

МОЛЕКУЛЯРНАЯ БИОЛОГИЯ

УДК 577.32

**СОВМЕСТНОЕ ВЛИЯНИЕ АМИНОКИСЛОТНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ
ГИСТОНА H1 И НУКЛЕОТИДНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ ДНК
НА СТРУКТУРУ ХРОМАТОСОМЫ: АНАЛИЗ МЕТОДАМИ
МОЛЕКУЛЯРНОГО МОДЕЛИРОВАНИЯ****Т.К. Горковец, Г.А. Армеев, К.В. Шайтан, А.К. Шайтан****Кафедра биоинженерии, биологический факультет, Московский государственный университет имени М.В. Ломоносова,
Россия, 119234, г. Москва, Ленинские горы, д. 1, стр. 12***e-mail: alex@intbio.org*

Хроматосома, состоящая из нуклеосомного ядра, линкерных участков ДНК и линкерного гистона (ЛГ), является важным структурным элементом хроматина. Существует два экспериментально подтвержденных типа связывания ЛГ с нуклеосомой и линкерной ДНК, которые различаются своей геометрией – связывание “на диаде” и “вне диады”. Показано, что на тип связывания гистона и конформацию хроматосомы влияет аминокислотная последовательность ЛГ. Однако при связывании ЛГ изменяется в том числе и геометрия линкерной ДНК. Взаимовлияние этих факторов и молекулярные основы, определяющие тип связывания ЛГ и нуклеосомы, остаются неясными. В данной статье мы применили методы молекулярного моделирования, включая моделирование по гомологии, анализ атом-атомных взаимодействий и расчет энергии деформации ДНК, для изучения совместного влияния аминокислотной последовательности ЛГ и нуклеотидной последовательности ДНК на конфигурацию хроматосомы. Были проанализированы известные кристаллические и ЯМР-структуры хроматосомы на предмет атом-атомных взаимодействий ЛГ и ДНК, а также энергии деформации ДНК в этих структурах для различных последовательностей ДНК. Для различных вариантов ЛГ H1 анализ проводился с использованием методов моделирования по гомологии. Были обнаружены зависящие от последовательности различия в энергии изгиба линкерной ДНК для двух различных конформаций хроматосомы, а также предложены нуклеотидные последовательности, предпочтительные для этих структур. В результате анализа было показано, что нуклеотидная последовательность ДНК наряду с аминокислотной последовательностью ЛГ оказывает влияние на тип связывания с нуклеосомой. Сформулированы гипотезы для экспериментальной проверки, согласно которым тип связывания ЛГ может меняться при изменении нуклеотидной последовательности ДНК.

Ключевые слова: *хроматосома, нуклеосома, хроматин, линкерный гистон, линкерная ДНК, гибкость ДНК, моделирование по гомологии*

Структурными единицами хроматина являются нуклеосомы – ДНК-белковые комплексы, содержащие гистоны. Ядро нуклеосомы образовано октамером гистонов H2A, H2B, H3 и H4, вокруг которого располагается двойная спираль ДНК длиной около 146 пар оснований [1]. Нуклеосома обладает осью псевдосимметрии второго порядка, называемой также диадной осью (рисунок, А). Эта ось проходит через центр нуклеосомной ДНК, называемый диадой.

Следующим уровнем компактизации является хроматосома [2], образованная при связывании линкерного гистона (ЛГ) H1 и нуклеосомы, включающей линкерные участки ДНК, выходящие за пределы ядра нуклеосомы (рисунок, А). Таким образом, в хроматосому входит октамер гистонов H2A, H2B, H3 и H4, нуклеосомная ДНК, ЛГ и линкерные участки ДНК.

ЛГ представлены во многих эукариотических организмах несколькими вариантами, различающимися длиной аминокислотной последовательности и ее составом [3]. В различных клетках и тканях могут экспрессироваться различные варианты ЛГ. Как правило, у наиболее просто устроенных организмов встречается всего один вариант ЛГ, тогда как, например, у человека известно 11 вариантов [4], некоторые из которых экспрессируются только в половых клетках. Также стоит отметить, что присутствие того или иного варианта ЛГ зависит не только от типа клетки, но и от стадии клеточного цикла, в которой она находится. В дальнейшем под термином “ЛГ” мы будем понимать все линкерные гистоны, не принимая во внимание их принадлежность к конкретному виду или типу.

Большая часть ЛГ содержит около 200 аминокислотных остатков. ЛГ состоят из трех доменов: короткий и неупорядоченный N-конец, за которым

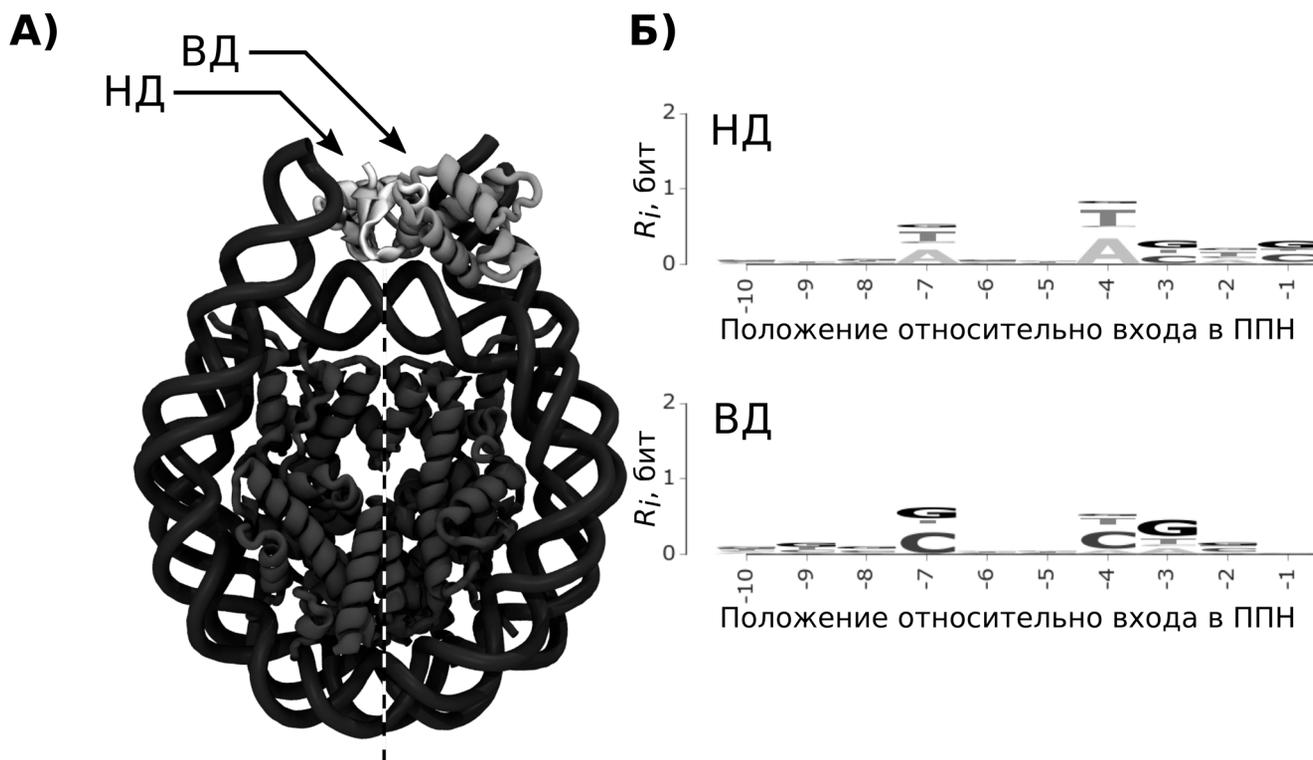


Рисунок. А – Внешний вид структуры хроматосомы. Белок и ДНК показаны в виде отображения вторичной структуры. Линкерные гистоны выделены оттенками серого. Диадная ось показана пунктирной линией. Б – Визуализация частоты встречаемости нуклеотидов для последовательностей, склонных к формированию структуры “на диаде” (НД, “on-dyad”) (сверху) и “вне диады” (ВД, “off-dyad”) (снизу). Визуализация выполнена в виде логотипов последовательностей

идет глобулярный домен, образованный 70–80 аминокислотными остатками и имеющий консервативную третичную структуру, и длинный С-концевой домен из примерно 100 аминокислотных остатков. С-конец является неорганизованным и содержит много остатков лизина. Согласно экспериментальным данным [5], глобулярный домен ЛГ Н1 связывается с нуклеосомой наравне с полноразмерным ЛГ Н1. Ввиду большой степени неупорядоченности концевых доменов отсутствуют кристаллические структуры полноразмерного ЛГ, но существуют кристаллические структуры глобулярного домена [6]. Известно, что ЛГ Н1 содержит ряд аминокислотных остатков, играющих важную роль в формировании контактов с нуклеосомой [7]. Эти остатки выделены в два сайта связывания: первый – Н25, R47, K69, K73, R74, K85, образующий контакты с большой бороздкой ДНК около диады нуклеосомы, и второй – R42, R94, K97, связывающийся с линкерной ДНК.

В данный момент областью для обсуждений является конфигурация связывания ЛГ с нуклеосомой. Ранее существовали различные модели такого связывания, некоторые из них даже предполагали встраивание ЛГ между витками нуклеосомальной ДНК и октамером гистонов [8]. Подобные модели не нашли впоследствии экспериментального подтверждения [5, 9] в отличие от моделей, в которых предполагается связывание ЛГ в области диады нуклеосомы и линкерных участков ДНК.

Имеющиеся на текущий момент экспериментальные данные [10–13] предполагают два возможных расположения ЛГ в хроматосоме, условно называемые как “на диаде” (НД, “on-dyad”) и “вне диады” (ВД, “off-dyad”). На данный момент получена кристаллическая структура хроматосомы с глобулярным доменом ЛГ Н5 *Gallus gallus* (Н5 является историческим названием ЛГ Н1 у *G. gallus*) в конфигурации НД, а также модель структуры хроматосомы с глобулярным доменом ЛГ Н1 *Drosophila melanogaster* в конфигурации ВД [12]. Важно отметить, что структура хроматосомы в конфигурации ВД была построена с помощью методов молекулярного докинга на основании структуры тетра-нуклеосомы (pdb-код 1zbb) с использованием данных ядерного магнитного резонанса (ЯМР).

Глобулярный домен гистона Н5 *G. gallus* продемонстрировал связывание НД с нуклеосомой, тогда как глобулярный домен Н1 *D. melanogaster* – связывание ВД. Анализ их последовательностей показал отличия, вероятно, играющие роль в предпочтительности того или иного типа связывания. Так, глобулярный домен Н5 *G. gallus* содержит положительно заряженные аминокислотные остатки (R47, K55, R74, K97) в позициях, соответствующих нейтрально заряженным аминокислотным остаткам в глобулярном домене Н1 *D. melanogaster* (L68, T76, S96, A119), и нейтрально заряженные аминокислотные остатки (Q51, V80, V87) в позициях, соответствующих положительно заряженным остат-

ками H1 *D. melanogaster* (K72, K102, K109) [12, 13]. Несмотря на то, что некоторые из ключевых остатков не образуют прямых контактов с нуклеосомой, при их замене в глобулярном домене H5 на аминокислотные остатки, характерные для H1, ЛГ показал изменение типа связывания с НД на ВД [13].

Другим фактором, который предположительно может оказывать влияние на конфигурацию хроматосомы, является нуклеотидная последовательность ДНК – как линкерной, так и входящей в состав нуклеосомы. По данным исследования позиционирования нуклеосом выдвигались предположения, что ЛГ предпочтительно связывается с АТ-богатыми участками ДНК [14].

Несмотря на то, что в полученной кристаллической структуре хроматосомы ЛГ H1 показал связывания НД, по данным, полученным методами молекулярного моделирования [15, 16], глобулярный домен ЛГ H5 может также демонстрировать тип связывания ВД. Одним из объяснений может служить тот факт, что нуклеотидные последовательности ДНК, используемые при моделировании, отличались от нуклеотидной последовательности, для которой была разрешена кристаллическая структура НД конформации комплекса нуклеосомы с глобулярным доменом H5 [12].

В описанных выше работах рассматривается связывание ЛГ с одиночной нуклеосомой *in vitro*, тогда как в клетке представлены структуры более высокого уровня компактизации – фибриллы диаметром 30 нм, представляющие собой цепочку из нуклеосом и ЛГ, связанных с линкерной ДНК. Согласно последним данным криоэлектронной микроскопии, в структуре фибриллы диаметром 30 нм, образованной 12 нуклеосомами, ЛГ демонстрирует связывание ВД, что может быть обусловлено геометрией линкерной ДНК в структуре фибриллы [10, 17]. Таким образом, нельзя исключить влияние ДНК, вызванное “узнаванием” ЛГ специфической формы ДНК при связывании [18].

Исходя из вышесказанного, весьма вероятно, что на тип связывания ЛГ с нуклеосомой и линкерной ДНК влияет не только аминокислотная последовательность самого гистона, но и нуклеотидная последовательность ДНК, а также структура хроматиновой фибриллы. Известно также, что при образовании комплекса белок-ДНК на силу и тип связывания оказывают влияние как прямые взаимодействия белка с парами оснований и сахарофосфатным остовом, так и геометрия ДНК за счет энергии ее деформации [18]. Однако, несмотря на наличие структур хроматосомы в различных конфигурациях, совместное влияние этих факторов ранее не изучалось.

В данной работе методами молекулярного моделирования, а именно анализа атом-атомных взаимодействий, моделирования по гомологии и оценки энергии деформации ДНК, было проведено изучение совместного влияния аминокислотной последовательности ЛГ и нуклеотидной последователь-

ности ДНК на тип конфигурации хроматосомы. Проведенный анализ позволил предсказать наиболее предпочтительные для того или иного типа связывания нуклеотидные последовательности линкерной ДНК и сформулировать гипотезы для экспериментальной проверки.

Материалы и методы

Анализ аминокислотных последовательностей различных гистоновых вариантов. В работе были использованы аминокислотные последовательности ЛГ H1, размещенные в базе данных HistoneDB [4]. Выравнивание доступных аминокислотных последовательностей проводилось программой Muscle [19]. Для дальнейшего анализа последовательностей использовался набор инструментов Biopython [20].

Моделирование по гомологии. Моделирование по гомологии осуществлялось в программе MODELLER [21, 22]. В качестве белков-шаблонов использовались кристаллическая структура хроматосомы в конфигурации НД с pdb-кодом 4qlc и модель хроматосомы в конфигурации ВД, полученная на основании данных ЯМР [12]. Для каждого исследуемого гистона с помощью моделирования по гомологии было построено по 10 моделей, из которых на основании DOPEScore были отобраны наилучшие структуры для анализа контактов. Дополнительная оптимизация структур, полученных методом моделирования по гомологии, не проводилась. Анализ контактов проводился с помощью программного пакета Chimera [23].

Расчет энергии деформации ДНК. Расчет зависимости деформационной энергии линкерных участков ДНК от ее последовательности производили в пространстве обобщенных переменных Tilt, Roll, Twist, Shift, Slide, Rise; переход от атомных координат к обобщенным осуществляли при помощи программы 3DNA [24]. Вычисление деформационной энергии участков ДНК производили в соответствии с формулой:

$$E = E_0 + \frac{1}{2} \sum_{i=1}^6 \sum_{j=1}^6 f_{ij} \Delta\theta_i \Delta\theta_j,$$

где E_0 – минимальная деформационная энергия (не зависит от конформации, принята за 0), суммирование осуществляется по обобщенным переменным для каждых соседних нуклеотидных пар, f_{ij} – коэффициенты жесткости для отклонения обобщенных переменных от равновесной конфигурации. Для расчетов применяли набор упругих коэффициентов и средних значений для нуклеотидных пар, описанный в работе [25].

Деформационную энергию определяли для всех возможных последовательностей ДНК каждого из четырех участков ДНК (по два на каждую модель связывания ЛГ) длиной по 12 н.п. каждый. Участок ДНК, рядом с которым располагается ЛГ в модели ВД, был назван входом в последовательность, позиционирующую нуклеосомы (ППН), противополо-

ложный участок был назван выходом из ППН. Для каждого варианта последовательности была рассчитана разница в деформационной энергии ДНК между моделями НД и ВД. Полученное распределение разниц энергий визуализировали при помощи графического представления консервативности нуклеотидов — логотипа последовательности (sequence logo) для 5% структур с наибольшей (положительной) разницей энергий (предпочтительные для модели НД) и 5% структур с наименьшей (отрицательной) разницей энергий (предпочтительные для модели ВД).

Результаты и обсуждение

Настоящее исследование состояло из нескольких взаимодополняющих компонент. На основе анализа экспериментальных данных были выделены ключевые остатки ЛГ, способствующие тому или иному типу связывания, на основании которых была построена классификационная модель, которая была применена к ЛГ человека: Н1.1, Н1.2, Н1.3, Н1.4, Н1.5, Н1.0 (Н1°), TS Н1.6 (Н1Т), TS Н1.7 (Н1Т2, HANP1), ОО Н1.8 (Н1_{oo}), TS Н1.9 (HILS1), Н1.10, Н1.11, а также к ЛГ Н1 *Xenopus laevis*. В данной модели вероятность того или иного типа связывания для произвольного ЛГ оценивалась следующим образом: (1) проводилось выравнивание последовательности исследуемого ЛГ с последовательностями ЛГ с известным типом связывания, (2) в ключевых для связывания позициях проводился подсчет количества совпадающих типов аминокислотных остатков. Данная модель является предсказательной и не позволяет в полной мере оценить точность предложенной классификационной модели, пока не будет накоплено дополнительных экспериментальных данных.

Для экспериментально определенных структур хроматосомы и структур с различными вариантами ЛГ Н1, построенных с помощью моделирования по гомологии, были проанализированы контакты между ЛГ и ДНК. Был проведен оригинальный анализ конформации ДНК в различных структурах хроматосомы и анализ пространства последовательностей линкерных ДНК с точки зрения энергии их изгиба и контактов с ЛГ.

Определение ключевых остатков ЛГ, влияющих на тип связывания. Аминокислотная последовательность ЛГ Н1 имеет ряд ключевых позиций, мутации в которых приводят к изменению типа связывания [12]. Таким образом, на основании аминокислотных остатков, расположенных в этих позициях, для ЛГ может быть определен предполагаемый тип связывания, что позволяет предложить классификацию ЛГ по типу их связывания с нуклеосомой на основании аминокислотных последовательностей.

В экспериментально определенной структуре хроматосомы в конфигурации НД и ВД присутствуют контакты непосредственно с парами оснований и линкерной ДНК, а также нуклеосомной

ДНК. Данные контакты представлены как водородными связями, так и Ван-дер-Ваальсовыми взаимодействиями.

Некоторые из аминокислотных остатков, расположенных в ключевых позициях (например, аминокислотные остатки, соответствующие K55, Q51, V87 ЛГ Н5) не образуют прямых контактов с ДНК в известных структурах. Однако, несмотря на отсутствие контактов с ДНК, данные остатки предположительно могут взаимодействовать с ДНК посредством электростатического потенциала, поскольку в гистонах Н1 и Н5 они имеют различные заряды.

Классификация ЛГ на основании наличия ключевых остатков. На основании предложенной выше классификации для 3 ЛГ человека предполагаемый тип связывания был определен как НД, для 7 гистонов человека тип связывания не может быть определен однозначно, поскольку они демонстрируют сходство в ключевых позициях и с гистонами Н5 *G. gallus* (тип связывания НД), и с гистонами Н1 *D. melanogaster* (тип связывания ВД), для 1 гистона человека предполагаемый тип связывания может быть определен однозначно как ВД. Для ЛГ *X. laevis* тип связывания был определен как НД.

Энергия деформации ДНК в различных конфигурациях хроматосомы. Для определения зависимости типа связывания ЛГ от нуклеотидной последовательности линкерной ДНК была рассчитана разница деформационных энергий линкерных участков ДНК между моделями НД и ВД для всех возможных вариантов последовательностей. Также была произведена Z-оценка положения последовательности из оригинальных моделей в распределении рассчитанных деформационных энергий ($Z = 0,75$), исходя из которой оригинальная последовательность, использованная для экспериментального получения структур, склонна к формированию структуры НД.

Как видно из рисунка (Б), большинство позиций в последовательностях нуклеотидов линкерных участков ДНК не значимы для определения геометрии ДНК в рамках моделей НД и ВД, кроме нуклеотидов в позициях -7, -4 и -3. В последовательностях, предпочтительных для модели НД, в этих позициях располагаются А/Т, А/Т, G/С, во время как для модели ВД более предпочтительны С/Г, С/Т, G/Т. Преобладание в линкерной ДНК тимидинов было показано ранее Цуи и Журкин [14].

Также по разнице в деформационных энергиях были найдены предпочтительные (оптимальные) последовательности линкерной ДНК (для которых энергия изгиба максимально способствует той или иной конформации) для моделей НД (CCGTCCCCGTC-ППН-ACGCCGCGG) и ВД (GACGCCCGAC-ППН-GTGATGCTGC).

Анализ совместного влияния аминокислотной последовательности ЛГ Н1 и нуклеотидной последовательности ДНК. На основании предложенной ранее классификации для различных вариантов

ЛГ H1 с помощью моделирования по гомологии были построены структурные модели хроматосом в конфигурациях НД и ВД. В полученных моделях был проведен анализ контактов между ЛГ H1 и ДНК.

И в тех, и в других моделях присутствуют контакты как между аминокислотными остатками ЛГ H1 и сахаро-фосфатным остовом ДНК, так и непосредственно с азотистыми основаниями, что говорит в пользу предположений о влиянии нуклеотидной последовательности ДНК на тип связывания ЛГ и нуклеосомы за счет эффекта прямого считывания (direct readout) последовательности ДНК белком. Также количество этих контактов изменяется в зависимости от варианта ЛГ, для которого построена модель. Контакты между ЛГ и азотистыми основаниями ДНК могут быть обеспечены за счет образования как водородных связей, так и Ван-дер-Ваальсовых взаимодействий.

Также для ЛГ H1 *D. melanogaster* и H5 *G. gallus* с помощью моделирования по гомологии были построены модели хроматосом в конфигурациях, противоположных экспериментальным: для ЛГ H1 использовалась конфигурация НД, для ЛГ H5 – конфигурация ВД. Данные модели показали уменьшение количества контактов между ЛГ и ДНК, что согласуется с определенными экспериментальными типами связывания.

На основании расчета энергии деформации ДНК были построены модели хроматосомы в конфигурации НД и ВД с предпочтительными (оптимальными) для них последовательностями, определенными по разнице в энергии деформации ДНК, для которых также были проанализированы контакты между ЛГ и ДНК. Анализ показал, что при замене исходной нуклеотидной последовательности в экспериментальных структурах ВН и НД на оптимальные определенные в нашей работе последовательности, способствующие исходному типу связывания, количество контактов ЛГ H1 с парами оснований либо не меняется (для связывания ВД – сохраняется три контакта), либо увеличивается (для связывания НД – увеличивается с двух контактов до трех). В то же время, перекрестное сравнение влияния оптимальных последовательностей ДНК

в НД и ВД структурах показало, что структуры с оптимальными последовательностями ДНК, соответствующими своему типу связывания, имеют такое же или большее число контактов между ЛГ и парами оснований, чем структуры, в которых используется последовательность ДНК, соответствующая альтернативному типу связывания. Таким образом, оптимальные последовательности линкерной ДНК для различных конформаций хроматосомы, предложенные выше, могут достигать своей избирательности как за счет механизмов непрямого считывания (indirect readout) последовательности ДНК белком, так и прямого считывания (direct readout) – взаимодействия белка с парами оснований.

На основании вышесказанного можно сделать предположение, что нуклеотидная последовательность, а также геометрия и жесткость ДНК при изгибе являются важным фактором, определяющим тип конфигурации хроматосомы. Это предположение на данный момент согласуется косвенным образом с экспериментальными данными [10, 17]. Вероятно, аминокислотная последовательность ЛГ не является первостепенным фактором, определяющим тип связывания ЛГ H1 с нуклеосомой.

Таким образом, можно сформулировать гипотезу, согласно которой один и тот же ЛГ, в зависимости от нуклеотидной последовательности ДНК, а также от геометрии линкерной ДНК, может демонстрировать различные типы связывания. Такие пары последовательностей ДНК предложены в данной работе. В дальнейшем эта гипотеза может быть проверена путем оценки расстояний между нуклеотидами, методом измерения эффективности Фёрстеровского переноса энергии (spFRET), с использованием различных нуклеотидных последовательностей для каждого ЛГ.

Авторы выражают благодарность Я. Бай (Y. Bai) за предоставленную модель хроматосомы. Работа выполнена с использованием оборудования Центра коллективного пользования сверхвысокопроизводительными вычислительными ресурсами МГУ имени М.В. Ломоносова [26] при финансовой поддержке Российского научного фонда (проект №14-24-00031, соглашение №14-24-00031-п).

СПИСОК ЛИТЕРАТУРЫ

1. Luger K., Mäder A.W., Richmond R.K., Sargent D.F., Richmond T.J. Crystal structure of the nucleosome core particle at 2.8 Å resolution // *Nature*. 1997. Vol. 389. N 6648. P. 251–260.
2. Simpson R.T. Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones // *Biochemistry*. 1978. Vol. 17. N 25. P. 5524–5531.
3. Lyubitelev A.V., Nikitin D.V., Shaytan A.K., Studitsky V.M., Kirpichnikov M.P. Structure and functions of linker histones // *Biochemistry (Mosc.)*. 2016. Vol. 81. N 3. P. 213–223.
4. El Kennani S., Adrait A., Shaytan A.K., Khochbin S., Bruley C., Panchenko A.R., Landsman D., Pflieger D., Govin J. MS_HistoneDB, a manually curated resource for proteomic analysis of human and mouse histones // *Epigenetics Chromatin*. 2017. Vol. 10:2.
5. Syed S.H., Goutte-Gattat D., Becker N., Meyer S., Shukla M.S., Hayes J.J., Everaers R., Angelov D., Bednar J., Dimitrov S. Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome // *Proc. Natl. Acad. Sci. U. S. A.* 2010. Vol. 107. N 21. P. 9620–9625.
6. Ramakrishnan V., Finch J.T., Graziano V., Lee P.L., Sweet R.M. Crystal structure of globular domain of histone H5 and its implications for nucleosome binding // *Nature*. 1993. Vol. 362. N 6417. P. 219–223.
7. Brown D.T., Izard T., Misteli T. Mapping the interaction surface of linker histone H1(0) with the nucleosome

of native chromatin *in vivo* // Nat. Struct. Mol. Biol. 2006. Vol. 13. N 3. P. 250–255.

8. Pruss D., Bartholomew B., Persinger J., Hayes J., Arents G., Moudrianakis E.N., Wolffe A.P. An asymmetric model for the nucleosome: a binding site for linker histones inside the DNA gyres // Science. 1996. Vol. 274. N 5287. P. 614–617.

9. Zhou Y.B., Gerchman S.E., Ramakrishnan V., Travers A., Muyldermans S. Position and orientation of the globular domain of linker histone H5 on the nucleosome // Nature. 1998. Vol. 395. N 6700. P. 402–405.

10. Bednar J., Garcia-Saez I., Boopathi R. et al. Structure and dynamics of a 197 bp nucleosome in complex with linker histone H1 // Mol. Cell. 2017. Vol. 66. N 3. P. 384–397

11. Zhou B.-R., Feng H., Kato H., Dai L., Yang Y., Zhou Y., Bai Y. Structural insights into the histone H1-nucleosome complex // Proc. Natl. Acad. Sci. U. S. A. 2013. Vol. 110. N 48. P. 19390–19395.

12. Zhou B.-R., Jiang J., Feng H., Ghirlando R., Xiao T.S., Bai Y. Structural mechanisms of nucleosome recognition by linker histones // Mol. Cell. 2015. Vol. 59. N 4. P. 628–638.

13. Zhou B.-R., Feng H., Ghirlando R., Li S., Schwieters C.D., Bai Y. A Small number of residues can determine if linker histones are bound on or off dyad in the chromatosome // J. Mol. Biol. 2016. Vol. 428. N 20. P. 3948–3959.

14. Cui F., Zhurkin V.B. Distinctive sequence patterns in metazoan and yeast nucleosomes: implications for linker histone binding to AT-rich and methylated DNA // Nucleic Acids Res. 2009. Vol. 37. N 9. P. 2818–2829.

15. Öztürk M.A., Pachov G.V., Wade R.C., Cojocar V. Conformational selection and dynamic adaptation upon linker histone binding to the nucleosome // Nucleic Acids Res. 2016. Vol. 44. N 14. P. 6599–6613.

16. Pachov G.V., Gabdoulline R.R., Wade R.C. On the structure and dynamics of the complex of the nucleosome and the linker histone // Nucleic Acids Res. 2011. Vol. 39. N 12. P. 5255–5263.

17. Song F., Chen P., Sun D., Wang M., Dong L., Liang D., Xu R.-M., Zhu P., Li G. Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units // Science. 2014. Vol. 344. N 6182. P. 376–380.

18. Rohs R., Jin X., West S.M., Joshi R., Honig B., Mann R.S. Origins of specificity in protein–DNA recognition // Annu. Rev. Biochem. 2010. Vol. 79. P. 233–269.

19. Edgar R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput // Nucleic Acids Res. 2004. Vol. 32. N 5. P. 1792–1797.

20. Cock P.A., Antao T., Chang J.T., Chapman B.A., Cox C.J., Dalke A., Friedberg I., Hamelryck T., Kauff F., Wilczynski B., de Hoon M.J.L. Biopython: freely available Python tools for computational molecular biology and bioinformatics // Bioinformatics 2009. Vol. 25. N 11. P. 1422–1423.

21. Webb B., Sali A. Comparative protein structure modeling using MODELLER // Curr. Protoc. Bioinformatics. 2016. Vol. 54. P. 5.6.1–5.6.37.

22. Martí-Renom M.A., Stuart A.C., Fiser A., Sánchez R., Melo F., Sali A. Comparative protein structure modeling of genes and genomes // Annu. Rev. Biophys. Biomol. Struct. 2000. Vol. 29. P. 291–325.

23. Pettersen E.F., Goddard T.D., Huang C.C., Couch G.S., Greenblatt D.M., Meng E.C., Ferrin T.E. UCSF Chimera — a visualization system for exploratory research and analysis // J. Comput. Chem. 2004. Vol. 25. N 13. P. 1605–1612.

24. Lu X., Olson W.K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures // Nucleic Acids Res. 2003. Vol. 31. N 17. P. 5108–5121.

25. Olson W.K., Gorin A.A., Lu X.-J., Hock L.M., Zhurkin V.B. DNA sequence-dependent deformability deduced from protein–DNA crystal complexes // Proc. Natl. Acad. Sci. U. S. A. 1998. Vol. 95. N 19. P. 11163–11168.

26. Воеводин В., Жуматий С., Соболев С., Антонов А., Брызгалов П., Никитенко Д., Стефанов К., Воеводин В. Практика суперкомпьютера “Ломоносов” // Открытые системы. СУБД. 2012. № 7 (183). С. 36–39.

Поступила в редакцию
21.01.2018

Принята в печать
15.03.2018

MOLECULAR BIOLOGY

JOINT EFFECT OF HISTON H1 AMINO ACID SEQUENCE AND DNA NUCLEOTIDE SEQUENCE ON THE STRUCTURE OF CHROMATOSOMES: ANALYSIS BY MOLECULAR MODELING METHODS

T.K. Gorkovets, G.A. Armeev, K.V. Shaitan, A.K. Shaytan*

Bioengineering Department, Faculty of Biology, Moscow State University,
Russia, 119234, Moscow, 1–12 Leninskie Gory

*e-mail: alex@intbio.org

A chromosome consisting of a nucleosome core, linker DNA and linker histone (LH), is an important structural element of chromatin and plays role in the replication and transcription regulation. There are two experimentally confirmed modes of LH binding to the nucleosome and linker DNA, which differ in their geometry: binding on-dyad and off-dyad. It was shown that the LH amino acid sequence influences the type of histone binding and the conformation of the chromatosome. However, the geometry of linker DNA bound with LH also changes. Thus, the mutual influence of these factors and the molecular basis determining the type of LH binding to nucleosomes remain unclear. In this study, we applied molecular modeling methods, including homology modeling, atom-atom interaction analysis and DNA deformation energy analysis to study the joint effect of the LH amino acid sequence and the DNA nucleotide sequence on the

configuration of the chromatosome. We analyzed the known crystal and NMR structures of the chromatosome for the atom-atom interactions of LH and DNA as well as the energy of DNA deformation in these structures for various DNA sequences. For various LH H1 variants, the analysis was carried out using homology modeling methods. Sequence-dependent differences in the bending energy of the linker DNA for two different conformations of the chromatosome were found, and nucleotide sequences preferred for these structures were also proposed. As a result of the analysis, it was shown that the DNA nucleotide sequence along with the LH amino acid sequence influences the type of binding to the nucleosome. It is assumed that the contribution of the DNA nucleotide sequence and its geometry can be determinative in comparison with the LH amino acid sequence in some cases. Hypotheses for experimental verification have been formulated, according to which the type of LH binding can change with different DNA nucleotide sequences.

Keywords: *chromatosome, nucleosome, chromatin, linker histone, linker DNA, DNA flexibility, homology modeling*

Сведения об авторах

Горковец Татьяна Константиновна – аспирант кафедры биоинженерии биологического факультета МГУ. Тел.: 8-495-939-57-38; e-mail: gorkovets@molsim.org

Армеев Григорий Алексеевич – аспирант, мл. науч. сотр. кафедры биоинженерии биологического факультета МГУ. Тел.: 8-495-939-57-38; e-mail: armeev@intbio.org

Шайтан Константин Вольдемарович – докт. физ-мат. наук, проф. кафедры биоинженерии биологического факультета МГУ. Тел.: 8-495-939-23-74; e-mail: shaytan49@yandex.ru

Шайтан Алексей Константинович – канд. физ-мат. наук, вед. науч. сотр. кафедры биоинженерии биологического факультета МГУ. Тел.: 8-495-939-57-38; e-mail: alex@intbio.org