

Комплексы белки–ДНК: специфичность и механизмы узнавания

Е. П. Борискина, М. Ю. Ткаченко, А. В. Шестопалова

Институт радиопизики и электроники им. А. Я. Усикова НАН Украины
Ул. Академика Проскуры, 12, Харьков, Украина, 61085

shestop@ire.kharkov.ua

Белково-нуклеиновое узнавание – важный аспект регуляции экспрессии генов, репликации ДНК и ее компактизации. Результаты изучения механизмов узнавания показывают, что последовательность нуклеотидов ДНК считается регуляторными белками и позволяет им селективно связываться с конкретными участками ДНК. Обзор посвящен процессам, происходящим при специфичном и неспецифичном комплексообразовании белков и ДНК. Основное внимание уделено прямому и непрямому механизмам сиквенс-специфичного узнавания. Приведены примеры комплексов, иллюстрирующие многообразие способов белково-нуклеинового узнавания.

Ключевые слова: типы белково-нуклеинового комплексообразования, специфичность белково-нуклеиновых взаимодействий, механизмы прямого и непрямого узнавания.

Введение. Образование комплексов белков с нуклеиновыми кислотами происходит на всех этапах репликации ДНК, экспрессии генов и ее регуляции. Часть таких комплексов высокоспецифичны: белок распознает «свой» сайт, отобрав его из многочисленных конкурирующих и перекрывающихся неспецифичных последовательностей, и связывается с ним [1]. Но белки могут взаимодействовать с ДНК, не проявляя однозначного сродства к определенной последовательности, а распознавая ДНК как таковую [2], например, в процессе компактизации при образовании нуклеосом [3].

Комплексы белков с ДНК характеризуются сродством и специфичностью. Сродство – это свойство белка связываться с ДНК, определяемое, в частности, наличием стерически и химически комплементарных поверхностей обеих молекул в сайтах связывания [4]; специфичность – селективное свя-

зывание белка на конкретных сайтах ДНК, она обусловлена отличием степени сродства белка к последовательности ДНК сайта-мишени и других участков [5].

Практически все белки способны образовывать и специфичные, и неспецифичные комплексы [6–8]. Но константы ассоциации, количественно определяющие сродство, отличаются на несколько порядков при связывании со «своей» и «чужими» последовательностями: для специфичных комплексов они составляют 10^{10} – 10^{12} M^{-1} (не менее 10^6 M^{-1}), для неспецифичных – 10^4 M^{-1} и ниже.

Диапазон изменений свободной энергии связывания для специфичных комплексов лежит в пределах -16 ÷ -7 ккал/моль, для неспецифичных – он равен -8 ÷ -5 ккал/моль [8, 9]. Основным признаком специфичности принято считать изменение теплоемкости, обусловленное модификациями структуры белка (дополнительным экранированием гидрофобных групп) [2, 6–12]: специфичное связывание

характеризуется большим отрицательным изменением ($-1,1 \div -1,3$ ккал · моль⁻¹ · К⁻¹), а неспецифичное – близким к 0 [10–12].

В то же время с помощью структурного анализа выявлены принципиальные различия в типах взаимодействий и контактов, параметрах двойной спирали ДНК, количестве связанной с комплексами воды даже у специфичных (с точки зрения термодинамики и кинетики) комплексов. Поэтому в обзоре [2] предложена классификация, включающая четыре типа комплексов: сиквенс-специфичные, неспецифичные, несиквенс-специфичные и структурно-специфичные.

В представленном обзоре нами рассмотрены перечисленные выше типы комплексов, белково-нуклеиновые взаимодействия и механизмы узнавания. Основное внимание уделено сиквенс-специфичным взаимодействиям, с помощью которых белки распознают «свои» последовательности ДНК. Проанализированы примеры комплексов, демонстрирующие многообразие способов белково-нуклеинового узнавания.

Для понимания основ специфичного белково-нуклеинового узнавания необходим анализ роли обоих участников как равноправных партнеров [13]. Поскольку ряд публикаций посвящен именно белкам (например, [14–18]), то в нашем обзоре мы хотим оценить вклад другой составляющей – ДНК – в узнавание белками «своих» сайтов. В последние годы эта проблема также обсуждается во многих статьях ([19–21] и ссылки в них), в которых, в частности, проанализировано, насколько структурная адаптация двойной спирали ДНК определяется связыванием с белком и/или является свойством, присущим конкретным последовательностям. Для объяснения этих эффектов необходим сравнительный анализ структуры как свободной ДНК, так и в комплексах с белками [22]. Поэтому в настоящем обзоре при рассмотрении примеров комплексов приведены результаты выполненных нами сравнительных оценок структурных параметров ДНК в этих комплексах и канонических форм двойной спирали.

Сиквенс-специфичное комплексообразование. В комплексах этого типа белок связывается с одной или несколькими последовательностями

ДНК, проявляющими наибольшее сродство к данному белку [23]. Более 30 лет назад высказано предположение о том, что сиквенс-специфичное белково-нуклеиновое узнавание обусловлено способностью белков образовывать водородные (Н-) связи со специфичными группами оснований в местах контактов [24]. Такой механизм получил название прямого (Direct Readout) [25, 26].

Более поздние исследования показали, что специфичные группы оснований не составляют достаточной базы для «кода узнавания» ДНК белками [27] и прямые взаимодействия – лишь часть сиквенс-специфичного узнавания [13, 28]. Важными факторами узнавания являются также структурные особенности и способность к деформации фрагмента ДНК, с которым связывается белок, непосредственно зависящие от последовательности [25, 26, 29, 30]: глобальный изгиб ДНК [31, 32], топология и заряд желобков [33, 34], упорядоченная структура воды [33], локальная геометрия сахарофосфатного остова [35], гибкость ДНК [19, 36, 37]. Механизм узнавания участка ДНК по его структуре и/или способности к деформации, обеспечивающий его адаптацию в сайте связывания и, следовательно, оптимальное взаимодействие с белком, получил название непрямого (Indirect Readout) [38].

Необходимо подчеркнуть, что деление на прямой и непрямо механизмы достаточно условно, поскольку в реальных условиях они зачастую используются совместно, обеспечивая безошибочное узнавание и надежное связывание.

В настоящее время прямой и непрямо механизмы рассматривают шире, учитывая способность белков распознавать линейную последовательность нуклеотидов – уникальную химическую основу узнавания (base readout) и зависящую от последовательности форму сайта связывания (shape readout). Первый механизм реализуется в большом и/или малом желобках, второй – при узнавании глобальных и/или локальных особенностей структуры конкретных последовательностей ДНК [13].

Согласно возможной схеме связывания белков с ДНК [24], на первом этапе образуется неспецифичный комплекс за счет электростатических взаимодействий с последующими «слайдингом» молекулы белка вдоль цепи ДНК, узнаванием «своей» по-

следовательности и формированием специфического комплекса (второй этап), которые сопровождаются структурной адаптацией обеих молекул [39–42].

Для описания этих процессов необходимо выяснить, как белки различают специфичные и неспецифичные сайты ДНК, какие взаимодействия и атомные группы ДНК ответственны за образование комплексов.

Наиболее полный анализ видов межатомных взаимодействий представлен в работе [43]: в комплексах выделены все белково-нуклеиновые контакты, которые разделены по видам взаимодействий и по возможным местам их образования.

Оказалось, что в среднем в комплексах 47 % всех контактов происходят с фосфатами, 29 % – с дезоксирибозами, а с основаниями – лишь 24 %.

Самым распространенным видом белково-нуклеиновых взаимодействий являются, как и предполагалось, Н-связи. Основным местом таких контактов оказались не основания, а атомы кислорода фосфатных групп (62 % Н-связей). Ван-дер-ваальсовы взаимодействия (22 %) чаще возникают с атомами сахарофосфатного остова ДНК (62 %), а на гидрофобные взаимодействия приходится 19 % всех контактов белок–ДНК и это – основной вид взаимодействий для дезоксирибоз (63 %). Местом реализации электростатических взаимодействий (8 %) должны быть атомы кислорода фосфатных групп, но, как оказалось, лишь 16 из 47 % взаимодействий белков с фосфатными группами – электростатические. Остальные представляют собой Н-связи и другие виды контактов.

Представленные данные – это результат анализа структур разных белково-нуклеиновых комплексов, использующих оба механизма узнавания: прямой и непрямой.

Прямое узнавание. При прямом механизме белки узнают и выделяют определенные пары оснований ДНК, образуя комплементарные Н-связи между боковыми цепями аминокислотных остатков и донорно-акцепторными группами оснований в желобках двойной спирали. Энергия Н-связей зависит от взаимной ориентации и расстояния между донором и акцептором в зоне контакта [44]. Учитывая имеющиеся пространственные ограничения для деформации взаимодействующих молекул, сама

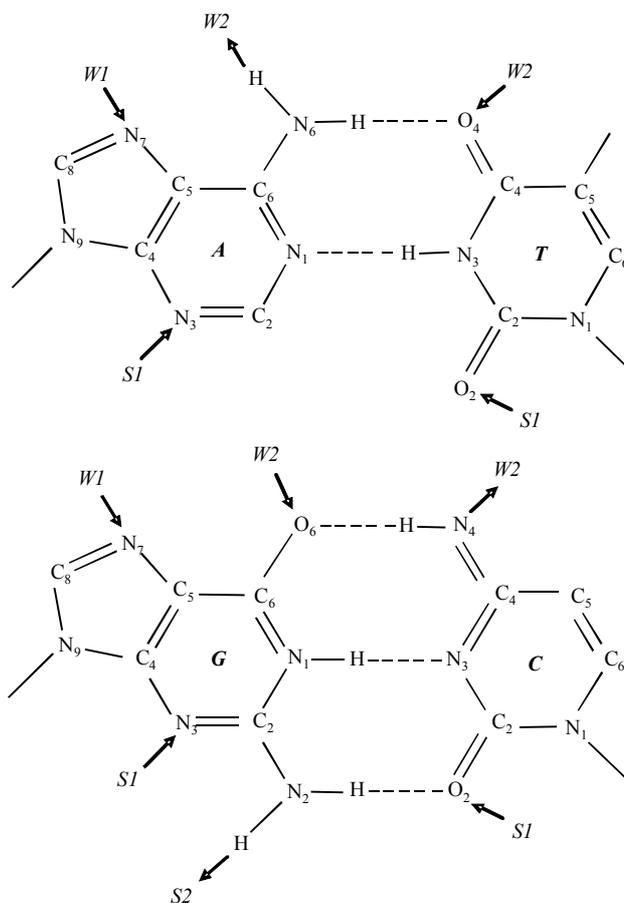


Рис. 1. Потенциальные сайты образования Н-связей в большом (W) и малом (S) желобках

возможность образования Н-связей и их энергия могут быть факторами отбора в процессе узнавания белком «своей» последовательности ДНК [28].

Потенциальные положения акцепторов и доноров Н-связей оснований в желобках ДНК впервые описаны в работе [24] и дополнены результатами более поздних исследований [11, 25]. В большом желобке (рис. 1) взаимодействия аминокислотных остатков и групп ДНК происходят, в основном, в положениях W2. Именно в этих позициях все четыре основания включают атомы, способные образовывать Н-связи. В положениях W1 только аденин и гуанин имеют возможности для формирования Н-связей, у пиримидинов этого положения нет. Поскольку пурины содержат больше сайтов для образования Н-связей, чем пиримидины, их роль в прямом узнавании значительно важнее [28]. Наибольшим потенциалом для образования Н-связей обладает гуанин, затем (в порядке его уменьшения) –

аденин, цитозин и тимин. Таким образом, в большом желобке формируется зависящий от последовательности уникальный набор донорно-акцепторных групп [24].

Взаимодействия в малом желобке (рис. 1) происходят, как правило, в положении S1. У аденина, цитозина и тимина в малом желобке содержится по одному акцептору, у гуанина – один акцептор и один донор.

В целом, большой желобок характеризуется большим количеством сайтов, образующих Н-связи и участвующих в других видах взаимодействий, например, гидрофобных [43]. Большой желобок В-ДНК стерически комплементарен α -спиралям белков [25]. Поэтому последние могут внедряться в большой желобок и формировать специфичные Н-связи с основаниями, узнавая «свой» участок ДНК по прямому механизму.

Проанализировав геометрическое расположение донорных и акцепторных групп в желобках ДНК, авторы работы [24] пришли к выводу, что одной Н-связи недостаточно для однозначного отбора пары оснований в процессе прямого узнавания. Во-первых, пары оснований в желобках располагаются достаточно близко друг к другу, поэтому взаимодействующая боковая цепь аминокислоты может «дотянуться» до донора или акцептора Н-связи и в неспецифичной для нее паре оснований. Во-вторых, различие в энергиях Н-связей с участием атомов кислорода и азота (например, O2 пиримидинов и N3 пуринов) не столь велико, чтобы служить фактором отбора в процессе узнавания.

Поэтому для однозначного распознавания комплементарной пары нуклеотидов необходимы, как минимум, две Н-связи в местах контактов [15, 24], что и является основой высокой специфичности. Так, боковые цепи аминокислот, образуя парные Н-связи, способны однозначно различить пары А–Т и G–C в большом желобке и пары G–C – в малом. А формирование двух Н-связей на одной функциональной группе обеспечивает механизм фиксации положения этих связей относительно друг друга с большей точностью, чем при образовании двух одиночных Н-связей. Анализ структуры белково-нуклеиновых комплексов выявил, что 36,9 % Н-связей между остатками аминокислот и основания-

ми – одиночные, 33,8 % – парные и 34,1 % – сложные (Н-связи образуются между аминокислотными остатками и основаниями, расположенными вдоль цепи); часть неординарных Н-связей являются смешанными [15].

Как правило, взаимодействующие с ДНК аминокислоты отдают предпочтение не определенному основанию (паре), а определенному положению атомов в малом или большом желобке [28]. Но лизин и аргинин чаще контактируют с гуанином в большом желобке, и именно эти взаимодействия считаются наиболее важными для узнавания гуанина. Глутамин и аспарагин предпочитают взаимодействовать с аденином в большом желобке [15, 45, 46]. Контакты остатков глутаминовой и аспарагиновой кислот с ДНК довольно редки, учитывая общий отрицательный заряд молекулы ДНК, и они чаще возникают в позиции W2 большого желобка с цитозином, имеющим наибольший положительный заряд [47].

В белково-нуклеиновых комплексах обнаружены связи C–H...O [43, 48–50], формирующиеся между OH-группами аминокислотных остатков (например, серина и треонина) и C5H-группами цитозина и метильными группами тимина в большом желобке, участвуя тем самым в прямом узнавании этих оснований [47, 49]. Их количество сопоставимо с таковым классических Н-связей [47].

Неспецифичные Н-связи наиболее часто образуются в местах контактов белков с группами сахарофосфатного остова ДНК без значимых предпочтений при взаимодействии аминокислот с нуклеотидами [28]. Н-связи пептидного остова белка и неспецифичных групп оснований или сахарофосфатного остова ДНК формируются в основном между пиримидинами и аминокислотами компактных размеров, которые не содержат групп, способных образовывать Н-связи: глицин, аланин, валин, лейцин [28].

Взаимодействия аминокислот, имеющих положительно заряженную боковую цепь или частичный положительный заряд на аминогруппе, с π -системой оснований (как правило, гуанина) также встречаются довольно часто. При этом аминокислотный радикал встраивается между соседними парами оснований с образованием контактов типа

«катион–π-система» с одной из пар и Н-связи – с другой [51, 52].

Гидрофобные взаимодействия боковых цепей неполярных аминокислот с основаниями чаще возникают при интеркаляции участка белка в малый желобок и сопровождаются изгибами ДНК в сторону большого желобка [53, 54].

Электростатические взаимодействия происходят между положительно заряженными аминокислотными цепями и атомами кислорода фосфатов, не только стабилизируя уже сформировавшийся комплекс, но и участвуя в пре-ориентации пептидных цепей по отношению к молекуле ДНК [55–57]. К электростатическим взаимодействиям следует отнести образование солевых мостиков между положительно заряженными группами боковых цепей аминокислотных остатков и кислородами фосфатных групп [57–59].

Таким образом, на основании результатов анализа кристаллических структур комплексов белков и ДНК получены доказательства существования нескольких видов взаимодействий, определяющих формирование и стабильность комплексов: Н-связи между аминокислотными остатками и специфичными группами оснований и/или неспецифичными группами сахарофосфатного остова, ван-дер-ваальсовы, гидрофобные и электростатические взаимодействия [60].

Специфичное белково-нуклеиновое узнавание, реализующееся с преобладанием прямого механизма, присуще многим белкам: эндонуклеазам, репрессорам, некоторым транскрипционным факторам [14, 15, 17, 18]. Эти белки проявляют высокую специфичность при связывании с конкретными последовательностями ДНК – промоторами или энхансерами. Менее распространенным является специфичное связывание в малом желобке за счет высокоспецифичных гидрофобных взаимодействий с основаниями. Например, в комплексе ДНК с транскрипционным фактором LEF-1 (lymphoid enhancer factor-1) (рис. 2, а, см. вклейку) [61] остаток метионина специфически встраивается только в АА-динуклеотидный шаг ДНК в сайте связывания (рис. 2, б, см. вклейку) [2, 13], при этом формируются специфичные Н-связи в большом желобке (рис. 2, в, см. вклейку).

Но для многих регуляторных белков высоко-специфичное узнавание «своего» сайта-мишени происходит без образования комплементарных Н-связей, вместо этого используются стерическое соответствие взаимодействующих поверхностей обеих молекул и/или ван-дер-ваальсовы контакты между боковыми цепями или остовом белка и основаниями [62].

Непрямое узнавание. При непрямом механизме узнавания участок ДНК распознается либо по неканонической В-форме двойной спирали, существующей в свободной ДНК, либо по способности ДНК к деформации и/или переходу в альтернативную конформацию в комплексе [21, 38]. Преимущество в энергии неканонической структуры над канонической в первом случае и энергетические затраты, необходимые для перевода ДНК в альтернативную конформацию, – во втором, зависят от последовательности [62, 63].

Систематический анализ данных рентгеноструктурного анализа различных фрагментов ДНК позволил сделать вывод о том, что способность двойной спирали ДНК принимать альтернативные конформации является ее «внутренним» свойством, присущим определенным последовательностям ДНК [29, 64–67]. В работах [68, 69] показана способность различных по нуклеотидному составу коротких фрагментов ДНК как сохранять структуру классических А- или В-форм двойной спирали, так и принимать промежуточные состояния или претерпевать конформационные переходы. Например, шаги GrG и CrG проявляют тенденцию к спонтанному переходу в А-форму в свободной ДНК [70–73], то есть могут быть конформационным «сигналом» к комплексообразованию.

В работах [43, 53] определяли частоту вовлечения в белково-нуклеиновые взаимодействия нуклеотидов с А- или В-конформациями сахарофосфатного остова. Оказалось, что хотя фракция А-подобных нуклеотидов небольшая (12 %), среднее количество белково-нуклеиновых контактов на один А-подобный нуклеотид в 1,6 раза выше, чем на В-подобный (2,5 и 1,6 контакта на нуклеотид соответственно).

Анализ площадей доступной поверхности В- и А-подобных нуклеотидов, изменение которых за-

висит от переключения конформации дезоксирибозы с С2'-эндо (В-форма) на С3'-эндо (А-форма), показал увеличение доступности гидрофобных атомов А-подобных нуклеотидов в малом желобке [53]. Именно поэтому взаимодействия аминокислот с двумя формами ДНК в малом желобке распределены по-разному: А-подобные нуклеотиды чаще связываются с неполярными аминокислотными остатками, а В-подобные – с положительно заряженными. Когда А-подобные нуклеотиды формируют кластеры вдоль цепи, взаимодействия с гидрофобными поверхностями белков происходят кооперативно и преимущественно с β -слоями [74]. Кроме того, в этих сайтах часто образуются изгибы ДНК [53], как правило, в пиримидин-пуриновых (YR) шагах, имеющих наибольшую склонность к структурным перестройкам. Учитывая, что возможность для распознавания ДНК посредством специфичных Н-связей в малой бороздке невысока [24], переходы ДНК в А-форму и возникновение гидрофобных взаимодействий могут быть важными факторами непрямого узнавания.

Конформация В' [75] формируется в участках последовательных АрА- и/или АрТ- шагов (А-трактов), имеющих узкий малый желобок с концентрацией отрицательно заряженных фосфатов большей, чем в канонической В-ДНК. Это объясняется тем, что в малом желобке А-трактов формируется уникальная гидратная оболочка, включающая до четырех слоев молекул воды, у первого из которых ограничены трансляционные и ротационные степени свободы, поскольку в отсутствие С–G-пар в малом желобке практически нет доноров протона [37, 76]. Такая гидратация А-трактов приводит к сужению малого желобка и увеличению его электростатического потенциала [34], облегчая узнавание белками этих участков ДНК [77]. В работах [78–80] показано, что А-тракты часто локализуются вблизи промоторных участков генов бактерий и высших организмов и/или участков инициации репликации и рекомбинации, являясь мишенями для образования комплексов с белками, причастными к регуляции этих важных биологических процессов.

Еще одним фактором непрямого узнавания белками ДНК является сиквенс-зависимая способность двойной спирали ДНК к деформации при ее взаи-

модействии с белком. Количественная оценка деформации ДНК выполняется на основе учета изменений структурных параметров, например, Twist, Roll и Tilt; Shift, Slide и Rise [30, 81] и производится по ее энергии [19, 82, 83]. Для каждого шага оснований вычисляются эмпирические значения энергии деформации любого из шести параметров, базируясь на разбросе их значений в имеющихся базах данных структур свободной ДНК и белково-нуклеиновых комплексов [20] либо обнаруженных при расшифровке траекторий молекулярной динамики коротких фрагментов ДНК [82–84]. Примечательно, что эти два подхода дают близкие результаты. Так, оказалось, что способность к деформации двойной спирали ДНК является функцией ее последовательности [82]. Наиболее легко деформируются пиримидин-пуриновые шаги (YR) [84], затем следуют шаги RR и RY. Самый гибкий тетрамер – TTAG, а наиболее жесткий – AATT [83].

Различия в энергии В→А-перехода для разных последовательностей используются белками – регуляторами транскрипции, взаимодействующими с ТАТА-боксом (последовательность типа 5'-ТАТА (AA)-3'), который находится в промоторных участках многих генов. В процессе комплексообразования эти участки переходят в А-форму, чтобы облегчить связывание β -слоя белка в малом желобке ДНК. Энергия перехода сиквенс-специфична и служит фактором отбора сайта-мишени [74].

Таким образом, чувствительность определенных фрагментов ДНК к взаимодействиям с функциональными группами белков может быть обусловлена зависимым от последовательности пространственным и энергетическим «кодом», определяемым способностью конкретного участка ДНК к деформации [85–87].

Примером связывания белка с ТАТА-боксом может служить комплекс ДНКазы I с ДНК (рис. 3, а, см. вклейку). ДНКазы I при образовании комплекса узнает сиквенс-зависимые структурные изменения двойной спирали ДНК. Белок связывается с ДНК в малом желобке с центральными парами оснований ТАТА-бокса, образуя ван-дер-ваальсовы контакты и Н-связи с основаниями (из них только одна специфична) и сахарофосфатным остовом. При этом происходит расширение малого желобка на 3 Å и изгиб

Таблица 1

Параметры, характеризующие упаковку сахаров сайта связывания ДНКазы I: угол δ и конформация дезоксирибоз

Пара	Цепь I		Цепь II	
	δ , °	Конформация	δ , °	Конформация
G:C	80,1	C3'-эндо	116,6	O4'-эндо
G:C	97,6	C3'-эндо	85,4	C3'-эндо
T:A	94,7	C3'-эндо	129,8	C1'-экзо
A:T	76,3	C3'-эндо	97,0	C3'-эндо
T:A	100,5	O4'-эндо	128,7	C1'-экзо
A:T	127,5	C1'-экзо	136,9	C1'-экзо
C:G	121,8	C1'-экзо	153,9	C2'-эндо

Примечание. А-ДНК: $\delta = 81 \pm 7^\circ$; В-ДНК: $\delta = 125 \pm 17^\circ$.

по направлению к большому желобку более чем на 20° [14]. С помощью программы 3DNA/compDNA [88] нами рассчитаны параметры, характеризующие конформацию дезоксирибоз фрагмента ДНК (табл. 1). Как следует из этих данных, часть нуклеотидов в сайте связывания имеет конформацию сахарофосфатного остова, соответствующую А- или промежуточным формам ДНК.

По механизму непрямого узнавания образуется комплекс белка Zif268 (транскрипционный фактор TFIIIA, семейство «цинковые пальцы») с ДНК [88]. «Цинковые пальцы» – относительно небольшие ДНК-связывающиеся домены, входящие в состав различных белков в виде тандемов [13, 14, 16, 90].

Согласно спектрам кругового дихроизма [89], фрагмент ДНК в комплексе (рис. 3, б, см. вклейку) принимает промежуточную конформацию между А- и В-формами. Рассчитанные нами по программе 3DNA/compDNA [88] значения углов Twist, Roll и Inclination характерны для А-формы, а величины параметров Slide и смещения по оси x – промежуточные между А- и В-формами (табл. 2). В работе [89] показано, что такая конформация сайта-мишени является оптимальной для образования комплекса и служит фактором отбора при узнавании, а связывание Zif268 с канонической В-ДНК менее эффективно.

В комплексе репрессора trp (регуляторный белок, контролирующей инициацию транскрипции оперона биосинтеза триптофана и некоторых других оперонов *Escherichia coli*) с ДНК [38, 91] также используется

непрямое узнавание (рис. 4, а, см. вклейку): образуются прямые и опосредованные молекулами воды Н-связи с сахарофосфатным остовом и только две Н-связи с основаниями – гуанинами в крайних положениях. Результаты моделирования методом молекулярной динамики [92] продемонстрировали, что в местах контакта ДНК и белка некоторые нуклеотиды имеют альтернативную ВII-конформацию и, что более важно, обнаружена корреляция между VI/VII-конформациями и количеством контактов белка с нуклеотидами в этих конформациях. Таким образом, VI/VII-переходы сахарофосфатного остова ДНК могут служить основой специфичности непрямого узнавания.

В работе [37] изучен комплекс репрессора бактериофага P22 (P22R) с ДНК (рис. 4, б, см. вклейку). P22R способен однозначно распознавать «свой» сайт связывания среди шести похожих [70], используя для узнавания и прямой механизм, и не прямой – различия в энергиях В→В'-перехода разных нуклеотидных шагов [37].

Структура P22R относится к типу «спираль–изгиб–спираль». При образовании комплекса α -спирали белка встраиваются в большие желобки, образуя Н-связи со специфичными группами оснований. Но специфичность комплекса (рис. 4, б, см. вклейку) определяется не только связыванием в местах контакта α -спиралей с ДНК, но и нуклеотидной последовательностью в промежутке между этими зонами, где нет специфичных связей с основаниями и узнавание происходит по непрямому ме-

Таблица 2
Структурные параметры сайта связывания белка Zif268

Шаг	Twist (А-ДНК: 31,1°; В-ДНК: 36,0°)	Roll (А-ДНК: 8,0°; В-ДНК: 0,6°)	Inclination (А-ДНК: 14,6°; В-ДНК: 2,1°)	Slide (А-ДНК: -1,53 Å; В-ДНК: 0,23 Å)	Смещение по оси x (А-ДНК: -4,2 Å; В-ДНК: 0,1 Å)
AC:GT	31,40	2,10	3,95	-0,50	-1,39
CG:CG	33,60	6,00	10,29	0,30	-0,48
GC:GC	30,70	-2,10	-4,10	-0,20	-0,09
CC:GG	35,20	11,00	17,73	-0,30	-2,12
CC:GG	30,60	3,20	6,19	-0,90	-2,44
CA:TG	34,10	3,40	5,87	0,10	-0,73
AC:GT	27,90	5,50	11,26	-0,60	-2,60
CG:CG	36,10	7,80	12,28	-0,20	-1,43
GC:GC	25,90	5,60	12,45	-0,10	-1,68

ханизму. ДНК при связывании α -спиралей претерпевает существенную деформацию в промежутке между участками контакта: В→В'-переход. Замена даже одной пары оснований меняет энергетический барьер такого перехода. В результате сродство репрессора Р22 к «своей» и измененным последовательностям ДНК различается [70] и белок распознает сайт-мишень.

Итак, можно сформулировать основные отличия прямого и непрямого механизмов узнавания [13, 21, 28, 37, 93].

Прямой механизм предполагает высокое стереохимическое соответствие определенных последовательностей ДНК местам контактов на поверхности белковых молекул и возможность образования специфичных Н-связей с атомами оснований, преимущественно направленных в большой желобок двойной спирали. То есть при прямом узнавании основной вклад в специфичность белково-нуклеиновых комплексов вносят прямые контакты между боковыми группами белков и специфичными группами оснований ДНК.

Непрямой механизм основан на структурной адаптации последовательностей ДНК в сайтах связывания с белком. Специфичность комплексов при непрямом узнавании определяются взаимодействиями, в результате которых образуется альтернативная структура конкретной последовательности ДНК, узнаваемая белком. При этом специфичные

группы оснований могут не вступать в прямые контакты с белком.

При реализации прямого механизма узнавания параметры, используемые для его количественного описания, – это отличия в энергиях связывания белка с различными последовательностями, а непрямого – в энергиях деформации конкретных нуклеотидных последовательностей.

В работе [18] сделана попытка оценить вклады прямого (Н-связи и гидрофобные взаимодействия) и непрямого (деформация ДНК) узнавания в специфичность связывания белков различных семейств с ДНК. Специфичность оценивали по энергетическим параметрам: энергиям парных атомных взаимодействий (прямое узнавание) и энергиям деформации конкретных фрагментов ДНК (непрямое узнавание). Основным выводом, сформулированный авторами, следующий: оба механизма узнавания используются всеми белками девяти исследованных семейств. Относительная величина вклада непрямого узнавания в специфичность связывания в среднем составляет $\approx 20\%$; максимальный вклад принадлежит семейству рестриктаз и подсемейству С2Н2/С2НСZF (более 50%: в комплексах свыше половины контактов неспецифичны); минимальный – семейству HE (homing endonucleases). При прямом узнавании вклад Н-связей (прямых и опосредованных водой) в специфичность равен в среднем 72%, гидрофобных – существенно меньше.

Рестриктазы используют не прямое узнавание как предварительный («prescreening») механизм отбора «своего» сайта по склонности конкретных последовательностей к деформации [94]. Минимальный вклад непрямого узнавания (и максимальный – прямого) у белков HE-семейства согласуется с их биологическими функциями [95] и структурными данными: в сайтах связывания этих комплексов обнаружено более 14 специфичных контактов.

Анализ результатов позволил авторам работы [18] предложить следующие правила, определяющие специфичность белково-нуклеинового связывания: необходимо учитывать наличие и доступность атомных групп обеих молекул, способных реализовать специфичные взаимодействия в сайтах связывания (прямое узнавание), и склонность конкретных последовательностей ДНК к деформации и/или конформационным перестройкам (непрямое узнавание).

Согласно этим правилам, исследованные комплексы можно разделить на три группы: у двух из них преобладает прямой механизм, но с различным числом специфичных межатомных контактов и их вкладом в специфичность связывания. В третьей группе основной вклад в специфичность вносит не прямое узнавание.

Неспецифичное комплексообразование. В комплексах этого типа происходит связывание сиквенс-специфичного белка с неспецифичной ему последовательностью ДНК, а стабилизация комплексов осуществляется за счет электростатических взаимодействий. Поскольку ДНК является полианионом, то электростатические взаимодействия должны играть существенную роль при формировании белково-нуклеиновых комплексов независимо от конкретной нуклеотидной последовательности в сайтах связывания. Это предположение подтверждается влиянием ионной силы на устойчивость неспецифичных комплексов [2, 9, 12]. Кроме того, в неспецифичных комплексах функцию посредников выполняют молекулы воды. Они занимают «мостиковые» положения между донорно-акцепторными группами белков и ДНК, дополнительно стабилизируя образующийся комплекс [96, 97].

Неспецифичное комплексообразование часто является первой стадией специфичного, где проис-

ходят обратимые конформационные изменения и переключение с неспецифичного комплекса, в котором доминируют электростатические взаимодействия с остовом ДНК, на высокоспецифичный, характеризующийся взаимодействиями с основаниями в сайте связывания [7].

В качестве примера можно привести комплексы эндонуклеазы *VamHI* с неспецифичным 5'-ATGAA-TCCATA-3' (PDB-индекс 2BAM) и специфичным 5'-ATGGATCCATA-3' (PDB-индекс 3BAM) сайтами ДНК (рис. 5, см. вклейку), нуклеотидные последовательности которых отличаются лишь одной парой [98].

Эндонуклеаза *VamHI* входит в рестрикционно-модификационную систему прокариотов и участвует в защите бактерий от бактериофагов. Ее биологическая роль заключается в быстром нахождении и разрезании чужеродной фаговой ДНК. На первом этапе эндонуклеаза неспецифично связывается с ДНК и «сканирует» ее в поисках «своего» сайта. Затем она специфично узнает последовательность из шести пар оснований (GGATCC) и делает двухнитчатый разрез.

В неспецифичном комплексе структура белка подобна таковой свободного фермента. При специфичном связывании эндонуклеаза *VamHI* претерпевает серию конформационных изменений, включающих вращение субъединиц и разупорядочение α -спирали, расположенной у С-конца белка [14, 28, 98, 99]. Процесс узнавания сопровождается деформацией двойной спирали ДНК и появлением излома в центре сайта связывания, которые необходимы для облегчения образования специфичных Н-связей (семь прямых и более 10 опосредованных молекулами воды) между группами оснований и боковыми цепями аминокислотных остатков [98].

Несиквенс-специфичное комплексообразование. Исследование нуклеосом и комплексов белков, содержащих НМГ-боксы, с ДНК позволило выявить наличие сродства и формирование стабильных комплексов без сиквенс-специфичности, то есть с минимальным предпочтением к конкретной последовательности ДНК. Такой тип комплексообразования определен как несиквенс-специфичный [2]. Тот факт, что этот вид связывания характерен для гистонов и других хромосомных белков, не случаен, поскольку функцией таких белков как раз

и является упаковка и компактизация ДНК с произвольной последовательностью.

При формировании левозакрученной суперспирали нуклеосомы ДНК претерпевает существенные структурные перестройки (рис. 6, а, см. вклейку), связанные с накоплением конформационных изменений в индивидуальных шагах пар. Основные спиральные параметры, значения которых изменяются, – это Roll, Tilt, Twist, Slide [100–102]. Такие изменения анизотропны. Величина изгиба в желобки (Roll) превышает величину изгиба в сторону сахарофосфатного остова (Tilt); в местах изгиба в большой желобок параметр Twist меньше, а в малый желобок – больше, чем у канонической В-ДНК; изменения параметра Slide периодичны – каждые 5–6 пар оснований. При этом отрицательные значения отмечены при изгибе ДНК в большой желобок, а положительные – в малый. Значения углов ϵ , ζ и β сахарофосфатного остова коррелируют со значениями спиральных параметров Roll–Slide–Twist [103]. Суперспираль стабилизируется прямыми и опосредованными молекулами воды контактами гистоновых белков с сахарофосфатным остовом ДНК и неспецифичными группами оснований в малом желобке (рис. 6, б, см. вклейку). Именно благодаря наличию водной оболочки, выполняющей роль «смазки» в местах контактов [101], у ДНК сохраняется возможность реализации локальных сиквенс-зависимых конформаций, которые могут способствовать процессу узнавания регуляторными белками конкретных сайтов-мишеней непосредственно на нуклеосомной ДНК.

В чем принципиальные отличия этого типа комплекса от специфичного? Связывание белков происходит в малом желобке с образованием Н-связей либо с группами сахарофосфатного остова, либо с неспецифичными группами оснований, направленными в малый желобок [31] (как правило, это Н-связи с акцепторными группами – N3 пуринов и O2 пиримидинов [101]) и опосредованными существенно большим числом молекул воды [104]. В стабилизации участвуют солевые мостики, возникающие между аминогруппами гистонов и кислородами фосфатных групп ДНК в нуклеосомах, и гидрофобные взаимодействия в малом желобке в комплексах с НМГ-белками [2, 105].

В комплексах упомянутого типа ДНК существенно деформирована по сравнению с линейной В-формой, но изменение структуры двойной спирали происходит по-разному: в нуклеосомах больше меняются размеры малого желобка и Roll, а у НМГ-содержащих белков – размеры обоих желобков, Twist и Roll [2]. Но в любом комплексе этого типа ДНК деформируется таким образом, чтобы максимально реализовать контакты с сахарофосфатным остовом и неспецифичными группами оснований малого желобка, при этом структура белка практически не меняется [2].

Структурно-специфичное комплексообразование. В некоторых комплексах узнавание происходит при наличии определенных структурных изменений ДНК, например, триплексов и квадруплексов или повреждений/деформаций, вызванных связыванием с ДНК биологически активных веществ [2]. При этом последовательность ДНК может быть произвольной, а белок распознает определенную деформированную структуру. Это тип структурно-специфичных комплексов, который стабилизируется неспецифичными Н-связями (прямыми и опосредованными молекулами воды), электростатическими и гидрофобными взаимодействиями. Примером здесь могут служить комплексы НМГ1 с ДНК, модифицированной цисплатином [2, 106–108].

Принципиальное отличие такого типа комплексов от рассмотренных выше состоит в следующем. В структурно-специфичных комплексах, как правило, происходит асимметричное связывание с образованием гидрофобных контактов между определенными аминокислотными остатками и участками локально деформированной ДНК (кинков). Связывание в этом виде комплексов несиквенс-специфично: степень сродства для взаимодействия одного белка с несколькими последовательностями практически не отличается. Специфичность связывания проявляется при контактах с различными по топологии фрагментами ДНК [109].

Выводы. Приведенные в обзоре данные являются результатом анализа структур комплексов белков с ДНК, которые можно разделить на четыре типа, исходя из особенностей взаимодействия и контактов в комплексах и использования при ком-

плексообразовании обоих механизмов узнавания – прямого и непрямого [2, 13, 18, 34].

Конкретные комплексы имеют различную специфичность связывания, зависящую от реализации либо одного из описанных выше механизмов, либо от обоих одновременно, что свидетельствует о многообразии ситуаций при формировании комплексов ДНК с белками, выполняющими как различные, так и сходные функции [13, 14, 18, 110].

На основе проведенного анализа литературы можно сделать некоторые обобщения.

Специфичность связывания определяется узнаванием оснований (прямой механизм) или структуры конкретной последовательности ДНК (непрямой механизм).

Прямой механизм предполагает высокое стереохимическое соответствие определенных последовательностей ДНК местам контактов на поверхности белковых молекул и возможность образования специфичных Н-связей с сиквенс-специфичными атомами оснований, преимущественно направленных в большой желобок. Непрямой механизм основан на структурной адаптации последовательностей ДНК в сайтах связывания. В этом случае участок ДНК распознается или по его неканонической структуре, существующей в свободной ДНК, или по способности ДНК к деформации и/или переходу в альтернативную конформацию, которые являются функцией последовательности.

Основные параметры, применяемые для количественного описания прямого механизма, – это отличия в энергиях связывания белка с различными последовательностями; для непрямого – отличия в энергиях деформации конкретных участков ДНК.

Любой белок при взаимодействии со «своей» последовательностью использует комбинацию механизмов, вклады которых существенно отличаются как для различных семейств ДНК-связывающих белков, так и в рамках одного семейства.

Способность отдельных фрагментов ДНК к связыванию с функциональными группами белков обусловлена зависимым от последовательности пространственным и энергетическим «кодом», определяемым возможностью конкретного фрагмента ДНК участвовать в специфичных взаимодействиях и/или деформироваться. Детальное описание зави-

симости структуры от последовательности ДНК и определение соотношения внутренних (конформационные перестройки ДНК) и внешних (деформация ДНК при связывании) факторов, влияющих на формирование комплексов, необходимо для понимания специфичности белково-нуклеиновых взаимодействий и требует дальнейших исследований.

В настоящее время возросли возможности анализа структур комплексов [111], что позволит более полно охарактеризовать физико-химические основы специфичности белково-нуклеинового комплексообразования.

O. P. Boryskina, M. Yu. Tkachenko, A. V. Shestopalova

Protein-DNA complexes: specificity and DNA readout mechanisms

O. Ya. Usikov Institute for Radio Physics and Electronics
NAS of Ukraine
12, Akademika Proskury Str., Kharkiv, Ukraine, 61085

Summary

Protein-nucleic acid recognition is essential in a number of cellular processes, in particular, gene regulation, DNA replication and compaction. Studies on the recognition mechanisms show that DNA sequence carries information which is read out by proteins that selectively bind to specific DNA sites. The review is focused on the processes taking place during formation of specific and nonspecific complexes of proteins and DNA. Special attention is paid to direct and indirect mechanisms of sequence-specific recognition. Several examples of protein-nucleic acid complexes are given to illustrate the variety of recognition mechanisms.

Keywords: protein-nucleic acid complexes, specificity of protein-nucleic acid interaction, mechanisms of direct and indirect readout.

O. П. Борискіна, М. Ю. Ткаченко, Г. В. Шестопалова

Комплекси білки–ДНК: специфічність і механізми впізнання

Резюме

Білково-нуклеїнове впізнання – важливий аспект регуляції експресії генів, реплікації ДНК та її компактизації. Результати вивчення механізмів впізнання показують, що послідовність нуклеотидів ДНК зчитується регуляторними білками і надає їм можливість селективно зв'язуватися з певними ділянками ДНК. В огляді розглянуто процеси, які відбуваються за специфічного та неспецифічного білково-нуклеїнового комплексоутворення. Основну увагу приділено прямому та непрямому механізмам сиквенс-специфічного впізнання. Наведено приклади комплексів, які демонструють різноманітність способів білково-нуклеїнового впізнання.

Ключові слова: типи білково-нуклеїнового комплексоутворення, специфічність білково-нуклеїнових взаємодій, механізми прямого та непрямого впізнання.

СПИСОК ЛИТЕРАТУРЫ

1. von Hippel P. H. Proteine-DNA recognition: new perspectives and underlying themes // *Science*.—1994.—**263**, N 5148.—P. 769–770.
2. Murphy F., Churchill M. Nonsequence-specific DNA recognition: a structural perspective // *Structure*.—2000.—**8**, N 4.—P. R83–R89.
3. Luger K., Mader A., Richmond R., Sargent D., Richmond T. Crystal structure of the nucleosome core particle at 2.8 Å resolution // *Nature*.—1997.—**389**, N 6648.—P. 251–260.
4. Yeh C., Chen F., Wang J., Cheng T., Hwang M., Tzou W. Directional shape complementarity at the protein-DNA interface // *J. Mol. Recognit.*—2003.—**16**, N 4.—P. 213–222.
5. Benos P., Lapedes A., Stormo G. Is there a code for protein-DNA recognition? Probab(istical)ly... // *Bioessays*.—2002.—**24**, N 5.—P. 466–475.
6. Rhodes D., Schwabe J. W., Chapman L., Fairall L. Towards an understanding of protein-DNA recognition // *Phil. Trans. R. Soc. Lond. B. Biol. Sci.*—1996.—**351**, N 1339.—P. 501–509.
7. Oda M., Nakamura H. Thermodynamic and kinetic analyses for understanding sequence-specific DNA recognition // *Genes Cells*.—2000.—**5**, N 5.—P. 319–326.
8. Privalov P. L. Thermodynamic problems in structural molecular biology // *Pure Appl. Chem.*—2007.—**79**, N 8.—P. 1445–1462.
9. Jen-Jacobson L., Engler L., Ames J., Kurpiewski M., Grigorescu A. Thermodynamic parameters of specific and nonspecific protein-DNA binding // *Supramol. Chem.*—2000.—**12**, N 2.—P. 143–160.
10. Spolar R. S., Record M. T. Jr. Coupling of local folding to site-specific binding of proteins to DNA // *Science*.—1994.—**263**, N 5148.—P. 777–784.
11. Liu Ch. C., Richard A. J., Datta K., LiCata V. J. Prevalence of temperature-dependent heat capacity changes in protein-DNA interactions // *Biophys. J.*—2008.—**94**, N 8.—P. 3258–3265.
12. Kalodimos C. G., Biris N., Bonvin A. M., Levandoski M. M., Guennegues M., Boelens R., Kaptein R. Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes // *Science*.—2004.—**305**, N 5682.—P. 386–389.
13. Rohs R., Jin X., West S., Joshi R., Honig B., Mann R. Origins of specificity in protein-DNA recognition // *Annu. Rev. Biochem.*—2010.—**79**.—P. 233–269.
14. Luscombe N. M., Austin S. E., Berman H. M., Thornton J. M. An overview of the structures of protein-DNA complexes // *Genome Biol.*—2000.—**1**, N 1.—r001.1–r001.10.
15. Garvie C. W., Wolberger C. Recognition of specific DNA sequences // *Mol. Cell*.—2001.—**8**, N 5.—P. 937–946.
16. Krishna S. S., Majumdar I., Grishin N. V. Structural classification of zinc fingers: survey and summary // *Nucl. Acids Res.*—2003.—**31**, N 2.—P. 532–550.
17. Pingoud A., Fuxreiter M., Pingoud V., Wende W. Type II restriction endonucleases: structure and mechanism // *Cell. Mol. Life Sci.*—2005.—**62**, N 6.—P. 685–707.
18. Contreras-Moreira B., Sancho J., Angarica V. E. Comparison of DNA binding across protein superfamilies // *Proteins*.—2010.—**78**, N 1.—P. 52–62.
19. 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. USA*.—1998.—**95**, N 19.—P. 11163–11168.
20. Jones S., van Heyningen P., Berman H. M., Thornton J. M. Protein-DNA interactions: a structural analysis // *J. Mol. Biol.*—1999.—**287**, N 5.—P. 877–896.
21. Koudelka G. B., Mauro S. A., Ciubotaru M. Indirect readout of DNA sequence by proteins: the roles of DNA sequence-dependent intrinsic and extrinsic forces // *Prog. Nucl. Acids Res. Mol. Biol.*—2006.—**81**, N 1.—P. 143–177.
22. Locasale J. W., Napoli A. A., Chen S., Berman H. M., Lawson C. L. Signatures of protein-DNA recognition in free DNA binding sites // *J. Mol. Biol.*—2009.—**386**, N 4.—P. 1054–1065.
23. Ladbury J. E., Wright J. G., Sturtevant J. M., Sigler P. B. A thermodynamic study of the Trp repressor-operator interaction // *J. Mol. Biol.*—1994.—**238**, N 5.—P. 669–681.
24. Seeman N. C., Rosenberg J. M., Rich A. Sequence-specific recognition of double helical nucleic acids by proteins // *Proc. Natl Acad. Sci. USA*.—1976.—**73**, N 3.—P. 804–808.
25. Gromiha M., Siebers J. G., Selvaraj S., Kono H., Sarai A. Intermolecular and intramolecular readout mechanisms in protein-DNA recognition // *J. Mol. Biol.*—2004.—**337**, N 2.—P. 285–294.
26. Sarai A., Kono H. Protein-DNA recognition patterns and predictions // *Annu. Rev. Biophys. Biomol. Struct.*—2005.—**34**.—P. 379–398.
27. Matthews B. W. Protein-DNA interaction. No code for recognition // *Nature*.—1988.—**335**, N 6188.—P. 294–295.
28. Coulocheri S. A., Pigis D. G., Papavassiliou K. A., Papavasiliou A. G. Hydrogen bonds in protein-DNA complexes: where geometry meets plasticity // *Biochimie*.—2007.—**89**, N 11.—P. 1291–1303.
29. Aeling K., Opel M., Steffen N., Tretyachenko-Ladokhina V., Hatfield G., Lathrop R., Senear D. Indirect recognition in sequence-specific DNA binding by *E. coli* integration host factor: the role of DNA deformation energy // *J. Biol. Chem.*—2006.—**281**, N 51.—P. 39236–39248.
30. Paillard G., Lavery R. Analyzing protein-DNA recognition mechanisms // *Structure*.—2004.—**12**, N 1.—P. 113–122.
31. Ohndorf U. M., Rould M. A., He Q., Pabo C. O., Lippard S. J. Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins // *Nature*.—1999.—**399**, N 6737.—P. 708–712.
32. Grosschedl R., Giese K., Pagel J. HMG domain proteins: architectural elements in the assembly of nucleoprotein structures // *Trends Genet.*—1994.—**10**, N 3.—P. 94–100.
33. Bewley C. A., Gronenborn A. M., Clore G. M. Minor groove-binding architectural proteins: structure, function, and DNA recognition // *Annu. Rev. Biophys. Biomol. Struct.*—1998.—**27**.—P. 105–131.
34. Rohs R., West S. M., Sosinsky A., Liu P., Mann R. S., Honig B. The role of DNA shape in protein-DNA recognition // *Nature*.—2009.—**461**, N 7268.—P. 1248–1253.
35. Kopka M. L., Lavelle L., Han G. W., Ng H. L., Dickerson R. E. An unusual sugar conformation in the structure of an RNA/DNA decamer of the polypurine tract may affect recognition by RNase H // *J. Mol. Biol.*—2003.—**334**, N 4.—P. 653–665.
36. Zhang Y., Xi Z., Hegde R. S., Shakked Z., Crothers D. M. Predicting indirect readout effects in protein-DNA interactions // *Proc. Natl Acad. Sci. USA*.—2004.—**101**, N 22.—P. 8337–8341.
37. Watkins D., Mohan S., Koudelka G. B., Williams L. D. Sequence recognition of DNA by protein-induced conformational transitions // *J. Mol. Biol.*—2010.—**396**, N 4.—P. 1145–1164.
38. Otwinowski Z., Schevitz R. W., Zhang R. G., Lawson C. L., Joachimiak A., Marmorstein R. Q., Luisi B. F., Sigler P. B. Crystal structure of trp repressor/operator complex at atomic resolution // *Nature*.—1988.—**335**, N 6188.—P. 321–329.
39. Gowers D. M., Wilson G. G., Halford S. E. Measurement of the contributions of 1D and 3D pathways to the translocation of a protein along DNA // *Proc. Natl Acad. Sci. USA*.—2005.—**102**, N 44.—P. 15883–15888.

40. *Hu T., Grosberg A. Y., Shklovskii B. I.* How proteins search for their specific sites on DNA: the role of DNA conformation // *Biophys. J.*–2006.–**90**, N 8.–P. 2731–2744.
41. *Ferreiro D. U., Sanchez I. E., de Prat Gay G.* Transition state for protein-DNA recognition // *Proc. Natl Acad. Sci. USA.*–2008.–**105**, N 31.–P. 10797–10802.
42. *Halford S. E.* An end to 40 years of mistakes in DNA-protein association kinetics? // *Biochem. Soc. Trans.*–2009.–**37**, pt 2.–P. 343–348.
43. *Lejeune D., Delsaux N., Charlotiaux B., Thomas A., Bras-seur R.* Protein-nucleic acid recognition: statistical analysis of atomic interactions and influence of DNA structure // *Proteins.*–2005.–**61**, N 2.–P. 258–271.
44. *Buck M., Karplus M.* Hydrogen bond energetics: a simulation and statistical analysis of N-methyl acetamide (NMA), water, and human lysozyme // *J. Phys. Chem. B.*–2001.–**105**, N 44.–P. 11000–11015.
45. *Luscombe N. M., Lascowski R. A., Thornton J. M.* Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level // *Nucl. Acids Res.*–2001.–**29**, N 13.–P. 2860–2874.
46. *Mandel-Gutfreund Y., Schueler O., Margalit H.* Comprehensive analysis of hydrogen bonds in regulatory protein DNA-complexes: in search of common principles // *J. Mol. Biol.*–1995.–**253**, N 2.–P. 370–382.
47. *Suzuki M.* A framework for the DNA-protein recognition code of the probe helix in transcription factors: the chemical and stereochemical rules // *Structure.*–1994.–**2**, N 4.–P. 317–326.
48. *Mandel-Gutfreund Y., Margalit H., Jernigan R. L., Zhurkin V. B.* A role for CH...O interactions in protein-DNA recognition // *J. Mol. Biol.*–1998.–**277**, N 5.–P. 1129–1140.
49. *Mandel-Gutfreund Y., Margalit H.* Quantitative parameters for amino acid-base interaction: implications for prediction of protein-DNA binding sites // *Nucl. Acids Res.*–1998.–**26**, N 10.–P. 2306–2312.
50. *Treger M., Westhof E.* Statistical analysis of atomic contacts at RNA-protein interfaces // *J. Mol. Recognit.*–2001.–**14**, N 4.–P. 199–214.
51. *Biot C., Wintjens R., Rooman M.* Stair motifs at protein-DNA interfaces: nonadditivity of H-bond, stacking, and cation- π -interactions // *J. Am. Chem. Soc.*–2004.–**126**, N 20.–P. 6220–6221.
52. *Rooman M., Lievin J., Buisine E., Wintjens R.* Cation- π /H-bond stair motifs at protein-DNA interfaces // *J. Mol. Biol.*–2002.–**319**, N 1.–P. 67–76.
53. *Tolstorukov M. Y., Jernigan R. L., Zhurkin V. B.* Protein-DNA hydrophobic recognition in the minor groove is facilitated by sugar switching // *J. Mol. Biol.*–2004.–**337**, N 1.–P. 65–76.
54. *Zhang Z., Gong Y., Guo L., Jiang T., Huang L.* Structural insights into the interaction of the crenarchaeal chromatin protein Cren7 with DNA // *Mol. Microbiol.*–2010.–**76**, N 3.–P. 749–759.
55. *Drozdov-Tikhomirov L. N., Linde D. M., Poroikov V. V., Alexandrov A. A., Skurida G. I.* Molecular mechanisms of protein-protein recognition: whether the surface placed charged residues determine the recognition process // *J. Biomol. Struct. Dyn.*–2001.–**19**, N 2.–P. 279–284.
56. *Norberg J.* Association of protein-DNA recognition complexes: electrostatic and nonelectrostatic effects // *Arch. Biochem. Biophys.*–2003.–**410**, N 1.–P. 48–68.
57. *Gurlie R., Duong T. H., Zakrzewska K.* The role of DNA-protein salt bridges in molecular recognition: a model study // *Biopolymers.*–1999.–**49**, N 4.–P. 313–327.
58. *Saecker R. M., Record M. T. Jr.* Protein surface salt bridges and paths for DNA wrapping // *Curr. Opin. Struct. Biol.*–2002.–**12**, N 3.–P. 311–319.
59. *Torrado M., Revuelta J., Gonzalez C., Corzana F., Bastida A., Asensio J. L.* Role of conserved salt bridges in homeodomain stability and DNA binding // *J. Biol. Chem.*–2009.–**284**, N 35.–P. 23765–23779.
60. *Spyrakakis F., Cozzini P., Bertoli C., Marabotti A., Kellogg G., Mozzarelli A.* Energetics of the protein-DNA-water interaction // *BMC Struct. Biol.*–2007.–**7**.–P. 4–21.
61. *Giese K., Amsterdam A., Grosschedl R.* DNA-binding properties of the HMG-domain of the lymphoid-specific transcriptional regulator LEF-1 // *Genes Dev.*–1991.–**5**, N 12B.–P. 2567–2578.
62. *Watkins D., Hsiao C., Woods K. K., Koudelka G. B., Williams L. D.* P22 c2 repressor-operator complex: mechanisms of direct and indirect readout // *Biochemistry.*–2008.–**47**, N 8.–P. 2325–2338.
63. *Becker N. B., Wolff L., Everaers R.* Indirect readout: detection of optimized subsequences and calculation of relative binding affinities using different DNA elastic potentials // *Nucl. Acids Res.*–2006.–**34**, N 19.–P. 5638–5649.
64. *Hud N. V., Plavec J.* A unified model for the origin of DNA sequence-directed curvature // *Biopolymers.*–2003.–**69**, N 1.–P. 144–158.
65. *Hays F. A., Teegarden A., Jones Z. J., Harms M., Raup D., Watson J., Cavaliere E., Ho P. S.* How sequence defines structure: a crystallographic map of DNA structure and conformation // *Proc. Natl Acad. Sci. USA.*–2005.–**102**, N 20.–P. 7157–7162.
66. *Svozil D., Kalina J., Omelka M., Schneider B.* DNA conformations and their sequence preferences // *Nucl. Acids Res.*–2008.–**36**, N 11.–P. 3690–3706.
67. *Boryskina O. P., Tkachenko M. Yu., Shestopalova A. V.* Variability of DNA structure and protein-nucleic acid recongnition // *Biopolym. Cell.*–2010.–**26**, N 5.–P. 360–372.
68. *Vargason J. M., Henderson K., Ho P. S.* A crystallographic map of the transition from B-DNA to A-DNA // *Proc. Natl Acad. Sci. USA.*–2001.–**98**, N 13.–P. 7265–7270.
69. *Dickerson R. E., Ng H. L.* DNA structure from A to B // *Proc. Natl Acad. Sci. USA.*–2001.–**98**, N 13.–P. 6986–6988.
70. *Wu L., Koudelka G. B.* Sequence-dependent differences in DNA structure influence the affinity of P22 operator for P22 repressor // *J. Biol. Chem.*–1993.–**268**, N 25.–P. 18975–18981.
71. *Weston S. A., Lahm A., Suck D.* X-ray structure of the DNase I-d(GGTATACC)₂ complex at 2.3 Å resolution // *J. Mol. Biol.*–1992.–**226**, N 4.–P. 1237–1256.
72. *Heinemann U., Alings C., Bansal M.* Double helix conformation, groove dimensions and ligand binding potential of a G/C stretch in B-DNA // *EMBO J.*–1992.–**11**, N 5.–P. 1931–1939.
73. *Wahl M. C., Sundaralingam M.* Crystal structures of A-DNA duplexes // *Biopolymers.*–1997.–**44**, N 1.–P. 45–63.
74. *Samanta S., Chakrabarti J., Bhattacharya D.* Changes in thermodynamic properties of DNA base pairs in protein-DNA recognition // *J. Biomol. Struct. Dyn.*–2010.–**27**, N 4.–P. 429–442.
75. *Arnott S., Hukins D. W.* Optimised parameters for A-DNA and B-DNA // *Biochem. Biophys. Res. Commun.*–1972.–**47**, N 6.–P. 1504–1509.
76. *Woods K. K., Lan T., McLaughlin L. W., Williams L. D.* The role of minor groove functional groups in DNA hydration // *Nucl. Acids Res.*–2003.–**31**, N 5.–P. 1536–1540.
77. *Haran T. E., Mohanty U.* The unique structure of A-tracts and intrinsic DNA bending // *Q. Rev. Biophys.*–2009.–**42**, N 1.–P. 41–81.

78. *Perez-Martin J., Rojo F., de Lorenzo V.* Promoters responsive to DNA bending: a common theme in prokaryotic gene expression // *Microbiol. Rev.*—1994.—**58**, N 2.—P. 268–290.
79. *Hagerman P. J.* Sequence-directed curvature of DNA // *Annu. Rev. Biochem.*—1990.—**59**.—P. 755–781.
80. *Gimenes F., Takeda K. I., Fiorini A., Gouveia F. S., Fernandez M. A.* Intrinsically bent DNA in replication origins and gene promoters // *Genet. Mol. Res.*—2008.—**7**, N 2.—P. 549–558.
81. *Olson W. K., Bansal M., Burley S. K., Dickerson R. E., Gerstein M., Harvey S. C., Heinemann U., Lu X. J., Neidle S., Shakked Z., Sklenar H., Suzuki M., Tung C. S., Westhof E., Wolberger C., Berman H. M.* A standard reference frame for the description of nucleic acid base-pair geometry // *J. Mol. Biol.*—2001.—**313**, N 1.—P. 229–237.
82. *Lankas F., Spomer J., Langowski J., Cheatham T. E. 3rd.* DNA basepair step deformability inferred from molecular dynamics simulations // *Biophys. J.*—2003.—**85**, N 5.—P. 2872–2883.
83. *Fujii S., Kono H., Takenaka S., Go N., Sarai A.* Sequence-dependent DNA deformability studied using molecular dynamics simulations // *Nucl. Acids Res.*—2007.—**35**, N 18.—P. 6063–6074.
84. *Beveridge D. L., Barreiro G., Byun K. S., Case D. A., Cheatham T. E. 3rd, Dixi S. B., Giudice E., Lankas F., Lavery R., Maddocks J. H., Osman R., Seibert E., Sklenar H., Stoll G., Thayer K. M., Varnai P., Young M. A.* Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. I. Research design and results on d(C_pG) steps // *Biophys. J.*—2004.—**87**, N 6.—P. 3799–3813.
85. *Chen Y., Kortemme T., Robertson T., Baker D., Varani G.* A new hydrogen-bonding potential for the design of protein-RNA interactions predicts specific contacts and discriminates decoys // *Nucl. Acids Res.*—2004.—**32**, N 17.—P. 5147–5162.
86. *Lavery R.* Recognizing DNA // *Q. Rev. Biophys.*—2005.—**38**, N 4.—P. 339–344.
87. *Temiz N. A., Camacho C. J.* Experimentally based contact energies decode interactions responsible for protein-DNA affinity and the role of molecular waters at the binding interface // *Nucl. Acids Res.*—2009.—**37**, N 12.—P. 4076–4088.
88. *Lu X., Olson W. K.* 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures // *Nucl. Acids Res.*—2003.—**31**, N 17.—P. 5108–5121.
89. *Elrod-Erickson M., Rould M. A., Nekludova L., Pabo C. O.* Zif268 protein-DNA complex refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions // *Structure.*—1996.—**4**, N 10.—P. 1171–1180.
90. *Gamsjaeger R., Liew C. K., Loughlin F. E., Crossley M., Mackay J. P.* Sticky fingers: zinc-fingers as protein-recognition motifs // *Trends Biochem. Sci.*—2007.—**32**, N 2.—P. 63–70.
91. *Haran T. E., Joachimiak A., Sigler P. B.* The DNA target of the trp repressor // *EMBO J.*—1992.—**11**, N 8.—P. 3021–3030.
92. *Wellenzohn B., Flader W., Winger R. H., Hallbrucker A., Mayer E., Liedl K. R.* Indirect readout of the trp-repressor-operator complex by B-DNA's backbone conformation transitions // *Biochemistry.*—2002.—**41**, N 12.—P. 4088–4095.
93. *Locasale J. W., Napoli A. A., Chen S., Berman H. M., Lawson C. L.* Signatures of protein-DNA recognition in free DNA binding sites // *J. Mol. Biol.*—2009.—**386**, N 4.—P. 1054–1065.
94. *Little E. J., Babic A. C., Horton N. C.* Early interrogation and recognition of DNA sequence by indirect readout // *Structure.*—2008.—**16**, N 12.—P. 1828–1837.
95. *Jurica M. S., Stoddart B. L.* Homing endonucleases: structure, function and evolution // *Cell Mol. Life Sci.*—1999.—**55**, N 10.—P. 1304–1326.
96. *Sidorova N. Y., Rau D. C.* Differences in water release for the binding of *EcoRI* to specific and nonspecific DNA sequences // *Proc. Natl Acad. Sci. USA.*—1996.—**93**, N 22.—P. 12272–12277.
97. *Jayaram B., Jain T.* The role of water in protein-DNA recognition // *Annu. Rev. Biophys. Biomol. Struct.*—2004.—**33**.—P. 343–361.
98. *Viadiu H., Aggarwal A. K.* The role of metals in catalysis by the restriction endonuclease *BamHI* // *Nat. Struct. Biol.*—1998.—**5**, N 10.—P. 910–916.
99. *Selvaraj S., Kono H., Sarai A.* Specificity of protein-DNA recognition revealed by structure-based potentials: symmetric/asymmetric and cognate/non-cognate binding // *J. Mol. Biol.*—2002.—**322**, N 5.—P. 907–915.
100. *Richmond T. J., Davey C. A.* The structure of DNA in the nucleosome core // *Nature.*—2003.—**423**, N 6936.—P. 145–150.
101. *Davey C. A., Sargent D. F., Luger K., Maeder A. W., Richmond T. J.* Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution // *J. Mol. Biol.*—2002.—**319**, N 5.—P. 1097–1113.
102. *Tolstorukov M. Y., Colasanti A. V., McCandlish D. M., Olson W. K., Zhurkin V. B.* A novel roll-and-slide mechanism of DNA folding in chromatin: implications for nucleosome positioning // *J. Mol. Biol.*—2007.—**371**, N 3.—P. 725–738.
103. *Ong M. S., Richmond T. J., Davey C. A.* DNA stretching and extreme kinking in the nucleosome core // *J. Mol. Biol.*—2007.—**368**, N 4.—P. 1067–1074.
104. *Lundback T., Hansson H., Knapp S., Ladenstein R., Hard T.* Thermodynamic characterization of non-sequence-specific DNA-binding by the Sso7d protein from *Sulfolobus solfataricus* // *J. Mol. Biol.*—1998.—**276**, N 4.—P. 775–786.
105. *McAfee J. G., Edmondson S. P., Zegar I., Shriver J. W.* Equilibrium DNA binding of Sac7d protein from the hyperthermophile *Sulfolobus acidocaldarius*: fluorescence and circular dichroism studies // *Biochemistry.*—1996.—**35**, N 13.—P. 4034–4045.
106. *Dunham S. U., Lippard S. J.* DNA sequence context and protein composition modulate HMG-domain protein recognition of cisplatin-modified DNA // *Biochemistry.*—1997.—**36**, N 38.—P. 11428–11436.
107. *Imamura T., Izumi H., Nagatani G., Ise T., Nomoto M., Iwamoto Y., Kohno K.* Interaction with p53 enhances binding of cisplatin-modified DNA by high mobility group 1 protein // *J. Biol. Chem.*—2001.—**276**, N 10.—P. 7534–7540.
108. *Wozniak K., Blasiak J.* Recognition and repair of DNA-cisplatin adducts // *Acta Biochim. Pol.*—2002.—**49**, N 3.—P. 583–596.
109. *Waldmann T., Baack M., Richter N., Gruss C.* Structure-specific binding of the proto-oncogene protein DEK to DNA // *Nucl. Acids Res.*—2003.—**31**, N 23.—P. 7003–7010.
110. *Ashworth J., Baker D.* Assessment of the optimization of affinity and specificity at protein-DNA interfaces // *Nucl. Acids Res.*—2009.—**37**, N 10.—e73.
111. *Tkachenko M. Y., Boryskina O. P., Shestopalova A. V., Tolstorukov M. Y.* ProtNA-ASA: Protein-nucleic acid structural database with information on accessible surface area // *Int. J. Quant. Chem.*—2010.—**110**, N 1.—P. 230–232.

Figures to article by O. P. Boryskina, M. Yu. Tkachenko, A. V. Shestopalova

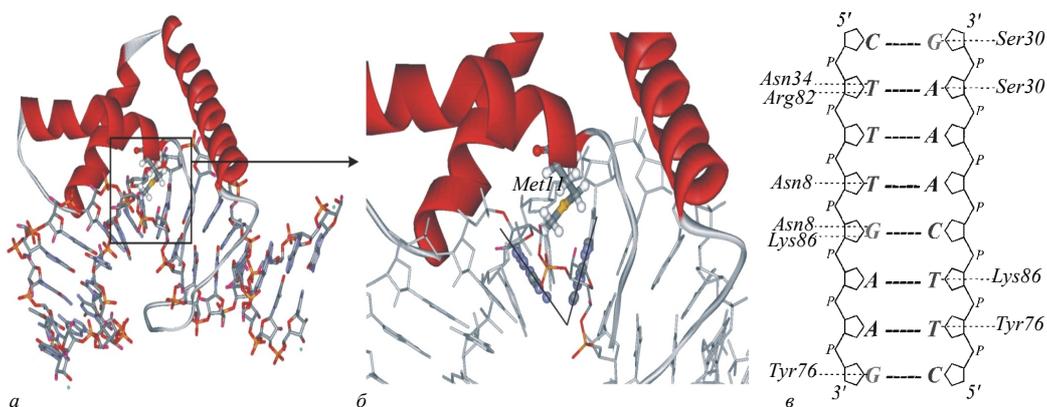


Рис. 2. Структура комплекса ДНК с транскрипционным фактором LEF-1 (PDB-индекс 2LEF): *а* – общий вид комплекса; *б* – Met11 белка специфично встроен в шаг АА; *в* – специфичные Н-связи между аминокислотами и основаниями (пунктир) в сайте связывания

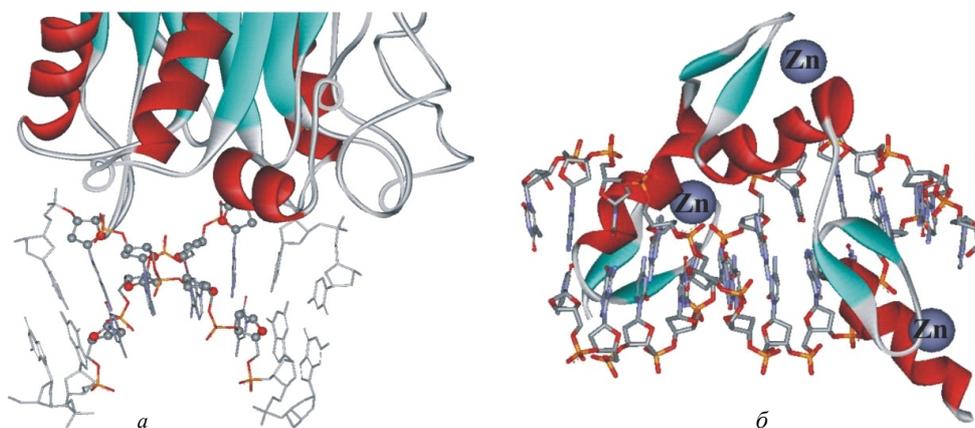


Рис. 3. Структуры комплексов ДНК с ДНКазой I (PDB-индекс 1DNK) (*а*) и белком Zif268 (PDB-индекс 1AAY) (*б*)

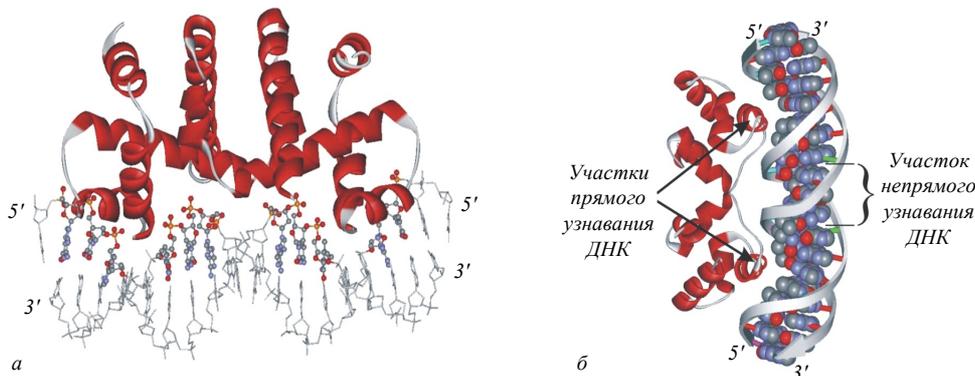


Рис. 4. Структуры комплексов ДНК с репрессором trp (PDB-индекс 1TRO, выделены нуклеотиды, образующие Н-связи с белком) (*а*) и репрессором бактериофага P22 (*б*)

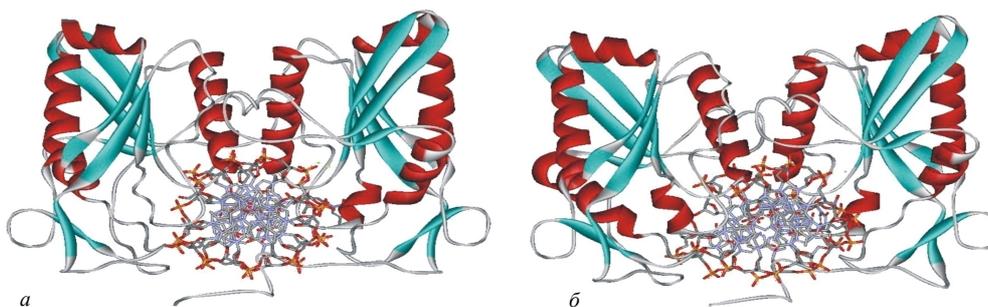


Рис. 5. Структуры комплексов эндонуклеаза BamHI–ДНК с неспецифичным (*а*) и специфичным сайтами ДНК (*б*)

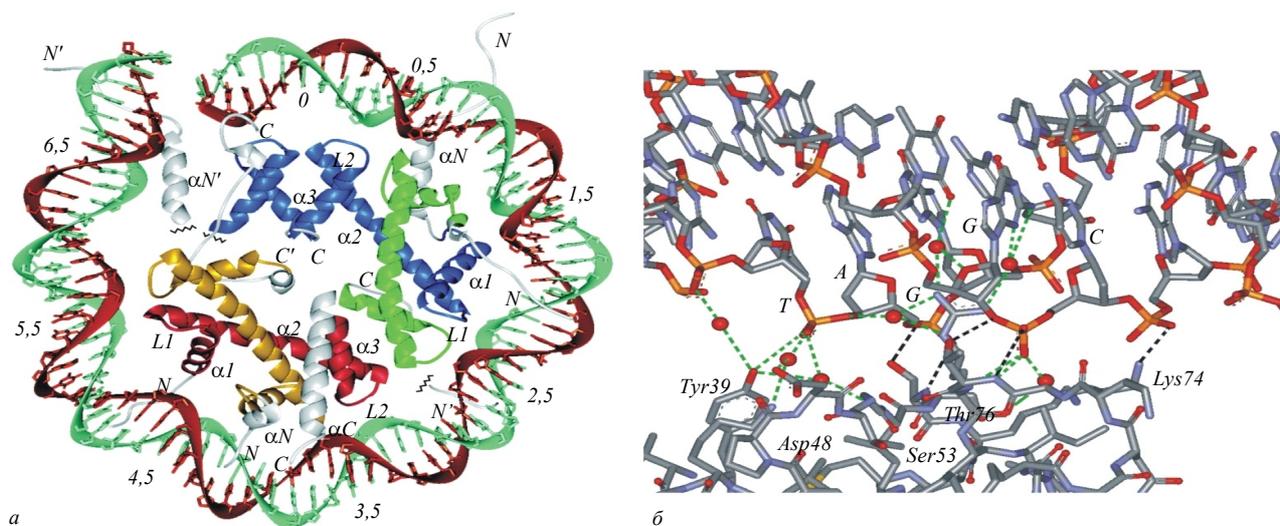


Рис. 6. Структура нуклеосомы (PDB-индекс 1KX5, сайты взаимодействия ДНК и гистонов обозначены цифрами) (а) и сайт связывания ДНК (сайт 5,5, выделены молекулы воды, занимающие мостиковые положения между донорно-акцепторными группами гистонического димера и сахарофосфатным остовом ДНК) (б)

Figures to article by S. L. Yefimova et al.

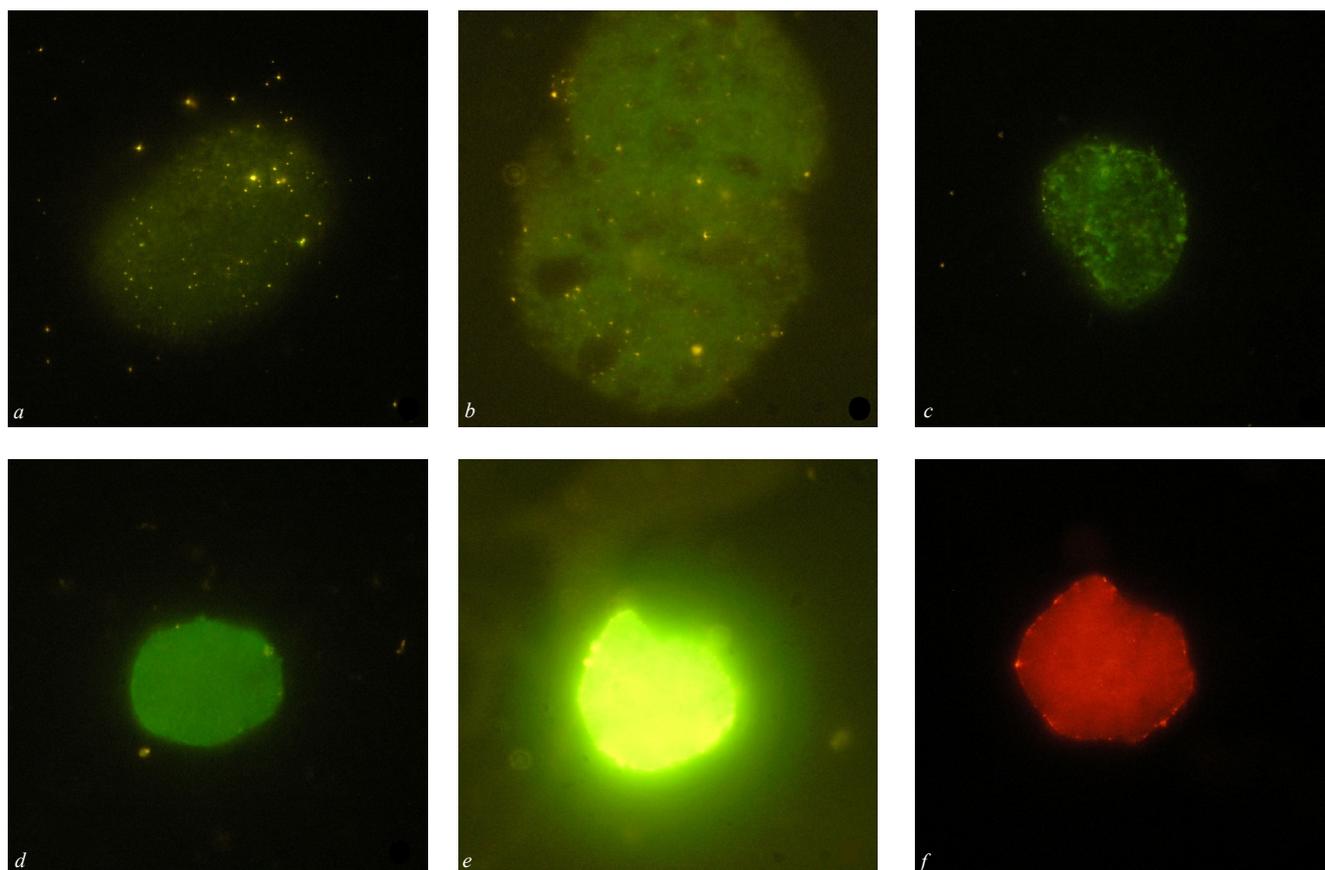


Fig. 4. Fluorescence images (original magnification $\times 100$) of rat hepatocytes incubated with FRET liposomes during different time periods: a – 0 min; b – 1 h; c – 2 h; d – 3 h; e – 20 h (excitation with BP 460–490 filter); f – 20 h (excitation with BP 510–550 filter)