) (на 87 %), так и для поли(dA) (на 53 %), однако τ_β при изотопном обмене H₂O ↔ D₂O остаются неизменными. В отличие от поли(A) два разных релаксационных процесса не наблюдаются для поли(C) [13].

В настоящее время неизвестно, почему релаксационные свойства одноцепочечных поли(A) и поли(dA) по-разному проявляются в УФ-спектрах поглощения. Porschke полагает, что электронные переходы, вносящие вклад в поглощение при 285 и 240-270 нм, чувствительны к конформационным изменениям полимеров в разной степени. Здесь интересно отметить, что положительная полоса в спектре КД *DS*-формы поли(A) ($\lambda_{\max} \sim 280$ нм) смещена в длинноволновую область по сравнению с полосой в спектре *DS_p*-формы ($\lambda_{\max} \sim 260$ нм), а в случае одноцепочечного поли(A) положительная полоса с $\lambda_{\max} \sim 264$ нм имеет плечо при 280 нм [27, 246]. УФ-спектры поглощения обеих двухцепочечных форм поли(A) (*DS* и *DS_p*) одинаковы [27].

Одноцепочечный поли(dA) по сравнению с другими полинуклеотидами имеет необычный спектр КД с небольшими положительными полосами при 282 и 264 нм, первая из этих полос более интенсивная [27]. По данным ЯМР, основания в поли(dA) сильнее взаимодействуют между собой, чем в поли(A) [247]. По мнению авторов работы [43], вероятно, именно вследствие того, что одноцепочечный поли(dA) уже в значительной степени структурирован, он не переходит в двухцепочечные формы при значениях pH выше рК мономера в отличие от поли(A). В 0,1 М NaCl рК переходов поли(dA) и поли(A) в двухцепочечное состояние составляют соответственно 4,4 и 5,9 [43].

Таким образом, более структурированный одноцепочечный поли(dA) характеризуется по сравнению с одноцепочечным поли(A) значительно более высоким отношением временных констант релаксации τ_β/τ_α и более высокой интенсивностью полос спектра КД при 280 нм, чем при 260 нм, а спектр КД двуспиральной *DS*-формы поли(A) имеет положительную полосу с $\lambda_{\max} \sim 280$ нм в отличие от спектров одноцепочечной формы поли(A) и *DS_p*-формы, имеющих полосы с $\lambda_{\max} \sim 260$ нм. Отсюда можно сделать вывод о том, что внутримолекулярные взаимодействия, стабилизирующие *DS*-форму поли(A), могут быть подобны тем, которые стабилизируют одноцепочечный поли(dA).

В пользу существования двух форм одноцепочечного поли(A) свидетельствуют не только описанные выше результаты, полученные с помощью техники температурного скачка [12, 13], но и данные изучения этого полимера методом флуоресцентной спектроскопии [15]. Согласно модели, предложенной авторами работы [15], одноцепочечный поли(A) существует в двух формах с разным стэкингом оснований. Основная и минорная формы поли(A) имеют в спектре люминесценции полосы с $\lambda_{\max} = 400$ нм (возбуждение светом с $\lambda_{\text{ex}} = 260$ нм) и 420 нм ($\lambda_{\text{ex}} = 320$ нм) соответственно. Авторами показано, что электронное взаимодействие между основаниями в минорной форме сильнее, чем в основной.

Вероятнее всего, в работах [12, 13] и [15], выполненных разными методами, регистрируются одни и те же две формы одноцепочечного поли(A).

4.2.4.2. Гипотеза о структуре *DS*-формы поли(A). Природа сил, которые определяют стабильность односпирального поли(A) в настоящее время неясна. При анализе модели одноцепочечной спирали поли(A), полученной на основе рентгеноструктурного анализа [34], привлекает внимание такая ее структурная особенность, как наличие довольно близких контактов между аминокруппами соседних оснований (рис. 4.27 *a*). На основе этого наблюдения было предположено, что односпиральный поли(A) может быть стабилизирован не только стэкинг-взаимодействиями, но и так

называемыми вертикальными водородными связями ...N6H...N6H...N6H... [248, 249] (рис. 4.27 б).

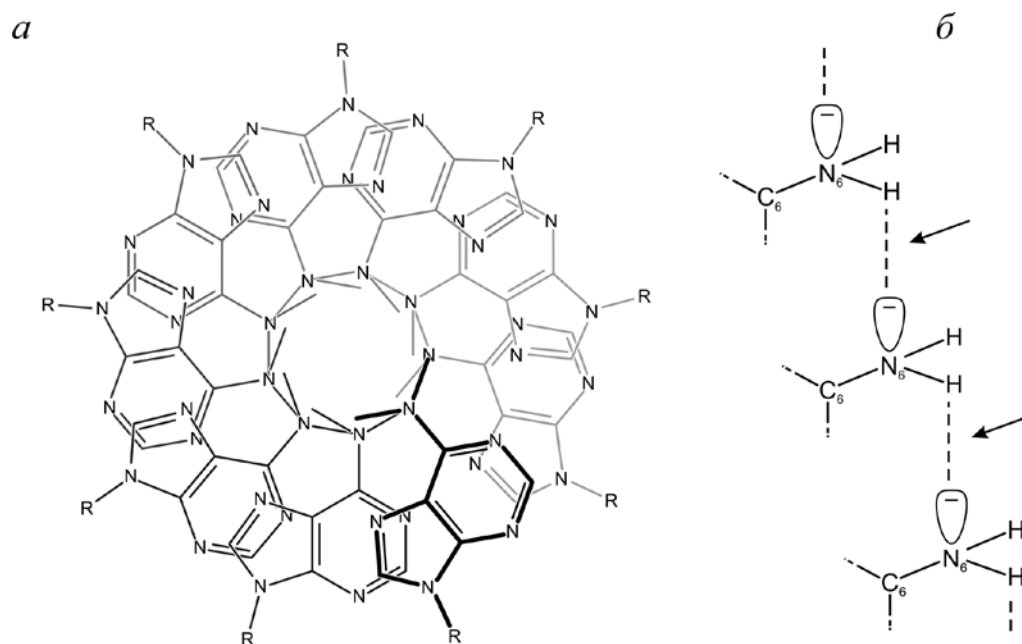


Рис. 4.27. Одноцепочечный поли(А), вид вдоль оси (а). Адаптировано из [34]; R – сахарофосфатный остаток. Вертикальные водородные связи в поли(А) (отмечены стрелками) (б)

Вертикальные Н-связи образуются (вследствие структурной нежесткости аминного фрагмента молекулы аденина) при выходе одного из атомов водорода аминогруппы из плоскости кольца основания и сближении его со свободной парой электронов атома азота аминогруппы соседней молекулы. В принципе такие связи могут образовываться в любых гомополи- и олигомерах, содержащих соответствующим образом расположенные адениновые основания, в частности, в поли(dA).

Мы предположили, что две конформации одноцепочечного поли(А) отличаются друг от друга как раз наличием или отсутствием вертикальных водородных связей [250]. Поскольку при наличии этих связей процесс релаксации является более сложным (происходит восстановление не только стэкинга, но и разрушенных водородных связей), то можно предположить, что

вертикальные связи образуются в той конформации поли(А), которая характеризуется бóльшей временной константой релаксации τ_B , наблюдаемой при 285 нм. Она соответствует минорной конформации поли(А) в работе [15]. Модель одноцепочечного поли(А), стабилизированного вертикальными водородными связями, представлена на рис. 4.28.

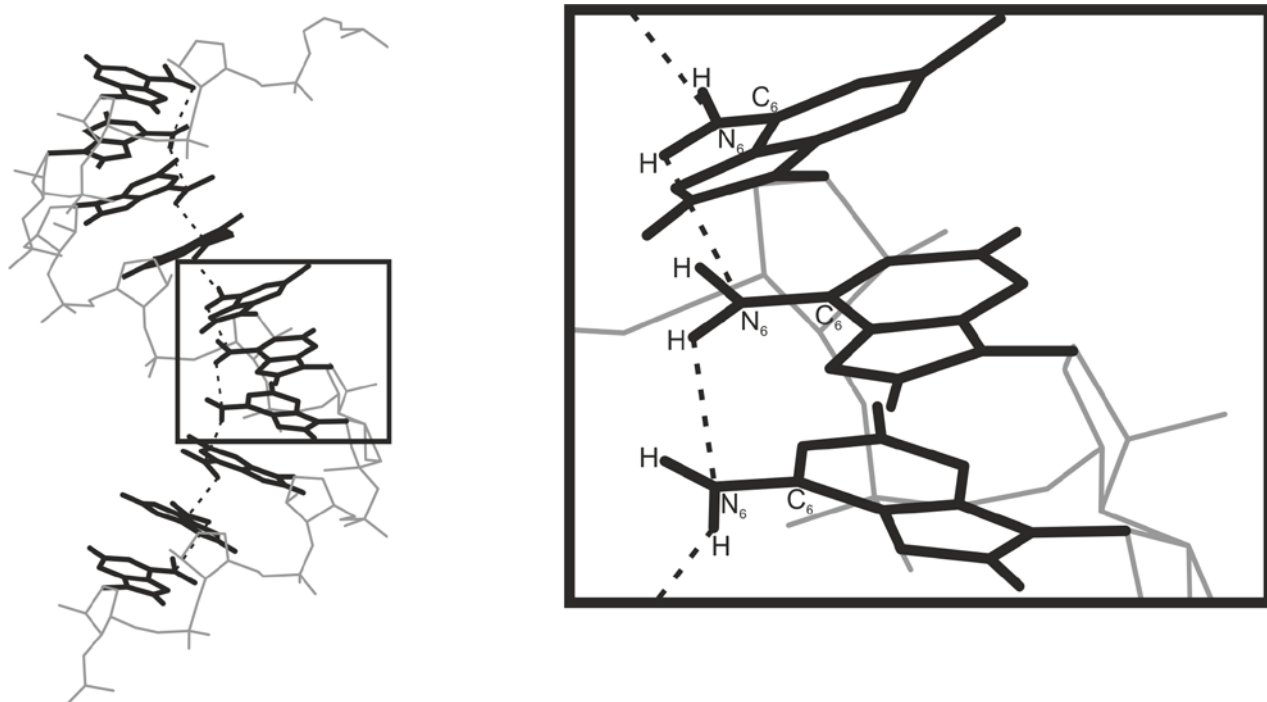


Рис. 4.28. Модель одноцепочечного поли(А), стабилизированного вертикальными водородными связями между аминогруппами соседних оснований (обозначены пунктиром). За основу взята модель поли(А) из работы [34]

Вертикальные водородные связи, кооперативно усиливая взаимодействие между адениновыми основаниями в одиночной цепи, могут усиливать и взаимодействие между парами адениновых оснований в двойной спирали. Скорее всего, что именно образование таких кооперативных (взаимоусиливающих) связей стабилизирует *DS*-форму поли(А), основания в которой связаны между собой только двумя горизонтальными водородными

связями. Эта кооперативность обусловлена, вероятно, тем, что втягивание аминного азота как акцептора в Н-связь, как известно [251, 252], усиливает протонодонорные свойства аминогруппы и наоборот. Предлагаемая гипотеза о стабилизации гомополимеров с адениновыми остатками кооперативными вертикальными водородными связями позволяет объяснить ряд экспериментальных фактов, изложенных и проанализированных выше.

По-видимому, одноцепочечный поли(dA) более структурирован, чем поли(A) по той причине, что конформация, характеризующаяся временной константой релаксации τ_β , является преобладающей в случае этого полимера, в сравнении с поли(A). Двух времен релаксации не наблюдается в случае одноцепочечного поли(C), поскольку аминогруппы цитозина в этом полимере не расположены благоприятным образом для устойчивого образования между основаниями вертикальных водородных связей. Изотопный эффект растворителя при обмене $\text{H}_2\text{O} \leftrightarrow \text{D}_2\text{O}$ отсутствует при конформации, характеризующейся временной константой релаксации τ_β , в отличие от конформации, характеризующейся τ_α , из-за наличия вертикальных водородных связей. Поли(6-MeA) не образует двойной спирали, на наш взгляд, из-за отсутствия как водородных связей с фосфатами, так и вертикальных водородных связей. Дальнейшая модификация этого аналога за счет присоединения аминогруппы к атому C2 может приводить к формированию вертикальных связей между этими аминогруппами (не участвующими в образовании классических горизонтальных водородных связей) и, как следствие, – к стабилизации двойной спирали.

Еще одно доказательство, свидетельствующее в пользу образования вертикальных водородных связей, можно найти в работе [253], посвященной изучению рамановского спектра (КР) одноцепочечного поли(A). Авторы показали, что интенсивность ряда полос, относящихся к плоскостным колебаниям аденинового кольца, по-разному изменяется с увеличением температуры раствора. По мнению Scovell [38], комментирующего эти

результаты, различные профили плавления поли(А) при разных частотах спектра КР должны отражать разные конформационные характеристики полимера. Он обратил внимание на необычную температурную зависимость полосы при 725 см^{-1} , поскольку кооперативный характер этой кривой, выходящей на плато при $40\text{-}50\text{ }^\circ\text{C}$, противоречит литературным данным, согласно которым разрушение одноцепочечной спирали поли(А) при нагреве не является кооперативным процессом (см. например [254]). В связи с вышеизложенным, можно сделать вывод о том, что температурная зависимость интенсивности рамановской полосы при 725 см^{-1} [253] отражает «плавление» вертикальных водородных связей в минорной конформации поли(А), которые должны разрушаться кооперативно. Следует отметить, что в рамановских спектрах поли(С) и поли(У) отсутствует полоса на частоте вблизи 725 см^{-1} [255]. Таким образом, ряд экспериментальных фактов, приведенных в литературе, свидетельствует в пользу представленной нами гипотезы о стабилизации олиго- и гомополимеров, содержащих адениновые основания, вертикальными водородными связями $\dots\text{N6H}\dots\text{N6H}\dots\text{N6H}\dots$ между экзоциклическими аминогруппами.

4.2.5. Специфические особенности структурных переходов в поли(А).

Проведенные в настоящей работе исследования структурных переходов в поли(А) с помощью новых методов (электрофореза и протонной буферной емкости) позволили уточнить существующие в литературе представления об этих переходах, получить новые данные о структурных особенностях разных форм поли(А) и выдвинуть ряд гипотез о механизмах их образования.

Предложена и обоснована гипотеза о том, что существование двух различных форм одноцепочечного поли(А) и двух различных двухцепочечных форм этого полимера связано с наличием или отсутствием вертикальных водородных связей $\dots\text{N6H}\dots\text{N6H}\dots\text{N6H}\dots$, образующихся между атомами водорода и азота экзоциклических аминогрупп соседних оснований. Таким образом, в обеих двухспиральных формах поли(А) – (DS и DS_P), а не только в DS_P -форме, каждое основание участвует в образовании трех водородных

связей. Один из атомов водорода аминогруппы N6H2 участвует в образовании водородной связи с атомом N7 противоположной цепи в обеих формах поли(А), а второй атом образует водородную связь либо с фосфатом противоположной цепи (DS_P -форма), либо вертикальную водородную связь с азотом экзоциклической аминогруппы соседнего основания в этой же цепи, усиливая взаимодействия между парами адениновых оснований в двойной спирали (DS -форма).

Следует подчеркнуть, что структурная модель DS -формы поли(А) в литературе до сих пор отсутствует. Потому любая новая информация об этой форме является важной. В работе экспериментально определены следующие новые характеристики DS -формы поли(А):

- эта форма образуется с высокой степенью кооперативности при разных ионных силах и температурах;
- она частично диссоциирует в присутствии дикатиона эфира аргинина;
- при увеличении степени протонирования DS -форма претерпевает конформационный переход;
- она имеет меньшую электрофоретическую подвижность, чем DS_P -форма.

Три первые характеристики DS -формы могут быть объяснены на основе нашей гипотезы о присутствии в ней вертикальных водородных связей. Вертикальные связи образуются кооперативно. Избыточный положительный заряд в макромолекуле поли(А) в присутствии дикатиона аргинина или при высокой степени протонирования гомополимера уменьшает степень структурной жесткости аминного фрагмента молекулы аденина, способствуя соответственно частичной диссоциации этой формы или конформационному переходу.

Анализ кривых буферной емкости поли(А) показал, что кооперативный переход из одноцепочечной спирали в двухцепочечное состояние при подкислении раствора полимера сопровождается не только кооперативным образованием DS -формы, но и кооперативным образованием DS_P -формы.

Параметр кооперативности образования DS -формы в любых условиях (т.е. при разных ионных силах и температурах) значительно выше, чем DS_P -формы. Последовательное кооперативное образование сначала DS -, а затем DS_P -формы наблюдается только в определенных условиях: при низкой температуре и высокой ионной силе (рис. 4.26 б). При $T = 25-26$ °C и $I = 50-100$ мМ обе формы образуются практически одновременно (с разной степенью кооперативности), а при низкой ионной силе появление DS_P -формы даже предшествует образованию DS -формы (Рис. 4.25 а).

О кооперативном образовании обеих кислых форм поли(А) при титровании изоионных растворов сообщалось в работе Кузнецова и Воронцовой [39], но авторы считают, что эти формы образуются последовательно как в бессолевых условиях, так и в присутствии соли. Однако в этой работе проведен анализ недифференцированных кривых потенциометрического титрования, а такой анализ является менее точным и достоверным, чем анализ кривых титрования методом буферной емкости. Таким образом, DS_P -форма поли(А) образуется в два этапа: на первом этапе кооперативно, а затем ее количество постепенно увеличивается за счет диссоциации DS -формы. При подкислении растворов поли(А) DS -форма присутствует наряду с DS_P -формой в широком интервале рН.

«Замороженная» форма поли(А), как это следует из электрофоретических экспериментов, действительно представляет собой объемистые структуры, существование которых предположил Vobst [41], исследуя поли(А) методом электронного парамагнитного резонанса. Эта форма появляется одновременно с образованием двухцепочечных форм и требуется принимать специальные меры для того, чтобы значительно уменьшить ее количество. По-видимому, полностью предотвратить образование «замороженной» формы невозможно, и в этом смысле существование определенного количества одноцепочечной формы в кислых растворах поли(А) (в виде «замороженной» формы), действительно, можно считать свойством самой системы, как это предполагают Кузнецов и соавт.

[39]. Количество «замороженной» формы увеличивается с увеличением ионной силы раствора и с увеличением длины макромолекул поли(А).

Мы предположили, что подобно тому, как участок РНК, содержащий кластеры G-трактов, может сворачиваться с образованием внутримолекулярного квадруплекса, так и поли(А) может сворачиваться с образованием внутримолекулярной двойной спирали. Одновременное образование при протонировании аденинов как межмолекулярных, так и внутримолекулярных дуплексов (самоассоциатов) приводит, по нашему мнению, к появлению «замороженной» формы поли(А). Ее образованию способствуют условия, облегчающие сворачивание макромолекулы: повышение ионной силы, уменьшающее электростатическое расталкивание фосфатных групп, и увеличение длины полимера.

На рис. 4.29 показана обобщающая схема структурных переходов в поли(А) при протонировании адениновых остатков (при средних и высоких ионных силах).

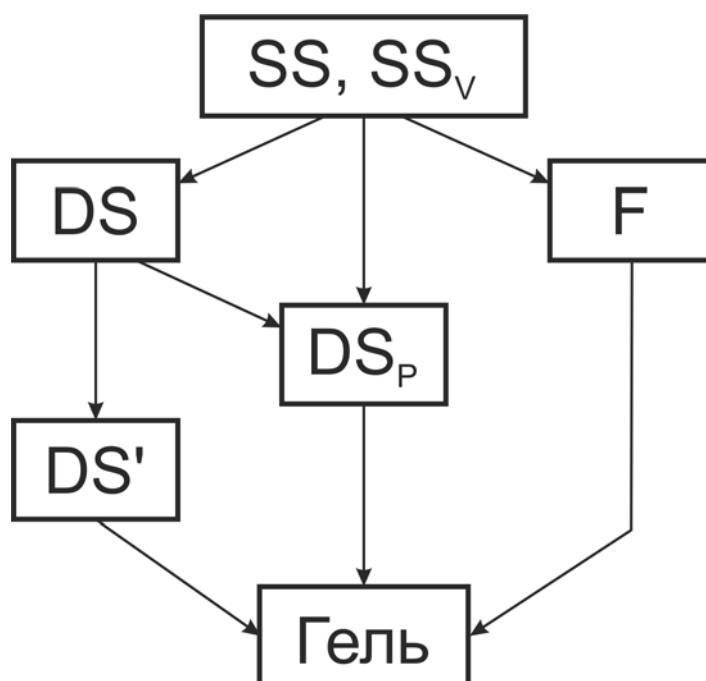


Рис. 4.29. Схема конформационных переходов в поли(А) при его протонировании ($I = 50-100$ мМ NaCl). Пояснения в тексте

Одноцепочечная конформация существует в двух формах – SS (single stranded) и SS_V ; символ V (vertical) относится к форме, в которой основания участвуют в образовании вертикальных водородных связей. При подкислении растворов практически одновременно образуются две различные двухцепочечные формы – DS и DS_P и «замороженная» форма F (frozen). Формы DS и DS_P образуются кооперативно, но с разной степенью кооперативности. При дальнейшем подкислении DS -форма постепенно переходит в DS_P -форму, а DS -форма при $pH \sim 5,0$ кооперативно переходит в DS' -форму. В условиях наших экспериментов (2-3 мМ поли(А) при pH , близких к 4,1, растворы поли(А) превращаются в прозрачные гели.

Следует отметить, что полиадениловая кислота, по-видимому, может образовывать наибольшее количество разных форм по сравнению со всеми остальными гомополимерами.

4.3. Структурные переходы в поли(С) и поли(dС)

4.3.1. Зависимость протонной буферной емкости поли(С) от pH .

Насколько нам известно, в литературе приведено лишь несколько кривых потенциометрического титрования поли(С), в частности, в работах [45, 47, 51, 52]. Три из них (при 0,1 М Na^+ или K^+) были использованы нами для вычисления зависимости протонной буферной емкости поли(С) от pH при комнатной температуре. Кривые буферной емкости, рассчитанные по данным работ Hartman & Rich [45], а также Klump [47], практически совпадают. Зависимость, построенная по данным работы Аполонника и соавт. [52], имеет такой же характер, как и две предыдущие, но значительно (на $\sim 0,9$ единиц pH) смещена в щелочную область. Поскольку, как будет показано ниже, значение pK перехода поли(С) из одноцепочечной формы в двухцепочечную, рассчитанное по данным этой работы, не согласуется с величинами, полученными в работах других авторов, мы приводим здесь только зависимости, построенные по данным работ [47] (рис. 4.30) и [45] (рис. 4.31).

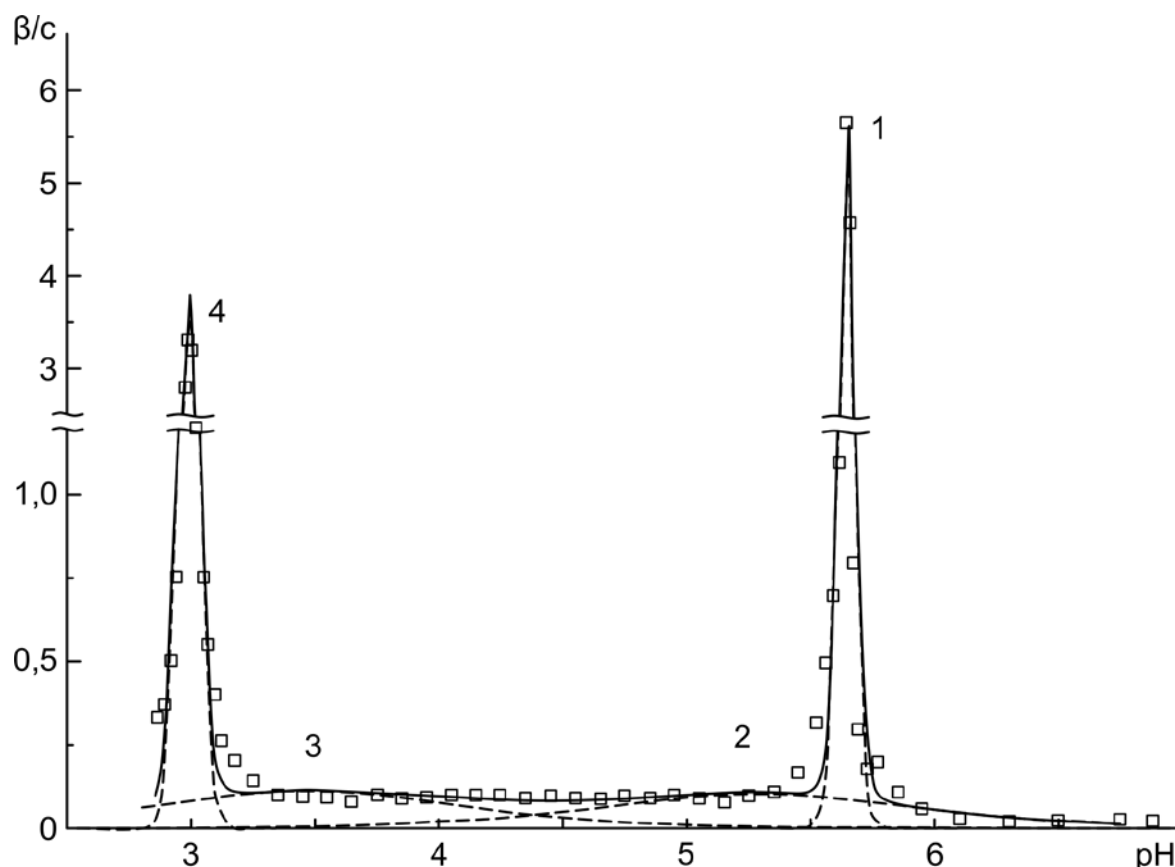


Рис. 4.30. Кривая протонной буферной емкости поли(С), рассчитанная по данным работы [47], и ее разложение на составляющие. Пунктирные линии – базисные функции, сплошная линия – суммарная кривая. $T = 25\text{ }^{\circ}\text{C}$; $0,1\text{ M NaCl}$; $0,5\text{-}3,0\text{ mM}$ поли(С); c – общая концентрация поли(С)

Как видно на рисунках 4.30 и 4.31 кривая буферной емкости поли(С) имеет два узких пика, соответствующих двум участкам с высокой буферной емкостью, наблюдаемых на кривых потенциометрического титрования [45, 47, 51, 52], и может быть удовлетворительно представлена в виде суперпозиции четырех базисных функций вида (2.8). Поскольку значения параметров пиков рассчитывали методом подбора (параграф 2.3), определенные трудности возникали при подборе значений γ_i для очень узких пиков, подобных пикам 1 и 4 на рисунках 4.30 и 4.31, а также при подборе значений pK_i для невысоких широких пиков.

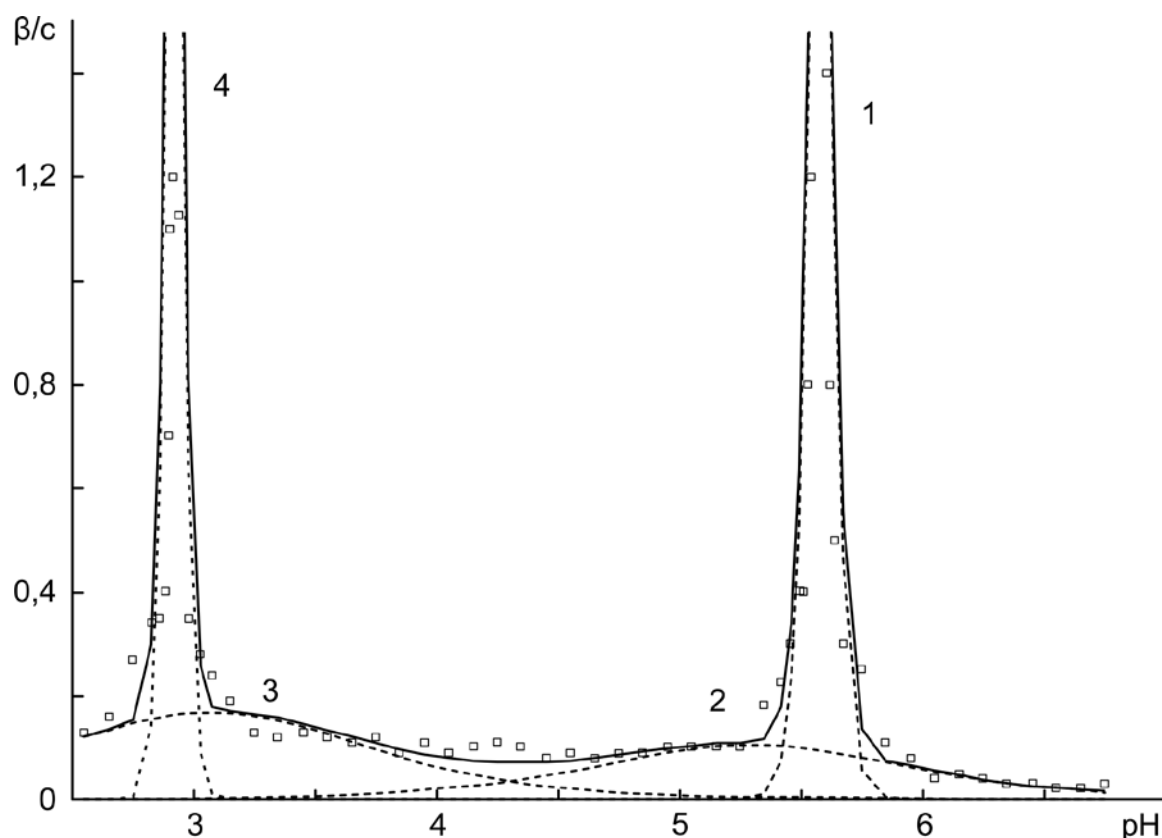


Рис. 4.31. Кривая протонной буферной емкости поли(С), рассчитанная по данным работы [45], и ее разложение на составляющие. Пунктирные линии – базисные функции, сплошная линия – суммарная кривая. $T = 25\text{ }^{\circ}\text{C}$; $0,1\text{ M NaCl}$; $3,6\text{ mM}$ поли(С); c – общая концентрация поли(С).

Учитывая, что, согласно теории [202], узкие ($\gamma_i > 1$) пики на кривых рН-зависимости буферной емкости биополимеров соответствуют конформационным переходам, мы относим узкий пик 1 на кривых буферной емкости поли(С) к кооперативному образованию двойных спиралей из одноцепочечных структур. Таким образом, рК первого конформационного перехода поли(С) (pK_1) равен 5,6 в $0,1\text{ M NaCl}$ (рисунки 4.30 и 4.31). Именно в интервале рН, соответствующем пику 1, регистрируется конформационный переход поли(С) в двухспиральное состояние методом УФ-спектроскопии [45]. Параметры всех пиков приведены в таблице 4.2.

Таблица 4.2.

Параметры пиков разложения кривых буферной емкости поли(С) по базисным функциям β_i

Номер пика, i	pK_i		γ_i		c_i/c	
	[45]	[47]	[45]	[47]	[45]	[47]
1	$5,58 \pm 0,01$	$5,64 \pm 0,01$	13 ± 2	30 ± 5	$0,30 \pm 0,1$	$0,34 \pm 0,01$
2	$5,30 \pm 0,05$	$5,30 \pm 0,1$	1	1	$0,18 \pm 0,02$	$0,18 \pm 0,1$
3	$3,07 \pm 0,05$	$3,5 \pm 0,1$	1	1	$0,29 \pm 0,02$	$0,18 \pm 0,1$
4	$2,92 \pm 0,01$	$2,99 \pm 0,01$	20 ± 5	20 ± 5	$0,25 \pm 0,05$	$0,31 \pm 0,01$

Примечание. Определение параметров дано в параграфе 2.3; c – общая концентрация поли(С).

В первом конформационном переходе, как следует из этой таблицы, кооперативно протонируется треть цитозиновых остатков гомополимера ($c_1/c = 0,30-0,34$). Все известные нам из литературы значения pK образования двойных спиралей поли(С) согласуются между собой с учетом условий эксперимента. Они располагаются на одной и той же кривой зависимости pK от ионной силы при комнатной температуре (рис. 4. 32) [43, 45, 47, 256-259]. Значение pK резко возрастает с уменьшением ионной силы и достигает величины 6,0-6,3 при концентрации NaCl, равной 10 мМ [43, 257]. В 0,1 М NaCl оно равно 5,7 [45, 47, 258].

В то же время значение pK_1 в 0,1 М KCl, рассчитанное нами по данным работы Аполлоника и соавт. [52], значительно завышено и равно 6,5. Кривая буферной емкости поли(С), построенная по данным этой работы, хорошо раскладывается на два узких и два широких пика, и расстояние по шкале рН между ее двумя узкими пиками примерно соответствует расстоянию между пиками 1 и 4 на рисунках 4.30 и 4.31. Однако пики 1 и 4 шире, а пики 2 и 3 выше, чем соответствующие пики на рисунках 4.30 и 4.31.

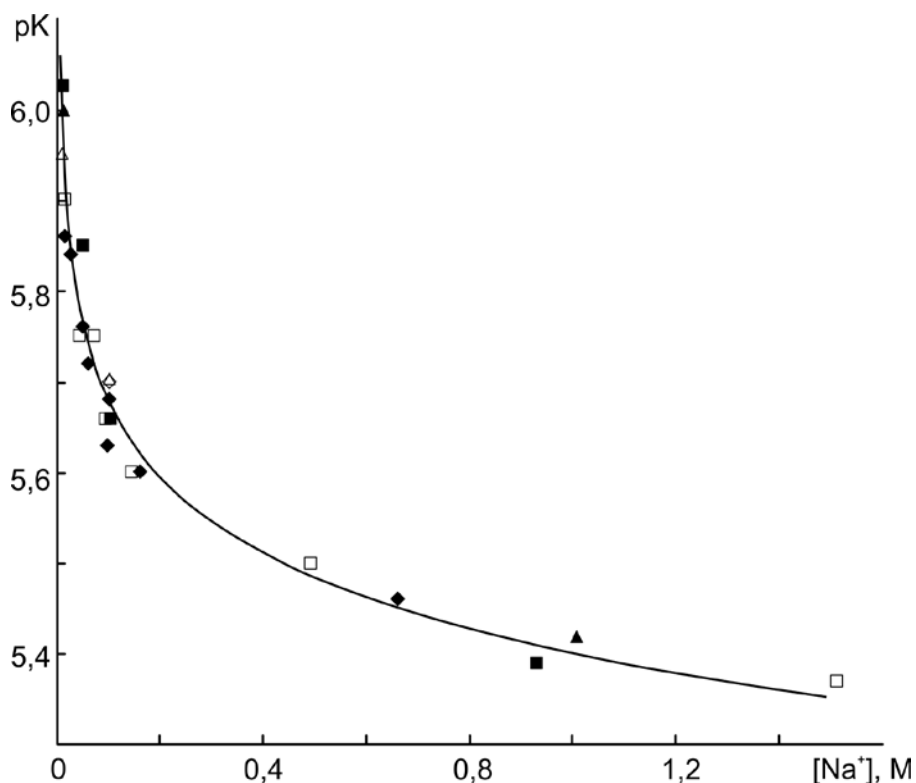


Рис. 4.32. Зависимость рК первого конформационного перехода в поли(С) от ионной силы, построенная по литературным данным: □ – [256], T = 25 °С; ■ – [257]; △ – [43], T = 20 °С; ◇ – [45, 47], T = 25 °С; ◆ – [258], T = 20 °С; ▲ – [259], T = 20 °С

Пик 2 на кривой протонной буферной емкости поли(С) мы интерпретируем как процесс некооперативного присоединения протонов к двойным спиральям поли(С) по положению атома азота N3 непротонированных нуклеотидных оснований, еще не вовлеченных в спаривание по схеме, представленной на рис. 1.2, и частично экспонированных в раствор. Протонирование поли(С) в области рН, предшествующей структурному переходу из одноцепочечной формы в двухцепочечную, можно объяснить, существованием в этой области определенного количества непротонированных двойных спиралей, стабилизированных высокой концентрацией соли и/или высокой концентрацией полимера ([52] и ссылки в этой работе). Процесс некооперативного протонирования происходит также

наряду с образованием двойных спиралей и продолжается после их формирования, поскольку при завершении конформационного перехода (рН 5,4) степень протонирования поли(С) составляет не 50 %, а примерно 40 %.

Широкий пик 3 на кривой протонной буферной емкости поли(С) соответствует процессу некооперативного присоединения протонов к двойной спирали поли(С) иной природы, чем в случае пика 2. Этот процесс, по-видимому, и подготавливает второй конформационный переход полимера (узкий пик 4). Согласно спектроскопическим и вискозиметрическим данным, в интервале рН между двумя структурными переходами полимер остается в двухспиральной форме, и его конформация не претерпевает сколько-нибудь заметных структурных изменений [44, 46]. При рН=4,6 (рис. 4.30 и 4.31) степень протонирования поли(С) достигает 50 %. Это так называемая полупротонированная двойная спираль, обладающая максимальной термодинамической стабильностью и имеющая максимальную T_m [47].

Мы, как и Klump [47], полагаем, что присоединение протонов к двойной полупротонированной спирали поли(С) изнутри сильно затруднено. Это связано с тем, что ориентированные навстречу друг другу неподделенные электронные пары атомов азота N3 в паре Cyt:Cyt(N3H)⁺ образуют протонную ловушку, вследствие чего протопфильность атома N3 спаренного непротонированного Cyt оказывается намного ниже, чем свободного, где он доступен протонированию. Кроме того, по квантовохимическим расчетам [260], энергия комплексообразования пары Cyt:Cyt(N3H)⁺ в изолированном состоянии (41,7 ккал/моль) в 1,7 раза превышает таковую для Уотсон-Криковской пары Gua:Cyt (23,7 ккал/моль) с подобными Н-связями. Это свидетельствует о том, что процесс полного раскрытия пары Cyt:Cyt(N3H)⁺ при комнатной температуре является крайне маловероятным событием с термодинамической точки зрения. Кстати, согласно квантовохимическим расчетам [261], основной вклад в энергию комплексообразования изолированной пары Cyt:Cyt(N3H)⁺ вносит водородная связь N3H...N3 и при удалении протона из протонной ловушки пара диссоциирует.

Согласно квантовохимическим расчетам [262, 263], наиболее вероятными местами протонирования изолированного Cyt считают атомы N3 и O2, причем атом кислорода – слегка более предпочтительным [263]. При переходе в водный раствор протофильность атома N3 Cyt несколько превышает протофильность атома O2 [263]. С учетом этого мы относим пик 3 на кривой буферной емкости поли(С) к протонированию двойной спирали по положению атома кислорода O2 непротонированных цитозиновых остатков с *цис*-ориентацией гидроксила относительно гликозидной связи.

В пользу такой интерпретации свидетельствует также появление полосы при 1440 см^{-1} , соответствующей, по нашему мнению, валентному колебанию протонированной карбонильной группы в ИК-спектрах поли(С) при низких рН [45, 48] (значения частот определены из спектрограммы на рис. 2 в работе [48]). Такое отнесение согласуется с расчетом колебательного спектра протонированного по месту O2 изолированного цитозина, а также с частотой аналогичного колебания, наблюдаемого в Рамановском спектре водного раствора протонированного по атому O2 Cyt и его кристалла на частотах 1418 и 1408 см^{-1} соответственно [263].

Реакция протонирования двойной спирали поли(С) по атому O2 электронейтральных цитозиновых остатков с неизбежностью приводит к ее распаду при высокой степени протонирования ($\sim 70\%$). Во-первых, протонирование атома кислорода O2 электронейтрального Cyt если не разрушает, то сильно дестабилизирует пару $\text{Cyt}(\text{O2H cis})^+:\text{Cyt}(\text{N3H})^+$. Вполне уместна здесь аналогия с Уотсон-Криковской парой Gua:Cyt, протонирование которой по этому же атому O2 (с *цис*-ориентированным гидроксилом относительно гликозидной связи N1H), согласно квантовохимическим расчетам [264], приводит к ее преддиссоциации. Кроме того, протонирование атомов O2 возмущает равновесную конформацию сахарофосфатного остова из-за сильного кулоновского отталкивания присоединенных водородов с близлежащими атомами рибозы H1'.

Во-вторых, вышеупомянутое протонирование дестабилизирует также стэкинг-взаимодействия вследствие сильного электростатического расталкивания протонов, присоединенных к атомам N3 и O2 соседних по цепи цитозиновых остатков. В полупротонированной двойной спирали поли(С) протоны при атомах азота N3 располагаются, очевидно, вдоль каждой цепи через основание, т. е. зигзагообразно, что позволяет минимизировать их кулоновское расталкивание.

Предполагается, что разрыв всех трех межмолекулярных Н-связей при протонировании полупротонированной пары $\text{Cyt}:\text{Cyt}(\text{N3H})^+$ завершается переносом протона, который разрушает пару, на соседний атом N3 Cyt^+ [265]. Перенос протона происходит в две стадии (рис. 4.33).

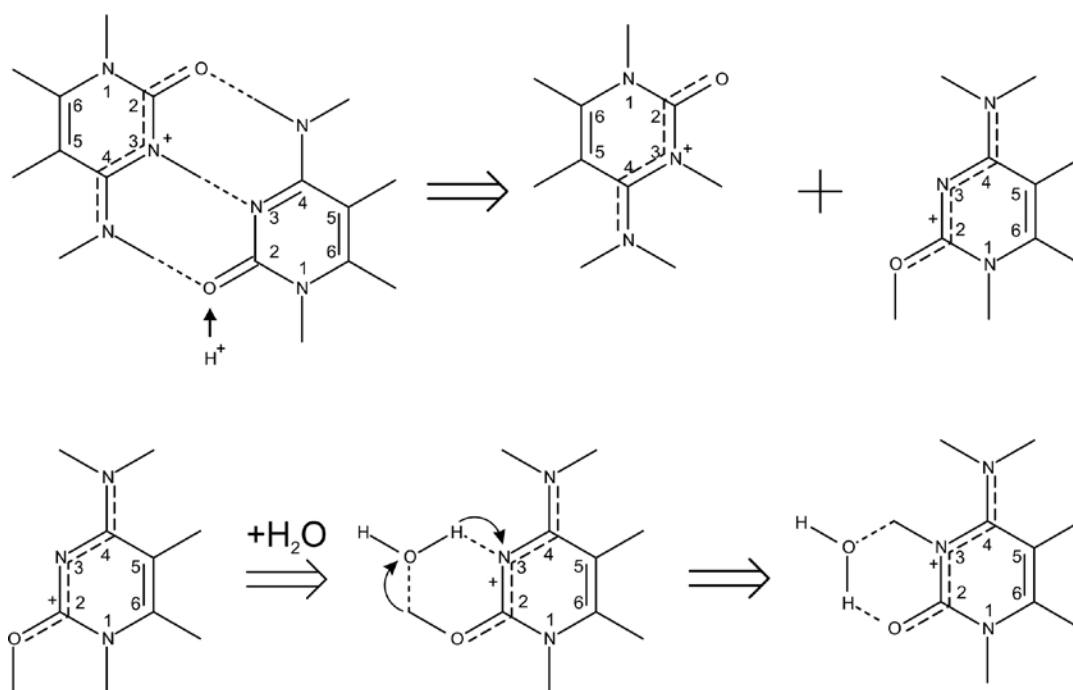


Рис. 4.33. Молекулярный механизм структурного перехода двухцепочечных поли(С) и поли(dС) в одноцепочечное состояние. Показана только одна молекула воды, которая контролирует процесс внутримолекулярного переноса протона, ранее присоединенного к атому O2 в *цис*-положении по отношению к соседней связи N1H, на атом N3 цитозинового остатка

Сначала происходит поворот гидроксильной группы O2H вокруг связи C2-O2 из *цис*- в *транс*-положение, а затем перенос ее протона с помощью молекулы воды на атом N3. Очевидно, что заключительная стадия процесса является лимитирующей с энергетической и кинетической точки зрения, поскольку барьеры внутримолекулярного переноса протона в нуклеотидных основаниях, которые ускоряются с помощью молекулы воды, образующей для этого подходящий мостик Н-связей [266], значительно больше, чем барьер поворота гидроксила в *транс*-положение.

Что же касается точки зрения Klump [47] на процесс протонирования поли(С) (см. параграф 1.1.2.), то она не имеет достаточного квантовохимического обоснования. Следует отметить, что относительная концентрация редкой иминной формы Cyt в водном растворе оценивается как 10^{-5} [267], а энергия комплексообразования пары Cyt:Cyt (imino) [260] – на 22 ккал/моль ниже, чем аналогичная величина для пары Cyt:Cyt(N3H)⁺. Следовательно, «сценарий» образования двойной спирали поли(С) через образование пары Cyt:Cyt (imino) с последующим ее протонированием по аминному атому азота N4 редкой таутомерной формы Cyt является нереальным.

Узкий пик 4 на кривой протонной буферной емкости поли(С) соответствует кооперативному протонированию одноцепочечных участков, образующихся при распаде двойных цепей поли(С). Кооперативное присоединение протонов обусловлено кооперативным процессом распада двойных спиралей. Считается (см. краткий обзор литературы, приведенный в работе [51]), что нуклеотидные остатки в составе одноцепочечных полинуклеотидов протонируются примерно в том же диапазоне рН, что и свободные мономеры. Таким образом, поскольку рК ионизации цитидин-5'-фосфата (CMP) в 0,1 М NaCl при 20 °С равен 4,42 [268], то цитозиновые остатки в составе одноцепочечных макромолекул в интервале рН 2,7-3,1, в котором происходит разрушение двойной спирали, должны быть практически полностью протонированы. Поэтому одноцепочечные структуры,

образующиеся при разрушении двойной спирали поли(С), протонируются (по положению атома азота N3) сразу же после своего появления, в том числе и вследствие перехода протона при атоме O2 на соседний атом N3.

Однако не все авторы относят участок протонирования поли(С), соответствующий пику 4, к распаду двойной спирали. Сулейманова, Аполонник и Кузнецов [51, 52] предлагают схему протонирования поли(С), отличную от таковых в работах [45, 47]. Они считают, что в 0,1 М КС1 при степени протонирования поли(С) $\sim 50\%$ сосуществуют полностью протонированные двойные спирали, куда входит 50% всех оснований, и одноцепочечные непротонированные структуры, которые также содержат 50% оснований. Эти одноцепочечные участки полностью протонируются при дальнейшем понижении рН, а двойные спирали при этом остаются стабильными.

Такое представление, на наш взгляд, является ошибочным. Оно не имеет достаточных физико-химических предпосылок и не подтверждается другими экспериментальными данными. Так, анализ спектров КД [43] и ОРД [269] поли(С) при степени ее протонирования в интервале 40-70% свидетельствует об отсутствии сколько-нибудь заметного количества одноцепочечных участков в ее структуре. С другой стороны, при рН 3 спектр КД поли(С) значительно отличается от спектра двухспиральной формы и соответствует неструктурированному состоянию полимера [43]. Кроме того, схема протонирования поли(С), предлагаемая авторами работ [51, 52], не может объяснить резкого уменьшения температуры плавления двухспиральных участков полимера в интервале рН, соответствующем $\alpha = 50-70\%$ [44-47].

Klump [47] рассматривает процесс протонирования поли(С) как четыре последовательные стадии. Третьей стадией, по его мнению, является реакция протонирования двойных спиралей поли(С), а четвертой – их кооперативный распад. Такая точка зрения согласуется с нашими представлениями. Однако мы считаем некорректным его мнение о начальных стадиях процесса, согласно

которому сначала протонируются одноцепочечные структуры до $\alpha \sim 40\%$, а затем протонированные одноцепочечные поли(С) превращаются в двойные спирали. Мы полагаем, что значения рК протонирования одноцепочечных участков поли(С) близки к значениям рК ионизации соответствующего мономера, и если такие участки в полимере присутствуют, то они должны протонироваться примерно в интервале рН 3,4-5,4 (при комнатной температуре в 0,1 М КС1) [45, 268]), т. е. уже после того, как образовались двойные спирали. Но в этом интервале рН, как уже упоминалось выше, одноцепочечные участки поли(С) не регистрируются. Однако они присутствуют в интервале рН между двумя узкими пиками на кривой протонной буферной емкости поли(С) при высоких температурах, когда происходит плавление двойных спиралей.

Результаты спектроскопических исследований поли(С) при высоких температурах [46, 270] достаточно убедительно свидетельствуют, по нашему мнению, о близости значений рК протонирования мономеров и одноцепочечных участков. В этих работах показано, что плавление двойных спиралей поли(С), степень протонирования которых не превышает 50%, сопровождается диссоциацией протонов, а при более высоком уровне ионизации высвобождается значительно меньшее количество протонов. Эти результаты легко объяснить, учитывая значения степени протонирования СМР в зависимости от рН и температуры. На рис. 4.34 показана кривая буферной емкости этого нуклеотида в 0,1 М NaCl при температуре 83 °С, соответствующей завершению процесса плавления двухспирального поли(С) при рН = 4,0 [46], т. е. при $\alpha > 50\%$ (рК СМР в этих условиях равен 3,88 [268]). При рН 5,0 ($\alpha < 50\%$) поли(С) практически полностью расплавлен при 70 °С [46], рК ионизации СМР при этой же температуре равен 3,94 [268]. Кривая буферной емкости мономера при этом значении рК будет иметь точно такую же форму, как и при рК, равном 3,88, но со сдвигом по шкале рН на 0,06 единицы в щелочную область.

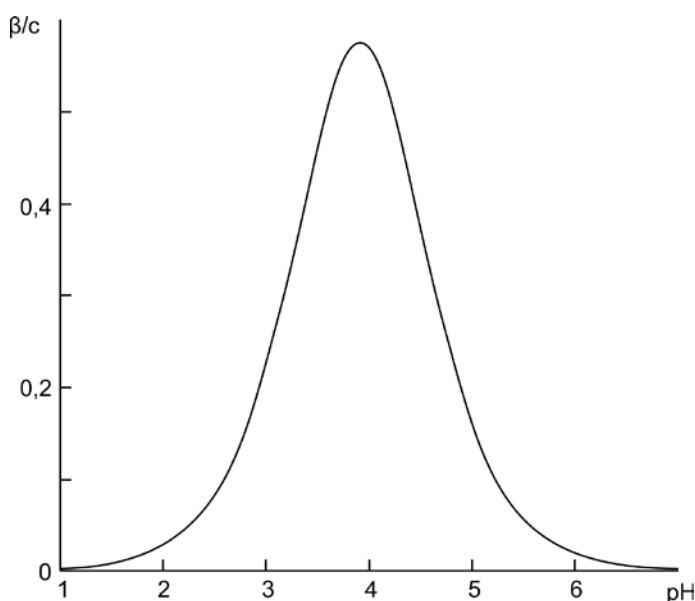


Рис. 4.34. Расчетная рН-зависимость протонной буферной емкости СМР ($pK = 3,88$; $\gamma = 1$); c – общая концентрация поли(С)

На рис. 4.34 видно, что количество протонированных молекул СМР при $pH = 4,0$ (43 %) значительно больше, чем при $pH = 5,0$ (8 %), поскольку степень протонирования исследуемого вещества прямо пропорциональна площади под его кривой протонной буферной емкости [202]. Для поли(С) при этих же значениях pH , $\alpha = 56$ и 46 % соответственно [45]. Следовательно, в том случае, когда одноцепочечные участки поли(С) протонируются при том же значении pK , что и мономер, они оказываются при распаде двойных спиралей протонированными на 13 % больше при $pH = 4,0$ и на 38 % больше при $pH = 5,0$ по сравнению со степенью ионизации в этих условиях. Избыточные протоны диссоциируют, причем тем в большей степени, чем выше pH , и этот процесс, по нашему мнению, регистрируют методами УФ и Рамановской спектроскопии [46, 270].

Следует отметить, что Klump [47] на основании данных калориметрических и спектроскопических исследований делает вывод о том, что при плавлении двойной спирали поли(С) протоны в интервале pH , соответствующем степени протонирования полимера 40-70 %, не высвобождаются, однако его аргументы при этом являются косвенными, а не прямыми, как в работах [46, 270].

В заключение кратко охарактеризуем поэтапные реакции протонирования полицитидиловой кислоты, влекущие за собой ее структурные перестройки:

- кооперативное образование двойных спиралей из одноцепочечных структур, индуцируемое протонированием атомов азота N3 оснований;
- некооперативное протонирование двойных спиралей по положению атомов азота N3 оснований, не вовлеченных в протонные ловушки;
- внешнее некооперативное протонирование двойных спиралей по положению атома кислорода O2 электронейтральных цитозинового остатков, дестабилизирующее двойную спираль;
- кооперативная диссоциация двойных спиралей, сопровождающаяся полным протонированием образовавшихся одноцепочечных структур по положению атома азота N3, в том числе и вследствие перехода протонов с атома O2 на соседний атом N3 цитозинового остатков.

4.3.2. Зависимость протонной буферной емкости поли(dC) от pH.

Кривая протонной буферной емкости поли(dC), рассчитанная по данным Inman [53], приведена на рис. 4.35. Она удовлетворительно описывается суперпозицией трех базисных функций, параметры которых приведены в таблице 4.3. Первые два пика, по аналогии с поли(C), мы относим соответственно к кооперативному образованию двойных спиралей и к их некооперативному протонированию по положению атома азота N3 цитозинового остатков, не вовлеченных в протонные ловушки. Третий пик соответствует процессу протонирования полимера, характеризуемому параметром кооперативности $\gamma < 1$.

Как видно из таблицы 4.3, переход поли(dC) из одноцепочечной структуры в двухцепочечную, так же как и переход поли(C), сопровождается протонированием примерно трети оснований. Он менее кооперативный, но этот факт можно объяснить меньшей концентрацией соли, при которой проводилось титрование поли(dC) (0,05 M), так как кооперативность перехода уменьшается с уменьшением ионной силы раствора [259].

Потенциометрическая кривая поли(dC) в работе Inman [53], по данным которой рассчитывалась кривая буферной емкости, получена в области pH выше 3,1 ($\alpha = 76\%$). Конформационный переход, соответствующий распаду двойной спирали, в этой области не зарегистрирован.

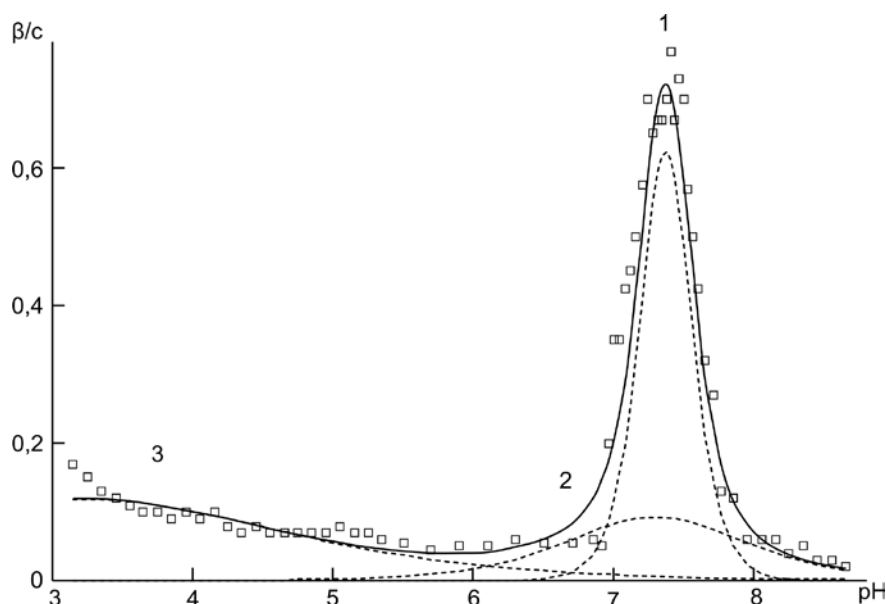


Рис. 4.35. Кривая протонной буферной емкости поли(dC), рассчитанная по данным работы [53], и ее разложение на составляющие. Пунктирные линии – базисные функции, сплошная линия – суммарная кривая. $T = 25\text{ }^\circ\text{C}$; $0,05\text{ M NaCl}$; c – общая концентрация поли(C)

Таблица 4.3.

Параметры пиков разложения кривой буферной емкости поли(dC) по базисным функциям β_i

Номер пика, i	pK_i	γ_i	c_i/c_0
1	$7,37 \pm 0,02$	$3,5 \pm 0,2$	$0,310 \pm 0,03$
2	$7,30 \pm 0,15$	1	$0,16 \pm 0,03$
3	$3,27 \pm 0,1$	$0,45 \pm 0,01$	$0,46 \pm 0,02$

Примечание. Определение параметров дано в параграфе 2.3; c – общая концентрация поли(C).

До сих пор мы в соответствии с выводами авторов работ [45, 47, 53], по данным которых были рассчитаны кривые буферной емкости поли(C) и поли(dC), соотносили первый конформационный переход в этих полинуклеотидах с образованием двойной спирали. Однако учитывая результаты более поздних работ, опубликованных после открытия i-мотива ДНК, логично предположить, что при протонировании поли(dC) образуется не двойная спираль, а i-мотив [54], поскольку именно эту структуру образуют полупротонированные dC₈ [55], dC₁₂ [56] и, по всей вероятности, dC₁₈ [271]. В работе [55] было показано, что i-мотив, образованный олигомером d(C₈) остается стабильным даже при pH 2,4. В связи с этим можно предположить, что и поли(dC) будет стабилен в этой области, т. е. интервал существования недиссоциированной протонированной структуры поли(dC), по сравнению с таковым для поли(C), может быть расширен не только в сторону щелочных, но и кислых pH.

Что касается поли(C), то в настоящее время в литературе нет прямых доказательств образования i-мотива при протонировании этого полинуклеотида. Cohen et al. [271], основываясь на результатах исследования динамики возбужденных состояний полупротонированных dC₁₈ и поли(C), а также их УФ- и КД-спектров, предположили, что эти полинуклеотиды могут иметь структуру, идентичную i-мотиву ДНК. Tsankov et al. [56], исследуя ИК-спектры и спектры колебательного кругового дихроизма (ВКД) dC₁₂ при разных pH, установили, что полоса ИК-поглощения 1694 см⁻¹ и соответствующий ей отрицательный ВКД-куплет являются показателем i-мотива. Petrovic & Polavarapu [272], исследовавших этим же методом структурные переходы в поли(C), выявили, что хотя pH-зависимое поведение полосы 1693 см⁻¹ подобно для dC₁₂ и поли(C), однако в поведении других pH-зависимых полос имеется ряд различий. По этой причине авторы работы [272], не исключая возможности образования i-мотива в поли(C), все-таки считают, что pH-зависимые структурные изменения в этом полинуклеотиде связаны с его переходом из односпиральной в двухспиральную форму.

Таким образом, в результате структурного перехода поли(С) из одноцепочечного состояния в двухцепочечное могут образовываться двойные спирали, *i*-мотив или смесь этих форм. Специфическая форма кривой буферной емкости поли(С) – очень узкие пики, соответствующие конформационным переходам, и низкие широкие пики, относящиеся к некооперативным процессам протонирования, – не позволяют однозначно сделать вывод о наличии одной или двух кислых форм. Причина, по-видимому, в том, что протонирование ведет к образованию тех же самых $Cyt \cdot Cyt^+$ пар оснований как в случае двойной спирали поли(С), так и в случае *i*-мотива и механизм распада этих форм тоже одинаков.

4.3.3. Обнаружение методом электрофореза двух различных форм протонированного поли(С).

4.3.3.1. Электрофорез поли(С) в трис-ацетатном буфере. В наших ранних работах при исследовании методом электрофореза гомополирибонуклеотидов при кислых рН мы обнаружили, что поли(С), подобно поли(А), мигрирует двумя зонами, одну из которых можно было бы отнести к двуспиральной форме полимера, но природа второй зоны была для нас неясна, и эти результаты не были опубликованы. В связи с открытием *i*-мотива ДНК мы продолжили исследование кислых форм поли(С) методом электрофореза. На рис. 4.36 приведены данные электрофореза поли(С) при разных значениях рН и ионной силе 13-14 мМ, являющиеся результатом одного эксперимента, т. е. составы растворителей в трубках с агарозными гелями были различными (см. параграф 2.2). Как показано в параграфе 4.2.1., результаты такого варианта электрофореза не отличаются значительно от полученных стандартным способом.

Как следует из данных потенциометрического титрования (не показано), рК перехода поли(С) из одноцепочечной спирали в двухцепочечную при $I = 14$ мМ составляет 6,10, что согласуется с литературными данными, суммированными на рис. 4.32. На рис. 4.36 видно, что при рН выше рК перехода наблюдается одна зона поли(С).

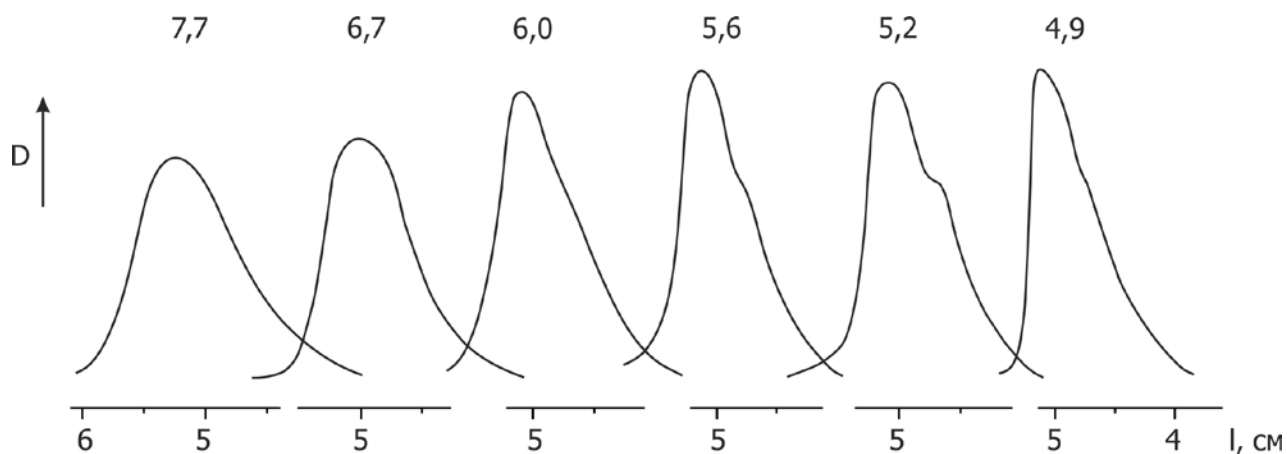


Рис. 4.36. Денситограммы поли(С) фирмы «Reanal» при разных значениях рН в трис- CH_3COOH буфере. $I = 13\text{-}14$ мМ; 0,8 %-ная агароза; $t = 85$ мин. Цифрами над зонами обозначены значения рН

При более низких рН наблюдаются две перекрывающиеся зоны, причем эти денситограммы относятся к той области рН, где переход в двойную спираль уже завершен, а распад дуплексов еще не начался. Из рис. 4.36 следует, что обе кислые формы поли(С) одновременно присутствуют в растворе при всех значениях рН ниже pK перехода, и с уменьшением величины рН их относительное количество значительно не изменяется.

Следует отметить, что в эксперименте, результаты которого представлены на рис. 4.36, на гели, приготовленные при разных значениях рН, наносили поли(С), растворенный при рН 7,6, т. е. структурные переходы полимера из нейтральной формы в кислую происходили непосредственно при вхождении образца в агарозный гель. Если же поли(С) растворяли в кислых растворах, и в них же проводили электрофорез, то относительное количество медленнее мигрирующей формы увеличивалось, как это видно из сравнения кривых *a* и *b* (рис. 4.37), где приведены результаты экспериментов, в которых на гель, приготовленный при $I = 6,7$ мМ и рН 5,5 наносили либо слабощелочной, либо кислый раствор поли(С).

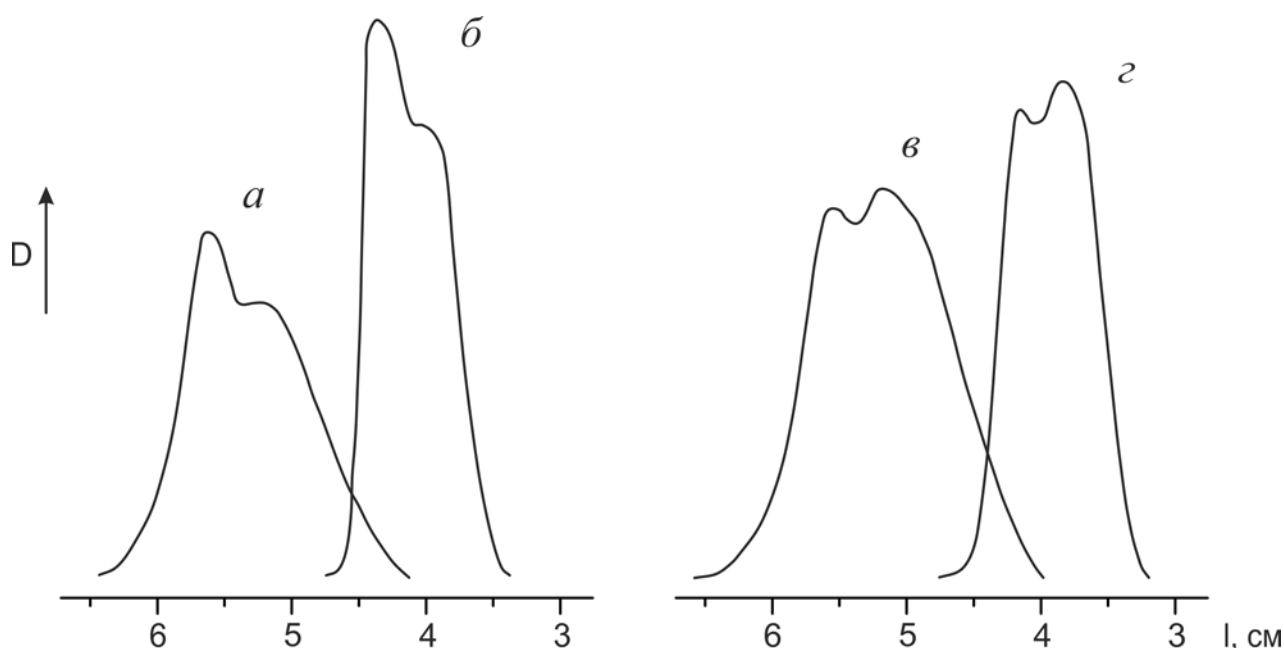


Рис. 4.37. Денситограммы поли(С) фирмы «Reanal» при разных ионных силах в трис- CH_3COOH буфере; *a, в* – $I = 6,7$ мМ; *б, г* – $I = 25,9$ мМ. 0,8 %-ная агароза; $\text{pH} \sim 5,5$; $t = 112$ мин; *a, б, г* – образец приготовлен при $\text{pH} 7,5$; *в* – образец приготовлен при $\text{pH} 5,5$; *г* – образец прогрет

Еще одним способом увеличения относительного количества медленнее мигрирующей формы в случае образца, растворенного в слабощелочной среде, является его перевод в кислый раствор и кратковременный (10 мин) прогрев при температуре $75-80$ °С непосредственно перед электрофорезом. Денситограммы непрогретого и прогретого образцов поли(С) при $I = 25,9$ мМ и $\text{pH} 5,5$ показаны как пример на рисунках 4.37 *б* и 4.37 *г* соответственно. На рисунках 4.37 *a*, 4.36 (3-я зона справа), и 4.37 *б* приведены денситограммы поли(С) при одном и том же значении $\text{pH} (\sim 5,5)$ и при разных ионных силах – 6,7 мМ, 14,0 мМ и 25,9 мМ соответственно. Две перекрывающиеся зоны наблюдаются во всех этих экспериментах.

Как следует из вышеизложенного, при pH ниже pK перехода из нейтральной формы в кислую наблюдаются две зоны поли(С). Какова же их природа? Учитывая литературные данные об *i*-мотиве РНК (см. параграфы

1.2.3. и 4.3.2.) можно предположить, что при электрофорезе кислых растворов поли(С) наблюдаются две формы полимера – двух- и четырехцепочечная. Однако прежде чем делать такое заключение, следует рассмотреть и другие возможные причины регистрации методом электрофореза двух протонированных форм поли(С). Следует отметить, что появление второй формы не является специфической особенностью какого-либо препарата поли(С), поскольку две зоны при кислых рН наблюдались с препаратами разных фирм («Reanal», «Calbiochem», «Serva» и др.).

Медленнее мигрирующая зона протонированного поли(С) могла бы относиться к «замороженной» форме полимера, подобной той, которая образуется при протонировании поли(А) (см. параграф 4.2.2.). Как указано в этом параграфе, «замороженная» форма легко образуется, когда препарат поли(А) непосредственно помещают в кислый раствор. Предотвратить ее образование можно, в частности, кратковременным прогревом кислого раствора поли(А) или нанесением слабощелочного раствора поли(А) на кислые гели. Для примера на рис. 4.38 представлены денситограммы препаратов поли(А) фирмы «Calbiochem», растворенных при рН 5,5 (рис. 4.38 а) и 7,5 (рис. 4.38 б) и нанесенных на гели, приготовленные при рН 5,5. В соответствии с результатами, изложенными в параграфе 4.2.2., образец, растворенный при кислом рН, в отличие от растворенного при рН 7,5, мигрирует во время электрофореза размытой зоной, соответствующей агрегатам самых разных размеров.

Согласно нашему предположению (параграф 4.2.2.), «замороженная» форма поли(А) возникает при его протонировании вследствие одновременного образования двухцепочечных фрагментов как между участками одной и той же молекулы (при сворачивании ее «самой на себя»), так и между этой молекулой и другими. Количество «замороженной» формы поли(А) увеличивается с увеличением гибкости макромолекул, т. е. с увеличением их длины и ионной силы растворителя.

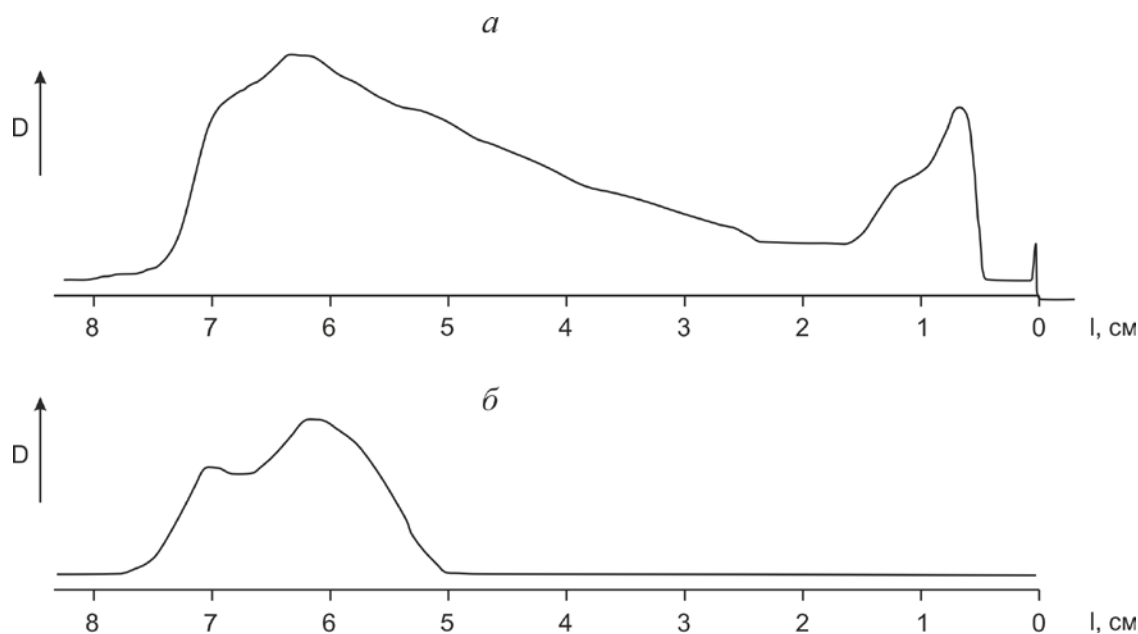


Рис. 4.38. Денситограммы поли(А)_с в трис-СН₃СООН буфере; *a* – препарат приготовлен при рН 5,5; *t* = 136 мин; *б* – препарат приготовлен при рН 7,5; *t* = 120 мин. *I* = 8,3 мМ; рН=5,5; 0,8 %-ная агароза

Учитывая механизм возникновения «замороженной» формы поли(А), образование второй формы поли(С) при протонировании полинуклеотида нельзя связать с «замороженной» формой. Во-первых, две формы поли(С) существуют в условиях, в значительной мере препятствующих образованию «замороженной» формы поли(А), т. е. при прогреве или нанесении слабощелочного образца на кислый гель. Во-вторых, в наших экспериментах мы не наблюдали при электрофорезе поли(С), растворенного в кислой среде, размытых зон, аналогичных зонам «замороженной» формы поли(А).

Важно отметить, что если в случае образцов поли(А) фирмы “Reanal”, с относительно короткими по длине макромолекулами (коэффициент седиментации 4,6-5,6 S), размытые зоны уже явно видны при *I* = 26 мМ (рис. 4.19), то в случае образцов поли(С) этой же фирмы (коэффициент седиментации 4 S) размытых зон не наблюдали и при самой высокой ионной силе (74 мМ), при которой проводили электрофорез.

Возможно, в одном из экспериментов мы наблюдали «замороженную» форму поли(С), когда препарат фирмы «Calbiochem», был приготовлен в кислом растворе при $I = 40$ мМ. Коэффициент седиментации этого препарата не указан, но его электрофоретическая подвижность близка по величине к таковой препарата поли(С) фирмы “Serva” с коэффициентом седиментации 8-13 S. При электрофорезе препарата фирмы «Calbiochem» отмечено небольшое уширение медленнее мигрирующей зоны полимера, исчезающее в случае предварительного прогрева образца или же когда образец сначала растворяли в кислой среде при $I = 10$ мМ, а затем переводили в условия более высокой ионной силы.

Таким образом, хотя мы и не можем однозначно утверждать, что поли(С) не способен образовывать «замороженную» форму ни при каких условиях, тем не менее наличие двух перекрывающихся зон при электрофорезе поли(С) в кислой среде не связано с этой формой. Неэффективное образование «замороженной» формы в случае поли(С) может быть обусловлено тем, что макромолекулы этого полимера более жесткие, чем макромолекулы поли(А) [273, 274]. В частности, оценка персистентной длины одноцепочечных поли(А) и поли(С) посредством растяжения индивидуальных макромолекул показала, что r поли(С) (3,1 нм) в 1,7 раза больше, чем поли(А) (1,8 нм) [274].

Следует однако отметить, что в некоторых случаях при длительном хранении кислых растворов поли(С) полинуклетид начинал выпадать в осадок.

4.3.3.2. Электрофорез поли(С) в трис-фосфатном буфере. Описанные выше электрофоретические эксперименты проводились в трис-ацетатных буферных растворах. При кислых рН часть молекул уксусной кислоты, входящей в состав этих растворов, содержит недиссоциированные (нейтральные) карбоксильные группы. Согласно литературным данным [20, 275], нейтральная карбоксильная группа способна стабильно связываться с цитозином и цитидином и может конкурировать с гуанином за образование Н-

связей с цитозином. Как уже отмечалось выше, водородные связи в паре $\text{Cyt} \cdot \text{Cyt}^+$ подобны таковым в паре $\text{Cyt} \cdot \text{Gua}$.

Следовательно, нужно было ответить на вопрос – не препятствует ли недиссоциированная уксусная кислота образованию двойной спирали поли(С) и не относится ли медленнее мигрирующая зона этого полимера к его одноцепочечной форме. Для этого мы провели электрофорез поли(С) в трис-фосфатном буфере в разных условиях. На рис. 4.39 приведены денситограммы поли(С) при $\text{pH} = 5,3$ и разных значениях ионной силы: 8,5 (рис. 4.39 *а*) и 44,2 мМ (рис. 4.39 *б*).

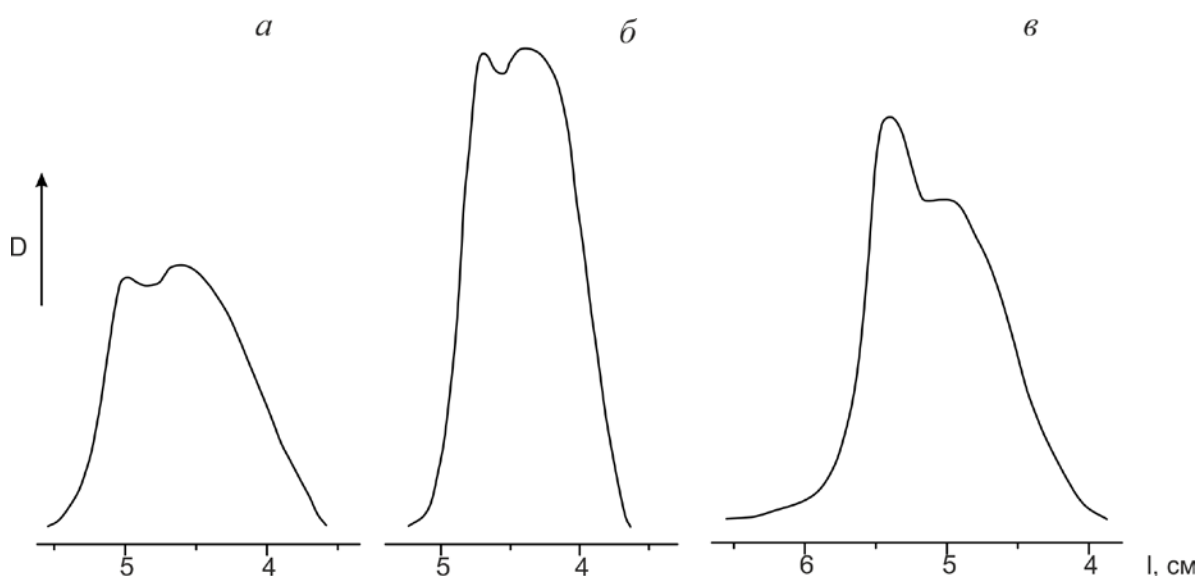


Рис. 4.39. Денситограммы поли(С) фирмы «Reanal» в трис-фосфатном буфере (*а* и *б*) и в 7,0 мМ NaH_2PO_4 (*в*). 0,8 %-ная агароза; *а* – $I = 8,5$ мМ; $t = 83$ мин; $\text{pH} = 5,3$; *б* – $I = 44,2$ мМ; $t = 92$ мин; $\text{pH} = 5,3$; *в* – $t = 94$ мин; $\text{pH} = 4,8$

Как следует из рисунка 4.39, в фосфатном буфере, так же как и в ацетатном, при электрофорезе протонированного поли(С) наблюдаются две перекрывающиеся зоны. Две зоны отмечаются и при более высокой ионной силе (74 мМ, не показано). Более того, две зоны регистрируются и в том случае, когда электрофорез проведен в отсутствие трис и ЭДТА, как это

следует из рис. 4.39 *в*, на котором показана денситограмма поли(С) в 7,0 мМ NaH_2PO_4 (рН = 4,8).

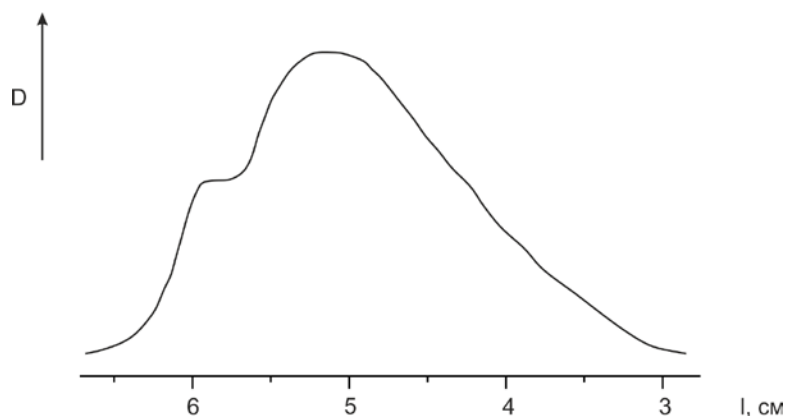
Важно отметить, что разделение на зоны других гомополирибонуклеотидов также не зависит от того, в каком буфере – ацетатном или фосфатном – проводили электрофорез. Однако подвижность полирибонуклеотидов в фосфатном буфере выше, чем в ацетатном. Например, в экспериментах, представленных на рис. 4.37 *а* и 4.39 *а*, узкая зона поли(С) мигрировала на расстояние, равное примерно половине длины трубки с агарозой, на 30 мин быстрее в фосфатном буфере, чем в ацетатном буфере.

Таким образом, регистрация методом электрофореза двух кислых форм поли(С) не обусловлена специфическим влиянием компонентов электрофоретического буфера на полинуклеотид.

4.3.3.3. Модель образования *i*-мотива поли(С). Проведенное нами исследование структурных переходов в поли(С) методом электрофореза выявило, что в интервале рН между двумя рК переходов, первый из которых соответствует (как думали раньше) образованию двойной спирали поли(С), а второй – ее распаду, поли(С) сосуществует в двух формах. Контрольные эксперименты показали, что появление двух форм не является особенностью какого-либо препарата поли(С), оно не обусловлено специфическим действием растворителя и медленнее мигрирующая форма не является «замороженной» формой, подобной той, которую легко образует поли(А) в кислых растворах.

Полученные нами результаты в совокупности с изложенными выше литературными данными [56, 271, 272] позволяют с большой долей вероятности предположить, что регистрируемые методом электрофореза протонированные формы поли(С) относятся к двуспиральной форме полимера и *i*-мотиву. При этом, дуплексу соответствует быстрее мигрирующая, а квадруплексу – медленнее мигрирующая форма. Дуплексы поли(С) имеют меньшие размеры и являются более гибкими, чем квадруплексы, и вследствие этого хуже разделяются по размеру в гелях агарозы: им соответствуют узкие (быстрее мигрирующие) зоны, тогда как более объемным и жестким

четырёхцепочечным структурам отвечают широкие зоны. Этот эффект заметнее выражен в случае препаратов поли(С), являющихся более высокомолекулярными, чем препараты фирмы «Reanal», таких например, как препараты фирмы «Calbiochem» (рис. 4.40).



4.40. Денситограмма поли(С) фирмы «Calbiochem» в трифосфатном буфере. $I = 8,5$ мМ; $t = 103$ мин; $pH = 5,3$

В данном случае большая ширина медленнее мигрирующей зоны может быть обусловлена не только гетерогенностью образца, но и образованием специфических четырехцепочечных структур. Например, при образовании *i*-мотива из дуплексов разной длины, на свободных концах могут образовываться другие *i*-мотивы. Образование длинных линейных структур из *i*-мотивов наблюдали для dC_7 (при спаривании цепей со сдвигом) [276] и олигонуклеотидов (C_nXC_m), содержащих олиго(С)-тракты разной длины, разделенные А/Т/С-линкером [277].

Учитывая тот факт, что при всех значениях pH ниже pK первого конформационного перехода в растворе присутствуют обе кислые формы поли(С), мы предположили, что *i*-форма, скорее всего, образуется лишь в том интервале pH , когда в растворе одновременно присутствуют, как одно-, так и двухцепочечные молекулы полинуклеотида, т. е. во время перехода поли(С) из нейтральной в кислую форму. Предлагаемая нами схема образования квадруплекса представлена на рис. 4.41. Согласно этому механизму, при наличии затравочного количества дуплексов, последующие дуплексы формируются непосредственно на них; основания одиночных цепей

интеркалируют между парами оснований уже имеющейся двойной спирали, приводя к образованию новых пар и четырехцепочечной структуры.

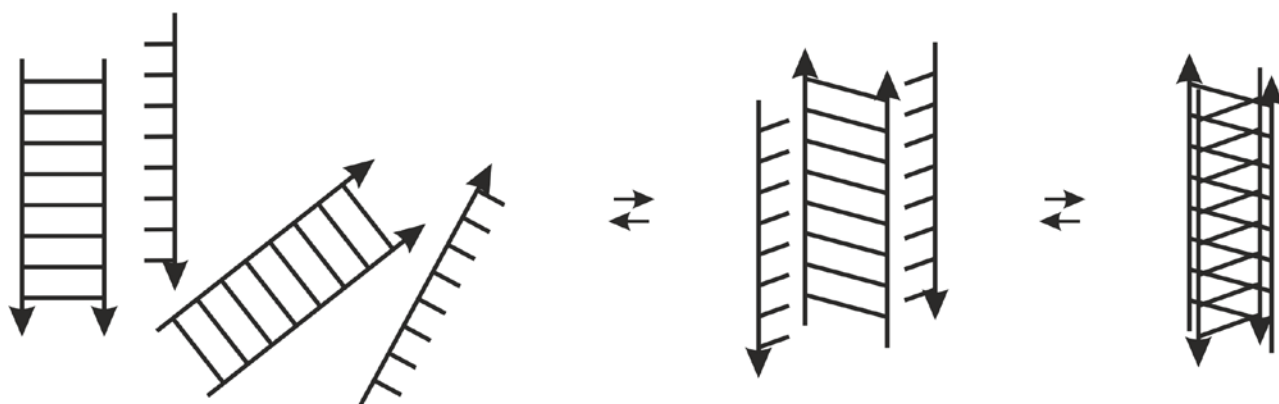


Рис. 4.41. Схема образования i-формы поли(С)

Небольшие различия между денситограммами, показанными на рис. 4.37 *а* и *в*, а также на рис. 4.37 *б* и *г*, на которых представлены образцы поли(С), приготовленные различными способами, можно, по-видимому, объяснить тем, что при подкислении поли(С) переходное состояние, в котором одновременно существуют и одноцепочечная, и двухцепочечная формы полинуклеотида, длится менее короткое время при электромиграции слабощелочного раствора поли(С) в кислый гель (рис. 4.37. *а* и *б*), чем в случае приготовления препарата в кислой среде (рис. 4.37. *в*) или медленном охлаждении прогретого образца (рис. 4.37 *г*), во время которого происходит реассоциация расплавленных дуплексов и квадруплексов.

Наше предположение о том, что «сборка» i-мотива поли(С) производится на дуплексе, находится в соответствии с данными работы Leroу [278], опубликованной позднее, чем наши результаты. Исследуя скорости образования и диссоциации $[TC_{3-5}]_4$ тетрамеров, Leroу показал, что образование i-мотива является реакцией третьего порядка; результат несовместимый с моделью, предполагающей ассоциацию двух предварительно образованных дуплексов. Он предположил, что образование тетрамера

происходит путем последовательного образования дуплекса и триплекса в качестве промежуточных состояний, причем лимитирующей стадией, скорее всего, является образование триплекса. В работе также сделано предположение о том, что стабильность дуплекса может увеличиваться с увеличением длины олиго(dC)-тракта.

Существование триплекса, как промежуточного состояния при образовании *i*-мотива экспериментально подтверждено в работе Dhakal et al. [279], исследовавших механическую стабильность *i*-мотива, образованного дезоксинуклеотидом (TGTCCCCACACCCC)₂ на уровне одиночной молекулы. В данном случае образовывался внутримолекулярный *i*-мотив. Авторы представили модель триплекса, отличающуюся от тетраплекса только отсутствием четвертого интеркалированного тракта (находящегося в свободном состоянии) (рис. 4.42).

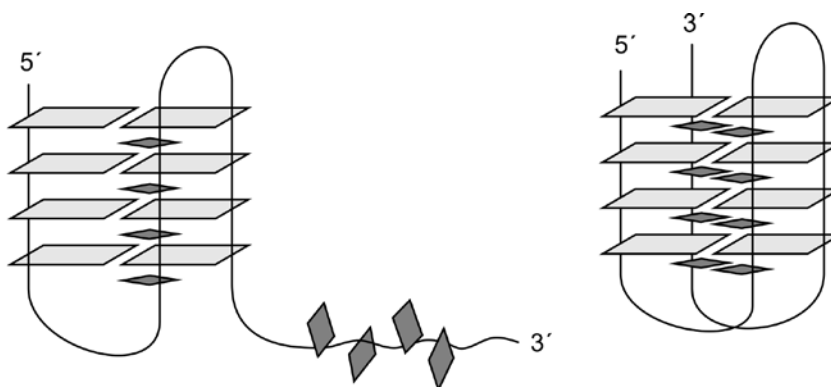


Рис. 4.42. Модели триплекса и *i*-мотива, образованных последовательностью (TGTCCCCACACCCC)₂, адаптировано из [279]

При pH 5,5 содержание триплекса в этом исследовании составляло 6,1 %, а *i*-мотива – 23 %. Следует отметить, что в этой работе, в отличие от работы Leroy [278], включение третьего тракта в *i*-мотив облегчалось его пространственной близостью к сформированному дуплексу, так как этот тракт принадлежал к той же самой молекуле.

Следует ли, учитывая данные работ [278, 279], относить быстрее мигрирующую зону протонированного поли(С) к триплексам, а не к дуплексам? По-видимому, в настоящее время однозначно ответить на этот вопрос не представляется возможным, поскольку, несмотря на сходство многих физико-химических характеристик *i*-мотивов полидезоксинуклеотидов и полирибонуклеотидов, между ними имеются и существенные различия, в частности *i*-мотивы ДНК значительно стабильнее, чем РНК.

Таким образом, результаты электрофореза дают новую информацию о структурных переходах в поли(С), которую следует учитывать при дальнейших исследованиях. В частности, они впервые показали, что при кислых рН сосуществуют две различные формы поли(С). Этот факт не учитывался ни в работах [56, 271, 272], в которых поли(С) исследовался спектроскопическими методами, ни в работе [49], где этот гомополимер исследовался методом рентгеноструктурного анализа.

4.4. Структурные переходы в поли(G)

4.4.1. Электрофорез поли(G) при разных ионных силах. В предыдущих исследованиях гомополирибонуклеотидов (параграф 4.1.) мы проводили электрофорез при ионных силах от 10 до 60 мМ и при рН от 7,8 до 4,9 в растворах, не содержащих солей двухвалентных металлов. В этих условиях наблюдалась одна узкая симметричная зона поли(G), за исключением эксперимента, проведенного при наиболее низкой ионной силе ($I = 10,2$ мМ) (рис. 4.43). В этом эксперименте наблюдалось небольшое уширение нижней части лидирующей границы зоны. Одной из причин этого уширения могло быть появление при $I \sim 10$ мМ новой зоны полимера, перекрывающейся с основной зоной, характерной для более высоких значений ионной силы. Для выяснения природы асимметрии электрофоретической зоны поли(G) при $I \sim 10$ мМ, мы провели более детальные исследования этого полинуклеотида методом электрофореза при низких ионных силах раствора.

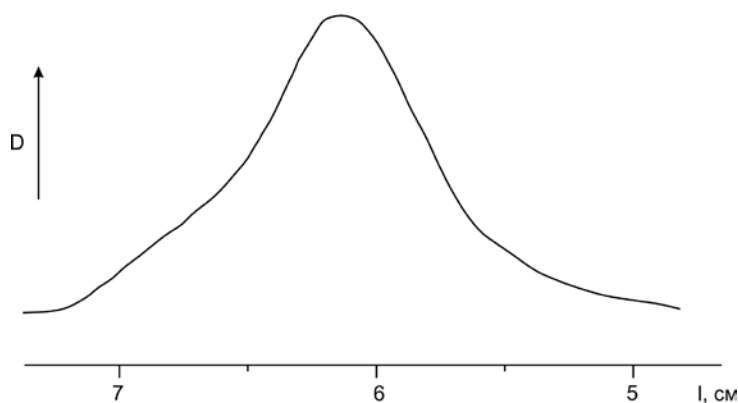


Рис. 4.43.
Денситограмма
поли(G) при $I = 10,2$
мМ; 0,8 %-ная
агароза

На рис. 4.44 представлены денситограммы диализованного препарата поли(G) при ионных силах 5,6 мМ (*a*) и 14,4 мМ (*б*). Использован вариант электрофореза (см. параграф 2.2), позволяющий в одном эксперименте проводить электрофорез полимеров при разных значениях ионной силы и рН.

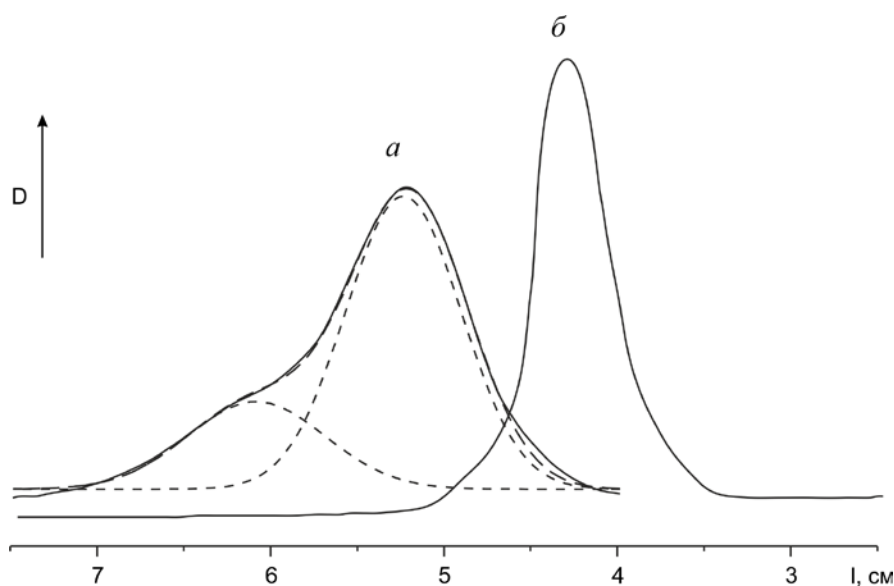


Рис. 4.44. Денситограммы поли(G) при разных ионных силах; *a* – $I = 5,6$ мМ, *б* – $I = 14,4$ мМ; 0,8 %-ная агароза, рН $\sim 7,0$

Как видно на рис. 4.44 при $I = 14,4$ мМ зона поли(G) узкая и довольно симметричная. Такие же зоны поли(G) наблюдались и в наших предыдущих исследованиях при I от ~ 12 до 60 мМ. Форма зоны *a* отличается от формы зоны *б*. Лидирующая граница этой зоны имеет заметное плечо. Асимметрия

зоны значительно более явная, чем на рис. 4.43 (при $I = 10,2$ мМ). Полученные результаты можно объяснить тем, что при низких ионных силах (< 10 мМ) четырехцепочечный поли(G) частично диссоциирует с образованием одноцепочечной формы, которая мигрирует с большей скоростью, чем квадруплекс. В условиях эксперимента зоны этих форм полимера частично перекрываются.

На рисунке 4.44 пунктиром показано разложение результирующей зоны *a* на составляющие зоны. Следует отметить, что в эксперименте, показанном на рис. 4.44, на гели, приготовленные как при $I = 14,4$ мМ, так и при $I = 5,6$ мМ, наносили поли(G), растворенный в диализном буфере (при ионной силе $14,2$ мМ). Появление плеча у зоны *a* свидетельствует о том, что диссоциация квадруплекса при уменьшении ионной силы произошла во время электрофореза, скорее всего, сразу же после вхождения образца в гель. Денситограммы с таким же плечом у быстрее мигрирующей зоны получены и с образцом, предварительно выдержанным в растворе с $I = 5,6$ мМ в течение 20 ч перед проведением электрофореза. Эти результаты хорошо согласуются с данными авторов работы [32], которые сообщают о том, что при изменении ионной силы две формы поли(G) легко переходят одна в другую либо в результате диализа, либо при непосредственном добавлении соли в раствор. В нашем случае роль диализа играет электромиграция полимера, растворенного в одних условиях, в гель, приготовленный в других условиях.

На рис. 4.45 *б* представлены результаты электрофореза поли(G), предварительно прогретого при $I = 5,6$ мМ в течение 15 минут при температуре $90-93$ °С и охлажденного на протяжении 1 ч при комнатной температуре, по сравнению с электрофорезом непрогретого образца (рис. 4.45. *а*). Видно, что после прогрева относительное количество быстрее мигрирующей формы полимера заметно увеличилось (примерно в 1,7 раза). Этот результат можно объяснить тем, что образовавшиеся в результате плавления квадруплекса одноцепочечные макромолекулы поли(G) не успели за время охлаждения полностью реассоциироваться в квадруплекс, поскольку

образование четырехцепочечной формы полимера является реакцией четвертого порядка, скорость которой зависит от концентрации полимера и одновалентных катионов в растворе [280].

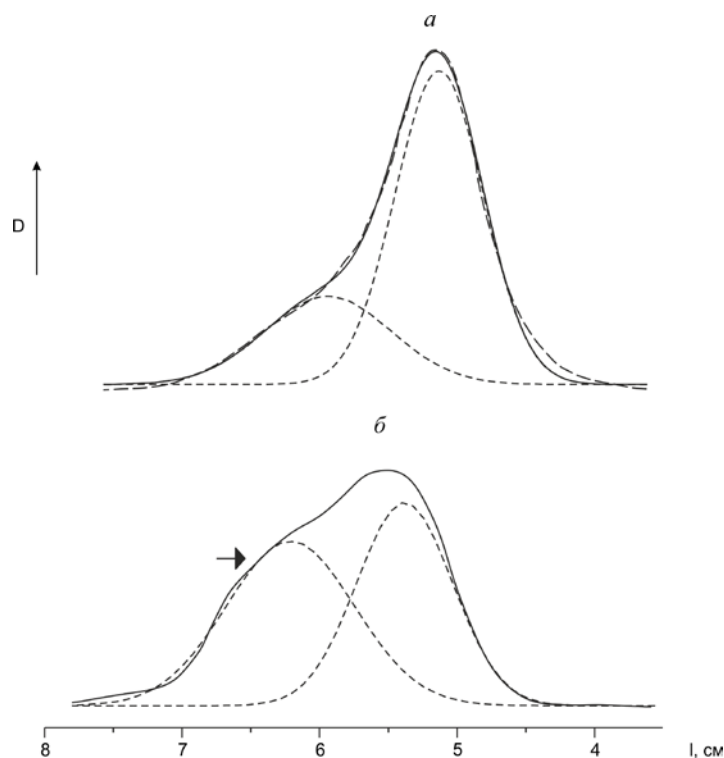


Рис. 4.45. Денситограммы поли(G) при $I = 5,6$ мМ: *a* – образец не прогрет, *б* – образец прогрет в течение 15 мин при температуре 90-93 °С. 0,8 %-ная агароза. Стрелкой обозначена точка, в которой снят спектр поглощения, показанный в следующем параграфе

Полученный результат свидетельствует в пользу того, что плечо зоны поли(G) образуется за счет появления при низких ионных силах одноцепочечной формы полимера, а не из-за каких-либо иных причин. Как упоминалось выше, электрофоретические эксперименты выполняли с диализованным препаратом поли(G). Подобные результаты получены и в опытах с использованием недиализованного препарата. С таким образцом был проведен дополнительный эксперимент. Учитывая, что скорость реассоциации G-квадруплекса возрастает при увеличении ионной силы [280], поли(G),

растворенный в 0,1 мМ ЭДТА и прогретый в течение 15 мин при температуре 100 °С, наносили после охлаждения на трубки с агарозой, приготовленной при $I = 14,2$. Однако, как видно на рис. 4.46, при этой ионной силе, как и при $I = 5,6$ мМ (рис. 4.45), реассоциация квадруплекса неполная: зона \bar{b} (прогретый образец) имеет небольшое плечо по сравнению с зоной a .

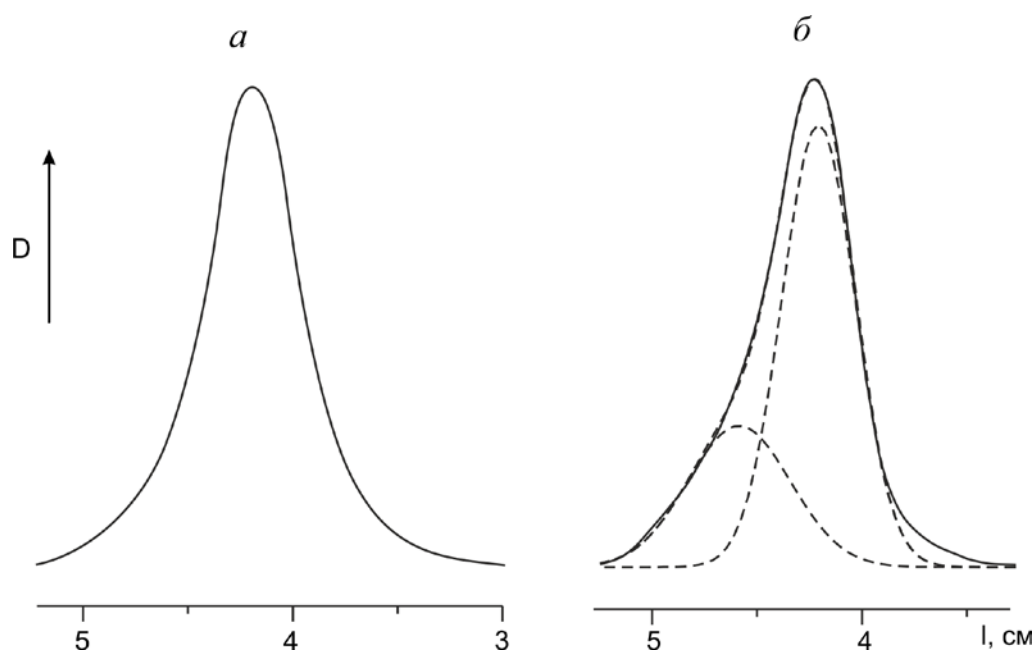


Рис. 4.46. Денситограммы поли(G), растворенного в 0,1 мМ ЭДТА, pH 6,5, при $I = 14,2$ мМ; a – образец не прогрет; \bar{b} – образец прогрет в течение 15 мин при температуре 100 °С. 1,0 %-ная агароза. D – оптическая плотность, l – расстояние от старта

Наличие одной зоны поли(G) при $I \sim 14$ мМ могло быть связано не с наличием одной формы поли(G) при этой ионной силе, а с тем, что полимеры хуже разделяются по размерам в агарозных гелях при высоких ионных силах, чем при низких из-за уменьшения размеров. Результаты эксперимента, показанного на рис. 4.46 \bar{b} (прогретый образец), исключают такой вариант – одноцепочечная и четырехцепочечная формы поли(G) разделяются при электрофорезе при $I \sim 14$ мМ, хотя и хуже, чем при более низких ионных

силах. Таким образом, исследование поли(G) методом электрофореза показало, что при ионной силе, равной или выше 14 мМ, он находится в четырехцепочечной форме, а при $I = 4-6$ мМ 20-30 % молекул полинуклеотида находятся в одноцепочечной форме.

Сравним полученные нами результаты с литературными данными, описанными в параграфе 1.1.3. В работах [33, 58, 59] поли(G) в H_2O , 2мМ NaCl или 1мМ трис, содержащем 1 мМ ЭДТА, уже находится в четырехцепочечной форме, тогда как в наших экспериментах при концентрации одновалентных катионов, равной 4-6 мМ, в этой форме находится лишь 70-80 % молекул поли(G). Такое противоречие можно объяснить, во-первых, тем, что чувствительность некоторых из использованных в [33, 58, 59] методов, могла быть недостаточно высокой, чтобы зарегистрировать одноцепочечную форму. Во-вторых, коммерческие препараты поли(G) выпускаются в виде соли калия или натрия и, по-видимому, они представляют собой квадруплексы. Растворение коммерческого препарата в воде [33] могло не приводить к диссоциации квадруплексов, поскольку экранирующие катионы не удалялись из раствора.

С другой стороны, авторы работы [31] диализовали поли(G), чтобы максимально удалить экранирующие моно- или дикатионы и исследовать структурные переходы этого полинуклеотида в зависимости от концентрации катионов Na^+ и Mg^{++} иммунологическим методом. Зависимость количества осажденного антителом поли(G) от концентрации NaCl, приведенная в этой работе, преобразована нами [281] в зависимость относительной концентрации четырехцепочечной формы поли(G) от концентрации соли (рис. 4.47).

Из этой зависимости следует, что поли(G) в H_2O или 5мМ трис, в основном, находится в одноцепочечной форме и при низких ионных силах, используемых нами (4-6 мМ), концентрация четырехцепочечной формы не превышает 20 %, а при $I = 14,4$ мМ – 40 %, хотя по нашим данным, она составляет соответственно 70-80 и 100 %. Возможно, что при концентрации полинуклеотида, используемой авторами [31] (~ 0,3 мМ), реакция не достигала

равновесия за время инкубации полинуклеотида и антитела, поскольку скорость образования квадруплекса существенно зависит от концентрации полинуклеотида [280]. В наших экспериментах на гели наносили растворы поли(G) в концентрации $\sim 1,7$ мМ.

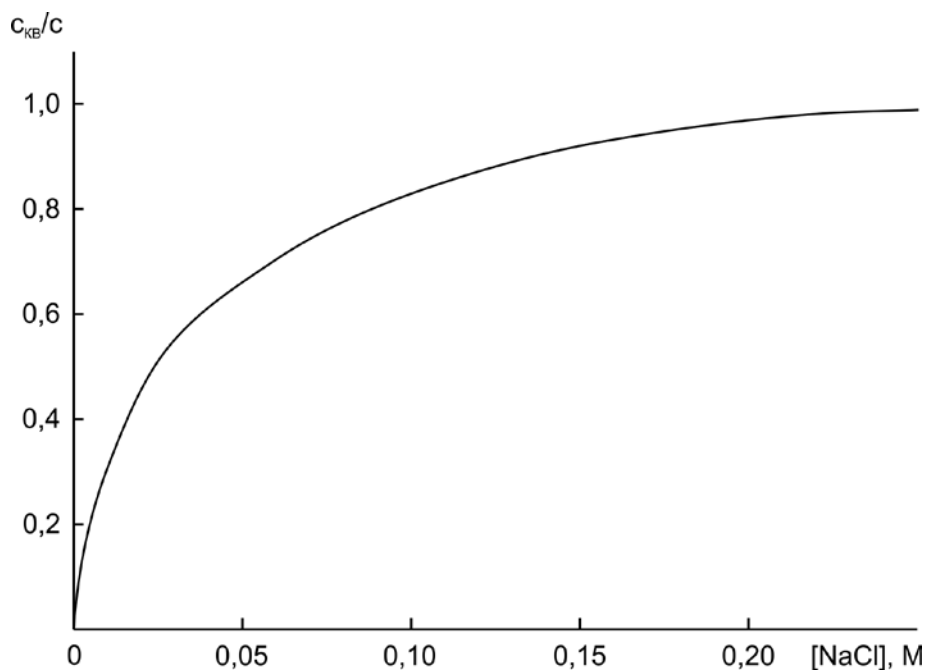


Рис. 4.47. Зависимость относительной концентрации четырехцепочечной формы поли(G) ($c_{кв}/c$) от концентрации соли по данным [31]; c – общая концентрация поли(G)

В работе [32], структурные переходы в поли(G) изучались методом спектроскопии комбинационного рассеяния. Переход поли(G) из одноцепочечной формы в четырехцепочечную происходил почти полностью, в 0,02 N NaCl. В 0,01 N соли переход был неполным, а в H_2O полимер находился, в основном, в одноцепочечной форме. Результаты этой работы согласуются с нашими данными.

4.4.2. Электронные спектры поглощения толуидинового голубого, связанного с гомополирибонуклеотидами в агарозных гелях. При исследованиях гомополинуклеотидов методом электрофореза, описанных в

предыдущих параграфах, агарозные гели после проведения электрофореза окрашивали толуидиновым голубым О (рис. 4.48 *a*, вставка) в слабощелочной среде вне зависимости от того, при каких условиях проводили электрофорез. В таких условиях поли(А), поли(С) и поли(У) находятся в одноцепочечной форме, а поли(Г) (при достаточно высокой концентрации соли) – в четырехцепочечной; при этом, как было указано в параграфе 2.2, поли(А) окрашивается в сиреневый цвет, поли(У) и поли(С) – в розовый, а поли(Г) – в фиолетовый. Окрашивание в разные цвета свидетельствует о том, что ТГ взаимодействует с гомополимерами по-разному. Для того чтобы выяснить, как ТГ взаимодействует с разными формами поли(Г) и других гомополинуклеотидов, гели были окрашены ТГ, растворенном в тех же буферных растворах, в которых проводили электрофорез, т.е. при разных рН и ионных силах. Электронные спектры поглощения окрашенных гелей записывали в разных местах геля. Полученные результаты представлены на рис. 4.48 *б* и суммированы в табл. 4.4.

Электронный спектр поглощения ТГ, связанного с быстрее мигрирующей формой поли(Г), (рис. 4.48, спектр 2) снят в точке, показанной стрелкой на рис. 4.45 (зона *б*). Соответствующий спектр в случае медленнее мигрирующей формы (рис. 4.48, спектр 3) снят в максимуме зоны *a* (рис. 4.45). Как видно на рис. 4.48 *б*, спектры значительно отличаются друг от друга: λ_{\max} спектра 2 смещен примерно на 20 нм в длинноволновую область по сравнению с λ_{\max} спектра 3. Эти данные еще раз подтверждают наш вывод о том, что при низких ионных силах поли(Г) сосуществует в двух разных формах.

Сравнение спектра ТГ, связанного с одноцепочечным поли(Г), со спектрами красителя, связанного с другими гомополимерами, показывает, что этот спектр (2), наиболее близок к спектру ТГ, связанному с поли(А) (1), λ_{\max} которого в наименьшей степени сдвинут в коротковолновую область по сравнению с λ_{\max} свободного красителя (рис. 4.48 *a*). Следует отметить, что

спектры ТГ в видимой области в H_2O и в растворах, используемых в данной работе, имеют одинаковую форму.

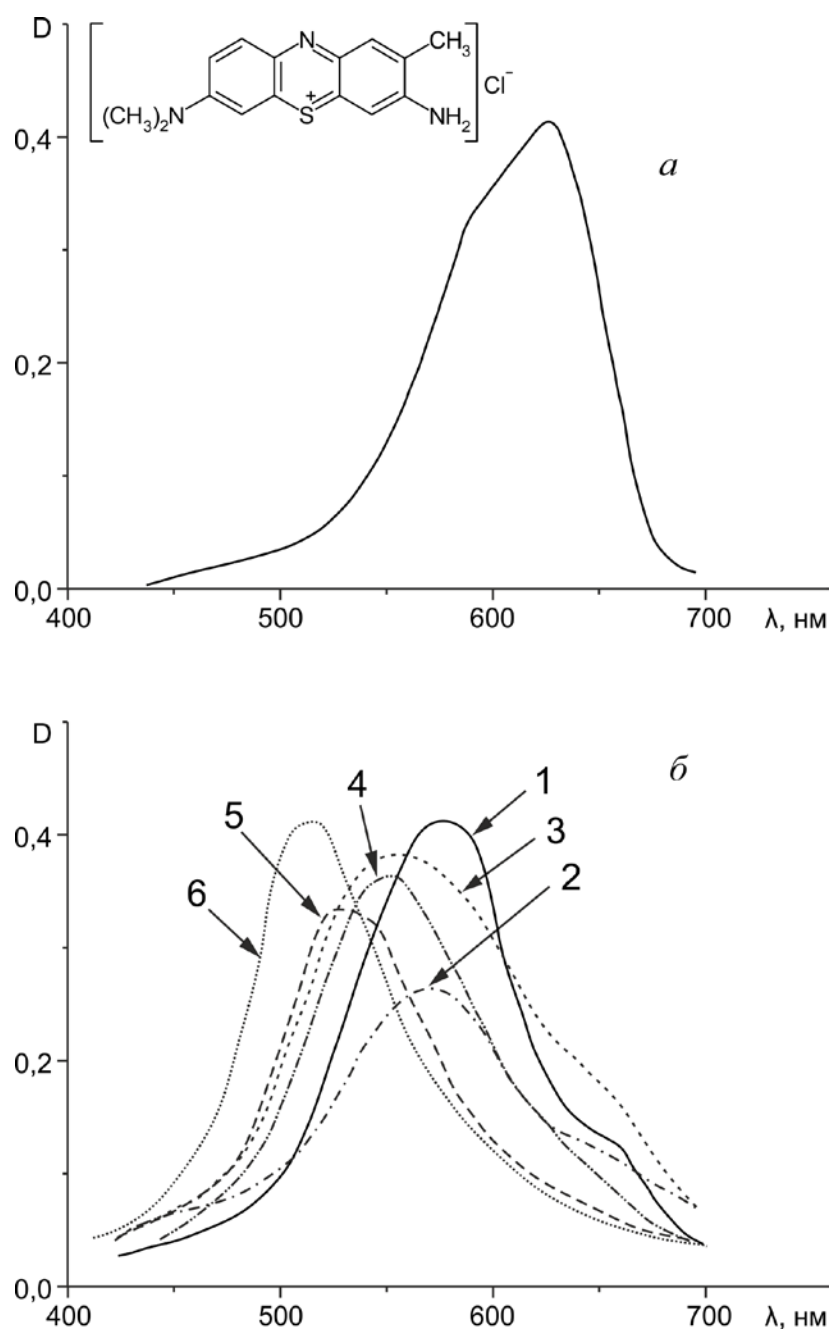


Рис. 4.48. Электронные спектры поглощения ТГ (а) и ТГ, связанного с гомополирибонуклеотидами в агарозных гелях (б). $I = 4-6$ мМ. 1 – поли(А), рН 4,9; 2 – одноцепочечный поли(Г), рН 7,0; 3 – четырехцепочечный поли(Г), рН 7,0; 4 – поли(С), рН 5,7; 5 – поли(У), рН 7,6; 6 – поли(С), рН 7,6

Таблица 4.4

**λ_{\max} спектров поглощения толуидинового голубого, связанного с
гомополирибонуклеотидами в агарозных гелях**

Гомополимер	Форма	λ_{\max} , нм	Номер спектра на рис. 4.48 б
Поли(А)	одноцепочечный	577	Не приведен
	двухцепочечный	577	1
Поли(Г)	одноцепочечный	570	2
	четырёхцепочечный	548	3
Поли(С)	четырёхцепочечный	547	4
Поли(У)	одноцепочечный	532	5
Поли(С)	одноцепочечный	511	6

Примечание. Ошибка в определении λ_{\max} в случае одного и того же гомополимера от эксперимента к эксперименту составляет не более ± 3 нм.

Поли(А) при слабокислых рН переходит в две различные двухспиральные формы (см. параграф 4.2.). Однако спектры ТГ, связанного с поли(А), снятые после электрофореза в щелочном и кислом буферных растворах, не отличаются друг от друга в пределах ошибки эксперимента (спектр 1 на рис. 4.48 б соответствует двухцепочечной форме поли(А)). По-видимому, этот краситель взаимодействует с одно- и двухцепочечными формами поли(А) одинаковым образом. λ_{\max} спектра поглощения ТГ, связанного с четырехцепочечной формой поли(Г) (рис. 4.48, спектр 3), близок к λ_{\max} спектра ТГ, связанного с поли(С) в кислой среде (спектр 4), однако формы спектров различны. При низких ионных силах i-мотив поли(С) значительно преобладает по сравнению с быстрее мигрирующей зоной (рисунки 4.37 в и 4.40), и спектр 4, показанный на рис. 4.48, относится, по всей видимости, к ТГ, связанному с четырехцепочечной формой поли(С).

Таким образом, спектр ТГ, связанного с одноцепочечной формой поли(G), близок к спектрам красителя, связанного с одноцепочечной и двухцепочечной формами другого пуринового гомополимера – поли(A), а спектр ТГ, связанного с четырехцепочечной формой поли(G), близок к спектру ТГ, связанного с четырехцепочечной формой поли(C). Спектры красителя, связанного с одноцепочечными формами пиримидиновых гомополимеров поли(U) (спектр 5) и поли(C) (спектр 6), близки между собой, но заметно отличаются от спектров ТГ, связанного с другими гомополинуклеотидами. Их λ_{\max} в наибольшей степени сдвинуты в коротковолновую область по сравнению с λ_{\max} свободного ТГ. В целом разница в λ_{\max} для ТГ, связанного с поли(A) и одноцепочечным поли(C) составляет ~ 70 нм.

Окрашенные зоны всех гомополирибонуклеотидов и их форм, представленных на рис. 4.48 б, отличаются по цвету визуально: поли(A) выглядит сиренево-голубым, поли(G) и i-форма поли(C) имеют разные оттенки сиреневого и фиолетового цветов, а поли(U) и одноцепочечный поли(C) – разные оттенки розового цвета. В восприятие цвета, по-видимому, вносит вклад также рассеяние света на агарозном геле, поскольку комплексы ТГ с гомополинуклеотидами в гелях и растворах выглядят по-разному. В частности, величины λ_{\max} спектров поглощения ТГ в растворе в присутствии щелочной и кислых форм поли(C) близки к соответствующим величинам λ_{\max} в агарозных гелях, но растворы полимера в присутствии ТГ при разных рН имеют лишь разные оттенки сиреневого цвета, а цвет гелей при переходе от кислых форм поли(C) к щелочной меняется с сиреневого на розовый.

Насколько нам известно, взаимодействие гомополирибонуклеотидов с ТГ ранее не изучали, однако были исследованы взаимодействия поли(A) с другими катионными красителями планарной структуры – метиленовым синим [282] и акридиновым оранжевым [283]. Авторы указанных работ показали, что красители связываются с поли(A) разными способами в зависимости от условий эксперимента и формы полимера. К ним относятся

интеркаляция между основаниями, интеркаляция между парами оснований как изолированных мономеров красителя, так и взаимодействующих между собой на соседних местах связывания, связывание с полинуклеотидом полуинтеркалированных димеров красителя, а также внешнее электростатическое связывание димеров красителя с фосфатными группами. Учитывая разнообразие спектров ТГ, связанного с разными гомополимерами и их формами, можно предположить, что краситель взаимодействует с поли(А), поли(С), поли(U) и поли(Г) разными способами, включающими, возможно, в отдельных случаях образование водородных связей.

4.5. Выводы к разделу 4

Методом электрофореза зарегистрировано специфическое взаимодействие производных заряженных аминокислот с гомополимерами. Устойчивость комплексов гомополимеров с эфирами лизина и аргинина убывает в ряду: поли(Г) > поли(А) ≥ поли(С) > поли(U). Эфир лизина уменьшает эффективный размер поли(U), по-видимому, индуцируя образование коротких шпилек в этом гомополимере. Эфир аргинина вызывает частичный переход *DS*-формы поли(А) в *DS_p*-форму.

Структурные переходы в гомополимерах впервые исследованы методом электрофореза и протонной буферной емкости. Подтверждено существование третьей двухспиральной формы поли(А), которая образуется вследствие конформационного перехода *DS*-формы в форму *DS'* при рН ~ 5,0. Определены параметры всех четырех базисных функций в разложении кривых буферной емкости поли(А) в зависимости от ионной силы и температуры.

Показано, что образованию «замороженной» формы поли(А) способствуют условия, облегчающие сворачивание макромолекулы: повышение ионной силы раствора и увеличение длины полимера. Предложена и обоснована гипотеза о том, что поли(А) может сворачиваться с образованием внутримолекулярных двойных спиралей. Одновременное образование при

кислых рН как межмолекулярных, так и внутримолекулярных двойных спиралей приводит к появлению «замороженной формы» поли(А).

Предложена и обоснована гипотеза о том, что в стабилизацию одноцепочечной и двухцепочечной спиралей поли(А) вносят вклад вертикальные водородные связи ...N6-H...N6-H...N6H....

Впервые зарегистрирован дополнительный процесс некооперативного протонирования поли(С). Он отнесен к некооперативному протонированию электронейтральных цитозиновых оснований по атому кислорода с *цис*-ориентацией присоединенного протона относительно гликозидной связи. Эта реакция приводит к уменьшению термодинамической стабильности спиралей поли(С) и, в конечном счете, к их диссоциации. Диссоциация двойных спиралей сопровождается переносом протонов с атомов O2 на соседние атомы N3 цитозиновых остатков в результате двухстадийного процесса с участием молекул воды. Определены параметры всех базисных функций в разложении кривых буферной емкости поли(С) и поли(dС).

Методом электрофореза впервые обнаружено, что при кислых рН поли(С) присутствует в двух разных формах. Предположено, что медленнее мигрирующая зона протонированного поли(С) относится к его четырехцепочечной форме (*i*-мотиву), а быстрее мигрирующая зона – к двухцепочечной (или трехцепочечной) форме. Исследована зависимость появления двух зон протонированного поли(С) от рН и ионной силы. Предложен молекулярный механизм образования *i*-мотива из дуплексов и одиночных цепей.

Методом электрофореза показано, что в отсутствие ионов двухвалентных металлов структурный переход поли(Г) из одноцепочечной формы в четырехцепочечную завершается при $I \sim 14$ мМ, а при $I \sim 10,2$ мМ в этой форме находятся 94 ± 2 % макромолекул. Впервые выявлено, что ТГ по-разному взаимодействует не только с разными гомополирибонуклеотидами, но и с разными формами этих полимеров, что подтверждает регистрацию методом электрофореза двух разных форм поли(Г).

РАЗДЕЛ 5

СТРУКТУРНЫЕ МОДЕЛИ УЧАСТИЯ САМОАССОЦИАТОВ ГОМОПОЛИМЕРОВ В БИОЛОГИЧЕСКИХ ПРОЦЕССАХ

5.1. Возможное вовлечение внутримолекулярных двойных спиралей олиго(А) в биологические процессы

Из изложенного выше материала следует, что поли(А) образует в кислых растворах ряд форм, в частности, две двухспиральные формы (DS и DS_p) и «замороженную» форму; поли(С) в этих условиях образует i-мотив и двухцепочечную (или трехцепочечную) форму. Поли(G) при физиологических значениях рН и ионных сил выше 10 мМ существует в виде квадруплексов. Образуются ли разные формы гомополирибонуклеотидов в клетке и принимают ли они участие в биологических процессах?

Самоассоциаты поли(А) и поли(С) индуцируются протонированием оснований. Может ли протонирование происходить в клетке? *A priori* отрицать такую возможность нельзя, хотя бы уже по той причине, что локальные физико-химические условия, в первую очередь рН, в соответствующих микрокомпартаментах, в которых локализованы РНК в клетке, неизвестны. Главный аргумент в пользу существования самоассоциатов поли(А) и поли(С) заключается в том, что необходимые условия для протонирования гомополинуклеотидных участков в РНК могут быть созданы взаимодействующими с ними белками.

Протонирование гомополинуклеотида может быть вызвано переносом протона от кислых аминокислотных остатков белка вдоль межмолекулярных водородных связей. Возможность такого процесса следует, в частности, из ряда модельных экспериментов. Например, методом рентгенографии показано, что в тройном кристаллическом комплексе Ade, Ura и карбоксиэтила образуется хугстиновская пара Ade : Ura, а карбоксильная группа своим

гидроксильной группой протонирует атом N1 Ade [284]. Межмолекулярный перенос протона по водородной связи от карбоксильной группы аминокислот на нуклеотидные основания характерен также и для многих других комплексов (см. [20] и приведенную там библиографию).

В пользу такого переноса свидетельствуют также данные работы [25] об образовании двойных спиралей поли(А) при нейтральных рН в присутствии углеродных нанотрубок, модифицированных путем присоединения карбоксильной или гидроксильной группы. Сами авторы полагают, что причиной самоассоциации поли(А), возможно, являются благоприятные электростатические взаимодействия между заряженной парой оснований $A \cdot A^+$ и карбоксильной или гидроксильной группой на трубках. Ранее было показано, что модифицированные нанотрубки могут также индуцировать образование *i*-мотива в теломерной ДНК человека при физиологических значениях рН [285].

Протонирование оснований под действием белка наблюдали кроме того и в надмолекулярных системах. Протонирование цитозиновых и адениновых остатков РНК в составе тимовирусных вирионов было зафиксировано методом Рамановской спектроскопии [286, 287]. Предполагается, что оно обусловлено близкими контактами между основаниями и кислыми аминокислотными остатками белка оболочки [19].

Таким образом, рибонуклеиновые кислоты могут протонироваться в клетке, вследствие чего их гомополимерные тракты могут образовывать двойные спирали или *i*-мотив. Какова функциональная значимость процесса образования самоассоциатов гомополинуклеотидов в клетке? В первую очередь рассмотрим возможные биологические функции внутримолекулярных двойных спиралей полиадениловой кислоты. Модели их участия в биологических процессах, предложенные нами, опубликованы в работах [227, 230, 288-291].

5.1.1. Стабилизация ARE-содержащих мРНК белками семейства Hu/ELAV. В таблице 5.1 перечислены биологические процессы, в которых могут принимать участие внутримолекулярные двойные спирали олиго(A).

Таблица 5.1.

Возможное участие внутримолекулярных дуплексов олиго(A) в биологических процессах

Процесс	Белки, участвующие в процессе. (Выделены поли(A)-связывающие белки (PABs))	Мотивы PABs, специфически взаимодействующие с поли(A)
Стабилизация мРНК GAP-43, содержащей ARE	<u>HuD</u>	RRM3
Стабилизация мРНК казеина, содержащей ARE	<u>HuR</u> , <u>PABPC</u>	RRM3 RRM1+RRM2
Стабилизация мРНК <i>c-fos</i> , содержащей элемент mCRD	<u>PABPC</u> , PAIP1, UNR, AUF1, NSAP1	RRM1+RRM2
Стабилизация мРНК α globin, содержащей C-богатый элемент	<u>PABPC</u> , α CP1, α CP2, AUF1	RRM1+RRM2
Саморегуляция синтеза PABC	<u>PABPC</u> , IMP1, UNR	RRM1+RRM2
Контроль длины поли(A)-хвоста мРНК высших эукариотов	PAP, CPSF, <u>PABPN1</u>	RRM
Контроль длины поли(A)-хвоста мРНК дрожжей	Pap1, CPF, CF I, <u>Nab2</u> , <u>Pab1p</u>	ZnF5+ZnF6+ZnF7, в основном RRM2

Рассмотрим сначала процессы с участием белков семейства Hu/ELAV (2-я и 3-я строки таблицы 5.1). Белки этого семейства выполняют различные функции в метаболизме мРНК [292]. HuR экспрессируется во многих тканях, места его связывания найдены в тысячах разных транскриптов. Другие белки, HuB, HuC и HuD, специфичны для нейронов. Белки Hu/ELAV имеют три мотива RRM и специфически связываются как с AU-богатым участком (ARE) в мРНК, так и с поли(А)-хвостом (через третий RRM), причем возможно одновременное связывание с ARE и с поли(А)-трактом [6]. Уникальной особенностью этих белков является то, что они связываются преимущественно с длинными молекулами поли(А) (≥ 70 нт), узнавая, как предполагают авторы, какую-то вторичную или третичную структуру.

В работе [293] показано, что повышенная экспрессия белка HuR приводит к стабилизации мРНК-конструктов, содержащих ARE класса I и II. Предполагается, что для проявления этого эффекта важно взаимодействие между RRM3 и поли(А)-хвостом мРНК. Учитывая эти факты, мы предполагаем, что HuR, связанный с ARE, индуцирует образование внутримолекулярного дуплекса в поли(А)-хвосте, который препятствует функционированию поли(А)-нуклеаз (рис. 5.1).

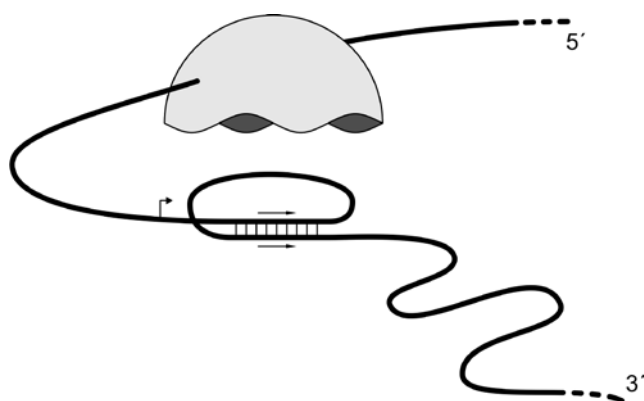


Рис. 5.1. Модельная схема образования внутримолекулярных двойных спиралей олиго(А) в процессе стабилизации ARE-содержащих мРНК ELAV-подобными белками. Начало поли(А)-хвоста отмечено стрелкой

С ARE-содержащими мРНК связываются многие белки. Одни из них причастны к ускорению распада мРНК, а другие к их стабилизации ([131] и ссылки в этой работе). Какие именно белки связываются с данным ARE – зависит от многих факторов: внутриклеточной локализации и концентрации белков, кинетических параметров их связывания с РНК и др. Предлагаемая нами модель стабилизации ELAV-подобными белками ARE-содержащих мРНК прежде всего относится к тем случаям, когда в процесс стабилизации вовлечены длинные поли(А)-хвосты.

В пользу этой модели свидетельствуют, например, результаты работы [294], где показано, что белок HuD при повышенных уровнях экспрессии стабилизирует мРНК GAP-43, содержащие только длинные поли(А)-хвосты (по крайней мере, размером в 150 нт) и его сродство к таким мРНК в 8 раз выше, чем к мРНК, имеющим поли(А)-хвосты длиной в 30 нт. Аналогичные результаты получены и в случае мРНК β -казеина [295]. Во время лактации эта мРНК приобретает длинный поли(А)-хвост, что приводит к ее стабилизации за счет образования рибонуклеопротеинового комплекса, включающего ARE, поли(А)-хвост, HuR и PABPC. Механизмы стабилизации мРНК в обоих случаях неизвестны. По мнению авторов работы [131] все экспериментальные результаты на момент написания ими статьи согласуются с тем, что контакты Hu-белков с длинными поли(А)-хвостами имеют другую природу, чем контакты с ARE и, вероятно, связаны с наличием ультраструктуры в поли(А)-хвостах. Авторы ссылаются на нашу гипотезу, о том, что ультраструктура может представлять собой двойные олиго(А)-спирали.

Внутримолекулярные двойные олиго(А)-спирали могут участвовать также в процессах стабилизации мРНК, не связанных с Hu-белками, например, в стабилизации мРНК c-fos (см. параграф 1.2.2) и мРНК α -глобина [296] (4-я и 5-я строки в табл. 5.1), детальные механизмы стабилизации которых не известны. В стабилизации мРНК α -глобина участвуют белки α CP1 и α CP2, связанные с С-богатым элементом (CRE), ARE-связывающий белок AUF1 и PABPC. Этот комплекс одновременно блокирует связывание с CRE

эндонуклеазы, специфичной для эритроидных клеток, и стабилизирует мРНК, ингибируя процесс ее деаденилирования.

В процессах стабилизации мРНК *c-fos* и α -глобина участвует PABPC. Мы предполагаем, что этот белок, взаимодействующий с поли(A) посредством доменов RRM 1 (RNA recognition motif 1) и RRM 2 [297], индуцирует образование внутримолекулярных дуплексов в поли(A) только в комплексе с другими белками, которые могут служить или донорами для протонирования адениновых оснований, или вызывать изменения в структуре PABPC, приводящие к экспонированию его собственных кислых аминокислотных остатков. Интересно отметить, что среди партнеров PABPC в процессах, описанных в таблице 5.1, дважды встречаются белки UNR и AUF1 (строки 4-6).

5.1.2. Саморегуляция синтеза цитоплазматического поли(A)-связывающего белка. Образование внутримолекулярных двойных спиралей олиго(A) может также происходить в процессе саморегуляции синтеза белка PABPC (табл. 5.1, строка 6). Авторы работы [298] показали, что при избытке PABPC в цитоплазме он связывается с консервативной А-богатой областью (ARS), расположенной в 5'-нетранслируемой области своей мРНК, и ингибирует свою трансляцию. ARS длиной 50-70 нт состоит из нескольких олиго(A)-трактов, чередующихся с короткими последовательностями других нуклеотидов. Позднее было показано, что рибосомные частицы 40 S задерживаются перед ARS, связанным с PABPC, и не достигают кодона инициации трансляции [299].

Мы предположили, что препятствием для продвижения 40 S рибосомных частиц является двойная олиго(A)-спираль, образованная внутри ARS в результате его самосворачивания, а донорами протонов для адениновых остатков могут служить определенные белки, взаимодействующие с PABPC. Patel et al. [11] идентифицировали два дополнительных белка, участвующих в процессе саморегуляции синтеза PABPC: IMP1 (insulin-like growth factor II mRNA binding protein 1) и UNR, причем IMP1 может, как и PABPC,

связываться с ARS независимо. Для связывания белков с ARS существенно наличие пиримидиновых линкеров между олиго(А)-трактами. Patel et al. [11] полагают, что нуклеопротеиновый комплекс, включающий ARS, PABPC, IMP1 и UNR, может служить платформой для присоединения дополнительных белков и образования еще большего по размеру комплекса, представляющего собой более эффективное препятствие, чем комплекс ARS-PABPC, на пути такой большой молекулярной машины, как рибосомная субъединица 40 S.

Эти данные позволяют предположить, что именно IMP1 и/или UNR являются донорами протонов для образования дуплекса поли(А) в предложенной нами модели. Эффективность препятствия, вероятнее всего, связана не столько с размерами белкового комплекса, сколько с изменением структуры ARS, т.е. с образованием двойной олиго(А)-спирали. Большие белковые комплексы не обязательно служат препятствием для продвижения рибосомы. В частности, комплекс из пяти белков, собранный на сигнальном элементе mCRD в мРНК *c-fos* (см. параграф 1.2.2) смещается рибосомой во время трансляции. С другой стороны, вторичная структура 5'-нетранслируемой области влияет на эффективность инициации трансляции [300], и дуплексы поли(А) могут усиливать ингибирующее действие комплекса белков, связанных с ARS. Кроме того, олиго(А)-дуплекс с параллельно ориентированными цепями и неканоническими парами оснований может представлять неподходящий субстрат для геликазы eIF4A и других геликаз, вовлеченных в инициацию трансляции, в отличие от РНК-дуплексов с антипараллельными цепями.

Недавно Kikuchi et al. [301] на основании того факта, что фрагменты РНК, соответствующие ARS, могут быть закристаллизованы, предположили, что А-повторы в этой области имеют специфическую регулярную структуру, включающую водородные связи между повторами, отличную от структуры олиго(А), не содержащего вставок из других оснований. По их мнению, регулярная структура способствует трансляции, а связывание PABPC при высокой концентрации с этой структурой разрушает ее, что приводит к

подавлению трансляции. В свете вышеизложенного такой механизм представляется маловероятным. Скорее всего, условия, в которых проводилась кристаллизация в работе [301], в частности, слегка кислый pH (6,5), способствовали образованию двойных спиралей между А-повторами в исследуемых фрагментах.

5.1.3. Контроль длины поли(А)-хвостов мРНК эукариотов. Две последние строки в табл. 5.1. относятся к процессам терминации полиаденилирования про-мРНК в эукариотах. Ядерный поли(А)-связывающий белок PABPN1 может рассматриваться как потенциальный донор протонов, поскольку его N-концевой домен содержит 36 остатков кислых аминокислот [302]. Этот белок вместе с PAP и CPSF осуществляет вторую стадию процесса полиаденилирования про-мРНК высших эукариотов – синтез поли(А)-хвоста [10]. В системе *in vitro* как и *in vivo*, PAP в присутствии CPSF и PABPN1 синтезирует поли(А)-хвост длиной около 250 нт за один акт присоединения, т.е. процессивно. При достижении этой длины процессивная реакция добавления адениновых остатков по неизвестной причине прекращается. Полимераза при этом диссоциирует и процесс становится дистрибутивным.

Kühn et al. [303] предложили следующую модель терминации процессивного синтеза поли(А). CPSF, связанный с гексамером AAUAAA, привлекает PAP к 3'-ОН концу расщепленного транскрипта. После синтеза олиго(А)-хвоста длиной ~ 12 нт к комплексу полиаденилирования присоединяется молекула PABPN1 и при продолжении синтеза новые молекулы PABPN1 связываются с поли(А)-хвостом, «покрывая» его полностью, причем поли(А)-хвост заворачивается сам на себя таким образом, чтобы поддерживался контакт между PAP и CPSF. Связывание молекул PABPN1 с поли(А)-последовательностью длиной ~ 250 нт приводит, как было показано ранее в работе [304], к образованию сферической частицы. При этом расстояние между PAP и CPSF остается близким. Однако сферическая частица не способна вмещать в себя следующие молекулы PABPN1, что ведет к разрушению контакта между PAP и CPSF. Дальнейший синтез, проводимый

PAP только при участии PABPN1, является значительно менее процессивным, чем при кооперативном взаимодействии полимеразы с CPSF и PABPN1.

Kerwitz et al. [305] показали, что только одна особая молекул PABPN1 из всех молекул, связанных с поли(A)-хвостом, участвует в стимуляции синтеза и предположили, что такой молекулой является та, что расположена рядом с PAP, то есть, согласно предложенной модели [303], во время образования сферической частицы каждый раз новая молекула PABPN1 участвует в процессивной реакции полиаденилирования.

Такой механизм процессивного синтеза кажется нам менее вероятным, чем если бы синтез осуществлялся постоянным комплексом, состоящим из CPSF, PAP и одной из молекул PABPN1, присоединившихся к поли(A)-хвосту в числе первых. Мы полагаем, что 250 нт – это примерно та длина, при которой синтезируемый поли(A)-хвост может с высокой вероятностью свернуться с образованием внутримолекулярной двойной спирали под действием кислых аминокислотных остатков присоединенных к нему молекул белка PABPN1 (рис. 5.2). Резкий переход части одноцепочечного поли(A)-хвоста в двойную спираль будет способствовать распаду комплекса белков, осуществляющих процессивное полиаденилирование, например в результате «выдергивания» из этого комплекса молекулы PABPN1.

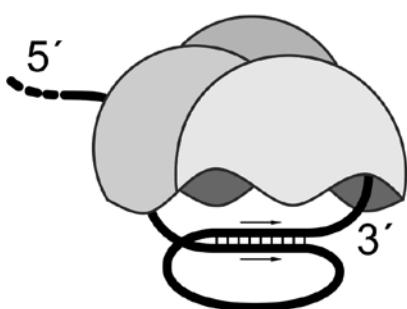


Рис. 5.2. Модельная схема образования внутримолекулярной двойной спирали олиго(A) в процессе терминации синтеза поли(A)-хвоста про-мРНК эукариотов

Следует отметить, что природа контактов между PAP и CPSF или PABPN1 в процессе полиаденилирования про-мРНК высших эукариотов в настоящее время полностью не выяснена, также как не выяснена и структура комплекса PABPN1-поли(A) [10].

В недавней работе Sagawa et al. [306] показано, что *in vivo* в синтезе поли(А)-хвостов кроме РАР, CPSF и РАВРН1 участвует также многофункциональный белок нуклеоплазмин (NPM1), способный взаимодействовать с CPSF. Согласно одной из моделей, предложенных авторами, NPM1, связываясь с транскриптом на участке, расположенном непосредственно перед гексамером AAUAAA, облегчает разрушение контакта между РАР и CPSF при терминации процессивного синтеза. Мы не исключаем также, что этот белок может стабилизировать комплекс, состоящий из РАР, CPSF и одной и той же молекулы РАВРН1 во время процессивного синтеза. Кроме того, в работе [307] показано, что наряду с РАВРН1 с поли(А)-хвостами транскриптов связывается также цитоплазматический белок РАВРС. Его точные функции в ядре в настоящее время не установлены, но он тоже может оказывать влияние на процесс полиаденилирования.

По нашему мнению сферические частицы, обнаруженные авторами работы [304] методами электронной и силовой микроскопии, представляют собой внутримолекулярные двойные спирали поли(А), образованные за счет переноса протонов от РАВРН1 на адениновые основания. В этой работе показано, что РАВРН1, связанный с поли(А), образует как нити (максимального диаметра 7 нм), так и компактные частицы (максимального диаметра 21 ± 2 нм). Частица размером 21 нм образуется лишь на поли(А) длиной не менее чем 200-300 нт, соответствующей длине вновь синтезированных поли(А)-хвостов про-мРНК высших эукариотов. На длинных молекулах поли(А) (> 1000 нт) образуется несколько сферических частиц.

В пользу нашей гипотезы о внутримолекулярных дуплексах, наблюдаемых в работе [304], свидетельствуют следующие факты. Во-первых, равновесие между нитями и компактными частицами зависит от ионной силы. Частицы образуются преимущественно при умеренной и высокой ионной силе (т. е. когда молекулы полимера более гибкие). Во-вторых, образование «шишкообразных» структур в поли(А) наблюдали методом силовой атомной

микроскопии [239] и в отсутствии поли(А)-связывающих белков, в условиях благоприятных для образования двойных спиралей (см. параграф 4.2.2.).

Следует отметить, что длина, при которой поли(А) будет образовывать внутримолекулярную двойную спираль под действием того или иного белка – донора протонов, по-видимому, зависит не только от количества кислых аминокислотных остатков в белке, но и от их доступности и структуры белково-нуклеинового комплекса. Причем донором протонов, как уже упоминалось, могут являться белки-партнеры, взаимодействующие с поли(А)-связывающим белком. В случае PABPN1 необходимая степень протонирования поли(А) достигается, по-видимому, когда белок покрывает участок поли(А)-хвоста длиной 250 нт; в случае других белков или белковых комплексов длина поли(А)-тракта может быть другой.

В частности, реакция полиаденилирования про-мРНК дрожжей осуществляется при участии других сигнальных последовательностей на транскрипте и других белковых факторов, и поли(А)-хвосты имеют длину 70-80 нт (см. параграф 1.2.2). В контроле длины поли(А)-хвостов про-мРНК дрожжей принимают участие два поли(А)-связывающих белка: Pab1p и Nab2p (nuclear abundant poly(A) RNA binding protein), отличные по структуре от PABPN1. В частности, Pab1p содержит четыре консервативных мотива RRM, PABPN1 содержит только один такой мотив [4], а Nab2p не содержит мотивы RRM и связывается с РНК преимущественно через мотивы «цинковый палец» типа CCCH [308]. В системе *in vitro* оба дрожжевые PABPs по отдельности способны контролировать длину поли(А)-хвостов транскриптов. Роль этих белков в процессе терминации полиаденилирования про-мРНК дрожжей *in vivo* в настоящее время неизвестна [308, 309]. Интересно, что оба белка могут привлекать нуклеазы для деградации поли(А)-хвостов про-мРНК: Pab1p связывается с комплексом деаденилаз PAN, а Nab2p взаимодействует с субъединицей ядерной экзосомы Rrp6p. Заслуживает внимание также тот факт [310], что в отличие от Pab1p, молекулы которых располагаются на поли(А) линейным образом, Nab2p образует комплекс с поли(А) с иной топологией и

механизм защиты поли(А) от нуклеазы из микрококков при связывании с Nab2p более подобен механизму защиты при связывании с PABPN1, чем с Pab1p или PABPC.

Мы предполагаем, что механизм терминации полиаденилирования про-мРНК дрожжей так же, как и в случае высших эукариотов, включает образование олиго(А)-дуплекса, разрушающего комплекс белков, осуществляющих процессивный синтез. Причем в случае дрожжей дуплекс, по-видимому, может образовываться при меньшей длине поли(А) (~ 70 нт), чем в случае высших эукариотов. Функцию протонирования адениновых остатков может выполнять какой-либо из белковых факторов полиаденилирования. Альтернативно, образование внутримолекулярного дуплекса в поли(А)-хвосте про-мРНК дрожжей может играть роль в его защите от деградации, препятствуя действию нуклеаз. Такой дуплекс может быть образован под действием Nab2p, поскольку более высокая организация комплекса поли(А) с этим белком, чем с Pab 1p (и с его аналогом у высших эукариотов PABPC), обнаруженная в работе [310], вероятнее всего, связана именно с образованием внутримолекулярной двойной спирали.

В заключение подчеркнем, что в пользу существования внутримолекулярных двойных спиралей в поли(А) свидетельствует ряд экспериментальных фактов: уменьшение радиуса вращения протонированных молекул поли(А) при низкой ионной силе и низкой концентрации полимера [35], образование «шишкообразных» структур в поли(А) в подкисленной воде, наблюдаемое методом силовой атомной микроскопии [239], образование «сжавшейся» (S) формы поли(dA) [241]. Кроме того, наличие внутримолекулярных олиго(А)-дуплексов хорошо согласуется со свойствами «замороженной» формы поли(А), выявленными в литературе и наших экспериментах. Поэтому, учитывая чрезвычайно высокую эффективность образования «замороженной» формы в длинных молекулах поли(А) и тот факт, что протонирование этого гомополимера может осуществляться за счет белков, нельзя исключать возможность образования внутримолекулярных

двойных спиралей в трактах олиго(А) природных РНК в клетке. А если такую возможность допускать, то можно удовлетворительно объяснить, как происходят процессы, перечисленные в табл. 5.1. Механизм всех этих процессов, как уже упоминалось, до сих пор не установлен.

5.2. Молекулярные модели участия G-квадруплексов РНК в биологических процессах

5.2.1. Структурные модели области димеризации геномной РНК ВИЧ-1.

5.2.1.1. Модели DLS геномных РНК ВИЧ-1_{Mal} и ВИЧ-1_{Lai}. Существует несколько представлений о механизмах димеризации генома вируса ВИЧ-1 (см. параграф 1.3), однако единое мнение по этому вопросу отсутствует. Мы предложили новую структурную модель димеризации геномной РНК ВИЧ-1, которая включает как участки, образованные шпильками DIS, так и четырехцепочечные участки, образованные G-богатыми трактами [311, 312].

Изоляты ВИЧ-1 подразделяются на 4 группы (M, N, O и P). Группа M, ответственная за пандемию СПИДа, подразделяется на 9 субтипов (A, B, C, D, F, G, H, J, K) и десятки рекомбинантных форм, основными из которых являются CRF01_AE (CRF – circulating recombinant form) и CRF02_AG [313]. Субтипы B, C и A являются наиболее распространенными в мире. Мы впервые провели широкомасштабное исследование структуры шпилек DIS в общей сложности для ~1300 изолятов ВИЧ-1 разных субтипов и групп [209, 314, 315]. В соответствии с литературными данными [316], было обнаружено, что в шпильках DIS геномных РНК ВИЧ-1 встречаются только три вида палиндромов (GCGCGC, GUGCAC и GUGCGC). Палиндром GUGCGC встречается редко (в 2 % изолятов); предполагают, что он является переходным между первыми двумя [316].

Мы подразделили все варианты шпилек DIS на пять основных групп по последовательности апикальной области (рис. 5.3). К группе I мы отнесли

шпильку с апикальной областью AAGCGCGCA (палиндром подчеркнут) и стеблем без замен оснований (по отношению к реферативному геному HXB2), а также подварианты этой шпильки с мутациями. Основную шпильку этой группы (без мутаций) мы назвали DIS_{Lai} .

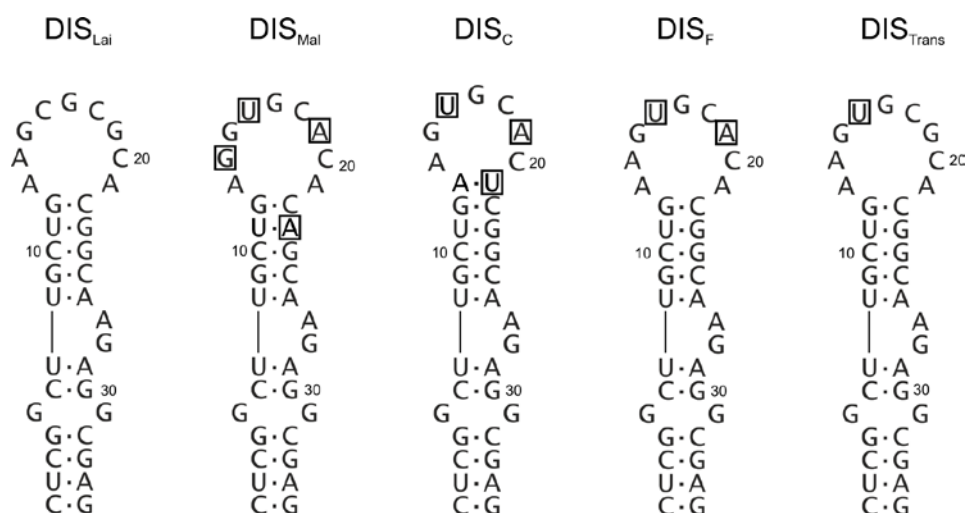


Рис. 5.3. Основные варианты шпилек DIS геномных РНК ВИЧ-1. Замещенные основания (по сравнению с реферативным геномом HXB2) изображены в квадратах

Шпильки DIS группы I встречаются только в изолятах субтипов В и D. К группе II мы отнесли шпильки с апикальной областью AGGUGCACA и мутацией G23A в стебле. Такие шпильки (DIS_{Mal} и подварианты) характерны для A/G-содержащих субтипов. К группе III относятся шпильки с апикальной областью AGUGCAC и дополнительной парой оснований A·U в верхней части стебля (DIS_C и подварианты), которые встречаются в основном в изолятах ВИЧ-1 субтипа С. К группе IV мы отнесли шпильки с апикальной областью AAGUGCACA (DIS_F и подварианты), характерные для изолятов субтипа F и, наконец, к группе V относятся шпильки с переходным палиндромом и апикальной областью AAGUGCGCA (DIS_{Trans} и подварианты), встречающиеся в изолятах субтипа В.

В общей сложности кроме пяти основных вариантов шпильки DIS нами были выявлены еще 13 подвариантов, которые встречались с частотой выше 0,5 % (ошибка секвенирования геномных последовательностей и представления их в GenBank [317]). Ни один из подвариантов, кроме DIS_{Lai}^{21G} с заменой A21G в апикальной петле, ранее в литературе не упоминался. Важно отметить, что частота мутаций в какой-либо позиции шпильки DIS зависит от того, к какой группе относится шпилька. Например, мутации в позиции 23 происходят в DIS_{Lai} , DIS_{Mal} , DIS_C and DIS_F с частотой 36, 3, 3 и 31 % соответственно, т. е. замена оснований в этой позиции допустима в шпильках DIS групп I и IV, но не является толерантной для DIS групп II и III. Мы предполагаем, что мутации в определенной позиции основных вариантов шпильки DIS различным образом влияют на их характеристики, такие как стабильность, степень доступности палиндрома в апикальной области для димеризации, сродство к нуклеокапсидному белку и соответствие между конформациями стебля шпильки и апикальной петли.

На рис. 5.4. показана одна из предполагаемых структур DLS для геномной РНК HIV-1_{Mal}, которая является представителем изолятов ВИЧ-1 A/G-содержащих субтипов. Двухцепочечный участок (260-294 нт) соответствует линейному дуплексу, в который превращается метастабильный димер «целующиеся петли» при созревании вирионов (рис. 1.12). G-квадруплексы образованы двумя молекулами РНК ВИЧ-1, свернутыми в параллельно направленные шпильки. В результате 5'- и 3'-концы геномных РНК также направлены параллельно.

Как видно на рис. 5.4, наряду с G-тетрадами, квадруплексы содержат и другие тетрады. Тетрады G·C·G·C, A·T·A·T и G·T·G·T были найдены в антипараллельных G-квадруплексах ДНК [148, 152, 157]. Возможно, что и G-квадруплексы РНК тоже могут включать тетрады G·C·G·C, A·U·A·U, и G·U·G·U. Другие тетрады – A·A·A·A, C·C·C·C и T·T·T·T — найдены в настоящее время только в параллельных квадруплексах [153, 154, 156], однако модельные построения и исследования методом молекулярной динамики

показали, что они также могут быть включены и в антипараллельные квадруплексы [318].

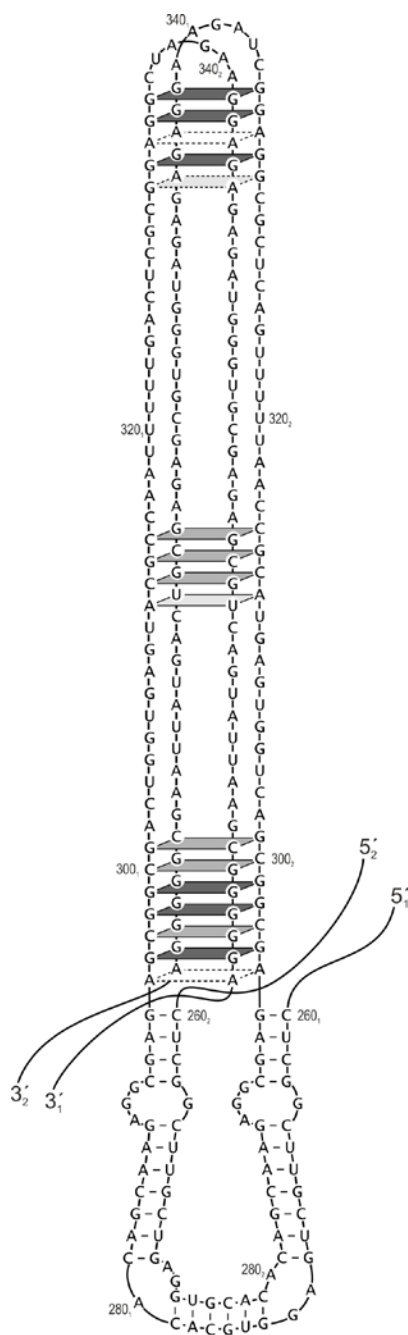


Рис. 5.4. Структурная модель области димеризации геномной РНК ВИЧ-1_{Mal}. Здесь и ниже разные тетрады выделены по-разному, нуклеотиды пронумерованы только на одной из геномных РНК. См. также текст относительно тетрад, обозначенных пунктирной линией

Хотя образование тетрады G·A·G·A допускается теоретически, ее включение в G-квадруплексы представляется маловероятным [319], поскольку она имеет V-подобную структуру и стабильнее, чем две отдельные пары A·G всего лишь на 3,5 ккал/моль. Эта тетрада и тетрада A·A·A·A обозначены на рис. 5.4 пунктиром. Возможность включения смешанных тетрад и тетрад из

других оснований в квадруплексы расширяет количество разнообразных четырехцепочечных структур, которые могут образовывать ДНК и РНК.

Следует отметить, что по данным электронной микроскопии область димеризации мономерных РНК в вирионе охватывает примерно первые 600 нуклеотидов, при этом РНК контактируют в нескольких местах [164, 170]. Кроме того, вне области димеризации они ориентированы параллельно. Как видно на рис. 5.4, наша модель удовлетворяет этим требованиям. Предположение о том, что именно образование квадруплексов обуславливает параллельную ориентацию мономерных РНК в димере, высказывалось ранее в работах [167, 320].

Наличие многочисленных G-богатых трактов, расположенных за шпилькой DIS, и возможность стабилизации G-тетрад тетрадами, образованными разными основаниями, обеспечивает сохранение структуры DLS, состоящей из линейного дуплекса и G-квадруплексов, при различных мутациях. По-видимому, лишь обширные и множественные мутации G-богатых трактов могут предотвратить ее образование.

Мы предполагаем, что квадруплексы образуются во время созревания вирионов либо наряду с образованием линейного дуплекса или сразу же после его образования под действием нуклеокапсидного белка. Увеличение локальной концентрации G-богатых трактов при сближении мономерных РНК вследствие образования димера должно способствовать образованию межмолекулярных квадруплексов. Примечательно, что *in vitro* белок NC_p15 может стимулировать димеризацию фрагмента 311-415 нт РНК ВИЧ-1_{Mal} [321], не содержащего шпильки DIS и образующего, согласно [167], межмолекулярные квадруплексы. Участие нуклеокапсидного белка в образовании G-квадруплексов может быть, например, связано с необходимостью расплести двухспиральные участки, включающие G-богатые тракты.

Мы полагаем, что дополнительная связь между мономерами РНК ВИЧ-1 посредством G-квадруплексов важна для стабилизации димера, образованного

шпильками DIS, и сохранения контакта между двумя геномами в процессе инициации обратной транскрипции, которая начинается с сигнального элемента PBS (рис.1.9), расположенного недалеко от DIS, и происходит в направлении к 5'-концу генома. Затем происходит перенос растущей цепочки ДНК на 3'-конец генома и синтез снова продолжается в направлении к 5'-концу. В результате область димеризации, включающая DIS и G-квадруплексы, служит матрицей для синтеза ДНК в последнюю очередь, обеспечивая контакт между геномными РНК во время всего синтеза минус цепи ДНК.

Вклад в стабилизацию димера и сохранение пространственной близости между геномными РНК, необходимой для рекомбинации, могут вносить также, как предполагают авторы работ [322, 323], дополнительные контакты в отдаленных областях генома. Они показали, что G-богатые последовательности, расположенные вблизи центрального полипуринового тракта, а также на 3' конце генома, могут сворачиваться в квадруплексы, что указывает на их возможную роль в димеризации. Однако эти контакты, в отличие от контакта в области шпильки DIS, будут разрушаться в процессе синтеза минус цепи ДНК.

G-богатые области, расположенные за шпилькой DIS, присутствуют во многих изолятах ВИЧ-1 и в некоторых других ретровирусах. На рис. 5.5. показано расположение G-богатых трактов в геноме ВИЧ-1_{Lai}, представителя изолятов ВИЧ-1 субтипа В. Эти тракты находятся в шпильках SD и Psi, а также в 5'-концевой части гена *gag*. Мы предсказали несколько возможных структурных моделей области димеризации геномной РНК ВИЧ-1_{Lai}, содержащей G-квадруплексы, образованные трактами I-VI и линейный дуплекс, образованный шпильками DIS. Одна из моделей показана на рис. 5.6.

Однако следует отметить, что эксперименты, подтверждающие образование линейного дуплекса посредством «расплетания» стеблей двух шпилек DIS и последующего межмолекулярного спаривания (рис. 1.12), были выполнены только с короткими РНК, например [324].

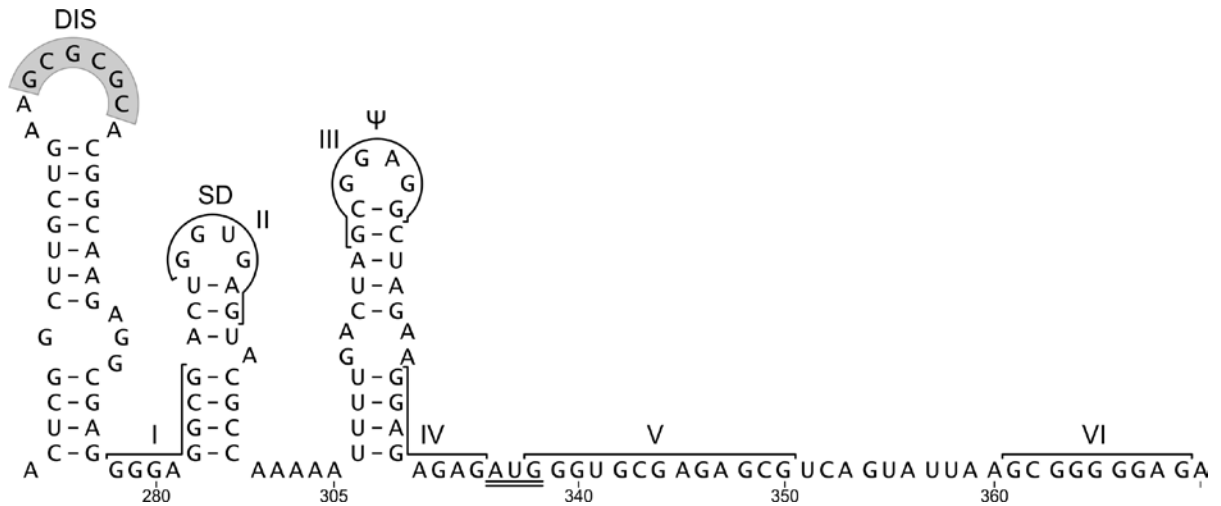


Рис. 5.5. G-богатая область РНК ВИЧ-1_{Lai}, расположенная за шпилькой DIS. Палиндром затенен. G-богатые тракты отмечены и пронумерованы. Кодон инициации трансляции гена *gag* подчеркнут двойной линией

Поскольку в настоящее время образование линейного дуплекса в вирионе пока еще экспериментально не установлено [325], на рис. 5.6. б представлена также модель димеризации, в которой шпильки DIS образуют комплекс «целующиеся петли». Четырехцепочечные домены, как и в случае геномной РНК ВИЧ-1_{Mal}, наряду с G-тетрадами содержат смешанные тетрады. Мы специально не изображали все потенциальные тетрады в представленных моделях, поскольку малоизвестно, в каких комбинациях тетрады из разных оснований могут быть последовательно включены в состав квадруплексов. В частности, на рис. 5.6. не обозначены последовательно расположенные тетрады U·U·U·U, A·A·A·A, G·G·G·G и A·A·A·A в верхней части верхнего G-квадруплекса.

Russel et al. [326] изучали *in vivo* роль отдельных G-богатых трактов в димеризации гРНК ВИЧ-1-1_{BH-10}, вводя мутации в соответствующие последовательности провирусной ДНК. Область, охватывающая шпильки DIS, SD и Psi в гРНК ВИЧ-1_{BH-10}, отличается от таковой в ВИЧ-1_{Lai} только в позиции 281 (рис. 5.5).

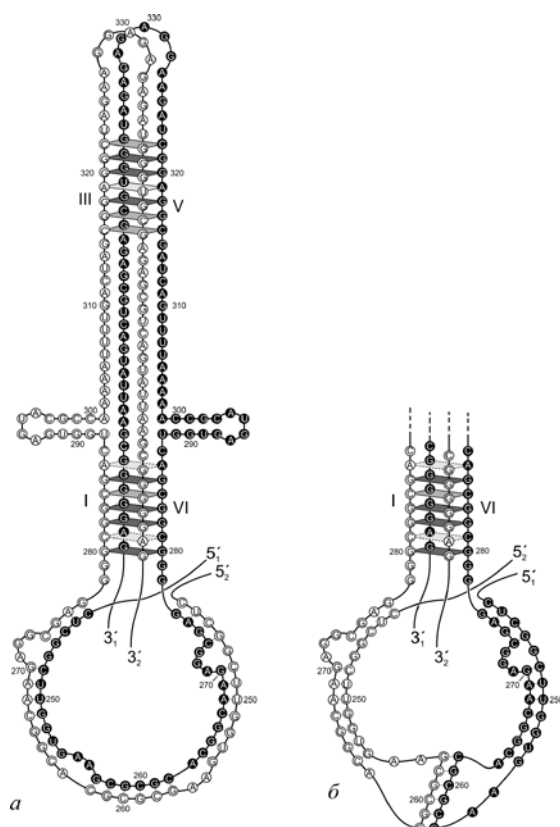


Рис. 5.6. Структурная модель области димеризации геномной РНК ВИЧ-1_{Lai} (а) и альтернативная структура нижней части модели (б)

Авторы показали, что некоторые мутации в шпильке Psi или замещения двух G (нт 333, 335) в GA-богатой области уменьшают процент димеризованных вирусных РНК примерно в 2 раза. С другой стороны, мутации, не нарушающие структуры шпильки Psi (мутации в апикальной петле или компенсаторные мутации в парах оснований в стебле), почти не изменяли степень димеризации генома по сравнению с геномом дикого типа. Также не влияли на димеризацию мутации в двух олиго(G) трактах (нт 363-367 и нт 405-409).

Авторы работы [326] пришли к выводу, что шпилька Psi и GA-богатые участки требуются для эффективной димеризации. Они предположили, исходя из анализа вторичной структуры фрагментов РНК, что мутации в GA-богатых трактах могли влиять на структуру шпильки DIS. Можно также предположить, что эти мутации разрушали дуплекс U5-AUG, необходимый для поддержания структуры димера. Представленная нами модель структуры области димеризации гРНК ВИЧ-1_{Lai} не находится в противоречии с результатами Russell et al. [326]. В частности, мутации в G-богатых трактах III и VI,

введенные в работе [326], не предотвратили полностью возможность образования G-квадруплексов этими трактами. Мутированные тракты могли образовывать G-квадруплексы с четырьмя тетрадами, включая тетрады A·A·A·A и C·C·C·C.

Для того чтобы выяснить вопрос, является ли возможное образование четырехцепочечных доменов в области димеризации общим свойством геномных РНК ВИЧ-1, мы проанализировали последовательности, охватывающие шпильки DIS, SD и Psi, и их вторичную структуру во всех геномах ВИЧ-1, представленных в базе данных GenBank к моменту подготовки публикации (2004 г). В общей сложности были проанализированы соответствующие фрагменты в 350 геномах ВИЧ-1 разных субтипов и рекомбинантов. Вторичную структуру предсказывали с помощью программы «mfold» [205].

Шпилька, подобная шпильке DIS, была найдена в предсказанных структурах почти всех исследованных геномов ВИЧ-1, в некоторых случаях ее структура была незначительно изменена. Шпилька SD оказалась менее консервативной. Она отсутствовала в свернутых структурах приблизительно в 7 % исследованных геномов. Как видно на рис. 5.5, эта шпилька содержит часть G-богатого тракта I и тракт II, в тракт I также входит линкер между шпильками DIS и SD. Линкер довольно консервативен и его длина не превышает 6 нт. В случаях, когда из-за мутаций SD- или DIS-подобные шпильки не образуются, мы подразумеваем под линкером последовательность нуклеотидов между трактами, соответствующими этим шпилькам. Линкер между шпильками SD и Psi значительно более вариабелен. Он состоит в основном из A и U оснований и его длина варьирует от 1 до 11 нт.

Вследствие консервативности SD-подобной шпильки и линкера между шпильками DIS и SD, большинство исследованных геномных РНК содержат G-богатый тракт длиной 3-11 нт (в основном 6-8 нт) в области, расположенной непосредственно за шпилькой DIS (тракт I). Примерно в 70 % изолятов ВИЧ-1 этот тракт представляет собой последовательности GCGGCG, GGGCGGCG

или GCGGCGG, как это следует из таблицы 5.2, в которой представлены наиболее часто встречающиеся последовательности трактов I-VI в гРНК изолятов ВИЧ-1 группы М.

Таблица 5.2

**G-богатые тракты, расположенные за шпилькой DIS, в геномных РНК
ВИЧ-1 группы М**

Тракт	GRS	% геномов, содержащих GRS	Тракт	GRS	% геномов, содержащих GRS
I	gcGGCG	31	V	gggugcgagagcg	96
	gggcGGCG	22		другие	4
	gcGGCGG	15	VI	gcggggg	22
	ggGGCGG	5		gaggggg	18
	ggg, GCG	5		gcggggggag	16
	gggc, GGCGG	3		guggggg	11
	ggGGCG	3		gaggcgg	6
	другие	16		gaggcg	6
II	GGUGAG	98	gagggg	4	
	другие	2	gcggcgg	3	
III	GCGGAGG	93	gcggggggagg	2	
	GAGG	4	ggcg	2	
	другие	3	другие	10	
IV	GGAGagag	96			
	другие	4			

Примечание. Нуклеотиды, напечатанные прописными буквами, локализованы в областях, соответствующим шпилькам SD или Psi.

G-богатые тракты II и III охватывают апикальные петли шпилек SD и Psi соответственно (рис. 5.5). Почти все исследованные геномные РНК содержат

G-богатые тракты в этих областях. Верхняя часть G-богатого тракта IV (рис. 5.5) находится в нижней части стебля шпильки Psi, последовательность которого оказалась высококонсервативной. Тракт IV включает также четыре нуклеотида, расположенных за шпилькой Psi (рис. 5.5). Филогенетический анализ показал, что практически все исследованные геномы ВИЧ-1 содержали тракт IV (таблица 5.2).

G-богатые тракты V и VI локализованы в области генома, транслируемой в белок GAG. Мы обнаружили, что последовательность первых пяти кодонов, включающая тракт V (рис. 5.5), оказалась довольно консервативной. Некоторые изменения в этой области по сравнению с таковой в геноме ВИЧ-1_{Lai} были найдены только в 23 геномах. Однако все измененные последовательности тоже содержали G-богатый тракт. Тракт VI локализован в области, охватывающей кодоны 9-12 (рис 5.5). Эта область значительно более переменная, чем область, прилегающая к старт-кодону. Тем не менее, все гРНК за исключением восьми содержали в этой области G-богатый тракт, который примерно в 70 % всех исследованных геномах имел последовательность GNGGGGG, где N представляет собой C, A или U (таблица 5.2). Таким образом, G-богатые тракты I-VI были найдены в большинстве исследованных геномов ВИЧ-1. Поскольку их последовательности переменны, мы также выяснили вопрос о том, могут ли структуры области димеризации, показанные на рис. 5.4 и 5.6, образовываться во всех исследованных геномах. Оказалось, что структуры, в которых тракты I и III взаимодействуют соответственно с трактами VI и V, образуя несколько тетрад (рис 5.6), могут быть предсказаны почти для всех изолятов ВИЧ-1 группы M.

Модели, в которых нижний квадруплекс также образуется из трактов I и VI, но верхний квадруплекс образуется из трактов III and IV (рис 5.4) тоже могут быть предсказаны для большинства изолятов. Те гРНК, которые не могут быть свернуты в структуры, показанные на рис. 5.4 и 5.6, все же способны образовывать структуры области димеризации, содержащие

дуплексы и квадруплексы. В этих структурах G-богатые тракты объединяются в квадруплексы в других комбинациях.

Примеры моделей DLS, показанных на рис. 5.4 и 5.6, относятся к антипараллельным квадруплексам. На рис. 5.7 приведен пример модели с параллельными квадруплексами.

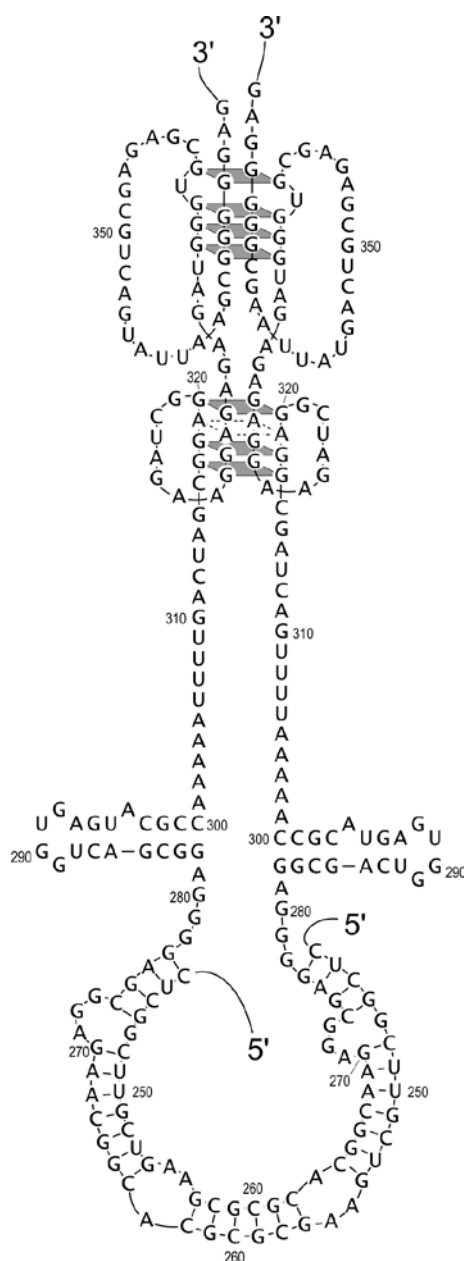


Рис. 5.7. Альтернативная структурная модель области димеризации геномной РНК ВИЧ-1_{Lai}

Эта модель представлена в связи с недавними исследованиями, показавшими, что в отличие от полиморфных G-квадруплексов ДНК, G-богатые последовательности РНК сворачиваются преимущественно в

параллельные квадруплексы, хотя, как отмечают Bugaut & Balasubramanian [144], эти данные получены преимущественно методом КД для довольно коротких синтетических РНК с близко расположенными G-трактами. Добавим также, что исследованные квадруплексы РНК не включали тетрады, состоящих из других оснований или смешанных тетрад.

Наличие множества разнообразных G-богатых последовательностей, расположенных на небольшом расстоянии за шпилькой DIS, позволяет представлять модели DLS с верхней частью, содержащей как антипараллельные, так и параллельные G-квадруплексы. Множественные G-богатые тракты, расположенные за палиндромными последовательностями, участвующими в димеризации, встречаются также и в геномных РНК других ретровирусов, например, в гРНК вируса саркомы мышей Молони [327]. G-квадруплексы, образованные этими трактами, также могут участвовать в димеризации геномной РНК.

5.2.1.2. Возможное участие шпильки TAR в димеризации геномной РНК ВИЧ-1. На рисунках 5.4, 5.6 и 5.7 показана область димеризации гРНК ВИЧ-1, расположенная за доменом PBS. Однако контакты между геномными РНК могут осуществляться в области, расположенной до шпильки DIS, в частности, посредством димеризации шпилек TAR [178-180] (см. параграф 1.3). Мы провели филогенетическое исследование структуры шпильки TAR в 1679 изолятов ВИЧ-1 разных субтипов и рекомбинантов [328] и обнаружили, что последовательность палиндрома GGGAGCUCUC, расположенного в верхней части шпильки TAR, является высококонсервативной (таблица 5.3).

Редкие мутации (G32A и G33A) встречаются лишь в первом и втором положении палиндрома с частотой 3 % и 5 % соответственно. Мутация G32A разрушает только две концевые пары в дуплексе палиндрома длиной 10 пар оснований, а мутация G33A (специфическая для изолятов ВИЧ-1 субтипа D) приводит к замене пары G-U на A-U. Консервативной является также вторичная структура шпильки TAR, несмотря на высокую частоту мутаций в ряде позиций, в частности в позициях 11, 13, 48 и др.

Таблица 5.3

**Частота мутаций в каждой позиции шпильки TAR для 1679 изолятов
ВИЧ-1**

Позиция	% изолятов	Позиция	% изолятов	Позиция	% изолятов
1	4,0	21	1,7	44-45	0,2
2	1,0	22	10,3	45	2,3
3	0,2	23	0,1	45-46	0,1
4	0,1	24	21,4	46	1,1
5	0,4	25	25,8	46-47	0,1
6	0,1	26	1,1	47	6,8
7	0,7	27	0,1	48	69,1
8	0,4	28	0,9	48-49	0,1
8-9	0,2	29	0,5	49	4,9
9	1,6	30	0,2	50	25,1
10	0,4	31	17,3	50-51	0,1
10-11	0,1	32	2,9	51	19,6
11	58,0	33	5,4	51-52	0,1
12	0,8	34	0,9	52	5,8
12-13	0,3	34-35	0,1	52-53	0,8
13	34,5	35	0,5	53	5,7
13-14	0,1	36	0,7	53a54	0,4
14	1,6	37	0,3	54	1,8
15	5,2	38	0,2	54-55	0,1
15-16	0,1	39	0,4	55	0,9
16	1,5	40	0,2	56	0,5
17	1,1	41	0,3	56-57	0,1
18	0,4	42	0,2	57	0,1
19	0,3	43	0,5	59	0,6
20	0,3	44	8,3		

Примечание. Позиции бальджа (23-25) и палиндрома GGGAGCUCUC (32-41) выделены.

Варианты шпильки TAR, встречающиеся с частотой $\geq 5\%$, (определенные нами как основные варианты), показаны на рис. 5.8.

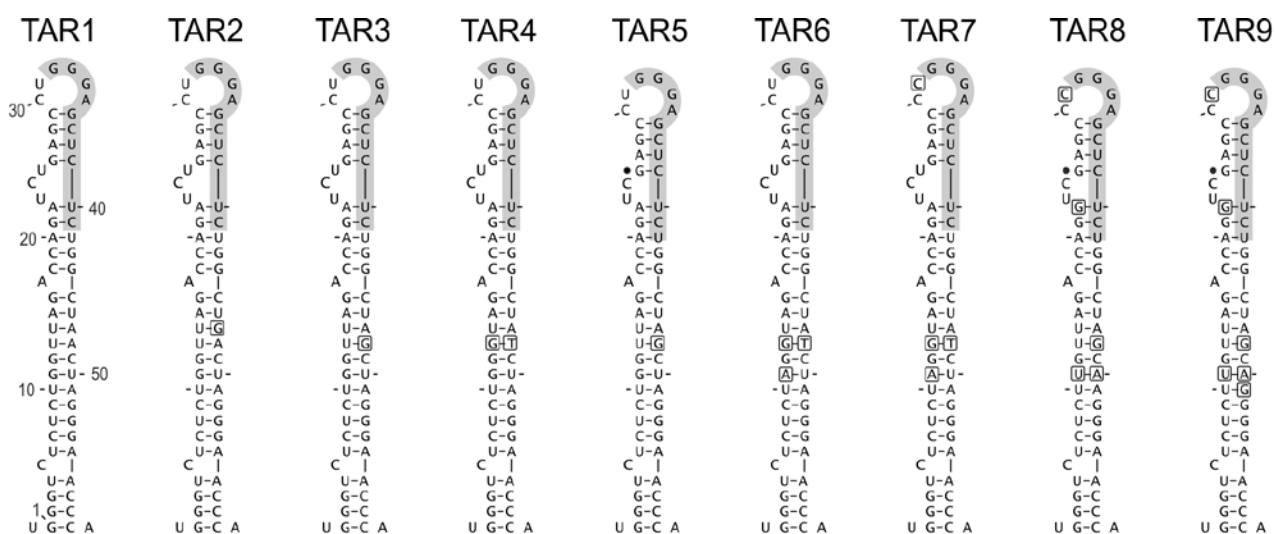


Рис. 5.8. Основные варианты TAR про-мРНК ВИЧ-1 разных субтипов. Замещенные основания (по сравнению с реферативным геномом НХВ2) изображены в квадратах, делеции отмечены черным кружком. Палиндромный тракт GGGAGCUCUC затенен. *TAR1-TAR3* – субтип В, *TAR4* – субтип С, *TAR5* – субтип А, *TAR6* и *TAR7* – субтип С, *TAR8* и *TAR9* – субтип 01_АЕ

Основные варианты встречаются всего лишь в 35 % всех исследованных нами изолятов ВИЧ-1. Все они имеют в 5'-цепи стебля по два бальджа из одного нуклеотида и бальдж из двух-трех нуклеотидов в верхней части. В сумме подобную структуру имеют шпильки TAR в 60 % изолятов. Еще 35 % изолятов ВИЧ-1 имеют шпильки TAR с небольшими изменениями в стебле, в частности, содержат внутренние петли вместо бальджей из одного нуклеотида. В результате, несмотря на высокую гетерогенность последовательности шпильки, подавляющее большинство изолятов ВИЧ-1 содержит шпильки TAR с подобной вторичной структурой, палиндромом GGGAGCUCUC и бальджем из двух-трех нуклеотидов, локализованным вблизи палиндрома.

В параграфе 1.3. сообщалось о том, что вопрос об участии палиндрома GGGAGCUCUC в шпильке TAR в димеризации ВИЧ-1 является спорным и по мнению Jalalirad et al. [180] основную роль в этом процессе играет бальдж в стебле длиной 3 нт. Мы предположили, что один из многочисленных дестабилизирующих белков, специфически связывающихся с бальджем шпильки TAR, например Vif (viral infectivity factor) [329], способствует расплетанию верхней части шпильки и высвобождению нижней части палиндрома из двойной спирали, тем самым облегчая образование дуплекса «целующиеся петли» со шпилькой TAR второго генома. Делеция бальджа будет препятствовать специфическому взаимодействию дестабилизирующего белка (или белков) со шпилькой TAR и соответственно образованию димера, поэтому наличие бальджа и будет главным фактором в димеризации генома ВИЧ-1 с участием шпильки TAR, как это продемонстрировано в работе [180]. Частичное экспонирование палиндрома TAR в апикальной петле, в отличие от полного экспонирования такового в шпильке DIS, может быть связано с тем, что участие TAR в образовании контакта между геномами необходимо только на определенной стадии репликационного цикла вируса и является регулируемым процессом.

5.2.2. Структурная модель поли(А)-сайта про-мРНК вируса SV40 L

Как уже упоминалось в параграфе 1.2.4., поли(А)-сайт про-мРНК SV40 L содержит «нижний» G-богатый вспомогательный элемент (рис. 5.9), с которым связывается белок гяРНП Н/Н', стимулируя реакцию полиаденилирования этого транскрипта. Hans & Alwine [330] определили вторичную структуру поли(А)-сайта про-мРНК SV40 L с помощью ферментативного зондирования. Они показали, что большая часть основного U-богатого «нижнего» элемента и часть гексамера AAUAAA вовлечены во вторичную структуру, в то время как G-богатый AuhDSE находится в основном в одноцепочечном состоянии.



Рис. 5.9. Последовательность сайта полиаденилирования про-мРНК SV40 L, расположенная непосредственно за местом расщепления транскрипта. Одной линией подчеркнут DSE, двойной линией – G-богатый AuxDSE. Пунктиром указано место связывания CstF

Замещая части «нижней» области поли(А)-сайта другими последовательностями они выявили, что наличие вторичной структуры в этой области коррелирует с эффективностью процесса полиаденилирования. Механизм корреляции остался неизвестным. Авторы предположили, что вторичная структура может способствовать взаимодействию CstF с DSE. Схема вторичной структуры поли(А)-сайта про-мРНК SV40 L не была представлена в работе [330] из-за расхождения между структурой, предсказанной с помощью имеющихся в то время компьютерных программ, и экспериментальными данными нуклеазного зондирования.

Мы определили вторичную структуру поли(А)-сайта про-мРНК SV40 L [72] с помощью программы «mfold» (см. параграф 2.4), и сравнили предсказанные структуры с экспериментальными результатами Hans & Alwine [330]. На рис. 5.10 показана вторичная структура фрагмента 2653-2731 нт, охватывающего гексамер AAUAAA, U-богатый основной «нижний» сигнал (URE) и GRS. Структура содержит две шпильки, первая из них включает гексамер AAUAAA, а вторая URE и GRS. На рисунке также показаны менее оптимальные варианты шпилек I и II (I' и II') с бóльшей свободной энергией.

Hans & Alwine [330] использовали РНКазу V₁, неспецифически расщепляющую двухцепочечные РНК, а также РНКазы, расщепляющие РНК по местам локализации неспаренных оснований гуанина (РНКазы T1) и аденина и урацила (РНКазы PhyM). Данные ферментативного зондирования, приведенные в работе [330], находятся в соответствии с нашей моделью

вторичной структуры, представленной на рис. 5.10, с учетом того, что варианты шпилек I и II, находятся в термодинамическом равновесии. В качестве примера на рис. 5.10 указаны все места, на которых происходит более или менее эффективное расщепления нуклеазой T1; размер треугольников, которыми отмечены эти места, коррелирует с интенсивностью расщепления.

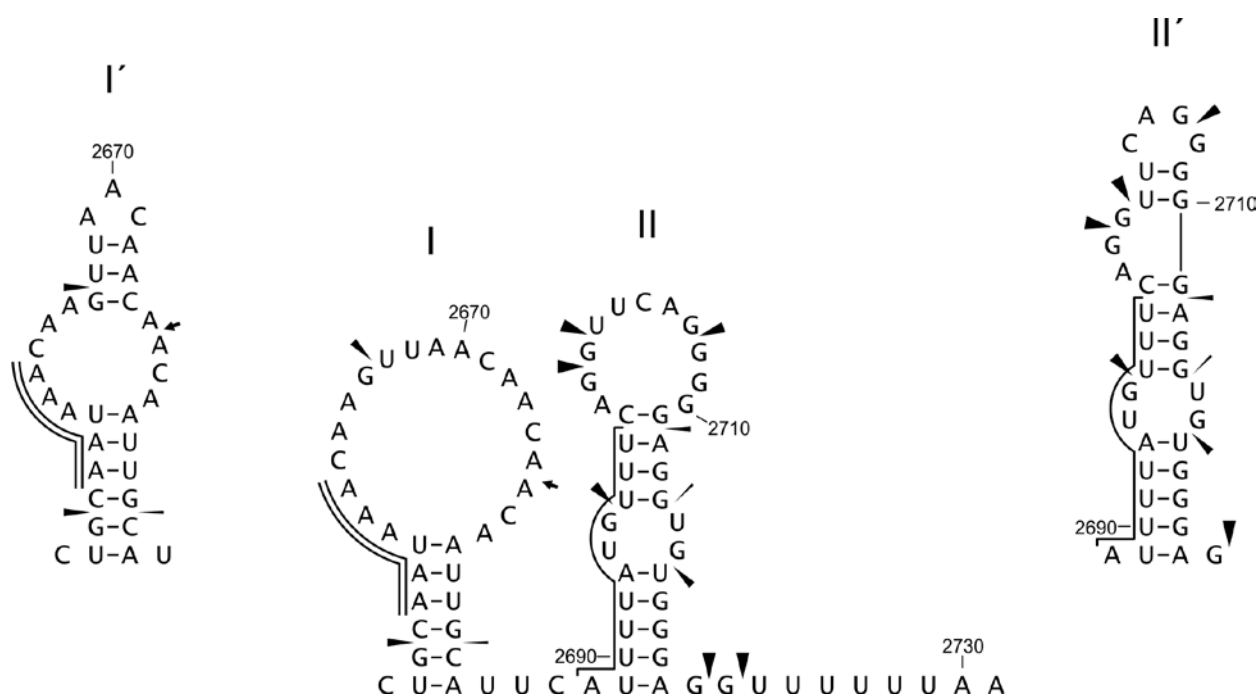


Рис. 5.10. Вторичная структура основного сайта полиаденилирования про-мРНК SV40 L. Гексамер AAUAAA и URE отмечены линиями. Место расщепления и присоединения поли(А)-хвоста отмечено стрелкой. Места, подвергающиеся воздействию РНКазы T1, обозначены треугольниками. Термодинамические стабильности шпилек I, I', II и II' равны -3,1, -2,5, -4,3 и -2,4 ккал/моль соответственно

Единственное несоответствие между экспериментальными данными и предсказанной вторичной структурой относится к участку 2711-2714 нт в 3'-цепи стебля шпильки II, который нечувствителен к РНКазе V₁. Однако нечувствительность некоторых фрагментов РНК к этой нуклеазе не означает, что они находятся в одноцепочечной форме. Согласно [331], РНКазы V₁ узнает

сегмент полинуклеотидной цепи длиной 4-6 нт, находящийся примерно в спиральной конформации и необязательно образующий пары оснований. Конформация сегмента GAGG в верхней части шпильки II может быть субоптимальной по отношению к нуклеазе V_1 , в отличие от конформации комплементарного сегмента UUUC. В соответствии с нашей моделью сегмент GAGG плохо расщепляется обеими нуклеазами, специфическими для одноцепочечных областей. Таким образом, наши предсказания вторичной структуры поли(А)-сайта про-мРНК SV40 L и экспериментальные результаты Hans & Alwine [330] позволяют предположить, что URE и GRS могут присутствовать соответственно в 5' и 3' цепях одной и той же шпильки (шпилька II).

Следует отметить, что полный G-богатый участок в «нижней» области про-мРНК SV40 L GGGGGAGGUGUGGGAGG длиннее на три нуклеотида, чем элемент GRS с которым связывается белок гяРНП Н/Н'. Этот фрагмент и GRS имеют потенциал для сворачивания во внутримолекулярные квадруплексы с двумя тетрадами. Известно, что гяРНП белок Н специфически связывается с поли(G) при концентрации NaCl, равной 2 М [332]. Поскольку в этих условиях гомополинуклеотид находится полностью в четырехцепочечной форме (раздел 4.4.), это означает что, на самом деле, белок специфически связывается с G-квадруплексами. Мы предположили, что белки Н и Н' также могут специфически связываться с квадруплексами, образуемыми G-богатыми трактами клеточных про-мРНК.

Таким образом, G-богатые тракты могут служить вспомогательными элементами поли(А)-сайтов про-мРНК и выполнять все функции, присущие этим элементам (см. параграф 1.2.1.), а именно: влиять на сборку комплекса расщепления и полиаденилирования, предотвращать пространственное смещение основных факторов полиаденилирования с места сборки, а также играть структурную роль, обеспечивая такое сворачивания сайта полиаденилирования, при котором основные сигнальные элементы будут доступны для взаимодействия с соответствующими белками. Все эти функции

может выполнять G-квадруплекс, образуемый GRS поли(А)-сайта про-мРНК SV40 L. В частности, в случае образования G-квадруплекса, основной сайт полиаденилирования сворачивается в структуру, представленную на рис. 5.11.

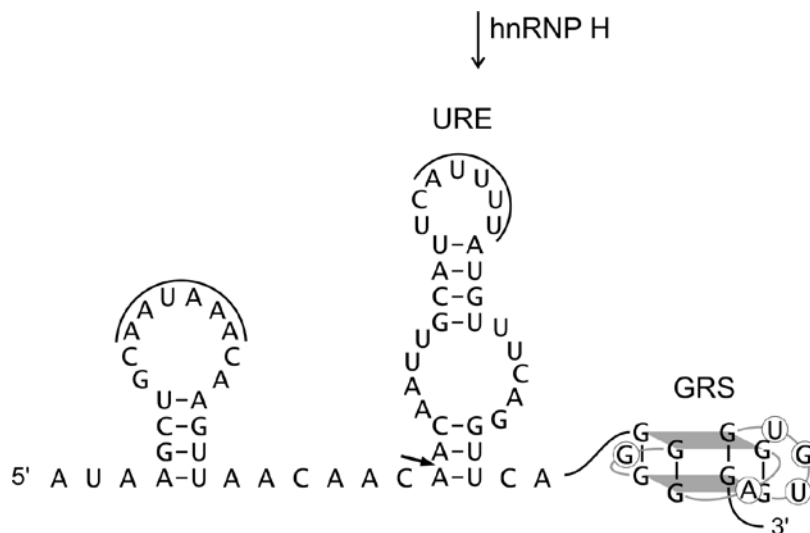


Рис. 5.11. Вторичная структура основного поли(А)-сайта про-мРНК SV40 L в случае сворачивания GRS в G-квадруплекс. Гексамер AAUAAA и URE отмечены линиями. Показан один из потенциальных G-квадруплексов с наименьшей общей длиной петель

На рис 5.11 показана субоптимальная структура, близкая по энергии к оптимальной структуре. В оптимальной структуре URE экспонирован, а гексамер AAUAAA частично вовлечен в образование пар оснований. Таким образом, при образовании G-квадруплекса гексамер AAUAAA и URE в поли(А)-сайте про-мРНК SV40 L экспонированы и доступны для взаимодействия с CPSF и CstF.

Способность GRS поли(А) сайта про-мРНК SV40 L образовывать G-квадруплексы была позднее экспериментально подтверждена в работе [8] методом электрофореза, однако детальная структура квадруплекса не была представлена. В этой работе также было показано, что белок гяРПН Н/Н' может связываться с G-квадруплексом из про-мРНК SV40 L, а также с G-

квадруплексом из про-мРНК р53. Не исключено, что белки гяРНП Н и Н' могут индуцировать превращение G-богатых участков в квадруплексы и/или стабилизировать образующиеся четырехцепочечные структуры.

5.2.3. Структурная модель комплексного поли(А)-сайта про-мРНК ВИЧ-1

5.2.3.1. Вторичная структура основного поли(А)-сайта про-мРНК ВИЧ-1. Как уже сообщалось в параграфе 1.2.2.3., в литературе отсутствуют модели 3'-нетранслируемой области про-мРНК ВИЧ-1 и расположенного в ней комплексного поли(А)-сайта. Известно только, что в 3'-НТО, так же как и в 5'-НТО, могут образовываться шпильки TAR и polyA (рис.1.8). В соответствии с нашим описанием области основного «нижнего» сигнала поли(А)-сайтов про-мРНК высших эукариотов и их вирусов (параграф 1.2.1.), мы предположили, что тракты UGUGU и GUUGUGU, расположенные на расстоянии 6 и 19 нт от места расщепления транскрипта, являются «нижними» элементами поли(А)-сайта про-мРНК ВИЧ-1. Второй тракт состоит из двух перекрывающихся U/GU-пентамеров (GUUGU и UGUGU).

Предсказав вторичную структуру основного поли(А)-сайта 414 про-мРНК ВИЧ-1 разных субтипов [333], мы выявили, что в большинстве случаев эта область сворачивается в три шпильки: TAR, polyA и шпилька длиной 26 нт, в апикальной петле которой находится второй U/GU-богатый тракт. Эту шпильку мы назвали шпилькой DSE. Шпилька TAR была включена в исследуемый участок по двум причинам. Во-первых, она пространственно сближает USE и основной поли(А)-сайт [334], а во-вторых, мутации, дестабилизирующие эту шпильку, могут приводить к снижению эффективности полиаденилирования про-мРНК ВИЧ-1 вследствие удлинения и стабилизации соседней шпильки polyA [335].

Позже мы предсказали вторичную структуру основного USE, полного (комплексного) поли(А)-сайта и почти полного 3'-НТО про-мРНК для более 1400 изолятов ВИЧ-1 разных субтипов и рекомбинантов [98].

Преимущественное сворачивание последовательности основного поли(А)-сайта в шпильки TAR, polyA и DSE наблюдалось и в этом случае, для значительно увеличенного количества исследованных геномов ВИЧ-1. Вторичная структура основного поли(А)-сайта про-мРНК ВИЧ-1, на примере изолята НХВ2 субтипа В показана на рис. 5.12.

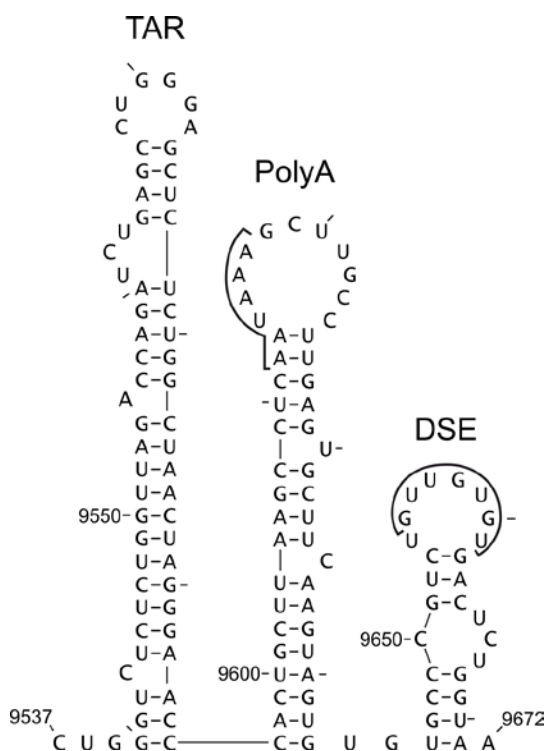


Рис. 5.12. Локальная вторичная структура основного поли(А)-сайта про-мРНК ВИЧ-1 НХВ2 (реферативный геном). Указаны гексамер AAUAAA и U/GU-богатый «нижний» элемент. Здесь и далее нумерация нуклеотидов в соответствии с нумерацией в геноме НХВ2 (номер K03455 в GenBank)

В параграфе 5.2.1.2 мы указали, что шпильки TAR в подавляющем большинстве исследованных нами геномов ВИЧ-1 имеют похожую вторичную структуру, несмотря на чрезвычайно высокое разнообразие их последовательностей. Это наблюдение справедливо также и для остальных структурных элементов поли(А)-сайта про-мРНК ВИЧ-1. На рисунке 5.13 показаны найденные нами основные варианты шпилек polyA, которые встречаются с частотой более 5 % [336]. Последовательности шпилек polyA и DSE значительно более консервативны, чем шпильки TAR. Примерно 70 % всех исследованных геномных РНК ВИЧ-1 содержат один из 7 основных вариантов шпильки polyA и 79 % содержат один из 7 основных вариантов

шпильки DSE. Все основные варианты шпильки polyA имеют идентичную верхнюю часть с гексамером AAUAAA, который частично вовлечен в образование пар оснований, а их стебли содержат различные небольшие дефекты (бальджи и внутренние петли, рис. 5.13). Только в 4 % исследованных геномов мутации ведут к значительным изменениям структуры шпильки polyA.

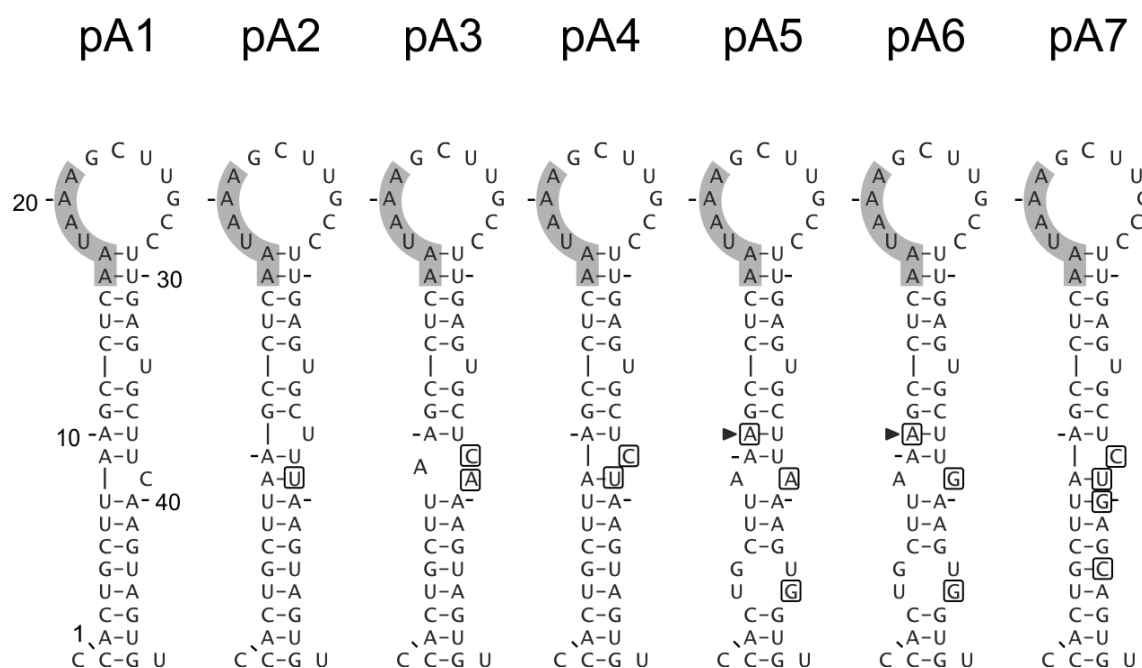


Рис. 5.13. Основные варианты шпилек polyA про-мРНК ВИЧ-1 разных субтипов. Здесь и далее замещенные основания (по сравнению с реферативным геномом HXB2) изображены в квадратах, вставки отмечены треугольником, а делеции (на других рисунках) отмечены черным кружком. Гексамер AAUAAA выделен серым цветом. *pA1* – в основном субтипы В, А; *pA2* – субтипы А, В, С; *pA3* – в основном субтип В; *pA4* – в основном субтип С; *pA5* и *pA6* – субтип 01_АЕ; *pA7* – субтип С

На рисунке 5.14 показаны основные варианты шпилек DSE, которые встречаются с частотой более 5 %. В этих шпильках наиболее вариабельной областью является апикальная петля. Наиболее частые мутации встречаются в

позициях 12, 13, 15 и 16. Комбинация этих мутаций, специфическая для изолятов ВИЧ-1 субтипа 01_AE, не разрушает U/GU-богатый DSE (рис. 5.14, шпильки *DSE6* и *DSE7*). Она приводит лишь к замене тракта GUUGUGU, на сигнальный пентамер GUGUU. В шпильке *DSE5* вследствие мутации образуется тракт GUUGUUU, состоящий из трех перекрывающихся сигнальных пентамеров GUUGU, UUGUU и UGUUU. Последние два пентамера представляют собой URE типа «четыре U из пяти» (см. параграф 1.2.1).

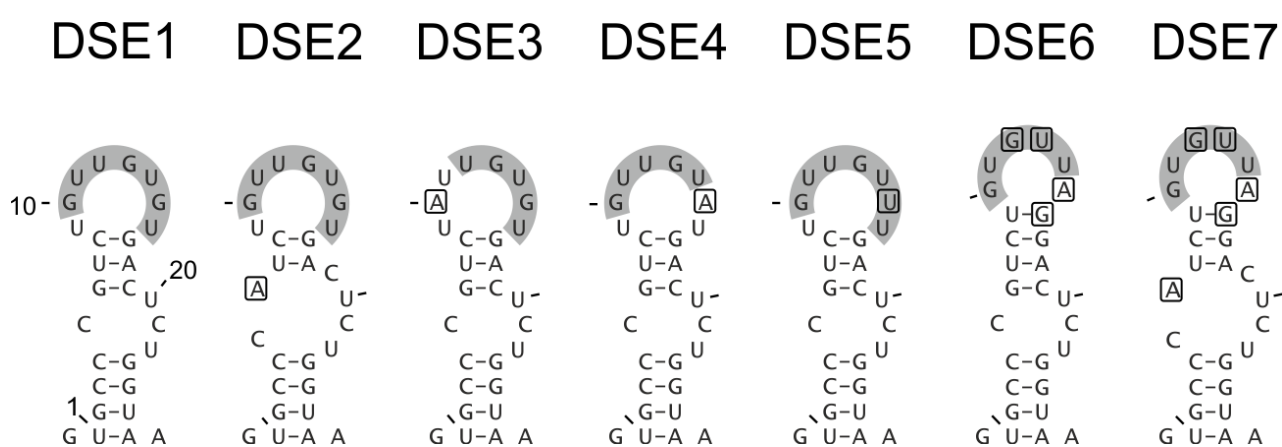


Рис. 5.14. Основные варианты шпилек DSE про-мРНК ВИЧ-1 разных субтипов. U/GU-богатый DSE выделен серым цветом. *DSE1* – субтипы А, В, С; *DSE2* – в основном субтипы В, С; *DSE3* – субтипы С и В; *DSE4* и *DSE5* – в основном субтип А; *DSE6* и *DSE7* – субтип 01_AE

Таким образом, в большинстве изолятов ВИЧ-1 мутации лишь незначительно влияют на вторичную структуру элементов, образующих основной сайт полиаденилирования. Также они мало влияют на вторичную структуру и самого поли(А)-сайта. Частота образования локальной структуры, состоящей из шпилек TAR, polyA и DSE, в геномах ВИЧ-1 разных субтипов приведена в таблице 5.4. Данные, относящиеся к изолятам ВИЧ-1 субтипов А, В, С, 01_AE и 02_AG, составляющих большие группы, приведены в верхних строках таблицы, а в нижней строке приведены суммарные данные для

изолятов субтипов D, F, G и других, а также для 100 изолятов, субтип которых неизвестен.

Таблица 5.4

Частота образования локальных структур фрагментов комплексного поли(А)-сайта в изолятах ВИЧ-1 (%)

Субтип	Количество изолятов ВИЧ-1	TAR, polyA, DSE	USE, TAR, polyA	домен USE, TAR, polyA	
		(9537-9672 нт)	(9516-9644 нт)	(9493-9644 нт)	
		$\Delta\Delta G \leq 2$ ккал/моль	$\Delta\Delta G \leq 2$ ккал/моль	$\Delta\Delta G \leq 2$ ккал/моль	$\Delta\Delta G \leq 3$ ккал/моль
A	137-139	96,4	95,0	89,9	91,4
B	548-568	97,3	97,9	62,4	78,5
C	269-273	94,8	91,6	82,0	85,7
01_AE	89-90	82,0	88,9	71,1	75,6
02_AG	88-95	95,8	97,9	73,9	83,0
Others	323-327	97,5	92,3	84,3	86,4

Примечание. Количество изолятов представлено в виде интервала, так как оно незначительно различается в разных колонках

Как следует из таблицы 5.4, локальная структура, состоящая из шпилек TAR, polyA и DSE, со свободной энергией, которая равна оптимальной или отличается от нее не более чем на 2 ккал/моль, образуется в 82-97 % всех свернутых фрагментов геномов ВИЧ-1.

5.2.3.2. Вторичная структура основного вспомогательного элемента комплексного поли(А)-сайта про-мРНК ВИЧ-1. Для того чтобы определить локальную вторичную структуру основного USE поли(А)-сайта про-мРНК ВИЧ-1 (тракт CAGCUGCUUUUUGCCUGU), мы свернули фрагменты исследуемых геномов, охватывающие этот тракт и последовательности

шпилек TAR и polyA. Было обнаружено, что в большинстве случаев USE (включая еще один нуклеотид, непосредственно прилегающий к его 5'-концу) сворачивается в шпильку длиной 19 нт, апикулярная петля которого состоит из U-богатой части элемента USE. Мы назвали эту шпильку шпилькой USE. На рис. 5.15 для примера показана локальная вторичная структура области, включающей шпильки USE, TAR и polyA, для про-мРНК ВИЧ-1 НХВ2 субтипа В.

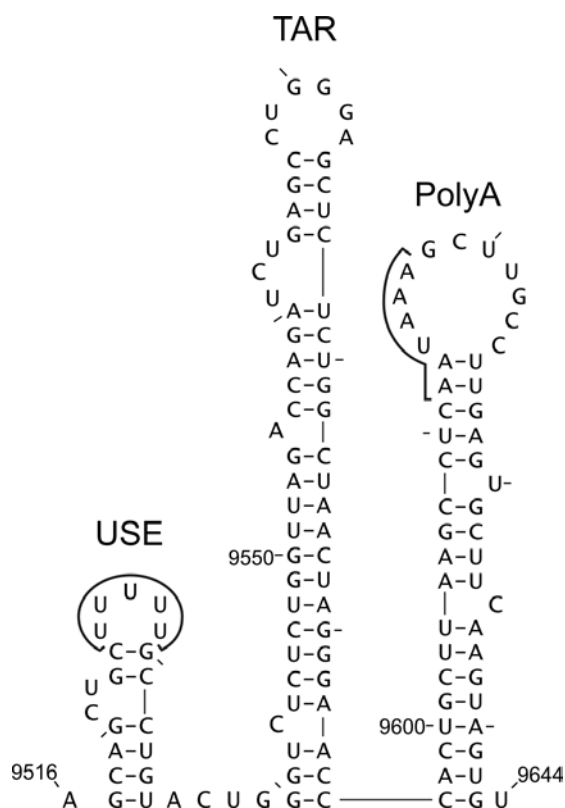


Рис. 5.15. Локальная вторичная структура области, включающей шпильки USE, TAR и polyA, для про-мРНК ВИЧ-1 НХВ2 субтипа В. U-богатый тракт USE и гексамер AAUAAA отмечены линиями

Частота образования локальной структуры, состоящей из шпилек USE, TAR и polyA, в геномах ВИЧ-1 разных субтипов приведена в таблице 5.4. Локальная структура, состоящая из этих шпилек, со свободной энергией, которая равна оптимальной или отличается от нее не более, чем на 2 ккал/моль, образуется в 89-98 % всех свернутых фрагментов геномов ВИЧ-1.

Следует отметить, что вторичная структура определенного фрагмента РНК зависит от того, в состав какого большего фрагмента он входит. В длинной РНК необходимо выявлять минимальный по размеру участок

(включающий исследуемый фрагмент), сворачивание которого будет происходить автономно. Оказалось, что в случае комплексного поли(А)-сайта про-мРНК ВИЧ-1 можно предположительно определить границы такого самодостаточного домена. В данном случае 5'-границей независимо сворачивающегося домена может служить G-богатый участок с высоким потенциалом для образования G-квадруплексов. Такой участок, как видно из рисунка 1.10, приведенного в параграфе 1.2.2, образуют минорные вспомогательные элементы поли(А)-сайта про-мРНК ВИЧ-1. Данная G-богатая область, как правило, содержит девять G-повторов. Первые два из них расположены относительно далеко от соседних олиго(G)-трактов, а остальные семь G-повторов находятся на близком расстоянии друг от друга (рис. 5.16, верхняя часть).

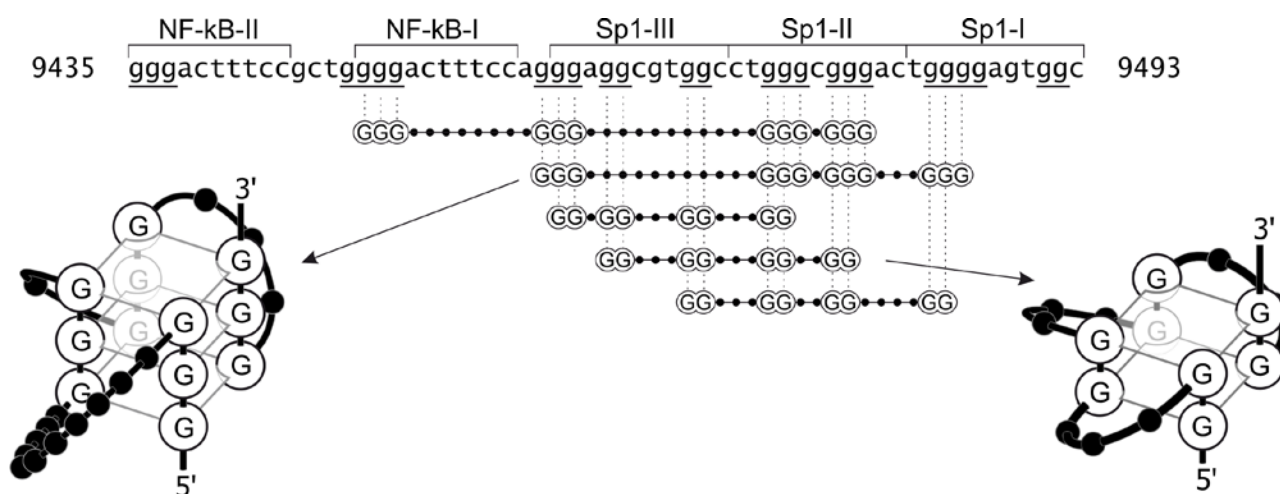


Рис. 5.16. G-богатая область с потенциалом для образования G-квадруплексов в про-мРНК НХВ2 ВИЧ-1. G-повторы подчеркнуты. Местоположение мест связывания для белков NF-кВ и Sp1 [337, 338] указано скобками. Показаны некоторые тракты с потенциалом для образования внутримолекулярных G-квадруплексов и приведены примеры квадруплексов с двумя или тремя G-тетрадами

В соответствии с литературными данными, изложенными в параграфе 1.2.5, в этой области может быть образован ряд G-квадруплексов, включающих разные комбинации прилегающих друг к другу олиго(G)-трактов, например, 12 различных G-квадруплексов с двумя тетрадами и общей длиной петель ≤ 9 нт, или шесть G-квадруплексов с тремя тетрадами и длиной отдельных петель, не превышающей 11 нт. На рис. 5.16 (нижняя часть) показаны некоторые G-богатые тракты, способные образовывать G-квадруплексы, а также гипотетические схемы квадруплексов с двумя и тремя G-тетрадами.

Для примера G-квадруплекс с двумя G-тетрадами, содержащий последний G-повтор в GRS, включен в структурные модели фрагментов 3'-НТО про-мРНК ВИЧ-1. Общая длина петель в этом случае равна 10 нт, что превышает условно принятую нами длину 9 нт для стабильных G-квадруплексов с двумя G-тетрадами (см. параграф 2.4), но, поскольку G-квадруплексы в ядре могут быть стабилизированы белком гЯРПП Н, нельзя исключить участие последнего G-повтора в GRS в образовании четырехцепочечной структуры.

G-богатая область, расположенная в 3'-НТО про-мРНК ВИЧ-1, довольно консервативна, поскольку она состоит из мест связывания для факторов транскрипции Sp1 в промоторной области провирусной ДНК ВИЧ-1 и для NF-κB – в энхансерной области (рис. 5.16). Провирусные ДНК ВИЧ-1 всех субтипов содержат по три места связывания для Sp1 и, по крайней мере, одно место для связывания NF-κB [338]. Наши филогенетические исследования показали, что мутации в G-богатой области могут лишь незначительно уменьшать количество G-повторов или предотвращать образование индивидуальных G-квадруплексов, но в любом случае, оставшиеся G-повторы способны образовывать различные G-квадруплексы. Мы предполагаем, что с G-квадруплексами в G-богатой области про-мРНК ВИЧ-1, взаимодействует ядерный белок гЯРПП Н, стимулируя реакцию полиаденилирования

транскрипта. Интересно, что о взаимодействии белка гЯРПП Н с областью про-мРНК ВИЧ-1, содержащей G-богатый минорный USE, сообщалось ранее в работе [339].

Учитывая высокую вероятность образования G-квадруплексов в GRS про-мРНК ВИЧ-1, мы предположили, что во время транскрипции область генома, расположенная за этой областью, будет сворачиваться независимо, и выбрали для предсказания вторичной структуры комплексного поли(А)-сайта участок транскрипта от конца GRS до конца шпильки DSE (9493-9672 нт). На первой стадии мы проанализировали локальную вторичную структуру основного USE в контексте «верхней» части поли(А)-сайта, от последнего G-повтора до конца шпильки polyA (9493-9644 нт).

В большинстве случаев этот фрагмент сворачивался с образованием трех структурных элементов: укороченной шпильки TAR, шпильки polyA и расположенного перед шпилькой TAR домена, содержащего в верхней части две короткие шпильки. Мы назвали этот домен доменом USE. В качестве примера на рис. 5.17 представлена вторичная структура области 9493-9644 нт, охватывающей домен USE, шпильку TAR и шпильку polyA, в про-мРНК AD8 ВИЧ-1 субтипа В. В области генома ВИЧ-1 (9493-9719 нт), исследуемой в данной работе, последовательности про-мРНК ВИЧ-1 AD8 и реферативной последовательности HXB2 идентичны, за исключением позиции 9506. В геноме AD8 в этой позиции основание С заменено на G. Поскольку такая замена наблюдается в 99 % всех исследованных нами геномов, мы выбрали для примера геном ВИЧ-1 с этой заменой.

Первая короткая шпилька (длиной 16 нт) в домене USE содержит в апикальной петле тракт UAUAA, соответствующий промоторному элементу ТАТАА. Вторая короткая шпилька (длиной 9 нт) представляет собой верхнюю часть шпильки USE, которая содержит в апикальной области U-богатый тракт. На рис. 5.17. б-д представлены также субоптимальные структуры домена USE в про-мРНК AD8.

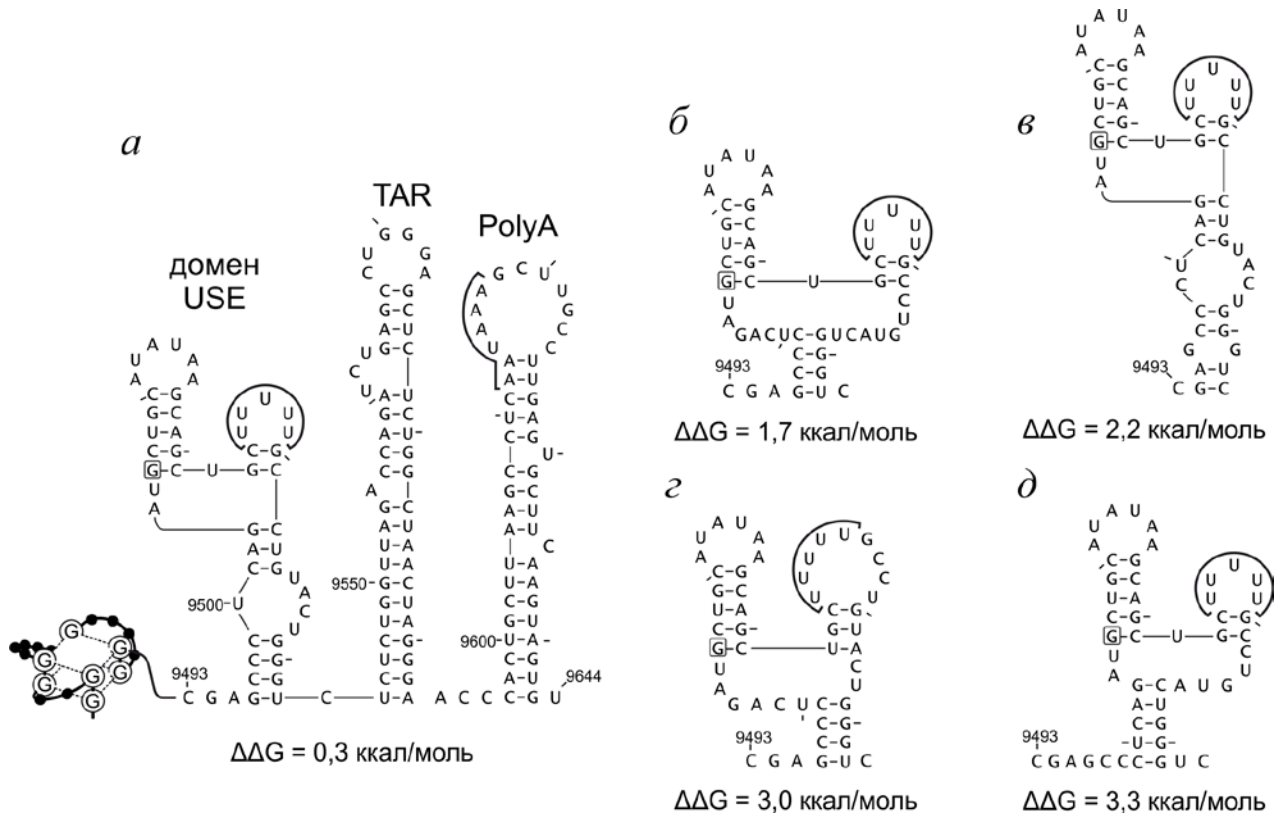


Рис. 5.17. Вторичная структура фрагмента поли(А)-сайта про-мРНК AD8 ВИЧ-1 субтипа В (номер AF004394 в GenBank), охватывающего домен USE, шпильку TAR и шпильку polyA (а). U-богатый тракт и гексамер AAUAAA отмечены линиями. Свободная энергия (ΔG) оптимальной структуры, не содержащей домен USE, равна -60,9 ккал/моль. Субоптимальные структуры домена USE про-мРНК AD8 (б-д)

Следует отметить, что обе короткие шпильки, содержащиеся в верхней части домена USE, присутствуют в модели вторичной структуры всей геномной РНК NL4-3 ВИЧ-1, определенной в работе [104] с помощью технологии SHAPE, но они входят в состав другого домена, включающего еще одну короткую шпильку и шпильку TAR. Короткая шпилька содержит два 3'-концевых олиго(G) тракта из G-богатой области. Однако в этой модели не учтена возможность образования G-квадруплексов на отдельных участках генома с высоким потенциалом для образования четырехцепочечных структур.

Более того, результаты зондирования области короткой шпильки, включающей олиго(G) тракты, значительно хуже согласуются с предсказаниями одно и двухцепочечных участков в РНК, чем результаты зондирования всех остальных трех шпилек. В более поздней работе [105] представлена пересмотренная модель вторичной структуры геномной РНК NL4-3 ВИЧ-1, предсказанная с использованием улучшенной программы. В новой структуре две короткие шпильки, входящие в нашей модели в состав домена USE, находятся в линкерной области большого домена, состоящего из многих структурных элементов. В этом случае результаты химического зондирования находятся в более хорошем согласии с нашей моделью.

Следует отметить, что G-богатая область в 3'-НТО про-мРНК ВИЧ-1 может более эффективно сворачиваться в G-квадруплексы в составе первичного транскрипта (про-мРНК), чем в составе геномной РНК. Так как транскрипция происходит в ядре хозяйской клетки, то в первом случае, ядерный белок гяРНК Н может стабилизировать образующиеся квадруплексы или даже стимулировать их образование. Во втором случае, G-богатая область может образовывать разные структуры с близкой свободной энергией, включая и G-квадруплексы.

Филогенетические исследования показали, что последовательность домена USE, как и в случае шпильки TAR, значительно более гетерогенна, чем последовательности шпилек polyA и DSE. Только 43 % всех исследованных нами геномов ВИЧ-1 содержали основные варианты домена USE, встречающиеся с частотой более 5 %. Эти варианты показаны на рис. 5.18. Как видно на рисунке, все варианты имеют подобные верхние части, состоящие из двух коротких шпилек, а нижняя часть представлена в двух разных конформациях. Мутации в домене USE в основном происходят в одноцепочечных участках и поэтому лишь незначительно влияют на его вторичную структуру.

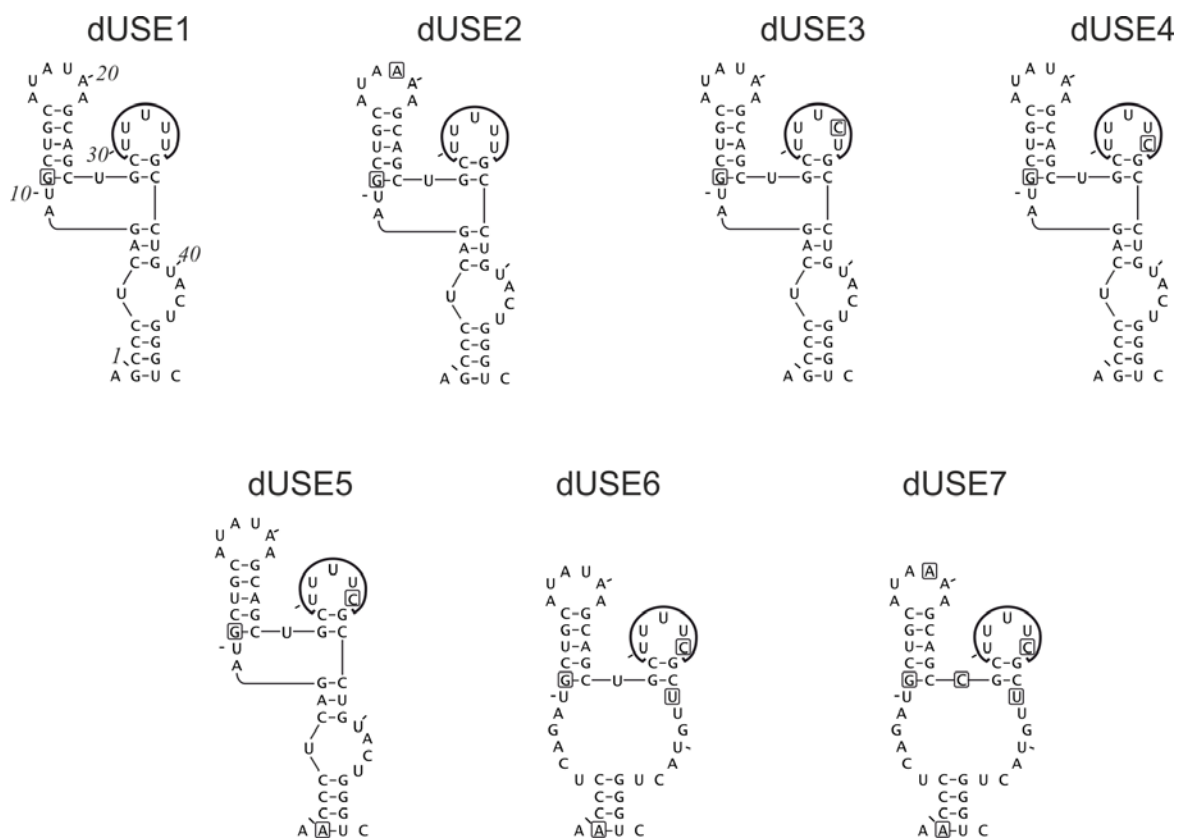


Рис. 5.18. Основные варианты домена USE про-мРНК ВИЧ-1 разных субтипов; dUSE1-dUSE3 – субтип В; dUSE 4 – субтип С; dUSE 5 – субтипы А и С; dUSE 6 – субтипы С и А; dUSE 7 – субтип 01_AE. U-богатый тракт отмечен линией.

Следует отметить, что U-богатый тракт USE довольно часто содержит замены U на C в четвертой или пятой позиции (рис. 5.18, dUSE3-dUSE7). Замена в пятой позиции встречается в изолятов ВИЧ-1 разных субтипов, а замена в четвертой позиции характерна для субтипа В.

Мы выявили, что при сворачивании фрагмента 9493-9644 нт локальная структура, состоящая из домена USE, шпильки TAR и шпильки polyA, с ΔG , равной оптимальной или отличающейся от нее не более, чем на 2 ккал/моль, образуется в 62-90 % всех исследованных про-мРНК ВИЧ-1, а с $\Delta\Delta G \leq 3$ ккал/моль – в 76-91 % (Таблица 5.4). Одноцепочечный U-богатый тракт, доступный для CPSF, может также присутствовать в других структурах этого

фрагмента, в частности в их линкерных частях или петлях, или в укороченной шпильке USE длиной 9 нт, не входящей в состав домена USE. Однако полная шпилька USE длиной 19 нт, представленная на рис. 5.15, практически не встречается при сворачивании фрагмента 9493-9644 нт. В общей сложности одноцепочечный U-богатый тракт присутствует в структурах с $\Delta\Delta G \leq 2$ ккал/моль в 76-94 % геномов ВИЧ-1, а с $\Delta\Delta G \leq 3$ ккал/моль в 81-95 %.

Следует отметить, что структуры, состоящие из шпилек TAR, polyA и DSE или USE, TAR и polyA являются оптимальными (а не субоптимальными) структурами в 80-90 % геномов ВИЧ-1 всех субтипов за исключением субтипа 01_AE (70 % и 30 % в первом и во втором случаях соответственно). Однако структура, состоящая из домена USE, шпильки TAR и шпильки polyA оптимальна только в 50 % геномов субтипов А и С, 30 % геномов субтипов В и 02_AG, 40 % геномов других субтипов и только в 6 % геномов субтипа 01_AE.

5.2.3.3. Вторичная структура комплексного поли(А)-сайта про-мРНК ВИЧ-1. При сворачивании области про-мРНК ВИЧ-1, охватывающей комплексный поли(А)-сайт (9493-9672 нт), две альтернативные структуры с экспонированными U-богатым трактом USE и U/GU-богатым DSE чаще всего встречаются в оптимальных структурах или в близких по энергии субоптимальных. На рис. 5.19 эти структуры показаны для про-мРНК AD8 субтипа В. На рис. 5.20 линейная и доменная структуры показаны для про-мРНК 93TN68 субтипа А. Эта про-мРНК является представителем группы изолятов ВИЧ-1 А/G-содержащих субтипов (А, G, 01_AE, 02_AG и других). Первая типичная структура (рисунки 5.19 а и 5.20 а) состоит из последовательно расположенных домена USE и шпилек TAR, polyA и DSE. Мы назвали ее линейной структурой. Замена основания С на G в позиции 9506, найденная почти во всех исследованных геномах ВИЧ-1, стабилизирует домен USE на 5,3 ккал/моль, тем самым способствуя образованию, как этого домена, так и всей линейной структуры (величина ΔG домена USE равна -13,0 ккал/мол).

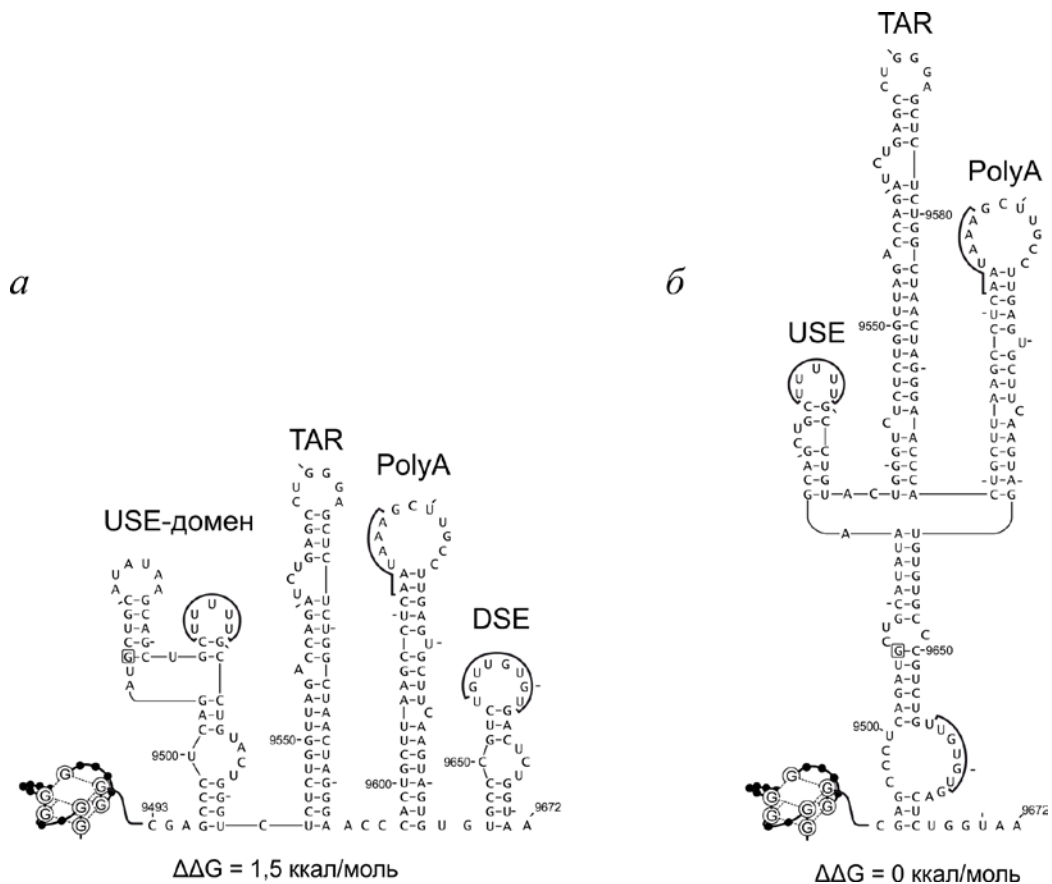


Рис. 5.19. Линейная (а) и доменная (б) структуры комплексного поли(А)-сайта про-мРНК AD8 ВИЧ-1 (номер AF004394 в GenBank). U-богатый тракт USE, гексамер AAUAAA и U/GU-богатый DSE отмечены линиями. Величина ΔG оптимальной структуры равна -67,7 ккал/моль

Вторая типичная структура (рисунки 5.19 б и 5.20 б) представляет собой домен, верхняя часть которого состоит из шпилек USE, TAR и polyA, а нижняя часть является нерегулярным дуплексом, во внутренней петле которого находится «нижний» элемент поли(А)-сайта про-мРНК ВИЧ-1 – тракт UGUGU. Эту структуру мы назвали доменной структурой.

При сворачивании области про-мРНК ВИЧ-1, охватывающей домен USE, шпильки TAR, polyA и DSE (9493-9672 нт), мы нашли одну из двух типичных структур комплексного поли(А)-сайта (или обе) с $\Delta\Delta G \leq 2$ ккал/моль в 65-76 % исследованных геномах, а с $\Delta\Delta G \leq 3$ ккал/моль – в 73-84

%, за исключением изолятов субтипа 01_АЕ, для которых эти величины составляют 30 % и 43 % соответственно (таблица 5.5, колонки 3 и 4).

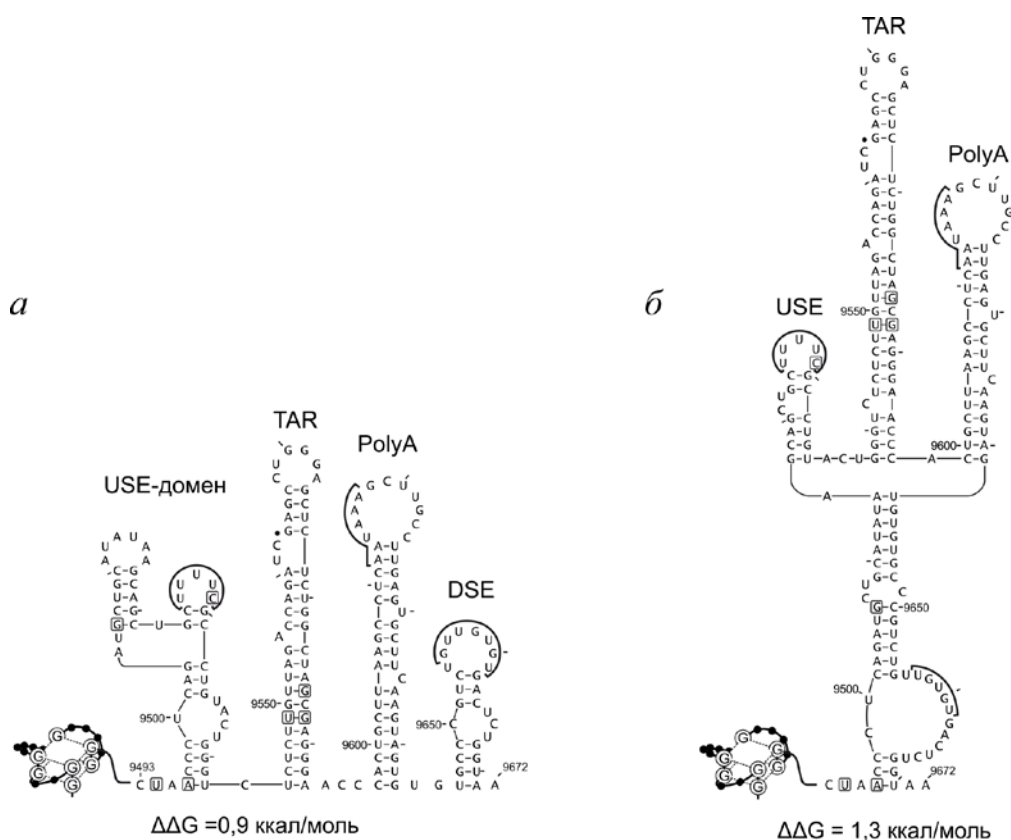


Рис. 5.20. Линейная (*a*) и доменная (*б*) структуры комплексного поли(А)-сайта про-мРНК 93TN68 ВИЧ-1 (номер AF196743 в GenBank). Остальные обозначения как на рис. 5.19. Величина ΔG оптимальной структуры равна -66,2 ккал/моль

Следует отметить, что U-богатый тракт USE и U/GU-богатые DSEs в одноцепочечном состоянии встречаются также в других структурах фрагмента 9493-9672 нт, отличных от линейной и доменной структур. Таким образом, структуры комплексного поли(А)-сайта с экспонированными U-богатым трактом USE и U/GU-богатым DSE образуются в большинстве исследованных геномов ВИЧ-1. Мы предположили, что линейная и доменная структуры, примеры которых показаны на рисунках 5.19 и 5.20, являются двумя основными структурами комплексного поли(А)-сайта про-мРНК ВИЧ-1.

Таблица 5.5

Частота встречаемости линейной и/или доменной структур комплексного поли(А)-сайта (фрагмент 9493-9672 нт) в изолятах ВИЧ-1 (%)

Субтип	Количество изолятов	«mfold»		«UNAFold»	
		$\Delta\Delta G \leq 2$ ккал/моль	$\Delta\Delta G \leq 3$ ккал/моль	$\Delta\Delta G \leq 2$ ккал/моль	$\Delta\Delta G \leq 3$ ккал/моль
А	136	76,5	82,4	69,1	80,1
В	525	69,3	83,8	72,6	86,7
С	269	69,1	75,5	62,4	79,2
01_АЕ	89	30,3	42,7	48,3	55,1
02_АГ	89	65,2	74,2	47,2	75,3
Другие	313	65,5	72,8	54,6	70,6

Предсказания вторичной структуры фрагментов комплексного поли(А)-сайта про-мРНК ВИЧ-1 проводились с помощью программы «mfold». Результаты сворачиваний хранятся в созданной нашей группой базе данных CESSHIV-1 (см. 2.4). Эта база данных создана как биоинформационный ресурс для исследования вторичной структуры контрольных элементов, расположенных в нетранслируемой и транслируемой областях геномной РНК и про-мРНК ВИЧ-1. В частности, база данных содержит секцию, посвященную вторичной структуре поли(А)-сайтов про-мРНК ВИЧ-1 разных субтипов и рекомбинантных форм. В ее семи подсекциях приведены оптимальные и субоптимальные структуры различных фрагментов комплексного поли(А)-сайта с указанием замен оснований, вставок и делеций (по сравнению с реферативной последовательностью про-мРНК НХВ2 ВИЧ-1).

Кроме программы «mfold» мы также использовали для предсказания вторичной структуры комплексного поли(А)-сайта про-мРНК ВИЧ-1 более современную программу «UNAFold». Результаты обеих программ находятся в хорошем согласии для линейных и/или доменных структур с $\Delta\Delta G \leq 3$

ккал/моль, за исключением изолятов ВИЧ-1 субтипа 01_АЕ. Количество этих структур для субтипа 01_АЕ, предсказанное программой «UNAFold», значительно выше, чем предсказанное «mfold» (55 % и 43 % соответственно, таблица 5.5).

Так как обе программы, «mfold» и «UNAFold», иногда не представляют результаты сворачивания для некоторых структур с близкими свободными энергиями, то некоторые интересующие нас линейные и доменные структуры поли(А)-сайтов про-мРНК ВИЧ-1 могли быть пропущены этими программами. Поэтому в тех случаях, когда при сворачивании последовательностей комплексного поли(А)-сайта мы не находили ни одной основной структуры с $\Delta\Delta G \leq 3$ ккал/моль, мы проверяли наличие таких структур в этом диапазоне ΔG , применяя соответствующие ограничения (см. параграф 2.4). Применение ограничений выявило небольшое количество пропущенных структур. Вклад этих дополнительных структур в данные, представленные в таблице 5.5, составляет в случае программы «mfold» 0-8 %, в зависимости от субтипа геномов, а в случае «UNAFold» – 0-5 %.

В большинстве исследованных про-мРНК ВИЧ-1 основные структуры комплексного поли(А)-сайта (доменные и/или линейные) образуются также и при сворачивании более длинного фрагмента (9493-9719 нт), охватывающего участок от конца G-богатой области до конца транскрипта (рисунки 5.21 и 5.22). Область, расположенная за поли(А)-сайтом, сворачивается в шпильку длиной 45 нт. Эта шпилька является частью домена PBS, содержащего сигнальные элементы, необходимые для инициации обратной транскрипции геномной РНК ВИЧ-1 (рис. 1.9). Поскольку шпилька длиной 45 нт имеет в литературе различные наименования, мы назвали ее для простоты шпилькой U5.

Примерно в 10 % изолятов ВИЧ-1 А/G-содержащих субтипов фрагмент 9493-9719 нт сворачивается в другую линейную структуру с одноцепочечными U-богатым трактом USE и U/GU-богатым DSE. В этой структуре область, расположенная за шпилькой polyA, сворачивается в две шпильки (рис. 5.22 а).

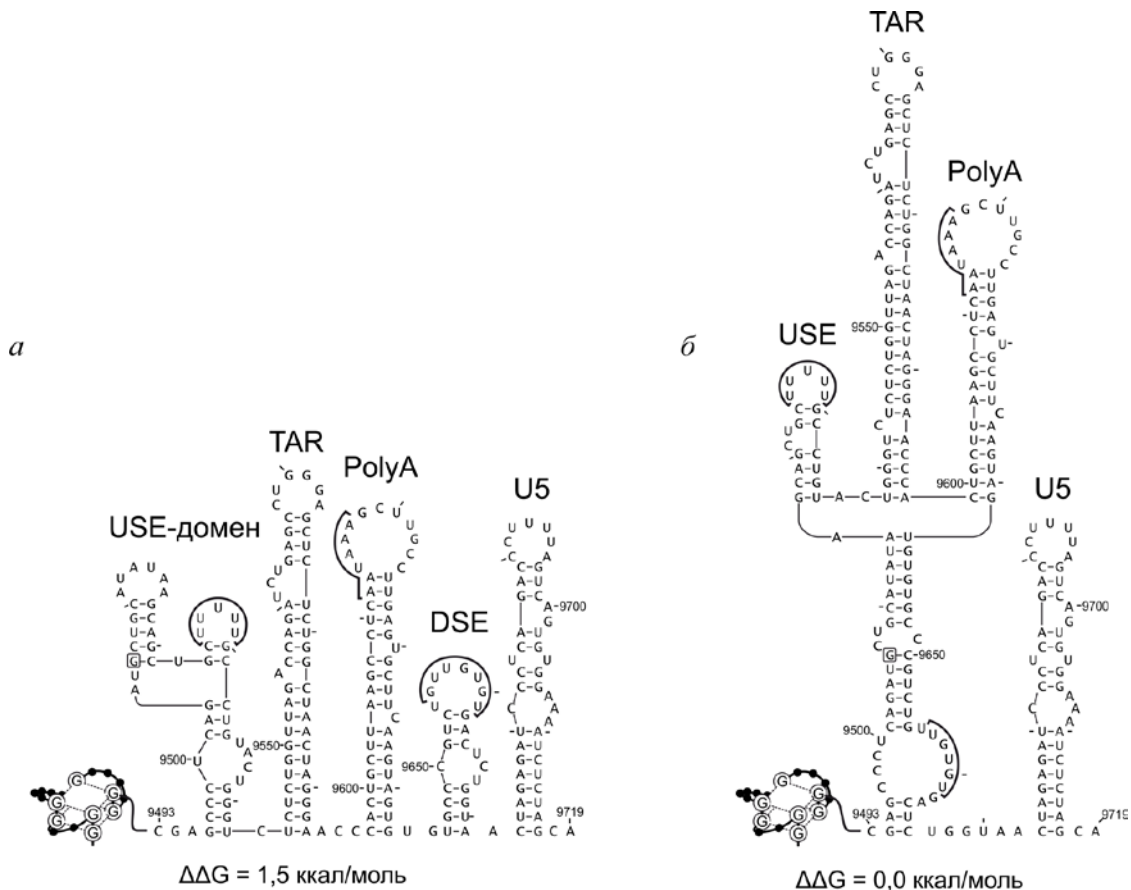


Рис. 5.21. Линейная (а) и доменная (б) структуры 3'-концевой части про-мРНК AD8 ВИЧ-1 (номер AF004394 в GenBank). Обозначения как на рис. 5.19. Величина ΔG оптимальной структуры равна -80,8 ккал/моль

Первая шпилька с апикальной петлей UAA содержит U/GU-богатый DSE во внутренней петле. Мы назвали эту шпильку длиной 48 нт шпилькой UAA(48). Вторая шпилька является укороченной формой (23 нт) шпильки с A-богатой апикальной петлей, найденной в домене PBS геномных РНК ВИЧ-1 A/G-содержащих субтипов [340]. Мы назвали эту шпильку шпилькой A-loop(23). Фрагмент 9493-9719 нт охватывает почти всю 3'-нетранслируемую область про-мРНК ВИЧ-1 за исключением короткого тракта длиной 17 нт перед G-богатой областью (рис. 1.10.). Следовательно, вторичные структуры, показанные на рис 5.21 и 5.22, могут рассматриваться как структурные модели 3'-НТО про-мРНК ВИЧ-1.

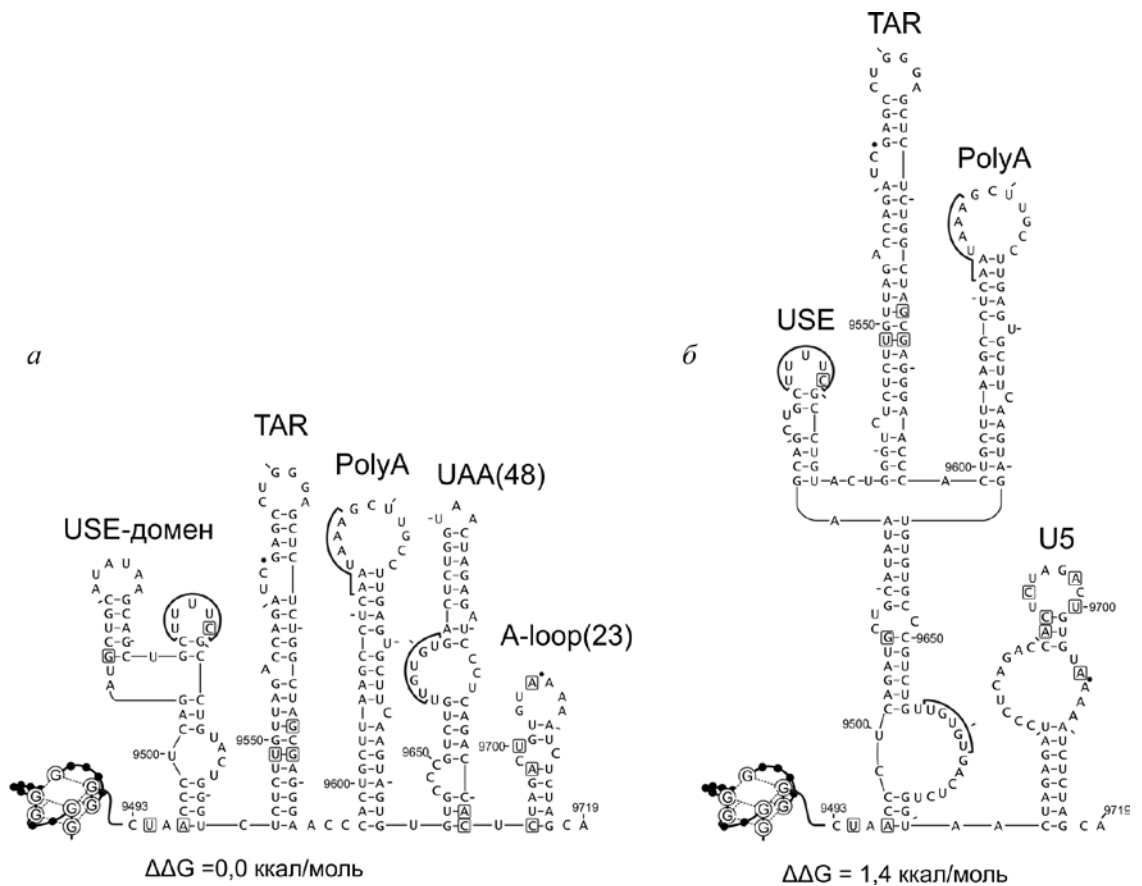


Рис. 5.22. Линейная (а) и доменная (б) структуры 3'-концевой части про-мРНК 93TN68 ВИЧ-1 (номер AF196743 в GenBank). Обозначения как на рис. 5.19. Величина ΔG оптимальной структуры равна -75,8 ккал/моль

Шпилька U5 непосредственно вовлечена в инициацию обратной транскрипции на 5'-конце генома ВИЧ-1 и, возможно, играет вспомогательную роль в процессе полиаденилирования на 3'-конце транскрипта, поддерживая функциональную структуру поли(А)-сайта с экспонированными U-богатым трактом USE и U/GU-богатым DSE. В изолятах ВИЧ-1 A/G-содержащих субтипов подобную структурную роль могут играть шпильки UAA(48 или 39) и A-loop(23).

В заключение перечислим основные доводы в пользу предложенной нами модели поли(А)-сайта про-мРНК ВИЧ-1 (две структуры). Во-первых, обоснован выбор 5'-границы для предсказания вторичной структуры комплексного поли(А)-сайта. Он основан на том, что G-богатая область,

охватывающая минорные USEs, имеет высокий потенциал для образования G-квадруплексов. Следовательно, участок, расположенный за этими USEs, может при транскрипции 3'-концевой области провирусной ДНК сворачиваться автономно. Во-вторых, предполагаемые линейная и доменная структуры включают хорошо известные шпильки TAR и polyA, образующиеся на обоих концах транскрипта [81]. В-третьих, найденные нами новые элементы – домен USE и шпилька DSE в линейной структуре или шпилька USE и внутренняя петля в доменной структуре содержат U-богатый тракт USE и U/GU-богатый DSE в одноцепочечной форме, доступной для аппарата полиаденилирования, что свидетельствует о функциональной компетентности модели. В-четвертых, линейная и/или доменная структуры образуются в большинстве из исследованных нами 1421 генома ВИЧ-1 разных субтипов и рекомбинантных форм (55-87 %) в оптимальных структурах или структурах с близкой свободной энергией.

Следует также отметить, что совсем недавно (декабрь 2014 г.) в работе Perrone et al. [341] было приведено экспериментальное доказательство образования различных G-квадруплексов в G-богатой области генома ВИЧ-1, соответствующей GRS, показанной на рис. 5.16. Было также показано, что 3' концевой G-повтор из двух G-остатков также участвует в образовании G-квадруплекса. В этом случае, как предполагают Perrone et al. [342] в своей предыдущей работе, где они экспериментально доказали образование G-квадруплексов в этой же самой G-богатой области в провирусной ДНК, четыре последних G-повтора (рис. 5.16) могут участвовать в образовании квадруплекса из трех G-тетрад с одним балджаем в четвертом тракте (GUGG). В соответствии с литературными данными G-квадруплексы РНК оказались значительно стабильнее соответствующих квадруплексов ДНК. В частности, T_m G-квадруплекса из четырех последних G-повторов равна 71,2 °C для РНК и всего 43,4 °C для ДНК в присутствии 100 mM K^+ [341]. Таким образом, работы Perrone et al. [341, 342] полностью подтверждают правильность нашего выбора

5' границы для независимого сворачивания последовательности РНК, соответствующей комплексному сайту полиаденилирования про-мРНК ВИЧ-1.

5.3. Выводы к разделу 5

Учитывая наличие физических предпосылок протонирования адениновых трактов РНК в клетке, предложены следующие модели участия внутримолекулярных двойных спиралей олиго(А) в биологических процессах:

- образование внутримолекулярного дуплекса в поли(А)-хвостах ARE-содержащих мРНК под действием Hu/ELAV-белков препятствует действию поли(А)-нуклеаз, стабилизируя эти мРНК;

- PABPC в комплексе с UNR и IMP1, индуцируя образование внутримолекулярной двойной спирали в А-богатом участке, расположенном в 5' НТО своей мРНК, препятствует продвижению рибосомной частицы по направлению к иницирующему кодону, что приводит к ингибированию синтеза этого белка;

- процессивная реакция полиаденилирования про-мРНК высших эукариотов прекращается при достижении поли(А)-хвостом длины ~ 250 нт вследствие образования в нем внутримолекулярной двойной спирали за счет переноса протонов от чрезвычайно кислого N-концевого домена PABPN1 на адениновые основания синтезируемого поли(А)-хвоста. Образование дуплекса приводит к диссоциации белкового комплекса полиаденилирования и к переходу от процессивной реакции синтеза к дистрибутивной.

Предложены структурные модели области димеризации генома ВИЧ-1, включающие шпильки DIS и G-квадруплексы, стабилизирующие димер генома. Подобные модельные структуры могут образовываться практически во всех исследованных 350 геномах ВИЧ-1 разных субтипов. В большинстве случаев они могут включать квадруплексы, состоящие из различных наборов G-богатых трактов.

Результаты филогенетического исследования ~ 1700 геномов ВИЧ-1 показали высокую консервативность палиндрома GGGAGCUCUC в верхней части шпильки TAR, что свидетельствует в пользу гипотезы димеризации этой шпильки через взаимодействие палиндромов. Предположено, что контакт между геномами ВИЧ-1 через шпильки TAR осуществляется только на определенной стадии репликационного цикла вируса и является регулируемым процессом.

Впервые представлены структурные модели поли(А)-сайтов, включающие G-квадруплексы, для двух вирусных про-мРНК. В про-мРНК вируса SV40 L образование G-квадруплекса приводит к реорганизации структуры поли(А)-сайта, в результате которой основные сигнальные элементы становятся полностью доступными для белков аппарата полиаденилирования.

В случае про-мРНК ВИЧ-1 в области минорных USEs выявлен GRS с высоким потенциалом для образования многочисленных G-квадруплексов. Предположив, что последний G-повтор в GRS служит 5'-границей для автономного сворачивания расположенного за ним фрагмента РНК, мы впервые предсказали вторичную структуру комплексного поли(А)-сайта про-мРНК ВИЧ-1 (две структуры), а также структуру почти всего 3'-НТО.

Впервые проведено широкомасштабное филогенетическое исследование структурных элементов комплексного поли(А)-сайта про-мРНК ВИЧ-1. Данные об общих и специфических чертах домена USE, шпилек TAR, polyA и DSE, а также двух структур комплексного поли(А)-сайта могут быть полезными при конструировании противовирусных препаратов, направленных на ингибирование сигнальных элементов в геноме ВИЧ-1.

РАЗДЕЛ 6

БАЗА ПОЛИ(А)-САЙТОВ ПРО-МРНК ЧЕЛОВЕКА

Анализируя нуклеотидные последовательности поли(А)-сайтов 95 про-мРНК высших эукариотов и их вирусов, приведенные в работе [343], мы обратили внимание на тот факт, что 19 % этих последовательностей содержали кластеры из G-повторов ($n \geq 3$), расположенные немного дальше основного U/GU-богатого элемента [65]. Некоторые из этих GRS, состоящие из четырех и более кластеров (в отдельных случаях из трех), имели потенциал для сворачивания в G-квадруплексы. Учитывая, что в поли(А)-сайте про-мРНК SV40 L подобный кластер играет роль вспомогательного элемента, стимулирующего процесс полиаденилирования, мы предположили, что такую же роль могут играть G-богатые участки и в других транскриптах. Позднее, как уже упоминалось в параграфе 1.2.4., Arhin et al. [135] нашли GRS в трети из 219 исследованных поли(А)-сайтов про-мРНК высших эукариотов. Четыре из 13 GRS, приведенных для примера в этой работе, могли потенциально сворачиваться в G-квадруплексы.

Для того чтобы выяснить более детально, как часто тракты, способные сворачиваться с образованием G-квадруплексов или i-мотивов, встречаются в области поли(А)-сайтов про-мРНК высших эукариотов, мы создали базу поли(А)-сайтов про-мРНК человека [71]. Также, мы хотели узнать, насколько часто в этих про-мРНК встречаются вспомогательные элементы поли(А)-сайтов.

База поли(А)-сайтов про-мРНК человека, созданная нами, содержит 244 последовательности ДНК, соответствующие областям про-мРНК, расположенным на расстоянии 200 нт до места расщепления (или до 5'-конца экзона) и на расстоянии 200 нт за этим местом. Последовательности генома относятся к разным хромосомам и экспрессируются в разных тканях. Для того чтобы определить сайт расщепления, мы сравнивали последовательности

геномных ДНК и соответствующих мРНК, взятые из GenBank. Были отобраны только последовательности, отмеченные кодом «Evidence Code C», т.е. подтвержденные модели генов, основанные на выравнивании мРНК, или мРНК совместно с ESTs (expressed sequence tags) с геномными последовательностями. Кроме того, отбирались только мРНК со статусом «Reviewed», т.е. с установленным местом расщепления.

6.1. Организация основных сайтов полиаденилирования про-мРНК высших эукариотов

Используя созданную нами базу данных, мы определили расстояния между местом расщепления и обоими сигнальными элементами основного поли(А)-сайта – поли(А)-сигналом и DSE. Наиболее часто встречающиеся значения (рис. 6.1) хорошо согласуются с литературными данными [70, 73]. Кроме того мы определили среднее расстояние между самими сигнальными элементами.

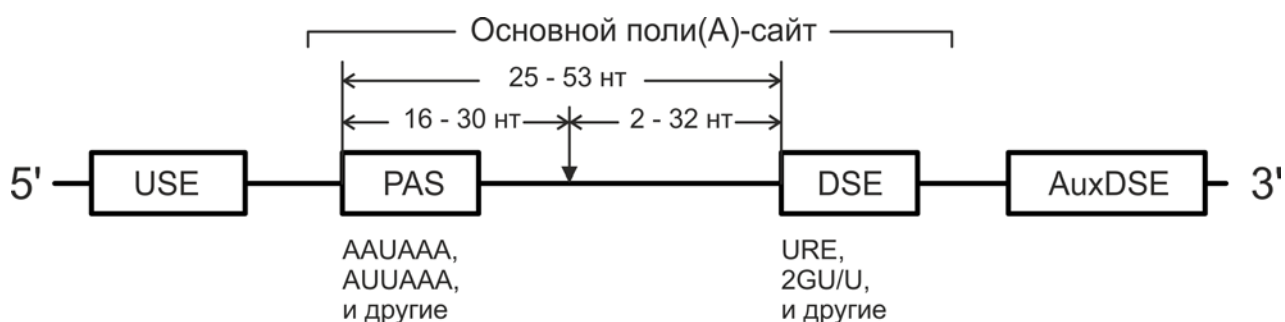


Рис. 6.1. Организация поли(А)-сайта про-мРНК человека. PAS – поли(А)-сигнал

На рисунках 6.2 и 6.3 показан фрагмент нашей базы для 37 про-мРНК человека. На рис. 6.2 приведены последовательности длиной 70 нт, расположенные до места расщепления, а на рис. 6.3 – последовательности такой же длины, расположенные непосредственно за местом расщепления. На

рисунках 6.2 и 6.3 представлены последовательности для транскриптов из нашей базы поли(А)-сайтов про-мРНК человека с нечетными номерами. Основные сигнальные элементы (PAS и DSE) выделены.

```

A      GTGACCAGATGCCCTACTCAGCTTCCATCACCCCTAGCAATATGTATTAAAGTCTGAAGTGTGGCCATGG
ACRBP  GTCCCAGAGAGGGCCACGGTGGGAGCTGCGCCCTCCTTAAAAGATGACTTTACATAAAAATGTTGATCTTC
ACTA2  TCATTCCTAGCCAAAAGCTCTGACTCGTTACCTATGTGTTTTTTTAATAAAACTGAAATAGGCTACTGGTAA
ACTG2  ACAGGTTCTCCAAGGATCCCCCTCGAGACTACTCTGTTACCAGTCATGAAACATTAAAACCTACAAGCCTT
ACY1   ACCTCCTCTTCCCCCTTCCAAATAATAAAGTCTATGGACAGGGCTGTCTCTGAAGTACTAACACAAGGAC
AIF1   ATATGGAACAGAAAGACAAAATTGTAAGCCAGAGTCAACAAATTAAATAAATTACCCCCCTCTCCAGATC
ALDH3A2 CAGATATACCTTATTGGAGACAATCCTTTGATCATAAATTCTCCCCAACTATAAATCATTTTATGTCTTT
ALDOB  GTCTTAGCAGTGGTAGGTTGGGAAGGAGATAGCTGCAACCAAAAAAGAAATAAATATTTCTATAAACCTTC
ANXA8  ACTGGGTACCTGGAGATTCTGAAGTGCCTTTGCTGTGGTTTTTCAAAATAATAAAGATTTGTATTCAACTC
ATP6V0B GTGTCCCCACCTCCACCCTCAACCCATCTTCCTAGTGTGTTGTGAAATAAACTTGGTATTTGTCTGGGTC
ATP6V1E1 GTAAATGTGATTTAAAATCTAAGCCATGAATATGCTTTTATTTTATTAAAAGAGTTATGCGGATTTAATGTG
B3GALT5 GACTTTACTTTGTACTCAGAGAAGAGGGCTCACATGGCTGTGTCACATATAAATGTTGGACTAAACTCTT
BACE2  GCCACTATCACATATGCTTACTCTTGCTTAAAATTAATAAATCATGTTTTGATGAGAAAAAACTATTCTA
BAT3   CCTCATCAGGGACCGTTTTCCCCCTCTTCTTACAGTATTTAAGAAATAAAAGTCGGATTTTTCTGGC
BBP    TAATCTTTTACAAAGGCAAGTTGCCACTTGTCATTTTTGTCTGAAAAATAAAAGTATAACTTATTCAC
CA5A   GAAGGAAGCGTTGTTTCCCAAGTTTCACAATGTGATTGTACGTGACTTCTGAAATTAAAAAGAGAGAATG
CA7    GTCCCCACCCAGTCACCAAAGCCACCTACATGACAGTCCATCCCTGTTGAATTAATAAATTAATGTATCC
CD97   GTTGTAATGTGTTTTTATCTGTTAAAATTTTTTCAAGTGTGACACTTAAAATTAAACACATGCATACAG
CDC10  TCATATATTTGCATTTCTGTATTTTTTGTATTGTAAAAAATTCACATAATAAACGATGTTGTGATGT
CDH16  CTCCAAATGTCAGGGTGTGTTGCCAATAATAAAGCCCCAGAGAACGGGCTGGGCCCTATGGGATTGGTA
CDKN3  AATGAAACCACCAGTGTATCAACTGAATGTAAATGTACATGTGCAGATATTCCTAAAAGTTTTATTGAC
CLIC4  TATCTTTGGAACATCAGCACCAGTATATTGCTGGCAGCTATTGTATTAAAAAATAAAGTATATTTTCACT
CLTCL1 GGTACCGGACCTCATAAAAGGACTCTCAGCCTCTGGCAATCATAAATATTAAAGTGCGTTTTATCCAGGC
CST11  TGTGAGAGAATCCTGCAATGGGCAATAATCCTTAATATAGGCAATAAATGTTTTGTCAACATATTCTCTTA
CYP21A2 CTCCGCTGCAGAGGATTGAGGCTTAATTCTGAGCTGGCCCTTCCAGCCAATAAATCAACTCCAGCTCCC
DDX16  CTATACATTTAATAACCTATTAAATAAAATTATTTTTGGAATAAAGCCTGTGGGAACATTTGGGATCT
DEEPEST CCCCAAATACCAAGACCAACTGGCATAGAGCCAACCTGAGATAAATGCTATTTAAATAAAGTGTATTTAATG
DNAJA2 TTTTTTTTTTTTTTACCATCTTGTGAAAGTTTCTGAAACTCGATAATAAAAAGCGGTTGGTGTAAATT
DPM1   TCTCAATTTTCGTTTTTCATTTTGTGCTGATTGAGACCTATAAATAAATGTATATTTTTTTTTTGCATAAAGT
DUSP4  CATCTTGAGAAGCAGGCGGGTTGGGTGGGAGGAGGAAGAAAGGAAGAATTAGGTTTGAATTGCTTTTTT
EFEMP1 AAATGTCACCTGCTTTACTTTGATGTATCATATTTTTTAAATAAAAAATAAATATTCSTTTAGAAGATCACTC
ENO1   TCCGGGGTGGCCACAGGCTAGATCCCCGGTGGTTTTTGTGCTCAAAATAAAAAGCCTCAGTGACCCATGAG
F13A1  CTGGCTTAACTTAGTATATTATTATTAATTACAATGTAATAGAAGCTTAAAATAAAGTAAACTGATTAT
FANCC  TTTTATAGCTSTTTAAAATCTGTAATTTGTTGTCTATTTTTTTCATCTAAATAAAACTTCAGTTTGCACCT
FBXL3A TTAATTTGTTTATACTGTAAAGTATATGTTAAATGCTTTTATCATTGAGAATAAAAAGTTACTAATGC
FBXL7  CAAATCACTGCGTGTGTATGAATCTAGAAAGCCTTAATTTACTACCAAGAAATAAAGCAATATGTTTCGT
FEZ2   TCATCAATATCATGAGCACTTTTGTGGTCCAGTCAAATCAATGGCATTAATAAATTTTTTAAGAAGT

```

Рис. 6.2. Фрагмент базы поли(А)-сайтов про-мРНК человека. Область про-мРНК, расположенная до места расщепления транскрипта

На рисунках 6.2 и 6.3 видно, что в большинстве случаев в области, расположенной за 16-35 нт до места расщепления, может быть найден основной поли(А)-сигнал (гексамер AAUAAA или его производные), а в области за местом расщепления (+1/+30 нт) имеется тот или иной U/GU-богатый основной нижний сигнал.

```

A      AAACCTCAGTGTCCAGTGTACTGCGGTGGAAAGGGGCTTGGAGGGTGGAGTGGCTCACAGGGGACACGAC
ACRBP  AGCCTGTGTGGCCTGTTCTGGTTATGATCCCCATCCTCAGCACTGCTACCCCTGCCCCGTGGCAGTCAGTC
ACTA2  CTTCTTGCCTTGGCCTCCGATCAATACTCGACTGATACCATTTGTTAGGTTGACCTCACCCGTTACAGGA
ACTG2  ACTTCTCTGTGTGGGGCTCTTTTTTCCTGGGCTATGTCTCATAACACAGTGTAAAGGACTTTTCCACACATT
ACY1   ACTCGTGGAGCAAGAATTTTCCTTTTTCCTGGGGACATGTTACCATCTCCATTTACAGATGAGGAAACTG
AIF1   AAGTCAGCTTAGTTTTTATTTTGGGTGATTTTTTCCCTGGGTTTGGGAAGGAGAGACAGGTCTTGAGGGAA
ALDH3A2 ATCATAGATGGTATTTGTCAAACTTTTTTTTTCTCATTTTAAAATTCTTAGCTGAGAAGATTTCTCAAAC
ALDOB  AGCTGCTATCGGGTTTCACCTTTCTGTCTCTTGCTGTCCAAAGACTCAGTGTTAATTTCATTACTTTTGACTC
ANXA8  AAAGTATCTTTGGATTGTTTAGCAGGGACCAGTGTAGGACCAATTTAATTTGTCAGAAAACAGAGCTTTGGG
ATP6V0B AGTGCAGCTTCTGTTGCCCTCTTGCGTCAACCGTTCACAGAGGAAGGGTAAAGAGGCCTGTACGCCTCAGC
ATP6V1E1 ATTTCTAGTGTAAGGCACTACAAATTTAGGTTTTGACAGAGTCAATTCTTACTGAGGCACCTAGAAAAG
B3GALT5 ACTTGAACCTCAGGAAAAGTACTTGTCTGCAAAATGGTTTTGACCATGTTCAATTGCAAAAGTACGTCTGA
BACE2  TTTCACTAGCTTAGTTGTCTCTTTTTCCAAATCTTCTCTGGAAGTAGGTTGGCTATTACCCTGTTGGGAA
BAT3   TGCTTTTCTCTCTACATTGTCTCCATTAGGTAGTTGTGTCCCTTAATCTTGTGTGAACTTTTTTAAATTAAT
BBP    ATTTTTAGTGCTTTTGTTTATTGAGAAAATCTTTCTTAAATGTTGAATGTAGCTGTCTTACCCATTTTGATCC
CA5A   AGGTTATTGTTTAGTCAAAGTACGTTTCAGCTGCAAGTGAAGAAAGATGTGTTAAGTACCCTGGGGAAC
CA7    ATGCAACAAATACCAATTGAGTGCTTGTGCTGTCTCTAGTACCCTCCTAGGCTTTGGGGAGCAATGGTGG
CD97   AAGATGGCCTTGCCCTGCTGGTGGTGATTCAGAGGCCGGGCCAGGGTGTCCGTGAAAGGCCAAGATGACC
CDC10  AATATTGTGTGAGGTCTTAAATATCCTACAGTCGATGTACAAGAGTAGAGTATGTTTGGGAAGAAACTTT
CDH16  TGTTGTGTCTCAGTGTCTATTGTGTGCATGTGTGTGTGTGAGTGTGCACACACCTGCTTTTTCATCCCTT
CDKN3  AAAACTCGTTGTGTCTCTTTCATTCTTATAGGAATAGGAAAATGGGATAGGAAAATGGTATGCAATT
CLIC4  ATCATAAAGGATTTCTTTTTTCCCCCCTCATGAAAATAAACAACAACCTGGGGTAAAAGTTTTGAGACTA
CLTCL1 AAAGGTGCTCTGCTTGGTGCTGGTATGGGTAGGAGGGACGGGCCAGTTTCTTAGGCCCCGCAGTCCATTA
CST11  AGATCGTGATTCACCTTTTTTCTTTTTCTTTTTATCCAAATTCTTCAGCTCATAGTAATAATCTGAAAG
CYP21A2 TCTGCGAGGCTGGCATGATTGTTCCATTTCAACCAGCCACTCAGTCCCTTGCCCTGTTACACTGTGGGGCT
DDX16  AGAGAAAAGTGATATGTGAAATCTATCTCATATAGTCAGTTAAACTTTATTTTTACAAGTTAAATTACA
DEEPEST AATTTCTCCAAGCTTACGGAACTTACTTTTTGGGTGTAGAGGGGGCGGATGATGAAGTCAGGGAAGAAGC
DNAJA2 ATCTTTTTGTGTCACATTTTTAGAAGGAAAAACATAAAAGAATGTATCCTTAGTACTGGTTCTTAAACAG
DPM1   ATTGCTGCCTTCATTAGAGTATGTGAATGTAGAATTTTTTACATGGGGAATAATTTGAAATAGCTTTTT
DUSP4  AAAAAAAAAAGAAAAAGAAAAAAAAAGACAGCATCTCACTATGTTGCCAAGGCTCATCTCAAGCTCTTGGGC
EFEMP1 TATCTTTGGAGGTTTTTCAGTGTAGTCAGTAGCCCCAATATAATTTTTATATCTGGGATAAACATCAGT
ENO1   AATACTCCGTGTGCCTTGTGTATGTCTGGAACAATCTGGGTCTGTCCCTTAGTGTTCAGGGGGCCCTGGCGA
F13A1  ATTTGCATCTGTGCAGAAGTGAATGCTTTTAATTTTCAATATAACACACTACTAATCACAACTTCTTGAG
FANCC  AATGTTGTATCCATTAATAATTAACAATCCTTCTCTCAGGAAAAAGCTCCCCCGATGGCCCCAGACAGTGG
FBXL3A TGTGTATTTTACTTTTTAAAACTACACAAACATTTGACTTTAAACTTAAAAATGTTTCAAATCCATAT
FBXL7  AATCAGCCTCAGCTTCATTTTTAATAACCTTCCAGAGGAGAGTGCTGTTTGGTTGGGGGCCCTCAAAG
FEZ2   AAAAAATACATGTTTGAGAGCCTTATATTTTGGGAGTTTTTGTGTGTGTATTTTTTTTTTTCTCTTACAGTGCTAG

```

Рис. 6.3. Фрагмент базы поли(А)-сайтов про-мРНК человека. Область про-мРНК, расположенная за местом расщепления транскрипта

Это общая черта основных поли(А)-сайтов про-мРНК высших эукариотов. С другой стороны, видно, что все поли(А)-сайты специфические. Каждый индивидуальный поли(А)-сайт можно охарактеризовать, по крайней мере, десятью параметрами, описывающими его последовательность, структуру и организацию. Скорее всего, поли(А)-сайты, характеризуемые различными параметрами, направляют реакцию полиаденилирования с различной эффективностью. Прежде всего, все про-мРНК в нашей базе поли(А)-сайтов можно разделить на 6 групп, исходя из последовательности поли(А)-сигнала – гексамера AAUAAA, с которым связывается белок CPSF. Про-мРНК с каноническим гексамером AAUAAA мы отнесли к группе I.

Про-мРНК с единичными заменами в каноническом гексамере подразделили на три группы. К группе II отнесли транскрипты с гексамером AUUAAA, направляющим реакцию полиаденилирования с относительно высокой эффективностью (77 %) по сравнению с каноническим гексамером [70]. Про-мРНК группы III имеют гексамер с единичными заменами в первой, второй или пятой позициях. Эти сигналы направляют процесс полиаденилирования с эффективностью 10-29 % (за исключением AAUAGA) [70]. Небольшую группу IV составляют про-мРНК с заменами в третьей, четвертой или шестой позициях гексамера. Транскрипты с двумя заменами в каноническом гексамере, о которых мало известно из литературы, мы отнесли к группе V. Остальные про-мРНК из нашей базы данных мы включили в группу VI. Доля про-мРНК с гексамерами разного типа указана в таблице 6.1.

Как видно из таблицы 6.1, подавляющее большинство про-мРНК (91 %) имеют гексамер типа I-III в области -35/-16 нт. Хотя транскрипты группы VI не содержат гексамер типа I-V вблизи места расщепления, многие из них содержат гексамер типа I-III (очевидный сигнал полиаденилирования) в пределах «верхней» области, исследованной в работе. Эти сигналы могут быть функциональными, если последовательность, разделяющая гексамер и «нижний» элемент, образует шпильку и пространственно совмещает основные элементы поли(А)-сайта.

Таблица 6.1

Основной поли(А)-сигнал в области - 35/-16 нт поли(А)-сайтов 244 про-мРНК человека

Тип элемента	Элемент	Доля про-мРНК, содержащих данный элемент (%)
I	AAUAAA	68,8
II	AUUAAA	14,7
III	NAUAAA ANUAAA AAUANA	7,8
IV	AANAAA AAUNAA AAUAAN	1,2
V	Замещение любых двух оснований в AAUAAA	2,9
VI	Другой сигнал, отличный от гексамера типа I-V	4,5

Распределение транскриптов по типу гексамера, показанное в таблице 6.1 для 244 про-мРНК, хорошо совпадает с позднее опубликованным распределением, полученным Tian et al. [344] для примерно 6400 конститутивных генов человека. По данным этой работы 70 % генов имеют в области -40/-1 нт поли(А)-сайта гексамер AAUAAA, 15 % – AUUAAA, 12 % – гексамер другого типа и 4 % не содержат известных гексамеров. На рис. 6.2 выделены не только гексамеры типа I-III, находящиеся вблизи места расщепления, но и расположенные в других местах исследуемой области, так как некоторые из таких гексамеров могут быть элементами альтернативных сигналов полиаденилирования. Мы нашли, что примерно 60 про-мРНК в нашей базе данных содержат более одного гексамера типа I-III и 70 % из этих

транскриптов имеют гипотетические «нижние» элементы, расположенные недалеко от гексамеров.

В отличие от основного поли(А)-сигнала и его вариантов, локализованных в основном в узкой области -35/-16 нт, гипотетические U/GU-богатые «нижние» элементы фактически распределены в пределах всего участка длиной 70 нт, представленного на рис. 6.3. В области +1/+32 нт, оптимальной для осуществления реакции полиаденилирования, их незначительно больше, чем на участке +33/+70 нт. Элементы значительно различаются по длине. Кроме того, нередко в одном и том же поли(А)-сайте встречаются два и более элемента. Как уже сообщалось в параграфе 1.2.1., анализируя литературные данные, а также свою базу поли(А)-сайтов про-мРНК человека, мы пришли к выводу, что «нижняя» область поли(А)-сайтов высших эукариотов состоит из различных простых элементов (пентамеров), расположенных на разных расстояниях друг от друга, в том числе и перекрывающихся между собой с образованием длинных трактов. К этим пентамерам относятся URE типа «четыре U из пяти», а также все комбинации из двух GU и одного U. Результаты анализа встречаемости различных «нижних» элементов в нашей базе поли(А)-сайтов приведены в таблице 6.2.

Как видно из таблицы, в про-мРНК поли(А)-сайтов человека преобладают основные «нижние» элементы типа URE (66-70 %). Однако GU-богатые элементы тоже составляют заметную долю (27-31 %). Интервал между значениями связан с тем, что в ряде случаев U/GU-богатый элемент можно отнести к обоим типам. Например, в поли(А)-сайте про-мРНК DNAJA2 (рис. 6.3) основной «нижний» сигнал UUCUUUUGUGU включает три перекрывающихся элемента типа «четыре U из пяти»: UUCUU, UCUUU и CUUUU, а также GU-богатый элемент UGUGU. В общей сложности «нижние» элементы двух типов (URE и 2GU/U) встречаются в 80 % исследованных про-мРНК.

Таблица 6.2

«Нижняя» область (+1/+32 нт) поли(А)-сайтов в 244 про-мРНК человека

Элемент	Вариант	Доля про-мРНК, содержащих данный элемент (%)
URE	NUUUU UNUUU UUNUU UUUNU UUUUN	66-70
2GU/U	GUUGU UGUGU GUGUU	27-31
Другие элементы	-	20

Наличие часто встречающихся двух сигнальных пентамеров в одном длинном U/GU-тракте или в разных близкорасположенных трактах находится в соответствии с современными данными о структуре комплекса белков, осуществляющих процесс полиаденилирования про-мРНК высших эукариотов. Согласно новой модели [345, 346], этот комплекс содержит стабильное ядро, состоящее из факторов CPSF, CstF и CFI, с которыми взаимодействуют остальные факторы полиаденилирования, а также различные белки, участвующие в координации процесса полиаденилирования с другими клеточными процессами. Два фактора, CstF и CFI, могут быть представлены в белковом комплексе в виде димеров. Предполагается, что одновременное связывание CstF и/или CFI с двумя сигнальными элементами способствует повышению специфичности и прочности РНК-белкового взаимодействия.

Фактор CFI специфически связывается с элементом UGUAN [84]. В работе [347] проведено детальное исследование влияния элементов UGUAN на процессы полиаденилирования про-мРНК PAPOLA с каноническим

гексамером AAUAAA и PAPOLG с неканоническим поли(А)-сигналом (AAAGAGAAA). Первая про-мРНК содержит 6 элементов UGUAN, расположенных в области -131/-1 нт, а вторая – 7 таких элементов, расположенных в области -119/-1 нт. Исследования показали, что связывание CFI с этими элементами повышает эффективность сборки комплекса полиаденилирования на обоих транскриптах, а также эффективность реакций расщепления и полиаденилирования. Yang et al. [348] показали, что димер CFI способен связываться с двумя элементами UGUAN, находящимися на разных расстояниях друг от друга, при этом участок РНК между элементами образует петлю. Авторы предположили, что связывание CFI с различными комбинациями трактов UGUAN может влиять на выбор альтернативных сайтов полиаденилирования.

Исследуя последовательности про-мРНК человека, представленные в нашей базе поли(А)-сайтов в области -200/-1 нт, мы обнаружили, что элементы UGUAN встречаются примерно в половине транскриптов, причем частота встречаемости незначительно варьирует среди мРНК, содержащих гексамеры разного типа (таблица 6.3). Элементы, расположенные в непосредственной близости от гексамера, встречаются всего лишь в 13 % про-мРНК. В области -70/-1 нт тракты UGUAN содержатся в общей сложности (включая близко расположенные тракты) в 28 % транскриптов. Количество элементов UGUAN в одной про-мРНК варьирует от 1 до 5. Чаще всего транскрипты содержат один (54 %) или два (28 %) элемента.

Полученные данные свидетельствуют о том, что CFI, по-видимому, стимулирует процесс полиаденилирования двумя способами. В первом случае, этот фактор совместно с CPSF и CstF участвует в сборке стабильного РНК-белкового комплекса на основном сайте полиаденилирования. Во втором случае, при наличии трактов UGUAN осуществляет дополнительную стабилизацию комплекса полиаденилирования за счет взаимодействия между этими трактами и субъединицей CFI25.

Таблица 6.3

Элементы UGUA в 244 про-мРНК человека

Тип гексамера	Количество про-мРНК	% про-мРНК, содержащих элемент UGUAN
I	168	55,9
II	36	47,2
III	19	52,6
IV-VI	21	38,1

Анализ последовательностей поли(А)-сайтов про-мРНК человека показывает, что даже основные поли(А)-сайты, состоящие из гексамера AAUAAA (или его вариантов) и U/GU-богатой «нижней» области, значительно различаются между собой [71, 349]. В таблице 6.4 мы для примера охарактеризовали восемь первых комплексных поли(А)-сайтов про-мРНК человека, представленных в нашей базе данных. Как видно из таблицы, эффективный или малоэффективный гексамер (например, типа I или III) может сочетаться с эффективным или менее эффективным «нижним» сигналом (например, содержащим несколько U/GU-пентамеров или один пентамер); каждый из этих сигналов может быть расположен оптимально или неоптимально по отношению к месту расщепления; оптимальным или неоптимальным может быть расстояние между сигналами. В частности, поли(А)-сайт про-мРНК ACTA2 содержит гексамер типа I и один «нижний» элемент URE, а поли(А) сайт про-мРНК ABCB7 содержит менее оптимальный гексамер типа II и два UREs. Расстояние между основными сигнальными элементами в поли(А)-сайте про-мРНК ACTA2 является оптимальным, а в поли(А)-сайте про-мРНК ACRBP, который содержит гексамер типа III и один тракт 2GU/U, – неоптимальным. Из таблицы 6.4 также видно, что поли(А)-сайты разных про-мРНК различаются по наличию вспомогательных сигнальных элементов.

Таблица 6.4

**Характеристика комплексных поли(А)-сайтов транскриптов из
базы поли(А)-сайтов про-мРНК человека**

Параметр	Про-мРНК			
	А	АВСВ7	АСRBP	АСТА2
Тип основного поли(А) сигнала	II	II	III	I
Локализация поли(А)-сигнала	оптимал	оптимал	оптимал	оптимал
Количество U/GU-трактов	0	2	1	1
Тип и длина U/GU-тракта	-	URE, 6 нт; URE, 8 нт	2GU/U, 5нт	URE, 5 нт
Локализация U/GU-трактов	-	неоптимал, оптимал	неоптимал	неоптимал
Расстояние между поли(А)-сигналом и U/GU-трактами	-	неоптимал, оптимал	неоптимал	оптимал
Альтернативные поли(А)-сайты	0	0	0	0
USE	4	4	4	3
G-квадруплексы	1 «нж»	0	0	0
i-мотив	0	0	1 «вх»	0
Другие параметры	-	-	-	-

Примечание. Оптимал – оптимальный, неоптимал – неоптимальный; «нж» - «нижний», «вх» - «верхний»

Продолжение таблицы 6.4

Параметр	Про-мРНК			
	ACTG2	ACY1	ADAM20	ADAMTS3
Тип основного поли(А) сигнала	II	VI	II	I
Локализация поли(А)-сигнала	оптималь	-	оптималь	оптималь
Количество U/GU-трактов	2	2	2	1
Тип и длина U/GU-тракта	2GU/U, 5 нт; URE, 9 нт	URE, 6 нт; URE, 6 нт	2GU/U, 7 нт; URE, 9 нт	URE, 6 нт
Локализация U/GU-трактов	неоптималь, оптималь	оптималь, оптималь	неоптималь, оптималь	оптималь
Расстояние между поли(А)-сигналом и U/GU-трактами	оптималь, оптималь	-	неоптималь, оптималь	оптималь
Альтернативные поли(А)-сайты	0	0	2	0
USE	1	2	3	4
G-квадруплексы	0	0	0	0
i-мотив	0	1 «ВХ»	0	0
Другие параметры	-	-	олиго(U), 19 нт	-

В настоящей работе не предсказывалась вторичная структура про-мРНК представленных в базе поли(А)-сайтов, поскольку для таких предсказаний требуется определенная дополнительная информация. Очевидно, что характеристика вторичной структуры элементов сайтов полиаденилирования

привела бы к дополнительной дифференциации этих сайтов по доступности сигнальных элементов для белковых факторов аппарата полиаденилирования.

6.2. Вспомогательные элементы сайтов полиаденилирования про-мРНК высших эукариотов

В обзоре литературы мы сообщили о том, что вспомогательные элементы можно разделить на 3 группы по механизму их воздействия на процесс полиаденилирования про-мРНК. USE и AuxDSE первой группы служат местами связывания для вспомогательных белков. Вспомогательные элементы второй группы служат местами связывания для основных белковых факторов аппарата полиаденилирования. USE и AuxDSE, относящиеся к третьей группе, играют структурную роль. Распределение поли(A)-сайтов про-мРНК из нашей базы данных по наличию некоторых вспомогательных элементов полиаденилирования показано в таблице 6.5.

USEs про-мРНК SV40 L и ряда клеточных транскриптов описываются консенсусом UAU₂₋₅GUNA [350]. USEs про-мРНК COX-2 также имеют значительное сходство с этой последовательностью [77]. Как отмечено в параграфе 1.2.1, с этими USEs специфически связываются четыре белка: p54^{nrb}, PTB, PSF и U1A, позитивно влияя на процесс полиаденилирования.

Мы выявили, что 8 % транскриптов в нашей базе поли(A)-сайтов имеют тракты, полностью описываемые консенсусом UAU₂₋₅GUNA или идентичные соответствующим трактам в мРНК COX-2 [77] и COL1A1 [350]. Если учесть транскрипты, которые имеют последовательности, близкие к консенсусу UAU₂₋₅GUNA (с малочисленными заменами оснований, встречающимися в известных USEs этого типа), то количество мРНК с данным USE составит 16 %. Чаще всего встречается тракт UAUUGUGA.

В нашей базе поли(A)-сайтов мы выявили 19 про-мРНК, содержащих места для связывания белка PTB (polypyrimidine tract binding protein), содержащего четыре мотива RRM [351]. Консенсусная последовательность

для взаимодействия РТВ с РНК представляет собой длинный полипиримидиновый тракт (15-25 нт), включающий элементы UCUU и/или CUCUCU. Этот белок может участвовать в процессе полиаденилирования, например, связываясь с другими белками или играя структурную роль.

Таблица 6.5

Вспомогательные элементы в базе поли(А)-сайтов про-мРНК человека

Элемент	Область, нт	Белок, связывающийся с трактом	% про-мРНК, содержащих данный элемент
UAU ₂₋₅ GUNA	-40/-200	p54 ^{nrb} , РТВ, PSF, U1A	8-16
UCUU, CUCUCU	-200/+200	РТВ	8
Три олиго(С)-тракта (≥ 3 нт)	-200/+200	αCPs	11
U-богатые тракты	-70/-1	CPSF	23
GGGA/U	-70/-1	гяРНП H, H'	23
GGGGGC	+1/+70		46
G-квадруплексы	-70/-1	гяРНП H, H'	3,7
	+1/+70		16
i-мотивы (≥ 5 пар C ⁺ -C)	-70/-1	-	3,7
	+1/+70		1,6

Четыре из найденных нами про-мРНК наряду с консенсусной последовательностью для связывания РТВ содержали также элемент UAU₂₋₅GUNA, с которым связывается комплекс белков, включающий и РТВ (таблица 6.5). Возможно, в этом случае предварительное взаимодействие РТВ со своим местом связывания облегчает сборку комплекса белков на UAU₂₋₅GUNA-тракте. В 11 других про-мРНК были найдены G-богатые тракты,

являющиеся местами связывания для белка гяРНП Н, причем в трех из этих мРНК G-богатые тракты могли образовывать G-квадруплексы. В этом случае, согласно [140], РТВ может стимулировать процесс полиаденилирования, непосредственно привлекая гяРНП Н к G-богатым трактам. Три про-мРНК содержали по два РТВ-связывающих тракта. Возможно, в этих транскриптах взаимодействие двух полипиримидиновых трактов с двумя или даже одной молекулой РТВ необходимо для сближения участков РНК, содержащих эти тракты, и экспонирования последовательности, расположенной между ними [351].

В параграфе 1.2.3 сообщалось о том, что С-богатые последовательности, включающие не менее трех коротких олиго(С)-трактов, могут связываться с белками α CPs и, в частности, служить в качестве USEs для ряда про-мРНК [121]. В нашей базе поли(А)-сайтов 28 транскриптов содержат С-богатые последовательности длиной 14-54 нт (чаще всего 17-30 нт), которые состоят из трех или более олиго(С)-трактов длиной не менее 3 нт, разделенных короткими линкерами из разных оснований. Десять таких последовательностей расположены в пределах 30-60 нт до основного или альтернативного сигналов полиаденилирования. Они могут выполнять функции USEs. Например, к ним относится последовательность ACCCGAGCCCCCGCGAUGCCCU (С-повторы подчеркнуты) в про-мРНК BAD, расположенная за 47 нт до PAS. Остальные последовательности, также могут служить в качестве USEs или участвовать в других процессах, например, в стабилизации мРНК.

Как упоминалось в параграфе 1.2.1 U-богатые USE служат дополнительным местом связывания для одного из основных факторов полиаденилирования – белка CPSF (вторая группа вспомогательных элементов). В случае про-мРНК ВИЧ-1 предварительное связывание CPSF с таким элементом способствует эффективной сборке аппарата полиаденилирования на поли(А)-сайте, в котором гексамер AAUAAA экспонирован не полностью. Консенсусная последовательность для U-богатых

USE не определена. Вероятно, их характерной особенностью является наличие олиго(U)-трактов. В нашей базе поли(A)-сайтов про-мРНК человека примерно половина транскриптов содержит олиго(U)-тракты длиной от 5 до 15 нуклеотидов в области, расположенной за 200 нт до места расщепления транскрипта. В области -70/-1 нт такие тракты встречаются в 23% про-мРНК. Чаще всего встречаются тракты длиной 5-8 нт.

Известные «нижние» вспомогательные элементы поли(A)-сайтов высших эукариотов менее многочисленны, чем «верхние». К ним, в частности, относятся G-богатые тракты, описанные в параграфе 1.2.4. Белок гяРНК Н/Н' связываясь с G-богатыми трактами стимулирует реакцию полиаденилирования. Места связывания гяРНК Н/Н' с РНК (GGGA/U или GGGGGC) встречаются в транскриптах из нашей базы данных очень часто. В областях -200/+200 и -70/+70 нт они найдены соответственно в 91 и 56 % транскриптов. G-богатые тракты встречаются чаще за местом расщепления, чем до него. В области -70/+70 нт количество мРНК с G-богатыми трактами, расположенными за местом расщепления, в два раза выше, чем количество мРНК с трактами, расположенными до этого места. В области -200/+200 нт превышение составляет всего лишь 1,2 раза. Количество G-богатых трактов в одном транскрипте также зависит от исследуемой области. В области -70/+70 нт только 7 % транскриптов с G-богатыми трактами содержат по одному такому тракту, остальные включают от 2 до 12 трактов (чаще всего по 2-8 трактов). В области -200/+200 нт количество транскриптов, содержащих только один G-богатый тракт, в 2,7 раза выше. Наличие большого количества G-богатых трактов в окрестности места расщепления транскрипта и их неравномерное распределение свидетельствуют в пользу того, что белок гяРНК Н/Н' играет большую роль в регулировании процесса полиаденилирования.

Следует отметить, что стимулирование реакции полиаденилирования G-богатыми трактами коррелирует со стабильностью комплексов, образованных этими трактами с белком гяРНК Н [135]. С другой стороны было показано, что

гяРПП Н стимулирует полиаденилирование про-мРНК р53, связываясь с G-богатыми трактами, структурированными в G-квадруплекс [8]. Учитывая, что GRS в про-мРНК SV40 L, образующий высокостабильный комплекс с белком гяРПП Н, способен сворачиваться в G-квадруплекс, можно предположить, что этот белок связывается с G-квадруplexами более прочно, чем с отдельными G-богатыми трактами и основную роль в стимулировании полиаденилирования играют именно те GRS, которые имеют потенциал для образования четырехцепочечных структур.

Поиск последовательностей, способных сворачиваться в G-квадруплексы, в нашей базе поли(A)-сайтов показал, что такие участки содержатся в 35 % про-мРНК, причем примерно половина из этих транскриптов содержат последовательности с потенциалом для образования квадруплексов, содержащих 3 или 4 тетрады, которые являются значительно более стабильными, чем структуры с двумя тетрадами. На рис. 6.4 показаны G-квадруплексы, которые могут быть образованы некоторыми транскриптами из нашей базы поли(A)-сайтов про-мРНК человека.

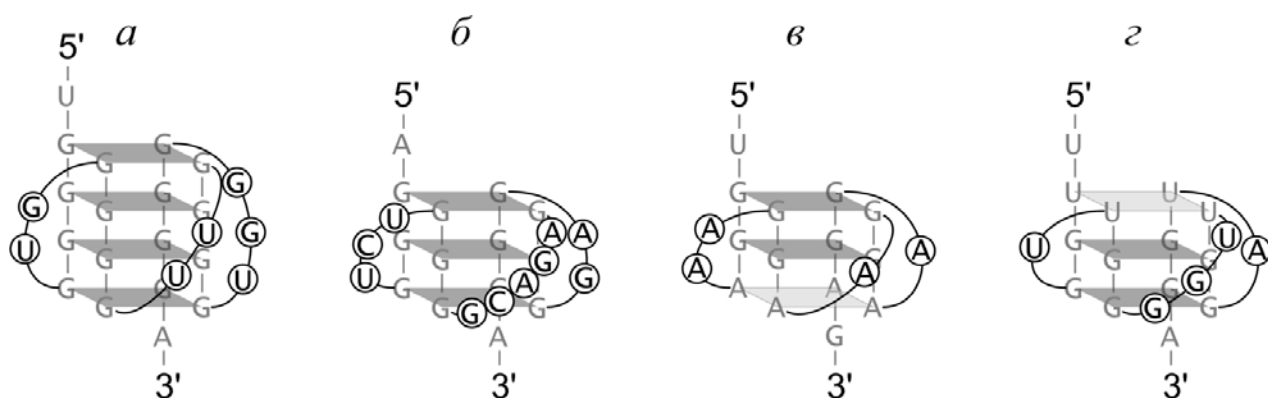


Рис. 6.4. Модели G-квадруплексов, образованных G-богатыми трактами некоторых транскриптов про-мРНК из базы поли(A)-сайтов про-мРНК человека. Разные тетрады выделены по-разному. Про-мРНК: *a* – CST, *б* – GAPD, *в* – DDX16, *г* – MPP2

Например, высокостабильные G-квадруплексы с четырьмя и тремя тетрадами могут быть образованы в про-мРНК CST и GAPD (рис. 6.4 а и б соответственно). Мы нашли в нашей базе ряд про-мРНК, G-богатые участки в которых могут образовывать квадруплексы, включающие другие тетрады, например квадруплексы с тетрадами A·A·A·A (рис. 6.4 в) или U·U·U·U (рис. 6.4 г). Аналогично G-богатым трактам, последовательности, способные сворачиваться в G-квадруплексы, встречаются в нашей базе поли(A)-сайтов в области, расположенной за местом расщепления транскрипта, чаще, чем в области, расположенной до этого места. Превышение в области -70/+70 нт более значительное (4,3 раза), чем в области -200/+200 нт (1,8 раза). Повышенное содержание кластеров G-повторов в области, расположенной непосредственно за местом расщепления, вряд ли случайно и может служить доказательством того, что G-квадруплексы играют определенную роль в процессе полиаденилирования про-мРНК высших эукариотов, возможно, как вспомогательные элементы поли(A)-сайтов.

Роль G-квадруплексов как вспомогательных элементов поли(A)-сайтов про-мРНК высших эукариотов была недавно подтверждена в работе [145]. Авторы провели биоинформационный анализ 33694 3'-нетранслируемых областей про-мРНК человека и нашли, что 15 % исследованных последовательностей содержат один или более тракт, которые способны сворачиваться в G-квадруплексы. Исследование отдельных G-квадруплексов показало, что они могут увеличивать эффективность полиаденилирования альтернативных поли(A)-сайтов. Следует отметить, что авторы работы [145] использовали для поиска квадруплексов часто применяемый алгоритм $N_{1-7}G_x-N_{1-7}G_x$, где $x \geq 3$ и N – любое основание [352, 353]. Этот алгоритм, как следует из параграфа 1.2.5, недооценивает число возможных квадруплексов: во-первых, стабильные G-квадруплексы РНК могут иметь не три, а только две тетрады, во-вторых, петли в структурах с числом тетрад ≥ 3 могут быть значительно длиннее, чем 7 нт и, в-третьих, G-квадруплексы могут включать смешанные тетрады. Кроме того, как показали недавние исследования [161],

G-квадруплексы могут содержать между тетрадами различные бальджи. Поэтому число различных G-квадруплексов в 3'-НТО про-мРНК высших эукариотов может быть значительно больше. Эти структуры могут участвовать не только в процессе полиаденилирования, но и в других процессах [145].

В заключение кратко обсудим роль кластеров С-трактов, способных сворачиваться в *i*-мотив. Как уже сообщалось в параграфе 1.1.2, для образования такой структуры С-тракты должны быть протонированы. Протонирование С-богатых областей РНК может осуществляться, как и в случае поли(А), под действием клеточных белков (см. параграф 5.1). Кроме того, образованию *i*-мотива в РНК и ДНК может способствовать эффект молекулярной «скученности» (crowding) [354].

В нашей базе поли(А)-сайтов 80 транскриптов содержат последовательности, состоящие из четырех и более олиго(С)-повторов длиной 2-8 нт, которые формально могут сворачиваться в *i*-мотив с четырьмя и более парами оснований С⁺-С и петлями длиной 2-7 нт. В 30 из этих про-мРНК могут образовываться *i*-мотивы с 5-9 парами С⁺-С и такой же длиной петель. Для примера на рис. 6.5 показаны две модели *i*-мотива для С-богатого тракта CCCCAGCCCACGUCCCCCACGACCCCCGCCCCCCCACCUCCCCCACGACCC С (С-повторы подчеркнуты), расположенного на расстоянии 147 нт за местом расщепления про-мРНК MLC1SA.

Количество С-богатых последовательностей с потенциалом для образования *i*-мотива, расположенных до места расщепления в 1,6 раз превышает количество таких последовательностей, расположенных за местом расщепления, а для *i*-мотивов с 5-9 парами С⁺-С – в 1,9 раза. Подобно G-квадруплексам *i*-мотивы РНК могут играть структурную роль, способствуя оптимальному сворачиванию последовательностей поли(А)-сайтов, при котором сигнальные элементы становятся доступны для соответствующих белков аппарата полиаденилирования. С другой стороны, если последовательность, образующая *i*-мотив, может служить местом связывания для белков α CPs, образование *i*-мотива будет препятствовать связыванию

этого белка с про-мРНК и стимуляции полиаденилирования, поскольку белки α CPs связываются с экспонированными олиго(С)-трактами [114]. Насколько нам известно, белки, взаимодействующие с i-мотивами РНК, пока еще не выявлены.

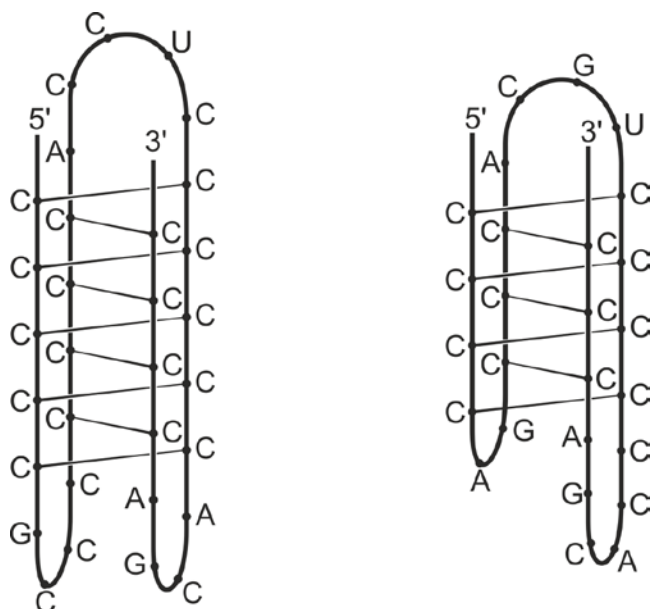


Рис. 6.5. Структурные модели i-мотива в про-мРНК MLC1SA

В небольших количествах мы нашли также и другие известные вспомогательные элементы сайтов полиаденилирования, функционирующие как в области, расположенной до места расщепления про-мРНК, так и за ним. Таким образом, анализ литературных данных и созданной нами базы поли(А)-сайтов про-мРНК человека показывает, что наличие вспомогательных элементов полиаденилирования не является специфической особенностью вирусных про-мРНК. USE и AuxDSE широко распространены и в про-мРНК высших эукариотов. Учитывая большое разнообразие основных поли(А)-сайтов про-мРНК высших эукариотов и наличие многочисленных вспомогательных элементов можно сделать вывод о том, что процесс полиаденилирования транскриптов является одним из главных уровней на которых происходит регуляция экспрессии генов наряду с такими процессами, как транскрипция, трансляция, контроль времени жизни мРНК и др.

В заключение следует обратить внимание на следующий факт. Как уже упоминалось выше, образование G- и C тетраплексов изменяет структуру области РНК, окружающей эти мотивы, и наличие таких участков следует учитывать при ее сворачивании. Однако не только тетраплексы влияют на вторичную структуру РНК. Таким свойством обладают и некоторые белки, специфически связанные с определенными одноцепочечными трактами. Например, белки α CPs, РТВ, а также гяРНП А/В и гяРНП F/Н [115, 351, 355] могут димеризоваться при связывании с отдаленно расположенными трактами и в результате экспонировать последовательность, расположенную между ними. Соответственно, сворачивание такой последовательности будет происходить автономно, а также изменится структура прилегающих областей, что также необходимо учитывать при определении вторичной структуры РНК.

6.3. Выводы к разделу 6

Создана база поли(А)-сайтов 244 про-мРНК человека. Показано, что как основные, так и комплексные сайты полиаденилирования значительно различаются между собой. Обнаружено, что различные вспомогательные элементы широко распространены в поли(А)-сайтах про-мРНК высших эукариотов. Найдено, что в области -70/+70 нт G-квадруплексы могут быть образованы примерно в 20 % про-мРНК человека, а i-мотивы – в 5 %. Полученные результаты свидетельствуют о том, что регуляция терминации транскрипции, т. е. полиаденилирования, является не менее важным процессом, чем регуляция ее инициации.

ЗАКЛЮЧЕНИЕ

Диссертационная работа посвящена исследованию структурных переходов в гомополирибонуклеотидах и их биологическому значению. Исследования были начаты в 80-е годы прошлого столетия. В это время синтетические гомополирибонуклеотиды поли(А), поли(У), поли(С) и поли(Г) еще рассматривались, в основном, как модельные полинуклеотиды для изучения природных РНК [3]. Несмотря на то, что наличие 3'-концевых поли(А)-последовательностей в мРНК разных организмов уже было установлено, их конкретные функции еще не были хорошо изучены, а работы, посвященные возможному участию G-квадруплексов ДНК и РНК в некоторых биологических процессах, только начали появляться. Интенсивное исследование синтетических гомополирибонуклеотидов различными методами выявило многие особенности их структуры, однако оставалось немало спорных и невыясненных вопросов. Требовалось дальнейшее изучение физико-химических свойств этих макромолекул с помощью новых методов и подходов, что могло бы оказаться полезным для выявления их функций в биологических процессах. Основные выводы из проведенных нами исследований представлены ниже.

Влияние ионной силы раствора на эффективные размеры полинуклеотидов. Поскольку электростатические эффекты, зависящие от ионных условий в растворе, во многом определяют структуру нуклеиновых кислот, их реактивность и способность связывать различные лиганды [213], на первом этапе работы было проведено исследование этих эффектов в полинуклеотидах на примере ДНК с помощью метода осмометрии. Для измерения осмотического давления ДНК был сконструирован прибор, обладающий достаточно малым временем установления равновесия. С его использованием мы впервые измерили осмотическое давление ДНК и экспериментально определили значения эффективного диаметра двойной спирали ДНК при промежуточных ионных силах раствора (0,01-0,1 М). При

уменьшении ионной силы в указанных пределах d_{eff} ДНК увеличивается примерно в три раза, а персистентная длина, согласно литературным данным, изменяется лишь незначительно, т. е. основной вклад в изменение объемных характеристик полинуклеотидов вносит именно изменение величины d_{eff} .

Для исследования структурных переходов в гомополирибонуклеотидах в диссертационной работе впервые использованы методы аналитического электрофореза и протонной буферной емкости. В обоих случаях измерения проводили при разных концентрациях соли в растворе. В электрофоретических экспериментах наблюдали значительное увеличение эффективных размеров (коэффициентов задержки) гомополирибонуклеотидов при уменьшении ионной силы, подобно зависимости d_{eff} ДНК от I . Исследование физико-химических характеристик гомополирибонуклеотидов при разных ионных силах оказалось в ряде случаев полезным для выявления различных форм этих макромолекул. Например, две различные двухспиральные формы поли(А) (DS и DS_p) обнаруживаются методом электрофореза только при достаточно низких ионных силах (< 30 мМ). DS_p -форма мигрирует быстрее, чем DS -форма при низких I , но при высоких I электрофоретические подвижности этих форм, по-видимому, близки, и они мигрируют в виде одной зоны.

Основные исследования в диссертационной работе проведены в растворах солей, содержащих моновалентные катионы. Однако степень нейтрализации отрицательно заряженных фосфатных групп нуклеиновых кислот зависит не только от концентрации соли, но и от величины заряда катионов. В связи с проблемой белково-нуклеинового узнавания электрофоретические эксперименты были проведены в присутствии эфиров различных аминокислот, однако чувствительность метода электрофореза оказалась достаточной только для регистрации взаимодействия между гомополимерами и дикатионами заряженных аминокислот. Наши исследования показали, что эфиры этих аминокислот оказывают специфическое действие на гомополимеры. Эфир лизина при больших

концентрациях вызывает уменьшение размеров поли(U), по-видимому, связанное с образованием коротких шпилек, а эфир аргинина индуцирует частичный переход *DS*-формы поли(A) в *DS_P*-форму. Кроме того, устойчивость комплексов производных заряженных аминокислот с гомополимерами зависит от вида полимера. Она убывает в ряду поли(G) > поли(A) ≥ поли(C) > поли(U), что согласуется с литературными данными [235, 236].

Структурные переходы в поли(A). Проведенное нами исследование поли(A) методом электрофореза подтвердило существование при кислых рН двух его различных двухспиральных форм, обнаруженное ранее другими методами. Как упоминалось выше, разделение двух форм наблюдается только при достаточно низких ионных силах ($I < 30$ мМ). Однако по литературным данным две кислые формы поли(A) сосуществуют в широком интервале ионных сил [27, 28, 39] и они находятся в динамическом равновесии [29]. Возможно, скорость перехода одной формы в другую возрастает с увеличением ионной силы. При $I > 30$ мМ обмен происходит достаточно быстро, вследствие чего наблюдается одна зона поли(A) с промежуточной электрофоретической подвижностью. А при более низкой ионной силе из-за значительного увеличения эффективных размеров обеих форм переход одной формы в другую может быть затруднен и наблюдается их разделение.

Метод протонной буферной емкости оказался наиболее подходящим из всех методов, примененных для изучения поли(A), для детального исследования картины переходов в этом гомополимере. В литературе нет единого мнения о порядке появления отдельных форм поли(A) в процессе протонирования. Например, согласно работам Adler et al. [27] и Janik et al. [28], в полупротонированном состоянии (при концентрации соли в растворе примерно 0,1 М) поли(A) находится в основном в *DS*-форме, которая образуется кооперативно и при дальнейшем протонировании постепенно переходит в *DS_P*-форму. В то же время по данным Lerner & Kearns [29]

примерно в таких же условиях поли(А) представляет собой промежуточную структуру, сахарофосфатный остов которой существует в двух конформациях, медленно переходящих одна в другую.

Анализ методом протонной буферной емкости кривых потенциометрического титрования поли(А), полученных в наших экспериментах и представленных в литературе, показал, что обе формы поли(А) образуются кооперативно, но с различной степенью кооперативности. В случае *DS*-формы параметр кооперативности γ значительно выше, чем в случае *DS_P*-формы при разных условиях. Порядок появления различных двухспиральных форм зависит от ионной силы раствора и температуры. При $I = 0,1$ М обе формы образуются практически одновременно. При высокой ионной силе и низкой температуре образование *DS_P*-формы происходит после образования *DS*-формы, а при низкой ионной силе и комнатной температуре сначала появляется *DS_P*-форма, а затем *DS*-форма. Из результатов наших исследований следует, что хотя при низких рН ($\sim 4,0$) поли(А) находится преимущественно в *DS_P*-форме, в широком интервале рН ($> 1,0$) обе формы присутствуют в сравнимых количествах. Поэтому логичнее рассматривать *DS*-форму как самостоятельную структуру, а не как промежуточное образование при протонировании поли(А).

Способность поли(А) образовывать две различные двухспиральные формы является уникальной среди нуклеиновых кислот. Ее физико-химическая природа до сих пор не установлена. Структурная модель *DS_P*-формы создана Rich et al. еще в 1961 г. [36], однако модель *DS*-формы в литературе до сих пор отсутствует. Некоторые параметры этой формы были получены ранее методом рентгеноструктурного анализа [26], но модель структуры двойной спирали не была представлена. В диссертационной работе мы предложили механизм образования *DS*-формы поли(А). Гипотеза о структуре двойной спирали *DS*-формы была выдвинута на основании того факта, что в модели одноцепочечной спирали поли(А) [34] наблюдаются довольно близкие контакты между аминоклуппами соседних оснований,

которые могут быть обусловлены наличием так называемых вертикальных водородных связей ...N6H...N6H...N6H... [248]. Мы предположили, что именно такие связи стабилизируют *DS*-форму поли(А). Главное отличие двух двухспиральных форм состоит в том, что в *DS_p*-форме один из протонов аминогруппы аденина участвует в образовании водородной связи с фосфатной группой, а в *DS*-форме этот же протон участвует в образовании вертикальной водородной связи, т. е. в обеих формах адениновые остатки участвуют в образовании трех водородных связей (две связи образуются между адениновыми остатками). Гипотеза об образовании вертикальных водородных связей между аминогруппами аденина позволила объяснить многие литературные данные, в частности, существование двух различных форм одноцепочечного поли(А).

Новые характеристики *DS*-формы поли(А), выявленные в работе, находятся в соответствии с нашей гипотезой о стабилизации этой формы вертикальными водородными связями. Кооперативный характер этих связей согласуется с высококооперативным образованием *DS*-формы при разных ионных силах и температурах. Частичный переход этой формы в *DS_p*-форму в присутствии дикатиона эфира аргинина или структурный переход в *DS'*-форму при увеличении степени протонирования поли(А) могут быть обусловлены внесением избыточного положительного заряда, который уменьшает степень структурной жесткости аминного фрагмента молекулы аденина, способствуя соответственно частичной диссоциации *DS*-формы или ее конформационному преобразованию.

Методом электрофореза были получены важные данные об еще одной кислой форме поли(А) – так называемой «замороженной» форме. Наличие этой формы, к сожалению, не учитывается во многих работах, посвященных исследованию поли(А), хотя предотвратить ее образование при кислых рН можно только с помощью специально принятых мер. Но даже в этих случаях нельзя быть уверенным в ее полном отсутствии. О том, что при подкислении поли(А) образуются агрегаты, сообщалось уже в ранних работах (например

[35]), однако природа этих образований не была раскрыта. Электрофорез поли(А) при кислых рН наглядно показал, что «замороженная» форма представляет собой структуры разных размеров, включая такие, которые не способны проникнуть в гель, и эта форма не диссоциирует на отдельные двухцепочечные молекулы во время электрофореза. Кроме того, мы выявили, что количество «замороженной» формы поли(А) в растворе увеличивается при увеличении длины молекул и повышении ионной силы, т. е. при увеличении гибкости полимерной цепи.

Учитывая литературные и наши данные о свойствах «замороженной» формы поли(А), мы предположили, что, вероятнее всего, она возникает при кислых рН в результате одновременного образования как межмолекулярных, так и внутримолекулярных двойных спиралей олиго(А). При наличии дуплексов обоих типов в одной и той же цепи поли(А) преобразование объемной структуры в межмолекулярные двойные спирали должно быть значительно затруднено и ее диссоциация будет происходить лишь при прогреве или подщелачивании раствора.

Наше предположение о том, что поли(А) может образовывать внутримолекулярные спирали при самосворачивании, является вполне адекватным, если учесть структурные особенности других гомополимеров. Известно, что тракты олиго(Г) могут образовывать не только межмолекулярные, но и внутримолекулярные G-квадруплексы [146], а тракты олиго(С) образуют i-мотивы как из четырех разных молекул, так и внутри одной и той же молекулы при ее сворачивании [57]. Мы высказали предположение о самосворачивании поли(А) для того, чтобы объяснить появление агрегатов поли(А) при кислых рН, обнаруживаемое методом электрофореза. А Kim et al. [241] недавно высказали предположение о самосворачивании одноцепочечного поли(dА) при кислых рН для объяснения появления компактной формы этого полимера в разбавленных растворах.

Учитывая, что самоассоциаты поли(А) образуются чрезвычайно легко, т. е. их образование является специфической особенностью данного

гомополирибонуклеотида, следует предполагать, что они могут иметь определенное биологическое значение. Предложенные нами модели вовлечения олиго(А)-дуплексов в различные биологические процессы обсуждаются ниже.

Структурные переходы в поли(С). Так же как и в случае поли(А), структурные переходы в поли(С) вызываются протонированием оснований. Однако в поли(А) присоединяющиеся к основаниям протоны не участвуют в образовании водородных связей между ними, а в поли(С) – участвуют, причем при степени протонирования полимера 50 % все водородные связи между основаниями уже образованы. Вследствие этого структурные переходы в поли(С) значительно отличаются от переходов в поли(А). Единое мнение о механизме протонирования полупротонированной двойной спирали поли(С) в литературе отсутствует.

Использование метода протонной буферной емкости для анализа кривых потенциометрического титрования поли(С) позволило детально изучить динамику структурных переходов в этом полинуклеотиде. В частности, мы выявили, что при подкислении раствора происходят два различных процесса некооперативного протонирования поли(С). На основании результатов, полученных методом протонной буферной емкости, а также анализа литературы мы предложили новый механизм структурных перестроек в поли(С). Протонирование атомов азота N3 цитозиновых оснований индуцирует высококооперативный процесс образования двойных спиралей поли(С) из одноцепочечных структур. После этого продолжается некооперативное протонирование дуплексов по положению атомов азота N3, еще не вовлеченных в образование пар оснований. Затем происходит некооперативное протонирование двойных спиралей по положению атома кислорода O2 электронейтральных цитозиновых остатков. Эта реакция снижает термодинамическую стабильность дуплексов поли(С) и при дальнейшем увеличении степени протонирования приводит к их диссоциации.

Кооперативная диссоциация двойных спиралей сопровождается полным протонированием одноцепочечных структур по положению атомов азота N3.

При исследовании структурных переходов в поли(C) методом электрофореза мы впервые обнаружили наличие двух различных форм протонированного полинуклеотида. Это важное наблюдение, поскольку сосуществование двух кислых форм поли(C) не принималось во внимание в литературе ни при анализе различных спектров этого полимера [56, 271, 272], ни при анализе дифрактограмм [49]. Быстрее мигрирующую форму поли(C) мы отнесли к двойной спирали, а медленнее мигрирующую форму идентифицировали как четырехцепочечный i-мотив. Учитывая, что соотношение двух форм практически не зависело от pH, мы предположили, что i-мотив образуется только во время перехода поли(C) из нейтральной в кислую форму, когда в растворе одновременно присутствуют и одноцепочечные, и двухцепочечные структуры. Четырехцепочечная структура формируется на основе дуплексов посредством интеркаляции одиночных цепей в двойные спирали. Две разные кислые формы поли(C) не регистрируются методом протонной буферной емкости, вероятно, вследствие того, что i-мотив состоит из двух интеркалированных двойных спиралей и механизмы протонирования оснований в отдельной двухцепочечной структуре или в дуплексах, входящих в состав четырехцепочечной структуры являются одинаковыми.

Структурные переходы в поли(G). Длинные поли(G)-участки встречаются в природных РНК редко. Насколько нам известно, они найдены лишь в составе «минус»-цепей геномных РНК пикорнавирусов [110]. Основную роль в биологических системах играют кластеры олиго(G)-трактов, способные, как и поли(G), образовывать G-квадруплексы. Однако поли(G) наряду с другими гомополирибонуклеотидами часто используется как инструмент для выявления специфичности РНК-связывающих белков. При этом обычно не указывается, в какой форме, одноцепочечной или четырехцепочечной, находится этот полинуклеотид. При нейтральных pH

структурные переходы в поли(G) индуцируются изменением концентрации соли, но литературные данные о том, при каких именно концентрациях соли происходит переход поли(G) из односпиральной формы в G-квадруплекс, являются противоречивыми.

Проведенное нами исследование структурных переходов в поли(G) методом электрофореза показало, что при концентрации моновалентных катионов > 10 mM поли(G) уже находится в форме G-квадруплекса. Следовательно, если специфическое связывание белка с поли(G) обнаруживается при средних и высоких ионных силах, это означает, что он специфически связывается именно с G-квадруплексами. Однако это не означает, что белок будет иметь высокое сродство также и к одноцепочечным олиго(G)-трактам в РНК.

Нами впервые получены электронные спектры поглощения красителя толуидинового голубого, связанного с гомополирибонуклеотидами, в агарозных гелях. Показано, что форма спектра зависит как от гомополимера, так и от его форм, что подтверждает регистрацию различных форм поли(C) и поли(G) методом электрофореза. В частности, $\Delta\lambda_{\max}$ спектров поглощения ТГ в комплексе с одноцепочечной и четырехцепочечной формами поли(G) составляет примерно 20 нм.

Таким образом, гомополирибонуклеотиды (и гомоолигонуклеотиды соответствующей длины) имеют высокий потенциал к разнообразным изменениям. Учет их структурного многообразия может оказаться полезным при выяснении детальных механизмов биологических процессов, происходящих с участием многочисленных белков, специфически связывающихся с этими макромолекулами. Следует также отметить, что поскольку гомополирибонуклеотиды рассматриваются как потенциальные терапевтические агенты или мишени, детальные сведения об их структурных характеристиках, полученные в данной работе, могут быть полезными при создании соответствующих лекарственных препаратов.

Молекулярные модели участия двойных спиралей олиго(А) в биологических процессах. Учитывая, что адениновые основания в поли(А) могут быть протонированы не только в кислой среде, но и при физиологических условиях, посредством переноса протонов от кислых аминокислотных остатков белков, мы предположили, что двойные спирали поли(А) могут образовываться и в клетке. Анализируя литературу, мы выявили ряд биологических процессов с участием поли/олиго(А)-последовательностей, в которых олиго(А)-тракты могли бы функционировать в форме внутримолекулярной двойной спирали.

В частности, неизвестны механизмы стабилизации ARE-содержащих мРНК (например мРНК GAP-43 и мРНК β -казеина), белками, относящими к семейству Hu/ELAV. Уникальной особенностью этих белков является преимущественное связывание с длинными молекулами поли(А) [6]. Мы предположили, что Hu-белки индуцируют образование внутримолекулярного дуплекса в поли(А)-хвосте, который препятствуют действию поли(А)-нуклеаз.

В другом случае, не полностью выяснен механизм процесса саморегуляции белка PABPC. При избытке PABPC в цитоплазме он вместе с другими белками связывается с А-богатым участком, расположенным в 5' НТО своей мРНК. Связывание белков с А-богатыми трактами препятствует продвижению 40 S рибосомной частицы по направлению к иницирующему кодону. Мы предположили, что в этом случае донорами протонов для образования внутримолекулярной двойной спирали олиго(А) являются белки, взаимодействующие с PABPC, а задержка движения рибосомной частицы происходит вследствие того, что внутримолекулярный дуплекс с параллельно ориентированными цепями, по-видимому, не является субстратом для функционирования геликаз, участвующих в процессе инициации трансляции.

В третьем случае, до сих пор не полностью выяснено, почему процессивный синтез поли(А)-хвоста про-мРНК высших эукариотов прекращается, когда его длина достигает примерно 250 нт. Причем этот эффект может быть воспроизведен в системе *in vitro* в присутствии всего трех

белков: PAP, CPSF и PABPN1 [10]. Мы предположили, что при достижении именно этой длины создаются условия для образования внутримолекулярной двойной спирали в поли(А)-хвосте за счет переноса протонов от чрезвычайно кислого N-концевого домена PABPN1. Образование дуплекса приводит к диссоциации белкового комплекса и к переходу от процессивной реакции полиаденилирования к дистрибутивной.

Впервые наши модели образования внутримолекулярных двойных спиралей олиго(А) в природных РНК и их участия в биологических процессах были опубликованы в 1998 г. Анализ современной литературы показал, что эти модели не потеряли своей актуальности и в настоящее время и что дуплексы олиго(А) могут быть вовлечены также в другие биологические процессы.

Молекулярные модели участия G-квадруплексов в биологических процессах. В диссертационной работе рассмотрен вопрос о возможном участии G-квадруплексов в димеризации генома ВИЧ-1 и в процессах полиаденилирования про-мРНК высших эукариотов и их вирусов. Вопрос о роли G-квадруплексов в димеризации генома ВИЧ-1 в настоящее время полностью не выяснен. В ранних работах было высказано предположение о том, что димеризация генома ВИЧ-1 осуществляется посредством образования межмолекулярных G-квадруплексов G-богатыми последовательностями, расположенными в 5'-концевой области гена *gag* (например [166, 167]). Позднее выяснилось, что инициация димеризации происходит через взаимодействие шпилек DIS [168, 170], однако контакты между двумя геномами могут быть локализованы и в других местах [164].

Проведя филогенетическое исследование нескольких сотен геномных последовательностей ВИЧ-1 разных субтипов, мы выявили, что за шпилькой DIS расположен ряд консервативных G-богатых трактов с потенциалом для образования димерных G-квадруплексов. Мы представили структурные модели области димеризации генома ВИЧ-1, в которых дуплекс, образованный шпильками DIS, стабилизирован G-квадруплексами, образованными

консервативными G-богатыми трактами. G-квадруплексы в этих моделях обеспечивают параллельную ориентацию областей геномных РНК, расположенных до и за DLS, как это наблюдается при исследовании димерных РНК ВИЧ-1 методом электронной микроскопии.

Мы предположили, что димерные G-квадруплексы включаются в DLS только на поздней стадии димеризации геномной РНК ВИЧ-1. Их образование происходит под действием нуклеокапсидного белка в зрелой вирусной частице. При этом димерные G-квадруплексы могут образовываться не только вблизи шпильки DIS, но и, как предполагают авторы работ [322, 323], в других местах генома. G-квадруплексы, образованные непосредственно за шпилькой DIS, будут способствовать сохранению контакта между геномами во время всего синтеза минус цепи ДНК. Сохранение пространственной близости между геномами, обеспеченное образованием димерных G-квадруплексов в разных местах генома, необходимо также для процесса рекомбинации.

Мы первые высказали предположение о том, что G-квадруплексы могут принимать участие в процессе полиаденилирования про-мРНК высших эукариотов и их вирусов и предложили механизм их участия в полиаденилировании про-мРНК SV40 L. В поли(А)-сайте этой про-мРНК имеется G-богатый участок, расположенный непосредственно за основным U-богатым «нижним» элементом. С этим GRS взаимодействует белок мяРНК Н, стимулируя процесс полиаденилирования [134]. Учитывая, что GRS может сворачиваться в G-квадруплекс, мы предположили, что стимуляция синтеза поли(А)-хвоста в этом случае может быть, связана не только со стабилизацией под действием белка мяРНК Н комплекса белков, взаимодействующих с основными сигналами полиаденилирования, но и со структурными перестройками транскрипта. Предсказание вторичной структуры участка про-мРНК SV40 L, охватывающего основной поли(А)-сайт, показало, что гексамер AAUAAA и URE, частично или полностью вовлеченные в образование пар оснований в отсутствие G-квадруплекса, переходят при его образовании в одноцепочечное состояние, что облегчает их узнавание белками аппарата

полиаденилирования. Белок мяРПП H, возможно, индуцирует образование G-квадруплекса и/или стабилизирует эту структуру. Таким образом, возможность образования квадруплексов G-богатыми трактами в РНК следует учитывать при компьютерном предсказании вторичной структуры различных регуляторных элементов в про-мРНК/мРНК. Подобную структурную роль могут играть также i-мотивы, образованные C-богатыми трактами.

В диссертационной работе впервые представлена структурная модель комплексного поли(A)-сайта про-мРНК ВИЧ-1 и модель почти полной 3' нетранслируемой области этого транскрипта. Создание модели основано на выявлении в геномной РНК ВИЧ-1 консервативного G-богатого участка в 3' НТО с большим потенциалом для образования различных G-квадруплексов. GRS расположен в начале 3' НТО и является минорным вспомогательным элементом поли(A)-сайта про-мРНК ВИЧ-1 [100].

Мы предположили, что при сворачивании этого участка в G-квадруплекс расположенная за ним область транскрипта может сворачиваться автономно, т. е. последний нуклеотид GRS служит 5'-границей комплексного поли(A)-сайта. Определив таким образом участок для сворачивания, мы предсказали вторичную структуру поли(A)-сайта для более 1400 изолятов ВИЧ-1 разных субтипов и рекомбинантов и выявили две альтернативные структуры (линейную и доменную), которые чаще всего встречаются в оптимальных структурах сворачиваемого участка или в близких по энергии субоптимальных. Мы представили эти структуры в качестве моделей комплексного поли(A)-сайта про-мРНК ВИЧ-1. Они содержат известные шпильки TAR и polyA, а также найденные нами новые структурные элементы: домен USE и шпильку DSE. U-богатый тракт в домене USE и U/GU-богатый элемент в шпильке DSE находятся в одноцепочечном состоянии и доступны для взаимодействия с клеточными белками аппарата полиаденилирования. Особенностью предложенных моделей является то, что мы впервые предложили для G-квадруплексов роль определителей границ

самодостаточных доменов при предсказании вторичной структуры длинных РНК.

Следует также отметить, что, в процессе создания моделей области димеризации генома ВИЧ-1 и комплексного поли(А)-сайта про-мРНК этого вируса были впервые проведены широкомасштабные филогенетические исследования отдельных структурных элементов про-мРНК ВИЧ-1, таких как домен USE и шпильки TAR, polyA, DSE и DIS. Данные об общих и специфических чертах этих элементов могут быть использованы при конструировании терапевтических антиретровирусных агентов, направленных на подавление функций сигнальных элементов, существенных для репликации вируса.

База поли(А)-сайтов про-мРНК человека. Для проверки нашего предположения о том, что G-квадруплексы могут играть роль вспомогательных элементов сайтов полиаденилирования высших эукариотов, мы создали и проанализировали базу поли(А)-сайтов про-мРНК человека. База содержит 244 геномные последовательности, охватывающие участки, расположенные за 200 нт до места расщепления транскрипта и 200 нт за этим местом. В базу включены последовательности, содержащие только точно определенные сайты полиаденилирования.

Анализ базы показал, что даже основные поли(А)-сайты, охватывающие только поли(А)-сигнал (гексамер AAUAAA и его производные) и «нижний» элемент DSE, значительно различаются между собой. Они различны по последовательности и количеству основных элементов, их расположению относительно места расщепления транскрипта и взаимному расположению и, вероятно, по степени их экспонирования (вторичная структура в работе не определялась). Все эти параметры в значительной степени влияют на эффективность реакции полиаденилирования. Кроме того, оказалось, что сайты полиаденилирования содержат многочисленные вспомогательные элементы, которые в основном были ранее обнаружены в вирусных про-мРНК. Мы также выявили, что не менее 16% транскриптов содержат в области

длиной 70 нт, расположенной сразу за местом расщепления, G-богатые тракты с потенциалом для образования G-квадруплексов, а ~ 4 % транскриптов содержат C-богатые тракты с потенциалом для образования i-мотива в области длиной 70 нт, расположенной до места расщепления. G- и C-тетраплексы могут, в частности, играть структурную роль в процессе полиаденилирования про-мРНК и других процессах.

Таким образом, каждый комплексный поли(A)-сайт про-мРНК человека фактически является специфическим. Это означает, что процесс терминации транскрипции является не менее сложным, чем процесс ее инициации и, что полиаденилирование наряду с транскрипцией, сплайсингом и другими процессами является одним из главных уровней, на которых происходит регуляция экспрессии генов.

ВЫВОДЫ

Впервые методами аналитического электрофореза и протонной буферной емкости изучены структурные переходы в гомополирибонуклеотидах – выявлены новые характеристики переходов, предложены и обоснованы механизмы образования ряда форм, а также представлены структурные модели участия гомоолигонуклеотидных трактов в ряде биологических процессах. Полученные результаты важны для дальнейшего выяснения детальных механизмов биологических процессов, основанных на взаимодействии широко распространенных клеточных белков с регуляторными трактами гомополи(олиго) нуклеотидов про-мРНК и мРНК.

1. С целью выявления влияния ионной силы на эффективные размеры полинуклеотидов впервые экспериментально определено значение эффективного диаметра двойной спирали ДНК $d_{эфф}$ при промежуточных ионных силах I . При уменьшении I от 100 до 10 мМ величина $d_{эфф}$ увеличивается примерно в 3 раза. Измерения проведены с помощью сконструированного нами осмометра, обладающего достаточно малым временем установления равновесия.
2. Методом аналитического электрофореза впервые показано что:
 - DS -форма поли(А) имеет меньшую электрофоретическую подвижность, чем форма DS_p при $I < 30$ мМ, а при более высоких I значения подвижностей сближаются. Дикатион эфира аргинина способствует переходу формы DS в DS_p -форму;
 - эффективность образования «замороженной» формы поли(А) возрастает при увеличении ионной силы и размера макромолекул, т.е. в условиях, при которых повышается их гибкость;
 - поли(С) при кислых рН существует в двух различных формах, соотношение которых не меняется с изменением рН;

- поли(G) при ионной силе ниже 10 мМ присутствует в двух формах, которые были дополнительно зарегистрированы методом спектроскопии при исследовании электронных спектров поглощения толуидинового голубого, связанного с гомополимерами в агарозных гелях. Медленнее мигрирующая форма отнесена к G-квадруплексам, а быстрее мигрирующая – к одноцепочечной форме;
3. Методом протонной буферной емкости:
- определены параметры структурных переходов (pK , концентрация способных к ионизации атомных групп и параметр кооперативности) в поли(A) и поли(C) при разных ионных силах и температурах;
 - зарегистрирован третий структурный переход в поли(A), наличие которого было предположено ранее на основании данных ЯМР-спектроскопии;
 - впервые зарегистрирован дополнительный процесс протонирования двойной спирали поли(C), относящийся к протонированию атомов кислорода в непротонированных основаниях, который индуцирует ее распад.
4. На основании полученных в работе экспериментальных результатов, выдвинуты и обоснованы следующие предположения:
- две разные односпиральные формы поли(A) и две его разные двухспиральные формы отличаются наличием или отсутствием вертикальных водородных связей $\dots N_6-H\dots N_6-H\dots N_6-H\dots$ между соседними основаниями;
 - появление «замороженной» формы поли(A) обусловлено одновременным образованием как межмолекулярных, так и внутримолекулярных двойных спиралей в протонированном гомополимере;
 - медленнее мигрирующая кислая форма поли(C) является *i*-мотивом. Она образуется только во время структурного перехода путем формирования тетраплексов из двойных спиралей и одиночных цепей.

5. Впервые предложены и обоснованы молекулярные модели участия внутримолекулярных двойных спиралей олиго(A) в процессах стабилизации ARE-содержащих мРНК белками, относящихся к семейству Hu/ELAV, саморегуляции синтеза цитоплазматического поли(A)-связывающего белка, терминация полиаденилирования про-мРНК и др.
6. Впервые предложены и обоснованы следующие структурные модели с участием G-квадруплексов:
 - модель области димеризации геномной РНК ВИЧ-1, в которой димер, образованный шпильками DIS, стабилизирован квадруплексами, содержащими как G-тетрады, так и тетрады из смешанных оснований.
 - модель поли(A)-сайта про-мРНК вируса SV-40, в которой G-квадруплекс играет структурную роль, обеспечивая доступность основных сигнальных элементов для белкового аппарата полиаденилирования;
 - модель комплексного поли(A)-сайта про-мРНК ВИЧ-1, а также почти всего 3'-НТО транскрипта. В этом случае G-квадруплексы выполняют функцию 5'-границы независимо сворачивающегося домена, охватывающего эти области.
7. Впервые проведено широкомасштабное (~ 1500 геномов) филогенетическое исследование структурных элементов сайта полиаденилирования про-мРНК ВИЧ-1. Эти данные могут быть полезными при конструировании терапевтических противовирусных агентов, направленных на подавление функций сигнальных элементов, существенных для репликации вирусного генома.
8. Впервые создана и проанализирована база сайтов полиаденилирования 244 про-мРНК человека в области -200/+200 нт. Установлено, что ~ 35 % про-мРНК содержат в этой области последовательности, способные образовывать G-квадруплексы, а ~ 12 % транскриптов включают тракты,

способные образовывать i-мотив. Кроме того, показано, что большинство транскриптов содержат различные вспомогательные элементы сайтов полиаденилирования (USEs и AuxDSEs). Эти результаты свидетельствуют о том, что регуляция терминации транскрипции, т. е. полиаденилирования, является не менее важным процессом, чем регуляция ее инициации.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Jacobs E. The role of RNA structure in posttranscriptional regulation of gene expression / E. Jacobs, J. D. Mills, M. Janitz // *J. Genet. Genomics*. – 2012. – Vol. 39, N 10. – P. 535-543.
2. Mortimer S. A. Insights into RNA structure and function from genome-wide studies / S. A. Mortimer, M. A. Kidwell, J. A. Doudna // *Nat. Rev. Genet.* – 2014. – Vol. 15, N 7. – P. 469–479.
3. Зенгер В. Принципы структурной организации нуклеиновых кислот / В. Зенгер. – Москва: Мир, 1987. – 584 с.
4. Eliseeva I. A. Poly(A)-binding proteins: structure, domain organization, and activity regulation / I. A. Eliseeva, D. N. Lyabin, L. P. Ovchinnikov // *Biochemistry (Mosc)*. – 2013. – Vol. 78, N 13. – P. 1377-1391.
5. Makeyev A. V. The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms / A. V. Makeyev, S. A. Liebhaber // *RNA*. – 2002. – Vol. 8, N 3. – P. 265–278.
6. Ma W.-J. The Elav-like proteins bind to AU-rich elements and to the poly(A) tail of mRNA / W.-J. Ma, S. Chung, H. Furneaux // *Nucleic Acids Res.* – 1997. – Vol. 25, N 18. – P. 3564–3569.
7. Caputi M. Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family / M. Caputi, A. M. Zahler // *J. Biol. Chem.* – 2001. – Vol. 276, N 47. – P. 43850–43859.
8. Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3'-end processing and function during DNA damage / [A. Decorsière, A. Cayrel, S. Vagner, S. Millevoi] // *Genes Dev.* – 2011. – Vol. 25, N 3. – P. 220–225.
9. Brook M. The role of mammalian poly(A)-binding proteins in co-ordinating mRNA turnover / M. Brook, N. K. Gray // *Biochem. Soc. Trans.* – 2012. – Vol. 40, N 4. – P. 856–864.

10. Eckmann C. R. Control of poly(A) tail length / C. R. Eckmann, C. Rammelt, E. Wahle // *Wiley Interdiscip. Rev. RNA*. – 2011. – Vol. 2, N 3. – P. 348–361.
11. Patel G. P. IMP1 interacts with poly(A)-binding protein (PABP) and the autoregulatory translational control element of PABP-mRNA through the KH III-IV domain / G. P. Patel, J. Bag // *FEBS J*. – 2006. – Vol. 273, N 24. – P. 5678–5690.
12. Porschke D. The dynamics of nucleic-acid single-strand conformation changes oligo- and polyriboadenylic acids / D. Porschke // *Eur. J. Biochem.* – 1973. – Vol. 39, N 1. – P. 117–126.
13. Porschke D. Molecular states in single-stranded adenylate chains by relaxation analysis / D. Porschke // *Biopolymers*. – 1978. – Vol. 17, N 2. – P. 315–323.
14. Dewey T. G Laser temperature-jump study of stacking in adenylic acid polymers / T. G. Dewey, D. H. Turner // *Biochemistry*. – 1979. – Vol. 18, N 26. – P. 5757–5762.
15. Kononov A. I. Luminescence excitation spectra reveal low-lying excited states in stacked adenine bases / A. I. Kononov, M. N. Bukina // *J. Biomol. Struct. Dyn.* – 2002. – Vol. 20, N 3. – P. 465–471.
16. Arnott S. Structure of the single-stranded polyribonucleotide polycytidylic acid / S. Arnott, R. Chandrasekaran, A. G. W. Leslie // *J. Mol. Biol.* – 1976. – Vol. 106, N 3. – P. 735–748.
17. Broido M. S. ^1H NMR evidence for a left-handed helical structure of poly(ribocytidylic acid) in neutral solution / M. S. Broido, D. R. Kearns // *J. Am. Chem. Soc.* – 1982. – Vol. 104, N 19. – P. 5207–5216.
18. Broyde S. The coil form of poly(rU): A model composed of minimum energy conformers that matches experimental properties / S. Broyde, B. Hingerty // In: “Nucleic acid: geometry and dynamics”; edited by R. H. Sarma. – NY: Pergamon Press, 1980. – P. 367-382.

19. Protonatable hairpins are conserved in the 5'-untranslated region of tymovirus RNAs / [K. Hellendoorn, P. J. Michiels, R. Buitenhuis, C. W. Pleij] // *Nucleic Acids Res.* – 1996. – Vol. 24, N 24. – P. 4910–4917.
20. Model considerations on physico-chemical nature of protein-nucleic acid contacts through amino acid carboxylic groups: spectroscopic data / [S. P. Samijlenko, I. M. Kolomiets, I. V. Kondratyuk, A. V. Stepanyugin] // *Biopolym. Cell.* – 1998. – Vol. 14, N 1. – P. 47–53.
21. Denman R. B. Déjà vu all over again: FMRP binds U-rich target mRNAs / R. B. Denman // *Biochem. Biophys. Res. Commun.* – 2003. – Vol. 310, N 1. – P. 1–7.
22. Adepoju L. J. Antitumor activity of polyuridylic acid in human soft tissue and bone sarcomas / L. J. Adepoju, J. D. Geiger // *J. Surg. Res.* – 2010. – Vol. 164, N 1. – e107–114.
23. Giri P. Molecular recognition of poly(A) targeting by protoberberine alkaloids: *in vitro* biophysical studies and biological perspectives / P. Giri, G. Suresh Kumar // *Mol. BioSyst.* – 2010. – Vol. 6, N 1. – P. 81–88.
24. Research progress of RNA quadruplex / [X. Ji, H. Sun, H. Zhou et al.] // *Nucleic Acid Ther.* – 2011. – Vol. 21, N 3. – P. 185–200.
25. Self-assembly of single-stranded RNA on carbon nanotube: polyadenylic acid to form a duplex structure / [C. Zhao, Y. Peng, Y. Song et al.] // *Small.* – 2008. – Vol. 4, N 5. – P. 656–661.
26. Finch J. T. Two double helical forms of polyriboadenylic acid and the pH-dependent transition between them / J. T. Finch, A. Klug // *J. Mol. Biol.* – 1969. – Vol. 46, N 3. – P. 597–598.
27. Adler A. J. Polyriboadenylic and polydeoxyriboadenylic acids. Optical rotatory studies of pH-dependent conformations and their relative stability / A. J. Adler, L. Grossman, G. D. Fasman // *Biochemistry.* – 1969. – Vol. 8, N 9. – P. 3846–3859.

28. Janik B. Polarography of polynucleotides. II. Conformations of poly(adenylic acid) at acidic pH / B. Janik, R. G. Sommer, A. M. Bobst // *Biochim. Biophys. Acta.* – 1972. – Vol. 281, N 2. – P. 152–168.
29. Lerner D. B. Proton and phosphorus NMR investigation of the conformational states of acid polyadenylic double helix / D. B. Lerner, D. R. Kearns // *Biopolymers.* – 1981. – Vol. 20, N 4. – P. 803–816.
30. Guéron M. The i-motif in nucleic acids / M. Guéron, J.-L. Leroy // *Curr. Opin. Struct. Biol.* – 2000. – Vol. 10, N 3. – P. 326–331.
31. Souleil C. Immunochemistry of polyribonucleotides. Study of polyriboinosinic and polyriboguanilyc acids / C. Souleil, J. Panijel // *Biochemistry.* – 1968. – Vol. 7, N 1. – P. 7–13.
32. Raman studies of nucleic acids. IX: a salt-induced structural transition in poly(rG) / [J. Rice, L. Lafleur, G. C. Medeiros, G. J. Thomas Jr.] // *J. Raman Spectrosc.* – 1973. – Vol. 1, N 2. – P. 207–215.
33. Структура полирибогуаниловой кислоты в растворе / [Е. А. Лесник, И. М. Кочкина, А. С. Тихоненко, Я. М. Варшавский] // *Молекул. биол.* – 1980. – Т. 14, № 4. – С. 820–829.
34. Saenger W. A structural model for the polyadenylic acid single helix / W. Saenger, J. Riecke, D. Suck // *J. Mol. Biol.* – 1975. – Vol. 93, N 4. – P. 529–534.
35. Steiner R. F. Polynucleotides. VI. The influence of various factors upon the structural transition of polyriboadenylic acid at acid pH's / R. F. Steiner, R. F. Beers Jr. // *Biochim. Biophys. Acta.* – 1959. – Vol. 32, N 1. – P. 166–176.
36. The molecular structure of polyadenylic acid / [A. Rich, D. R. Davies, F. H. Crick, J. D. Watson] // *J. Mol. Biol.* – 1961. – Vol. 3, N 1. – P. 71–86.
37. Bobst A. M. Effect of the methylation of the 2'-hydroxyl groups in polyadenylic acid on its structure in weakly acidic and neutral solutions and on its capability to form ordered complexes with polyuridylic acid / A. M.

- Bobst, F. Rottman, P. A. Cerutti // *J. Mol. Biol.* – 1969. – Vol. 46, N 2. – P. 221–234.
38. Scovell W. M. Structural and conformational studies of polyriboadenylic acid in neutral and acid solution / W. M. Scovell // *Biopolymers.* – 1978. – Vol. 17, N, 4. – P. 969–984.
39. Кузнецов И. А. Демонстрация трех протонированных форм поли(А): потенциометрическое и кондуктометрическое исследование изоионных растворов / И. А. Кузнецов, О. В. Воронцова // *Молекул. биол.* – 1984. – Т. 18, № 4. – С. 1032–1041.
40. Holcomb D. N. Temperature dependence of the hydrogen ion equilibria in poly(riboadenylic acid) / D. N. Holcomb, S. N. Timasheff // *Biopolymers.* – 1968. – Vol. 6, N 4. – P. 513–529.
41. Bobst A. M. Studies on spin-labeled polyriboadenylic acid / A. M. Bobst // *Biopolymers.* – 1972. – Vol. 11, N 7. – P. 1421–1433.
42. Morgan R. S. “Alkaline” polyadenylic acid / R. S. Morgan, R. Byrne // *J. Mol. Biol.* – 1959. – Vol. 1, N 2. – P. 188–189.
43. Antao V. P. CD spectral comparisons of the acid-induced structures of poly[d(A)], poly[r(A)], poly-[d(C)], and poly[r(C)] / V. P. Antao, D. M. Gray // *J. Biomol. Struct. Dyn.* – 1993. – Vol. 10, N 5. – P. 819–839.
44. Akinrimisi E. O. Properties of helical polycytidylic acid / E. O. Akinrimisi, C. Sander, P. O. P. Ts'o // *Biochemistry.* – 1963. – Vol. 2, N 2. – P. 340–344.
45. Hartman K. A. The tautomeric form of helical polyribocytidylic acid / K. A. Hartman, A. Rich // *J. Amer. Chem. Soc.* – 1965. – Vol. 87, N 9. – P. 2033–2039.
46. Guschlbauer W. Protonated polynucleotide structures. I. The thermal denaturation of polycytidylic acid in acid solution / W. Guschlbauer // *Proc. Natl Acad. Sci. USA.* – 1967. – Vol. 57, N 5. – P. 1441–1448.
47. Klump H. A calorimetric study of the helix-coil transition of poly(cytidylic acid) in acid solution / H. Klump // *Biochim. Biophys. Acta.* – 1975. – Vol. 383, N 1. – P. 1–8.

48. Chou C. H. Raman spectral studies of nucleic acids. XVI. Structures of polyribocytidylic acid in aqueous solution / C. H. Chou, G. J. Thomas Jr. // *Biopolymers*. – 1977. – Vol. 16, N 4. – P. 765–789.
49. Langridge R. Molecular structure of helical polycytidylic acid / R. Langridge, A. Rich // *Nature*. – 1963. – Vol. 198, N 4882. – P. 725–728.
50. Structure and hydration of polycytidylic acid from the data of infrared spectroscopy, EHF dielectrometry and computer modeling / [V. Maleev, M. Semenov, V. Kashpur et al.] // *J. Mol. Struct.* – 2002. – Vol. 605, N 1. – P. 51–61.
51. Сулейманова Р. С. Кислотно-основные свойства изоионных растворов полирибозитидиловой кислоты / Р. С. Сулейманова, Н. В. Аполонник, И. А. Кузнецов // *Биофизика*. – 1987. – Т. 32, № 3. – С. 413–416.
52. Аполонник Н. В. Кондуктометрическое подтверждение существования полностью протонированной двойной спирали полирибозитидиловой кислоты / Н. В. Аполонник, Р. С. Сулейманова, И. А. Кузнецов // *Молекул. биол.* – 1987. – Т. 21, № 2. – С. 428–433.
53. Inman R. B. Transitions of DNA homopolymers / R. B. Inman // *J. Mol. Biol.* – 1964. – Vol. 9, N 3. – P. 624–637.
54. Gehring K. A tetrameric DNA structure with protonated cytosine · cytosine base pairs / K. Gehring, J.-L. Leroy, M. Guéron // *Nature*. – 1993. – Vol. 363, N 6429. – P. 561–565.
55. Benevides J. M. Raman signature of the four-stranded intercalated cytosine motif in crystal and solution structures of DNA deoxycytidylates d(CCCT) and d(C8) / J. M. Benevides, C. Kang, G. J. Thomas Jr. // *Biochemistry*. – 1996. – Vol. 35, N 18. – P. 5747–5755.
56. Vibrational circular dichroism signature of hemiprotonated intercalated four-stranded i-DNA / [D. Tsankov, M. Krasteva, V. Andrushchenko et al.] // *Biophys. Chem.* – 2006. – Vol. 119, N 1. – P. 1–6.
57. Snoussi K. The RNA i-motif / K. Snoussi, S. Nonin-Lecomte, J.-L. Leroy // *J. Mol. Biol.* – 2001. – Vol. 309, N 1. – P. 139–153.

58. Fresco J. R. Polynucleotides. V. Helix-coil transition of polyriboguanilyc acid / J. R. Fresco, J. Massoulié // *J. Am. Chem. Soc.* – 1963. – Vol. 85, N 9 – P. 1352–1353.
59. Лесник Е. А. Исследование влияния ионной силы и температуры на структурное состояние полирибогуаниловой кислоты в растворе методом кинетики $^1\text{H} \rightarrow ^3\text{H}$ -обмена / Е. А. Лесник, Р. Н. Маслова, Я. М. Варшавский // *Молекул. биол.* – 1981. – Т. 15, № 1. – С. 161–166.
60. Thrierr J. C. A study of polyuridylic acid. / J. C. Thrierr, M. Dourlent, M. Leng // *J. Mol. Biol.* – 1971. – Vol. 58, N 3. – P. 815–830.
61. Michelson A. M. Polynucleotides. VIII. Base stacking in polyuridylic acid / A. M. Michelson, C. Monny // *Proc. Natl Acad. Sci. USA.* – 1966. – Vol. 56, N 5. – P. 1528–1534.
62. Young P. R. Secondary structure in polyuridylic acid. Non-classical hydrogen bonding and the function of the ribose 2'-hydroxyl group / P. R. Young, N. R. Kallenbach // *J. Mol. Biol.* – 1978. – Vol. 126, N 3. – P. 467–479.
63. Szer W. Ordered state of poly-uridylic acid above room temperature / W. Szer // *J. Mol. Biol.* – 1966. – Vol. 16, N 2. – P. 585–587.
64. Carroll D. Complexes of polylysine with polyuridylic acid and other polynucleotides / D. Carroll // *Biochemistry.* – 1972. – Vol. 11, N 3. – P. 426–433.
65. Зарудная М. И. Полиаденилирование про-мРНК. 1. Образование 3'-концов мРНК позвоночных / М. И. Зарудная // *Biopolym. Cell.* – 2001. – Vol. 17, N 2. – P. 93-108.
66. Зарудная М. И. Полиаденилирование про-мРНК. 2. Образование 3'-концевых поли(А)-последовательностей мРНК дрожжей, растений, прокариотов и вирусов / М. И. Зарудная // *Biopolym. Cell.* – 2001. – Vol. 17, N 3. – P. 185-202.
67. Chan S. Pre-mRNA 3'-end processing complex assembly and function / S. Chan, E.-A. Choi, Y. Shi // *WIREs RNA.* – 2011. – Vol. 2, N 3. – P. 321–335.

68. Ryan K. Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease / K. Ryan, O. Calvo, J. L. Manley // *RNA*. – 2004. – Vol. 10, N 4. – P. 565–573.
69. Tian B. Signals for pre-mRNA cleavage and polyadenylation / B. Tian, J. H. Graber // *WIREs RNA*. – 2012. – Vol. 3, N 3. – P. 385–396.
70. Sheets M. D. Point mutations in AAUAAA and the poly(A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation *in vitro* / M. D. Sheets, S. C. Ogg, M. P. Wickens // *Nucleic Acids Res.* – 1990. – Vol. 18, N 19. – P. 5799–5805.
71. Auxiliary elements of mammalian pre-mRNAs polyadenylation signals / [M. I. Zarudnaya, A. L. Potyahaylo, I. M. Kolomiets, D. M. Hovorun] // *Biopolym. Cell.* – 2002. – Vol. 18, N 6. – P. 500–517.
72. Downstream elements of mammalian pre-mRNA polyadenylation signals: primary, secondary and higher-order structures / [M. I. Zarudnaya, I. M. Kolomiets, A. L. Potyahaylo, D. M. Hovorun] // *Nucleic Acids Res.* – 2003. – Vol. 31, N 5. – P. 1375–1386.
73. Chen F. Cleavage site determinants in the mammalian polyadenylation signal / F. Chen, C. C. MacDonald, J. Wilusz // *Nucleic Acids Res.* – 1995. – Vol. 23, N 14. – P. 2614–2620.
74. Millevoi S. Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation / S. Millevoi, S. Vagner // *Nucleic Acids Res.* – 2010. – Vol. 38, N 9. – P. 2757–2774.
75. Proudfoot N. J. Ending the message: poly(A) signals then and now / N. J. Proudfoot // *Genes Dev.* – 2011. – Vol. 25, N 17. – P. 1770–1782.
76. Зарудная М. И. Молекулярные механизмы сопряжения процессов терминации транскрипции и полиаденилирования про-мРНК / [М. И. Зарудная, А. Л. Потягайло, Н. Э. Дзержинский, Д. Н. Говорун] // *Ukr. Biochem. J.* – 2001. – Vol. 73, N 2. – P. 28–32.

77. Alternative polyadenylation of cyclooxygenase-2 / [T. Hall-Pogar, H. Zhang, B. Tian, C. S. Lutz] // *Nucleic Acids Res.* – 2005. – Vol. 33, N 8. – P. 2565–2579.
78. Specific *trans*-acting proteins interact with auxiliary RNA polyadenylation elements in the COX-2 3'-UTR / [T. Hall-Pogar, S. Liang, L. K. Hague, C. S. Lutz] // *RNA.* – 2007. – Vol. 13, N 7. – P. 1103–1115.
79. Involvement of the carboxyl terminus of vertebrate poly(A) polymerase in U1A autoregulation and in the coupling of splicing and polyadenylation / [S. I. Gunderson, S. Vagner, M. Polycarpou-Schwarz, I. W. Mattaj] // *Genes & Dev.* – 1997. – Vol. 11, N 6. – P. 761–773.
80. The human papillomavirus type 31 late 3' untranslated region contains a complex bipartite negative regulatory element / [S. A. Cumming, C. E. Repellin, M. McPhillips et al.] // *J. Virol.* – 2002. – Vol. 76, N 12. – P. 5993–6003.
81. The ability of the HIV-1 AAUAAA signal to bind polyadenylation factors is controlled by local RNA structure / [B. I. F. Klasens, M. Thiesen, A. Virtanen, B. Berkhout] // *Nucleic Acids Res.* – 1999. – Vol. 27, N 2. – P. 446–454.
82. Sittler A. The secondary structure of the adenovirus-2 L4 polyadenylation domain: evidence for a hairpin structure exposing the AAUAAA signal in its loop / A. Sittler, H. Gallinaro, M. Jacob // *J. Mol. Biol.* – 1995. – Vol. 248, N 3. – P. 525–540.
83. Chen F. Auxiliary downstream elements are required for efficient polyadenylation of mammalian pre-mRNAs / F. Chen, J. Wilusz // *Nucleic Acids Res.* – 1998. – Vol. 26, N 12. – P. 2891–2898.
84. Yang Q. Structural biology of poly(A) site definition / Q. Yang, S. Doublé // *WIREs RNA.* – 2011. – Vol. 2, N 5. – P. 732–747.
85. Mandel C. R. Protein factors in pre-mRNA 3'-end processing / C. R. Mandel, Y. Bai, L. Tong // *Cell. Mol. Life Sci.* – 2008. – Vol. 65, N 7–8. – P. 1099–1122.

86. Зарудная М. И. Действительно ли до сих пор не обнаружена эндонуклеаза, расщепляющая про-мРНК в реакции полиаденилирования? / М. И. Зарудная, Д. Н. Говорун // Ukr. Biochem. J. – 2001. – Vol. 73, N 5. – P. 128–134.
87. Zarudnaya M. I. What nuclease cleaves pre-mRNA in the process of polyadenylation? / M. I. Zarudnaya, I. M. Kolomiets, D. M. Hovorun // IUBMB Life. – 2002. – Vol. 54, N 1. – P. 27-31.
88. The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins / [S. M. Barabino, W. Hübner, A. Jenny et al.] // Genes Dev. – 1997. – Vol. 11, N 13. – P. 1703–1716.
89. Arabidopsis mRNA polyadenylation machinery: comprehensive analysis of protein-protein interactions and gene expression profiling / [A. G. Hunt, R. Xu, B. Addepalli et al.] // BMC Genomics. – 2008. – Vol. 9:220.
90. Hunt A. G. Plant polyadenylation factors: conservation and variety in the polyadenylation complex in plants / A. G. Hunt, D. Xing, Q. Q. Li // BMC Genomics. – 2012. – Vol. 13:641.
91. Kapp L. D. The molecular mechanics of eukaryotic translation / L. D. Kapp, J. R. Lorsch // Annu. Rev. Biochem. – 2004. – Vol. 73. – P. 657–704.
92. Polyadenylation of genomic RNA and initiation of antigenomic RNA in a positive-strand RNA virus are controlled by the same *cis*-element / [M. J. van Ooij, C. Polacek, D. H. Glaudemans et al.] // Nucleic Acids Res. – 2006. – Vol. 34, N 10. – P. 2953–2965.
93. Direct evidence that the poly(A) tail of influenza A virus mRNA is synthesized by reiterative copying of a U track in the virion RNA template / [L. L. M. Poon, D. C. Pritlove, E. Fodor, G. G. Brownlee] // J. Virol. – 1999. – Vol. 73, N 4. – P. 3473–3476.
94. A polyadenylation-specific RNA-contact site on the surface of the bifunctional vaccinia virus RNA modifying protein VP39 that is distinct from

- the mRNA 5' end-binding "cleft" / [L. Deng, L. Johnson, J. M. Neveu et al.] // *J. Mol. Biol.* – 1999. – Vol. 285, N 4. – P. 1417–1427.
95. Lu C. Characterization of small nontranslated polyadenylylated RNAs in vaccinia virus-infected cells / C. Lu, R. Bablanian // *Proc. Natl Acad. Sci. USA.* – 1996. – Vol. 93, N 5. – P. 2037–2042.
96. Kelly T. J. SV40 DNA replication / T. J. Kelly // *J. Biol. Chem.* – 1988. – Vol. 263, N 34. – P. 17889–17892.
97. Carswell S. Efficiency of utilization of the simian virus 40 late polyadenylation site: effects of upstream sequences / S. Carswell, J. C. Alwine // *Mol. Cell. Biol.* – 1989. – Vol. 9, N 10. – P. 4248–4258.
98. Structural model of the complete poly(A) region of HIV-1 pre-mRNA / [M. I. Zarudnaya, A. L. Potyahaylo, I. M. Kolomiets, D. M. Hovorun] // *J. Biomol. Struct. Dyn.* – 2013. – Vol. 31, N 10. – P. 1044–1056.
99. Ashe M. P. The HIV-1 5' LTR poly(A) site is inactivated by U 1 snRNP interaction with the downstream major splice donor site / M. P. Ashe, L. H. Pearson, N. J. Proudfoot // *EMBO J.* – 1997. – Vol. 16, N 18. – P. 5752–5763.
100. The human immunodeficiency virus type 1 polyadenylation signal: a 3' long terminal repeat element upstream of the AUAAA necessary for efficient polyadenylation / [A. Valsamakis, S. Zeichner, S. Carswell, J. C. Alwine] // *Proc. Natl Acad. Sci. USA.* – 1991. – Vol. 88, N 6. – P. 2108–2112.
101. DeZazzo J. D. Involvement of long terminal repeat U3 sequences overlapping the transcription control region in human immunodeficiency virus type 1 mRNA 3' end formation / J. D. DeZazzo, J. E. Kilpatrick, M. J. Imperiale // *Mol. Cell. Biol.* – 1991. – Vol. 11, N 3. – P. 1624–1630.
102. Abbink T. E. A novel long distance base-pairing interaction in human immunodeficiency virus type 1 RNA occludes the Gag start codon / T. E. Abbink, B. Berkhout // *J. Biol. Chem.* – 2003. – Vol. 278, N 13. – P. 11601–11611.

103. High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states / [K. A. Wilkinson, R. J. Gorelick, S. M. Vasa et al.] // *PLoS Biol.* – 2008. – Vol. 6, N 4: e96.
104. Architecture and secondary structure of an entire HIV-1 RNA genome / [J. M. Watts, K. K. Dang, R. J. Gorelick et al.] // *Nature.* – 2009. – Vol. 460, N 7256. – P. 711–716.
105. Comparison of SIV and HIV-1 genomic RNA structures reveals impact of sequence evolution on conserved and non-conserved structural motifs / [E. Pollom, K. K. Dang, E. L. Potter et al.] // *PLoS Pathogens.* – 2013. – Vol. 9, N 4: e1003294.
106. Gee A. H. Structural differentiation of the HIV-1 polyA signals / A. H. Gee, W. Kasprzak, B. A. Shapiro // *J. Biomol. Struct. Dyn.* – 2006. – Vol. 23, N 4. – P. 417–428.
107. Circularization of the HIV-1 RNA genome / [M. Ooms, T. E. Abbink, C. Pham, B. Berkhout] // *Nucleic Acids Res.* – 2007. – Vol. 35, N 15. – P. 5253–5261.
108. CPSF recognition of an HIV-1 mRNA 3'-processing enhancer: multiple sequence contacts involved in poly(A) site definition / [G. M. Gilmartin, E. S. Fleming, J. Oetjen, B. R. Graveley] // *Genes Dev.* – 1995. – Vol. 9, N 1. – P. 72–83.
109. Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly (A) polymerase / [I. Kaufmann, G. Martin, A. Friedlein et al.] // *EMBO J.* – 2004. – Vol. 23, N 3. – P. 616–626.
110. Sequence and location of the poly C tract in aphtho- and cardiovirus RNA / [D. N. Black, P. Stephenson, D. J. Rowlands, F. Brown] // *Nucleic Acids Res.* – 1979. – Vol. 6, N 7. – P. 2381–2390.
111. Carocci M. The encephalomyocarditis virus / M. Carocci, L. Bakkali-Kassimi // *Virulence.* – 2012 – Vol. 3, N 4. – P. 351-367.

112. The 3' end of the foot-and-mouth disease virus genome establishes two distinct long-range RNA-RNA interactions with the 5' end region / [P. Serrano, M. R. Pulido, M. Sáiz, E. Martínez-Salas] // J. Gen. Virol. – 2006. – Vol. 87, N 10. – P. 3013–3022.
113. Murray K. E. Poly(rC) binding proteins mediate poliovirus mRNA stability / K. E. Murray, A. W. Roberts, D. J. Barton // RNA. – 2001. – Vol. 7, N 8. – P. 1126–1141.
114. Thisted T. Optimized RNA targets of two closely related triple KH domain proteins, heterogeneous nuclear ribonucleoprotein K and α -CP-2KL, suggest distinct modes of RNA recognition / T. Thisted, D. L. Lyakhov, S. A. Liebhaber // J. Biol. Chem. – 2001. – Vol. 276, N 20. – P. 17484-17496.
115. Structure of a construct of a human poly(C)-binding protein containing the first and second KH domains reveals insights into its regulatory mechanisms / [Z. Du, S. Fenn, R. Tjhen, T. L. James] // J. Biol. Chem. – 2008. – Vol. 283, N 42. – P. 28757–28766.
116. Gamarnik A. V. Interactions of viral protein 3CD and poly(rC) binding protein with the 5' untranslated region of the poliovirus genome / A. V. Gamarnik, R. Andino // J. Virol. – 2000. – Vol. 74, N 5. – P. 2219-2226.
117. Poly(rC)-binding protein 2 interacts with the oligo(rC) tract of coxsackievirus B3 / [R. Zell, Y. Ihle, S. Seitz et al.] // Biochem. Biophys. Res. Commun. – 2008. – Vol. 366, N 4. – P. 917-921.
118. Isolation and characterization of a folate receptor mRNA-binding *trans*-factor from human placenta. Evidence favoring identity with heterogeneous nuclear ribonucleoprotein E1 / [X. Xiao, Y.-S. Tang, J. Y. Mackins et al.] // J. Biol. Chem. – 2001. – Vol. 276, N 44. – P. 41510-41517.
119. Identification of a C-rich element as a novel cytoplasmic polyadenylation element in *Xenopus* embryos / [L. Paillard, D. Maniey, P. Lachaume et al.] // Mech. Dev. – 2000. – Vol. 93, N 1–2. – P. 117–125.

120. Ji X. An RNA-protein complex links enhanced nuclear 3' processing with cytoplasmic mRNA stabilization / X. Ji, J. Kong, S. A. Liebhaber // EMBO J. – 2011. – Vol. 30, N 13. – P. 2622-2633.
121. α CP Poly(C) binding proteins act as global regulators of alternative polyadenylation / [X. Ji, J. Wan, M. Vishnu et al.] // Mol. Cell. Biol. – 2013. – Vol. 33, N 13. – P. 2560-2573.
122. The 3' untranslated region complex involved in stabilization of human α -globin mRNA assembles in the nucleus and serves an independent role as a splice enhancer / [X. Ji, J. Kong, R. P. Carstens, S. A. Liebhaber] // Mol. Cell. Biol. – 2007. – Vol. 27, N 9, P. 3290-3302.
123. Inability of RNA to form the i-motif: implications for triplex formation // [L. Lacroix, J.-L. Mergny, J.-L. Leroy, C. Hélène] Biochemistry. – 1996. – Vol. 35, N 26. – P. 8715-8722.
124. Brooks T. A. Making sense of G-quadruplex and i-motif functions in oncogene promoters / T. A. Brooks, S. Kendrick, L. Hurley // FEBS J. – 2010. – Vol. 277, N 17. – P. 3459–3469.
125. Day H. A. i-Motif DNA: structure, stability and targeting with ligands / H. A. Day, P. Pavlou, Z. A. Waller // Bioorg. Med. Chem. – 2014. – Vol. 22, N 16. – P. 4407–4418.
126. Solution equilibria of cytosine- and guanine-rich sequences near the promoter region of the *n-myc* gene that contain stable hairpins within lateral loops / [S. Benabou, R. Ferreira, A. Aviñó et al.] // Biochim. Biophys. Acta. – 2014. – Vol. 1840, N 1. – P. 41–52.
127. I-motif structures formed in the human c-MYC promoter are highly dynamic—insights into sequence redundancy and i-motif stability / [J. Dai, E. Hatzakis, L. H. Hurley, D. Yang] // PLoS One. – 2010. – Vol. 5, N 7: e11647.
128. Mullen T. E. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5' / T. E. Mullen, W. F. Marzluff // Genes Dev. – 2008. – Vol. 22, N 1. – P. 50–65.

129. Scott D. D. RNA decay *via* 3' uridylation / D. D. Scott, C. J. Norbury // *Biochim. Biophys. Acta.* – 2013. – Vol. 1829, N 6–7. – P 654–665.
130. Gallouzi I. E. A distinctively novel exoribonuclease that really likes U / I. E. Gallouzi, J. Wilusz // *EMBO J.* – 2013. – Vol. 32, N 13. – P. 1799–1801.
131. Park-Lee S. Characterization of the interaction between neuronal RNA-binding protein HuD and AU-rich RNA / S. Park-Lee, S. Kim, I. A. Laird-Offringa // *J. Biol. Chem.* – 2003. – Vol. 278, N 41. – P. 39801–39808.
132. Simon R. Further analysis of cytoplasmic polyadenylation in *Xenopus* embryos and identification of embryonic cytoplasmic polyadenylation element-binding proteins / R. Simon, J. D. Richter // *Mol. Cell. Biol.* – 1994. – Vol. 14, N 12. – P. 7867–7875.
133. CPEB3 and CPEB4 in neurons: analysis of RNA-binding specificity and translational control of AMPA receptor GluR2 mRNA / [Y.-S. Huang, M.-C. Kan, C.-L. Lin, J. D. Richter] // *EMBO J.* – 2006. – Vol. 25, N 20. – P. 4865–4876.
134. Bagga P. S. DSEF-1 is a member of the hnRNP H family of RNA-binding proteins and stimulates pre-mRNA cleavage and polyadenylation *in vitro* / P. S. Bagga, G. K. Arhin, J. Wilusz // *Nucleic Acids Res.* – 1998. – Vol. 26, N 23. – P. 5343–5350.
135. Downstream sequence elements with different affinities for the hnRNP H/H' protein influence the processing efficiency of mammalian polyadenylation signals / [G. K. Arhin, M. Boots, P. S. Bagga et al.] // *Nucl. Acids Res.* – 2002. – Vol. 30, N 8. – P. 1842–1850.
136. Fogel B. L. Efficient polyadenylation of Rous sarcoma virus RNA requires the negative regulator of splicing element / B. L. Fogel, L. M. McNally, M.T. McNally // *Nucleic Acids Res.* – 2002. – Vol. 30, N 3. – P. 810–817.
137. hnRNP F influences binding of a 64-kilodalton subunit of cleavage stimulation factor to mRNA precursors in mouse B cells / [K. L. Veraldi, G.

- K. Arhin, K. Martincic et al.] // *Mol. Cell. Biol.* – 2001. – Vol. 21, N 4. – P. 1228–1238.
138. The high kinetic stability of a G-quadruplex limits hnRNP F qRRM3 binding to G-tract RNA / [B. Samatanga, C. Dominguez, I. Jelesarov, F. H. Allain] // *Nucleic Acids Res* – 2013. – Vol. 41, N 4. – P. 2505–2516.
139. Caputi M. SR proteins and hnRNP H regulate the splicing of the HIV-1 *tev*-specific exon 6D / M. Caputi, A. M. Zahler // *EMBO J.* – 2002. – Vol. 21, N 4. – P. 845–855.
140. A physical and functional link between splicing factors promotes pre-mRNA 3' end processing / [S. Millevoi, A. Decorsière, C. Loulergue et al.] // *Nucleic Acids Res.* – 2009. – Vol. 37, N 14. – P. 4672–4683.
141. Shafer R. H. Biological aspects of DNA/RNA quadruplexes / R. H. Shafer, I. Smirnov // *Biopolymers.* – 2001. – Vol. 56, N 3. – P. 209–227.
142. Christiansen J. A guanosine quadruplex and two stable hairpins flank a major cleavage site in insulin-like growth factor II mRNA / J. Christiansen, M. Kofod, F. C. Nielsen // *Nucleic Acids Res.* – 1994. – Vol. 22, N 25. – P. 5709–5716.
143. Millevoi S. G-quadruplexes in RNA biology / S. Millevoi, H. Moine, S. Vagner // *WIREs RNA.* – 2012. – Vol. 3, N 4. – P. 495–507.
144. Bugaut A. 5'-UTR RNA G-quadruplexes: translation regulation and targeting / A. Bugaut, S. Balasubramanian // *Nucleic Acids Res.* – 2012. – Vol. 40, N 11. – P. 4727–4741.
145. Beaudoin J. D. Exploring mRNA 3'-UTR G-quadruplexes: evidence of roles in both alternative polyadenylation and mRNA shortening / J.-D. Beaudoin, J.-P. Perreault // *Nucleic Acids Res.* – 2013. – Vol. 41, N 11. – P. 5898–5911.
146. Simonsson T. G-quadruplex DNA structures – variations on a theme / T. Simonsson // *Biol. Chem.* – 2001. – Vol. 382, N 4. – P. 621–628.

147. A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex / [G. W. Collie, S. M. Haider, S. Neidle, G. N. Parkinson] // *Nucleic Acids Res.* – 2010. – Vol. 38, N 16. – P. 5569–5580.
148. Structures of guanine-rich and cytosine-rich quadruplexes formed *in vitro* by telomeric, centromeric, and triplet repeat disease DNA sequences / [D. J. Patel, S. Bouaziz, A. Kettani, Y. Wang] // In: “Oxford handbook of nucleic acid structure”; edited by S. Neidle. – Oxford: Oxford University press, 1999. – P. 389–453.
149. Field- and phage-induced dipolar couplings in a homodimeric DNA quadruplex: relative orientation of G·(C-A) triad and G-tetrad motifs and direct determination of C2 symmetry axis orientation / [H. M. Al-Hashimi, A. Majumdar, A. Gorin et al.] // *J. Am. Chem. Soc.* – 2001. – Vol. 123, N 4. – P. 633–640.
150. A double chain reversal loop and two diagonal loops define the architecture of a unimolecular DNA quadruplex containing a pair of stacked G (*syn*)·G (*syn*)·G (*anti*)·G (*anti*) tetrads flanked by a G·(T-T) triad and a T·T·T triple / [V. Kuryavyi, A. Majumdar, A. Shallop et al.] // *J. Mol. Biol.* – 2001. – Vol. 310, N 1. – P. 181–194.
151. A diamond-shaped zipper-like DNA architecture containing triads sandwiched between mismatches and tetrads / [V. Kuryavyi, A. Kettani, W. Wang et al.] // *J. Mol. Biol.* – 2000. – Vol. 295, N 3. – P. 455–469.
152. Dimeric DNA quadruplex containing major groove-aligned A·T·A·T and G·C·G·C tetrads stabilised by inter-subunit Watson-Crick A·T and G·C pairs / [N. Zhang, A. Gorin, A. Majumdar et al.] // *J. Mol. Biol.* – 2001. – Vol. 312, N 5. – P. 1073–1088.
153. Patel P. K. NMR studies on truncated sequences of human telomeric DNA: observation of a novel A-tetrad / P. K. Patel, A. S. R. Koti, R. V. Hosur // *Nucleic Acids Res.* – 1999. – Vol. 27, N 19. – P. 3836–3843.

154. Patel P. K., NMR observation of a novel C-tetrad in the structure of the SV 40 repeat sequence GGGCGG / P. K. Patel, N. S. Bhavesh, R. V. Hosur // *Biochem. Biophys. Res. Commun.* – 2000. – Vol. 270, N 3. – P. 967–971.
155. Patel P. K. NMR observation of T-tetrads in a parallel stranded DNA quadruplex formed by *Saccharomyces cerevisiae* telomere repeats / P. K. Patel, R. V. Hosur // *Nucleic Acids Res.* – 1999. – Vol. 27, N 12. – P. 2457–2464.
156. Cheong C. Solution structure of an unusually stable RNA tetraplex containing G- and U-quartet structures / C. Cheong, P. B. Moore // *Biochemistry.* – 1992. – Vol. 31, N 36. – P. 8406–8414.
157. Jing N. Structure-activity of tetrad-forming oligonucleotides as a potent anti-HIV therapeutic drug / N. Jing, M. E. Hogan // *J. Biol. Chem.* – 1998. – Vol. 273, N 52. – P. 34992–34999.
158. A dimeric RNA quadruplex architecture comprised of two G:G(:A):G:G(:A) hexads, G:G:G:G tetrads and UUUU loops / [H. Liu, A. Matsugami, M. Katahira, S. Uesugi] // *J. Mol. Biol.* – 2002. – Vol. 322, N 5. – P. 955-970.
159. How long is too long? Effects of loop size on G-quadruplex stability / [A. Guédin, J. Gros, P. Alberti, J.-L. Mergny] // *Nucleic Acids Res.* – 2010. – Vol. 38, N 21. – P. 7858–7868.
160. Zhang A. Y. A sequence-independent analysis of the loop length dependence of intramolecular RNA G-quadruplex stability and topology / A. Y. Zhang, A. Bugaut, S. Balasubramanian // *Biochemistry.* – 2011. – Vol. 50, N 33. – P. 7251–7258.
161. Mukundan V. T. Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences / V. T. Mukundan, A.T. Phan // *J. Am. Chem. Soc.* – 2013. – Vol. 135, N 13. – P. 5017–5028.
162. Pandey S. Effect of loops and G-quartets on the stability of RNA G-quadruplexes / S. Pandey, P. Agarwala, S. Maiti // *J. Phys. Chem. B.* – 2013. – Vol. 117, N 23. – P. 6896–6905.

163. An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES / [M. J. Morris, Y. Negishi, C. Pázsint et al.] // *J. Am. Chem. Soc.* – 2010. – Vol. 132, N 50. – P. 17831–17839.
164. Greatorex J. Retroviral RNA dimer linkage / J. Greatorex, A. Lever // *J. Gen. Virol.* – 1998. – Vol. 79, Pt 12. – P. 2877–2882.
165. Mechanisms and factors that influence high frequency retroviral recombination / [K. Delviks-Frankenberry, A. Galli, O. Nikolaitchik et al.] // *Viruses.* – 2011 – Vol. 3, N 9. – P. 1650–1680.
166. Sundquist W. I. Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA / W. I. Sundquist, S. Heaphy // *Proc. Natl Acad. Sci. USA.* – 1993. – Vol. 90, N 8. – P. 3393–3397.
167. Awang G. Mode of dimerization of HIV-1 genomic RNA / G. Awang, D. Sen // *Biochemistry.* – 1993. – Vol. 32, N 42. – P. 11453–11457.
168. Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization *in vitro* / [E. Skripkin, J.-C. Paillart, R. Marquet et al.] // *Proc. Natl Acad. Sci. USA* – 1994. – Vol. 91, N 11. – P. 4945–4949.
169. Laughrea M. A 19-nucleotide sequence upstream of the 5' major splice donor is part of the dimerization domain of human immunodeficiency virus 1 genomic RNA / M. Laughrea, L. Jetté // *Biochemistry.* – 1994. – Vol. 33, N 45. – P. 13464–13474.
170. Dimerization of human immunodeficiency virus type 1 RNA involves sequences located upstream of the splice donor site / [R. Marquet, J.-C. Paillart, E. Skripkin et al.] // *Nucleic Acids Res.* – 1994. – Vol. 22, N 2. – P. 145–151.
171. NCp7 activates HIV-1_{Lai} RNA dimerization by converting a transient loop-loop complex into a stable dimer / [D. Muriaux, H. De Rocquigny, B.-P. Roques, J. Paoletti] // *J. Biol. Chem.* – 1996. – Vol. 271, N 52. – P. 33686–33692.

172. Ennifar E. Polymorphism of bulged-out residues in HIV-1 RNA DIS kissing complex and structure comparison with solution studies / E. Ennifar, P. Dumas // *J. Mol. Biol.* – 2006. – Vol. 356, N 3. – P. 771-782.
173. The conformation of the mature dimeric human immunodeficiency virus type 1 RNA genome requires packaging of Pol protein / [M. Shehu-Xhilaga, M. Hill, J. A. Marshall et al.] // *J. Virol.* – 2002. – Vol. 76, N 9. – P. 4331-4340.
174. Berkhout B. Role of the DIS hairpin in replication of human immunodeficiency virus type 1 / B. Berkhout, J. L. B. van Wamel // *J. Virol.* – 1996. – Vol. 70, N 10. – P. 6723-6732.
175. Sakuragi J.-I. Human immunodeficiency virus type 1 RNA outside the primary encapsidation and dimer linkage region affects RNA dimer stability in vivo / J.-I. Sakuragi, A. T. Panganiban // *J. Virology* – 1997. – Vol. 71, N 4. – P. 3250-3254.
176. Deficient dimerization of human immunodeficiency virus type 1 RNA caused by mutations of the U5 RNA sequences / [R. S. Russell, J. Hu, M. Laughrea et al.] // *Virology*. – 2002. – Vol. 303, N 1. – P. 152-163.
177. Role of the trans-activation response element in dimerization of HIV-1 RNA / [E. S. Andersen, S. A. Contera, B. Knudsen et al.] // *J. Biol. Chem.* – 2004. – Vol. 279, N 21. – P. 22243-22249.
178. Song R. Role of the 5' TAR stem-loop and the U5-AUG duplex in dimerization of HIV-1 genomic RNA / R. Song, J. Kafaie, M. Laughrea // *Biochemistry*. – 2008. – Vol. 47, N 10. – P. 3283-3293.
179. Pallesen J. Structure of the HIV-1 5' untranslated region dimer alone and in complex with gold nanocolloids: support of a TAR-TAR-containing 5' dimer linkage site (DLS) and a 3' DIS-DIS-containing DLS / J. Pallesen // *Biochemistry*. – 2011. – Vol. 50, N 28. – P. 6170-6177.
180. Jalalirad M. Dominant role of the 5' TAR bulge in dimerization of HIV-1 genomic RNA, but no evidence of TAR-TAR kissing during in vivo virus

- assembly / M. Jalalirad, J. Saadatmand, M. Laughrea // *Biochemistry*. – 2012. – Vol. 51, N 18. – P. 3744–3758.
181. Purzycka K. J. The *in vitro* loose dimer structure and rearrangements of the HIV-2 leader RNA / K. J. Purzycka, K. Pachulska-Wieczorek, R. W. Adamiak // *Nucleic Acids Res.* – 2011. – Vol. 39, N 16. – P. 7234–7248.
182. Secondary structure of the mature *ex virio* Moloney murine leukemia virus genomic RNA dimerization domain / [C. Gherghe, C. W. Leonard, R. J. Gorelick, K. M. Weeks] // *J. Virol.* – 2010. – Vol. 84, N 2. – P. 898–906.
183. Kim C. H. A retroviral RNA kissing complex containing only two G.C base pairs / C. H. Kim, I. Tinoco Jr. // *Proc. Natl Acad. Sci. U S A.* – 2000. – Vol. 97, N 17. – P. 9396–9401.
184. Borochoy N. Dependence of DNA conformation on the concentration of salt / N. Borochoy, H. Eisenberg, Z. Kam // *Biopolymers.* – 1981. – Vol. 20, N 1. – P. 231–235.
185. Clewell D. B. Nature of Col E 1 plasmid replication in *Escherichia coli* in the presence of the chloramphenicol / D. B. Clewell // *J. Bacteriol.* – 1972. – Vol. 110, N 2. – P. 667–676.
186. Liu L. F. Type II DNA topoisomerases: enzymes that can unknot a topologically knotted DNA molecule via a reversible double-strand break / L. F. Liu, C. C. Liu, B. M. Alberts // *Cell.* – 1980. – Vol. 19, N 3. – P. 697–707.
187. A rapid procedure for purification of EcoRI endonuclease / [J. Sümegi, D. Breedveld, P. Hossenlopp, P. Chambon] // *Biochem. Biophys. Res. Commun.* – 1977. – Vol. 76, N 1. – P. 78–85.
188. Yarmola E. G. Osmotic pressure of DNA solutions and effective diameter of the double helix / E. G. Yarmola, M. I. Zarudnaya, Yu. S. Lazurkin // *J. Biomol. Struct. Dyn.* – 1985. – Vol. 2, N 5. – P. 981–993.

189. Post C. B. Excluded volume of an intermediate-molecular-weight DNA. A Monte Carlo analysis / C. B. Post // *Biopolymers*. – 1983. – Vol. 22, N 4. – P. 1087–1096.
190. Расчет характеристической вязкости синтетических и биологических полиэлектролитов различной жесткости / [С. В. Слоницкий, Э. В. Фрисман, А. К. Валеев, А. М. Ельяшевич] // *Молекул. биол.* – 1980. – Т. 14, № 3. – С. 484–495.
191. Schaefer D. W. Dynamics of semiflexible polymers in solution / D. W. Schaefer, J. F. Joanny, P. Pincus // *Macromolecules*. – 1980. – Vol. 13, N 5. – P. 1280–1289.
192. Бирштейн Т. М. Диаграмма состояния раствора полужестких макромолекул / Т. М. Бирштейн // *Высокомогл. соед., сер. А.* – 1982. – Т. 24, № 10. – С. 2110–2114.
193. Stigter D. Interactions of highly charged colloidal cylinders with applications to double-stranded DNA / D. Stigter // *Biopolymers*. – 1977. – Vol. 16, N 7. – P. 1435–1448.
194. Flory P. J. Principles of polymer chemistry / P. J. Flory. – NY: Cornell Univ. Press, 1953. – P. 495–540.
195. Onsager L. The effects of shape on the interaction of colloidal particles / L. Onsager // *Ann. N. Y. Acad. Sci.* – 1949. – Vol. 51, N 4. – P. 627–659.
196. Ярмола Е. Г. Зависимость осмотического давления растворов ДНК и эффективного диаметра от ионной силы / Е. Г. Ярмола, М. И. Зарудная, Ю. С. Лазуркин // *Биофизика*. – 1986. – Т. 31, № 2. – С. 338–340.
197. Остерман Л. А. Методы исследования белков и нуклеиновых кислот / Л. А. Остерман. – Москва: Наука, 1981. – 288 с.
198. Гааль Э. Электрофорез в разделении биологических макромолекул / Э. Гааль, Г. Медьеши, Л. Верецкеи. – Москва: Мир, 1982. – 448 с.

199. Rodbard D. Unified theory for gel electrophoresis and gel filtration / D. Rodbard, A. Chrambach // *Proc. Natl Acad. Sci. U S A.* – 1970. – Vol. 65, N 4. – P. 970–977.
200. RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical reexamination / [H. Lehrach, D. Diamond, J. M. Wozney, H. Boedtker] // *Biochemistry.* – 1977. – Vol. 16, N 21. – P. 4743–4751.
201. Tietz D. Computer simulation of the variable agarose fiber dimensions on the basis of mobility data derived from gel electrophoresis and using the Ogston theory / D. Tietz, A. Chrambach // *Anal Biochem.* – 1987. – Vol. 161, N 2. – P. 395–411.
202. Опанасенко В. К. Буферная емкость полипротонных объектов / В. К. Опанасенко, С. М. Герц, А. Д. Макаров // *Биохимия.* – 1978. – Vol. 43, № 8. – С. 1357–1368.
203. Изучение взаимодействия ДНК с протонами среды методом буферной емкости / [Б. И. Сухоруков, М. М. Монтрель, В. К. Опанасенко, Е. К. Золотарева] // *Молекул. биол.* – 1983. – Т. 17, № 5. – С. 1009–1018.
204. Опанасенко В. К. Определение и анализ зависимости буферной емкости хлоропластов от pH среды / В. К. Опанасенко // *Физиол. растений.* – 1980. – Т. 27, № 1. – С. 195–202.
205. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction / M. Zuker // *Nucleic Acids Res.* – 2003. – 31, N 13. – P. 3406–3415.
206. Markham N. R. UNAFold: software for nucleic acid folding and hybridization / N. R. Markham, M. Zuker // *Methods Mol. Biol.* – 2008. – Vol. 453, P. 3–31.
207. Phillips C. Identification of a stem-loop structure important for polyadenylation at the murine IgM secretory poly(A) site / C. Phillips, C. B. Kyriakopoulou, A. Virtanen // *Nucleic Acids Res.* – 1999. – Vol. 27, N 2. – P. 429–438.

208. McBride M. S. The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures / M. S. McBride, A. T. Panganiban // *J. Virol.* – 1996. – Vol. 70, N 5. – P. 2963–2973.
209. Extensive structural analysis of HIV-1 genomic RNA region encompassing DIS, SD and Psi hairpins / M. Zarudna, I. Kolomiets, A. Potyahaylo, D. Hovorun // 5th European Conference on Clinical and Social Research on AIDS and Drugs, Vilnius, Lithuania. – 2009. – P. 57.
210. Database of the secondary structures of HIV-1 genome control elements / A. Potyahaylo, I. Kolomiets, M. Zarudnaya, D. Hovorun // 34th FEBS Congress, Prague, Czech Republic. *FEBS J.* – 2009. – Vol. 276 (Suppl. 1). – P. 119.
211. Bon M. McGenus: a Monte Carlo algorithm to predict RNA secondary structures with pseudoknots / M. Bon, C. Micheletti, H. Orland // *Nucleic Acids Res.* – 2013. – Vol. 41, N 3. – P. 1895–1900.
212. Kikin O. QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences / O. Kikin, L. D'Antonio, P. S. Bagga // *Nucleic Acids Res.* – 2006. – Vol. 34 (Web Server issue):W676-682.
213. Sharp K. A. Salt effects on nucleic acids / K. A. Sharp, B. Honig // *Curr. Opin. Struct. Biol.* – 1995. – Vol. 5, N 3. – P. 323–328.
214. Brian A. A. Thermodynamics and equilibrium sedimentation analysis of the close approach of DNA molecules and a molecular ordering transition / A. A. Brian, H. L. Frisch, L. S. Lerman // *Biopolymers.* – 1981. – Vol. 20, N 6. – P. 1305–1328.
215. Rybenkov V. V. Probability of DNA knotting and the effective diameter of the DNA double helix / V. V. Rybenkov, N. R. Cozzarelli, A. V. Vologodskii // *Proc. Natl Acad. Sci. U S A.* – 1993. – Vol. 90, N 11. – P. 5307–5311.
216. Shaw S.Y. Knotting of a DNA chain during ring closure / S.Y. Shaw, J. C. Wang // *Science.* – 1993. – Vol. 260, N 5107. – P. 533–536.

217. Nicolai T. The ionic strength dependence of the second virial coefficient of low molar mass DNA fragments in aqueous solutions / T. Nicolai, M. Mandel // *Macromolecules*. – 1989. – Vol. 22, N 1. – P. 438–444.
218. Vologodskii A. Modeling of long-range electrostatic interactions in DNA / A. Vologodskii, N. Cozzarelli // *Biopolymers*. – 1995. – Vol. 35, N 3. – P. 289–296.
219. Manning G. S. The persistence length of DNA is reached from the persistence length of its null isomer through an internal electrostatic stretching force / G. S. Manning // *Biophys J*. – 2006. – Vol. 91, N 10. – P. 3607–3616.
220. Sobel E. S. Effects of Na^+ on the persistence length and excluded volume of T7 bacteriophage DNA / E. S. Sobel, J. A. Harpst // *Biopolymers*. – 1991. – Vol. 31, N 13. – P. 1559–1564.
221. Cairney K. L. Flow birefringence of T7 phage DNA: dependence on salt concentration / K. L. Cairney, R. E. Harrington // *Biopolymers*. – 1982. – Vol. 21, N 5. – P. 923–934.
222. Rizzo V. Flow dichroism of T7 DNA as a function of salt concentration / V. Rizzo, J. Schellman // *Biopolymers*. – 1981. – Vol. 20, N 10. – P. 2143–2163.
223. Зарудная М. И. Определение констант связывания гомополирибонуклеотидов с производным лизина методом электрофореза в агарозных гелях / М. И. Зарудная, Н. В. Желтовский // *Молекуляр. биология*. – 1989. – Т. 23, № 1. – С. 215–224.
224. Зарудная М. И. Исследование взаимодействия гомополирибонуклеотидов с дикатионом эфира лизина методом аффинного электрофореза / М. И. Зарудная, Н. В. Желтовский // *Молекуляр. биология*. – 1992. – Т. 26, № 1. – С. 110–117.
225. Зарудная М. И. Электрофоретическое исследование конформационных переходов в поли(А) при кислых рН / М. И. Зарудная, Н. В. Желтовский // *Молекуляр. биология*. – 1995. – Т. 29, № 5. – С. 1040–1047.

226. Электрофоретическое исследование конформационных переходов в поли(G) под действием моновалентных катионов / [М. И. Зарудная, А. В. Степанюгин, А. Л. Потягайло, Д. Н. Говорун] // *Biopolym. Cell.* – 2007. – Vol. 23, N 2. – P. 122–129.
227. Внутримолекулярные двухцепочечные спирали полиадениловой кислоты и их возможное биологическое значение / [М. И. Зарудная, А. В. Степанюгин, А. Л. Потягайло, Д. Н. Говорун] // *Ukr. Biochem. J.* – 2008. – Vol. 80, N 4. – P. 74–81.
228. Обнаружение методом электрофореза двух различных форм протонированного поли(C) / [М. И. Зарудная, А. В. Степанюгин, А. Л. Потягайло, Д. Н. Говорун] // *Biopolym. Cell.* – 2008. – Vol. 24, N 1. – P. 69–77.
229. Зарудная М. И. Исследование конформационных переходов в поли(A) методом буферной ёмкости / М. И. Зарудная // *Молекуляр. биология.* – 1998. – Т. 32, № 3. – С. 508–514.
230. Зарудна М. І. Структурні переходи в поліаденіловій кислоті. Можливі молекулярні механізми функціонування полі(A)-хвостів мРНК / М. І. Зарудна, Д. М. Говорун // *Доповіді НАН України.* – 1998. – №. 12. – С. 155–160.
231. Зарудная М. И. Конформационные переходы поли(C) и поли(dC): исследование методом протонной буферной емкости / М. И. Зарудная, А. Л. Потягайло, Д. Н. Говорун // *Biopolym. Cell.* – 2000. – Vol. 16, N 6. – P. 495–504.
232. Structural transitions in polycytidylic acid: proton buffer capacity data / [M. I. Zarudnaya, S. P. Samijlenko, A. L. Potyahaylo, D. M. Hovorun] // *Nucleosides, Nucleotides, Nucleic Acids.* – 2002. – Vol. 21, N 2. – P. 125–137.
233. Pinder J. C. Gel electrophoretic analysis of poly(riboadenylic acid) / J. C. Pinder, W. B. Gratzner // *Biochim. Biophys. Acta.* – 1974. – Vol. 349, N 1. – P. 47–52.

234. Pitha J. Affinity gel electrophoresis of polynucleotides / J. Pitha // *Anal. Biochem.* – 1975. – Vol. 65, N 1–2. – P. 422–426.
235. Брусков В. И. Специфичность взаимодействия аргинина и лизина с полинуклеотидами и их компонентами / В. И. Брусков, В. Н. Бушуев // *Биоорган. Химия.* – 1975. – Т. 1, № 11. – С. 1606–1615.
236. Porschke D. The binding of Arg- and Lys-peptides to single stranded polyribonucleotides and its effect on the polymer conformation / D. Porschke // *Biophys. Chem.* – 1979. – Vol. 10, N 1. – P. 1–16.
237. Кантор Ч. Биофизическая химия, Т. 3 / Ч. Кантор, П. Шиммель. – Москва: Мир, 1985. – 536 с.
238. Steiner R. F. Some properties of enzymatically produced polynucleotides / R. F. Steiner, R. F. Beers Jr. // *J. Polym. Sci.* – 1958. – Vol. 30, N 121– P. 17–28.
239. Visualization of poly(A)-binding protein complex formation with poly(A) RNA using atomic force microscopy / [B. L. Smith, D. R. Gallie, H. Le, P. K. Hansma] // *J. Struct. Biol.* – 1997. – Vol. 119, N 2. – P. 109–117.
240. Atomic force microscopy of long and short double-stranded, single-stranded and triple-stranded nucleic acids / [H. G. Hansma, I. Revenko, K. Kim, D. E. Laney] // *Nucleic Acids Res.* – 1996. – Vol. 24, N 4. – P. 713–720.
241. Kim S. Self-assembly of polydeoxyadenylic acid studied at the single-molecule level / S. Kim, J. Choi, T. Majima // *J. Phys. Chem. B.* – 2011. – Vol. 115, N 51. – P. 15399–15405.
242. Petrovic A. G. Structural transitions in polyriboadenylic acid induced by the changes in pH and temperature: vibrational circular dichroism study in solution and film states / A. G. Petrovic, P. L. Polavarapu // *J. Phys. Chem. B.* – 2005. – Vol. 109, N 49. – P. 23698–23705.
243. Zimmerman S. B. An ordered precipitate of polyadenylic acid formed by freezing at acidic pH: comparison of X-ray diffraction and other properties of the precipitate with those of fibers or direct acid-precipitates / S. B.

- Zimmerman, N. F. Coleman // *Biopolymers*. – 1972. – Vol. 11, N 9. – P. 1943-1960.
244. Griffin B. E. Synthesis and properties of some methylated polyadenylic acids / B. E. Griffin, W. J. Haslam, C. B. Reese // *J. Mol. Biol.* – 1964. – Vol. 10, N 2. – P. 353–356.
245. Ikeda K. Poly 2-amino-6-N-methyladenylic acid: synthesis, characterization and interaction with polyuridylic acid / K. Ikeda, J. Frazier, H. T. Miles // *J. Mol. Biol.* – 1970. – Vol. 54, N 1. – P. 59–84.
246. Brahms J. Circular dichroism investigations of the two conformations of polyriboadenylic acid / J. Brahms // *Nature*. – 1964. – Vol. 202, N 4934. – P. 797–798.
247. Alderfer J. L. A proton magnetic resonance study of polydeoxyriboadenylic acid / J. L. Alderfer, S. L. Smith // *J. Am. Chem. Soc.* – 1971. – Vol. 93, N 26. – P. 7305–7314.
248. Говорун Д. М. Фізико-хімічні механізми біомолекулярного впізнавання: автореф. дис. на здобуття наукового ступеня докт. біол. наук: спец. 03.00.03 «молекулярна біологія» / Д. М. Говорун. – Київ, – 1999. – 34 с.
249. The ...N6H...N6H... vertical hydrogen bonds contribute to the helical structure of single-stranded poly(A) / D. M. Hovorun, A. L. Potyahaylo, M. I. Zarudna // XIV International School-Seminar “Spectroscopy of molecules and crystals”, Odessa, Ukraine. – 1999. – P. 146.
250. Зарудная М. И. Возможный механизм стабилизации одной из двух форм двуспирального поли(А) / М. И. Зарудная // *Biopolym. Cell*. – 2008. – Vol. 24, N 4. – P. 333–339
251. Билобров В. М. Водородная связь. Внутримолекулярные взаимодействия / В. М. Билобров. – Киев: Наукова думка, 1991. – 320 с.
252. Билобров В. М. Водородная связь. Межмолекулярные взаимодействия / В. М. Билобров. – Киев: Наукова думка, 1993. – 520 с.

253. Small E. W. Conformational dependence of the raman scattering intensities from polynucleotides / E. W. Small, W. L. Peticolas // *Biopolymers*. – 1971. – Vol. 10, N 1. – P. 69–88.
254. Leng M. A study of polyadenylic acid at neutral pH / M. Leng, G. Felsenfeld // *J. Mol. Biol.* – 1966. – Vol. 15, N 2. – P. 455–466.
255. Raman studies of nucleic acids. XII. Conformations of oligonucleotides and deuterated polynucleotides / [B. Prescott, R. Gamache, J. Livramento, G. J. Thomas Jr.] // *Biopolymers*. – 1974. – Vol. 13, N 9. – P. 1821–1845.
256. Poly-2'-deoxy-2'-fluoro-cytidylic acid: enzymatic synthesis, spectroscopic characterization and interaction with poly-inosinic acid / [W. Guschlbauer, M. Blandin, J. L. Drocourt, M. N. Thang] // *Nucl. Acids Res.* – 1977. – Vol. 4, N 6. – P. 1933–1943.
257. Thiele D. Polynucléotides protonés. VII. Transitions thermiques entre différents complexes de l'acide polyinosinique et de l'acide polycytidylique en milieu acide / D. Thiele, W. Guschlbauer // *Biopolymers*. – 1969. – Vol. 8, N 3. – P. 361–378.
258. Aronsson G. Electrostatic potential modulation on polynucleotides as a function of ionic content: potentiometric determinations / G. Aronsson, F. Travers // *Nucl. Acids Res.* – 1976. – Vol. 3, N 5. – P. 1373–1385.
259. Конформационное состояние и особенности молекулярно-ионного превращения полицитидиловой кислоты, иммобилизованной в мультислойных лэнгмюровской и полиэлектrolитной пленках / [Б. И. Сухоруков., Г. Б. Сухоруков., Л. И. Шабарчина, М. М. Монтрель] // *Биофизика*. – 2000. – Т. 45, № 1. – С. 40–50.
260. Sponer J. Hydrogen bonding and stacking of DNA bases: a review of quantum-chemical ab initio studies / J. Sponer, J. Leszczynski, P. Hobza // *J. Biomol. Struct. Dyn.* – 1996. – Vol. 14, N 1. – P. 117–135.
261. Міщук Я. Р. Вивчення фізико-хімічної природи елементарних актів білково-нуклеїнового та нуклеїново-нуклеїнового впізнавання на низькомолекулярних модельних системах: автореф. дис. на здобуття

- наук. ступеня канд. біол. наук: спец. 03.00.03 «молекулярна біологія» / Я. Р. Мішук. – Київ, – 1993. – 21 с.
262. Говорун Д. М. Газофазні кислотно-лужні властивості канонічних нуклеотидних основ / Д. М. Говорун, І. В. Кондратюк // Доповіді НАН України. – 1998. – № 1. – С. 207–212.
263. Florian J. IR and Raman spectra, tautomeric stabilities, and scaled quantum mechanical force fields of protonated cytosine / J. Florián, V. Baumruk, J. Leszczyński // J. Phys. Chem. – 1996. – Vol. 100, N 13. – P. 5578-5589.
264. Войтюк А. А. Влияние протонирования оснований нуклеиновых кислот на энергию образования Уотсон-Криковских пар / А. А. Войтюк, А. А. Близнюк // Молекул. биол. – 1988. – Т. 22, № 4. – С. 1080–1086.
265. Потягайло А. Л. Молекулярні механізми структурних переходів полі(С) і полі(дС), спричинених протонуванням, та їхнє квантово-хімічне обґрунтування / А. Л. Потягайло, М. І. Зарудна, Д. М. Говорун // Доповіді НАН України. – 2003. – № 2. – С. 196–199.
266. Gorb L. Intramolecular proton transfer in mono- and dihydrated tautomers of guanine: an ab initio post Hartree–Fock study / L. Gorb, J. Leszczyński // J. Am. Chem. Soc. – 1998. – Vol. 120, N 20. – P. 5024–5032.
267. DNA base modification: Ionized base pairs and mutagenesis / [L. C. Sowers, B. R. Shaw, M. L. Veigl, W. D. Sedwick] // Mutat. Res. – 1987. – Vol. 177, N 2. – P. 201–218.
268. Wróbel A. Conformation of acid forms of poly(C): temperature and ionic strength dependence of protonation of cytidine and cytidine-5'-phosphate / A. Wróbel, A. Rabczenko, D. Shugar // Acta biochim. pol. – 1970. – Vol. 17, N 4. – P. 339–349.
269. Green G. Comparative study of polyribonucleotides in aqueous and glycol solutions / G. Green, H. R. Mahler // Biochemistry. – 1970. – Vol. 9, N 2. – P. 368–387.

270. O'Connor T. pH-dependent Raman spectra and thermal melting profiles for polycytidylic acid / T. O'Connor, W. M. Scovell // *Biopolymers*. – 1981. – Vol. 20, N 11. – P. 2351–2367.
271. Cohen B. Ultrafast excited-state dynamics of RNA and DNA C tracts / B. Cohen, M. H. Larson, B. Kohler // *Chem. Phys.* – 2008. – Vol. 350, N 1–3. – P. 165–174.
272. Petrovic A. G. Structural transitions in polyribocytidylic acid induced by changes in pH and temperature: vibrational circular dichroism study in solution and film states / A. G. Petrovic, P. L. Polavarapu // *J. Phys. Chem. B*. – 2006. – Vol. 110, N 45. – P. 22826–22833.
273. Cooperative disordering of single-stranded polynucleotides through copper crosslinking / [J. M. Rifkind, Y. A. Shin, J. M Heim., G. L. Eichhorn] // *Biopolymers*. – 1976. – Vol. 15, N 10. – P. 1879–1902.
274. Stretching of homopolymeric RNA reveals single-stranded helices and base-stacking / [Y. Seol, G. M. Skinner., K. Visscher et al.] // *Phys. Rev. Lett.* – 2007. – Vol. 98, N 15: 158103.
275. Изучение комплексов цитозина с карбоксильной группой аминокислот методом спектроскопии ЯМР / [И. В. Кондратюк, И. Н. Коломиец, С. А. Самойленко, Н. В. Желтовский] // *Biopolym. Cell*. – 1989. – Vol. 5, N 6. – P. 21–25.
276. The I-tetraplex building block: rational design and controlled fabrication of robust 1D DNA scaffolds through non-Watson-Crick interactions / [H. B. Ghodke, R. Krishnan, K. Vignesh et al.] // *Angew Chem. Int. Ed.* – 2007. – Vol. 46, N 15. – P. 2646–2649.
277. Laisné A. [C7GC4]₄ association into supra molecular i-motif structures / A. Laisné, D. Pompon, J.-L. Leroy // *Nucleic Acids Res.* – 2010. – Vol. 38, N 11. – P. 3817–3826.
278. Leroy J.-L. The formation pathway of i-motif tetramers / J.-L. Leroy // *Nucleic. Acids Res.* – 2009. – Vol. 37, N 12. – P. 4127–4134.

279. Coexistence of an ILPR i-motif and a partially folded structure with comparable mechanical stability revealed at the single-molecule level / [S. Dhakal, J. D. Schonhofs, D. Koirala et al.] // *J. Am. Chem. Soc.* – 2010. – Vol. 132, N 26. – P. 8991–8997.
280. Kinetics of tetramolecular quadruplexes / [J.-L. Mergny, A. De Cian, A. Ghelab et al.] // *Nucl. Acids Res.* – 2005. – Vol. 33, N 1. – P. 81–94.
281. Зарудная М. И. Самоассоциаты гомополимерных трактов клеточных РНК: физические механизмы образования и функционирования / М. И. Зарудная, Д. Н. Говорун // *Фізика живого.* – 1999. – Т. 7, № 2. – С. 38–52.
282. Antony T. Spectroscopic studies on the binding of methylene blue to poly(riboadenylic acid) / T. Antony, M. Atreyi, M. V. R. Rao // *J. Biomol. Struct. Dyn.* – 1993. – Vol. 11, N 1. – P. 067–081.
283. Imae T. Interaction between acridine orange and polyriboadenylic acid / [T. Imae, S. Hayashi, S. Ikeda, T Sakaki] // *Int. J. Biol. Macromol.* – 1981. – Vol. 3, N 4. – P. 259–266.
284. Fujita S. Crystal structure of adenine-1-(2-carboxyethyl)uracil (1:1) complex. A model for interactions of amino acid side chains with nucleic acid base pair / S. Fujita, A. Takenaka, Y. Sasada // *Bull. Chem. Soc. Jpn.* – 1983. – Vol. 56, N 8. – P. 2234–2237.
285. Carboxyl-modified single-walled carbon nanotubes selectively induce human telomeric i-motif formation. / [X. Li, Y. Peng, J. Ren, X. Qu] // *Proc. Natl Acad. Sci. U S A.* – 2006. – Vol. 103, N 52. – P. 19658–19663.
286. Studies of virus structure by laser-Raman spectroscopy. Turnip yellow mosaic virus and capsids / [K. A. Hartman, P. E. McDonald-Ordzie, J. M. Kaper et al.] // *Biochemistry.* – 1978. – Vol. 17, N 11. – P. 2118–2123.
287. Protein-RNA interactions in belladonna mottle virus investigated by laser Raman spectroscopy / [B. Prescott, K. Sitaraman, P. Argos, G. J. Thomas Jr.] // *Biochemistry.* – 1985. – Vol. 24, N 5. – P. 1226–1231.

288. Zarudnaya M. I. Double helical forms of poly(A): possible involvement in biological processes / M. I. Zarudnaya // *Biopolym. Cell.* – 1999. – Vol. 15, N 3. – P. 195–199.
289. Zarudnaya M. I. Structural transitions in polyadenylic acid and hypothesis on biological role of its double-stranded forms / M. I. Zarudnaya, D. M. Hovorun // *Ukr. Biochem. J.* – 1999. – Vol. 71, N 4. – P. 15–20.
290. Zarudnaya M. I. Hypothetical double-helical poly(A) formation in a cell and its possible biological significance / M. I. Zarudnaya, D. M. Hovorun // *IUBMB Life.* – 1999. – Vol. 48, N 6. – P. 581–584.
291. Zarudnaya M. I. The double-helical forms of poly(A) can be formed in a cell and participate in biological processes / M. I. Zarudnaya, D. M. Hovorun // *Comments Mol. Cell. Biophys.* – 2000. – Vol. 10, N1. – P. 53–62.
292. Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR / [S. Lebedeva, M. Jens, K. Theil et al.] // *Mol. Cell.* – 2011. – Vol. 43, N 3. – P. 340–352.
293. Fan X. C. Overexpression of HuR, a nuclear-cytoplasmic shuttling protein, increases the *in vivo* stability of ARE-containing mRNAs / X. C. Fan, J.A. Steitz // *EMBO J.* – 1998. – Vol. 17, N 12. – P. 3448–3460.
294. Poly(A) tail length-dependent stabilization of GAP-43 mRNA by the RNA-binding protein HuD / [A. C. Beckel-Mitchener, A. Miera, R. Keller, N. I. Perrone-Bizzozero] // *J. Biol. Chem.* – 2002. – Vol. 277, N 31. – P. 27996–28002.
295. Stability of casein mRNA is ensured by structural interactions between the 3'-untranslated region and poly(A) tail via the HuR and poly(A)-binding protein complex / [K. Nagaoka, T. Suzuki, T. Kawano et al.] // *Biochim. Biophys. Acta.* – 2006. – Vol. 1759, N 3–4. – P. 132–140.
296. Waggoner S. A. Regulation of α -globin mRNA stability / S. A. Waggoner, S. A. Liebhaber // *Exp. Biol. Med. (Maywood).* – 2003. – Vol. 228, N 4. – P. 387–395.

297. Poly(A)-binding protein binds to A-rich sequences via RNA-binding domains 1+2 and 3+4 / [T. Khanam, R. S. Muddashetty, A. Kahvejian et al] // *RNA Biol.* – 2006. – Vol. 3, N 4. – P. 170–177.
298. de Melo Neto O. P. Autoregulation of poly(A)-binding protein synthesis in vitro / O. P. de Melo Neto, N. Standart, C. M. de Sa // *Nucleic Acids Res.* – 1995. – Vol. 23, N 12. – P. 2198–2205.
299. Bag J. Feedback inhibition of poly(A)-binding protein mRNA translation. A possible mechanism of translation arrest by stalled 40 S ribosomal subunits / J. Bag // *J. Biol. Chem.* – 2001. – Vol. 276, N 50. – P. 47352–47360.
300. Alekhina O. M. Translation initiation in eukaryotes: versatility of the scanning model / O. M. Alekhina, K. S. Vassilenko // *Biochemistry (Mosc).* – 2012. – Vol. 77, N 13. – P. 1465–1477.
301. Crystallization of oligonucleotides containing A-rich repeats suggests a structural contribution to the autoregulation mechanism of PABP translation / [K. Kikuchi, S. Shimizu, Y. Sato et al.] // *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* – 2012. – Vol. 68, Pt 2. – P. 185–189.
302. Isolation of genomic and cDNA clones encoding bovine poly(A) binding protein II / [A. Nemeth, S. Krause, D. Blank et al.] // *Nucleic Acids Res.* – 1995. – Vol. 23, N 20. – P. 4034–4041.
303. Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor / [U. Kühn, M. Gündel, A. Knoth et al.] // *J. Biol. Chem.* – 2009. – Vol. 284, N 34. – P. 22803–22814.
304. The nuclear poly(A) binding protein, PABP2, forms an oligomeric particle covering the length of the poly(A) tail / [R. W. Keller, U. Kühn, M. Aragón et al.] // *J. Mol. Biol.* – 2000. – Vol. 297, N 3. – P. 569–583.
305. Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA / [Y. Kerwitz, U. Kühn, H. Lilie et al.] // *EMBO J.* – 2003. – Vol. 22, N 14. – P. 3705–3714.

306. Nucleophosmin deposition during mRNA 3' end processing influences poly(A) tail length / [F. Sagawa, H. Ibrahim, A. L. Morrison et al.] // *EMBO J.* – 2011. – Vol. 30, N 19. – P. 3994–4005.
307. Hosoda N. Evidence that poly(A) binding protein C1 binds nuclear pre-mRNA poly(A) tails / N. Hosoda, F. Lejeune, L. E. Maquat // *Mol. Cell. Biol.* – 2006. – Vol. 26, N 8. – P. 3085–3097.
308. Soucek S. The long and the short of it: the role of the zinc finger polyadenosine RNA binding protein, Nab2, in control of poly(A) tail length / S. Soucek, A. H. Corbett, M. B. Fasken // *Biochim. Biophys. Acta.* – 2012. – Vol. 1819, N 6. – P. 546–554.
309. Rrp6p controls mRNA poly(A) tail length and its decoration with poly(A) binding proteins / [M. Schmid, M. B. Poulsen, P. Olszewski et al.] // *Mol. Cell.* – 2012. – Vol. 47, N 2. – P. 267–280.
310. Viphakone N. Molecular dissection of mRNA poly(A) tail length control in yeast / N. Viphakone, F. Voisinet-Hakil, L. Minvielle-Sebastia // *Nucleic Acids Res.* – 2008. – Vol. 36, N 7. – P. 2418–2433.
311. Зарудная М. И. Структурная модель области димеризации генома вируса иммунодефицита человека HIV-1 / М. И. Зарудная, А. Л. Потягайло, Д. Н. Говорун // *Biopolym. Cell.* – 2003. – Vol. 19, N 1. – P. 37–42.
312. Dimer linkage structure in retroviruses: models that include both duplex and quadruplex domains / [M. I. Zarudnaya., I. M. Kolomiets, A. L. Potyahaylo, D. M. Hovorun] // *Ukr. Biochem. J.* – 2005. – Vol. 77, N 2. – P. 5–15.
313. Lau K. A. Current trends of HIV recombination worldwide / K. A. Lau, J. J. Wong // *Infect. Dis. Rep.* – 2013. – Vol. 5 (Suppl 1): S:e4.
314. The primary and secondary structures of HIV-1 genomic RNA region encompassing DIS, SD and Psi hairpins: *in silico* study / [M. I. Zarudnaya, I. M. Kolomiets, A. L. Potyahaylo, D. M. Hovorun] // In: "Trends in RNA Research"; edited by P.A. McNamara. – New York: Nova Science Publishers, 2006. – P. 159–189.

315. Diversity of DIS, SD and Ψ hairpins in HIV-1 isolates of group M: *in silico* study / [M. I. Zarudnaya, A. L. Potyahaylo, I. N. Kolomiets, D. M. Hovorun] // Ukr. Biochem. J. – 2007. – Vol. 79, N 1. – P. 68–84.
316. Variant effects of non-native kissing-loop hairpin palindromes on HIV replication and HIV RNA dimerization: role of stem-loop B in HIV replication and HIV RNA dimerization / [M. Laughrea, N. Shen, L. Jetté, M. A. Wainberg] // Biochemistry. – 1999. – Vol. 38, N 1. – P. 226–234.
317. Stem of SL1 RNA in HIV-1: structure and nucleocapsid protein binding for a 1 x 3 internal loop / [Y. Yuan, D. J. Kerwood, A. C. Paoletti et al.] // Biochemistry. – 2003. – Vol. 42, N 18. – P. 5259–5269.
318. Chowdhury S. Modelling studies on neurodegenerative disease-causing triplet repeat sequences d(GGC/GCC)_n and d(CAG/CTG)_n / S. Chowdhury, M. Bansal // J. Biosci. – 2001. – Vol. 26, N 5. – P. 649–665.
319. Gu J. Structures and properties of mixed DNA bases tetrads: nonempirical ab Initio HF and DFT studies / J. Gu, J. Leszczynski // J. Phys. Chem. – 2000. – Vol. 104, N 9. – P. 1898–1904.
320. Clever J. L. Requirements for kissing-loop-mediated dimerization of human immunodeficiency virus RNA / J. L. Clever, M. L. Wong, T. G. Parslow // J. Virol. – 1996. – Vol. 70, N 9. – P. 5902–5908.
321. *Cis* elements and *trans*-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1 / [J.-L. Darlix, C. Gabus, M.-T. Nugeyre et al.] // J. Mol. Biol. – 1990. – Vol. 216, N 3. – P. 689–699.
322. Piekna-Przybylska D. Mechanism of HIV-1 RNA dimerization in the central region of the genome and significance for viral evolution / D. Piekna-Przybylska, G. Sharma, R. A. Bambara // J. Biol. Chem. – 2013. – Vol. 288, N 33. – P. 24140–24150.
323. U3 region in the HIV-1 genome adopts a G-quadruplex structure in its RNA and DNA sequence / [D. Piekna-Przybylska, M. A. Sullivan, G. Sharma, R. A. Bambara] // Biochemistry. – 2014. – Vol. 53, N 16. – P. 2581–2593.

324. Rist M. J. Mechanism of nucleocapsid protein catalyzed structural isomerization of the dimerization initiation site of HIV-1 / M. J. Rist, J. P. Marino // *Biochemistry*. – 2002. – Vol. 41, N 50. – P. 14762–14770.
325. Moore M. D. HIV-1 RNA dimerization: It takes two to tango / M. D. Moore, W.-S. Hu // *AIDS Rev.* – 2009. – Vol. 11, N 2. – P. 91-102.
326. Sequences downstream of the 5' splice donor site are required for both packaging and dimerization of human immunodeficiency virus type 1 RNA / [R. S. Russell, J. Hu, V. Bériault et al.] // *J. Virol.* – 2003. – Vol. 77, N 1. – P. 84–96.
327. Functional characterization of the dimer linkage structure RNA of Moloney murine sarcoma virus / [H. Ly, D. P. Nierlich, J. C. Olsen, A. H. Kaplan] // *J. Virol.* – 2000. – Vol. 74, N 21. – P. 9937–9945.
328. Phylogenetic study on structural elements of HIV-1 poly(A) region. 2. USE domain and TAR hairpin / [M. I. Zarudnaya, A. L. Potyahaylo, I. M. Kolomiets, D. M. Hovorun] // *Biopolym. Cell.* – 2014. – Vol. 30, N 1. – P. 29–36.
329. Sequence and structure requirements for specific recognition of HIV-1 TAR and DIS RNA by the HIV-1 Vif protein / [S. Freisz, J. Mezher, L. Hafirassou et al.] // *RNA Biol.* – 2012. – Vol. 9, N 7. – P. 966-977.
330. Hans H. Functionally significant secondary structure of the simian virus 40 late polyadenylation signal / H. Hans, J. C. Alwine // *Mol. Cell. Biol.* – 2000. – Vol. 20, N 8. – P. 2926–2932.
331. Lowman H. B. On the recognition of helical RNA by cobra venom V₁ nuclease / H. B. Lowman, D. E. Draper // *J. Biol. Chem.* – 1986. – Vol. 261, N 12. – P. 5396–5403.
332. Matunis M. J. The hnRNP F protein: unique primary structure, nucleic acid-binding properties, and subcellular localization / M. J. Matunis, J. Xing, G. Dreyfuss // *Nucleic Acids Res.* – 1994. – Vol. 22, N 6. – P. 1059–1067.
333. Вторичная структура основного поли(А) сигнала про-мРНК вируса иммунодефицита человека / [М. И. Зарудная, А. Л. Потягайло, В. В.

- Отенко, И. Н. Коломиец, Д. Н. Говорун] // Доповіді НАН України – 2011. – № 4. – С. 170–176.
334. Gilmartin G. M. Activation of HIV-1 pre-mRNA 3' processing *in vitro* requires both an upstream element and TAR / G. M. Gilmartin, E. S. Fleming, J. Oetjen // EMBO J. – 1992. – Vol. 11, N 12. – P. 4419–4428.
335. Destabilization of the TAR hairpin leads to extension of the polyA hairpin and inhibition of HIV-1 polyadenylation / [M. M. Vrolijk, A. Harwig, B. Berkhout, A. T. Das] // Retrovirology. – 2009. – Vol. 6:13.
336. Phylogenetic study on structural elements of HIV-1 poly(A) region. 1. PolyA and DSE hairpins / [M. I. Zarudnaya, A. L. Potyahaylo, I. M. Kolomiets, D. M. Hovorun] // Biopolym. Cell. – 2013. – Vol. 29, N 6. – P. 454–462.
337. A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter / [L. A. Pereira, K. Bentley, A. Peeters et al.] // Nucleic Acids Res. – 2000. – Vol. 28, N 3. – P. 663–668.
338. New findings on transcription regulation across different HIV-1 subtypes / [E. Ramírez de Arellano, V. Soriano, J. Alcamil, A. Holguín] // AIDS Rev. – 2006. – Vol. 8, N 1. – P. 9–16.
339. The G-rich auxiliary downstream element has distinct sequence and position requirements and mediates efficient 3' end pre-mRNA processing through a *trans*-acting factor / [P. S. Bagga, L. P. Ford, F. Chen, J. Wilusz] // Nucleic Acids Res. – 1995. – Vol. 23, N 9. – P. 1625–1631.
340. Isel C. Initiation of HIV Reverse Transcription / C. Isel, C. Ehresmann, R. Marquet // Viruses. – 2010. – Vol. 2, N 1 – P. 213–243.
341. Anti-HIV-1 activity of the G-quadruplex ligand BRACO-19 / [R. Perrone, E. Butovskaya, D. Daelemans, et al.] // J. Antimicrob. Chemother. – 2014. – Vol. 69, N 12. – P. 3248–3258.
342. A dynamic G-quadruplex region regulates the HIV-1 long terminal repeat promoter / [R. Perrone, M. Nadai, I. Frasson et al.] // J. Med. Chem. – 2013. – Vol. 56, N 16. – P. 6521–6530.

343. The consensus sequence YGTGTTY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini / [J. McLauchlan, D. Gaffney, J. L. Whitton, J. B. Clements] // *Nucleic Acids Res.* – 1985. – Vol. 13, N 4. – P. 1347–1368.
344. A large-scale analysis of mRNA polyadenylation of human and mouse genes / [B. Tian, J. Hu, H. Zhang, C. S. Lutz] // *Nucleic Acids Res.* – 2005. – Vol. 33, N 1. – P. 201–212.
345. Molecular architecture of the human pre-mRNA 3' processing complex / [Y. Shi, D. C. Di Giammartino, D. Taylor et al.] // *Mol. Cell.* – 2009. – Vol. 33, N 3. – P. 365–376.
346. Shi Y. An up-close look at the pre-mRNA 3'-end processing complex / Y. Shi, S. Chan, G. Martinez-Santibañez // *RNA Biol.* – 2009. – Vol. 6, N 5. – P. 522–525.
347. Venkataraman K. Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition / K. Venkataraman, K. M. Brown, G. M. Gilmartin // *Genes Dev.* – 2005. – Vol. 19, N 11. – P. 1315–1327.
348. Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping / [Q. Yang, M. Coseno, G. M. Gilmartin, S. Doublé] // *Structure.* – 2011. – Vol. 19, N 3. – P. 368–377.
349. Mammalian pre-mRNA polyadenylation signals / [M. Zarudnaya, A. Potyahaylo, I. Kolomiets, D. Novorun]: Тези доповідей Установчого з'їзду Українського товариства клітинної біології (25-28 квітня 2004). – Львів, 2004. – С. 371
350. Upstream elements present in the 3'-untranslated region of collagen genes influence the processing efficiency of overlapping polyadenylation signals / [B. J. Natalizio, L. C. Muniz, G. K. Arhin et al.] // *J. Biol. Chem.* – 2002. – Vol. 277, N 45. – P. 42733–42740.

351. Romanelli M. G. New insights into functional roles of the polypyrimidine tract-binding protein / M. G. Romanelli, E. Diani., P. M. Lievens // *Int. J. Mol. Sci.* – 2013. – Vol. 14, N 11. – P. 22906–22932.
352. Huppert J. L. Prevalence of quadruplexes in the human genome / J. L. Huppert, S. Balasubramanian // *Nucleic Acids Res.* – 2005. – Vol. 33, N 9. – P. 2908–2916.
353. Todd A. K. Highly prevalent putative quadruplex sequence motifs in human DNA / A. K. Todd, M. Johnston, S. Neidle // *Nucleic Acids Res.* – 2005. – Vol. 33, N 9. – P. 2901–2907.
354. The effect of molecular crowding on the stability of human c-MYC promoter sequence I-motif at neutral pH / [J. Cui, P. Waltman, V. H. Le, E. A. Lewis] // *Molecules.* – 2013. – Vol. 18, N 10. – P. 12751-12767.
355. Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing / [R. Martinez-Contreras, J.-F. Fisette, F. U. Nasim et al.] // *PLoS Biol.* – 2006. – Vol. 4, N 2:e21.