

Расшифровка роли последовательности ДНК в коконденсации белка с ДНК

Рохит Кумар Сингх^{1,2}, Пинаки Суэйн^{1,2}, Махипал Ганджи³, Сандип Чуби^{1,2}

¹Институт математических наук, кампус CIT, Тарамани, Ченнаи 600113, Индия

²Национальный институт Хоми Бхабхи, комплекс учебных заведений, Анушактинагар, Мумбаи 400094, Индия

³Кафедра биохимии, Индийский институт науки, Бангалор, Индия

Абстрактный

Геном эукариот организован в ядре клетки посредством трёхмерной компактизации. Физические принципы, управляющие организацией генома *in vivo*, остаются менее изученными. Фазовое разделение белка и ДНК оказалось перспективным механизмом для перестройки хроматина и компактизации генома. Исследования *in vitro* пролили свет на биофизические принципы образования конденсатов белок-ДНК, обусловленных взаимодействиями белок-белок и белок-ДНК. Однако роль последовательности ДНК и ее влияние на конденсацию белка с ДНК остаются неясными. В данной работе, основанной на экспериментальных данных, представлена простая полимерная модель белок-опосредованной конденсации ДНК, которая явно учитывает влияние последовательности ДНК на связывание с белками. Используя моделирование крупнозернистой броуновской динамики, мы показываем, что в случае гомогенной ДНК в равновесии образуется только один конденсат. В противоположность этому, гетерогенность последовательности ДНК может приводить к сосуществованию нескольких конденсатов, что приводит к образованию структур, напоминающих жемчужные ожерелья. Интересно, что мы наблюдаем, что сродство к связыванию белка с ДНК на границе раздела определяет капиллярные силы, возникающие в конденсатах белок-ДНК. Чтобы продемонстрировать полезность нашей модели, мы сравниваем результаты моделирования с опубликованными данными по коконденсации Dps, Sox2 и HP1. Мы обнаружили, что, хотя Dps демонстрирует связывание, независимое от последовательности, гетерогенность последовательности ДНК определяет коконденсацию Sox2 и HP1 с ДНК. В целом, разработанная здесь модель может быть использована для получения механистического понимания роли последовательности ДНК в коконденсации белка с ДНК и прокладывает путь к более глубокому пониманию организации генома.

1 Введение

Геном эукариот пространственно организован в ядре клетки посредством трёхмерной (3D) компактизации. Такая пространственная организация генома кодирует пространственно-временную регуляцию экспрессии генов, репликации ДНК и т.д. [1–4]. Геном организован иерархически, демонстрируя чёткие, но повсеместно встречающиеся архитектурные особенности на разных масштабах длины, такие как петли, домены и компартменты, образуемые хроматином [4, 5]. Выявление физических принципов регуляции и ремоделирования хроматина *in vivo* остаётся центральной задачей регуляторной биологии.

Недавние эксперименты *in vitro* показывают, что факторы транскрипции и ДНК могут подвергаться совместной конденсации. через коллективное поведение тысяч белков и ДНК, управляемое белок-белковыми и

*Эти авторы внесли равный вклад в эту работу.

Взаимодействие белка с ДНК [6–9]. Полученный коконденсат белка с ДНК равен определённому количеству ДНК, оказывая силу на окружающую свободную ДНК, известную как капиллярная сила [9–12]. Силы, возникающие из физики взаимодействия белково-ДНК-конденсатов, могут служить как механизм регуляции и перестройки хроматина. Эта способность обещает влиять на трехмерная архитектура генома.

Экспериментальные исследования изучили физику взаимодействия белково-ДНК-соконденсатов и их влияние на капиллярные силы с помощью оптического пинцета и анализа на основе покровных стекол, как показано на рисунке 1 A, B. Например, Куэйл и др. использовали флуоресценцию полного внутреннего отражения (TIRF) Микроскопия для исследования коконденсации ДНК λ -фага и белка-первопроходца фактора транскрипции FoxA1 [10]. Результаты показали, что коконденсация FoxA1 и ДНК может генерировать силы порядка 0,2 пиконьютона (пН). Другое исследование показало, что коконденсация PARP1-ДНК генерируют капиллярные силы в субпиконьютонном диапазоне [11]. Интересно, что более недавнее исследование продемонстрировано, что плюрипотентный фактор Sox2 способен генерировать капиллярные силы до 7 пН [9]. Хотя сочетание теории и экспериментов дало ценные физические данные, понимание формирования коконденсатов белка и ДНК, всестороннее понимание физика межфазных соконденсатов и их влияние на результирующие капиллярные силы остаются в младенчестве. В частности, влияние гетерогенности последовательности ДНК на коконденсацию белка с ДНК. и капиллярные силы остаются менее изученными. Цель данной работы — разработать модельную модель, которая позволит изучать физику коконденсатов белка и ДНК в наличие гетерогенности последовательности ДНК.

Для этой цели мы использовали простую модель коконденсации белка и ДНК, которая явно учитывает гетерогенность последовательности ДНК для связывания белка. Более того, наша модель специально разработана разработан для захвата двух типов систем *in vitro*, используемых при изучении коконденсации белка и ДНК: Анализы с использованием пинцета и покровного стекла. Используя моделирование крупнозернистой броуновской динамики, мы демонстрируем, что в случае однородного полимера ДНК в равновесии образуется только один конденсат. Напротив, в зависимости от гетерогенности последовательности ДНК, могут образовываться множественные конденсаты разных размеров. сосуществуют на одной и той же ДНК в равновесии. Наше исследование показывает, что сродство связывания с белками ДНК на границе раздела определяет величину капиллярных сил. Это наблюдение открывает механизм Регулирование глобальных воздействий на геном посредством изменений на уровне пар оснований ДНК. Подводя итог, можно сказать, что наша модель моделирования представляет собой ценный инструмент для использования существующих систем *in vitro* для получения механистическое понимание биофизики совместной конденсации белка и ДНК.

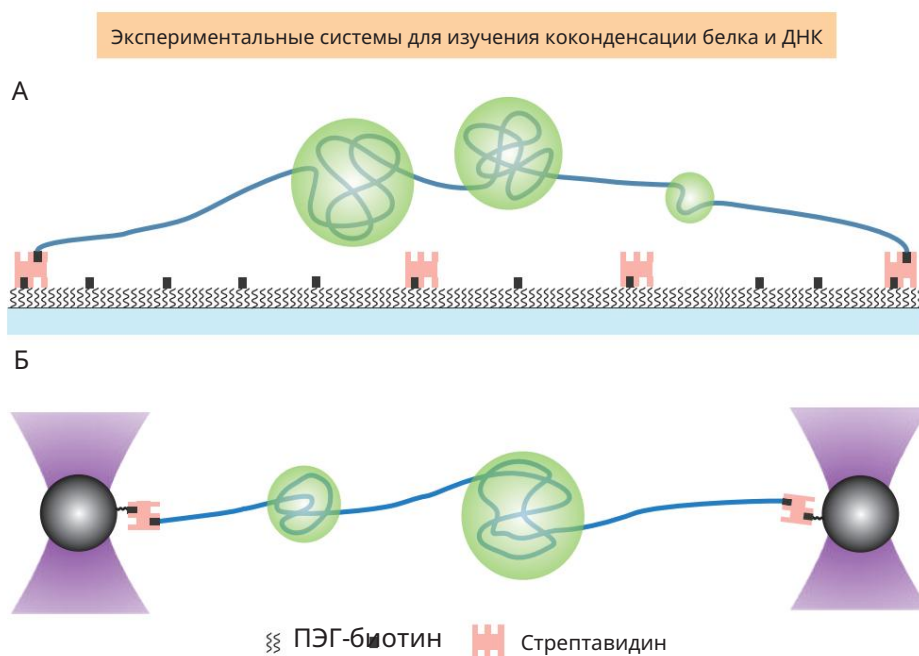


Рисунок 1: Экспериментальные системы для изучения коконденсации белка и ДНК. Схема экспериментальной установки, в которой концы одной молекулы ДНК (синего цвета) прикреплены к предметному стеклу с использованием взаимодействия биотина и стрептавидина. Белки (зелёные) разделяются с ДНК, образуя коконденсаты белок-ДНК. В Схема анализа с помощью оптического пинцета: молекула ДНК (синяя) удерживается между двумя оптически захваченными шариками (чёрными) посредством взаимодействия биотина и стрептавидина.

2 модели

Чтобы изучить влияние последовательности ДНК на коконденсацию комплексов белок-ДНК и в связи с последующим возникновением капиллярных сил мы предлагаем простую полимерную модель связывания белка с ДНК. Эта модель подробно учитывает как белок-белковые взаимодействия, так и взаимодействия белка с ДНК, зависящие от последовательности.

Основываясь на исследованиях *in vitro*, показанных на рисунке 1, мы считаем ДНК длиной 5 кб полуциркулярной. Полимер, состоящий из 500 мономеров, закреплённых с обоих концов внутри кубовидного блока. В соответствии с Согласно предыдущим исследованиям [13], персистентная длина ДНК составляет 150 пар оснований, что соответствует 15 мономерам в нашей модели. ДНК контактирует с белками. Без потери В общем, мы предполагаем, что каждый мономер соответствует участку из 10 занятых оснований связанной молекулой белка. Действительно, различные прокариотические и эукариотические белки занимают 10 оснований на ДНК [14, 15]. Мы предполагаем, что белки представляют собой сферические гранулы с одинаковыми размерами к мономерам. Гетерогенность последовательности ДНК представлена различными расположениями мономеров вдоль контура ДНК, где каждый мономер проявляет различную связывающую способность к Белки. Чтобы проанализировать влияние последовательности ДНК на коконденсацию белка с ДНК, сначала мы вводим нулевую модель гомогенной ДНК. В этой модели каждый мономер имеет одинаковую связь. сродство к белкам 2 кБТ (см. рис. 2A)[7]. В противоположность этому, мы рассматриваем три различные гетерогенные модели ДНК. 1) Гетерогенная ДНК I. В этой модели ДНК представлена как

Блок-сополимер АБК, основанный на дифференциальном притяжении к белкам. ДНК состоит из высокоаффинной области, охватывающей 250 мономеров в центре, с аффинностью связывания белок-ДНК 2,25 кБ. Эта высокоаффинная область окружена с обеих сторон низкоаффинными областями, каждая из которых содержит 125 мономеров, с аффинностью связывания белок-ДНК 1,75 кБ. Это гарантирует, что средняя аффинность связывания белок-ДНК останется такой же, как и в гомогенном случае (см. рис. 2B). 2)

Гетерогенная ДНК II, в этой модели пять блоков высокоаффинных областей по 2,25 т.п.н., каждая из которых охватывает 50 мономеров вдоль контура ДНК, перемежаются блоками низкоаффинных областей по 1,75 т.п.н., охватывающими 50 мономеров. Средняя аффинность связывания белок-ДНК для этой модели идентична гомогенному случаю (см. Рисунок 4A).

3) Частичная λ -ДНК, для моделирования биологически реалистичной ДНК мы рассматриваем последние 5 т.п.н. лямбда-ДНК (см. Рисунок 5A и раздел «Методы»). Для простоты мы предполагаем, что каждый белок имеет предпочтение связываться с АТ-богатыми областями. Действительно, в бактериях и высших организмах существует множество белков, которые широко связываются с АТ- или GC-богатыми последовательностями [16–18]. Мы назначаем соответствующую энергию взаимодействия каждому мономеру на основе содержания АТ в каждой 10-основной последовательности (подробности см. в Методах и SI).

После экспериментов мы настраиваем нормализованное расстояние между концами ДНК, а объем определяется как Концентрация белка. Нормализованное расстояние между концами (R') и ρ отношение расстояния контурная длина ДНК (s). Общая концентрация белка ($\rho\rho$) — это начальная концентрация белков, введенных в область моделирования, в единицах мкМ. Мы используем моделирование крупнозернистой броуновской динамики для изучения конденсации белка с ДНК, зависящей от последовательности. Подробное обсуждение методов моделирования и различных параметров см. в разделе «Методы».

3 результата

3.1 Последовательность ДНК управляет капиллярными силами, исходящими от белков-ДНК-кокконденсаты

Мы стремимся раскрыть влияние гетерогенности последовательностей ДНК на образование кокконденсатов белок-ДНК и возникающие в результате этого капиллярные силы. В частности, мы рассматриваем две модели: гомогенную ДНК, называемую гомо-ДНК (рис. 2A), и гетерогенную ДНК, называемую гетеро-ДНК (рис. 2B). Подробное обсуждение выбора параметров см. в разделе «Методы», таблицах 1 и 2 и разделе «Модель».

Во-первых, мы сохраняем нормализованное расстояние между концами (R'_c) фиксированы и систематически увеличивают концентрацию белка в объеме ($\rho\rho$). Мы наблюдаем образование кокконденсата белок-ДНК в равновесии для гомогенной ДНК (рис. 2E). Такая кокконденсация происходит при концентрации ниже концентрации насыщения для разделения фаз в объеме, что согласуется с предыдущими теоретическими [19] и экспериментальными исследованиями [20]. Однако интересно отметить, что мы наблюдаем образование одного конденсата вблизи одного из связанных концов, как показано на рис. 2E (рис. S1A и S1B). Мы наблюдали это явление последовательно при различных параметрах модели (рис. S2A). Более того, когда сродство связывания белок-белок ниже сродства связывания белок-ДНК, возникают вытянутые конденсаты (рис. S2B). По мере усиления белок-белковых взаимодействий форма конденсатов становится более сферической. В то время как при более высоких концентрациях белка вдоль ДНК образуются множественные кокконденсаты, эти конденсаты со временем сливаются и укрупняются, образуя единый конденсат в равновесии. Давление Лапласа способствует этому процессу укрупнения, поскольку единый конденсат минимизирует площадь поверхности по сравнению с несколькими более мелкими конденсатами в

коконденсат. Напротив, в случае гетерогенной ДНК I конденсат образуется в равновесии в области высокого сродства ДНК, как показано на рисунках 2F и 2H. Хотя в начале моделирования вдоль ДНК образуются несколько конденсатов, в конечном итоге система сходится к одному конденсату, как показано на кимографе на рисунке 2J (рис. **S3A** и S3B).

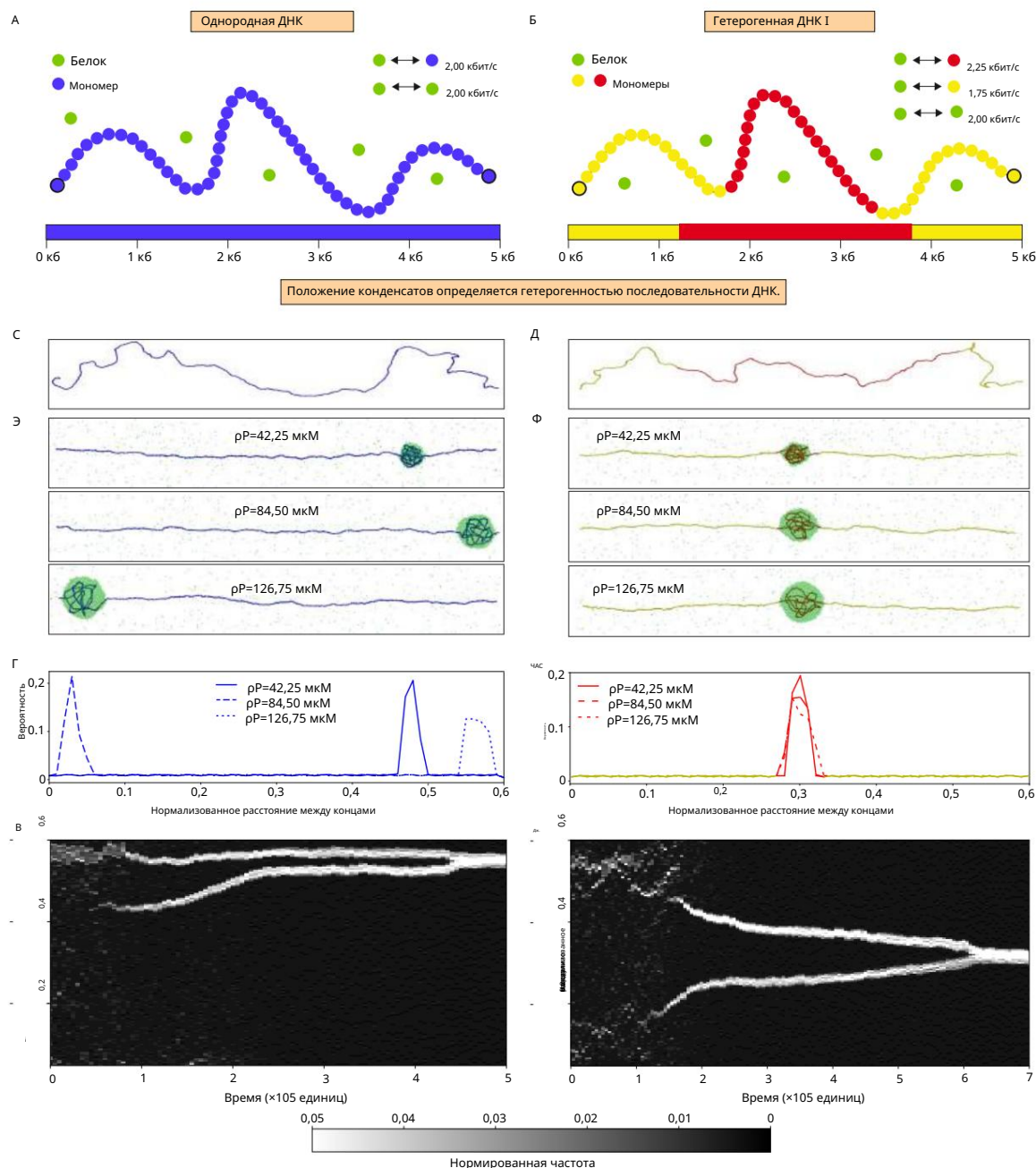


Рисунок 2: Гетерогенность последовательности ДНК определяет положение конденсатов.

Схема представляет собой гомогенную ДНК; каждый мономер (синий) связывается с белками (зеленый) с помощью идентичного сродства связывания (2 кБТ). Показана модель гетерогенной ДНК ВА (гетерогенная ДНК I); Модель учитывает эффект гетерогенности последовательности ДНК путем введения двух типов мономеров; Мономеры (красные) в центре связывают белки с силой сродства 2,25 кБт, тогда как мономеры (желтые) на периферии имеют силу взаимодействия 1,75 кБт. Белки связываются друг с другом с силой 2,00 кБт. Подробности параметров см. в таблицах 1 и 2. С, D. Снимки Показаны конфигурации ДНК для гомогенной и гетерогенной ДНК I при отсутствии белка. Е и F, показаны моментальные снимки коконденсатов белка и ДНК в равновесии для двух моделей. Для построения графиков мы зафиксировали нормализованное расстояние между концами ($R'_e = 0,6$) и варьировали ρ_P . G, H Вероятность нахождения каждого мономера внутри конденсата (ширина полосы = 5 σ) показана как функция положения вдоль ДНК для гомогенной (синяя) и гетерогенной ДНК соответственно. H Линии, выделенные желтым и красным, отражают ДНК с разным сродством. I и J – репрезентативные кимографы для определения положения конденсатов вдоль контура (ширина полосы = 2 σ) в зависимости от времени гомогенная ДНК ($R'_e = 0,6$, $\rho_P = 109,9 \text{ мкМ}$) и гетерогенная ДНК I ($R'_e = 0,6$, $\rho_P = 93,01 \text{ мкМ}$) показаны.

Далее мы исследуем влияние конденсации белка-ДНК на капиллярные силы (как функцию). Для однородного концентрации белка (pp) и нормализованного расстояния от конца до конца ДНК (R') мы наблюдаем незначительное увеличение капиллярных сил на голой ДНК по мере увеличения pp и R' (рис. 3А и 3В). Чтобы подтвердить, что силы на голой ДНК действительно возникают из интерфейса, мы вычисляем полную потенциальную энергию (см. Методы) для каждого мономера вдоль контура для однородной ДНК при $R' = 0,2$ и $pp = 84,50$ мкМ (рис. S1D). Потенциальная энергия резко меняется для мономеров, присутствующих на интерфейсе. Для гетерогенной ДНК I демонстрирует нелинейный отклик относительно R' .

В начале pp приводит к небольшому увеличению капиллярной силы, за которым следует резкий скачок при $R' = 0,5$ (см. рис. 3А). С другой стороны, по мере увеличения pp наблюдается обратная тенденция (рис. 3Б). Капиллярная сила изначально остается постоянной при увеличении pp , а затем постепенно уменьшается после $pp = 59,15$ мкМ. Поведение капиллярных сил для гетерогенной ДНК I определяется сродством интерфейсной ДНК. Когда R' превышает определенный порог (здесь 0,5), ДНК с более высоким сродством (показана красным) становится доступной на границе конденсата (между компактизированной ДНК и чистой ДНК), что приводит к увеличению сил, действующих на чистую ДНК (рисунок S3D). Напротив, с увеличением pp объем конденсата линейно увеличивается (рисунок S3C), что приводит к тому, что интерфейс конденсата становится доступным для ДНК с низким сродством (1,75 кБт) (показана желтым). В результате происходит уменьшение капиллярных сил, действующих на чистую ДНК (рисунки 3В и S3E). В целом, наши результаты показывают, что сродство ДНК к связыванию с белками определяет капиллярные силы. В следующих разделах мы рассмотрим, как различные виды гетерогенности ДНК влияют на капиллярные силы.

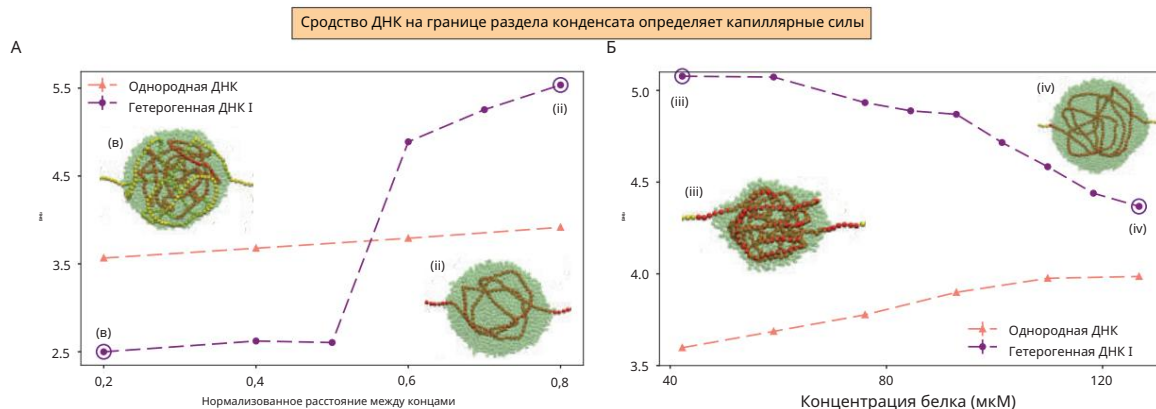


Рисунок 3: Сродство к ДНК на границе раздела фаз определяет капиллярные силы. Графики А и Б: зависимость капиллярной силы от R' при $pp = 84,50$ мкМ и от pp при $R' = 0,6$ для гомогенной (лососевой) и гетерогенной ДНК I (фиолетовой) соответственно.

На снимках показаны конденсаты гетерогенной ДНК I, где красным цветом обозначены мономеры с высоким сродством, а жёлтым — мономеры с низким сродством.

Белки показаны зелёным цветом. Снимки (слева направо) показывают конденсаты для следующих наборов параметров: (i) $R' = 0,2$ и $pp = 84,50$ мкМ, (ii) $R' = 0,8$ и $pp = 84,50$ мкМ, (iii) $R' = 0,6$ и $pp = 42,25$ мкМ и (iv) $R' = 0,6$ и $pp = 126,75$ мкМ.

3.2 Гетерогенность последовательности ДНК приводит к образованию жемчужин, напоминающих ожерелье структуры

Для дальнейшего изучения влияния гетерогенности ДНК на коконденсацию белка с ДНК мы рассмотрим модель «Гетерогенная ДНК II», показанную на рис. 4А (подробнее см. в разделе «Модель»). Модель «Гетерогенная ДНК II» позволяет нам исследовать влияние более распределённой гетерогенной последовательности ДНК на конденсацию белка с ДНК.

Подобно наблюдениям с однородной ДНК и гетерогенной ДНК I (рис. 2E-2J), с гетерогенной ДНК II мы наблюдаем образование множественных конденсатов в начале моделирования. Эффект был выражен при более высоких концентрациях белка (ρ) и меньших нормализованных расстояниях между концами (R'). Конденсаты разных размеров могут сосуществовать вдоль ДНК в равновесии, образуя e . Интересно, что в отличие от предыдущих моделей ДНК, мы обнаружили, что несколько структур, напоминающие жемчужные ожерелья (см. рис. 4В и рис. S4А). Примечательно, что в то время как однородная ДНК приводит к единственному конденсату в равновесии из-за огрубления и слияния, для однородной ДНК II наличие локальных областей высокого сродства вдоль ДНК может остановить огрубление. Более того, количество конденсатов изменяется немонотонно как функция R' и ρ (см. рис. S4B и S4C).

e

В случае объемного фазового разделения капли либо сливаются, либо укрупняются за счет диффузии молекул белка из-за разницы в давлении Лапласа. Такое укрупнение приводит к образованию единого конденсата в равновесии. Однако в нашей системе капли также взаимодействуют друг с другом посредством капиллярных сил, действующих через ДНК, что приводит к перетягиванию каната. Как мы наблюдали в предыдущем разделе на примере гетерогенной ДНК II, эти капиллярные силы зависят от сродства ДНК к поверхности раздела. Очевидно, для сосуществования множества конденсатов вдоль молекулы ДНК система должна достичь механического равновесия. Систематическое исследование объемов капель и их среднего сродства к ДНК на границе раздела показывает, что в сценариях с конденсатами различных размеров меньшие конденсаты демонстрируют большее среднее сродство к ДНК на границе раздела по сравнению с их более крупными аналогами (случаи 1 и 3 на рис. 4С). Конденсаты схожего объёма демонстрируют схожее среднее сродство к ДНК на границе раздела (случай 2 на рис. 4С). Следовательно, при изменении ρ и $R'e$, в зависимости от сложного взаимодействия положения конденсатов и их среднего сродства к ДНК на границе раздела, мы получаем разное количество конденсатов вдоль ДНК.

Далее мы рассмотрим капиллярные силы как функцию $R'e$ и ρ . Капиллярные силы демонстрируют немонотонное поведение в обоих случаях (рис. 4D и 4E) для гетерогенной ДНК II. Такое немонотонное поведение обусловлено блочной архитектурой ДНК, основанной на средних значениях межфазного сродства (см. метод 5.4) для различных условий, демонстрирующих наличие одиночных и множественных конденсатов в равновесии. Для дальнейшей демонстрации этого мы вычислили коэффициент корреляции Пирсона между капиллярными силами и межфазным сродством при различных условиях (см. метод 5.5). Как показано на рис. 4F, мы обнаружили высокую степень корреляции между ними.

Ключевым выводом этого раздела является то, что гетерогенность последовательности ДНК может существенно влиять на количество и положение коконденсатов белок-ДНК.

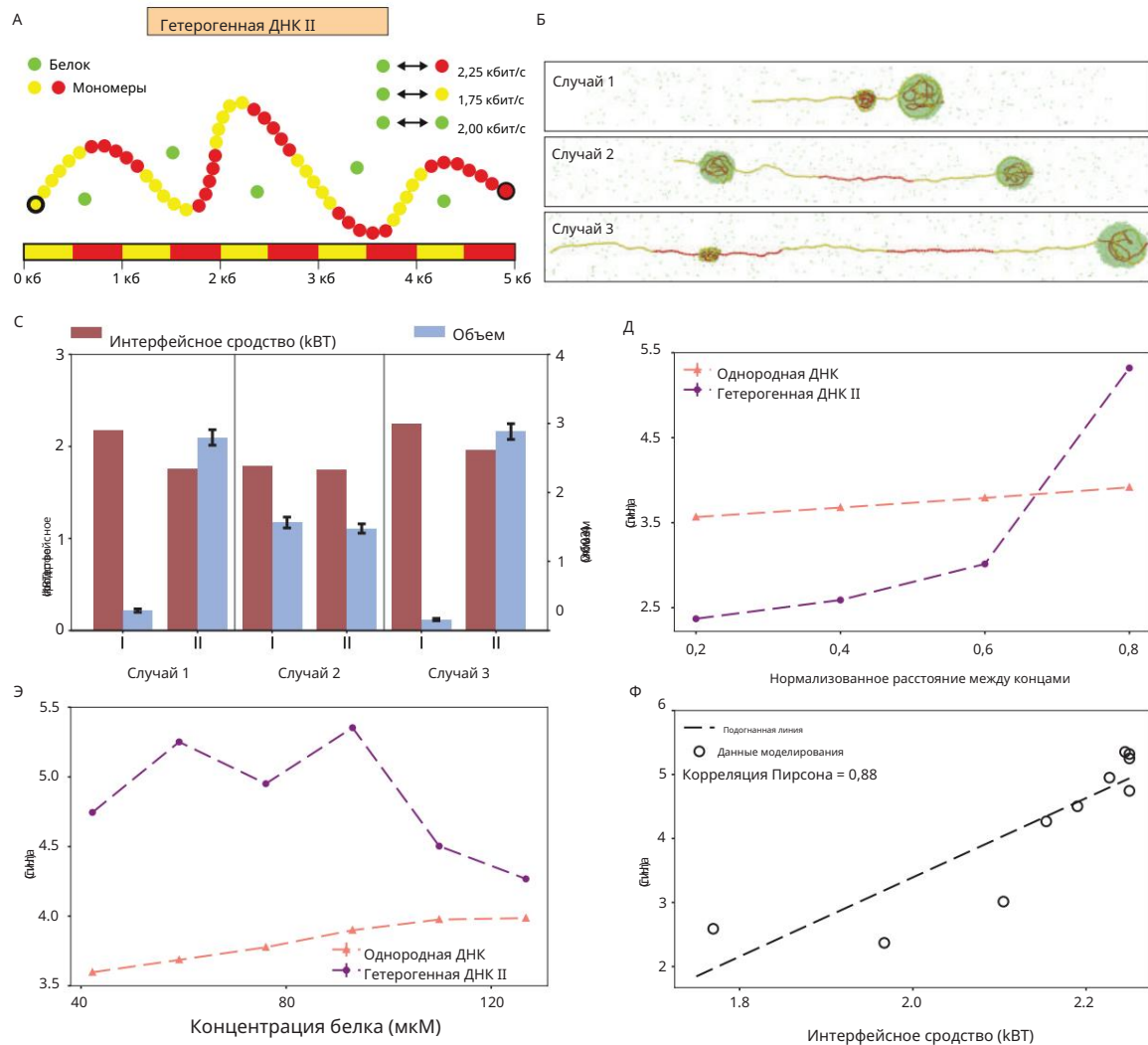


Рисунок 4: Гетерогенность ДНК приводит к сосуществованию нескольких белково-ДНК-конденсатов. Схематически представлена гетерогенная ДНК-II. Высокоаффинные мономеры (2,25 кБт, красные).

распределены по контуру в виде блоков (длина = 50 мономеров), перемежающихся с блоками низкоаффинных мономеров (1,75 кБт, жёлтый) одинаковой длины. Белки (зелёные) взаимодействуют друг с другом. другой с силой взаимодействия 2,00 кБт. В. Снимки показаны для случаев с несколькими конденсатами, сосуществующими в равновесии; соответствующие значения параметров: для случая 1, $R' = 0,2$ и $pp = 84,50$ мкМ, для случая 2 $R' = 0,4$ и $pp = 84,50$ мкМ, а для случая 3 $R' = 0,6$ и $pp = 84,50$ мкМ (сверху вниз). Высокоаффинные мономеры, низкоаффинные мономеры и белки показаны красным, желтовато-низкоаффинным и зеленым цветом соответственно. С. Гистограмма показывает средние значения межфазного сродства (бордовый) и объемы. (синий) для двух конденсатов (I и II), показанных на снимках для всех трёх случаев. D, E Капиллярная сила отображается как функция R' при $pp = 84,50$ мкМ и как функция pp при $R' = 0,6$ для гетерогенного ДНК II (фиолетовая) и однородная ДНК (лососевая) соответственно. F, Корреляция между силами и Для всех симуляций были построены графики средних значений интерфейсного сродства для гетерогенной ДНК II.

3.3 Изменения в локальном средстве ДНК на границе раздела регулируют глобальную капиллярные силы.

Чтобы глубже понять, как последовательность ДНК влияет на силы, создаваемые коконденсатами белок-ДНК, далее мы изучим биологически реалистичную последовательность ДНК, частичную λ -ДНК (рис. 5A). Подробное обсуждение этой модели см. в разделе «Модель» (таблица 3).

Частичная λ -ДНК может приводить к сосуществованию нескольких коконденсатов белок-ДНК, аналогично модели гетерогенной ДНК II. Мы обнаружили, что количество конденсатов демонстрирует немонотонное поведение относительно ρ , и конденсатов R' (рис. 5B и 5SA) мы вычислили объёмы и e . Для моделирования, в котором мы наблюдали множественные средние значения сродства к границе раздела этих конденсатов (рис. 5D). В обоих случаях, когда в равновесии присутствуют несколько конденсатов, конденсаты меньшего размера образуют интерфейсы с более высоким средним сродством к границе раздела по сравнению с конденсатами большего размера. Эти результаты согласуются с результатами модели гетерогенной ДНК II.

Примечательно, что результаты, полученные в рамках как гетерогенной модели II, так и частичной модели λ -ДНК, свидетельствуют о том, что для одновременного поддержания нескольких конденсатов вдоль ДНК в равновесии, меньшие конденсаты должны содержать интерфейсную ДНК с более высоким средним сродством.

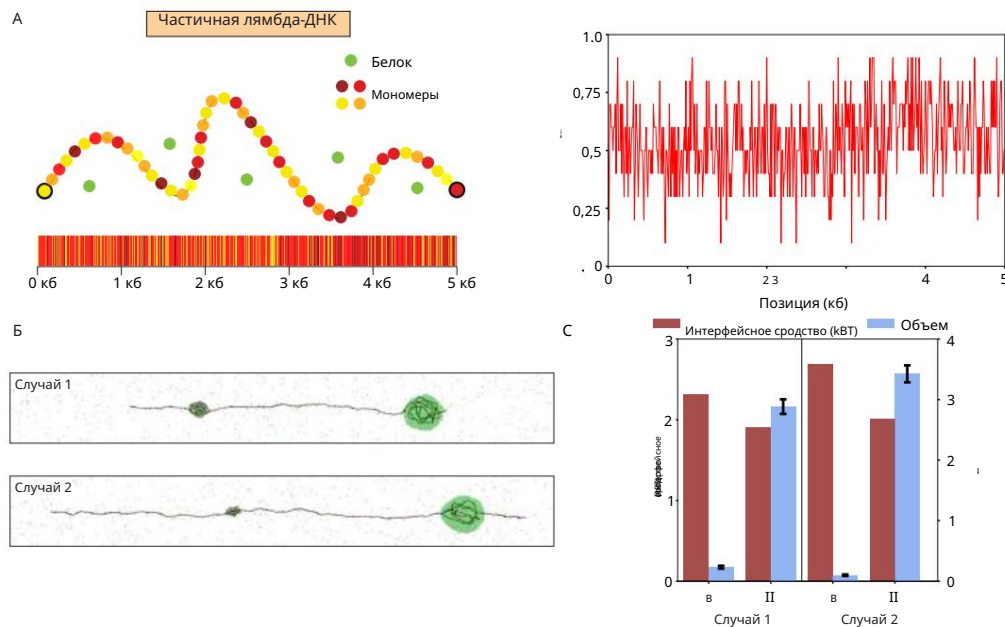


Рисунок 5: Гетерогенность последовательности в частичной λ ДНК порождает множественные конденсаты в Равновесие. А (слева), графическое изображение гетерогенной ДНК III в виде последних 5 кб фага λ . ДНК: Одиннадцать видов мономеров (от желтого до бордового) введены по контуру на основе Содержание АТ в 10 п.н. в частичной последовательности ДНК λ . Сродство связывания мономера с белком варьируется в зависимости от диапазон от 0 до 4 кБТ в зависимости от содержания АТ соответствующего мотива 10 п.н., который представляет Мономер (см. раздел «Модель» и SI). Белки (зелёные) связываются друг с другом с аффинностью 2 кБт. А (справа) показано содержание АТ в каждом из 10 пар оснований на фрагменте ДНК λ . В – Моментальные снимки двух конденсатов. для двух случаев показаны: случай 1, $R' = 0,4$ и $pp = 84,50$ мкМ, и случай 2, $R' = 0,6$ и $pp = 92,95$ мкМ (сверху вниз). С. Гистограмма показывает среднее межфазное сродство (бордовый) и объём (синий) для два конденсата (I и II) для двух случаев, показанных на снимках В.

Капиллярная сила демонстрирует немонотонное поведение в зависимости от pp и R' (см. рисунок 6А) и 6Б). Чтобы исследовать немонотонное поведение капиллярной силы, мы рассмотрим связанные с ней случаи с самой низкой ($pp = 59,15$ мкМ) и самой высокой ($pp = 76,05$ мкМ) концентрацией белка, как показано на рисунке 6А. Для этого мы анализируем сродство мономеров и вероятности их образования, находясь на границе раздела в обоих случаях (рис. 6С и 6Д). При изменении pp от 59,15 мкМ до 76,05 мкМ, мы наблюдаем изменение капиллярной силы на 2 пН. Хотя сегменты ДНК внутри конденсат остается практически идентичным (снимки на рис. 6С и 6Д), последовательности на границе раздела меняются, что приводит к существенному различию капиллярных сил. Сравнение интерфейсов показывает, что включение мотива с высоким сродством в интерфейс для $pp = 76,05$ мкМ приводит к закреплению конденсата вдоль контура ДНК, что приводит к уменьшению ширины интерфейса (рис. 6Д (внизу) панель)). Роль средней аффинности интерфейсной ДНК действительно показывает высокую корреляцию 0,91 с капиллярными силами для различных симуляций (рис. 6Е). Другими словами, небольшие изменения в сродстве интерфейсной ДНК приводит к значительному увеличению глобальных капиллярных сил.

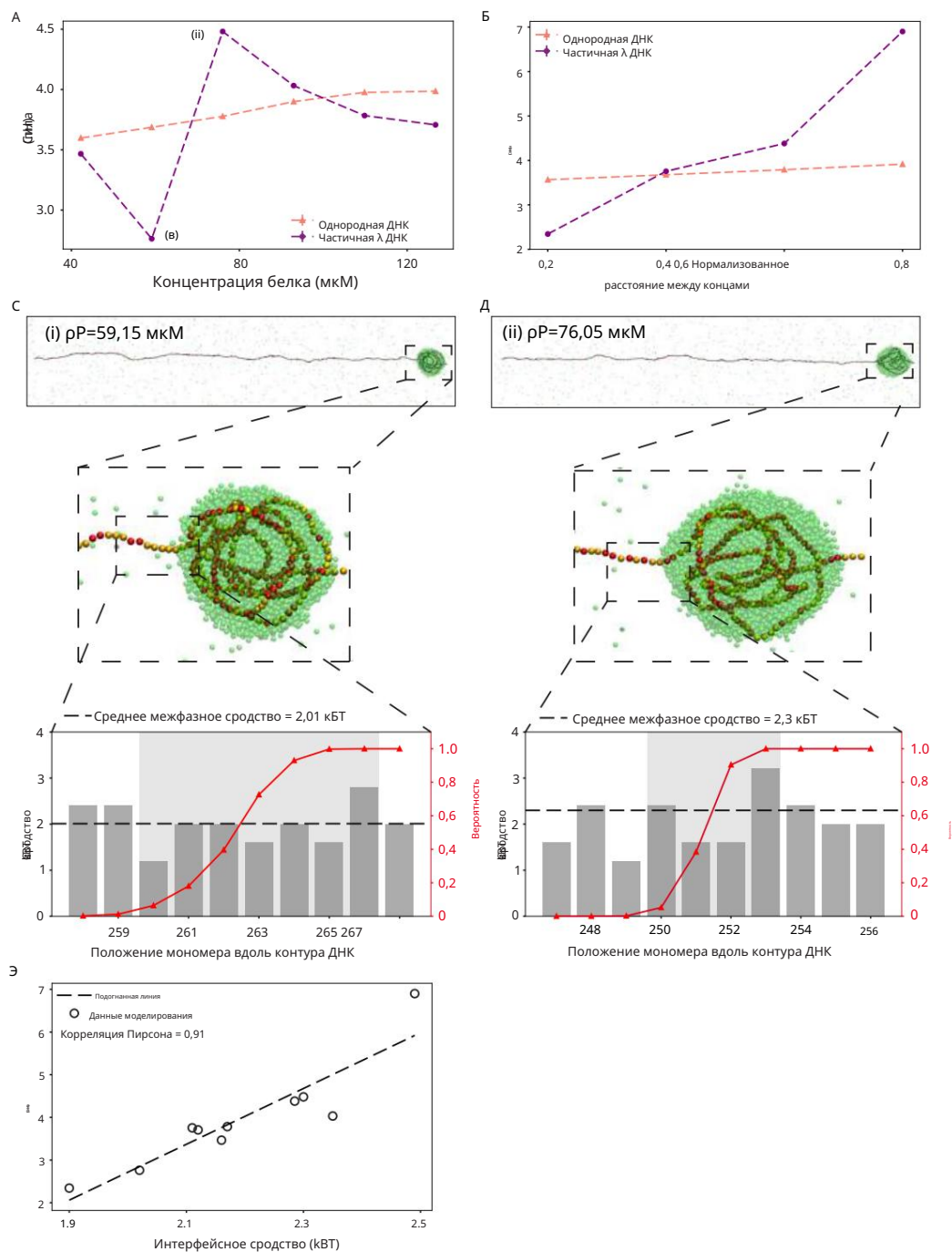


Рисунок 6: Локальное сродство ДНК на границе раздела определяет глобальные капиллярные силы. А и В представляют изменения капиллярных сил в зависимости от pP и R' соответственно. Немонотонное поведение в Силы объясняются с помощью рассмотрения двух условий (i. $pP = 59,15$) и (ii. $76,05$ мкМ) в А, С, D (вверху) панели), показаны моментальные снимки конденсатов для (i) и (ii); мономеры показаны желтым цветом красный (символизирует сродство связывания мономера с белком в порядке возрастания), а белки показаны зеленым цветом. С, D (нижняя панель), столбчатые диаграммы отображают сродство связывания мономера с белком; серый Затенённая область показывает границу раздела. Показаны вероятности нахождения мономеров внутри конденсата. в ред. Е, Корреляция между капиллярными силами и средним сродством к поверхности раздела для всех случаев с один или несколько конденсатов в равновесии.

3.4 HP1 и Sox2 демонстрируют коконденсацию, зависящую от последовательности ДНК, тогда как Dps демонстрирует коконденсацию, независимую от последовательности

Результаты, полученные в ходе нашего моделирования, могут быть использованы для интерпретации недавних экспериментальных данных [9, 10, 20] по коконденсации белков и ДНК. В частности, наши результаты могут пролить свет на роль последовательности ДНК в коконденсации. Для этого мы анализируем пространственный профиль и количество коконденсатов вдоль ДНК в равновесии для набора белков: ассоциированного с нуклеоидом белка Dps, который конденсирует ДНК, соединяя различные её части [21], фактора транскрипции Sox2, связывающего нуклеосому [9], и HP1 [20], который опосредует образование гетерохроматина.

Сначала рассмотрим модель нулевой гомогенности ДНК. Гомогенная модель предсказывает образование единого конденсата, находящегося в равновесии вблизи одного из связанных концов, как показано на рисунке 2. В соответствии с предсказаниями этой модели, недавнее исследование показало, что Dps образует один конденсат на одном из связанных концов. Соответствие между предсказаниями однородной модели и экспериментальными данными подразумевает, что Dps участвует в независимой от последовательности коконденсации с ДНК. Это открытие согласуется с более ранними исследованиями, которые продемонстрировали, что Dps связывается с ДНК независимо от последовательности ([21]). Далее мы рассмотрим совместную конденсацию Sox2-ДНК. В отличие от предсказаний однородной модели, экспериментальные данные выявили противоположную тенденцию: Sox2 образует множественные конденсаты в различных местах вдоль ДНК. Эти наблюдения опровергают однородную модель, предполагая, что последовательность ДНК диктует совместную конденсацию Sox2. Эти открытия согласуются с предыдущими исследованиями, которые показали связывание Sox2, зависящее от последовательности ДНК.

Далее мы рассмотрим коконденсацию HP1 с ДНК. До сих пор связывание HP1 с ДНК считалось относительно независимым от последовательности [22]. Основываясь на этих предыдущих исследованиях, мы предполагаем, что модель гомогенной ДНК эффективно описывает конденсацию HP1 с ДНК. Вопреки нашим ожиданиям, HP1 образует два конденсата, один из которых расположен вблизи связанного конца, а другой – в середине ДНК, тем самым опровергая нашу гипотезу [20]. Очевидно, что модель гетерогенной ДНК с гетерогенностью последовательностей может объяснить сосуществование нескольких конденсатов в разных местах ДНК. Наш анализ показывает, что HP1 проявляет дифференциальную аффинность связывания с ДНК.

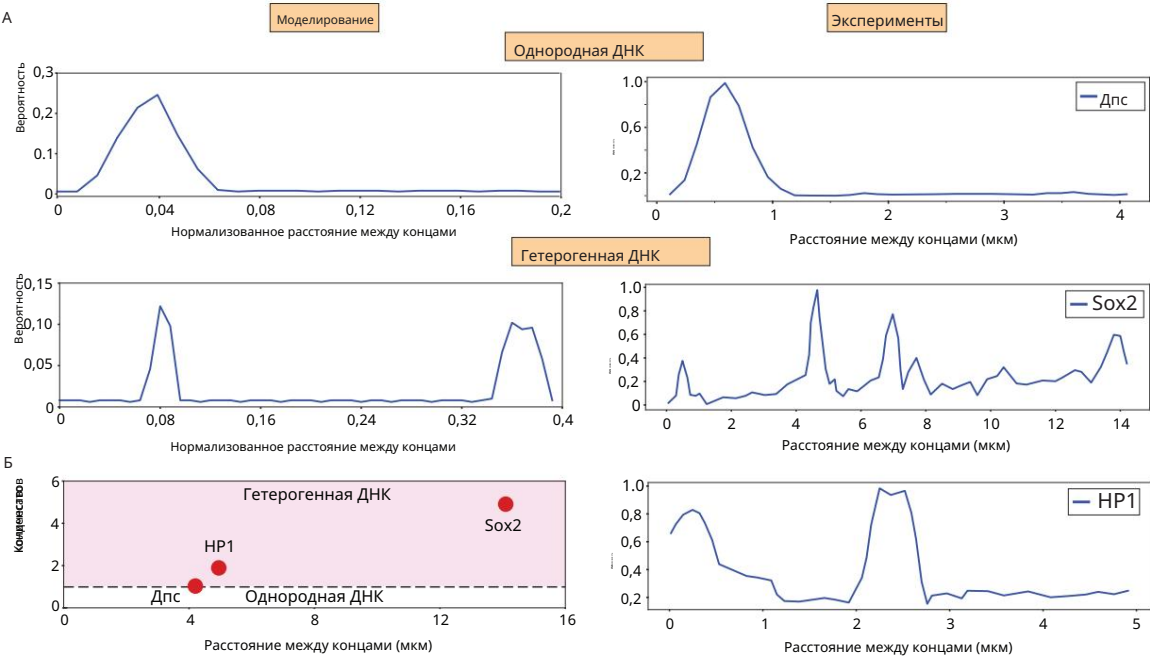


Рисунок 7: Сравнение результатов моделирования и эксперимента. Графики слева показывают: Вероятности нахождения мономеров внутри конденсата для гомогенной и гетерогенной ДНК III Соответственно (сверху вниз). Для сравнения, справа, распределение интенсивности белка вдоль ДНК была построена для Dps (взято из рис. 4 из [21]), Sox2 (взято из рис. 1b из [9]) и HP1 (взято из рис. 4B из [20]) соответственно (сверху вниз). Мы извлекаем данные, используя imageJ и digitizeit (см. раздел «Методы»). В Диаграмма рассеяния показывает количество конденсатов (красный). как функция расстояния от конца до конца для Dps, Sox2 и HP1, прямая линия, изображающая конденсат Число единиц отражает независимую от последовательности коконденсацию, тогда как область, закрашенная розовым цветом, показывает коконденсация белка и ДНК, зависящая от последовательности.

4 Обсуждение

Последние данные свидетельствуют о том, что фазовое разделение хроматина и ДНК может играть решающую роль в Организация эукариотического генома [6, 23–29]. Однако исследование физических принципов функционирования таких фазово-разделённых тел *in vivo* оказалось сложной задачей. В качестве альтернативы, биохимические исследования *in vitro* с очищенными факторами транскрипции (TF) и ДНК пролили свет на конденсацию, опосредованную белками ДНК [8–10]. Серия таких исследований *in vitro* показала, что белок-белок и белок-ДНК взаимодействия приводят к конденсации ДНК[21, 30, 31]. Однако роль последовательности ДНК, таким образом, Отец остался неясным. Моделирование, представленное в данной статье, позволяет определить роль Последовательность ДНК при коконденсации белка и ДНК.

Мы обнаружили, что для однородной ДНК, даже если изначально образуются множественные конденсаты, они имеют тенденцию грубеть со временем, в конечном итоге приводя к образованию единого конденсата — явление, похожее на явление Оствальда созревание. Напротив, множественные коконденсаты могут сосуществовать вдоль гетерогенной ДНК, проявляя Структуры, напоминающие жемчужное ожерелье. При разделении фаз в объёме капли взаимодействуют посредством диффузии белка.

Молекулы. Однако в нашей модели капли также взаимодействуют друг с другом посредством капиллярных сил, друг друга через ДНК, что приводит к эффекту «перетягивания каната». Среднее количество и положение различных Конденсаты в равновесии определяются сложным взаимодействием этих эффектов. Хотя недавние теоретические исследования изучали конденсацию ДНК, опосредованную белками, эти модели в первую очередь были сосредоточены на конденсацию свободных полимеров ДНК[19]. Результаты нашего моделирования показывают, что связывание ДНК на обоих концах и включение гетерогенности последовательности ДНК порождает новые физические идеи, имеющие решающее значение для интерпретации существующих и будущих данных, полученных в исследованиях *in vitro*.

Пересмотрев данные о коконденсации белка и ДНК из предыдущих экспериментальных исследований [9, 10, 20, 21], мы обнаружили, что бактериальный белок Dps, ассоциированный с нуклеотидом, уплотняет ДНК в последовательности, независимой от [21]. Напротив, пионер-фактор Sox2 и гетерохроматиновый белок HP1 демонстрируют последовательность-зависимое связывание с ДНК [9, 20]. Наши результаты для Dps согласуются с результатами предыдущих экспериментов. исследования. Аналогично, известно, что Sox2 проявляет высокую специфичность к последовательностям; sox2 распознаёт 7 пар оснований (п.н.) последовательностей ДНК (СТТТГТТ). Поэтому наше открытие влияния последовательности ДНК на коконденсацию Sox2 неудивительно. Примечательно, что наш повторный анализ недавней статьи о Конденсация ДНК, опосредованная HP1, выявила зависимость связывания HP1 от последовательности [20]. В частности, существование множественных конденсаций подразумевает, что последовательность ДНК играет важную роль в контролируя конденсацию HP1 с ДНК. Это наблюдение ставит под сомнение общепринятое понимание связывания HP1 с ДНК, которое обычно считается относительно независимым от последовательности. Более того, в оригинальной статье авторы предположили, что появление компактизированных структур ДНК в центре молекулы ДНК происходит исключительно за счет больших конформационных флуктуаций, происходящих в середине молекулы[20]. Напротив, результаты нашего моделирования для гомополимера показывают, что образуется только один конденсат, который, как правило, располагается близко к одному из двух Связанные концы. Несомненно, необходимы дополнительные эксперименты для проверки зависимости от последовательности Связывание HP1 с ДНК и его роль в конденсации ДНК.

Биомолекулярные конденсаты в клеточной форме образуются в результате фазового разделения [32–36]. Часто образование этих Конденсаты включают в себя различные поверхности внутри клеток, такие как ДНК, микротрубочки, цитоскелет, мембраны и т.д. [37–39]. Интерфейсы, связанные с этими структурами, приводят к образованию капилляров. Силы. Капиллярные силы и их физиологические аспекты остаются относительно малоизученными. Наши Результаты показывают, что капиллярные силы, возникающие из-за гетерогенности межфазной ДНК в системе белок-ДНК Сококонденсаты могут генерировать силы до пиконьютонов. Интересно, что изменения в средстве на уровне мотива даже один кВТ из-за мутаций или паттернов метилирования может привести к двукратным изменениям в глобальном Силы. Эти силы способны формировать трёхмерную структуру генома. Более того, зависящая от последовательности коконденсация ДНК и белков потенциально может образовывать пространственно гетерогенные хромосомные территории. Будущие исследования могут пролить свет на возможность такого механизма.

В целом, конденсация ДНК, опосредованная белками, все чаще признается критически важным явлением для потенциальной организации генома и ответственным за другие клеточные процессы[8, 10]. Тем не менее, комплексное теоретическое понимание пока находится на ранней стадии. В связи с этим, Наша модельная структура позволяет раскрыть физические принципы, управляющие зависимой от последовательности совместной конденсацией белка и ДНК, а также интерпретировать экспериментальные наблюдения.

Благодарности

MG выражает благодарность за финансирование в рамках программы промежуточной стипендии DBT/Wellcome India Alliance (IA/I/21/2/505928). SC получает поддержку в рамках программы повторного поступления Ramalingaswami (BT/HRD/35/02/2006) – программы повторного поступления Департамента биотехнологии Министерства науки и технологий Правительства Индии и Департамента атомной энергии (DAE) Правительства Индии через проект Арех в Институте математических наук (IMSc) в Ченнаи.

5 методов

5.1 Методология моделирования

Для моделирования системы белок-ДНК мы используем молекулярную динамику (МД) Верле-скорости. схема, где температура поддерживается kBT = 1 с помощью термостата Ланжевена

$$p_i = F_i - \gamma_i p_i + f_i,$$
 (5)

где F_i — сила, действующая на частицу i из-за взаимодействия с другими частицами, f_i — случайная сила, генерируемая на частице i из-за растворителя, а γ_i - коэффициент трения, который поддерживается ¹ в проведенных симуляциях. равным 0,1 т. Мы устанавливаем шаг интегрирования $\delta t = 0,01$ т, где τ - единица измерения шкалы времени. Моделирование выполняется с использованием пакета ESPResSo [40]. Во-первых, мы генерируем линейную модель пружинной бусины с $N_m = 500$ в ящике с размерами 20, 20, 600 (единицы σ) и прикладываем силу 0,7 kBT / σ к первому и последнему мономеру в противоположном направлении, чтобы уменьшить R' . Мы моделируем полимер для $1,5 \times 10^6$ т, где конфигурации сохраняются через каждые 1000т. Затем мы вычисляем R' для всех конфигураций и берем конфигурации, соответствующие значениям R' , используемым в моделировании. Далее мы вводим соответствующие конфигурации _е для различных значений R' в блоке 80, 80 и 600 (единицы σ), чтобы минимизировать взаимодействие полимера со стенкой; концы полимера фиксируются, а длина персистентности полимера увеличивается с 15 до 50; мы моделируем систему на протяжении 106 временных шагов. Затем, для достижения равновесия, мы уменьшаем длину персистентности обратно до 15 и моделируем на протяжении 3×10^6 т. Для моделирования с белками минимальное время производства составляет 3×10^6 т и варьируется от моделирования к моделированию. Конфигурации сохраняются каждые 5000т для анализа, чтобы избежать смещения, вызванного временной корреляцией.

5.2 Параметры модели

В данном исследовании сила взаимодействия ДНК-белок поддерживалась в диапазоне от 0 до 4 кБт для различных моделей ДНК. Действительно, различные исследования показывают, что сродство связывания белка с ДНК у эукариот составляет от 0 до 5 кБт[7]. Мы установили силу взаимодействия между белками равной 2 кБт , что согласуется с результатами предыдущих исследований [7, 41]. Мы варьировали концентрацию белка (ρ_p) в диапазоне от 42,25 до 126,75 мкМ, что обусловлено исследованием in vitro, изучавшим коконденсацию HP1 и ДНК [20].

5.3 Кластерный анализ

Для идентификации конденсатов в наших симуляциях мы используем алгоритм пространственной кластеризации на основе плотности (DBSCAN)[42] из Scikit-learn. В этом алгоритме переменная минимальных точек установлена равной 6,0, поскольку система трёхмерная, а переменная ϵ рассчитывается из графика k-расстояний[43]. Для расчёта ϵ сначала вычисляется расстояние между частицей и её шестым ближайшим соседом для всех частиц и упорядочивается по возрастанию для построения графика k-расстояний. Затем, чтобы уменьшить шум на кривой,

Мы подгоняем полином третьей степени к окну длиной 199 с помощью фильтра Савицкого-Голея[44]. Наконец, мы вычисляем точку максимальной кривизны кривой и присваиваем соответствующее расстояние ϵ в DBSCAN[45]. Эти вычисления были выполнены для 500 конфигураций, расположенных на расстоянии более 5000 т друг от друга.

5.4 Расчет силы

Сначала мы моделируем полимер в отсутствие белков для различных нормализованных расстояний между концами для 3×10^6 т в отсутствие белков. Затем мы вычисляем длины связей для всех 499 связей и строим графики их распределений для первой и последней связей, а также других связей для различных расстояний между концами. Для этого мы используем 500 конфигураций, расположенных на расстоянии 103 т друг от друга. Как и ожидалось, распределение длин связей для связанных концов идентично остальным 497 связям (рисунок S6A). Затем для расчета силы, действующей на полимер из-за растяжения в ансамбле с постоянным расстоянием, мы используем $F_i = K(l_i - l_0)$, где K — гармонический потенциал связи, l_i — длина связи i , а l_0 — средняя длина связи, рассчитанная для полимера со свободными концами. Сила рассчитывается по всем связям полимера и усредняется для каждой конфигурации. Как и прежде, это делается для 500 конфигураций, расположенных на расстоянии 103 т друг от друга. Затем мы выбираем 500 списков сил, используя бутстреп, чтобы вычислить среднее значение и стандартное отклонение средних значений для этих списков. Чтобы выразить силу в пН, мы умножаем $k_B T / \sigma$ на 4.114×10^{-4} . Мы систематически проводим эти расчёты для различных нормализованных расстояний между концами, и сила отображается как функция R' значений (рис. S6B). Мы наблюдаем незначительное изменение сил до $R' = 0.8$, за которым следует резкий рост при $R' = 1$ (рис. S6B), что согласуется с расширением сил в ансамбле с постоянным расстоянием для полугибких червеобразных цепей[46].

Для расчета силы, обусловленной коконденсатами белок-ДНК, сначала идентифицируются конденсаты с помощью DBSCAN. Затем мы вычисляем длину связи мономеров вне конденсата и вычисляем силу, используя формулу $F_i = K(l_i - l_0)$, где K — гармонический потенциал связи, l_i — длина связи i , а l_0 — средняя длина связи, рассчитанная для полимера при заданном значении R' , упомянутом ранее. Для каждой конфигурации рассчитывается средняя сила, и аналогичный анализ проводится для последних 500 конфигураций, расположенных на расстоянии 5×10^3 т друг от друга для каждого случая. Выполняется процедура бутстрэппинга (см. выше), и для каждого случая рассчитываются среднее значение и стандартное отклонение.

5.5 Межфазное сродство

Для каждой конфигурации мономер и белки внутри конденсата обнаруживаются с помощью DBSCAN. Далее мы присваиваем мономеру оценку 0 или 1 в зависимости от его наличия или отсутствия в конденсате. Это делается для всех мономеров в соответствии с их положением вдоль контура ДНК, чтобы создать массив для каждой конфигурации. Далее мы усредняем массивы, сгенерированные из последних 500 конфигураций моделирования для каждого мономера, чтобы получить массив вероятностей. Мы считаем мономеры частью интерфейса, если соответствующие вероятности расположены в порядке возрастания (левый интерфейс) или убывания (правый интерфейс) и лежат в диапазоне от 0,1 до 1, включая последний (левый интерфейс) или первый (правый интерфейс) мономер с вероятностью 1. Обоснование этого условия заключается в том, чтобы получить все мономеры, образующие интерфейс в любой из 500 конфигураций. Далее мы вычисляем средневзвешенное сродство к интерфейсу I_a , определяемое как

$$I_a = \frac{\epsilon_1 P_1 + \epsilon_2 P_2 + \dots + \epsilon_n P_n}{P_2 + \dots + P_n} \quad (7)$$

где ϵ_i — сила взаимодействия мономера с белками (ϵ_{MP}) для мономера i , а P_i — вероятность нахождения мономера i внутри конденсата. Интерфейсы, включающие связанные мономеры, игнорируются.

5.6 Корреляционный анализ

Мы вычисляем корреляцию между силой и средним сродством к поверхности раздела, используя коэффициенты корреляции Пирсона для гетерогенных моделей ДНК II и частичной λ-ДНК. Для расчета корреляции используется среднее значение среднего сродства к поверхности раздела конденсатов.

5.7 Расчет объема

Для вычисления объема конденсата мы используем максимальное расстояние между частицами, принадлежащими конденсату (определенное с помощью DBSCAN), рассчитанное по всем трем осям. Радиус конденсата для каждой оси принимается равным половине максимального расчетного расстояния вдоль этой оси.

Объем тогда определяется как

$$V_c = \pi \times r_x \times r_y \times r_z$$

(8)

Где V_c — объём конденсата, r_x , r_y и r_z — радиусы конденсата по трём осям. В ходе расчётов мы задаём пороговое значение в 25 частиц в кластере, которые можно считать конденсатом, чтобы избежать временных агрегатов в системе. Наши результаты устойчивы к изменениям этого порогового значения.

5.8 Расчет потенциальной энергии

Мы рассчитываем потенциальную энергию для каждого мономера вдоль контура, вычисляя их расстояния до белков и других мономеров. Для расчета вклада потенциальной энергии от взаимодействий мономер-белок мы используем потенциал ЛДж, используемый для взаимодействий мономер-белок в моделировании (см. «Подробнее о модели»). Аналогично, для расчета вклада потенциальной энергии от взаимодействий мономер-мономер мы используем гармонические и угловые потенциалы в качестве вкладов от связанных взаимодействий, а потенциал ЛДж между мономерами в качестве вкладов от не связанных взаимодействий. Мы суммируем все вклады потенциальной энергии для каждого мономера, чтобы получить общую потенциальную энергию. Это делается для каждого мономера в более чем 500 конфигурациях и усредняется для каждого мономера на контуре (рисунок).

S1D).

5.9 Анализ экспериментальных данных

Для анализа экспериментально полученных изображений коконденсатов белка с ДНК в различных исследованиях мы используем imageJ[47]. Мы извлекаем профили интенсивности белка для Sox2 и HP1, рассматривая интенсивность линии вдоль ДНК. Для Sox2 мы рассматриваем изображение на рисунке 1b из [9], а для HP1 — на рисунке 4B из [20]. Пики в профилях интенсивности подсчитываются для оценки количества конденсатов. Для Dps мы оцифровываем профиль интенсивности белка с рисунка 4 [21] и строим его зависимость от положения вдоль ДНК (мкм) (рисунок 7).

Ссылки

[1] Томас Кремер и Кристоф Кремер. «Хромосомные территории, ядерная архитектура и регуляция генов в клетках млекопитающих». Nature Reviews Genetics 2.4 (2001), 292–301.

[2] Мари́ке Оуделаар и Дуглас Р. Хиггс. «Взаимосвязь между структурой и функцией генома». Nature Reviews Genetics 22.3 (2021), 154–168.

[3] Юхун Фань и др. «Истощение гистона H1 у млекопитающих изменяет глобальную структуру хроматина, но вызывает специфические изменения в регуляции генов». Cell 123.7 (2005), 1199–1212.

- [4] Йоб Деккер и Эдит Херд. «Структурное и функциональное разнообразие топологически ассоциированных доменов». Письма FEBS 589.20 (2015), 2877–2884.
- [5] Кристофер Л. Вудкок и Раджарши П. Гхош. «Структура и динамика хроматина высшего порядка». Перспективы Колд-Спринг-Харбор в биологии 2.5 (2010), a000596.
- [6] Адам Г. Ларсон и др. «Образование капель жидкости HP1α предполагает роль разделения фаз в гетерохроматине». Nature 547.7662 (2017), 236–240.
- [7] Хосе А. Морин и др. «Последовательно-зависимая поверхностная конденсация пионерского фактора транскрипции на ДНК». Nature Physics 18.3 (2022), 271–276.
- [8] Роман Ренгер и др. «Коконденсация белков с одно- и двухцепочечной ДНК». Труды Национальной академии наук 119.10 (2022), e2107871119.
- [9] Туан Нгуен и др. «Хроматин блокирует воздействие пионерского фактора транскрипции Sox2 на ДНК». Nature Communications 13.1 (2022), 3988.
- [10] Томас Куэйл и др. «Генерация силы при совместной конденсации белка и ДНК». Nature Physics 17.9 (2021), 1007–1012.
- [11] Николас А. В. Белл и др. «Измерения отдельных молекул показывают, что PARP1 конденсирует ДНК путем стабилизации петли». Science Advances 7.33 (2021), eabf3641.
- [12] Тенг Ван и др. «Хеликаза при синдроме Блума сжимает одноцепочечную ДНК в фазово-разделённые конденсаты». Angewandte Chemie International Edition 61.39 (2022), e202209463.
- [13] Эрнан Г. Гарсия и др. «Биологические последствия сильно изогнутой ДНК: другая жизнь макромолекулярной знаменитости». Биополимеры 85 (2007), 115–30.
- [14] Александр Дж. Стюарт, Шридхар Ханненхалли и Джошуа Б. Плоткин. «Почему сайты связывания факторов транскрипции имеют длину в десять нуклеотидов». Genetics 192.3 (2012), 973–985.
- [15] Ариэль А. Аптекманн и др. «Специфичность факторов транскрипции ограничивает количество ДНК-связывающих мотивов». Plos one 17.1 (2022), e0263307.
- [16] Ана И. Прието и др. «Геномный анализ связывания ДНК и регуляции генов гомологичными нуклеоид-ассоциированными белками IHF и HU в Escherichia coli K12». Nucleic acids research 40.8 (2012), 3524–3537.
- [17] Блэр Р.Г. Гордон и др. «Lsr2 — это белок, ассоциированный с нуклеоидом, который воздействует на богатые АТ последовательности и гены вирулентности Mycobacterium tuberculosis». Труды Национальной академии наук 107.11 (2010), 5154–5159.
- [18] Марк Дж. Соломон, Франсуа Штраус и Александр Варшавский. «Группа белков с высокой подвижностью млекопитающих распознаёт любой участок из шести пар оснований АТ в дуплексной ДНК». Труды Национальной академии наук 83.5 (1986), 1276–1280.
- [19] Йенс-Уве Зоммер, Хольгер Мерлиц и Хельмут Шиссель. «Конденсация с участием полимеров: механизм образования гетерохроматина и эпигенетической памяти». Macromolecules 55.11 (2022), 4841–4851.
- [20] Мадлен М. Кинен и др. «Белки HP1 уплотняют ДНК в механически и позиционно стабильные фазово-разделённые домены». elife 10 (2021), e64563.
- [21] Снеха Шаху и др. «Мостикивые контакты ДНК позволяют Dps из E. coli конденсировать ДНК». bioRxiv (2024), 2024-01.
- [22] Тао Чжао и др. «Гетерохроматиновый белок 1 связывается с нуклеосомами и ДНК in vitro». Журнал биологической химии 275.36 (2000), 28332–28338.

- [23] Пайел Гхатак и др. «Раскрытие роли Dps в организации микобактериального нуклеоида». *PloS one* 6.1 (2011), e16019.
- [24] Брайан А. Гибсон и др. «Организация хроматина путем внутреннего и регулируемого разделения фаз». *Ячейка* 179.2 (2019), 470–484.
- [25] Лян Ма и др. «Коонденсация между фактором транскрипции и коактиватором p300 модулирует кинетику транскрипционного взрыва». *Molecular Cell* 81.8 (2021), 1682–1697.
- [26] Раджеш Шарма и др. «Жидкая конденсация фактора репрограммирования KLF4 с ДНК обеспечивает механизм организации хроматина». *Nature communications* 12.1 (2021), 5579.
- [27] Сынусу Ким и Джей Шендуре. «Механизмы взаимодействия между факторами транскрипции и 3D геном». *Molecular Cell* 76.2 (2019), 306–319.
- [28] Сенг Чуан Тан и др. «Супер-усилители, фазово-разделенные конденсаты и 3D-геномная организация в области рака». *Рак* 14.12 (2022), 2866.
- [29] Омар Л. Кантидзе и Сергей В. Разин. «Слабые взаимодействия в органах хроматина высшего порядка ация». *Исследование нуклеиновых кислот* 48.9 (2020), 4614–4626.
- [30] Нагараджа Чаппиди и др. «Совместная конденсация PARP1-ДНК приводит к сборке участка репарации ДНК к предотвратить расхождение разорванных концов ДНК». *Клетка* (2024).
- [31] Шравани К. Рамисетти и др. «Исследования in vitro конденсации ДНК с помощью мостикового белка в условиях скопления молекул». *Международный журнал биологических макромолекул* 103 (2017), 845–853.
- [32] Салман Ф. Банани и др. «Биомолекулярные конденсаты: организаторы клеточной биохимии». *Nature обзоры Молекулярная клеточная биология* 18.5 (2017), 285–298.
- [33] Йоханнес Сёдинг и др. «Механизмы активной регуляции биомолекулярных конденсатов». *Тенденции в клеточной биологии* 30.1 (2020), 4–14.
- [34] Кристин Д. Китинг и Рохит В. Паппу. «Разделение фаз жидкость–жидкость: распространённый и универсальный способ организации водных растворов». *J. Phys. Chem. B* 125 (2021), 12399–12400.
- [35] Саймон Альберти, Эми Гладфелтер и Таня Миттаг. «Соображения и проблемы при изучении разделения фаз жидкость–жидкость и биомолекулярных конденсатов». *Cell* 176 (2019), 419–434.
- [36] Клиффорд П. Брэнгвинн, Питер Томпа и Рохит В. Паппу. «Физика полимеров внутриклеточных фазовые переходы». *Nature Physics* 11 (2015), 899–904.
- [37] Бернардо Гувейя и др. «Капиллярные силы, создаваемые биомолекулярными конденсатами». *Nature* 609 (2022), 255–264.
- [38] Тина Укмар-Годец, Сюзанна Вегманн и Маркус Цвекштеттер. «Биомолекулярная конденсация белка tau, ассоциированного с микротрубочками». *Семинары по клеточной биологии и биологии развития*. Том 99. Elsevier. 2020, 202–214. ..
- [39] Макс Копперс, Назмийе Озкан и Джинни Г. Фарьяс. «Сложные взаимодействия между мембраносвязанными органеллами, биомолекулярными конденсатами и цитоскелетом». *Frontiers in Cell and Developmental Biology* 8 (2020), 618733.
- [40] Ханс-Йорг Лимбах и др. «ESPResso — расширяемый пакет моделирования для исследования систем мягкой материи». *Computer Physics Communications* 174.9 (2006), 704–727.
- [41] Андреа М. Кьяриелло, Федерико Корбери и Марио Салерно. «Взаимосвязь между фазовым разделением и коммуникацией генов-энхансеров: теоретическое исследование». *Biophysical Journal* 119.4 (2020), 873–883.

- [42] Мартин Эстер и др. «Алгоритм на основе плотности для обнаружения кластеров в больших пространственных базах данных» с шумом». кдд. Т. 96. 34. 1996, 226–231.
- [43] Йорг Сандер и др. «Кластеризация на основе плотности в пространственных базах данных: алгоритм gdbscan и его применение». Анализ данных и обнаружение знаний 2 (1998), 169–194.
- [44] Авраам Савицкий и Марсель Дж. Э. Голей. «Сглаживание и дифференциация данных с помощью упрощенного Метод наименьших квадратов». Аналитическая химия 36.8 (1964), 1627–1639.
- [45] Пэн Лю, Дун Чжоу и Найцзюнь Ву. «VDBSCAN: пространственная кластеризация приложений с шумом на основе переменной плотности». Международная конференция по сервисным системам и обслуживанию 2007 г. управление. IEEE. 2007, 1–4.
- [46] Питер Цифра и Томас Блеха. «Растяжение самовзаимодействующих червеобразных макромолекул». Поли-подробнее 48.8 (2007), 2444–2452.
- [47] Кэролайн А. Шнайдер, Уэйн С. Расбанд и Кевин В. Элишейри. «NIH Image to ImageJ: 25 годы анализа изображений». Nature methods 9.7 (2012), 671–675.
- [48] Джон Д. Уикс, Дэвид Чандлер и Ханс К. Андерсен. «Роль сил отталкивания в определении равновесной структуры простых жидкостей». Журнал химической физики 54.12 (1971), 5237–5247.

Дополнительная информация

Подробности модели

Мы используем минимальную модель для изучения коконденсации белка и ДНК на концевой ДНК. Мы моделируем ДНК как крупнозернистый полугибкий полимер, состоящий из 500 мономеров, а белки (P) – как сферические частицы того же размера, что и мономеры. Каждый мономер (M) в модели соответствует 10 парам оснований (п.н.) ДНК, что подразумевает, что каждый белок связывается с 10 парами оснований ДНК. Гармонические пружины связывают соседние мономеры, тогда как гармонический угловой потенциал между последовательными связями моделирует полугибкость ДНК. Потенциал длины связи определяется выражением:

$$U(r) = kb \left(\frac{r - r_0}{r_0} \right)^2. \quad (1)$$

Здесь kb и r_0 — энергия связи и равновесная длина связи соответственно. Полагаем $kb = 100 \text{ kBT}$ и $r_0 = \sigma$. Потенциал угла связи определяется по формуле:

$$U(\theta) = -k\theta \left(\frac{\theta - \theta_0}{\theta_0} \right)^2. \quad (2)$$

Здесь $k\theta$ и θ_0 — энергия валентного угла и равновесный валентный угол. Мы принимаем $k\theta = 15 \text{ kBT}$ и $\theta_0 = 180^\circ$. $k\theta$ относится к персистентной длине (l_p) ДНК следующим образом:

$$l_p = \frac{k\theta}{kBT}. \quad (3)$$

В исследовании рассматриваются четыре различные последовательности ДНК. В каждой из последовательностей ДНК моделируется как самоизбегающий полимер с только отталкивательным взаимодействием между мономерами. Для моделирования отталкивания между мономерами мы используем потенциал Уикса-Чандлера-Андерсона (WCA)[48], который представляет собой чисто отталкивательную форму потенциала Леннарда-Джонса. Для моделирования невалентных взаимодействий между мономером и белком мы используем потенциал Л.Дж. с притягивающим хвостом.

$$U_{nb}(r_{ij}) = \begin{cases} 4\epsilon_{ij} \left(\frac{\sigma}{r_{ij}} \right)^{12} - \frac{\sigma}{r_{ij}}^6 + \frac{1}{4} r_{ij} - r_{cut} & r_{ij} \leq r_{cut} \\ 0 & r_{ij} > r_{cut} \end{cases} \quad (4)$$

где r_{ij} — расстояние между частицами i и j .

Однако детали взаимодействий мономер-белок и белок-белок различаются для разных последовательностей. Для гомогенной ДНК (рис. 2A), состоящей только из одного типа мономеров (A), мы имеем три взаимодействующие пары: AA, AP и PP. Как указано выше, мы сохраняем чисто отталкивательное взаимодействие между парами AA для моделирования объёмного исключения мономеров. Для начала мы сохраняем силу притяжения мономера к белку (ϵ_{AP}), а также силу притяжения белок-белок (ϵ_{PP}) равной 2 kBT . В таблице 1 приведены параметры невалентного взаимодействия для гомогенных систем соконденсатов ДНК-белок.

Для гетерогенной ДНК I (рисунок 2B) и гетерогенной ДНК II (рисунок 3A) мы предполагаем, что ДНК состоит из двух типов мономеров: A и B. Мономеры A притягивают белки P с силой $1,75 \text{ kBT}$, тогда как мономеры B притягивают белки P с силой $2,25 \text{ kBT}$. В обеих последовательностях мы сохраняем одинаковое количество (по 250 каждого) мономеров A и B, так что среднее сродство ДНК к белку составляет 2 kBT , что соответствует показателю для гомогенной ДНК. Однако две модели гетерогенной ДНК различаются с точки зрения гетерогенности последовательности. Гетерогенная ДНК I (рисунок 2B) имеет

ij	ϵ_{ij} (kBT)	r_{cut} (σ)
AA 1.00 AP 2.00		$2\frac{1}{6}$
PP 2.00		2.5
		2.5

Таблица 1: Параметры несвязанного взаимодействия для всех пар частиц в гомогенной системе коконденсата ДНК-белок: А представляет мономеры, а Р представляет белки.

архитектура типа -AAAAA-BBBBBB-AAAAA- с центральными 250 В-мономерами, имеющими более сильные сродство к белкам Р. Центральная область В ограничена двумя участками мономеров А, каждый из которых длиной 125, имеющий более слабое сродство к белкам Р. Гетерогенная ДНК II (рис. 3А) имеет Архитектура, подобная блок-сополимеру (-AABBAABBAABB-), где за 50 мономерами А следуют 50 мономеров В. Параметры взаимодействия для гетерогенных моделей I и II суммированы в Таблица 2.

ij	ϵ_{ij} (kBT)	r_{cut} (σ)
AA 1.0 2		$\frac{1}{6}$
AB 1.0		$2\frac{1}{6}$
BB 1.0 AP 1.75		$2\frac{1}{6}$
BP 2.25		2.5
		2.5
PP	2.0	2.5

Таблица 2: Параметры не связанного взаимодействия для всех пар частиц. А и В – мономеры, а Р представляет белки. Мономер А имеет более слабое сродство к белкам (1,75 кБт), а мономер В имеет более сильное сродство к белкам (2,25 кБт).

Наша четвертая последовательность ДНК — это последний участок ДНК λ-фага длиной 5 кб (частичная ДНК λ). Мы предполагаем, что белок Р связывается с богатым АТ участком последовательности. Поскольку один мономер в нашей модели соответствует 10 б/с, каждый мономер может иметь разное сродство к белку Р, на основе содержания АТ в базовой последовательности из 10 пар оснований. Если все 10 пар оснований представляют собой аденин или тиаминового нуклеотида, то он имеет 100% содержание АТ, тогда как если ни одна из последовательностей не является А или Т, то он содержит 0% АТ-контента. Таким образом, АТ-контент последовательности длиной 10 бит/с может иметь только 11 возможных вариантов. Для этого мы предполагаем, что наша система состоит из 11 типов мономеров (А, В, С, D, E, F, G, H, I, J, K), каждый из которых соответствует разному содержанию АТ в последовательности 10 бит/с. Используя В этом формализме мы находим содержание АТ в частичной λ ДНК как функцию длины последовательности (Рис. 5В). Далее мы присваиваем этим 11 мономерам различные значения сродства к мономеру-белку по шкале от 0 до 4 кБт. Параметры взаимодействия мономера с белком суммированы в таблице (3). сохраним $\epsilon_P = 2 \text{ kBT}$, как в предыдущих последовательностях.

Содержание АТ	(%) мономера (М)	εMP (кВТ)
0	А	0.0
10	Б	0,4
20	С	0,8
30	Д	1.2
40	Э	1.6
50	Ф	2.0
60	Г	2.4
70	ЧАС	2.8
80	В	3.2
90	Д*	3.6
100	К	4.0

Таблица 3: Содержание АТ в 10 пар оснований одиннадцати видов мономеров в частичной модели λ ДНК и их сродство взаимодействия с белком Р.

5.10 Дополнительные цифры

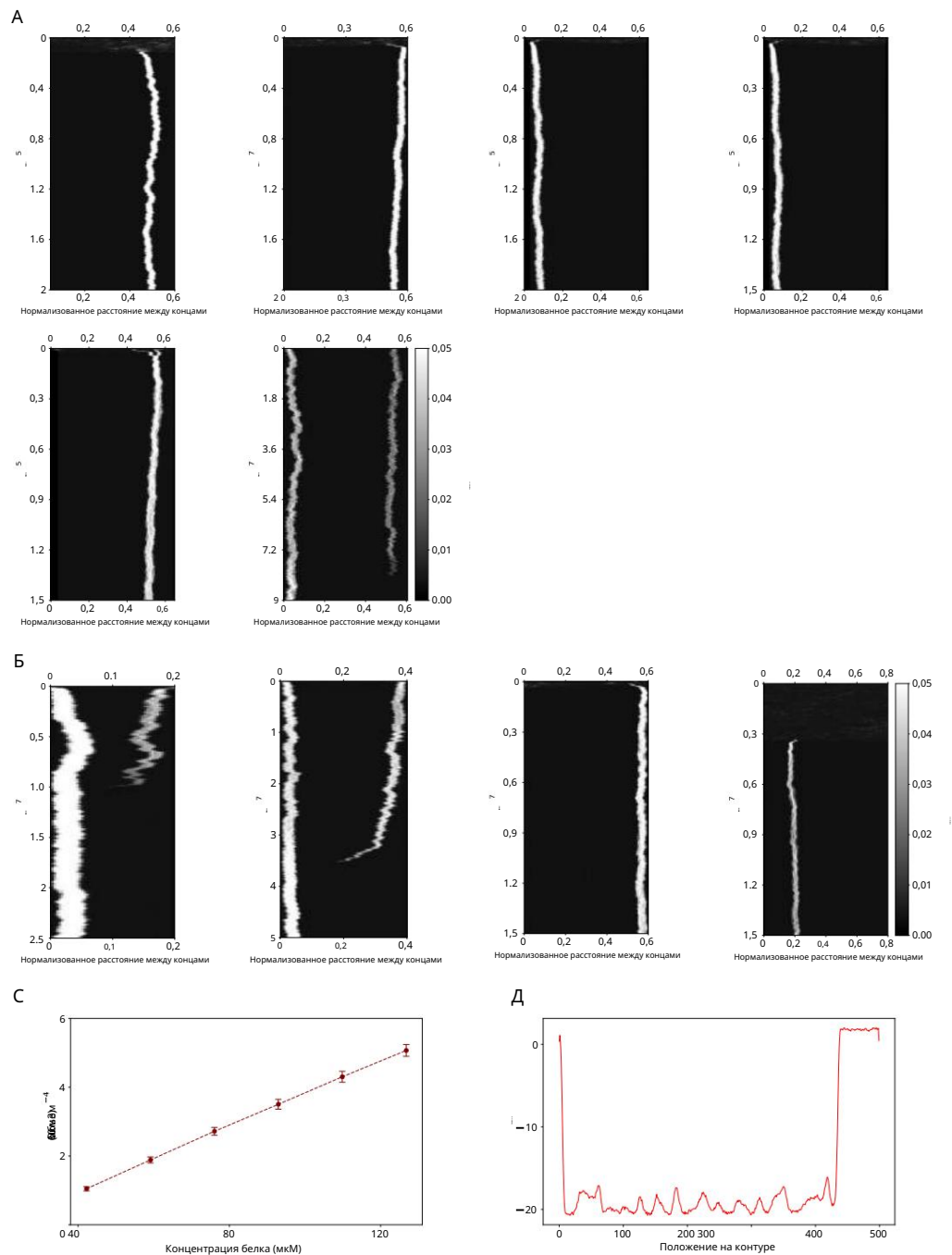


Рисунок S1: Гомогенная ДНК ко-конденсируется с белками с образованием единого конденсата вблизи одного из связанных концов в состоянии равновесия. А Типичные кимограммы.

Распределение плотности ДНК вдоль контура (ширина столбца = 2σ) как функция времени для однородной ДНК при различных $pp = 42,25, 59,15, 76,05, 92,95, 109,85$ и $126,75$ (в мкМ) (слева направо) соответственно для постоянного $R' = 0,6$. В Типичные кимограммы для однородной ДНК при различных $R' = 0,2, 0,4, 0,6$ и $0,8$ (слева направо) соответственно для $pp = 84,50$ мкМ. С Объем конденсата отображен как функция pp . D Полная потенциальная энергия показана для каждого мономера вдоль контура для однородной ДНК при $R' = 0,2$ и $pp = 84,50$ мкМ (см. Методы расчета потенциалов).

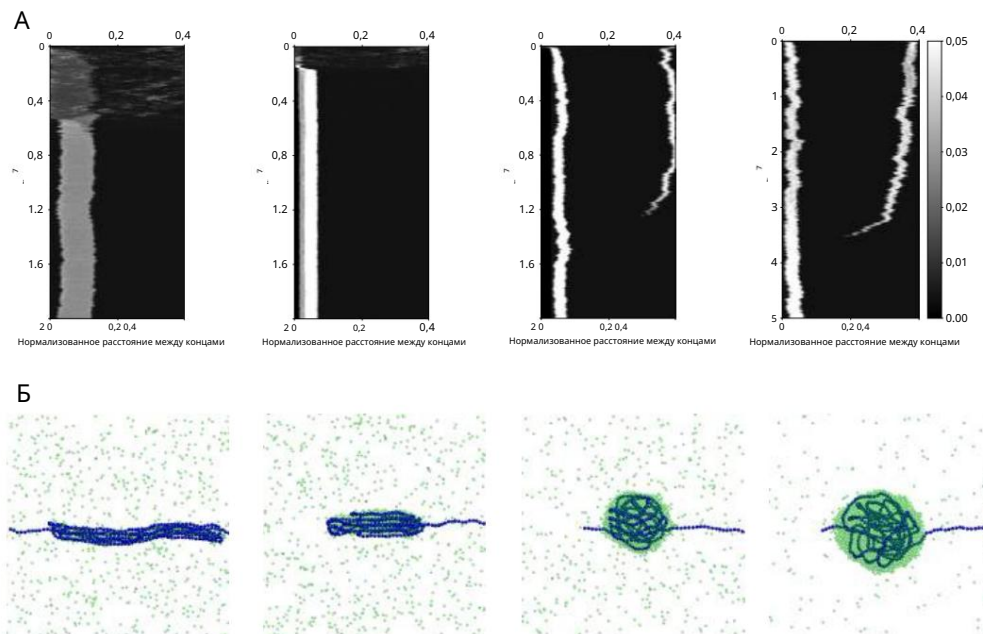


Рисунок S2: Коконденсация с гомогенной ДНК приводит к образованию одного конденсата при различной силе белок-белкового взаимодействия. А. Типичные кимограммы положения конденсатов вдоль контура (ширина столбца = 2σ) в зависимости от времени для гомогенной ДНК ($R'e = 0,5$, $pp = 85,50$ мкМ) при ϵP P ; 1, 1,5, 1,8 и 2 кВТ (слева направо) соответственно. В. Снимки (i), (ii), (iii) и (iv) представляют конденсаты при ϵP P = 1, 1,5, 1,8 и 2 кВТ (слева направо) соответственно. Гомогенные мономеры ДНК показаны синим цветом, а белки — зелёным.

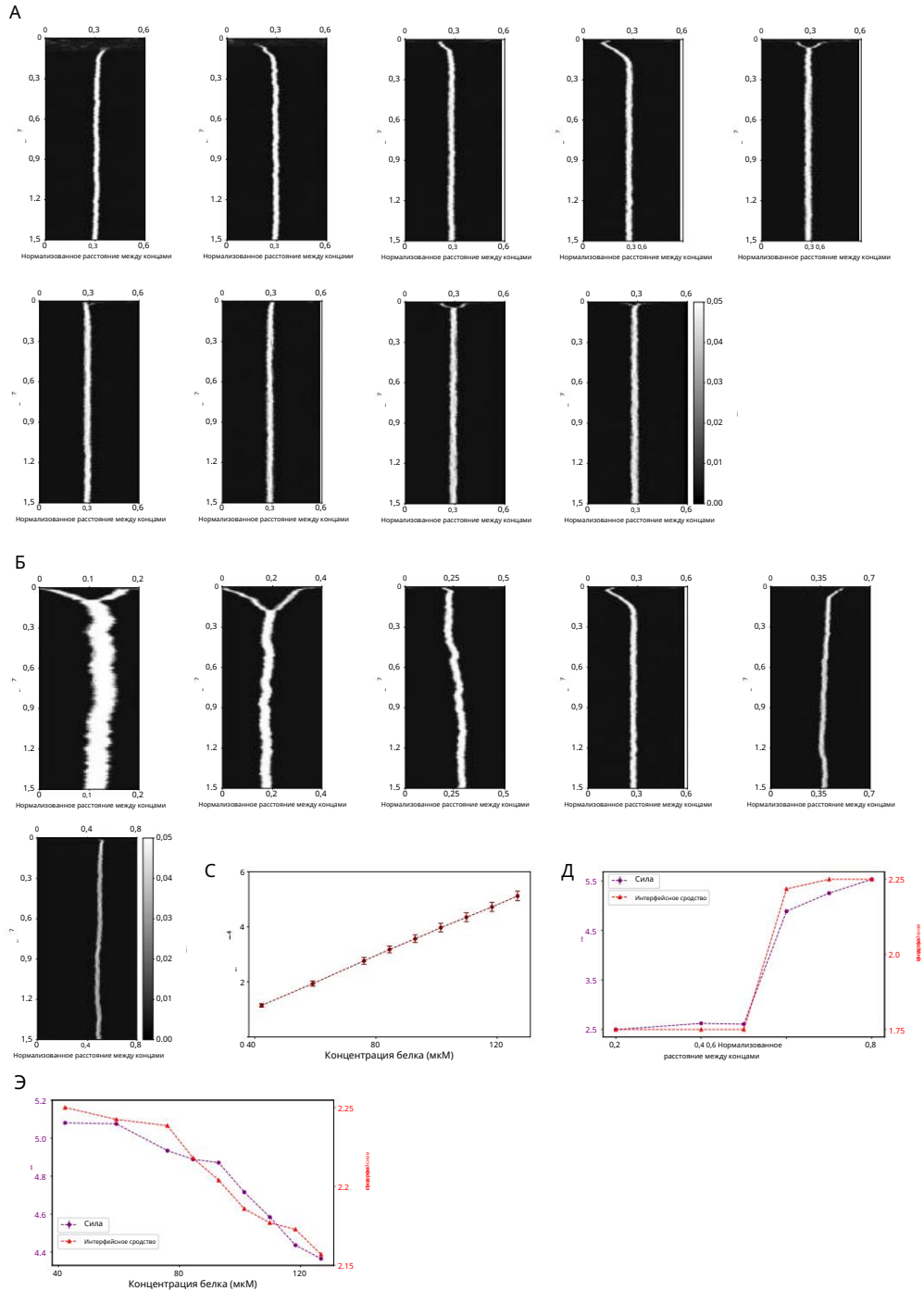


Рисунок S3: Положения конденсата определяются гетерогенностью последовательности ДНК. А. Типичные кимограммы распределения плотности ДНК вдоль контура (ширина столбца = 2σ) в зависимости от времени для гетерогенной ДНК I при различных значениях $pp = 42,25, 59,15, 76,05, 92,95, 101,40, 109,85, 118,30$ и $126,75$ (в мкМ) (слева направо) соответственно для $R' = 0,6$. В. Типичные кимограммы для гетерогенной ДНК при различных значениях $R' = 0,2, 0,4, 0,5, 0,6, 0,7$ и $0,8$ соответственно для $pp = 84,50$ мкМ. С Объем конденсата представлен в виде функции pp для гетерогенной ДНК I (см. Методы расчета объема). D Сила (фиолетовый) и среднее сродство к границе раздела (красный) в виде функции R' при $pp = 84,50$ мкМ. E Сила (фиолетовый) и среднее сродство к границе раздела (красный) представлены в виде функции pp при $R' = 0,6$ для гетерогенной ДНК I. F Общая потенциальная энергия для каждого мономера представлена вдоль контура для гетерогенной ДНК I при $R' = 0,2$ и $pp = 84,50$ мкМ.

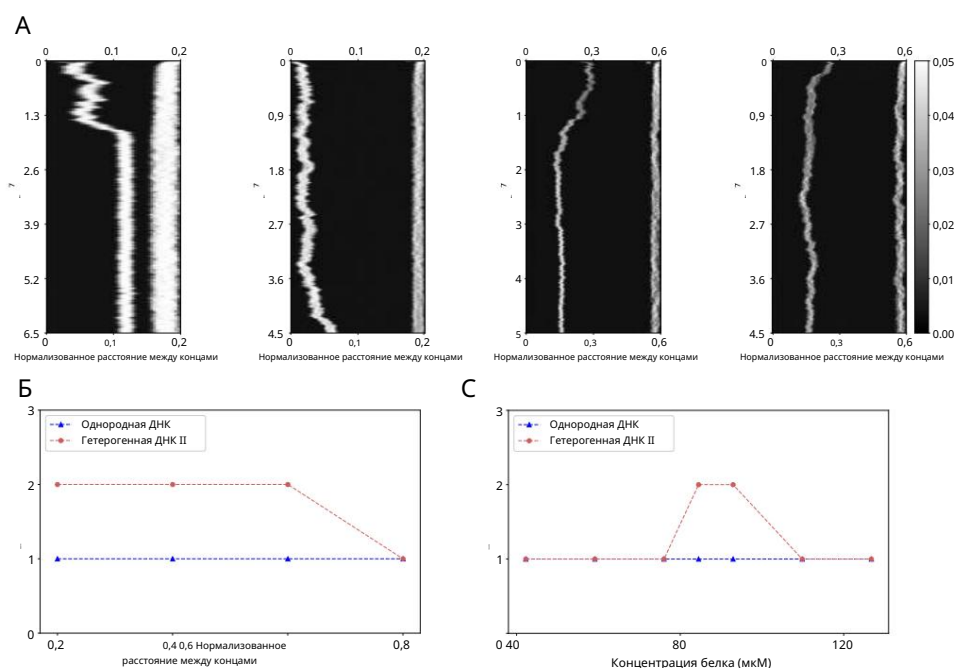


Рисунок S4: Гетерогенность ДНК приводит к сосуществованию нескольких коконденсатов в состоянии равновесия. А. Типичные кимограммы распределения плотности ДНК вдоль контура (ширина столбца = 2σ) как функции времени для гетерогенной ДНК II для 4 различных случаев, приводящих к образованию нескольких конденсатов в состоянии равновесия: случай 1: $R' = 0,2$ и $pp = 84,50$ мкМ, случай 2: $R' = 0,4$ и $pp = 84,50$ мкМ, случай 3: $R' = 0,6$ и $pp = 84,50$ мкМ, случай 4: $R' = 0,6$ и $pp = 92,95$ мкМ (слева направо). В. Число конденсатов в состоянии равновесия представлено в виде функции R' при $pp = 84,50$ мкМ для гомогенной ДНК (показано синим цветом) и гетерогенной ДНК II (показано бордовым цветом). С. Число конденсатов в состоянии равновесия представлено в виде функции pp при $R' = 0,6$ для гомогенной ДНК (показано синим цветом) и гетерогенной ДНК II (показано бордовым цветом).

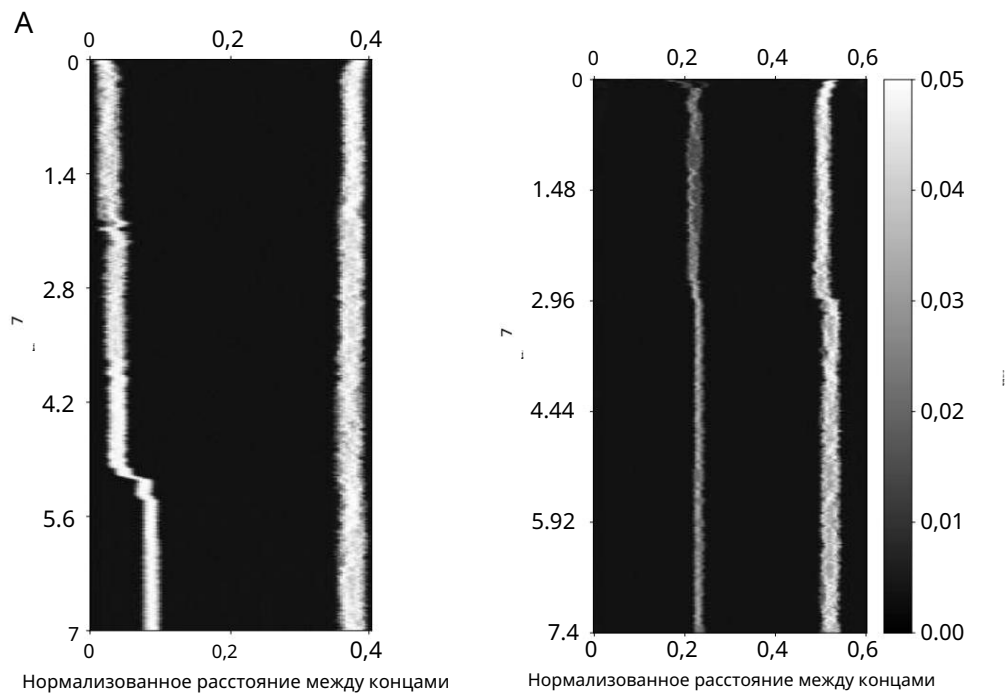
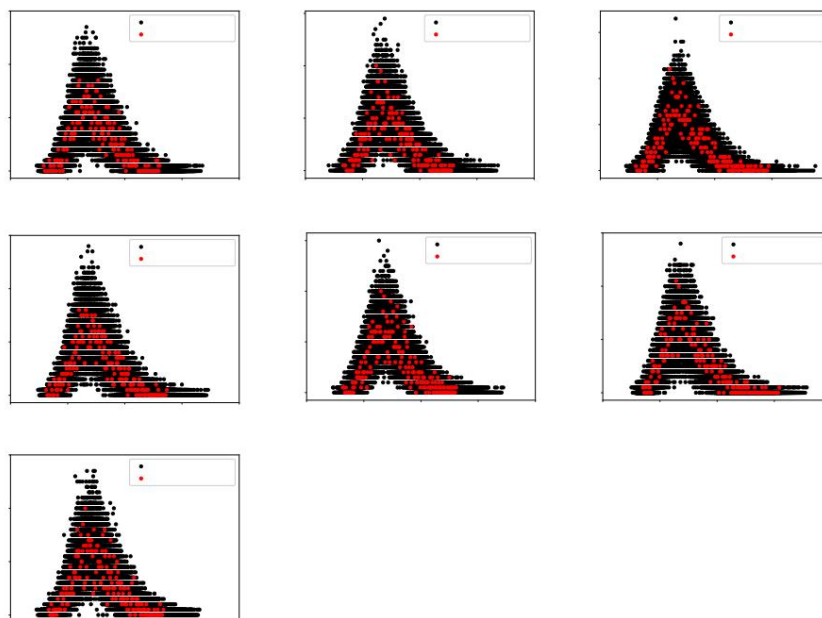


Рисунок S5: Частичная ДНК λ может поддерживать множественные конденсаты. А. Типичные кимограммы распределения плотности ДНК вдоль контура (ширина столбца = 2σ) как функции времени для гетерогенной ДНК III (частичная ДНК λ) для двух случаев, приводящих к множественным конденсатам в равновесии: случай 1: $R' e = 0,4$ и $pp = 84,50$ мкМ, случай 2: $R' e = 0,6$ и $pp = 92,95$ мкМ (слева направо).

А



Б

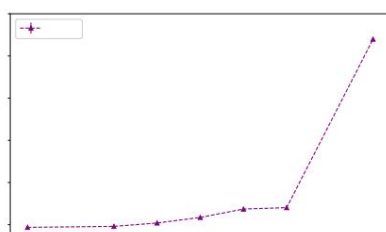


Рисунок S6: Распределение длин связей и кривая силы растяжения для полугибкого полимера. А Распределение длин связей для первой и последней связи (показано красным) и других связей (показано серым цветом) для полимера был построен график для различных нормализованных расстояний между концами $R'_e = 0,2, 0,4, 0,5, 0,6, 0,7, 0,8$ и 1 в отсутствие белков. На графике представлена сила V , рассчитанная по всем связям. как функция нормализованного расстояния между концами для полугибкого полимера (см. Методы определения силы расчеты).