

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
имени М.В.ЛОМОНОСОВА ХИМИЧЕСКИЙ ФАКУЛЬТЕТ

*На правах рукописи*

**Савицкая Виктория Юрьевна**

**Особенности взаимодействия белков систем эксцизионной  
репарации ДНК с G-богатыми фрагментами регуляторных  
областей генома эукариот и прокариот**

1.4.9. Биоорганическая химия

**ДИССЕРТАЦИЯ**

на соискание ученой степени кандидата химических наук

Научный руководитель:

доктор химических наук, профессор

Кубарева Елена Александровна

**Москва 2026**

## Оглавление

Список использованных сокращений .....	5
I. Введение .....	7
I.1. Актуальность .....	7
I.2. Степень проработанности проблемы .....	7
I.3. Цели и задачи работы .....	9
I.4. Объект исследования .....	10
I.5. Предмет исследования .....	10
I.6. Научная новизна .....	10
I.7. Теоретическая значимость работы .....	11
I.8. Практическая значимость работы .....	11
I.9. Методология диссертационного исследования .....	12
I.10. Личный вклад автора .....	12
I.11. Положения, выносимые на защиту .....	13
I.12. Апробация работы .....	13
I.13. Публикации .....	13
I.14. Структура и объем .....	14
II. Эксцизионная репарация ДНК прокариот и эукариот: сравнение механизмов, роль в поддержании целостности генома и реализации генетического разнообразия .....	15
II.1. Эксцизионная репарация ДНК в клетках различных организмов .....	15
II.1.1. Основные типы и общий механизм эксцизионной репарации ДНК .....	15
II.1.2. Эксцизионная репарация оснований и нуклеотидов .....	15
II.1.3. Репарация неканонических пар нуклеотидов или «мисматчей» .....	22
II.2. Роль систем репарации ДНК в патогенезе бактерий .....	26
II.2.1. Репарационные пути в клетках бактерии <i>Neisseria gonorrhoeae</i> .....	27
II.2.2. Ключевые факторы вирулентности <i>Neisseria gonorrhoeae</i> .....	28
II.2.3. Роль белков систем репарации ДНК в адаптации и изменчивости <i>Neisseria gonorrhoeae</i> .....	31

II.2.4. Неканонические структуры ДНК и их влияние на патогенез и репарацию .....	33
II.3. Роль систем репарации ДНК в клетках человека .....	36
II.3.1. Организация клеточного цикла человека и функционирование систем репарации ДНК .....	36
II.3.2. Клиническое значение мутаций в генах белков систем эксцизионной репарации ДНК человека .....	38
II.3.3. Роль белков систем эксцизионной репарации ДНК человека в поддержании генетического разнообразия .....	41
II.3.4. Влияние GC-состава ДНК на эффективность репарации ДНК .....	44
II.3.5. Взаимодействие белков эксцизионной репарации ДНК человека с G-квадруплексами .....	45
III. Материалы и методы .....	50
IV. Особенности взаимодействия белков систем репарации ДНК с G-богатыми фрагментами регуляторных областей генома эукариот и прокариот.....	64
IV. 1. Новые аспекты функционирования системы MMR по метилнезависимому пути .....	65
IV. 1.1. Выделение белков ngMutL <sup>wt</sup> и ngMutL <sup>cf</sup> .....	66
IV.1.2. Сравнение вторичной структуры ngMutL <sup>wt</sup> и ngMutL <sup>cf</sup> .....	67
IV.1.3. Характеристика и сравнение АТФазной активности белков ngMutL <sup>wt</sup> и ngMutL <sup>cf</sup> ...	68
IV. 1.4. Характеристика и сравнение эндонуклеазной активности белков ngMutL <sup>wt</sup> и ngMutL <sup>cf</sup> .....	69
IV.1.5. Эффективность гидролиза линейного субстрата белком ngMutL в присутствии β-субъединицы ДНК-полимеразы III .....	71
IV.1.6. Анализ ДНК-связывающей активности ngMutL .....	75
IV.1.7. Анализ взаимодействия ngMutL с β-субъединицей ДНК-полимеразы III.....	76
IV. 2. Биоинформатический анализ влияния полиморфизмов белков MMR на антигенную вариацию пилина .....	78
IV.2.1. Оценка полиморфизма генов <i>ngmutS</i> и <i>ngmutL</i> в популяции <i>Neisseria gonorrhoeae</i> .	78
IV.2.2. Связь между полиморфизмом белков MMR и антигенной вариацией пилина .....	79
IV.3. Взаимодействие белка ngMutL с G4-мотивом регуляторной последовательности перед <i>pilE</i> .....	80

IV.3.1. Характеристика модельных субстратов, содержащих G4-мотив перед <i>pile</i> .....	80
IV.3.2. Связывание ngMutL с одно- и двуцепочечными ДНК, содержащими G4-мотив регуляторной последовательности перед <i>pile</i> .....	84
IV.3.3. Гидролиз белком ngMutL ДНК, содержащих G4-мотив регуляторной последовательности <i>pile</i> .....	85
IV.3.4. Влияние АТФ на связывание G4-содержащей ДНК белком ngMutL.....	87
IV.4. Эффективность узнавания неканонической пары внутри GC-богатой последовательности при репарации «мисматчей» у прокариот .....	88
IV.5. Влияние GC-богатой последовательности на эффективность эксцизионной репарации эукариот.....	95
IV.5.1. Консервативные последовательности PQS в промоторах <i>TERT</i> млекопитающих ....	96
IV.5.2. Характеристики G-квадруплексной структуры, сформированной G-богатой последовательностью промотора <i>hTERT</i> .....	98
IV.5.3. Дизайн модельных фрагментов промотора <i>hTERT</i> .....	101
IV.5.4. Влияние F-сайтов на топологию и термическую стабильность G-квадруплексной структуры, формируемой фрагментом промотора <i>hTERT</i> .....	102
IV.5.5. Влияние G4-мотивов промотора <i>hTERT</i> на функционирование APE1 .....	104
IV.5.6. Влияние GC-богатого состава промотора <i>hTERT</i> на способность белка APE1 узнавать повреждение .....	108
V. Заключение .....	113
VI. Выводы.....	115
Список литературы .....	116

## Список использованных сокращений

а.о.	аминокислотный остаток
АДФ	аденозин-5'-дифосфат
АТФ	аденозин-5'-трифосфат
БСА	бычий сывороточный альбумин
БФС	бромфеноловый синий
ДНК	дезоксирибонуклеиновая кислота
ДСН	додецилсульфат натрия
ДТТ	1,4-дитиотреит, <i>трео</i> -2,3-дигидрокси-1,4-димеркаптобутан
ИПТГ	изопропил- $\beta$ -D-тиогалактопиранозид
КД	круговой дихроизм
ПААГ	полиакриламидный гель
Трис	<i>трис</i> -(гидроксиметил)аминометан
ЭДТА	этилендиаминтетраацетат, динатриевая соль этилендиаминтетрауксусной кислоты
$\beta$ -«зажим»	$\beta$ -субъединица ДНК-полимеразы III
т.п.н.	тысяч пар нуклеотидов
п.н.	пара нуклеотидов
ПСА	персульфат аммония
н.о.	нуклеотидный остаток
АР-сайт	апурин/апиримидиновый сайт
АРЕ1	апурин/апиримидиновая эндонуклеаза 1
BER	<u>B</u> ase <u>E</u> xcision <u>R</u> epair, эксцизионная репарация оснований
DMS	диметилсульфат
F-сайт	Остаток 2-оксиметил-3-окситетрагидрофурана или 1,2-дидезоксирибозы
HEPES	N-2-гидроксиэтилпиперазин-N'-2-этансульфоновая кислота
GHKL	<u>G</u> yrase, <u>h</u> eat-shock protein 90, histidine <u>k</u> inase, Mut <u>L</u> , семейство АТФаз
$K_d^{app}$	кажущаяся константа диссоциации
LB	бактериальная питательная среда Luna-Bertam
LOS	липоолигосахариды

MAC	мембраноатакующий комплекс
MMR	<u>M</u> ismatch <u>R</u> epair, репарация неканонических пар нуклеотидов («мисматчей»)
MutL-CTD	<u>C</u> - <u>T</u> erminal <u>D</u> omain of MutL, С-концевой домен белка MutL
MutL-NTD	<u>N</u> - <u>T</u> erminal <u>D</u> omain of MutL, N-концевой домен белка MutL
NER	<u>N</u> ucleotide <u>E</u> xcision <u>R</u> epair, эксцизионная репарация нуклеотидов
Ni-NTA	никель-нитрилтриуксусная кислота
OGG1	8-оксогуанин-ДНК-гликозилаза
8-охоG	8-оксо-2'-дезоксигуанозин
PCNA	<u>P</u> roliferating <u>c</u> ell <u>n</u> uclear <u>a</u> ntigen, ядерный антиген пролиферирующих клеток
PDB	<u>P</u> rotein <u>D</u> ata <u>B</u> ank, база данных пространственных структур макромолекул
PMSF	фенилметансульфонилфторид
RER	<u>R</u> ibonucleotide <u>E</u> xcision <u>R</u> epair; эксцизионная репарация рибонуклеотидов
RFC	<u>C</u> ellular <u>r</u> eplication <u>f</u> actor, фактор репликации клеток
RPA	<u>R</u> eplication <u>P</u> rotein <u>A</u> , репликативный белок А
PQS	<u>P</u> utative <u>G</u> - <u>Q</u> uadruplex <u>S</u> equences, последовательности, потенциально образующие G-квадруплекс
SSB	<u>S</u> ingle- <u>S</u> trand <u>B</u> inding protein, белок, связывающий одноцепочечную ДНК
TEMED	N,N,N',N'-тетраметилэтилендиамин
TRCF	<u>T</u> ranscription <u>R</u> epair <u>C</u> oupling <u>F</u> actor; фактор связи транскрипции и репарации
TSS	<u>T</u> ranscription <u>S</u> tart <u>S</u> ite, точка инициации транскрипции
VSP	<u>V</u> ery <u>S</u> hort <u>P</u> atch repair; репарация очень коротких участков

## **I. Введение**

### **I.1. Актуальность**

Поддержание стабильности генома является ключевым условием нормальной жизнедеятельности клетки. Нарушения функционирования систем репарации ДНК приводят к накоплению мутаций в генах, кодирующих белки, которые обеспечивают основные биологические процессы клетки. Кроме того, мутации нередко затрагивают регуляторные области генов, контролирующие клеточный цикл, апоптоз и пролиферацию у эукариот и адаптацию к внешним условиям прокариот. Особый интерес представляют GC-богатые области геномов, являющиеся «горячими» точками мутаций из-за высокого окислительно-восстановительного потенциала гуанина, а также способности цитозина к дезаминированию. Многие G-богатые мотивы склонны к формированию G-квадруплексных (G4) структур, которые являются неканонической формой ДНК. Анализ баз данных геномов как прокариот, так и эукариот, показывает большое количество потенциальных G-квадруплекс-формирующих последовательностей (PQS - putative quadruplex sequences) [1]. Считается, что G4 обладают широким спектром биологических функций, связанных с эпигенетической регуляцией. В патогенных для человека бактериях есть примеры G4, выступающих в роли факторов, необходимых для антигенной изменчивости, которая способствует «ускользанию» организма от иммунного ответа человека [2]. PQS часто встречаются в промоторных областях геномов эукариот, поэтому G4 могут являться регуляторным элементом транскрипции генов, а нуклеотидные замены dG в PQS, способные регулировать стабильность G4, нередко ассоциируют с онкогенезом. Таким образом, G-квадруплексы являются важной структурной формой организации ДНК, и их изучение в настоящее время является актуальной задачей [3].

Предположительно, GC-богатые области и формируемые ими в определенных условиях G4 являются барьером для репарации ДНК, поэтому мутации в GC-богатых областях генома появляются с высокой частотой [4]. Системные исследования особенностей взаимодействия таких участков генома с белками, ответственными за репарацию ДНК, отсутствуют. Научная проблема, на решение которой направлена данная работа, состоит в выяснении влияния GC-богатого состава ДНК на функционирование белков из систем репарации ДНК прокариот и эукариот. Решение этой проблемы важно с фундаментальной точки зрения для понимания причин возникновения мутаций в PQS, а также могут быть полезны для поиска новых подходов к терапии различных заболеваний.

### **I.2. Степень проработанности проблемы**

Интерес представляет GC-богатая регуляторная последовательность, локализуемая перед промотором гена белка пилина (*pilE*) патогена *Neisseria gonorrhoeae* (*N. gonorrhoeae*).

Бактерия использует гомологичную рекомбинацию для осуществления антигенной вариации, чтобы обойти иммунный ответ организма человека. Поверхностный белок PilE является основным компонентом пилей IV-типа, с помощью которых патоген перемещается и прикрепляется к эпителиальным клеткам организма-хозяина. По этой причине *pilE* - одна из мишеней антигенной вариации для блокирования ответа иммунной системы человека на этапе адгезии клеток патогена к эпителиальным клеткам. Перед промотором *pilE* обнаружена G-богатая последовательность, способная *in vitro* формировать стабильный параллельный G4. Полагают, что именно стабилизация G4 инициирует гомологичную рекомбинацию гена *pilE* в клетках *N. gonorrhoeae* [5–9]. На этот процесс оказывают влияние белки системы репарации неканонических пар нуклеотидов или «мисматчей» (MMR) патогена: MMR-дефицитные штаммы *N. gonorrhoeae* демонстрируют повышенный уровень антигенной вариации и, как следствие, более эффективную адгезию к эпителиальным клеткам [9]. MMR – это пострепликативная система, консервативная для прокариот и эукариот, которая восстанавливает комплементарность пар нуклеотидов, возникающую после репликации ДНК. Ключевыми белками MMR *N. gonorrhoeae* являются MutS и MutL. MutS узнает «мисматч» в ДНК и привлекает эндонуклеазу MutL, которая вносит одноцепочечный разрыв в дочернюю цепь ДНК, инициируя процесс репарации. Механизм функционирования гомологов MutL из метилнезависимой системы MMR, которая реализуется в клетках большинства организмов (в том числе *N. gonorrhoeae*), до сих пор неясен. Также неизвестно, могут ли белки MMR в клетках *N. gonorrhoeae* влиять на антигенную вариацию через связывание с G-квадруплексом перед *pilE*. Взаимодействие G4 с белком MutL ранее изучено не было. Интерес к данной проблеме обусловлен возрастающей резистентностью патогена к антибиотикам, а также высокой степенью изменчивости антигенов.

Важная задача исследования - изучить ограничения работы MMR не только в условиях образования G4, но и в условиях дуплексной ДНК при возникновении повреждения в GC-богатом контексте.

Помимо генома прокариот, G4-образующие последовательности встречаются в промоторах онкогенов, мутации в которых часто коррелируют с развитием онкологических заболеваний. Одним из примеров является промотор гена обратной транскриптазы теломеразы человека (*hTERT*) – фермента, активного в стволовых, половых и раковых клетках, поддерживающего длину теломерных повторов [10]. Рассматриваемый участок некодирующей цепи промотора *hTERT* имеет GC-богатую последовательность, в определенных условиях *in vitro* формирующую G-квадруплексную структуру, роль которой не вполне ясна. Мутации в промоторе *hTERT* встречаются при множественных видах рака и часто представлены заменами G→A в матричной G-богатой цепи в позициях -124, -146 и двойной заменой -138/-139 относительно стартового

кодона. Их называют «драйверными», так как данные мутации образуют новые участки связывания факторов семейства ETS, действие которых повышает экспрессию гена *hTERT* [11]. Считается, что описанные выше «драйверные» мутации возникают случайно под действием условий окружающей среды, а специфичность их появления достигается за счет селекции: если такая мутация возникла, то потенциал к делению клетки становится неограниченным, клетка накопит мутации и переродится в опухолевую. Активация теломеразы, позволяющая поддерживать длину теломер и избегать запрограммированной смерти клетки после определенного количества делений, происходит в клетках всех типов опухолей. Однако существуют опухолевые поражения, в которых мутации в промоторе гена *hTERT* не обнаружены. Это дает основание предполагать, что кроме селекции за счет повышенного потенциала к делению существует специфическое нарушение в функционировании одной или нескольких систем репарации, реализующееся в типах клеток, в которых частота возникновения «драйверных» мутаций гена *hTERT* при перерождении приближается к 100%. Это наблюдается в случае рака мочевого пузыря, при глиобластомах и меланомах. Механизмы, с помощью которых факторы окружающей среды вызывают последующее развитие рака таких типов, недостаточно изучены. Высокий мутагенный потенциал GC-богатых областей ДНК способствует возникновению повреждений: наиболее распространенной модификацией является апурин/апириимидиновый участок (AP-сайт), который репарируется с помощью фермента апурин/апириимидиновой эндонуклеазы 1 (APE1) [12]. APE1 является компонентом системы эксцизионной репарации оснований (BER). Молекулярный механизм BER достаточно хорошо охарактеризован в клетках человека. В то же время влияние GC-богатого окружения и G4 на эффективность работы APE1 изучено недостаточно.

### 1.3. Цели и задачи работы

Целью данной работы является биохимическая характеристика взаимодействия ключевых белков систем репарации MMR прокариот и BER человека с GC-богатыми последовательностями и образуемыми ими G-квадруплексными структурами *in vitro* на примере регуляторных областей генов *pilE* и *hTERT*.

Для достижения цели сформулированы следующие задачи:

1. установить влияние G-квадруплекса, потенциально формирующегося регуляторной G-богатой последовательностью перед геном *pilE*, на функционирование *in vitro* MutL *Neisseria gonorrhoeae*;
2. охарактеризовать способность белка MutS *Cereibacter sphaeroides* узнавать неканоническую пару нуклеотидов в GC-богатой (82%) последовательности ДНК;
3. выявить эволюционную консервативность G4-мотивов G-богатой области промотора

- генов обратной транскриптазы теломеразы (*TERT*) млекопитающих;
4. подтвердить формирование G-квадруплексной структуры *in vitro* в 68-звенном G-богатом участке промотора *TERT* человека и установить влияние стабильного аналога AP-сайта - остатка 1,2-дидезоксирибозы (F-сайта) в положениях, соответствующих «драйверным» мутациям, на стабильность G4;
  5. охарактеризовать эффективность связывания и гидролиза ферментом APE1 человека модельных одно- и двуцепочечных фрагментов ДНК, содержащих 68-звенную G-богатую последовательность промотора *hTERT* с одним или двумя аналогами AP-сайта, локализованными в положениях, соответствующих «драйверным» мутациям.

#### **I.4. Объект исследования**

Рекомбинантные белки из системы MMR прокариот: MutL из *Neisseria gonorrhoeae* и MutS из *Cereibacter sphaeroides*; белок из системы BER человека - APE1.

#### **I.5. Предмет исследования**

Взаимодействие белков систем репарации MMR и BER с GC-богатыми участками (и образуемыми ими G-квадруплексными структурами) регуляторных областей прокариот и эукариот на примере G4-мотива перед промотором гена поверхностного белка пилина *Neisseria gonorrhoeae* и промоторной области гена обратной транскриптазы теломеразы человека.

#### **I.6. Научная новизна**

Новизна нашей идеи заключается в предположении, что антигенная вариация *pilE Neisseria gonorrhoeae* регулируется белком MutL из MMR путем связывания с G4. В ходе предварительных исследований нами была детально охарактеризована эндонуклеазная активность MutL *N. gonorrhoeae* на линейном ДНК-субстрате. Впервые продемонстрировано измененное функционирование белка MutL в области G4: сродство белка к G4 заметно выше, тогда как катализ гидролиза фосфодиэфирных связей MutL реализуется до и после, но не внутри G4. Впервые показано, что MutS из *Cereibacter sphaeroides* не способен эффективно узнавать «мисматч» в GC-богатом окружении ДНК, что подтверждает высокий мутагенный потенциал GC-богатых участков и делает их потенциальным ограничением для MMR.

Научная новизна второй части работы заключается в выявлении причинно-следственных связей закрепления «драйверных» мутаций в промоторе *hTERT* и ограничений функционирования белка APE1 в GC-богатом контексте ДНК. В первую очередь, с помощью спектральных методов анализа было подтверждено образование G4 параллельной топологии G-богатой последовательностью промотора *hTERT*. Замена dG на стабильный аналог AP-сайта в

положениях, соответствующих «драйверным» мутациям, снижает термостабильность G4. Замены dC на стабильный аналог AP-сайта в комплементарной C-богатой цепи промотора *hTERT* в положениях -146 и -139/-139 приводят к образованию смеси ДНК-дуплексов и G4, то есть способствуют стабилизации G-квадруплексов в противоположной G-богатой цепи. Эффективность APE1-индуцируемого гидролиза ДНК, содержащей F-сайт, существенно снижается в условиях формирования G4, в то время как нуклеотидное окружение повреждения не влияет на этот процесс. Наши результаты свидетельствуют в пользу того, что закрепление «драйверных» мутаций в последовательности промотора *hTERT* может быть результатом сниженной активности APE1 в случае: (а) возникновения повреждения в положении -124 G-богатой цепи в условиях образования G-квадруплексной структуры, (б) возникновения повреждений в положениях -146 и -138/-139 C-богатой цепи в условиях дуплексной ДНК.

### **I.7. Теоретическая значимость работы**

Результаты работы существенно расширяют имеющиеся представления о функционировании систем репарации ДНК в GC-богатых регионах генома. Показано снижение способности белков MMR и BER узнавать и устранять повреждения в таких последовательностях.

Неспособность MutS узнавать «мисматч» в GC-богатом окружении, говорит о ранее не изученном вкладе в мутагенный потенциал подобных последовательностей. Изменение эндонуклеазной активности MutL в G4 указывает на возможность белка принимать участие в регуляции процессов рекомбинации. Работа вносит вклад в формирование теоретических представлений о молекулярных механизмах ограничения эффективности репарации ДНК в GC-богатых областях генома и роли взаимодействия белков системы MMR с нетипичными вторичными структурами ДНК, что выходит за рамки классического представления об этой системе репарации.

Активность APE1 как эндонуклеазы напрямую зависит от стабильности G4-структуры, образованной матричной цепью промотора *hTERT*, тогда как дуплетная GC-богатая последовательность не оказывает такого влияния. Результаты свидетельствуют в пользу гипотезы о том, что снижение эффективности BER в PQS промотора *TERT* человека может быть причиной закрепления «драйверных» мутаций в этой области и подтверждают вклад неканонических вторичных структур ДНК в процесс накопления соматических мутаций в регуляторных областях генома.

### **I.8. Практическая значимость работы**

Полученные данные об изменении функционирования белков MMR прокариот в GC-

богатых областях генома могут быть полезны для поиска новых подходов к терапии инфекционных заболеваний. Альтернативные мишени необходимы для рационального применения антибиотиков в связи с растущей тенденцией к развитию множественной резистентности патогенов. Методология, использованная для исследования консервативных G4-мотивов *TERT*-промоторов млекопитающих может быть использована для выявления PQS в других регионах генома, включая промоторы менее изученных онкогенов. Создание «библиотек» GC-богатых последовательностей регуляторных областей генома, склонных к накоплению мутаций, значительно упростит диагностику онкологических заболеваний на ранних стадиях. Данные о влиянии AP-сайтов на стабильность G4 вносят вклад в понимание роли этих модификаций как эпигенетических регуляторов, что может быть использовано для разработки новых методов таргетной терапии различных заболеваний, направленных на стабилизацию или дестабилизацию G4. Обнаруженное нами снижение эффективности репарации повреждений в GC-богатых областях генома стимулирует к изучению стратегий, направленных на регуляцию активности ферментов из систем репарации ДНК. Это может способствовать разработке новых противоопухолевых препаратов.

### **I.9. Методология диссертационного исследования**

При проведении экспериментов использовали современные методы биоорганической химии и молекулярной биологии. Рекомбинантные белки MutS и MutL были выделены хроматографически на Ni-NTA-агарозе с последующей очисткой эксклюзионной хроматографией. АТФазная активность MutL изучена колориметрическим методом. Эффективность связывания белков с ДНК определялась методом «торможения в геле» или путем измерения интерферометрии биослоя. Анализ ферментативной активности белков проводили с помощью электрофоретического разделения продуктов реакции в полиакриламидном геле (ПААГ) в денатурирующих условиях. Структурные особенности модельных ДНК, представляющих собой GC-богатые участки исследуемых регуляторных последовательностей, определялись современными спектральными методами (спектроскопия кругового дихроизма и УФ-спектроскопия). Для биоинформатического анализа использовали базы данных, содержащие геномную информацию различных организмов.

### **I.10. Личный вклад автора**

Автором сформулированы задачи и план экспериментальной работы. Лично получены препараты рекомбинантного белка MutL, проведена характеристика его активности. Спроектированы модельные ДНК различной длины с G4-мотивом перед *pilE*, исследованы их структура и взаимодействие с MutL. Для оценки влияния GC-богатого окружения на узнавание

«мисматча» белком MutS исследована ДНК-связывающая способность белка с ДНК разного GC-состава и разработан подход к анализу зависимости эффективности комплексообразования от концентрации белка с использованием трех различных моделей. Спроектированы модельные ДНК различной длины, содержащие 68-звенную GC-богатую последовательность промотора *hTERT* с остатками 1,2-дидезоксирибозы (аналогами AP-сайта), локализованными в положениях, соответствующих «драйверным» мутациям. Спектральными методами анализа установлено влияние аналогов AP-сайта на термостабильность G4 модельных ДНК. Охарактеризована эффективность связывания и гидролиза ферментом APE1 человека модельных одно- и двуцепочечных фрагментов ДНК с одним или двумя аналогами AP-сайта. Проведен биоинформатический анализ баз данных для оценки консервативности G4-мотивов в промоторах *hTERT* млекопитающих. Вклад Савицкой В.Ю. в большинстве опубликованных работ является определяющим.

#### **I.11. Положения, выносимые на защиту**

1. Белок MutL из системы MMR *Neisseria gonorrhoeae* эффективно связывает G-квадруплекс, но не вносит одноцепочечный разрыв внутри G4-мотива.
2. Белок MutS из системы MMR *Cereibacter sphaeroides* не способен эффективно узнавать неканоническую пару в GC-богатом окружении ДНК-дуплекса.
3. 68-звенный участок G-богатой матричной цепи промотора *hTERT* образует G-квадруплекс параллельной топологии *in vitro*, стабильность которого снижается при наличии одного или двух остатков 1,2-дидезоксирибозы (стабильного аналога AP-сайта) в положениях, соответствующих «драйверным» мутациям.
4. Фермент APE1 эффективно связывается с G-квадруплексной структурой G-богатого фрагмента промотора *hTERT*, содержащей остаток 1,2-дидезоксирибозы, однако эндонуклеазная активность белка снижена в условиях формирования G4.

#### **I.12. Апробация работы**

Результаты исследований были представлены на 11 международных и всероссийских конференциях в виде устных и стендовых докладов. Работа была доложена на заседании кафедры химии природных соединений химического факультета МГУ имени М.В. Ломоносова.

#### **I.13. Публикации**

Результаты работы опубликованы в 6 статьях в российских и международных рецензируемых научных журналах, а также представлены в сборниках тезисов научных конференций.

#### **I.14. Структура и объем**

Текст диссертации состоит из списка сокращений, введения, обзора литературы, материалов и методов, результатов, обсуждения, выводов, списка литературы, который включает 214 источников. Объем диссертации 127 страниц, материал иллюстрирован 46 рисунками и содержит 13 таблиц.

## **II. Эксцизионная репарация ДНК прокариот и эукариот: сравнение механизмов, роль в поддержании целостности генома и реализации генетического разнообразия**

*(Обзор литературы)*

### **II.1. Эксцизионная репарация ДНК в клетках различных организмов**

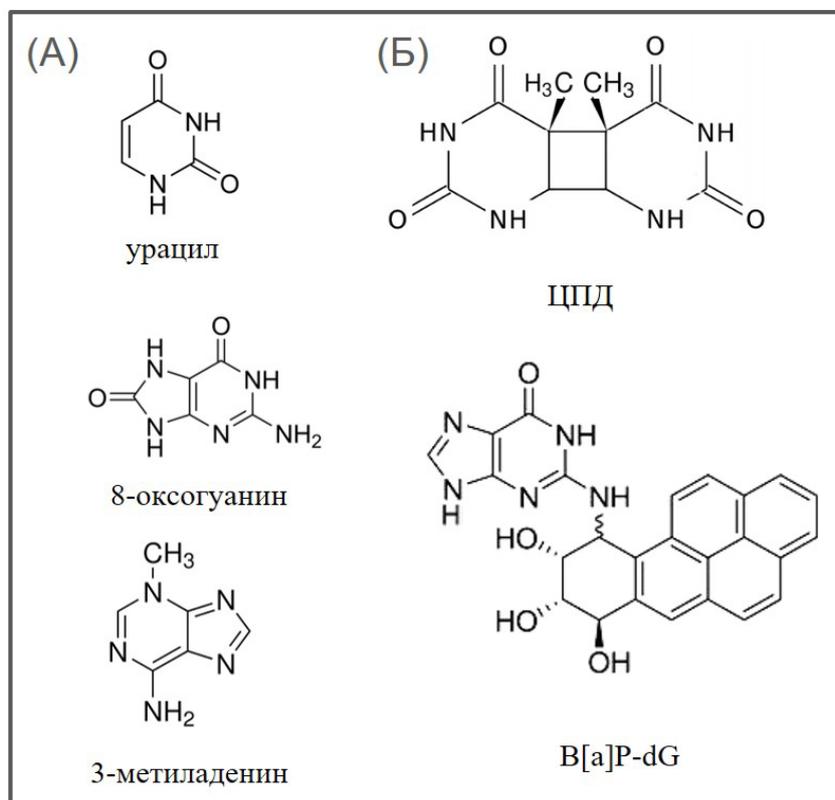
#### **II.1.1. Основные типы и общий механизм эксцизионной репарации ДНК**

ДНК всех клеток организма постоянно подвергается множеству модификаций, источниками которых могут быть рентгеновские лучи, токсины, вирусы, бактерии, активные формы кислорода (ROS) и др. [13]. Поскольку такие модификации представляют значительную угрозу для нормального функционирования организма, специализированные ферменты и белковые комплексы различных систем репарации ДНК узнают и исправляют повреждения. Процесс репарации ДНК реализуется в клетках всех живых организмов и направлен на поддержание стабильности генома и сохранение целостности генетической информации [14]. В зависимости от типа повреждения ключевыми белками могут служить гликозилазы, эндонуклеазы и другие белки, способные специфически узнавать и удалять повреждения, а также вспомогательные белки, координирующие взаимодействия [15–17]. Дисфункция белков систем репарации ДНК приводит к накоплению мутаций, что может стать причиной различных заболеваний, включая рак, в случае человека [18]. Репарация повреждений ДНК играет ключевую роль в поддержании генетической стабильности не только в клетках эукариот, но и прокариот, способствуя развитию адаптации клеток к условиям окружающей среды [19]. Механизмы процессов репарации ДНК у высших и низших организмов часто схожи, а аминокислотные последовательности ключевых ферментов этих систем обладают высокой степенью консервативности, что подчеркивает важность процесса репарации ДНК для выживания и адаптации всех организмов [20]. Тем не менее, у эукариот системы репарации ДНК устроены сложнее, чем у прокариот, что связано с эволюционным развитием высших организмов [21]. Как у эукариот, так и у прокариот существуют системы эксцизионной репарации оснований (BER), нуклеотидов (NER) и «мисматчей» (MMR), которые устраняют повреждения, возникающие по разным причинам, таким как окислительный стресс, действие ультрафиолетового излучения, ошибки полимераз и др. [22]. Далее рассмотрим более подробно системы эксцизионной репарации ДНК, функционирующие в большинстве живых организмов [23].

#### **II.1.2. Эксцизионная репарация оснований и нуклеотидов**

BER – это эволюционно консервативный путь репарации ДНК, исправляющий необъемные

модификации, искажающие структуру молекулы ДНК [24]. Примерами являются повреждения, возникшие в результате окисления пуриновых или пиримидиновых оснований (окисление гуанина гидроксильными радикалами с образованием 8-оксогуанина), дезаминирования (превращение аденина, гуанина и цитозина в гипоксантин, ксантин и урацил, соответственно) или алкилирования (метилирование S-аденозил-L-метионином с преимущественным образованием 3-метиладенина) (Рис. 1А) [25].



**Рисунок 1.** Примеры модификаций, узнаваемых и репарируемых ферментами BER и NER. (А) Модифицированные формы цитозина (урацил), гуанина (8-оксогуанин) и аденина (3-метиладенин). (Б) Циклобутан-пиримидиновый димер (ЦПД) и B[a]P-Gua: (+)-транс-бенз[а]пирен-модифицированный гуанин.

NER – более сложная система, исправляющая повреждения больших участков ДНК, включающих одно или несколько поврежденных оснований. Примером могут быть фотоповреждения по типу циклобутан-пиримидиновых димеров, такие как пиримидин-(6,4)-пиримидиновые димеры, преимущественно образующиеся в ходе ковалентных «сшивок» тимина и цитозина (Рис. 1Б) [26–30].

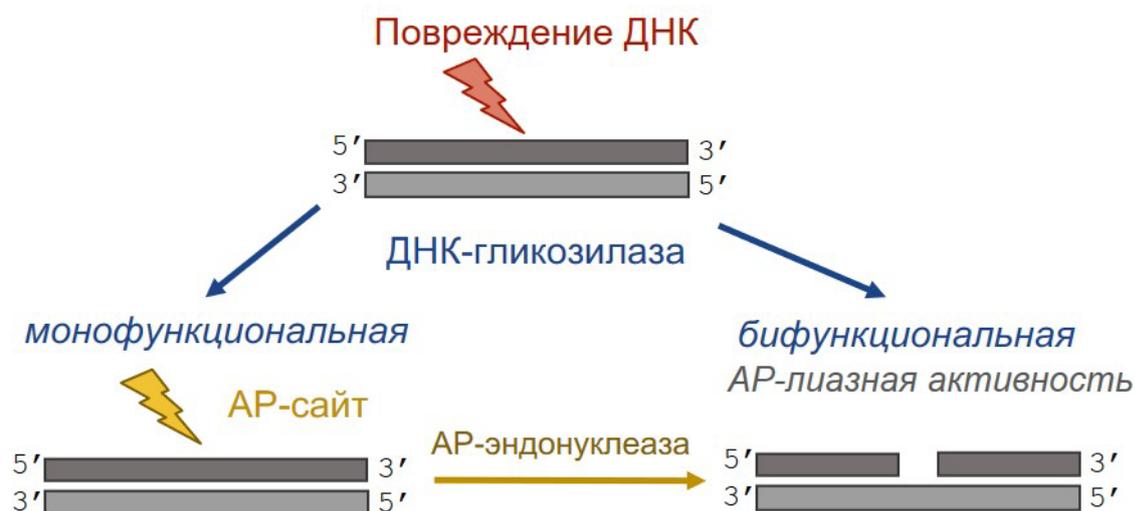
В клетках эукариот белки эксцизионной репарации функционируют как в ядре, так и в митохондриях, защищая организм от неконтролируемого деления клеток, процессов старения и нейродегенеративных заболеваний в случае млекопитающих [29,31]. В отличие от эукариот бактериальные клетки не имеют обособленного ядра, поэтому BER и NER прокариот осуществляется в цитоплазме, где расположена их ДНК. Мутации в генах белков систем BER и

NER бактерий могут приводить к снижению устойчивости организма к действию УФ-света и повышенной чувствительности к окислительному и кислотному стрессам, что может снижать их выживаемость в неблагоприятных условиях [32,33].

Ферменты и белковые комплексы, участвующие в процессах BER и NER в клетках прокариот и эукариот, отличаются, однако основные этапы репарации у обоих типов организмов схожи. В клетках эукариот существуют гомологи прокариотических ферментов, которые выполняют аналогичные функции бактериальных ферментов BER и NER, однако часто репарация проходит с участием вспомогательных белков, адаптированных к более сложной клеточной структуре эукариот [28,34].

### II.1.2.1. Механизм эксцизионной репарации оснований

Инициация BER заключается в специфическом узнавании и удалении поврежденного основания ДНК-гликозилазой [35–37]. Фермент сканирует ДНК путем последовательных актов связывания и диссоциации в поисках специфичного к нему повреждения. После обнаружения соответствующего основания образуется каталитически активный комплекс фермента с субстратом. Далее осуществляется гидролиз N-гликозидной связи с образованием апурин/апириμιдинового участка (AP-сайта) (Рис. 2) [38].



**Рисунок 2.** Инициация BER. Монофункциональные ДНК-гликозилазы узнают и удаляют модифицированное основание, что приводит к образованию AP-сайта, далее AP-эндонуклеаза расщепляет углеводофосфатный остов ДНК. В случае бифункциональной ДНК-гликозилазы ее AP-лиазной активности достаточно для удаления AP-сайта.

Существует два типа ДНК-гликозилаз: монофункциональные и бифункциональные. Монофункциональные гликозилазы обладают только гликозилазной активностью, поэтому если в ходе репарации была задействована монофункциональная гликозилаза, то следующий шаг реализуется с помощью апурин/апириμιдиновой эндонуклеазы (AP-эндонуклеазы), которая

расщепляет углеводофосфатный остов ДНК с 5'-стороны от AP-сайта, оставляя 3'-концевой гидроксил (ОН) и 5'-дезоксирибозофосфат (dRp) на образующихся фрагментах ДНК [39]. Бифункциональные гликозилазы могут провести репарацию без участия AP-эндонуклеазы, так как в дополнение к гликозилазной активности они обладают AP-лиазной функцией [40]. После расщепления N-гликозидной связи AP-сайт атакуется аминогруппой бифункциональной ДНК-гликозилазы с образованием основания Шиффа (имина). AP-лиазная функция фермента активируется, и происходит катализ расщепления углеводофосфатного остова ДНК с 3'-стороны от повреждения по механизму  $\beta$ -элиминирования [35].

У эукариот и у прокариот существуют гомологичные ДНК-гликозилазы, которые действуют в зависимости от типа повреждения. В таблице 1 приведены примеры известных гомологов, участвующих в BER эукариот и прокариот в ответ на определенное повреждение ДНК.

**Таблица 1.** Примеры гомологичных ДНК-гликозилаз прокариот и эукариот [40].

<i>Прокариоты</i>	<i>Эукариоты</i>	<i>Пример узнаваемой модификации ДНК</i>
Udg (урацил-ДНК-гликозилаза) [48]	UNG (урацил-ДНК-гликозилаза) [40]	Дезоксиуридин (dU) [41]
Fpg (MutM) (формамидопиримидин-ДНК-гликозилаза) [49]	OGG1 (8-оксогуанин-ДНК-гликозилаза) [37]	8-Оксогуанин (8-охоG)
MutY [49]	MUTY [50]	Аденин (A) в паре с 8-охоG
AlkA (алкил-ДНК-гликозилаза) [51]	MPG (3'-метиладенин-ДНК-гликозилаза) [52]	3'-Метиладенин
Nei [53]	NEIL [54]	Формамидопиримидины (FapyG и FapyA)

AP-эндонуклеаза, существующая в клетках прокариот и эукариот, удаляет AP-сайты, которые остаются после действия ДНК-гликозилаз. При расшифровке кристаллических структур комплекса человеческой AP-эндонуклеазы с ДНК, содержащей 1,2-дидезоксирибозу (F-сайт, аналог AP-сайта), было показано, что фермент выворачивает AP-сайт из двойной спирали, размещая его в активном центре и изгибая ДНК на  $35^\circ$  [41,42]. Для катализируемого AP-эндонуклеазой расщепления ДНК необходим двухвалентный катион металла, наиболее оптимальным считается ион  $Mg^{2+}$  [43]. В клетках всех организмов AP-

эндонуклеаза обладает аминокислотными последовательностями, консервативными в высокой степени, а мутации в генах, кодирующих этот фермент, часто неблагоприятно влияют на функционирование клеток [44,45]. В клетках млекопитающих AP-сайт является типичным геномным повреждением, которое накапливается в количестве более 10000 в день в одной клетке [46]. Полиморфизм гена, кодирующего AP-эндонуклеазу в клетках человека, ассоциирован с различными онкологическими заболеваниями, включая гепатоцеллюлярную карциному [47].

После удаления AP-сайта BER может пойти двумя путями: «короткозаплаточный», в котором заменяется нуклеотид, и «длиннозаплаточный» – в котором заменяется несколько нуклеотидов в цепи ДНК. У бактерий «короткозаплаточный» путь репарации является доминирующим, а «длиннозаплаточный» механизм изучен гораздо хуже. Основную роль в «заполнении бреши», образовавшейся в ходе удаления AP-сайта, играет ДНК-полимераза  $\beta$  (POL $\beta$ ), затем ДНК-лигаза восстанавливает целостность молекулы [19]. В случае эукариот выбор между «короткозаплаточным» и «длиннозаплаточным» путем BER зависит от 5'-dRp-конца [37]. Преобладающим также является «короткозаплаточный» путь BER, реализуемый с помощью фосфодиэстразной и полимеразной активностей POL $\beta$ . Затем ДНК-лигаза III (LIG3), работающая в комплексе с белками XRCC1 (от X-ray repair cross complementing 1) и PARP1 (от Poly [ADP-ribose] polymerase 1), лигирует одноцепочечный разрыв [12,55]. «Длиннозаплаточный» путь оказывается более предпочтительным, например, при низких концентрациях АТФ. Тогда ресинтез новой цепи осуществляется ДНК полимеразой  $\delta$  или  $\epsilon$ , добавляющей 2-8 комплементарных нуклеотидов в репарационный разрыв. В результате образуется 5'-одноцепочечный фрагмент ДНК, не содержащий участков комплементарности (5'-«flap» участок), удаляемый PCNA-зависимой ((Proliferating Cell Nuclear Antigen - ядерный антиген пролиферирующих клеток) FEN1-эндонуклеазой (flap structure-specific endonuclease 1). Разрыв ДНК ликвидируется PCNA-зависимой ДНК-лигазой I [56].

### **II.1.2.2. Механизм эксцизионной репарации нуклеотидов**

Механизм NER для прокариот и эукариот также похож и заключается в том, что белки этой системы узнают поврежденный участок за счет объемности модификации, инициируя его удаление (вырезается обычно участок длиной в 12–30 нуклеотидов, таблица 2). Существуют два пути NER: общегеномная репарация (GG-NER) и репарация, сопряженная с транскрипцией (TC-NER). TC-NER устраняет повреждения в транскрибируемой цепи «активных» генов и инициируется, когда такие повреждения препятствуют работе РНК-полимеразы [57]. В процессе общегеномной NER устраняются модификации, возникающие во всем геноме, включая нетранскрибируемые участки и области конденсированного хроматина у эукариот [58]. Эти два

пути отличаются стадией инициации и процессом обнаружения повреждения, однако затем репарация происходит по одному механизму.

**Таблица 2.** Сравнение белков NER прокариот и эукариот [14,19].

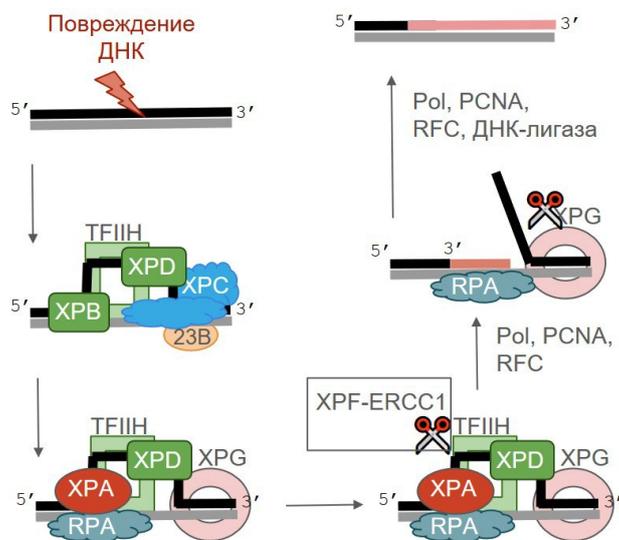
Прокариоты	Функции	Эукариоты	Функции
UvrA	узнает повреждения	ХРС в комплексе с Rad23B	узнает повреждения
UvrB	расщепляет фосфодиэфирные связи в ДНК и «выпрямляет» поврежденный участок	ХРА	стабилизирует поврежденный участок ДНК и помогает в образовании предрепарационного комплекса
		TFIIH (ХРВ и ХРД)	«разворачивают» ДНК вокруг поврежденного участка, обеспечивая доступ для вырезания поврежденных фрагментов
UvrC	гидролиз ДНК рядом с повреждением с двух сторон с удалением модифицированного участка	ХРФ	гидролиз ДНК с 5'-конца от повреждения
		ХРГ	гидролиз ДНК с 3'-конца от повреждения
UvrD	хеликаза, которая помогает удалить поврежденный участок после расщепления ДНК	ERCC1	участвует в хеликазной активности и помогает в процессе удаления поврежденных участков после их вырезания из ДНК

У бактерий NER начинается с узнавания повреждения ДНК белковым комплексом UvrA-UvrB [27]. При обнаружении повреждения UvrA высвобождается из комплекса, оставляя UvrB, связанный с поврежденным участком ДНК. Далее к комплексу присоединяется UvrC, и образующийся комплекс UvrBC, обладающий эндонуклеазной активностью, инициирует два разрыва в поврежденной цепи ДНК: один – с 5'-стороны от повреждения на расстоянии около 8 нуклеотидов, и второй – с 3'-стороны на расстоянии 4–5 нуклеотидов. В результате удаляется олигонуклеотид длиной 12–13 звеньев [59]. Наконец, содержащий повреждение фрагмент ДНК высвобождается хеликазой UvrD [60]. Образовавшаяся ОН группа на 3'-конце неповрежденного фрагмента ДНК используется ДНК-полимеразой I для репаративного синтеза ДНК, затем разрыв лигируется ДНК-лигазой. В клетках эукариот за реализацию NER отвечают белки с аналогичными функциями, устроенные сложнее (Таблица 2).

В эукариотических клетках, включая клетки человека, процесс NER более сложен [61]. В этом механизме задействовано свыше тридцати белков, которые постепенно образуют комплексы на участке поврежденной ДНК (Рис. 3). Несмотря на сложность NER, на сегодняшний день определены основные ферменты и белковые факторы, участвующие в процессе NER эукариот, а также изучены механизмы их действия [58]. Исследования показали, что ХРС-RAD23В является начальным фактором узнавания повреждений ДНК в клетках человека по механизму GG-NER [62]. Считается, что при сканировании ДНК в поисках повреждений ХРС-RAD23В, вступая в акты ассоциации-диссоциации с молекулой ДНК, формирует множество короткоживущих комплексов [63]. В случае обнаружения повреждения RAD23В высвобождается из ДНК-белкового комплекса, а фактор ХРС формирует стабильный комплекс с ДНК, к которому привлекаются последующие белки NER [64]. Для узнавания некоторых повреждений могут также потребоваться дополнительные белковые факторы, например, UVDDDB (от UV-damaged DNA-binding protein) [65].

Далее транскрипционный фактор ТФИИН взаимодействует с комплексом ХРС-ДНК. ТФИИН состоит из множества субъединиц, включая ХРВ и ХРД, проявляющих хеликазную и АТРазную активности, функции которых значительно отличаются в процессах транскрипции и репарации. Также в состав ТФИИН входит комплекс CDK-активирующей киназы (циклин-зависимой киназы, англ. «cyclin-dependent kinase») САК [66]. В процессах транскрипции ХРД взаимодействует с САК, что подавляет «ненужные» репарационные активности. Для «переключения» ТФИИН на репарацию фактор ХРА (от англ. «xeroderma pigmentosum complementation type A») связывается с комплексом ТФИИН и высвобождает САК из комплекса с другими субъединицами [67]. Субъединица ХРД становится способной взаимодействовать с ДНК и осуществлять проверку наличия в ней повреждения [68]. Остановка ХРД на повреждении позволяет сформировать предрепаационный комплекс путем привлечения ХРА и ХРГ. Эндонуклеаза ХРГ на этом этапе не гидролизует ДНК. На разных этапах NER фактор ХРА контактирует с белками ТФИИН, RPA, ХРС, DDB2, ERCC1-XPF (от англ. «excision repair cross complementing-group 1» и «xeroderma pigmentosum type F») и PCNA [69]. RPA служит для защиты ДНК от деградации нуклеазами [70]. Когда образуется стабильный комплекс факторов ТФИИН, ХРА, RPA и ХРГ, привлекается эндонуклеаза XPF-ERCC1, которая гидролизует ДНК с 5'-стороны на расстоянии 15–25 нуклеотидов от повреждения. Образованная при этом свободная 3'-ОН группа используется ДНК-полимеразами при репликативной застройке «бреши». Конформационные изменения комплекса факторов приводят к активации ХРГ, которая вносит в ДНК разрыв с 3'-стороны на расстоянии 3–9 нуклеотидов от повреждения [58]. Свободный 5'-фосфат необходим для лигирования разрыва в ДНК при завершении процесса репарации [71]. Завершают процесс

репарации NER ДНК-полимеразы со вспомогательными белками (ДНК-полимеразы  $\delta$ ,  $\epsilon$  и факторы RPA, PCNA и RFC у эукариот), встраивая в ДНК «правильный» нуклеотид, комплементарный основанию матрицы. ДНК-лигаза соединяет концы, восстанавливая целостность молекулы [72].

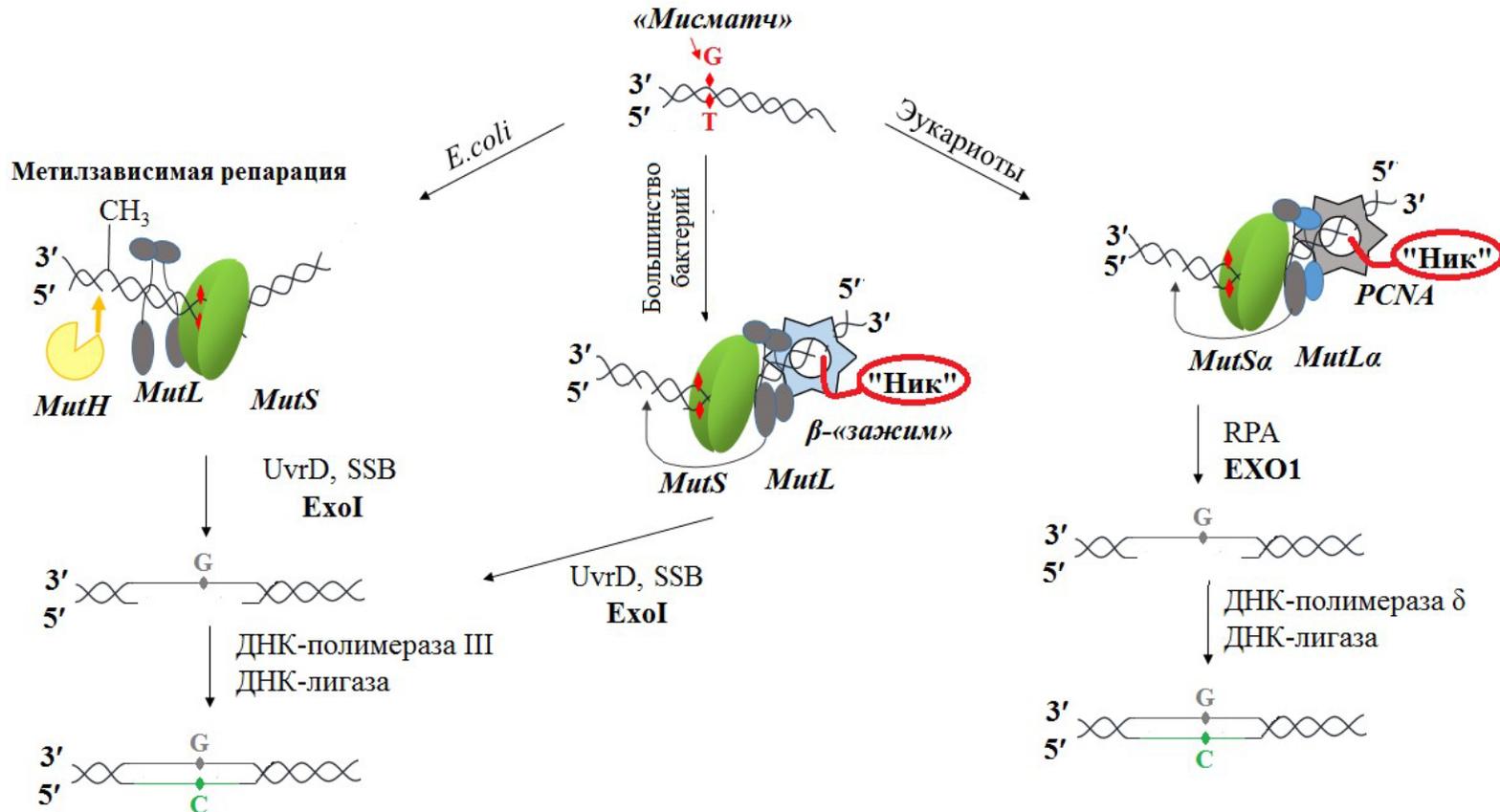


**Рисунок 3.** Схематичный механизм реализации репарации NER в клетках эукариот.

### II.1.3. Репарация неканонических пар нуклеотидов или «мисматчей»

Не менее распространенной в клетках прокариот и эукариот является система репарации «мисматчей» (MMR). MMR – это система, исправляющая ошибки «дочерней» цепи ДНК, возникающие при репликации, такие как неканонические пары нуклеотидов («мисматчи»). Несмотря на эволюционно консервативный механизм MMR, в разных организмах имеются свои особенности [73]. На сегодняшний день наиболее детально процесс MMR изучен в организме *E. coli*, в клетках которого реализуется метилзависимый механизм удаления «мисматча» [74]. Ключевыми белками системы являются MutS - ДНК-сканирующий белок, узнающий «мисматч», MutL - белок, координирующий взаимодействия при образовании комплекса MutSL (комплекс MutS с MutL) на ДНК, и MutH - никующая эндонуклеаза, вносящая разрыв в дочернюю цепь, после образования тройного комплекса MutSLH на ДНК (Рис. 4, левая панель) [75,76]. В поисках неканонической пары ДНК, MutS «скользит» по молекуле, частично изгибая ДНК [77]. При обнаружении мисматча MutS прочно связывается с ДНК, изгибая ее на  $\sim 45-60^\circ$ . АТФазная активность MutS обеспечивает его участие во всех ключевых стадиях MMR [78]. Связывание АТФ переводит MutS в конформацию «скользящего зажима», позволяя белку перемещаться вдоль ДНК и взаимодействовать с MutL, инициируя дальнейшую репарацию [79,80]. Основная роль белка MutL в метилзависимом пути MMR - координация взаимодействий, образование тройного комплекса MutSLH и активация эндонуклеазы MutH [81]. За гидролиз дочерней цепи

ДНК, где, как правило, и возникает «мисматч» в ходе репликации, отвечает никующая эндонуклеаза MutH, которая обладает высокой специфичностью. MutH узнает последовательность 5'-Gm6ATC-3'/3'-CTAG↓-5' и вносит разрыв в неметилированную цепь монометилированной последовательности (где m6A – N6-метил-2'-дезоксиаденозин; стрелкой указано место гидролиза) [82]. Перед началом репликации ДНК остатки аденозина обеих цепей метилируются ферментом Dam-метилтрансферазой по сайтам 5'-Gm6ATC-3'/3'-CTAG↓-5', а сразу после репликации – некоторое время метилирована лишь материнская цепь [83]. Одноцепочечный разрыв служит сигналом для эксцизионной репарации, при которой удаляется фрагмент цепи ДНК, содержащей «мисматч». К ДНК в месте разрыва присоединяется ДНК-хеликаза UvrD, которая расплетает ДНК до неканонической пары [84]. В зависимости от расположения разрыва (5' или 3'-сторона от «мисматча»), включаются специфические экзонуклеазы (RecJ/ExoVII или ExoI/ExoVII/ExoX, соответственно) для удаления дочерней цепи с «мисматчем». SSB-белок стабилизирует оставшийся одноцепочечный участок, ДНК-полимераза III (Pol III) синтезирует корректную последовательность, и ДНК-лигаза ликвидирует разрыв [85]. Метилзависимый путь репарации «мисматчей» реализуется лишь в клетках *E. coli* и некоторых других γ-протеобактериях. В клетках эукариот и у большинства других прокариот механизм действия MMR является метилнезависимым, отсутствует MutH, а метилирование не влияет на процесс дискриминации цепей ДНК [86]. Бактериальные клетки, в которых MMR метилнезависимая, обладают гомодимерными MutS и MutL, последний из которых функционирует как никующая экзонуклеаза (Рис. 4, центральная панель) [87]. В случае эукариот за MMR отвечают гетеродимерные белки MutSα (MSH2–MSH6), MutSβ (MSH2–MSH3) и MutLα (MLH1 и PMS2), MutLβ (MLH1 и PMS1) (Рис. 4, правая панель) [88]. В отличие от MutL белок MutS из метилнезависимой MMR в большей степени консервативен с MutS *E. coli*. Поэтому первый этап репарации, состоящий из сканирования ДНК и обнаружения неканонической пары, похож независимо от того, является ли MMR метилзависимой или нет [89,90]. Гомологи MutL проявляют эндонуклеазную функцию, которая активируется после узнавания «мисматча» и формирования комплекса MutSL [91]. За реализацию одноцепочечного разрыва отвечает эндонуклеазный мотив MutL, высоко консервативный среди гомологов белка в клетках разных организмов [92]. Для гидролиза ДНК требуется координация ионов двухвалентных металлов. Ион Mg<sup>2+</sup> является оптимальным, однако для некоторых организмов наиболее эффективно эндонуклеазная функция реализуется в присутствии Mn<sup>2+</sup> и Zn<sup>2+</sup> [86].



**Рисунок 4.** Механизмы MMR у *E. coli* (слева), других бактерий (в центре) и эукариот (справа). У бактерий гомодимерный MutS узнает «мисматчи» и образует комплекс с ДНК. Затем MutS привлекает MutL. Аналогично, у эукариот гетеродимерный MutS $\alpha$  привлекает MutL $\alpha$  после узнавания «ошибки». У  $\gamma$ -протеобактерий (включая *E. coli*) комплекс MutSL активирует эндонуклеазу MutH, которая катализирует гидролиз неметилированной дочерней цепи ДНК (метилзависимая MMR). У большинства прокариот и всех эукариот MutH отсутствует. У этих организмов гомологи MutL (гомодимер MutL у бактерий и гетеродимер MutL $\alpha$  у эукариот) обладают эндонуклеазной активностью и селективно гидролизуют вновь синтезированную цепь (метилнезависимый MMR).

Белки MutL в метилнезависимом пути MMR не имеют сайт-специфичности, поэтому могут гидролизовать ДНК в любом месте как «дочерней», так и «материнской» цепи ДНК. Это обстоятельно требует строгой регуляции активности MutL дополнительными факторами, чтобы избежать двуцепочечных разрывов, наиболее губительных для клетки. Механизм узнавания дочерней цепи у большинства организмов до сих пор неизвестен, однако предполагается, что белки MMR различают «материнскую» и «дочернюю» цепи с помощью специфических маркеров, образующихся во время репликации [93,94]. Например, во время репликации в эукариотах такими маркерами могут быть ошибочно включенные в ДНК рибонуклеотиды вместо дезоксирибонуклеотидов. После репликации эти рибонуклеотиды удаляются с помощью РНКазы, что позволяет системе MMR узнавать образовавшиеся одноцепочечные разрывы («ники») и исправлять некомплементарную пару [95,96]. Одноцепочечные разрывы также образуются в процессе репликации, особенно это касается «отстающей» цепи ДНК, синтезируемой фрагментами Оказаки. Важную роль в дискриминации цепей приписывают PCNA у эукариот или  $\beta$ -«зажим» у прокариот ( $\beta$ -субъединица ДНК-полимеразы III). Эти белки локализуются на дочерней цепи ДНК во время репликации и служат платформой для связывания других белков, участвующих в репарации, усиливая их процессивность [97,98]. Полагают, что взаимодействие MutSL с  $\beta$ -«зажимом» способствует корректной ориентации белкового комплекса для внесения одноцепочечного разрыва в «дочернюю» цепь.

После внесения одноцепочечного разрыва в ДНК в клетках прокариот включаются хеликаза (преимущественно UvrD), экзонуклеаза, ДНК-полимераза и ДНК-лигаза для завершения процесса репарации. В клетках эукариот завершение MMR осуществляется двумя путями: EXO1-зависимым и EXO1-независимым [99]. В EXO1-зависимом механизме 5'-3'-экзонуклеаза EXO1 удаляет участок дочерней цепи, содержащий ошибку, начиная с «ника» с 5'-стороны от «мисматча» [100]. Если «ник» расположен с 3'-стороны от некомплементарной пары, эндонуклеаза MutLa эукариот создает разрыв на 5'-стороне, обеспечивая начало EXO1-зависимой репарации. В EXO1-независимом пути MutLa вносит разрыв с 5'-стороны от «мисматча», а ДНК-полимераза  $\delta$  (Pol  $\delta$ ) синтезирует новый фрагмент ДНК, вытесняя ошибочную цепь с помощью своей экзонуклеазной активности [101]. Скорость EXO1-независимой MMR ниже, поэтому предпочтительным является EXO1-зависимый механизм.

Несмотря на различия в молекулярной организации BER, NER и MMR у прокариот и эукариот реализуются по единому принципу эксцизионной репарации: специфическое узнавание дефекта, локальное удаление поврежденного участка ДНК, ресинтез и последующее лигирование. Специфика каждого пути определяется типом повреждения и набором факторов, обеспечивающих его узнавание.

## II.2. Роль систем репарации ДНК в патогенезе бактерий

С развитием науки и медицины ученые стали чаще обращать внимание на проблему адаптации патогенных микроорганизмов, а именно на их устойчивость к антибиотикам, возникающую из-за различных факторов: работы эффлюксных систем, мутаций в генах-мишенях, компенсаторных изменений, приобретения плазмид и др. [102–104]. Бактерии также используют механизмы, затрудняющие узнавание патогена иммунной системой, такие как антигенная изменчивость, регуляция экспрессии поверхностных белков или имитацию гликопротеинов клеток организма-хозяина [105]. Несмотря на разработку новых препаратов и вакцин быстрое развитие механизмов адаптации микроорганизмов остается серьезной проблемой. Экспрессия и изменчивость антигенов, а также синтез генов, ответственных за устойчивость к лекарственным препаратам, контролируются на уровне ДНК [106]. В этих процессах часто участвуют белки рекомбинации и репарации, поэтому фундаментальные исследования, направленные на изучение механизмов работы белков систем репарации в клетках бактерий, важно для разработки новых эффективных вакцин, иммунотерапии и усовершенствования антибиотикотерапии.

Рассмотрим роль белков эксцизионной репарации в адаптации микроорганизмов на примере бактерии *Neisseria gonorrhoeae*. Этот организм является распространенным патогеном, способным вызывать серьезные заболевания у человека [106]. Бактерия демонстрирует способность развивать широкую и множественную устойчивость к различным антибиотикам, что осложняет лечение инфекционных заболеваний. Механизмы патогенеза, экологическая ниша, пути передачи, условия выживания и способы адаптации определяют характер модификаций, которые чаще возникают в ДНК патогенных организмов, и, как следствие, обуславливают развитие и специфику механизмов систем репарации, а также влияют на адаптацию систем репарации ДНК к специфическим условиям жизни каждой бактерии и роль белков репарации в патогенезе. *N. gonorrhoeae* представляет собой грамотрицательную бактерию, характеризующуюся наличием тонкого слоя пептидогликана и внешней мембраны, содержащей липополисахариды. Данный микроорганизм обитает исключительно в организме человека. *N. gonorrhoeae* поражает преимущественно слизистые оболочки мочеполовой системы и передается главным образом половым путем. По потребности в кислороде бактерия относится к облигатным аэробам. Существенной клинической особенностью *N. gonorrhoeae* является высокая устойчивость к антибиотикам, включая пенициллины, фторхинолоны и цефалоспорины, что существенно осложняет терапию гонококковой инфекции [107].

### II.2.1. Репарационные пути в клетках бактерии *Neisseria gonorrhoeae*

*N. gonorrhoeae* является патогеном человека и обитает в основном в организме хозяина [107]. Основные повреждения ДНК клетки патогена возникают в результате воздействия веществ, присутствующих в организме человека. *N. gonorrhoeae* обладает следующими системами репарации ДНК: прямое исправление повреждений, эксцизионная репарация оснований, эксцизионная репарация нуклеотидов, репарация «мисматчей», репарация очень коротких участков (VSP), репарация с участием систем рекомбинации (Таблица 3).

**Таблица 3.** Системы репарации ДНК *N. gonorrhoea* [109].

Система репарации ДНК	Ключевые ферменты, которые вовлечены в процесс репарации	Возможная причина возникновения мутации	Пример нарушения структуры ДНК
Система прямой репарации ДНК	O <sup>6</sup> -алкилуанин-ДНК-алкилтрансфераза II (Ogt)	алкилирующие агенты	O <sup>4</sup> -метилтимин и O <sup>6</sup> -метилгуанин
BER	ДНК-гликозилазы (Nth, Tag, Ung, MutM, MutY), AP-эндонуклеаза (Xth)	активные формы кислорода	7,8-дигидро-8-оксогуанин
NER	UvrABCD; Mfd	химические агенты (H <sub>2</sub> O <sub>2</sub> ), УФ-излучение	пиримидиновые димеры
MMR	MutS, MutL	«ошибки» ДНК-полимеразы	неканонические пары нуклеотидов
VSP	Vsr-эндонуклеазы; ДНК-полимераза I; MutS, MutL	Дезаминирование 5-метилцитозина	неканонические G/T-пары
Репарация с участием систем рекомбинации	RecABCD; RecGORQN; RuvABC	излучение различного типа, химические агенты	двухцепочечные разрывы

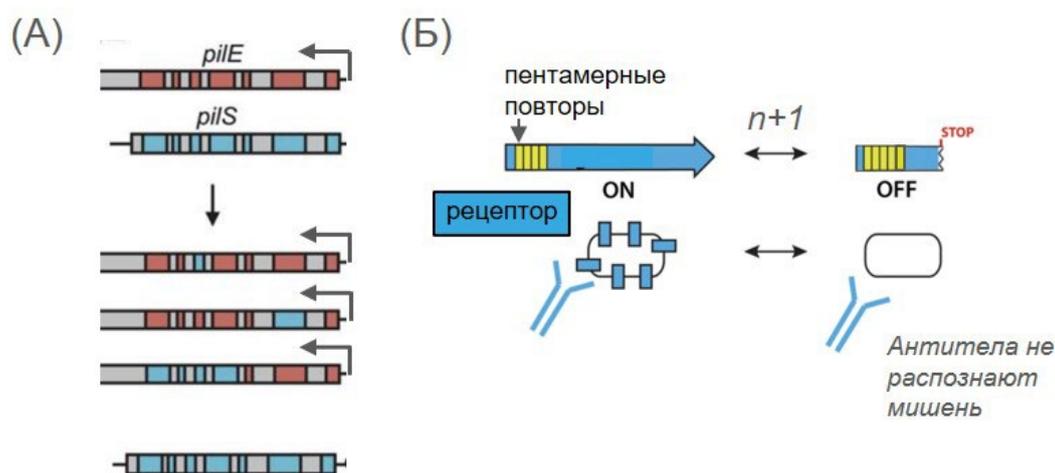
Считается, что у *N. gonorrhoeae* отсутствует система SOS-репарации, которая активируется в ответ на серьезные повреждения ДНК, такие как остановка репликационной вилки [108]. Вместо транскрипционного репрессора LexA, который подвергается автопротеолизу в классическом варианте SOS-ответа и активирует транскрипцию генов белков репарации, в клетках *N. gonorrhoeae* был выявлен его функциональный аналог – белок NG1427. Предполагается, что NG1427 может участвовать в похожем на SOS-репарацию пути, отличающемся, однако, по механизму активации.

В следующих разделах рассмотрим, как именно связана активность белков из систем репарации ДНК гонококка с их вирулентностью и способностью к адаптации. Автор диссертации является соавтором обзорной статьи [109], которая содержит информацию, описанную в разделах ниже.

### II.2.2. Ключевые факторы вирулентности *Neisseria gonorrhoeae*

Факторы вирулентности бактерий – это молекулы, обычно компоненты поверхности клетки, которые обеспечивают микроорганизму способность заражать организм хозяина и вызывать заболевание. Их изменчивость обеспечивает патогену возможность адаптироваться к условиям организма-хозяина и уклоняться от иммунного ответа. Роль белков систем репарации ДНК в обеспечении этих адаптационных механизмов активно обсуждается в последние годы.

*N. gonorrhoeae* обладает рядом факторов вирулентности, которые позволяют ей успешно колонизировать слизистые оболочки, уклоняться от иммунного ответа и выживать в стрессовых условиях организма-хозяина [110]. Первым этапом патогенеза является прикрепление бактерий к эпителию слизистой оболочки. Важную роль в этом процессе играют пили IV-го типа (IV type pili – Tfp) – белковые нити бактерии, которые позволяют *N. gonorrhoeae* закрепляться на поверхности клетки-хозяина и перемещаться по ней [106]. Пили преимущественно состоят из белка пилина (PilE) и дополнительных белков (PilC), они являются необходимым фактором вирулентности бактерии, поэтому претерпевают антигенную вариацию, что приводит к экспрессии различных вариантов белка пилина. Ген *pilE* и кодируемый им белок можно разделить на переменные и консервативные области. Процесс антигенной вариации пилей заключается в переносе переменной кодирующей последовательности из неэкспрессируемого «молчащего» локуса (*pilS*) в локус экспрессии (*pilE*) без изменения донорского локуса. Последовательности *pilS* не имеют промотора, но значительно схожи с экспрессируемой последовательностью *pilE*, являясь источником переменной генетической информации (Рис. 5А) [5].

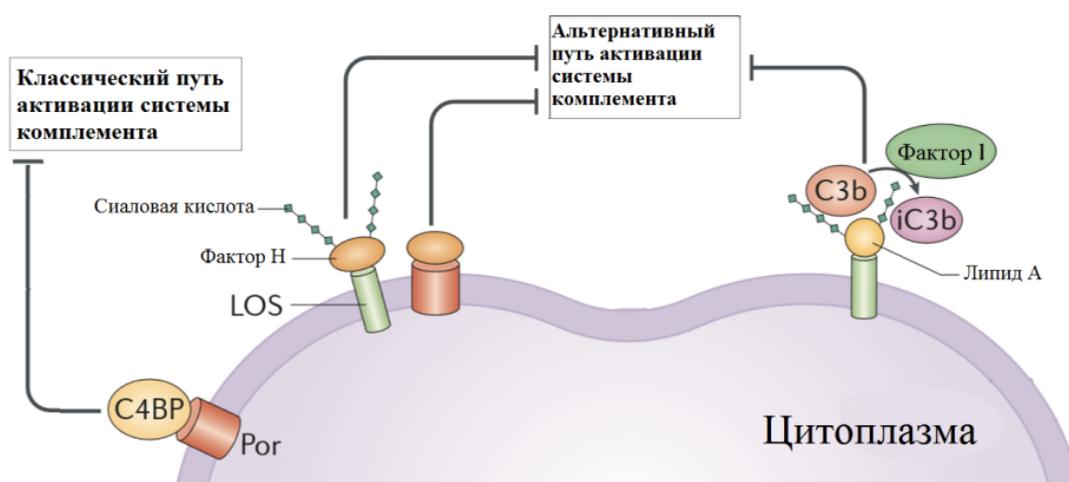


**Рисунок 5.** Механизмы адаптации *N. gonorrhoeae*. **(А)** Антигенная вариация поверхностного белка пилина *N. gonorrhoeae*. Графическое изображение гена *pilE* (красный) и *pilS* (голубой) с гомологичными областями, представленными серым цветом. **(Б)** Фазовая вариация поверхностного белка Ора *N. gonorrhoeae*. Изменение количества пентамерных повторов в гене, кодирующем сигнальный пептид. В результате экспрессия Ора подавляется и антитела организма-хозяина не связываются с рецепторами клеток патогена.

Другим важным фактором являются поверхностные белки Ора, участвующие в адгезии и инвазии путем связывания с рецепторами СЕАСАМ иммунных клеток человека. Это оказывает влияние на нейтрофилы, изменяя процессы фагоцитоза и нарушая выработку активных форм кислорода. Тем не менее, экспрессия Ора не является необходимым фактором для адгезии и инвазии, поэтому контролируется механизмом фазовой вариации, включающим изменение числа пентамерных повторов (5'-СТСТТ-3') в гене, кодирующем сигнальный пептид (Рис. 5Б) [111]. Такая изменчивость может обеспечить *N. gonorrhoeae* способность сохраняться в организме хозяина продолжительное время и повторно инфицировать.

Как и все грамотрицательные бактерии *N. gonorrhoeae* имеет липоолигосахариды (LOS) на поверхности внешней мембраны клетки, которые также необходимы для развития инфекции [110]. По структуре LOS *N. gonorrhoeae* подобны широко распространенным бактериальным липоолигосахаридам, однако не содержат полисахаридных цепей O-антигена, которые легко бы распознавались иммунной системой организма-хозяина при инфицировании. Более того, LOS *N. gonorrhoeae* подвергаются вариации фаз: в результате сдвига рамки считывания в генах гликозилтрансферазы липоолигосахариды бактерии приобретают различные углеводные фрагменты [112]. Например, экспрессия сиалированных липоолигосахаридов на поверхности клетки бактерии позволяет *N. gonorrhoeae* имитировать клетки организма-хозяина. У человека фрагменты сиаловой кислоты присутствуют на поверхности клеток и взаимодействуют с компонентами иммунной системы (с фактором Н – сывороточным гликопротеином системы комплемента) для предотвращения атаки на собственные клетки в результате активации альтернативного пути системы комплемента. Кроме того, как и другие бактериальные

липоолигосахариды, LOS *N. gonorrhoeae* содержит липид А, который, связываясь с белком каскада системы комплемента человека (C3b), вызывает деградацию последнего через опосредованный фактор I [113]. В результате этого каскад инактивируется, терминальный мембраноатакующий комплекс (МАС) не образуется, клетки патогена не подвергаются лизису, эффективно размножаясь в организме человека. Пориновые белки (PorA и PorB) – белки внешней мембраны клетки *N. gonorrhoeae*, являются наиболее распространенными белками в гонококках. Штаммы с преобладанием PorA вызывают более сложное течение заболевания, тогда как штаммы с преобладанием PorB на поверхности мембраны клеток патогена с большей вероятностью вызывают неосложненную инфекцию слизистых оболочек. Основная функция порина – перенос ионов и питательных веществ через внешнюю мембрану клетки бактерии. Для ускользания от иммунного ответа Por эффективно связывается с фактором Н, подавляя ответ системы комплемента человека. Более того, известно, что PorA взаимодействует с регуляторным белком системы комплемента эффективнее, чем PorB. С другой стороны, порин способен связываться с другим регуляторным участником системы комплемента – белком С4Вр, функции которого аналогичны функциям фактора Н, однако С4Вр защищает клетки хозяина от опсонизации в результате активации классического пути системы комплемента. Таким образом, *N. gonorrhoeae* эффективно колонизируется на поверхности эпителиальных клеток хозяина, избегая ответа системы комплемента человека (Рис. 6) [114].



**Рисунок 6.** Различные механизмы адаптации *N. gonorrhoeae* в ответ на активацию системы комплемента человека.

Помимо уклонения от иммунитета *N. gonorrhoeae* обладает высокой устойчивостью к повреждающим агентам благодаря ферментам антиоксидантной защиты и системам репарации ДНК. Иммунные клетки атакуют бактерию, производя активные формы кислорода и азота, которые повреждают ДНК. Для противодействия этому *N. gonorrhoeae* экспрессирует каталазы, супероксиддисмутазу, глутатионпероксидазу и метионин-сульфоксидредуктазу, снижающие

уровень окислительного стресса.

Представленные выше факторы вирулентности обеспечивают *N. gonorrhoeae* высокую степень адаптивности в условиях организма-хозяина. Центральным звеном такой адаптации выступает изменчивость белков наружной мембраны, антигенная и фазовая вариация которых лежат в основе иммунного уклонения. Все больше данных свидетельствует о том, что в управление этими процессами включаются белки систем репарации ДНК – как MMR, так и рекомбинационной и эксцизионной репарации. В следующем разделе рассмотрены молекулярные механизмы и экспериментальные данные, подтверждающие это участие.

### **II.2.3. Роль белков систем репарации ДНК в адаптации и изменчивости *Neisseria gonorrhoeae***

У *N. gonorrhoeae* важную роль не только поддержания стабильности генома, но и вирулентности, играют белки системы MMR [115]. Репарация в клетках *N. gonorrhoeae* идет по метилнезависимому механизму (см. раздел 2.1. «Репарационные пути в клетках бактерии *Neisseria gonorrhoeae*») [116]. Было обнаружено, что в клетках этой бактерии хеликаза UvrD не является необходимым участником MMR, хотя в клетках штаммов *N. gonorrhoeae* с хорошо представленной UvrD за расплетание ДНК в процессе MMR отвечает именно она [33]. В настоящее время в литературе появляются новые данные о влиянии белков системы MMR на различные процессы адаптации и изменчивости *N. gonorrhoeae* и его родственного вида *Neisseria meningitidis*, системы репарации ДНК которых аналогичны, как и процессы фазовой и антигенной вариации [117,118]. Так, продемонстрировано, что снижение активности белков MMR способствует видовой изменчивости и адаптации бактерий к условиям окружающей среды [119]. В работе показано, что нокауты генов *mutS* или *mutL* в штаммах *N. meningitidis* привели к резкому увеличению частоты появления спонтанной мутации устойчивости к рифампицину. Также описана природная клеточная линия, экспрессирующая ген мутантной формы белка MutL (32 аминокислотные замены по сравнению с ферментом дикого типа), которая обладает высокой частотой фазовых вариаций и быстро приобретает устойчивость к этому антибиотику. Изолированные клетки *N. meningitidis*, вырабатывающие мутантную форму MutS, также показывали высокую скорость изменчивости по данному признаку. Показано влияние белков MMR на вариацию поверхностного белка PorA. Штамм *N. meningitidis*, лишенный белка MutS, демонстрировал снижение экспрессии гена, кодирующего PorA. Это позволило патогену выработать устойчивость к антибактериальной активности PorA-специфического моноклонального антитела, являющегося основным компонентом ряда вакцин, снижая их эффективность [120]. По некоторым данным белки системы MMR *N. gonorrhoeae* могут принимать участие в регуляции гомологичной рекомбинации *pilE/pilS*, лежащей в основе

антигенной вариации пилина [121]. Мутантные линии патогена с нокаутами генов *mutS* или *mutL* демонстрируют повышенную частоту антигенной вариации *pileE* [122]. Предполагаемые механизмы влияния белков MMR на антигенную вариацию пилей обсуждаются далее (см. раздел 2.4. «Неканонические структуры ДНК и их влияние на патогенез и репарацию»).

Значительную роль в процессе антигенной вариации *N. gonorrhoeae* играют белки системы рекомбинационной репарации, которые также участвуют в восстановлении двойных разрывов ДНК [123]. В клетках *N. gonorrhoeae* реализуются два пути гомологичной рекомбинации: RecBCD- и RecF-подобный [7]. Как и другие бактериальные гомологи комплекс RecBCD из *N. gonorrhoeae* иницирует рекомбинацию, связываясь с «тупым» концом поврежденной ДНК [124]. Субъединицы RecB и RecD являются хеликазами и расплетают ДНК до определенной последовательности, называемой Chi-сайтом (5'-GCTGGTGG-3'). Далее, эндонуклеаза RecB вносит одноцепочечный разрыв в цепь ДНК, содержащую Chi-сайт, и продолжает расплетать двойную спираль. Комплекс ферментов RecBCD привлекает рекомбиназу RecA на образовавшийся 3'-конец одноцепочечной ДНК, затем RecBCD распадается. Одноцепочечный фрагмент ДНК в комплексе с RecA связывается с гомологичной последовательностью, а вторая цепь ДНК-дуплекса вытесняется с образованием структуры Холлидея [125]. Разрушение этой структуры происходит с помощью белков RuvABC и RecG, гены которых также обнаружены в геноме *N. gonorrhoeae* [126].

Мутации в *recA* могут нарушать процессы гомологичной рекомбинации, что снижает способность бактерии к репарации ДНК и может уменьшать эффективность антигенной изменчивости, необходимой для уклонения от иммунного ответа хозяина. Известно, что дефекты в последовательности гена *recB* (но не *recD*) гонококков приводят к нарушению репарации двуцепочечных разрывов ДНК, вызванных ионизирующим излучением [127]. В работе [7] показано, что выживаемость клеток *N. gonorrhoeae*, содержащих делеции генов *ruvA*, *ruvC* и *recG*, после воздействия на них H<sub>2</sub>O<sub>2</sub> была резко снижена по сравнению с клетками дикого типа.

Независимо от активности комплекса RecBCD гомологичная рекомбинация *N. gonorrhoeae* может происходить по так называемому RecF-подобному пути, в котором участвуют хеликаза RecQ, эндонуклеаза RecJ и белки RecO и RecR. Последние привлекают рекомбиназу RecA на одноцепочечный фрагмент репарируемой ДНК [128].

Механизм антигенной вариации пилина *N. gonorrhoeae* активно изучается. Известно, что в его реализации задействованы рекомбиназа RecA, ферменты RecF-подобного пути рекомбинации (RecOR, RecJ, RecQ), а также комплекс белков RuvABC и RuvG для разрушения структуры Холлидея [129]. Показано, что в результате делеций генов *recA*, *recO* и *recR* мутантные клеточные линии *N. gonorrhoeae* полностью теряли способность к антигенной вариации [130]. Нокауты генов *ruvA*, *ruvB*, *ruvC*, *recQ* или *recG* приводили к резко сниженной частоте этого

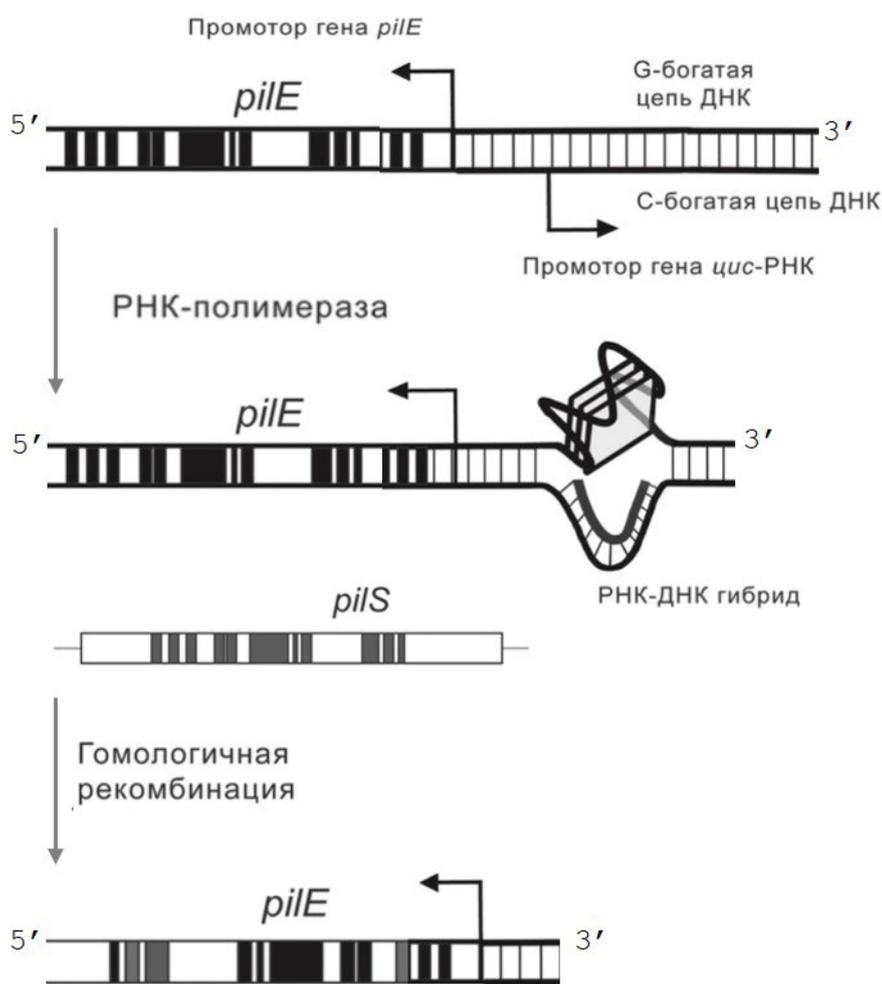
процесса [131]. Однако штаммы с инактивированными *recB* и *recD* демонстрировали ту же способность к вариации пилина, что и штаммы дикого типа [132]. Это свидетельствует в пользу того, что антигенная вариация пилей организма *N. gonorrhoeae* обусловлена гомологичной RecF-подобной рекомбинацией и не зависит от активности комплекса RecBCD [133].

Пути эксцизионной репарации оснований и нуклеотидов в ДНК *N. gonorrhoeae* служат преимущественно для устранения повреждений, вызванных различными экзогенными факторами, позволяя патогену выживать в среде организма-хозяина в ходе его инфицирования. Тем не менее, окислительные повреждения – не самые частые модификации ДНК данного патогена. Кроме своей основной роли, белки NER вносят вклад в изменчивость пилина. Несмотря на то, что они не вовлечены в антигенную вариацию *pilE*, показано, что нокаут *uvrD* (но не *uvrABC*) значительным образом увеличивает вариацию белка PilC, входящего в состав пилей IV-го типа [121]. Структурный белок PilC претерпевает фазовые вариации, однако по RecA-независимому пути. В *N. gonorrhoeae* присутствуют два идентичных гена *pilC1* и *pilC2*, которые способны «переключаться» между собой за счет сдвига рамки считывания в областях, кодирующих сигнальный пептид [134].

#### **II.2.4. Неканонические структуры ДНК и их влияние на патогенез и репарацию**

Одним из малоизученных, но потенциально значимых факторов, влияющих на репарацию и адаптацию патогенов, являются неканонические структуры ДНК. G-квадруплексы (G4) – одна из таких структур, формирующаяся в богатых гуанином участках генома и способная оказывать влияние как на экспрессию генов, так и на работу белков репарации. В клетках *N. gonorrhoeae* эти структуры играют ключевую роль в регуляции антигенной вариации пилей. Как упоминалось выше антигенная вариация *PilE* – ключевого белка пилей IV-го типа происходит через перенос изменчивых последовательностей из неактивных «молчащих» локусов (*pilS*) в активный локус *pilE* без изменения самих *pilS* с помощью белков RecF-подобного пути рекомбинации. Для инициации этого процесса необходима короткая G-богатая последовательность, расположенная примерно на расстоянии 350 пар нуклеотидов (п.н.) с 3'-конца от промотора *pilE* [135]. Эта последовательность способна образовывать параллельный G4 *in vitro* (Рис. 7) [136].





**Рисунок 8.** Стабилизация параллельного G-квадруплекса для инициации антигенной вариации *pilE* патогена *N. gonorrhoeae*. Гибридизация С-богатой последовательности ДНК с комплементарной *цис*-РНК стабилизирует структуру G4.

Для последующего гомологичного обмена, помимо инициации разрыва, необходимо локальное расплетение двойной спирали вблизи G4. Одним из кандидатов на эту роль является хеликаза ResQ, нокаут гена которой приводит к снижению частоты антигенной вариации [8]. Однако точечная замена в хеликазном мотиве (D307A), нарушающая способность ResQ расплетать G4, но сохраняющая активность в отношении обычной ДНК, не приводит к снижению вариативности. Это указывает на то, что ResQ, вероятно, не участвует напрямую в разрушении G4, но облегчает процесс рекомбинации, расплетая соседние с G4 участки дуплекса.

Таким образом, G-квадруплекс в *N. gonorrhoeae* может представлять собой функциональный регуляторный элемент, влияющие на активацию рекомбинационных белков и антигенную изменчивость. Это подчеркивает важную роль пространственной организации ДНК и структурных элементов генома в реализации вирулентных свойств бактерии.

Помимо *N. gonorrhoeae*, формирование G4-структур продемонстрировано и у других бактерий, включая *E. coli* [141]. Они локализуются рядом с локусами транспортной РНК и могут

участвовать в регуляции вирулентных генов. В геноме *Helicobacter pylori* обнаружены консервативные G-богатые мотивы, способные формировать G4 *in vitro*, расположенные в генах [142], кодирующих белки-транспортеры ионов Ni<sup>2+</sup> (*nixA*, *niuB1/2*, *niuD*). В *Streptococcus pneumoniae* идентифицированы три G4-мотива в кодирующих областях трех жизненно важных генов (*recD*, *pmrA*, *hdsS*) [143]. Предполагается участие этих мотивов в регуляции репарации, управлении экспрессией генов и реакцией на антибиотики.

Таким образом, понимание причин формирования и разрушения G4 в геноме бактерий может открыть новые подходы к ингибированию антигенной изменчивости и, как следствие, к созданию более эффективных терапевтических стратегий. Особое значение приобретает изучение взаимодействия белков с G-квадруплексами, в частности, белков систем репарации ДНК, функции которых могут изменяться вблизи G4-структур. Учитывая вовлеченность G4 в регуляцию антигенной изменчивости, исследование вопроса о том, как белки репарации (например, MutL, RecQ, RecA) взаимодействуют с такими участками, может пролить свет на механизмы переключения между стабильностью и изменчивостью генома *N. gonorrhoeae*. Это особенно важно в контексте противомикробной терапии и разработки препаратов, направленных на ограничение эволюционного потенциала патогенов.

### **II.3. Роль систем репарации ДНК в клетках человека**

Если в бактериях системы репарации совмещают функции поддержания геномной стабильности и адаптивной изменчивости, то в клетках человека репарационные механизмы жестко координируются с фазами клеточного цикла и преимущественно направлены на предотвращение накопления повреждений. Тем не менее, аналогичные принципы используются и для обеспечения генетического разнообразия, например, в процессе мейотической рекомбинации, который также опирается на работу белков репарации. Эукариотические клетки обладают более сложной структурой и многоуровневыми механизмами контроля за стабильностью генома на каждом этапе жизни, что особенно важно в контексте предотвращения канцерогенеза.

#### **II.3.1. Организация клеточного цикла человека и функционирование систем репарации ДНК**

Жизненный цикл эукариотической клетки состоит из нескольких последовательных фаз, каждая из которых имеет свои особенности как с точки зрения биохимических процессов, так и с точки зрения риска повреждения генома. Цикл деления клетки включает интерфазу и митоз [144]. Интерфаза, в свою очередь, делится на три фазы: G1 (период роста и подготовки к репликации), S (фаза синтеза, в ходе которой происходит репликация ДНК) и G2 (финальная

подготовка к митозу). Митоз – это фаза деления клетки, в ходе которой происходит равномерное распределение реплицированных хромосом между дочерними клетками [145]. Стоит отметить, что упаковка генетического материала в каждой фазе цикла деления не одинакова. Это обстоятельство значительно влияет на возможность возникновения повреждений в ДНК и активации тех или иных путей репарации [146,147]. У человека наиболее часто делятся клетки эпителия, костного мозга, стволовые, эмбриональные и половые клетки. Напротив, клетки нервной ткани делятся значительно реже, и многие из них не вступают в деление после завершения эмбрионального периода.

Фаза G1 (период роста клетки) является первой фазой интерфазы, в которой клетка активно готовится к репликации ДНК. На этом этапе хроматин находится преимущественно в конденсированном состоянии, что характеризуется высокой плотностью упаковки и формированием гетерохроматина [147]. Это способствует снижению доступности ДНК для внешних факторов и ограничивает вероятность возникновения мутаций. Однако даже в гетерохроматине возможны повреждения ДНК, поскольку низкомолекулярные окислительные агенты, такие как активные формы кислорода (АФК), способны проникать в эти области и вызывать окислительные модификации. Например, остатки 8-охоG, могут образовываться в результате воздействия окислителей и являться причиной нестабильности генома (см. раздел 1.2.1. «Механизм эксцизионной репарации оснований») [148]. При переходе фазы G1 в фазу S клетка должна пройти через проверочную точку («чекпоинт») G1, чтобы обеспечить контроль за целостностью генома. Эта стадия включает серию молекулярных механизмов, оценивающих повреждения ДНК и активирующих соответствующие пути репарации. Одним из ключевых белков, участвующих в этом процессе, является p53, который может инициировать репарацию, замедлить цикл или даже активировать апоптоз, если повреждения трудно исправляемые [149]. Белок p53 также может индуцировать экспрессию p21 – ингибитора циклин-зависимой киназы, что приводит к задержке клеточного цикла и предоставляет время для восстановления ДНК. На этом этапе также могут быть активированы киназы, такие как ATM и ATR, которые фосфорилируют и активируют эффекторные киназы, запуская каскад сигнальных событий, ведущих к задержке клеточного цикла и активации механизмов репарации ДНК. Активация ATM и ATR приводит к фосфорилированию и ингибированию фосфатазы Cdc25, что препятствует активации циклин-зависимой киназы и замедляет клеточный цикл [150]. Если клетка успешно исправляет повреждения, она может перейти в фазу S, где начинается репликация ДНК.

В фазе S клетка начинает активный процесс репликации ДНК, что сопровождается деконденсацией хроматина и переходом его в более открытое состояние. Этот процесс требует активной работы систем репарации, так как именно на этом этапе ДНК наиболее подвержена повреждениям. Деконденсация хроматина увеличивает доступность ДНК не только для

репликационных комплексов, но и для различных факторов, таких как окислители, ультрафиолетовое излучение или химические агенты, что создает угрозу для стабильности генома [151]. В процессе репликации ДНК часто возникают «мисматчи» из-за ошибок ДНК-полимераз при синтезе «дочерней» цепи на матрице исходной молекулы [152]. При расхождении цепей ДНК могут возникать промежуточные вторичные структуры (например, G4), которые могут приводить к остановке репликативной вилки, возникновению одноцепочечных или двуцепочечных разрывов [153]. Таким образом, активность белков систем репарации ДНК крайне важна в S-фазе клеточного деления, когда происходит репликация ДНК.

Фаза G2 – это заключительная подготовка клетки к митозу, в ходе которой осуществляется коррекция ошибок, возникших в процессе репликации ДНК. К концу G2 хроматин снова начинает конденсироваться, но в начале фазы сохраняет значительную степень деконденсации для возможности осуществления репарации [154]. Клетка проверяет правильность и целостность реплицированной ДНК с помощью специализированных белков и комплексов, таких как циклин-зависимые киназы и белки, участвующие в сигнальных путях апоптоза. На этом этапе также активно работают системы репарации, которые исправляют ошибки репликации или другие повреждения ДНК, возникшие в фазах G1 и S. Стоит отметить, что при переходе из фазы G2 в митоз клетка обязана пройти контрольную точку G2/M, регулируемую циклин-зависимыми киназами, обеспечивая проверку на наличие геномных повреждений [155].

В митозе клетка делится, и все процессы, связанные с делением хромосом, требуют особенно строгой координации. Все повреждения ДНК должны быть устранены до начала митоза, так как в митозе репарация ДНК ограничена, а наличие повреждений может привести к мутациям или неправильному распределению хромосом между «дочерними» клетками. В случае серьезных повреждений, которые не могут быть устранены, активируются сигнальные пути, ведущие к апоптозу, чтобы предотвратить деление клетки с дефектной ДНК [156].

Таким образом, активация систем репарации ДНК преимущественно происходит в интерфазе, когда клетка активно контролирует и исправляет повреждения генома. Репарационные механизмы у эукариот играют не только роль в рутинном восстановлении повреждений, но и имеют решающее значение в предотвращении канцерогенеза. Мутации в генах белков репарации могут выступать как драйверы опухолевого роста, а особенности строения ДНК – влиять на вероятность возникновения ошибок.

### **П.3.2. Клиническое значение мутаций в генах белков систем эксцизионной репарации ДНК человека**

Мутации в генах, кодирующих белки систем эксцизионной репарации ДНК, таких как BER, NER и MMR, широко ассоциируются с развитием различных типов рака. Нарушение механизма

репарации поврежденной ДНК способствует мутагенезу и хромосомной нестабильности, повышая риск канцерогенеза (Таблица 4).

В частности, полиморфизмы гена, кодирующего 8-оксогуанин-ДНК-гликозилазу человека (OGG1), которая отвечает за репарацию окислительных повреждений, ассоциированы с риском развития рака различного типа, включая рак легких, печени и молочной железы. Это подтверждает связь между нарушением репарации окислительных повреждений и развитием онкологических заболеваний. Замена Ser326Cys в последовательности OGG1 часто рассматривается как фактор повышенного риска развития колоректального рака, рака легких, рака мочевого пузыря и других видов рака [157]. Установлено, что гомозиготы Cys/Cys имеют повышенный риск возникновения рака легких, особенно при совместном действии факторов, таких как курение [164]. В мета-анализах было обнаружено множество других генетических вариантов OGG1, включая «миссенс»- и «нонсенс»-замены, а также варианты в интронах и 3'-нетранслируемой области (3'-UTR).

Кроме того, в раковых клетках различного типа были обнаружены мутации и в *APEX1* – гене, кодирующем эндонуклеазу человека, участвующую не только в гидролизе фосфодиэфирных связей, но и в других процессах (см. раздел 1.2.1. «Механизм эксцизионной репарации оснований»). Замена Asp148Glu в *APEX1* рассматривается в контексте онкологически значимых вариантов, но для включения ее в стандартные клинические панели онкомаркеров для BER необходимы дальнейшие исследования [158]. Связь с раком молочной железы и толстой кишки обсуждается в мета-анализах, но данные неоднозначны. В раковых клетках различных типов были выявлены полиморфизмы в *XRCC1* – гене, кодирующем белок, который играет ключевую роль в системе эксцизионной репарации оснований, выступая в качестве молекулярного каркаса для сборки репарационного комплекса. Полиморфизмы Gln399Arg и Arg194Trp *XRCC1* широко изучаются в контексте онкологической предрасположенности. Замена Gln399Arg в этом белке ассоциируется с изменением эффективности репарации ДНК и обсуждается в связи с повышенным риском рака легких, молочной железы и желудка, однако данные мета-анализов остаются противоречивыми из-за различий в популяциях и методах исследований [159]. Есть исследования, предполагающие влияние этих полиморфизмов в развитии острого миелоидного лейкоза. Включение *XRCC1* Gln399Arg и Arg194Trp в клинические панели онкомаркеров требует дополнительных подтверждений, в том числе функциональные исследования и крупные когортные анализы [165]. Среди компонентов системы NER особое внимание привлекает полиморфизм *ERCC2 (XPD)* Lys751Gln, широко изученный в контексте канцерогенеза. Этот вариант ассоциирован с повышенным риском развития различных видов рака, включая рак легких у курящих пациентов европейского происхождения [160]. Экспериментальные данные показали, что данная аминокислотная замена снижает

эффективность удаления бензо[а]пирен–индуцированных повреждений ДНК, что указывает на ее функциональную значимость и вклад в снижение эффективности репарационного пути [166]. Благодаря стабильной ассоциации с повышенным онкологическим риском и функциональной обоснованности XPD Lys751Gln рассматривается как потенциальный биомаркер индивидуальной чувствительности к канцерогенам и может быть включен в состав предиктивных панелей, ориентированных на оценку персонализированных рисков.

**Таблица 4.** Клиническое значение мутаций/полиморфизмов в генах белков систем эксцизионной репарации.

Система репарации	Ген или белок	Мутация или полиморфизм	Тип	Повышенный риск/ ассоциированные опухоли
BER	OGG1 [157]	Ser326Cys	«миссенс»	Рак легких, колоректальный рак, рак мочевого пузыря, рак молочных желез
	APE1 [158]	Asp148Glu	«миссенс»	Рак молочных желез и толстой кишки
	XRCC1 [159]	Gln399Arg	«миссенс»	Рак легких, молочных желез и желудка
		Arg194Trp	«миссенс»	Острый миелоидный лейкоз
NER	ERCC2 (XPD) [160]	Lys751Gln **	«миссенс»	Рак легких у курящих
MMR	MLH1 [161]	Arg687Trp *	«миссенс»	Синдром Линча (колоректальный рак)
	<i>MSH2</i> [162]	<i>A942+3T</i> *	сплайс-сайт	
	<i>MSH2</i> [163]	<i>C2152T</i> *	«нонсенс»	

\* входят в клинические панели онкомаркеров

\*\* могут входить в расширенные или исследовательские панели онкомаркеров

Нарушения в системе репарации «мисматчей», отвечающей за исправление ошибок, возникающих в ходе репликации ДНК, также играют важную роль в возникновении канцерогенеза. В частности, мутации в генах белков MSH2 и MLH1 (а также MSH6 и PMS2), которые являются ключевыми компонентами этого пути, представлены различными и множественными вариантами и часто наблюдаются при синдроме Линча - наследственном заболевании, характеризующемся высоким риском развития рака толстой кишки, эндометрия, а также ряда других опухолей, включая рак молочной железы. Одним из наиболее изученных примеров наследуемых мутаций синдрома Линча является «миссенс»-вариант гена *MLH1*

C2059T (в белке - Arg687Trp) [161]. Эта мутация документирована как «мутация основателя»: у пациентов из Швеции и Финляндии встречается всегда с одним и тем же участком окружающей ДНК (гаплотипом). Это значит, что она произошла у одного предка и передается от него. С другой стороны, мутации *MLH1* C2059T у носителей из Германии и Франции встречаются и на других гаплотипах, что говорит о независимом, повторном возникновении мутации в этих популяциях. В базе данных ClinVar она классифицирована как патогенная. Тестирование на наличие *MLH1* C2059T пациентов североευропейского происхождения позволяет повысить эффективность диагностики и снизить затраты, исключив необходимость полного экзомного скрининга на первом этапе.

Мутация гена *MSH2* A942+3T является одной из наиболее часто описываемых патогенных вариантов, связанных с синдромом Линча [162]. Она приводит к дефекту сплайсинга и пропуску экзона 5 при процессинге пре-мРНК, что нарушает структуру и функцию белка *MSH2*. Мутация занесена в базы ClinVar и InSiGHT как клинически значимая и патогенная. Она признана «мутацией основателя» в популяции Ньюфаундленда, а также обнаружена в ряде латиноамериканских семей. Благодаря высокой частоте и диагностической значимости, вариант включен в панели для целевого скрининга при подозрении на синдром Линча.

Замена C2152T в гене *MSH2* представляет собой «нонсенс»-мутацию, вызывающую преждевременную остановку трансляции. Она зарегистрирована как «мутация основателя» в португальских и бразильских семьях, и ассоциирована с синдромом Линча. Этот вариант часто сопровождается ранним началом заболевания, микросателлитной нестабильностью и развитием множественных опухолей [163].

Таким образом, мутации в белках систем репарации BER, NER и MMR имеют важное значение в развитии генетической нестабильности и часто играют роль драйверов канцерогенеза. Поиск и функциональная характеристика таких мутаций не только позволяет глубже понять механизмы развития рака, но и открывает перспективы для разработки методов ранней диагностики и индивидуализированного подбора терапии.

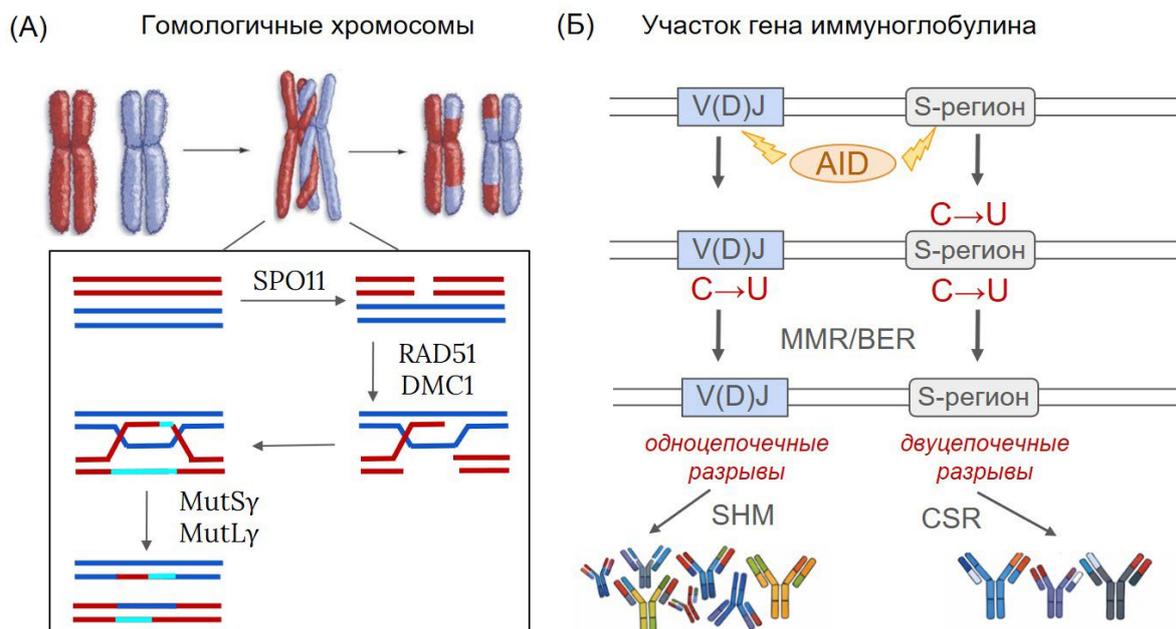
### **II.3.3. Роль белков систем эксцизионной репарации ДНК человека в поддержании генетического разнообразия**

Хотя основная функция систем эксцизионной репарации ДНК (BER, NER, MMR) заключается в предотвращении мутагенеза и поддержании стабильности генома, отдельные компоненты этих систем, а именно - специфические формы белков MMR, также играют ключевую роль в обеспечении генетического разнообразия в клетках зародышевой линии, участвуя в мейотической рекомбинации (Рис. 9А) [167]. Этот процесс сопровождается обменом участками ДНК между гомологичными хромосомами и происходит в клетках зародышевой

линии в ходе мейоза. Мейотическая рекомбинация необходима для формирования генетического разнообразия потомства, а также для корректного расхождения гомологичных хромосом на поздних стадиях мейоза.

В ходе профазы I мейоза в клетках человека программированные двуцепочечные разрывы ДНК индуцируются ферментом SPO11 [168]. Эти разрывы являются стартовым сигналом для начала гомологичной рекомбинации, в процессе которой происходит обмен участками между гомологичными хромосомами – кроссинговер, критически важный для правильного расхождения хромосом и формирования генетически разнообразных гамет. Формирование 3'-одноцепочечных выступающих концов в местах двуцепочечных разрывов необходимо для инициации гомологичной рекомбинации, и осуществляется с участием специализированных белков, обеспечивающих удаление нуклеотидов и формирование подходящего субстрата для последующего поиска гомологичной последовательности. Поиск гомологичного донора и внедрение цепи происходит при участии белков рекомбинации RAD51 и DMC1. Далее идет формирование промежуточных рекомбинационных структур, таких как D-петли и структуры Холлидея. Именно на этом этапе подключаются гомологи MMR: гетеродимер MSH4–MSH5 (MutS $\gamma$ ), стабилизирующий D-петлю и структуры Холлидея [169], предотвращая их распад, и гетеродимер MLH1–MLH3 (MutL $\gamma$ ), который действует как эндонуклеаза, способствующая разрешению структуры Холлидея с протеканием кроссинговера [170]. Завершение рекомбинации происходит с участием белков EXO1 и BLM, которые регулируют длину и направление обмена.

Таким образом, MutS $\gamma$  и MutL $\gamma$  – это специализированные формы белков, аналогичных белкам систем репарации «мисматчей» (MMR), адаптированные для участия в мейозе. Их активность строго ограничена клетками зародышевой линии и фазами мейоза, что отличает их от соматической формы MMR, где доминируют комплексы MSH2–MSH6 (MutS $\alpha$ ), MSH2–MSH3 (MutS $\beta$ ), MLH1–PMS2 (MutL $\alpha$ ) и др. Снижение активности MLH1, MLH3, MSH4 и MSH5 ассоциированы с нарушением мейотической рекомбинации и могут приводить к анеуплоидиям, субфертильности, преждевременному истощению яичников и ухудшению качества сперматозоидов. Один из наиболее изученных вариантов – MSH5 Pro29Ser [171]. Аминокислотная замена снижает способность MSH5 взаимодействовать с MSH4, что нарушает формирование гетеродимера MutS $\gamma$ . Генетические исследования в популяциях китайцев выявили, что носители аллеля Ser имеют повышенный риск азооспермии или олигозооспермии. Это доказывает функциональную значимость полиморфизма и обосновывает его перспективу в скрининге мужского бесплодия.



**Рисунок 9.** Участие белков систем эксцизионной репарации в поддержании генетического разнообразия. **(А)** MMR - разрешение структур Холлидея и D-петель в процессе кроссинговера в ходе профазы I мейоза в зародышевой линии клеток человека. **(Б)** MMR и BER - инициация разрывов в ДНК для реализации соматической гипермутации (SHM) и классовой перестройки иммуноглобулинов (CSR) в активированных В-клетках.

Другой значимый пример – *MLH3* C2531T (Pro844Leu) [172]. Исследователи из Ирана связывали этот полиморфизм с бесплодием у женщин: мутантный генотип ассоциировался с двукратным повышением риска бесплодия по сравнению с нативным. Аналогично, у мужчин с азооспермией или олигозооспермией носители Т-аллеля имели почти двукратный риск нарушения сперматогенеза. Рассматриваемый полиморфизм позволяет прогнозировать репродуктивную функцию и может использоваться для направленного генотипирования при диагностике бесплодия. Данное направление активно изучается: понимание молекулярных механизмов работы описанных ферментов может быть полезным для диагностики нарушений фертильности и разработки новых репродуктивных технологий.

Помимо участия в мейотической рекомбинации системы репарации играют важную роль в соматических механизмах адаптации. Особое значение здесь имеет BER и MMR, белки которых вовлечены в соматическую гипермутацию (SHM) и классовую перестройку иммуноглобулинов (CSR) в активированных В-клетках (Рис. 9Б) [173]. SHM инициируется ферментом AID (активационно-индуцируемая цитидиндезаминаза), который дезаминирует цитозин до урацила (U) в вариабельных областях V(D)J-экзонов. Неканонические пары U/G распознаются компонентами систем BER или MMR. В случае BER фермент UNG удаляет урацил, а эндонуклеазы APEX1/APE2 гидролизуют AP-сайт. В случае NER белок MutS узнает «мисматч», привлекая MutL для внесения одноцепочечного разрыва. Последующее вовлечение низкоточных

ДНК-полимераз способствует накоплению точечных мутаций. Этот механизм лежит в основе вариабельности антител [174].

При CSR иммуноглобулинов возникновение повреждений инициируется по тому же принципу, однако AID действует на цитозины в S-регионах, богатых повторами и интенсивно транскрибируемых («switch»-области, расположенные перед консервативными экзонами тяжелой цепи антител) [175]. В этих областях формирование R-петель приводит к длительному «выпетливанию» одноцепочечной ДНК и высокой плотности AID-индуцированных остатков урацила в обеих цепях ДНК. В результате репарации U/G белками BER и MMR происходит накопление одноцепочечных разрывов в обеих цепях ДНК, которые конвертируются в двуцепочечные разрывы. Последующее действие белков системы репарации двуцепочечных разрывов, преимущественно пути NHEJ, обеспечивает перестройку консервативных участков тяжелой цепи иммуноглобулинов [176].

Таким образом, системы эксцизионной репарации ДНК у человека играют ключевую роль в обеспечении геномной стабильности, предотвращая накопление мутаций и развитие злокачественных трансформаций. Однако в отдельных случаях, например, в клетках зародышевой линии и в процессе созревания В-лимфоцитов, элементы этих систем также вовлечены в формирование генетического разнообразия – через участие в мейотической рекомбинации и соматической гипермутации иммуноглобулинов. У человека белки систем репарации задействованы в процессах, поддерживающих генетическое разнообразие в специализированных клеточных типах и в строго ограниченных условиях. В отличие от эукариот в клетках прокариот функции репарации и поддержания разнообразия менее разделены (например, MMR *N. gonorrhoeae* может одновременно участвовать в контроле ошибок репликации и регуляции антигенной вариации).

#### **П.3.4. Влияние GC-состава ДНК на эффективность репарации ДНК**

К генетической нестабильности могут приводить не только мутации в генах, кодирующих белки систем репарации ДНК. Существенное влияние на эффективность этих систем оказывают и другие молекулярные факторы, среди которых особое значение имеет нуклеотидный состав ДНК. GC-богатые участки генома, широко распространенные в регуляторных областях, представляют собой уязвимые зоны, в которых высокая термодинамическая стабильность двойной спирали, обусловленная тремя водородными связями между гуанином и цитозином, затрудняет нормальное течение репликации, транскрипции и репарации. Эти участки способны служить барьерами для репликационной вилки и транскрипционных комплексов, что повышает риск возникновения одно- и двуцепочечных разрывов, а также требует дополнительного участия белков репарации.

GC-области особенно чувствительны к окислительному стрессу: G легко окисляется до 8-охоG, устраняемого системой BER с участием фермента OGG1 [177]. Остатки C в составе C-трактов склонны к дезаминированию, что может приводить к формированию «мисматчей», ликвидируемых системой MMR [178]. Повреждения и структурные затруднения, возникающие в GC-богатых регионах, активируют дополнительные репарационные каскады, что увеличивает нагрузку на клетку и может способствовать хромосомной нестабильности и канцерогенезу. Известно, что нуклеотидное окружение повреждения, которое должно быть распознано и удалено специфической системой репарации ДНК, может влиять на эффективность этого процесса.

Так, эффективность OGG1 значительно снижается, если 8-охоG локализуется в GC-богатых областях генома [179]. Причиной тому считают усиленное межплоскостное  $\pi$ -взаимодействие гетероциклов, которое препятствует «выворачиванию» (flipping out) окисленного основания, что необходимо для узнавания и удаления модификации ферментом [180].

Для другого участника BER - APE1 влияние GC-богатого окружения AP-сайта четко не установлено, хотя показано, что в случае формирования G-богатой последовательностью стабильной G4-структуры, эффективность удаления повреждения APE1 изменяется [181]. Было также установлено, что полиаденозиновые и политимидиновые олигонуклеотиды, содержащие AP-сайт, чрезвычайно плохо расщепляются AP-эндонуклеазой [182].

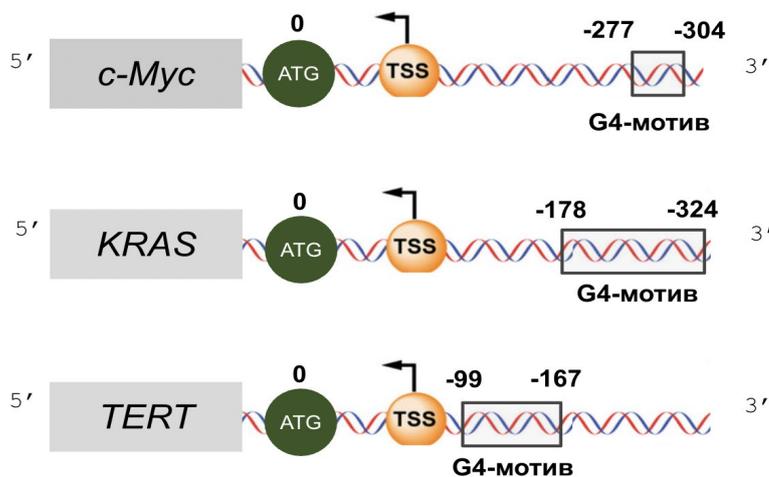
Эффективность белков MMR в случае, когда «мисматч» расположен в GC-богатом окружении ранее не изучали. Однако в литературе есть данные о зависимости АТФазной активности гомолога MutSa человека от ближайшего окружения «мисматча», что также может снижать или повышать эффективность репарации [183]. В большинстве случаев АТФазная активность hMSH2-hMSH6 повышается при фланкировании некомплементарной пары 3'-пуриновыми нуклеотидами и снижается при фланкировании 3'-пиримидиновыми, что особенно может быть заметно в GC-богатых регионах генома.

Как уже упоминалось, некоторые GC-богатые последовательности при определенных условиях могут формироваться вторичные структуры, такие как G-квадруплексы, оказывающие значительное влияние на процессы репликации, транскрипции и репарации ДНК. И если влияние GC-богатого окружения повреждения в ДНК на функции белков систем репарации изучено мало, то взаимодействие этих белков с G4 является предметом интереса многих исследователей.

### **II.3.5. Взаимодействие белков эксцизионной репарации ДНК человека с G-квадруплексами**

В геноме человека плотность на мегабазу GC-богатых последовательностей, вероятно

способных в определенных условиях к формированию G4, на порядок выше, чем у прокариот. Показано их наличие в промоторах онкогенов *c-MYC*, *KRAS*, *TERT* (Рис. 10) и в теломерах [2]. Эти структуры играют важную роль в регуляции транскрипции и репликации, однако могут создавать препятствия для работы систем репарации ДНК.



**Рисунок 10.** Схематичное изображение GC-богатых промоторных регионов генов *c-MYC*, *KRAS*, *TERT*. TSS - точка начала транскрипции, ATG - стартовый кодон. Выделены G4-мотивы, указано положение G4-мотивов относительно стартового кодона.

В последние годы накопились данные о взаимодействии ключевых белков репарации BER и MMR с G4-структурами, которые имеют важное значение для понимания механизмов поддержания геномной стабильности, а также свидетельствуют в пользу дополнительных функций белков систем репарации вблизи G4.

Исследования показали, что OGG1 не способна эффективно удалять 8-охоG из G-тетрад теломерных G4-структур как с антипараллельной (стабилизированных  $\text{Na}^+$ ), так и с гибридной или параллельной топологией (в присутствии  $\text{K}^+$ ). Хотя OGG1 демонстрирует сопоставимое сродство к G4- и дуплексной ДНК, содержащим 8-охоG, каталитическая активность фермента в случае G4 практически отсутствует. Это связано с невозможностью ключевых конформационных изменений: в отличие от репарации 8-охоG в дуплексе (где происходит выворачивание поврежденного основания и изгиб ДНК), при взаимодействии с G4 таких структурных перестроек не наблюдается [184]. Интересно, что прокариотический аналог OGG1 – гликозилаза Fpg – сохраняет ограниченную способность удалять 8-охоG из G4, хотя эффективность реакции снижается более чем на 40% по сравнению с дуплексом. Эти данные свидетельствуют о том, что жесткая архитектура G4, особенно в промоторных областях и теломерах, может служить барьером для репарации 8-охоG, способствуя накоплению мутаций в G4-богатых участках генома. В работе [185] показано, что 8-охоG накапливается в G4-мотиве *KRAS* при окислительном стрессе: OGG1 связывается с этой областью, но неэффективно удаляет повреждение из G4-структуры.

Для изучения эндонуклеазной активности APE1 человека в G4-структурах обычно используют олигодезоксирибонуклеотиды, содержащие тетрагидрофуран (1,2- дидезоксирибоза, стабильный аналог AP-сайта). Присутствие тетрагидрофурана в различных позициях G4 часто приводит к снижению температуры плавления квадруплекса. Показано, что APE1 гидролизует AP-сайт в составе G4 теломерных повторов в 200 раз медленнее, чем в двуцепочечной ДНК, несмотря на сопоставимую эффективность связывания фермента с обоими типами субстратов [186]. APE1 взаимодействует с таким G4 через остатки лизина (K27/31/32/35) в неструктурированном N-концевом домене [187]. Ацетилирование этих лизинов снижает положительный заряд белка и подавляет его эндонуклеазную активность. Кристаллографические данные демонстрируют, что APE1 узнает AP-сайт в дуплексе не только через контакты с поврежденной цепью (включая фланкирующие нуклеотиды), но и с комплементарной цепью, что требует значительного изгиба ДНК для позиционирования повреждения в каталитическом центре фермента. Эти структурные требования не выполняются в случае G4, что объясняет низкую активность APE1 в отношении квадруплексов.

Несмотря на низкую эффективность, APE1 гидролизует G4 с тетрагидрофураном как в петле, так и в G-тетраде, причем в последнем случае активность фермента в 5 раз выше [188]. Кинетический анализ показал, что лимитирующей стадией процесса является образование каталитически активного комплекса после первичного связывания APE1 с G4. Эффективность расщепления AP-сайта также зависит от топологии G4: активность APE1 выше в параллельных теломерных G4, чем в гибридных; антипараллельные ( $\text{Na}^+$ -стабилизированные) теломерные G4 расщепляются эффективнее, чем  $\text{K}^+$ -координированные гибридные [189].

Современные исследования выявили, что 8-охоG в составе G4-мотивов, вызванные окислительным повреждением гуаниновых остатков, могут выполнять функцию эпигенетических маркеров. Наиболее изученным примером является промоторная область гена *VEGF*, где образование 8-охоG в некодирующей цепи усиливает экспрессию репортерного гена (находящегося под контролем этого промотора) через BER-зависимый механизм. G4-мотив промотора *VEGF* имеет пятый G-тракт, который выступает в роли запасного при повреждении одного из основных четырех G-трактов. Это способствует «выпетливанию» G-тракта с окислительным повреждением, тем самым делая его доступным для репарации системой BER. Ключевыми участниками процесса выступают белки OGG1 и APE1: предположительно, удаление 8-охоG ферментом OGG1 и образование AP-сайта приводят к дестабилизации дуплекса и формированию G4-структуры за счет «запасного» G-тракта и «выпетливания» поврежденной области. Связывание APE1 с AP-сайтом способствует привлечению транскрипционных факторов (HIF-1 $\alpha$ , AP-1) [2].

Многочисленные исследования, в том числе наши, продемонстрировали способность

гомологов MutS (в том числе и MutS $\alpha$  человека) из системы MMR специфически связываться с G-квадруплексными структурами [190–192]. Интересно, что сродство этих белков к G4-ДНК обычно в несколько раз превышает их сродство к «природному» субстрату - ДНК с G/T-парой. В отличие от канонического узнавания «мисматчей» в дуплексах, где критическую роль играет мотив Phe-X-Glu (с участием фенилаланина в межплоскостных взаимодействиях и глутамата в образовании водородных связей [193]), белок способен связываться с G4 даже при мутации Phe36Ala, хотя и с 20%-ным снижением сродства. Взаимодействие MutS с G4 практически не влияет на АТФазную активность белка, в то время как для узнавания G/T-пар эффективный гидролиз АТФ критически важен [191]. Эти данные свидетельствуют о принципиально ином механизме узнавания G4 белками MMR по сравнению с «мисматчами». Предположительно, связывание G4-MutS не является сигналом к активации MMR, однако может выполнять некую альтернативную роль. В работе [194] выявили, что MutS $\beta$  (MSH2-MSH3) и MutL $\beta$  (MLH1-PMS1), специфически связываясь с G4, образуют функциональный комплекс с хеликазой FANCI, разрушающей G4. Этот механизм играет ключевую роль в устранении транскрипционно-репликационных конфликтов (TRCs), где G4-структуры, стабилизированные в R-петлях, блокируют репликацию. Важно, что взаимодействие белков MMR с G4 не активирует классический путь репарации «мисматчей», а направлено на поддержание репликационной процессивности через разрушение G4-структур, что подчеркивает их альтернативную функцию в защите генома.

Тем не менее, системное изучение взаимодействия гомологов MutL с G4-образующими последовательностями ДНК практически отсутствуют, что делает данное направление исследований актуальным.

Описанные наблюдения расширяют наши представления о многофункциональности систем BER и MMR, предполагая их участие в процессах, выходящих за рамки классической репарации повреждений ДНК. Однако точные молекулярные механизмы и физиологическое значение такого взаимодействия требуют дальнейшего изучения.

\*\*\*

Анализ данных литературы показывает, что эксцизионная репарация ДНК, включая системы BER, NER и MMR, является универсальным и высококонсервативным механизмом поддержания геномной стабильности у прокариот и эукариот. В то же время, накопленные экспериментальные данные свидетельствуют о том, что эффективность функционирования этих систем существенно зависит от локального нуклеотидного контекста и пространственной организации ДНК. Особое внимание в последние годы привлекают GC-богатые последовательности и формируемые ими G4-структуры, обладающие высокой термодинамической стабильностью и способностью к образованию различных топологических

состояний. Эти структуры широко представлены в регуляторных областях генов, включая промоторы, теломерные участки и сайты репликационного старта, и вовлечены в ключевые клеточные процессы – транскрипцию, репликацию и рекомбинацию.

Ряд исследований показал, что G4 способны снижать эффективность узнавания и удаления поврежденных нуклеотидов, включая 8-охоG и неканонические пары нуклеотидов, за счет ограничения доступа ферментов к повреждению, стабилизации дефектного нуклеотида в составе тетрад и конкуренции с другими белковыми факторами, взаимодействующими с G4. Эти эффекты зафиксированы как для ферментов BER (OGG1, APE1), так и для ключевых белков MMR (MutS $\alpha$ , MutS $\beta$ , MutL $\alpha$ ). Несмотря на то, что для ряда белков репарации установлено их прямое связывание с G4-мотивами и снижение каталитической активности в таких участках, молекулярные механизмы этих взаимодействий остаются недостаточно изученными, а данные часто фрагментарны и зависят от конкретной модели и типа исследованной G4-структуры.

Недостаточная эффективность BER и MMR в GC-богатых и G4-содержащих участках ДНК может способствовать накоплению мутаций, формированию нестабильных геномных зон и, как следствие, играть значимую роль как в адаптационных процессах бактерий, так и в развитии различных патологических состояний у человека, включая опухолевую трансформацию. В связи с этим актуальным является детальное исследование функционала белков BER и MMR в контексте

GC-богатых последовательностей и G4-структур, с целью выявления специфических молекулярных особенностей их взаимодействия с неканоническими элементами ДНК и определения факторов, лимитирующих эффективность репарации в таких участках. Полученные результаты могут не только расширить фундаментальные представления о работе систем репарации, но и предложить новые подходы к направленной модуляции этих процессов.

### III. Материалы и методы

**Реактивы:** ДМСО, АДФ, АТФ, ампициллин, трео-2,3-дигидрокси-1,4- димеркаптобутан (дитиотреит, ДТТ), додецилсульфат натрия (ДСН), Трис (трис- (гидроксиметил)аминометан), N-2-гидроксиэтилпиперазин-N'-2-этансульфоновая кислота (HEPES), ЭДТА, красители бромфеноловый синий (БФС), ксиленцианол (КЦ) («Reanal», Венгрия), кумасси бриллиантовый синий G-250. Глицин («Хеликон», Россия). Хлориды магния, калия, кальция, марганца, цинка, кобальта, никеля. Персульфат аммония (ПСА), бромид этидия (Sigma-Aldrich, США), N,N,N',N'-тетраметилэтилендиамин (TEMED), акриламид, N,N'-метиленабисакриламид, изопропил-β-D-тиогалактопиранозид (ИПТГ), триптон, дрожжевой экстракт, агар, агароза, гидрофосфат натрия, β-меркаптоэтанол, имидазол, фенилметансульфонилфторид (PMSF), этанол, ацетат натрия, формаид, молибдат аммония, цитрат натрия, соляная кислота, тетраметил-4,4-диаминотрифенилметан (малахитовый зеленый), уксусная кислота, Ni-NTA (никель-нитрилотриуксусная кислота)-агароза (Novagen, Германия).

**Ферменты и белки:** Бычий сывороточный альбумин (БСА) (Thermo Fisher Scientific, США), протеиназа К (~600–1000 ед. акт./мкл), наборы белков–маркеров молекулярной массы PageRuler™ (10–200 кДа), ДНК-маркеры O'GeneRuler™ 1 kb, 250-10000 п.н., ДНК-маркеры O'GeneRuler™ 20-200 п.н. протеиназа К (~600–1000 ед. акт./мкл), наборы белков–маркеров молекулярной массы PageRuler™ (10–200 кДа), апурин-апиримидиновая (AP) эндонуклеаза 1 (APE1) (New England Biolabs, са# M0282S), APE1-6His любезно предоставленная Жарковым Д.О. (Институт химической биологии и фундаментальной медицины СО РАН, г. Новосибирск).

#### **Буферные растворы:**

ТАЕ: 40 mM Трис-CH<sub>3</sub>COOH (pH 7,5), 1 mM ЭДТА;

ТВЕ: 50 mM Трис-HCl, 50 mM борная кислота, 1 mM ЭДТА (pH 8,3);

ТГ: 25 mM Трис-HCl, 250 mM глицин (pH 8,3), 0,1% (m/V) ДСН;

STE: 10 mM Трис-HCl (pH 8,0), 100 mM NaCl, 100 мкМ ЭДТА (для промывания осадков клеток);

LAP: 50 mM Трис-HCl (pH 6,8), 2,5% (m/V) ДСН, 10% (V/V) глицерина, 0,5% (V/V) β-меркаптоэтанола, 0,01% (m/V) БФС;

А: 10 mM HEPES-КОН (pH 8,0), 300 mM KCl, 1 mM β-меркаптоэтанол, 5 mM имидазол, 10% (V/V) глицерина (для связывания MutL с Ni-NTA-агарозой);

Б: 10 mM HEPES-КОН (pH 8,0), 300 mM KCl, 1 mM β-меркаптоэтанол, 60 mM имидазол, 10% (V/V) глицерина (для промывки Ni-NTA-агарозы от неспецифически связанных белков);

В: 10 mM HEPES-КОН (pH 8,0), 300 mM KCl, 1 mM β-меркаптоэтанол, 300 mM имидазол, 10% (V/V) глицерина (для элюирования MutL с Ni-NTA-агарозы);

Г: 10 мМ HEPES-КОН (рН 8,0), 300 мМ КСl, 1 мМ β-меркаптоэтанол, 50% (V/V) глицерина (для хранения MutL при -20°C);

Д: 10 мМ HEPES-КОН (рН 8,0), 300 мМ КСl, 1 мМ β-меркаптоэтанол, 10% (V/V) глицерина (для хранения MutL при -80°C);

PBS: 10 мМ Na<sub>2</sub>HPO<sub>4</sub>, 1,76 мМ K<sub>2</sub>HPO<sub>4</sub>, 2,7 мМ КСl, 137 мМ NaCl (рН 7,4).

***Питательные среды:***

LB: 10 г триптона, 5 г NaCl, 5 г дрожжевого экстракта (рН 7,2) (на 1 л воды);

LB-агар: 10 г триптона, 5 г NaCl, 5 г дрожжевого экстракта (рН 7,2), 15 г агара (на 1 л воды).

Среду автоклавировали в автоклаве Tuttnauer 2540 МК (США), после охлаждения добавляли необходимое количество антибиотика.

***Клетки:*** B121(DE3)pLysS, XL1 Blue (Novagen, Германия).

***Плазмиды:***

Плазмидная конструкция рЕТ15b, которая содержит ген белка репрессора lac-оперона, промотор и терминатор T7 РНК-полимеразы, ген устойчивости к ампициллину, и гены, кодирующие белки ngMutL<sup>wt</sup> или ngMutL<sup>cf</sup> (содержащий замены С171I, С251L, С528А, С600S и С631S), которые содержат нуклеотидную последовательность с шестью остатками His в N-концевой области белков, были любезно предоставлены Rao D.N. (Индийский научный институт, г. Бангалор). Плазида рUC-MMR для оценки эффективности гидролиза ДНК белками MutL<sup>wt</sup> и MutL<sup>cf</sup> любезно предоставлена проф. П. Фридрихсом (Институт биохимии Университета имени Ю. Либиха, г. Гиссен, Германия).

***Мини-колонки и микро-колонки для гель-фильтрации:*** NAP-5, Ultra MicroSpin G-50 (GE Healthcare Life Sciences, США).

***Колонки для аффинной хроматографии:*** HisTrap™ HP (GE Healthcare, США).

***Сорбенты для аффинной хроматографии:*** Ni<sup>2+</sup>-NTA-агароза (Qiagen, Германия).

**Олигонуклеотиды:** олигонуклеотиды, используемые для приготовления ДНК-дуплексов, синтезированы компанией СИНТОЛ (Россия) амидофосфитным твердофазным методом. В таблицах 5-9 представлены последовательности ДНК, используемых в экспериментах для выявления особенностей взаимодействия белков систем репарации ДНК с GC-богатыми фрагментами регуляторных областей генома эукариот и прокариот. ДНК содержали флуоресцентную метку тетраметилродамин (TAMRA) на 5'- или 3'-конце.

**Таблица 5.** Олигонуклеотидные последовательности 76-звенных контрольных ДНК.

ДНК дуплекс	Последовательность (5'-3')
76R/76M	5'-ATAGGACGCTGACACTGGTGCTTGGCAGCTGAGCCATATGCTCGAGTAACGCTCATAGGATCCAAGCGCGAAAGGA-3' TAMRA 3'-TATCCTGCGACTGTGACCACGAACCGTCTGACTCGGTATACGAGTCATTGCGAGTATCCTAGGTTGCGGCTTTCCT-5

**Таблица 6.** Обозначения и последовательность олигонуклеотидов, используемых в экспериментах по установлению влияния G4 на функции MutL.

ДНК	Последовательность (5'-3')
19pilG4	AGGGTGGGTGGGTGGGGA-TAMRA
19T	AGTGTGTGTTGTGTGTGGA-TAMRA
19M_A	TCCACACACAACACACACT
41pilG4	ACGCGTTAGAATAGGGTGGGTGGGTGGGGAATTTTCTATT-TAMRA
41M	AATAGAAAATTCCCCACCCAACCCACCCATTCTAACGCGT
95G4	ATAGGACGCTGACACTGGTGCTTGGCAGCTGAGCCATATTTGGGTGGGTGGGTGGGTGCTCGAGTAACGCTCATAGGATCCAAGCGCGAAAGGA-TAMRA
95pilG4	GTCGGAATTTGAGATTTTGAATTTACGCGTTAGAATAGGGTGGGTGGGTGGGGAATTTTCTATTTTAAAAAGCTCCGTTTCTTGGAAGC-TAMRA
95M	GCTTTCCAAGAAAACGGAGCTTTTAAAAAATAGAAAATTCCCCACCCAACCCACCCATTCTAACGCGTAAATTCAAAAATCTCAAATTCGAC

**Таблица 7.** Обозначения и последовательность олигонуклеотидов, используемых в экспериментах по установлению влияния GC-богатого состава на эффективность узнавания «мисматча» белком MutS.

ДНК	Последовательность (5'-3')
ssR (53%GC)	5'-CAGGCTTAGGCTGGAGCTGAAGCTTAGCTTAGGATCATCGAGGATC <sub>c</sub> AGTCCAGTGC <sub>A</sub> ATTCAGCGGTACCCAATTCGCCCTATAGGCATCCAGGT-TAMRA-3'
ssR (53%GC C→T)	5'-CAGGCTTAGGCTGGAGCTGAAGCTTAGCTTAGGATCATCGAGGATCTAGTCCAGTGC <sub>A</sub> ATTCAGCGGTACCCAATTCGCCCTATAGGCATCCAGGT-TAMRA-3'
ssR (82%GC)	5'-CTCCTCCGCGCGCCGCCAGCCGCCGCTCGCCACCTCCGCGCGCTCCGCTCTCCTCGCCACGCTCCTCCGCGCGCTCCACGGCCTTCCGCC-TAMRA-3'
ssR (82%GC C→T)	5'-CTCCTCCGCGCGCCGCCAGCCGCCGCTCGCCACCTCCGCGCGCTCCGCTCTCCTCGCCACGCTCCTCCGCGCGCTCCACGGCCTTCCGCC-TAMRA-3'

При синтезе фрагментов ДНК с F-сайтом (стабильным циклическим аналогом AP-сайта, 1,2-дидезоксирибозой) использовали коммерческий амидофосфит dSpacer CE Phosphoramidite (Glen Research, США, са# 129821-76-7).

**Таблица 8.** Используемые в работе олигонуклеотиды. Замена G→F отмечена красным цветом.

ДНК	Последовательность (5'-3')
96-G4	5'- GCCGCGAAAGGAAGGGAGGGCTGGGAGGGCCCGAGGGGGCTGGGCCGGGGACCCGGGAGGGGTGGGACGGGGCGGGGTCCGCGCGGAGGAG-3'-TAMRA
96-RF	5'- ACCTGGATGCCCTATAGGGCGAATTGGGTACCGCTGAATGCACTGGACT <b>F</b> GATCCTCGATGATCCTAAGCTAAGCTTCAGCTCCAGCCTAAGCCTG-3'-TAMRA
96-G124F	5'- GCCGCGAAAGGAAGGGAGGGCTGGGAGGGCCCGA <b>F</b> GGGGCTGGGCCGGGGACCCGGGAGGGGTGGGACGGGGCGGGGTCCGCGCGGAGGAG-3'-TAMRA
96-G138F	5'- GCCGCGAAAGGAAGGGAGGGCTGGGAGGGCCCGAGGGGGCTGGGCC <b>F</b> GGACCCGGGAGGGGTGGGACGGGGCGGGGTCCGCGCGGAGGAG-3'-TAMRA
96-G139F	5'- GCCGCGAAAGGAAGGGAGGGCTGGGAGGGCCCGAGGGGGCTGGGCC <b>F</b> GACCCGGGAGGGGTGGGACGGGGCGGGGTCCGCGCGGAGGAG-3'-TAMRA
96-G138/139F	5'- GCCGCGAAAGGAAGGGAGGGCTGGGAGGGCCCGAGGGGGCTGGGCC <b>F</b> FACCCGGGAGGGGTGGGACGGGGCGGGGTCCGCGCGGAGGAG-3'-TAMRA
96-G146F	5'- GCCGCGAAAGGAAGGGAGGGCTGGGAGGGCCCGAGGGGGCTGGGCCGGGGACCC <b>F</b> AGGGGTGGGACGGGGCGGGGTCCGCGCGGAGGAG-3'-TAMRA
30-RF	5'- TGAATTGCACTGGACT <b>F</b> GATCCTCGATGAT-3'-TAMRA
30-G124F	5'- CCGGA <b>F</b> GGGGCTGGGCCGGGGACCCGGGAG-3'-TAMRA
30-G138F	5'- CCGGAGGGGGCTGGGCC <b>F</b> GACCCGGGAG-3'-TAMRA
30-G139F	5'- CCGGAGGGGGCTGGGCC <b>F</b> ACCCGGGAG-3'-TAMRA
30-G138/139F	5'- CCGGAGGGGGCTGGGCC <b>F</b> FACCCGGGAG-3'-TAMRA
30-G146F	5'- CCGGAGGGGGCTGGGCCGGGGACCC <b>F</b> AG-3'-TAMRA

Таблица 9. Используемые в работе ДНК-дууплексы. Замена G→F или C→F отмечена красным цветом.

ДНК-дууплекс	Последовательность
96-RF/96C	5' - ACCTGGATGCCTATAGGGCGAATTGGGTACCGCTGAATTGCACTGGACT <b>F</b> GATCCTCGATGATCCTAAGCTAAGCTTCAGCTCCAGCCTAAGCCTG-3' - TAMRA 3' - TGGACCTACCGATATCCCGCTTAACCCATGGCGACTTAACGTGACCTGACCTAGGAGCTACTAGGATTCGATTCTGAAGTCGAGGTCGGATTCGGAC-5'
96-G124F/96C	5' - GCCGCGGAAAGGAAGGGGAGGGGCTGGGAGGGCCCGGA <b>F</b> GGGGCTGGGCCGGGGACCCGGGAGGGGTCGGGACGGGGCGGGTCCGCGCGGAGGAG-3' - TAMRA 3' - CGGCGCCTTTCCTTCCCCTCCCGACCCCTCCCGGGCCTCCCGACCCGGCCCTGGGCCCTCCCGAGCCCTGCCCGCCCCAGGCGCGCCTCCTC-5'
96-G138F/96C	5' - GCCGCGGAAAGGAAGGGGAGGGGCTGGGAGGGCCCGAGGGGCTGGGCCGG <b>F</b> GACCCGGGAGGGGTCGGGACGGGGCGGGTCCGCGCGGAGGAG-3' - TAMRA 3' - CGGCGCCTTTCCTTCCCCTCCCGACCCCTCCCGGGCCTCCCGACCCGGCCCTGGGCCCTCCCGAGCCCTGCCCGCCCCAGGCGCGCCTCCTC-5'
96-G139F/96C	5' - GCCGCGGAAAGGAAGGGGAGGGGCTGGGAGGGCCCGAGGGGCTGGGCCGGG <b>F</b> ACCCGGGAGGGGTCGGGACGGGGCGGGTCCGCGCGGAGGAG-3' - TAMRA 3' - CGGCGCCTTTCCTTCCCCTCCCGACCCCTCCCGGGCCTCCCGACCCGGCCCTGGGCCCTCCCGAGCCCTGCCCGCCCCAGGCGCGCCTCCTC-5'
96-G138/139F/96C	5' - GCCGCGGAAAGGAAGGGGAGGGGCTGGGAGGGCCCGAGGGGCTGGGCCGG <b>FF</b> ACCCGGGAGGGGTCGGGACGGGGCGGGTCCGCGCGGAGGAG-3' - TAMRA 3' - CGGCGCCTTTCCTTCCCCTCCCGACCCCTCCCGGGCCTCCCGACCCGGCCCTGGGCCCTCCCGAGCCCTGCCCGCCCCAGGCGCGCCTCCTC-5'
96-G146F/96C	5' - GCCGCGGAAAGGAAGGGGAGGGGCTGGGAGGGCCCGAGGGGCTGGGCCGGGGACCCGG <b>F</b> AGGGGTCGGGACGGGGCGGGTCCGCGCGGAGGAG-3' - TAMRA 3' - CGGCGCCTTTCCTTCCCCTCCCGACCCCTCCCGGGCCTCCCGACCCGGCCCTGGGCCCTCCCGAGCCCTGCCCGCCCCAGGCGCGCCTCCTC-5'
96-C124F/96G	5' - CTCCTCCGCGCGGACCCCGCCCGTCCCGACCCCTCCCGGGTCCCGGCCAGCCCC <b>F</b> TCCGGGCCCTCCAGCCCTCCCTTTCCTTCCGCGGC-3' - TAMRA 3' - GAGGAGGCGCGCCTGGGGCGGGGACGGGCTGGGGAGGGCCAGGGGCCGGGTTCGGGGAGGCCCGGGAGGGTCGGGGAGGGGAAGGAAAGGCGCCG-5'
96-C138F/96G	5' - CTCCTCCGCGCGGACCCCGCCCGTCCCGACCCCTCCCGGGT <b>F</b> CCGGGCCAGCCCTCCGGGCCCTCCAGCCCTCCCTTTCCTTCCGCGGC-3' - TAMRA 3' - GAGGAGGCGCGCCTGGGGCGGGGACGGGCTGGGGAGGGCCAGGGGCCGGGTTCGGGGAGGCCCGGGAGGGTCGGGGAGGGGAAGGAAAGGCGCCG-5'
96-C139F/96G	5' - CTCCTCCGCGCGGACCCCGCCCGTCCCGACCCCTCCCGGGT <b>F</b> CCGGGCCAGCCCTCCGGGCCCTCCAGCCCTCCCTTTCCTTCCGCGGC-3' - TAMRA 3' - GAGGAGGCGCGCCTGGGGCGGGGACGGGCTGGGGAGGGCCAGGGGCCGGGTTCGGGGAGGCCCGGGAGGGTCGGGGAGGGGAAGGAAAGGCGCCG-5'
96-C138/139F/96G	5' - CTCCTCCGCGCGGACCCCGCCCGTCCCGACCCCTCCCGGGT <b>FF</b> CCGGGCCAGCCCTCCGGGCCCTCCAGCCCTCCCTTTCCTTCCGCGGC-3' - TAMRA 3' - GAGGAGGCGCGCCTGGGGCGGGGACGGGCTGGGGAGGGCCAGGGGCCGGGTTCGGGGAGGCCCGGGAGGGTCGGGGAGGGGAAGGAAAGGCGCCG-5'
96-C146F/96G	5' - CTCCTCCGCGCGGACCCCGCCCGTCCCGACCCCT <b>F</b> CCGGGTCCCCGGCCAGCCCTCCGGGCCCTCCAGCCCTCCCTTTCCTTCCGCGGC-3' - TAMRA 3' - GAGGAGGCGCGCCTGGGGCGGGGACGGGCTGGGGAGGGCCAGGGGCCGGGTTCGGGGAGGCCCGGGAGGGTCGGGGAGGGGAAGGAAAGGCGCCG-5'

Концентрацию водных растворов олигонуклеотидов определяли на спектрофотометре NanoDrop ND-1000 (PepLab, США) по закону Бугера-Ламберта-Бера:  $A = \varepsilon \cdot l \cdot c$ , где  $c$  – концентрация, М;  $A$  – оптическая плотность;  $\varepsilon$  – молярный коэффициент экстинкции,  $\text{см}^{-1} \cdot \text{М}^{-1}$ ;  $l$  – длина оптического пути, см. Молярные коэффициенты экстинкции ( $\varepsilon_{260}$ ) олигонуклеотидов рассчитывали, используя интернет-ресурс (<https://eu.idtdna.com>). Растворы ДНК-дууплексов готовили смешиванием в воде эквимольного количества соответствующих олигонуклеотидов (с 5% избытков немеченной цепи) с последующим нагреванием до 90°C и медленным охлаждением в течение нескольких часов.

**Выделение плазмидных ДНК.** Клетки *E. coli* штамма XL1 Blue с плазмидой pUC-MMR инокулировали в 25 мл среды LB, содержащей ампициллин (150 мкг/мл), и культивировали 16–18 ч при 37 °С (инкубатор Innova® 40, New Brunswick, Германия). Из полученной культуры отбирали 10 мл, осаждали клетки центрифугированием и ресуспендировали в 500 мл свежей среды LB с антибиотиком. Культивирование продолжали при интенсивной аэрации до достижения оптической плотности  $A_{600} = 1,2–1,5$  О.Е. (примерно 8–9 ч); через 4 ч после начала выращивания дополнительно вносили ампициллин до конечной концентрации 50 мкг/мл.

Клеточную массу собирали центрифугированием (4500 об/мин, 30 мин, 4 °С, центрифуга J6-НС, Beckmann, США). Осадки промывали буфером STE и повторно центрифугировали в тех же условиях. Полученную биомассу хранили при –20 °С. Плазмидную ДНК выделяли с помощью набора GeneJET Plasmid Miniprep Kit (ThermoFisher Scientific, США). Концентрацию ДНК определяли спектрофотометрически на приборе NanoDrop ND-1000 (PecLab, США), используя закон Бугера–Ламберта–Бера.

**Трансформация компетентных клеток *E. coli* плазмидой, содержащей ген белков MutL<sup>wt</sup> и MutL<sup>cf</sup>** (метод «теплового шока»). К 100 мкл компетентных клеток *E. coli* (штамм XL1 Blue), размороженных на льду, добавляли 100 нг плазмидной ДНК (1 мкл раствора). Смесь инкубировали на льду в течение 30 мин, затем подвергали тепловому шоку при 42 °С (1 мин) и снова переносили в лёд на 2 мин. Далее вносили 800 мкл среды LB и инкубировали клетки 1 ч при 37 °С для восстановления. Клетки осаждали центрифугированием (4200 об/мин, 5 мин, 25 °С) и высевали на чашки Петри с агаризованной средой LB, содержащей 150 мкг/мл ампициллина. Культивирование вели 16–18 ч при 37 °С.

Для длительного хранения трансформированных клеток готовили глицериновые стоки. Для этого единичную колонию с чашки засевали в 5 мл жидкой среды LB с ампициллином (150 мкг/мл) и выращивали 16–18 ч при 37 °С. Полученную биомассу ресуспендировали в стерильном 50% водном растворе глицерина, замораживали в жидком азоте и хранили при –80 °С.

**Выращивание трансформированных клеток и индукция биосинтеза белков MutL<sup>wt</sup> и MutL<sup>cf</sup>.** Для биосинтеза белков MutL использовали клетки *E. coli* (штамм BL21(DE3)pLysS). Колонию трансформированных клеток (или биомассу из «глицеринового стока») переносили в 10 мл среды LB (содержащей 150 мкг/мл ампициллина) и выращивали в течение 16–18 ч при 37°С. На следующий день полученную культуру помещали в 2 л среды LB (150 мкг/мл ампициллина) и инкубировали при интенсивной аэрации и 37 °С до  $A_{600} = 0,8$  О.Е. Затем добавляли водный раствор ИПТГ до концентрации 1 мМ и инкубировали клетки 3 ч при 28 °С. После этого отделяли клетки от среды центрифугированием в центрифужных стаканах объемом 500 мл (30 мин, 4000 об./мин, 4 °С), ресуспендировали клеточный осадок в 35 мл буфера STE,

центрифугировали (в пробирке объемом 50 мл) 15 мин при 4200 об./мин и 4 °С, удаляли жидкость над осадком и замораживали биомассу при –18 °С.

**Анализ набора клеточных белков до и после индукции.** Отбирали раствор выращенных клеток с  $A_{600} = 0,3$  О.Е., центрифугировали 1 мин при 4000 об./мин и 25 °С, удаляли жидкость над осадком. Клетки ресуспендировали в 20 мкл буфера LAP, выдерживали 5 мин при 90 °С, затем помещали пробирки в баню со льдом. Полученные образцы анализировали методом электрофореза в 8%-ном полиакриламидном геле (ПААГ) с 0,1% (m/V) ДСН. Для визуализации зон белка гель окрашивали раствором PageBlue™ (Thermo Fisher Scientific, США) на основе кумасси бриллиантового синего G-250.

**Выделение белков MutL<sup>wt</sup> и MutL<sup>cf</sup>.** Замороженные клетки размораживали на льду в течение 30 мин, после чего добавляли буфер А и PMSF до конечной концентрации 1 мМ. Клетки ресуспендировали и разрушали ультразвуком (15 циклов по 30 с с интервалами 30 с) в сосуде, охлаждаемом льдом. В полученный лизат повторно вносили PMSF до 1 мМ и отделяли неразрушенные фрагменты центрифугированием (18 000 об/мин, 30 мин, 4 °С).

Для подготовки сорбента к 50 мл буфера А добавляли 500 мкл суспензии Ni-NTA-агарозы. Смесь инкубировали при медленном перемешивании 30 мин при 4 °С, затем центрифугировали (800 об/мин, 5 мин, 4 °С) и удаляли надосадочную жидкость. К осадку агарозы приливали клеточный лизат и инкубировали 1 ч при 4 °С с осторожным перемешиванием. Сорбент с иммобилизованным белком осаждали центрифугированием (800 об/мин, 5 мин, 4 °С) и удаляли супернатант. Осадок промывали 35 мл буфера Б, содержащего 1 мМ PMSF, при осторожном перемешивании с последующим центрифугированием в тех же условиях. Надосадочную жидкость отбирали, оставляя около 10 мл, осторожно ресуспендировали осадок в этом объеме и помещали (при 4 °С) суспензию в хроматографическую колонку объемом 15 мл. После оседания сорбента позволяли жидкости стечь, к сорбенту добавляли 500 мкл буфера В с 1 мМ PMSF, перемешивали пипетированием и оставляли на 5 мин, затем собирали элюат в пробирку объемом 1,5 мл. Элюцию буфером В повторяли 3 раза. Определяли концентрацию белка в растворе спектрофотометрически при длине волны 280 нм.

**Эксклюзионная хроматография.** Для проведения эксклюзионной хроматографии использовали колонку Enrich SEC 650. Колонку уравнивали буфером Д для выделения MutL, который затем служил элюентом. Объем наносимого на колонку препарата – 500 мкл. Расход элюента – 1 мл/мин. За ходом разделения следили по поглощению при 280. Концентрацию белка во фракциях определяли на приборе NanoDrop ND-1000 (PecqLab, США). Затем отбирали аликвоты по 5-10 мкл в отдельные пробирки и замораживали их в жидком азоте. Пробирки с аликвотами раствора белка хранили при –80 °С и размораживали в бане со льдом непосредственно перед использованием. Концентрацию препаратов белков определяли на

спектрофотометре NanoDrop ND-1000 (PerqLab, США) по закону Бугера-Ламберта-Бера:  $A = \epsilon \cdot l \cdot c$ , где  $c$  – концентрация, М;  $A$  – оптическая плотность;  $\epsilon$  – молярный коэффициент экстинкции,  $\text{см}^{-1} \cdot \text{М}^{-1}$ ;  $l$  – длина оптического пути, см. Для определения концентрации белков использовали теоретически рассчитанный молярный коэффициент экстинкции ( $\epsilon_{280}$ ) для ngMutL –  $19160 \text{ см}^{-1} \cdot \text{М}^{-1}$ . Указанные в дальнейшем концентрации ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup> рассчитаны на димер белка.

**Электрофорез по Лэмбли** использовался для анализа чистоты выделенных белков. Электрофорез проводили в электродном буфере ТГ в пластинах размером 10×8 см. Разделяющий гель (8%): 7,73% (m/V) акриламида, 0,27% (m/V) N,N'-метиленабисакриламида, 0,1% (m/V) ДСН, 375 мМ Трис-НСl (рН 8,8). Концентрирующий гель (4%): 3,87% (m/V) акриламида, 0,13% (m/V) N,N'-метиленабисакриламида, 0,1% (m/V) ДСН, 125 мМ Трис-НСl (рН 6,5). Электрофорез проводили при напряжении 200 В, силе тока 80 мА, напряженности поля 20 В/см в камере Mini-PROTEAN® TetraSystem (Bio Rad, США). Пробы наносили на гель в объеме 20 мкл в буфере LAR. Перед нанесением на гель пробы прогревали в течение 5 мин при 95 °С, затем помещали в баню со льдом и после центрифугировали 5 мин при 14500 об./мин. Для визуализации белковых зон использовали краситель кумасси G-250.

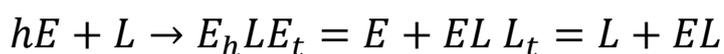
**Спектры кругового дихроизма белка** дикого типа и его мутантной формы регистрировали в диапазоне длин волн 200 – 250 нм при комнатной температуре. Исследуемые образцы находились в буфере PBS, содержащем 50 мМ КСl. Спектры кругового дихроизма регистрировали на спектрометре Chirascan (Applied Photophysics Ltd., Великобритания) в кювете с длиной оптического пути 0,02 см.

**Гидролиз плазмидной ДНК белками MutL<sup>wt</sup> и MutL<sup>cf</sup>.** Плазмиду pUC-MMR (концентрация 10 нМ) инкубировали с белками MutL<sup>wt</sup> и MutL<sup>cf</sup> (концентрация 200-1000 нМ) в буфере для реакции гидролиза, содержащем 100 мМ КСl, 5 мМ MgCl<sub>2</sub> или 5 мМ MnCl<sub>2</sub> в зависимости от эксперимента, при 37 °С в течение 1 часа. При проведении гидролиза в присутствии ZnCl<sub>2</sub> его концентрация варьировалась от 5 до 1000 нМ. Реакцию останавливали добавлением к смеси 50 мМ ЭДТА. Для деградации белков в реакционную смесь добавляли протеиназу К (10 ед. акт.), инкубировали 20 мин при 55 °С. Реакционные смеси анализировали методом электрофореза в 1%-ном агарозном геле, содержащем бромид этидия, при напряженности поля 10 В/см. Анализируемые образцы наносили на гель в специальном буфере (6X DNA Loading Dye, ThermoFisher Scientific, США). Визуализацию флуоресцентных зон в геле, соответствующих ДНК, и обсчет данных проводили на приборе ChemiDoc MP (Bio-Rad, США), используя компьютерную программу Image Lab.

**Гидролиз линейных ДНК-дуплексов белками MutL<sup>wt</sup> и MutL<sup>cf</sup>.** Линейные ДНК-дуплексы (концентрация 10 нМ) различной длины (в зависимости от эксперимента), содержащие метку TAMRA на одном из концов цепи ДНК, инкубировали с белками ng $\beta$  (до 0,75 мкМ) и MutL<sup>wt</sup> или

MutL<sup>cf</sup> (до 0,75 мкМ на димер) в буфере для реакции гидролиза, содержащем 1 мМ ДТТ, 0,5 мг/мл БСА, 10% глицерина (m/V), 100 мМ КСl, 5 мМ MgCl<sub>2</sub> или 5 мМ CaCl<sub>2</sub> и/или 5 мМ MnCl<sub>2</sub> (в зависимости от эксперимента), при 37 °С в течение 60 мин. При проведении гидролиза в присутствии АТФ его концентрацию варьировали от 0,5 до 5 мМ. Реакцию останавливали добавлением к смеси 50 мМ ЭДТА. Для деградации белков в реакционную смесь добавляли протеиназу К (10 ед. акт.), инкубировали 20 мин при 55 °С. Реакционные смеси анализировали методом электрофореза в 10%-ном ПААГ (18×22×0,15 см), содержащем 9,66% (m/V) акриламида и 0,34% (m/V) N,N'-метиленабисакриламида и 7 М мочевины, в денатурирующих условиях в ТВЕ-буфере в течение 4 ч при напряжении от 200 до 400 В. Фотографии гелей получали с помощью прибора Turboon FLA 9500 (GE Healthcare, Великобритания). Префорез и электрофорез проводили в течение 4 ч, используя напряжение от 200 до 400 В.

**Проверка ДНК-связывающей активности белков MutL<sup>wt</sup> и MutL<sup>cf</sup>.** Для проверки ДНК-связывающей активности белков MutL<sup>wt</sup> и MutL<sup>cf</sup> использовали ДНК (концентрация 10 нМ), содержащие на одной из цепей флуоресцентную метку TAMRA. Смеси объемом 20 мкл в буфере для реакции, содержащие 10 нМ ДНК, 1 мМ ДТТ, 100 мМ КСl, 0,5 мг/мл БСА инкубировали с белками MutL<sup>wt</sup> или MutL<sup>cf</sup> (до 750 нМ) в течение 10 мин во льду. Анализ формирования ДНК-белковых комплексов проводили методом гель-электрофореза в 4%-ом ПААГ (18×22×0,15 см), содержащем 5,81% (m/V) акриламида и 0,19% (m/V) N,N'-метиленабисакриламида, в неденатурирующих условиях в ТВЕ буфере при 4 °С (метод «торможения в геле») в течение 4 ч при напряжении 80 В. Префорез и электрофорез проводили в течение 4 ч при температуре 4 °С при силе тока 15 мА и напряжении 100 В (напряженность 4,5 В/см). Фотографии гелей получали с помощью прибора Turboon FLA 9500 (GE Healthcare, Великобритания). Параметры связывания рассчитывали в программе Origin 2018b (9.55) с применением схемы Хилла, где Y<sub>max</sub> – максимальная степень комплексообразования, K<sub>d</sub> – константа диссоциации комплексов MutL-ДНК, h – коэффициент Хилла.

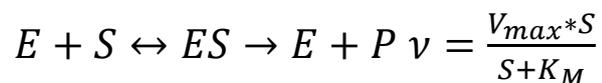


$$K_d = \frac{E * L}{EL} \quad Y = \frac{Y_{max} * E^h}{(E^h + K_d^h)}$$

**Приготовление реагента для количественного определения фосфат-иона.** 0,0812% малахитового зеленого, 2,32% поливинилового спирта, 5,72% молибдата аммония в 6 Н НСl смешивали с водой в соотношении 2:1:1:2.

**Определение АТФазной активности белков MutL<sup>wt</sup> и MutL<sup>cf</sup>.** Смеси объемом 60 мкл в буфере для реакции, содержащие 50 нМ ДНК-дуплекса I, 5 мМ MgCl<sub>2</sub>, 50 мМ КСl, различные

концентрации АТФ (100 - 1000 мкМ), инкубировали с белками **MutL<sup>wt</sup>** и **MutL<sup>cf</sup>** (0,8 мкМ) при 37 °С в течение 2,5 ч. Каждые 30 мин из реакционной смеси отбирали пробы по 10 мкл и останавливали реакцию добавлением 50 мМ ЭДТА. После этого добавляли 80 мкл реагента для идентификации фосфат-иона и тщательно перемешивали. Затем добавляли 10 мкл 34% раствора цитрата натрия и перемешивали до стабилизации окраски, выдерживали растворы примерно 15 мин, измеряли поглощение растворов при длине волны 590 нм на планшетном ридере Victor X5 (PerkinElmer, США). С помощью калибровочной прямой, которая представляла из себя зависимость поглощения раствора, содержащего известное количество фосфат-иона, от концентрации фосфат-иона, определяли количество фосфата в исследуемых растворах. Для получения линейных калибровочных зависимостей в раствор добавляли 5 мкМ «фонового» фосфат-иона. Кинетические параметры рассчитывали в программе Origin 2018b (9.55) с применением схемы Михаэлиса-Ментен, где  $V_{max}$  – максимальная скорость реакции,  $K_M$  – константа Михаэлиса.



$$V_{max} = E_0 * k_{cat} K_M = \frac{k_{-1} + k_2}{k_1}$$

**Проверка ДНК-связывающей активности белка MutS.** Для проверки ДНК-связывающей активности белка MutS использовали ДНК-дуплексы (концентрация 10 нМ), длиной 96 п.н., которые содержат на одной из цепей флуоресцентную метку TAMRA. Смеси объемом 20 мкл в реакционном, содержащие 10 нМ ДНК-дуплекса, 1 мМ ДТТ, 100 мМ KCl, 0,5 мг/мл БСА инкубировали с белком MutS (10-500 нМ) в течение 10 мин при 37 °С. Анализ формирования ДНК-белковых комплексов проводили методом гель-электрофореза в 6%-ом ПААГ в неденатурирующих условиях в TBE буфере при 4 °С (метод «торможения в геле») в течение 4 ч при напряжении 80 В. Фотографии гелей получали с помощью прибора Typhoon FLA 9500 (GE Healthcare, Великобритания). Для определения параметров связывания был написан код Python, позволяющий охарактеризовать комплексообразование тремя моделями: односубстратное (один сайт), двусубстратное (два сайта) и модель Хилла.

Константы диссоциации ( $K_d^{app}$ ) определяли на основании анализа зависимости концентрации комплекса белок-ДНК ([PL]) от концентрации белка при фиксированной концентрации ДНК. Эксперименты проводили в трех независимых повторах для каждой точки титрования. Концентрации выражали в молярных единицах (М). Для каждой концентрации белка рассчитывали среднее значение [PL] и стандартную ошибку среднего (SEM) по трем повторам. Аппроксимацию экспериментальных данных выполняли методом нелинейной регрессии с

использованием алгоритма наименьших квадратов (Python). Рассматривали несколько моделей связывания:

*Модель односубстратного связывания (1:1):*

$$[PL] = \frac{L_0 \cdot P}{K_d^{app} + P}$$

где  $L_0$  - концентрация ДНК,  $P$  - концентрация белка,  $K_d^{app}$  - константа диссоциации.

*Модель Хилла:*

$$[PL] = \frac{L_0 \cdot P^n}{K_d^{app} + P^n}$$

где  $n$  - коэффициент кооперативности.

*Модель двух независимых сайтов связывания:*

$$[PL] = L_0 \left( \frac{f_1 \cdot P}{k_{d1}^{app} + P} + \frac{(1 - f_1) \cdot P}{k_{d2}^{app} + P} \right)$$

$k_{d1}^{app}$  и  $k_{d2}^{app}$  - константы диссоциации двух сайтов,  $f_1$  - доля первого сайта.

Подбор параметров выполняли методом нелинейной регрессии с ограничением числа итераций ( $\text{maxfev} = 5000$ ). Качество аппроксимации оценивали по коэффициенту детерминации ( $R^2$ ), рассчитываемому как:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

где  $SS_{res}$  - сумма квадратов остатков,  $SS_{tot}$  – общая сумма квадратов отклонений от среднего значения. Погрешности параметров определяли как стандартные ошибки, рассчитанные из диагональных элементов ковариационной матрицы. По результатам аппроксимации предпочтение отдавали модели с наибольшим  $R^2$  при минимальном числе параметров.

**Спектры кругового дихроизма ДНК.** Олигонуклеотиды, содержащие и не содержащие мотив G4, нагревали в реакционном буфере (в зависимости от эксперимента) до 95 °C, выдерживали при этой температуре 20 мин, затем медленно охлаждали до комнатной температуры. Двухцепочечные ДНК «отжигали» в реакционном буфере по стандартной методике. Спектры КД регистрировали в кварцевой кювете с длиной оптического пути 10 мм в интервале температур от 30 до 85 °C с последовательным повышением температуры приблизительно на 5 °C при средней скорости нагрева 0,5 °C/мин на КД-спектрометре Chirascan (Applied Photophysics Ltd., Великобритания). Концентрация ДНК была выбрана для достижения оптической плотности 0,3–0,6  $OE_{260}$ , что дает оптимальное соотношение сигнал/шум. Измерения проводились в диапазоне длин волн 230–360 нм при скорости сканирования 30 нм/мин при постоянном потоке сухого азота. Все спектры КД были скорректированы по базовой линии с

учетом вклада сигнала, вызванного соответствующим буфером. Спектры обрабатывали с помощью программного обеспечения Origin 9.85 (Electronic Arts, Redwood City, CA, USA). Профили плавления КД выявили температурную зависимость сигнала КД при 265 нм.

**Гидролиз олигонуклеотидов и ДНК-дуплексов белком APE1.** Флуоресцентно-меченные препараты ДНК с концентрацией 100 нМ длиной 30 или 96 н.о. или п.н. (в зависимости от эксперимента), содержащие флуоресцентную метку TAMRA на 3'-конце ДНК, инкубировали с 10 ед. акт. (0,36 пмоль, 36 нМ) APE1 (New England Biolabs) в реакционном буфере, содержащем 1 мМ ДТТ, 0,5 мг/мл БСА, 5 мМ MgCl<sub>2</sub> при 37 °С в течение 60 мин. Реакцию останавливали инактивацией фермента при 60 °С в течение 20 мин. Реакционные смеси анализировали методом электрофореза в 20%-ном ПААГ в денатурирующих условиях (ТВЕ, 7 М мочевины) в течение 1 ч при напряжении 800 В. Фотографии гелей получали с помощью прибора Typhoon FLA 9500 (GE Healthcare, Великобритания).

**Анализ комплексообразования APE1 с ДНК.** Так как фермент APE1, который был нам любезно предоставлен проф. Жарковым Д.О. (Институт химической биологии и фундаментальной медицины СО РАН, г. Новосибирск), содержит дополнительную аминокислотную последовательность с шестью остатками His, для анализа комплексообразования использовали биосенсоры с иммобилизованной на нем Ni-NTA-агарозой (ForteBio, США) и интерферометр BLItz (ForteBio, США). Биосенсор «насыщали» белком APE1 (3,5 мкМ) в течение 5 мин, затем проводили стадию ассоциации ДНК, содержащий модификацию F, в концентрациях (10, 20, 50 и 100 нМ) в течение 5 мин. Эксперимент проводили при постоянном перемешивании (1000 об./мин) в реакционном буфере. Константу диссоциации комплексов рассчитывали в программе Origin 2021b (9.85) по следующим формулам:

$$R = R_{eq}(1 - e^{-(k_a \cdot C + k_d)(t-t_0)}), \text{ интегральное уравнение скорости ассоциации};$$

$R = R_0 \cdot e^{-k_d(t-t_0)} + R_{(t \rightarrow \infty)}$ , интегральное уравнение скорости диссоциации, где:  
 $R$  – отклик спектрометра в произвольный момент времени, сдвиг длины волны [нм];  
 $R_{eq}$  – отклик спектрометра, соответствующий стадии равновесия [нм];  
 $R_0$  – отклик спектрометра, соответствующий начальной стадии диссоциации [нм];  
 $R_{(t \rightarrow \infty)}$  – отклик спектрометра, соответствующий  $t \rightarrow \infty$  [нм];  
 $C$  – введенная концентрация аналита [нМ].

## Биоинформатические методы

### Докинг и молекулярная динамика.

Комплексообразование С-концевого домена ngMutL с β-субъединицей ДНК-полимеразы III исследовали методом молекулярного докинга с использованием пакета Schrodinger. В качестве исходных моделей использовали структуру β-«зажима» (Swiss-model, Q5FAJ1), а также PDB-

структуру CTD-MutL<sup>wt</sup> (3NCV). Для моделирования варианта CTD-MutL<sup>cf</sup> внесли точечные замены C532A, C604S и C635S в обе цепи гомодимера CTD-MutL<sup>wt</sup>. Структуры подвергали подготовке с учетом «protonation state» и оптимизации геометрии. Для каждого комплекса было рассчитано 30 наиболее вероятных конформаций связывания с β-«зажимом». Полученные комплексы дополнительно подвергали оптимизации в водной среде методом молекулярной динамики в модуле Desmond для оценки стабильности комплекса и количества контактных взаимодействий между CTD-MutL и β-«зажимом».

**Влияния дефицита MMR *Neisseria gonorrhoeae* на антигенную вариацию *pilE*.** Для локусов генов *ngmutS* и *ngmutL*, а также для дальнейшего анализа аллелей мы использовали плагин PubMLST. Используя последовательности генов *ngmutS* и *ngmutL* штамма *N. gonorrhoeae* MS11 (NCBI: WP\_003688085.1 и WP\_003688688.1) в качестве консенсуса, было обнаружено, что локусы NEIS2138 и NEIS1378, соответственно, представляют собой нуклеотидные последовательности. Все назначенные аллели *ngmutS* и *ngmutL* изолятов *N. gonorrhoeae*, представленные в базе данных, были собраны. Плагин PubMLST использовался для анализа полиморфизмов собранных аллелей *ngmutS* и *ngmutL*. Далее нуклеотидные последовательности этих аллелей были экспортированы в формат FASTA для трансляции в белковые последовательности и выравнивания в Jalview с использованием алгоритма ClustalW. Для поиска изолятов *N. gonorrhoeae*, в геноме которых отсутствует ген *ngmutS* или *ngmutL*, использовался плагин PubMLST «Gene present» (Наличие гена). В списке «Loci» были выбраны локусы гена NEIS2138 или NEIS1378 для *ngmutS* или *ngmutL*, соответственно. «Параметры и опции» не менялись: минимальный процент идентичности и выравнивания составлял 70% и 50% от общей длины выравнивания соответственно. Все идентификаторы изолятов, в геноме которых присутствовал ген *ngmutS* или *ngmutL*, были собраны и далее проанализированы с помощью плагина «Dataset» (Набор данных) для поиска тех изолятов, у которых аллель не был назначен.

Анализируемые изоляты *N. gonorrhoeae* были разделены на три группы: (1) содержащие ген *ngmutS* или *ngmutL* с назначенным аллелем; (2) содержащие ген *ngmutS* или *ngmutL* с неназначенным аллелем; (3) без интересующего гена. В зависимости от типа гена – *ngmutS* (S) или *ngmutL* (L) – эти группы были обозначены как 1S или 1L, 2S или 2L и 3S или 3L, соответственно. Мы проанализировали изменчивость *pilE* в каждой группе с помощью плагина «BLAST» базы данных PubMLST. Нуклеотидная последовательность *pilE* штамма MS11 *N. gonorrhoeae* использовалась в качестве консенсусной (UniProtKB/Swiss-Prot: P02974.2). Затем нуклеотидные последовательности гена *pilE* каждой группы были экспортированы в FASTA для выравнивания с помощью алгоритма ClustalW. Результаты выравнивания были восстановлены в Jalview и проанализированы в Python для определения частоты вариаций для каждой позиции *pilE*.

**Выявление консервативных G4-мотивов в области GC-богатого региона *TERT* промотора различных млекопитающих.** Мы загрузили 1000 п.н. промоторных областей *TERT*, предшествующих TSS (кодирующим цепям), для 158 млекопитающих из базы данных COG от 13 августа 2022 г. ([https://github.com/nooroka/conserved\\_quadruplexes](https://github.com/nooroka/conserved_quadruplexes), дата обращения: 13 августа 2022 г.). Во многих случаях координаты TSS не отображаются явно в записях GenBank, и нам пришлось определять их как самый близкий 5'-конец доступной мРНК для конкретного гена. С помощью программы QGRS, ищущего мотивы G4, соответствующие паттерну  $G_{3+} L_{1-30} G_{3+} L_{1-30} G_{3+} L_{1-30} G_{3+}$ , были найдены G4-мотивы в промоторных областях *TERT* у млекопитающих разных отрядов. Рассматривались только неперекрывающиеся мотивы G4. Максимальная длина последовательности составляла 45 нуклеотидов, минимальная длина G-тракта – три последовательных гуанозина, а длина петли – от 1 до 30 нуклеотидов. Координаты мотивов G4 определялись на кодирующей цепи ДНК. Для достижения нашей цели – поиска консервативных мотивов G4 и нуклеотидных замен в них – мы выбрали картограф QGRS.

#### IV. Особенности взаимодействия белков систем репарации ДНК с G-богатыми фрагментами регуляторных областей генома эукариот и прокариот

(Обсуждение результатов)

Системы эксцизионной репарации ДНК (BER и MMR) обеспечивают ключевой вклад в поддержание стабильности генома как у прокариот, так и у эукариот. Однако их эффективность во многом зависит от нуклеотидного контекста вокруг повреждения и структурных особенностей ДНК. Особый интерес представляют GC-богатые участки, которые локализованы преимущественно в регуляторных областях генома. Такие участки распространены в геномах всех организмов. Многие из этих участков содержат повторяющиеся остатки dG (G-тракты), разделенные короткими вставками. Такие участки получили обозначение PQS (putative G-quadruplex sequences), так как в одноцепочечном состоянии они склонны к формированию G-квадруплексов. Сама по себе GC-богатая последовательность обладает большим мутагенным потенциалом за счет окисления гуанина и спонтанного дезаминирования цитозина. В GC-богатых областях генома высокая термодинамическая стабильность и жесткость двойной спирали ухудшают доступность участков узнавания для белков, препятствуя необходимым локальным деформациям ДНК и правильной локализации ДНК-лиганда в центре связывания белка. В мутагенный потенциал GC-богатых последовательностей ДНК вносят вклад стабилизируемые ими вторичные структуры. Формирование G4 возможно в процессах, ассоциированных с расхождением цепей ДНК, например, репликации, транскрипции и репарации повреждений. Показано, что эти структуры играют значимую роль в регуляции биологических процессов, а также участвуют в контроле репликации и транскрипции [2].

В нашей работе на первом этапе влияние GC-богатых последовательностей на функционирование систем репарации ДНК прокариот было изучено на примере последовательности, локализованной перед промотором гена *pilE* патогена *N. gonorrhoeae* и способной *in vitro* формировать стабильный параллельный G4, предположительно необходимый для инициации гомологичной рекомбинации (см. раздел II.2.2. «Ключевые факторы вирулентности *Neisseria gonorrhoeae*»). Белки системы MMR *N. gonorrhoeae* «сдерживают» этот процесс, но для понимания самого механизма регуляции данных все еще недостаточно (см. раздел II.2.3. «Роль белков систем репарации ДНК в адаптации и изменчивости *Neisseria gonorrhoeae*»).

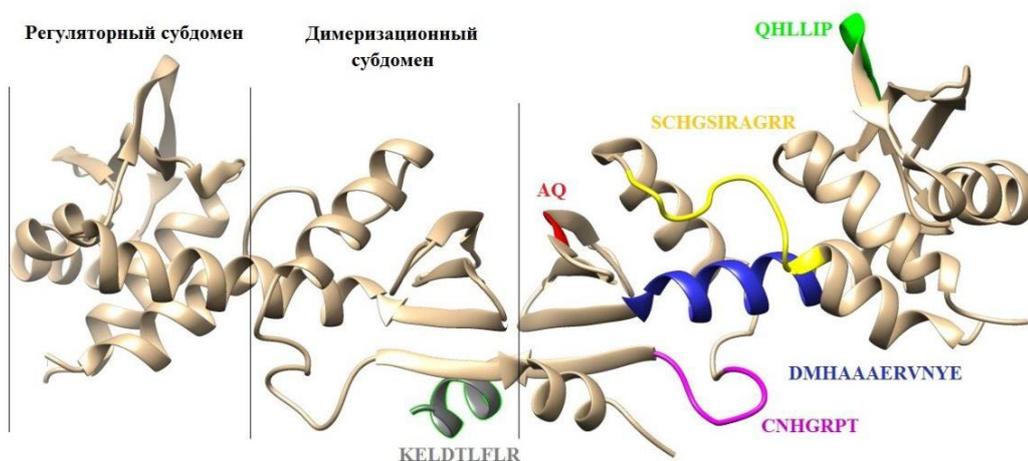
В клетках *N. gonorrhoeae* MMR функционирует по метилнезависимому принципу, ключевыми участниками процесса являются MutS (ngMutS) и MutL (ngMutL), обладающие специфической для данного организма активностью. Многие аспекты метилнезависимого пути репарации «мисматчей» остаются невыясненными, что подчеркивает необходимость

осуществления комплексного исследования функционала MMR. Стоит отметить, что не хватает данных о механизмах действия гомологов MutL, тогда как белки MutS структурно и функционально более консервативны от организма к организму, включая *E. coli*. Именно поэтому в рамках первого этапа данной работы было проведено комплексное исследование активности ключевого белка системы репарации «мисматчей» – ngMutL.

#### IV. 1. Новые аспекты функционирования системы MMR по метилнезависимому пути

Согласно базе данных National Center for Biotechnology Information (NCBI, США), ngMutL, как и другие бактериальные гомологи, является гомодимером в растворе и содержит 658 аминокислотных остатков (а.о.) в мономере белка. Белок состоит из двух доменов – N-концевого (NTD - 366 а.о.) и С-концевого (CTD - 148 а.о.), которые соединены неструктурированным линкером из 144 а.о. [195]. Гомологи MutL, включая ngMutL, многофункциональны и обладают ключевыми активностями: АТФазной, ДНК-связывающей и эндонуклеазной.

Домены, ответственные за реализацию АТФазной и ДНК-связывающей активности локализируются в NTD-ngMutL. Хотя кристаллическая структура полноразмерного ngMutL не решена, наличие рентгеноструктурного анализа N-концевого домена гомологов (например, esMutL - PDB-код: 1B63) позволяет сервису AlphaFold Database получить модель NTD-ngMutL. Область, ответственная за эндонуклеазную функцию белка ngMutL, находится в С-концевом домене (CTD), его кристаллическая структура была получена в работе Namadurai et al. (PDB-код: 3NCV) [196].



**Рисунок 11.** Кристаллическая структура С-концевого домена ngMutL (PDB: 3NCV). Пять консервативных мотивов эндонуклеазного центра белка указаны красным, желтым, синим, фиолетовым и серым цветами. Металлсвязывающий мотив обозначен синим цветом.

В каталитическом центре белка расположены пять консервативных мотивов (Рис. 11), включая металлсвязывающий - DMHAAAERVNYE (на рис. 11 обозначен синим), который играет

важную роль в координации ионов металлов при гидролизе ДНК, и мотив QHLLIP, ответственный за взаимодействие ngMutL с  $\beta$ -субъединицей ДНК-полимеразы III для усиления процессивности в ходе MMR (на рис. 11 показан зеленым цветом) [195]. Кроме того, ngMutL-CTD содержит димеризационный субдомен и регуляторный (внешний) субдомен, обращенный к растворителю. Некоторые а.о. консервативных мотивов критически важны для реализации эндонуклеазной функции гомологов MutL и функционирования системы MMR. Например, двойная замена аминокислотных остатков D491N и E497Q металлсвязывающего мотива DMHAAAERVNYE ngMutL существенно снижает эффективность гидролиза ДНК [92]. Многие гомологи MutL теряют активность *in vivo* при замене консервативных а.о. цистеинов в мотивах, аналогичных SCHGSIRAGRR и CPHGRPT ngMutL (Рис. 12), однако причина инактивации до конца неясна [204].

PMS2 <i>H. sapiens</i>	RVKQMFASRA	CRKSV	MIGTALNTSEM	KKLITHMGEMDHPWN	CPHGRPTMRHIANLGVISQ	861
PMS1 <i>S. cerevisiae</i>	KIRSMFAMRA	CRSSIM	IGKPLNKKTMTRVVHNLSELDK	PWNC	CPHGRPTMRHLMELRDWSS	897
<i>T. thermophilus</i>	SLKALLARLAC	LPAVKAGHPLG	EAQGOALLDALLACETPWAC	CPHGRPVLLALKEEDLIRR	514	
<i>B. subtilis</i>	LREEAAIMM	SKGSIKANRH	LRNDEIKALLDDL	RSTSDPFT	CPHGRPTIIHHSTYEMEKM	622
<i>N. gonorrhoeae</i>	HENHILATM	SCHGSV	RAGRQLTLP	EMNALLRDM	ENTPRSNQCNHGRPTWVKLTLKELD	653

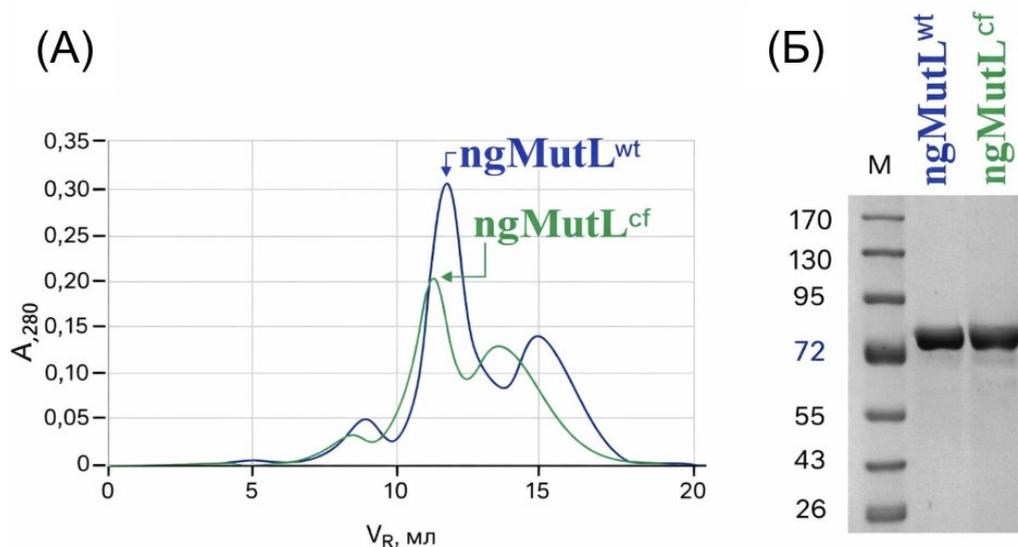
**Рисунок 12.** Сравнение аминокислотных последовательностей гомологов MutL из разных организмов. Представлен фрагмент, содержащий два консервативных мотива С-концевого домена, включающий консервативные остатки цистеина.

Считается, что остатки цистеина часто включены в металлсвязывающие мотивы такие, как, например, цинковый палец [197]. Каталитический центр ngMutL содержит два консервативных остатка цистеина (С604 и С635), роль которых во взаимодействии с ионами марганца и цинка не вполне изучена.

#### IV. 1.1. Выделение белков ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup>

Для исследования функциональных особенностей ngMutL и установления роли консервативных а.о. цистеина были выделены в достаточном количестве препараты белков ngMutL дикого типа (ngMutL<sup>wt</sup>) и его мутантной формы (ngMutL<sup>cf</sup>), не содержащей а.о. цистеина (Рис. 13). В мутантной форме ngMutL<sup>cf</sup> все имеющиеся в диком типа белка 5 остатков Cys были заменены на другие аминокислотные остатки (С171I, С251L, С528А, С600S и С631S). В ходе широкомасштабного сравнения последовательностей гомологов MutL из разных организмов [198] выяснилось, что на месте С171 ngMutL в 40% исследуемых аминокислотных последовательностей гомологов находится I, на месте С251 – L, С528 – А (в 30% случаев). Два консервативных аминокислотных остатка цистеина были заменены на Ser. Для экспрессии рекомбинантных белков ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup> в клетках *E. coli* штамма BL21(DE3) использовали плазмидную конструкцию pET15b, которая содержит ген *lacI* белка-репрессора *lac*-оперона, промотор и терминатор T7 РНК-полимеразы, ген устойчивости к ампициллину, и гены, кодирующие белки ngMutL<sup>wt</sup> или ngMutL<sup>cf</sup> с шестью остатками His в N-концевой области белков.

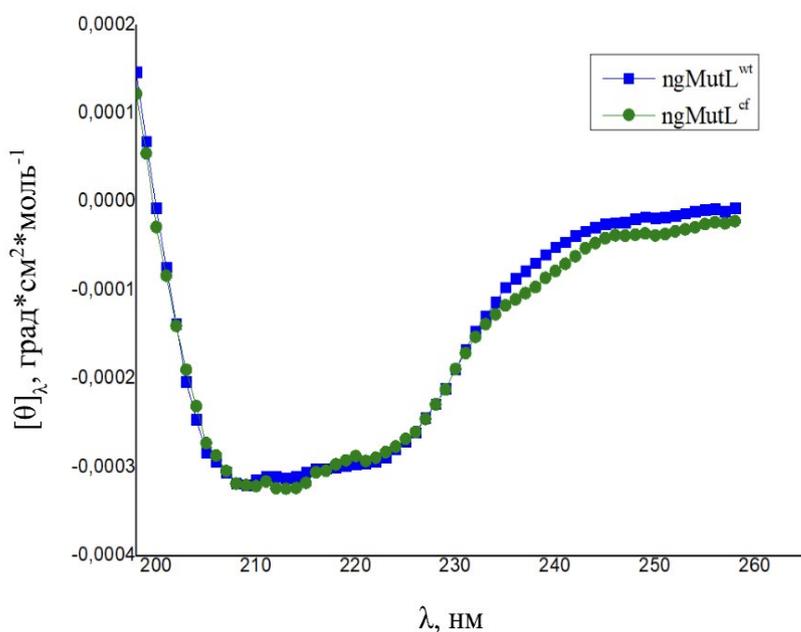
Белки выделяли методом аффинной хроматографии на Ni-NTA-агарозе (Novagen, Германия), дополнительную очистку белков MutL, проводили методом эксклюзионной хроматографии на колонке Enrich SEC 650 (Biorad, США). Фракции, поглощающие при 280 нм (Рис. 13А), собирали и анализировали методом гель-электрофореза по Лэммли (Рис. 13Б). Видно, что молекулярная масса выделенных белков составляет примерно 72 кДа, что соответствует рассчитанной молекулярной массе MutL.



**Рисунок 13.** Очистка белка ngMutL методом эксклюзионной хроматографии. **(А)** Хроматограмма разделения белков ngMutL<sup>wt</sup> (кривая синего цвета) и ngMutL<sup>cf</sup> (кривая зеленого цвета); **(Б)** Анализ препаратов полученных белков после эксклюзионной хроматографии: **М** – смесь белков-маркеров молекулярной массы 10-180 кДа (ThermoFisher Scientific, США); Электрофорез в 8%-ном ПААГ с ДСН. Гель окрашен раствором кумасси бриллиантового синего G-250.

#### IV.1.2. Сравнение вторичной структуры ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup>

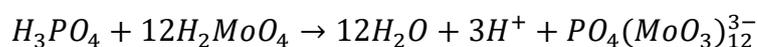
Для оценки влияния замен аминокислотных остатков Cys на структуру белка ngMutL<sup>cf</sup> использовали метод кругового дихроизма (КД). Спектры КД для белка дикого типа и его мутантной формы регистрировали на спектрометре Chirascan (Applied Photophysics Ltd., Великобритания) в диапазоне длин волн 200–250 нм при комнатной температуре. Исследуемые образцы находились в буфере PBS (см. раздел III «Материалы и методы»), содержащем 50 мМ KCl. Следует отметить, что для полноразмерного белка ngMutL подобный эксперимент проводился впервые. Результаты показали, что спектры КД белков ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup> не отличаются значительно и имеют по два максимума в области 208 и 222 нм, что характерно для белков с преобладанием  $\alpha$ -спиралей в структуре (Рис. 14). Таким образом, можно полагать, что замена всех аминокислотных остатков цистеина в ngMutL<sup>wt</sup> не приводит к существенному изменению вторичной структуры белка ngMutL<sup>cf</sup>, поэтому полученный препарат можно использовать в дальнейших сравнительных экспериментах.



**Рисунок 14.** Спектры КД белков ngMutL<sup>wt</sup> (кривая синего цвета) и ngMutL<sup>cf</sup> (кривая зеленого цвета).

#### IV.1.3. Характеристика и сравнение АТФазной активности белков ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup>

Связывание и гидролиз АТФ в АТФазном центре MutL способствует изменению конформации белка, что необходимо для взаимодействия с другими участниками системы репарации и реализации его эндонуклеазной активности [199]. Сравнительную характеристику АТФазной активности исследуемых белков определяли колориметрическим методом с использованием красителя малахитового зеленого. Метод был успешно использован ранее для определения каталитических констант Михаэлиса-Ментен для белка дикого типа [200]. В составе реакционной смеси присутствовали ионы молибдата (VI) в кислой среде и малахитовый зеленый, характеризующийся желтой окраской. При внесении фосфат-ионов происходит образование фосфомолибдатного комплекса, не обладающего собственным цветом, который связывает молекулы красителя. Этот процесс сопровождается изменением спектральных свойств раствора: окраска смещается с желтой к изумрудной. Количественная оценка реакции проводится по изменению оптической плотности при длине волны 590 нм, что отражает концентрацию фосфат-ионов в реакционной среде.



Препараты белков ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup> использовали в концентрации 0,8 мкМ в присутствии 50 нМ 76-звенного ДНК-дуплекса (см. раздел III. «Материалы и методы»), так как присутствие ДНК увеличивает эффективность гидролиза АТФ ферментом [200]. Концентрацию

АТФ варьировали от 100 до 1000 мкМ. АТФазная активность белков ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup> была охарактеризована кинетическими параметрами, рассчитанными с помощью применения схемы Михаэлиса-Ментен (Таблица 10).

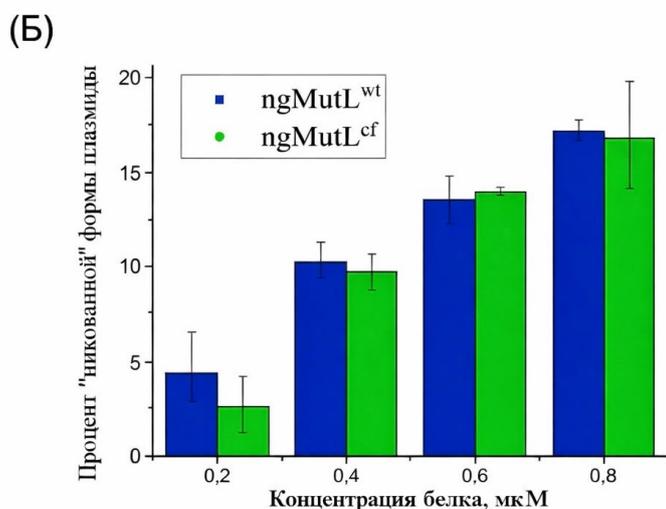
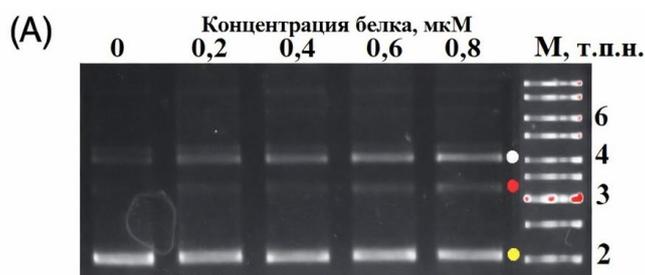
**Таблица 10.** Параметры реакции гидролиза АТФ белками.

Белок	$K_M$ , мкМ	$V_{max}$ , МЕ/мг $\times 10^3$
ngMutL <sup>wt</sup>	377±16	7,18±0,14
ngMutL <sup>cf</sup>	354±16	5,7±0,12

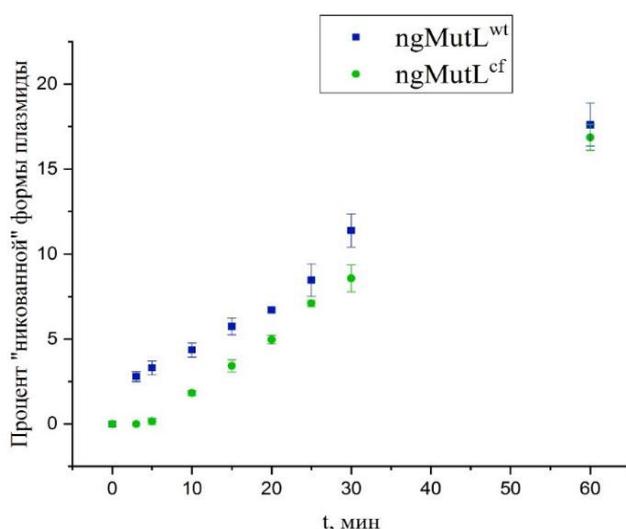
Полученные данные свидетельствуют о том, что замены а.о. цистеина ngMutL не оказывают влияния на сродство белка к АТФ ( $K_M$  реакции гидролиза АТФ белками ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup> одинаковы в пределах погрешности эксперимента), однако приводят к снижению каталитической активности, что проявляется в уменьшении максимальной скорости гидролиза примерно на 20%.

#### IV. 1.4. Характеристика и сравнение эндонуклеазной активности белков ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup>

Белок ngMutL катализирует образование одноцепочечных разрывов ДНК в присутствии ионов  $Mn^{2+}$ , что обеспечивается С-концевым доменом, содержащим два консервативных остатка Cys (С600 и С631) [200]. В первую очередь гидролитическую активность ферментов оценивали с использованием в качестве субстрата плазмидной ДНК рUC-MMR. Так как гомологи MutL не обладают специфичностью к «мисматчу», то во всех экспериментах, связанных с внесением разрыва в ДНК, были использованы субстраты без «мисматчей». Реакцию гидролиза ДНК белками MutL проводили в присутствии 5 мМ  $MnCl_2$  и 50 мМ КСl при 37 °С в течение 60 мин. Продукты гидролиза рUC-MMR белками ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup> анализировали методом электрофореза в агарозном геле (Рис. 15А). В результате гидролиза ферментами суперскрученная форма плазмиды превращалась в «никованную» – содержащую одноцепочечный разрыв и обладающую меньшей электрофоретической подвижностью. Полученные данные показывают, что даже при замене консервативных Cys бесцистеиновая форма ngMutL сохраняет способность вносить одноцепочечные разрывы в плазмидную ДНК. Эффективность гидролиза за 60 мин сопоставима с активностью белка дикого типа (Рис. 15Б). Сравнение кинетики реакции показало, что ngMutL<sup>wt</sup> вносит разрывы с первых минут, тогда как активность ngMutL<sup>cf</sup> проявляется после 5-й минуты (Рис. 16). При этом в диапазоне 10–30 мин эффективность гидролиза у мутантного белка была немного ниже, но к 60-й минуте различия нивелировались.



**Рисунок 15. Сравнительный анализ эффективности гидролиза** плазмиды pUC-MMR белками ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup> при 37 °С в течение 60 мин. (А) Гидролиз плазмиды pUC-MMR белком ngMutL<sup>wt</sup>. Репрезентативная электрофореграмма в 1%-ном агарозном геле, содержащем бромид этидия. М – маркер длины ДНК в тысячах пар нуклеотидов (т.п.н.). Точками белого, красного и желтого цвета указано положение «никованной», линейной и суперскрученной форм плазмиды, соответственно. (Б) Процент «никованной» формы ДНК, полученной в результате реакции гидролиза плазмиды pUC-MMR белками ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup>. Погрешности соответствуют 95%-ным доверительным интервалам (n = 3).



**Рисунок 16. Эффективность образования «никованной» формы ДНК** в результате реакции гидролиза плазмиды pUC-MMR белками ngMutL<sup>wt</sup> (синий) и ngMutL<sup>cf</sup> (зеленый) при 37 °С в зависимости от времени реакции. Погрешности соответствуют 95% доверительным интервалам (n = 3).

Таким образом, замены цистеинов не оказывает существенного влияния на эндонуклеазную функцию, и ngMutL<sup>cf</sup> может применяться в дальнейших экспериментах. Это указывает на отсутствие критической роли этих остатков для реализации эндонуклеазной активности ngMutL *in vitro* и координации ионов металлов.

#### IV.1.5. Эффективность гидролиза линейного субстрата белком ngMutL в присутствии $\beta$ -субъединицы ДНК-полимеразы III

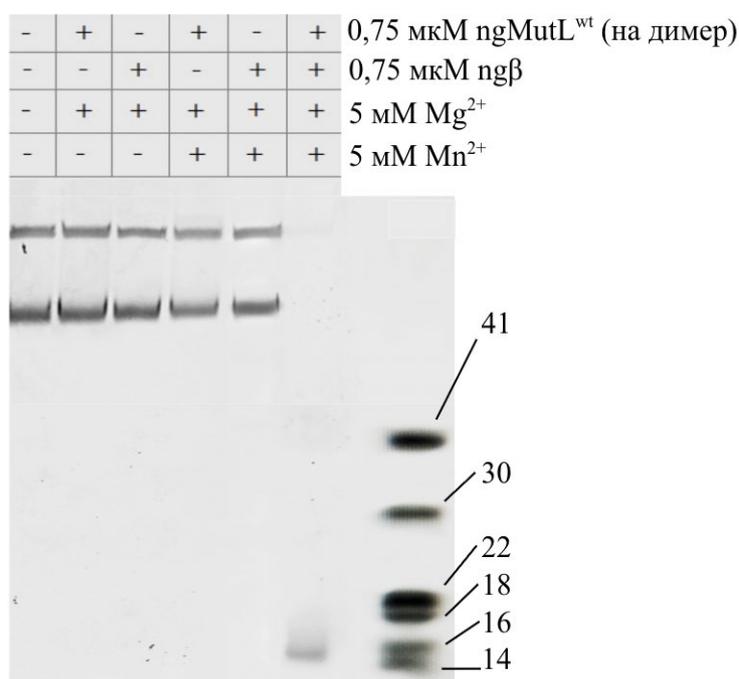
У большинства бактериальных MutL выявлен консервативный мотив QXXIXP (X - любой а.о.), обеспечивающий взаимодействие с  $\beta$ -субъединицей ДНК-полимеразы III, что повышает эффективность гидролиза ДНК; такие белки относят к подсемейству I [201]. Считается, что это взаимодействие способствует сближению каталитического центра С-концевого домена MutL с субстратом. Существуют представители подсемейств II (например, MutL из *Thermus thermophilus*) и III (например, MutL из *Aquifex aeolicus*), лишённые консервативного мотива или содержащие его в изменённой форме, они демонстрируют  $\beta$ -«зажим»-независимую эндонуклеазную активность (Рис. 17) [202].

Организм		Подсемейство
<i>B. subtilis</i>	AQERIKYEYFREKVGVEVEPEVQEMIVPLTFHYSTNE 503	I
<i>N. gonorrhoeae</i>	AERVNYEKMKRORQENGNLQSQHLLIPVTFAASHEE 531	I
<i>T. thermophilus</i>	AHERILYEDLLKRVAEGPKPLPR---PLLVLPAPEE 400	II
<i>A. aeolicus</i>	EERINYEKCLKDE----- 367	III

**Рисунок 17.** Сравнение аминокислотных последовательностей гомологов MutL из разных организмов. Представлен фрагмент, содержащий консервативный мотив QXXIXP (показан желтым) С-концевого домена, обеспечивающий взаимодействие с  $\beta$ -субъединицей ДНК-полимеразы III.

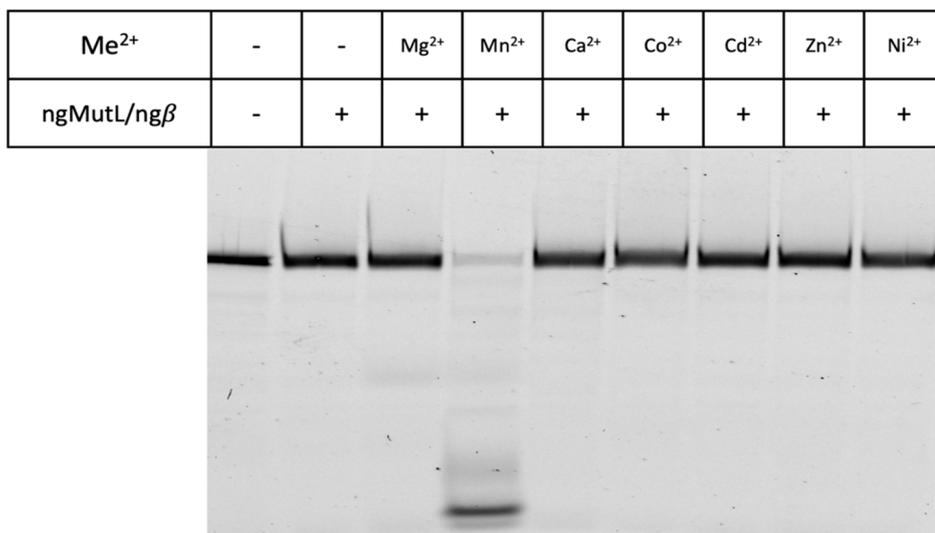
Исследуемый белок ngMutL содержит консервативный мотив QHLLIP, поэтому для исследования гидролиза белками линейного ДНК-дуплекса необходимо было вносить в систему  $\beta$ -субъединицу ДНК-полимеразы III из *N. gonorrhoeae* (ng $\beta$ ). Для характеристики гидролиза ДНК мы использовали линейный ДНК-дуплекс длиной 76 п.н., содержащий флуоресцентную метку - TAMRA на 3'-конце в «верхней» цепи (см. Раздел III. «Материалы и методы»). Мы проанализировали гидролиз линейного ДНК-дуплекса белком ngMutL<sup>wt</sup> (Рис. 18). Препарат ng $\beta$  был получен ранее научным сотрудником НИИ им. Белозерского МГУ Монаховой М.В. Условия эксперимента были выбраны в соответствии с данными работы [203]. Установлено, что ngMutL<sup>wt</sup> практически полностью гидролизует указанный выше субстрат в присутствии 5 мМ MgCl<sub>2</sub> (для эффективного гидролиза АТФ), 5 мМ MnCl<sub>2</sub>, 0,5 мМ АТФ и ng $\beta$  (димер) в эквимольном

количестве по отношению к димеру ngMutL<sup>wt</sup> при 37°C за 90 мин. На рис. 18 представлен репрезентативный результат электрофореза в 10%-ном ПААГ в денатурирующих условиях, где наблюдается зона, соответствующая коротким продуктам гидролиза линейного 76-звенного ДНК-дуплекса белком ngMutL<sup>wt</sup>. В качестве контрольных экспериментов проводили гидролиз субстрата белками ngMutL<sup>wt</sup> и ngβ отдельно в присутствии ионов Mg<sup>2+</sup> или Mn<sup>2+</sup>. Отсутствие продукта реакции гидролиза в контрольных пробах свидетельствует об отсутствии неспецифических нуклеаз в препарате фермента, а также демонстрирует отсутствие гидролиза линейного субстрата белком ngMutL<sup>wt</sup> без ngβ. Исходя из этого, можно сделать однозначный вывод, что для эффективной реализации эндонуклеазной функции ngMutL на линейных двутяжевых ДНК-субстратах необходимо присутствие β-«зажима». Таким образом, исследуемый белок является β-«зажим»-зависимой эндонуклеазой, что было показано нами впервые.



**Рисунок 18.** Гидролиз линейного 76-звенного ДНК-дуплекса (10 нМ), содержащего флуоресцентную метку TAMRA на 3'-конце «верхней» цепи, при 37 °С в течение 90 мин в присутствии 0,5 мМ АТФ. М – маркер длины ДНК (н.о. - нуклеотидные остатки). Представлен репрезентативный результат электрофореза в 10%-ном ПААГ в денатурирующих условиях (7 М мочевины, ТВЕ).

Изучение гидролиза линейных ДНК-субстратов белком ngMutL в присутствии β-«зажима» ранее не проводилось, поэтому было необходимо подобрать оптимальные условия протекания реакции. Известно, что для реализации эндонуклеазной функции большинства гомологов MutL из различных организмов необходимы ионы двухвалентного марганца, однако имеются индивидуальные особенности. Мы провели сравнительную характеристику эффективности гидролиза линейного 76-звенного ДНК-дуплекса белком ngMutL<sup>wt</sup> (0,75 мкМ на димер белка) в присутствии эквимольного количества ngβ (димер), 0,5 мМ АТФ и 5 мМ солей различных металлов, указанных на рис.19.

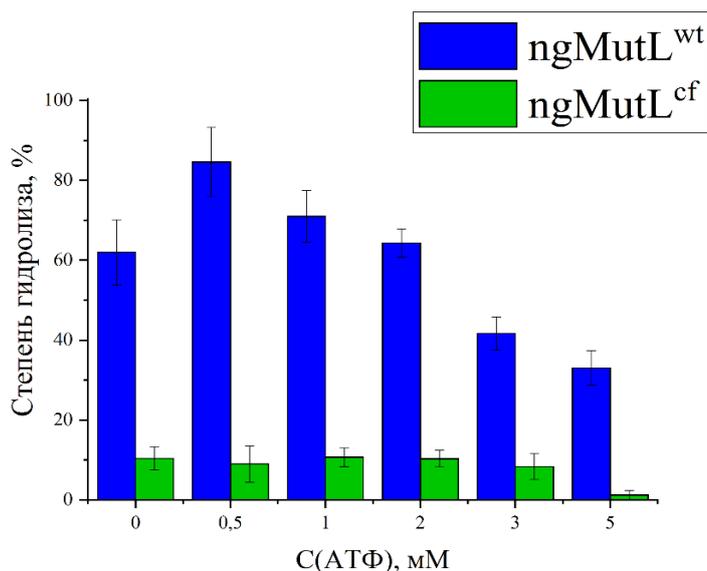


**Рисунок 19.** Влияние ионов различных двухвалентных металлов на гидролиз ДНК-дуплекса длиной 76 п.н. (10 нМ). Реакционную смесь инкубировали в присутствии ngMutL<sup>wt</sup> (0,75 мкМ на димер), ngβ (0,75 мкМ на димер), 0,5 мМ АТФ и 5 мМ Me<sup>2+</sup> в течение 90 мин при 37 °С. Представлен репрезентативный результат электрофореза в 10%-ном ПААГ в денатурирующих условиях (7 М мочевины, буфер ТВЕ).

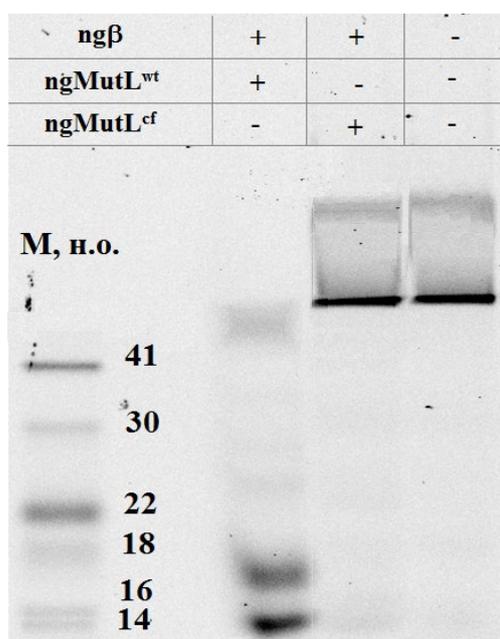
Белок ngMutL<sup>wt</sup> проявляет наибольшую эндонуклеазную активность в присутствии Mn<sup>2+</sup>, подобно другим бактериальным гомологам MutL. Mg<sup>2+</sup> активирует эндонуклеазную функцию ngMutL, но гораздо менее эффективно, чем Mn<sup>2+</sup>. Ионы Ca<sup>2+</sup> не способствуют расщеплению дуплекса длиной 76 п.н., опосредованному ngMutL; что также верно для Cd<sup>2+</sup>, Co<sup>2+</sup>, Ni<sup>2+</sup> и Zn<sup>2+</sup>. Комбинация 5 мМ MgCl<sub>2</sub> и 5 мМ MnCl<sub>2</sub> – является оптимальной, так как Mg<sup>2+</sup> необходим для АТФазной активности ngMutL.

Ранее показано, что повышение концентрации АТФ в реакционной смеси от 0,5 до 5 мМ снижает образование «никванной» формы плазмиды pUC-MMR белком ngMutL<sup>wt</sup> [200]. Согласно нашим исследованиям эффективность гидролиза линейного 76-звенного дуплекса была максимальной при 0,5–2 мМ АТФ, тогда как 5 мМ АТФ вызывал выраженное ингибирование (Рис. 20). В отсутствие АТФ гидролиз ДНК происходил, но менее эффективно, вероятно из-за недостаточных конформационных перестроек белка для взаимодействия с β-«зажимом». Снижение активности ngMutL при высоких концентрациях АТФ может объясняться хелатированием ионов металлов, необходимых для катализа. Для последующих экспериментов была выбрана концентрация АТФ, равная 0,5 мМ. Нами проведено сравнение эффективности гидролиза линейного субстрата белками ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup> в концентрации 0,75 мкМ на димер (Рис. 21). В выбранных условиях, а именно в присутствии 5 мМ MgCl<sub>2</sub>, 5 мМ MnCl<sub>2</sub>, 0,5 мМ АТФ и эквимольном количестве димера ngβ при 37 °С в течение 90 мин, эффективность гидролиза линейного 76-звенного ДНК-дуплекса белком ngMutL<sup>cf</sup> резко снижена по сравнению с эффективностью расщепления субстрата белком дикого типа: степень гидролиза субстрата

составляет не более 15%. Кроме того, в диапазоне концентраций АТФ от 0 до 5 мМ повышение эндонуклеазной активности бесцистеинового варианта ngMutL не наблюдалось (Рис. 20 и 21).



**Рисунок 20.** Зависимость гидролиза линейного 76-звенного ДНК-дуплекса, содержащего флуоресцентную метку TAMRA на 3'-конце «верхней» цепи, при 37 °С в течение 90 мин от концентрации АТФ (0-5 мМ). Степень гидролиза ДНК (в %) белками ngMutL<sup>wt</sup> (синие столбцы) и ngMutL<sup>cf</sup> (зеленые столбцы). Реакции проводили в присутствии 5 мМ MnCl<sub>2</sub>. Погрешности соответствуют 95%-ным доверительным интервалам (n = 3).

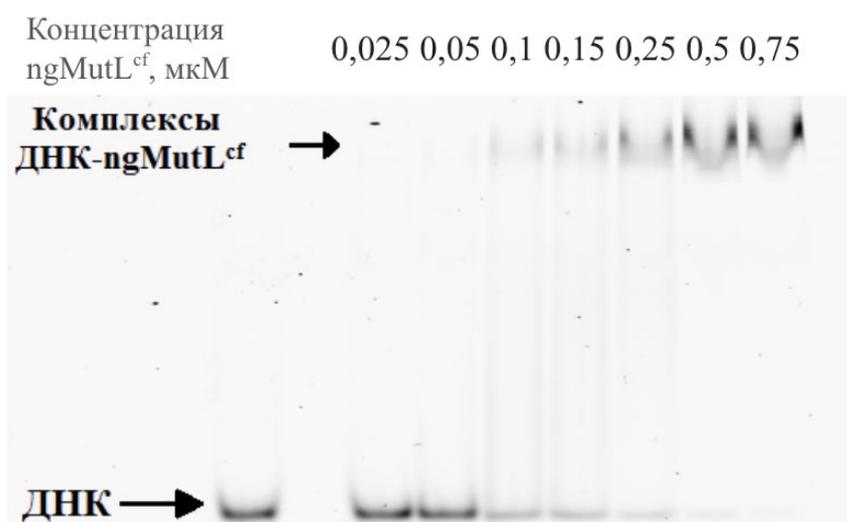


**Рисунок 21.** Гидролиз линейного 76-звенного ДНК-дуплекса, содержащего флуоресцентную метку TAMRA на 3'-конце «верхней» цепи, при 37°С в течение 90 мин димерами ngMutL<sup>wt</sup> (0,75 мкМ) и ngMutL<sup>cf</sup> (0,75 мкМ) в присутствии эквимольной концентрации димера ngβ, 5 мМ MgCl<sub>2</sub>, 5 мМ MnCl<sub>2</sub> и 0,5 мМ АТФ. М – маркер длины ДНК (н.о.). Представлен репрезентативный результат электрофореза в 10%-ном ПААГ в денатурирующих условиях (7 М мочеви́на, TBE).

Таким образом, замена всех а.о. Cys в ngMutL приводит к резкому снижению способности фермента гидролизовать линейный субстрат в присутствии  $\beta$ -«зажима», причем этот эффект не зависит от концентрации АТФ.

#### IV.1.6. Анализ ДНК-связывающей активности ngMutL

Замена всех Cys, включая консервативные, снижает гидролиз линейного субстрата в присутствии  $\beta$ -«зажима», возможно, из-за менее эффективного связывания ДНК ngMutL<sup>cf</sup>. Ранее в нашей лаборатории было показано, что ngMutL<sup>wt</sup> эффективно связывает дуплекс длиной 76 п.н., и были подобраны условия такого комплексообразования [200]. Для проверки гипотезы было проведено сравнение эффективности связывания препаратов ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup> с 76-звенным ДНК-дуплексом ДНК, меченным флуорофором TAMRA на 3'-конце (см. раздел III. «Материалы и методы»). Образование комплекса между ферментом и субстратом регистрировали методом «торможения в геле» (Рис. 22). При 100 мМ KCl ДНК-связывающая активность бесцистеинового варианта ngMutL<sup>cf</sup> достигает почти 100% в диапазоне 1–0,75 мкМ, что сопоставимо с активностью белка дикого типа.



**Рисунок 22.** Анализ комплексообразования димера ngMutL<sup>cf</sup> (0,025 - 0,75 мкМ) с линейным 76-звенным ДНК-дуплексом, содержащим флуоресцентную метку TAMRA на 3'-конце «верхней» цепи. Представлен репрезентативный результат электрофореза в 4%-ном ПААГ в неденатурирующих условиях в буфере TBE.

Комплексообразование ngMutL<sup>wt</sup> и ngMutL<sup>cf</sup> с 76-звенным ДНК-дуплексом охарактеризовано кажущимися константами диссоциации с помощью модели Хилла (см. раздел III. «Материалы и методы»), равными в пределах погрешности эксперимента (Таблица 11).

**Таблица 11.** Константы диссоциации комплексов исследуемых белков с ДНК.

Фермент	ngMutL <sup>wt</sup>	ngMutL <sup>cf</sup>
K <sub>d</sub> <sup>app</sup> , нМ	117±7	118±6
Коэффициент Хилла	2,9±0,2	3,1±0,4

Таким образом, замена всех аминокислотных остатков Cys в белке ngMutL не влечет за собой снижение ДНК-связывающей активности белка, следовательно, а.о. цистеина не вовлечены во взаимодействие с ДНК-субстратом. Другой причиной резкого снижения эффективности гидролиза линейного субстрата белком ngMutL<sup>cf</sup> может быть изменение характера его взаимодействия с молекулярным партнером β-«зажимом».

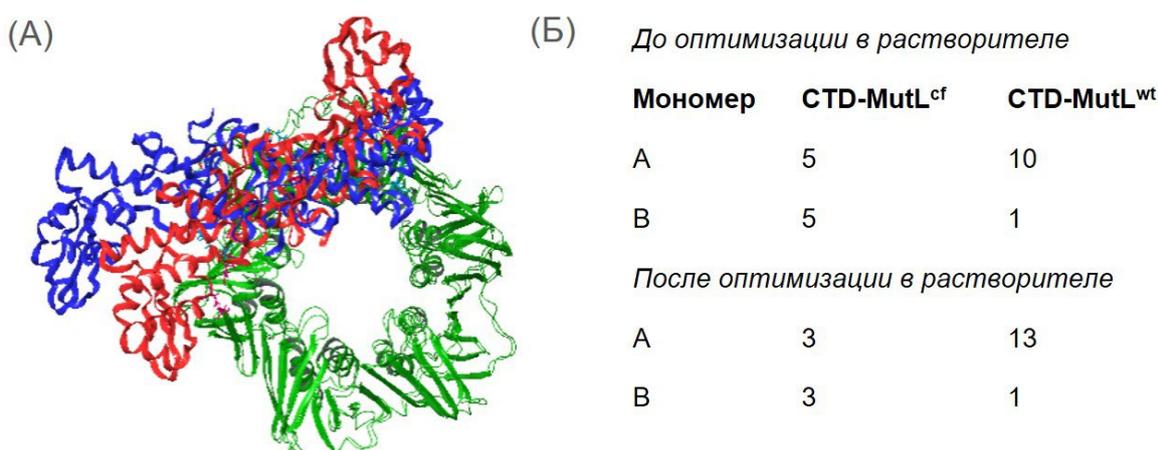
#### IV.1.7. Анализ взаимодействия ngMutL с β-субъединицей ДНК-полимеразы III

β-субъединица ДНК-полимеразы III представляет собой гомодимер в растворе, образует кольцевую структуру и «скользит» по цепи ДНК, привлекая туда белковых партнеров. Основная роль β-«зажима» - усиливать процессивность ДНК-полимеразы III, однако он напрямую связывается и с другими белками, необходимыми для поддержания стабильности генома, например, с MutS и MutL. Белок-белковое взаимодействие ферментов с β-«зажимом» необходимо, например, для увеличения времени пребывания фермента на ДНК или для координации и регуляции его активности. На поверхности С-концевого домена β-«зажима» имеется гидрофобная область, с которой взаимодействует большинство белковых партнеров. В свою очередь, β-«зажим»-зависимые ферменты содержат достаточно консервативный β-связывающий мотив, который взаимодействует с гидрофобной областью β-«зажима» (мотив QHLLIP в исследуемом белке ngMutL, см. раздел IV.1.5. «Эффективность гидролиза линейного субстрата белком ngMutL в присутствии β-субъединицы ДНК-полимеразы III»). И хотя данный мотив не содержит а.о. цистеина, их замена может косвенно влиять на белок-белковое взаимодействие, изменяя локальное расположение мотивов С-концевого домена ngMutL относительно друг друга.

Комплексообразование С-концевого домена MutL с β-«зажимом» было исследовано нами методом молекулярного докинга с помощью пакета программ Schrodinger. В качестве моделей были выбраны swiss model β-«зажима» (Q5FAJ1), структура CTD-MutL<sup>wt</sup> (PDB: 3NCV) и CTD-MutL<sup>cf</sup>. На первом этапе провели подготовку имеющихся PDB-структур, а также внесли точечные замены C532A, C635S, C604S в структуру CTD-MutL<sup>wt</sup> обеих цепей гомодимера для получения структуры CTD-MutL<sup>cf</sup>. Для CTD-MutL<sup>wt</sup> и CTD-MutL<sup>cf</sup> получены 30 наиболее вероятных вариантов комплексообразования с β-«зажимом». Показано, что связывание ngMutL с ngβ не нарушается в ходе введения замен, однако геометрия комплексов отличалась: в случае CTD-

MutL<sup>wt</sup> комплексообразование с  $\beta$ -«зажимом» реализуется в основном за счет одного из мономеров MutL, тогда как у CTD-MutL<sup>cf</sup> оба мономера вовлечены в формирование комплекса (Рис. 23А и Б - до оптимизации в растворителе). Количество взаимодействий с  $\beta$ -«зажимом» составило 11 и 10 для CTD-MutL<sup>wt</sup> и CTD-MutL<sup>cf</sup>, соответственно. Стоит отметить, что ни один из остатков цистина CTD-MutL<sup>wt</sup> не вовлечен в формирование комплекса *in silico*.

Полученные комплексы дополнительно оптимизировали в растворителе (воде) методом молекулярной динамики для проверки их устойчивости. В динамике количество взаимодействий  $\beta$ -«зажима» с CTD-MutL<sup>cf</sup> оказалось равным 6, в то время как для CTD-MutL<sup>wt</sup> - 14 (Рис. 23Б). Это говорит о значительном снижении прочности комплекса ng $\beta$  с CTD-MutL<sup>cf</sup> в присутствии растворителя.



**Рисунок 23.** Результат молекулярного докинга ng $\beta$  с CTD-MutL. (А) Модель комплекса ng $\beta$  (зеленый) с CTD-MutL<sup>wt</sup> (синий) и CTD-MutL<sup>cf</sup> (красный). (Б) Количество взаимодействий CTD-MutL<sup>cf</sup> и CTD-MutL<sup>wt</sup> с ng $\beta$  на каждый мономер (А и В).

Таким образом, снижение эффективности гидролиза ДНК-дуплекса комплексом ngMutL-ng $\beta$  в результате замен консервативных а.о. цистеина в белке MutL, может быть следствием менее эффективного комплексообразования между  $\beta$ -«зажимом» и CTD-MutL<sup>cf</sup> по сравнению с белком дикого типа. Хотя остатки цистеина CTD-MutL не участвуют непосредственно в комплексообразовании с  $\beta$ -«зажимом», они обеспечивают конформационную жесткость домена, необходимую для формирования энергетически выгодного взаимодействия белков. Их замена сопровождается значительным снижением устойчивости комплекса в условиях растворителя, что указывает на роль Cys в стабилизации функционального комплекса MutL-ng $\beta$ .

\*\*\*\*

Таким образом, в ходе предварительного исследования функционирования ngMutL было установлено, что белок представляет собой  $\beta$ -«зажим»-зависимую эндонуклеазу, то есть эндонуклеазная активность этого фермента существенно усиливается в присутствии  $\beta$ -

субъединицы ДНК-полимеразы III. Продемонстрировано, что гидролиз ngMutL линейного субстрата реализуется только в присутствии двухвалентных катионов  $Mg^{2+}$  и  $Mn^{2+}$ . Также АТФ в диапазоне концентраций до 2 мМ способствует эффективному гидролизу ДНК. Остатки цистеина значительным образом не вовлечены в АТФазную, ДНК-связывающую или эндонуклеазную активности, но могут быть необходимы для стабильного взаимодействия ngMutL с  $\beta$ -субъединицей ДНК-полимеразы III в клеточных условиях. Результаты исследования опубликованы в работе с участием автора диссертации [200].

#### **IV. 2. Биоинформатический анализ влияния полиморфизмов белков MMR на антигенную вариацию пилина**

В ходе комплексного исследования свойств ngMutL были подобраны оптимальные условия функционирования белка, которые использовались в дальнейшем для выявления влияния GC-богатой последовательности на функционирование MMR на примере области, содержащей G4-мотив и локализующейся перед промотором гена *pilE* патогена *N. gonorrhoeae*. Было предложено несколько вариантов механизма регуляции антигенной вариации *pilE* посредством MMR: 1) повышение эффективности антигенной вариации за счет общего увеличения мутаций из-за дисфункции MMR в таких областях; 2) регуляция белками MMR стабильности G4 и, как следствие, ингибирование чрезмерной вариативности *pilE*.

##### **IV.2.1. Оценка полиморфизма генов *ngmutS* и *ngmutL* в популяции *Neisseria gonorrhoeae***

Для оценки общего вклада системы MMR в регуляцию антигенной изменчивости был проведен анализ полиморфизма генов *ngmutS* и *ngmutL* в популяции *N. gonorrhoeae* с использованием базы данных PubMLST – крупнейшего ресурса, содержащего информацию о многоочаговом типировании последовательностей в геномах и аллелях штаммов *Neisseria*. В результате анализа 15632 клинических изолятов, содержащих данные по локусу NEIS2138 (*ngmutS*), было выявлено 342 аллельных варианта, включающих 867 полиморфных позиций в гене длиной 2697 нуклеотидов. Несмотря на наличие мутаций подавляющее большинство участков в последовательности *ngmutS* оказались высококонсервативными: доля позиций с вариативностью менее 10% составляет значительную часть. Самые распространенные аллели охватывают 13,4% всех изолятов, что свидетельствует о сильной селективной консервативности. Аналогичный анализ был проведен для локуса NEIS1378, кодирующего белок ngMutL, на выборке из 15975 изолятов. Обнаружено 310 аллельных вариантов, содержащих 460 полиморфных позиций в гене длиной 2017 нуклеотидов. Наиболее вариабельным оказался участок в пределах 1370–1425 нуклеотидов, соответствующий линкерному домену белка

ngMutL, что может указывать на меньшую структурную значимость данного участка. Тем не менее, в целом *ngmutL* также демонстрирует высокую степень консервативности: доля переменных позиций составила 14,2% по всей популяции.

#### IV.2.2. Связь между полиморфизмом белков MMR и антигенной вариацией пилина

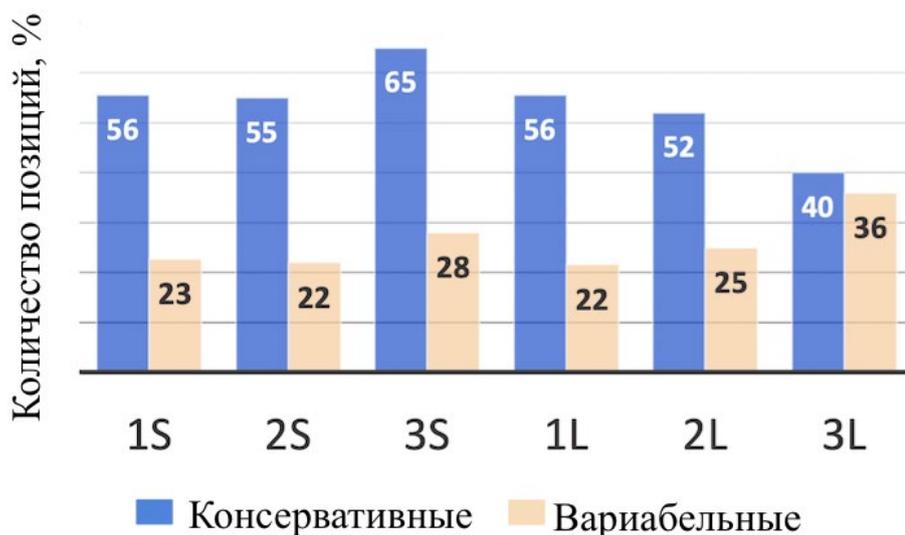
Для изучения влияния отсутствия или функционального дефекта компонентов MMR на изменчивость *pilE* – основного антигена, подвергающегося гомологичной рекомбинации – геномы изолятов, представленных в базе данных PubMLST, были классифицированы на группы:

1. 1S / 1L – штаммы с присутствием соответствующего гена (*ngMutS/ngMutL*, соответственно) и назначенным аллелем;
2. 2S / 2L – штаммы, у которых ген присутствует, но аллель не определен;
3. 3S / 3L – штаммы, у которых *ngmutS/ngmutL* отсутствуют.

Так как последовательности *pilE* демонстрируют высокую изменчивость и не всегда типированы в базе данных, для определения частоты вариаций была проведена реконструкция *pilE* из всех геномов с помощью BLAST (по консенсусной последовательности штамма MS11 – одного из наиболее распространенных штаммов *N. gonorrhoeae*), сравнение нуклеотидной и аминокислотной последовательности в программе Jalview и расчет частоты вариаций. Было установлено, что частота полиморфизмов *pilE* значительно увеличена в группах 3S и 3L (где отсутствуют соответствующие гены MMR, рис. 24), по сравнению с группами (1) и (2). В частности, доля высоковариабельных позиций (>25%) возрастала, а доля консервативных позиций снижалась (особенно в группе 3L). Интересно, что одновременное отсутствие обоих генов MMR наблюдалось крайне редко (только у двух изолятов). Статистическая обработка данных (t-критерий Стьюдента) подтвердила значимые различия в количестве полиморфных сайтов между группами с отсутствием MMR и остальными ( $p \ll 0,05$ ), в то время как между группами с присутствием гена, но неопределенным аллелем, и референсными (1S–2S, 1L–2L) статистически значимой разницы не выявлено. Это указывает на то, что именно отсутствие белков MutS и MutL, а не вариации в пределах аллелей, ассоциировано с усилением антигенной вариации.

Таким образом, природные MMR-дефицитные штаммы действительно склонны к повышенной антигенной вариации *pilE*.

## Частота вариаций *pilE*



**Рисунок 24.** Полиморфизм гена *pilE* в группах всех изолятов *N. gonorrhoeae*. Количество позиций гена *pilE* с определенными нуклеотидными заменами в каждой группе изолятов *N. gonorrhoeae* как отношение консервативных (<10%) или переменных (>25%) замен к общей длине выравнивания.

### IV.3. Взаимодействие белка ngMutL с G4-мотивом регуляторной последовательности перед *pilE*

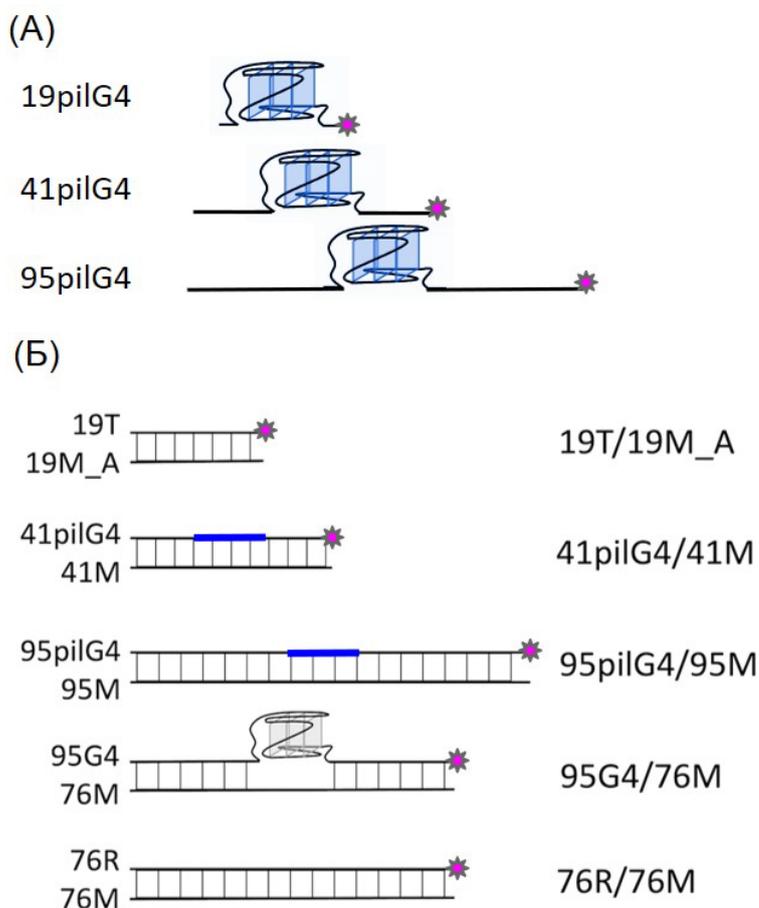
Известно, что гомологи MutS обладают повышенным сродством к G4 по сравнению с ДНК канонического строения [122], однако для MutL подобных исследований не проводилось. Для проверки гипотезы о возможности белков MMR регулировать стабильность G4 и, как следствие, ингибировать чрезмерную антигенную вариацию *pilE* *N. gonorrhoeae*, было проведено экспериментальное исследование взаимодействия белка ngMutL с участком ДНК, содержащим G4-мотив перед промотором *pilE*.

#### IV.3.1. Характеристика модельных субстратов, содержащих G4-мотив перед *pilE*

Для анализа взаимодействия ngMutL с G4-мотивом были синтезированы олигонуклеотиды различной длины, содержащие GC-богатую последовательность, соответствующую регуляторному участку перед *pilE* *N. gonorrhoeae* (см. раздел II.2.4. «Неканонические структуры ДНК и их влияние на патогенез и репарацию»). Все последовательности, их сокращенные обозначения, а также соответствующие комплементарные цепи представлены в таблице 6 (см. раздел III. «Материалы и методы»).

Для характеристики взаимодействия ngMutL с G4-мотивом были использованы модельные одноцепочечные ДНК-фрагменты, обозначенные как 19pilG4, 41pilG4 и 95pilG4, которые

различались по длине фланкирующих G4 участков (Рис. 25А). В качестве контрольных ДНК-лигандов использовали ДНК-дуплексы, образованные в ходе добавления комплементарной С-богатой цепи к соответствующим олигонуклеотидам с последующей гибридизацией (см. раздел III. «Материалы и методы»). Они обозначены как 41pilG4/41М и 95pilG4/96М (Рис. 25Б).

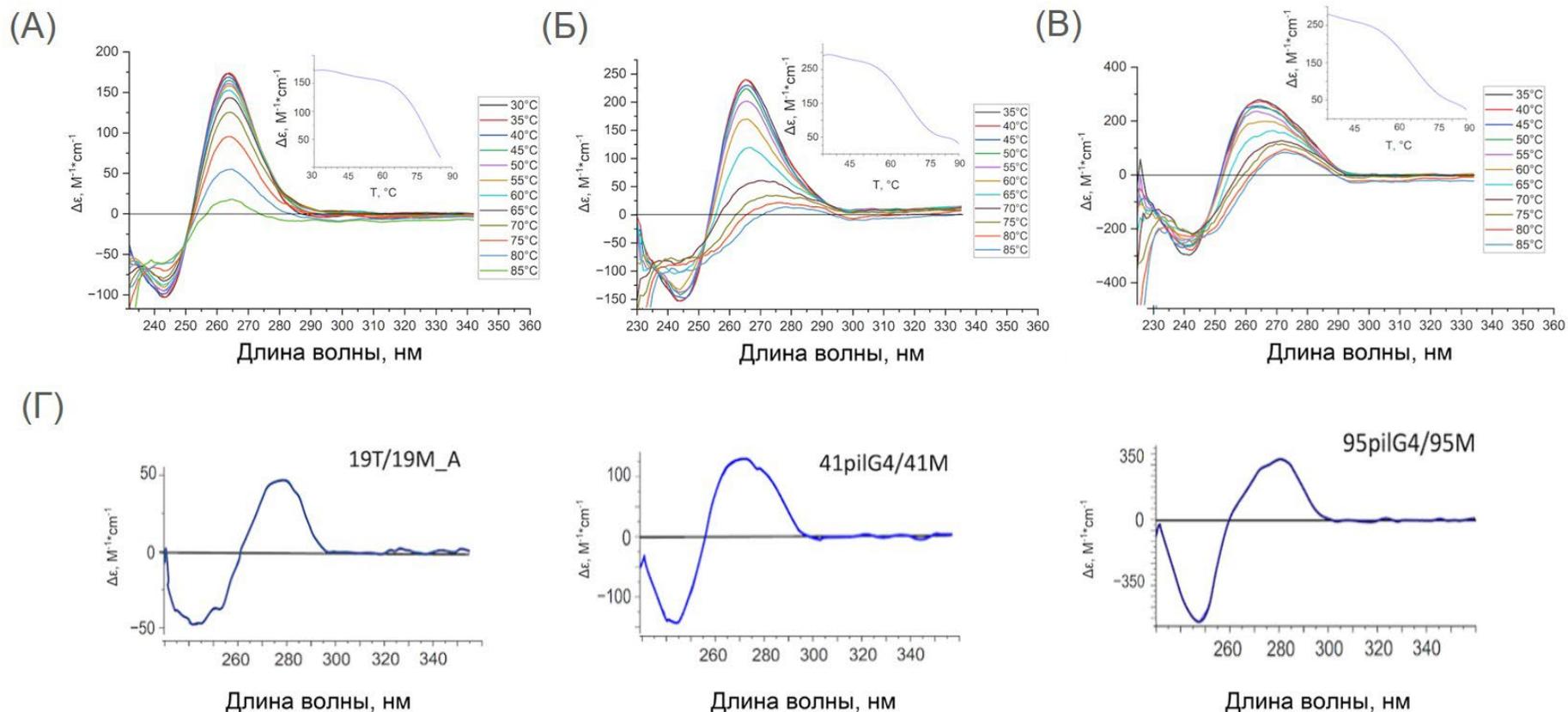


**Рисунок 25.** Структура и сокращенные обозначения модельных ДНК, использованных в исследовании. (А) Одноцепочечные модельные ДНК, содержащие G4-мотив регуляторного участка *pilE*. (Б) Двухцепочечные модельные ДНК. Наличие и локализация G4-мотива обозначены синим цветом. TAMRA на 3'-конце ДНК показана розовой звездочкой.

Для формирования стабильного 19-звенного дуплекса, содержащего мотив  $(GGGT)_4$ , было необходимо ввести четыре замены  $G \rightarrow T$  (олигонуклеотид 19Т), чтобы полностью подавить возможное формирование G4 (Таблица 6, см. раздел III. «Материалы и методы»). Комплементарный олигонуклеотид 19М\_А позволил получить стабильный контрольный дуплекс 19Т/19М\_А (Рис. 25Б). Модельный ДНК-дуплекс 95G4/76М образован 95-звенным олигонуклеотидом, содержащим  $(GGGT)_4$ -мотив, и 76-звенным олигонуклеотидом, комплементарным фланкирующим областям G4-мотива 95-звенного ДНК-фрагмента. Ранее было показано, что конструкция 95G4/76М способна стабильно формировать параллельную G4-структуру внутри дуплекса и сохранять структурную топологию в условиях, идентичных

физиологическим [192]. Данная система была выбрана как модель для последующего анализа гидролиза ДНК белком ngMutL.

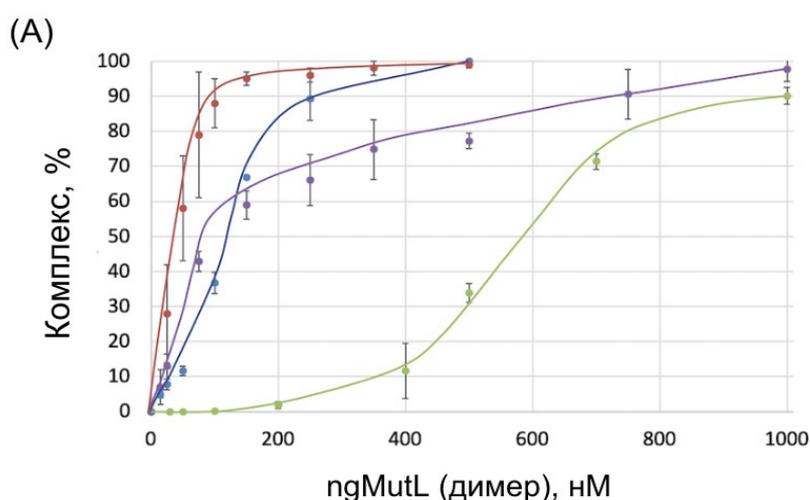
Для оценки стабильности образуемых G4-структур (19pilG4, 41pilG4 и 95pilG4) использовалась спектроскопия КД. Полученные спектры демонстрировали типичную для параллельной G4-структуры топологию: положительный пик при 265 нм и отрицательный – около 240 нм (Рис. 26А-В). Температуру плавления ( $T_{пл}$ ) олигонуклеотидов определяли по температурной зависимости молярной эллиптичности при 265 нм как температуру середины кооперативного перехода, соответствующую точке перегиба сигмоидальной кривой денатурации. Увеличение длины участков, фланкирующих G4, сопровождалось снижением температуры плавления G4:  $T_{пл}$  для 19pilG4 составила  $77 \pm 1$  °С, для 41pilG4 –  $66 \pm 1$  °С, и для 95pilG4 –  $63 \pm 1$  °С. Это связано с ростом энтропийного вклада от одноцепочечных участков, фланкирующих G4. Как и ожидалось, добавление комплементарной С-богатой цепи к ДНК-фрагменту с G4-мотивом (во всех случаях кроме 19pilG4) приводило к полной перестройке спектра КД в спектр, типичный для В-формы ДНК с максимумом при  $\sim 280$  нм. Это свидетельствует об образовании дуплекса даже в присутствии 100 мМ KCl, стабилизирующего G4 (Рис. 26Г).



**Рисунок 26.** Спектры КД олигонуклеотидов 19pilG4 (А), 41pilG4 (Б) и 95pilG4 (В), зарегистрированные при различных температурах в 20 мМ буфере HEPES (pH 7,3), содержащем 5 мМ KCl и 140 мМ NaCl (концентрация олигонуклеотидов ~ 2 мкМ). Температура повышалась от 30 до 85 °С с шагом 5 °С; цветные линии показывают спектры КД при различных температурах. На вставках показаны «кривые плавления» при длине волны 265 нм. (Г) Спектры КД моделей двухцепочечной ДНК (концентрация ~ 2 мкМ), зарегистрированные при 37 °С в 20 мМ буфере HEPES (pH 8,0), содержащем 100 мМ KCl. Именно этот буфер впоследствии использовался для анализа взаимодействия ngMutL с ДНК, содержащей *pilE* G4.

### IV.3.2. Связывание ngMutL с одно- и двуцепочечными ДНК, содержащими G4-мотив регуляторной последовательности перед *pilE*

Поскольку ДНК-связывающая активность белка ngMutL, являющегося ключевым компонентом системы MMR, существенно зависит от длины ДНК-субстрата, мы провели сравнительный анализ его взаимодействия с G4-содержащими олигонуклеотидами и соответствующими дуплексами различной длины. Образование комплекса между ngMutL и ДНК регистрировалась методом «торможения» в геле. На рис. 27А представлены графики зависимости связывания белка с одно- и двуцепочечными фрагментами длиной 41 и 95 звеньев, содержащими последовательность G4-мотива, от концентрации ngMutL.



(Б)

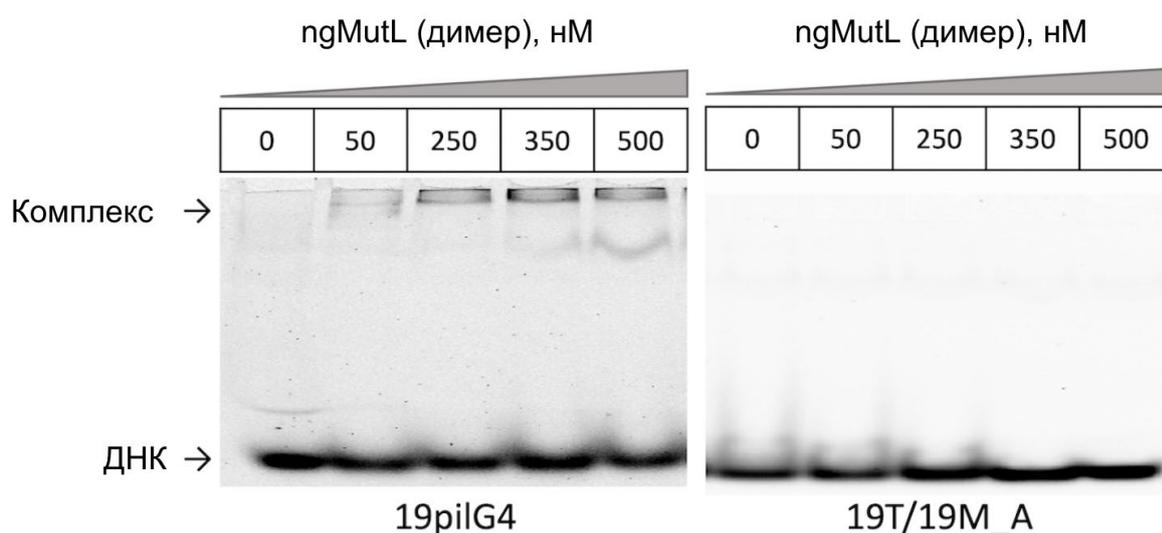
ДНК	$K_d^{app}$ , нМ
● 95pilG4/95M	123±8
● 95pilG4	41±1
● 41pilG4/41M	650±85
● 41pilG4	135±30

**Рисунок 27.** Связывание ngMutL с одноцепочечными и двухцепочечными модельными ДНК (20 нМ). Реакционные смеси инкубировали в течение 10 мин в бане со льдом с 15–1000 нМ ngMutL (на димер) в 20 мМ буфере HEPES (pH 8,0), содержащем 100 мМ KCl, 0,5 мг/мл БСА и 1 мМ ДТТ. Комплексообразование анализировали в 6%-ном ПААГ в буфере TAE при 4 °С. (А) Зависимость комплексообразования в % от концентрации ngMutL. (Б) Значения  $K_d^{app}$  комплексов ngMutL-ДНК. Расчет проводился как минимум для трех независимых экспериментов. Погрешности соответствуют 95%-ным доверительным интервалам ( $n = 3$ ).

Взаимодействие охарактеризовано значениями кажущейся константы диссоциации ( $K_d^{app}$ , рис. 27Б). Видно, что при увеличении длины фланкирующих участков существенно возрастает сродство ngMutL к одноцепочечному ДНК-лиганду:  $K_d^{app}$  для 95pilG4 почти в 3 раза ниже, чем

для 41pilG4. Это подтверждает, что длина окружающих G4 участков важна для эффективного связывания белка. Интересно, что аналогичная зависимость сохраняется и для соответствующих дуплексных форм:  $K_d^{app}$  для 95-звенного дуплекса составила  $123 \pm 8$  нМ, тогда как для 41-звенного –  $650 \pm 85$  нМ. Более того, независимо от длины, ngMutL демонстрирует повышенное сродство к одноцепочечным G4-структурам, по сравнению с дуплексами. Это подчеркивает, что G4 сам по себе является более предпочтительной мишенью связывания для белка, чем каноническая B-форма ДНК.

Для короткой модели 19pilG4 не удалось достоверно рассчитать  $K_d^{app}$  из-за крайне слабого связывания. Тем не менее, при 25-кратном избытке белка наблюдалось образование комплекса (~30%), тогда как для 19-звенного дуплекса (19T/19M\_A) комплексообразование отсутствовало (Рис. 28).

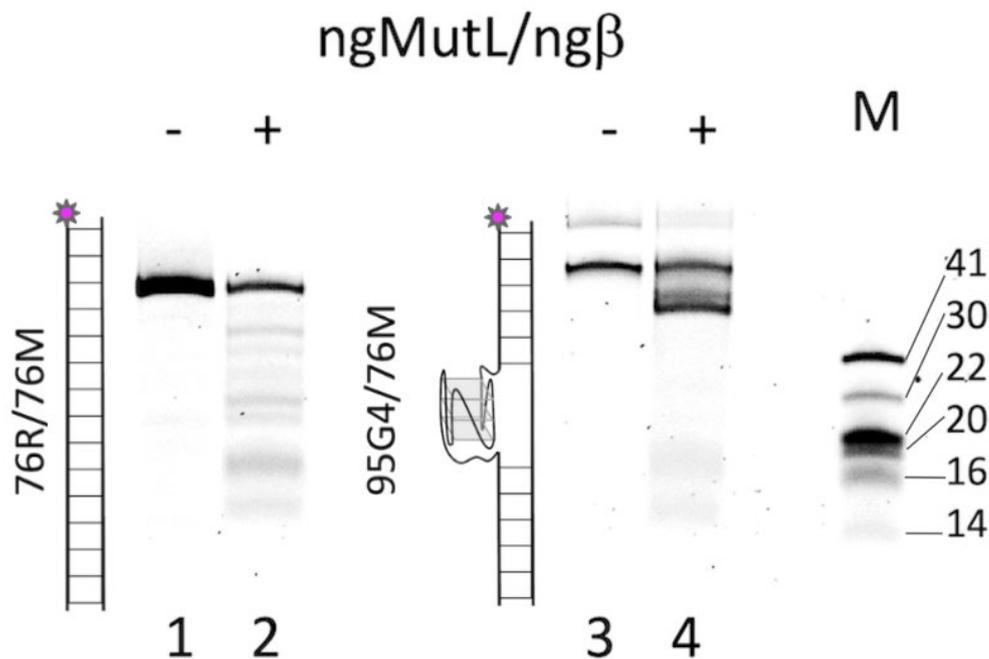


**Рисунок 28.** Электрофореграммы комплексообразования ngMutL с 19pilG4 и с двуцепочечным 19T/19M\_A. ДНК (20 нМ) инкубировали в течение 10 мин в бане со льдом с ngMutL в 20 мМ буфере HEPES (pH 8,0), содержащем 100 мМ KCl, 0,5 мг/мл БСА и 1 мМ ДТТ, а затем анализировали в 6%-ном ПААГ в буфере TAE при 4 °С. Над дорожками геля указана концентрация ngMutL в расчете на димер.

Таким образом, повышенное сродство ngMutL к G4-структурам сохраняется при увеличении длины фланкирующих дуплексных участков, что свидетельствует о биологической значимости наблюдаемого эффекта и его потенциальной реализуемости в условиях геномного контекста. При образовании G4 в GC-богатом участке рядом с промотором *pilE*, ngMutL способен взаимодействовать с такой структурой, стабилизируя ее и препятствуя дальнейшему продвижению репликационно-рекомбинационной машины. Это может быть одним из механизмов подавления повышенной антигенной изменчивости.

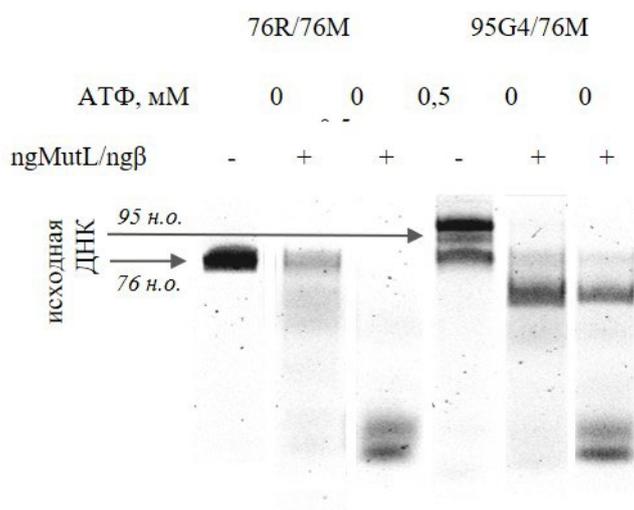
#### IV.3.3. Гидролиз белком ngMutL ДНК, содержащих G4-мотив регуляторной последовательности *pilE*

У бактерий, включая *N. gonorrhoeae*, белок MutL функционирует как «никующая» эндонуклеаза, активность которой зависит от взаимодействия с  $\beta$ -субъединицей ДНК-полимеразы III, что было показано нами впервые (см. раздел IV.1.5. «Эффективность гидролиза линейного субстрата белком ngMutL в присутствии  $\beta$ -субъединицы ДНК-полимеразы III»). В данной части исследования была оценена способность ngMutL расщеплять модельную дуплексную ДНК, содержащую одиночный G4-мотив d(GGGT)<sub>4</sub> (конструкция 95G4/76M), стабилизированный за счет отсутствия комплементарного участка в противоположной цепи. Для сравнения использовался 76-звенный дуплекс с произвольной последовательностью (76R/76M). Эндонуклеазная активность ngMutL оценивалась в присутствии эквимольного количества ng $\beta$  при наиболее оптимальных условиях (0,5 мМ АТФ, 5 мМ MgCl<sub>2</sub>, 5 мМ MnCl<sub>2</sub>, соотношение белок:ДНК составляло 25:1), при которых наблюдался эффективный, но не 100%-ый гидролиз контрольного субстрата для детекции продуктов различной длины (что доказывает неспецифичность внесения разрыва). Гидролиз оценивался по накоплению укороченных 3'-ТАМРА-меченных продуктов, разделенных в ПААГ в денатурирующих условиях (Рис. 29). Результаты показали, что дуплекс 76R/76M в данных условиях гидролизуется эффективно и неспецифично, тогда как степень расщепления 95G4/76M была существенно ниже. При этом расщепление происходило только до и после участка, содержащего G4, но не внутри него.



**Рисунок 29.** Гидролиз ДНК (10 нМ) эндонуклеазой ngMutL. Реакционные смеси инкубировали в присутствии 0,5 мМ АТФ, 5 мМ MgCl<sub>2</sub> и 5 мМ MnCl<sub>2</sub> в течение 90 мин при 37 °С, а затем анализировали в 12%-ном ПААГ, содержащем 7 М мочевины; использовали 0,25 мкМ ngMutL (на димер) в присутствии эквимольного количества димера ng $\beta$ . Представлены электрофореграммы продуктов расщепления ДНК, меченных 3'-ТАМРА: 76R/76M (дорожки 1–2) и 95G4/76M (дорожки 3–4). Длины ДНК-маркеров М (в н.о.) указаны справа.

На следующем этапе был проведен сравнительный анализ гидролиза *ngMutL* субстратов 76R/76M и 95G4/76M в условиях 100% расщепления контрольного субстрата (76R/76M). Этот эффект был достигнут при соотношении ДНК:белок = 1:75. Добавление в реакционную смесь АТФ в концентрации 0,5 мМ способствовало увеличению эффективности расщепления как канонического дуплекса, так и G4-содержащей ДНК (Рис. 30). В присутствии АТФ продукты гидролиза представляли собой короткие фрагменты (до 20 нуклеотидов), тогда как без него преобладали более длинные. По-видимому, гидролиз сопряжен с АТФ-зависимыми конформационными изменениями: *ngMutL* «скользит» вдоль ДНК, повторно взаимодействуя с субстратом после каждого акта расщепления. Аналогичный эффект наблюдался и при расщеплении G4-содержащего субстрата 95G4/76M: в присутствии АТФ усиливалось накопление коротких фрагментов, но сам G4 оставался интактным. Таким образом, даже при высоких концентрациях фермента и в присутствии всех необходимых кофакторов, *ngMutL* не расщепляет G4-структуру, что демонстрирует ее устойчивость к ферментативному гидролизу.

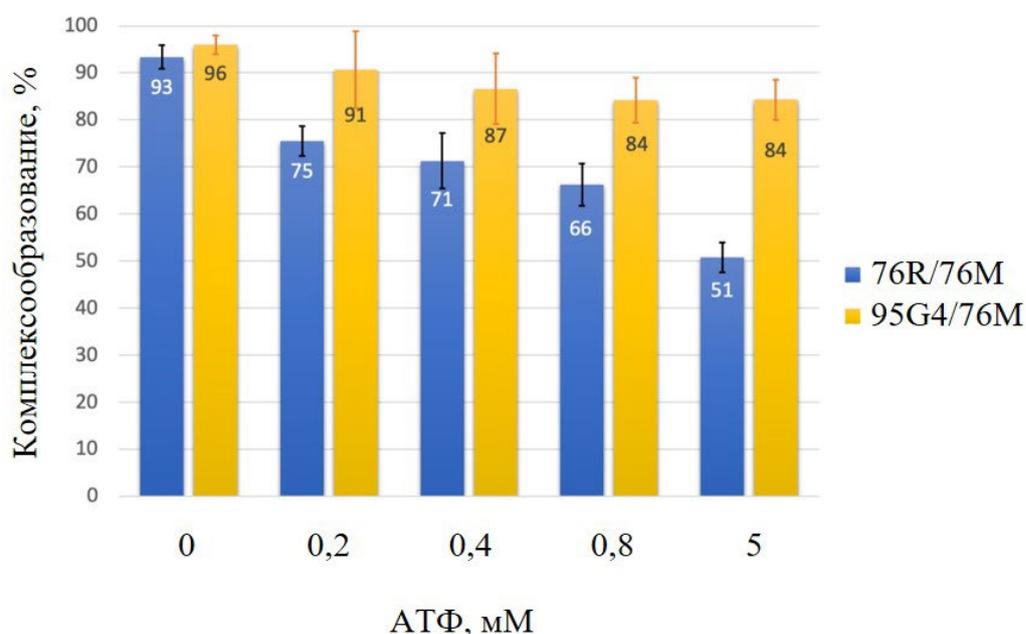


**Рисунок 30.** Влияние АТФ на гидролиз 76R/76M и 95G4/76M (10 нМ) эндонуклеазой *ngMutL*. Реакционные смеси инкубировали в присутствии 0 или 0,5 мМ АТФ, 5 мМ  $MgCl_2$  и 5 мМ  $MnCl_2$ , 0,75 мкМ *ngMutL* (на димер) и эквимольного количества димера *ngβ* в течение 90 мин при 37 °С, а затем анализировали в 12%-ном ПААГ, содержащем 7 М мочевины. Представлена электрофореграмма продуктов расщепления ДНК, меченных 3'-ТАМРА.

#### IV.3.4. Влияние АТФ на связывание G4-содержащей ДНК белком *ngMutL*

Дополнительные эксперименты по связыванию *ngMutL* с субстратами показали, что значительное повышение концентрации АТФ снижает сродство белка к каноническому дуплексу 76R/76M, но на связывание с 95G4/76M влияет значительно меньше (Рис. 31). Это говорит о том, что взаимодействие *ngMutL* с G4 менее чувствительно к АТФ-зависимым регуляторным механизмам, чем с В-формой ДНК. При высоких концентрациях АТФ связывание *ngMutL* с дуплексной ДНК снижается, что связано с конформационными перестройками белка,

индуцированными избытком АТФ. Эти изменения могут уменьшать доступность центра связывания ngMutL с двойной спиралью и снижать эффективность контакта белка с дуплексом. В то же время G-квадруплекс представляет собой «выпетленную» структуру, контактные поверхности которой для MutL остаются доступными независимо от АТФ-индуцируемых перестроек. Это объясняет, почему высокая концентрация АТФ значительно меньше влияет на связывание белка с G4 и подтверждает его повышенное сродство к квадруплексам. Таким образом, в ходе этой части работы мы показали, что ngMutL обладает высоким сродством к G4, однако не проявляет способности к их гидролизу. Вероятно, взаимодействие MutL *N. gonorrhoeae* с G4-мотивом выполняет регуляторную функцию, потенциально блокируя доступ других белков, участвующих в инициации рекомбинации *pilE/pilS*, лежащей в основе уклонения патогена от иммунного ответа человека. Результаты исследования опубликованы в работе с участием автора диссертации [204].



**Рисунок 31.** Влияние АТФ на образование комплекса ngMutL с 76R/76M (синий) и 95G4/76M (желтый) (концентрация ДНК-фрагментов 20 нМ). Реакционные смеси (20 мкл) инкубировали в течение 10 мин в бане со льдом с 500 нМ ngMutL (на димер) в 20 мМ буфере HEPES (рН 8,0), содержащем 100 мМ КСl, 0,5 мг/мл БСА, 5 мМ MgCl<sub>2</sub> и 1 мМ ДТТ. Образцы анализировали в 6%-ном ПААГ в буфере ТАЕ в течение 2–3 ч при температуре 4 °С. Погрешности соответствуют 95%-ным доверительным интервалам (n = 3).

#### IV.4. Эффективность узнавания неканонической пары внутри GC-богатой последовательности при репарации «мисматчей» у прокариот

Полученные ранее результаты, демонстрирующие высокую эффективность связывания ngMutL с G-квадруплексами при отсутствии гидролитической активности, позволяют

предположить, что участие этого белка в процессах, опосредованных G4, выходит за пределы классической функции в системе MMR.

Взаимодействие гомологов MutS различных организмов с G4 довольно хорошо изучено, в том числе в работах нашей научной группы [192,205]. Известно, что все гомологи MutS обладают повышенным сродством к G4 по сравнению с природным субстратом («мисматчем»). В нашей лаборатории было обнаружено, что MutS из *E.coli* связывает модельный субстрат 95G4/76M в два раза более эффективно, чем контрольный 76-звенный дуплекс (76/76-GT), содержащий G/T-пару. Для гомолога из системы репарации «мисматчей» из *Cereibacter sphaeroides* данный эффект выражен сильнее:  $K_d^{app}$  MutS-ДНК (дуплекс с GT-парой) более, чем в 10 раз превышает  $K_d^{app}$ , характеризующую взаимодействие белка с G4-ДНК (Таблица 12). Взаимодействие с G4 наблюдали и для гомологов MutS человека [190,206].

**Таблица 12.** Значения кажущихся констант диссоциации комплексов MutS с модельными ДНК [192].

Белок MutS из	ДНК	$K_d^{app}$ , нМ
<i>Escherichia coli</i>	76/95G4	31±4
	76/76-GT	70±10
<i>Cereibacter sphaeroides</i>	76/95G4	23±3
	76/76-GT	360±60

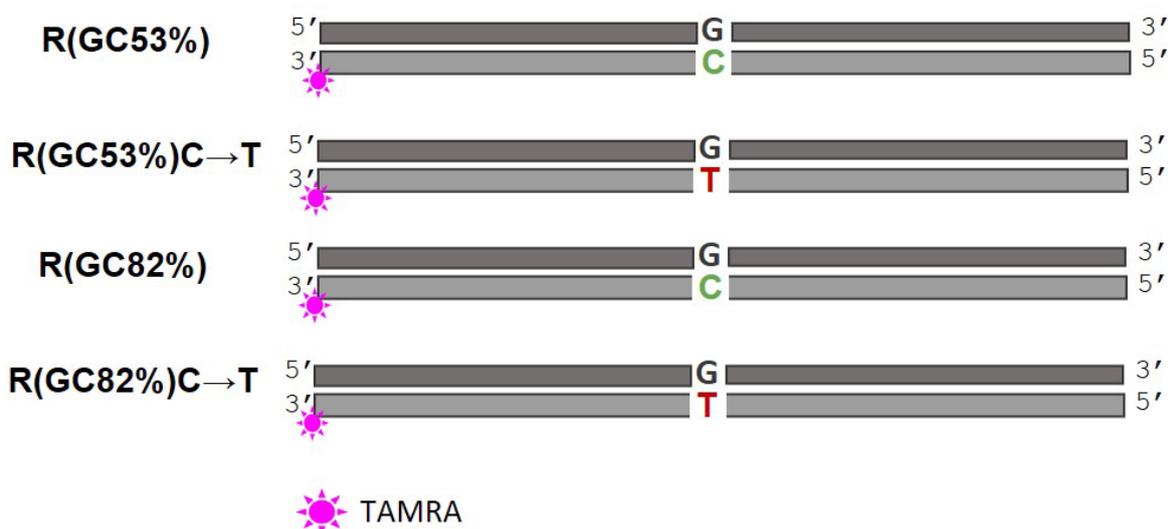
Таким образом, эффективное комплексообразование ngMutS с G4 не ставится под сомнение. Однако клеточные линии *N. gonorrhoeae*, содержащие ген мутантной формы MutS(Phe36Ala), который частично сохраняет способность связывать G-квадруплексы, но значительно хуже узнает «мисматчи», демонстрировали повышенную антигенную вариацию пилина [122]. Этот эксперимент свидетельствует в пользу заключения, что MMR-опосредованная регуляция антигенной вариации пилина происходит не за счет взаимодействия MutS с G4, а за счет спонтанного накопления мутаций. Тем не менее, авторы исследования подчеркивают, что MutS(Phe36Ala) лишь частично сохраняет способность связывать G4 по сравнению с MutS дикого типа, поэтому утверждение требует дополнительных доказательств.

Несмотря на регуляторные функции G4 эта структура может существовать в клетке ограниченное время в процессах, связанных с расхождением цепей ДНК. В остальном G4-мотивы

включены в дуплексную ДНК и представляют GC-богатые участки, которые обладают высоким мутагенным потенциалом. Такие участки склонны к накоплению «мисматчей» в ходе спонтанного дезаминирования цитозина и метилцитозина с образованием некомплементарных U/G- и T/G-пар. Последние в отсутствие репарации могут приводить к закреплению мутаций в следующем раунде репликации и формированию мутантного фенотипа, характерного не только для *N. gonorrhoeae*, но и других прокариот. Влияние GC-богатого состава ДНК-дуплекса на способность белка MutS узнавать «мисматч» ранее не изучалось.

В работе в качестве объекта исследования был выбран гомолог MutS из *Cereibacter sphaeroides* (*C. sphaeroides*, ранее – *Rhodobacter sphaeroides*, штамм 2.4.1). *C. sphaeroides* широко используется как модельный организм для изучения фотосинтеза и регуляции метаболизма у прокариот. Этот вид характеризуется высокой «пластичностью» генома: бактерия обладает двумя хромосомами и несколькими плазмидами, что позволяет быстро адаптироваться к изменяющимся условиям среды. Уникальной особенностью является способность *C. sphaeroides* переключаться между различными типами дыхания и фотосинтетического метаболизма: она может расти аэробно, анаэробно и в условиях фототрофии [207]. Для данной работы немаловажно, что геном некоторых штаммов *C. sphaeroides* (в том числе штамма 2.4.1) содержит около 70% GC-пар, что делает его интересным объектом для изучения репарационных систем в контексте устойчивых вторичных структур ДНК и специфики узнавания повреждений. Использование гомолога MutS именно из этого организма позволяет проверить гипотезу о влиянии GC-богатого окружения на узнавание «мисматчей» в условиях, когда сам белок эволюционно адаптирован к высокому GC-составу генома.

Для оценки эффективности узнавания «мисматча» в GC-богатом окружении ДНК использовали рекомбинантный белок MutS из *C. sphaeroides* (csMutS), который был выделен и очищен научным сотрудником НИИ ФХБ им. Белозерского МГУ к.х.н. Монаховой М.В. Модельные субстраты представляли собой 96-звенные ДНК-дуплексы, полученные в ходе гибридизации олигонуклеотидов ДНК (см. раздел III «Материалы и методы», таблица 7): с GC-составом 53% и G/T-парой (**R(GC53%)C→T**), с GC-составом 53% и с канонической C/T-парой (**R(GC53%)**), с GC-составом 82% и G/T-парой (**R(GC82%)C→T**) и с GC-составом 82% и C/T-парой (**R(GC82%)**). Их схематичное изображение приведены на рисунке 32: розовой звездочкой указана флуоресцентная метка TAMRA.



**Рисунок 32.** Схематичное изображение модельных 96-звенных ДНК-дуплексов с различным GC-составом (53% или 82%), содержащих или не содержащих G/T-пару.

Связывание ДНК белком csMutS проводили в присутствии 5 мМ MgCl<sub>2</sub> и 100 мМ KCl при 37 °С в течение 15 мин. В реакции использовали 1 мМ негидролизующий аналог АТФ (аденозин-5'-(β,γ-имидо)трифосфат, АТФγS) для предотвращения изменения конформации белка и, как следствие, его диссоциации с ДНК после связывания. Таким образом, мы исключили влияние ДНК последовательности на АТФазную активность csMutS, чтобы последнее не оказывало эффект на связывание белка с ДНК. ДНК-белковый комплекс анализировали методом «торможения» в 6%-ном ПААГ в неденатурирующих условиях в буфере TBE. Комплексообразование было охарактеризовано кажущимися константами диссоциации ( $K_d^{app}$ , нМ, таблица 13) с помощью трех моделей с последующим выбором наиболее подходящей по значению  $R^2$  (см. раздел III «Материалы и методы, Проверка ДНК-связывающей активности белка MutS»). Сравнение параметров связывания csMutS с ДНК различного GC-состава показывает, что характер взаимодействия определяется не только наличием «мисматча», но и нуклеотидным окружением. Для субстрата с GC-составом 53% наличие некомплементарной пары существенно повышает эффективность комплексообразования. Если для немодифицированной ДНК R(GC53%) связывание описывается моделью одного участка с относительно высокой кажущейся константой диссоциации ( $349 \pm 46$  нМ), то введение замены C→T приводит к появлению более специфического взаимодействия. Применение модели двух сайтов позволяет разделить вклад специфического узнавания и неспецифического взаимодействия с ДНК: одна из полученных констант ( $39 \pm 5$  нМ) характеризует узнавание «мисматча», тогда как вторая ( $118 \pm 14$  нМ) отражает взаимодействие с фланкирующими участками. Это согласуется с представлением о том, что MutS первоначально сканирует ДНК в режиме неспецифического связывания и стабилизируется при обнаружении искажений двойной

спирали. Модель Хилла, дающая интегральную оценку взаимодействия ( $154 \pm 18$  нМ), не позволяет разделить эти компоненты, однако подтверждает повышение общего сродства к субстрату с «мисматчем».

**Таблица 13.** Значения кажущихся констант диссоциации комплексов MutS с модельными ДНК.

ДНК	$K_d^{app}$ , нМ	$R^2$	Модель
<b>R(GC53%)C→T</b>	39±5 118±14	0,999	Два участка связывания
	154±18	0,999	Связывание по Хиллу (n 1,3±0,1)
<b>R(GC53%)</b>	349±46	0,997	Один участок связывания
<b>R(GC82%)C→T</b>	72±9	0,993	Связывание по Хиллу (n 1,88±0,27)
<b>R(GC82%)</b>	88±11	0,996	Связывание по Хиллу (n 1,76±0,14)

Иная картина наблюдается при GC-составе ДНК, равном 82%. Как при наличии, так и при отсутствии «мисматча» взаимодействие корректно описывается только моделью Хилла, а различия в кажущихся константах диссоциации ( $72 \pm 9$  и  $88 \pm 11$  нМ) находятся в пределах погрешности. Это указывает на то, что высокий GC-состав снижает эффективность узнавания «мисматча» белком csMutS. Полученные результаты хорошо согласуются с современной моделью непрямого узнавания, согласно которой MutS сканирует ДНК, проверяя ее гибкость и способность к изгибу на протяжении примерно двух витков спирали (~20 п.н.) [208]. Высокая жесткость GC-богатой ДНК, обусловленная тремя водородными связями в GC-парах и усиленным межплоскостным взаимодействием, даже при наличии «мисматча» не дает достаточного искажения структуры для дискриминации, и белок MutS связывается с любым GC-богатым дуплексом примерно с одинаковой эффективностью.

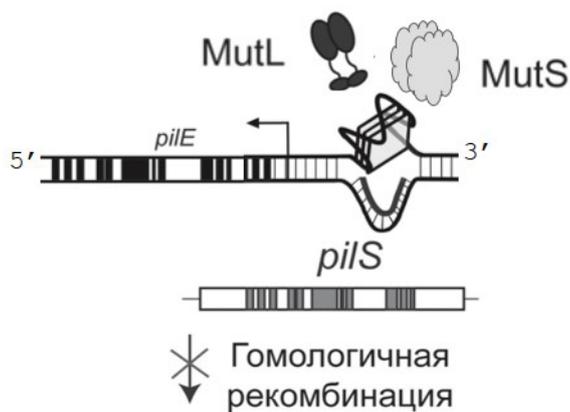
Учитывая высокую консервативность аминокислотной последовательности и структуры гомологов MutS, можно предположить, что GC-богатый состав ДНК может снижать эффективность репарации «мисматчей» в целом.

\*\*\*

В рамках этой части работы нами были выявлены новые аспекты функционирования ключевых белков метилнезависимой системы репарации «мисматчей» прокариот на примере

ngMutL и csMutS. Во-первых, установлены ранее не описанные особенности ngMutL: высококонсервативные остатки цистеина в белке играют важную роль при взаимодействии ngMutL с  $\beta$ -субъединицей ДНК-полимеразы III, которая усиливает процессивность к ДНК. Во-вторых, впервые показано высокое сродство ngMutL к G4 параллельной топологии, образованного G-богатой регуляторной последовательностью перед промотором *pilE*. При этом, ngMutL не гидролизует фосфодиэфирные связи внутри G4, но при определенных условиях одноцепочечные разрывы наблюдаются во фланкирующих областях до и после G4. Это позволяет рассматривать ngMutL как потенциальный регулятор процесса антигенной вариации пилей. В-третьих, впервые продемонстрировано, что в GC-богатом контексте csMutS теряет способность дискриминировать ДНК, содержащую и не содержащую «мисматч»: белок связывается с ДНК с высокой эффективностью ( $K_d \sim 80$  нМ) независимо от наличия повреждения, тогда как в GC-умеренном контексте (53%) дискриминация четко выражена ( $K_d$  39 нМ против 350 нМ). Данный результат получен на примере гомолога из бактерии *Cereibacter sphaeroides* (штамм 2.4.1), хромосомы которой имеют GC-состав 69%, что делает этот белок эволюционно адаптированным к функционированию в GC-богатом окружении. Тот факт, что даже в оптимальных условиях наблюдается потеря дискриминации «мисматчей», указывает на фундаментальный характер ограничений MMR в GC-богатых областях генома. Совокупность этих данных показывает, что функционирование MMR существенно зависит от локальной последовательности и структуры ДНК, демонстрируя необходимость дополнения классических моделей репарации с учетом влияния GC-богатых регионов и стабилизации G4 в определенных условиях. Полученные результаты не только уточняют молекулярные механизмы работы систем репарации в прокариотах, но и открывают перспективы для изучения их роли в изменчивости патогенов (Рис. 33).

### Альтернативная роль белков MMR в контексте G4



### Накопление мутаций из-за снижения эффективности MMR



**Рисунок 33.** Предполагаемые механизмы функционирования белков MMR в GC-богатых участках генома.

#### IV.5. Влияние GC-богатой последовательности на эффективность эксцизионной репарации эукариот

Биоинформатические исследования показали, что в ДНК человека идентифицировано 95 семейств PQS [209], часто локализованных в регуляторных элементах генома. Стабилизация G4-структур даже на доли секунды, может модулировать репликацию ДНК, регулировать экспрессию генов, рекомбинацию, и репарацию. Считается, что PQS способствуют нестабильности генома, блокируя продвижение репликативной вилки и повышая вероятность мутаций, делеций и рекомбинационных событий [210]. Высокая мутагенность PQS связана как с вероятностью стабилизации G4, так и с повышенной чувствительностью GC-богатых ДНК к окислительным повреждениям. Примером PQS-содержащего регуляторного участка в геноме человека является GC-богатый промотор гена обратной транскриптазы теломеразы человека (*hTERT*), отвечающей за сохранение длины теломер в стволовых, половых и раковых клетках. Активация транскрипции гена в других типах клеток ассоциируется с развитием онкологии.

Некодирующая матричная цепь ДНК промотора *hTERT* характеризуется высокой концентрацией GC-повторов. Интересующая нас область, локализованная в пределах позиций -167 и -99 до старт-кодона, способна при определенных условиях образовывать тандемный G-квадруплекс параллельной топологии (см. раздел II.3.5. «Взаимодействие белков эксцизионной репарации ДНК человека с G-квадруплексами»). Данная последовательность промотора *hTERT* включает 68-звенный GC-богатый фрагмент с 12 G-трактами [205]. В промоторе *hTERT* при различных типах рака часто обнаруживаются мутации, преимущественно замены G→A в матричной цепи в позициях -124 и -146, а также реже встречающаяся двойная замена в позициях -138/-139 относительно стартового кодона. Считается, что эти изменения создают новые участки связывания факторов ETS, усиливая транскрипцию *hTERT* [211].

Рассматриваемый 68-звенный участок матричной цепи промотора *hTERT* содержит 66% остатков dG, а кодирующей – 66% остатков dC, поэтому вероятность возникновения в нем поврежденных оснований вследствие окислительного стресса или спонтанного дезаминирования высока. За восстановление таких повреждений в ДНК отвечает система эксцизионной репарации оснований (BER), которая реализуется через удаление дефектного основания специфической ДНК-гликозилазой с формированием AP-сайта, последующее расщепление которого обеспечивается AP-эндонуклеазой APE1. Влияние GC-богатого состава на эффективность работы BER остается мало изученным, что обуславливает актуальность исследования. Нами было высказано предположение, что причиной закрепления указанных выше тканеспецифичных мутаций в промоторе *hTERT* может являться сниженная активность ключевого фермента BER - APE1 в GC-богатой области этого промотора. Таким образом, целью второй части работы

являлось изучение взаимодействия фермента APE1 человека с модельными фрагментами промотора *hTERT*, включающими замены G→F в матричной цепи (F – стабильный циклический аналог AP-сайта, 1,2-дидезоксирибоза) и C→F в кодирующей цепи в соответствующих «драйверным» мутациям положениях: -124, -146, -138/-139. Тетрагидрофурановый F-аналог AP-сайта используют в модельных системах для изучения BER как адекватную замену нестабильному AP-сайту, поскольку F химически стабилен и имитирует отсутствие основания в ДНК, но не подвергается спонтанному гидролизу [212]. Основной задачей работы являлось выявление влияния положения модификаций на гидролитическую и ДНК-связывающую активность APE1 в различных условиях.

#### **IV.5.1. Консервативные последовательности PQS в промоторах *TERT* млекопитающих**

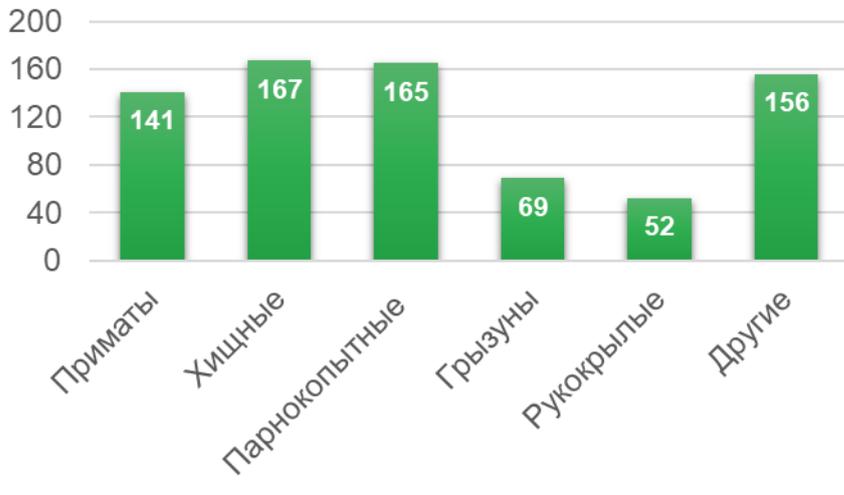
С применением биоинформатических методов был проведен сравнительный анализ областей с PQS в промоторах гена *TERT* у 141 вида млекопитающих, относящихся к 20 отрядам, из базы данных NCBI Orthologs (<https://www.ncbi.nlm.nih.gov/>, дата обращения: 13 августа 2022 г.). Мы исследовали последовательности промоторного региона, охватывающего область длиной 1000 п.н. до предполагаемого сайта начала транскрипции (TSS). Для поиска потенциальных G4-мотивов в промоторных областях *TERT* использовали QGRS Mapper (см. раздел III. «Материалы и методы»). Сервис позволяет осуществить поиск G4 по фиксированному шаблону (G<sub>3+</sub> L<sub>1-30</sub> G<sub>3+</sub> L<sub>1-30</sub> G<sub>3+</sub> L<sub>1-30</sub> G<sub>3+</sub>, где G<sub>n</sub> соответствует G-трактам с n последовательно расположенными dG (n ≥ 3), а L – полупроизвольная последовательность из 1–30 нуклеотидов, соответствующая петлям квадруплекса). Такой подход делает результаты удобными для межвидового сравнения.

Исследование показало, что некодирующая цепь обогащена G4-мотивами в большей степени, чем кодирующая. Причем данный эффект наиболее выражен в областях на расстоянии 0–400 п.н. от TSS, в то время как удаленные от TSS участки содержат значительно меньше таких структур. Суммарное количество найденных G4-мотивов оказалось наименьшим в промоторных областях *TERT* грызунов и рукокрылых, тогда как у приматов, хищных, парнокопытных и других это значение было значительно выше (Рис. 34А). Несмотря на то, что PQS распространены в промоторных зонах *TERT* представителей разных отрядов млекопитающих, доля консервативных

G4-мотивов варьируется. Большинство из них локализованы преимущественно на некодирующей цепи, при этом распределение оказалось специфичным для каждого отряда (Рис. 34Б).

(А)

### Количество G4 мотивов



(Б)

### Количество консервативных G4-мотивов



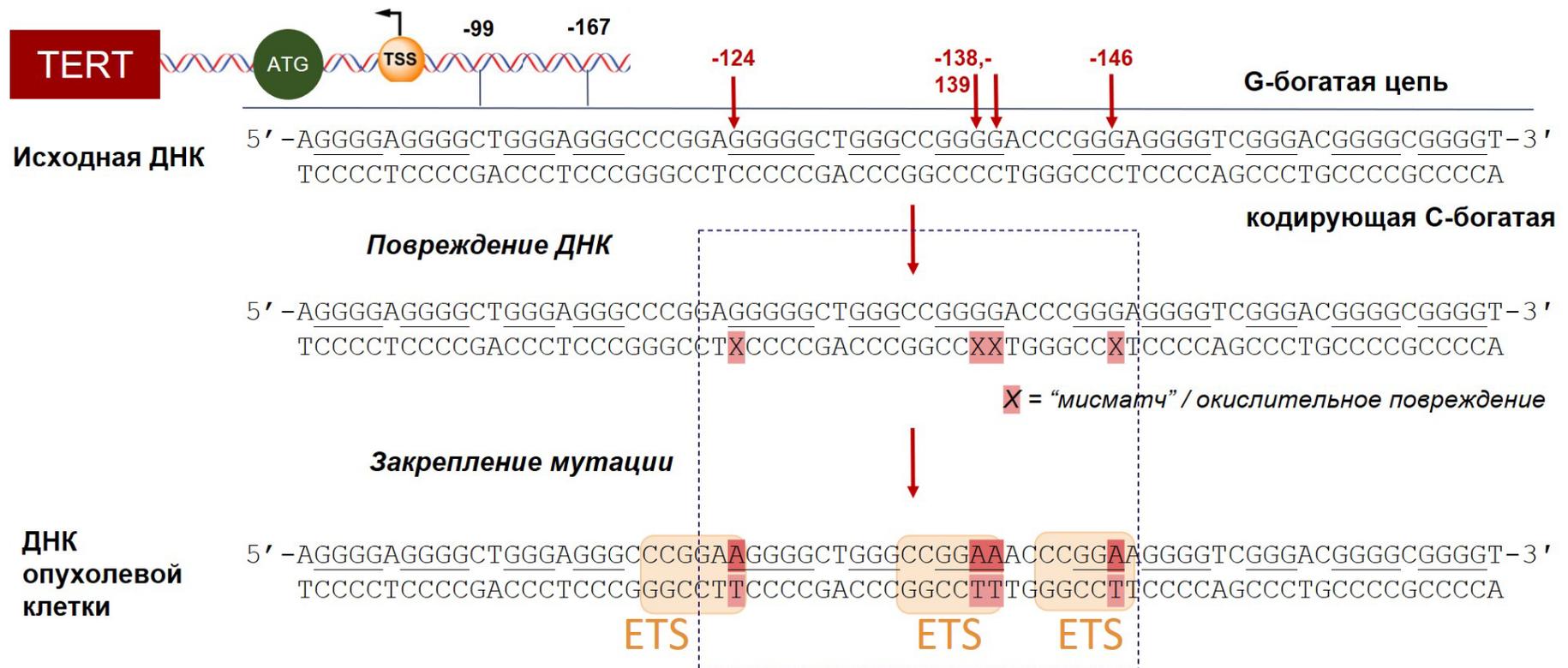
**Рисунок 34.** Количество G4-мотивов в последовательностях промоторных регионов *TERT* разных отрядов. (А) Количество найденных мотивов G4 для всех изученных отрядов млекопитающих. (Б) Число консервативных G4-мотивов в кодирующих (синие столбцы) и не кодирующих, матричных (коричневые столбцы) цепях промоторных областей *TERT* в зависимости от типа отряда млекопитающих.

Таким образом, G4-мотивы промоторных областей *TERT* сохраняются в ходе эволюции млекопитающих, но их частота, распределение по цепям и длина имеют специфические для отряда особенности. Эволюционная консервативность и локализация PQS в области промотора *TERT* приматов служит косвенным подтверждением функциональной значимости этих последовательностей. К ним относится интересующая нас 68-звенная GC-богатая область промотора гена обратной транскриптазы теломеразы человека. Исследование опубликовано в работе с участием автора диссертации [213].

#### IV.5.2. Характеристики G-квадруплексной структуры, сформированной G-богатой последовательностью промотора *hTERT*

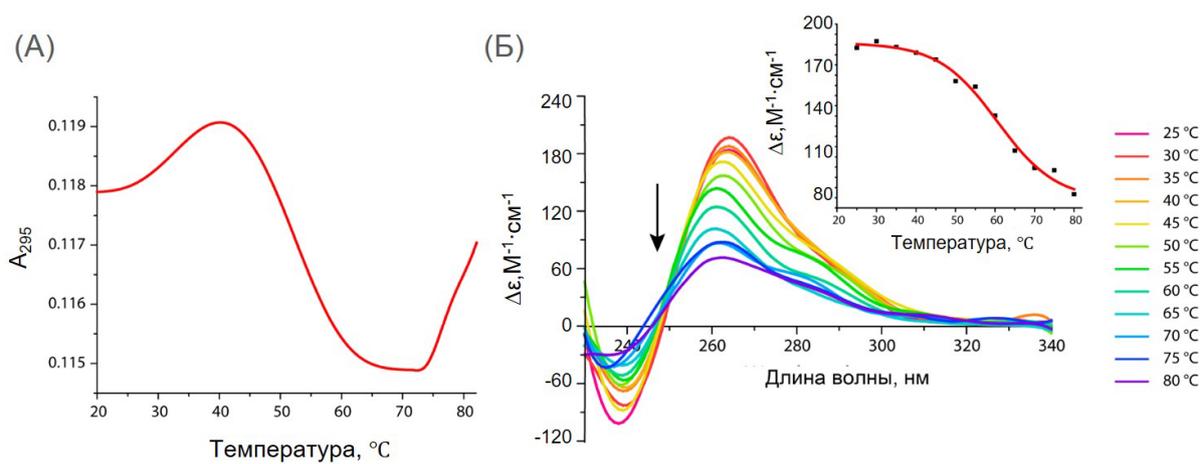
Аномальная экспрессия *TERT* наблюдается в случае 85–90% от всех изученных на сегодняшний день видов рака. Этиология онкологических процессов основана на накоплении различных мутаций, локализованных как в соматических клетках, так и в зародышевых линиях. Функционально-генетические исследования выявили несколько точечных мутаций в области промотора *hTERT* (примерно от -180 до +1 от стартового кодона), получивших название «драйверных» (Рис. 35).

Замены С→Т в кодирующей цепи промотора *hTERT* чаще всего встречаются в позициях -124 и -146 п.н. относительно стартового кодона АТГ гена *hTERT*. Эти мутации наблюдаются в 60–80% случаев уротелиальной карциномы, 71% случаев меланомы, 83% случаев глиобластомы и различных других видах рака. Менее распространенная двойная замена СС→ТТ в позициях -138, -139 п.н. от стартового кодона встречается преимущественно в геномах клеток опухолей кожи. Онкологически значимые замены G→А (аналогичные описанным выше, но локализованные в G-богатой некодирующей цепи) кластеризуются в 68-нуклеотидном GC-богатом регионе промотора *hTERT*, содержащем двенадцать G-трактов из трех или более последовательных dG, которые в определенных условиях позволяют формировать G-квадруплексные структуры [205]. Для анализа структурных особенностей G-богатого фрагмента промотора *hTERT* в качестве модельной системы был выбран одноцепочечный 96-нуклеотидный фрагмент некодирующей цепи промотора *hTERT*, содержащий 68-звенную G-богатую последовательность (включая двенадцать последовательных G-трактов) и фланкирующие 14-звенные области (обозначен 96-G4, см. III. «Материалы и методы», таблица 8).



**Рисунок 35.** Специфические для опухолевых клеток «драйверные» мутации в первичной структуре промотора *hTERT*, которые генерируют *de novo* сайты связывания факторов транскрипции ETS. Мутации локализованы в 68-звенном GC-богатом участке промоторного региона, содержащем двенадцать G-трактов, потенциально способных образовывать G4 в определенных условиях (G-тракты подчеркнуты).

Спектры УФ-поглощения при 295 нм выявили характерные признаки формирования G-квадруплексной структуры в 96-звенном фрагменте промотора *hTERT* (Рис. 36А). «Плавление» структуры (зависимость оптической плотности от температуры) сопровождалось гипохромным эффектом и сигмоидальной формой кривой, при этом температура плавления составила  $54 \pm 1$  °С в буфере, содержащем 20 мМ КСl, после длительной процедуры «отжига». Увеличение поглощения на низкотемпературных и высокотемпературных участках связано с разрушением взаимодействий в 14-звенных одноцепочечных фланкирующих областях. При снижении концентрации КСl до 5 мМ зависимости оптической плотности от температуры в диапазоне 20–85 °С не наблюдалось.



**Рисунок 36.** Оценка термической стабильности и топологии структуры G4, образованной одноцепочечным 96-звенным фрагментом промотора *hTERT* (96-G4), спектральными методами. **(А)** Зависимость изменения УФ-поглощения при 295 нм от температуры в диапазоне от 20 до 85 °С. **(Б)** Спектры КД, зарегистрированные при различной температуре, возрастающей по направлению стрелки; значения температуры варьировались от 25 до 80 °С с интервалом 5 °С. Все измерения проводили в 8 мМ калий-фосфатном буфере (рН 7,1), содержащем 20 мМ КСl.

Для подтверждения формирования G4 и определения его топологии регистрировали КД-спектры в 8 мМ калий-фосфатном буфере (рН 7,1), содержащем 20 мМ КСl (Рис. 36Б). Спектры демонстрировали характерные признаки параллельной топологии G4: положительный пик при 265 нм и отрицательный – около 240 нм, что исключает наличие анти-параллельной конформации. С помощью КД-спектров также была охарактеризована температура плавления олигонуклеотида 96-G4 ( $T_{пл}$  около 59 °С). Близкие значения этой величины, полученные методом УФ-и КД-спектроскопии, подтверждают надежность определения параметров структуры G4.

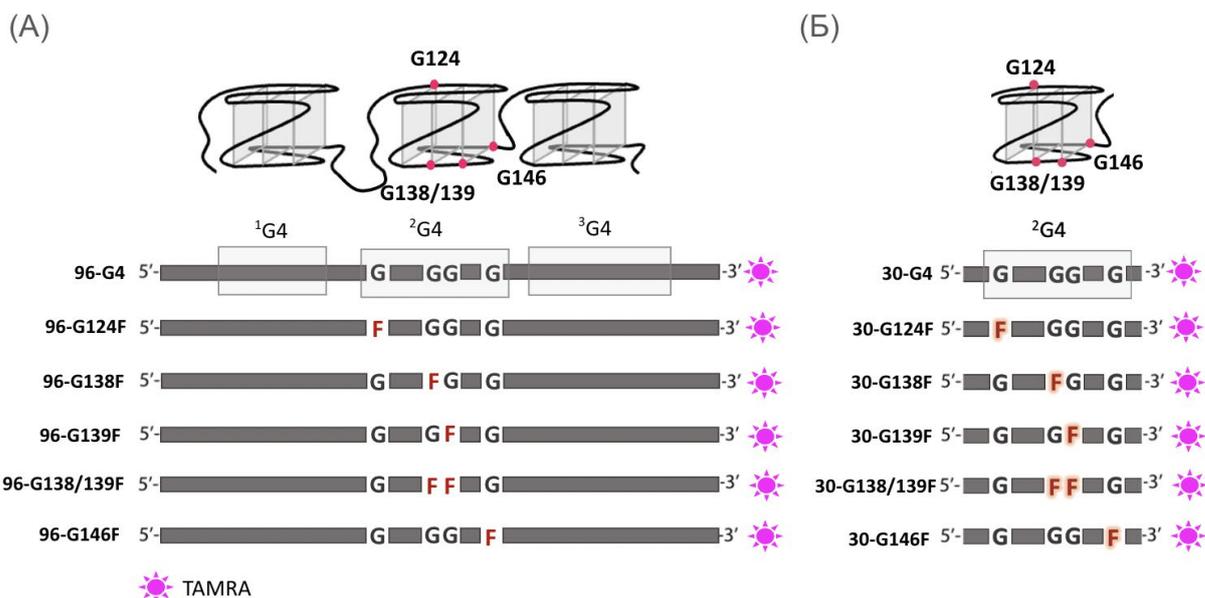
Дополнительные сведения о формировании G-квадруплексов были получены в нашей лаборатории методом химического футпринтинга с использованием диметилсульфата (DMS), метилирующего N7-положение гуанина [205]. Так как этот атом недоступен в составе G4, снижение реакционной способности остатков гуанина по отношению к DMS служит показателем их участия в формировании квадруплексной структуры. Действительно, в одноцепочечном

фрагменте промотора *hTERT* было показано значительное снижение «чувствительности» остатков гуанина к модифицирующему агенту, входящих в квадруплексные мотивы G1, G2 и G3 (Рис. 36А), по сравнению с их двуцепочечными аналогами. Эти данные свидетельствуют о вовлечении гуанинов из промотора *hTERT* в образование Хугстиновских пар и подтверждают формирование тандемной G-квадруплексной структуры.

#### IV.5.3. Дизайн модельных фрагментов промотора *hTERT*

Апуриновые/апиримидиновые сайты представляют собой одни из наиболее распространенных повреждений ДНК в клетках человека. Особый интерес вызывает роль AP-сайтов в GC-богатых регионах, способных формировать G4. Высокая окисляемость остатков гуанина делает GC-богатые последовательности предрасположенными к образованию AP-сайтов, что может изменять стабильность G4. С другой стороны, формирование G-квадруплексов может ингибировать активность APE1, отвечающей за репарацию AP-сайтов. Это в свою очередь приведет к накоплению повреждений и будет способствовать закреплению мутаций.

Мы оценили влияние тандемной G4-структуры, сформированной 96-звенными одноцепочечными фрагментами промотора *hTERT*, на эффективность функционирования APE1. В качестве субстратов были использованы 96-звенные фрагменты, содержащие замены G→F в положениях, соответствующих «драйверным» (-124, -138, -139, -146 и -138/139) и флуоресцентную метку TAMRA на 3'-конце (Рис. 37А, полная последовательность приведена в разделе III. «Материалы и методы» в таблице 9). В качестве контрольных ДНК были использованы олигонуклеотидные последовательности, представляющие собой: а) 96-звенный фрагмент промотора *hTERT*, не содержащий модификацию F (96-G4, таблица 8, рис. 37А), б) произвольная олигонуклеотидная последовательность, содержащая модификацию (96-RF, таблица 8). С учетом того, что все повреждения локализуются в G-трактах 5, 7 и 8 (центральный <sup>2</sup>G4-мотив, рис. 37), дополнительно были созданы укороченные по сравнению с 96-звенными 30-звенные ДНК-фрагменты без фланкирующих мотивов <sup>1</sup>G4 и <sup>3</sup>G4 (содержащие только центральный <sup>2</sup>G4-мотив) с флуоресцентной меткой TAMRA на 3'-конце с аналогичными заменами G→F (Рис. 37Б). Эти укороченные модели позволяют оценить влияние соседних немодифицированных G4 на структуру субстратов и активность APE1. В качестве контроля использовали 30-звенные фрагменты с интактными G4-мотивами (30-G4) и соответствующие фрагменты с F, но неспособные образовывать G4 (30-RF) (таблица 8).

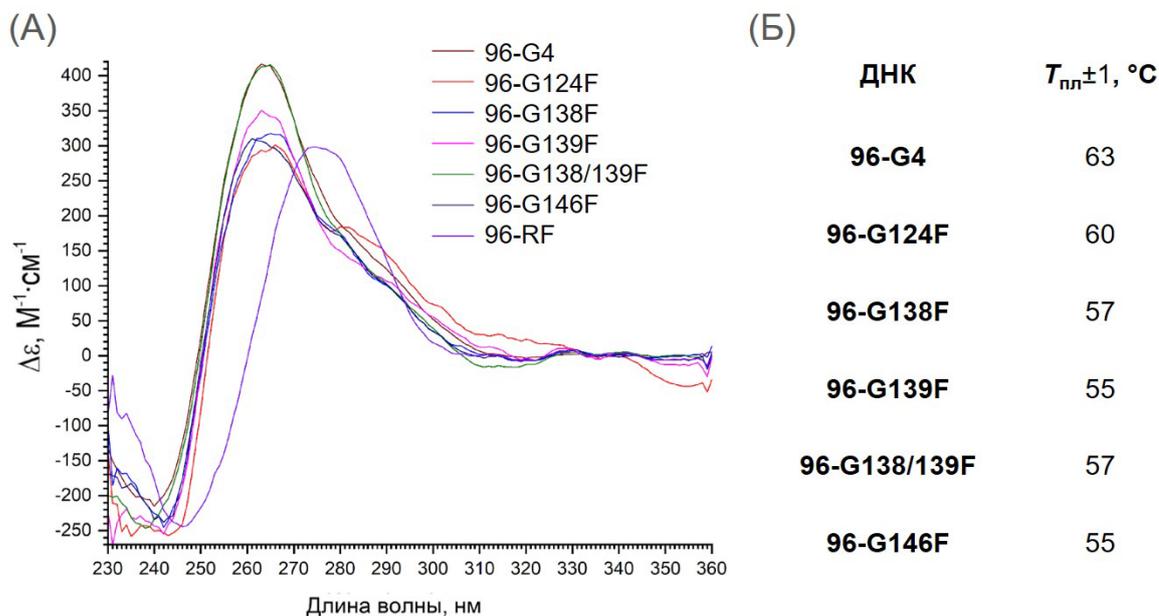


**Рисунок 37.** Схематическое изображение синтетических одноцепочечных ДНК, идентичных последовательности G-богатой цепи промотора *hTERT* и предполагаемая структура образуемых промотором *hTERT* последовательных параллельных квадруплексов. **(А)** Модельные 96-звенные фрагменты, содержащие замены G→F в положениях, соответствующих «драйверным». **(Б)** 30-звенные олигонуклеотиды, содержащие только центральный мотив <sup>2</sup>G4 без фланкирующих мотивов <sup>1</sup>G4 и <sup>3</sup>G4 с аналогичными заменами G→F. Все олигонуклеотиды содержали флуоресцентную метку TAMRA на 3'-конце (розовая звездочка).

#### IV.5.4. Влияние F-сайтов на топологию и термическую стабильность G-квадруплексной структуры, формируемой фрагментом промотора *hTERT*

Топологию и термическую стабильность G4-структур, сформированных 96-звенными одноцепочечными ДНК с заменами G→F в различных позициях, анализировали с помощью спектроскопии КД в 20 мМ буфере Трис-НСl (рН 7,5), 1 мМ ЭДТА и 100 мМ КСl (буфер, используемый в дальнейшем для исследования АРЕ1-индуцированного гидролиза модельных ДНК). Фрагмент 96-G4 демонстрировал КД-спектр с максимумом около 265 нм и минимумом около 240 нм, что характерно для G4 параллельной топологии. КД-спектр 96-звенного олигонуклеотида без G4-мотива (96-RF), но с модификацией, показывал положительный пик при 280 нм, что указывает на отсутствие G4-структуры (Рис. 38А). Спектры КД 96-звенных модифицированных последовательностей промотора *hTERT* показали, что замены G→F внутри G4-мотива способствуют формированию антипараллельной (гибридной) топологии G4, проявляющейся положительным «плечом» при длинах волн >280 нм; единственным исключением был 96-G138/139F, демонстрирующий параллельную топологию. Значения  $T_{пл}$  модифицированных аналогов 96-звенного *hTERT* промотора (Рис. 38Б), полученные на основании температурной зависимости молярной эллиптичности при 265 нм, оказались ниже по сравнению с  $T_{пл}$  для немодифицированного фрагмента 96-G4 (разница составила от 3 до 8 °С). Дестабилизация G4 была менее выражена при замене G124F, вероятно, из-за возможной

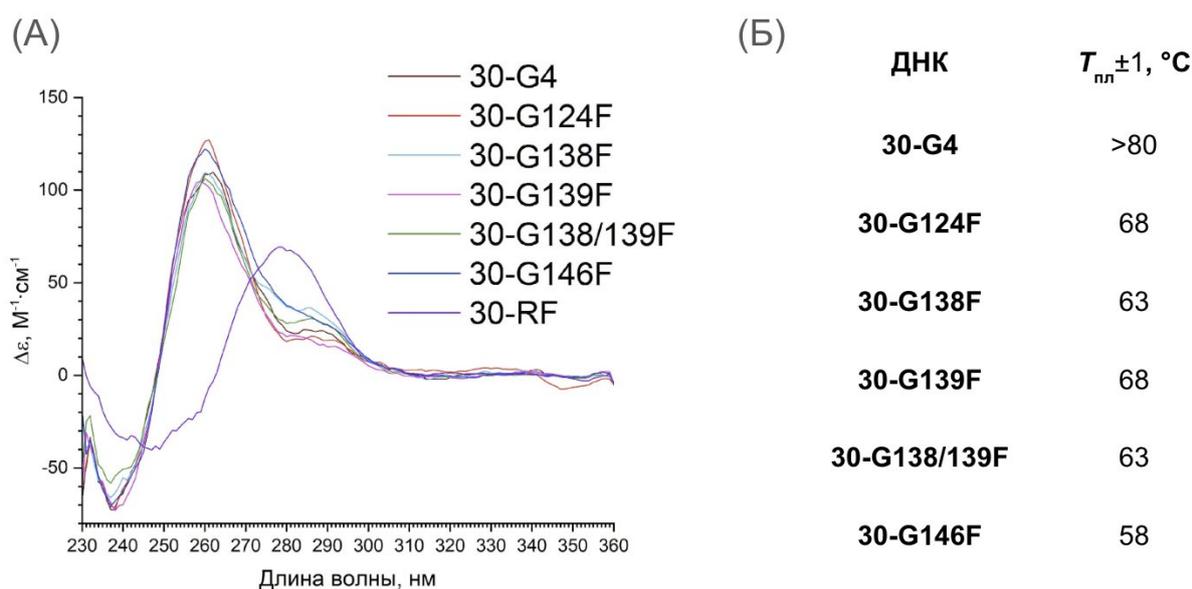
стабилизации структуры за счет привлечения соседнего dG из длинного G-тракта с пятью остатками dG, тогда как замена G146F наоборот оказывала более выраженное дестабилизирующее влияние на G4, предположительно, из-за локализации модификации в G-тракте, состоящем всего из трех остатков dG.



**Рисунок 38.** Влияние аналога AP-сайта (F) в составе модельных олигонуклеотидов на топологию и термическую стабильность G4 *hTERT*. (А) Спектры КД 96-звенных G-богатых олигонуклеотидов промотора *hTERT*, содержащих замены G→F при 30 °C в 20 мМ буфере Трис-НСl (рН 7,5), 1 мМ ЭДТА и 100 мМ КСl. (Б) Значения  $T_{пл}$ , определенные из данных температурных зависимостей молярной эллиптичности при 265 нм для G-квадруплексных структур соответствующих модельных фрагментов.

Для выяснения влияния фланкирующих G4 на стабильность центрального квадруплекса были изучены спектры КД укороченных 30-звенных аналогов 96-звенных олигонуклеотидов, содержащих только центральный <sup>2</sup>G4. КД-спектр 30-звенного контрольного олигонуклеотида без G4-мотива (30-RF) с F демонстрировал положительный пик при 280 нм, указывая на отсутствие G4-структуры (Рис. 39А). Спектры 30-G4 и 30-G139F однозначно свидетельствуют о формировании параллельной G4-структуры. Все остальные F-содержащие 30-звенные варианты с G4-мотивом центрального квадруплекса *hTERT* давали положительные «плечи» при длинах волн >280 нм, что указывает на присутствие антипараллельных G4 (Рис. 39А). Кривые «плавления», построенные на основе КД-спектров при различных температурах, показали, что 30-звенные олигонуклеотиды формируют значительно более стабильные G4-структуры по сравнению с 96-звенными аналогами; разрушение немодифицированного 30-G4 было неполным даже при 80 °C. В то же время аналоги 30-G4 с F демонстрировали заметно меньшую термостабильность относительно немодифицированного 30-G4 с  $\Delta T_{пл}$  от >12 до >22 °C (Рис. 39Б). Результаты КД показывают, что величина дестабилизирующего эффекта и изменения

топологии  ${}^2G4$  в 30-звенном олигонуклеотиде зависят от положения аналога AP-сайта внутри G4. Замена dG на F в наиболее коротком трехзвенном G-тракте (30-G146F) оказала наибольшее дестабилизирующее влияние ( $\Delta T_{пл} > 22$  °C) аналогично эффекту, наблюдаемому для 96-G146F. Замены G→F в первой или последней позиции 4- или 5-звенных G-трактов (30-G139F и 30-G124F) вызывали меньшую дестабилизацию ( $\Delta T_{пл} > 12$  °C), так как поврежденные G-тракты сохраняли как минимум три последовательных гуанина для формирования G-тетрад. Замена dG на F в середине 4х-нуклеотидного G-тракта (30-G139F) или двойная замена (30-G138/139F) оставляла только два последовательных гуанина для формирования тетрад, что сильнее дестабилизировало  ${}^2G4$  ( $\Delta T_{пл} > 17$  °C).



**Рисунок 39.** Влияние положение аналога AP-сайта (F) на топологию и термостабильность центрального  ${}^2G4$  *hTERT*. (A) Спектры КД 30-звенных фрагментов промотора *hTERT*, содержащих замены G→F при 25 °C в 20 мМ буфере Трис-НСl (рН 7,5), 1 мМ ЭДТА и 100 мМ КСl. (Б) Значения  $T_{пл}$ , рассчитанные на основе зависимостей молярной эллиптичности от температуры для 30-звенных олигонуклеотидов при 265 нм.

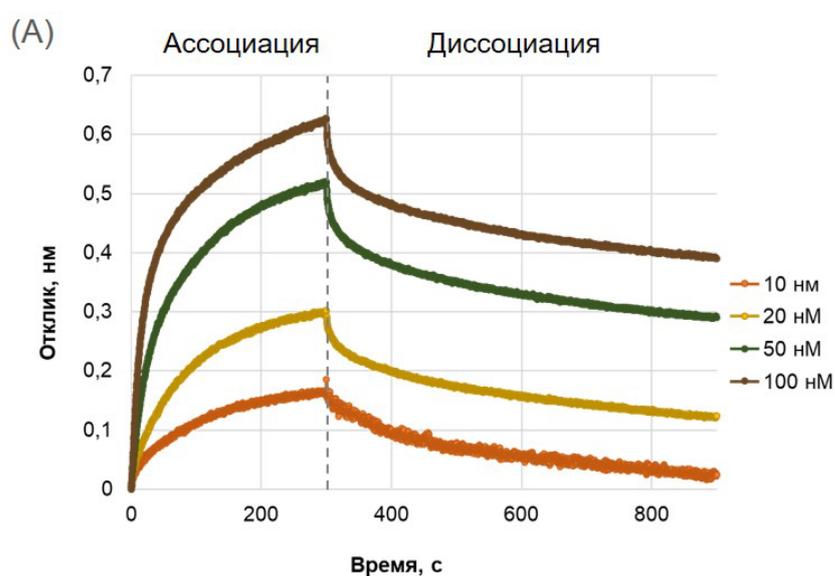
Таким образом, G→F в G-трактах центрального  ${}^2G4$  *hTERT* промотора не разрушают структуру квадруплекса полностью, но оказывают дестабилизирующее влияние, которое зависит от положения модификации. Наличие фланкирующих G4 мотивов ( ${}^1G4$  и  ${}^3G4$ ) снижают этот эффект, способствуя стабилизации tandemной структуры.

#### IV.5.5. Влияние G4-мотивов промотора *hTERT* на функционирование APE1

##### IV.5.5.1 Эффективность связывания APE1 с 96-звенными фрагментами промотора *hTERT*

Для изучения взаимодействия APE1 с модельными субстратами с последовательностью GC-богатой области промотора *hTERT* в реальном времени была использована интерферометрия

слоя биомолекул (BLI). Препарат фермента был любезно предоставлен академиком Жарковым Д.О. (Институт химической биологии и фундаментальной медицины СО РАН, г. Новосибирск). Фермент, содержащий дополнительную последовательность с шестью остатками гистидина, в концентрации 3,5 мкМ иммобилизовали на поверхности биосенсора с Ni-NTA-агарозой в течение 5 мин. Ассоциацию фермента с 96-звенными олигонуклеотидами, содержащими и не содержащими G4 и замену/ы G→F, регистрировали в течение 300 с, затем проводили диссоциацию комплекса и анализировали в том же буфере на протяжении 600 с (см. «Материалы и методы», Рис. 40А).



(Б)

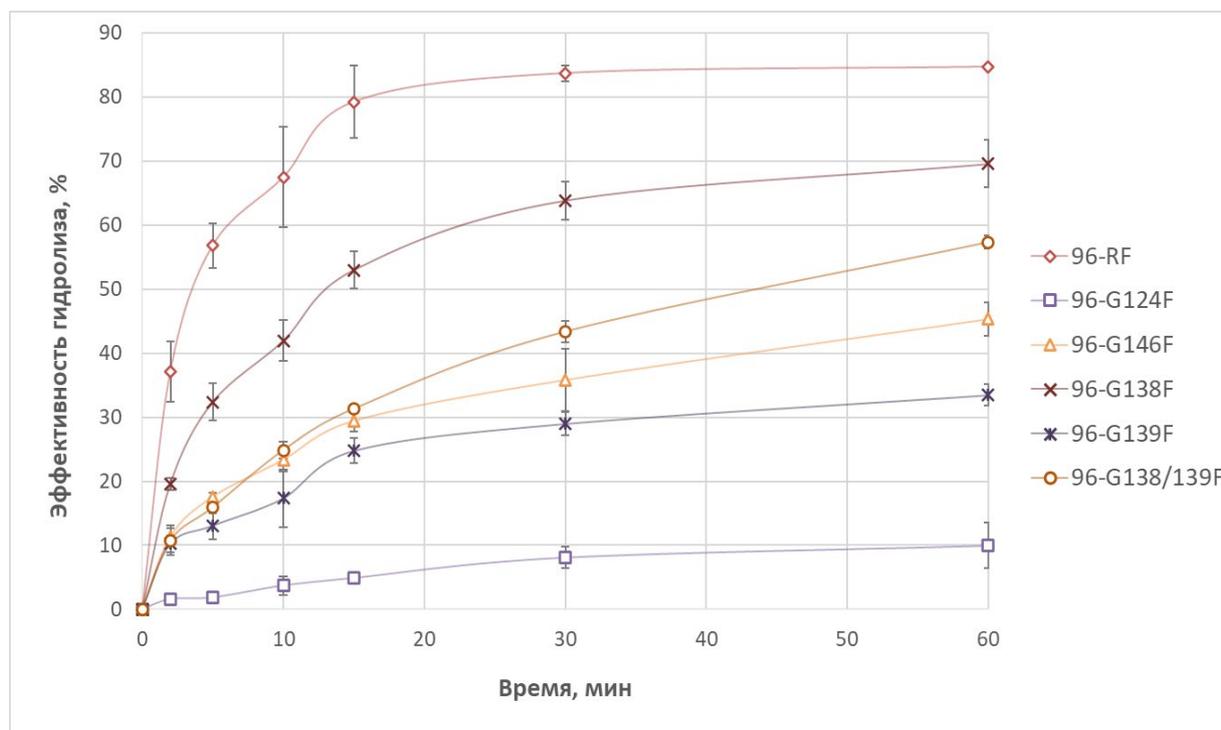
ДНК	$K_d$ , нМ
96-G4	$4,2 \pm 1,4$
96-G124F	$2,3 \pm 1,8$
96-G138F	$3,6 \pm 1,1$
96-G139F	$3,2 \pm 0,4$
96-G138/139F	$2,8 \pm 0,8$
96-G146F	$3,2 \pm 1,1$
96-RF	$56 \pm 3$

**Рисунок 40.** Связывание фермента APE1 с 96-звенными фрагментами промотора *hTERT*, содержащими и не содержащими G4 и замену/ы G→F. (А) Репрезентативные кривые ассоциации-диссоциации белково-нуклеинового комплекса, полученные методом измерения интерференции биослоя для 96-G138F. (Б) Константы диссоциации комплексов APE1-ДНК, определенные методом измерения интерференции биослоя с помощью программного обеспечения BLItz Pro 1.3.0.5 software (ForteBio).

Связывание APE1 проходило эффективно с 96-G124F, 96-G138F, 96-G139F, 96-G138,139F и 96-G146F, при этом значения констант диссоциации ( $K_d$ ) составляли около 3 нМ и не зависели от положения F. Для немодифицированного 96-G4 комплекс с APE1 характеризовался схожим значением  $K_d$  ( $4,2 \pm 1,4$  нМ). Напротив, сродство фермента к 96-RF, лишённого G4-мотива, было заметно более слабым:  $K_d$  оказалась на порядок выше по сравнению с субстратами, содержащими G4-мотивы (Рис. 40Б).

#### IV.5.5.2 Эффективность гидролиза 96-звенных фрагментов промотора *hTERT* белком APE1

Для оценки влияния G4-структуры промотора *hTERT* на эндонуклеазную активность APE1 в качестве субстратов использовали описанные ранее одноцепочечные 96-звенные олигонуклеотиды с G4-мотивом промотора *hTERT* с заменами G→F (Рис. 41). Реакцию проводили при 37 °С в 20 мМ буфере Трис-НСl (рН 7,5), содержащем 100 мМ КСl в присутствии препарата APE1 (New England Biolabs, са# M0282S) в количестве 10 ед. акт., согласно протоколу, что соответствовало 36 нМ концентрации рекомбинантного APE1, используемого при исследовании эффективности связывания модельных ДНК.



**Рисунок 41.** Эффективность гидролиза 96-звенных фрагментов промотора *hTERT*, содержащих и не содержащих G4 и замену/ы G→F, катализируемого APE1. Реакционные смеси содержали 20 мМ Трис-НСl (рН 7,5), 100 мМ КСl, 5 мМ MgCl<sub>2</sub>, 1 мМ DTT, 0,5 мг/мл BSA, 100 нМ одноцепочечный ДНК-фрагмент, меченный TAMRA с 3'-конца, и 10 ед. акт. (36 нМ) APE1. Реакции проводили при 37 °С в течение 60 мин и останавливали прогреванием при 60 °С в течение 20 мин. Погрешности соответствуют 95%-ным доверительным интервалам (n = 3).

Установлено, что эффективность APE1-индуцированного гидролиза 96-звенных

фрагментов промотора *hTERT* с квадруплексными структурами была снижена от 1,5 до 10 раз по сравнению с контрольным олигонуклеотидом 96-RF, не формирующим G4. Конечный выход продуктов гидролиза зависел от положения модификации в G-тракте G4-мотива промотора *hTERT*. Локализация G→F в пятизвенном G-тракте (96-G124F), в меньшей степени дестабилизирующая G4-структуру по данным КД, в большей степени снижала активность APE1 (выход ~12% после 60 мин инкубации). Эффективность гидролиза модельных субстратов увеличивалась в ряду 96-G124F < 96-G139F ~ 96-G146F ~ 96-G138F < 96-G138,139F < 96-RF. Эта закономерность сохранялась на всем промежутке времени от 2 до 60 мин.

#### **IV.5.5.3 Эффективность гидролиза 30-звенных фрагментов промотора *hTERT* белком APE1**

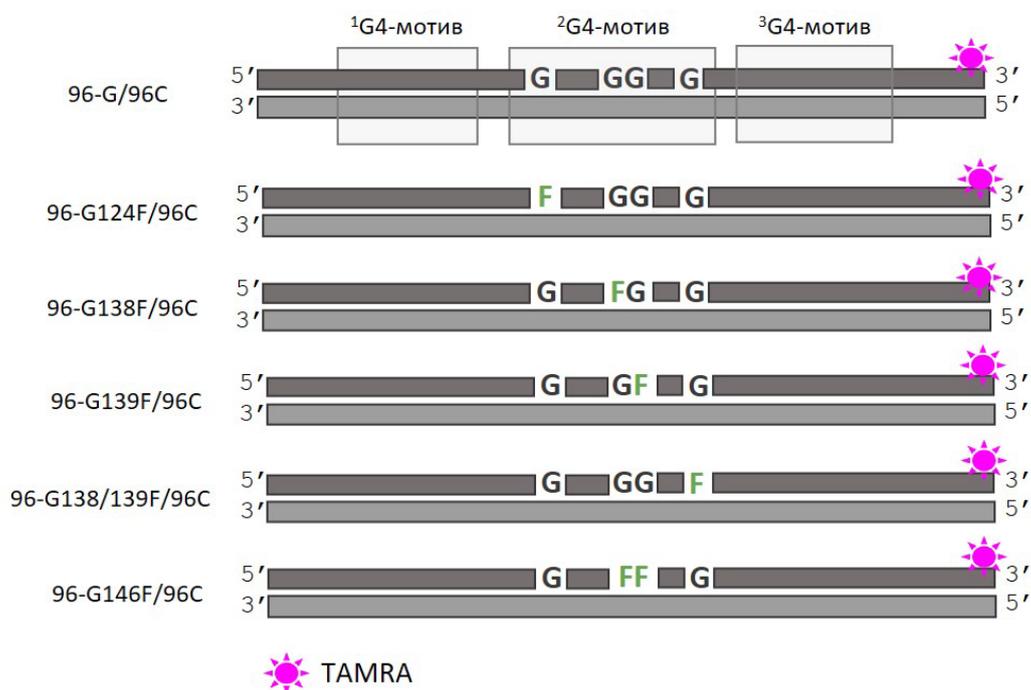
Для детального анализа влияния положения аналога AP-сайта внутри G4 на способность APE1 расщеплять поврежденную ДНК мы исследовали гидролиз 30-звенных фрагментов промотора *hTERT*, включающих центральный G4 (<sup>2</sup>G4) и содержащих изучаемые нами варианты замен G→F. Условия эксперимента соответствовали тем, что применялись при изучении гидролиза 96-звенных одноцепочечных ДНК-фрагментов белком APE1. Эффективность гидролиза модельных субстратов уменьшалась в ряду: 30-G146F < 30-G138F ~ 30-G138,139F << 30-G139F ~ 30-G124F. В отличие от 96-звенных модельных фрагментов, где соседние G4-мотивы частично маскировали эффект дестабилизации AP-сайтом центрального <sup>2</sup>G4, в 30-звенных ДНК-фрагментах влияние повреждения на устойчивость <sup>2</sup>G4 проявлялось гораздо сильнее. Нами была обнаружена отчетливая обратная корреляция между стабильностью G4 и эффективностью расщепления ДНК белком APE1 (коэффициент Спирмена  $\rho = -0,973$ ,  $p < 0,05$ ). Выход продуктов ферментативной реакции варьировался от 3–7% у наиболее стабильных структур (30-G124F и 30-G139F,  $T_{пл}$  68 °C; минимальное дестабилизирующее действие замены G→F, локализованной в терминальных положениях 5- и 4-звенных G-трактов), до 50% у олигонуклеотидов 30-G138F и 30-G138,139F ( $T_{пл}$  63 °C; одинарная или двойная замена в центральной части 4-нуклеотидного тракта), и достигал 88% в случае наименее стабильного варианта 30-G146F ( $T_{пл}$  58 °C). Таким образом, чем сильнее дестабилизирует квадруплексную структуру аналог AP-сайта внутри G4-мотива 30-звенных фрагментов промотора *hTERT*, тем выше активность APE1.

Таким образом, наличие AP-сайта дестабилизирует G4, однако степень этой дестабилизации зависит от локализации модификации и длины фланкирующих последовательностей. Образование и стабильность G-квадруплекса негативно влияет на активность APE1 человека (ключевого фермента BER) при удалении повреждения, локализованного в центральном <sup>2</sup>G4-мотиве промотора *hTERT*. Эффективность гидролиза модельного субстрата с заменой G→F в -124 положении белком APE1 значительно снижена по

сравнению с другими субстратами независимо от длины фланкирующих последовательностей. Таким образом, снижение активности BER в PQS промотора *hTERT* в условиях стабилизации G4 может быть одной из причин закрепления тканеспецифичной мутации в положении -124, которая является «драйверной».

#### IV.5.6. Влияние GC-богатого состава промотора *hTERT* на способность белка APE1 узнавать повреждение

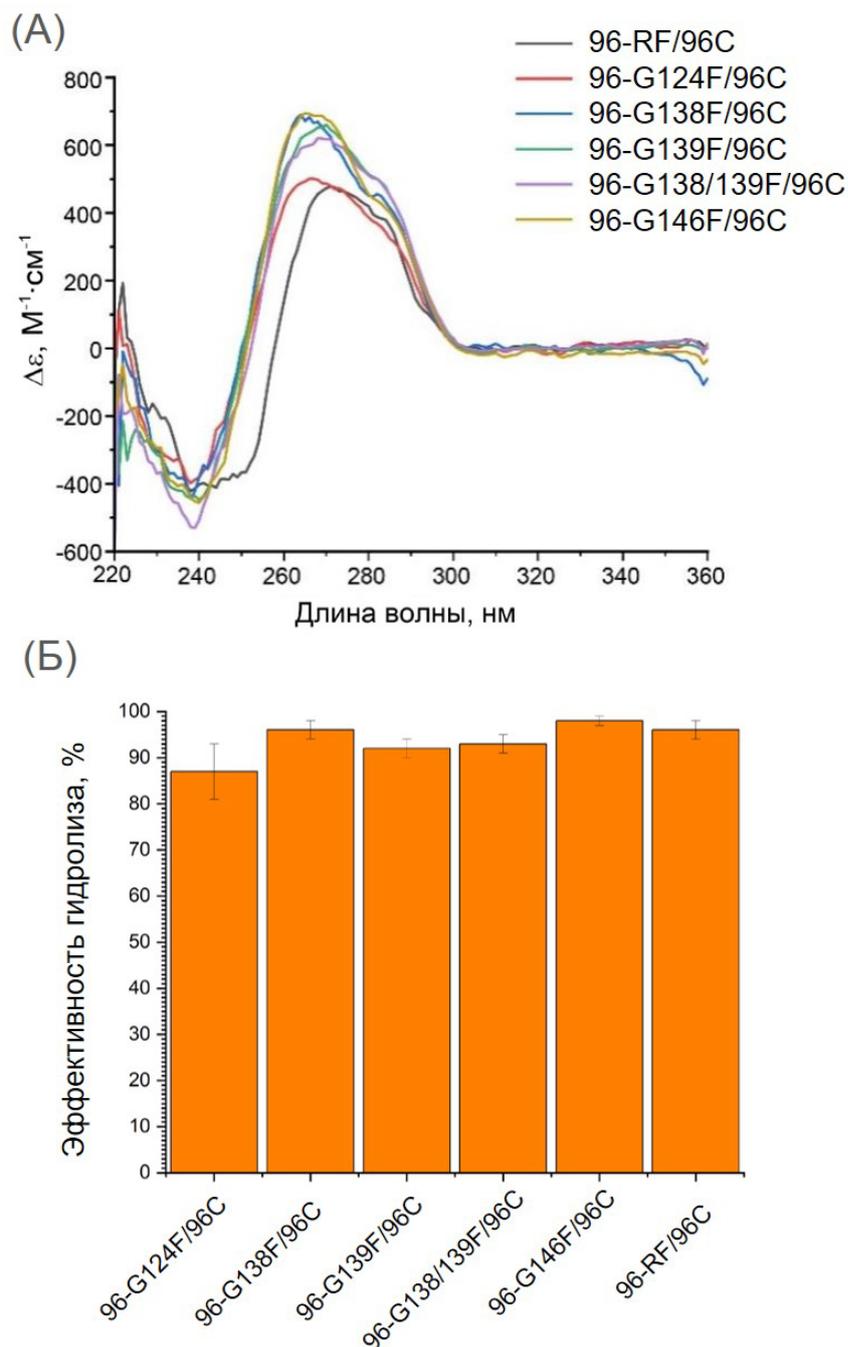
Влияние GC-богатого региона промотора *hTERT* изучали на ДНК-дуплексах, которые формировали гибридизацией 96-звенных одноцепочечных ДНК-фрагментов, содержащих и не содержащих G4 и с замену/ы G→F, с комплементарной C-богатой цепью (в которой напротив F расположен dC). Последовательности приведены в разделе III. «Материалы и методы» в таблице 9 (флуоресцентная метка TAMRA находится на 3'-конце G-цепи и обозначена розовой звездочкой, рис. 42).



**Рисунок 42.** Модельные ДНК-дуплексы, которые формировали гибридизацией 96-звенных одноцепочечных ДНК-фрагментов, содержащих и не содержащих G4 и с замену/ы G→F, в положениях, соответствующих «драйверным» мутациям, с комплементарной C-богатой цепью (в которой напротив F расположен dC).

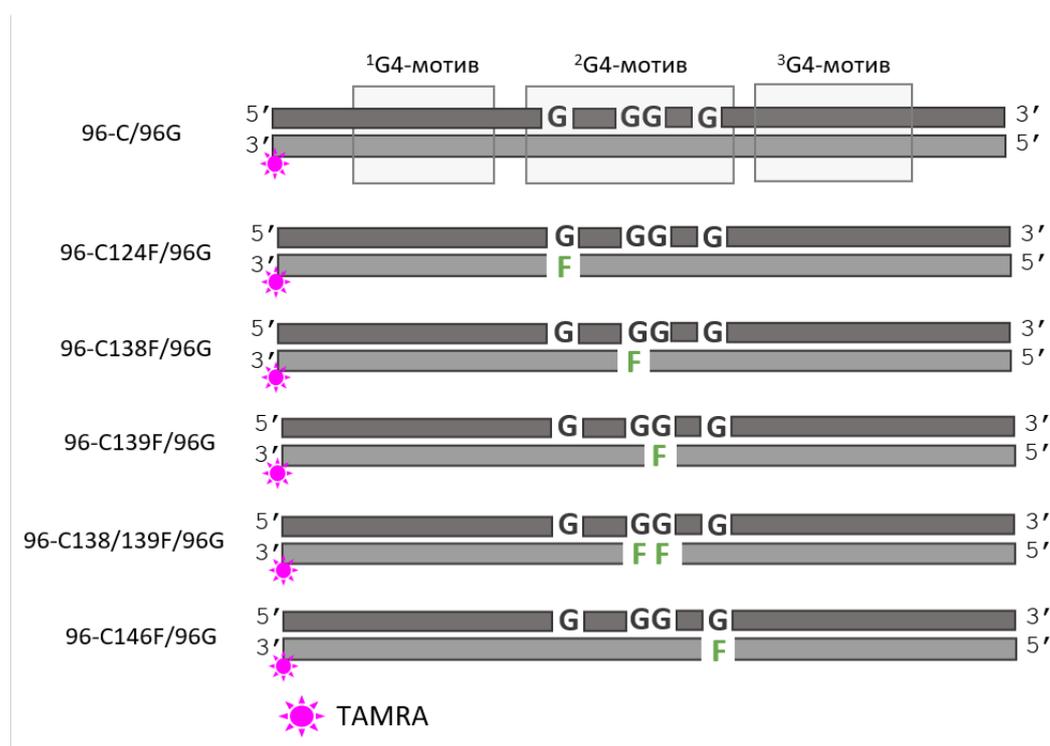
Спектры КД, полученные для этих ДНК, были типичными для двутяжевой ДНК с максимумом при ~ 280 нм: (Рис. 43А). APE1-индуцированный гидролиз этих субстратов проходил эффективно, независимо от положения F, и достигал 85-100%, что говорит об отсутствии влияния нуклеотидного окружения на этот процесс (Рис. 43Б). Таким образом,

определяющим фактором для активности эндонуклеазы APE1 является не последовательность ДНК, а формирование и стабильность G4-структуры.



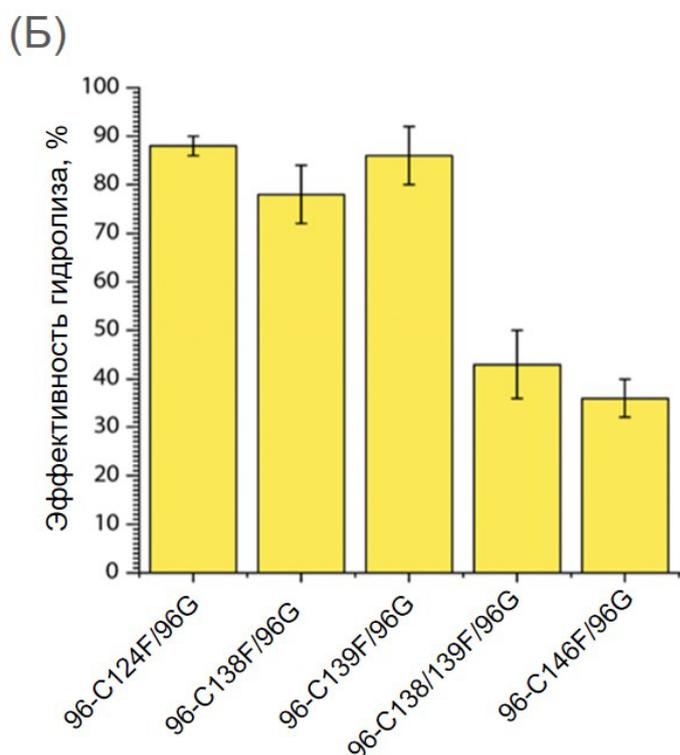
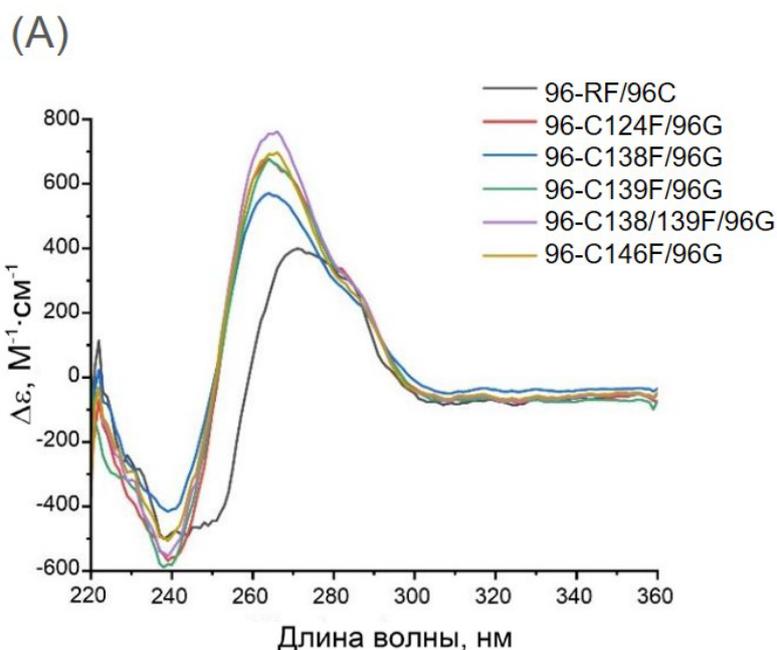
**Рисунок 43.** Влияние замен G→F на структуру двухцепочечных 96-звенных фрагментов промотора *hTERT*, содержащих и не содержащих G4 и их гидролиз ферментом APE1. **(А)** Спектры КД двухцепочечных 96-звенных GC-богатых фрагментов промотора *hTERT*, которые различаются по положению F, полученные при 30 °C в 20 мМ буфере Трис-НСl (pH 7,5), 1 мМ ЭДТА и 100 мМ КСl. **(Б)** Эффективность расщепления субстратов ферментом APE1 в течение 60 мин. Реакционные смеси содержали 20 мМ Трис-НСl (pH 7,5), 100 мМ КСl, 5 мМ MgCl<sub>2</sub>, 1 мМ DTT, 0,5 мг/мл BSA, 100 нМ одноцепочечного ДНК-субстрата, меченного TAMRA по 3'-концу модифицированной цепи, и 10 ед. (36 нМ) APE1. Погрешности соответствуют 95%-ным доверительным интервалам (n = 3).

Мы исходили из гипотезы, что накопление AP-сайтов в G4-мотивах промотора *hTERT* может быть одной из причин повышенной частоты мутаций в этом участке генома. Однако подобные повреждения могут образовываться и в комплементарной С-богатой цепи – например, вследствие дезаминирования цитозина с превращением в урацил с последующим удалением последнего урацил-ДНК-гликозилазой с образованием AP-сайта в рамках пути BER. Мы предположили, что наличие F в С-богатой цепи *hTERT* промотора может способствовать формированию G4 из-за частичной дестабилизации дуплекса. Учитывая наши предыдущие данные о повышенном сродстве фермента APE1 к G4, мультиквадруплексная структура может действовать как «ловушка» для белка, снижая эффективность удаления AP-сайта из комплементарной С-богатой цепи. Для проверки этой гипотезы были использованы двухцепочечные 96-звенные модельные фрагменты *hTERT*-промотора с модификациями F в С-богатой цепи в позициях, соответствующих «драйверным» мутациям и флуоресцентной меткой TAMRA на 3'-конце С-цепи, отмеченной розовой звездочкой на рисунке 44. Их последовательность представлена в разделе III «Материалы и методы» (Таблица 9).



**Рисунок 44.** Модельные двуцепочечные 96-звенные модельные фрагменты *hTERT*-промотора с модификациями F в С-богатой цепи в позициях, соответствующих «драйверным» мутациям.

Спектры КД (Рис. 45А) 96-звенных дуплексов с заменами С→F в позициях, соответствующих драйверным мутациям промотора *hTERT*, оказались характерными для смеси GC-богатых дуплексов и G4-структур (виден характерный пик при 260 нм), образованных G-богатой цепью промоторного региона *hTERT*.



**Рисунок 45.** Влияние замены С→F в С-цепи промотора *hTERT* на структуру двуцепочечных 96-звенных субстратов и их гидролиз APE1. (A) Спектры КД двуцепочечных 96-звенных GC-богатых фрагментов промотора *hTERT*, которые различаются по положению замены С→F, полученные при 30 °С в 20 мМ буфере Трис-НСl (рН 7,5), 1 мМ ЭДТА и 100 мМ КСl. (Б) Расщепление модельных ДНК, содержащих замену/ы С→F в С-цепи двуцепочечного фрагмента промотора *hTERT*. Реакционные смеси содержали 20 мМ Трис-НСl (рН 7,5), 100 мМ КСl, 5 мМ MgCl<sub>2</sub>, 1 мМ DTT, 0,5 мг/мл BSA, 100 нМ одноцепочечной ДНК-субстрата, меченного 3'-ТАМРА, и 10 ед. акт. (36 нМ) APE1. Погрешности соответствуют 95% доверительным интервалам (n = 3).

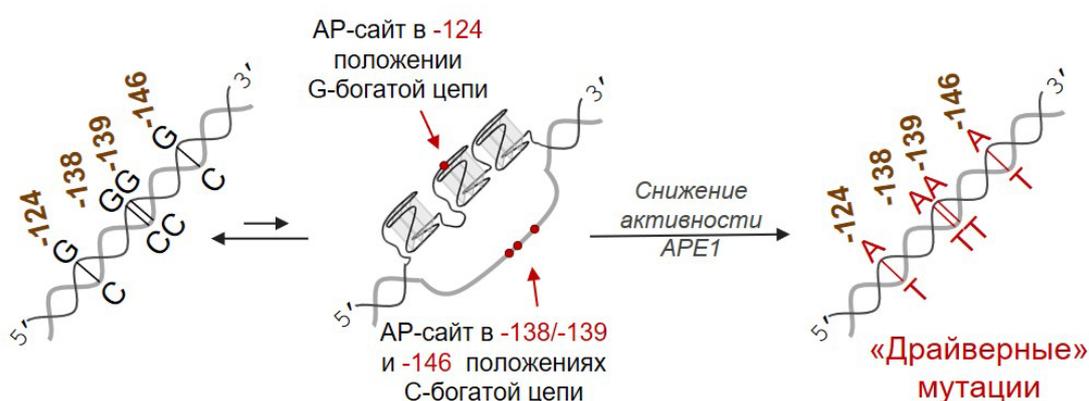
Далее мы исследовали, насколько эффективно APE1 расщепляет аналог AP-сайта в

указанных двуцепочечных ДНК-фрагментах, моделирующих участок промотора *hTERT*. Реакции проводили в буфере, содержащем 20 мМ Трис-НСl (рН 7,5), 100 мМ КСl, 5 мМ MgCl<sub>2</sub>, 1 мМ DTT, 0,5 мг/мл BSA, 100 нМ 3'-ТАМРА-меченого субстрата и 10 ед. акт.(36 нМ) АРЕ1, при 37 °С в течение 60 мин с последующей инактивацией фермента прогреванием до 60 °С в течение 20 мин. В случае дуплексов с единичной заменой С→F в С-богатой цепи фрагмента промотора *hTERT* эффективность расщепления определялась локализацией АР-сайта. Для дуплексов 96-С124F/96G, 96-С138F/96G, 96-С139F/96G она достигала 80–90%, тогда как для 96-С138/139F/96G и 96-С146F/96G не превышала 40% (Рис. 45Б). Отчетливо видно, что по сравнению с двуцепочечными 96-звенными модельными фрагментами, содержащими АР-сайт в G-богатой цепи, картина сильно изменилась. В случае замен С146F и двойной С138/139F этот эффект может быть более значительным ввиду того, что замена С146F находится напротив G146, входящего в состав G-тракта из трех dG, облегчая формирование G4, а двойная замена С138/139F, несомненно, оказывает еще большее дестабилизирующее влияние на дуплекс, что приводит к стабилизации квадруплекса. С другой стороны, замены С124F, С138F и С139F, по всей видимости, не оказывают такого значительного влияния на смещение равновесия в сторону формирования G4. Во-первых, С138F и С139F находятся в центральной области модельного субстрата, а длинные фланкирующие последовательности, несмотря на дестабилизирующую дуплекс модификацию, препятствуют формированию G4 (этого нельзя сказать о наличии сразу двух таких модификаций). Во-вторых, замена С124F, находящаяся напротив G124, расположенного в G-тракте из пяти dG, не так сильно облегчает формирование G4, как модификация С146F, расположенная напротив G146, входящего в G-тракте из трех dG.

\*\*\*

В рамках этой части работы исследовано функционирование белка АРЕ1 человека в GС-богатом PQS-мотиве промотора *hTERT*. Показано, что 68-звенная G-богатая последовательность промотора в одноцепочечном состоянии *in vitro* формирует G4 параллельной топологии, стабильность которого снижается в результате замены одного или двух остатков G в положениях, соответствующих «драйверным» мутациям (-124, -146 и -139/-139 относительно стартового кодона), на аналог апурин/апиридинового сайта - остаток 1,2-дидезоксирибозы. Эффективность расщепления модельных субстратов АРЕ1 в условиях, стабилизирующих G4, снижена по сравнению с ДНК, не способной к формированию G4; десятикратное снижение наблюдается для субстрата с заменой G→F в позиции -124. При этом связывание ДНК-фрагментов промотора *hTERT* с АРЕ1 в 10 раз эффективнее, чем с ДНК, не содержащей G4-мотив. Локализация замен G→F в условиях дуплекса не влияет на активность АРЕ1. Введение замен С→F в С-богатую цепь промотора *hTERT* способствует частичной стабилизации G4 на G-

цепи даже в присутствии комплементарной цепи, а эффективность гидролиза ДНК с модификациями F в положениях -146 и -138/-139 белком APE1 в два раза ниже по сравнению с другими модельными субстратами. Полученные данные позволяют предложить следующую гипотезу о механизмах закрепления «драйверных» мутаций в промоторе *hTERT*: (1) закрепление мутации в позиции -124 промотора *hTERT* может происходить по причине снижения эффективности BER в результате дисфункции APE1 в области PQS в условиях образования G4 (например, в процессах расхождения цепей ДНК), (2) закрепление «драйверных» мутаций в позициях -146 и -138/139 промотора *hTERT* может происходить по причине возникновения AP-сайтов в С-цепи, способствующих локальной стабилизации G4, и ингибированию APE1 даже в условиях существования ДНК-дуплекса (Рис. 46) [214].



**Рисунок 46.** Схематичное изображение гипотезы о влиянии неканонической структуры, формируемой G-богатой цепью фрагмента *hTERT*-промотора, на закрепление «драйверных» мутаций.

## V. Заключение

В рамках проведенного исследования выявлены новые функциональные особенности ключевых белков систем репарации ДНК прокариот и эукариот в ходе их взаимодействия с GC-богатыми областями генома. Результаты позволяют на новом уровне понять то, как системы поддержания геномной стабильности могут быть одновременно защитным и регуляторным факторами.

Показано, что MutL *Neisseria gonorrhoeae* представляет собой «никующую» эндонуклеазу с ранее ограничено описанными свойствами: ее активность на линейных дуплексных системах зависит от концентрации двухвалентных катионов магния и марганца и регулируется  $\beta$ -субъединицей ДНК-полимеразы III. Выявлено, что белок MutL способен специфически связываться с G-квадруplexами в регуляторных областях генома *Neisseria gonorrhoeae*, а именно с параллельным G4 перед *pilE*, необходимым для инициации антигенной вариации гена белка пилина. Этот факт открывает новое направление в изучении механизмов регуляции адаптации патогена через вторичные структуры ДНК посредством MMR. Неспособность белка

MutS узнавать «мисматч» в GC-богатом регионе дуплексной ДНК, показанная на примере гомолога из *Cereibacter sphaeroides*, подтверждает высокий мутагенный потенциал таких регионов.

Анализ геномов позволил выявить консервативные G4-мотивы в области промотора гена обратной транскриптазы теломеразы для различных отрядов млекопитающих, включая приматов. Эволюционная консервативность PQS подтверждает их функциональную значимость в регуляторных областях генома.

Спектральными методами на модельных ДНК, содержащих G-богатую последовательность промотора гена *hTERT* человека, продемонстрировано образование G4 параллельной топологии *in vitro*.

Особое внимание заслуживает исследование роли апуриновых/апиримидиновых участков в GC-богатых последовательностях регуляторных областей, так как их возникновение может являться причиной закрепления «драйверных» мутаций в положениях -124, -146 и -138/-139 п.н. относительно стартового кодона. Наши данные показывают, что «искусственные» AP-сайты (остатки 1,2-дидезоксирибозы) в G-богатой (некодирующей) цепи промотора *hTERT* снижают термическую стабильность G4, тогда как эти же модификации в некоторых положениях комплементарной кодирующей C-цепи, наоборот, могут стабилизировать G-квадруплексы в противоположной цепи. Это указывает на то, что AP-сайты могут выступать в роли потенциальных эпигенетических модуляторов экспрессии генов, находящихся под контролем GC-богатых промоторов, таких, как *hTERT*.

Мы установили, что в условиях образования G4-структур эффективность действия ключевого белка системы эксцизионной репарации оснований - апурин/апиримидиновой эндонуклеазы 1 человека при удалении стабильного аналога AP-сайта - остатков 1,2 - дидезоксирибозы протекает менее эффективно. Тем не менее, GC-богатое нуклеотидное окружение модификации в дуплексной ДНК не оказывает заметного влияния на активность фермента. Закрепление «драйверных» мутаций в GC-богатом регионе *hTERT* промотора может быть результатом дисфункции APE1, и выражается в сниженной эффективности гидролиза ДНК при локализации повреждения в положении -124 G-богатой цепи в условиях образования G-квадруплексной структуры и в случае возникновения повреждения в положениях -146 и -138/-139 C-богатой цепи *hTERT* промотора.

Таким образом, настоящее исследование впервые комплексно демонстрирует взаимосвязь между структурными особенностями GC-богатых областей генома и ограничениями в работе систем репарации ДНК прокариот и эукариот.

## VI. Выводы

1. Белок MutS из *Cereibacter sphaeroides* не дискриминирует «мисматч» G/T в ДНК-дуплексе с составом GC-пар 82%: кажущиеся константы диссоциации комплексов белка с ДНК, содержащей и не содержащей некомплементарную пару, составляют  $72 \pm 9$  и  $88 \pm 11$  нМ, соответственно. MutS узнает «мисматч» G/T в ДНК-дуплексе с составом GC-пар 53%: кажущиеся константы диссоциации комплексов белка с дуплексом, содержащим пару G/T, составляет  $39 \pm 5$  нМ, а с каноническим дуплексом -  $349 \pm 46$  нМ.
2. Белок MutL из *Neisseria gonorrhoeae* связывается с G-квадруплексом, образованным G-богатой последовательностью перед геном *pilE*, в 3-5 раз более эффективно по сравнению с ДНК-дуплексом. Фермент не гидролизует межнуклеотидные связи внутри G-квадруплекса, но вносит одноцепочечные разрывы во фланкирующие его дуплексные фрагменты.
3. Анализ баз данных подтвердил эволюционную консервативность G-квадруплексных мотивов в промоторных областях гена обратной транскриптазы теломеразы (*TERT*) 141 вида млекопитающих, включая приматов.
4. 68-звенный G-богатый фрагмент некодирующей цепи промотора *TERT* человека *in vitro* образует G-квадруплекс параллельной топологии, стабильность которого снижается в результате замены одного или двух остатков dG в положениях, соответствующих «драйверным» мутациям (-124, -146 и -139/-139 относительно стартового кодона), на аналог апурин/апиримидинового сайта - остаток 1,2-дидезоксирибозы.
5. Показано, что апуриновая/апиримидиновая эндонуклеаза 1 эффективно связывается с G-квадруплексной структурой G-богатого фрагмента некодирующей цепи промотора *TERT* человека, содержащей остаток 1,2-дидезоксирибозы, но практически не гидролизует ДНК с модификацией в положении -124. Наблюдается двукратное снижение каталитической активности этого фермента в случае введения модификации в положения -146 и -138/-139 C-богатой кодирующей цепи в составе ДНК-дуплекса.

## Список литературы

1. Zhang Z. H., Qian S. H., Wei D., Chen Z.-X. *In vivo* dynamics and regulation of DNA G-quadruplex structures in mammals // *Cell Biosci.* 2023. T. 13, № 1. C. 117.
2. Pavlova A. V., Kubareva E. A., Monakhova M. V., Zvereva M. I., Dolinnaya N. G. Impact of G-quadruplexes on the regulation of genome integrity, DNA damage and repair // *Biomolecules.* 2021. T. 11, № 9. C. 1284.
3. Hänsel-Hertsch R., Di Antonio M., Balasubramanian S. DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential // *Nat. Rev. Mol. Cell Biol.* 2017. T. 18, № 5. C. 279–284.
4. Bryan T. M. Mechanisms of DNA Replication and Repair: Insights from the study of G-quadruplexes // *Molecules.* 2019. T. 24, № 19. C. 3439.
5. Obergfell K. P., Seifert H. S. The pilin N-terminal domain maintains *Neisseria gonorrhoeae* transformation competence during pilus phase variation // *PLoS Genet* / под ред. Søgaard-Andersen L. 2016. T. 12, № 5. C. e1006069.
6. Schook P. O. P., Stohl E. A., Criss A. K., Seifert H. S. The DNA-binding activity of the *Neisseria gonorrhoeae* LexA orthologue NG1427 is modulated by oxidation // *Molecular Microbiology.* 2011. T. 79, № 4. C. 846–860.
7. Stohl E. A., Seifert H. S. *Neisseria gonorrhoeae* DNA recombination and repair enzymes protect against oxidative damage caused by hydrogen peroxide // *J. Bacteriol.* 2006. T. 188, № 21. C. 7645–7651.
8. Voter A. F., Callaghan M. M., Tippana R., Myong S., Dillard J. P., Keck J. L. Antigenic variation in *Neisseria gonorrhoeae* occurs independently of RecQ-mediated unwinding of the *pilE* G quadruplex // *J Bacteriol* / под ред. Stock A. M. 2020. T. 202, № 3.
9. *pilE*.
10. Huang F. W., Hodis E., Xu M. J., Kryukov G. V., Chin L., Garraway L. A. Highly Recurrent *TERT* Promoter Mutations in Human Melanoma // *Science.* 2013. T. 339, № 6122. C. 957–959.
11. Gandolfi G., Ragazzi M., Frasoldati A., Piana S., Ciarrocchi A., Sancisi V. *TERT* promoter mutations are associated with distant metastases in papillary thyroid carcinoma // *European Journal of Endocrinology.* 2015. T. 172, № 4. C. 403–413.
12. Kim Y.-J., M. Wilson Iii D. Overview of base excision repair biochemistry // *CMP.* 2012. T. 5, № 1. C. 3–13.
13. Jackson S. P., Bartek J. The DNA-damage response in human biology and disease // *Nature.* 2009. T. 461, № 7267. C. 1071–1078.
14. Chatterjee N., Walker G. C. Mechanisms of DNA damage, repair, and mutagenesis // *Environ. and Mol. Mutagen.* 2017. T. 58, № 5. C. 235–263.
15. Jiricny J. The multifaceted mismatch-repair system // *Nat. Rev. Mol. Cell Biol.* 2006. T. 7, № 5. C. 335–346.
16. Scott T. L., Rangaswamy S., Wicker C. A., Izumi T. Repair of oxidative DNA damage and cancer: recent progress in DNA base excision repair // *Antioxidants & Redox Signaling.* 2014. T. 20, № 4. C. 708–726.
17. San Filippo J., Sung P., Klein H. Mechanism of eukaryotic homologous recombination // *Annu. Rev. Biochem.* 2008. T. 77, № 1. C. 229–257.
18. Hoeijmakers J. H. J. DNA damage, aging, and cancer // *N. Engl. J. Med.* 2009. T. 361, № 15. C. 1475–1485.
19. Wozniak K. J., Simmons L. A. Bacterial DNA excision repair pathways // *Nat. Rev. Microbiol.* 2022. T. 20, № 8. C. 465–477.
20. Galande A. A., Perween N., Saijo M., Ghaskadbi S. S., Ghaskadbi S. Analysis of the conserved NER helicases (XPB and XPD) and UV-induced DNA damage in Hydra // *Biochimica et Biophysica Acta (BBA) - General Subjects.* 2018. T. 1862, № 9. C. 2031–2042.

21. M. Taylor A. R. Lehmann E. Review: Conservation of eukaryotic DNA repair mechanisms // International Journal of Radiation Biology. 1998. T. 74, № 3. C. 277–286.
22. Aravind L. Conserved domains in DNA repair proteins and evolution of repair systems // Nucleic Acids Research. 1999. T. 27, № 5. C. 1223–1242.
23. DiRuggiero J., Robb F. T. Early Evolution of DNA repair mechanisms // The Genetic Code and the Origin of Life. Boston, MA: Springer US, 2004. C. 169–182.
24. Gohil D., Sarker A. H., Roy R. Base excision repair: mechanisms and impact in biology, disease, and medicine // IJMS. 2023. T. 24, № 18. C. 14186.
25. Guliaev A. B., Singer B. DNA Damage: Alkylation // Encyclopedia of biological chemistry. Elsevier, 2004. C. 609–613.
26. Spivak G. Nucleotide excision repair in humans // DNA Repair. 2015. T. 36. C. 13–18.
27. Kraithong T., Hartley S., Jeruzalmi D., Pakotiprapha D. A Peek inside the machines of bacterial nucleotide excision repair // IJMS. 2021. T. 22, № 2. C. 952.
28. Gillet L. C. J., Schärer O. D. Molecular mechanisms of mammalian global genome nucleotide excision repair // Chem. Rev. 2006. T. 106, № 2. C. 253–276.
29. Malik S. S., Mubarak S., Baig M., Masood N., Chaudhry N. Genetic polymorphism in ERCC5 and breast cancer risk // Mol. Biol. Res. Commun. 2019. T. 8, № 1.
30. Paul D., Mu H., Zhao H., Ouerfelli O., Jeffrey P. D., Broyde S., Min J.-H. Structure and mechanism of pyrimidine–pyrimidone (6-4) photoproduct recognition by the Rad4/XPC nucleotide excision repair complex // Nucleic Acids Research. 2019. T. 47, № 12. C. 6015–6028.
31. Krokan H. E., Bjoras M. Base Excision Repair // Cold Spring Harbor Perspectives in Biology. 2013. T. 5, № 4. C. a012583–a012583.
32. Gonzalez K., Faustoferri R. C., Quivey Jr R. G. Role of DNA base excision repair in the mutability and virulence of *Streptococcus mutans* // Molecular Microbiology. 2012. T. 85, № 2. C. 361–377.
33. LeCuyer B. E., Criss A. K., Seifert H. S. Genetic characterization of the nucleotide excision repair system of *Neisseria gonorrhoeae* // J. Bacteriol. 2010. T. 192, № 3. C. 665–673.
34. Kisker C., Kuper J., Van Houten B. Prokaryotic nucleotide excision repair // Cold Spring Harbor Perspectives in Biology. 2013. T. 5, № 3. C. a012591–a012591.
35. Drohat A. C., Maiti A. Mechanisms for enzymatic cleavage of the N-glycosidic bond in DNA // Org. Biomol. Chem. 2014. T. 12, № 42. C. 8367–8378.
36. Baute J., Depicker A. Base Excision Repair and its Role in maintaining genome stability // Critical Reviews in Biochemistry and Molecular Biology. 2008. T. 43, № 4. C. 239–276.
37. Moor N. A., Lavrik O. I. Protein–protein interactions in DNA base excision repair // Biochemistry Moscow. 2018. T. 83, № 4. C. 411–422.
38. Diatlova E. A., Mechetin G. V., Zharkov D. O. Distinct mechanisms of target search by endonuclease VIII-like DNA glycosylases // Cells. 2022. T. 11, № 20. C. 3192.
39. Demple B., Herman T., Chen D. S. Cloning and expression of APE, the cDNA encoding the major human apurinic endonuclease: definition of a family of DNA repair enzymes. // Proc. Natl. Acad. Sci. U.S.A. 1991. T. 88, № 24. C. 11450–11454.
40. Beard W. A., Horton J. K., Prasad R., Wilson S. H. Eukaryotic base excision repair: new approaches shine light on mechanism // Annu. Rev. Biochem. 2019. T. 88, № 1. C. 137–162.
41. Freudenthal B. D., Beard W. A., Cuneo M. J., Dyrkheeva N. S., Wilson S. H. Capturing snapshots of APE1 processing DNA damage // Nat. Struct. Mol. Biol. 2015. T. 22, № 11. C. 924–931.
42. Hoitsma N. M., Whitaker A. M., Beckwitt E. C., Jang S., Agarwal P. K., Van Houten B., Freudenthal B. D. AP-endonuclease 1 sculpts DNA through an anchoring tyrosine residue on the DNA intercalating loop // Nucleic Acids Research. 2020. C. gkaa496.
43. Erzberger J. P., Wilson D. M. The role of Mg<sup>2+</sup> and specific amino acid residues in the catalytic reaction of the major human abasic endonuclease: new insights from EDTA-resistant incision of acyclic abasic site analogs and site-directed mutagenesis // Journal of Molecular Biology. 1999. T. 290, № 2. C. 447–457.
44. Kanchan S., Sharma P., Chowdhury S. Evolution of endonuclease IV protein family: an in silico analysis // Biotech. 2019. T. 9, № 5. C. 168.

45. Sak S. C., Harnden P., Johnston C. F., Paul A. B., Kiltie A. E. APE1 and XRCC1 protein expression levels predict cancer-specific survival following radical radiotherapy in bladder cancer // *Clinical Cancer Research*. 2005. T. 11, № 17. C. 6205–6211.
46. Kladova O. A., Alekseeva I. V., Sapparbaev M., Fedorova O. S., Kuznetsov N. A. Modulation of the apurinic/aprimidinic endonuclease activity of human APE1 and of its natural polymorphic variants by base excision repair proteins // *IJMS*. 2020. T. 21, № 19. C. 7147.
47. Yang Z., Zhao J. Effect of *APE1* and *XRCC1* gene polymorphism on susceptibility to hepatocellular carcinoma and sensitivity to cisplatin // *Int. J. Clin. Exp. Med.* 2015. T. 8, № 6. C. 9931–9936.
48. Chembazhi U. V., Patil V. V., Sah S., Reeve W., Tiwari R. P., Woo E., Varshney U. Uracil DNA glycosylase (UDG) activities in *Bradyrhizobium diazoefficiens*: characterization of a new class of UDG with broad substrate specificity // *Nucleic. Acids Research*. 2017. T. 45, № 10. C. 5863–5876.
49. Michaels M. L., Cruz C., Grollman A. P., Miller J. H. Evidence that MutY and MutM combine to prevent mutations by an oxidatively damaged form of guanine in DNA. // *Proc. Natl. Acad. Sci. U.S.A.* 1992. T. 89, № 15. C. 7022–7025.
50. Parker A. R., Eshleman J. R. Human *MutY*: gene structure, protein functions and interactions, and role in carcinogenesis // *Cellular and Molecular Life Sciences (CMLS)*. 2003. T. 60, № 10. C. 2064–2083.
51. Mielecki D., Grzesiuk E. Ada response - a strategy for repair of alkylated DNA in bacteria // *FEMS Microbiol. Lett.* 2014. T. 355, № 1. C. 1–11.
52. Fu D., Samson L. D. Direct repair of 3,N4-ethenocytosine by the human ALKBH2 dioxygenase is blocked by the AAG/MPG glycosylase // *DNA Repair*. 2012. T. 11, № 1. C. 46–52.
53. Prakash A., Doublie S., Wallace S. S. The Fpg/Nei family of DNA glycosylases // *Progress in Molecular Biology and Translational Science*. Elsevier, 2012. T. 110. C. 71–91.
54. Albelazi M. S., Martin P. R., Mohammed S., Mutti L., Parsons J. L., Elder R. H. The biochemical role of the human NEIL1 and NEIL3 DNA glycosylases on model DNA replication forks // *Genes*. 2019. T. 10, № 4. C. 315.
55. Petermann E. ATФ-dependent selection between single nucleotide and long patch base excision repair // *DNA Repair*. 2003. T. 2, № 10. C. 1101–1114.
56. Nickson C. M., Parsons J. L. Monitoring regulation of DNA repair activities of cultured cells in-gel using the comet assay // *Front. Genet.* 2014. T. 5.
57. Fousteri M., Mullenders L. H. Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects // *Cell Res.* 2008. T. 18, № 1. C. 73–84.
58. Petrusheva I. O., Evdokimov A. N., Lavrik O. I. Molecular mechanism of global genome nucleotide excision repair // *Acta Naturae*. 2014. T. 6, № 1. C. 23–34.
59. Sancar A., Tang M. Nucleotide excision repair // *Photochem & Photobiology*. 1993. T. 57, № 5. C. 905–921.
60. He F., Bravo M., Fan L. Helicases required for nucleotide excision repair: structure, function and mechanism // *The Enzymes*. Elsevier, 2023. T. 54. C. 273–304.
61. Лаврик О. И. Репарация ДНК как ключевой механизм обеспечения стабильности генома и преодоления болезней человека (от редактора-составителя) // *Молекул. биол.* 2021. T. 55, № 2. C. 179–180.
62. Sugawara K., Okamoto T., Shimizu Y., Masutani C., Iwai S., Hanaoka F. A multistep damage recognition mechanism for global genomic nucleotide excision repair // *Genes Dev.* 2001. T. 15, № 5. C. 507–521.
63. Maltseva E. A., Rechkunova N. I., Sukhanova M. V., Lavrik O. I. Poly(ADP-ribose) Polymerase 1 modulates interaction of the nucleotide excision repair factor XPC-RAD23B with DNA via Poly(ADP-ribosylation) // *Journal of Biological Chemistry*. 2015. T. 290, № 36. C. 21811–21820.
64. Bergink S., Toussaint W., Luijsterburg M. S., Dinant C., Alekseev S., Hoeijmakers J. H. J., Dantuma N. P., Houtsmuller A. B., Vermeulen W. Recognition of DNA damage by XPC coincides

- with disruption of the XPC–RAD23 complex // *Journal of Cell Biology*. 2012. T. 196, № 6. C. 681–688.
65. Kusakabe M., Onishi Y., Tada H., Kurihara F., Kusao K., Furukawa M., Iwai S., Yokoi M., Sakai W., Sugawara K. Mechanism and regulation of DNA damage recognition in nucleotide excision repair // *Genes and Environ*. 2019. T. 41, № 1. C. 2.
  66. Coin F., Oksenysh V., Mocquet V., Groh S., Blattner C., Egly J. M. Nucleotide excision repair driven by the dissociation of CAK from TFIIH // *Molecular Cell*. 2008. T. 31, № 1. C. 9–20.
  67. Kokic G., Chernev A., Tegunov D., Dienemann C., Urlaub H., Cramer P. Structural basis of TFIIH activation for nucleotide excision repair // *Nat. Commun*. 2019. T. 10, № 1. C. 2885.
  68. Rudolf J., Rouillon C., Schwarz-Linek U., White M. F. The helicase XPD unwinds bubble structures and is not stalled by DNA lesions removed by the nucleotide excision repair pathway // *Nucleic. Acids Research*. 2010. T. 38, № 3. C. 931–941.
  69. Sugitani N., Sivley R. M., Perry K. E., Capra J. A., Chazin W. J. XPA: A key scaffold for human nucleotide excision repair // *DNA Repair*. 2016. T. 44. C. 123–135.
  70. Chen R., Wold M. S. Replication protein A: Single-stranded DNA's first responder: Dynamic DNA-interactions allow replication protein A to direct single-strand DNA intermediates into different pathways for synthesis or repair // *BioEssays*. 2014. T. 36, № 12. C. 1156–1161.
  71. Sancar A. DNA excision repair // *Annu. Rev. Biochem*. 1996. T. 65, № 1. C. 43–81.
  72. Dalhus B., Laerdahl J. K., Backe P. H., Bjørås M. DNA base repair – recognition and initiation of catalysis // *FEMS Microbiol. Rev*. 2009. T. 33, № 6. C. 1044–1078.
  73. Putnam C. D. Evolution of the methyl directed mismatch repair system in *Escherichia coli* // *DNA Repair*. 2016. T. 38. C. 32–41.
  74. Pukkila P. J., Peterson J., Herman G., Modrich P., Meselson M. Effects of high levels of DNA adenine methylation on methyl-directed mismatch repair in *Escherichia Coli* // *Genetics*. 1983. T. 104, № 4. C. 571–582.
  75. Acharya S., Foster P. L., Brooks P., Fishel R. The coordinated functions of the *E. coli* MutS and MutL proteins in mismatch repair // *Molecular Cell*. 2003. T. 12, № 1. C. 233–246.
  76. Ban C. Structural basis for MutH activation in *E.coli* mismatch repair and relationship of MutH to restriction endonucleases // *The EMBO Journal*. 1998. T. 17, № 5. C. 1526–1534.
  77. Nirwal S., Kulkarni D. S., Sharma A., Rao D. N., Nair D. T. Mechanism of formation of a toroid around DNA by the mismatch sensor protein // *Nucleic. Acids Research*. 2018. T. 46, № 1. C. 256–266.
  78. Lamers M. H., Winterwerp H. H. K., Sixma T. K. The alternating ATPase domains of MutS control DNA mismatch repair // *EMBO J*. 2003. T. 22, № 3. C. 746–756.
  79. Obmolova G., Ban C., Hsieh P., Yang W. Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA // *Nature*. 2000. T. 407, № 6805. C. 703–710.
  80. Lamers M. H., Perrakis A., Enzlin J. H., Winterwerp H. H. K., De Wind N., Sixma T. K. The crystal structure of DNA mismatch repair protein MutS binding to a G·T mismatch // *Nature*. 2000. T. 407, № 6805. C. 711–717.
  81. Mardenborough Y. S. N., Nitsenko K., Laffeber C., Duboc C., Sahin E., Quessada-Vial A., Winterwerp H. H. K., Sixma T. K., Kanaar R., Friedhoff P., Strick T. R., Lebbink J. H. G. The unstructured linker arms of MutL enable GATC site incision beyond roadblocks during initiation of DNA mismatch repair // *Nucleic. Acids Research*. 2019. T. 47, № 22. C. 11667–11680.
  82. Au K. G., Welsh K., Modrich P. Initiation of methyl-directed mismatch repair // *J. Biol. Chem*. 1992. T. 267, № 17. C. 12142–12148.
  83. Campbell J. L., Kleckner N. The rate of Dam-mediated DNA adenine methylation in *Escherichia coli* // *Gene*. 1988. T. 74, № 1. C. 189–190.
  84. Robertson A. B., Pattishall S. R., Gibbons E. A., Matson S. W. MutL-catalyzed ATP hydrolysis is required at a post-UvrD loading step in methyl-directed mismatch repair // *Journal of Biological Chemistry*. 2006. T. 281, № 29. C. 19949–19959.
  85. López De Saro F. J., O'Donnell M. Interaction of the  $\beta$  sliding clamp with MutS, ligase, and DNA polymerase I // *Proc. Natl. Acad. Sci. U.S.A*. 2001. T. 98, № 15. C. 8376–8380.

86. Monakhova M. V., Penkina A. I., Pavlova A. V., Lyaschuk A. M., Kucherenko V. V., Alexeevski A. V., Lunin V. G., Friedhoff P., Klug G., Oretskaya T. S., Kubareva E. A. Endonuclease Activity of MutL Protein of the *Rhodobacter sphaeroides* Mismatch Repair System // *Biochemistry Moscow*. 2018. T. 83, № 3. С. 281–293.
87. De Saro F. J. L., Marinus M. G., Modrich P., O'Donnell M. The  $\beta$  sliding clamp binds to multiple sites within MutL and MutS // *Journal of Biological Chemistry*. 2006. T. 281, № 20. С. 14340–14349.
88. Drotschmann K., Hall M. C., Shcherbakova P. V., Wang H., Erie D. A., Brownell F. R., Kool E. T., Kunkel T. A. DNA Binding properties of the yeast Msh2-Msh6 and Mlh1-Pms1 heterodimers // *Biological Chemistry*. 2002. T. 383, № 6.
89. Fishel R. Mismatch Repair // *Journal of Biological Chemistry*. 2015. T. 290, № 44. С. 26395–26403.
90. Han X.-P., Yang X.-W., Liu J. MutS and MutL sliding clamps in DNA mismatch repair // *Genome Instab. Dis.* 2022. T. 4, № 1. С. 1–11.
91. Fukui K., Nishida M., Nakagawa N., Masui R., Kuramitsu S. Bound nucleotide controls the endonuclease activity of mismatch repair enzyme MutL // *Journal of Biological Chemistry*. 2008. T. 283, № 18. С. 12136–12145.
92. Correa E. M. E., Martina M. A., De Tullio L., Argaraña C. E., Barra J. L. Some amino acids of the *Pseudomonas aeruginosa* MutL D(Q/M)HA(X)2E(X)4E conserved motif are essential for the *in vivo* function of the protein but not for the *in vitro* endonuclease activity // *DNA Repair*. 2011. T. 10, № 11. С. 1106–1113.
93. Reyes G. X., Schmidt T. T., Kolodner R. D., Hombauer H. New insights into the mechanism of DNA mismatch repair // *Chromosoma*. 2015. T. 124, № 4. С. 443–462.
94. Монахова М. В., Милакина М. А., Трикин Р. М., Орецкая Т. С., Кубарева Е. А. Особенности функционирования белка MutL из системы репарации "мисматчей" различных организмов // *Биоорганическая химия*. 2020. T. 46, № 6. С. 875–890.
95. Ghodgaonkar M. M., Lazzaro F., Olivera-Pimentel M., Artola-Borán M., Cejka P., Reijns M. A., Jackson A. P., Plevani P., Muzi-Falconi M., Jiricny J. Ribonucleotides Misincorporated into DNA Act as Strand-Discrimination Signals in Eukaryotic Mismatch Repair // *Molecular Cell*. 2013. T. 50, № 3. С. 323–332.
96. Berglund A.-K., Navarrete C., Engqvist M. K. M., Hoberg E., Szilagyi Z., Taylor R. W., Gustafsson C. M., Falkenberg M., Clausen A. R. Nucleotide pools dictate the identity and frequency of ribonucleotide incorporation in mitochondrial DNA // *PLoS Genet* / под ред. Larsson N.-G. 2017. T. 13, № 2. С. e1006628.
97. Putnam C. D. Strand discrimination in DNA mismatch repair // *DNA Repair*. 2021. T. 105. С. 103161.
98. Kunkel T. A., Erie D. A. Eukaryotic Mismatch Repair in Relation to DNA Replication // *Annu. Rev. Genet.* 2015. T. 49, № 1. С. 291–313.
99. Kadyrov F. A., Genschel J., Fang Y., Penland E., Edelman W., Modrich P. A possible mechanism for exonuclease 1-independent eukaryotic mismatch repair // *Proc. Natl. Acad. Sci. U.S.A.* 2009. T. 106, № 21. С. 8495–8500.
100. Genschel J., Modrich P. Mechanism of 5'-Directed Excision in Human Mismatch Repair // *Molecular Cell*. 2003. T. 12, № 5. С. 1077–1086.
101. Longley M. J., Pierce A. J., Modrich P. DNA Polymerase  $\delta$  Is Required for Human Mismatch Repair *in Vitro* // *Journal of Biological Chemistry*. 1997. T. 272, № 16. С. 10917–10921.
102. Cho H., Misra R. Mutational Activation of Antibiotic-Resistant Mechanisms in the Absence of Major Drug Efflux Systems of *Escherichia coli* // *J Bacteriol* / под ред. Comstock L. E. 2021. T. 203, № 14.
103. Webber M. A. The importance of efflux pumps in bacterial antibiotic resistance // *Journal of Antimicrobial Chemotherapy*. 2003. T. 51, № 1. С. 9–11.

104. Li X.-Z., Mehrotra M. Role of Plasmid-Encoded Drug Efflux Pumps in Antimicrobial Resistance // *Efflux-Mediated Antimicrobial Resistance in Bacteria* / под ред. Li X.-Z., Elkins C. A., Zgurskaya H. I. Cham: Springer International Publishing, 2016. С. 595–623.
105. Belay W. Y., Getachew M., Tegegne B. A., Teffera Z. H., Dagne A., Zeleke T. K., Abebe R. B., Gedif A. A., Fenta A., Yirdaw G., Tilahun A., Aschale Y. Mechanism of antibacterial resistance, strategies and next-generation antimicrobials to contain antimicrobial resistance: a review // *Front. Pharmacol.* 2024. Т. 15. С. 1444781.
106. Quillin S. J., Seifert H. S. *Neisseria gonorrhoeae* host adaptation and pathogenesis // *Nat Rev Microbiol.* 2018. Т. 16, № 4. С. 226–240.
107. Lim K. Y. L., Mullally C. A., Haese E. C., Kibble E. A., McCluskey N. R., Mikucki E. C., Thai V. C., Stubbs K. A., Sarkar-Tyson M., Kahler C. M. Anti-Virulence Therapeutic Approaches for *Neisseria gonorrhoeae* // *Antibiotics.* 2021. Т. 10, № 2. С. 103.
108. Black C. G., Fyfe J. A. M., Davies J. K. Absence of an SOS-like system in *Neisseria gonorrhoeae* // *Gene.* 1998. Т. 208, № 1. С. 61–66.
109. Савицкая В. Ю., Монахова М. В., Якушкина Ю. В., Боровикова И. И., Кубарева Е. А. Бактерия *Neisseria gonorrhoeae*: системы репарации ДНК и их роль в патогенезе // *Биохимия.* 2022. Т. 87, № 9. С. 965–982.
110. Hill S. A., Masters T. L., Wachter J. Gonorrhea – an evolving disease of the new millennium // *MIC.* 2016. Т. 3, № 9. С. 371–389.
111. Sadarangani M., Pollard A. J., Gray-Owen S. D. Opa proteins and CEACAMs: pathways of immune engagement for pathogenic *Neisseria* // *FEMS Microbiol Rev.* 2011. Т. 35, № 3. С. 498–514.
112. Wetzler L. M., Barry K., Blake M. S., Gotschlich E. C. Gonococcal lipooligosaccharide sialylation prevents complement-dependent killing by immune sera // *Infect Immun.* 1992. Т. 60, № 1. С. 39–43.
113. Edwards J. L., Apicella M. A. The role of lipooligosaccharide in *Neisseria gonorrhoeae* pathogenesis of cervical epithelia: lipid A serves as a C3 acceptor molecule // *Cell Microbiol.* 2002. Т. 4, № 9. С. 585–598.
114. Ram S., Cullinane M., Blom A. M., Gulati S., McQuillen D. P., Monks B. G., O'Connell C., Boden R., Elkins C., Pangburn M. K., Dahlbäck B., Rice P. A. Binding of C4b-Binding Protein to Porin // *The Journal of Experimental Medicine.* 2001. Т. 193, № 3. С. 281–296.
115. Davidsen T., Tuven H. K., Bjørås M., Rødland E. A., Tønjum T. Genetic Interactions of DNA Repair Pathways in the Pathogen *Neisseria meningitidis* // *J Bacteriol.* 2007. Т. 189, № 15. С. 5728–5737.
116. Монахова М. В., Милакина М. А., Трикин Р. М., Орецкая Т. С., Кубарева Е. А. Особенности функционирования белка MutL из системы репарации «мисматчей» различных организмов // *Биоорганическая химия.* 2020. Т. 46, № 6. С. 563–579.
117. Boyce K. J. Mutators Enhance Adaptive Micro-Evolution in Pathogenic Microbes // *Microorganisms.* 2022. Т. 10, № 2. С. 442.
118. Vigué L., Eyre-Walker A. The comparative population genetics of *Neisseria meningitidis* and *Neisseria gonorrhoeae* // *PeerJ.* 2019. Т. 7. С. e7216.
119. Richardson A. R., Stojiljkovic I. Mismatch repair and the regulation of phase variation in *Neisseria meningitidis* // *Molecular Microbiology.* 2001. Т. 40, № 3. С. 645–655.
120. Tauseef I., Ali Y. M., Bayliss C. D. Phase Variation of PorA, a Major Outer Membrane Protein, Mediates Escape of Bactericidal Antibodies by *Neisseria meningitidis* // *Infect Immun* / под ред. Camilli A. 2013. Т. 81, № 4. С. 1374–1380.
121. Criss A. K., Bonney K. M., Chang R. A., Duffin P. M., LeCuyer B. E., Seifert H. S. Mismatch Correction Modulates Mutation Frequency and Pilus Phase and Antigenic Variation in *Neisseria gonorrhoeae* // *J Bacteriol.* 2010. Т. 192, № 1. С. 316–325.
122. Rotman E., Seifert H. S. *Neisseria gonorrhoeae* MutS Affects Pilin Antigenic Variation through Mismatch Correction and Not by *pilE* Guanine Quartet Binding // *J Bacteriol* / под ред. Gourse R. L. 2015. Т. 197, № 10. С. 1828–1838.

123. Cahoon L. A., Seifert H. S. Focusing homologous recombination: pilin antigenic variation in the pathogenic *Neisseria* // *Molecular Microbiology*. 2011. T. 81, № 5. C. 1136–1143.
124. Michel B., Boubakri H., Baharoglu Z., Lemasson M., Lestini R. Recombination proteins and rescue of arrested replication forks // *DNA Repair*. 2007. T. 6, № 7. C. 967–980.
125. Dillingham M. S., Kowalczykowski S. C. RecBCD Enzyme and the Repair of Double-Stranded DNA Breaks // *Microbiol Mol Biol Rev*. 2008. T. 72, № 4. C. 642–671.
126. Beyene G. T., Balasingham S. V., Frye S. A., Namouchi A., Homberset H., Kalayou S., Riaz T., Tønjum T. Characterization of the *Neisseria meningitidis* Helicase RecG // *PLoS ONE* / под ред. Marinus M. G. 2016. T. 11, № 10. C. e0164588.
127. Salvatore P., Bucci C., Pagliarulo C., Tredici M., Colicchio R., Cantalupo G., Bardaro M., Del Giudice L., Massardo D. R., Lavitola A., Bruni C. B., Alifano P. Phenotypes of a Naturally Defective *recB* Allele in *Neisseria meningitidis* Clinical Isolates // *Infect Immun*. 2002. T. 70, № 8. C. 4185–4195.
128. Morimatsu K., Kowalczykowski S. C. RecFOR Proteins Load RecA Protein onto Gapped DNA to Accelerate DNA Strand Exchange // *Molecular Cell*. 2003. T. 11, № 5. C. 1337–1347.
129. Killoran M. P., Kohler P. L., Dillard J. P., Keck J. L. RecQ DNA helicase HRDC domains are critical determinants in *Neisseria gonorrhoeae* pilin antigenic variation and DNA repair // *Molecular Microbiology*. 2009. T. 71, № 1. C. 158–171.
130. Mehr I. J., Seifert H. S. Differential roles of homologous recombination pathways in *Neisseria gonorrhoeae* pilin antigenic variation, DNA transformation and DNA repair // *Molecular Microbiology*. 1998. T. 30, № 4. C. 697–710.
131. Sechman E. V., Kline K. A., Seifert H. S. Loss of both Holliday junction processing pathways is synthetically lethal in the presence of gonococcal pilin antigenic variation // *Molecular Microbiology*. 2006. T. 61, № 1. C. 185–193.
132. Helm R. A., Seifert H. S. Pilin Antigenic Variation Occurs Independently of the RecBCD Pathway in *Neisseria gonorrhoeae* // *J Bacteriol*. 2009. T. 191, № 18. C. 5613–5621.
133. Ozer E. A., Prister L. L., Yin S., Ward B. H., Ivanov S., Seifert H. S. PacBio Amplicon Sequencing Method To Measure Pilin Antigenic Variation Frequencies of *Neisseria gonorrhoeae* // *mSphere* / под ред. Young V. B. 2019. T. 4, № 5. C. e00562-19.
134. Fussenegger M., Rudel T., Barten R., Ryll R., Meyer T. F. Transformation competence and type-4 pilus biogenesis in *Neisseria gonorrhoeae* – areview // *Gene*. 1997. T. 192, № 1. C. 125–134.
135. Kline K. A., Criss A. K., Wallace A., Seifert H. S. Transposon Mutagenesis Identifies Sites Upstream of the *Neisseria gonorrhoeae pilE* Gene That Modulate Pilin Antigenic Variation // *J Bacteriol*. 2007. T. 189, № 9. C. 3462–3470.
136. Cahoon L. A., Seifert H. S. An Alternative DNA Structure Is Necessary for Pilin Antigenic Variation in *Neisseria gonorrhoeae* // *Science*. 2009. T. 325, № 5941. C. 764–767.
137. Cahoon L. A., Seifert H. S. Transcription of a cis-acting, Noncoding, Small RNA Is Required for Pilin Antigenic Variation in *Neisseria gonorrhoeae* // *PLoS Pathog* / под ред. Nassif X. 2013. T. 9, № 1. C. e1003074.
138. Maizels N., Davis L. Initiation of homologous recombination at DNA nicks // *Nucleic Acids Research*. 2018. T. 46, № 14. C. 6962–6973.
139. Prister L. L., Ozer E. A., Cahoon L. A., Seifert H. S. Transcriptional initiation of a small RNA, not R-loop stability, dictates the frequency of pilin antigenic variation in *Neisseria gonorrhoeae* // *Molecular Microbiology*. 2019. T. 112, № 4. C. 1219–1234.
140. Linke R., Limmer M., Juranek S., Heine A., Paeschke K. The Relevance of G-Quadruplexes for DNA Repair // *IJMS*. 2021. T. 22, № 22. C. 12599.
141. Lyu B., Song Q. The intricate relationship of G-Quadruplexes and bacterial pathogenicity islands // *eLife*. 2024. T. 12. C. RP91985.
142. Shankar U., Mishra S. K., Jain N., Tawani A., Yadav P., Kumar A. Ni<sup>+2</sup> permease system of *Helicobacter pylori* contains highly conserved G-quadruplex motifs // *Infection, Genetics and Evolution*. 2022. T. 101. C. 105298.

143. Mishra S. K., Jain N., Shankar U., Tawani A., Sharma T. K., Kumar A. Characterization of highly conserved G-quadruplex motifs as potential drug targets in *Streptococcus pneumoniae* // *Sci Rep*. 2019. T. 9, № 1. C. 1791.
144. Molecular biology of the cell. 4th ed / под ред. Alberts B. New York: Garland Science, 2002.
145. Murray A. W., Kirschner M. W. Cyclin synthesis drives the early embryonic cell cycle // *Nature*. 1989. T. 339, № 6222. C. 275–280.
146. Liu B., K.H Yip R., Zhou Z. Chromatin Remodeling, DNA Damage Repair and Aging // *Curr. Genomics*. 2012. T. 13, № 7. C. 533–547.
147. Doenecke D. Chromatin Dynamics During the Cell Cycle // *Chromatin Regulation and Dynamics*. Elsevier, 2017. C. 117–139.
148. Poetsch A. R. The genomics of oxidative DNA damage, repair, and resulting mutagenesis // *Computational and Structural Biotechnology Journal*. 2020. T. 18. C. 207–219.
149. Bartek J., Lukas J. Mammalian G1- and S-phase checkpoints in response to DNA damage // *Current Opinion in Cell Biology*. 2001. T. 13, № 6. C. 738–747.
150. Zhou B.-B. S., Elledge S. J. The DNA damage response: putting checkpoints in perspective // *Nature*. 2000. T. 408, № 6811. C. 433–439.
151. Gospodinov A., Dzhokova S., Petrova M., Ugrinova I. Chromatin regulators in DNA replication and genome stability maintenance during S-phase // *Advances in Protein Chemistry and Structural Biology*. Elsevier, 2023. T. 135. C. 243–280.
152. Brown K. D., Rathi A., Kamath R., Beardsley D. I., Zhan Q., Mannino J. L., Baskaran R. The mismatch repair system is required for S-phase checkpoint activation // *Nat Genet*. 2003. T. 33, № 1. C. 80–84.
153. Bartek J., Lukas C., Lukas J. Checking on DNA damage in S phase // *Nat Rev Mol Cell Biol*. 2004. T. 5, № 10. C. 792–804.
154. Smits V. A. J., Alonso-de Vega I., Warmerdam D. O. Chromatin regulators and their impact on DNA repair and G2 checkpoint recovery // *Cell Cycle*. 2020. T. 19, № 17. C. 2083–2093.
155. Stark G. R., Taylor W. R. Analyzing the G2/M Checkpoint // *Checkpoint Controls and Cancer*. New Jersey: Humana Press, 2004. T. 280. C. 051–082.
156. Li J. J., Li S. A. Mitotic kinases: The key to duplication, segregation, and cytokinesis errors, chromosomal instability, and oncogenesis // *Pharmacology & Therapeutics*. 2006. T. 111, № 3. C. 974–984.
157. Okasaka T., Matsuo K., Suzuki T., Ito H., Hosono S., Kawase T., Watanabe M., Yatabe Y., Hida T., Mitsudomi T., Tanaka H., Yokoi K., Tajima K. hOGG1 Ser326Cys polymorphism and risk of lung cancer by histological type // *J Hum Genet*. 2009. T. 54, № 12. C. 739–745.
158. Peng Q., Lu Y., Lao X., Chen Z., Li R., Sui J., Qin X., Li S. Association between OGG1 Ser326Cys and APEX1 Asp148Glu polymorphisms and breast cancer risk: a meta-analysis // *Diagn Pathol*. 2014. T. 9, № 1. C. 108.
159. Floris M., Sanna D., Castiglia P., Putzu C., Sanna V., Pazzola A., De Miglio M. R., Sanges F., Pira G., Azara A., Lampis E., Serra A., Carru C., Steri M., Costanza F., Bisail M., Muroli M. R. MTHFR, XRCC1 and OGG1 genetic polymorphisms in breast cancer: a case-control study in a population from North Sardinia // *BMC Cancer*. 2020. T. 20, № 1. C. 234.
160. Zhan P., Wang Q., Wei S.-Z., Wang J., Qian Q., Yu L.-K., Song Y. ERCC2/XPD Lys751Gln and Asp312Asn Gene Polymorphism and Lung Cancer Risk: A Meta-Analysis Involving 22 Case-Control Studies // *Journal of Thoracic Oncology*. 2010. T. 5, № 9. C. 1337–1345.
161. Von Salomé J., Liu T., Keihäs M., Morak M., Holinski-Feder E., Berry I. R., Moilanen J. S., Baert-Desurmont S., Lindblom A., Lagerstedt-Robinson K. Haplotype analysis suggest that the *MLH1* c.2059C > T mutation is a Swedish founder mutation // *Familial Cancer*. 2018. T. 17, № 4. C. 531–537.
162. Dominguez-Valentin M., Nilbert M., Wernhoff P., López-Köstner F., Vaccaro C., Sarroca C., Palmero E. I., Giraldo A., Ashton-Prolla P., Alvarez K., Ferro A., Neffa F., Caris J., Carraro D. M., Rossi B. M. Mutation spectrum in South American Lynch syndrome families // *Hered Cancer Clin Pract*. 2013. T. 11, № 1. C. 18.

163. Pinheiro M., Francisco I., Pinto C., Peixoto A., Veiga I., Filipe B., Santos C., Maia S., Silva J., Pinto P., Santos R., Claro I., Lage P., Lopes P., Ferreira S., Rosa I., Fonseca R., Rodrigues P., Henrique R., Chaves P., Pereira A. D., Brandão C., Albuquerque C., Teixeira M. R. The nonsense mutation *MSH2* c.2152C→T shows a founder effect in Portuguese Lynch syndrome families // *Genes Chromosomes & Cancer*. 2019. T. 58, № 9. C. 657–664.
164. Hung R. J., Hall J., Brennan P., Boffetta P. Genetic Polymorphisms in the Base Excision Repair Pathway and Cancer Risk: A HuGE Review // *American Journal of Epidemiology*. 2005. T. 162, № 10. C. 925–942.
165. Saad A. M., Abdel-Megied A. E. S., Elbaz R. A., Hassab El-Nabi S. E., Elshazli R. M. Genetic variants of *APEX1* p.Asp148Glu and *XRCC1* p.Gln399Arg with the susceptibility of hepatocellular carcinoma // *Journal of Medical Virology*. 2021. T. 93, № 11. C. 6278–6291.
166. Xiao S., Cui S., Lu X., Guan Y., Li D., Liu Q., Cai Y., Jin C., Yang J., Wu S., Van Der Straaten T. The *ERCC2/XPD* Lys751Gln polymorphism affects DNA repair of benzo[a]pyrene induced damage, tested in an *in vitro* model // *Toxicology in Vitro*. 2016. T. 34. C. 300–308.
167. Dai J., Sanchez A., Adam C., Ranjha L., Reginato G., Chervy P., Tellier-Lebegue C., Andreani J., Guérois R., Ropars V., Le Du M.-H., Maloisel L., Martini E., Legrand P., Thureau A., Cejka P., Borde V., Charbonnier J.-B. Molecular basis of the dual role of the Mlh1-Mlh3 endonuclease in MMR and in meiotic crossover formation // *Proc. Natl. Acad. Sci. U.S.A.* 2021. T. 118, № 23. C. e2022704118.
168. Li W., Ma H. Double-stranded DNA breaks and gene functions in recombination and meiosis // *Cell Res*. 2006. T. 16, № 5. C. 402–412.
169. He W., Rao H. B. D. P., Tang S., Bhagwat N., Kulkarni D. S., Ma Y., Chang M. A. W., Hall C., Bragg J. W., Manasca H. S., Baker C., Verhees G. F., Ranjha L., Chen X., Hollingsworth N. M., Cejka P., Hunter N. Regulated Proteolysis of MutSγ Controls Meiotic Crossing Over // *Molecular Cell*. 2020. T. 78, № 1. C. 168-183.e5.
170. Rogacheva M. V., Manhart C. M., Chen C., Guarne A., Surtees J., Alani E. Mlh1-Mlh3, a Meiotic Crossover and DNA Mismatch Repair Factor, Is a Msh2-Msh3-stimulated Endonuclease // *Journal of Biological Chemistry*. 2014. T. 289, № 9. C. 5664–5673.
171. Ji G., Long Y., Zhou Y., Huang C., Gu A., Wang X. Common variants in mismatch repair genes associated with increased risk of sperm DNA damage and male infertility // *BMC Med*. 2012. T. 10, № 1. C. 49.
172. Pashaiefar H., Sheikha M. H., Kalantar S. M., Jahaninejad T., Zaimy M. A., Ghasemi N. Analysis of *MLH3* C2531T polymorphism in Iranian women with unexplained infertility // *Iran J Reprod Med*. 2013. T. 11, № 1. C. 19–24.
173. Stratigopoulou M., Van Dam T. P., Guikema J. E. J. Base Excision Repair in the Immune System: Small DNA Lesions With Big Consequences // *Front. Immunol*. 2020. T. 11. C. 1084.
174. Yu K. AID function in somatic hypermutation and class switch recombination // *ABBS*. 2022. T. 54, № 6. C. 759–766.
175. Zan H., Casali P. AID- and Ung-dependent generation of staggered double-strand DNA breaks in immunoglobulin class switch DNA recombination: A post-cleavage role for AID // *Molecular Immunology*. 2008. T. 46, № 1. C. 45–61.
176. Oliveira T. T., Coutinho L. G., De Oliveira L. O. A., Timoteo A. R. D. S., Farias G. C., Agnez-Lima L. F. *APE1/Ref-1* Role in Inflammation and Immune Response // *Front. Immunol*. 2022. T. 13. C. 793096.
177. Amouroux R., Campalans A., Epe B., Radicella J. P. Oxidative stress triggers the preferential assembly of base excision repair complexes on open chromatin regions // *Nucleic Acids Research*. 2010. T. 38, № 9. C. 2878–2890.
178. Flynn A., Waszak S. M., Weischenfeldt J. Somatic CpG hypermutation is associated with mismatch repair deficiency in cancer // *Mol Syst Biol*. 2024. T. 20, № 9. C. 1006–1024.
179. Sassa A., Beard W. A., Prasad R., Wilson S. H. DNA Sequence Context Effects on the Glycosylase Activity of Human 8-Oxoguanine DNA Glycosylase // *Journal of Biological Chemistry*. 2012. T. 287, № 44. C. 36702–36710.

180. Allgayer J., Kitsera N., Von Der Lippen C., Epe B., Khobta A. Modulation of base excision repair of 8-oxoguanine by the nucleotide sequence // *Nucleic Acids Research*. 2013. T. 41, № 18. C. 8559–8571.
181. Roychoudhury S., Pramanik S., Harris H. L., Tarpley M., Sarkar A., Spagnol G., Sorgen P. L., Chowdhury D., Band V., Klinkebiel D., Bhakat K. K. Endogenous oxidized DNA bases and APE1 regulate the formation of G-quadruplex structures in the genome // *Proc. Natl. Acad. Sci. U.S.A.* 2020. T. 117, № 21. C. 11409–11420.
182. Fan J., Matsumoto Y., Wilson D. M. Nucleotide Sequence and DNA Secondary Structure, as Well as Replication Protein A, Modulate the Single-stranded Abasic Endonuclease Activity of APE1 // *Journal of Biological Chemistry*. 2006. T. 281, № 7. C. 3889–3898.
183. Jones M., Wagner R., Radman M. Repair of a Mismatch Is Influenced by the Base Composition of the Surrounding Nucleotide Sequence // *Genetics*. 1987. T. 115, № 4. C. 605–610.
184. Kuznetsova A. A., Fedorova O. S., Kuznetsov N. A. Lesion Recognition and Cleavage of Damage-Containing Quadruplexes and Bulged Structures by DNA Glycosylases // *Front. Cell Dev. Biol.* 2020. T. 8. C. 595687.
185. Cogoi S., Ferino A., Miglietta G., Pedersen E. B., Xodo L. E. The regulatory G4 motif of the *Kirsten ras (KRAS)* gene is sensitive to guanine oxidation: implications on transcription // *Nucleic Acids Research*. 2018. T. 46, № 2. C. 661–676.
186. Broxson C., Hayner J. N., Beckett J., Bloom L. B., Tornaletti S. Human AP endonuclease inefficiently removes abasic sites within G4 structures compared to duplex DNA // *Nucleic Acids Res.* 2014. T. 42, № 12. C. 7708–7719.
187. Burra S., Marasco D., Malfatti M. C., Antoniali G., Virgilio A., Esposito V., Demple B., Galeone A., Tell G. Human AP-endonuclease (Ape1) activity on telomeric G4 structures is modulated by acetylable lysine residues in the N-terminal sequence // *DNA Repair*. 2019. T. 73. C. 129–143.
188. Davletgildeeva A. T., Kuznetsova A. A., Fedorova O. S., Kuznetsov N. A. Activity of Human Apurinic/Apyrimidinic Endonuclease APE1 Toward Damaged DNA and Native RNA With Non-canonical Structures // *Front. Cell Dev. Biol.* 2020. T. 8. C. 590848.
189. Zhou J., Fleming A. M., Averill A. M., Burrows C. J., Wallace S. S. The NEIL glycosylases remove oxidized guanine lesions from telomeric and promoter quadruplex DNA structures // *Nucleic Acids Research*. 2015. T. 43, № 8. C. 4039–4054.
190. Bai G., Endres T., Kühbacher U., Mengoli V., Greer B. H., Peacock E. M., Newton M. D., Stanage T., Dello Stritto M. R., Lungu R., Crossley M. P., Sathirachinda A., Cortez D., Boulton S. J., Cejka P., Eichman B. F., Cimprich K. A. HLTf resolves G4s and promotes G4-induced replication fork slowing to maintain genome stability // *Molecular Cell*. 2024. T. 84, № 16. C. 3044-3060.e11.
191. Ehrat E. A., Johnson B. R., Williams J. D., Borchert G. M., Larson E. D. G-quadruplex recognition activities of *E. coli* MutS // *BMC Molecular Biol.* 2012. T. 13, № 1. C. 23.
192. Pavlova A. V., Monakhova M. V., Ogloblina A. M., Andreeva N. A., Laptev G. Yu., Polshakov V. I., Gromova E. S., Zvereva M. I., Yakubovskaya M. G., Oretskaya T. S., Kubareva E. A., Dolinnaya N. G. Responses of DNA Mismatch Repair Proteins to a Stable G-Quadruplex Embedded into a DNA Duplex Structure // *IJMS*. 2020. T. 21, № 22. C. 8773.
193. Perevoztchikova S. A., Romanova E. A., Oretskaya T. S., Friedhoff P., Kubareva E. A. Modern aspects of the structural and functional organization of the DNA mismatch repair system // *Acta Naturae*. 2013. T. 5, № 3. C. 17–34.
194. Isik E., Shukla K., Pospisilova M., König C., Andrs M., Rao S., Rosano V., Dobrovolna J., Krejci L., Janscak P. MutSβ-MutLβ-FANCD1 axis mediates the restart of DNA replication after fork stalling at cotranscriptional G4/R-loops // *Sci. Adv.* 2024. T. 10, № 6. C. eadk2685.
195. Duppatla V., Bodda C., Urbanke C., Friedhoff P., Rao D. N. The C-terminal domain is sufficient for endonuclease activity of *Neisseria gonorrhoeae* MutL // *Biochemical Journal*. 2009. T. 423, № 2. C. 265–277.
196. Namadurai S., Jain D., Kulkarni D. S., Tabib C. R., Friedhoff P., Rao D. N., Nair D. T. The C-Terminal Domain of the MutL Homolog from *Neisseria gonorrhoeae* Forms an Inverted Homodimer // *PLoS ONE* / под ред. Ahmed N. 2010. T. 5, № 10. C. e13726.

197. Laitaoja M., Valjakka J., Jänis J. Zinc Coordination Spheres in Protein Structures // *Inorg. Chem.* 2013. T. 52, № 19. С. 10983–10991.
198. Monakhova M., Ryazanova A., Kunetsky V., Li P., Shilkin E., Kisil O., Rao D. N., Oretskaya T., Friedhoff P., Kubareva E. Probing the DNA-binding center of the MutL protein from the *Escherichia coli* mismatch repair system via crosslinking and Förster resonance energy transfer // *Biochimie.* 2020. T. 171–172. С. 43–54.
199. Ban C., Yang W. Crystal Structure and ATPase Activity of MutL // *Cell.* 1998. T. 95, № 4. С. 541–552.
200. Монахова М. В., Милакина М. А., Савицкая В. Ю., Романова Е. А., Rao D. N., Кубарева Е. А. Белок MutL из системы репарации мисматчей бактерии *Neisseria gonorrhoeae*: взаимодействие с АТФ и ДНК // *Молекул. биол.* 2021. T. 55, № 2. С. 289–304.
201. Pillon M. C., Miller J. H., Guarné A. The endonuclease domain of MutL interacts with the  $\beta$  sliding clamp // *DNA Repair.* 2011. T. 10, № 1. С. 87–93.
202. Fukui K., Baba S., Kumasaka T., Yano T. Structural Features and Functional Dependency on  $\beta$ -Clamp Define Distinct Subfamilies of Bacterial Mismatch Repair Endonuclease MutL // *Journal of Biological Chemistry.* 2016. T. 291, № 33. С. 16990–17000.
203. Pillon M. C., Babu V. M. P., Randall J. R., Cai J., Simmons L. A., Sutton M. D., Guarné A. The sliding clamp tethers the endonuclease domain of MutL to DNA // *Nucleic Acids Res.* 2015. T. 43, № 22. С. 10746–10759.
204. Savitskaya V. Yu., Strekalovskikh V. V., Snyga V. G., Monakhova M. V., Arutyunyan A. M., Dolinnaya N. G., Kubareva E. A. pilE G-Quadruplex Is Recognized and Preferentially Bound but Not Processed by the MutL Endonuclease from *Neisseria gonorrhoeae* Mismatch Repair Pathway // *IJMS.* 2023. T. 24, № 7. С. 6167.
205. Pavlova A. V., Savitskaya V. Yu., Dolinnaya N. G., Monakhova M. V., Litvinova A. V., Kubareva E. A., Zvereva M. I. G-Quadruplex Formed by the Promoter Region of the *hTERT* Gene: Structure-Driven Effects on DNA Mismatch Repair Functions // *Biomedicines.* 2022. T. 10, № 8. С. 1871.
206. Fernandez A., Zhou T., Lei Y., Liu N., Esworthy S., Shen C., Liu H., Hess J. D., Yuan H., Shi G., Zhou M., Shen L., Zhang S., Kosiyatrakul S., Gaur V., Sommers J. A., Srivastava N., Edelmann W., Li G.-M., Brosh Jr R. M., Chai W., Lee M. Y. W. T., Zhang D., Schildkraut C., Zheng L., Shen B. DNA2 and MSH2 cooperatively repair stabilized G4 and allow efficient telomere replication // *Nat. Commun.* 2025. T. 16, № 1. С. 8519.
207. Brzezinski P., Öjemyr L. N., Ädelroth P. Intermediates generated during the reaction of reduced *Rhodobacter sphaeroides* cytochrome c oxidase with dioxygen // *Biochimica et Biophysica Acta (BBA) - Bioenergetics.* 2013. T. 1827, № 7. С. 843–847.
208. Jayaraj A., Thayer K. M., Beveridge D. L., Hingorani M. M. Molecular dynamics of mismatch detection—How MutS uses indirect readout to find errors in DNA // *Biophysical Journal.* 2023. T. 122, № 15. С. 3031–3043.
209. Neupane A., Chariker J. H., Rouchka E. C. Structural and Functional Classification of G-Quadruplex Families within the Human Genome // *Genes.* 2023. T. 14, № 3. С. 645.
210. Maffia A., Ranise C., Sabbioneda S. From R-Loops to G-Quadruplexes: Emerging New Threats for the Replication Fork // *IJMS.* 2020. T. 21, № 4. С. 1506.
211. Wang K., Liu T., Ge N., Liu L., Yuan X., Liu J., Kong F., Wang C., Ren H., Yan K., Hu S., Xu Z., Björkholm M., Fan Y., Zhao S., Liu C., Xu D. *TERT* promoter mutations are associated with distant metastases in upper tract urothelial carcinomas and serve as urinary biomarkers detected by a sensitive castPCR // *Oncotarget.* 2014. T. 5, № 23. С. 12428–12439.
212. Kroeger K. M. A comprehensive comparison of DNA replication past 2-deoxyribose and its tetrahydrofuran analog in *Escherichia coli* // *Nucleic Acids Research.* 2004. T. 32, № 18. С. 5480–5485.
213. Panova V. V., Dolinnaya N. G., Novoselov K. A., Savitskaya V. Yu., Chernykh I. S., Kubareva E. A., Alexeevski A. V., Zvereva M. I. Conserved G-Quadruplex-Forming Sequences in Mammalian *TERT* Promoters and Their Effect on Mutation Frequency // *Life.* 2023. T. 13, № 7. С. 1478.

214. Savitskaya V. Yu., Novoselov K. A., Dolinnaya N. G., Monakhova M. V., Snyga V. G., Diatlova E. A., Peskovatskova E. S., Golyshev V. M., Kitaeva M. I., Eroshenko D. A., Zvereva M. I., Zharkov D. O., Kubareva E. A. Position-Dependent Effects of AP Sites Within an *hTERT* Promoter G-Quadruplex Scaffold on Quadruplex Stability and Repair Activity of the APE1 Enzyme // IJMS. 2025. T. 26, № 1. C. 337.