

Федеральное государственное бюджетное учреждение «Федеральный научно-клинический центр физико-химической медицины имени академика Ю.М. Лопухина Федерального медико-биологического агентства»

На правах рукописи

Павлова Юлия Ивановна

**G-КВАДРУПЛЕКСЫ СУПЕРЭНХАНСЕРОВ И ПРИГРАНИЧНЫХ  
УЧАСТКОВ ХРОМАТИНОВЫХ ДОМЕНОВ КАК РЕГУЛЯТОРЫ  
ТРАНСКРИПЦИИ**

1.5.3 – Молекулярная биология

Диссертация на соискание ученой степени кандидата биологических наук

Научный руководитель: Варижук Анна Михайловна,  
доктор химических наук

Москва – 2025

## ОГЛАВЛЕНИЕ

СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ .....	5
ВВЕДЕНИЕ.....	7
1. ОБЗОР ЛИТЕРАТУРЫ .....	11
1.1 G-квадруплексы в регуляции транскрипции .....	11
1.1.1 Картирование G4 в геноме.....	11
1.1.2 Промоторные G4 как активаторы транскрипции .....	13
1.1.3 G4 как эпигенетические регуляторы.....	15
1.1.4 G4 как участники энхансер-промоторной коммуникации.....	16
1.2 Биоконденсаты на G-квадруплексах.....	20
1.2.1 Склонные к образованию биоконденсатов белки в интерактоме G4 .....	22
1.2.2 G4 в ядерных конденсатах, связанных с процессингом РНК.....	24
1.2.3 G4 в гетерохроматине и конденсатах шелтерина .....	27
1.2.4 G4 в конденсатах, связанных с инициацией или элонгацией транскрипции и репарацией.....	28
1.2.5 Открытые вопросы.....	30
2. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ.....	32
2.1 G4 как участники демаркации транскрипционных (суб)доменов за счет привлечения CTCF .....	32
2.1.1 G4 колокализованы с сайтами связывания CTCF в геноме.....	32
2.1.2 G4 взаимодействуют с CTCF <i>in vitro</i> .....	35
2.2 G4 как фактор поддержания низкой нуклеосомной плотности.....	38
2.2.1 G4 частично сложены в составе модельной дуплексной конструкции с сайтом позиционирования нуклеосомы .....	38
2.2.2 G4 снижают эффективность сборки немодифицированного гистонового октамера и аналога с ацетилированными гистонами H3 на фланкирующем дуплексе.....	44
2.2.3 Известные G4-лиганды ингибируют сборку гистонового октамера на G4-фланкирующем дуплексе .....	53
2.3 G4 суперэнхансеров с низкой нуклеосомной плотностью как регуляторы транскрипции	57

2.3.1 G4 обогащены в SE человека и исключают нуклеосомы, но не BRD4 .....	57
2.3.2 G4 из SE проявляет умеренное сродство к BRD4 <i>in vitro</i> .....	64
2.3.3 Подобно ацетилированным нуклеосомам, G4 из SE способствует образованию биоконденсатов в растворах BRD4 .....	68
2.3.4 Биоконденсаты BRD4-G4 нечувствительны к ингибитору BRD4, но чувствительны к G4-лиганду.....	71
2.3.5 Гены, активируемые G4-богатыми SE, демонстрируют повышенную чувствительность к G4-лиганду.....	72
3. ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ .....	77
3.1 Методы исследования роли G4 в демаркации транскрипционных (суб)доменов и привлечении CTCF .....	77
3.1.1 Обработка клеточных культур G4-лигандами и иммунопреципитация хроматина (ChIP).....	77
3.1.2 ChIP-seq и анализ данных .....	78
3.1.3 ChIP с кПЦР.....	78
3.1.4 Биоинформатический анализ распределения сайтов связывания CTCF относительно G4.....	79
3.1.5 Олигонуклеотиды, рекомбинантный белок и низкомолекулярные лиганды .....	79
3.1.6 Спектроскопия кругового дихроизма и электрофорез.....	80
3.1.7 Микромасштабный термофорез с CTCF .....	80
3.2 Методы исследования G4 в составе конструкции с нуклеосомой.....	81
3.2.1 Получение ДНК-конструкции .....	81
3.2.2 Характеризация фрагментов конструкции методами спектроскопии кругового дихроизма (КД) и плавления .....	82
3.2.3 Подтверждение фолдинга G4 методом флуориметрии с тиофлавином Т (ThT).....	82
3.2.4 Подтверждение фолдинга G4 методом спектроскопии ядерного магнитного резонанса (ЯМР) .....	83
3.2.5 Сборка нуклеосом .....	83
3.2.6 Подтверждение сборки нуклеосом методом атомно-силовой микроскопии (АСМ)....	84
3.2.7 Электрофоретический анализ влияния лигандов на нуклеосому с G4.....	84

3.3 Методы исследования G4 суперэнхансеров в контексте их влияния на транскрипцию.....	86
3.3.1 Биоинформатический анализ SE с G4 и без G4 .....	86
3.3.2 Микромасштабный термофорез и поверхностный плазмонный резонанс с BRD4 .....	88
3.3.3 Сборка биоконденсатов и флуоресцентная микроскопия .....	89
3.3.4 Анализ экспрессии генов .....	90
ЗАКЛЮЧЕНИЕ .....	93
БЛАГОДАРНОСТИ .....	94
ЛИТЕРАТУРА.....	95

## СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

- BD – бромодомен (bromdomain)
- BET – семейство белков с бромодоменом (BD) и экстра терминальным (ET) доменом
- CAGE – кэп-анализ экспрессии генов (Cap analysis of gene expression)
- CBMPP – метод профилирования ДНК/РНК-белковых взаимодействий на основе одновременного взаимодействия нуклеиновых кислот с белком и фотоактивируемой пробой (co-binding-mediated protein profiling)
- CGI – CpG-островки (CpG islands)
- ChIP-seq – иммунопреципитация хроматина и секвенирование (chromatin immunoprecipitation and sequencing)
- ChIA-PET – метод анализа контактов хроматина секвенированием спаренных концов (Chromatin Interaction Analysis by Paired-End Tag Sequencing)
- CTCF – CCCTC-связывающий фактор (CCCTC-binding factor)
- CTD – С-концевой домен (C-terminal domain)
- CTE – С-концевая область (C-terminal extension)
- CSB – белок синдрома Коккейна типа В (Cockayne syndrome protein B)
- Cut&Tag – метод расщепления под мишенями и маркировки (Cleavage Under Targets and Tagmentation)
- DEAD-бокс – мотив аспарагин-глутамин-аланин-аспарагин
- DNMT1 – ДНК-метилтрансфераза 1 (DNA-methyltransferase 1)
- eCLIP – метод детекции ДНК/РНК-белковых взаимодействий и картирования сайтов связывания белков в нуклеиновых кислотах на основе кросс-сшивки и иммунопреципитации (enhanced cross-linking and immunoprecipitation)
- EMSA – анализ электрофоретической подвижности (Electrophoretic Mobility Shift Assay)
- FAIRE-seq – метод кросс-сшивки ДНК с белками хроматина при помощи формальдегида и для выделения регуляторных элементов с секвенированием (Formaldehyde-Assisted Isolation of Regulatory Elements sequencing)
- G4 – гуаниновый квадруплекс (G-quadruplex)
- HAT – ацетилтрансфераза гистонов (histone acetyl-transferase)
- hnRNP A1 – гетерогенный рибонуклеопротеин А1 (heterogeneous ribonucleoprotein A1)
- НО – гистоновый октамер (histone octamer)
- HR – шарнирная область (hinge region)
- IDR – конформационно неупорядоченные участки (intrinsically disordered regions)
- LCD – домены низкой сложности (low-complexity domains)

LLPS – разделение фаз «жидкость-жидкость» (liquid-liquid phase separation)  
MST – микромасштабный термофорез (microscale thermophoresis)  
NDR – области низкой нуклеосомной плотности (nucleosome depleted regions)  
NCL – нуклеолин (nucleolin)  
NPM – нуклеофосмин (nucleophosmin)  
NTE – N-концевая область (N-terminal extension)  
ORC – комплекса распознавания орижина репликации (origin recognition complex)  
PAR – поли(АДФ-рибоза)  
PARP1 – поли(АДФ-рибоза) полимеразы 1  
PDS – пиридостатин (pyridostatin)  
PIC – преинициаторный комплекс (preinitiation complex)  
RGG – аргинин-глицин-глицин-богатый мотив  
RRM – РНК-распознающий мотив (RNA-recognizing motif)  
ScFv – одноцепочечный вариабельный фрагмент (single-chain variable fragment)  
SE – суперэнхансер (superenhancer)  
SPR – поверхностный плазмонный резонанс (surface plasmon resonance)  
ThT – тиофлавин Т (thioflavin T)  
TSS – сайт начала транскрипции (transcription start site)  
АСМ – атомно-силовая микроскопия  
ГМВ – гигантские мембранные везикулы  
дцДНК – двухцепочечная ДНК  
КД – круговой дихроизм  
ОДН – олигодезоксирибонуклеотиды  
ОТ с кПЦР – количественная полимеразная цепная реакция с обратной транскрипцией  
оцДНК – одноцепочечная ДНК  
ПААГ – полиакриламидный гель  
п.н (т.п.н) – пара нуклеотидов (тысяча пар нуклеотидов)  
ТАД – топологически-ассоциированный домен  
ТФ – транскрипционные факторы  
ЯМР – ядерный магнитный резонанс

## ВВЕДЕНИЕ

**Актуальность темы исследования.** G-богатые нуклеиновые кислоты могут образовывать неканонические вторичные структуры – G-квадруплексы (G4). Ядро таких структур составляют стопки гуаниновых тетрад, в которых остатки гуанина соединены хугстиновскими водородными связями. Участие G4 в контроле транскрипции протоонкогенов общепризнано, и разработка противоопухолевых средств на основе G4-лигандов ведется несколько десятилетий с умеренным успехом (два соединения проходят клинические испытания). При этом в качестве мишеней рассматриваются исключительно промоторные G4, привлекающие транскрипционные факторы. Этот подход правомерен, но ограничен. По мере совершенствования методов картирования G4 становится очевидно, что помимо промоторов они перепредставлены в других цис-регуляторных элементах генома. Большая часть из них совпадает с границами хроматиновых петель, т.е. с горячими точками хромосомных контактов в транскрипционно активных хроматиновых доменах.

Хромосомные контакты в активном хроматине соответствуют взаимодействиям энхансер-промотор. Они включают ближние и дальние взаимодействия – следствие экструзии ДНК через когезин, ограниченной конвергентными сайтами позиционирования CCCTC-связывающего белка-инсулятора (CTCF). Среди энхансеров выделяются наиболее протяженные, богатые сайтами связывания транскрипционных факторов и маркированные бромодоменсодержащим белком BRD4. Они получили название суперэнхансеров (SE) и имеют две отличительные особенности. Первая – механизм аккумуляции транскрипционных факторов, основанный на эффекте разделения фаз «жидкость-жидкость» при участии BRD4. Вторая особенность SE – способность контактировать одновременно с несколькими промоторами, что обуславливает их доминирующий вклад в активацию транскрипции. Границы областей (хроматиновых доменов), в пределах которых возможны контакты энхансер-промотор, маркированы кластерами CTCF.

На момент начала диссертационного исследования были опубликованы данные по обогащению G4 в сайтах позиционирования CTCF, содержащих или не содержащих консенсусный CCCTC-богатый мотив связывания. Поскольку помимо наличия этого мотива вероятность привлечения CTCF зависит от локальной нуклеосомной плотности и CpG-метилирования, наблюдаемая колокализация CTCF с G4 в теории объяснялась ингибиторным действием G4 на работу ДНК-метилтрансферазы и нуклеосомную сборку. Впоследствии гипотеза о роли метилирования не подтвердилась, а влияние G4 на нуклеосомную укладку требовало экспериментальной проверки.

В то же время появились прямые свидетельства роли G4 в привлечении CTCF и перераспределении областей контактов энхансер-промотор вокруг сайта инсерции G4, хотя

механизмы оставались неясны [1]. Их прояснение было необходимо для дальнейшего развития перспективного направления 3D-геномики – контроля транскрипции за счет ремоделирования хроматина путем таргетирования неклассических вторичных структур ДНК. В частности, актуальными задачами были проверка непосредственного связывания G4-CTCF и влияния на него G4-лигандов, а также прояснение альтернативных механизмов участия G4 в ремоделировании хроматина, связанных с нуклеосомной плотностью.

Также на момент начала диссертационного исследования в рамках магистерской работы диссертантом были получены данные по колокализации геномных G4 с BRD4, что указывало на их вероятную повышенную представленность и роль в SE. Колокализация G4 с BRD4 в областях с низкой нуклеосомной плотностью представлялась парадоксальной, поскольку консенсусный механизм позиционирования данного белка на хроматине предполагает узнавание им ацетилированных фрагментов гистонов. Учитывая значимость SE для активации транскрипции, установление функции SE G4 (суперэнхансерных квадруплексов) и влияния на них G4-лигандов представлялось существенным для прогнозирования эффектов этих лигандов.

**Научная новизна и практическая значимость.** Диссертационное исследование было направлено на углубление представлений о G4-опосредованной регуляции транскрипции с учетом предполагаемой роли G4 в декомпактизации хроматина, SE-зависимой активации транскрипции и демаркации транскрипционно активных хроматиновых доменов. Решение данной научной проблемы создает задел для совершенствования стратегий управления транскрипцией с помощью G4-лигандов в рамках противоопухолевой терапии.

**Целью работы** было прояснение механизмов участия непромоторных G4 в регуляции транскрипции. Для достижения этой цели были поставлены следующие **задачи**:

- 1) Сравнить сродство CTCF к геномным G4 и консенсусному мотиву связывания в рамках проверки гипотезы прямого участия G4 в формировании CTCF-зависимых хроматиновых петель и демаркации хроматиновых доменов.
- 2) Охарактеризовать влияние G4-лигандов на позиционирование CTCF в геноме в рамках уточнения механизмов действия этих лигандов.
- 3) Оценить влияние G4 во фланкирующем участке сайта позиционирования нуклеосомы на сборку и устойчивость гистонового октамера в рамках прояснения ассоциации G4 с декомпактизацией хроматина.
- 4) Оценить представленность G4 в SE; сравнить сродство BRD4 к G4 и консенсусному сайту связывания в ацетилированных гистонах, а также разделить фазы в этих смесях для прояснения механизма удержания BRD4 в свободных от нуклеосом участках SE.

- 5) Охарактеризовать влияние G4-лигандов на связывание BRD4 с G4 и разделение фаз в растворах этих смесей, а также на экспрессию генов под контролем G4-богатых SE для уточнения механизмов действия лигандов.

**Положения, выносимые на защиту:**

- G4 имеют высокое сродство к CTCF, и обработка клеток G4-лигандами PDS и PhenDC3 повышает частоту связывания CTCF с G4.
- G4 затрудняют сборку нуклеосом и дестабилизируют примыкающие к ним нуклеосомы, в том числе с ацетилированными гистонами.
- G4 в свободных от нуклеосом участках SE удерживают BRD4 и могут способствовать образованию конденсатов по механизму разделения фаз «жидкость-жидкость» *in vitro*.
- G4-лиганды PDS и SOP1812 ингибируют образование конденсатов G4-BRD4 *in vitro* и подавляют экспрессию генов под контролем G4-богатых SE.

**Вклад автора.** Автором лично или при непосредственном участии выполнено большинство экспериментов, включая получение данных методом КД-спектроскопии, флуориметрии с тиофлавином Т, микромасштабного термофореза, электрофореза, флуоресцентной микроскопии, иммунопреципитации хроматина, анализа выживаемости клеток, ПЦР, биоинформатического анализа и ChIP-seq. В работе активно использовались методы анализа биоконденсатов, разработанные при непосредственном участии автора в рамках ГЗ Шаперон 17.002.24.800.

**Публикации и апробация работы.** По материалам диссертации опубликовано 4 статьи в международных периодических изданиях. Основные положения диссертационной работы были представлены на следующих российских и международных конференциях: The International Society of Nucleosides, Nucleotides and Nucleic Acids, 26-28 August 2021, Virtual forum; INTERACTION WINS, (Innovative ThERApEutiC Targets In nOn-canonical Nucleic acids structures Winter InterNational School 2021), 23-26 November 2021, Virtual winter school; Итоговая научно-практическая конференция, посвященная присвоению ФГБУ «Федеральный научно-клинический центр физико-химической медицины ФМБА России» имени академика РАН Ю. М. Лопухина, 21-22 декабря 2021 г, Москва; Научная конференция молодых ученых ФГБУ ФНКЦ ФБМА России, 20-21 апреля 2022 г, Москва; III объединенный научный форум физиологов, биохимиков и молекулярных биологов. 3-7 октября 2022 г, Сочи; III межвузовская студенческая конференция. Студенческий биохимический форум – 2023. 12–13 марта 2023 г, Москва; 65-я Всероссийская научная конференция МФТИ в честь 115-летия Л. Д. Ландау, 3–8 апреля 2023 г, Москва; Биокатализ – 2023. 13-ая Международная научная конференция. 25-29 июня 2023 г, Суздаль; Научная конференция молодых учёных ФБГУ ФНКЦ ФХМ им. М.Ю. Лопухина ФБМА России, 21-22 мая 2024 г, Москва.

**Структура и объем диссертации.** Диссертационная работа состоит из введения, обзора литературы, полученных результатов и их обсуждения, описания материалов и методов исследования, выводов и списка цитируемой литературы, включающего 261 источник. Работа изложена на 112 страницах машинописного текста, содержит 25 рисунков, 7 таблиц.

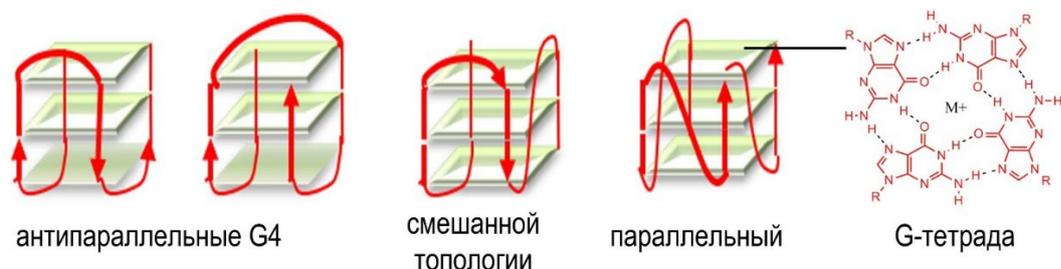
## 1. ОБЗОР ЛИТЕРАТУРЫ

### 1.1 G-квадруплексы в регуляции транскрипции

#### 1.1.1 Картирование G4 в геноме

G-богатые нуклеиновые кислоты могут складываться в неканонические структуры – G-квадруплексы (G4), содержащие стопки гуаниновых тетрад, в которых остатки гуанина соединены хугстиновскими водородными связями [2]. Общий мотив мономолекулярных G4 включает четыре G-тракта (G<sub>3+</sub>), разделенные последовательностями произвольного состава, образующими петли (Рисунок 1). Помимо мономолекулярных G4, возможно образование межмолекулярных G4, в том числе гибридных, образованных цепями ДНК и РНК. Взаимная ориентация G-трактов (параллельные/антипараллельные), анти/син-конформации гуанозинов в тетрадах и ориентация петель (пропеллерные/латеральные/диагональные) определяют тип топологии G4. Выявление G4-мотивов с короткими петлями и достаточно протяженными G-трактами (не менее трех остатков гуанинов подряд) – основа биоинформатического предсказания G4 в геномной ДНК.

Мономолекулярные G4 были предсказаны [3–7] в геномах различных организмов. В геноме человека обнаружено до нескольких сотен тысяч G4-мотивов; точное значение зависит от строгости критериев. Однако, лишь в части из них (по разным оценкам – от нескольких тысяч до десятков тысяч, в зависимости от клеточной линии) реализуется G4-фолдинг. Это обусловлено влиянием эпигенетических факторов, торсионного стресса и проч. на равновесие дцДНК-G4, а также активностью G4-специфичных хеликаз, которые раскручивают большую часть таких структур в ходе транскрипции и репликации. В этой связи активно развиваются методы экспериментального подтверждения G4-фолдинга – от простейших вариантов экспресс-детекции *in tube* до полногеномного картирования.



**Рисунок 1** – G-квадруплексные (G4) структуры. Схематическое изображение.

Для *in vitro* детекции G4 в составе выделенных из природных источников или синтетических фрагментов ДНК используют метод диметилсульфатного футпринтинга,

спектроскопию кругового дихроизма, гель-электрофорез, а также флуориметрию с G4-специфичными зондами (низкомолекулярными флуорогенами). Некоторые выводы о наличии G4 в составе протяженной ДНК позволяют сделать метод удлинения праймера, метод остановки полимеразы, метод *in vitro* транскрипции [8,9].

Для детекции G4 в хроматине подходят метод иммунопреципитации хроматина и секвенирования (ChIP-seq) или иммуногистохимия [10,11] с антителами к G4. Активно используются антитела BG4, 1H6 [12], HF2, а также специфичный к параллельным G4 одноцепочечный вариабельный фрагмент (ScFv) D1 [13] и синтетическое антитело малого размера G4P [14]. Относительно недавно для картирования G4 была адаптирована технология расщепления под мишенями и маркировки (Cut&Tag) [15]. Как и классический вариант ChIP-seq, она предусматривает выявление геномных фрагментов, узнаваемых антителом (обычно BG4), но благодаря селективному маркированию этих фрагментов за счет связывания антитела и слитой с белком А транспозазы в теории обеспечивает более высокую точность [16]. Альтернативой ChIP-seq и CUT&Tag является метод G4assess, не требующий использования антител. Он предполагает стабилизацию и выделение G4 из открытого хроматина с помощью низкомолекулярных лигандов [17]. Хотя методы на основе использования антител признаны наиболее надежными, они имеют свои ограничения. Среди G4, картированных методом ChIP-seq с антителами BG4 и D1, перепредставлены G4 с длинными петлями, которые в экспериментах *in tube* показывают достаточно низкую термодинамическую устойчивость. Это может быть отнесено к возможности антител стабилизировать G4 или к разительному отличию условий для стабильности G4 *in tube* и *in cellulo*.

В некоторых случаях целесообразно проверить результаты картирования непосредственной визуализацией G4 в ДНК/РНК; для решения этой задачи подходит атомно-силовая микроскопия (АСМ) [18]. Задачи внутриклеточного имиджинга G4-ДНК/РНК решаются с помощью флуорогенных (light-up) проб [10,19–21], в том числе способных визуализировать G4 не только в открытом хроматине, но также в стресс-гранулах и ядрышках [20]. Помимо низкомолекулярных light-up проб такую функцию может выполнять белковый зонд [22]. Несмотря на быстрый прогресс в этой области, в целом на данный момент вопрос внутриклеточного распределения РНК/ДНК-G4 и их фолдинга/расплетения исследован недостаточно подробно.

Распределение G4 в геноме, напротив, изучено относительно полно. Иммунопреципитация хроматина выявила ассоциацию G4 с "горячими точками" повреждения ДНК [23] и обилие G4 в регуляторных областях, включая промоторы и 5'-UTR [24,25], теломеры [13], границы хроматиновых доменов [26], энхансеры, консервативные сайты связывания транскрипционных факторов, сайты, гиперчувствительные к нуклеазам (NHS) [9] и др. Такое

распределение указывает на значимость G4 для фундаментальных клеточных процессов [27]: репарации, транскрипции, репликации, поддержания целостности генома и ремоделирования хроматина и т.д. [28–32]. При изучении участия G4 в транскрипции наибольшее внимание уделяется промоторным G4 [25,33].

### ***1.1.2 Промоторные G4 как активаторы транскрипции***

В геноме человека у 42.7% генов обнаруживается как минимум один G4-мотив на расстоянии не более 1 т.п.н. перед сайтом начала транскрипции (TSS). Также G4-мотивы часто находятся на расстоянии до 500 п.н. после TSS. Обогащение промоторных участков и 5'-нетранслируемых областей потенциальными квадруплексными сайтами указывает на вероятность участия G4 в регуляции транскрипции. Влияние на транскрипцию должно проявляться более явно при стабилизации G4 – например, за счет действия экзогенных низкомолекулярных лигандов. Одним из наиболее специфичных пан-квадруплексных лигандов (т.е. действующих на все G4, но не на иные вторичные структуры) является бис(хинолинил)пиримидиновое производное пиридостатин (PDS). С использованием данного лиганда было показано, что стабилизация G4 оказывает разнонаправленное действие на гены, т.е. может повышать или понижать экспрессию [34]. Снижение экспрессии при наличии G4 в темплатной цепи обычно связывают с торможением РНК-полимеразы II на G4, что подтверждается колокализацией ChIP-seq-пиков РНК-полимеразы II с G4. Однако, эта колокализация может быть интерпретирована двояко: либо как остановка РНК-полимеразы II на G4, либо как активная транскрипция. Исследования, выполненные методами удлинения праймера, остановки полимеразы и *in vitro* транскрипции (с минимальным набором транскрипционных факторов), показывают, что G4 действительно представляют собой барьер для ДНК/РНК-полимераз, т.е. останавливают репликацию и транскрипцию. Кроме варианта с G4 на темплатной цепи, остановке полимеразы может способствовать межмолекулярный ДНК-РНК G4.

Несмотря на это, роль G4 в транскрипции нельзя определить как сугубо негативную. Более того, G4 обычно встречаются именно в активно транскрибируемых областях [25,35] (особенно в нетемплатной цепи) и скорее ассоциированы с высоким уровнем транскрипции [36,37], что нельзя объяснить исключительно упрощением G4-фолдинга при локальном расплетении дуплекса. Препятствуя переотжигу цепей, G4 повышают вероятность повторных раундов транскрипции с темплатной цепи, пока она остается одноцепочечной [29].

Сравнение экспрессии одних и тех же генов в клеточных линиях с различным G4-профилем показывает, что сложенным G4 чаще соответствует более высокий уровень

транскрипции, и он тем выше, чем ближе промоторные G4 к сайту старта транскрипции (TSS). В ряде опухолевых клеточных линий (например, в клетках липосаркомы) промоторные G4 неизменно ассоциированы с высоким уровнем транскрипции генов в открытом хроматине [25].

Это наблюдение также может быть интерпретировано двояко, а именно (1) сложенные G4 привлекают транскрипционные факторы или (2) взаимодействие с этими транскрипционными факторами способствует фолдингу G4. Не исключено, что реализуются оба варианта.

Среди транскрипционных факторов, ассоциированных с G4, наиболее известны представители семейства «цинковые пальцы», такие как SP1 и MAZ. Для SP1 сродство к G4, сопоставимое со сродством к конесенсусному дуплексному мотиву связывания, было показано на примере G4 промотора *c-KIT* [37]; для MAZ – на примере G4 промотора *KRAS* [38]. В обоих случаях связывание транскрипционных факторов с промоторным G4 активировало генную экспрессию. Помимо белков, обладающих специфичностью к дцДНК-мотивам, в числе G4-связывающих транскрипционных факторов отмечены связывающие оцДНК, такие как PC4 (hSub1) [39] и РНК-связывающие, такие как нуклеолин (NCL) [40]. Про NCL известно следующее: при взаимодействии с G4 промотора *VEGF* он усиливает экспрессию гена, а при взаимодействии с промотором *c-MYC* ингибирует (предположительно, за счет вытеснения другого транскрипционного фактора) [41].

G4 участвуют не только в инициации транскрипции. Так, при элонгации образующиеся в R-петлях G4 стабилизируют эти петли и модулируют вероятность их повторного формирования на следующих раундах транскрипции [42,43].

Необходимое условие для положительного влияния G4 на экспрессию генов и репликацию – корректная работа хеликаз: поскольку G4 вызывают остановку как РНК-, так и ДНК-полимераз [44], их своевременное расплетение [45,46] требуется для того, чтобы избежать повреждений ДНК [46,47]. Хеликазы, которые эффективно расплетают G4, относятся преимущественно к семейству RecQ и имеют характерный С-концевой домен, ответственный за узнавание нуклеиновых кислот – домен RQC [48]. Поломка в хеликазе такого типа BLM приводит к нестабильности генома, повышению частоты рекомбинации и развитию синдрома Блума [49]. Мутации, приводящие к потере функций хеликазы WRN, способствуют развитию синдрома Вернера.

Роль G4 в геномных перестройках также отчасти объясняет ассоциацию их повышенного содержания с риском развития онкозаболеваний [50]. Ввиду недостатка активности ферментов системы репарации, онкоклетки в большей степени чувствительны к повреждениям генома. Кроме того, промоторные и непромоторные G4 регулируют экспрессию ряда ключевых онкогенов и генов, ответственных за развитие резистентности к химиотерапевтическим препаратам, что делает их перспективной терапевтической мишенью [51]. Точные механизмы

влияния G4 на экспрессию генов требуют дальнейшего прояснения. В следующих разделах рассмотрены механизмы, в которых задействованы непромоторные G4, и учитывается связь экспрессии с 3D-структурой хроматина.

### ***1.1.3 G4 как эпигенетические регуляторы***

По-видимому, G4 регулируют множество процессов, связанных с организацией хроматина и поддержанием эпигенетического статуса [52–54]. G4 вносят вклад в эпигенетическую регуляцию [32,55], определяя метилом ДНК [56] и способствуя образованию хроматиновых петель посредством таких архитектурных белков хроматина семейства «цинковые пальцы» как CCCTC-связывающий фактор (CTCF) и Yin Yang 1 (YY1). Первый известен как основной белок-инсулятор. Конвергентная пара CTCF ограничивает экструзию ДНК через когезиновый комплекс и служит основанием хроматиновой петли, конфигурация которой определяет 3D-контакты, включая контакты энхансер-промотор, а кластеры CTCF разграничивают и изолируют друг от друга области этих контактов – хроматиновые домены [26]. Второй архитектурный белок хроматина, YY1 выполняет сходную функцию за счет димеризации, но обеспечивает преимущественно 3D-контакты промоторов генов – регуляторов развития [57]. Пики G4, полученные в результате иммунопреципитации G4-хроматина и секвенирования (G4-ChIP-seq), демонстрируют статистически значимую колокализацию с сайтами связывания обоих белков, но наиболее выражена она в случае CTCF [26]. Однако роль G4 в рекрутировании CTCF до сих пор не объяснена.

Как и другие транскрипционные факторы с доменами вида «цинковый палец», CTCF обладает специфичностью к последовательности ДНК. В CTCF 11 доменов «цинковый палец» расположены тандемно, все они вносят вклады в специфичность, но эти вклады неравноценны, а сайты связывания белка в геноме довольно разнообразны [58]. Мотив, который наилучшим образом описывает эти сайты, представляет собой 20-мерный участок с ядром CCGCGNGGNGGCAG в 14 п.н. [59]. Другой консенсусный мотив, NCANNAG(G/A)NGGC(G/A)(C/G)(T/C), был обнаружен в экспериментах по иммунопреципитации хроматина с последующим расщеплением экзонуклеазой [60]. Он присутствует в большинстве ChIP-seq-пиков CTCF и частично соответствует последовательности из пяти триплетов, предсказанной на основе специфичности связывания ДНК доменами CTCF вида «цинковый палец» № 3-7 (CCAGCAGGGGGGCGCT) [61]. Ни одна из этих последовательностей не соответствует консенсусному G4-мотиву  $G_3+N_{1-7}G_3+N_{1-7}G_3+N_{1-7}G_3+$ , но все они G-богаты и могут перекрываться с G4-сайтами.

Связь между нуклеосомной укладкой и CTCF до сих пор не была всесторонне проанализирована. Умеренная плотность нуклеосом и их регулярное позиционирование [62,63] необходимы для того, чтобы CTCF мог получить доступ к своим сайтам связывания в линкерной (межнуклеосомной) ДНК. Можно предположить, что участки высокой нуклеосомной плотности окажутся CTCF-обедненными. Важно отметить, что CTCF сам по себе является фактором ремоделирования хроматина, способным репозиционировать нуклеосомы [64]. Таким образом, рекрутирование CTCF в ДНК регулируется несколькими взаимосвязанными путями, и G4 могут вносить свой вклад в каждый из них. Например, G4 могут способствовать ремоделированию хроматина благодаря своему сродству к архитектурным белкам и модификаторам хроматина, включая субъединицы репрессивного комплекса поликомб 2 [65,66].

Связь между образованием G4 и нуклеосомной плотностью представляется более определенной. G4 преобладают в регионах, обедненных нуклеосомами. Это подтверждено на нормальной линии кератиноцитов человека методом кросс-шивки ДНК с белками хроматина при помощи формальдегида и выделения регуляторных элементов с секвенированием (FAIRE-seq). Около 98% пиков G4 ДНК, определенных методом ChIP-seq с BG4, перекрываются областями низкой нуклеосомной плотности (nucleosome depleted regions, NDR) [24].

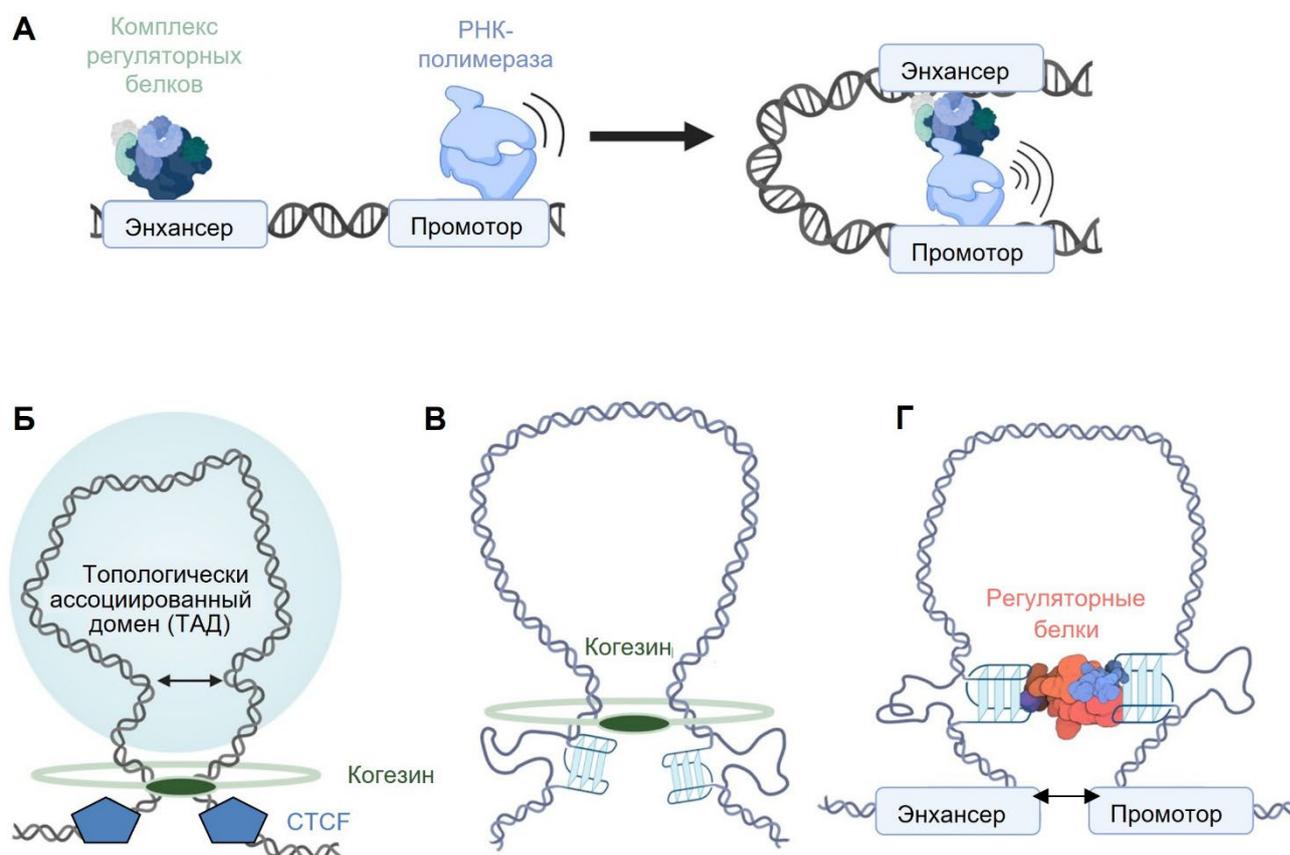
Эти наблюдения можно объяснить следующим образом. С одной стороны, освобождение от гистонов повышает вероятность раскручивания дуплекса и формирования G4. С другой стороны, G4 нарушают паттерн контактов ДНК с гистонами и, предположительно, исключают нуклеосомы [67,68]. Насколько наличие G4 сказывается на позиционировании нуклеосом на планирующих участках дДНК еще предстоит выяснить. Межнуклеосомные G4 всё же встречаются в некоторых типах геномных областей, характеризующихся регулярным расположением нуклеосом и умеренной нуклеосомной плотностью (например, вокруг сайтов начала транскрипции [3] или сайтов связывания CTCF [26]). Такие G4 могут влиять на модификации гистонов в близлежащих нуклеосомах напрямую, т.е. рекрутируя факторы ремоделирования хроматина [55,69], или косвенно – вызывая торможение полимеразы на этапе репликации, нарушая тем самым своевременную рециркуляцию гистонов и вызывая потерю гистонного кода [70]. Кроме того, G4 могут влиять на межнуклеосомные взаимодействия, опосредованные линкерными гистонами и определяющие (де)компактизацию хроматина, предположительно через связывание с линкерным гистоном или его аналогами [71].

#### ***1.1.4 G4 как участники энхансер-промоторной коммуникации***

Участие G4 в энхансер-промоторной коммуникации подтверждают следующие данные: обогащение G4 в энхансерах, выявленное методом ChIP-seq [72]; корреляция между G4 и

дальними контактами ДНК в хроматиновых доменах или на их границах, выявленная методом определения конформации хромосом (3C, chromosome conformation capture) [26]; корреляция между G4 и зависимыми от РНК-полимеразы II контактами ДНК, выявленная с помощью анализа взаимодействия хроматина секвенированием спаренных концов (Chromatin Interaction Analysis by Paired-End Tag Sequencing, ChIA-PET) и др. [73–75].

Как было отмечено выше, вклад G4 в энхансер-промоторную коммуникацию может осуществляться за счет привлечения белков укладки хроматина, включая CTCF [69,76], который ограничивает экструзию ДНК через когезиновый комплекс [77] (Рисунок 2 А,Б), или YY1 [57], который стабилизирует хроматиновые петли за счет димеризации [78]. Недавние исследования 3D-архитектуры генома с высоким разрешением позволяют предположить, что эти белки не являются основными для поддержания контактов энхансер-промотор [79]: горячие точки таких контактов, обогащенные G4, сохраняются в отсутствие белка-инсулятора. Гипотетические механизмы, объясняющие не зависящий от CTCF вклад G4 в энхансер-промоторную коммуникацию, представлены на Рисунке 2В,Г. Далее рассмотрены более подробно молекулярные основы энхансер-промоторной коммуникации и биофизические аспекты этой коммуникации – фазовые переходы в местах контактов.



**Рисунок 2** – G4 в энхансер-промоторной коммуникации. (А) Общая схема энхансер-промоторной коммуникации: выпетливание хроматина и сближение сконцентрированных на энхансере регуляторных белков с РНК-полимеразой II, удерживаемой на промоторе основными транскрипционными факторами, обеспечивает сборку полного инициаторного комплекса и

активную транскрипцию гена под данным промотором. **(Б)** CTCF/когезин-зависимое разграничение областей энхансер-промоторной коммуникации – хроматиновых доменов, также известных как топологически ассоциированные домены (ТАДы), или субдоменов. При наличии нескольких промоторов/энхансеров в пределах домена возможны любые парные контакты за счет экстрюзии ДНК через когезин. CTCF ограничивает экстрюзию; единичные конвергентные сайты связывания CTCF маркируют горячие точки контактов, а кластеры CTCF изолируют промоторы в пределах ТАДа от прочих энхансеров. **(В)** Гипотетическое когезин-зависимое CTCF-независимое формирование хроматиновых петель при участии G4: механическое ограничения экстрюзии ДНК в G4-сайтах способствует сближению промоторов и энхансеров, содержащих G4 или фланкированных ими. **(Г)** Аналогичное сближение промоторов и энхансеров при G4 без участия когезина, но с формированием «мостиковых контактов» через G4-связывающие белки [28].

Коммуникация энхансер-промотор поддерживается медиаторным комплексом [80], который взаимодействует с РНК-полимеразой II и другими факторами преинициаторного комплекса (PIC) и коактиваторами. В результате этих взаимодействий образуются многокомпонентные белковые ассоциаты или конденсаты [81], предположительно путем разделения фаз «жидкость-жидкость» (LLPS) [82]. Они удерживаются вместе за счет взаимодействий между структурированными белковыми доменами и внутренне неупорядоченными областями (IDR). Важно отметить, что медиаторный комплекс не контактирует с хроматином напрямую. Вместо этого он связывается с транскрипционными факторами в энхансерах/промоторах, которые распознают определенные последовательности ДНК или модификации гистонов (как правило, через структурированные домены) и образуют множественные транзистентные контакты с медиаторным комплексом (как правило, через IDR) [80]. Таким образом, позиционирование медиаторного комплекса в энхансерах/промоторах одновременно зависит от наличия транскрипционных факторов и способствует их дальнейшему накоплению. В протяженных энхансерах этот процесс (сродни положительной обратной связи) приводит к образованию крупных транскрипционных «хабов». Вопрос о том, всегда ли в местах контактов энхансер-промотор наблюдается LLPS, остается дискуссионным [80]. Однако, в суперэнхансерах (SE), т.е. длинных энхансерах, содержащих несколько сайтов связывания транскрипционных факторов и взаимодействующих с несколькими промоторами, LLPS можно считать доказанным [83].

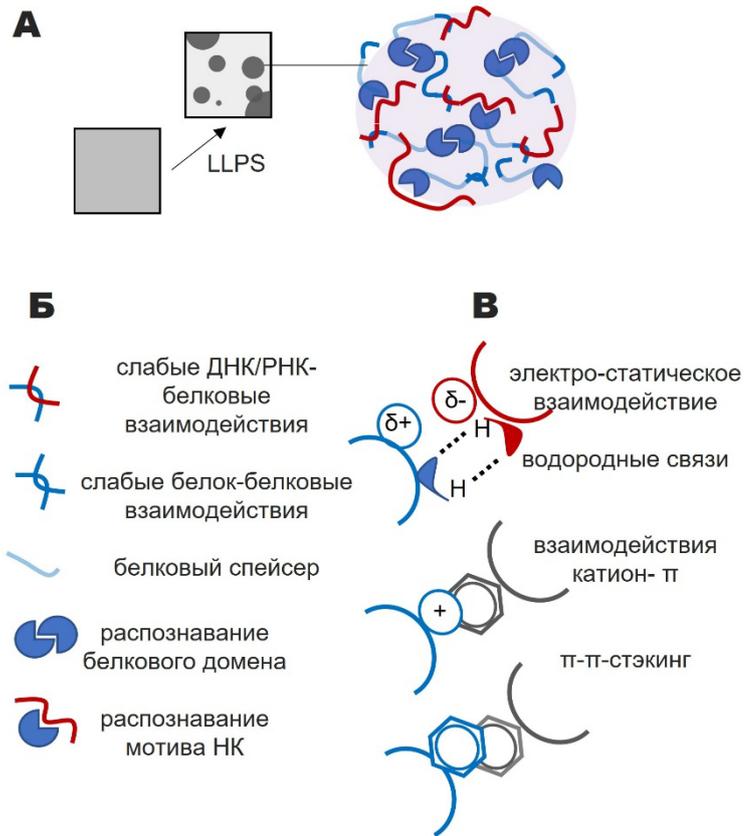
Из всех транскрипционных факторов, участвующих в рекрутировании субъединиц медиаторного комплекса, BRD4, член семейства белков с бромодоменами (BD) и экстратерминальным (ET) доменом (BET), выделяется повышенной представленностью в SE [84] и склонностью к LLPS [83]. BRD4 взаимодействует с субъединицами медиаторного комплекса благодаря своему С-концевому IDR [83] и распознает ацетилированные хвосты гистонов через один или оба своих тандемных BD [85]. Также не исключен альтернативный вариант связывания, при котором первый BD взаимодействует с ацетилированным гистоном, а второй BD слабо

связывается с ДНК. Этот вариант был продемонстрирован для гомолога BRD4 BRDT, но в теории может быть реализован со всеми белками семейства BET [86]. В любом случае, размещение гистонового хвоста в кармане связывания ацетилованного лизина по крайней мере одного BD необходимо для связывания BRD4 с хроматином, а образующиеся комплексы обладают низкой или умеренной стабильностью при биологически значимых концентрациях, в зависимости от положения ацетилованного остатка на гистоновом хвосте [85]. Поскольку в промоторах и энхансерах наблюдаются разные корреляции между модификациями гистонов и обогащением BRD4 [87], в дело могут вступать и другие факторы, помимо ацетилирования хроматина. Кроме того, причинно-следственная связь между BRD4 и ацетилированием хроматина неоднозначна, учитывая, что бромодомены BRD4 опосредуют рекрутирование ацетилтрансфераз гистонов (HAT) [88]. Наконец, BRD4 проявляет собственную HAT-активность [89]. Его основными мишенями являются гистоны H3 и H4, и ацетилирование первого из них по K122 приводит к разборке нуклеосом [89].

Если BRD4 ремоделирует хроматин, уменьшая нуклеосомную плотность, то молекулярная основа его поддержания в связанном с хроматином состоянии становится неясной. Более того, многие промоторы, энхансеры и SE изначально имеют низкую нуклеосомную плотность из-за большого количества G4, которые, как уже упоминалось, исключают нуклеосомы [17,24,90,91]. Даже если оставшиеся нуклеосомы ацетилированы, их низкая плотность несовместима с поддержанием кластеров BRD4 в G4-богатых SE, что необходимо для LLPS. Таким образом, наше представление о позиционировании BRD4 в энхансерах/промоторах в целом и в G4-богатых SE в частности – неполное. Поскольку для G4 была показана колоколизация с некоторыми белками семейства BET [92], и для модельных G4 показано сродство к BET-белку [66], можно предположить непосредственное участие G4 в поддержании уровня BRD4 на SE [93]. Эта гипотеза проверялась в части 2.3 данной работы. Она напрямую связана с вопросом участия G4 в образовании конденсатов при SE и в целом влияния G4 на LLPS.

## 1.2 Биоконденсаты на G-квадруплексах

Биомакромолекулярные конденсаты, также известные как немембранные органеллы, собираются спонтанно в результате разделения фаз «жидкость-жидкость» (liquid-liquid phase separation, LLPS) в растворах биополимеров [94–96] – как правило, нуклеиновых кислот и белков с доменами низкой сложности (LCD). Обозначение LCD отражает особенности первичной структуры, а именно повышенное содержание одинаковых или однотипных аминокислотных остатков. Оно является устоявшимся, хотя применимость термина «домен» спорна ввиду отсутствия у LCD определенной вторичной/третичной структуры. Более корректным представляется родственное понятие конформационно неупорядоченных участков (IDR) [94,97]. Белки с IDR и нуклеиновые кислоты склонны к слабым мультивалентным взаимодействиям (Рисунок 3), которые в определенном диапазоне концентраций полимера и при определенных значениях внешних параметров (температура, ионная сила и проч.) энергетически более выгодны, чем взаимодействия растворитель-биополимер, способствуя разделению раствора на две фазы: обедненную полимером и конденсированную жидкую фазу [98–101]. В цитоплазме или нуклеоплазме возникающее при LLPS увеличение локальной концентрации макромолекул может способствовать сборке многокомпонентных комплексов, облегчать распознавание ферментом субстрата и увеличивать скорость реакции [102]. В то же время, LLPS-опосредованная компартментализация цитоплазмы или нуклеоплазмы позволяет обратимо изолировать избыточные или токсичные биополимеры от основной среды [103].



**Рисунок 3** – Состав биомолекулярных конденсатов и предполагаемые контакты G4 внутри конденсатов. (А) Схематическое изображение образования конденсата путем разделения фаз жидкость-жидкость (LLPS). (Б) Типичные компоненты конденсатов и их взаимодействия. (В) Основные типы контактов, характерные для LLPS.

В последнее время G4-РНК стали рассматриваться как вероятные активаторы LLPS и иных фазовых переходов (образования гелей и нерастворимых белковых агрегатов) в цитоплазме [104]. Доказательства их вклада в сборку стрессовых гранул и токсичных агрегатов, связанных с нейродегенеративными заболеваниями [105,106], усилили интерес к лигандам к G4-РНК как кандидатам в лекарственные препараты. Основные достижения в этой области обсуждались в недавних публикациях [107,108]. В ядре активаторами LLPS также могут выступать G4-ДНК: формирование G4 расширяет спектр возможных слабых взаимодействий с IDR белков. В случае дцДНК основным типом контактов с IDR являются электростатические взаимодействия, поскольку нуклеиновые основания в дуплексе стерически недоступны. В случае G4 внешние гуаниновые тетрады открыты для контактов «π-π» и «катион-π» с ароматическими и положительно заряженными аминокислотными остатками. Стабилизация G4 и экранирование внешней тетрады низкомолекулярными лигандами должны сказаться не только на узнавании G4 специфичными к ним конформационно упорядоченными белками (например, представителями семейства «цинковые пальцы»), но и на слабых взаимодействиях с IDR-содержащими белками. Примеры таких белков и значимость их взаимодействий с G4 для образования ядерных биоконденсатов рассмотрены ниже.

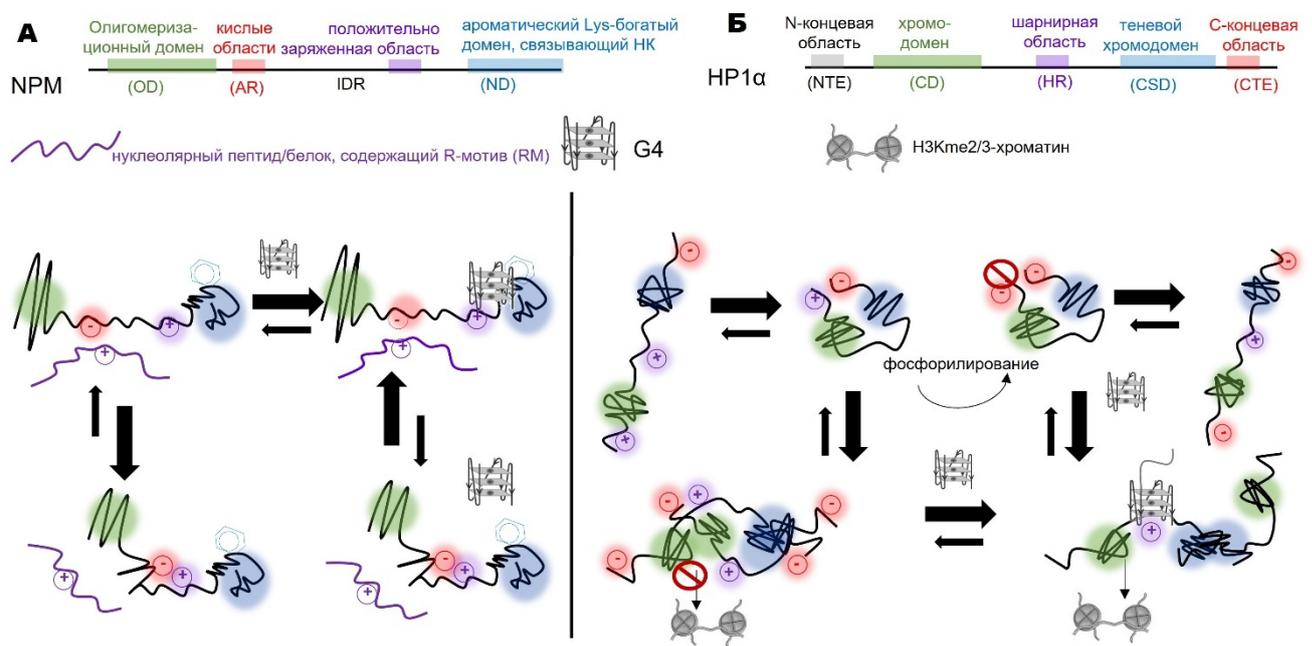
### 1.2.1 Склонные к образованию биоконденсатов белки в интерактоме G4

Значительное число белков, взаимодействующих с G4, было выявлено с помощью невысокопроизводительных методов, которые не полностью учитывают внутриклеточные условия [109]. Ранние исследования интерактома G4 основывались на профилировании G4 на микрочипах с белками человека [66] или аффинной очистке G4-связывающих белков из ядерных экстрактов с последующей их идентификацией и количественным анализом методами на основе масс-спектрометрии [65]. Функциональный анализ кандидатов на взаимодействие с G4, полученных в этих экспериментах, показал обогащение РНК-связывающими белками, факторами транскрипции или факторами процессинга РНК, хеликазами и факторами ремоделирования хроматина [110,111]. Большинство из них обобщены в базе данных белков, взаимодействующих с G4 [112]. Недавно разработанные альтернативные методы, основанные на кросс-сшивке и иммунопреципитации (eCLIP) или совместном связывании фотоактивируемых низкомолекулярных зондов (CBMPP) и G4-распознающих белков с G4 *in situ*, значительно расширили список известных G4-связывающих белков [113,114]. В дополнение к вышеупомянутым перекрывающимся функциональным категориям, eCLIP и CBMPP выявили преобладание в G4-интерактоме белков, связанных с ответом на повреждение ДНК и инфекционным ответом.

Установленные или предсказанные активаторы LLPS можно найти в каждой категории G4-связывающих белков. К ним относятся хеликазы с DEAD-боксом (консервативным мотивом Asp-Glu-Ala-Asp), такие как DDX1 и DDX24 [113], DDX5 [115], Dbp1 [116] и Dbp2 [117]. Эти хеликазы перемещаются между ядром, где они контролируют транскрипцию, и цитоплазмой, где они узнают вирусные нуклеиновые кислоты и участвуют в противовирусном ответе, например, способствуя сборке стрессовых гранул по механизму LLPS за счет своих IDR [118,119]. Помимо G4-ДНК, хеликазы с DEAD-боксами распознают РНК в зависимости от последовательности. Они считаются основными АТФ-зависимыми регуляторами всех РНК-содержащих биомолекулярных конденсатов [120,121]. Другие G4-распознающие активаторы LLPS из категории РНК-связывающих белков включают ингибитор активатора плазминогена SERBP1 [122], нуклеофосмин [123], нуклеолин (NCL) [124], а также гетерогенные ядерные рибонуклеопротеины A1 (hnRNP A1) [125,126] и FUS [127,128].

SERBP1 участвует в транскрипции, ремоделировании хроматина и сборке промиелоцитарных лейкозных телец в ядре, а в цитоплазме стабилизирует мРНК и контролирует трансляцию [129]. Этот белок был идентифицирован как G4-связывающий методами pull-down [122] и eCLIP [114]. Он содержит повторяющиеся мотивы Arg-Gly/Arg-Gly-Gly (RG/RGG), которые участвуют как в распознавании G-богатых последовательностей, так и в LLPS [130].

Нуклеофосмин, многофункциональный белок, играющий ключевую роль в биогенезе рибосом [131], распознает G4-ДНК преимущественно ароматическим С-концевым доменом и прилежащим IDR [132], контролирующим склонность этого белка к гомотипическим и гетеротипическим взаимодействиям [133]. Последние включают контакты IDR с аргинин-богатыми мотивами нуклеолярных белков и предположительно способствуют сборке ядрышка вокруг рРНК [131] (Рисунок 4). HnRNP A1 распознает G4 с помощью RGG-мотива. Вместе с прилегающим N-концевым доменом UP1, включающим два РНК-распознающих мотива (RRM), RGG-мотив расплетает G4, и эта хеликазоподобная активность значима для поддержания теломер и активации транскрипции [125,126,134]. Это является одной из основных функций hnRNPA1 в ядре – наряду с регуляцией процессинга (в первую очередь – сплайсинга) пре-мРНК, в то время как в цитоплазме hnRNPA1 способствует сборке стрессовых гранул, которые в конечном итоге могут превратиться в патогенные твердые агрегаты [135,136]. NCL, один из основных нуклеолярных белков, имеет ряд общих функций и структурных особенностей с hnRNPA1 [137]. Он содержит четыре RRM и RGG-богатый С-концевой IDR [138]. Эти мотивы участвуют в связывании G4 [139].



**Рисунок 4** – Примеры влияния G4 на конформационные переходы белков-активаторов LLPS. (А) Нуклеофосмин (NPM). Связывание с G4, опосредованное центральным конформационно неупорядоченным участком (IDR) и N-концевым доменом (ND) NPM, предотвращает внутримолекулярные контакты NPM и облегчает межмолекулярные, а именно - взаимодействия NPM с Arg-богатым мотивом ядрышковых белков. Эти взаимодействия лежат в основе LLPS при сборке ядрышек. (Б) Белок гетерохроматина 1α (HP1α). В связывании HP1α с G4 участвует центральная «шарнирная» область (HR), что предотвращает контакты ее с С-концевым участком (CTE) HP1α и стабилизирует раскрытую конформацию HP1α. В раскрытой

конформации HP1 $\alpha$  образует комплекс с гистонами H3Kme2/3. Эти взаимодействия лежат в основе LLPS и гетерохроматинизации.

Белок FUS, один из немногих транскрипционных факторов, обогащенных на эндогенных G4-ДНК и эффективно вытесняемых G4-специфичным лигандом PDS [35], также связывает G4 RGG-мотивами. Таких мотивов в FUS три, и они имеют так называемую  $\beta$ -спиральную структуру, стабилизированную соседними Pro- и Arg-богатыми участками [127]. Помимо распознавания G4 и неспецифического связывания нуклеиновых кислот посредством электростатических взаимодействий, RGG-мотивы способствуют LLPS FUS [140]. Как RGG-мотивы, так и N-концевой LCD, богатый ароматическими аминокислотами, значимы для LLPS FUS [141,142], поскольку метилирование RGG [143] или фосфорилирование LCD [144] препятствуют разделению фаз FUS в клетке. Как и в случае с hnRNPA1, LLPS и переходы FUS из плотной жидкой фазы в твердую в основном связаны с патологическими процессами в цитоплазме. Ядерно-цитоплазматический транспорт ограничивает фазовые переходы FUS в нормальных условиях благодаря шапероноподобной активности импортина [145]. Однако при повреждении ДНК наблюдается внутриядерный переход FUS в плотную жидкую фазу, что, по-видимому, является неотъемлемой частью инициации ответа на повреждение [146].

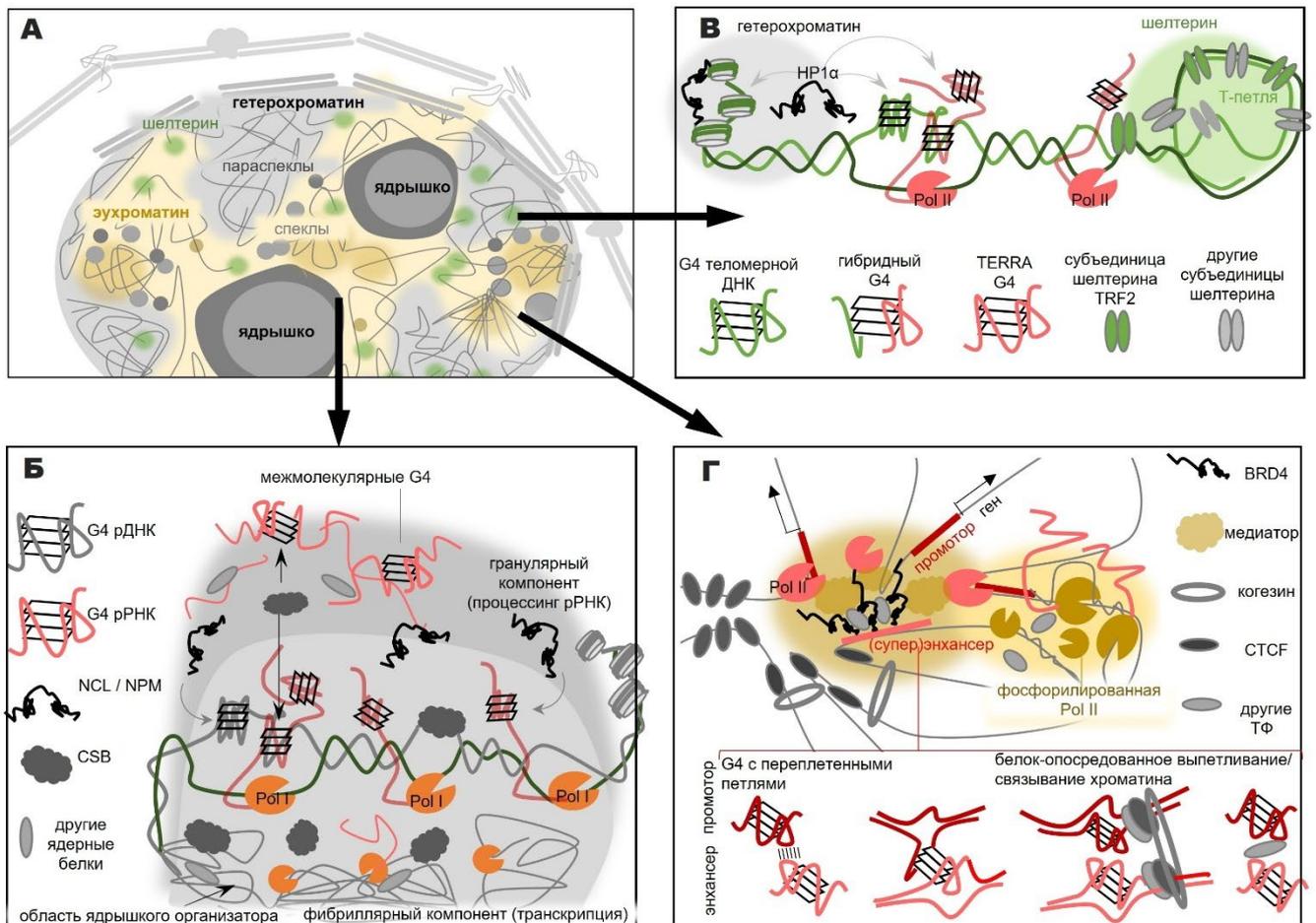
РНК-связывающие активаторы LLPS, переходящие из нуклеоплазмы в цитозоль и обратно и обладающие заметным сродством к G4, обычно содержат Arg-богатые домены, способные к слабым электростатическим/катион- $\pi$  взаимодействиям, и/или ароматические фрагменты, способные к слабым  $\pi$ - $\pi$  взаимодействиям, характерным для клеточных конденсатов [147,148]. Прямое влияние G4-ДНК на фазовые переходы большинства таких РНК-связывающих белков еще предстоит проверить. Одним из немногих белков с охарактеризованной чувствительностью фазовых переходов к G4 является SERBP1. Зависимость его конденсатов от G4-ДНК была косвенно проанализирована в модельной системе [122]. Имитирующие протоклетки гигантские мембранные везикулы (ГМВ) нагружали G4- или не-G4 ДНК. В первом случае наблюдалось обратимое разделение фаз. Последующие эксперименты показали, что в плотной фазе SERBP1 является основным G4-связывающим белком. G4-лиганд PDS ослаблял G4-зависимое LLPS SERBP1. Хотя биологическую значимость модели ГМВ не стоит переоценивать, она стала первым прямым доказательством образования конденсатов с G4-ДНК в нативной среде.

### ***1.2.2 G4 в ядерных конденсатах, связанных с процессингом РНК***

Синтетические G4-ДНК обычно использовались в интерактивных исследованиях [65,66], анализе LLPS в ГМВ [122] и большинстве других экспериментов *in vitro* [71,75], моделирующих сборку ядерных конденсатов. Однако *in vivo* нельзя исключать роль G4-РНК. Некодирующие

транскрипты с G4-мотивом, по-видимому, регулируют ядерные конденсаты, связанные с процессингом РНК, такие как спеклы, параспеклы и, возможно, ядрышки [149–151] (Рисунок 5А).

Спеклы – это РНК-белковые гранулы в межхроматиновых областях. В них сконцентрированы факторы сплайсинга пре-мРНК, такие как Ser-Arg-богатые белки и представители семейства hnRNP, включая hnRNP A1 и FUS, а также нуклеолин и нуклеофосмин [152]. Параспеклы примыкают к спеклам и обогащены белками и РНК, связанными с ответом на стресс или дифференцировкой [153]. Характерный компонент параспеклов – некодирующий транскрипт NEAT1 – и характерный компонент спеклов – транскрипт NEAT2, также известный как MALAT1, содержат несколько G4-сайтов [149,150]. В MALAT1 эти сайты имеют решающее значение для привлечения склонного к LLPS фактора транскрипции и сплайсинга NONO [149,154], а мутации в MALAT1, затрагивающие G4, приводят к нарушению сплайсинга [155]. Помимо NONO, G4 транскрипта MALAT1 связывают нуклеофосмин и нуклеолин [155] и могут взаимодействовать с FUS и hnRNP A1 [156].



**Рисунок 5** – G4 в ядерных биомолекулярных конденсатах. (А) Схематическое изображение типичных ядерных конденсатов с G4-ДНК/РНК. (Б) Схема организации ядрышка. (В) Схема организации субтеломерного гетерохроматина и шелтерина. (Г) Предполагаемая организация конденсатов РНК-полимеразы II при активной транскрипции.

Ядрышки, конденсаты, в которых синтезируется рРНК, собираются вокруг так называемых ядрышковых организаторов (Рисунок 5Б) – областей геномной ДНК с кластерами G4-содержащих повторов рибосомных генов (рДНК) [156]. Транскрипция рДНК под действием РНК-полимеразы I требует привлечения фактора ремоделирования и репарации хроматина с высоким содержанием Gly-Arg – белка синдрома Коккейна типа В (CSB). Недавно было показано, что этот белок распознает и расплетает межмолекулярные G4 рДНК [75]. Предполагается, что CSB накапливается в R-петлях и предотвращает чрезмерное образование межмолекулярных G4, что объясняет его важность для поддержания транскрипции в ядрышках. Механизмы, лежащие в основе этого процесса, требуют дальнейшего изучения. Однако можно предположить, что компоненты ядрышек накапливаются на короткоживущих G4. Помимо CSB, такие G4 могут рекрутировать RGG- и IDR-содержащие активаторы LLPS-опосредованной сборки ядрышка NCL [157] и нуклеофосмин [158], а также Gly-Arg -богатые активаторы LLPS CAR1 и фибрилларин [159]. Взаимодействия между G4 и CSB или NCL могут быть нарушены G4-лигандами, такими как PDS, CX-5461 или CX-3543, что приводит к дисфункции ядрышка и нарушениям синтеза рРНК [74,75].

Следует еще раз подчеркнуть, что G4 рДНК часто упоминаются в качестве возможных мишеней для лигандов, поскольку большинство данных *in vitro* было получено с использованием синтетической ДНК. Однако окрашивание ядрышка специфичным для G4 бензотиазольным зондом тиофлавином Т (ThT) выявило уменьшение числа G4 после обработки РНКзой, но не ДНКазой, что указывает на преобладание РНК-, а не ДНК- G4 в ядрышке [160]. В основной части ядра клетки наблюдалась противоположная ситуация при окрашивании с помощью другого зонда на основе бензотиазола, IMT [161]. Обработка РНКзой практически не влияла на паттерн флуоресценции IMT, в то время как обработка ДНКзой или мочевиной устраняла их, что свидетельствует о преобладании ДНК-G4, а не РНК-G4.

Подводя итог этой части, можно сказать, что G4 в некодирующих РНК и рДНК способны рекрутировать установленные и предполагаемые активаторы LLPS, вероятно, увеличивая их локальную концентрацию до критической точки, необходимой для спонтанного фазового перехода. Поскольку в большинстве этих белков присутствуют Arg-богатые мотивы, G4 могут способствовать LLPS через слабые катион- $\pi$ -взаимодействия (Рисунок 3В). В ядрышках индуцированные G4 конформационные перестройки нуклеофосмина способствуют его взаимодействию с белками, содержащими Arg-богатые мотивы (Рисунок 4А). Образование короткоживущих межмолекулярных G4 также может вносить вклад в LLPS (Рисунок 5Б). Однако еще предстоит выяснить, какие именно структуры (меж- или внутримолекулярные ДНК, РНК или гибридные G4 в R-петлях) играют ключевую роль.

### 1.2.3 G4 в гетерохроматине и конденсатах шелтерина

Участки генома с высоким содержанием G4-мотивов могут попадать в области как открытого, так и закрытого хроматина – в зависимости от типа клеток [9,25]. Подтвержденные методом ChIP-seq G4 были обнаружены в гетерохроматине человека [162] и способны рекрутировать ответственные за подавление экспрессии генов белки репрессивного комплекса поликомб 2 [65], линкерный гистон H1 [71] и белок гетерохроматина HP1 $\alpha$  [163]. Это не обязательно противоречит признанной ассоциации между G4 и активной транскрипцией [24]. Гетерохроматинизация может быть инициирована для предотвращения повреждения ДНК при чрезмерном образовании G4 [164], что предполагает механизм, подобный петле отрицательной обратной связи. Если H1 и HP1 $\alpha$  накапливаются в G4-богатых сайтах, гетерохроматинизация, вероятно, происходит за счет их совместного разделения фаз, т.е. соконденсации [71,165].

*In vitro*-эксперименты с полинуклеосомами показали, что H1 с высокой вероятностью является ключевым фактором разделения фаз [166]. В отсутствие конкурентов (аналогов линкерных гистонов) или модификаций гистонов, привлекающих активаторы транскрипции, H1 связывает межнуклеосомную ДНК в сайтах, прилежащих к нуклеосоме, посредством электростатических взаимодействий, минимизирует отталкивание ДНК-ДНК и таким образом способствует конденсации хроматина. Конденсацию часто называют LLPS-опосредованной [167], поскольку конформационно неупорядоченные хвосты H1 и гистонов образуют временные контакты друг с другом и с ДНК [168]. Данный тип LLPS является несколько нетривиальным, т.к. вовлеченные в него нуклеосомы имеют четкую структуру, кардинально отличаясь от IDR. Однако *in silico* демонстрация пластичности нуклеосом подтверждает корректность употребления термина [169].

Хотя G4 и нуклеосомы исключают друг друга, недавние исследования с помощью атомно-силовой микроскопии (АСМ) предоставили косвенные доказательства образования G4 в линкерной (межнуклеосомной) ДНК и возможного влияния G4 на пластичность близлежащих нуклеосом [90]. Исследования в упрощенной (безгистоновой) модели выявили усиление H1-зависимого LLPS в присутствии ДНК-G4 параллельного типа по сравнению с поли-А ДНК [71]. HP1 $\alpha$  взаимодействует с H1 *in vitro* и должен играть важную вспомогательную роль в гетерохроматинизации *in vivo* [166]. В растворе ожидается динамическое равновесие между компактной (автоингибированной) мономерной/димерной и декомпактизованной (склонной к олигомеризации) формами HP1 $\alpha$  [170]. Компактные формы стабилизируются электростатическими взаимодействиями между С-концевой областью (СТЕ) и центральной неструктурированной "шарнирной" областью (hinge region, HR) или N-концевой областью (NTE), а димеры удерживаются вместе благодаря контактам N-концевого хромодомена (CD) и

С-концевого теневого хромодомена (CSD). Фосфорилирование HP1 $\alpha$  NTE и HR-опосредованное связывание HP1 $\alpha$  с G4 параллельного типа [163] сдвигает равновесие в сторону активной формы HP1 $\alpha$  и в конечном итоге способствует LLPS, что было продемонстрировано *in vitro* [170] (Рисунок 4Б). Предполагаемый механизм похож на тот, что был предложен для нуклеофосмина: G4 способствуют LLPS, повышая вероятность межмолекулярных белковых контактов (Рисунок 4А). Высокая аффинность HP1 $\alpha$  в отношении параллельных G4 в сравнении с антипараллельными G4 или случайными последовательностями была подтверждена *in vitro* и *in situ*. Экзогенные параллельные G4 в ядрах клеток конкурировали с геномными и разрушали скопления HP1 $\alpha$  [163].

Отдельно стоит рассмотреть расселение гетерохроматинизацию в субтеломерных областях. Она зависит от рекрутирования HP1 $\alpha$  к теломерной РНК (TERRA), а не к ДНК, и связана со сборкой шелтериновых комплексов [163,171] (Рисунок 4В). Шелтерин также обладает характерными признаками жидких конденсатов, что было подтверждено *in cellulo* с помощью оптогенетического подхода [172]. Субъединицами шелтеринового комплекса, способствующими LLPS, являются факторы связывания теломерных повторов TRF1 и TRF2 [172]. Они распознают теломерные последовательности и перестраивают их в Т-петли, обеспечивая защиту оцДНК от нуклеазного гидролиза. IDR и димеризационные домены TRF1/2 участвуют в многочисленных гомо/гетеротипических взаимодействиях, приводящих к образованию конденсатов, контролируемых доступом теломеразы к ДНК. Изначально предполагалось, что TRF2 узнает G4 в теломерной ДНК. Это предположение строилось на том, что TRF2 проявляет сродство к модельным G4 *in vitro*, обнаруживается на G4-сайтах по всему геному (например, в промоторных областях) и может быть вытеснен с этих сайтов G4-связывающими лигандами 360A [173] или PDS [174]. При обработке лигандом PDS разрушается и шелтерин. Тем не менее, согласно последним данным, взаимодействие TRF2 с теломерной ДНК определяется ее последовательностью, а не вторичной структурой [167]. *In vitro* исследования тройных комплексов ДНК-TRF2-TERRA показали, что Gly-Arg -богатый домен TRF2 взаимодействует с РНК G4, в то время как ДНК распознается отдельным ДНК-связывающим доменом как в квадруплексной, так и в дуплексной форме [175]. Таким образом, для сборки шелтерина достаточно сворачивания G4 только в TERRA [176].

#### ***1.2.4 G4 в конденсатах, связанных с инициацией или элонгацией транскрипции и репарацией***

Недавние исследования распределения G4 относительно основных типов ядерных конденсатов с помощью иммуноцитохимии [177] выявили значительную колокализацию с

конденсатами РНК-полимеразы II и ядерными спеклами. Авторы исследования интерпретируют конденсаты РНК-полимеразы II как транскрипционные фабрики, не уточняя степень фосфорилирования полимеразы, что может быть не вполне корректно. Транскрипционными фабриками называют неподвижные конденсаты активной РНК-полимеразы II, через которые предположительно протаскивается ДНК во время элонгации транскрипции [178]. Их удерживают конформационно неупорядоченные С-концевые участки молекул РНК-полимеразы II, в которых Ser-содержащие гептадные повторы фосфорилированы по остаткам Ser5 и Ser2. Транскрипционные фабрики отличны от конденсатов с нефосфорилированной или фосфорилированной исключительно по Ser5 полимеразой, наблюдаемых при SE и, возможно, прочих энхансерах, на этапе инициации транскрипции [179]. Конденсаты инициации транскрипции удерживаются не только за счет С-концевого участка РНК-полимеразы II, но и за счет IDR ассоциированных с ней белков, включая BRD4, которые рекрутируются к SE. Они часто колокализованы с ядерными спеклами [180], что обеспечивает своевременный процессинг транскриптов генов с высоким уровнем экспрессии, регулируемых SE [184].

Роль G4-богатых промоторов и энхансеров в образовании или поддержании конденсатов инициации или элонгации транскрипции может состоять в привлечении транскрипционных факторов, включая активаторы LLPS TAF15 [14], SP1 [181] и BET-белки [182]. Прямое связывание с G4 *in vitro* было продемонстрировано для BET-белка BRD3 [66,182], наиболее близкими гомологами которого являются BRD4 и BRD2. Белки BRD2 и BRD3 ремоделируют нуклеосомы для поддержания процессивности РНК-полимеразы II во время элонгации транскрипции [183]. Их частично неупорядоченные домены ET привлекают другие транскрипционные факторы. BRD4 и BRD2 также действуют на этапе инициации транскрипции. Например, BRD2 привлекает *STAT3* к активным энхансерам [184], а BRD4 рекрутирует медиаторный комплекс, соединяющий SE с РНК-полимеразой II [84].

Также активно обсуждается вклад G4 в образование конденсатов за счет поддержания хроматиновых петель для энхансер-промоторной коммуникации [185]. G4 могут способствовать «мостиковым» контактам ДНК-ДНК напрямую, складываясь в межмолекулярные структуры [28] или в пару G4 с переплетёнными петлями [186], либо косвенно – через рекрутирование YY1 [78]; CTCF [187] и MYC-ассоциированного белка MAZ [37], который работает совместно с CTCF, ограничивая подвижность когезина [188].

Все вышперечисленные гипотезы подразумевают, что образование G4 может предшествовать декомпактизации хроматина при активной транскрипции, а не только быть ее следствием [25]. То же самое относится и к репликации: с одной стороны, G4 формируются вследствие расплетения дцДНК при репликации, с другой – G4 сами способствуют репликации. Они обогащены в ориджинах репликации [189] и способствуют рекрутированию и LLPS-

опосредованной [190] сборке комплекса распознавания орижина (ORC) для последующего ORC-, Cdc6- и Cdt1-зависимого привлечения хеликаз и инициации репликации [191]. Интересно, что, несмотря на повышенное содержание G4 в S-фазе, о чем свидетельствуют результаты визуализации как с помощью малых молекул [161], так и с помощью антител [22], G4, по-видимому, выполняют свои основные регуляторные функции, включая сборку транскрипционных конденсатов и ORC, в G1-фазе. Во время репликации они препятствуют продвижению репликативной вилки и должны быть своевременно расплетены хеликазами, чтобы избежать повреждения ДНК или потери гистонового кода [23,46,70]. При нормальной работе хеликаз не обнаруживается значительной корреляции G4 с маркерами повреждений ДНК [177]. Особый случай – воздействие активных форм кислорода. Из-за низкого окислительно-восстановительного потенциала гуанина G4 принимают на себя основную нагрузку, становясь «горячими точками» окисления генома [192,193]. В то же время они, по-видимому, способствуют своевременной репарации путем рекрутирования и активации фактора сигнализации о повреждении и репарации – поли(АДФ-рибоза)полимеразы 1 (PARP1) [194,195].

Последние данные говорят в пользу PARP1-опосредованного образования конденсатов ремоделирования хроматина в местах повреждения [194]. Во-первых, PARP1 присоединяет поли(АДФ-рибозные) цепи (PAR) к себе (автоактивация) и близлежащим белкам, что приостанавливает репликацию или транскрипцию из-за нарушения соответствующих конденсатов [196,197]. PARилированные белки в местах повреждения ДНК рекрутируют FUS [146] и способствуют разделению фаз [198,199] с образованием конденсатов репарации. После завершения репарации происходит гидролиз PAR под действием гликогидролазы PAR (PARG), и FUS выходит из состава конденсатов, что приводит к разборке конденсатов репарации. Этот предполагаемый механизм был подтвержден в модельной системе *in vitro* и визуализирован с помощью АСМ [146].

### **1.2.5 Открытые вопросы**

Текущее понимание вклада G4 в сборку и функционирование ядерных макромолекулярных конденсатов основано на исследованиях G4-интерактома, *in vitro*-исследованиях LLPS и внутриклеточной визуализации с использованием антител к G4 и маркерам конденсатов, а также низкомолекулярных light-up проб. Основные результаты, полученные с помощью этих подходов, можно обобщить в следующих тезисах:

- G4 способствуют LLPS белков, ассоциированных с гетерохроматином, в искусственных системах, но биологическая значимость этих результатов еще требует проверки.

- G4 способствуют LLPS РНК-связывающих белков в имитированной внутриклеточной среде. Эти результаты указывают на вероятный вклад G4 в сборку ядерных конденсатов, богатых факторами процессинга РНК, включая ядрышки, спеклы и параспеклы.
- Морфология и/или функции спеклов/параспеклов нарушаются мутациями в G4 и G4-стабилизирующими лигандами. Морфология и функции шелтериновых комплексов также нарушаются под действием G4-лигандов. Эффекты этих лигандов объясняются их вмешательством в G4-белковые взаимодействия.
- Влияние G4 на LLPS белков, ассоциированных с инициацией и элонгацией транскрипции, практически не изучено, но показаны сродство к данным белкам и колокализация с ними, что стимулирует дальнейшие исследования транскрипционных конденсатов на G4-богатой ДНК и проверку их чувствительности к G4-лигандам.

Большинство G4-стабилизирующих малых молекул представляют собой плоские полиароматические структуры с положительно заряженными (обычно аминогруппами/гуанидиновыми группами) заместителями, приспособленные для формирования стэкинг-контактов с внешней тетрадой G4. Они могут конкурировать с G4-связывающими белками и нарушать  $\pi$ - $\pi$ /катион- $\pi$ -взаимодействия IDR таких белков с G4. Это наделяет G4-лиганды способностью к ремоделированию G4-содержащих ядерных конденсатов. Последствия такого ремоделирования будут системными на клеточном уровне и могут перевесить желаемые специфические эффекты – результат действия лигандов на G4-мишени в промоторах онкогенов, теломерах и т.д. Этот тезис в определенной степени подрывает предсказуемость таргетирования G4 малыми молекулами, но также указывает на новые возможности в управлении ядерными конденсатами.

## 2. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### 2.1 G4 как участники демаркации транскрипционных (суб)доменов за счет привлечения CTCF

В рамках работы над первой и второй задачами исследования проверялась связь между G4 и рекрутированием белка-инсулятора CTCF. Сначала для проверки колокализации с G4 на примере линии клеток острого миелоидного лейкоза (K-562) методом ChIP-seq были картированы сайты связывания CTCF в присутствии и в отсутствие G4-стабилизирующего лиганда. Зависимость частоты позиционирования CTCF от G4-лиганда была выборочно подтверждена методом кПЦР. Далее взаимодействие CTCF с G4 подтверждали непосредственно на рекомбинантном белке с синтетическими G4-олигонуклеотидами. Результаты были сопоставлены с данными молекулярного моделирования, которое прояснило вероятный тип связывания G4-CTCF.

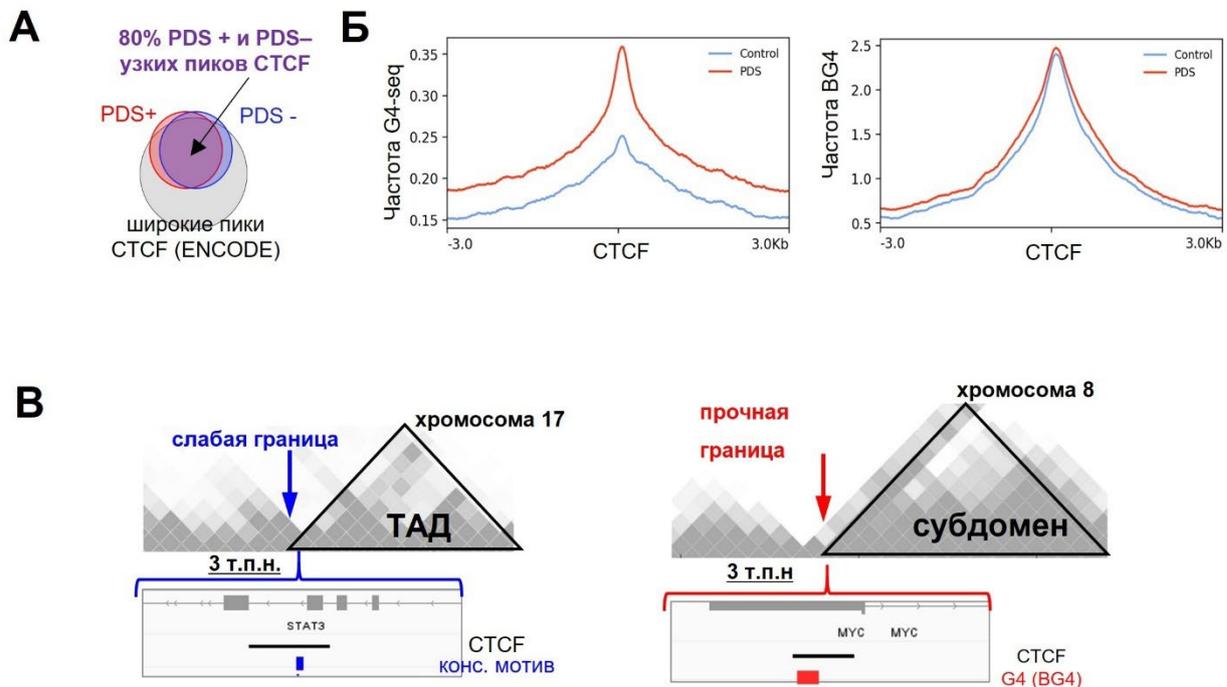
#### 2.1.1 G4 колокализованы с сайтами связывания CTCF в геноме

Ранее предполагалось, что G4-фолдинг негативно сказывается на вероятности связывания CTCF с G/C-богатыми сайтами [55]. Однако, этот вывод основывался на единичном примере из экзона 1 гена обратной транскриптазы теломеразы человека (hTERT), где образование G4 нарушает связывание CTCF, что приводит к репрессии транскрипции [200]. Примечательно, что сайт связывания CTCF в гене hTERT необычен тем, что *in vitro* он принимает либо конформацию шпильки, либо G4 в зависимости от статуса метилирования CpG. Конкуренция G4/шпилька может быть специфичным для hTERT случаем, а не репрезентативным примером.

Чтобы получить более широкую картину, было проанализировано распределение CTCF относительно G4 на полногеномном уровне в нативных условиях и условиях, благоприятствующих G4. Для этого клетки K-562 были обработаны стабилизатором G4 пиридостатином (PDS) [201] в концентрации 10 мкМ или физиологическим раствором с той же концентрацией ДМСО без лиганда, и после 24 ч инкубации были выполнены эксперименты ChIP-seq с антителом к CTCF.

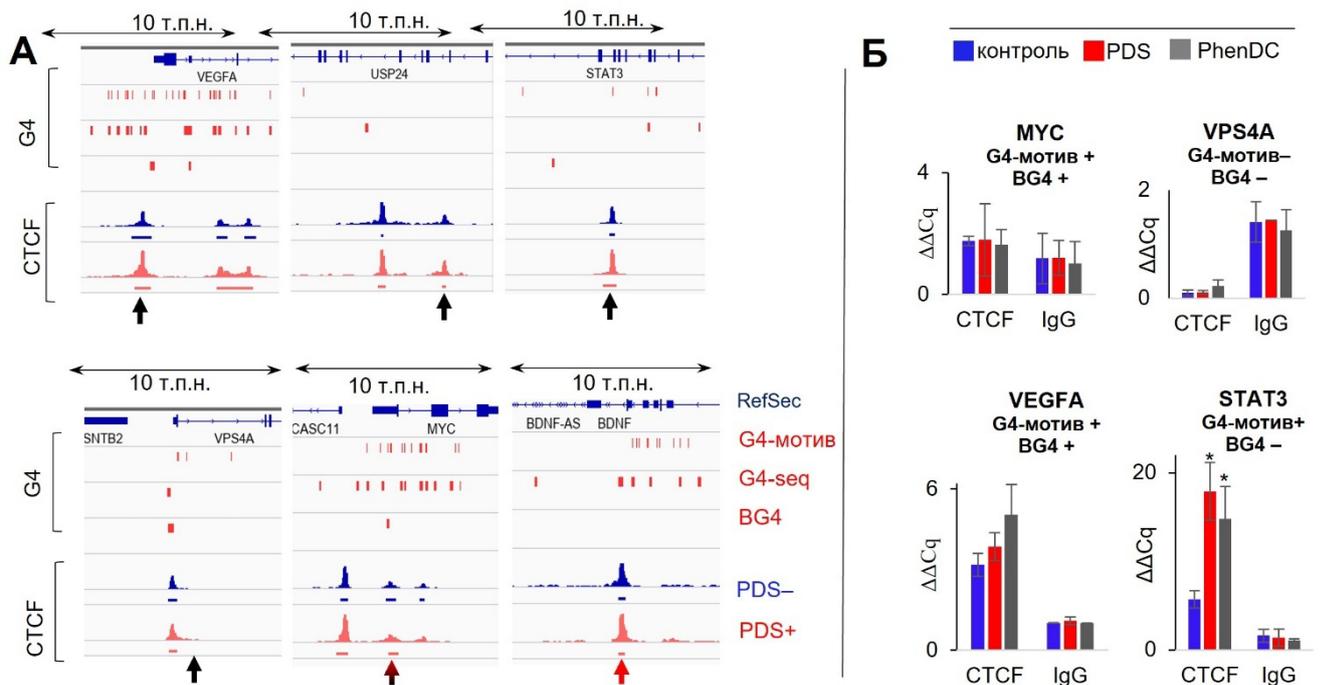
Полученные профили распределения CTCF в клетках, обработанных и не обработанных PDS, различались незначительно и были близки к опубликованным ранее (Рисунок 6А) [26]. Пики CTCF часто колокализовались с G4-мотивами, G4-seq-пиками и пиками G4-ChIP-seq (BG4) или были фланкированы ими (Рисунок 6Б). Статистическая значимость обогащения ChIP-seq пиков CTCF G4-сайтами в отсутствие лиганда была показана ранее [26] с помощью

пермутационного теста, поэтому дополнительных тестов не проводилось. В присутствии лиганда изменения носили преимущественно количественный характер: менялось соотношение пиков, но не общий вид профиля. Качественные изменения, а именно появление новых пиков CTCF после инкубации с PDS, наблюдались в основном в области пиков G4-seq. Точный тест Фишера показал значимое различие процента колокализации G4-seq с CTCF в клетках, обработанных PDS, и контрольных клетках ( $p\text{-value} < 1 \times 10^{-300}$ ). Напротив, пики антитела BG4 часто встречались в сайтах связывания CTCF независимо от PDS (Рисунок 6Б). Это можно интерпретировать следующим образом: подтвержденные с помощью антитела G4 стабильны в отсутствие лиганда, тогда как G4-seq пики соответствуют менее устойчивым структурам, которые удерживают CTCF только в присутствии лиганда.



**Рисунок 6** – Локализация CTCF относительно G4-сайтов в клетках, обработанных и не обработанных пиридостатином (PDS). (А) Диаграмма Венна, обобщающая пересечение ChIP-seq пиков CTCF, полученных в присутствии и в отсутствие лиганда (PDS+ и PDS-) с ранее опубликованными ChIP-seq пиками CTCF. (Б) Распределение G4-seq и пиков BG4 относительно PDS+ или PDS-(контроль) пиков CTCF. (В) Примеры хроматиновых доменов в хромосомах 17 (координаты: 40,300,000-40,700,000, сборка генома: hg19) и 8 (координаты: 128,600,000-129,000,000, сборка генома: hg19). Первый содержит приграничный кластер CTCF в областях гена *STAT3* (chr17:40,473,500-40,476,500) с консенсусным мотивом связывания (синий) и G4-мотивами без подтвержденных G4-структур (левая панель). Второй содержит кластер CTCF в промоторной области гена *MYC* (chr8:128,746,836-128,749,835) без консенсусного дуплексного мотива, но с подтвержденным G4 (BG4) в клетках линии K-562. Последовательность консенсусного дуплекса из *STAT3*: CCACCAGGGGGCAGT. Последовательность подтвержденного G4 из *MYC*: TGGGGAGGGTGGGGAGGGTGGGGAAGG.

Далее было проверено, приведет ли длительная инкубация с избытком G4-стабилизирующих лигандов к увеличению обогащения CTCF в G4-сайтах. Для этого клетки K-562 были обработаны 20 мкМ PDS. PhenDC3, более мощный G4-стабилизатор, чем PDS [202], был протестирован параллельно в концентрации 8 мкМ. Предполагается, что и PDS, и PhenDC3 являются пан-квадруплексными стабилизаторами [202,203]. Изменения обогащения CTCF в G4-мотиве (*STAT3*), G4-кластерах, колокализованных с BG4 пиками (*MYC* и *VEGFA*), и сайтах без G4 (*USP24* и *VPS4A*) были оценены с помощью ChIP с кПЦР (сайты показаны на Рисунке 6В и Рисунке 7А; последовательности праймеров приведены в Таблице 1). Лиганды увеличивали обогащение CTCF в сайте с потенциальным G4 без BG4-пика (ген *STAT3*), оказывали незначительное влияние на сайты BG4 и не влияли на сайты без G4 (Рисунок 7Б). Эти данные свидетельствовали против распространенного предположения о том, что фолдинг G4 препятствует связыванию CTCF [55], что побудило нас изучить возможность комплексообразования G4-CTCF.



**Рисунок 7** – Обогащение CTCF на сайтах G4 в клетках, обработанных и не обработанных PDS. (А) Геномный браузер: области генов *VEGFA*, *USP24*, *STAT3*, *VPS4A*, *MYC* и *BDNF*, содержащих G4-мотивы, пики G4-seq, пики G4 ChIP-seq (BG4) и/или сайты связывания CTCF в клетках K-562, обработанных и не обработанных пиридостатином (PDS). Черными стрелками отмечены сайты, для которых проводился полуколичественный анализ зависимости CTCF-пиков от PDS методом ChIP с кПЦР. Сайт связывания CTCF без G4 в *USP24* был использован в качестве положительного контроля, а сайт в *VPS4A* без CTCF и G4 – в качестве отрицательного контроля. Красными стрелками отмечены сайты с G4, которые использовались в анализах связывания. (Б) Результаты ChIP с кПЦР. Гистограммы иллюстрируют нормализованное обогащение CTCF на G4-сайтах (*MYC*, *VEGFA* и *STAT3*)

относительно референсного сайта без G4 (*USP24*) в необработанных (синий) клетках K-562 и клетках, обработанных PDS (красный) или PhenDC3 (серый).

**Таблица 1. Детали ChIP с кПЦР: последовательности праймеров**

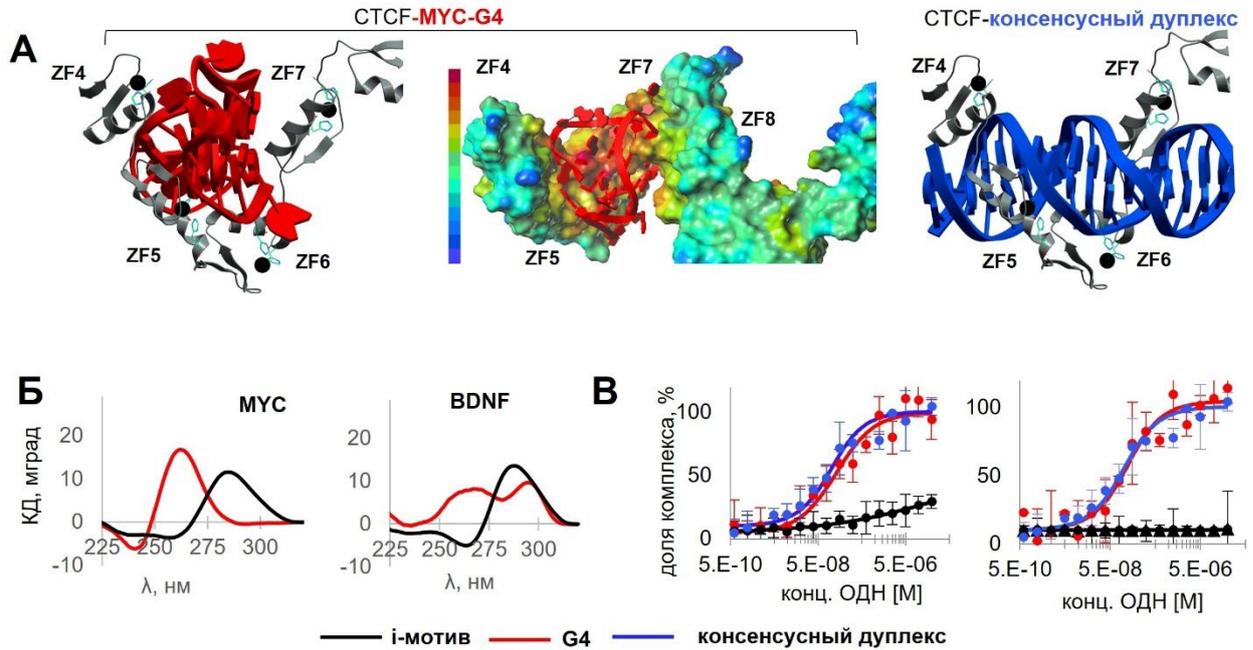
шифр	последовательность	положение в геноме
USP24_forward USP24_reverse	CCACTGGCTGTCCCTAGATCA TCCTACCTGCACTACCATGC	Chr 1, <i>USP24</i> интрон 21
STAT3_forward STAT3_reverse	TGAGTGAAACAGGGAGTCAAG AGAAGGACATCAGCGGTAAG	Chr 17, <i>STAT3</i> интрон и экзон 20
VEGFA_forward VEGFA_reverse	AGCCCATTCCTCTTTAGCC ACACACTCACTACCCACAC	Chr 6, <i>VEGFA</i> промотор
MYC_forward MYC_reverse	CTCAGAGGCTTGGCGGGAAAA CAGCGAGTTAGATAAAGCCCG	Ch8, <i>MYC</i> промотор
VGPS4A_forward VGPS4A_reverse	GGCAACCACTGCTAATCTACTT GCCACAAAGGACCACCTATT	<i>VPS4A</i> интрон 1

### 2.1.2 G4 взаимодействуют с CTCF *in vitro*

Сначала ведущим научным сотрудником Первого МГМУ им. И.М. Сеченова Цветковым В.Б. была проверена возможность прямых взаимодействий CTCF-G4 *in silico*. В качестве модельной структуры был выбран хорошо охарактеризованный G4 из элемента гиперчувствительности к нуклеазе гена *MYC* [204]. В клетках линии K-562 он перекрывается с ChIP-seq пиком CTCF и BG4 (Рисунок 7А). Докинг MYC-G4 на CTCF позволил выявить наиболее вероятные сайты связывания. Полученный комплекс был похож на ранее описанный комплекс CTCF-дуплекс [61] (Рисунок 8А). G4 поместился между цинковыми пальцами ZF4-ZF7 CTCF, которые "обхватывали" G/C-богатую В-ДНК в опубликованных ранее кристаллических структурах [61].

Далее сродство G4 к CTCF проверяли *in vitro*. Помимо MYC-G4 тестировали потенциальный квадруплексный сайт из фланкирующего ChIP-seq пик CTCF участка гена *BDNF* – BDNF-G4 (красные стрелки на Рисунке 7А; Таблица 2), а также комплементарные данным G4 C-богатые последовательности, способные формировать интеркалированные мотивы (i-мотивы, iM) при условии частичного протонирования цитозинового остатков [205]. Предпосылок к узнаванию i-мотивов белком CTCF нет; C-богатые последовательности MYC-iM и BDNF-iM использовались в качестве отрицательных контролей. В качестве положительного контроля использовали консенсусный CTCF-связывающий дуплекс с последовательностью, предсказанной на основе специфичности CTCF ZF3-7 [61]. Вторичные структуры всех олигонуклеотидов (ОДН) G4 и i-мотивов были подтверждены характерными сигнатурами кругового дихроизма (КД) – пиками при 265/295 и 288 нм, соответственно (Рисунок 8Б). Взаимодействие с CTCF было проанализировано методом микромасштабного термофореза (Рисунок 8В). Оба G4 связывали CTCF достаточно эффективно, значения константы диссоциации комплексов находились в

наномолярном диапазоне концентраций ( $K_d_{\text{MYC-G4}} = 140 \pm 60 \text{ нМ}$  и  $K_d_{\text{BDNF-G4}} = 120 \pm 30 \text{ нМ}$ ) и были сопоставимы со значением, полученным для контрольного дуплекса ( $K_d = 80 \pm 10 \text{ нМ}$ ), в то время как отрицательные контроли (*i*-мотивы) с CTCF не связывались (Рисунок 8В).



**Рисунок 8** – Анализ взаимодействий CTCF-G4. **(А)** *In silico* проверка связывания G4 с CTCF. Показана наилучшая по энергии связывания конформация комплекса, полученная при докинге MYC-G4 (красный) на CTCF. Для сравнения показан комплекс с консенсусным дуплексом (синий). На средней панели поверхность CTCF окрашена в соответствии с электростатическим потенциалом: от отрицательного (синий) к положительному (красный). Данные получены В.Б. Цветковым **(Б)** Подтверждение фолдинга G4 для *in vitro* проверки связывания с CTCF. Показаны спектры кругового дихроизма 2 мкМ растворов MYC-G4 и BDNF-G4 (красные) и контрольных олигонуклеотидов MYC-*i*М и BDNF-*i*М (черные) в псевдофизиологическом буферном растворе. **(В)** Анализ взаимодействия CTCF с G4 (красный), *i*-мотивами (черный) и контрольным консенсусным дуплексом (синий) методом микромасштабного термофореза.

**Таблица 2. G4, iM и контрольные дуплексы, протестированные на связывание с CTCF**

шифр	последовательность	ген
MYC-G4	TGGGGAGGGTGGGGAGGGTGGGGAAGG	MYC
MYC-iM	TCCCCACCTTCCCCACCCTCCCCACCCTCCCCA	
BDNF-G4	GGGGGGCGGGGGGGCGGGGGGGGGGG	BDNF
BDNF-iM	CCCCCCCCCGCCCCCGCCCCC	
дуплекс	GCCAGCAGGGGGCGCTA/ TAGCGCCCCCTGCTGGC	-

Таким образом, было показано, что G4 образуют комплексы с CTCF *in vitro*, что, вероятно, объясняет колокализацию с CTCF в геноме. Альтернативный гипотетический механизм участия G4 в позиционировании CTCF связан с регуляцией нуклеосомной укладки, а именно – поддержанием низкой локальной нуклеосомной плотности, необходимой для узнавания консенсусных мотивов связывания в дуплексе. Влияние G4 на локальную нуклеосомную плотность рассмотрено в следующем разделе.

## 2.2 G4 как фактор поддержания низкой нуклеосомной плотности

В данной части работы была сконструирована простая *in vitro* модель геномной ДНК с G4, которая была использована для выяснения влияния G4 и G4-стабилизирующих лигандов на нуклеосомную плотность. Было получено две конструкции дцДНК длиной 226 п.н., состоящие из сайта позиционирования нуклеосом и имитированный дцДНК-хвост. Одна ДНК-конструкция (контрольная) не содержала G4, а вторая содержала "сильный" (стабильный) G4-мотив. Фолдинг G4 был подтвержден оптическими методами и <sup>1</sup>H-ЯМР-спектроскопией. Также в полученной модели были исследованы три известных G4-лиганда, чтобы оценить их вероятные побочные эффекты, связанные с действием на нуклеосому.

### 2.2.1 G4 частично сложены в составе модельной дуплексной конструкции с сайтом позиционирования нуклеосомы

Для сборки мононуклеосом были получены конструкции дцДНК длиной 226 п.н. (Таблица 4), состоящие из ранее охарактеризованного сайта позиционирования нуклеосом длиной 167 п.н. [86] и ДНК-хвоста, имитирующую линкерную ДНК, со случайной последовательностью (конструкция (1) на Рисунке 9А) или G4-содержащего хвоста (конструкция (2) на Рисунке 9А). Последовательность позиционирования нуклеосомы, являлась частью так называемой консенсусной последовательности Widom-601 [206] и включала центральный сайт связывания гистонового октамера длиной 147 п.н. (Рисунок 9Б) [207] с короткими фланками. Средняя часть ДНК-хвоста длиной 59 п.н. содержала неструктурированный 15-мер TTCAACCAGTCTATG в контрольной конструкции (1) и G4-мотив GGGTGGTGGTGGTGGTGGTGGG в конструкции (2). Этот G4 был выбран в качестве репрезентативной квадруплексной структуры из-за его высокой термической стабильности [208], распространенности в геноме человека и сродства к транскрипционным факторам, что предполагает биологическую значимость [182].

**Таблица 4. Последовательности нуклеосомных конструкций и фрагментов для их сборки**

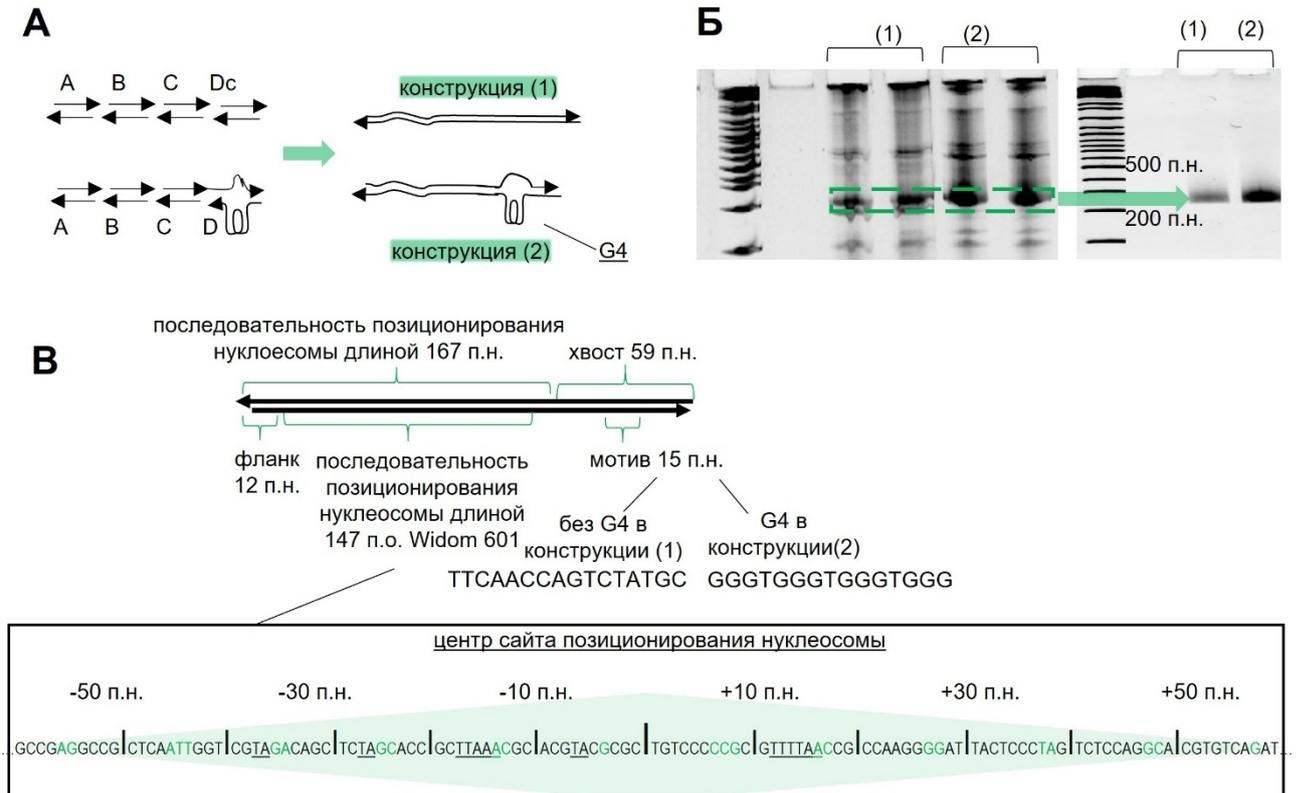
шифр дуплекса	шифр цепи	Последовательность, 5'-3'*
A	af	<i>tcggggccgccttgagagaatcccgggtgccgagccgctcaattggctgtagacagctct</i>
	ar	<i>ctgtctacgaccaattgagcggcctcggcaccgggattctccagggcgccccca</i>
Ac	afc	<i>ctaacgactctgttgagagaatcccgggtgccgagccgctcaattggctgtagacagctct</i>
	arc	<i>ctgtctacgaccaattgagcggcctcggcaccgggattctccagcagagtcgttag</i>
B	bf	<i>agcaccgcttaaacgcacgtacgcgctgtccccgcgttttaaccgccaagggga</i>
	br	<i>cttggcggttaaacgcggggacagcgcgtacgtgcgtttaagcgggtgctagag</i>

Bg	bfg	<i>agcaccgcttaaacgcacgtacgcgtgtcccgggtgggtgggtgggcaagggga</i>
	brg	<i>cttgcccaccaccaccgggacagcgcgtacgtgcttlaagcgggtctagag</i>
C	cf	<i>ttactccctagctccaggcacgtgacagatatacatcctgtgcatgactcca</i>
	cr	<i>gtacatgcacaggatgtatatactgacacgtgacctgggagactaggagtaatccc</i>
D	df	<i>atgtgatcagctgcacccaccaccaccagccaagtgtgatcagctgctcgg</i>
	dr	<i>ccgagcagctgatcacactggctgggtgggtgggtgggtgggtgacagctgatcacattgga</i>
Dc	dfc	<i>atgtgatcagctgcagcatagactggtgaaagccaagtgtgatcagctgctcgg</i>
	drc	<i>ccgagcagctgatcacactggcttcaaccagtctatgctgacagctgatcacattgga</i>
(1)**	forward	<u>tcggggccgcctggagaatcccgtgccgagcccgtcaattgctgtagacagcttagcaccgcttaaacgcacgtacgcgt</u> <u>gtccccgcgttfaaccgccaaggggattactccctagtctccaggcacgtgtagatatacatcctgtgcatgtactccaatgtgat</u> <u>agctgcagcatagactggtgaaagccaagtgtgatcagctgctcgg.</u>
(1)	reverse	<u>ccgagcagctgatcacactggcttcaaccagtctatgctgacagctgatcacattggagtacatgcacaggatgtatatactgacacgt</u> <u>gctggagactaggagtaatcccctggcggttaaacgcggggacagcgcgtacgtgctggttaagcgggtgtagagctgtctac</u> <u>gaccaatgagcggcctcggcaccgggattctccaggcggccccca</u>
(2)	forward	<u>tcggggccgcctggagaatcccgtgccgagcccgtcaattgctgtagacagcttagcaccgcttaaacgcacgtacgcgt</u> <u>gtccccgcgttfaaccgccaaggggattactccctagtctccaggcacgtgtagatatacatcctgtgcatgtactccaatgtgat</u> <u>agctgcacccaccaccaccagccaagtgtgatcagctgctcgg</u>
(2)	reverse	<u>ccgagcagctgatcacactggctGGGTGGGTGGGTGGGTgacagctgatcacattggagtacatgcacaggatgta</u> <u>tatactgacacgtgctggagactaggagtaatcccctggcggttaaacgcggggacagcgcgtacgtgctggttaagcgggtg</u> <u>tagagctgtctacgaccaatgagcggcctcggcaccgggattctccaggcggccccca</u>
(4)	forward	<u>tcggggccgcctggagaatcccgtgccgagcccgtcaattgctgtagacagcttagcaccgcttaaacgcacgtacgcgt</u> <u>gtcccGGGTGGGTGGGTGGGcaaggggattactccctagtctccaggcacgtgtagatatacatcctgtgcatgta</u> <u>ctccaatgtgatcagctgcagcatagactggtgaaagccaagtgtgatcagctgctcgg.</u>
(4)	reverse	<u>ccgagcagctgatcacactggcttcaaccagtctatgctgacagctgatcacattggagtacatgcacaggatgtatatactgacacgt</u> <u>gctggagactaggagtaatcccctggccaccaccaccaccgggacagcgcgtacgtgctggttaagcgggtgtagagctgtctac</u> <u>gaccaatgagcggcctcggcaccgggattctccaggcggccccca</u>
(5)	forward	<u>ctaagactctgctggagaatcccgtgccgagcccgtcaattgctgtagacagcttagcaccgcttaaacgcacgtacgcgtg</u> <u>tccccgcgttfaaccgccaaggggattactccctagtctccaggcacgtgtagatatacatcctgtgcatgtactccaatgtgatca</u> <u>gctgcacccaccaccaccagccaagtgtgatcagctgctcgg</u>
(5)	reverse	<u>ccgagcagctgatcacactggctTGGGTGGGTGGGTGGGTgacagctgatcacattggagtacatgcacaggatgt</u> <u>atatactgacacgtgctggagactaggagtaatcccctggcggttaaacgcggggacagcgcgtacgtgctggttaagcgggtg</u> <u>ctagagctgtctacgaccaatgagcggcctcggcaccgggattctccagcagagctgtag</u>
(6)	forward	<u>ctaagactctgctggagaatcccgtgccgagcccgtcaattgctgtagacagcttagcaccgcttaaacgcacgtacgcgtg</u> <u>tccccgcgttfaaccgccaaggggattactccctagtctccaggcacgtgtagatatacatcctgtgcatgtactccaatgtgatca</u> <u>gctgcagcatagactggtgaaagccaagtgtgatcagctgctcgg</u>
(6)	reverse	<u>ccgagcagctgatcacactggcttcaaccagtctatgctgacagctgatcacattggagtacatgcacaggatgtatatactgacacgt</u> <u>gctggagactaggagtaatcccctggcggttaaacgcggggacagcgcgtacgtgctggttaagcgggtgtagagctgtctac</u> <u>gaccaatgagcggcctcggcaccgggattctccagcagagctgtag</u>
G4-Widom	forward	<u>tcggggccgcctggagaatcccgtgccgagcccgtcaattgctgtagacagcttagcaccgcttaaacgcacgtacgcgt</u> <u>gtccccgcgttfaaccgccaaggggattactccctagtctccaggcacgtgtagatatacatcctgtgcatgtactccaatgtgatc</u> <u>agctgcacccctccctccctccctgccaagtgtgatcagctgctcgg</u>
G4-Widom	reverse	<u>ccgagcagctgatcacactggcaGGGAGGGAGGGAGGGatgacagctgatcacattggagtacatgcacaggatgta</u> <u>tatactgacacgtgctggagactaggagtaatcccctggcggttaaacgcggggacagcgcgtacgtgctggttaagcgggtg</u> <u>tagagctgtctacgaccaatgagcggcctcggcaccgggattctccaggcggccccca</u>

\* Последовательность Widom 601 длиной 147 п.н. подчеркнута;

«сильный» G4-мотив выделен жирным шрифтом заглавными буквами

\*\* также упоминается как Cntr-Widom

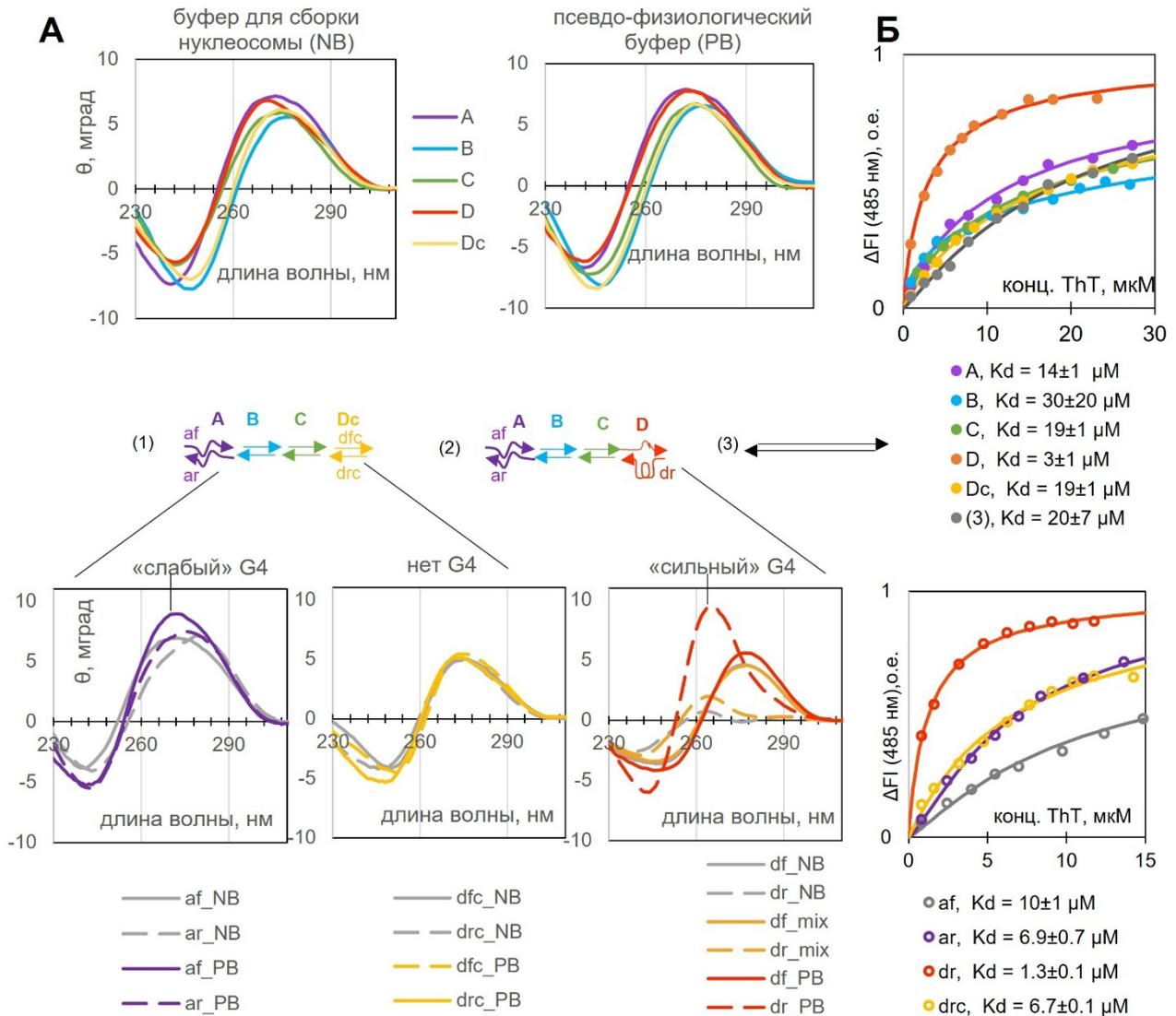


**Рисунок 9** – Дизайн и сборка ДНК-конструкций с G4 и без G4 (контроль). **(А)** Схематическое изображение конструкций дцДНК (1) и (2). Обе конструкции содержат сайт позиционирования нуклеосом длиной 167 п.н. (фрагменты А, В и С) и ДНК-хвост длиной 59 п.н. с G4/без G4 (фрагмент D/Dc). Конструкции были собраны из дуплексов с липкими концами А-Д, затем очищены и амплифицированы для последующей сборки нуклеосом. **(Б)** Сканы денатурирующего ПААГ, иллюстрирующие результаты сборки и очистки конструкций. **(В)** Основные сведения о последовательностях конструкций. На нижней панели подчеркнутые фрагменты последовательности являются сигналами позиционирования нуклеосом. Нуклеотиды, которые контактируют с гистонами и, предположительно, участвуют образованию водородных связей с гистонами, обозначены зеленым цветом.

Конструкции (1) и (2) получали путем фосфорилирования и лигирования синтетических дуплексов А-Д/Дс длиной 55/56 п.н (Таблица 4). Дуплексы А-С составляли сайт позиционирования нуклеосом длиной 167 п.н., а дуплексы D и Dc представляли хвосты с G4 и без G4, соответственно. Полученная дцДНК длиной 226 п.н. была амплифицирована, очищена и затем визуализирована с помощью денатурирующего ПААГ (Рисунок 9Б). Секвенирование по Сэнгеру подтвердило правильность сборки.

Существование сложеного G4 в окружении двухцепочечных фланков целевых конструкций (дуплексы А-Д/Дс) было проверено с помощью КД-спектроскопии в буфере для сборки нуклеосом, содержащем 250 мМ NaCl (Рисунок 10А, левая панель). Дуплексы В и С показали спектры КД, сходные с Dc и типичные для В-формы ДНК, с положительными и отрицательными пиками при 280 и 240 нм, соответственно. В случае D максимум был смещен к

270 нм, что указывает на G4-структуру. Дуплекс А напоминал D с основным максимумом, смещенным к 270 нм, и имел дополнительный локальный максимум при 280 нм.



**Рисунок 10** – Проверка фолдинга G4 в ДНК-конструкциях с помощью КД-спектроскопии (А) и анализа с флуоресцентным зондом (Б). КД-спектры двухцепочечных (верхняя панель) фрагментов ДНК-конструкций (A, B, C, D и Dc) были получены в буфере для сборки нуклеосом (NB) или в псевдофизиологическом буфере (PB). NB: 20 мМ Трис-НСl, рН 8, и 250 мМ NaCl; PB: 20 мМ Трис-НСl, рН 7.2, и 140 мМ KCl. КД-спектры одноцепочечных (нижняя панель) фрагментов ДНК-конструкций (af, ar, df, dr, dfc и drc) были получены в NB, PB или промежуточном ("mix") буфере, который представляет собой NB с 10 мМ KCl. Концентрация ОДН 1 мкМ. ThT-анализ с двухцепочечными (вверху) или одноцепочечными (внизу) фрагментами ДНК-конструкций проводили в PB, концентрация ОДН 0,7 мкМ. Все эксперименты проводились при температуре 20 °С.

Чтобы выяснить, сохраняются ли предполагаемые G4 в А и D в физиологических условиях, были проведены измерения КД в псевдофизиологическом буфере, содержащем 140 мМ KCl. Хотя ионы калия и натрия способны к G4-стабилизирующим координационным взаимодействиям с атомами O6 гуанина, калий несколько более благоприятен для поддержания

G4 благодаря эффективному межквартетному позиционированию [209]. Таким образом, ожидалось, что сигнатуры G4 будут более выражены в присутствии KCl. Действительно, спектры КД А и D псевдофизиологическом буфере были явно сдвинуты в сторону приближения к максимуму, характерному для G4 (265 нм), в то время как спектры В, С и Dc оставались типичными для дуплексов (Рисунок 10А, правая панель).

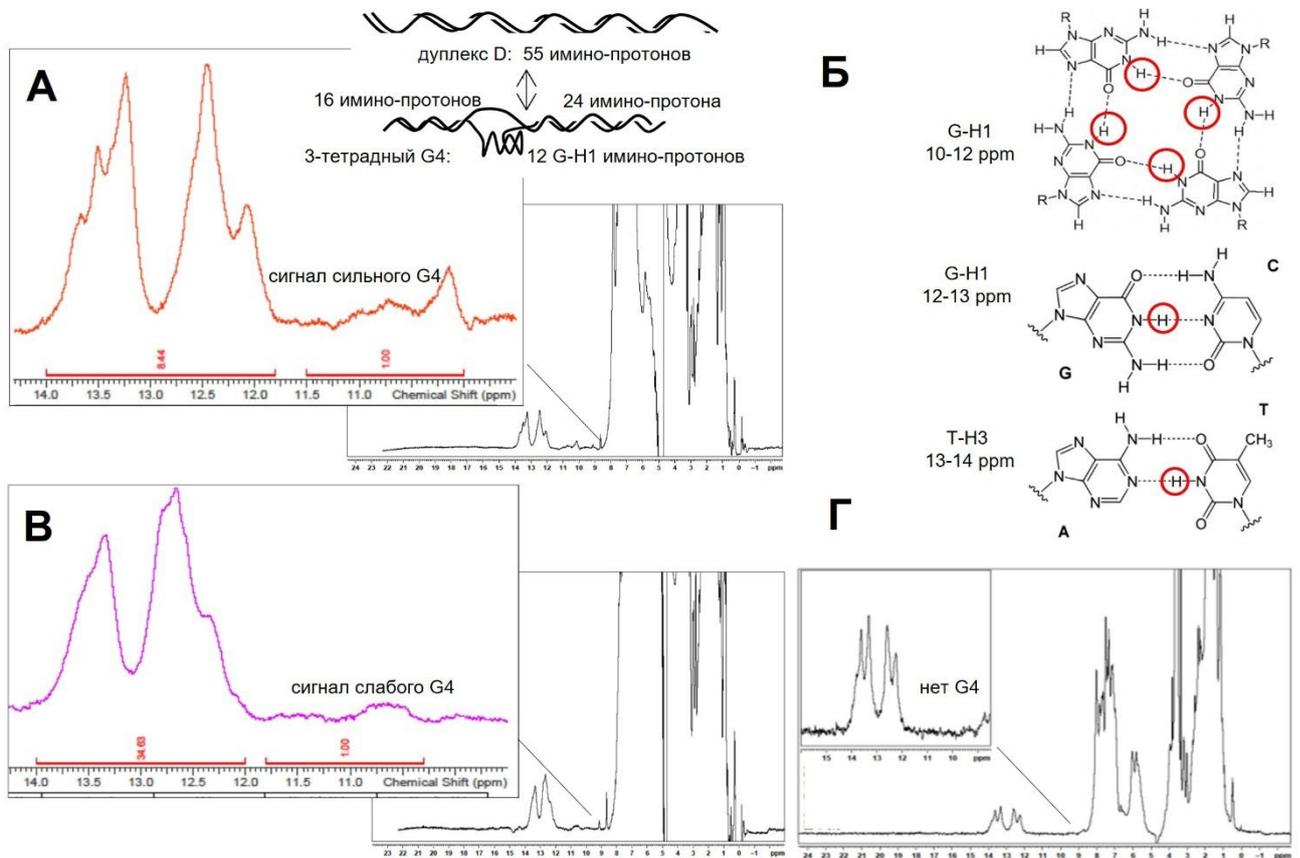
Результаты КД-спектроскопии оцДНК af, ar, df, dr, drc и drc из дуплексов А, D и Dc (Таблица 4) обобщены на нижней панели Рисунка 10А. Индексы "r" и "f" в шифрах ОДН обозначают прямую и обратную цепи, соответственно. Хвостовой фрагмент dr со встроенным G4-мотивом, который здесь и далее обозначается как "сильный" (термодинамически стабильный) G4, показал чувствительную к калию полосу КД при 265 нм, характерную для параллельных G4. Ни контрольный фрагмент drc, ни комплементарная нить df не имели признаков неканонических структур. В спектрах ar и af наблюдались незначительные или умеренные индуцированные калием пики КД вблизи 265 нм, что подтверждает наличие слабого (термодинамически нестабильного) G4 в А.

Для дополнительной проверки фолдинга G4 в двухцепочечных (Рисунок 10Б, вверху) и одноцепочечных (Рисунок 10Б, внизу) фрагментах конструкций (1) и (2) было использовано титрование G4-специфичным зондом тиофлавином Т (ThT) [210]. Предположительно G4-несодержащая контрольная ДНК размером 208 п.н. из набора для сборки нуклеосом (3) была проанализирована параллельно с (1) и (2).

Дуплексы В, С и Dc, а также контрольная ДНК (3) слабо связывались с ThT, значения  $K_d$  варьировались от 19 до 30 мкМ и вызывали лишь умеренное усиление флуоресценции ThT. Эффективность связывания была несколько выше для дуплекса А ( $K_d = 14 \pm 1$  мкМ) и значительно выше для дуплекса D ( $K_d = 3 \pm 1$  мкМ). Значения коэффициента Хилла ( $0,6 \pm 0,1$  для дуплекса В,  $0,8 \pm 0,1$  для дуплекса С и  $0,9 \pm 0,1$  для дуплексов D и Dc) свидетельствовали о наличии нескольких сайтов связывания лиганда в ДНК-мишени и указывали на незначительную отрицательную кооперативность. Одноцепочечный фрагмент ДНК ar превосходил af по эффективности связывания ( $K_d_{af} = 14 \pm 1$  мкМ против  $K_d_{ar} = 6,9 \pm 0,7$  мкМ), но оба уступали dr ( $K_d = 1,3 \pm 0,1$  мкМ). Эти результаты согласуются с данными КД-спектроскопии (Рисунок 10А). Они подтверждают отсутствие G4 в средней и хвостовой частях сайта позиционирования нуклеосом (В и С) в конструкциях (1) и (2), в хвостовой последовательности (Dc) контрольной конструкции (1) и в контрольной ДНК (3). Они также показывают наличие слабого G4 в начальном фрагменте сайта позиционирования нуклеосомы (А) и сильного G4 в хвостовой последовательности D.

Слабый G4 во фрагменте А и сильный G4 во фрагменте D были дополнительно подтверждены с помощью  $^1\text{H}$ -ЯМР-спектроскопии (Рисунок 11). В качестве отрицательного

контроля использовали хвостовой фрагмент Dc, не содержащий G4. Области сигнала иминовых протонов во всех спектрах содержали две четко выраженные группы сигналов при 12-13 ppm и 13-14 ppm. Такие химические сдвиги характерны для иминовых протонов гуанина (G-H1) и тимина (T-H3), участвующих в водородных уотсон-криковских связях [211]. В спектрах A и D, соответственно, между 10 и 12 ppm наблюдались дополнительные уширенные сигналы иминовых протонов незначительной и умеренной интенсивности. Такие химические сдвиги характерны для протонов G-H1, участвующих в хугстинговских связях [212,213]. Спектр контрольного дуплекса Dc не содержал признаков протонов, участвующих в хугстинговских связях. Таким образом, сигналы между 10-12 ppm и 12-14 ppm можно рассматривать как сигнатуры G4 и дуплекса, соответственно.



**Рисунок 11** – Проверка фолдинга G4 в ДНК-конструкциях с помощью  $^1\text{H}$ -ЯМР-спектроскопии. (А) Схема равновесия дцДНК $\leftrightarrow$ (G4 в дцДНК) для фрагмента D конструкции (2) и его  $^1\text{H}$ -ЯМР-спектр (области иминовых и ароматических протонов). (Б) Иминовые протоны в тетрадах G4 и уотсон-криковских парах с типичными химическими сдвигами. (В)  $^1\text{H}$ -ЯМР-спектр фрагмента A конструкций (1) и (2) (области иминовых и ароматических протонов) (Г) Спектр фрагмента Dc конструкции (1) (области иминовых и ароматических протонов). Условия: 20 mM Трис-НСl, рН 7,2, и 140 mM КСl). Концентрация ДНК: 0,1 mM. ЯМР-спектры получены Р.А. Новиковым

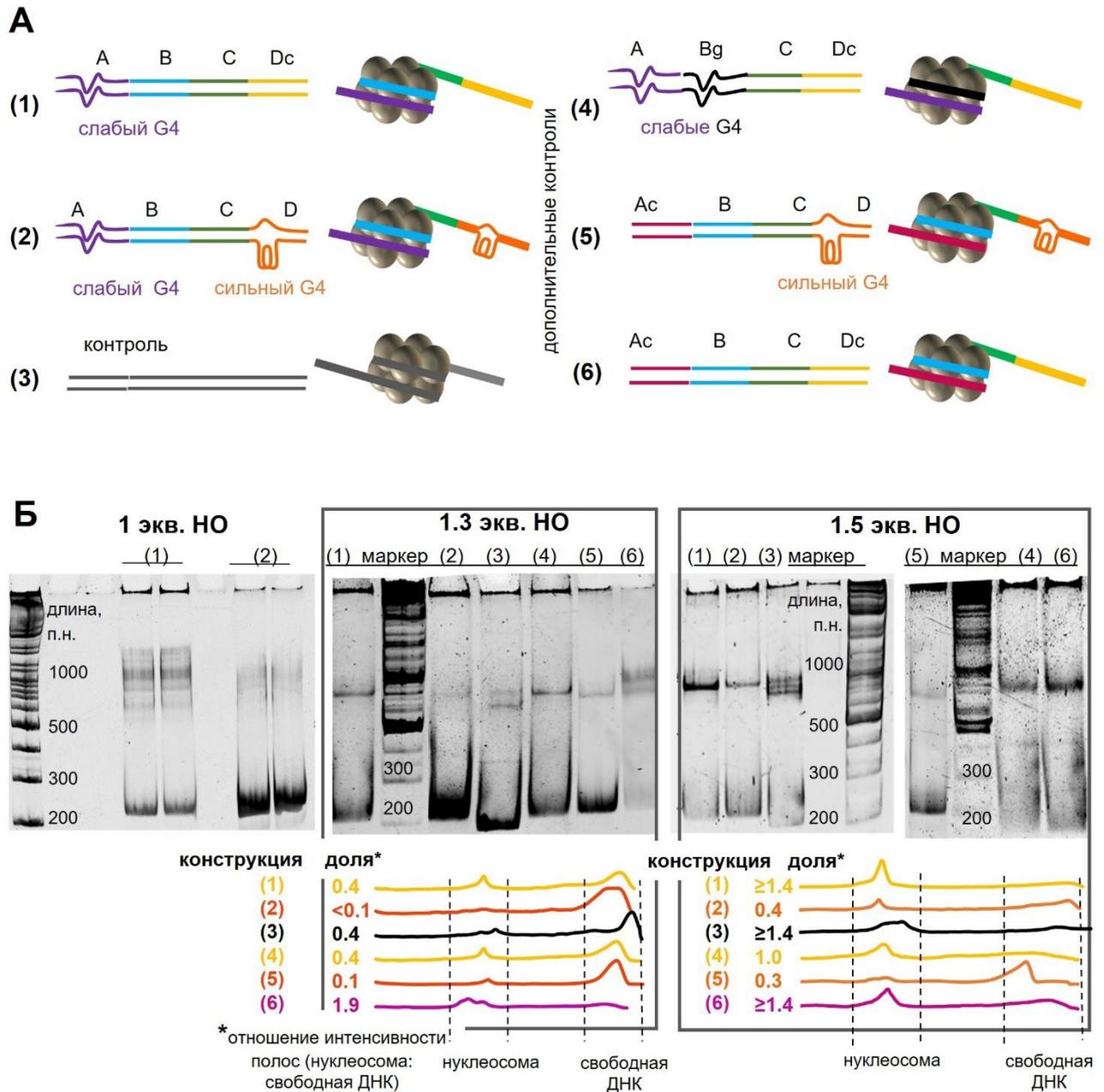
Конструкции со встроенными G4-мотивами, по-видимому, существуют как смеси правильно сложенных структур (G4 в дцДНК) и классических дуплексов (без G4 в дцДНК) (Рисунок 11А). Равновесие должно зависеть как от мотива G4, так и от фланков дцДНК. Чтобы

оценить долю правильно сложенных структур ( $\alpha$ ), были рассчитаны теоретические значения отношения G4:дуплекс (Рисунок 11Б) для A и D, которые сравнились с фактическими интегралами сигналов иминопотонов (Рисунок 11В).

Рассчитанные доли правильно свернутых G4-структур составили 47% для D и 10-15% для A. Следует отметить, что эти доли G4 не обязательно равны полученным при сборке нуклеосом из-за разницы в условиях эксперимента. Высокая ионная сила буфера для сборки нуклеосом приводила к уширению сигнала ЯМР, поэтому вместо него использовали псевдофизиологический буфер с умеренной ионной силой (140 мМ KCl). Кроме того, ЯМР требует относительно высоких концентраций ДНК (100 мкМ), что препятствует прямому сравнению с данными КД. Однако основной результат ЯМР-анализа ( $\alpha_D > \alpha_A$ ;  $\alpha_{Dc} = 0$ , Рисунок 11 согласуется с данными КД-спектроскопии (Рисунок 10А), а также ThT-анализа (Рисунок 10Б) и подтверждает образование слабого G4 в A и сильного G4 в D.

### ***2.2.2 G4 снижают эффективность сборки немодифицированного гистонового октамера и аналога с ацетилированными гистонами H3 на фланкирующем дуплексе***

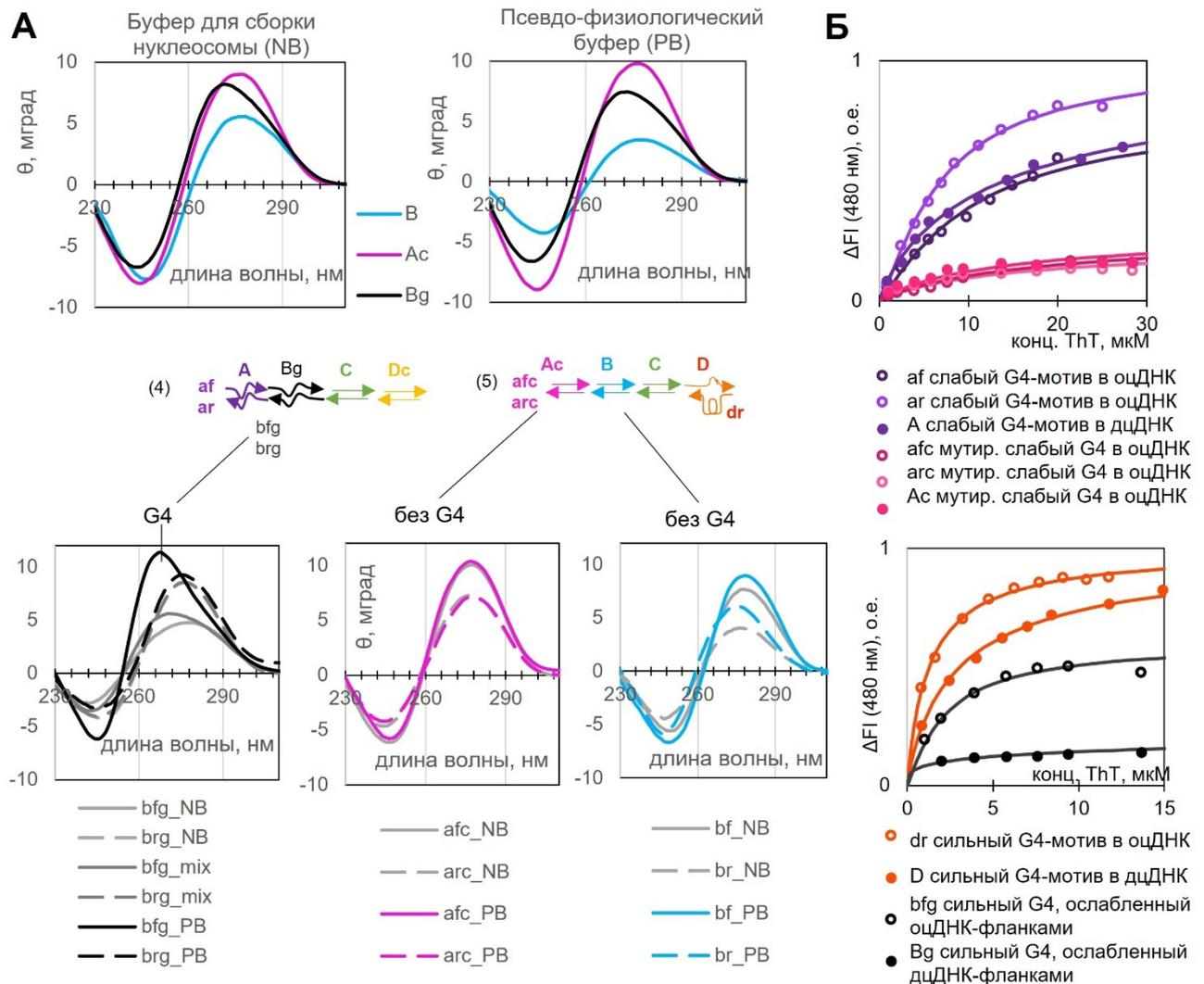
Сборку нуклеосом на конструкциях (1), (2), (3) (Рисунок 12А) проводили путем последовательного разбавления растворов смесей ДНК и гистонов в высокосолевым буфере [214], а результаты анализировали методом анализа смещения электрофоретической мобильности (EMSA) (Рисунок 12Б). При эквимольном соотношении гистоновых октамеров к ДНК нуклеосомы были склонны к проскальзыванию, о чем свидетельствовали диффузные полосы в полиакриламидном геле (ПААГ). Умеренный избыток гистонов повышал выход корректно собранных нуклеосом. Выход нуклеосом был выше для конструкции (1) по сравнению с G4-конструкцией (2), во всех случаях, т.е. при соотношении гистоновых октамеров и ДНК 1:1, 1,3:1 и 1,5:1. Таким образом, конструкция (1) оказалась более надежной матрицей для позиционирования гистоновых октамеров, чем (2). Контрольная ДНК (3) была сопоставима с (1) по выходу нуклеосом, но более диффузные полосы в ПААГ указывали на большее количество побочных конструкций образованных предположительно за счет проскальзывания гистонового октамера (НО).



**Рисунок 12** – Сборка контрольной и G4-содержащей модели нуклеосом. (А) Схематическое изображение конструкций дцДНК и соответствующих предполагаемых нуклеосомных структур. Набор конструкций включает контрольную без G4 (1) и G4-содержащую (2) ДНК длиной 226 п.н., контрольную ДНК длиной 208 п.н.(3) и дополнительные контроли (4)-(6) длиной 226 п.н. (Б) Сборка нуклеосом из смесей ДНК с октамерами гистонов (НО) в отношениях 1:1, 1:1.3 и 1:1.5. Концентрация ДНК: 150 нМ. Условия: 6% неденатурирующий ПААГ, стандартный ТВЕ-буфер, 200 В, прокраска SYBR Gold.

Хотя слабый G4-мотив во фрагменте (А) последовательности позиционирования нуклеосом был в основном развернут согласно данным КД, ЯМР и ThT-анализа, его незначительное влияние на сборку нуклеосом нельзя исключить. Поэтому был получен контрольный дуплекс Ac, который является мутантом А. Мутация была введена во фланкирующую область последовательности позиционирования нуклеосомы, чтобы нарушить

большинство предсказанных G4 в af и ar, оставив нетронутым основной сайт позиционирования нуклеосомы (Таблица 4). КД-спектроскопия и ThT-анализ подтвердили отсутствие G4 в Ac (Рисунок 13). Еще один дополнительный дуплекс, Bg (Таблица 4), был получен для того, чтобы проверить, усиливает ли перемещение сильного мотива G4 из хвостовой (D) в центральную (B) часть последовательности позиционирования нуклеосом его влияние на сборку нуклеосом. КД-спектр Bg содержал сигнатуру G4, а флуоресценция G4-специфичного зонда ThT была сопоставима с таковой для фрагмента A, но слабее, чем для D (Рисунок 13)



**Рисунок 13** – Проверка фолдинга G4 в дополнительных конструкциях ДНК с помощью спектроскопии кругового дихроизма и ThT-анализа. (А) Спектры кругового дихроизма (КД) двухцепочечных (вверху) или одноцепочечных (внизу) фрагментов конструкций. (Б). ThT - анализ. NB, буфер для сборки нуклеосом (20 мМ Трис-НСl, рН 8 и 250 мМ NaCl); PB, псевдофизиологический буфер (20 мМ Трис-НСl, рН 7,2 и 140 мМ KCl); и смесь (NB с добавлением 10 мМ KCl). ThT-анализ проводили в PB. Концентрации ДНК: 1 мкМ в (А) и 0,7 мкМ в (Б). Температура: 20 °С.

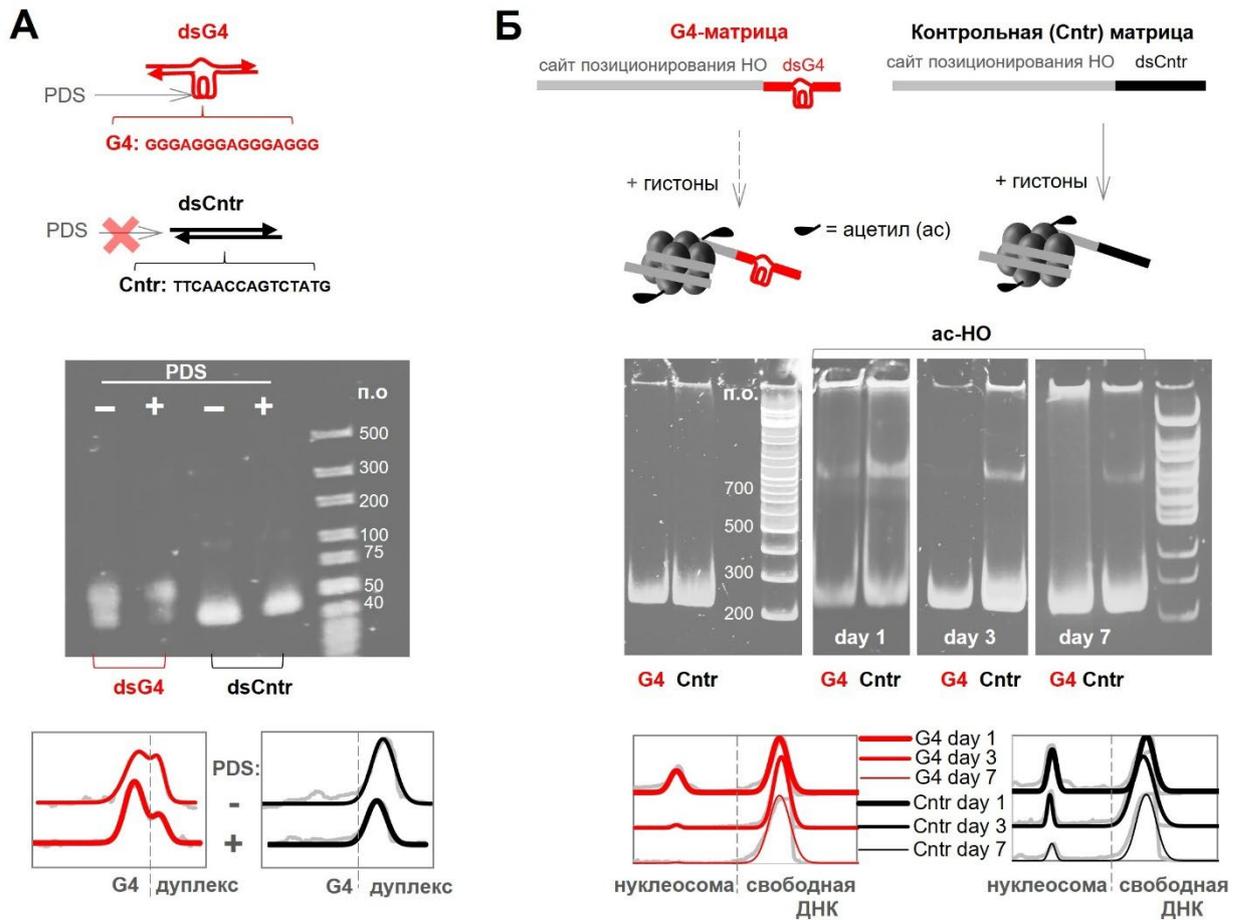
Таким образом, новый дуплекс Ас был аналогом дуплекса А без G4; Bg – аналог В, содержащий G4. С использованием комбинаций исходных дуплексов (А-D/Dс) и новых (Ас и Bg), были получены конструкции (4)-(6) (Таблица 4). Эффективность сборки нуклеосом на Bg-содержащей конструкции (4) была близка к таковой на конструкции (1). Ас-содержащая конструкция (5) с G4 в хвосте, имитирующем линкерную ДНК, показала эффективность сборки нуклеосом, близкую к (2), а конструкция без G4 (6) продемонстрировала самую высокую эффективность (Рисунок 12). Тот факт, что сильный G4-мотив в Bg не препятствует позиционированию гистонов, говорит о том, что он не может эффективно конкурировать с В-формой в присутствии длинных фланков и гистонового октамера, стабилизирующего В-ДНК.

Результаты EMSA, представленные на Рисунке 12, можно обобщить следующим образом:

- Сильный G4 в хвостовой последовательности снижал выход нуклеосом более чем на 50%, что было видно из попарного сравнения конструкций (2) и (5) с их аналогами (1) и (6) без G4 в линкерной ДНК;
- G4 в сайте позиционирования нуклеосом также были неблагоприятны, но не критичны (предположительно потому, что гистоны смещают равновесие G4-дуплекс в сторону дуплекса).
- В целом G4 и нуклеосомы склонны исключать друг друга, хотя правило исключения может быть нестрогим.

По поводу последнего пункта следует отметить, что сильный G4-мотив лишь частично сложен внутри дуплекса D даже в отсутствие гистонов (Рисунок 10,11), а доля нуклеосом ( $\leq 40\%$ , Рисунок 12) не превышает долю несложенного G4-мотива ( $\geq 50\%$ , Рисунок 11). Таким образом, вероятно, что нуклеосомы собираются на тех конструкциях (2) и (5), которые имеют сильные мотивы G4 в несложенном состоянии.

На сходной минимальной модели хроматина была проверена способность G4 к исключению ацетилированных нуклеосом. Последовательность (G4T)<sub>3</sub>G4 была заменена на (G3A)<sub>3</sub>G3 – G4, обогащенный в суперэнхансерах SE (подробнее о выборе G4 в разделе 2.3 данной работы), и оценивалась эффективность сборки НО из смеси гистонов, содержащей ацетилированный НЗ (НЗК18ас), с помощью полиакриламидного электрофореза (Рисунок 14).



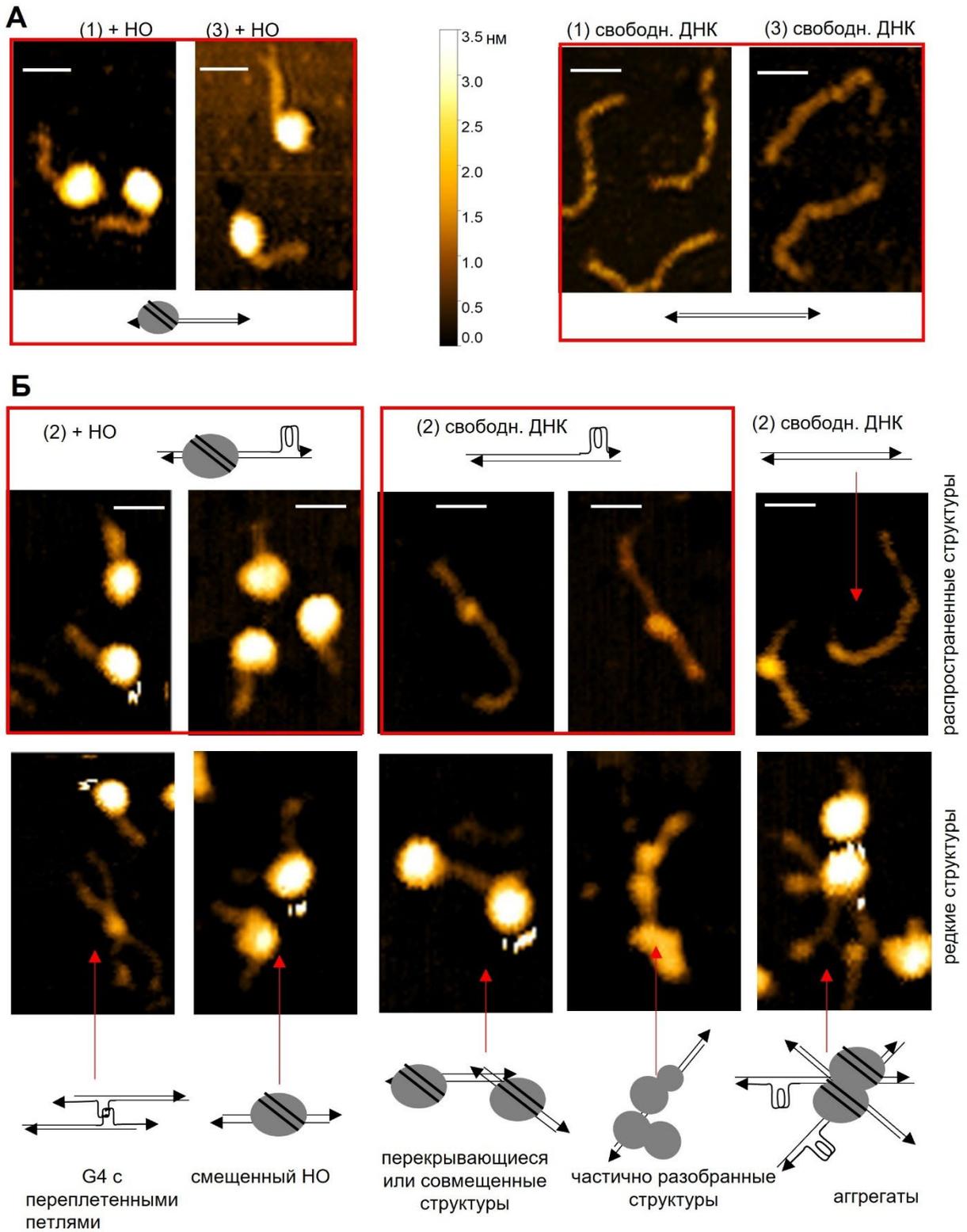
**Рисунок 14** – Влияние SE G4 на сборку нуклеосом. **(А)** Схематическое изображение дуплекса, содержащего SE G4 (dsG4), или контроля без G4 (dsCntr) и проверка их чувствительности к стабилизирующему G4-лиганду PDS с помощью электрофореза в ПААГ. Условия отжига: 1 мкМ dsG4/dsCntr и 0/10 мкМ PDS в 10 мМ буфере Трис-НСl (рН 8) с 125 мМ КСl и 20% ПЭГ. Условия электрофореза в ПААГ: 6% неденатурирующий ПААГ, 1xТВЕ с 10 мМ КСl и 10% ПЭГ, краситель SYBR Gold. Графики внизу иллюстрируют анализ доли сложенного G4 на основе AUC (эксперимент – серый; гауссова аппроксимация – красный/черный). **(Б)** Схематическое изображение фрагментов ДНК Widom с хвостом G4 (G4-Widom) или не-G4 контролем (Cntr-Widom) и анализ сборки гистонного октамера H3K18ac (ac-HO) на этих ДНК в ПААГ. Графики внизу иллюстрируют анализ доли нуклеосомы на основе AUC после инкубации в течение 1–7 дней (эксперимент – серый; гауссова аппроксимация – красный/черный).

Фолдинг характерного для SE G4 в составе ДНК-хвостов – усеченных (41 п.н.) дуплексов был проверен электрофоретически. Подвижность контрольного дуплекса (dsCntr) в ПААГ соответствовала подвижности 41-мерного дуплекса, в то время как G4-содержащий дуплекс (dsG4) демонстрировал дополнительную полосу с пониженной подвижностью, предположительно свидетельствующую о правильно сложенной структуре (Рисунок 14А). Соотношение полос было близко к 1:1, что согласуется с предыдущими оценками на основе ЯМР для аналогичной конструкции с (G<sub>3</sub>T)<sub>3</sub>G<sub>3</sub> (Рисунок 11А). Важно отметить, что PDS увеличил соотношение полос G4:дуплекс почти до 2:1, в то время как контроль не показал никаких

изменений в подвижности, вызванных PDS. В совокупности эти данные позволяют предположить, что SE G4, по крайней мере, частично сложен в хвосте нуклеосомы.

Нуклеосома показала подвижность в ПААГ, близкую к подвижности маркера ДНК размером 700 п.н., что согласуется с подвижностью конструкции с  $(G_3T)_3G_3$  (конструкция (2) на Рисунке 12). Интенсивность полосы нуклеосомы на ДНК с G4 была ниже, чем на контрольной ДНК (конструкция (1) на Рисунке 12), а отношение доли нуклеосом к свободной ДНК составляло ~ 0,4:1 и 0,5:1, соответственно (день 1). Через 3 дня инкубации образца в буфере для сборки нуклеосом соотношение снизилось до 0,03:1 (на G4-содержащей матрице) и 0,15:1 (на контрольной). Кроме того, инкубация образцов с SE-G4 и контрольных образцов в буфере для сборки в течение последующих 3 дней приводила к полной разборке нуклеосом с G4-хвостом, но не контрольных нуклеосом (Рисунок 14). Эти наблюдения согласуются с колокализацией G4-DNS на уровне всего генома [91] и частично объясняют низкую нуклеосомную плотность в G4-богатых SE.

С целью проверки флоретических данных по влиянию G4 на сборку гистоновых октамеров (НО), а также получения качественной информации о морфологии конструкций или их комплексов с НО, основные образцы были дополнительно проанализированы методом атомно-силовой микроскопии (АСМ) (Рисунок 15). АСМ подтвердила сборку нуклеосом на всех ДНК-конструкциях (Рисунки 15,16). Нуклеосомы визуализировались в виде зерен высотой около 4 нм, а их кажущийся диаметр составлял около 6 нм. В соответствии с дизайном конструкций, в большинстве случаев из ядра нуклеосомы выступал один хвост ДНК (Рисунок 15). Также часто встречалась свободная ДНК, а некоторые переплетенные или перекрывающиеся дуплексы ДНК выглядели как крестообразные структуры. В отличие от конструкций (1) и (3) (Рисунок 15А), конструкции (2) часто содержали узлы, которые могли быть сложенными G4 (Рисунок 15Б, верхняя панель). Помимо большого количества правильно собранных нуклеосом и нуклеосом со смещенным НО (Рисунок 15Б, верхняя панель), АСМ также выявила несколько редких объектов, таких как частично разобранные нуклеосомы и агрегаты (Рисунок 15Б, нижняя панель).

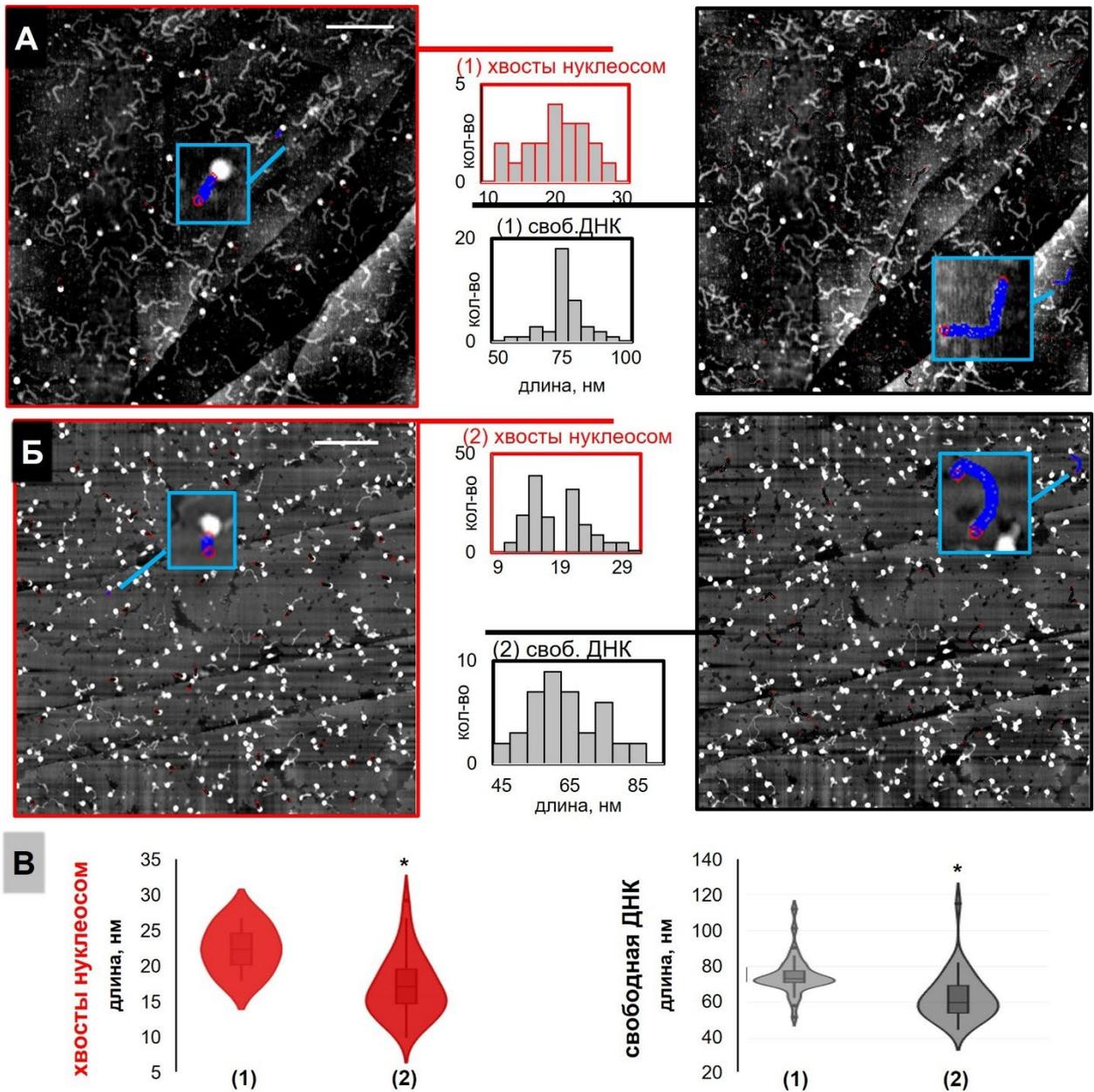


**Рисунок 15** – АСМ-изображения нуклеосом на контрольной (А) и G4-содержащей (Б) ДНК. Масштаб: 10 нм. Изображения получены Н.А. Бариновым.

Нуклеосомы были дополнительно охарактеризованы с точки зрения позиционирования гистонного октамера на ДНК. Рассматривалось два варианта позиционирования гистоновых октамеров: терминальный (визуализированный как зерно с одним хвостом, Рисунок 15Б, верхняя панель) и центральный (визуализированный как зерно с двумя хвостами, Рисунок 15Б, нижняя

панель). В контрольном образце НО+(3) терминальный и центральный варианты расположения гистоновых октамеров были равновероятны. В образцах НО+(1) и НО+(2) их соотношение было примерно равно 1,5:1 и 2:1, соответственно. Неизвестно, были ли нуклеосомы со смещенным НО изначально собраны на центральной части ДНК или же они сформировались как терминальные и впоследствии подверглись сдвигу. В любом случае, корректно собранные нуклеосомы встречались наиболее часто в случаях (1) и (2).

Одним из ключевых вопросов, связанных с морфологией корректно собранных нуклеосом в образце (2), было наличие сложенных G4 в ДНК-хвостах. Разрешение АСМ было недостаточным для визуализации тонкой структуры хвостов, но позволяло приблизительно оценить их длину (Рисунок 16). Фолдинг сильного G4 укоротил бы конструкцию (2) приблизительно на 4 нм (разница между дуплексом длиной 15 п.н. и 3-тетрадным G4). Был проведен анализ распределения длин свободной ДНК (2) и соответствующих хвостов нуклеосом, а также сравнение их с распределениями для конструкции (1). Средняя длина ДНК-хвостов в нуклеосомах с терминальным расположением гистонового октамера составила  $22 \pm 3$  нм и  $16 \pm 4$  нм для конструкций (1) и (2), соответственно (левая панель на Рисунке 16). Несмотря на значительный разброс длины, разница между образцами (1) и (2) была статистически значима (Рисунок 16). Анализ образца (2) в большом поле дал почти биномиальное распределение длины хвоста с максимумами вблизи 15 и 22 нм, предположительно относящимися к сложенным и развернутым хвостам G4, соответственно, при примерно эквимольном соотношении. Эти результаты согласуются с данными ЯМР (Рисунок 11) и позволяют предположить, что G4 частично сохраняются в хвостах правильно собранных конструкций.



**Рисунок 16** – АСМ-анализ морфологии нуклеосом: распределение длины ДНК в смесях ДНК-гистоновый октамер (1:1,5) для хвостов нуклеосом (левая панель **А-В**) или свободной ДНК без нуклеосом (правая панель **А-В**). (**А**) Конструкция (1). (**Б**) Конструкция (2). (**В**) Распределение длины хвостов (слева) и длины свободной ДНК (справа) в образцах (1) и (2). Масштаб: 200 нм. \*t-тест,  $p < 0.05$ .

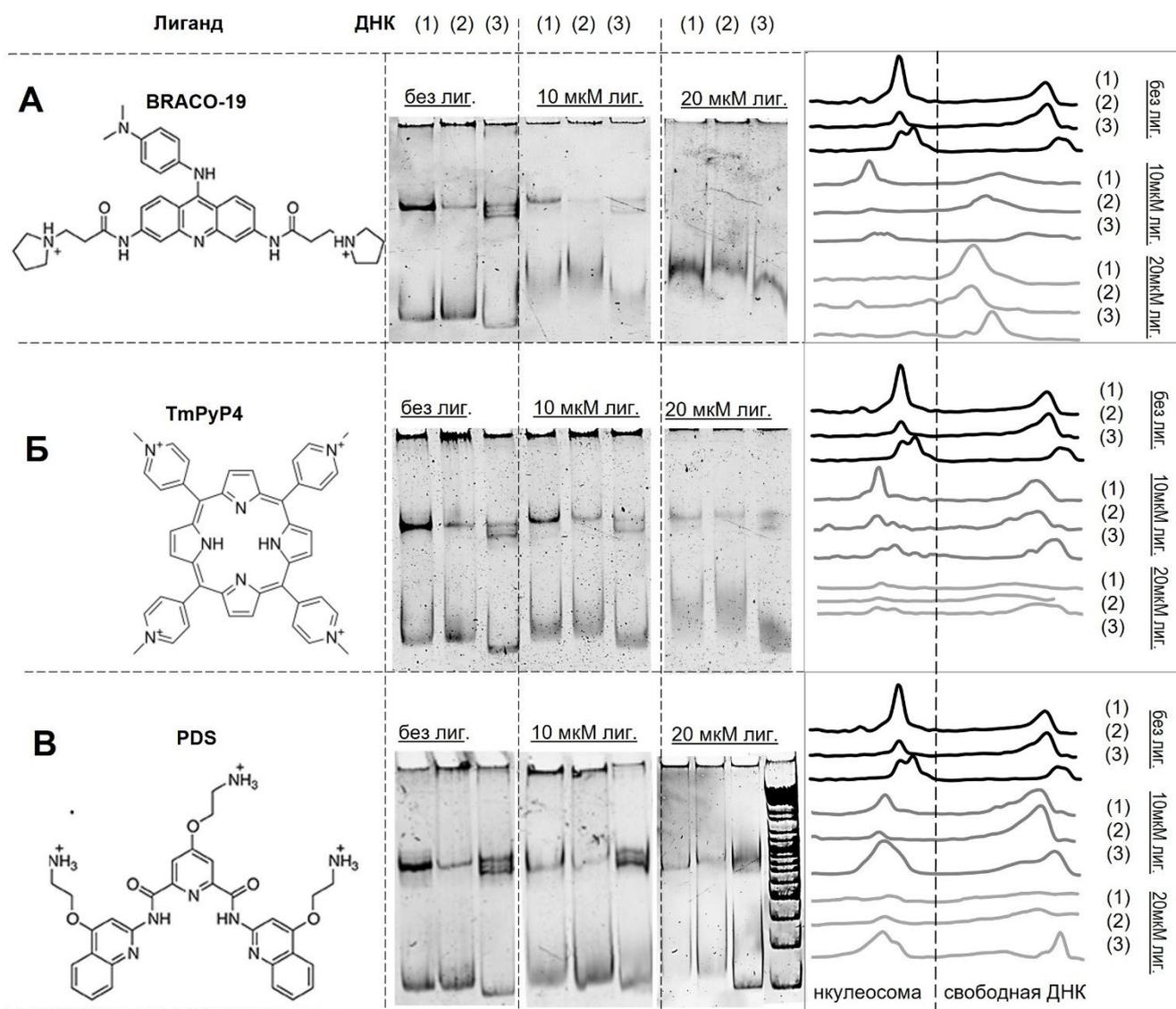
Анализ свободных от нуклеосом конструкций (правая панель на Рисунке 16) дал аналогичные результаты. Видимые длины ДНК  $75 \pm 9$  нм и  $60 \pm 10$  нм в основном соответствовали теоретическим значениям (примерно 76 нм и 72 нм) для (1) и (2) со сложным G4, соответственно. Увеличение разброса длин свободных конструкций по сравнению с ДНК-хвостами нуклеосом и уменьшение средних значений длины по сравнению с теоретическими значениями, вероятно, связано с изгибом ДНК, что затрудняет точные измерения.

Таким образом, АСМ-исследования подтвердили фолдинг сильного мотива G4 в свободной конструкции (2), подтвердили сборку НО на сайте позиционирования нуклеосом и косвенно подтвердили сохранение сильного G4 в хвосте собранной нуклеосомы (2). Хотя во всех образцах содержалась доля конструкций без НО и доля нуклеосом со сдвинутым НО, наиболее распространенные виды соответствовали предложенным моделям, показанным Рисунке 12А.

### ***2.2.3 Известные G4-лиганды ингибируют сборку гистонового октамера на G4-фланкирующем дуплексе***

Модели нуклеосом (1), (2), (3) были использованы для профилирования трех известных G4-стабилизирующих лигандов: BRACO-19, TmPyP4 и пиридостатин (PDS). BRACO-19 отличается значительным G4-стабилизирующим действием, но имеет недостаток – низкую селективность по отношению к G4 в сравнении с дцДНК [215]. TmPyP4 обычно рассматривается как пан-квадруплексный лиганд, но его универсальная стабилизирующая активность в последнее время оспаривается [216,217]. Более того, он также обладает ограниченной селективностью по отношению к G4 [218]. PDS – один из наиболее селективных лигандов, разработанных на сегодняшний день [174], и он широко используется для таргетирования геномных G4 [203].

Нуклеосомы инкубировали с лигандами в концентрациях, сопоставимых или ниже тех, которые обычно используются для модуляции геномных G4 в клетках [35,219], и визуализировали результаты с помощью EMSA (Рисунок 17). BRACO-19 воздействовал как на свободную ДНК, так и на нуклеосомы в концентрации 10 мкМ, вызывая особенно выраженный сдвиг полосы, соответствующей G4-конструкции (2). В концентрации 20 мкМ он разрушал все три нуклеосомы и вызывал выраженные сдвиги свободной ДНК всех конструкций. TmPyP4 оказывал незначительное действие при концентрации 10 мкМ, а при концентрации 20 мкМ вызывал частичную разборку нуклеосом и сдвиг полос свободной ДНК, в частности конструкций (1) и (2). PDS «размывал» полосы свободной ДНК с сильным G4 и незначительно уменьшал фракции корректно сложенных нуклеосом.



**Рисунок 17.** EMSA нуклеосом в присутствии G4-связывающих лигандов BRACO-19 (**A**), TmPyP4 (**Б**) и PDS (**В**). Концентрации: 150 нМ ДНК, 225 нМ гистонового октамера (НО) и 10 или 20 мкМ лиганда. Условия: 6% неденатурирующий ПААГ, стандартный TBE-буфер, 200 В, окрашивание SYBR Gold.

Таким образом, действие лигандов в целом соответствовало опубликованным ранее данным по их селективности [174,203,216]. Все лиганды связывали свободные ДНК-конструкции как с сильным, так и со слабым G4 (1) и (2). BRACO и TmPyP4 также связывали конструкцию ДНК, не содержащую G4 (3). Незначительное снижение выхода нуклеосом при обработке конструкций (1) и (2) G4-селективным лигандом PDS можно объяснить стабилизацией внутринуклеосомного слабого G4, что должно было сместить равновесие "свободная ДНК ↔ нуклеосома" в сторону свободной ДНК. Ярко выраженное нарушение всех трех нуклеосом наблюдалось только для наименее селективного лиганда BRACO-19. Это может быть следствием нецелевых взаимодействий лиганда с дцДНК и искажения геометрии дуплекса. Чтобы

разобраться в этом вопросе, Цветковым В.Б. был проведен *in silico* анализ, включающий докинг лигандов на нуклеосому с G4.

Для *in silico* анализа влияния лигандов В.Б. Цветковым была смоделирована нуклеосома (2) и проведена симуляция молекулярной динамики. Для всех трех лигандов дополнительно проводился докинг сначала на изолированный сильный G4, а затем на полную конструкции (2) с нуклеосомой и G4-хвостом. В двадцати лучших конформациях комплексов с изолированным G4 лиганды располагались поверх внешней G-тетрады. Модуль среднего значения скоров для двадцати конформаций и скор лучшей конформации уменьшались в следующем порядке: PDS > TmPyP4 > BRACO-19, что свидетельствует о том, что PDS является особенно эффективным G4-связывающим агентом.

Анализ, проведённый В.Б. Цветковым, показал, что BRACO, как правило, взаимодействует с бороздкой ДНК-дуплекса, обернутого вокруг гистонового октамера. В наилучшей конформации он находился довольно близко к месту контакта ДНК с гистонами. Двадцать лучших конформаций также включали те, в которых BRACO находился между ДНК и гистонами. Такой режим связывания может объяснить разрушение нуклеосом под действием BRACO, наблюдаемое в EMSA (Рисунок 17). Сравнение скоринговых функций, полученных для комплексов BRACO-нуклеосома и BRACO-G4, указывает на нуклеосому как основную мишень BRACO. TmPyP4 располагался на G4 в большинстве конформаций, в том числе и в наилучшей. Удивительно, что PDS тоже связывался с ядром нуклеосомы, однако, если BRACO, как правило, занимал пространство между ДНК и гистонами, то PDS в большинстве конформаций, включая самую выгодную, располагался в самом центре гистонового октамера. В то время как взаимодействия BRACO с нуклеосомами, вероятно, деформируют ядро нуклеосомы, PDS, по-видимому, помогает удерживать гистоновый октамер вместе, поскольку формирует водородные связи по крайней мере с четырьмя гистонами одновременно. Этот вывод согласуется со стабильностью нуклеосомы при обработке PDS по данным EMSA (Рисунок 12). Важно отметить, что с точки зрения энергии связывания комплекс PDS-нуклеосома был гораздо менее благоприятным, чем комплекс PDS-G4, полученный при докинге PDS к изолированному сильному G4.

Таким образом, результаты докинга и данные EMSA (Рисунок 17) подтверждают, что сильный G4 является основным сайтом связывания PDS и TmPyP4, в то время как BRACO может препятствовать взаимодействию ДНК с гистонами, разрушая нуклеосому.

Подводя итоги раздела, можно констатировать, что наличие G4 в участке, фланкирующем сайт позиционирования нуклеосомы, приводит к  $\geq 4$ -кратному снижению эффективности сборки гистонового октамера (Рисунок 12) – как немодифицированного, так и с ацетилированным

гистоном H3. Часть G4-лигандов способна дополнительно дестабилизировать нулеосомы, однако их эффект скорее определяется нежелательным сродством к дуплексу.

## 2. 3 G4 суперэнхансеров с низкой нуклеосомной плотностью как регуляторы транскрипции

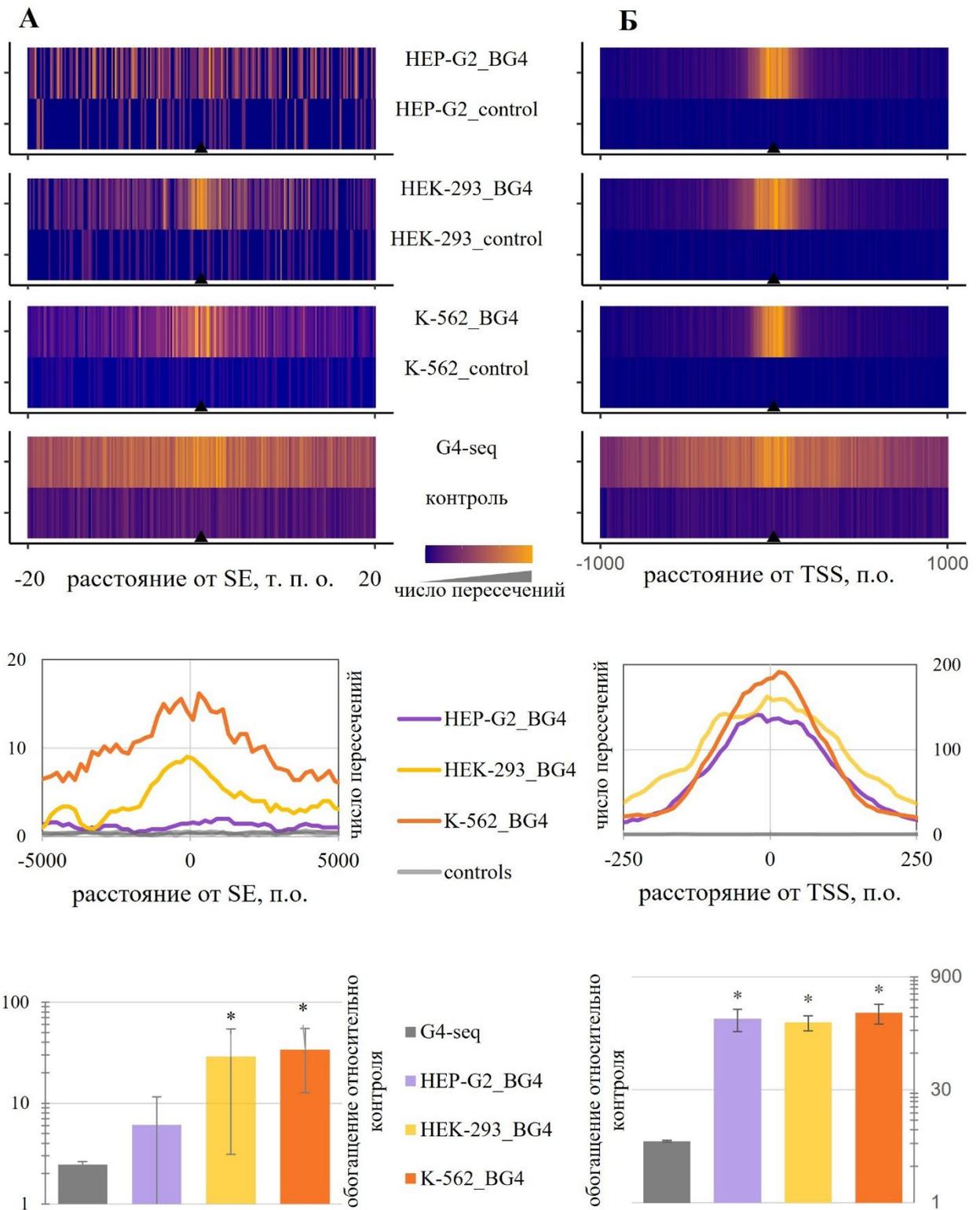
В текущем разделе представлены результаты работы по задачам 4 и 5. Эти задачи вытекают из показанной в предыдущем разделе способности G4 вытеснять ацетилированные гистоны с фланкирующих участков дцДНК, что ставит под сомнение способность специфичных к ацетилированным гистонам белкам удерживаться на G4-богатых сайтах. К таким белкам относится маркер суперэнхансеров (SE) BRD4, который обеспечивает контакты SE-промотор по механизму, зависящему от разделения фаз «жидкость-жидкость» (LLPS), т.е. через образование конденсатов. В рамках задачи 4 было проверено обогащение G4 в SE с использованием ранее опубликованных данных ChIP-seq. Далее был проведен анализ сродства консенсусного G4-мотива из G4-богатых SE (SE-G4) к BRD4 и его влияние на эффективность образования конденсатов *in vitro*. Наконец, в рамках задачи 5 была оценена чувствительность образующихся биоконденсатов к G4-лигандам, а также рассмотрено влияние лигандов на экспрессию генов под контролем G4-богатых SE.

### 2.3.1 G4 обогащены в SE человека и исключают нуклеосомы, но не BRD4

Распространенность G4 в цис-регуляторных сайтах человека [24], включая энхансеры [29], была отмечена для различных типов клеток. Однако значимость обогащения G4 в SE была четко продемонстрирована только в стволовых клетках [220]. Обилие G4 в SE клеток почки эмбриона человека HEK-293 и нескольких раковых клеточных линий было частично проанализировано, но основное внимание уделялось общим (неспецифичным) энхансерам. Чтобы восполнить этот пробел, было проведено комплексное исследование распределения G4 в клетках HEK-293, гепатоцеллюлярной карциноме HEP-G2 и клетках миелоидного лейкоза (K-562).

Во-первых, было проанализировано распределение потенциальных и подтвержденных сайтов G4 относительно установленных SE [221] (Рисунок 18). Потенциальные сайты G4 представляли собой пики G4-seq [8], а подтвержденные сайты G4 - пики ChIP-seq (HEP-G2 и K-562) или CUT&Tag (HEK-293), полученные ранее с использованием G4-специфического антитела BG4. [14,22]. Во всех клеточных линиях пересечения сайтов G4 с SE происходили значительно чаще, чем с контрольными (случайными) сайтами ( $p < 0,05$ ). Обогащение BG4 по сравнению с контролем в SE HEP-G2 ( $6 \pm 5$ ) было близко к обогащению G4-seq во всех SE ( $2,5 \pm 0,2$ ). В SE HEK-293 и K-562 обогащение BG4 по сравнению с контролем было значительным (30

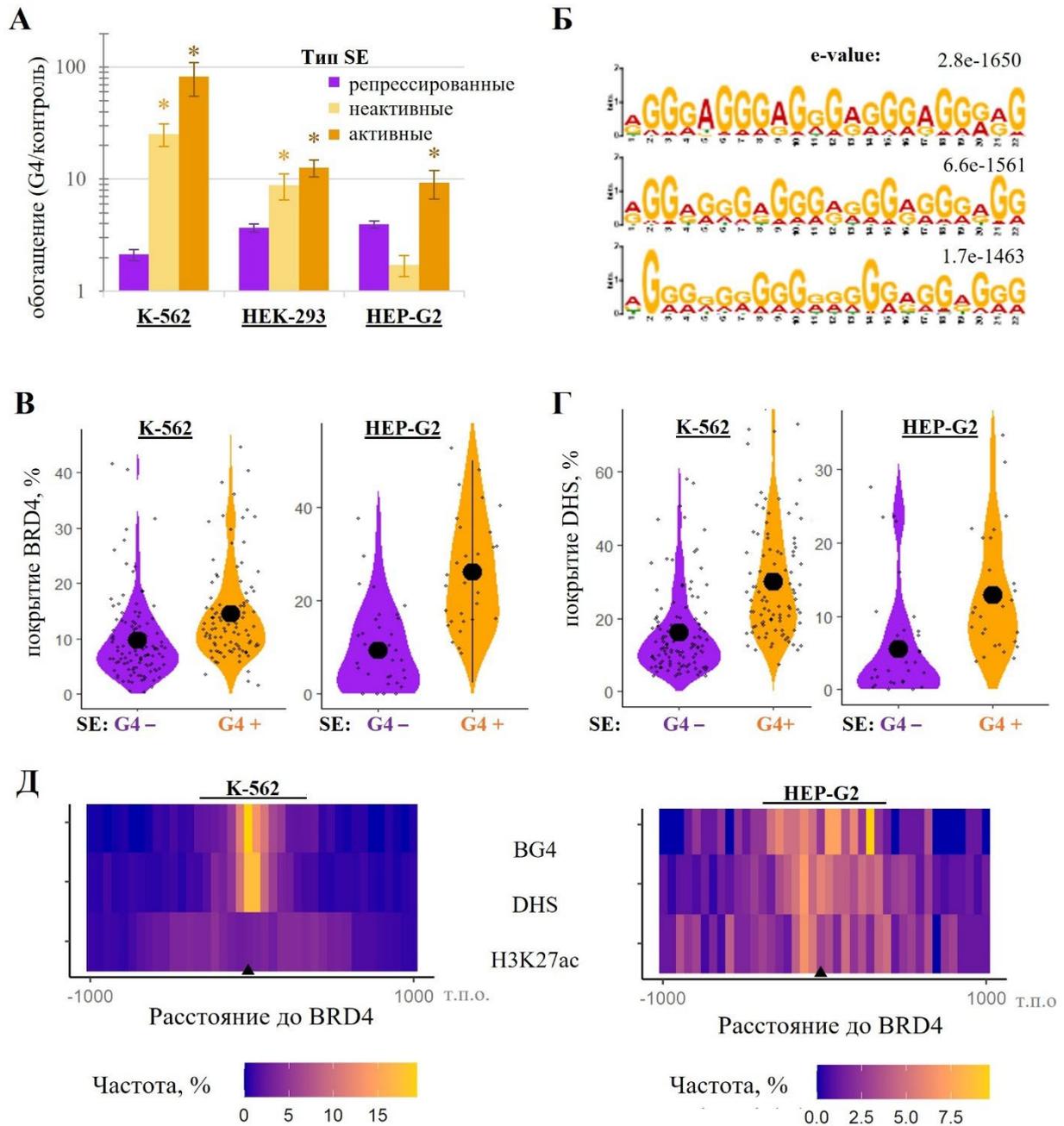
$\pm 20$ ), сравнимым с обогащением в промоторных областях ( $TSS \pm 1000$  п.н.), но примерно в 10 раз ниже, чем в коровых промоторах ( $TSS \pm 100$  п.н.).



**Рисунок 18** – Полногеномный анализ G4: сравнение представленности в SE (А) и промоторах (Б). Верхняя панель: распределение потенциальных (G4-seq) и подтвержденных (BG4) G4-сайтов относительно SE и сайтов старта транскрипции (TSS) в трех клеточных линиях. Каждый набор G4 сравнивали с контрольным набором сайтов той же длины, отобранных случайным

образом по всему геному. Средняя панель: профили распределения, сглаженные методом скользящего среднего (шаг: 5). Нижняя панель: обобщение обогащения G4 непосредственно в SE и промоторах ( $TSS \pm 1000$  п.о.); результаты пермутационного теста. \* Значимые различия между G4-seq и BG4 ( $p < 0.05$ ).

Поскольку наблюдаемые различия в обогащении между клеточными линиями могут быть связаны с различиями в активности SE, критерии включения SE были изучены более подробно. Канонические критерии включают наличие нескольких сайтов связывания транскрипционных факторов (ТФ) и гистоновых меток H3K27ac. В SEdb 2.0 выполнение этих нестрогих критериев обобщается как ранг SE. Более строгим критерием является наличие энхансерной РНК, выявляемой с помощью кэп-анализа экспрессии генов (CAGE) [222]. Для каждой интересующей нас клеточной линии CAGE-позитивные SE из SEdb 2.0 были обозначены как активные (401 из 686 в K-562, 310 из 538 в HEK-293 и 120 из 505 в HEP-G2). Остальные были обозначены как неактивные. SE, которые присутствовали в одной или двух клеточных линиях, но не в исследуемой клеточной линии, были обозначены как репрессированные. Анализ охвата G4 в активных/неактивных/репрессированных SE представлен на Рисунке 19А (пермутационные тесты). Обогащение подтвержденных G4 в HEK-293 и K-562 в неактивных SE было значительно выше ( $p < 0,05$ ), чем в репрессированных SE. Для активных SE обогащение было значительным во всех рассмотренных клеточных линиях (Рисунок 19А).



**Рисунок 19** – Полногеномный анализ G4: обогащение в активных SE, репрезентативные мотивы и распределение относительно сайтов связывания BRD4. (А) Обогащение G4 в SE (пермутационные тесты). Пересечение SE-G4 было рассчитано с использованием данных G4 ChIP-seq/Cut&Tag, полученных с помощью антитела BG4 и координат SE, специфичных для клеточной линии, полученных из SEdb 2.0. Для каждой клеточной линии пересечение SE с G4 сравнивалось с пересечением с контрольным набором того же размера со случайными сайтами той же средней длины. Результаты показаны как среднее значение  $\pm$  SD кратности обогащения G4 по сравнению с контрольными сайтами для репрессированных, неактивных и активных SE. Активные SE присутствуют в SEdb 2.0 и подтверждены CAGE. Неактивные SE присутствуют в SEdb 2.0, но не имеют пиков CAGE. Репрессированные SE отсутствуют в интересующей клеточной линии, но присутствуют в других клеточных линиях, согласно SEdb 2.0. \*Значительная разница между репрессированными и неактивными/активными SE ( $p < 0,05$ , тест Стьюдента). (Б) Мотивы, обогащенные в объединенном наборе G4-сайтов из G4- богатых SE, K-562 и HEK-293 (лучшие результаты MEME). (В) Сравнение активных SE без G4 и G4-богатых SE: нормализованное покрытие пиков ChIP-seq BRD4. (Г) Нормализованное покрытие DHS в

активных SE без G4 и G4-богатых SE. В (В) и (Г) разница между SE без G4 и G4-богатыми SE статистически значима во всех случаях ( $p < 0,05$ , тест Манна-Уитни). (Д) Распределение пиков BG4, DHS и H3K27ac относительно сайтов связывания BRD4 в активных SE.

Далее было проверено, отличаются ли типичные G4 в SE от G4, распространенных по всему геному. Из частично перекрывающихся наборов SE с наибольшим количеством или частотой пиков G4-seq/BG4/G4P (Таблица 5) были отобраны фрагменты с существенной (выше пороговой) склонностью к G4-фолдингу с помощью G4Hunter [223] и определены обогащенные мотивы с помощью алгоритма MEME. Мотив с наивысшим скором (Рисунок 19Б), далее обозначаемый как SE-G4, складывается в G4 параллельной топологии [224]. Так было показано, что G4 значительно обогащены в SE различных клеточных линий, и что параллельные структуры, по-видимому, особенно многочисленны.

**Таблица 5. G4-богатые SE, использованные для выявления обогащенного мотива (SE-G4)**

Тип SE	SE ID в SEdb 2.0 (K-562)	координаты*			ранг	кол-во G4-seq	частота G4-seq	кол-во BG4	частота BG4**	присутствие в HEK-293
		chr	start	end						
активные	SE_01_03900005	chr16	85556307	85656073	5	72	7.22	11	1.10	+
	SE_01_03900048	chr10	72409004	72517319	48	41	3.79	2	0.18	-
	SE_01_03900056***	chr1	27830278	27900086	56	57	8.17	3	0.43	-
	SE_01_03900071	chr3	73159466	73160498	71	1	9.69	1	9.69	-
	SE_01_03900080	chr19	45926050	45989734	80	26	4.08	10	1.57	-
	SE_01_03900094***	chr19	2034446	2062937	94	32	11.23	7	2.46	+
	SE_01_03900154***	chr17	38219164	38281051	154	51	8.24	3	0.48	+
	SE_01_03900294	chr19	18512233	18553632	294	53	12.80	4	0.97	+
SE_01_03900308	chr1	154971359	154990749	308	36	18.57	3	1.55	+	
неактивные	SE_01_03900348***	chr1	45476133	45480956	348	2	4.15	3	6.22	-
	SE_01_03900374	chr19	47728520	47733678	374	10	19.39	4	7.75	+
	SE_01_03900423	chr19	13206330	13215982	423	18	18.65	2	2.07	+
	SE_01_03900471	chr19	49630844	49665470	471	32	9.24	7	2.02	-
	SE_01_03900550	chr7	128466526	128473940	550	12	16.19	3	4.05	-
	SE_01_03900600	chr19	59065060	59072302	600	11	15.19	3	4.14	-
	SE_01_03900613	chr8	144988714	145031500	613	40	9.35	8	1.87	+
	SE_01_03900633	chr17	27050878	27056513	633	5	8.87	4	7.10	-
SE_01_03900648	chr16	28833244	28837373	648	2	4.84	3	7.27	-	
Тип SE	SE ID в SEdb 2.0 (HEK-293)	координаты*			ранг	кол-во G4-seq	частота G4-seq	кол-во BG4 (G4P)	частота BG4 (G4P)**	присутствие в K-562
активные	chrX	start	end							
активные	SE_02_43600001	chrX	39653394	39875900	1	111	4.99	0 (2)	0 (0.09)	-
	SE_02_43600005	chr15	70381020	70520450	5	63	4.52	2 (5)	0.14 (0.43)	-
	SE_02_43600011	chr17	79358647	79405136	11	46	9.89	2 (13)	0.43 (2.80)	-
	SE_02_43600012	chrX	39989857	40037874	12	34	7.08	1 (16)	0.21 (3.33)	+
	SE_02_43600014	chr9	139419483	139522273	14	85	8.27	1 (5)	0.10 (0.49)	-
	SE_02_43600020	chr13	100619662	100646569	20	21	7.80	3 (13)	1.11 (4.83)	-
	SE_02_43600025	chr18	46455559	46548567	25	58	6.24	0 (4)	0 (0.43)	-
	SE_02_43600031	chr15	96865843	96907297	31	37	8.93	0 (14)	0 (3.38)	-
SE_02_43600041	chr19	13123274	13208085	41	59	6.96	0 (2)	0 (0.24)	+	

	<b>SE_02_43600061</b>	<b>chr17</b>	<b>77763199</b>	<b>77788320</b>	<b>61</b>	<b>30</b>	<b>11.94</b>	<b>2 (14)</b>	<b>0.80 (5.57)</b>	<b>+</b>
неактивные	<i>SE_02_43600122</i>	<i>chrX</i>	<i>153233043</i>	<i>153240403</i>	<i>122</i>	<i>6</i>	<i>8.15</i>	<i>1 (9)</i>	<i>1.36 (12.23)</i>	<b>+</b>
	<i>SE_02_43600172</i>	<i>chr17</i>	<i>47071889</i>	<i>47077759</i>	<i>172</i>	<i>15</i>	<i>25.55</i>	<i>1 (3)</i>	<i>1.70 (5.11)</i>	<b>+</b>
	<i>SE_02_43600231</i>	<i>chr16</i>	<i>1818899</i>	<i>1824674</i>	<i>231</i>	<i>5</i>	<i>8.66</i>	<i>1 (6)</i>	<i>1.73 (10.39)</i>	<b>-</b>
	<i>SE_02_43600298</i>	<i>chr12</i>	<i>4378369</i>	<i>4385463</i>	<i>298</i>	<i>16</i>	<i>22.55</i>	<i>0 (5)</i>	<i>0 (7.05)</i>	<b>-</b>
	<i>SE_02_43600387</i>	<i>chr5</i>	<i>142781350</i>	<i>142785986</i>	<i>387</i>	<i>5</i>	<i>10.79</i>	<i>3 (6)</i>	<i>6.47 (12.94)</i>	<b>-</b>
	<i>SE_02_43600460</i>	<i>chr1</i>	<i>154971472</i>	<i>154990869</i>	<i>460</i>	<i>37</i>	<i>19.08</i>	<i>2 (7)</i>	<i>1.03 (3.61)</i>	<b>+</b>
	<i>SE_02_43600510</i>	<i>chr1</i>	<i>182358446</i>	<i>182362440</i>	<i>510</i>	<i>4</i>	<i>10.02</i>	<i>1 (4)</i>	<i>2.5 (10.02)</i>	<b>-</b>
	<i>SE_02_43600518</i>	<i>chr17</i>	<i>48942249</i>	<i>48945892</i>	<i>518</i>	<i>9</i>	<i>24.70</i>	<i>2 (2)</i>	<i>5.49 (5.49)</i>	<b>-</b>

\* сборка генома: hg19

\*\* среднее количество пиков на 10 kb;

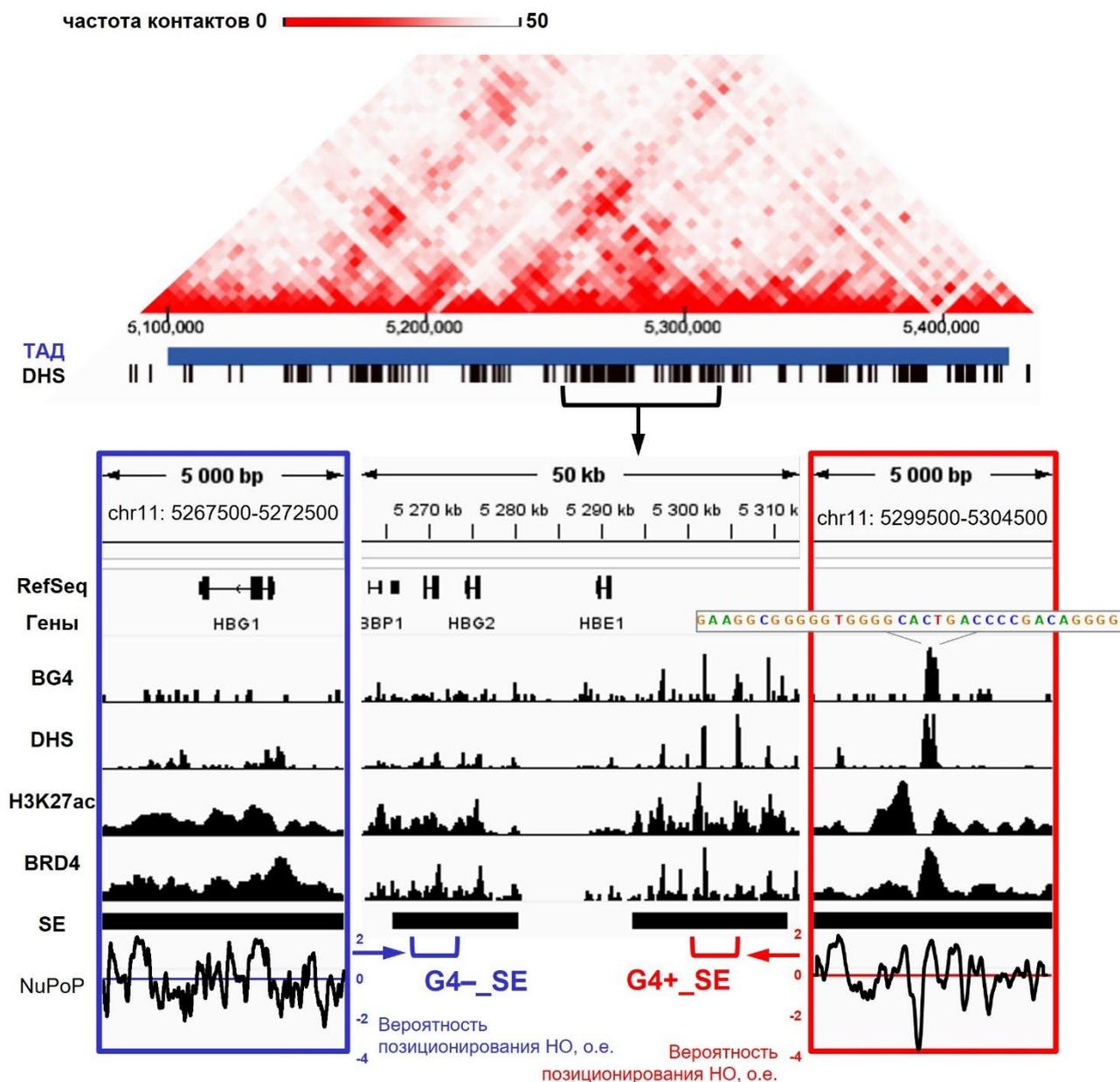
\*\*\*отобраны для анализа изменения экспрессии (специфичные – синий; консервативные – красный)

Чтобы предварительно оценить эффекты обогащения G4, активные SE были ранжированы по покрытию BG4. Набор SE верхнего квартиля назвали G4-богатыми SE, а набор SE с нулевым покрытием BG4 – SE без G4. Для данных наборов SE был проведен сравнительный анализ рангов, нуклеосомной плотности и обогащения BRD4. В анализ были включены клеточные линии K-562 и HEP-G2, для которых были доступны полные данные о распределении маркеров SE. G4-богатые SE и SE без G4 имели схожее распределение рангов во обеих клеточных линиях. В обеих клеточных линиях пики BRD4, гиперчувствительные к ДНКазе сайты (DHS), которые маркируют регионы, обедненные нуклеосомами, и пики H3K27ac, ассоциированные с BRD4, [87] на оставшихся нуклеосомах были значительно более многочисленными в активных G4-богатых SE, чем в активных SE без G4 ( $p < 0,05$ , двухсторонний t-тест, Рисунок 19B). Таким образом, было показано, что G4 могут исключать из SE нуклеосомы, но не BRD4, консенсусным сайтом которого признаны ацетилированные гистоны.

Дальнейшее изучение G4-богатых активных SE выявило колокализацию BRD4 с подтвержденными пиками G4 и сайтами гиперчувствительности к нуклеазам (DHS), в то время как пики H3K27ac группировались во фланкирующих областях ( $\pm 200-500$  п.н.), по крайней мере, в клетках линии K-562. В HEP-G2 общая тенденция была менее очевидной из-за небольшого количества G4-богатых SE (Рисунок 19B). Таким образом, было показано, что позиционирование BRD4 в G4-богатых SE может отличаться от канонического, наблюдаемого в SE без G4.

Пример различных типов SE и паттернов позиционирования BRD4 можно найти в локусе гена  $\beta$ -глобина в хромосоме 11 (Рисунок 20). Его центральная часть содержит гены эмбриональной (*HBE1*) и фетальной (*HBG2* и *HBG1*) субъединиц гемоглобина, которые демонстрируют стабильную и зависимую от условий экспрессию в K-562. [225]. Эмбриональный ген активируется вышележащим SE, также известным как область контроля локуса [226],

который содержит кластер DHS и кластер G4 с двумя достоверными пиками BG4 в K-562. Этот глобиновый SE далее называется G4-богатым. Фетальные гены перекрываются с другим DHS-кластером, содержащим SE, и активируются находящимся в нем регуляторным сайтом HGB-4kb [227]. Этот глобиновый SE не содержит достоверных пиков BG4.



**Рисунок 20** – Примеры G4-богатого SE и SE без G4 из  $\beta$ -глобинового локуса. 3D-контакты в пределах ТАДа, содержащего глобиновый локус (данные Hi-C), и распределение пиков BG4/DHS/H3K27ac/BRD4 (данные ChIP-seq и DNase-seq) в его центральной области, охватывающих фетальные (*HBG1* и *HBG2*) и эмбриональный (*HBE1*) гены глобина с соответствующими SE, показаны для линии клеток K-562. Боковые панели представляют собой репрезентативные фрагменты SE без G4 (синий) и G4-богатого SE (красный). Нижние графики иллюстрируют основанные на последовательностях предсказания сродства ДНК к октамерам гистонов (НО), полученные с использованием инструмента предсказания позиционирования нуклеосом (NuPoP). Серый прямоугольник на правой панели иллюстрирует G-богатый мотив из пика BG4, солокализованного с пиками BRD4 и DHS.

Покрытие пиков H3K27ac и BRD4 в SE глобина с высоким содержанием G4 (21% и 5%, соответственно) почти в 2 раза выше, чем в SE глобина без G4 (12% и 3%). Важно отметить, что, несмотря на схожее общее покрытие DHS (около 10%), SE демонстрируют различное распределение DHS. В SE без G4 оно достаточно хорошо согласуется с предсказаниями позиционирования нуклеосом на основе последовательности (NuPoP), тогда как в SE с G4 пики DHS немного смещены в сторону от предсказанных сайтов без нуклеосом (отрицательные пики в профиле оценки сродства гистонов) и, как правило, колокализуются с G4-сайтами, как, например, репрезентативный пик BG4 на Рисунке 20. Важно отметить, что этот пик также колокализован с пиком BRD4, фланкированным H3K18ac.

Подводя итог, можно сказать, что полногеномный анализ (Рисунок 19) и проверка локуса  $\beta$ -глобина в K-562 (Рисунок 20) показывают, что G4 несовместимы с нуклеосомами, в том числе ацетилированными, но это не препятствует рекрутированию или поддержанию BRD4, что говорит о возможном взаимодействии BRD4-G4. Прямое связывание BRD4 или других BET-белков со свободной от нуклеосом В-ДНК ранее не наблюдалось, тогда как G4-ДНК показала умеренное сродство к гомологам BRD4 [92]. В связи с этим взаимодействие BRD4-G4 было далее проверено *in vitro*, и был выполнен сравнительный анализ сродства BRD4 к G4 и ацетилированным гистонам.

### **2.3.2 G4 из SE проявляет умеренное сродство к BRD4 *in vitro***

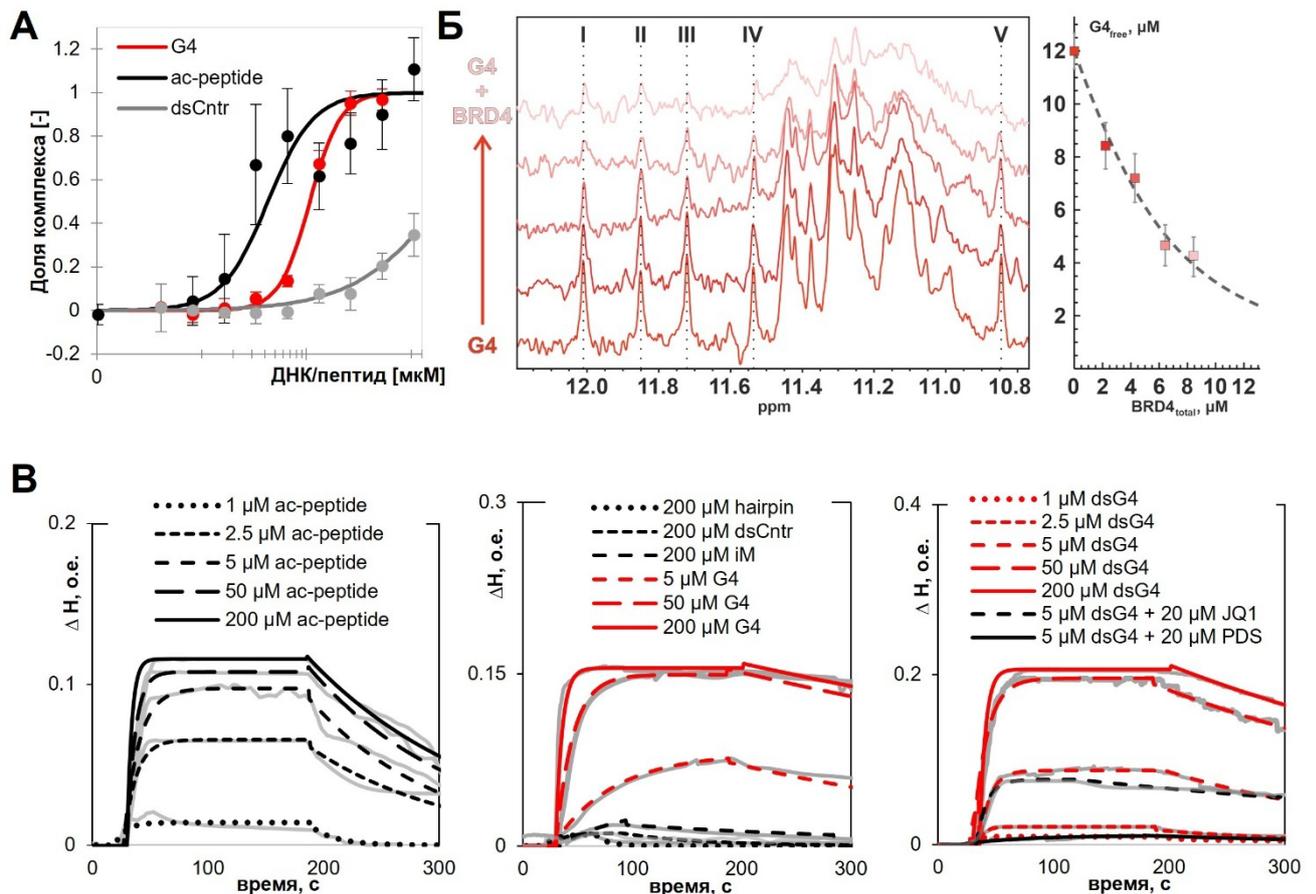
Чтобы проверить возможность поддержания BRD4 на свободных от нуклеосом участках G4, было проанализировано сродство BRD4 к изолированному суперэнхансерному G4 (SE-G4) и к тому же G4 с дуплексными фланками (dsG4). В качестве отрицательных контролей использовали альтернативные неканонические структуры ДНК, а именно стабильный интеркалированный мотив (iM) [205] и шпильку (ds26) [228], а также В-ДНК (dsCntr). Последовательности всех ОДН представлены в Таблице 6. В качестве положительного контроля использовали Кас-содержащий фрагмент H3K18ac, а именно 20-мерный пептид S10-S28 (ас-пептид). Анализ проводили с помощью методов поверхностного плазмонного резонанса (SPR) и микромасштабного термофореза (MST), которые чувствительны к размеру белка/комплекса и сольватационной оболочки, соответственно. Наличие большого неупорядоченного С-концевого домена (CTD) в BRD4 негативно сказывалось на соотношении сигнал/шум. Поэтому был использован частичный рекомбинантный белок, а именно N-концевой фрагмент, охватывающий тандемные BD (E49 – E460), которые обуславливают позиционирование BRD4 на хроматине. Этот частичный белок имел N-концевую полигистидиновую (His) метку, что позволило пометить его флуоресцентным красителем RED-NTA для проведения экспериментов по оценке связывания

методом MST (Рисунок 21А) с флуоресцентной детекцией в красной области спектра или иммобилизовать на  $Ni^{2+}$ -активированном чипе NTA-Au для оценки связывания методом SPR (Рисунок 21В).

**Таблица 6. Последовательности ОДН, использованные в анализах связывания с BRD4 и влияния на разделение фаз BRD4**

шифр	последовательность (5'→3')*
G4	<b>GGGAGGGAGGGAGGG</b>
Cntr	TTCAACCAGTCTATG
ds26	CAATCGGATCGAATTCGATCCGATTG
iM	CCCCCTCCCCCTCCCCCTCCCC
dsG4	Цепь 1: atcacacttgca <b>GGGAGGGAGGGAGGG</b> atgcagctgatca Цепь 2: tgatcagctgcatccctccctccctccctgccaagtgtgat
dsCntr	Цепь 1: atcacacttgcttcaaccagctctatgctgcagctgatca Цепь 2: tgatcagctgcagcatagactggtgaaagccaagtgtgat

\*SE-G4 жирным шрифтом

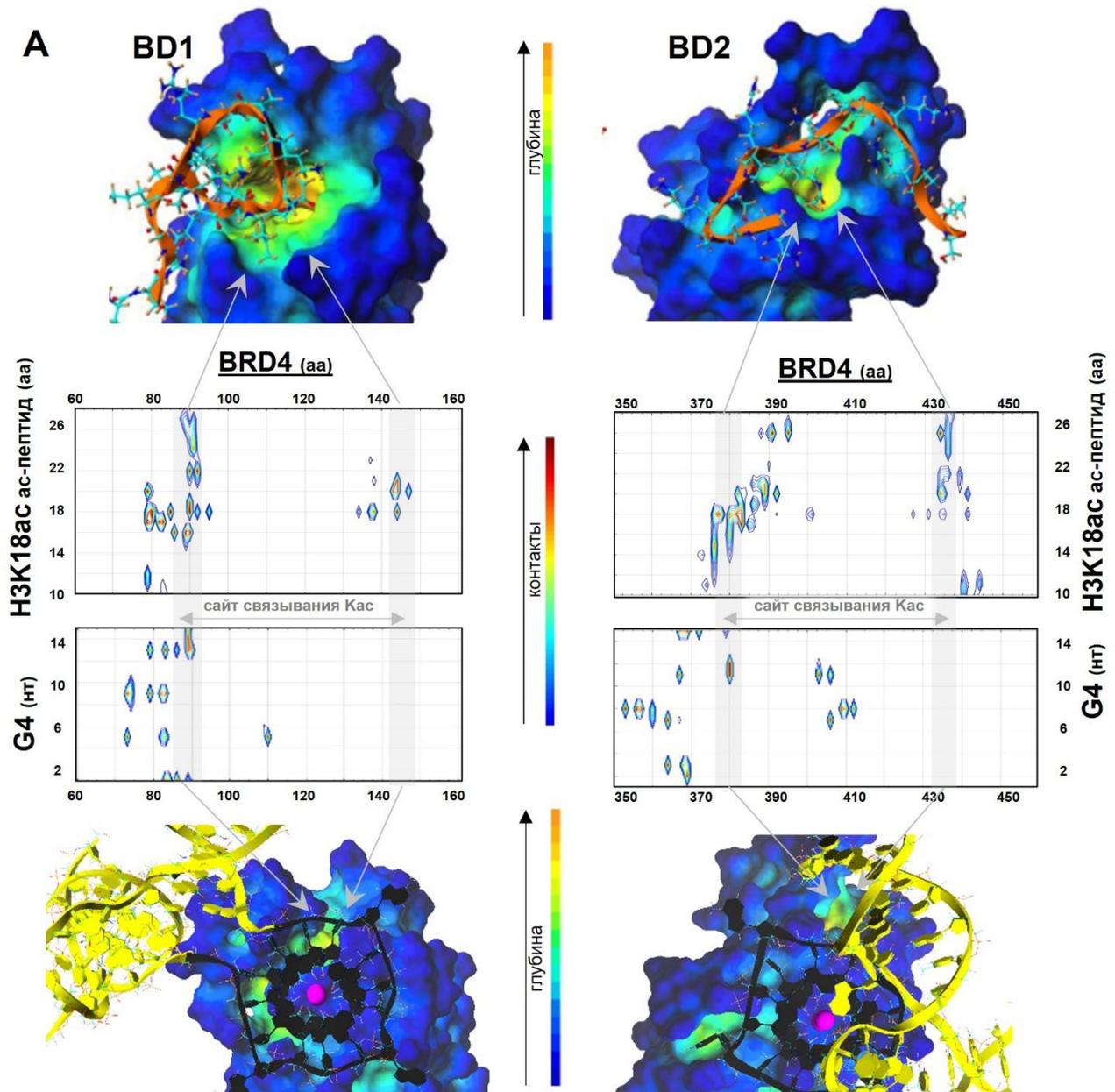


**Рисунок 21** – Анализ взаимодействия SE G4 с BRD4 *in vitro* методами на основе MST, ЯМР, SPR. (А) Анализ на основе микромасштабного термофореза (MST) взаимодействий BRD4 с G4, контрольной ДНК и фрагментом H3K18ac (ac-пептид). Условия: 50 нМ меченого BRD4 в PBS с 0,05% Tween-20. (Б, слева) Спектры ЯМР (700 МГц, 25°C) 24 мкМ раствора SE G4 в рабочем буфере на основе фосфата (20 мМ натрий-фосфата, pH 7,5 и 80 мМ KCl) с увеличивающимися

концентрациями BRD4 (0-8,4 мкМ) и незначительными примесями буфера для хранения белка. Интенсивности спектров были нормализованы для корректировки незначительной зависимости от буфера и разбавления. Конечные буферные условия: 6,1 мМ HEPES, 17,6 мМ натрий-фосфата, рН 7,5, 24 мМ NaCl, 70 мМ KCl, 1,2% глицерина, 4,4% D<sub>2</sub>O и 0,12 мМ DTT. ЯМР-спектры были получены С. Д. Орешковым в лаборатории структурной биологии ионных каналов ИБХ РАН. (Б, справа) Зависимость концентрации свободного мономолекулярного SE G4 от общей концентрации BRD4 в образце. Погрешности рассчитывали на основе спектральных соотношений сигнал/шум. Данные соответствуют простой модели связывания с  $K_d = 6 \pm 1$  мкМ. (В) Анализ на основе поверхностного плазмонного резонанса (SPR) взаимодействий BRD4 с ацетилированным пептидом (ас-пептидом) из H3K18ac (а.о. 10-28), G4, dsG4 и контрольных ОДН (1-200 мкМ в рабочем буфере SPR). Сенсограммы (серые) наложены на моноэкспоненциальную аппроксимацию (красный/черный).

SPR-анализ выявил умеренное сродство контрольного ас-пептида к BRD4 ( $K_d = 3 \pm 2$  мкМ), сопоставимое с данными для других ацетилированных пептидов из N-концевых областей H3/H4 [85], и несколько меньшее сродство SE G4 и dsG4 ( $K_d = 5 \pm 3$  мкМ). MST-анализ дал несколько более высокие значения  $K_d$  ( $4 \pm 3$  мкМ и  $10 \pm 1$  мкМ), но качественно совпал с SPR и с значениями  $K_d$ , полученные методом ЯМР-спектроскопии ( $6 \pm 1$  мкМ, Рисунок 21Б). Ни один из контрольных олигонуклеотидов (dsCntr, шпилька и iM) не показал значительного сродства к BRD4. Интересно, что связывание между dsG4-BRD4 исчезало в присутствии G4-лиганда PDS (4 экв.), но не в присутствии известного конкурента ацетилированных пептидов JQ1 (Рисунок 21В). Было выдвинуто предположение, что способ связывания G4 может отличаться от способа связывания ас-пептида. Данное предположение было проверено В.Б. Цветковым *in silico*.

Для выяснения вероятных режимов связывания G4 с BRD4 и для сравнения их с комплексами гистона с BRD4 В.Б. Цветковым было выполнено молекулярное моделирование. При этом были использованы доступные полноатомные модели бромодоменов BRD4 (PDB: 2UVW и 6U6L), а также разработанные В.Б. Цветковым оригинальные методы моделирования G4 и литературные данные по топологии G4 [224]. Было показано, что сайт связывания G4 смещен относительно консенсусного сайта связывания ацетилированного пептида, ацетилированных гистонов и их конкурентов – известных противоопухолевых ингибиторов BRD4 (Рисунок 22). Из этого следует, что такие ингибиторы BRD4 не должны влиять на взаимодействие BRD4-G4. Кроме того, было показано, что связывание BRD4 с G4 происходит по внешней гуаниновой тетраде. Та же тетрада отвечает за взаимодействие с G4-лигандами, что объясняет ингибирование взаимодействия BRD4-G4 лигандом (PDS).



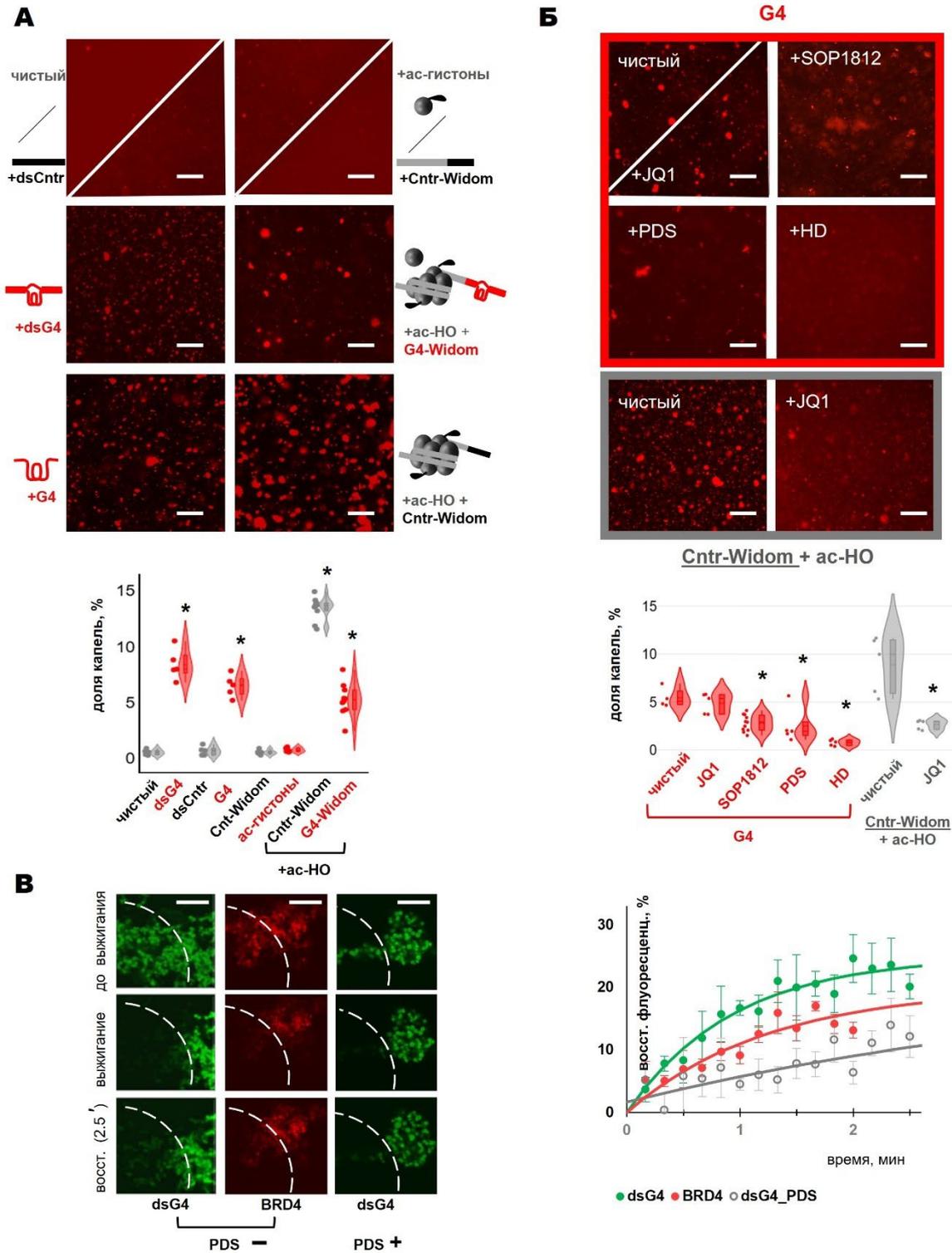
**Рисунок 22** – *In silico* анализ взаимодействий BRD4 с ацетилированным хвостом гистона и SE-G4. Показаны комплексы BRD4 BD1 (слева) и BD2 (справа) с ацетилированным пептидом (ас-пептидом) из H3K18ac (вверху) и SE G4 (внизу) – результаты моделирования молекулярной динамики (МД) и карты контактов. Модели были получены путем стыковки ас-пептида или dsG4 с BD1/2 с последующим моделированием МД. Цветовая схема: коричневый — пептидный остов; черный — G4; желтый — дуплексные фланки и противостоящий G4 С-богатый фрагмент. Карты контактов были получены на основе траекторий МД и иллюстрируют процент снимков с контактами между соответствующими остатками. Поскольку частота контактов между С-богатым фрагментом ДНК и дуплексными флангами была незначительной, карты контактов показаны только для G4, а нуклеотидные остатки пронумерованы от 5'- к 3'-концу. Нумерация аминокислот в ас-пептиде и в BD1/2 согласуется с таковой в H3K18ac и полноразмерном BRD4 соответственно. Данные получены В.Б. Цветковым.

Таким образом, выполненные В.Б. Цветковым исследования *in silico* показали, что dsG4 связывается с BD2 BRD4 через открытые ароматические поверхности гуаниновых остатков из

терминального квартета вне кармана связывания хвоста гистона. Это объясняет сильное ингибирующее действие лиганда PDS, нацеленного на терминальный квартет, и незначительное влияние конкурента гистонов JQ1 на взаимодействие BRD4-G4 в SPR-анализе (Рисунок 21B). Данные *in silico* также показали, что средство dsG4 уступает средству ас-пептида, но значительно превосходит средство dsCntr, что качественно согласуется с данными SPR (Рисунок 21B).

### **2.3.3 Подобно ацетилированным нуклеосомам, G4 из SE способствует образованию биоконденсатов в растворах BRD4**

Согласно опубликованным результатам моделирования LLPS BRD4, в основе формирования конденсатов лежат контакты BRD4 с ацетилированным гистоновым октамером (асНО) и контакты BRD4-BRD4 [229]. В первом случае взаимодействуют тандемные BD, а во втором – IDR, в частности С-концевые области. Предположительно, минимальная модель хроматина с двумя ацетилированными гистонами H3K18ас может быть достаточной для образования конденсата *in vitro*, хотя и при более высокой критической концентрации. Учитывая средство SE G4 к BRD4, продемонстрированное *in silico* и *in vitro*, было необходимо исследовать, может ли этот G4 также влиять на разделение фаз в растворах BRD4. Для наблюдения за разделением фаз с помощью флуоресцентной микроскопии использовали смесь меченного RED (15%) и немеченого полноразмерного белка BRD4 с неупорядоченной С-концевой областью, способствующей LLPS [230]. Для имитации молекулярного краудинга, способствующего LLPS в клетке [231,232], к растворам BRD4 добавляли 20% ПЭГ-400 перед инкубацией с G4, контрольной ДНК, H3K18ас или нуклеосомами. В отсутствие ДНК и гистонов в растворах BRD4 в микромолярных концентрациях (до 4 мкМ) не наблюдалось никаких признаков LLPS, за исключением нескольких протоконденсатов/агрегатов (суб)микронного диаметра. Добавление нуклеосом (ас-НО, собранных на Cntr-Widom и C4-Widom, Рисунок 14) вызывало сильные и умеренные LLPS, соответственно, в то время как ДНК без ацетилированных гистонов или ДНК без G4 не оказывали существенного влияния. ДНК-шпилька ds26 и dsCntr также не оказывали никакого влияния, в то время как G4 и dsG4 были довольно эффективными активаторами LLPS, не сильно уступая стабильной нуклеосоме (Рисунок 23А).



**Рисунок 23** – Влияние G4 на разделение фаз в растворах BRD4. (А) Изображения с флуоресцентного микроскопа смесей BRD4 с G4/контрольной ДНК, ацетилированными гистонами или нуклеосомами. Концентрация BRD4: 4 мкМ (15% RED-меченого). Другие компоненты: 0/4 мкМ dsCntr (вверху, слева); 4 мкМ dsG4 (в середине, слева); 4 мкМ G4 (внизу, слева); 300 нМ НЗК18ас/150 нМ Cntr-Widom (вверху, справа); нестабильная нуклеосома: 150 нМ G4-Widom и 300 нМ ацетилированного гистонного октамера (в середине, справа); стабильная нуклеосома: 150 нМ Cntr-Widom и 300 нМ ацетилированного гистонного октамера (внизу справа). Буфер: 1X PBS (рН 7,4) с добавлением 20% ПЭГ-400. Образцы с G4-Widom/Cntr-Widom также содержали 10 мМ Tris-HCl (рН 8) и 125 мМ NaCl из буфера сборки нуклеосом. Масштаб:

50 мкм. Графики-скрипки иллюстрирует оценку доли капель в каждом образце. \*Значительное увеличение по сравнению с пустым образцом (только BRD4),  $p < 0,0001$ , однофакторный дисперсионный анализ с апостериорным тестом Даннетта. (Б) Изображения с флуоресцентного микроскопа смесей BRD4 с G4 и с добавлением JQ1, PDS, HD (в красной рамке)/смесей BRD4 с контрольной нуклеосомой (Cntr-Widom) и с добавлением JQ1. Концентрация BRD4: 4 мкМ (15% RED-меченого). Другие компоненты: 4 мкМ G4; 150 нМ Cntr-Widom; 20 мкМ PDS, 5 мкМ JQ1, 10% HD. Так как лиганды были растворены в ДМСО, в пустые образцы было добавлено то же количество ДМСО. Буфер: 1X PBS (pH 7,4) с добавлением 20% ПЭГ-400. Образцы с Cntr-Widom также содержали 10 мМ Tris-HCl (pH 8) и 125 мМ NaCl из буфера сборки нуклеосом. Масштаб: 50 мкм. Графики-скрипки иллюстрирует оценку доли капель в каждом образце. \*Значительное увеличение по сравнению с пустым образцом G4/контрольной нуклеосомы с BRD4 (без лиганда),  $p < 0,0001$ , однофакторный дисперсионный анализ с апостериорным тестом Даннетта (для красных скрипок)/t-тест (для серых) (В) Восстановление флуоресценции после фотообесцвечивания в гелеобразных структурах BRD4-dsG4: репрезентативные снимки (обесцвеченные зоны отмечены пунктирными линиями) и кинетические кривые (среднее значение 3 измерений). Масштаб: 10 мкм.

Как и нуклеосомы, G4 и dsG4 индуцировали образование конденсатов диаметром до нескольких микрометров. Визуальный контроль подтвердил их плавучесть и постепенное слияние. В качестве репрезентативного параметра была выбрана объемная доля капель. В соответствии с предполагаемым жидким состоянием капли имели округлость  $80 \pm 10 \%$ , поэтому они были приняты за сферы и их объемная доля оценивалась на основе видимых радиусов, которые, в свою очередь, были рассчитаны по площади проекции на изображениях с микроскопа. Результаты представлены в виде скрипичных графиков на Рисунке 23А и демонстрируют значительное ( $p < 0,0001$ , однофакторный дисперсионный анализ с апостериорным тестом Даннетта) обогащение доли капель в смесях G4 ( $8 \pm 1\%$ ), dsG4 ( $6 \pm 1\%$ ), G4-Widom ( $5 \pm 1\%$ ) и Cntr-Widom ( $13 \pm 1\%$ ) с BRD4 по сравнению со свободным BRD4 ( $< 1\%$ ). Положительное влияние стабильной ацетилированной нуклеосомы на LLPS BRD4 соответствует консенсусной модели SE [83], тогда как эффект нестабильной нуклеосомы, предположительно связанный с G4-хвостом, ранее не наблюдался.

Для исследования колокализации BRD4 с dsG4 в конденсатах был получен FAM-меченый dsG4 и была исследована его смесь с RED-меченым BRD4 с помощью конфокальной флуоресцентной микроскопии. Конденсаты, визуализируемые в каналах флуоресценции RED-BRD4 и FAM-dsG4, были колокализованы. Восстановление флуоресценции после фотообесцвечивания (FRAP) было использовано для проверки соотношения их жидкого и твердого состояния (Рисунок 23В). Сигналы флуоресценции BRD4 (RED) и dsG4 (FAM) частично восстанавливались с константами скорости  $0,8 \pm 0,1 \text{ мин}^{-1}$  и  $1,1 \pm 0,1 \text{ мин}^{-1}$ , соответственно, что свидетельствует об ограниченной, но обнаруживаемой диффузии BRD4/dsG4 и обмене между конденсатами и основным раствором.

Таким образом, было показано, что SE G4, но не В-ДНК, способствует разделению фаз BRD4 при низкой микромолярной концентрации, близкой к константе связывания G4-BRD4. Поскольку связывание BRD4-G4 чувствительно к лигандам (Рисунок 21В), было проверено влияние лигандов на разделение фаз.

#### **2.3.4 Биоконденсаты BRD4-G4 нечувствительны к ингибитору BRD4, но чувствительны к G4-лиганду**

Влияние лигандов G4 и BRD4 на разделение фаз в смесях BRD4 с G4/dsG4 отслеживали с помощью флуоресцентной микроскопии. Нерастворимые в воде/слаборастворимые лиганды разводили в 0,5% ДМСО и сравнивали с действием 0,5% ДМСО на разделение фаз BRD4, так как ДМСО влияет на LLPS [233]. Конкурент ас-гистона JQ1 [234] вызвал трехкратное снижение средней доли конденсатов в смеси BRD4 с ацетилированной нуклеосомой (Рисунок 23Б). Однако он не оказывал существенного влияния на смесь BRD4 с G4 (Рисунок 23Б), что подтверждает различие между способами связывания BRD4 с гистоном и G4.

Разделение фаз в смеси BRD4-G4 почти полностью ингибировалось универсальным ингибитором LLPS, 1,6-гексанидиолом (HD) [235,236], что подтверждает вклад гидрофобных взаимодействий в образование конденсата. Значительное ингибирование разделения фаз также наблюдалось для двух G4-лигандов: пиридостатина (PDS), который обычно используется в качестве эталонного антипролиферативного агента, нацеленного на G4 [174], и SOP1812, который показал эффективность в нескольких моделях рака *in vivo* и в настоящее время проходит клинические испытания [237,238]. Эти эффекты могут быть связаны со способностью лигандов нарушать связывание G4-BRD4, предположительно путем экранирования внешних G-квартетов от ВD BRD4. Контрольный антипролиферативный агент, антиметаболит 5-фторурацил (5FU), который не связывает BRD4 или G4, не вызывал никаких изменений.

Интересно, что наряду с уменьшением доли капель, PDS и SOP1812 способствовали переходу конденсата в гелеобразные или твердые объекты неправильной формы: коэффициент соответствия сферической форме уменьшился с  $0,8 \pm 0,1$  до  $0,6 \pm 0,1$ . Этот эффект может быть связан с образованием транзистентных тройных комплексов (BRD4-G4-лиганд) с пониженной растворимостью. Оба эффекта были особенно выражены в случае PDS, поэтому они были дополнительно проверены с помощью смеси BRD4-dsG4.

Конфокальная микроскопия показала PDS-зависимое уменьшение доли конденсатов и колокализацию сигналов BRD4 (RED) и dsG4 (FAM) в оставшихся объектах неправильной формы. FRAP подтвердил снижение подвижности dsG4 в присутствии PDS (скорость восстановления  $0,2 \pm 0,1 \text{ мин}^{-1}$ , что в 5 раз медленнее, чем в отсутствие PDS, Рисунок 23В).

Подводя итоги, следует отметить, что разделение фаз в растворах BRD4-G4 было нечувствительно к конкуренту гистонов JQ1, но чувствительно к G4-лигандам PDS и SOP1812. Оба этих лиганда ингибировали LLPS и вызывали переход конденсатов в гель. Последний эффект может быть связан с артефактами сорбции, которые являются общим ограничением для *in vitro* анализов LLPS [239]. Еще одним ограничением в наших экспериментах является отсутствие дополнительных компонентов, которые включаются в конденсаты при SE наряду с BRD4 (медиаторный комплекс, энхансерная РНК и т.д.). В связи с этим наши результаты не являются прямым свидетельством внутриклеточных фазовых переходов, но указывают на вероятность таких переходов.

В клетках при нарушении баланса BRD4 или обработке JQ1 ранее наблюдались как растворение конденсатов BRD4, так и их переход в гель, и оба изменения сопровождалось репрессией транскрипции [240–242]. Если G4 способствуют разделению фаз при накоплении BRD4 в SE, то образующиеся конденсаты могут также реагировать на G4-лиганды, что предполагает возможное ингибирование коммуникации SE-промотор и репрессию соответствующих генов. Поэтому нашей следующей целью было проверить, связаны ли антипролиферативные эффекты PDS и SOP1812 хотя бы частично с SE.

### ***2.3.5 Гены, активируемые G4-богатыми SE, демонстрируют повышенную чувствительность к G4-лиганду***

Влияние PDS и других G4-лигандов на экспрессию генов ранее изучалось в полногеномных исследованиях [23,243,244]. Полученные результаты не позволяют отличить эффекты, связанные с промотором, от эффектов, связанных с SE. Поэтому, вместо того чтобы проводить еще один анализ полного транскриптома клеток, обработанных PDS, было выбрано несколько генов в пределах четко определенных ТАДов с подтвержденными G4-богатыми SE, и они сравнились с контрольными генами из ТАДов без SE.

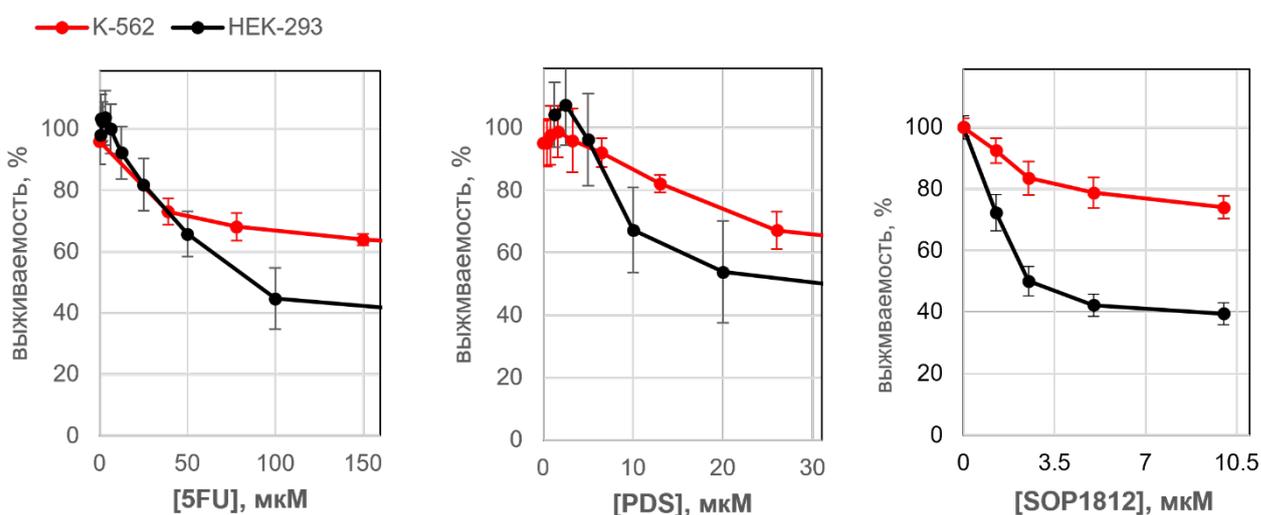
Из набора G4-богатых SE высокого ранга были выбраны два консервативных SE (SE\_01\_03900094 и SE\_01\_03900154, активные в клеточных линиях K-562 и HEK-293) и два SE, специфичных для каждого типа клеток (SE\_01\_03900056, активный в K-562 и репрессированный в HEK-293, и SE\_01\_03900348, активный в K-562 и репрессированный в HEK-293). Затем были подобраны гены, регулируемые этими SE, то есть те, которые расположены в том же ТАДе и демонстрирующие частоту контактов промотора с SE выше среднего, согласно представленным картам 3D генома K-562 на основе Hi-C [245]. Все эти гены содержали пики BG4 в своих промоторах, что согласуется с предыдущими исследованиями участия G4 в коммуникации промотор-SE [1,246]. Контрольные гены были отобраны из ТАДов, свободных от SE. В качестве

эталона использовали ген домашнего хозяйства *RPLP0*, не содержащий BG4 в промоторе. Десять отобранных генов перечислены в Таблице 7.

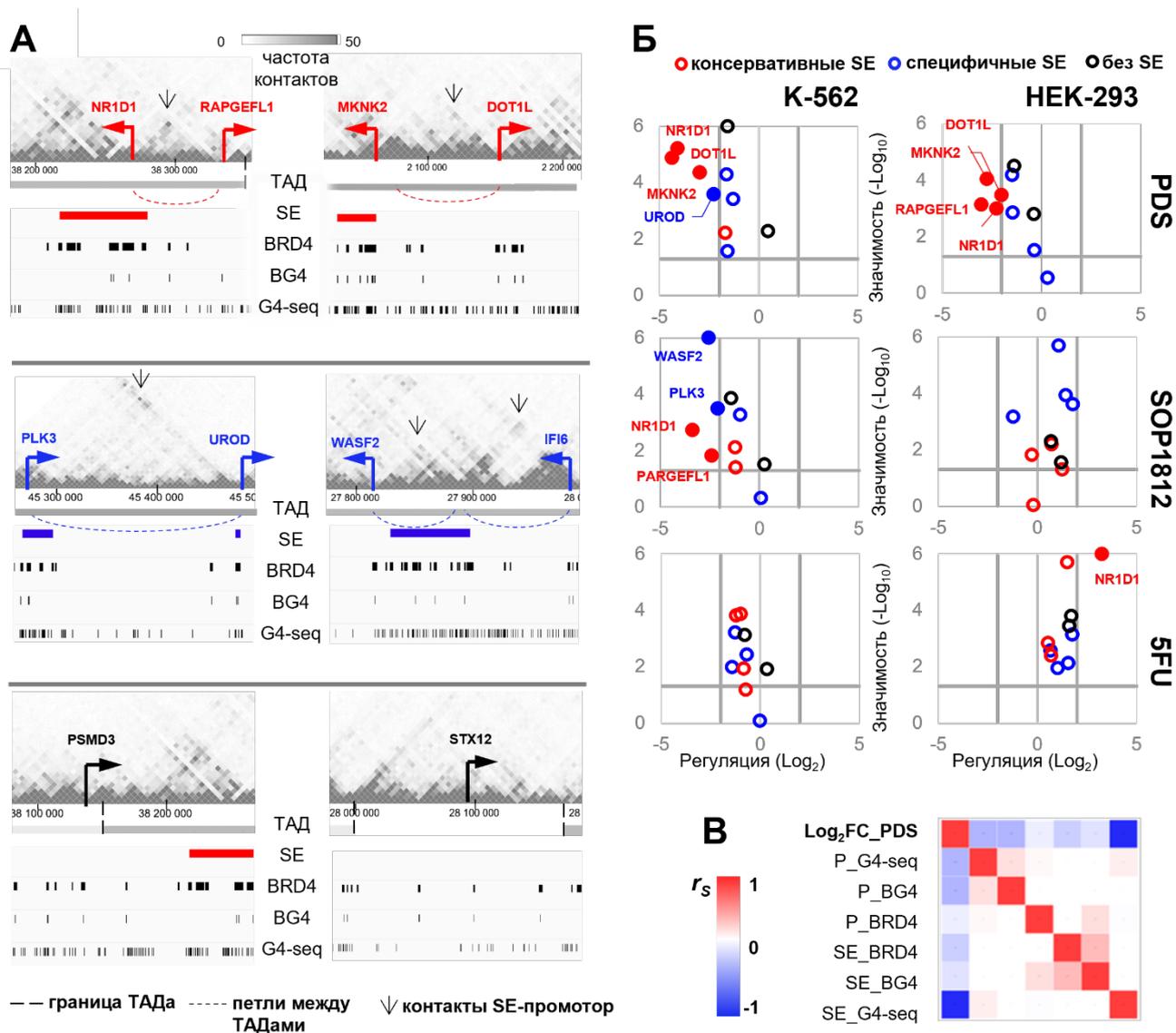
**Таблица 7. Гены, отобранные для оценки изменения экспрессии под действием G4-лигандов**

ген	Праймеры: forward (F); reverse (R)	log <sub>2</sub> FC в К-562 (HEK-293)*	p-value, К-562 (HEK-293)
<i>PLK3</i>	F: tgtcgccaacagaagact; R: ccgtagaagttcacctggaca	-1.30 (-0.35) -2.10 (1.78)	3.80E-04 (3.04E-02) 3.23E-04 (2.41E-04)
<i>IFI6</i>	F: tgggggtggaggcaggaag; R: gatcgagaccagctcatca	-1.58 (0.30) -0.97 (-1.23)	2.71E-02 (2.83E-01) 5.34E-04 (6.62E-04)
<i>WASF2</i>	F: gctggggacttctgggtatc; R: ttctgctggtggaggaggc	-1.62 (-1.46) -2.55 (1.06)	5.00E-05 (1.38E-03) 1.00E-06 (2.00E-06)
<i>UROD</i>	F: ccgttactaccagagtttagga; R: cgtgtcgggtaagggtgat	-2.27 (-1.49) 0.07 (1.41)	2.67E-04 (6.40E-05) 4.85E-01 (1.13E-04)
<i>RAPGEFL1</i>	F: ttgagaacctgacggaccc; R: aggcacgaagggaatcacag	-1.68 (-3.04) -2.42 (-0.29)	6.12E-03 (7.31E-04) 1.53E-02 (1.49E-02)
<i>MKNK2</i>	F: gcagaagaaccagccgaac; R: tgcaggccaagtgcagatc	-2.99 (-2.03) -1.2 (-0.21)	4.30E-05 (3.32E-04) 3.88E-02 (9.00E-01)
<i>NR1D1</i>	F: aacaacaacacaggtggcgtc; R: gggcttgggtgaggaggc	-4.09 (-2.26) -3.37 (1.23)	6.00E-06 (9.96E-04) 1.89E-03 (4.89E-02)
<i>DOT1L</i>	F: gacttggggagcgggtgtg; R: aatgcctctctccaatgtgtat	-4.37 (-2.76) -1.24 (0.71)	1.30E-05 (8.80E-05) 7.59E-03 (6.31E-03)
<i>STX12</i>	F: tcgtcatgtacataggtccc; R: ctttatctgagcagtggtgg	-1.58 (-1.38) -1.45 (0.68)	1.00E-06 (3.10E-05) 1.35E-04 (5.01E-03)
<i>PSMD3</i>	F: cgcctcaaccactatgttctg; R: ggacggaactgtaaatcagcc	0.48 (-0.40) 0.23 (1.20)	5.12E-03 (1.57E-03) 3.12E-02 (2.73E-02)
Референсный ген			
<i>RPLP0</i>	F: gcgggaaggctgtggtg; R: atatgaggcagcagtttctcca	N/A	N/A

\* PDS, черным; SOP1812, синим.



**Рисунок 24 – Цитотоксичность лигандов G4 (PDS и SOP1812) и контрольного соединения (5FU). Выживаемость клеток К-562 или HEK-293 оценивали с помощью реагента PrestoBlue после 48-часовой инкубации с 5FU, PDS или SOP1812.**



**Рисунок 25** – Влияние G4-лигандов на экспрессию генов под контролем G4-богатых SE. (А) Выбор генов, зависящих от консервативных (активных в обеих клеточных линиях; красный) или специфических (в K-562; синий) G4-богатых SE и контрольных, независимых от SE генов (черный). Для каждого набора генов показаны 3D-контекст, ближайшие SE, пики BRD4 и G4 (пики G4-seq и BG4) в клетках K-562. Гены под контролем SE — это те, которые перекрываются с SE или участвуют в дальних взаимодействиях с SE того же ТАДа, на что указывают данные Hi-C (выше средней частоты недиагональных контактов). Независимые от SE гены — это те, которые находятся в области без SE и не имеющие контактов с SE. (Б) Регуляция выбранных генов G4-лигандами PDS/SOP1812 или контрольным соединением 5FU в K-562 (слева) и HEK-293 (справа). Графики рассеяния иллюстрируют строгий анализ (порог  $\text{Log}_2\text{FC} = 2$ , пороговое значение  $p = 0,05$ ). (В) Корреляция (коэффициент Спирмана) между регуляцией генов PDS ( $\text{Log}_2\text{FC\_PDS}$ ) и выбранными характеристиками SE/промотора (P), а именно наивысшим скором пиков G4-seq/BG4/BRD4 в области промотора и покрытием G4-seq/BG4/BRD4 в соответствующем SE.

С помощью ПЦР с обратной транскрипцией (ОТ-ПЦР) сравнивалась экспрессия выбранных генов в клетках K652/HEK-293, обработанных PDS, SOP1812 или контрольным антипролиферативным агентом 5FU в концентрациях, соответствующих значениям IC30 (Рисунок 24).

3D-контакты отобранных генов и результаты ОТ-ПЦР представлены на Рисунке 25А,Б. За единственным исключением (повышение регуляции *NR1D1* в клетках HEK-293, обработанных 5FU), строгий анализ (порог  $\text{Log}_2\text{FC} = 2$ ) не выявил влияния контрольного лиганда 5FU на исследуемые гены. PDS не оказывал существенного влияния на не зависящие от SE гены и гены под контролем специфичных SE ни в одной из клеточных линий, за единственным исключением (снижение регуляции *UROD* в K-562). Однако он снизил уровень экспрессии всех генов под контролем консервативных SE, за исключением *RAPGEFL1* в K-562. SOP1812 не оказал значительного влияния на HEK-293, в то время как в K-562 он снизил уровень половины генов под контролем специфичных/консервативных SE. Таким образом, оба G4-лиганда, но не 5FU, снизили уровень по крайней мере некоторых SE-зависимых генов, причем в K-562 эта регуляция была более выраженной.

Эти результаты могут быть специфичны для выбранных ТАДов и не доказывают критическую роль G4-богатых SE в ответе клеток на G4-лиганды. Однако они ясно показывают, что не только промоторные G4 играют роль в лиганд-зависимой регуляции транскрипции. Для дальнейшего прояснения этого вопроса проводилось сравнение пиков BRD4/BG4/G4-seq в промоторах всех проанализированных генов, и не было обнаружено явной корреляции с изменениями экспрессии, вызванными лигандами (модуль коэффициента Спирмана составил  $< 0,4$ ). Положительная корреляция между G4-seq и BG4 также была слабой или отсутствовала как в промоторах, так и в SE, что свидетельствует о том, что лишь немногие потенциальные G4 складываются в отсутствие стабилизирующего лиганда. Тем не менее, отклик на PDS показал сильную корреляцию с покрытием пиков G4-seq в SE (Рисунок 25В). Коэффициент Спирмана ( $r_s$ ) составил  $-0,8$ , что указывает на эффективное снижение экспрессии генов, под контролем G4-seq-богатых SE (двухсторонний t-тест,  $p < 0,05$ ).

Таким образом, было показано, что гены, зависящие от G4-богатых SE, могут быть мишенью для G4-лигандов, таких как PDS или SOP1812. Одним из возможных механизмов действия лигандов является их способность препятствовать накоплению BRD4 на SE, возможно, в сочетании с вмешательством во взаимодействие транскрипционных факторов с промоторами [35].

Подводя итоги раздела, можно констатировать, что G4, обогащенные в SE человека, исключают ацетилированные нуклеосомы – основной сайт связывания BRD4, но сами способны связывать BRD4. Это позволяет поддерживать мультивалентность SE и количество BRD4 на SE. G4-ДНК суперэнхансеров, но не В-ДНК, способствует разделению фаз в растворах BRD4, и этот эффект сопоставим с эффектом ацетилированных нуклеосом. В отличие от конденсатов в смесях BRD4-нуклеосомы, конденсаты BRD4-G4 нечувствительны к ингибитору связывания гистонов, но чувствительны к G4-лиганду PDS. Эти данные свидетельствуют о том, что PDS может влиять

на функцию SE. Наблюдаемое снижение экспрессии генов, активируемых G4-богатыми SE, подтверждает и подразумевает, что эффекты, связанные с SE, могут перевешивать эффекты, связанные с промотором.

### 3. ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

#### 3.1 Методы исследования роли G4 в демаркации транскрипционных (суб)доменов и привлечении CTCF

##### 3.1.1 *Обработка клеточных культур G4-лигандами и иммунопреципитация хроматина (ChIP)*

Эксперименты были выполнены совместно с сотрудниками лаборатории клеточной биологии ФНКЦ ФХМ. Клетки K-562 культивировали в среде RPMI-1640 (Paneco, Россия) с 0,4%-й фетальной бычьей сывороткой (HyClone GE Healthcare, США). Анализ метаболической активности клеток проводили с помощью окрашивания трипановым синим. G4-стабилизирующие лиганды (PDS или PhenDC3) добавляли к суспензии клеток (1-1,25 млн/мл) до конечных концентраций 10 или 20 мкМ (PDS) и 8 мкМ (PhenDC3). После 24 ч (10 мкМ PDS) или 48 ч (20 мкМ PDS или 8 мкМ PhenDC3) инкубации клетки фиксировали и проводили иммунопреципитацию хроматина (ChIP) с использованием набора SimpleChIP® Plus Enzymatic Chromatin IP Kit с магнитными бусинами (Cell Signalling Technology, США) в соответствии с инструкцией производителя. Вкратце, клетки фиксировали формальдегидом и гасили глицином. Затем клетки инкубировали с буфером для лизиса. Расщепление хроматина проводили путем обработки микрококковой нуклеазой и останавливали действие нуклеазы обработкой ЭДТА. Следующим этапом была соникация лизата для получения фрагментов ДНК размером 200-500 п.н. Для иммунопреципитации хроматина (ChIP)~ 5-10 мкг хроматина инкубировали с 2-4 мкг антитела pAb к CTCF (Active Motif, США) в течение ночи при 4 °С. Нормальный кроличий IgG (Cell Signaling Technology, США) использовали в качестве отрицательного контроля. На следующий день в каждый образец добавляли магнитные бусины Protein G (Cell Signaling Technology, США) и инкубировали в течение 6 ч при 4 °С. Иммуобилизованные комплексы промывали два раза по 10 мин при 4 °С в слабосолевом (1X ChIP-буфер) и высокосолевым (1X ChIP-буфер с 350 мМ NaCl) растворах. Образцы инкубировали с РНКазой А (Cell Signaling Technology, США) в ТЕ-буфере (10 мМ Трис-НСl, рН 8.0, 1 мМ ЭДТА) в течение 30 мин при комнатной температуре. ДНК отделяли от магнитных бусин и инкубировали с протеиназой К в течение 2 ч при 65 °С. Затем ДНК очищали с помощью спин-колонок для очистки ДНК и анализировали методом секвенирования Illumina Next Generation или кПЦР.

### 3.1.2 ChIP-seq и анализ данных

Секвенирование было выполнено сотрудниками лаборатории геномных исследований и вычислительной биологии. Парные библиотеки были подготовлены в соответствии с рекомендациями производителя с использованием набора NEBNext Ultra II ДНК Library Prep Kit (New England Biolabs, США). Библиотеки индексировали с помощью набора NEBNext Multiplex Oligos kit for Illumina (96 индексированных праймеров, New England Biolabs, США). Распределение библиотек по размерам и их качество оценивали на биоанализаторе Agilent с использованием высокочувствительных чипов Agilent ДНК (Agilent Technologies). Количественное определение библиотек проводили с помощью набора Quant-iT ДНК Assay Kit, High Sensitivity (Thermo Scientific, США). Секвенирование ДНК проводили на платформе HiSeq 2500 (Illumina, США) в соответствии с рекомендациями производителя с использованием следующих наборов реагентов: HiSeq Rapid PE Cluster Kit v2, HiSeq Rapid SBS Kit v2 (на 200 циклов), HiSeq Rapid PE FlowCell v2 и 1% PhiX. Эксперимент повторяли дважды. Затем, обработка данных была выполнена сотрудниками лаборатории стабильности генома ИБГ РАН и сотрудником лаборатории системной биологии ФНКЦ ФХМ Р.И. Султановым. Прочтения для каждой биологической реплики были сопоставлены с геномом человека (сборка hg19) с помощью Bowtie2 (версия 2.2.3) с предустановкой '-very-sensitive' [247]. Не уникально сопоставленные прочтения, ПЦР-дубликаты и чтения с MAPQ < 30 были отфильтрованы с помощью 'samtools rmdup' и 'samtools view -h -F 256 -q 30'. Пики выделяли с помощью PePr (<https://github.com/shawnzhangyx/PePr>) с отсечением по *p-value* 0.05 и размером скользящего окна 100 п.н.[248]. Файлы bigWig были сгенерированы с помощью deepTools, версия 2.0 [249].

### 3.1.3 ChIP с кПЦР

Для выборочного анализа сайтов локализации CTCF методом кПЦР были подобраны праймеры, пересекающие или фланкирующие интересующие сайты CTCF (пересекающиеся с пиками BG4, G4-seq, G4-мотивами или не пересекающиеся ни с одним из них) с помощью инструментов NCBI primer-BLAST и Eurofins Genomics primer design (Таблица 1). Эксперименты проводили в QuantStudio5 (ThermoFisher Scientific, США) с использованием 96-луночных планшетов. В каждом эксперименте смешивали образцы, праймеры и SYBR Green PCR Master Mix – коммерчески доступный раствор ДНК-полимеразы, dNTP, MgCl<sub>2</sub> и стабилизаторов (ГенТерра, Россия) – чтобы получить конечный объем 25 мкл. Программа амплификации: 95° С в течение 5 мин (1 цикл), затем двухэтапная реакция 95° С в течение 10 с и 60° С (для пар праймеров STAT3 и VEGFA) или 62° С (для пар праймеров MYC, USP24 и VPS4A) в течение

30 с (40 циклов). Кривые амплификации анализировали с помощью программного обеспечения Thermofisher, а в Microsoft Excel рассчитывали значения  $\Delta\Delta Cq$ , чтобы выразить относительное обогащение ДНК из сайтов связывания CTCF в образцах, обработанных лигандом/не обработанных лигандом, по отношению к исходным:

$$\Delta\Delta Cq = 2^{(-\Delta Cq_{\text{sample}})/2^{(-\Delta Cq_{\text{input}})}};$$

$$\Delta Cq = Cq_{\text{анализируемого региона}} - Cq_{\text{контроля}}$$

В качестве контрольной области использовали сайт *USP24* без G4. Все эксперименты проводились в двух биологических и двух технических повторах.

### **3.1.4 Биоинформатический анализ распределения сайтов связывания CTCF относительно G4**

Поиск G4-мотивов и i-мотивов проводили с помощью программ imGQFinder (<https://github.com/RCPCM-GCB/ImGQFinder>) и G4/iM-Grinder [250]. Для картирования G4-мотивов в imGQFinder были выбраны следующие настройки: число тетрад  $\geq 3$ , максимальная длина петли: 7.

Данные G4-seq (GSM3003539) и G4-ChIP-seq с антителом BG4 (GSE107690) и ChIP-seq CTCF (ENCSR000AKO) были загружены из ENCODE или Gene Expression Omnibus.

Пересечения, оценку относительных расстояний и другие манипуляции с наборами геномных интервалов проводили с помощью программы Bedtools [251]. Статистическая значимость пересечений между сайтами G4 и сайтами, связанными с белком, проверялась с помощью тестов  $\chi^2$ , тестов Фишера и пермутационных тестов. Обогащенные мотивы в пиках, связанных с белками, и отдельные совпадающие последовательности были идентифицированы с помощью алгоритмов MEME [252] и FIMO [253].

### **3.1.5 Олигонуклеотиды, рекомбинантный белок и низкомолекулярные лиганды**

Олигонуклеотиды (ОДН) были получены от компании "Литех", Россия (чистота  $\geq 95\%$ , ВЭЖХ). Рекомбинантный мышинный CTCF с полигистидиновой (His) меткой был приобретен у MyBioSource, США. Для проведения микромасштабного термофореза (MST) CTCF был помечен с помощью набора His-Tag Labeling Kit RED-tris-NTA (NanoTemper Technologies, Германия) в соответствии с протоколом производителя. Лиганды к G4 – пиридостатин (PDS) и PhенDC3 были получены от Sigma-Aldrich (Германия).

### ***3.1.6 Спектроскопия кругового дихроизма и электрофорез***

Растворы ОДН (40 мкМ) в 140 мМ калий-фосфатном буфере (pH 7.2 или 6.7), содержащем 10 мМ NaCl, нагревали до 95 °С в течение пяти минут и быстро (для G4 и i-мотивов) или медленно охлаждали до комнатной температуры (для дуплексов), чтобы обеспечить правильное сворачивание. Затем эти образцы использовались в качестве исходных растворов для проведения анализа связывания методом микромасштабного термофореза. Для оценки фолдинга методом кругового дихроизма (КД) и электрофореза в неденатурирующем полиакриламидном геле предварительно отоженные образцы разбавляли соответствующими буферами до конечной концентрации 1-5 мкМ. Спектры КД регистрировали при комнатной температуре с помощью спектрофотометра Chirascan (Applied Photophysics, Великобритания) в кварцевой кювете с длиной оптического пути 10 мм. Электрофорез проводили в стандартном трис-борат-ЭДТА (ТВЕ) буфере (pH 8) с добавлением 10 мМ KCl, 20% полиакриламида в геле, в течение 2 ч при 200 В при комнатной температуре. В качестве контроля использовали низкомолекулярный маркер оцДНК с диапазоном длин 10-100 нуклеотидов (Affymetrix). Гель окрашивали SYBR Gold (Thermo Fisher Scientific, США) и полосы ДНК визуализировали с помощью сканера Gel Doc (Bio-Rad, США).

### ***3.1.7 Микромасштабный термофорез с CTCF***

Меченый CTCF смешивали с растворами немеченых ОДН до конечной концентрации белка 50 нМ и варьировали концентрацию ОДН (от 20 мкМ до 0,61 нМ). Смеси инкубировали при комнатной температуре в течение 15 мин до проведения измерений методом микромасштабного термофореза (MST). MST-кривые регистрировали с помощью прибора Monolith NT.115, оснащенного детектором RED, в стандартных капиллярах (NanoTemper, Германия) при 22 °С с мониторингом по флуоресценции меченого белка. Зависимость его нормализованной флуоресценции от концентрации немеченого олигонуклеотида анализировали с помощью программы MO.Affinity Analysis (NanoTemper, Германия). Для получения значений констант диссоциации экспериментальные данные анализировали с помощью модели  $K_d$ .

## 3.2 Методы исследования G4 в составе конструкции с нуклеосомой

### 3.2.1 Получение ДНК-конструкции

Целевые ДНК-конструкции, включали сайт позиционирования нуклеосомы длиной 167 п.н., основанный на последовательности длиной 147 п.н., известной как Widom-601 [207] и ДНК-хвост длиной 59 п.н., без G4 (1) или содержащий G4 (2). Конструкции были собраны из дуплексов A, B, C и Dc или D с липкими концами путем фосфорилирования и лигирования, затем амплифицированы и выделены из агарозного геля. Олигодезоксирибонуклеотиды (ОДН), образующие дуплексы A-D/Dc, были приобретены в компании "Литех" (Россия). Схематическое изображение конструкций показано на Рисунке 9А, а все последовательности и шифры ОДН приведены в Таблице 4. Заглавные буквы использовались для дуплексов, а строчные с индексами f/r (forward/reverse) - для одноцепочечных ОДН.

#### Фосфорилирование и лигирование

Полинуклеотидкиназа T4 (ThermoFisher Scientific, США) была добавлена к раствору всех ОДН, кроме 5'-концевых af и dr/dcr, в 1X киназном буфере A (ThermoFisher Scientific, США), с АТФ, до конечной концентрации 0,05 ед/мкл. Конечная концентрация каждого ОДН составляла 0,5 мкМ. Смесь инкубировали при 37°C в течение 30 мин и 10 мин при 75°C. Затем добавляли некинированные концевые ОДН af и dcr или af и dr в эквимолярном соотношении с другими ОДН, и смесь медленно отжигали. Для последующего лигирования к отожженной смеси ОДН (каждый в концентрации 0,1 мкМ) в 1X лигазном буфере добавляли ДНК-лигазу T4 (ThermoFisher Scientific, США) до конечной концентрации 0,125 ед/мкл. Смесь инкубировали при 4°C в течение 2 ч, а затем 10 мин при 70°C.

#### Амплификация

Лигированную дцДНК смешивали с Taq ПЦР Master Mix (GenTerra, Россия) и добавляли праймеры (прямой: TCGGGGGCCGCCCTG; обратный: CCGAGCAGCTGATCACAATT) до конечной концентрации 1 мкМ. Амплификацию проводили с помощью системы Quant studio (Thermo Fisher Scientific, США). Условия: предварительный нагрев при 97°C в течение 1 мин, затем 40 циклов денатурации при 97°C (10 с), отжига при 60°C (15 с) и амплификации при 72°C (20 с). После ПЦР в реакционную смесь добавляли 2%-й перхлорат лития (4X от изначального объема ПЦР-смеси). Смесь инкубировали в течение 3 ч при -20° С и центрифугировали при 14 000 g в течение 30 мин. Осадок промывали сухим ацетоном. Затем ДНК растворяли в загрузочном буфере для электрофореза и наносили на 2%-ный агарозный гель. Гель элюировали

стандартным ТАЕ-буфером и окрашивали EtBr. Целевую ДНК выделяли из геля с помощью набора Cleanup Standard (Evrogen, Россия).

### **3.2.2 Характеризация фрагментов конструкции методами спектроскопии кругового дихроизма (КД) и плавления**

КД-спектры протяженных фрагментов ДНК регистрировали для подтверждения фолдинга G4 в дуплексном окружении, а КД-спектры коротких G4-ОДН – для выяснения топологии G4. Для оценки термической стабильности ( $T_{1/2}$ ) G4-ОДН и их гистерезиса ( $T_{\text{плавление}} - T_{\text{отжиг}}$ ) регистрировали кривые КД плавления и отжига. Все измерения КД проводили на спектрофлуориметре Chrisacan (Applied Photophysics, Великобритания). Спектры растворов 1-4 мкМ ОДН в буфере для сборки нуклеосом (20 мМ Трис-НСl, pH 8, и 250 мМ NaCl) и псевдофизиологическом буфере (20 мМ Трис-НСl, pH 7.2, и 140 мМ KCl) регистрировали при 20°C, а изменения КД в максимуме, специфичном для G4 (265 нм), отслеживали при нагревании/охлаждении образца со скоростью 0,5°C/мин.

### **3.2.3 Подтверждение фолдинга G4 методом флуориметрии с тιοфлавином T (ThT)**

Для подтверждения фолдинга G4 в ДНК-фрагментах были проведены флуориметрические тесты с тιοфлавином T (ThT). ThT, G4-специфичный зонд, не флуоресцирует в свободном состоянии, но флуоресцирует в комплексах с ДНК (в первую очередь G4) [254]. Чтобы избежать образования межмолекулярных структур ДНК, таких как пара G4 с переплетенными петлями, была выбрана схема обратного титрования, которая предполагает увеличение концентрации ThT при фиксированной низкой концентрации ДНК. ThT добавлялся в 0,7 мкМ растворы ДНК в псевдофизиологическом буфере и наблюдалось усиление флуоресценции ThT 485 нм при возбуждении 450 нм. Изменения флуоресценции нормировали на максимальное значение, полученное для G4-ОДН dr; нормированные значения аппроксимировали уравнением для получения кривых связывания и расчета значений  $K_d$ :

$$\Delta FI (485\text{нм}) / \Delta FI (485\text{нм})_{\text{макс.}} = 1 / (1 + K_d^n / [\text{ThT}]^n),$$

где [ThT] - концентрация свободного лиганда, а n - коэффициент Хилла.

Все измерения флуоресценции проводили в трех повторах (стандартное отклонение не превышало 15%) при 20°C с использованием спектрофлуориметра Chrisacan (Applied Photophysics, Великобритания).

### 3.2.4 Подтверждение фолдинга G4 методом спектроскопии ядерного магнитного резонанса (ЯМР)

Для проверки образования G4 внутри дцДНК образцы дуплексов с G4-мотивами (A и D) и контрольный дуплекс без G4 (Dc) были приготовлены в концентрации 0,1 мМ в 0,5 мл буферного раствора H<sub>2</sub>O + D<sub>2</sub>O (10%), содержащего 20 мМ Трис-НСl (рН 7.2) и 140мМ КСl, их отжигали (нагревали до 90 °С в течение 3 мин, а затем быстро охлаждали для облегчения образования мономолекулярного G4) и инкубировали при 4 °С не менее 1 ч для облегчения образования дуплекса. Спектры ЯМР <sup>1</sup>H регистрировал Р.А. Новиков на спектрометре Bruker AVANCE III 400 (400 МГц 4 °С), используя предварительное насыщение или импульсно-полевую градиентную последовательность импульсов WATERGATE W5 (zggpw5 из библиотеки Bruker) для подавления H<sub>2</sub>O. Химические сдвиги <sup>1</sup>H были отнесены к внешнему стандарту, 2,2-диметил-2-силапентан-5-сульфонату натрия (DSS).

Отношение G4:дуплекс рассчитывалось с использованием уравнения:

$$\frac{\alpha \cdot N_{G4 \text{ in } dsDNA}^{H_{G4}}}{\alpha \cdot N_{G4 \text{ in } dsDNA}^{H_{ds}} + (1-\alpha) \cdot N_{no \text{ G4 in } dsDNA}^{H_{ds}}} = \frac{\int_{10}^{12} ppm}{\int_{12}^{14} ppm},$$

где  $N_{G4 \text{ in } dsDNA}^{H_{G4}}$  - число иминопотонов в G4 (до 12 для A, и равно 12 для D);

$N_{G4 \text{ in } dsDNA}^{H_{ds}}$  число иминопотонов в G4-фланкирующих дуплексных фрагментах (до 32 для A и до 40 для D);

$N_{no \text{ G4 in } dsDNA}^{H_{ds}}$  число иминопотонов в длинном дуплексе, полученном при разворачивании G4 и сопряжении оснований по Уотсону-Крику (56 для A и 55 для D);

$\int_{10}^{12} ppm$  интегральная интенсивность сигналов <sup>1</sup>H ЯМР в специфической для G4 области 10-12 ppm;

$\int_{12}^{14} ppm$  интегральная интенсивность сигналов <sup>1</sup>H ЯМР в специфической для дуплекса области 10-12 ppm.

### 3.2.5 Сборка нуклеосом

Сборку нуклеосом проводили методом последовательного разведения и с использованием набора Nucleosome Assembly Kit (EpiMark, США) с димерами и тетрамерами гистонов по протоколу производителя. Вкратце, концентрацию NaCl ступенчато уменьшали от 2М до 250 мМ в 20 мМ Трис-НСl буфере, рН 8, в четыре раунда последовательного разведения и инкубации,

что позволило собрать гистоновый октамер на ДНК-конструкциях (1) и (2), коммерчески доступной контрольной конструкции (3) и дополнительных контрольных ДНК-конструкциях (4), (5) и (6). Все нуклеосомы собирались параллельно. Для визуализации нуклеосомы загружали в 6%-й неденатурирующий полиакриламидный гель. Электрофорез проводили в стандартном буфере трис-борат-ЭДТА (ТВЕ) при 4 °С, гель прокрашивали SYBR Gold (Thermo Fisher Scientific, США) и визуализировали с помощью сканера Gel Doc (Bio-Rad, США). Относительную интенсивность полос свободной ДНК и нуклеосом рассчитывали при обработке скана геля в программе ImageJ. Для ключевых конструкций (1)-(3) было проведено два независимых эксперимента по сборке при каждом соотношении гистоновых октамеров и ДНК, и показаны репрезентативные сканы полиакриламидного геля.

Ацетилированные нуклеосомы собирали на 226-п.н. конструкциях дцДНК, содержащих сайт позиционирования нуклеосомы Widom-601 и ДНК-хвост длиной 59-п.н., имитирующий участок межнуклеосомной ДНК. Использовали две конструкции: без G4 (контрольную), (1), и новую с SE-специфичным мотивом G4 в ДНК-хвосте. Сборку нуклеосом на двух конструкциях дцДНК проводили параллельно методом последовательного разведения с небольшим изменением: немодифицированный гистон H3 заменяли на H3K18ac.

### ***3.2.6 Подтверждение сборки нуклеосом методом атомно-силовой микроскопии (АСМ)***

Эксперименты по АСМ проводил сотрудник лаборатории медицинских нанотехнологий Н.А. Баринов, как описано ранее [208]. На поверхность подложки наносили 5 мкл 150 нМ раствора нуклеосом в 20 мМ Трис-НСl буфере, рН 8 с 250 мМ NaCl, выдерживали в течение 5-15 с, а затем удаляли потоком азота. Все эксперименты проводились как минимум в двух экземплярах с использованием кантилеверов, изготовленных в лаборатории Д.В. Клинова [255], и многомодового АСМ-прибора с контроллером NTEGRA Prima (НТ-МДТ, Россия). АСМ-данные фильтровались и анализировались с помощью программы FemtoScan Online (АТС, Россия), для сглаживания АСМ-изображений применялись стандартные алгоритмы, алгоритмы улучшения разрешения не использовались.

Затем, высоту ОДН анализировали с помощью программы ДНК\_trace3.3.6.msi, а высоту/радиус нуклеосом - с помощью программы Gwyddion.

### ***3.2.7 Электрофоретический анализ влияния лигандов на нуклеосому с G4***

Нуклеосомы, собранные на G4-ДНК или контрольной ДНК, обрабатывали G4-лигандами, а комплексы визуализировали в ПААГ. Лиганды пиридостатин (PDS), BRACO-19 и TmPyP4

были приобретены у компании Sigma Aldrich (США). Каждый лиганд добавляли к раствору нуклеосомы, собранной из смеси 1,5:1 гистоновый октамер (НО)-ДНК (225 нМ гистоновый октамер и 150 нМ ДНК) в рабочем буфере (20 мМ Трис-НСl, рН 8, и 250 мМ NaCl) до конечной концентрации лиганда 10 мкМ или 20 мкМ. Смесь инкубировали при комнатной температуре в течение 15 минут и наносили на 6%-й неденатурирующий полиакриламидный гель. Гели готовили в стандартном TBE-буфере (рН 8), электрофорез проводили в течение 2 ч при 200 В при комнатной температуре со стандартным буфером 1×TBE, затем окрашивали SYBR Gold (Thermo Fisher Scientific, США) и анализировали с помощью сканера Gel Doc (Bio-Rad, США). Во всех анализах ПААГ относительная интенсивность полос ДНК анализировалась в программе ImageJ.

### 3.3 Методы исследования G4 суперэнхансеров в контексте их влияния на транскрипцию

#### 3.3.1 Биоинформатический анализ SE с G4 и без G4

##### Проверка обогащения G4 в SE

Распределение G4-сайтов относительно SE в геноме было исследовано в клеточных линиях K-562 и HEK-293 и HEP-G2. Потенциальные G4 - это те, которые были картированы с помощью G4-seq [8], а подтвержденные G4 - это те, которые были картированы с помощью ChIP-seq с антителом BG4 [22] в K-562 и HEP-G2 [35] или с помощью CUT&Tag с антителом BG4 в HEK-293 [256]. Обработанные данные G4-seq (434272 пика) и данные BG4 ChIP-seq (8955 достоверных пиков в K-562 и 8805 пиков в HEP-G2), а также необработанные данные CUT&Tag были загружены из Gene Expression Omnibus (GSM3003539, G4-seq; GSE107690, BG4 в K-562; GSE145090, BG4 в HEP-G2; GSE220882, BG4 в HEK-293).

Для получения пиков BG4 в HEK-293 необработанные данные CUT&Tag [256] (SRX20765805, SRX20765806 и SRX20765807) были обработаны с использованием платформы Galaxy [257]. Прочтения выравнивали по сборке генома человека hg19 (NCBI версия 37) с помощью minimap2 [258], а дубликаты удаляли с помощью инструментов Picard. Выделение пиков проводили с помощью алгоритма Model-based Analysis of ChIP-Seq (MACS2) [259] с порогом q-value 0,05, что позволило получить 11260 пиков BG4 с высокой достоверностью, большинство из которых (более 60 %) колокализовались ( $\pm 1000$  п.н.) с сайтами G4, картированными с помощью альтернативной методики, а именно ChIP-seq с белком G4P [14].

Интервалы SE для всех клеточных линий были загружены из SEdb 2.0 (<http://www.licpathway.net/sedb>, Sample\_01\_0039, K-562; Sample\_01\_0038, HEP-G2; Sample\_02\_436, HEK-293). Для определения активных/неактивных/репрессированных SE эти интервалы пересекались друг с другом, а пики энхансерной РНК картировались методом кэп-анализа экспрессии генов (CAGE) [222] с использованием пакета BEDTools [251]. Интервалы, присутствующие в наборе данных SEdb для одной или двух клеточных линий, но не для интересующей нас клеточной линии, и не содержащие пиков CAGE, были обозначены как репрессированные. CAGE-отрицательные и CAGE-положительные интервалы, присутствующие в SEdb для интересующей клеточной линии, обозначались как "неактивные" и "активные", соответственно.

Относительные расстояния между SE или сайтами начала транскрипции (TSS) и G4 или случайно выбранными (контрольными) сайтами рассчитывали с помощью инструментов GenomicDistributions из пакета Bioconductor R (версия 3.18). Для каждого набора сайтов G4

(потенциальных/подтвержденных в K-562/HEK-293/HEP-G2) использовался отдельный контрольный набор с тем же количеством интервалов длиной, соответствующей средней длине интервала из набора G4. Эти контрольные наборы были созданы с помощью инструмента RandomBed из пакета BEDTools [251].

Для проверки обогащения G4 в активных/неактивных/репрессированных SE проводились пермутационные тесты. Для каждого набора G4 с помощью инструмента Intersect из BEDTools рассчитывали долю перекрывающихся интервалов SE и сравнивали с соответствующим контрольным набором. Процедуру повторяли  $n$  раз ( $n \geq 20$ ), а статистическую значимость обогащения G4 по сравнению с контролем оценивали с помощью эмпирического  $p$ -value Дэвисона и Хинкли:  $p = (r+1)/(n+1)$ , где  $r$  - число реплик, в которых пересечение контроль-SE было больше, чем пересечение G4-SE [260].

#### Поиск репрезентативных G4-мотивов

Репрезентативные G4-мотивы, обогащенные в SE K-562/HEK-293, были определены с помощью трехэтапной процедуры. Сначала активные и неактивные SE были отсортированы по частоте или общему количеству пиков G4-seq/BG4/G4P. Были отобраны 5 SE с наибольшей частотой/количеством G4-пиков. Набор SE, полученный после удаления дубликатов, включал 18 SE. Далее эти SE были проанализированы с помощью программы G4Hunter [223] для отбора конкретных последовательностей со склонностью к образованию G4 выше порогового значения ( $\text{score} \geq 1.2$ ). Полученные наборы коротких последовательностей с высоким скором объединялись для получения набора предсказанных G4 из SE. Наконец, с помощью алгоритма MEME [252] были определены мотивы, обогащенные среди предсказанных G4 из SE. Для лидерного мотива MEME, наличие совпадающих/гомологичных последовательностей было проверено с помощью алгоритма FIMO [253].

#### Сравнение G4-богатых SE и SE без G4

Сравнительный анализ ключевых эпигенетических особенностей G4-богатых SE и SE без G4 был проведен для клеточных линий K-562 и HEP-G2. Хотя чувствительность к G4-стабилизирующим лигандам и некоторые другие особенности SE могут зависеть как от подтвержденных (ChIP-seq пики), так и от потенциальных (G4-seq пики) G4-сайтов, эпигенетический профиль должен быть связан в первую очередь с подтвержденными G4. Поэтому мы рассматривали только данные G4-ChIP-seq и обозначили SE K-562 с частотой пиков BG4, попадающих в верхний квартиль, как G4-богатые (171 из 687 K-562 SE), а те, в которых отсутствовали пики BG4, как SE без G4 (205 из 687 SE). G4-богатые SE, и SE без G4 сравнивали

по их рангам в SEdb 2.0 (<http://www.licpathway.net/sedb>), обогащению сайтами гиперчувствительными к ДНКазе (DHS), пиками BRD4 или гистоновыми метками. Координаты DHS, ранее картированных с помощью DNase-seq, были загружены из Gene Expression Omnibus (GSM816655 в K-562 и GSM816662 в HEP-G2). Координаты пиков BRD4 и H3K27ac в HEP-G2 также были загружены из Gene Expression Omnibus (GSM2635249, BRD4; GSM5866899, H3K27ac). В случае K-562 были доступны только широкие пики BRD4 и (GSM2635249) и необработанные данные H3K27ac (GSM5866899), поэтому данные были (повторно) обработаны с помощью MACS2 для выделения пиков, как описано выше. Полученные пики пересекались с G4-богатыми SE или SE без G4 с помощью пакета BEDTool [251], а результаты обобщались в виде графиков процента пересечения. Поскольку покрытие DHS/BRD4/H3K27ac демонстрировало асимметричное распределение в одном или обоих наборах SE, значимость различий между наборами SE проверяли с помощью *u*-теста Манна-Уитни.

Распределение пиков BRD4/G4/DHS/H3K18ac в репрезентативных G4-богатых SE и SE без G4 из локуса  $\beta$ -глобина было визуализировано в браузере IGV, а данные Hi-C [245]- в геномном 3D-браузере Encode (<http://3dgenome.fsm.northwestern.edu/>). Сродство ДНК к гистоновым октамерам в глобиновых SE рассчитывали с помощью пакета для предсказания позиционирования нуклеосом Bioconductor (<https://bioconductor.org/packages/release/bioc/html/NuPoP.html>).

### ***3.3.2 Микромасштабный термофорез и поверхностный плазмонный резонанс с BRD4***

#### *Микромасштабный термофорез (MST)*

Рекомбинантный фрагмент BRD4 человека (Glu49 - Glu460) с N-концевой полигистидиновой меткой (His), полученный от Bio-Techne, США, был помечен флуоресцентным красителем RED с помощью набора RED-tris-NTA (NanoTemper Technologies, Мюнхен, Германия) в соответствии с протоколом производителя и смешан с растворами немеченых ОДН в буфере PBST (137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub> и 1,8 mM KH<sub>2</sub>PO<sub>4</sub>, 0,05% Tween-20) до конечной концентрации белка 50 нМ. Ас-пептид, фрагмент гистона H3K18ac (чистота >90%, ЖХ-МС), был приобретен в Центре коллективного пользования Института биомедицинской химии имени Ореховича РАН. Смеси BRD4- ОДН /ас-пептид выдерживали при комнатной температуре в течение 15 мин. Затем регистрировали MST-кривые с помощью Monolith NT.115 (NanoTemper, Мюнхен, Германия) при 22 °C с мониторингом флуоресценции в режиме RED. Зависимость нормализованной флуоресценции от концентрации ОДН анализировали с помощью программы MO.Affinity Analysis (NanoTemper, Мюнхен, Германия) и аппроксимировали моделью Хилла. Эксперименты проводили в трех повторах.

### Поверхностный плазмонный резонанс (SPR)

Измерение кинетики связывания BRD4-ОДН в реальном времени проводили с помощью мини-прибора iMSPR (Icluebio, Южная Корея). Для иммобилизации BRD4 сенсорный чип NTA-Au NiHC1000 (Icluebio, Южная Корея) активировали путем нанесения 20 мкл 0,5 мМ NiCl<sub>2</sub>. Затем 300 мкл 100 нМ раствора BRD4 в буфере (10 мМ HEPES, pH 7,4, 150 мМ NaCl, 50 мкМ ЭДТА и 0,005% ПАВ P20) пропускали через чип со скоростью потока 30 мкл/мин. После этого вводили 1-200 мкМ раствор ОДН или ас-пептида со скоростью потока 30 мкл/мин при 24 °С. Эксперименты проводились в двух повторах для каждой концентрации ОДН/ас-пептида. Для dsG4 дополнительные эксперименты проводили в присутствии 4 экв. лиганда G4 PDS (Sigma-Aldrich, США) или ингибитора BRD4 JQ1 (Abcam, Великобритания). Сенсорные чипы регенерировали путем введения 50 мМ имидазола для удаления с поверхности связанного BRD4 и его комплексов с ОДН/ас-пептидом. Константы ассоциации/диссоциации ( $K_a/K_d$ ) и константа равновесия ( $K_D = K_d/K_a$ ) были рассчитаны с помощью программы iMSPR (Icluebio, Южная Корея).

### **3.3.3 Сборка биоконденсатов и флуоресцентная микроскопия**

#### Подготовка образцов

Полноразмерный BDR4 с полигистидиновой меткой (His) (BRD4-5928h, Creative BioMart, США) был помечен флуоресцентным красителем RED с помощью набора RED-tris-NTA (NanoTemper Technologies, Мюнхен, Германия) и смешан в отношении 0,15:1 с немеченым белком в буфере PBST (pH 7,4) с 10 мМ KCl и 20% ПЭГ-400, до конечной концентрации белка 4 мкМ. Затем добавляли ОДН, ДНК без H3K18ac или ДНК без гистонов в конечных концентрациях 4 мкМ, 150 нМ или 300 нМ, соответственно. Нуклеосомы добавляли к конечной концентрации ДНК 150 нМ и гистонового октамера (НО) 300 нМ. Ингибитор BRD4 (JQ1), G4-лиганды (PDS, SOP1812), добавляли в конечных концентрациях 5 мкМ и 20 мкМ, соответственно, в ДМСО, и сравнивали с образцом в ДМСО без лигандов. Конечная концентрация ДМСО во всех образцах составляла 0,5%. 1,6-гександиол (HD), известный ингибитор LLPS, выбранный в качестве положительного контроля, был добавлен до конечной концентрации 10 %. HD и PDS были получены от Sigma-Aldrich (США), JQ1 был получен от Abcam (Великобритания), а SOP1812 – от MedChemExpress, США.

### Флуоресцентная микроскопия смесей BRD4 с ДНК/гистонами/лигандами

Смеси BRD4 с ДНК/гистонами/лигандами или контрольные образцы инкубировали в течение 15 мин при комнатной температуре, затем помещали в рабочую камеру высотой 0,5 мм между покровным и предметным стеклами и визуализировали с помощью микроскопа Eclipse Ti2 (Nikon, Япония) в канале RED. Все эксперименты проводили в 2-3 биологических и 3 технических повторах. В каждом эксперименте объемную долю окрашенных капель рассчитывали, исходя из общей площади проекции капель и предполагая сферическую форму капель. Общая площадь проекции и средняя окружность капель на каждом изображении определялись с помощью программы Droplet\_Calc [261] с настройками по умолчанию.

### Конфокальная микроскопия и восстановление флуоресценции после фотообесцвечивания (FRAP)

Для проверки колокализации dsG4 с BRD4 при разделении фаз G4-содержащая цепь dsG4 была мечена по 5'-концу остатком 6-карбоксихлорофлуоресцеина (FAM), смешана с немеченой комплементарной цепью, и медленно отжигались в разбавленном NAB (10 мМ Трис-HCl, pH 8, и 125 мМ NaCl) с 20% ПЭГ-400. Затем (FAM)-dsG4 смешивали с (RED)-BRD4 (меченый: немеченый = 0,15:1) в том же буфере до конечных концентраций ДНК и белка 4 мкМ. Образцы готовили в присутствии или в отсутствии 10 мкМ PDS для проверки влияния PDS на разделение фаз. Образцы инкубировали в течение 15 мин, затем наносили на покровное стекло и инкубировали еще 15-30 мин для постепенного оседания конденсата под действием силы тяжести.

Конфокальную микроскопию смесей (FAM)-dsG4/(RED)-BRD4 проводили с помощью микроскопа Olympus FV3000 (Olympus Scientific Solutions, США) в каналах FAM и RED (длины волн лазера: 488 и 640 нм, соответственно). При проведении FRAP-анализа три сегмента конденсата/геля в каждом типе образца (с PDS или без него) выжигались лазером, после чего регистрировалась серия изображений с интервалом 10 с. Экспериментальные кривые восстановления флуоресценции усредняли и аппроксимировали моноэкспоненциальным уравнением.

### **3.3.4 Анализ экспрессии генов**

#### Культуры клеток и анализ цитотоксичности

Клетки линии НЕК-293 культивировали в лаборатории клеточной биологии ФГБУ ФНКЦ ФХМ в CO<sub>2</sub> инкубаторе с влажностью 5%, проверяли на наличие микоплазм и выращивали в 96-луночных планшетах в среде DMEM (Paneco, Россия), дополненной смесью 1%

пенициллина/стрептомицина (Paneco, Россия), 2 мМ глутамин (Paneco, Россия) и 10%-й фетальной бычьей сыворотки (HyClone GE Healthcare). Клетки линии K-562 выращивали в среде RPMI-1640 ("Панэко", Москва, Россия), дополненной 0,4%-м раствором фетальной бычьей сыворотки (HyClone GE Healthcare, США). В экспериментах с лигандами жизнеспособность клеток проверяли через 48 ч инкубации с возрастающими концентрациями PDS, SOP1812 или 5FU с помощью реагента PrestoBlue (ThermoFisher Scientific, США).

#### Выделение РНК и обратная транскрипция

Для анализа регуляции генов к суспензии клеток добавляли PDS и 5FU в колбах T25 (4\*10<sup>5</sup> клеток в 10 мл культуральной среды) до конечных концентраций, равных значениям IC30 (10 мкМ PDS, 1,25 мкМ SOP1812 или 50 мкМ 5FU для HEK-293 и 20 мкМ PDS, 10 мкМ SOP1812 или 100 мкМ 5FU для K-562). После 48 ч инкубации клетки обрабатывали лизирующим буфером. Тотальную РНК (без ДНК) выделяли и очищали с помощью набора HiRure Total RNA Kit (Magen Biotech, Китай) в соответствии с инструкциями производителя. Количество полученной РНК оценивали, а чистоту подтверждали измерением поглощения при 260 нм/280 нм с помощью прибора Tecan NanoQuant (Tecan, Швейцария). Обратную транскрипцию проводили следующим образом: из 5 мкг РНК на образец синтезировали кДНК с помощью обратной транскриптазы Magnus (Evrogen, Россия) с использованием 40 пмоль праймера случайного состава (Evrogen, Россия) в соответствии с инструкциями производителя. Синтез кДНК для всех образцов проводили параллельно.

#### Выбор целевых генов

Основной набор генов, отобранных для проверки SE-зависимой и промотор-зависимой регуляции генов методом ОТ-кПЦР, включал протоонкогены, кодирующие киназы и другие сигнальные белки (*PLK3*, *MKNK2*, *RAPGEFL1*, *IFI6* и *WASF2*), ядерный рецептор *NR1D1*, метилтрансферазу гистонов *DOT1L* и уропорфириногендекарбоксилазу *UROD* (чувствительный к химиотерапии белковый фактор). Все они содержали пики BG4 ChIP-seq в промоторных областях, которые пересекались с SE или совпадали с точками дальних контактов с SE. Контрольные гены, а именно синтаксин-кодирующий ген *SNTX12* и ген *PSMD3*, кодирующий субъединицу протеосомы, также содержали пики BG4 в промоторных областях, но не образовывали дальних контактов с SE или другими регуляторными сайтами с кластерами BRD4 в пределах ТАД. Ген *RPLP0*, кодирующий рибосомальную РНК, выбранный в качестве стандарта, имел промотор без G4. Границы ТАД и контакты внутри ТАД были проанализированы на основе доступных данных Hi-C [245] и визуализированы в геномном браузере Encode 3D

(<http://3dgenome.fsm.northwestern.edu/>). В дополнительный набор генов были включены гены с неявными 3D-контактами (частота контактов с SE ниже среднего), но с хорошо выраженными BG4-пиками в промоторах.

#### ОТ-кПЦР и корреляционный анализ

ОТ-кПЦР проводили совместно с сотрудником лаборатории молекулярной онкологии О.А. Ивановой с использованием qPCRmix-HS SYBR (Evrogen, Россия). Для каждой реакции 1,5 мкл кДНК смешивали с 5 пмоль соответствующих праймеров и 4 мкл qPCRmix-HS SYBR. Затем добавляли воду, свободную от нуклеаз, до общего объема 20 мкл и проводили ПЦР с использованием CFX96 Touch RT PCR Detection System (Bio-Rad, США). Программа амплификации включала 45 циклов (95°C в течение 3 мин, затем 95°C в течение 10 с, 62°C в течение 20 с и 72°C в течение 30 с на цикл). Результаты анализировали методом  $\Delta\Delta C_q$  с использованием *RPLP0* в качестве стандарта. Все эксперименты проводились в 2 биологических и 3 технических повторах.

Изменение экспрессии под действием лиганда, было представлено как  $\text{Log}_2FC$ . Гены с  $p < 0,05$  и  $\text{Log}_2FC > 2$  и  $\text{Log}_2FC < -2$  рассматривались как гены с повысившейся и снизившейся экспрессией, соответственно. После чего оценивали значимость корреляций между регуляцией экспрессии под действием PDS и характеристиками промотора (скор пиков BRD4/BG4/G4-seq). Затем был проведен аналогичный корреляционный анализ SE-характеристик. Скоры пиков BRD4/BG4/G4-seq были получены из соответствующих наборов данных ChIP-seq/G4-seq (GSM3003539, GSE107690 и GSE99178).

## ЗАКЛЮЧЕНИЕ

По итогам исследования были сделаны следующие выводы:

1. Сродство CTCF к геномным G4 сопоставимо со сродством к консенсусному мотиву связывания, что объясняет участие G4 в формировании CTCF-зависимых хроматиновых петель и демаркации транскрипционных доменов.
2. Обработка клеток G4-стабилизирующими лигандами увеличивает частоту позиционирования CTCF на G4, что подтверждает вклад G4 в CTCF-зависимые перестройки хроматина.
3. G4 препятствуют сборке нуклеосомы, что проясняет дополнительные механизмы участия G4 в активации транскрипции за счет освобождения ДНК для связывания с транскрипционными факторами.
4. G4 значимо перепредставлены в суперэнхансерах (SE); G4-богатые SE обеднены консенсусными сайтами связывания BRD4 (асНО), но сродство G4 к BRD4 сопоставимо с таковым для асНО, что обеспечивает удержание BRD4 в SE и сборку биоконденсатов, необходимую для множественных контактов SE с промоторами.
5. G4-лиганды препятствуют взаимодействию BRD4-G4 и ингибируют образование конденсатов, что частично объясняет повышенный отклик генов под контролем G4-богатых SE на G4-лиганды.

Значимость исследования обусловлена тем, что оно создает задел для корректировки стратегий применения G4-лигандов. Результаты, обобщенные в выводах 1-2, проясняют зависимость от белка-инсулятора CTCF реорганизацию хроматина при использовании противоопухолевых G4-лигандов. Из выводов 4-5 следует, что часть суперэнхансеров – перспективных мишеней противоопухолевых ингибиторов BRD4 – нечувствительны к классическим ингибиторам, но чувствительны к G4-лигандам. Отсюда вытекает целесообразность тестирования комбинаций ингибиторов BRD4 с G4-лигандами для подавления аномальной активации транскрипции в онкоклетках. Результаты подкреплены данными В.Б.Цветкова по моделированию влияния лигандов на взаимодействия G4-CTCF и G4-BRD4. Они указывают на конкуренцию между лигандом и BRD4 за G4 и отсутствием конкуренции между лигандом и CTCF. Таким образом, основное действие лигандов на непромоторные G4 – это укрепление инсуляторов и нарушение работы суперэнхансеров. Эти эффекты стоит учитывать при прогнозировании действия лигандов, которые проходят клинические испытания.

## БЛАГОДАРНОСТИ

Автор благодарит за вклад в работу сотрудников лаборатории структуры и функций биополимеров ФГБУ ФНКЦ ФХМ и. М.Ю. Лопухина ФМБА России М. С. Юдина, В. В. Северова, сотрудников лаборатории клеточной биологии А. Н. Богомазову, М. А. Лагарькову, А. В. Сурдину, М. Е. Богомякову, сотрудников лаборатории медицинских нанотехнологий Н. А. Барина, Д. В. Клинова, сотрудников лаборатории молекулярной онкологии О. М. Иванову, В. А. Шендер, сотрудника лаборатории системной биологии Р. И. Султанова, ведущего научного сотрудника Первого МГМУ им. И. М. Сеченова В. Б. Цветкова, сотрудников лаборатории стабильности генома ИБГ РАН, сотрудников лаборатории геномных исследований и вычислительной биологии и других сотрудников центра, сотрудников лаборатории структурной биологии ионных каналов ИБХ РАН С. Д. Орешкова, З. О. Шенкарева, сотрудника лаборатория химии diaзосоединений ИОХ РАН Р. А. Новикова. Особенно благодарю научного руководителя Анну Михайловну Варижук, заведующую лабораторией структуры и функций биополимеров ФГБУ ФНКЦ ФХМ им. М. Ю. Лопухина ФМБА России за инициацию исследования и огромную поддержку на всех его этапах.

В работе активно использовались методы анализа биоконденсатов, разработанные в рамках ГЗ Шаперон 17.002.24.800.

**ЛИТЕРАТУРА**

1. Roy S.S. и др. Artificially inserted strong promoter containing multiple G-quadruplexes induces long-range chromatin modification // *Elife*. 2024. Т. 13.
2. Burge S. и др. Quadruplex DNA: Sequence, topology and structure // *Nucleic Acids Res.* 2006.
3. Maizels N., Gray L.T. The G4 Genome // *PLoS Genet.* 2013. Т. 9, № 4. С. e1003468.
4. Huppert J.L., Balasubramanian S. Prevalence of quadruplexes in the human genome // *Nucleic Acids Res.* 2005. Т. 33, № 9. С. 2908–2916.
5. Kwok C.K., Merrick C.J. G-Quadruplexes: Prediction, Characterization, and Biological Application // *Trends Biotechnol.* 2017. Т. 35, № 10. С. 997–1013.
6. Puig Lombardi E., Londoño-Vallejo A. A guide to computational methods for G-quadruplex prediction // *Nucleic Acids Res.* 2020. Т. 48, № 1. С. 1–15.
7. Korsakova A., Phan A.T. Prediction of G4 formation in live cells with epigenetic data: a deep learning approach // *NAR Genom Bioinform.* 2023. Т. 5, № 3.
8. Marsico G. и др. Whole genome experimental maps of DNA G-quadruplexes in multiple species // *Nucleic Acids Res.* 2019. Т. 47, № 8. С. 3862–3874.
9. Kim N. The Interplay between G-quadruplex and Transcription // *Curr Med Chem.* 2019. Т. 26, № 16. С. 2898–2917.
10. Nadai M., Richter S.N. G-Quadruplex Visualization in Cells via Antibody and Fluorescence Probe. 2019. С. 383–395.
11. Hänsel-Hertsch R. и др. Genome-wide mapping of endogenous G-quadruplex DNA structures by chromatin immunoprecipitation and high-throughput sequencing // *Nat Protoc.* 2018. Т. 13, № 3. С. 551–564.
12. Henderson A. и др. Detection of G-quadruplex DNA in mammalian cells // *Nucleic Acids Res.* 2014. Т. 42, № 2. С. 860–869.
13. Liu H.-Y. и др. Conformation Selective Antibody Enables Genome Profiling and Leads to Discovery of Parallel G-Quadruplex in Human Telomeres // *Cell Chem Biol.* 2016. Т. 23, № 10. С. 1261–1270.
14. Zheng K. и др. Detection of genomic G-quadruplexes in living cells using a small artificial protein // *Nucleic Acids Res.* 2020. Т. 48, № 20. С. 11706–11720.
15. Kaya-Okur H.S. и др. CUT&Tag for efficient epigenomic profiling of small samples and single cells // *Nat Commun.* 2019. Т. 10, № 1. С. 1930.
16. Lyu J. и др. Genome-wide mapping of G-quadruplex structures with CUT&Tag // *Nucleic Acids Res.* 2022. Т. 50, № 3. С. e13–e13.

17. Esnault C. и др. G4access identifies G-quadruplexes and their associations with open chromatin and imprinting control regions // *Nat Genet.* 2023. Т. 55, № 8. С. 1359–1369.
18. Neaves K.J. и др. Direct visualization of G-quadruplexes in DNA using atomic force microscopy // *Nucleic Acids Res.* 2009. Т. 37, № 18. С. 6269–6275.
19. Pirota V. и др. Core-Extended Naphthalene Diimide Dyads as Light-Up Probes with Targeted Cytotoxicity Toward Tumor Cells // *Biomolecules.* 2025. Т. 15, № 2. С. 311.
20. Iida K. и др. Development of a novel light-up probe for detection of G-quadruplexes in stress granules // *Sci Rep.* 2022. Т. 12, № 1. С. 12892.
21. Di Antonio M. и др. Single-molecule visualization of DNA G-quadruplex formation in live cells // *Nat Chem.* 2020. Т. 12, № 9. С. 832–837.
22. Biffi G. и др. Quantitative visualization of DNA G-quadruplex structures in human cells // *Nat Chem.* 2013. Т. 5, № 3. С. 182–186.
23. Rodriguez R. и др. Small-molecule–induced DNA damage identifies alternative DNA structures in human genes // *Nat Chem Biol.* 2012. Т. 8, № 3. С. 301–310.
24. Hänsel-Hertsch R. и др. G-quadruplex structures mark human regulatory chromatin // *Nat Genet.* 2016. Т. 48, № 10. С. 1267–1272.
25. Lago S. и др. Promoter G-quadruplexes and transcription factors cooperate to shape the cell type-specific transcriptome // *Nat Commun.* 2021. Т. 12, № 1. С. 3885.
26. Hou Y. и др. Integrative characterization of G-Quadruplexes in the three-dimensional chromatin structure // *Epigenetics.* 2019. Т. 14, № 9. С. 894–911.
27. Spiegel J., Adhikari S., Balasubramanian S. The Structure and Function of DNA G-Quadruplexes // *Trends in Chemistry.* Cell Press, 2020. Т. 2, № 2. С. 123–136.
28. Robinson J. и др. DNA G-quadruplex structures: more than simple roadblocks to transcription? // *Nucleic Acids Res.* 2021. Т. 49, № 15. С. 8419–8431.
29. Valton A.-L., Prioleau M.-N. G-Quadruplexes in DNA Replication: A Problem or a Necessity? // *Trends in Genetics.* 2016. Т. 32, № 11. С. 697–706.
30. Linke R. и др. The Relevance of G-Quadruplexes for DNA Repair // *Int J Mol Sci.* 2021. Т. 22, № 22. С. 12599.
31. Pavlova A. V. и др. Impact of G-Quadruplexes on the Regulation of Genome Integrity, DNA Damage and Repair // *Biomolecules.* 2021. Т. 11, № 9. С. 1284.
32. Varizhuk A., Isaakova E., Pozmogova G. DNA G-Quadruplexes (G4s) Modulate Epigenetic (Re)Programming and Chromatin Remodeling // *BioEssays.* 2019. Т. 41, № 9.
33. Li G. и др. Integrative genomic analyses of promoter G-quadruplexes reveal their selective constraint and association with gene activation // *Commun Biol.* 2023. Т. 6, № 1. С. 625.

34. Lam E.Y.N. и др. G-quadruplex structures are stable and detectable in human genomic DNA // *Nat Commun.* 2013. Т. 4, № 1. С. 1796.
35. Spiegel J. и др. G-quadruplexes are transcription factor binding hubs in human chromatin // *Genome Biol.* 2021. Т. 22, № 1. С. 117.
36. Du Z., Zhao Y., Li N. Genome-wide analysis reveals regulatory role of G4 DNA in gene transcription // *Genome Res.* 2008. Т. 18, № 2. С. 233–241.
37. Cogoi S. и др. The KRAS Promoter Responds to Myc-associated Zinc Finger and Poly(ADP-ribose) Polymerase 1 Proteins, Which Recognize a Critical Quadruplex-forming GA-element // *Journal of Biological Chemistry.* 2010. Т. 285, № 29. С. 22003–22016.
38. Kerkour A. и др. High-resolution three-dimensional NMR structure of the KRAS proto-oncogene promoter reveals key features of a G-quadruplex involved in transcriptional regulation // *Journal of Biological Chemistry.* 2017. Т. 292, № 19. С. 8082–8091.
39. Gao J. и др. Yeast transcription co-activator Sub1 and its human homolog PC4 preferentially bind to G-quadruplex DNA // *Chemical Communications.* 2015. Т. 51, № 33. С. 7242–7244.
40. Hanakahi L.A., Sun H., Maizels N. High affinity interactions of nucleolin with G-G-paired rDNA // *Journal of Biological Chemistry.* 1999.
41. González V., Hurley L.H. The C-Terminus of Nucleolin Promotes the Formation of the c-*MYC* G-Quadruplex and Inhibits c-*MYC* Promoter Activity // *Biochemistry.* 2010. Т. 49, № 45. С. 9706–9714.
42. Lee C.-Y. и др. R-loop induced G-quadruplex in non-template promotes transcription by successive R-loop formation // *Nat Commun.* 2020. Т. 11, № 1. С. 3392.
43. Duquette M.L. и др. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA // *Genes Dev.* 2004. Т. 18, № 13. С. 1618–1629.
44. Belotserkovskii B.P. и др. Mechanisms and implications of transcription blockage by guanine-rich DNA sequences // *Proceedings of the National Academy of Sciences.* 2010. Т. 107, № 29. С. 12816–12821.
45. Mendoza O. и др. G-quadruplexes and helicases // *Nucleic Acids Res.* 2016. Т. 44, № 5. С. 1989–2006.
46. Lerner L.K., Sale J.E. Replication of G Quadruplex DNA // *Genes (Basel).* 2019. Т. 10, № 2. С. 95.
47. Besnard E. и др. Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs // *Nat Struct Mol Biol.* 2012. Т. 19, № 8. С. 837–844.
48. Fry M. Tetraplex DNA and its interacting proteins // *Frontiers in Bioscience.* 2007. Т. 12, № 8–12. С. 4336.

49. Hershman S.G. и др. Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in *Saccharomyces cerevisiae* // *Nucleic Acids Res.* 2008. Т. 36, № 1. С. 144–156.
50. Du R. и др. Targeting AURKA in Cancer: molecular mechanisms and opportunities for Cancer therapy // *Mol Cancer.* 2021. Т. 20, № 1. С. 15.
51. Robinson J. и др. G-quadruplex structures regulate long-range transcriptional reprogramming to promote drug resistance in ovarian cancer. 2024.
52. Bochman M.L., Paeschke K., Zakian V.A. DNA secondary structures: stability and function of G-quadruplex structures // *Nat Rev Genet.* 2012. Т. 13, № 11. С. 770–780.
53. Rhodes D., Lipps H.J. G-quadruplexes and their regulatory roles in biology // *Nucleic Acids Res.* 2015. Т. 43, № 18. С. 8627–8637.
54. Varshney D. и др. The regulation and functions of DNA and RNA G-quadruplexes // *Nat Rev Mol Cell Biol.* 2020. Т. 21, № 8. С. 459–474.
55. Mukherjee A.K., Sharma S., Chowdhury S. Non-duplex G-Quadruplex Structures Emerge as Mediators of Epigenetic Modifications // *Trends in Genetics.* 2019. Т. 35, № 2. С. 129–144.
56. Mao S.Q. и др. DNA G-quadruplex structures mold the DNA methylome // *Nat Struct Mol Biol.* Nature Publishing Group, 2018. Т. 25, № 10. С. 951–957.
57. Li L. и др. YY1 interacts with guanine quadruplexes to regulate DNA looping and gene expression // *Nat Chem Biol.* 2021. Т. 17, № 2. С. 161–168.
58. Ohlsson R., Renkawitz R., Lobanenko V. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease // *Trends in Genetics.* 2001. Т. 17, № 9. С. 520–527.
59. Kim T.H. и др. Analysis of the Vertebrate Insulator Protein CTCF-Binding Sites in the Human Genome // *Cell.* Elsevier B.V., 2007. Т. 128, № 6. С. 1231–1245.
60. Persikov A. V., Singh M. De novo prediction of DNA-binding specificities for Cys2His2 zinc finger proteins // *Nucleic Acids Res.* *Nucleic Acids Res.*, 2014. Т. 42, № 1. С. 97–108.
61. Hashimoto H. и др. Structural Basis for the Versatile and Methylation-Dependent Binding of CTCF to DNA // *Mol Cell.* Cell Press, 2017. Т. 66, № 5. С. 711-720.e3.
62. Rasim Barutcu A. и др. The connection between BRG1, CTCF and topoisomerases at TAD boundaries // *Nucleus.* Taylor and Francis Inc., 2017. Т. 8, № 2. С. 150–155.
63. Barisic D. и др. Mammalian ISWI and SWI/SNF selectively mediate binding of distinct transcription factors // *Nature.* Nature Publishing Group, 2019. Т. 569, № 7754. С. 136–140.
64. Owens N. и др. CTCF confers local nucleosome resiliency after dna replication and during mitosis // *Elife.* eLife Sciences Publications Ltd, 2019. Т. 8.

65. Makowski M.M. и др. Global profiling of protein–DNA and protein–nucleosome binding affinities using quantitative mass spectrometry // *Nat Commun.* 2018. Т. 9, № 1. С. 1653.
66. Vlasenok M. и др. Data set on G4 DNA interactions with human proteins // *Data Brief.* 2018.
67. Halder K., Halder R., Chowdhury S. Genome-wide analysis predicts DNA structural motifs as nucleosome exclusion signals // *Mol Biosyst.* The Royal Society of Chemistry, 2009. Т. 5, № 12. С. 1703–1712.
68. Foulk M.S. и др. Characterizing and controlling intrinsic biases of lambda exonuclease in nascent strand sequencing reveals phasing between nucleosomes and G-quadruplex motifs around a subset of human replication origins // *Genome Res.* 2015. Т. 25, № 5. С. 725–735.
69. Tikhonova P. и др. DNA G-Quadruplexes Contribute to CTCF Recruitment // *Int J Mol Sci.* 2021. Т. 22, № 13. С. 7090.
70. Schiavone D. и др. Determinants of G quadruplex-induced epigenetic instability in *REV* 1-deficient cells // *EMBO J.* 2014. Т. 33, № 21. С. 2507–2520.
71. Mimura M. и др. Quadruplex Folding Promotes the Condensation of Linker Histones and DNAs via Liquid–Liquid Phase Separation // *J Am Chem Soc.* 2021. Т. 143, № 26. С. 9849–9857.
72. Hou Y. и др. Novel Roles of G-quadruplexes on Enhancers in human chromatin // *bioRxiv.* Cold Spring Harbor Laboratory, 2021.
73. Yuan J., He X., Wang Y. G-quadruplex DNA contributes to RNA polymerase II-mediated 3D chromatin architecture // *Nucleic Acids Res.* 2023. Т. 51, № 16. С. 8434–8446.
74. Liano D. и др. Long-range DNA interactions: inter-molecular G-quadruplexes and their potential biological relevance // *Chem Commun (Camb).* Royal Society of Chemistry, 2022. Т. 58, № 92. С. 12753.
75. Liano D., Chowdhury S., Di Antonio M. Cockayne Syndrome B Protein Selectively Resolves and Interact with Intermolecular DNA G-Quadruplex Structures // *J Am Chem Soc.* 2021. Т. 143, № 49. С. 20988–21002.
76. Wulfridge P. и др. G-quadruplexes associated with R-loops promote CTCF binding // *Mol Cell.* 2023. Т. 83, № 17. С. 3064-3079.e5.
77. Davidson I.F. и др. CTCF is a DNA-tension-dependent barrier to cohesin-mediated loop extrusion // *Nature.* 2023. Т. 616, № 7958. С. 822–827.
78. Weintraub A.S. и др. YY1 Is a Structural Regulator of Enhancer-Promoter Loops // *Cell.* 2017. Т. 171, № 7. С. 1573-1588.e28.
79. Hsieh T.-H.S. и др. Enhancer–promoter interactions and transcription are largely maintained upon acute loss of CTCF, cohesin, WAPL or YY1 // *Nat Genet.* 2022. Т. 54, № 12. С. 1919–1932.

80. Richter W.F. и др. The Mediator complex as a master regulator of transcription by RNA polymerase II // *Nat Rev Mol Cell Biol.* 2022. Т. 23, № 11. С. 732–749.
81. Cho W.-К. и др. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates // *Science* (1979). 2018. Т. 361, № 6400. С. 412–415.
82. Lambert É. и др. From structure to molecular condensates: emerging mechanisms for Mediator function // *FEBS J.* 2023. Т. 290, № 2. С. 286–309.
83. Sabari B.R. и др. Coactivator condensation at super-enhancers links phase separation and gene control // *Science* (1979). 2018. Т. 361, № 6400.
84. Donati B., Lorenzini E., Ciarrocchi A. BRD4 and Cancer: going beyond transcriptional regulation // *Mol Cancer.* 2018. Т. 17, № 1. С. 164.
85. Filippakopoulos P. и др. Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family // *Cell.* 2012. Т. 149, № 1. С. 214–231.
86. Miller T.C.R. и др. A bromodomain–DNA interaction facilitates acetylation-dependent bivalent nucleosome recognition by the BET protein BRDT // *Nat Commun.* 2016. Т. 7, № 1. С. 13855.
87. Olp M.D., Jackson V., Smith B.C. Nucleosome scaffolding by Brd4 tandem bromodomains in acetylation-dependent chromatin compartmentalization // *bioRxiv.* Cold Spring Harbor Laboratory, 2019.
88. Wu T., Kamikawa Y.F., Donohoe M.E. Brd4's Bromodomains Mediate Histone H3 Acetylation and Chromatin Remodeling in Pluripotent Cells through P300 and Brg1 // *Cell Rep.* 2018. Т. 25, № 7. С. 1756–1771.
89. Devaiah B.N. и др. BRD4 is a histone acetyltransferase that evicts nucleosomes from chromatin // *Nat Struct Mol Biol.* 2016. Т. 23, № 6. С. 540–548.
90. Pavlova I. и др. Modeling G4s in chromatin context confirms partial nucleosome exclusion and reveals nucleosome-disrupting effects of the least selective G4 ligands // *Biochimie.* 2023. Т. 204. С. 8–21.
91. Esnault C. и др. G-quadruplexes are promoter elements controlling nucleosome exclusion and RNA polymerase II pausing // *bioRxiv.* Cold Spring Harbor Laboratory, 2023.
92. Pavlova I.I. и др. Transcription-facilitating histone chaperons interact with genomic and synthetic G4 structures // *Int J Biol Macromol.* 2020. Т. 160. С. 1144–1157.
93. Pavlova I.I. и др. G-quadruplexes as potential traps for superenhancer marker BRD4: ligand-sensitive binding and co-separation *in vitro* // *Nucleic Acids Res.* 2025. Т. 53, № 14.
94. Mao Y.S., Zhang B., Spector D.L. Biogenesis and function of nuclear bodies // *Trends in Genetics.* 2011. Т. 27, № 8. С. 295–306.

95. Rippe K. Liquid–Liquid Phase Separation in Chromatin // *Cold Spring Harb Perspect Biol.* 2022. T. 14, № 2. C. a040683.
96. Banani S.F. и др. Biomolecular condensates: organizers of cellular biochemistry // *Nat Rev Mol Cell Biol.* 2017. T. 18, № 5. C. 285–298.
97. Shin Y., Brangwynne C.P. Liquid phase condensation in cell physiology and disease // *Science (1979).* 2017. T. 357, № 6357.
98. Bergeron-Sandoval L.-P., Safae N., Michnick S.W. Mechanisms and Consequences of Macromolecular Phase Separation // *Cell.* 2016. T. 165, № 5. C. 1067–1079.
99. Lee J., Cho H., Kwon I. Phase separation of low-complexity domains in cellular function and disease // *Exp Mol Med.* 2022. T. 54, № 9. C. 1412–1422.
100. Dignon G.L., Best R.B., Mittal J. Biomolecular Phase Separation: From Molecular Driving Forces to Macroscopic Properties // *Annu Rev Phys Chem.* 2020. T. 71, № 1. C. 53–75.
101. Borchers W. и др. How do intrinsically disordered protein regions encode a driving force for liquid–liquid phase separation? // *Curr Opin Struct Biol.* 2021. T. 67. C. 41–50.
102. O’Flynn B.G., Mittag T. The role of liquid–liquid phase separation in regulating enzyme activity // *Curr Opin Cell Biol.* 2021. T. 69. C. 70–79.
103. Frottin F. и др. The nucleolus functions as a phase-separated protein quality control compartment // *Science (1979).* 2019. T. 365, № 6451. C. 342–347.
104. Zhang Y. и др. G-quadruplex structures trigger RNA phase separation // *Nucleic Acids Res.* 2019.
105. Kharel P. и др. Properties and biological impact of RNA G-quadruplexes: from order to turmoil and back // *Nucleic Acids Res.* 2020. T. 48, № 22. C. 12534–12555.
106. Ishiguro A., Ishihama A. Essential Roles and Risks of G-Quadruplex Regulation: Recognition Targets of ALS-Linked TDP-43 and FUS // *Front Mol Biosci.* 2022. T. 9.
107. Wang E. и др. G-Quadruplexes as pathogenic drivers in neurodegenerative disorders // *Nucleic Acids Res.* 2021. T. 49, № 9. C. 4816–4830.
108. Teng Y., Zhu M., Qiu Z. G-Quadruplexes in Repeat Expansion Disorders // *Int J Mol Sci.* 2023. T. 24, № 3. C. 2375.
109. Brázda V. и др. DNA and RNA Quadruplex-Binding Proteins // *Int J Mol Sci.* 2014. T. 15, № 10. C. 17493–17517.
110. Meier-Stephenson V. G4-quadruplex-binding proteins: review and insights into selectivity // *Biophys Rev.* 2022. T. 14, № 3. C. 635–654.
111. Sanchez-Martin V. DNA G-Quadruplex-Binding Proteins: An Updated Overview // *DNA.* 2023. T. 3, № 1. C. 1–12.

112. Mishra S.K. и др. G4IPDB: A database for G-quadruplex structure forming nucleic acid interacting proteins // *Sci Rep*. 2016. Т. 6, № 1. С. 38144.
113. Zhang X. и др. Chemical profiling of DNA G-quadruplex-interacting proteins in live cells // *Nat Chem*. 2021. Т. 13, № 7. С. 626–633.
114. Su H. и др. Photoactive G-Quadruplex Ligand Identifies Multiple G-Quadruplex-Related Proteins with Extensive Sequence Tolerance in the Cellular Environment // *J Am Chem Soc*. 2021. Т. 143, № 4. С. 1917–1923.
115. Wu G. и др. DDX5 helicase resolves G-quadruplex and is involved in *MYC* gene transcriptional activation // *Proceedings of the National Academy of Sciences*. 2019. Т. 116, № 41. С. 20453–20461.
116. Gao J. и др. G-quadruplex DNA inhibits unwinding activity but promotes liquid–liquid phase separation by the DEAD-box helicase Ded1p // *Chemical Communications*. 2021. Т. 57, № 60. С. 7445–7448.
117. Song Q.-X. и др. DEAD-box RNA helicase Dbp2 binds to G-quadruplex nucleic acids and regulates different conformation of G-quadruplex DNA // *Biochem Biophys Res Commun*. 2022. Т. 634. С. 182–188.
118. Ali M.A.M. DEAD-box RNA helicases: The driving forces behind RNA metabolism at the crossroad of viral replication and antiviral innate immunity // *Virus Res*. 2021. Т. 296. С. 198352.
119. Taschuk F., Cherry S. DEAD-Box Helicases: Sensors, Regulators, and Effectors for Antiviral Defense // *Viruses*. 2020. Т. 12, № 2. С. 181.
120. Overwijn D., Hondele M. DEAD-box ATPases as regulators of biomolecular condensates and membrane-less organelles // *Trends Biochem Sci*. 2023. Т. 48, № 3. С. 244–258.
121. Hondele M. и др. DEAD-box ATPases are global regulators of phase-separated organelles // *Nature*. 2019. Т. 573, № 7772. С. 144–148.
122. Liu X. и др. G-Quadruplex-Induced Liquid–Liquid Phase Separation in Biomimetic Protocells // *J Am Chem Soc*. 2021. Т. 143, № 29. С. 11036–11043.
123. Gallo A. и др. Structure of Nucleophosmin DNA-binding Domain and Analysis of Its Complex with a G-quadruplex Sequence from the c-MYC Promoter // *Journal of Biological Chemistry*. 2012. Т. 287, № 32. С. 26539–26548.
124. Dempsey L.A. и др. G4 DNA Binding by LR1 and Its Subunits, Nucleolin and hnRNP D, A Role for G-G pairing in Immunoglobulin Switch Recombination \* // *Journal of Biological Chemistry*. Elsevier, 1999. Т. 274, № 2. С. 1066–1071.

125. Ghosh M., Singh M. RGG-box in hnRNPA1 specifically recognizes the telomere G-quadruplex DNA and enhances the G-quadruplex unfolding ability of UP1 domain // *Nucleic Acids Res.* 2018. T. 46, № 19. C. 10246–10261.
126. Paramasivam M. и др. Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human KRAS promoter: implications for transcription // *Nucleic Acids Res.* 2009. T. 37, № 9. C. 2841–2853.
127. Yagi R., Miyazaki T., Oyoshi T. G-quadruplex binding ability of TLS/FUS depends on the  $\beta$ -spiral structure of the RGG domain // *Nucleic Acids Res.* 2018. T. 46, № 12. C. 5894–5901.
128. Takahama K. и др. G-Quadruplex DNA- and RNA-Specific-Binding Proteins Engineered from the RGG Domain of TLS/FUS // *ACS Chem Biol.* 2015. T. 10, № 11. C. 2564–2569.
129. Martini S. и др. A genetically-encoded crosslinker screen identifies SERBP1 as a PKC $\epsilon$  substrate influencing translation and cell division // *Nat Commun.* 2021. T. 12, № 1. C. 6934.
130. Baudin A. и др. Structural Characterization of the RNA-Binding Protein SERBP1 Reveals Intrinsic Disorder and Atypical RNA Binding Modes // *Front Mol Biosci.* 2021. T. 8.
131. Mitrea D.M. и др. Self-interaction of NPM1 modulates multiple mechanisms of liquid–liquid phase separation // *Nat Commun.* 2018. T. 9, № 1. C. 842.
132. Arcovito A. и др. Synergic Role of Nucleophosmin Three-helix Bundle and a Flanking Unstructured Tail in the Interaction with G-quadruplex DNA // *Journal of Biological Chemistry.* 2014. T. 289, № 31. C. 21230–21241.
133. Mitrea D.M. и др. Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA // *Elife.* 2016. T. 5.
134. Nishikawa T. и др. HnRNPA1 interacts with G-quadruplex in the TRA2B promoter and stimulates its transcription in human colon cancer cells // *Sci Rep.* 2019. T. 9, № 1. C. 10276.
135. Molliex A. и др. Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization // *Cell.* 2015. T. 163, № 1. C. 123–133.
136. Gui X. и др. Structural basis for reversible amyloids of hnRNPA1 elucidates their role in stress granule assembly // *Nat Commun.* 2019. T. 10, № 1. C. 2006.
137. Tonello F., Massimino M.L., Peggion C. Nucleolin: a cell portal for viruses, bacteria, and toxins // *Cellular and Molecular Life Sciences.* 2022. T. 79, № 5. C. 271.
138. Serin G. и др. Two RNA-binding Domains Determine the RNA-binding Specificity of Nucleolin \* // *Journal of Biological Chemistry.* Elsevier, 1997. T. 272, № 20. C. 13109–13116.
139. Masuzawa T., Oyoshi T. Roles of the RGG Domain and RNA Recognition Motif of Nucleolin in G-Quadruplex Stabilization // *ACS Omega.* American Chemical Society, 2020. T. 5, № 10. C. 5202–5208.

140. Ji Y., Li F., Qiao Y. Modulating liquid–liquid phase separation of FUS: mechanisms and strategies // *J Mater Chem B*. 2022. Т. 10, № 42. С. 8616–8628.
141. Murthy A.C. и др. Molecular interactions contributing to FUS SYGQ LC-RGG phase separation and co-partitioning with RNA polymerase II heptads // *Nat Struct Mol Biol*. 2021. Т. 28, № 11. С. 923–935.
142. Murray D.T. и др. Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains // *Cell*. 2017. Т. 171, № 3. С. 615-627.e16.
143. Qamar S. и др. FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation- $\pi$  Interactions // *Cell*. 2018. Т. 173, № 3. С. 720-734.e15.
144. Monahan Z. и др. Phosphorylation of the <sc>FUS</sc> low-complexity domain disrupts phase separation, aggregation, and toxicity // *EMBO J*. 2017. Т. 36, № 20. С. 2951–2967.
145. Hofweber M. и др. Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation // *Cell*. 2018. Т. 173, № 3. С. 706-719.e13.
146. Singatulina A.S. и др. PARP-1 Activation Directs FUS to DNA Damage Sites to Form PARG-Reversible Compartments Enriched in Damaged DNA // *Cell Rep*. 2019. Т. 27, № 6. С. 1809-1821.e5.
147. Vernon R.M. и др. Pi-Pi contacts are an overlooked protein feature relevant to phase separation // *Elife* / под ред. Shan Y. eLife Sciences Publications, Ltd, 2018. Т. 7. С. e31486.
148. Hong Y. и др. Hydrophobicity of arginine leads to reentrant liquid-liquid phase separation behaviors of arginine-rich proteins // *Nat Commun*. 2022. Т. 13, № 1. С. 7326.
149. Simko E.A.J. и др. G-quadruplexes offer a conserved structural motif for NONO recruitment to NEAT1 architectural lncRNA // *Nucleic Acids Res*. 2020.
150. Mou X., Liew S.W., Kwok C.K. Identification and targeting of G-quadruplex structures in *MALAT1* long non-coding RNA // *Nucleic Acids Res*. 2022. Т. 50, № 1. С. 397–410.
151. Mestre-Fos S. и др. G-Quadruplexes in Human Ribosomal RNA // *J Mol Biol*. 2019. Т. 431, № 10. С. 1940–1955.
152. Galganski L., Urbanek M.O., Krzyzosiak W.J. Nuclear speckles: molecular organization, biological function and role in disease // *Nucleic Acids Res*. 2017. Т. 45, № 18. С. 10350–10368.
153. Bond C.S., Fox A.H. Paraspeckles: nuclear bodies built on long noncoding RNA // *Journal of Cell Biology*. 2009. Т. 186, № 5. С. 637–644.
154. Bhatt P.R. и др. Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome // *Science* (1979). American Association for the Advancement of Science, 2021. Т. 372, № 6548. С. 1306–1313.

155. Editorial Expression of Concern on article ‘Alternative splicing modulation mediated by G-quadruplex structures in MALAT1 lncRNA’ // *Nucleic Acids Res.* 2022. Т. 50, № 7. С. 4199–4199.
156. Lindström M.S. и др. Nucleolus as an emerging hub in maintenance of genome stability and cancer pathogenesis // *Oncogene.* 2018. Т. 37, № 18. С. 2351–2366.
157. Cong R. и др. Interaction of nucleolin with ribosomal RNA genes and its role in RNA polymerase I transcription // *Nucleic Acids Res.* 2012.
158. Chiarella S. и др. Nucleophosmin mutations alter its nucleolar localization by impairing G-quadruplex binding at ribosomal DNA // *Nucleic Acids Res.* 2013. Т. 41, № 5. С. 3228–3239.
159. Guillen-Chable F. и др. Phase Separation of Intrinsically Disordered Nucleolar Proteins Relate to Localization and Function // *Int J Mol Sci.* 2021. Т. 22, № 23. С. 13095.
160. Zhang S. и др. Direct visualization of nucleolar G-quadruplexes in live cells by using a fluorescent light-up probe // *Biochimica et Biophysica Acta (BBA) - General Subjects.* 2018. Т. 1862, № 5. С. 1101–1106.
161. Zhang S. и др. Real-time monitoring of DNA G-quadruplexes in living cells with a small-molecule fluorescent probe // *Nucleic Acids Res.* 2018. Т. 46, № 15. С. 7522–7532.
162. Hoffmann R.F. и др. Guanine quadruplex structures localize to heterochromatin // *Nucleic Acids Res.* 2016. Т. 44, № 1. С. 152–163.
163. Roach R.J. и др. Heterochromatin protein 1 $\alpha$  interacts with parallel RNA and DNA G-quadruplexes // *Nucleic Acids Res.* 2020. Т. 48, № 2. С. 682–693.
164. Hanna R. и др. G-quadruplexes originating from evolutionary conserved L1 elements interfere with neuronal gene expression in Alzheimer’s disease // *Nat Commun.* 2021. Т. 12, № 1. С. 1828.
165. Larson A.G. и др. Liquid droplet formation by HP1 $\alpha$  suggests a role for phase separation in heterochromatin // *Nature.* 2017. Т. 547, № 7662. С. 236–240.
166. Shakya A. и др. Liquid-Liquid Phase Separation of Histone Proteins in Cells: Role in Chromatin Organization // *Biophys J.* 2020. Т. 118, № 3. С. 753–764.
167. Erdel F., Rippe K. Formation of Chromatin Subcompartments by Phase Separation // *Biophys J.* 2018. Т. 114, № 10. С. 2262–2270.
168. Sridhar A. и др. Emergence of chromatin hierarchical loops from protein disorder and nucleosome asymmetry // *Proceedings of the National Academy of Sciences.* 2020. Т. 117, № 13. С. 7216–7224.
169. Farr S.E. и др. Nucleosome plasticity is a critical element of chromatin liquid–liquid phase separation and multivalent nucleosome interactions // *Nat Commun.* 2021. Т. 12, № 1. С. 2883.

170. Her C. и др. Molecular interactions underlying the phase separation of HP1 $\alpha$ : role of phosphorylation, ligand and nucleic acid binding // *Nucleic Acids Res.* 2022. Т. 50, № 22. С. 12702–12722.
171. Wang J. и др. The proper connection between shelterin components is required for telomeric heterochromatin assembly // *Genes Dev.* 2016. Т. 30, № 7. С. 827–839.
172. Jack A. и др. Compartmentalization of telomeres through DNA-scaffolded phase separation // *Dev Cell.* 2022. Т. 57, № 2. С. 277-290.e9.
173. Mukherjee A.K. и др. Telomere repeat-binding factor 2 binds extensively to extra-telomeric G-quadruplexes and regulates the epigenetic status of several gene promoters // *Journal of Biological Chemistry.* 2019. Т. 294, № 47. С. 17709–17722.
174. Rodriguez R. и др. A Novel Small Molecule That Alters Shelterin Integrity and Triggers a DNA-Damage Response at Telomeres // *J Am Chem Soc.* 2008. Т. 130, № 47. С. 15758–15759.
175. Biffi G., Tannahill D., Balasubramanian S. An Intramolecular G-Quadruplex Structure Is Required for Binding of Telomeric Repeat-Containing RNA to the Telomeric Protein TRF2 // *J Am Chem Soc.* 2012. Т. 134, № 29. С. 11974–11976.
176. Mei Y. и др. TERRA G-quadruplex RNA interaction with TRF2 GAR domain is required for telomere integrity // *Sci Rep.* 2021. Т. 11, № 1. С. 3509.
177. Komůrková D., Svobodová Kovaříková A., Bártová E. G-Quadruplex Structures Colocalize with Transcription Factories and Nuclear Speckles Surrounded by Acetylated and Dimethylated Histones H3 // *Int J Mol Sci.* 2021. Т. 22, № 4. С. 1995.
178. Cook P.R., Marenduzzo D. Transcription-driven genome organization: a model for chromosome structure and the regulation of gene expression tested through simulations // *Nucleic Acids Res.* 2018. Т. 46, № 19. С. 9895–9906.
179. Ishov A.M., Gurumurthy A., Bungert J. Coordination of transcription, processing, and export of highly expressed RNAs by distinct biomolecular condensates // *Emerg Top Life Sci.* 2020. Т. 4, № 3. С. 281–291.
180. Chen Y., Belmont A.S. Genome organization around nuclear speckles // *Curr Opin Genet Dev.* 2019. Т. 55. С. 91–99.
181. Raiber E.-A. и др. A non-canonical DNA structure is a binding motif for the transcription factor SP1 in vitro // *Nucleic Acids Res.* 2012. Т. 40, № 4. С. 1499–1508.
182. Pavlova I.I. и др. Transcription-facilitating histone chaperons interact with genomic and synthetic G4 structures // *Int J Biol Macromol.* 2020. Т. 160.
183. Cheung K.L., Kim C., Zhou M.-M. The Functions of BET Proteins in Gene Transcription of Biology and Diseases // *Front Mol Biosci.* 2021. Т. 8.

184. Cheung K.L. и др. Distinct Roles of Brd2 and Brd4 in Potentiating the Transcriptional Program for Th17 Cell Differentiation // *Mol Cell*. 2017. Т. 65, № 6. С. 1068-1080.e5.
185. Monfils K., Barakat T.S. Models behind the mystery of establishing enhancer-promoter interactions // *Eur J Cell Biol*. 2021. Т. 100, № 5–6. С. 151170.
186. Williams J.D. и др. Characterization of long G4-rich enhancer-associated genomic regions engaging in a novel loop:loop ‘G4 Kissing’ interaction // *Nucleic Acids Res*. 2020. Т. 48, № 11. С. 5907–5925.
187. Ong C.-T., Corces V.G. CTCF: an architectural protein bridging genome topology and function // *Nat Rev Genet*. 2014. Т. 15, № 4. С. 234–246.
188. Xiao T., Li X., Felsenfeld G. The Myc-associated zinc finger protein (MAZ) works together with CTCF to control cohesin positioning and genome organization // *Proceedings of the National Academy of Sciences*. 2021. Т. 118, № 7.
189. Prorok P. и др. Involvement of G-quadruplex regions in mammalian replication origin activity // *Nat Commun*. 2019. Т. 10, № 1. С. 3274.
190. Nathanailidou P., Taraviras S., Lygerou Z. DNA Replication Control: Liquid-liquid Phase Separation Comes Into Play. // *J Mol Biochem*. 2020. Т. 9, № 1. С. 54–56.
191. Parker M.W. и др. A new class of disordered elements controls DNA replication through initiator self-assembly // *Elife*. 2019. Т. 8.
192. Saito I. и др. Photoinduced DNA Cleavage via Electron Transfer: Demonstration That Guanine Residues Located 5’ to Guanine Are the Most Electron-Donating Sites // *J Am Chem Soc*. 1995. Т. 117, № 23. С. 6406–6407.
193. Fleming A.M., Burrows C.J. Interplay of Guanine Oxidation and G-Quadruplex Folding in Gene Promoters // *J Am Chem Soc*. 2020. Т. 142, № 3. С. 1115–1136.
194. Cogoi S. и др. Structural polymorphism within a regulatory element of the human KRAS promoter: formation of G4-DNA recognized by nuclear proteins // *Nucleic Acids Res*. 2008. Т. 36, № 11. С. 3765–3780.
195. Edwards A.D. и др. G-Quadruplex loops regulate PARP-1 enzymatic activation // *Nucleic Acids Res*. 2021. Т. 49, № 1. С. 416–431.
196. Alemasova E.E., Lavrik O.I. A sePARate phase? Poly(ADP-ribose) versus RNA in the organization of biomolecular condensates // *Nucleic Acids Res*. 2022. Т. 50, № 19. С. 10817–10838.
197. Sugimura K. и др. PARP-1 ensures regulation of replication fork progression by homologous recombination on damaged DNA // *J Cell Biol*. 2008. Т. 183, № 7. С. 1203–1212.
198. Fu H. и др. Poly(ADP-ribosylation) of P-TEFb by PARP1 disrupts phase separation to inhibit global transcription after DNA damage // *Nat Cell Biol*. 2022. Т. 24, № 4. С. 513–525.

199. Rhine K. и др. Poly(ADP-ribose) drives condensation of FUS via a transient interaction // *Mol Cell*. 2022. Т. 82, № 5. С. 969-985.e11.
200. Li P.T. и др. Expression of the human telomerase reverse transcriptase gene is modulated by quadruplex formation in its first exon due to DNA methylation // *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology Inc., 2017. Т. 292, № 51. С. 20859–20870.
201. Rodriguez R. и др. A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres // *J Am Chem Soc*. 2008. Т. 130, № 47. С. 15758–15759.
202. De Cian A. и др. Highly efficient G-quadruplex recognition by bisquinolinium compounds // *J Am Chem Soc*. 2007. Т. 129, № 7. С. 1856–1857.
203. Müller S. и др. Pyridostatin analogues promote telomere dysfunction and long-term growth inhibition in human cancer cells // *Org Biomol Chem*. Royal Society of Chemistry, 2012. Т. 10, № 32. С. 6537–6546.
204. Phan A.T., Modi Y.S., Patel D.J. Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter // *J Am Chem Soc*. 2004. Т. 126, № 28. С. 8710–8716.
205. Wright E.P., Huppert J.L., Waller Z.A.E. Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH // *Nucleic Acids Res*. Oxford University Press, 2017. Т. 45, № 6. С. 2951–2959.
206. Lowary P.T., Widom J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning // *J Mol Biol*. 1998. Т. 276, № 1. С. 19–42.
207. Vasudevan D., Chua E.Y.D., Davey C.A. Crystal Structures of Nucleosome Core Particles Containing the ‘601’ Strong Positioning Sequence // *J Mol Biol*. 2010. Т. 403, № 1. С. 1–10.
208. Varizhuk A.M. и др. Polymorphism of G4 associates: From stacks to wires via interlocks // *Nucleic Acids Res*. 2018.
209. Bhattacharyya D., Mirihana Arachchilage G., Basu S. Metal Cations in G-Quadruplex Folding and Stability // *Front Chem*. 2016. Т. 4.
210. Renaud de la Faverie A. и др. Thioflavin T as a fluorescence light-up probe for G4 formation // *Nucleic Acids Res*. 2014. Т. 42, № 8. С. e65–e65.
211. Bhattacharya P.K., Cha J., Barton J.K. <sup>1</sup>H NMR determination of base-pair lifetimes in oligonucleotides containing single base mismatches // *Nucleic Acids Res*. 2002. Т. 30, № 21. С. 4740–4750.
212. Adrian M., Heddi B., Phan A.T. NMR spectroscopy of G-quadruplexes // *Methods*. 2012. Т. 57, № 1. С. 11–24.

213. Abou Assi H. и др. 2'-Fluoroarabinonucleic acid modification traps G-quadruplex and i-motif structures in human telomeric DNA // *Nucleic Acids Res.* 2017. Т. 45, № 20. С. 11535–11546.
214. Steger D.J., Workman J.L. Transcriptional Analysis of Purified Histone Acetyltransferase Complexes // *Methods.* 1999. Т. 19, № 3. С. 410–416.
215. Read M. и др. Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors // *Proceedings of the National Academy of Sciences.* 2001. Т. 98, № 9. С. 4844–4849.
216. Han F.X., Wheelhouse R.T., Hurley L.H. Interactions of TMPyP4 and TMPyP2 with Quadruplex DNA. Structural Basis for the Differential Effects on Telomerase Inhibition // *J Am Chem Soc.* 1999. Т. 121, № 15. С. 3561–3570.
217. Lejault P. и др. How to untie G-quadruplex knots and why? // *Cell Chem Biol.* Cell Press, 2021. Т. 28, № 4. С. 436–455.
218. Ruan T.L. и др. Lowering the overall charge on TMPyP4 improves its selectivity for G-quadruplex DNA // *Biochimie.* Elsevier B.V., 2017. Т. 132. С. 121–130.
219. Morris M.J. и др. The porphyrin TmPyP4 unfolds the extremely stable G-quadruplex in MT3-MMP mRNA and alleviates its repressive effect to enhance translation in eukaryotic cells // *Nucleic Acids Res.* *Nucleic Acids Res.*, 2012. Т. 40, № 9. С. 4137–4145.
220. Zyner K.G. и др. G-quadruplex DNA structures in human stem cells and differentiation // *Nat Commun.* 2022. Т. 13, № 1. С. 142.
221. Wang Y. и др. SEdb 2.0: a comprehensive super-enhancer database of human and mouse // *Nucleic Acids Res.* 2023. Т. 51, № D1. С. D280–D290.
222. Shiraki T. и др. Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage // *Proceedings of the National Academy of Sciences.* 2003. Т. 100, № 26. С. 15776–15781.
223. Brázda V. и др. G4Hunter web application: a web server for G-quadruplex prediction // *Bioinformatics.* 2019. Т. 35, № 18. С. 3493–3495.
224. Pokorná P. и др. Molecular dynamics simulations reveal the parallel stranded d(GGGA)<sub>3</sub>GGG DNA quadruplex folds via multiple paths from a coil-like ensemble // *Int J Biol Macromol.* 2024. Т. 261. С. 129712.
225. Kasai F., Mizukoshi K., Nakamura Y. Variable characteristics overlooked in human K-562 leukemia cell lines with a common signature // *Sci Rep.* 2024. Т. 14, № 1. С. 9619.
226. Philipsen S., Hardison R.C. Evolution of hemoglobin loci and their regulatory elements // *Blood Cells Mol Dis.* 2018. Т. 70. С. 2–12.

227. Shen Y. и др. Identification of a Novel Enhancer/Chromatin Opening Element Associated with High-Level  $\gamma$ -Globin Gene Expression // *Mol Cell Biol*. 2018. Т. 38, № 19.
228. Stellwagen E., Renze A., Stellwagen N.C. Capillary electrophoresis is a sensitive monitor of the hairpin–random coil transition in DNA oligomers // *Anal Biochem*. 2007. Т. 365, № 1. С. 103–110.
229. Strom A.R. и др. Interplay of condensation and chromatin binding underlies BRD4 targeting // *Mol Biol Cell*. 2024. Т. 35, № 6.
230. Wang R. и др. Bromodomain Protein Brd4 Associated with Acetylated Chromatin Is Important for Maintenance of Higher-order Chromatin Structure // *Journal of Biological Chemistry*. 2012. Т. 287, № 14. С. 10738–10752.
231. André A.A.M., Spruijt E. Liquid–Liquid Phase Separation in Crowded Environments // *Int J Mol Sci*. 2020. Т. 21, № 16. С. 5908.
232. Han X. и др. Roles of the BRD4 short isoform in phase separation and active gene transcription // *Nat Struct Mol Biol*. 2020. Т. 27, № 4. С. 333–341.
233. Oh K.-I. и др. Liquid–Liquid Phase Separation Produces Fast H-Bond Dynamics in DMSO–Water Mixtures // *J Phys Chem Lett*. 2020. Т. 11, № 5. С. 1903–1908.
234. Filippakopoulos P. и др. Selective inhibition of BET bromodomains // *Nature*. 2010. Т. 468, № 7327. С. 1067–1073.
235. Ulianov S. V и др. Suppression of liquid–liquid phase separation by 1,6-hexanediol partially compromises the 3D genome organization in living cells // *Nucleic Acids Res*. 2021. Т. 49, № 18. С. 10524–10541.
236. Alberti S., Gladfelter A., Mittag T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates // *Cell*. 2019. Т. 176, № 3. С. 419–434.
237. Ahmed A.A. и др. Asymmetrically Substituted Quadruplex-Binding Naphthalene Diimide Showing Potent Activity in Pancreatic Cancer Models // *ACS Med Chem Lett*. 2020. Т. 11, № 8. С. 1634–1644.
238. Neidle S. и др. Abstract 4069: The potent quadruplex-binding compound SOP1812 shows anti-tumor activity in patient-derived *in vivo* models of pancreatic cancer // *Cancer Res*. 2022. Т. 82, № 12\_Supplement. С. 4069–4069.
239. Gibson B.A. и др. In diverse conditions, intrinsic chromatin condensates have liquid-like material properties // *Proceedings of the National Academy of Sciences*. 2023. Т. 120, № 18.
240. Lovén J. и др. Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers // *Cell*. 2013. Т. 153, № 2. С. 320–334.

241. Wang C. и др. The C-terminal low-complexity domain involved in liquid–liquid phase separation is required for BRD4 function in vivo // *J Mol Cell Biol.* 2019. Т. 11, № 9. С. 807–809.
242. Wang C. и др. A natural product targets BRD4 to inhibit phase separation and gene transcription // *iScience.* 2022. Т. 25, № 1. С. 103719.
243. Yinzhu Hou и др. G-quadruplex inducer/stabilizer pyridostatin targets SUB1 to promote cytotoxicity of a transplatinum complex // *ChemRxiv.* 2021.
244. Müller S. и др. Pyridostatin analogues promote telomere dysfunction and long-term growth inhibition in human cancer cells // *Org Biomol Chem.* 2012. Т. 10, № 32. С. 6537.
245. Rao S.S.P. и др. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping // *Cell.* 2014. Т. 159, № 7. С. 1665–1680.
246. DeMeis J.D. и др. Long G4-rich enhancers target promoters via a G4 DNA-based mechanism // *Nucleic Acids Res.* 2025. Т. 53, № 2.
247. Langmead B., Salzberg S.L. Fast gapped-read alignment with Bowtie 2 // *Nat Methods.* 2012. Т. 9, № 4. С. 357–359.
248. Zhang Y. и др. PePr: A peak-calling prioritization pipeline to identify consistent or differential peaks from replicated ChIP-Seq data // *Bioinformatics.* Oxford University Press, 2014. Т. 30, № 18. С. 2568–2575.
249. Ramírez F. и др. deepTools2: a next generation web server for deep-sequencing data analysis // *Nucleic Acids Res.* Oxford University Press, 2016. Т. 44, № W1. С. W160–W165.
250. Belmonte-Reche E., Morales J.C. G4-iM Grinder: when size and frequency matter. G-Quadruplex, i-Motif and higher order structure search and analysis tool // *NAR Genom Bioinform.* 2020. Т. 2, № 1.
251. Quinlan A.R., Hall I.M. BEDTools: a flexible suite of utilities for comparing genomic features // *Bioinformatics.* 2010. Т. 26, № 6. С. 841–842.
252. Bailey T.L. и др. MEME SUITE: tools for motif discovery and searching // *Nucleic Acids Res.* 2009. Т. 37, № Web Server. С. W202–W208.
253. Grant C.E., Bailey T.L., Noble W.S. FIMO: scanning for occurrences of a given motif // *Bioinformatics.* 2011. Т. 27, № 7. С. 1017–1018.
254. Gabelica V. и др. Multiple and Cooperative Binding of Fluorescence Light-up Probe Thioflavin T with Human Telomere DNA G-Quadruplex // *Biochemistry.* 2013. Т. 52, № 33. С. 5620–5628.
255. Klinov D. и др. High-resolution atomic force microscopy of duplex and triplex DNA molecules // *Nanotechnology.* 2007. Т. 18, № 22. С. 225102.

256. Zanin I. и др. Genome-wide mapping of i-motifs reveals their association with transcription regulation in live human cells // *Nucleic Acids Res.* 2023. Т. 51, № 16. С. 8309–8321.
257. Abueg L.A.L. и др. The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update // *Nucleic Acids Res.* 2024. Т. 52, № W1. С. W83–W94.
258. Li H. Minimap2: pairwise alignment for nucleotide sequences // *Bioinformatics.* 2018. Т. 34, № 18. С. 3094–3100.
259. Feng J. и др. Identifying ChIP-seq enrichment using MACS // *Nat Protoc.* 2012. Т. 7, № 9. С. 1728–1740.
260. North B.V., Curtis D., Sham P.C. A Note on the Calculation of Empirical P Values from Monte Carlo Procedures // *The American Journal of Human Genetics.* 2002. Т. 71, № 2. С. 439–441.
261. Shtork A.S. и др. Quantitative Analysis of Biomolecular Condensates on a Modified Support // *Russ J Bioorg Chem.* 2025. Т. 51, № 1. С. 273–284.