

СТРУКТУРА И ФУНКЦИИ ИНТЕГРАЗЫ ВИЧ-1

© 2005 г. Ю. Ю. АГАПКИНА¹, Т. А. ПРИКАЗЧИКОВА²,
М. А. СМОЛОВ², М. Б. ГОТТИХ²

¹ *Институт проблем химической физики РАН, Черноголовка,
Московская область*

² *Химический факультет и Научно-исследовательский институт
физико-химической биологии им. А.Н. Белозерского Московского
государственного университета им. М.В. Ломоносова, Москва*

I. Введение. II. Структура. III. Функции. IV. Заключение.

I. ВВЕДЕНИЕ

Вирус иммунодефицита человека первого типа принадлежит к классу ретровирусов. Он поражает иммунную систему организма человека и вызывает синдром приобретенного иммунодефицита (СПИД) [1]. Репликативный цикл ВИЧ-1 условно можно разделить на две фазы: 1) раннюю, которая начинается с узнавания клетки-мишени зрелым вирионом и включает все процессы, ведущие к интеграции вирусной ДНК в хромосому зараженной клетки, и 2) позднюю, начинающуюся с регулируемой экспрессии интегрированного провирусного генома и заканчивающуюся отпочковыванием и созреванием вируса [2].

На ранней фазе частицы ВИЧ-1 связываются специфично только с клетками, несущими на поверхности белок CD4, ответственный за иммунное узнавание клетки. Вслед за проникновением вируса в клетку следуют процессы «раздевания» вируса и высвобождения внутриклеточного предтранскрипционного комплекса. Обратная транскрипция геномной РНК осуществляется вирусным ферментом – обратной транскриптазой – в цитоплазме. Результатом этого процесса является получение ДНК копии вирусной РНК. Эта двухцепочечная кДНК

Принятые сокращения: ВИЧ-1 – вирус иммунодефицита человека первого типа; СПИД – синдром приобретенного иммунодефицита; ИН – интегразы.
Адрес для корреспонденции: gottikh@belozersky.msu.ru.

Работа поддержана грантом РФФИ № 02-04-48797.

транспортируется в ядро в составе преинтеграционного комплекса, включающего ряд вирусных и клеточных белков. После транспорта в ядро кДНК ковалентно встраивается в геном зараженной клетки благодаря каталитической активности вирусного фермента — интегразы — с образованием провирусной ДНК. Таким образом, интеграция вирусной ДНК в клеточную является ключевой стадией репликативного цикла ВИЧ.

Понятно, что подавление репликации вируса разумно вести на ранних этапах его развития — обратной транскрипции и интеграции, — которые осуществляются вирусными ферментами. Направление, связанное с подавлением стадии обратной транскрипции ВИЧ-1, широко развито. На сегодняшний день комплексная терапия, основанная на использовании ингибиторов обратной транскриптазы и другого вирусного фермента — протеазы, рассматривается как наиболее действенная и доступная. К сожалению, она не приводит к полной остановке репликации вируса, а позволяет лишь подавлять ее до низкого уровня, продлевая жизнь ВИЧ-инфицированным пациентам [3, 4]. Кроме того, использование этой терапии ограничивается токсичностью многих ингибиторов обратной транскриптазы и протеазы, недостаточной селективностью их действия в клетках и восприимчивостью к ним организма инфицированного человека. Вирусная интеграза (ИН) не имеет гомологов в клетке, и именно по этой причине она является очень привлекательной фармакологической мишенью с точки зрения селективности и эффективности воздействия антивирусных препаратов. Подавление стадии интеграции позволило бы предотвратить репликацию вируса в *de novo* зараженных клетках и впоследствии элиминировать зараженные клетки, изолированные в тканях пациента, подвергающегося химиотерапии [5]. К сожалению, лекарственных препаратов, способных подавлять интеграцию, до настоящего времени не создано. Разработка эффективных ингибиторов ИН во многом затрудняется ограниченностью наших знаний о механизме действия и структуре этого фермента. Можно сказать, что ИН является наименее изученным ферментом ВИЧ-1.

ИН ВИЧ-1 относят к семейству полинуклеотидилтрансфераз, в которое входят такие ферменты, как транспозазы и ретровирусные ИН, а также РНКазы H и резолваза Ruv C [6, 7]. Ферменты данного семейства катализируют процессы, приводящие к перемещению ДНК или ее фрагментов внутри генома либо между геномами. Следует отметить, что данные процессы осуществляются без привлечения дополнительных источников энергии и, в конечном итоге, без изменения числа фосфодиэфирных связей. Кроме того, для ретровирусных интеграз характерно формирование очень устойчивых комплексов с

ДНК-субстратом. Однако катализируемые ими реакции протекают без образования ковалентной связи между белком и ДНК [8].

Для осуществления интеграции ИН необходимо связать сразу две молекулы ДНК: вирусную и клеточную. Соответственно, встраиваемую вирусную ДНК называют субстратом ИН, а клеточную ДНК — мишенью. На первой стадии интеграции, называемой 3'-концевой процессинг, ИН отщепляет динуклеотид GT с 3'-конца каждой из цепей вирусной ДНК. На второй стадии, называемой перенос цепи, фермент осуществляет встраивание субстратной ДНК в ДНК-мишень. Кроме того, показано, что рекомбинантная ИН может катализировать реакцию, обратную переносу цепи, которую называют дезинтеграцией. Взаимодействие ИН с ДНК-субстратом носит строго сиквенс-специфический характер, в то время как мишень связывается ферментом неспецифически. Все ретровирусные ИН высококонсервативны и обладают гомологией между собой, а также с множеством транспозаз [8].

II. СТРОЕНИЕ ИНТЕГРАЗЫ ВИЧ-1

ИН ВИЧ-1 содержит 288 аминокислотных остатков, молекулярная масса фермента составляет 32 кДа. Частичным протеолизом и направленным мутагенезом было показано, что в структуре фермента можно выделить три домена: короткий N-концевой, в состав которого входят аминокислотные остатки с 1 по 50; каталитический, который образуют аминокислотные остатки с 51 по 212; C-концевой, состоящий из аминокислотных остатков с 220 по 270 [9–12]. N-концевой и каталитический домены содержат участки, характерные для всех ретровирусных интеграз, в то время как C-концевой домен существенно менее консервативен. На рис. 1 схематически изображены три домена ИН и выделены несколько консервативных аминокислотных остатков в каждом домене.

Показано, что для проведения реакций 3'-концевого процессинга и переноса цепи необходимо наличие всех трех доменов, в то время как реакцию дезинтеграции может осуществить изолированный катали-



Рис. 1. Схематическое изображение первичной структуры ИН.

тический домен, и эффективность реакции в данном случае сравнима с эффективностью реакции, проводимой полным ферментом [13].

Пространственная структура ИН до сих пор точно не установлена. В значительной степени это обусловлено низкой растворимостью фермента, а также его высокой склонностью к агрегации [14]. Внесение в структуру отдельных доменов ИН точечных мутаций, повышающих растворимость, позволило определить структуры отдельных доменов, а также белков, состоящих из двух доменов: каталитического и С-концевого или N-концевого и каталитического.

СТРОЕНИЕ КАТАЛИТИЧЕСКОГО ДОМЕНА

Каталитический домен является наиболее устойчивым к протеолиту участком ИН [15], кроме того, он наиболее консервативен по первичной структуре. Три аминокислотных остатка Asp64, Asp116 и Glu152 формируют каталитический центр фермента. Они образуют так называемый D,D(35)E мотив, который характерен для каталитических доменов большинства известных ретровирусных интеграз и транспозаз [16].

Структура каталитического домена ИН ВИЧ-1 была определена с помощью рентгеноструктурного анализа [6]. Получить кристаллы каталитического домена ИН и определить его структуру позволила замена Phe185 на Lys. Как указывалось выше, ИН склонна к агрегации, поэтому часто ее выделяют в присутствии детергента, например, CHAPS или NP-40, который может влиять на структуру фермента, в результате чего снижается достоверность информации о строении белка. Поэтому важно отметить, что в указанной работе кристаллизацию проводили в отсутствие детергента. Внесенная аминокислотная замена не влияла на активность мутантного белка в реакции дезинтеграции. Согласно полученным данным, каталитический домен фермента в кристалле образует димер сферической формы, каждый мономер которого образует полусферу (рис. 2А). Структуру каждого мономера формируют β -лист, состоящий из пяти β -цепей, и шесть β -спиралей. Ход полипептидной цепи схематически представлен на рис. 2Б. Димер двух молекул каталитического домена ИН формируется за счет образования водородных связей и солевых мостиков между структурами $\beta 3$, $\alpha 1$, $\alpha 3$, $\alpha 5$ и $\alpha 6$ (рис. 3), при этом площадь перекрытия составляет 1300 \AA^2 .

Методом сайт-направленного мутагенеза была продемонстрирована исключительная важность каждой из аминокислот, принадлежащей к консервативной триаде D,D(35)E [17]. Например, замена Asp64 на Ile приводит к полной инактивации ИН в реакциях 3'-концевого процессинга, переноса цепи и дезинтеграции. При замене Asp116

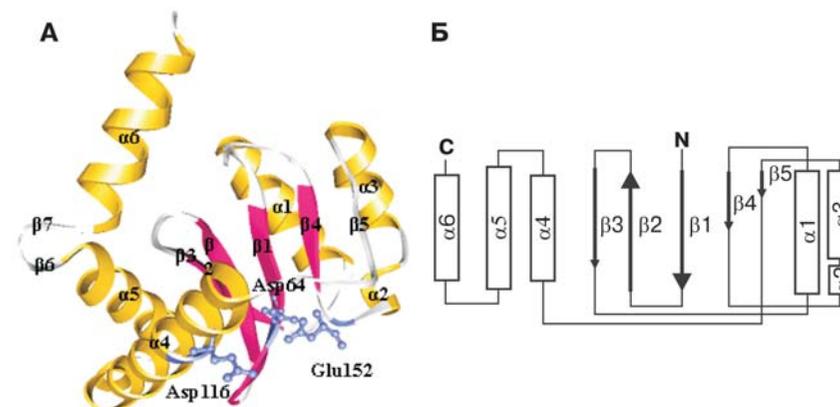


Рис. 2. Строение каталитического домена ИН ВИЧ-1 по данным работы [6].

А – строение мономера (PDB код 1BIU); Б – схематическое изображение полипептидной цепи.

на Glu детектируются лишь следовые количества продуктов этих реакций. В случае замены Glu152 на Val ИН сохраняет лишь невысокую активность в реакции 3'-концевого процессинга. Два из трех аминокислотных остатков, входящих в каталитическую триаду, а именно Asp64 и Asp116, расположены в середине $\beta 1$ -цепи и после $\beta 4$, соответственно (рис. 2А). Положение третьего остатка, Glu152, точно установить не удалось из-за высокой конформационной подвижности участка со 141 по 153 аминокислотный остаток [6]. Аминокислоты каталитической триады близко расположены в третичной структуре каталитического домена ИН ВИЧ-1, образуя активный центр фермента. Однако активные центры каждого мономера ИН располагаются на противоположных сторонах сферической поверхности димера на расстоянии около 35 \AA [6].

Подвижность в районе активного центра предполагает возможность конформационных изменений, которые могут быть вызваны связыванием субстрата. Такие примеры известны для некоторых ДНК-связывающих белков, например, ДНК-связывающий домен γ резолвазы принимает необходимую для каталитической активности белка конформацию

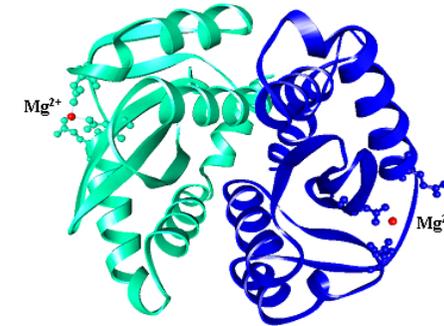


Рис. 3. Димер каталитического домена с ионами Mg^{2+} (PDB код 1BL3).

только после связывания ДНК-субстрата [18]. Возможно, по аналогии с этим ферментом, подвижная петля, расположенная в каталитическом домене ИН ВИЧ-1, также принимает необходимую для проявления каталитической активности фермента конформацию после связывания вирусной ДНК.

В целях дальнейшего уточнения структуры в работе [19] был закристаллизован каталитический домен ИН, содержащий дополнительную замену Trp131 на Lys, которая не приводила к потере каталитической активности ни в случае изолированного домена, ни в случае полного фермента. Кристаллизацию осуществляли в присутствии кофактора ИН – ионов Mg^{2+} . Определенная в данной работе структура практически полностью совпадает со структурой, описанной выше. Однако авторам удалось уточнить структуры двух участков, строение которых не было определено в работе [6]. Это участки, содержащие со 139 по 154 аминокислотный остаток и со 188 по 193. Показано, что первый из них задействован в образовании $\alpha 4$ -спирали, а второй формирует две дополнительные β -структуры, $\beta 6$ и $\beta 7$, расположенные между двумя $\alpha 5$ - и $\alpha 6$ -спиралями. Таким образом, третий участник каталитической триады, Glu152, входит в состав $\alpha 4$ -спирали, причем расположен он таким образом, что его боковая цепь направлена в сторону Asp64 и Asp116. Согласно данной структуре, каждый мономер каталитического домена связывает один ион Mg^{2+} , при этом в его координации задействованы Asp64, Asp116 (рис. 4) и две молекулы воды, в то время как Glu152 не принимает участия в связывании кофактора.

Несмотря на то, что структурные исследования показали связывание лишь одного иона магния в активном центре фермента, возможно,

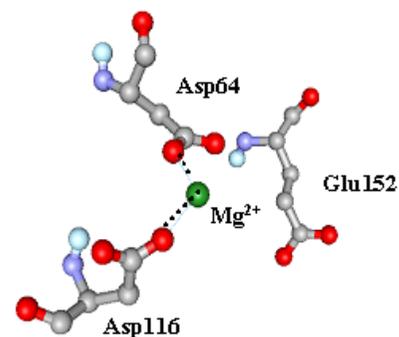


Рис. 4. Взаимное расположение аминокислотных остатков каталитического центра и иона Mg^{2+} в составе каталитического домена ИН ВИЧ-1 (PDB код 1BIU).

что в акте катализа могут участвовать два иона, причем второй из них привносится в активный центр фермента при связывании ДНК. Такое предположение можно сделать на основании данных о механизме действия транспозазы Tn5 [20], относящейся к тому же семейству, что и ИН ВИЧ-1.

Описанная структура мономера каталитического домена ИН ВИЧ-1 подобна доменам других белков семейства полинуклеотидилтрансфераз, таких как РНКазы H, транспозаза MuA и др. [6, 18, 21, 22, 23]. Сходство молекулярных механиз-

мов транспозиции ДНК и интеграции ретровирусной ДНК, продемонстрированное с помощью биохимических исследований [21, 24, 25], также отражается и на структурах их каталитических доменов.

СТРОЕНИЕ N-КОНЦЕВОГО ДОМЕНА

Роль N-концевого домена в процессах, которые осуществляет ИН, наименее ясна. Известно, что этот участок фермента содержит четыре аминокислотных остатка His12, His16, Cys40 и Cys43, так называемый NHSS мотив, который характерен для большинства ретровирусных интеграз [26]. Кроме того, известно, что ряд ферментов, содержащих подобный мотив, формируют Zn^{2+} -связывающий домен типа «цинковых пальцев», который наряду с координацией иона цинка выполняет и ДНК-связывающую функцию. ИН действительно связывает Zn^{2+} в стехиометрическом соотношении 1:1 [27]. Мутантный фермент, в котором либо отсутствовал N-концевой домен, либо были введены замены в NHSS мотив, терял способность как координировать ион цинка, так и осуществлять реакции 3'-концевого процессинга и переноса цепи [26, 27]. Это говорит о значимости функций, выполняемых данным доменом в структуре ИН.

Неструктурированный без цинка изолированный N-концевой домен в его присутствии формирует структуру, состоящую из α -спиралей. Такого влияния на полный фермент не зафиксировано, однако возможно, что, наряду с цинком, в стабилизации структуры этого домена в полноразмерном белке принимают участие и другие участки ИН [27].

Структура изолированного N-концевого домена была определена методом ЯМР-спектроскопии (рис. 5) [28]. Авторы показали, что в растворе N-концевой домен ИН образует димер. При этом мономеры, входящие в состав димера, могут находиться в двух структурно различных формах, E и D, способных переходить друг в друга. Каждый мономер состоит из четырех α -спиралей. Спирали 2, 3 и 4 идентичны для обеих форм и включают аминокислотные остатки 19–25, 30–39 и 41–45, соответственно. Различия в строении двух форм мономера связаны со структурной организацией $\alpha 1$ -спирали. Так, в E-форме ее образуют аминокислотные остатки со 2 по 14, тогда как в D-форме – со 2 по 8. Кроме того, четыре аминокислотных остатка (с 14 по 17), которые в случае E-формы входят в состав петли, соединяющей $\alpha 1$ и $\alpha 2$ -спирали, в случае D-формы образуют спиральный поворот.

Нижняя часть каждой субъединицы стабилизирована гидрофобным участком, который образуется при участии $\alpha 1$, $\alpha 2$ и $\alpha 3$ -спиралей. Структура верхней части мономера стабилизирована участием His12, His16, Cys40 и Cys43 в координации иона Zn^{2+} . Расположение иона цинка и цистеиновых аминокислотных остатков идентично для обеих форм, а положение His12 и His16 различно. Кроме того, в разных фор-

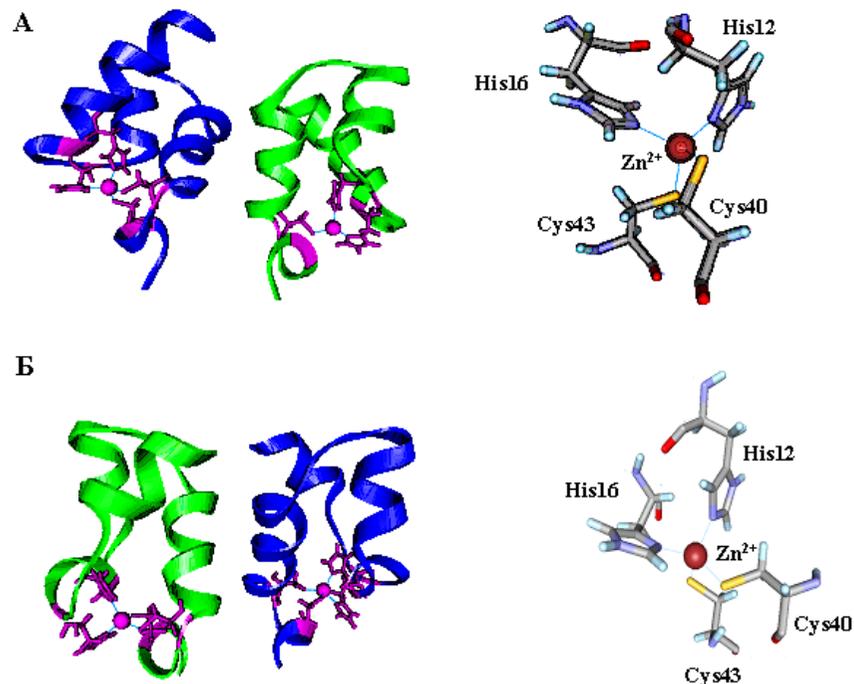


Рис. 5. Строение N-концевого домена ИН и взаимное пространственное расположение аминокислот, координирующих ион Zn^{2+} : А – E-форма; Б – D-форма (PDB код 1WJA).

мах мономера в координации иона Zn^{2+} принимают участие различные атомы азота имидазольного кольца. В случае E-формы His12 расположен в $\alpha 1$ -спирали, и в связывании цинка участвует ϵ -атом азота имидазольного кольца (рис. 5). В D-форме His12 находится в петле, соединяющей $\alpha 1$ и $\alpha 2$ -спирали, и для координации Zn^{2+} используется δ -атом азота имидазольного кольца. И в той, и в другой форме в координации иона Zn^{2+} принимают участие α -атомы азота имидазольного кольца His16, при этом в E-форме His16 расположен в петле между $\alpha 1$ и $\alpha 2$ -спиралями, а в D-форме он находится в спиральном повороте.

E-форма мономера N-концевого домена существует предпочтительно при низких температурах. При этом аминокислотные остатки с 9 по 14 также участвуют в образовании α -спирали, т.е. данный участок более структурирован, а в координации иона Zn^{2+} участвует менее предпочтительная таутомерная форма имидазольного кольца. При повышении температуры часть $\alpha 1$ -спирали плавится, имидазольное кольцо принимает обычную для себя таутомерную форму, Zn^{2+} координируется δ -атомом азота, E-форма мономера переходит в D-форму.

Формирование димера происходит за счет образования межмолекулярных связей между α -спиралями 1, 3 и 4. Отмечено формирование лишь гомодимеров, т.е. димеров, состоящих из мономеров либо в E-форме, либо в D-форме. Две субъединицы, входящие в состав димера, располагаются почти параллельно друг относительно друга. Все межсубъединичные связи между $\alpha 1$ и $\alpha 3$ -спиралями имеют гидрофобную природу. Надо отметить, что $\alpha 3$ -спирали двух субъединиц ориентированы друг относительно друга под углом около 60° . Их взаимное пространственное расположение стабилизировано гидрофобными взаимодействиями между Val31, Val32 и Ala38, а также, возможно, за счет образования солевого мостика между Glu35 одной субъединицы и Lys34 другой. В то же время, $\alpha 4$ -спирали двух мономеров располагаются под углом 110° , при этом Leu45 и Leu45' находятся напротив друг друга, кроме того, возможно образование водородной связи между амидными группами боковых цепей Gln44 и Gln44'.

Важно отметить структурное сходство N-концевого домена ИН ВИЧ-1 с некоторыми ДНК-связывающими доменами, такими как репрессор триптофанового оперона, содержащими так называемый фрагмент спираль–поворот–спираль [28]. Однако в то время как вторая спираль в подобных белках отвечает за узнавание ДНК, эквивалентная спираль в N-концевом домене ИН ВИЧ-1 участвует в формировании димера. Поэтому связывание ДНК димерной структурой N-концевого домена невозможно. Более того, распределение заряда на этой спирали неблагоприятно для связывания ДНК. Сходство с мотивом спираль–поворот–спираль ДНК-связывающих доменов некоторых белков может быть случайным и не является основанием для предположения, что N-концевой домен ИН связывает ДНК.

СТРОЕНИЕ С-КОНЦЕВОГО ДОМЕНА

В ряду ретровирусных интеграз С-концевой домен обладает наименьшей консервативностью по сравнению с N-концевым или каталитическим доменами. Данному домену приписывают функцию неспецифического связывания ДНК [12, 29, 30]. Показано, что участок С-концевого домена, содержащий с 220 по 270 аминокислотный остаток, способен связывать ДНК с эффективностью полного фермента [30].

Структура ДНК-связывающего домена ИН ВИЧ-1 была определена методом ЯМР-спектроскопии [31]. Несмотря на различия в первичной структуре, топологически С-концевой домен ИН имеет много общего со структурой домена Src homology 3 (SH3). Показано, что в растворе он образует димер (рис. 6). Полипептидная цепь каждого мономера формирует структурный мотив типа β -барреля, состоящий из пяти антипараллельных β -цепей (222–229, 232–245, 248–253, 256–262 и

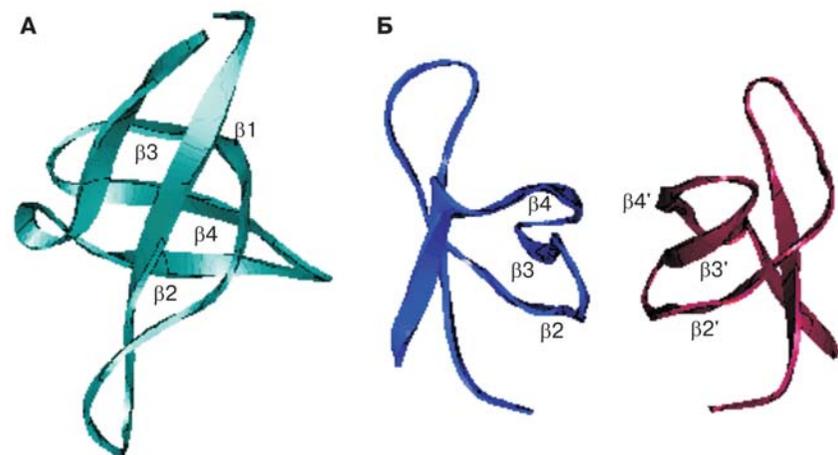


Рис. 6. Строение С-концевого домена ИН: А – структура мономера; Б – структура димера (PDB код 1IHV).

266–270 аминокислотные остатки, соответственно). Кор β -барреля образуют Val225, Tyr227, Ala239, Val249, Ile251, Val260 и Ala265.

В структуре мономера цепи $\beta 2$, $\beta 3$ и $\beta 4$ образуют антипараллельный β -лист. Перекрытие β -листов от двух субъединиц приводит к формированию димерной структуры, в которой мономеры ориентированы антипараллельно. Структура димера стабилизирована гидрофобными контактами между субъединицами, а также водородной связью между Glu246 одной субъединицы и Gln252 другой.

В результате образования двумя С-концевыми доменами димерной структуры происходит формирование большой седлообразной бороздки, ограниченной петлей, соединяющей $\beta 1$ и $\beta 2$ цепи. Размеры бороздки составляют $24 \times 23 \times 12$ Å. Она содержит положительно заряженные аминокислотные остатки Arg231, Lys244, Lys258, Arg262, Arg263, Lys264, Lys266, которые, по мнению авторов работы [31], могут участвовать в связывании ДНК. Это предположение подтверждается тем, что в случае близкой по структуре ИН ВИЧ-2 тройная замена Arg262Asp, Arg263Val и Lys264Glu приводит к тому, что эта ИН перестает связывать субстрат и, как следствие, осуществлять реакцию 3'-концевого процессинга [30]. Введение одной точечной мутации, Arg262Gly, в С-концевой домен ИН ВИЧ-1 снижает ее способность связывать ДНК на 50% [32]. Наряду с положительно заряженными аминокислотами, в связывании ДНК могут принимать участие и гидрофобные аминокислоты. Замена Leu234 на меньший по размеру Ala приводит к значительному снижению эффективности связывания ДНК [30].

СТРОЕНИЕ МУТАНТНЫХ БЕЛКОВ, СОДЕРЖАЩИХ ДВА ДОМЕНА ИНТЕГРАЗЫ

Все предыдущие разделы посвящены структуре изолированных доменов ИН: каталитического, N-концевого и С-концевого. Для кристаллизации этих белков либо использовались детергенты, либо в состав белков вводились точечные мутации, повышающие их растворимость и облегчающие кристаллизацию. Понятно, что строение доменов ИН в нативном ферменте может несколько отличаться от описанного выше. Однако можно предположить, что эти отличия не очень существенны или, по крайней мере, в организации отдельных доменов принимают участие те же элементы вторичной структуры, что и в полноразмерном белке. Тем не менее, знание структуры изолированных доменов не дает существенной информации об их расположении друг относительно друга и их взаимодействиях в структуре нативной ИН. В связи с этим представляют значительный интерес работы, в которых определены структуры белков, включающих два домена ИН.

С помощью РСА удалось определить структуру мутантного белка (ИН^{52–288}), содержащего два домена: каталитический и С-концевой, в которые были введены пять точечных мутаций: Cys56Ser, Trp131Asp, Phe139Asp, Phe185Lys, Cys280Ser (рис. 7) [33]. Строение отдельных доменов в составе ИН^{52–288}, в целом, подобно тому, которое было определено для изолированных доменов [19, 31]. Каталитический домен в этом белке, как и в случае изолированного домена, формирует симметричный димер. Два антипараллельных β -листа образуют С-концевой домен. В состав одного из них входят цепь $\beta 1$, N-концевая часть $\beta 2$ и цепь $\beta 5$, а в состав другого С-концевая часть цепи $\beta 2$ и цепи $\beta 3$ и

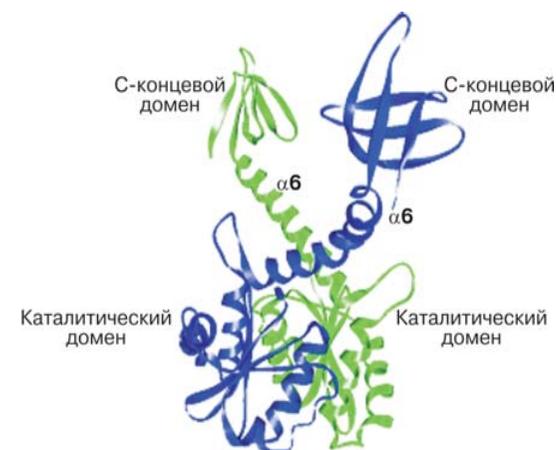


Рис. 7. Строение мутантного белка, содержащего каталитический и С-концевой домены ИН (PDB код 1EX4).

β4. Соединение двух доменов одного мономера осуществляется через α6-спираль, которая образована аминокислотными остатками 195–220 (рис. 7).

Двудоменный белок ИН^{52–288} формирует димер, в котором две β6-спирали, входящие в состав разных мономеров, ориентированы в пространстве под углом 90° относительно друг друга (рис. 7). В результате этого структура димера имеет Y-образную форму, в которой расстояние между центрами двух С-концевых доменов составляет ~55Å. Разделенные в пространстве они не могут образовывать димерную структуру, описанную в работе [31], которая, как предполагалось, играет важную роль в связывании ДНК.

В кристаллической структуре белка ИН^{52–288} зафиксированы димер–димерные контакты за счет взаимодействия С-концевых доменов соседних димеров. Однако большинство из них образуются через посредничество молекул детергента, в присутствии которого проводилась кристаллизация белка. Найден лишь один прямой контакт между двумя димерами в виде водородной связи между Arg224 и Trp235 [33].

Вторым двудоленным мутантом ИН, который удалось закристаллизовать, был белок (ИН^{1–212}), включающий каталитический и N-концевой домены и содержащий три точечные мутации: Trp131Asp, Phe139Asp и Phe185Lys [34]. Полный фермент, содержащий такие мутации, был активен в реакциях 3'-концевого процессинга и переноса цепи. Кристаллизацию проводили в фосфатном буфере в отсутствие детергента.

Согласно данным РСА, белок ИН^{1–212} в кристалле так же, как и в случае изолированных доменов, образует димер (рис. 8). Причем строение каталитического и N-концевого доменов на уровне мономера дву-

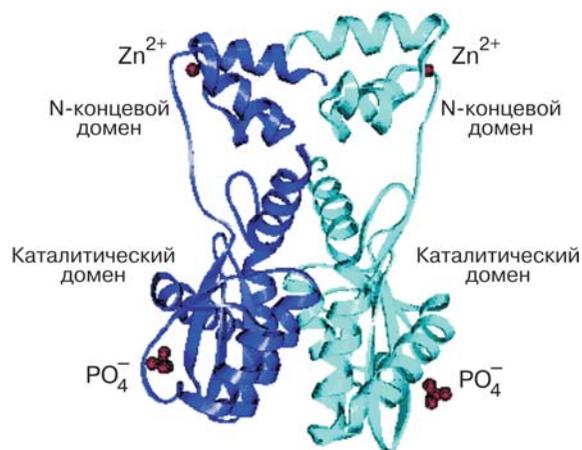


Рис. 8. Структура мутантного белка, содержащего каталитический и N-концевой домены ИН (PDB код 1K6Y).

доменного белка, в целом, совпадает с определенным ранее для изолированных доменов [19, 28]. Однако на уровне формирования димерной структуры белка ИН^{1–212} имеются различия по сравнению с димерами изолированных доменов. Так, в случае изолированного N-концевого домена, основную роль в образовании димера играют взаимодействия между α3-спиралями, а в структуре белка ИН^{1–212} димеризация N-концевых доменов происходит за счет контактов между α1-спиралями и α3-спиралями двух мономеров. При этом поверхность взаимодействия меньше и более гидрофобна. Однако такие аминокислоты, как Phe1, Leu2, Pro29, Val31 и Val32, участвуют в образовании димеров и для двудоменного белка, и для изолированного N-концевого домена. Участок с 47 по 55 аминокислоту, соединяющий два домена в белке ИН^{1–212}, неструктурирован. Несмотря на его гибкость и подвижность, он поддерживает расстояние между двумя доменами либо на уровне 9 Å, либо 13 Å, а также препятствует димеризации N-концевого домена в белке ИН^{1–212} в той мере, в какой она наблюдалась для изолированного домена.

Кристаллизация в фосфатном буфере позволила зафиксировать связывание фосфат-иона каталитическим доменом, на расстоянии около 7 Å от его центра. В координации фосфат-иона участвуют боковые цепи Thr66, His67, Lys159, атом азота пептидной связи His67 и молекулы воды, которые, возможно, замещаются ионами Mg²⁺ при их наличии. Связывание фосфат-иона вызывает конформационные изменения в структуре петли, содержащей Asp116 [34].

МУЛЬТИМЕРНАЯ ОРГАНИЗАЦИЯ ИНТЕГРАЗЫ

Авторы работы [10] использовали для проведения интеграции смесь двух мутантных интеграз, одна из которых была дефектна по N-концевому домену и содержала с 50 по 288 аминокислотный остаток (ИН^{50–288}), а другая – по С-концевому домену и включала с 1 по 173 аминокислотный остаток (ИН^{1–173}). Каждый из этих мутантов был каталитически неактивен. Однако оказалось, что смесь двух неактивных мутантных интеграз может осуществлять реакции 3'-концевого процессинга и переноса цепи, хотя и с меньшей эффективностью, чем фермент дикого типа. На основании этих результатов был сделан вывод о том, что ИН функционирует как мультимер.

Помимо этого оказалось, что при добавлении к ИН, содержащей летальные для ее активности мутации в каталитической триаде, мутантного белка ИН^{50–288}, имеющего в своем составе лишь каталитический и С-концевой домены, происходило восстановление активности ИН [10]. Из этого можно заключить, что для осуществления катализируемых ИН процессов, фермент задействует N-концевой и каталитический домены, принадлежащие разным мономерам (транс-

состояние). Аналогичные эксперименты с С-концевым доменом показали, что в катализе может принимать участие С-домен, принадлежащий как тому же мономеру, что и каталитический (цис-состояние), так и другому мономеру (транс-состояние) [10].

Из описанного выше можно сделать вывод, что в активном состоянии ИН находится, по меньшей мере, в виде димера. Этому выводу не противоречат и результаты структурных исследований, согласно которым и каждый домен, и совокупность двух доменов образуют димеры. Однако в димере каталитического домена активные центры расположены на противоположных сторонах глобулы и расстояние между ними составляет $\sim 35\text{--}45 \text{ \AA}$ [6], в то время как концы вирусной кДНК встраиваются в ДНК-мишень на расстоянии 5 нуклеотидных пар друг от друга ($\sim 15\text{--}20 \text{ \AA}$) [26, 33]. Если предположить, что такое расстояние между двумя каталитическими центрами димера сохраняется и в случае полного фермента в комплексе с ДНК, то один и тот же димер, скорее всего, не может осуществлять встраивание обоих концов вирусной кДНК в клеточную по пространственным соображениям [35]. Таким образом, для интеграции вирусной кДНК в ДНК-мишень необходима тетрамерная организация ИН.

Этот вывод подтверждают результаты работы [36]. Авторы изучали ИН, экспрессированную в клетках человека. Из ядерного экстракта ИН выделяется в комплексе с белком LEDGF/p75, стоксовский радиус комплекса $\sim 61 \text{ \AA}$. Данный комплекс легко диссоциирует с образованием устойчивой глобулярной структуры радиусом $\sim 41 \text{ \AA}$. По результатам гель-фильтрации и кросс-линкинга, данная структура представляет собой гомотетрамер ИН ВИЧ-1. Кроме того, авторы, проводя аналогию с транспозазами фага μ и Tn5, делают предположение, что тетрамер является активной формой ИН в реакции 3'-концевого процессинга, в то время как для переноса цепи необходима октамерная организация фермента [36].

Для изучения олигомерного состояния рекомбинантной ИН был также использован метод поляризации флюоресценции [37]. Фермент, который использовали для этих исследований, был выделен в присутствии детергента CHAPS. Выделенный в таких условиях белок был гомогенен и имел мономерное строение в интервале концентрации $40\text{--}200 \text{ нМ}$, то есть при тех концентрациях ИН, при которых осуществляются реакции 3'-процессинга и переноса цепи *in vitro*. Необходимо заметить, что этот фермент проявлял каталитическую активность в присутствии ионов Mn^{2+} , а в присутствии природного кофактора, ионов Mg^{2+} , был практически неактивен.

Дальнейшее изучение влияния детергента на олигомерное состояние рекомбинантной ИН проводилось с использованием фермента, выделенного тремя различными способами: IN_{CHAPS} — в присутствии

детергента CHAPS; IN_{dial} — в присутствии CHAPS, но очищенного от него диализом; IN_{Zn} — без детергента, в присутствии ионов Zn^{2+} [38]. Оказалось, что IN_{Zn} имеет в растворе тетрамерную организацию при концентрациях $\leq 200 \text{ нМ}$, в отличие от мономерного состояния IN_{CHAPS} . Причем IN_{Zn} была активна как в присутствии ионов Mn^{2+} , так и ионов Mg^{2+} — природного кофактора ИН. Кроме того, IN_{Zn} сохраняла активность в Mg^{2+} -содержащем буфере даже при последующем добавлении детергента. Следует отметить, что фермент IN_{dial} , очищенный после выделения от детергента, но несодержащий ионов цинка, также проявлял некоторую активность в присутствии ионов магния. Возможно, это связано с тем, что в растворе данный белок находится в состоянии равновесия между мономерной и димерной формами. Однако при вторичном добавлении CHAPS IN_{dial} становилась полностью неактивной в присутствии Mg^{2+} .

Помимо изучения олигомерного состояния ИН, выделенной в различных условиях, авторы работы [38] исследовали связывание ДНК данными ферментами. Оказалось, что IN_{CHAPS} , имеющая мономерное строение в растворе, способна связывать ДНК лишь в присутствии ионов марганца, но не магния. В то время как фермент IN_{Zn} , выделенный без детергента и имеющий тетрамерное строение, связывал ДНК в присутствии ионов обоих металлов. Эти результаты объясняют, почему IN_{Zn} была активна, а IN_{CHAPS} неактивна с природным кофактором. Возможно, что выделенный с детергентом белок проявляет свою активность только в присутствии Mn^{2+} вследствие того, что способен связывать ДНК только в присутствии данного иона.

Таким образом, на основании результатов различных исследований, можно предположить, что ИН функционирует в виде, по крайней мере, тетрамера. Однако точное мультимерное состояние ИН в реакциях 3'-концевого процессинга и переноса цепи еще не определено.

МОДЕЛИРОВАНИЕ СТРУКТУРЫ ПОЛНОРАЗМЕРНОГО ФЕРМЕНТА И ЕГО КОМПЛЕКСА С ДНК

Авторы работы [33] на основании своих данных по строению мутантного белка, содержащего каталитический и N-концевой домены ИН ВИЧ-1, а также данных работы [32], в которой была определена структура белка ($\text{IN}^{52\text{--}288}$), включающего каталитический и С-концевой домены, смоделировали строение полного фермента путем суперпозиции двух структур. В этой смоделированной структуре N-концевые домены вписываются в димерную структуру ИН между каталитическими доменами и перекрещивающимися С-концевыми доменами без стерических затруднений. Кроме того, авторы предполагают, что N-концевые домены могут дополнительно стабилизировать αB -спирали, соединяющие каталитические и С-концевые домены [33].

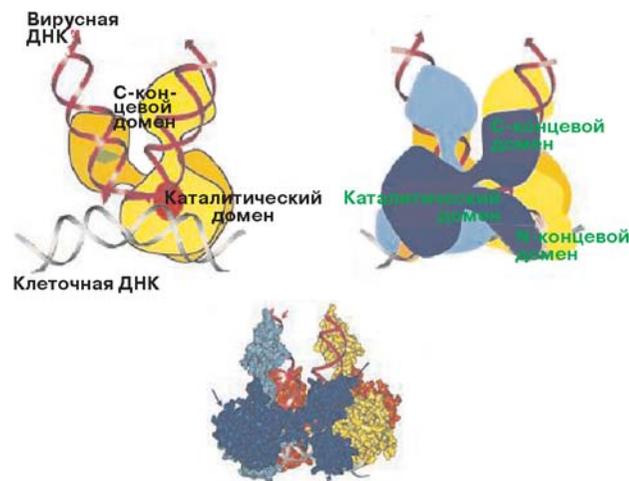


Рис. 9. Строение активного комплекса ИН с ДНК по данным работы [39].

Авторы работы [39], основываясь на результатах своей работы, а также используя результаты структурных исследований, предложили модель активного комплекса ИН с ДНК (рис. 9). Согласно данной модели, активный комплекс ИН – симметричный тетрамер, состоящий из двух димерных структур. В димере одна молекула (активная) ИН предоставляет каталитический домен, а другая – С-концевой домен для связывания ДНК (транс-взаимодействие). Таким образом, одна димерная структура осуществляет связывание и процессинг одного конца вирусной ДНК, а тетрамер, в целом, связывает и процессирует оба конца вирусной ДНК. Объединение двух димеров в тетрамерную структуру происходит таким образом, чтобы два активных мономера располагались друг напротив друга и имели возможность осуществить реакцию переноса цепи для обоих концов вирусной ДНК. В связывании клеточной ДНК, согласно данной модели, принимают участие каталитический домен активного мономера и N-концевой домен.

Моделирование структуры ИН в комплексе с вирусной и клеточной ДНК было проведено в работе [40]. Этот процесс авторы осуществляли поэтапно. На первом этапе они смоделировали структуру мономера в комплексе с ДНК (рис. 10). Согласно данной модели, в связывании вирусной ДНК принимает участие лишь каталитический домен, в связывании же клеточной ДНК участвуют все три домена. Однако авторы подчеркивают, что представленная структура не является единственно возможной. Так смещение N- и С-концевых доменов относительно каталитического домена приводит к альтернативной

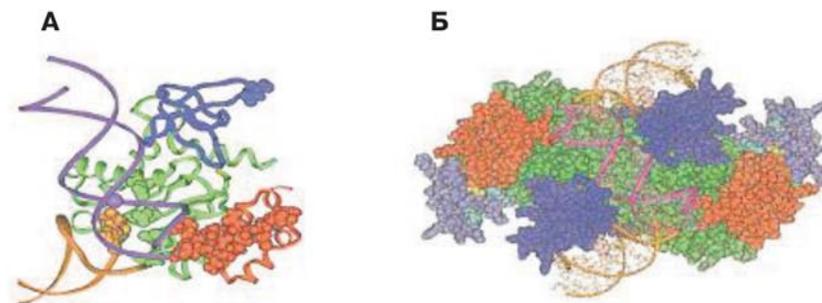


Рис. 10. Строение ИН в комплексе с вирусной и клеточной ДНК по данным работы [40].

Структура каталитического домена показана зеленым цветом, N-концевого – красным, С-концевого – синим. Вирусная ДНК показана желтым цветом, а клеточная ДНК – фиолетовым. А – структура мономера, Б – структура тетрамера.

структуре, обладающей, однако, теми же гидродинамическими параметрами. Моделирование структуры димера ИН авторы указанной работы проводили путем суперпозиции двух мономерных структур, а также структуры димера каталитического домена [19], полагая, что основную роль в образовании димерной структуры ИН выполняют контакты между двумя каталитическими доменами.

В основу моделирования тетрамерной формы ИН были положены следующие требования: во-первых, расстояние между двумя каталитическими центрами двух димеров должно быть не более 20–25 Å, так чтобы один тетрамер фермента мог осуществить реакцию переноса цепи для обоих концов вирусной ДНК; во-вторых, в образовании поверхности контакта между двумя димерами должны принимать участие аминокислотные остатки N- и С-концевых доменов, так как показана их роль в мультимеризации фермента; в-третьих, структура должна иметь ось симметрии второго порядка. Структура активного комплекса ИН с ДНК, смоделированная на основе этих требований, представлена на рис. 9.

В работе [41] была предложена модель взаимодействия димера ИН ВИЧ-1 с 27-звенным процессированным субстратом. В рамках этой модели вирусная ДНК взаимодействует со всеми тремя доменами ИН, а точнее с каталитическим и N-концевым доменами одного мономера и с С-концевым доменом другого мономера ИН. 3'-Конец вирусной ДНК, который встраивается в ДНК-мишень при интеграции, расположен вблизи аминокислотных остатков Asp64, Asp116 и Glu152, формирующих активный центр ИН. В рамках полученной модели с вирусной ДНК взаимодействуют аминокислоты Lys156, Lys159,

Lys160, Lys186 и Lys188 каталитического домена ИН, Arg20 из N-концевого домена и Ser230, Arg231, Trp243, Lys244, Lys263 и Lys264 из C-концевого домена.

В работе [42] смоделировали сайт связывания процессированной вирусной ДНК каталитическим доменом ИН. Однако в отличие от предыдущего исследования была построена модель ИН, содержащая два иона Mg^{2+} , необходимых для катализа. Согласно полученной модели, 3'-концевая гидроксильная группа образует водородные связи с Asp64 и ионом Mg^{2+} , который координируется этой же аминокислотой. Также 3'-концевой нуклеозид контактирует с Lys156. Кроме того, предполагаются взаимодействия между аминокислотой Lys159 и атомом H^+ углеводного остатка цитидина в непроцессированной цепи. Авторы также предполагают, что аминокислоты Gln62 и Gln148 участвуют в формировании мультимера ИН и в стабилизации вирусной ДНК.

III. ФУНКЦИОНИРОВАНИЕ ИНТЕГРАЗЫ

РЕАКЦИИ, КАТАЛИЗИРУЕМЫЕ ИНТЕГРАЗОЙ

Как отмечалось выше, ИН ВИЧ-1 осуществляет встраивание (интеграцию) вирусной кДНК в хромосому зараженной клетки. Интеграция протекает в несколько стадий. Начинается процесс в составе сложного нуклеопротеинового комплекса, называемого прединтеграционным комплексом (ПИК). До настоящего времени структурная организация комплекса до конца не установлена. Функциональный ПИК, способный катализировать интеграцию вирусной кДНК во внешнюю ДНК-мишень *in vitro*, был выделен в виде цитоплазматического экстракта из ВИЧ-инфицированных клеток [43]. Обнаружено, что он содержит вирусную кДНК, вирусный белок Vrg, нуклеокапсидный белок p7, обратную транскриптазу и ИН [44]. Кроме того, в состав ПИК входят клеточные белки, такие как HMG/1Y, Ini-1, гистоны и VAF [45, 46]. ПИК обладает уникальной способностью транспортироваться в ядро через ядерную мембрану неделящихся клеток. В транспорте, по-видимому, принимают участие вирусные белки: Vrg, матриксный белок p17 и ИН. Роль клеточных факторов в процессе интеграции остается малоизученной [47].

Процесс интеграции начинается в цитоплазме. Вирусная ДНК содержит на обоих концах прямые повторы, состоящие из нескольких сотен пар оснований. Каждый длинный концевой повтор (ДКП) содержит три участка U3, R и U5 (рис. 11). На расстоянии двух нуклеотидов от 3'-концов каждой цепи ДНК находится консервативный динуклеотид CA (рис. 11), который встречается в ДКП всех ретровирусных интеграз [48]. ИН в составе прединтеграционного комплекса узнает нуклеотидные последовательности, расположенные на концах участков

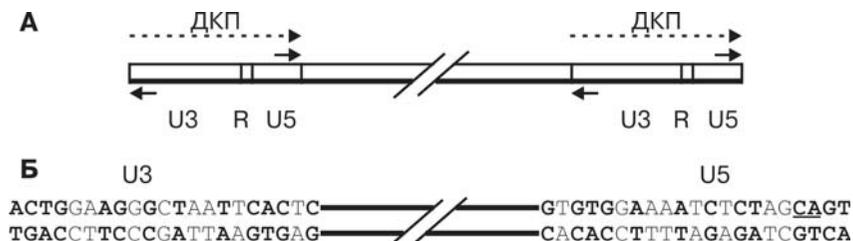


Рис. 11. Схематическое изображение ДНК ВИЧ-1.

А. Расположение участков U3, R и U5 длинных концевых повторов (ДКП) в молекуле ДНК.

Б. Концевые последовательности U3- и U5-фрагментов ДКП. Жирным шрифтом выделены нуклеотиды, совпадающие в U3- и U5-фрагментах ДКП. Консервативный динуклеотид CA подчеркнут.

U3 и U5 вирусной кДНК, связывается с ними и затем катализирует реакцию 3'-концевого процессинга [48]. Эта реакция представляет собой эндонуклеазное расщепление вирусной кДНК, в результате которого с 3'-конца каждой цепи удаляется динуклеотид GT (рис. 12). Расщепление субстрата происходит в результате нуклеофильной атаки молекулой воды фосфатной группы между 2-м и 3-м нуклеотидами. Затем ПИК транспортируется в ядро, где ИН катализирует реакцию переэтерификации, которая включает нуклеофильную атаку 3'-гидроксильных групп процессированных цепей ДНК на межнуклеотидные фосфаты обеих цепей клеточной ДНК с образованием ковалентного продукта [48].

Необходимо подчеркнуть, что реакция 3'-концевого процессинга сиквенс-специфична, т.е. ей подвергаются только последовательности ДНК, первичная структура которых аналогична U3 или U5 последовательностям ДНК ВИЧ-1. В то же время к последовательности клеточной ДНК не обнаружено никаких требований для того, чтобы она могла служить мишенью для интеграции. Встраивание вирусной кДНК, называемое переносом цепи, может осуществляться в любое место клеточной ДНК, однако, преимущественно направлено на определенные последовательности генома с нарушенным или ослабленным стэкингом [49]. В выборе участка встраивания, по-видимому, принимает участие сама ИН, а также ряд клеточных белков (LexA, Zif) [50, 51, 52]. Переэтерификации подвергаются межнуклеотидные связи, расположенные в различных цепях ДНК-мишени на расстоянии 5 пар нуклеотидов одна от другой.

Для завершения процесса интеграции требуются: процессинг 5'-концов вирусной ДНК, полимеразная достройка пяти недостающих нуклеотидов и лигирование, которые осуществляются при участии клеточных белков [53, 54].

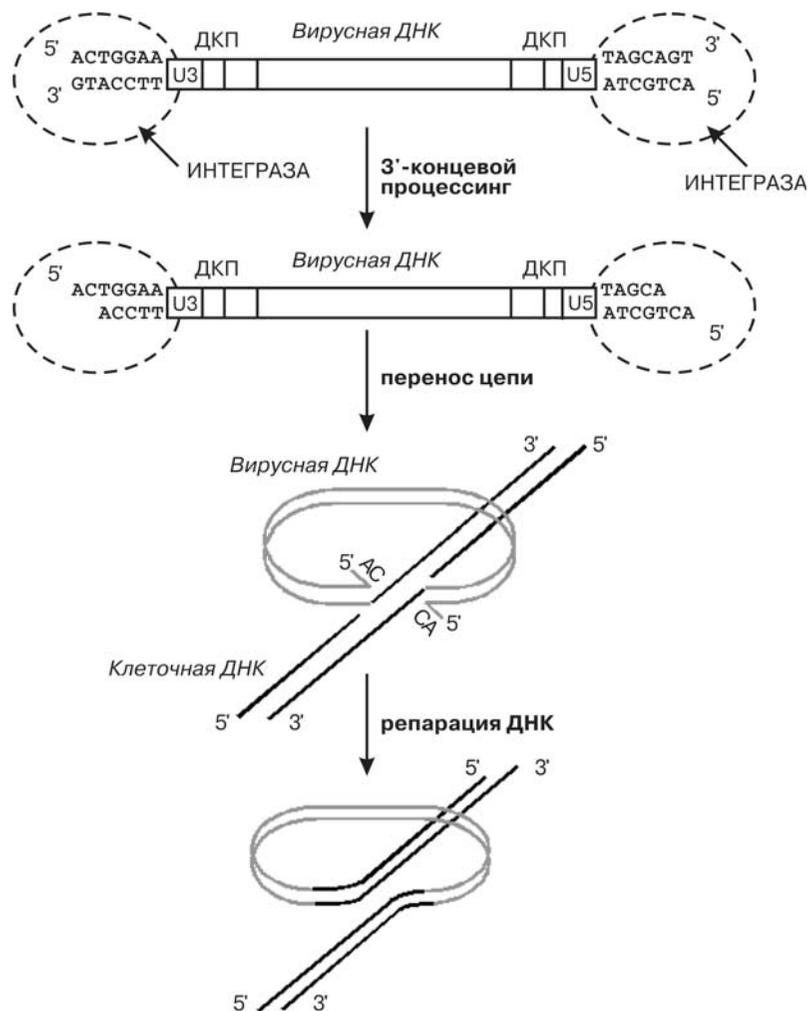


Рис. 12. Схема интеграции ДНК ВИЧ-1 в геном клетки.

Реакции, катализируемые интегразой *in vitro*

Реакции 3'-концевого процессинга и переноса цепи могут быть воспроизведены *in vitro* с использованием очищенного рекомбинантного фермента в присутствии кофактора — ионов Mn^{2+} или Mg^{2+} . В качестве субстратов ИН используют, как правило, ДНК-дуплексы длиной в 21 пару оснований, последовательности которых соответствуют фрагментам U3 или U5, содержащим сайты узнавания ИН [55,

56] (рис. 11). Эти дуплексы соответственно называют U3- и U5-субстратами ИН. В реакциях *in vitro* эти дуплексы могут выступать как в роли ДНК-субстрата (вирусной ДНК), так и в роли ДНК-мишени (клеточной ДНК), т.е. *in vitro* ИН способна встраивать ДНК-субстрат самого в себя.

Кроме реакций 3'-процессинга и переноса цепи *in vitro* в присутствии ионов Mn^{2+} , ИН может осуществлять и ряд других превращений ДНК. Например, в качестве атакующей группировки для расщепления межнуклеотидной связи между 2-м и 3-м нуклеотидами может быть использована гидроксильная группа спиртов, некоторых аминокислот или 3'-концевая гидроксильная группа самой вирусной ДНК [57–59]. Кроме того, ИН способна встраивать многоатомные спирты, такие как глицерин или этиленгликоль, в любую последовательность ДНК, используя их в качестве нуклеофилов для атаки ДНК-мишени вместо 3'-гидроксильных групп процессированных цепей вирусной ДНК [60]. Эта активность ИН называется неспецифический алкогольиз.

Также *in vitro* ИН способна осуществлять реакцию дезинтеграции, которая является обратной реакции переноса цепи [61]. В результате дезинтеграции происходит удаление встроенного вирусного фрагмента из Y-подобного субстрата, имитирующего ковалентный продукт интеграции, с восстановлением вирусного и клеточного фрагментов ДНК.

Механизм реакций, катализируемых интегразой ВИЧ-1

Известно, что введение межнуклеотидных тиофосфатных группировок в процессируемую цепь U5-субстрата вместо всех фосфодиэфирных связей способно полностью ингибировать 3'-процессинг [62]. Однако если вводить всего лишь одно тиофосфатное звено между консервативными СА и GT-динуклеотидами, то такой субстрат будет участвовать в реакциях 3'-процессинга и переноса цепи [59]. Причем в результате этих реакций с участием ДНК-дуплексов, содержащих R-изомер тиофосфатов, образуются продукты, содержащие S-изомер. Эти данные указывают на то, что обе эти реакции нуклеофильного замещения протекают по одностадийному S_N2 механизму с обращением конфигурации при атоме фосфора.

С помощью кинетических измерений также было установлено, что R-изомер является гораздо более предпочтительным по сравнению с S-изомером и в реакции 3'-процессинга, и при переносе цепи [63]. Соответственно для дезинтеграции благоприятной была S-тиофосфатная ДНК. Об этом судили по разнице скоростей расщепления Y-образных ДНК, содержащих фосфодиэфирные или тиофосфатные связи.

Из данных по стереоспецифичности катализируемых ИН реакций был сделан следующий вывод: 3'-гидроксил консервативного 2'-дезоксиденозина ДНК-субстрата и при 3'-процессинге, и при переносе

цепи занимает аксиальную позицию в структуре тригональной бипирамиды, характерной для переходного состояния у атома фосфора. Это вызвано тем, что в этих реакциях 3'-гидроксил соответственно является либо уходящей группой, либо нуклеофилом. По этой же логике другое аксиальное положение при процессинге занимает молекула воды, а при переносе цепи — 5'-концевой фрагмент атакуемой цепи ДНК-мишени (рис. 12). Соответственно экваториальную позицию при процессинге занимает динуклеотид GT, а при переносе цепи — 3'-концевой фрагмент ДНК-мишени. В связи с этим высказываются предположение, что динуклеотид GT и ДНК-мишень связываются в разных местах белка, причем один из этих сайтов связывает одноцепочечную ДНК (GT), а другой — двухцепочечную (ДНК-мишень) [63]. Существует, однако, альтернативное предположение, что между 3'-процессингом и переносом цепи в активном центре ИН происходит структурная реорганизация [59], поскольку в ходе процессинга фосфодиэфирная связь в ДНК-субстрате является атакуемой группировкой, а при интеграции образовавшаяся 3'-гидроксильная группа той же молекулы ДНК — уже атакующая.

Связывание ДНК интегразой

Стадия образования стабильного прединтеграционного комплекса, с которой начинается процесс интеграции, характерна для всех интеграз. Концы вирусной ДНК остаются в комплексе с ИН в течение длительного промежутка времени между стадиями 3'-процессинга и переноса цепи. Способность ИН связывать ДНК *in vitro* была показана с использованием синтетических ДНК-субстратов, моделирующих один из концов вирусной ДНК, методами хроматографии и седиментации [64], связывания на нитроцеллюлозных фильтрах [65, 66], торможения в геле [67], фотоиндуцируемого присоединения [45, 68–70] и др. [30, 71–74].

Связывание ИН с ДНК может происходить и в отсутствие ионов двухвалентного металла [69, 70], однако для образования стабильного комплекса фермента с ДНК металл-кофактор необходим [70, 72, 75]. Обычно такие комплексы формируются в присутствии ионов Mn^{2+} и Mg^{2+} , но возможно их образование также в присутствии Ca^{2+} [76].

Домены интегразы, отвечающие за связывание ДНК

В ходе осуществления процесса интеграции ИН взаимодействует с двумя молекулами ДНК: вирусной (субстратом) и клеточной (мишенью). К сожалению, картина этих взаимодействий до сих пор не ясна. Неизвестно, имеет ли ИН различные или перекрывающиеся сайты связывания вирусной и клеточной ДНК [26, 72, 74, 77–79,] или, возможно, она имеет один сайт связывания ДНК [80, 12]. В пользу

последнего предположения говорит тот факт, что *in vitro* ИН приблизительно с одинаковой аффинностью связывает любую молекулу ДНК, и обладающую, и необладающую последовательностью ДНК-субстрата [65, 81, 82]. До сих пор неизвестно, какие домены ИН участвуют в связывании ДНК-субстрата и ДНК-мишени.

Как уже отмечалось выше, N-концевой домен ИН содержит характерный НССС-мотив, напоминающий цинк-связывающий мотив некоторых ДНК-связывающих белков. Однако, скорее всего, этот домен не играет существенной роли в связывании ДНК. Так, было показано, что ИН с делегированным N-концевым доменом сохраняет способность связывать ДНК на уровне полного фермента [29]. Также была исследована активность химерных белков, в состав которых в различных комбинациях входили домены двух интеграз: ВИЧ-1 и вируса Visna [78]. Один из них содержал каталитический и C-концевой домены ИН вируса Visna, а N-концевой — ИН ВИЧ-1. Авторы показали, что данный фермент в качестве субстрата в реакциях 3'-концевого процессинга и переноса цепи предпочитал ДНК вируса Visna. На основании этого они сделали заключение, что для специфического узнавания вирусной ДНК нет необходимости в N-концевом домене. Потеря же активности ферментом при мутациях в N-концевом домене, возможно, связана с изменением мультимерности его активной формы, т.к. показано, что присутствие ионов цинка стимулирует мультимеризацию ИН [27].

Считается, что ДНК-связывающие свойства ИН обеспечиваются в первую очередь C-концевым доменом [12, 29, 30, 65, 66, 71, 78, 81]. На поверхности C-концевого домена находится большое количество положительно заряженных аминокислот, необходимых для неспецифического связывания ДНК за счет электростатических взаимодействий.

Специфическое связывание ДНК ВИЧ-1, очевидно, осуществляется каталитическим доменом ИН, который играет важную роль в узнавании динуклеотида CA рядом с расщепляемой при процессинге связью [83]. Существуют также данные, полученные из экспериментов с участием химерных интеграз, которые свидетельствуют о том, что каталитический домен может также участвовать в связывании ДНК-мишени [84, 85].

Участие аминокислотных остатков интегразы во взаимодействии с ДНК

Одним из наиболее информативных на настоящий день методов выявления контактов между ИН и вирусной ДНК или ДНК-мишенью является метод фотоиндуцируемых сшивок. Так для определения остатков лизина, контактирующих с вирусной ДНК, были использованы аналоги U5-субстрата, содержащие остатки 2'-дезоксидеокси-5-йоду-

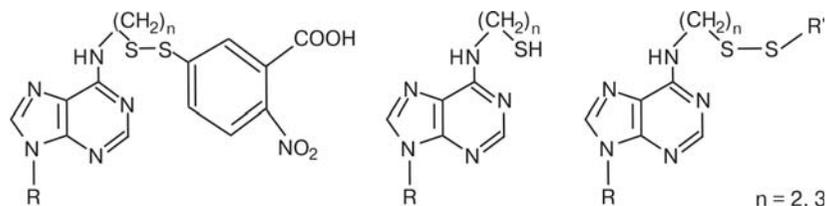


Рис. 13. Модифицированные нуклеозиды, содержащие группы, способные образовывать дисульфидную связь с остатками цистеина в ИН, использованные в работе [39].

ридина [86]. Под действием УФ-излучения с длиной волны 302 нм происходила избирательная активация этих остатков и их последующая реакция с аминокислотами ИН. Место присоединения ДНК к ИН определяли в результате ее протеолиза с последующим секвенированием связанного с ДНК пептида по Эдману. Оказалось, что Lys159 взаимодействует с 2'-дезоксидеозидом, введенным в четвертую позицию процессируемой цепи, что соответствует консервативному 2'-дезоксидеозиду. С помощью сайт-специфического мутагенеза была выявлена важность Lys159 для всех трех реакций, катализируемых ИН *in vitro*. Также предполагается, что 2'-деоксиаденозин из консервативного динуклеотида СА и соответственно атакуемый при процессинге фосфат расположены около Glu 152 из каталитической триады D,D(35)E [86].

Позднее, также с использованием метода фотоприсоединения, была показана возможность взаимодействия Tyr143 и Gln148 с 5'-концевым 2'-деоксиаденозином непроецессуемой цепи U5-субстрата, а также пептидов ИН⁵¹⁻⁶⁴ и ИН²⁴⁷⁻²⁷⁰ со вторым нуклеозидом непроецессуемой и с седьмым нуклеозидом процессируемой цепей, соответственно [87].

В работе [39] в С-концевом домене фермента 246-ю природную аминокислоту заменяли на остаток цистеина (мутация E246C). В U5-дуплекс вместо остатков 2'-деоксиаденозина вводили показанные на рис. 13 аналоги нуклеозидов, содержащих в гетероциклических основаниях группы, способные реагировать с Cys. Образование ковалентной связи между Cys246 и модифицированным U5-субстратом лучше всего происходило в том случае, когда реакционноспособный аналог нуклеотида вводился вместо 7-го 2'-деоксиаденозина в непроецессуемой цепи. На основании этого результата был сделан вывод, что в структуре комплекса ИН с ДНК ВИЧ-1 аминокислотный остаток Glu246 сближен с нуклеотидом в 7-м положении непроецессуемой цепи вирусной ДНК.

Для выяснения контактов вирусной ДНК с остатками лизина в ИН была синтезирована серия аналогов U5-субстрата, в которых

природные нуклеозиды процессируемой цепи в положениях с 1 по 9, а также в 13 заменялись на 2'-деоксиуридиновые звенья [74]. Полученные дуплексы обрабатывали урацил-ДНК-гликозилазой, в результате чего в составе субстрата возникали сайты, не содержащие гетероциклических оснований (АП-сайты). Известно, что такая модификация ДНК не приводит к изменению В-формы ее спирали [88]. Образующаяся в таких АП-сайтах альдегидная группа способна образовать основание Шиффа с ε-аминогруппой близкорасположенного остатка лизина. Наибольшая эффективность образования ковалентно связанных комплексов ИН с аналогами U5-субстрата наблюдалась при размещении АП-сайтов вблизи консервативного динуклеотида СА (1–5 позиции) и в 13 позиции, что указывало на то, что в этих положениях аминокислоты белка расположены наиболее близко к ДНК. Известно, что в непосредственной близости от активного центра ИН располагаются остатки лизинов 103, 111, 127, 136, 156 [6]. Использование ИН ВИЧ с мутацией K136E позволило выяснить, что Lys136 взаимодействует с нуклеотидами, расположенными в 3–6 положениях субстрата [74].

Более или менее подробную картину взаимодействия всех участников интеграционного процесса удалось получить, изучая реакцию дезинтеграции с использованием специальным образом сконструированных субстратов (рис. 14А) [79]. Такие Y-образные субстраты одновременно содержали последовательности, представляющие собой ДНК-субстрат и ДНК-мишень, в различные положения которых вводили фотоактивируемые группировки: 2'-деокси-5-йодцитидин или азидофенацильную. В результате было идентифицировано шесть пептидов из всех доменов ИН, участвующих во взаимодействии с дезинтеграционным субстратом (таблица). На рис. 14Б указано, с какими пептидами реагировали фотоактивируемые группы в субстрате. С участком дезинтеграционного субстрата, имитирующим встраиваемую вирусную ДНК (ДНК-субстрат), взаимодействовали только пептиды 2, 3, 5 и 6. Пептиды 1 и 4 взаимодействовали только с той частью дезинтеграционного субстрата, которая имитировала ДНК-мишень. Пептиды 2 и 3, включающие в себя Asp64 и Glu152 из каталитической триады D,D(35)E, взаимодействовали вблизи консервативного фрагмента СА и точки интеграции, что указывало на их участие в каталитическом акте. Интересна картина связывания пептидов С-концевого домена (4–6) с ДНК. Для пептида 4 были установлены контакты только с участком, имитирующим ДНК-мишень, в то время как пептид 5 связывался только с участком, имитирующим ДНК-субстрат. Для пептида 6 были обнаружены контакты с обоими участками. Эти данные подтверждают предположение о том, что С-концевой домен

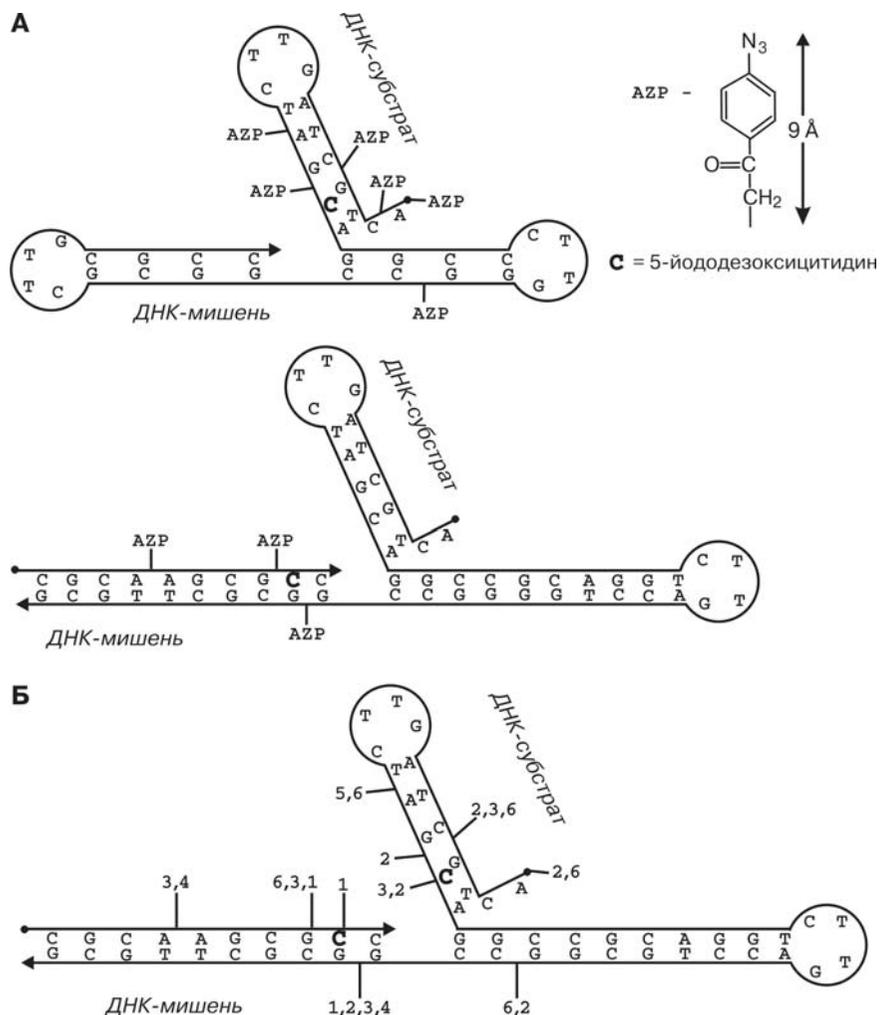


Рис. 14. Структуры дезинтеграционных субстратов, использованных в работе [79] для изучения контактов отдельных участков интегразы ВИЧ-1 с вирусной и клеточной ДНК.

А – места расположения фотоактивных группировок в структуре субстратов.

Б – места ковалентного присоединения пептидов, полученных после гидролиза ИН, по данным работы [79].

ИН принимает участие в связывании обеих молекул ДНК (субстрата и мишени) за счет электростатического взаимодействия заряженных остатков лизина и аргинина, которыми богат этот домен, с фосфатными группами ДНК [33, 86].

Таблица
Пептиды, связанные с дезинтеграционным субстратом в результате реакции фотоприсоединения

№ пептида	Аминокислоты	Домен ИН
1	1–11	N
2	49–69	Каталитический
3	139–152	Каталитический
4	213–246	C
5	247–270	C
6	271–288	C

ВЗАИМОДЕЙСТВИЕ ИНТЕГРАЗЫ С ВИРУСНОЙ ДНК

Взаимодействие ИН с вирусной ДНК происходит строго специфично. ИН узнает определенные последовательности, расположенные на концах фрагментов U3 и U5 из длинных концевых повторов. Не все нуклеотиды концевой последовательности вирусной ДНК в одинаковой мере участвуют во взаимодействии с ИН. Было показано, что 21-звенный дуплекс, аналогичный по нуклеотидной последовательности концевому повтору U5, процессировался лучше, чем 21-звенный дуплекс со структурой повтора U3 [80]. Большое количество работ посвящено изучению того, как влияет нуклеотидная последовательности ДНК-субстрата на протекание процессинга.

Установлено, что ИН взаимодействует главным образом с углеводородфосфатным остовом олигонуклеотидов за счет электростатических взаимодействий фосфатных групп ДНК с положительно заряженными аминокислотами ИН [82]. В этой же работе показано, что ИН взаимодействует с обеими цепями ДНК-субстрата, но преимущественно с процессируемой цепью. Авторы предполагают, что по аналогии с ферментами репарации и рестрикции, вторая цепь ДНК важна для того, чтобы ДНК-субстрат принял конформацию, оптимальную для катализа. Кроме того, предполагается, что ИН имеет отдельные сайты связывания для 3'-концевого динуклеотида ГТ и для консервативного динуклеотида СА. Возможно, 3'-концевой тимидин формирует специфические контакты с ИН, и после удаления динуклеотида ГТ в ходе реакции 3'-концевого процессинга происходит перестройка активного центра ИН, в результате которой осуществляются взаимодействия с динуклеотидом СА [82].

Дуплекс, выступающий в роли субстрата ИН ВИЧ-1, должен иметь, по крайней мере, 15 пар оснований [89, 80]. Однако более длинные субстраты лучше взаимодействуют с ИН. Например, добавление 13 неспецифических пар оснований к 15-звенной концевой последовательности

тельности вирусной ДНК стимулирует активность ИН ВИЧ-1 [57]. Также было показано, что при увеличении длины субстрата с 24 до 35 пар оснований увеличивается активность ИН в присутствии иона магния [77].

Для осуществления реакций 3'-процессинга и переноса цепи ДНК-субстрат должен быть двухцепочечным [89–91]. Однако не обязательно, чтобы все нуклеотиды в структуре субстрата образовывали комплементарные пары: для осуществления указанных реакций могут быть использованы дуплексы, содержащие в некоторых положениях некомплементарные пары оснований [92, 93]. В частности, оказалось, что нарушение комплементарности 2'-дезоксиаденина в 3-м положении ДНК-субстрата даже стимулирует процессинг [89, 92, 94, 95]. ИН ВИЧ-1 также лучше, чем полностью комплементарный субстрат, процессирует субстраты, в которых две концевые пары нуклеотидов некомплементарны [55, 57, 89]. Последний результат послужил основанием для гипотезы, согласно которой необходимым условием 3'-процессинга является предварительное разрушение комплементарных пар оснований на конце ДНК-субстрата [94]. Однако позднее было показано, что полное «расплетание» двойной спирали на конце ДНК не является обязательным, но вместе с тем ИН необходима локальная дестабилизация ДНК в районе расщепляемой связи [95].

Как известно, *in vivo* ИН ВИЧ-1 расщепляет межнуклеотидную связь, расположенную за 2 нуклеозида от 3'-конца вирусной ДНК. Однако такое расположение сайта расщепления не является обязательным для реакции процессинга *in vitro*. ИН ВИЧ-1 способна процессировать субстрат, содержащий вплоть до 5 пар оснований после консервативного динуклеотида СА [56, 57, 89]. Тем не менее, если концевая последовательность ДНК ВИЧ-1 располагается во внутренней части дуплекса, моделирующего субстрат ИН, эффективность процессинга такого субстрата существенно снижается [57, 90, 94, 96].

С использованием различным образом модифицированных аналогов U5-субстрата ИН установлена важность консервативного динуклеотида СА для всех реакций, катализируемых ИН ВИЧ-1 *in vitro* [56, 57, 80, 89, 92, 93, 97]. Положения, прилегающие к консервативному динуклеотиду СА, также играют немаловажную роль для узнавания субстрата ИН. Например, показана важность нуклеотидов в 5-м и 6-м положениях процессируемой цепи субстратов U5 или U3 [98].

Детальное исследование взаимодействия ИН с ДНК-субстратом было проведено в работе [87]. Были синтезированы аналоги U5-субстрата, в процессируемой цепи которых в положениях с первого по шестнадцатое были произведены все возможные замены нуклеотидов (например, А заменяли на G, C и T). Вторая цепь во всех дуплексах

являлась полностью комплементарной процессируемой цепи. Важно отметить, что эксперименты проводились с использованием как ионов Mn^{2+} , так и Mg^{2+} в качестве кофактора. На основании изучения начальных скоростей 3'-процессинга были выделены два важных района взаимодействия ДНК с белком: со второго по шестой нуклеотиды и с десятого по четырнадцатый нуклеотиды в присутствии Mg^{2+} . В случае ионов Mn^{2+} дальняя область узнавания отсутствовала [87].

Эти данные позволили сформулировать важную гипотезу, касающуюся строения активного центра ИН. Согласно ей, активный центр фермента в присутствии ионов Mg^{2+} более организован, чем в присутствии ионов марганца, в связи с чем его субстратная специфичность выше. В присутствии ионов Mn^{2+} структура активного центра становится «более расшатанной», поэтому ИН в этих условиях менее «чувствительна» к структуре субстрата, а также допускает взаимодействие с субстратом нуклеофилов, отличных от воды, например, метанола и глицерина [59]. Есть и другое подтверждение: только в присутствии ионов Mn^{2+} идет дезинтеграция, поскольку здесь не нужно столь точного узнавания субстрата, в частности консервативного динуклеотида СА [98].

Для выявления конкретных атомов в структуре азотистых оснований, с которыми образуют контакты аминокислотные остатки фермента, были использованы олигонуклеотиды, содержащие модифицированные гетероциклические основания [99]. Нуклеозиды, содержащие 2,6-диаминопурин, 7-дезааденин, инозин, N⁶-метиладенин, O⁶-метилгуанин, 5-метилцитозин и O⁴-метилтимин, вводились по очереди в обе цепи U5-дуплекса в положения с 3 по 5, как наиболее важные в 3'-процессинге. Нуклеозидные аналоги выбирались с таким расчетом, чтобы модифицированные группы располагались в большой или малой бороздках дуплекса. Авторы полагали, что такие модификации слабо искажали структуру ДНК, но при этом нарушали возможные водородные связи между гетероциклическими основаниями и белком. Однако необходимо отметить, что некоторые из использованных в работе [99] модификаций нуклеозидов (O⁶-метилгуанин, O⁴-метилтимин) приводят к нарушению водородных связей в комплементарных парах и, следовательно, к изменению структуры двойной спирали. К сожалению, авторы цитируемой работы не учитывали возможные последствия этих эффектов. Изменение скорости 3'-процессинга при введении модифицирующих группировок в структуру гетероциклических оснований они связывали только с изменением тех или иных контактов с ИН. Как правило, скорость 3'-процессинга модифицированных субстратов была ниже, чем для U5-субстрата. Единственной модификацией, увеличивающей скорость 3'-процес-

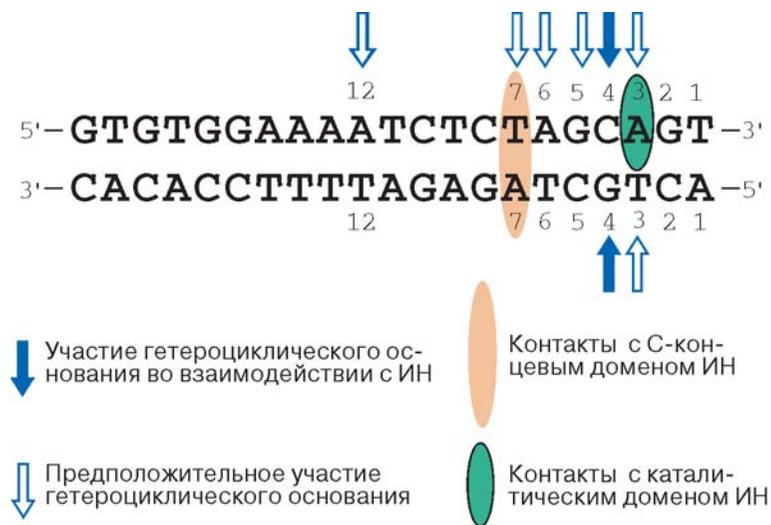


Рис.15. Схематичное представление описанных в литературе контактов ИН ВИЧ-1 с различными нуклеозидами U5-субстрата.

синга, была замена 3-го тимидина непротессуемой цепи на O⁴-метилтимидин. На основании полученных данных, авторы предположили, что основные контакты ИН осуществляются со стороны большой бороздки ДНК.

Таким образом, на основании данных химической модификации ДНК и зондирования контактов ДНК с ИН, можно представить себе следующую картину взаимодействия ИН с ДНК-субстратом (рис. 15). Однозначно показано, что для эффективного 3'-протессинга субстрат ИН должен содержать пару G/C в 4-м положении [87, 99]. Кроме того, данные многих работ указывают на важность нуклеотидов в положениях 3, 5–7 протессуемой цепи. Причем для 2'-дезоксиаденозина в положении 3 предполагается контакт с каталитическим доменом ИН [86], а для тимидина в положении 7 – контакт с С-концевым доменом [87]. Также предполагаются контакты ИН с положением 12 в протессуемой цепи субстрата [87]. В работе [99] обсуждается возможность гидрофобного контакта ИН с тимидином в 3-м положении непротессуемой цепи. Других данных о взаимодействии ИН с непротессуемой цепью субстрата нет.

ВЗАИМОДЕЙСТВИЕ ИНТЕГРАЗЫ С КЛЕТОЧНОЙ ДНК

Встраивание вирусной ДНК, по существу, представляет собой процесс, независимый от последовательности клеточной ДНК, хотя описаны некоторые предпочтения в выборе последовательностей *in*

vivo [100, 101]. На процесс узнавания ДНК-мишени могут влиять некоторые структурные особенности клеточной ДНК, а также строение прединтеграционного комплекса и клеточные факторы. Реакция интеграции *in vivo* существенно более сложный процесс, чем моделирование интеграции *in vitro* с использованием очищенной ИН и ДНК-субстратов. Например, ретровирусный прединтеграционный комплекс проявляет более высокую ферментативную активность, чем свободная ИН [102], а хроматин является лучшей мишенью для интеграции, чем свободная ДНК. Считается, что очищенная ИН использует для встраивания первый же ДНК-дуплекс, с которым она связывается, в то время как прединтеграционный комплекс, выделенный из инфицированных клеток, способен подыскивать определенную последовательность в структуре клеточной ДНК благодаря повторяющемуся связыванию, которое обеспечивается клеточными факторами [103].

Не все фосфатные группы в ДНК одинаково узнаются ИН *in vitro* [104, 105]. Предполагается, что структура ДНК-мишени [105], ее последовательность [105–107] или расположение по отношению к концу ДНК-мишени [57, 91, 107, 108] являются основными параметрами выбора места встраивания. Предпочтительный сайт встраивания в случае ИН ВИЧ-1 расположен в том месте, где наблюдается какой-либо изгиб ДНК, образованный, например, при закручивании вокруг нуклеосом, петлеобразованием или другими особенностями структуры ДНК [109, 110]. Вероятно, наблюдающаяся в этих случаях стимуляция переноса цепи происходит за счет улучшения связывания ИН с ДНК, при этом частота сайтов встраивания *in vitro* прямо коррелирует с количеством нарушений ДНК в большой бороздке [110, 111].

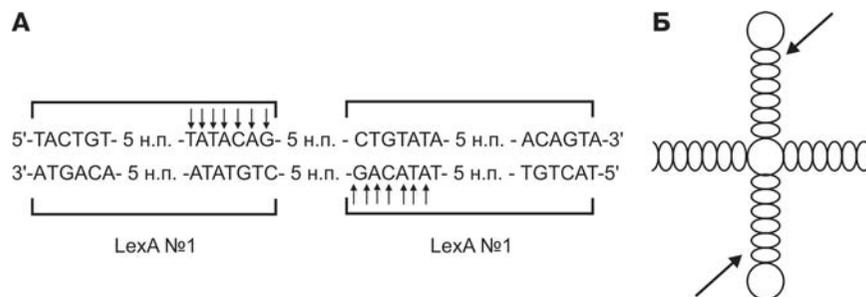


Рис. 16. Влияние структуры ДНК-мишени на направление интеграции в нее вирусной ДНК.

А – первичная структура ДНК-мишени, способной образовывать крестообразные структуры.

Б – предполагаемые ИН области для встраивания вирусной ДНК по данным работы [113].

Стимуляция переноса цепи в местах излома ДНК-мишени может быть соотнесена с механизмом интеграции, подразумевающим наличие одного активного центра. Излом ДНК может способствовать локальному нарушению комплементарных взаимодействий. Как уже упоминалось, в ходе процессинга ДНК-субстратов ИН ВИЧ-1 происходит нарушение комплементарных взаимодействий концевых пар нуклеотидов [94]. В связи с тем, что ИН имеет один активный центр для процессинга и переноса цепи, вероятно, подобное нарушение или расплетание ДНК-мишени может быть необходимо и для переноса цепи [96]. *In vivo* для более эффективного осуществления переноса цепи ИН может использовать накручивание ДНК на нуклеосому или искажения ДНК-мишени, такие как нуклеотидные замены или расплетенные участки.

Вероятно, перенос цепи не происходит случайным образом [112]. Некоторые закономерности выбора сайта интеграции были рассмотрены в работе [113]. В качестве ДНК-мишени использовалась плаزمиды, содержащая два последовательно расположенных LexA-оператора (рис. 16А), инвертированных друг относительно друга. В случае линеаризованной плазмиды интеграция вирусного генома происходила статистически по всей молекуле плазмидной ДНК, однако, если плазмиды оставалась в характерной для нее кольцевой форме, встраивание происходило строго в места, указанные на рисунке стрелками. Объяснение этому факту выглядит так: в сверхспирализованном состоянии, характерном для ковалентно-замкнутых плазмид, подобные нуклеотидные последовательности способны образовывать крестообразные структуры со шпильками на концах. Встраивание вирусной ДНК происходит в ближайшую к одноцепочечным участкам двухспиральную область (рис. 16Б). Замена нуклеотидных последовательностей двуспиральных участков около петель на другие, но также комплементарные, не влияла на эффективность переноса цепи.

Таким образом, для реакции переноса цепи очень существенной оказалась вторичная структура ДНК-мишени. Образующиеся за счет сверхспирализации ДНК петли являются сильным направляющим фактором реакции [113]. Это может быть связано с тем, что нарушение комплементарных взаимодействий нуклеотидов, вовлеченных в образование гидролизуемой фосфодиэфирной связи, является одной из необходимых стадий механизма реакции переноса цепи, аналогично тому, как это показано для транспозазы Tn10 [114]. В шпильчатой структуре это нарушение уже есть, что, разумеется, помогает протеканию реакции. Известно, что наиболее часто предпочитаемые места интеграции – это транскрипционно активные зоны [115], в которых часто встречаются шпильчатые структуры.

IV. ЗАКЛЮЧЕНИЕ

Поводя итог всему изложенному выше, можно сказать, что, несмотря на большое число исследований по структуре и механизму действия ИН ВИЧ-1, этот фермент по-прежнему остается очень «загадочным». До настоящего времени не удается понять, как точно происходит связывание вирусной и клеточной ДНК, какие структурные характеристики вирусной ДНК определяют сиквенс-специфичность реакции 3'-концевого процессинга, сколько субъединиц ИН входит в состав активного комплекса. Ясно только, что специфичность взаимодействия ИН со своим субстратом, в первую очередь, определяется наличием в его структуре консервативного динуклеотида СА. Кроме этого понятно, что для гидролиза фосфодиэфирных связей на обоих этапах интеграции необходимо нарушение комплементарных взаимодействий нуклеотидов, расположенных рядом с гидролизуемой связью. Отсутствие точной информации о структуре ИН ВИЧ-1 и ее взаимодействии с ДНК во многом тормозит направленный поиск ее ингибиторов, которые могли бы стать перспективными препаратами для лечения СПИД'а. Кроме того, надо отметить, что все ретровирусные интегразы являются достаточно консервативными по структуре и близкими по механизму действия ферментами, следовательно, получение новых знаний, касающихся ИН ВИЧ-1, позволит лучше понять весь механизм репликации ретровирусов, а также и механизм транспозиции генов.

ЛИТЕРАТУРА

1. Barre-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Dauguet, C., Axler-Blin C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., Montagnier L. (1983) *Science*, **220**, 868–871.
2. Turner, B.G., Summers, M.F. (1999) *J. Mol. Biology*, **285**, 1–32.
3. Ramratnam, B., Mittler, J.E., Zhang, L., Boden, D., Hurley, A., Fang, F., Macken, C.A., Perelson, A.S., Markowitz, M., Ho, D.D. (2000) *Nat. Med.*, **6** (1), 82–85.
4. Finzi, D., Blankson, J., Siliciano, J.D., Margolick, J.B., Chadwick, K., Pierson, T., Smith, K., Lisziewicz, J., Lori, F., Flexner, C., Quinn, T.C., Chaisson, R.E., Rosenberg, E., Walker, B., Gange, S., Gallant, J., Siliciano, R.F. (1999) *Nat. Med.*, **5**, 512–517.
5. Geist, A., BouHamdan, M., Nunnari, G., Pomerantz, R.J., Kulkosky, J. (1999) *Gene Ther. Mol. Biol.*, **4**, 133–141.
6. Dyda, F., Hickman, B., Jenkins, T.M., Engelman, A., Craigie, R., Davies, D.R. (1994) *Science*, **266** (5193), 1981–1986.
7. Ariyoshi, M., Vassilyev, D. G., Iwasaki, H., Nakamura, H., Shinagawa, H., Morikawa, K. (1994) *Cell*, **78** (6), 1063–1072.
8. Rice, P.A., Baker, T.A. (2001) *Nat. Struct. Biol.*, **8** (5), 302–307.
9. Kulkosky, J., Skalka, A.M. (1994) *Pharmacol. Ther.*, **61**, 185–203.
10. Engelman, A., Bushman, F.D., Craigie, R. (1993) *EMBO J.*, **12** (8), 3269–3275.
11. van Gent, D.C., Vink, C., Groeneger, A.A., Plasterk, R. (1993) *EMBO J.*, **12**, 3261–3267.

12. Vink, C., Groeneger, A.A., Plasterk, R. (1993) *Nucleic Acids Res.*, **21** (6), 1419–1425.
13. Bushman, F.D., Engelman, A., Palmer, I., Wingfield, P., Craigie, R. (1993) *Proc. Natl. Acad. Sci. USA*, **90** (8), 3428–3432.
14. Lee, S.P., Xiao, J., Knutson, J.R., Lewis, M.S., Han, M.K. (1997) *Biochemistry*, **36** (1), 173–180.
15. Engelman, A., Craigie, R. (1992) *J. Virol.*, **66** (11), 6361–6369.
16. Polard, P., Chandler, M. (1995) *Mol. Microbiol.*, **15** (1), 13–23.
17. Leavitt, A.D., Shiue, L., Varmus, H.E. (1993) *J. Biol. Chem.*, **268** (3), 2113–2119.
18. Yang, W., Steitz, T.A. (1995) *Structure*, **3**, 131–134.
19. Goldgur, Y., Dya, F., Hickman, A.B., Jenkins, T.M., Craigie, R., Davies, D.R. (1998) *Proc. Natl. Acad. Sci. USA*, **95** (16), 9150–9154.
20. Lovell, S., Goryshin, I. Y., Reznikoff, W.R., Rayment, I. (2002) *Nature Struct. Biol.*, **9** (4), 278–281.
21. Mizuuchi, K. (1997) *Genes Cells*, **2**, 1–12.
22. Andrade, M.D., Skalka, A.M. (1996) *J. Biol. Chem.*, **271**, 19633–19636.
23. Rice, P., Craigie, R., Davies, D.R. (1996) *Curr. Opin. Struct. Biol.*, **6**, 76–83.
24. Mizuuchi, K. (1992) *J. Biol. Chem.*, **267**, 21273–21276.
25. Mizuuchi, K., (1992) *Annu. Rev. Biochem.*, **61**, 1011–1051.
26. Vincent, K.A., Ellison, V., Chow, S.A., Brown P.O. (1993) *J. Virol.*, **67** (1), 425–437.
27. Zheng, R., Jenkins, T.M., Craigie, R. (1996) *Proc. Natl. Acad. Sci. USA*, **93** (24), 13659–13664.
28. Cai, M., Zheng, R., Caffrey, M., Graigie, R., Clore, G.M., M. Grannenborn, A.M. (1997) *Nature Struct. Biol.*, **4** (7), 567–577.
29. Woerner, A.M., Marcus-Sekura, C.J. (1993) *Nuc. Acids Res.*, **21** (15), 3507–3511.
30. Puras Lutzke, R.A., Vink C., Plasterk, R. (1994) *Nuc. Acids Res.*, **22** (20), 4125–4131.
31. Lodi, P.J., Ernst, J.A., Kuszewski, J., Hickman, A.B., Engelman, A., Craigie, R. Clore, G.M., Gronenborn, A.M. (1995) *Biochemistry*, **34** (31), 9826–9833.
32. Puras Lutzke, R.A., Plasterk, R. (1998) *J. Virol.*, **72** (6), 4841–4848.
33. Chen, J.C.-H., Krucinski, J., Miercke, L.J.W., Finer-Moore, J.S., Tang, A.H., Leavitt, A.D., Stroud, R.M. (2000) *Biochemistry*, **39** (15), 8233–8238.
34. Wang, J.-Y., Ling, H., Yang, W., Graigie, R. (2001) *EMBO J.*, **20** (24), 7333–7343.
35. Esposito, D., Craigie, R. (1999) *Adv. Virus Res.*, **52**, 319–333.
36. Cherepanov, P., Maertens, G., Proost, P., Devreese, B., Van Beeumen, J., Engelborgh, Y., De Clercq, E., Debyser, Z. (2003) *J. Biol. Chem.*, **278** (1), 372–381.
37. Deprez E., Tauc P., Leh H., Mouscadet J.-F., Auclair C., Brochon J.-C. (2000) *Biochemistry*, **39** (31), 9275–9284.
38. Leh, H., Brodin, P., Bischerour, J., Deprez, E., Tauc, P., Brochon, J.-C., LeCam, E., Couland, D., Auclair, C., Mouscadet, J.-F. (2000) *Biochemistry*, **39** (31), 9285–9294.
39. Gao, K., Butler, S.L., Bushman, F. (2001) *EMBO J.*, **20**, 3565–3576.
40. Podtelezhnikov, A.A., Gao, K., Bushman, F.D., McCammon, J.A. (2003) *Biopolymers*, **68** (1), 110–120.
41. De Luca, L., Pedretti, A., Vistoli, G., Barreca, M.L., Villa, L., Monforte, P., Chimirri, A. (2003) *Biochem. and Biophys. Res. Commun.*, **310**, 1083–1088.
42. Adesokan, A.A., Roberts, V.A., Lee, K.W., Lins, R.D., Briggs, J.M. (2004) *J. Med. Chem.*, **47** (4), 821–828.
43. Brown, P.O., Bowerman, B., Varmus, H.E., Bishop, J.M. (1987) *Cell*, **49** (3), 347–356.
44. Bukrinsky, M.I., Sharova, N., Dempsey, M.P., Stanwick, T.L., Bukrinskaya, A.G., Haggerty, S., Stevenson, M. (1992) *Proc. Natl. Acad. Sci. USA*, **89** (14), 6580–6584.

45. Miller, M.D., Farnet, C.M., Bushman, F.D. (1997) *J. Virol.*, **71** (7), 5382–5390.
46. Farnet, C.M., Bushman, F.D. (1997) *Cell*, **88** (4), 483–492.
47. Bushman, F.D. (1999) *Adv. Virus Res.*, **52**, 301–317.
48. Brown, P.O. (1990) *Curr. Top. Microbiol. Immunol.*, **157**, 19–48.
49. Carteau, S., Hoffmann, C., Bushman, F. (1998) *J. Virol.*, **72** (5), 4005–4014.
50. Appa, R.S., Shin, C.G., Lee, P., Chow, S.A. (2001) *J. Biol. Chem.*, **276** (49), 45848–45855.
51. Goulaouic, H., Chow, S.A. (1996) *J. Virology*, **70** (1), 37–46.
52. Bushman, F.D., Miller M.D. (1997) *J. Virology*, **71** (1), 458–464.
53. Coffin, J.M., Hughes, S.H., Varmus, H.E. (1998) Cold Spring Harbor: Cold Spring Harbor Lab. Press. 161–203.
54. Brin, E., Yi, J., Skalka, A., Leis, J. (2000) *J. Biol. Chem.*, **275** (50), 39287–39295.
55. Bushman, F.D., Craigie, R. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 3458–3462.
56. Bushman, F.D., Craigie, R. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1339–1343.
57. Vink, C., van Gent, D.C., Elgersma, Y., Plasterk, R.H. (1991) *J. Virol.*, **65**, 4636–4644.
58. Katzman, M., Mack, J.P.G., Skalka, A.M., Leis, J. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 4695–4699.
59. Engelman, A., Mizuuchi, K., Craigie, R. (1991) *Cell*, **67**, 1211–1221.
60. Katzman, M., Sudol, M. (1996) *J. Virol.*, **70**, 2598–2604.
61. Chow, S.A., Vincent, K.A., Ellison, V., Brown, P.O. (1992) *Science*, **255**, 723–726.
62. Tramontano, E., Colla, P.L., Cheng, Y.C. (1998) *Biochemistry*, **37**, 7237–7243.
63. Gerton, J.L., Herschlag, D., Brown, P.O. (1999) *J. Biol. Chem.*, **274**, 33480–33487.
64. Mumm, S.R., Grandgenett, D.P. (1991) *J. Virol.*, **65**, 1160–1167.
65. Khan, E., Mack, J.P.G., Katz, R.A., Kulkosky, J., Skalka, A.M. (1991) *Nucleic Acids Res.*, **19**, 851–860.
66. Schauer, M., Billich, A. (1992) *Biochem. Biophys. Res. Commun.*, **185**, 874–880.
67. van Gent, D. C., Elgersma, Y., Bolk, M.W.J., Vink, C., Plasterk, R.H.A. (1991) *Nucleic Acids Res.*, **19**, 3821–3827.
68. Yoshinaga, T., Kimura-Ohtani, Y., Fujiwara, T. (1994) *J. Virol.*, **68**, 5690–5697.
69. Hazuda, D.J., Wolfe, A.L., Hastings, J. C., Robbins, H.L., Graham, P.L., LaFemina, R.L., Emini, E.A. (1994) *J. Biol. Chem.*, **269**, 3999–4004.
70. Vink, C., Lutzke, R.A., Plasterk, R.H. (1994) *Nucleic Acids Res.*, **22**, 4103–4110.
71. Woerner, A.M., Klutch, M., Levin, J.G., Marcus-Sekura, C.J. (1992) *AIDS Res. Hum. Retroviruses*, **8**, 297–304.
72. Wolfe, A.L., Felock, P.J., Hastings, J. C., Blau, C.U., Hazuda, D.J. (1996) *J. Virol.*, **70**, 1424–1432.
73. Vora, A.C., Grandgenett, D.P. (1995) *J. Virol.*, **69**, 7483–7488.
74. Mazumder, A., Neamati, N., Pilon, A.A., Sunder, S., Pommier, Y. (1996) *J. Biol. Chem.*, **271**, 27330–27338.
75. Ellison, V., Brown, P.O. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 7316–7320.
76. Hazuda, D.J., Felock, P.J., Hastings, J.C., Pramanik, B., Wolfe, A.L. (1997) *J. Virol.*, **71**, 7005–7011.
77. Lee, S.P., Censullo, M.L., Kim, H.G., Han, M.K. (1995) *Biochemistry*, **34**, 10215–10223.
78. Katzman, M., Sudol, M. (1998) *J. Virol.*, **72** (3), 1774–1753.
79. Heuer, T.S., Brown, P.O. (1997) *Biochemistry*, **36** (35), 10655–10665.
80. LaFemina, R.L., Callahan, P.L., Cordingley, M.G. (1991) *J. Virol.*, **65**, 5624–5630.
81. Engelman, A., Hickman, A.B., Craigie, R. (1994) *J. Virol.*, **68**, 5911–5917.
82. Bugreev, D.V., Baranova, S., Zakharova, O.D., Parissi, V., Desjobert, C.,

- Sottofattori, E., Balbi, A., Litvak, S., Tarago-Litvak, L., Nevinsky, G.A.* (2003) *Biochemistry*, **42** (30), 9235–9247.
83. *Gerton, J.L., Brown, P.O.* (1997) *J. Biol. Chem.*, **272**, 25809–25815.
84. *Shibagaki, Y., Holmes, M.L., Appa, R.S., Chow, S.A.* (1997) *Virology*, **230**, 1–10.
85. *Katzman, M., Sudol, M.* (1995) *J. Virol.*, **69**, 5687–5696.
86. *Jenkins, T.M., Esposito, D., Engelman, A., Craigie, R.* (1997) *EMBO J.*, **16**, 6849–6859.
87. *Esposito, D., Craigie, R.* (1998) *EMBO J.*, **17**, 5832–5843.
88. *Withka, J.M., Wilde, J.A., Bolton, P.H., Mazumder, A., Gerlt, J.A.* (1991) *Biochemistry*, **30**, 9931–9940.
89. *Sherman, P.A., Dickson, M.L., Fyfe, J.A.* (1992) *J. Virol.*, **66**, 3593–3601.
90. *Katzman, M., Katz, R.A., Skalka, A.M., Leis, J.* (1989) *J. Virol.*, **63**, 5319–5327.
91. *Katzman, M., Sudol, M.* (1994) *J. Virol.*, **68**, 3558–3569.
92. *van den Ent, F.M.I., Vink, C., Plasterk, R.H.A.* (1994) *J. Virol.*, **68**, 7825–7832.
93. *Mazumder, A., Pommier, Y.* (1995) *Nucl. Acids. Res.*, **23**, 2865–2871.
94. *Scottoline, B.P., Chow, S., Ellison, V., Brown, P.O.* (1997) *Genes & Development*, **11**, 371–382.
95. *Agapkina, J., Smolov, M., Zubin, E., Mouscadet, J.-F., Gottikh, M.* (2004) *Eur. J. Biochem.*, **271**, 205–211.
96. *Leavitt, A.D., Rose, R.B., Varmus, H.E.* (1992) *J. Virol.*, **34**, 2359–2368.
97. *Sherman, P., Fyfe, J.A.* (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 5119–5123.
98. *Katzman, M., Sudol, M.* (1996) *J. Virol.*, **70**, 9069–9073.
99. *Wang, T., Balakrishnan, M., Jons-son, C.B.* (1999) *Biochemistry*, **38**, 3624–3632.
100. *Shih, C.-C., Stoye, J.S., Coffin, J.M.* (1988) *Cell*, **53**, 531–537.
101. *Withers-Ward, E.S., Kitamura, Y., Barnes, J.P., Coffins, J.M.* (1994) *Genes Dev.*, **8**, 1473–1487.
102. *Farnet, C.M., Wang, B., Lipford, J.R., Bushman, F.D.* (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 9742–9747.
103. *Miller, M.D., Bor, Y.-C., Bushman, F.* (1995) *Curr. Biol.*, **5**, 1047–1055.
104. *Kitamura, Y., Lee, Y.M.H., Coffin, J.M.* (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 5532–5536.
105. *Pryciak, P.M., Varmus, H.E.* (1992) *Cell*, **69**, 769–780.
106. *Bor, Y.-C., Miller, M.D., Bushman, F.D., Orgel, L.E.* (1996) *Virology*, **222**, 283–288.
107. *Fitzgerald, M.L., Grandgenett, D.P.* (1994) *J. Virol.*, **68**, 4314–4321.
108. *Hong, T., Murphy, E., Groarke, J., Drlica, K.* (1993) *J. Virol.*, **67**, 1127–1131.
109. *Muller, H.-P., Varmus, H.E.* (1994) *EMBO J.*, **13**, 4704–4714.
110. *Pruss, D., Reeves, R., Bushman, F.D., Wolffe, A. P.* (1994) *J. Biol. Chem.*, **269**, 25031–25041.
111. *Bushman, F.D.* (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 9233–9237.
112. *Craigie, R.* (1992) *Trends Genet.*, **8**, 187–190.
113. *Katz, R.A., Gravuer, K., Skalka, A.M.* (1998) *J. Biol. Chem.*, **273**, 24190–24195.
114. *Pribil, P.A., Haniford, D.B.* (2003) *J. Mol. Biol.*, **330**(2), 247–259.
115. *Kalpana, G.V., Marmon, S., Wang, W., Crabtree, G.R., Goff, S.P.* (1994) *Science*, **266**, 2002–2006.