

УДК 577.151.34; 577.151.64

# Изучение функциональных и аллостерических сайтов в суперсемействах белков

Д. А. Суплатов<sup>1,2</sup>, В. К. Швядас<sup>1,2\*</sup>

<sup>1</sup>Московский государственный университет им. М.В. Ломоносова, Научно-исследовательский институт физико-химической биологии им. А.Н. Белозерского, 119991, Москва, Ленинские горы, 1, стр. 40

<sup>2</sup>Московский государственный университет им. М.В. Ломоносова, факультет биоинженерии и биоинформатики, 119991, Москва, Ленинские горы, 1, стр. 73

\*E-mail: vyfas@belozersky.msu.ru

Поступила в редакцию 19.08.2015

**РЕФЕРАТ** К важнейшим факторам регуляции функциональных свойств белков (ферментов) относится их взаимодействие с различными низкомолекулярными соединениями, а также белок-белковые взаимодействия. К настоящему времени наиболее хорошо изучены молекулярные механизмы действия лигандов, которые связываются в функциональных сайтах белков (активных центрах ферментов), в то время как механизмы аллостерической регуляции или связывания в других центрах изучены недостаточно. Исследования последнего времени показывают, что аллостерия может быть свойством практически всех белков, и для систематического анализа организации и роли различных сайтов, установления взаимосвязи между их структурой, функцией и регуляцией необходимы новые подходы. Современные методы компьютерной биологии, биоинформатики и молекулярного моделирования позволяют вести поиск новых центров связывания регуляторных лигандов, изучать особенности их структурной организации, сходство различных сайтов в гомологичных белках, молекулярные механизмы аллостерии, а также взаимосвязи функции и регуляции. Установление эволюционных взаимосвязей между различными центрами связывания в суперсемействах белков, открытие новых функциональных, аллостерических и регуляторных сайтов с использованием вычислительных подходов должны улучшить наше понимание структурно-функциональных взаимосвязей в белках, предоставить новые возможности для создания лекарственных средств и дизайна более эффективных биокатализаторов.

**КЛЮЧЕВЫЕ СЛОВА** центры связывания, каталитический сайт, аллостерический сайт, функция, регуляция, структурно-функциональные взаимосвязи, биоинформатика.

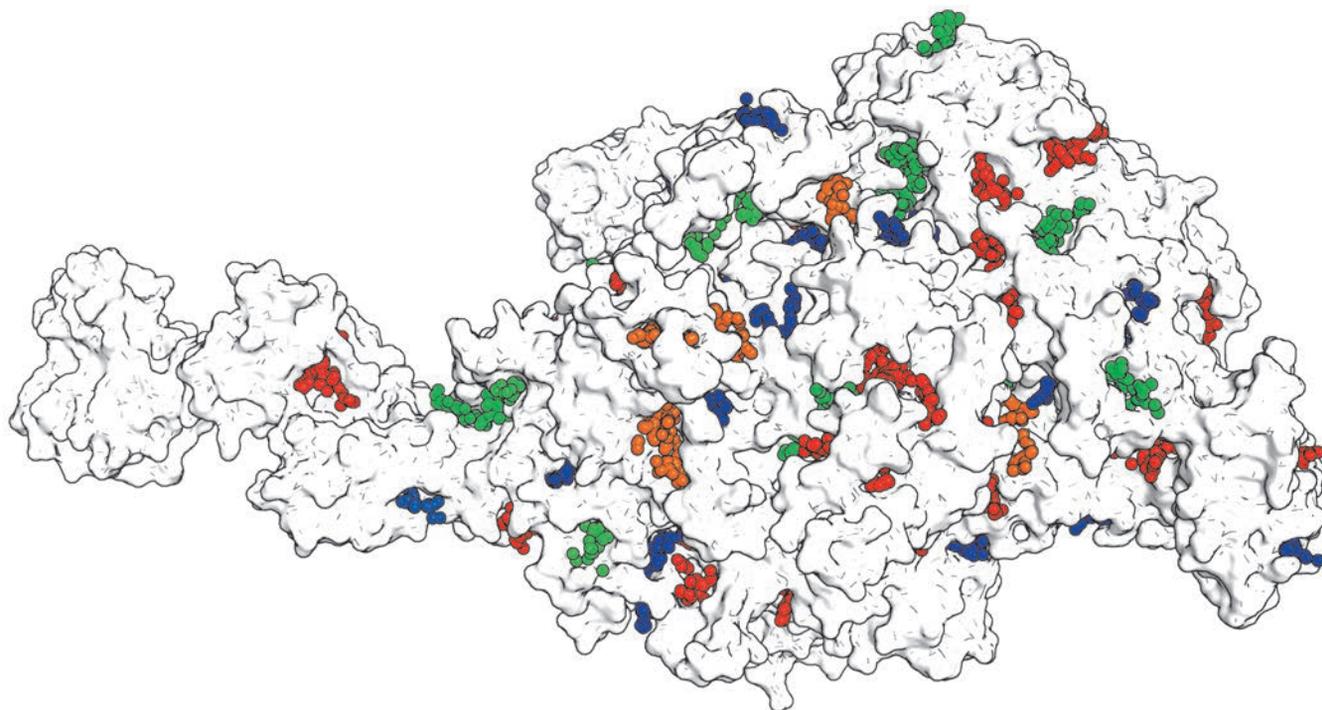
**СОКРАЩЕНИЯ** PDB – Protein Data Bank; РНКП – ДНК-зависимая РНК-полимераза; МД – молекулярная динамика; СПП – специфическая позиция подсемейства.

## ВВЕДЕНИЕ

Изучение взаимосвязи структуры и функции белка – одна из актуальных задач современной биохимии, сложность которой заключается в том, что сходство структуры не означает тождественности функций. При общей укладке полипептидной цепи белки могут обладать принципиально разными свойствами [1, 2], в то время как одну и ту же функцию могут выполнять белки с различной структурой [3]. Специфические белок-белковые взаимодействия и распознавание низкомолекулярных соединений лежат в основе функционирования живых систем. Для понимания молекулярных механизмов этих процессов и структурно-функциональных взаимосвязей в белках необходимо изучать структурную орга-

низацию конкретных сайтов, ответственных за эти взаимодействия и связывание различных лигандов (субстратов, ингибиторов, эффекторов) [4]. Анализ карманов и полостей на поверхности глобул, формирующих сайты связывания с уникальными свойствами, и их функциональная классификация должны привести к более глубокому пониманию общих закономерностей функционирования белков, предсказания функции новых ферментов, целенаправленного изменения свойств белков/ферментов дикого типа, а также дизайна лекарственных средств.

В большинстве работ изучение свойств белков фокусируется на анализе их функциональных сайтов – активных центров ферментов, каналов мембранных транспортных белков, ДНК- и белоксвязывающих



**Рис. 1.** Потенциальные сайты связывания низкомолекулярных эффекторов в структуре бактериальной РНКП. Скопление шариков одного цвета указывает на местонахождение потенциального сайта на поверхности фермента и соответствует центрам  $\alpha$ -сфер, которые заполняют пространство соответствующей полости связывания (см. «Приложение»). Рисунок подготовлен с помощью программы PyMol на основании кристаллографической структуры 1YNN из PDB

участков различных регуляторных белков. Однако за последние годы накоплено множество свидетельств явления аллостерии – регуляции функции белков посредством связывания низкомолекулярных эффекторов в регуляторных сайтах, топологически независимых от функциональных сайтов [5]. Эти сведения стимулируют интерес к изучению регуляции функции биологических макромолекул при связывании различных лигандов в аллостерических центрах. Созданы экспериментальные и компьютерные подходы к поиску новых регуляторных центров в структурах белков. Предпринимаются попытки понять взаимосвязь между функциональными и регуляторными сайтами и изучить молекулярные механизмы взаимного влияния центров связывания, находящихся на значительном расстоянии друг от друга [6]. Обнаружены низкомолекулярные ингибиторы, селективно связывающиеся в аллостерических сайтах ряда белков, ассоциированных с различными заболеваниями человека [7]. Однако особое внимание к этой теме связано не столько с уникальными особенностями функционирования конкретных белков, сколько с общей ролью этих процессов в регуляции живых систем. Есть основания полагать, что алло-

стерия представляет собой весьма универсальное явление, присущее большинству белков [8]. Интерес вызывает не только фундаментальное значение аллостерии, но и возможность практического использования этого явления в биотехнологии и биомедицине. За последнее время показано, что белки и ферменты наряду с весьма хорошо известными функциональными сайтами (активными центрами) и аллостерическими сайтами содержат значительное количество практически неизученных участков связывания. На *рис. 1* представлена структура ДНК-зависимой РНК-полимеразы – ключевого фермента синтеза РНК во всех живых организмах [9, 10]. Поверхность этого крупного полисубъединичного фермента покрыта большим количеством полостей – потенциальных сайтов связывания. Среди них можно увидеть как активный центр, содержащий каталитические остатки и участок связывания ДНК, так и несколько известных аллостерических сайтов, способных связывать различные низкомолекулярные лиганды [11, 12]. При этом роль остальных участков связывания, а их большинство, остается неизвестной. Насколько важны эти сайты для функционирования фермента? Какие сайты связывания играют физиологическую

роль, а какие можно приспособить для создания белка с новыми практически значимыми свойствами? Как оценить потенциал использования каждого отдельного сайта для регуляции функции?

В данном обзоре обсуждается изучение структурно-функциональных взаимосвязей в белках на основании анализа различных сайтов связывания в их структурах. Рассмотрены экспериментальные и компьютерные подходы к поиску и изучению новых центров связывания регуляторных лигандов, молекулярные механизмы аллостерии, а также взаимосвязь функции и регуляции в белках. Увеличение размеров публичных баз данных открывает доступ к новой информации о геномных последовательностях, структуре и функции большого числа белков. В этой связи новые возможности предоставляют методы биоинформатики, которые позволяют изучать свойства ферментов не по отдельности, а системно, в суперсемействах. Анализ структурной информации и экспериментальных данных о функциональных свойствах отдельных белков в их взаимосвязи с близкими, а также эволюционно удаленными гомологами, должен позволить лучше понять структурно-функциональные взаимосвязи в белках/ферментах и обнаружить новые механизмы регуляции их функциональных свойств.

### ЯВЛЕНИЕ АЛЛОСТЕРИИ

Аллостерию принято определять как процесс регуляции функции белка путем связывания эффектора – лиганда или другого белка – в сайте на поверхности структуры, который называют аллостерическим центром [6]. В свою очередь, слово «аллостерический» происходит от греческих корней *allos* (другой) и *stereos* (твердый) и может быть переведено как «иная форма», чтобы подчеркнуть взаимосвязь конформационных состояний между структурно удаленными сайтами в белках [8]. Известно, что аллостерическая регуляция метаболизма важна для функционирования живых клеток, а аллостерические эффекторы могут быть как ингибиторами, так и активаторами функции белков [13].

Исторически под аллостерией понимали кооперативный эффект в полисубъединичных белках, функционирующих на уровне четвертичной структуры. В 1965 году на основании 24 известных на тот момент случаев была сформулирована «согласованная» модель аллостерии, называемая также моделью MWC (по заглавным буквам фамилий авторов Monod–Wyman–Changeux) [14]. Скачкообразное увеличение сродства гемоглобина к кислороду по мере его насыщения, описываемое S-образной кривой, позволило предположить кооперативный эффект. Однако полученная в 1960 году с разрешением 5.5 Å

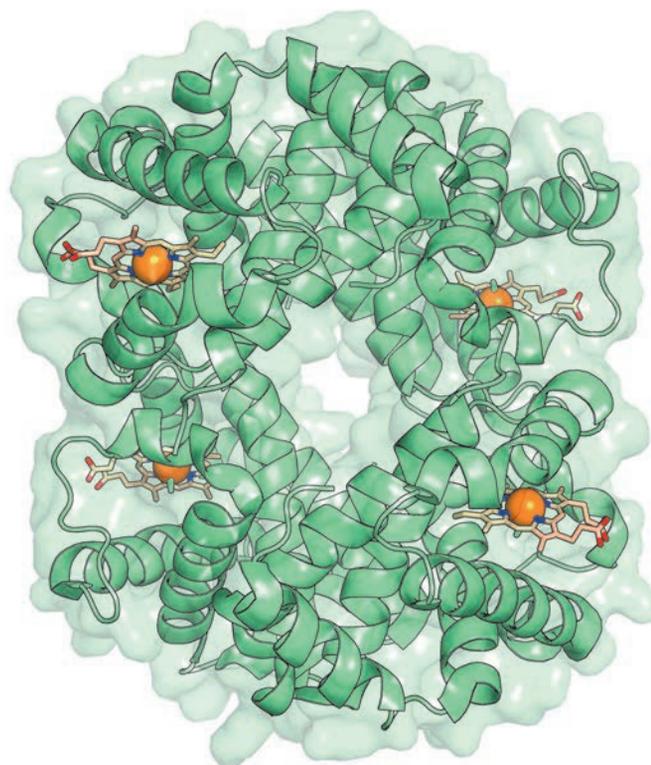


Рис. 2. Трехмерная структура гемоглобина человека. Молекулы гема (оранжевый) показаны в каждой субъединице тетрамера. Рисунок подготовлен с помощью программы PyMol на основании структуры 1GZX из PDB

структура гемоглобина показала, что молекулы гема, связывающие кислород, расположены на значительном расстоянии друг от друга в разных субъединицах белка (рис. 2) [15]. Это позволило сделать вывод о том, что в основе организации аллостерических белков лежит симметричное расположение субъединиц, причем существуют по меньшей мере два конформационных состояния каждой субъединицы – R (от англ. *relaxed*) и T (от англ. *tense*), которые, в случае гемоглобина, характеризуются высоким и низким сродством к кислороду соответственно. Переход от одной конформации к другой в результате связывания лиганда происходит во всех субъединицах согласованно и таким образом, что олигомерный белок не существует в гибридном состоянии (RT). Эта упрощенная модель позволила кинетически описать наблюдаемое увеличение сродства гемоглобина к кислороду по мере его насыщения [14]. Однако объяснить этот феномен на молекулярном уровне удалось только после проведения ряда структурных исследований [16, 17]. Стало понятно, что присоединение (отщепление) кислорода сопровождается значительными измене-

ниями пространственной организации функционального центра и разрывом (образованием) нескольких солевых мостиков, а возникающее при этом смещение субъединиц друг относительно друга приводит к тому, что эффект от связывания первой молекулы кислорода распространяется на весь тетрамер. Таким образом, присоединение и отщепление одной молекулы кислорода в одной субъединице инициирует соответствующий процесс в остальных субъединицах, что делает гемоглобин эффективным переносчиком кислорода по градиенту давления. Подобные кооперативные эффекты в гомоолигомерных белках/ферментах – один из наиболее известных примеров аллостерии. В этом случае аллостерическим центром по отношению к активному центру одной субъединицы является активный центр другой субъединицы, при этом связывание второй молекулы субстрата (или его аналога) приводит к аллостерическому влиянию на центр связывания первой молекулы субстрата и может не сопровождаться ее каталитическим превращением. В соответствии с «последовательной», или KNF-моделью (по заглавным буквам фамилий авторов Koshland–Nemethy–Filmer) [18], субъединицы в рамках мультимера изменяют свою конформацию поочередно, т.е. связывание лиганда меняет конформацию и свойства соответствующей субъединицы, что, в свою очередь, способно влиять на ее соседей. Иными словами, связывание лигандов последовательно приводит к изменению конформаций субъединиц. Такая модель описывает отрицательную кооперативность, которая наблюдается, например, в глицеральдегид-3-фосфатдегидрогеназе при связывании кофермента  $\text{NAD}^+$ , когда присоединение кофактора в активном центре одной субъединицы ослабляет его связывание в соседней вследствие перестройки внутри- и межсубъединичных контактов [19, 20]. Наблюдаемое свойство способствует поддержанию активности фермента на постоянном уровне вне зависимости от концентрации лиганда в среде. Хотя с целью выработки более общих закономерностей была предпринята попытка объединить модели MWC и KNF [21], дальнейшие исследования показали, что молекулярные механизмы аллостерической регуляции сложны и разнообразны, и ни одна из предложенных моделей не описывает исчерпывающе явление аллостерии.

В настоящее время можно считать общепринятым, что аллостерической регуляции подвержены не только полисубъединичные, но и мономерные белки, а аллостерическими лигандами служат прежде всего низкомолекулярные соединения, которые связываются в регуляторном центре, топографически независимом от функционального центра. Кроме того, к аллостерическим эффектам стали относить

регуляторное влияние, вызываемое белок-белковыми взаимодействиями, фосфорилированием и даже точечными мутациями. Многообразие конкретных примеров аллостерической регуляции белков и ферментов хорошо иллюстрирует ряд недавних статей [5, 7, 22]. Высказано предположение, что аллостерия может быть неотъемлемым свойством практически всех белков [8], исключение могут составить лишь структурные белки, жесткая конформация которых ограничивает их подвижность и возможности регуляции. Действительно, экспериментальные исследования все чаще обнаруживают аллостерию в ферментах, которые раньше не считались аллостерическими. Фосфофруктокиназа, катализирующая одну из ключевых стадий гликолиза, является примером белка со сложно организованной регуляцией, осуществляемой различными эффекторами. Аллостерия описана в ферментах этого суперсемейства – прокариотических [23] и эукариотических, причем последние имеют значительно большие размеры глобул вследствие дупликаций, вставок и мутаций предкового прокариотического гена, которые способствовали появлению новых аллостерических центров [24]. В то же время свойства фосфофруктокиназы из миксомицета *Dictyostelium discoideum* отличаются от свойств ее гомологов, и этот белок не считается аллостерическим. Однако делеция одного С-концевого остатка лейцина приводит к проявлению у мутантного белка аллостерических свойств, присущих другим представителям суперсемейства [25]. В качестве другого примера можно привести пируваткиназу [26]. У млекопитающих известно четыре изоформы этого фермента – L, R,  $M_1$  и  $M_2$ . Все изоформы, за исключением  $M_1$ , аллостерические и характеризуются положительной гомотропной кооперативностью по отношению к субстрату, а также положительной гетеротропной кооперативностью по отношению к фруктозо-1,6-дифосфату. При этом изоформы  $M_1$  и  $M_2$  образуются в результате альтернативного сплайсинга одного гена и отличаются на 23 аминокислотных остатка, локализованных в области межсубъединичных контактов, а также участвующих в образовании центра связывания фруктозо-1,6-дифосфата. Внесение двух точечных мутаций в структуру изоформы  $M_1$ , одной – в область межсубъединичного контакта [27], другой – в область центра связывания фруктозо-1,6-дифосфата [26], приводит к проявлению аллостерических свойств, характерных для других гомологов. Еще одним примером служит миоглобин, паралог<sup>1</sup> гемоглобина, существующий в виде трех основных конфор-

<sup>1</sup> Паралоги – эволюционно родственные белки, которые произошли вследствие дупликации предкового гена.

мационных состояний с разными каталитическими свойствами, так называемых таксономических под-состояний, внутри которых могут выделяться дополнительные конформационные вариации [28]. Предполагается, что миоглобин способен аллостерически контролировать протекание бимолекулярных реакций с двухатомными субстратами (например, NO и O<sub>2</sub>) за счет изменения геометрии консервативных полостей, примыкающих к активному центру. Интересно отметить, что во всех этих случаях аллостерическая регуляция обнаруживалась в белках, у гомологов которых это явление было описано ранее. Эти примеры говорят не только о широкой распространенности аллостерической регуляции, но и об общих механизмах этого явления в рамках суперсемейств белков. Рассмотренные случаи указывают также на возможность тонкой настройки аллостерии за счет нескольких точечных мутаций, но при этом подчеркивают сложную взаимосвязь функции и регуляции.

По современным представлениям структуру белка предлагается рассматривать в рамках сложных статистических ансамблей конформеров, которые, совершая локальные перегруппировки, находятся в постоянном движении [6, 29]. В этом контексте аллостерический эффект обусловлен перераспределением конформационных состояний [8]. Иными словами, связывание аллостерического эффектора приводит к популяционному сдвигу в сторону таких конформационных состояний макромолекулы, которые значительно отличаются функционально от нативного [30]. В свою очередь, в традиционно неаллостерических белках альтернативные конформа-

ции центров связывания и функционально важные конформационные переходы на сегодняшний день могут быть неизвестны. Но это не означает, что нельзя подобрать лиганд или особые условия среды, способные вызвать конформационный сдвиг и проявить аллостерические свойства в неаллостерических белках. В принципе, практически любое вещество, связавшись на поверхности белка, может вызвать популяционный сдвиг его конформаций, вопрос состоит лишь в эффективности такого сдвига и его влиянии на функцию [8]. Этот вопрос требует дальнейшего изучения на более широкой выборке белков. Однако конформационные изменения, связанные с аллостерической регуляцией, трудно выявить с помощью современных экспериментальных методов. Разработанные в последнее время подходы биоинформатики и компьютерной биологии предоставляют новые возможности для решения этой задачи.

### ПОИСК ЦЕНТРОВ СВЯЗЫВАНИЯ В БЕЛКАХ

Предсказание центров связывания в белках на основании информации об их структуре представляет собой новое направление в компьютерной биологии [31]. Для поиска полостей и карманов на поверхности структуры белков предложены разнообразные геометрические подходы, основанные на использовании пространственных критериев (таблица). Поскольку в большинстве случаев программы позволяют выявить несколько потенциальных сайтов, предпринимаются попытки их классифицировать с использованием различных геометрических параметров – размера, глубины и ориентации [32–34], а также различных статистических оценок, учиты-

Интернет-сервисы для предсказания сайтов связывания в структурах белков и их ранжирования по функциональной значимости

Название	Адрес в сети Интернет	Алгоритм поиска сайтов в структуре	Алгоритм ранжирования сайтов
Fpocket [35]	<a href="http://mobylye.rpbs.univ-paris-diderot.fr/">http://mobylye.rpbs.univ-paris-diderot.fr/</a> > Programs > Structure > Pockets > fpocket	Геометрический, с использованием диаграмм Вороного и построением $\alpha$ -сфер	Статистический, оценивается схожесть с известными сайтами связывания низкомолекулярных веществ
POCASA [33]	<a href="http://altair.sci.hokudai.ac.jp/g6/service/pocasa/">http://altair.sci.hokudai.ac.jp/g6/service/pocasa/</a>	Геометрический, с прокаткой сферических проб по поверхности белка	Геометрический, с учетом положения и размера сайта
pocketZebra [39]	<a href="http://biokinet.belozersky.msu.ru/pocketzebra">http://biokinet.belozersky.msu.ru/pocketzebra</a>	Геометрический, с использованием диаграмм Вороного и построением $\alpha$ -сфер	Биоинформатический, анализ специфических позиций подсемейств в суперсемействах белков
SiteHound [38]	<a href="http://scbx.mssm.edu/sitehound/sitehound-web/Input.html">http://scbx.mssm.edu/sitehound/sitehound-web/Input.html</a>	Энергетический, оценивается энергия взаимодействия аминокислот на поверхности с углеродной или фосфатной пробой	Энергетический, оценивается энергия взаимодействия аминокислот на поверхности с углеродной или фосфатной пробой
LIGSITE <sup>csc</sup> [41]	<a href="http://projects.biotec.tu-dresden.de/pocket/">http://projects.biotec.tu-dresden.de/pocket/</a>	Геометрический, на основе расчета поверхности Кюннолли	Биоинформатический, анализ консервативных позиций

вающих физико-химические свойства известных центров связывания лигандов [35, 36]. Альтернативу геометрическим методам представляют энергетические подходы, в которых предсказание и ранжирование сайтов основано на расчете энергии связывания небольших органических молекул (проб) на поверхности белка [37, 38]. Все эти подходы, основываясь на структурной информации, способны быстро и эффективно обнаружить полости и карманы потенциальных сайтов связывания, однако не дают представления об их функциональной значимости и структуре комплементарных лигандов.

В процессе эволюции от общего предка одни свойства белков сохраняются, в то время как другие претерпевают изменения в результате естественного отбора, что приводит к формированию функционального разнообразия. Так, гомологичные ферменты одного суперсемейства могут обладать общей структурной организацией и химизмом катализируемой реакции, но отличаться по другим функциональным свойствам (субстратная специфичность, энантио- и региоселективность, тип химического превращения), а также принципам их регуляции. Существенный рост публичных баз данных, содержащих геномную и структурную информацию о белках/ферментах, позволяет проводить масштабный сравнительный анализ в рамках суперсемейств – совокупностей как близких, так и эволюционно удаленных гомологов. Не все позиции в структуре белков одинаково подвержены изменчивости в процессе эволюции, что отражает различия в давлении отбора на остатки с разной функциональной ролью. Этот факт позволяет использовать биоинформатический анализ суперсемейств белков для поиска эволюционных закономерностей, характерных для аминокислотных остатков функциональных и регуляторных центров связывания [39] (таблица). Полностью консервативные позиции играют ключевую роль в функции, общей для всего суперсемейства, например, принимают участие в каталитическом механизме действия ферментов. Нужно отметить, что не всегда каталитические аминокислоты в суперсемействе полностью консервативны, кроме того, известны случаи миграции каталитических аминокислот в структурах гомологичных белков [2, 40]. Так, например, нуклеофил каталитической триады  $\alpha/\beta$ -гидролаз может быть представлен как серином, так и цистеином или аспаратом, а каталитическая кислота может располагаться по меньшей мере в двух альтернативных позициях основной полипептидной цепи. Тем не менее показано, что консервативность остатков в полостях на поверхности белков может служить критерием при аннотации функциональных центров [41–43], и это свойство может быть исполь-

зовано для характеристики широкого круга ферментов [44–46]. Фактически, описание любого нового белка с неизвестной функцией разумно начинать с его сравнительного анализа с ближайшими гомологами и расчета консервативных позиций в колонках множественного выравнивания. Роль наиболее консервативных остатков, обнаруженных при таком анализе, может быть проверена экспериментально посредством внесения точечных мутаций и оценке их влияния на каталитические свойства. Благодаря критерию консервативности предсказание функциональных центров в новом белке с неизвестной функцией возможно даже при отсутствии данных о его трехмерной структуре при условии, что доступна аннотация функциональных центров в его гомологах. На сегодняшний день такая информация о достаточно большом числе белков из различных суперсемейств собрана в публичных базах данных (см. следующую главу). Таким образом, комплексное использование геометрических структурных подходов и биоинформатического анализа делает поиск функциональных центров в белках более эффективным.

Сравнительное исследование белков позволило сделать вывод о том, что аллостерические сайты характеризуются меньшим содержанием консервативных позиций и более высоким содержанием переменных позиций [47]. Показано также, что мутагенез переменных позиций в аллостерических центрах приводит к изменению аллостерического эффекта, в то время как замена консервативных позиций в этих же центрах в основном приводит к потере каталитической функции. Эти результаты говорят о малой пригодности критерия консервативности при поиске регуляторных центров и свидетельствуют о важной роли переменных позиций в связывании аллостерических лигандов и регуляции функциональных свойств в суперсемействах белков. В этом отношении особый интерес представляют специфические позиции подсемейств (СПП) – консервативные внутри функциональных подсемейств, но различающиеся между ними [48, 49]. Присутствие СПП характерно как для каталитических, так и аллостерических сайтов и может быть эффективным критерием идентификации функциональных и регуляторных центров в структурах белков [39]. Поиск статистически значимых СПП может помочь понять разницу в строении участков связывания в родственных белках. Так, например, ДНК-зависимая РНК-полимераза (РНКП) является ключевым ферментом транскрипции, который присутствует во всех живых системах. Каталитическая основа бактериального фермента состоит из субъединиц  $\alpha_2\beta\beta'\omega$  и характеризуется высокой степенью сходства по структуре и функции с гомологами всех известных орга-

низмов. Подтверждено, что бактериальная РНКП служит мишенью для противомикробных средств [12]. Рифампицин – антибиотик «первого эшелона» при туберкулезе – селективно ингибирует транскрипцию в *Mycobacterium tuberculosis* за счет взаимодействия с аллостерическим центром, расположенным в  $\beta$ -субъединице фермента. Взаимодействие ингибитора с бактериальным ферментом напрямую блокирует элонгацию в патогене, не затрагивая при этом родственный фермент человека [50]. Биоинформатический анализ суперсемейства РНКП показывает, что селективность связывания рифампицина с ферментом бактерий обусловлена наличием специфических позиций подсемейств в соответствующем участке связывания, которые отличаются у прокариот и человека (рис. 3). Этот пример свидетельствует о том, что изучение роли СПП в формировании сайтов связывания регуляторных лигандов позволяет лучше понять молекулярные механизмы специфического узнавания аллостерических эффекторов и закономерности функциональной регуляции в семействах белков.

### ПУБЛИЧНЫЕ БАЗЫ ДАННЫХ

Возможность быстрого получения информации из открытых источников в сети Интернет является важным фактором развития современной науки. В этом контексте особое место занимают публичные базы данных.

База данных каталитических сайтов Catalytic Site Atlas (CSA) – один из основных источников информации о ферментах [51]. В основе CSA лежат экспериментальные данные о 1000 каталитически активных белков с разными свойствами, а методы поиска сходства по геномным последовательностям применяются для того, чтобы с высокой степенью значимости аннотировать каталитические остатки в других белках на основании критерия консервативности. В результате использования методов биоинформатики база данных CSA содержит данные об активных центрах нескольких десятков тысяч белков с известной структурой.

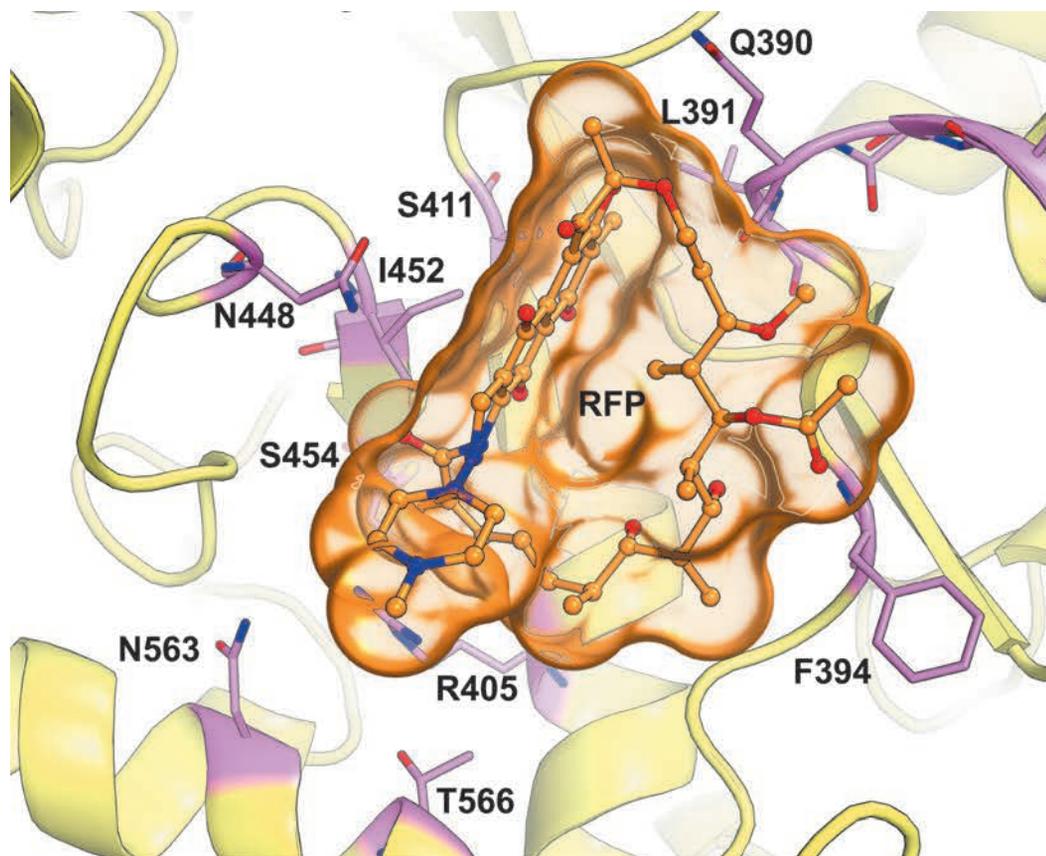
Объем знаний об аллостерических белках не столь значителен, что, по всей видимости, связано с недостатком структурной информации и сложностями при определении аллостерических сайтов. Однако и в этой области достигнут значительный прогресс. Если 50 лет назад (когда была предложена первая модель кооперативности) были известны 24 аллостерических белка, то сегодня насчитываются сотни таких примеров. Недавнее создание публичной базы данных аллостерических центров Allosteric Database (ASD) стало первой попыткой обобщить массив соответствующей опубликованной информации [52].

На сегодняшний день база ASD насчитывает почти 2000 сайтов, однако далеко не все сведения подкреплены кристаллографической информацией о структуре белка в комплексе с эффектором, а некоторые аннотации и вовсе неоднозначны. Тем не менее можно ожидать, что систематизация экспериментального материала о структуре и функции аллостерических центров будет продолжаться, в том числе и в рамках других публичных баз данных.

### ВЗАИМОСВЯЗЬ ФУНКЦИИ И РЕГУЛЯЦИИ

Не вызывает сомнений тот факт, что конформационные изменения в структуре белка, обусловленные связыванием лиганда в аллостерическом центре, в конечном итоге приводят к изменению функциональных свойств. Однако не много известно о конкретных молекулярных механизмах этого явления. Как объяснить кооперативность, наблюдаемую при связывании различных лигандов, и как предсказать взаимодействие между сайтами в белках, в которых аллостерия еще не описана? В последние годы предпринято несколько попыток понять взаимосвязь функции и регуляции [6]. В основе этих подходов лежат компьютерные методы поиска корреляций – структурных или эволюционных – между событиями, происходящими при связывании лигандов в топологически независимых центрах на поверхности белков. Рассмотрим несколько таких примеров.

Структурные изменения, возникающие при связывании лиганда, непосредственно связаны с конформационной подвижностью белковой глобулы. Молекулярная динамика зарекомендовала себя как эффективный метод изучения структурных изменений в белках [53, 54], в том числе согласованных флуктуаций атомов, происходящих в результате коллективного движения [55]. Так, расчет матриц ковариаций флуктуации атомов вдоль МД-траекторий был использован для изучения молекулярных механизмов аллостерической регуляции белка-супрессора опухолевого роста Хиппеля–Линдау (pVHL) [56]. pVHL в свободном состоянии нестабилен и существует в виде так называемой «расплавленной глобулы». Белок стабилизируется при взаимодействии с элонгинами С и В, которые вместе образуют лигандраспознающий компонент для связывания фактора транскрипции HIF в составе ферментного комплекса убиквитинлигазы E3 (рис. 4). С использованием МД показано, что интерфейс между двумя доменами  $\alpha$  и  $\beta$  в pVHL представляет собой наименее стабильный участок белка. Были отобраны такие аминокислотные остатки, подвижность которых вдоль МД-траекторий в наибольшей степени коррелирует с движением остатков из нестабильной междоменной области pVHL. Молекулярное моделирова-



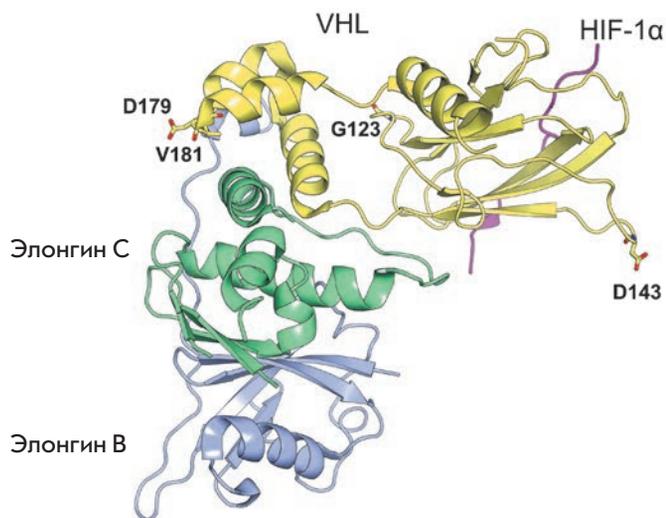
	390	411	447	455	562	568
PDB: 1YNN	<b>QLSQFK</b> DETNP LSSLRHKRRIS		ANIGLITSL		SNMQTQA	
RPOB_THEAQ	<b>QLSQFK</b> DETNP LSSLRHKRRIS		ANIGLITSL		SNMQTQA	
RPOB_NITEU	<b>QLSQFMDQ</b> TNPLSEVTHKRRIS		PNIGLINSL		SNMQRQA	
RPOB_AZOSE	<b>QLSQFMDQ</b> TNPLSEITHKRRVS		PNIGLINSL		ANMQRQA	
RPOB_GEOTN	<b>QLSQFMDQ</b> TNPLAELTHKRRLS		PNIGLINSL		ANMQRQA	
RPOB_CHAVU	<b>PLCQFMDQ</b> TNPLAEITHKRRIS		QNAGLVSSL		SNMQRQA	
RPOB_STRPN	<b>QLSQFMDQ</b> HNPLSELCHKRRIS		PNIGLINNL		ANMQRQA	
RPOB_MACCJ	<b>QLSQFMDQ</b> ANPLAELTHKRRIS		PNIGLINSL		ANMQRQA	
RPOB_MYCVP	<b>QLSQFMDQ</b> NNPLSGLTHKRRIS		PNIGLIGSL		ANMQRQA	
RPOB_MYCUA	<b>QLSQFMDQ</b> NNPLSGLTHKRRIS		PNIGLIGSL		ANMQRQA	
RPOB_MYCLE	<b>QLSQFMDQ</b> NNPLSGLTHKRRIS		PNIGLIGSL		ANMQRQA	
RPOB_MYCTA	<b>QLSQFMDQ</b> NNPLSGLTHKRRIS		PNIGLIGSL		ANMQRQA	
<hr/>						
RPA2_HUMAN	<b>GLCVVADK</b> LNFI RYLSHFRCVH		EPCGLM NHL		CQMGKQT	
RPB2_HUMAN	<b>GVSQVLN</b> R LTFAS TLSHLRRLN		HAVGLV KNL		SAMGKQA	
RPC2_HUMAN	<b>GVTQVLS</b> R LSYISALGMMTRIS		EACGLV KNL		CAMGKQA	
RPA2_MOUSE	<b>GLCVVADK</b> LNFLRYLSHFRCVH		APCGLL NHL		CQMGKQT	
RPB2_CAEL	<b>GVSQVLN</b> R LTYTATLSHLRRAN		QAVGLV KNL		SAMGKQA	
RPC2_DROME	<b>GVTQVLS</b> R LSYISALGMMTRVN		EACGLV KNL		CAMGKQA	
RPA2_RAT	<b>GLCVVADK</b> LNFI RYLSHFRCVH		APCGLL NHL		CQMGKQT	
RPB2_YEAST	<b>GVSQVLN</b> R YTSSTLSHLRRTN		QACGLV KNL		SAMGKQA	
RPOB_SULAC	<b>GVSQLLD</b> R TNWLSMLSHLRRVV		PNSGLV KNL		SAMAKQS	
RPB2_CANGA	<b>GVSQVLN</b> R YTSSTLSHLRRTN		QACGLV KNL		SAMGKQA	
NRPB2_ARATH	<b>GVSQVLN</b> R LTYASTLSHLRRLN		QACGLV KNL		SAMGKQA	

Рис. 3. Связывание рифампицина (RFP) в аллостерическом сайте β-субъединицы бактериальной РНКП. Показаны специфические позиции подсемейств (розовый), определенные с помощью биоинформатического анализа 271 РНКП из различных источников, и приведены соответствующие фрагменты выравнивания аминокислотных последовательностей. Интерактивная версия доступна в сети Интернет (см. «Приложение»). Рисунок подготовлен с помощью программы PyMol на основании структуры 1YNN из PDB

ние показало, что замена аминокислот в отобранных позициях на гомологичные остатки в аминокислотной последовательности более стабильного pVHL из *Caenorhabditis elegans* приводит к значительной стабилизации pVHL из человека как в свободном состоянии, так и в составе комплекса. При этом мутации остатков в областях, топологически независимых от центров связывания элонгина С и фактора транскрипции HIF, приводят к стабилизации комплекса pVHL–элонгин С и снижению энергии связывания HIF в pVHL. Авторы работы [56] делают вывод, что стабильность и эффективность связывания в pVHL могут регулироваться аллостерически с помощью лекарственных средств, мимикрирующих эффект внесенных мутаций.

Другой пример свидетельствует о пользе поиска эволюционных корреляций с использованием статистического анализа геномных последовательностей [57, 58]. Подход основан на предположении, что если два сайта в структуре белка функционально связаны, то соответствующие аминокислотные остатки в гомологичных белках должны коэволюционировать в процессе развития от общего предка и эта корреляция может быть выявлена с помощью статистического сравнения аминокислотных последовательностей. В таком случае корреляция встречаемости аминокислот в двух сайтах в структурах родственных белков может означать существование функциональной зависимости между ними. Описанный подход применили для анализа мембранного белка FesA – представителя семейства TonB-зависимых транспортных белков, функция которого заключается в перекачке железа через внешнюю мембрану грамотрицательных бактерий [59]. Ключевая стадия переноса железа – взаимодействие белка TonB, создающего протонный градиент, с консервативным N-концевым TonB-связывающим мотивом транспортного белка. Предполагалось, что это взаимодействие, протекающее на периплазматической стороне мембраны, вызывает конформационные изменения, которые приводят к связыванию сидерофора<sup>1</sup> на противоположной поверхности мембраны со стороны внеклеточного пространства и последующему импорту железа, однако конкретные механизмы этой аллостерической коммуникации между двумя центрами связывания, расположенными в разных компартментах на значительном расстоянии друг от друга, оставались неизвестными. Статистический анализ последовательностей TonB-зависимых транспортных белков обнаружил рассеянную, но связную сеть коррелирующих остатков, которая обеспечивает функцио-

<sup>1</sup> Сидерофор – низкомолекулярное соединение с высоким сродством к ионам железа (например, цитрат железа).



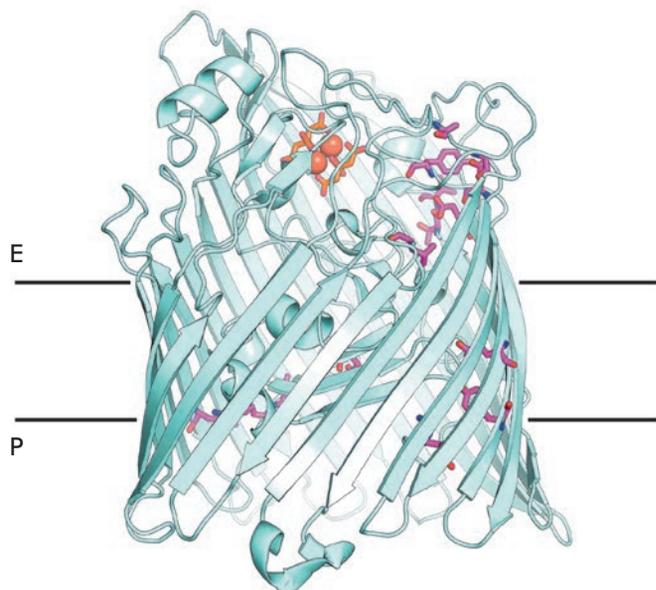
**Рис. 4.** Комплекс фактора транскрипции HIF (пурпурный) – белка-супрессора опухолевого роста Хиппеля–Линдау VHL (желтый) – элонгина С (зеленый) – элонгина В (синий). Домен  $\alpha$  белка VHL взаимодействует с элонгином С, в то время как домен  $\beta$  связывает HIF. Показаны аминокислотные остатки, движение которых в процессе МД-траекторий в наибольшей степени коррелирует с остатками из междоменной области pVHL. Рисунок подготовлен с помощью программы PyMol на основании структуры 1LM8 из PDB для иллюстрации результатов из [56]

нальную коммуникацию между периплазматическим и внеклеточным сайтами связывания в FesA (рис. 5), а направленный мутагенез обнаруженных остатков, ни один из которых не участвует непосредственно в образовании центров связывания TonB и сидерофора, привел к нарушению транспортной функции FesA.

Основываясь на известных к этому моменту результатах, можно сделать вывод, что поиск эволюционных и структурных корреляций может служить важным инструментом изучения молекулярных механизмов аллостерической регуляции в белках.

#### ЦЕНТРЫ СВЯЗЫВАНИЯ В BIOTECHNOLOGY И BIOMEDICINE

Сайты связывания субстратов/лигандов в ферментах/белках активно изучаются с целью создания новых биокатализаторов для промышленного использования (см., например, обзор [60]), а также ингибиторов, используемых в качестве лекарственных средств человека [61, 62]. В каждом случае применяются свои методы, однако в основе большинства успешных исследований лежит общий принцип, который можно назвать «стохастический анализ». Для дизай-



**Рис. 5.** Структура транспортного белка FecA внешней мембраны грамотрицательных бактерий. Периплазматическое пространство (P) отделено от внеклеточного пространства (E) схематическим изображением мембраны. Ионы железа в составе цитрата железа (оранжевый) показаны в центре связывания сидерофора. Показаны коррелирующие остатки (пурпурный), участвующие в образовании сети взаимодействий, которая обеспечивает функциональную коммуникацию между периплазматическим и внеклеточным сайтами связывания в FecA. Рисунок подготовлен с помощью программы PyMol на основании структуры 1KMP из PDB для иллюстрации результатов из [59]

на белков с новыми свойствами предложены так называемые подходы «направленной эволюции» [63, 64]. Эти методы сочетают в себе случайный мутагенез структуры белка, сопряженный с последующим скринингом для отбора наиболее предпочтительных фенотипов. Мутации вносят в структуру или определенную часть структуры белка произвольным образом, после чего эффект замен оценивают экспериментально и отбирают те из них, которые приводят к улучшению нужного свойства. В последние годы стохастические подходы стали более быстрыми и эффективными за счет использования различных методов статистического анализа и компьютерной биологии [65, 66]. Однако они по-прежнему требуют значительных ресурсов, экспериментального исследования больших библиотек мутантов, разработки эффективных методов скрининга и при этом способны изучить лишь очень небольшую часть возможных изменений в структуре белка. Эффективное

применение методов случайной эволюции затруднено высокой вероятностью отрицательных мутаций и низкой вероятностью проявления функционально полезных фенотипов. Аналогичным образом дизайн новых лекарственных средств в большинстве случаев основан на слепом экспериментальном скрининге огромных библиотек низкомолекулярных соединений в качестве потенциальных ингибиторов белка-мишени [67, 68]. Хотя структуру такого случайно обнаруженного ингибитора далее можно оптимизировать с использованием экспериментальных и компьютерных подходов, это в целом чрезвычайно трудоемкий и малоэффективный подход. Так, компания GlaxoSmithKline за период с 1995 по 2001 год провела 70 высокопроизводительных экспериментальных скрининговых исследований (стоимостью около \$1 млн каждый) белков-мишеней, отобранных из различных болезнетворных бактерий, с использованием собственных коллекций потенциальных ингибиторов, содержащих 260000–530000 соединений. По результатам шестилетних исследований для дальнейших испытаний отобрали всего пять лидерных компонентов с подтвержденными эффектами [69]. Обзор опубликованных данных показывает, что в период с 1996 по 2004 год подобные исследования проводились по крайней мере 34 различными компаниями на 60 мишенях и в целом считаются неудачными [70]. Высокие затраты на экспериментальные исследования в сочетании с низкой результативностью эмпирического поиска предопределили значительное падение интереса к этой методологии.

Несмотря на кажущуюся универсальность стохастических подходов, они, как правило, направлены на изучение функциональных центров. В большинстве случаев с целью изменения каталитических свойств случайные мутации вносят в структуру активного центра фермента [71]. Аналогично, большинство разрабатываемых лекарственных средств связываются с функциональными центрами белков (см., например, [72]). Примеров практического использования аллостерических сайтов значительно меньше. Так, показано, что введение единственной мутации в структуру глицеральдегид-3-фосфатдегидрогеназы приводит к разрушению солевого мостика вблизи активного центра и потере кооперативности при связывании  $NAD^+$  [73]. При этом полученный препарат фермента характеризуется двукратным увеличением удельной активности. Известны лекарственные средства, которые взаимодействуют с регуляторными сайтами белков. Так, рифампицины и миксопиранины связываются в полостях субъединиц  $\beta$  и  $\beta'$  РНКП, топологически независимых от активного центра, и прямо блокируют работу фермента [12]. Связывание ингибитора до-

рамапимода в аллостерическом центре р38 МАР-киназы человека и сопутствующие этому процессу конформационные перестройки приводят к блокировке связывания АТФ [74]. Ингибиторы обратной транскриптазы ВИЧ-1 эфавиренц, невирапин и далавидин связываются в аллостерическом сайте на значительном расстоянии от активного центра [75]. Обобщая имеющийся опыт, следует отметить, что благодаря более высокой вариабельности в рамках суперсемейства аллостерические сайты можно считать не менее привлекательными мишенями для селективного ингибирования, чем каталитические [76].

Низкая эффективность стохастических методов стимулирует развитие компьютерных подходов к дизайну эффективных биокатализаторов и поиску селективных ингибиторов ферментов. Показано, что биоинформатический анализ эволюционных взаимосвязей в функционально разнообразных суперсемействах позволяет обнаруживать не только ключевые «горячие точки» в структурах ферментов, но и конкретные аминокислотные замены, внесение которых позволяет получать мутанты с улучшенными свойствами [54, 77–79]. Применение вычислительных подходов к дизайну белков подробно рассмотрено в недавних обзорах [80–82]. Секвенирование полных геномов основных бактериальных патогенов, в том числе *M. tuberculosis* [83], положило начало применению компьютерной геномики в медицине. Использование геномных подходов позволяет не только составить список всех потенциальных белков-мишеней в конкретном организме, но и выявить наиболее перспективные объекты для последующего экспериментального исследования [84]. Важным преимуществом такого постгеномного анализа принято считать возможность выбора мишени, которая встречается во многих, либо, наоборот, всего в нескольких бактериальных геномах. Предполагалось, что таким путем можно создавать как лекарственные средства с широким спектром действия, так и высокоспецифичные ингибиторы определенных патогенов. Кроме того, сравнительный анализ геномов бактерий и животных позволяет исключать из списка возможных мишеней такие белки, гомологи которых есть в организме человека, что может позволить избежать ситуаций, когда созданный прототип токсичен не только для бактерий, но и для человека [85]. Однако, обобщая, следует отметить, что используемые до последнего времени постгеномные методы весьма поверхностно подходили к выбору мишеней, на которые действуют лекарственные средства. Так, например, предпочтительными мишенями при поиске новых антибиотиков считали белки, консервативные у бактерий и отсутствующие в организме челове-

ка. При этом не учитывали особенности структурной организации этих белков, в том числе строение конкретных участков связывания потенциальных ингибиторов. Исключение из списка потенциальных мишеней белков патогенных микроорганизмов, гомологи которых есть у человека (в попытке избежать токсичности ингибитора), в целом следует признать малообоснованным. Важнейшие метаболические пути, как правило, консервативны, и ключевые ферменты этих путей содержатся как в патогенных бактериях, так и в организме человека. В качестве примера можно привести уже упоминавшийся противотуберкулезный препарат рифампицин, который ингибирует репликацию бактерий, селективно связываясь с  $\beta$ -субъединицей РНКП, гомолог которой есть и в организме человека [50].

Подводя итог, можно отметить общую тенденцию отказа от малоэффективных стохастических подходов в пользу более рациональных и фокусированных стратегий. В этом контексте роль методов биоинформатики и молекулярного моделирования в биотехнологии и биомедицине продолжает неуклонно возрастать. Разработка новых подходов к систематическому анализу различных центров связывания в больших суперсемействах белков должна помочь, с одной стороны, установить взаимосвязь между структурой, функцией и регуляцией белков/ферментов, и, с другой стороны, обнаружить участки связывания новых субстратов, а также ингибиторов/эффекторов с ранее неизвестным механизмом действия.

## ЗАКЛЮЧЕНИЕ

Белок-белковые взаимодействия, а также взаимодействие белков (ферментов) с различными низкомолекулярными соединениями являются важнейшими факторами регуляции их функциональных свойств. В настоящее время наиболее хорошо изучены молекулярные механизмы действия лигандов, связывающихся в функциональных сайтах белков (активных центрах ферментов), в то время как механизмы аллостерического ингибирования или связывания в других малоизвестных сайтах в структурах белков изучены недостаточно. В этом контексте большой интерес представляют не только взаимодействия между функциональным и аллостерическим центрами, но также поиск и характеристика новых сайтов связывания, их роль в функционировании белка. Несмотря на первые шаги в сторону большего понимания взаимосвязи структуры, функции и регуляции, этот вопрос еще далек от разрешения и требует дальнейшего изучения. Анализ опубликованных данных позволяет сделать вывод о том, что роль методов биоинформатики и молекулярного

моделирования в изучении роли различных участков связывания в функционировании белка, в том числе аллостерических эффектов, будет неуклонно расти. Установление эволюционных взаимосвязей между различными сайтами связывания в рамках суперсемейств белков, открытие новых функциональных, аллостерических и регуляторных сайтов с использованием вычислительных подходов должны улучшить наше понимание структурно-функциональных взаимосвязей в белках и предоставить новые возможности как для создания новых лекарственных средств, так и для дизайна более эффективных биокатализаторов. ●

*Работа выполнена при финансовой поддержке  
Российского научного фонда  
(грант № 15-14-00069).*

## ПРИЛОЖЕНИЕ

### Выравнивание белков суперсемейства РНКП

Структура бактериальной РНКП (код структуры в PDB: 1YNN) была использована в качестве заправки. Для реконструкции выборки эволюционно удаленных РНКП различных организмов (бактерий, животных, человека и т.д.) использовали поиск по структурному сходству в базе данных PDB с помощью алгоритма superpose пакета CCP4 [86]. При парном структурном сравнении допускалось не менее 30% совпадений элементов вторичной структуры в каждом белке. Для каждого обнаруженного удаленного гомолога реконструировали эволюционно близкие белки с использованием алгоритма BLAST и базы данных Swiss-Prot [87]. Получившиеся выборки фильтровали для удаления избыточных (повторя-

ющихся) последовательностей на уровне 95% парной идентичности, а также слишком непохожих белков со сходством менее 0.5 бит в колонке [88] по отношению к соответствующему гомологу с известной структурой. Структуры белков выравнивали с помощью алгоритма Matt [89], выравнивание последовательностей – с использованием программы T-coffee [90]. Полученное выравнивание удаленных гомологов использовали в качестве основы для выравнивания близких гомологов. Полученное структурно-опосредованное множественное выравнивание суперсемейства РНКП содержит 271 последовательность белков.

### Поиск потенциальных сайтов связывания

Идентификацию карманов и полостей на поверхности структуры РНКП (код структуры в PDB: 1YNN), потенциально способных связывать низкомолекулярные лиганды, выполняли с помощью алгоритма Focket [35].

### Биоинформатический анализ

Поиск СПП и специфических сайтов связывания в суперсемействе РНКП выполняли с использованием алгоритмов Zebra [91] и pocketZebra [39].

### Структурный анализ

Визуализацию и анализ структурной информации о белках выполняли с использованием программы PyMol (Schrödinger LLC).

### Доступ

Интерактивный доступ к результатам биоинформатического анализа суперсемейства ферментов ДНК-зависимых РНК-полимераз предоставлен через интернет-сервис <http://biokinet.belozersky.msu.ru/pocketzebra> (в разделе «Примеры»).

## СПИСОК ЛИТЕРАТУРЫ

- Thornton J.M., Todd A.E., Milburn D., Borkakoti N., Orengo C.A. // Nat. Struct. Mol. Biol. 2000. V. 7. P. 991–994.
- Todd A.E., Orengo C.A., Thornton J.M. // Curr. Opin. Chem. Biol. 1999. V. 3. № 5. P. 548–556.
- Martin A.C., Orengo C.A., Hutchinson E.G., Jones S., Karmirantzou M., Laskowski R.A., Mitchell J., Taroni C., Thornton J.M. // Structure. 1998. V. 6. № 7. P. 875–884.
- Jones S., Thornton J.M. // Curr. Opin. Chem. Biol. 2004. V. 8. № 1. P. 3–7.
- Laskowski R.A., Gerick F., Thornton J.M. // FEBS Lett. 2009. V. 583. № 11. P. 1692–1698.
- Goodey N.M., Benkovic S.J. // Nat. Chem. Biol. 2008. V. 4. № 8. P. 474–482.
- Hardy J.A., Wells J.A. // Curr. Opin. Struct. Biol. 2004. V. 14. № 6. P. 706–715.
- Gunasekaran K., Ma B., Nussinov R. // Proteins. 2004. V. 57. № 3. P. 433–443.
- Campbell E.A., Pavlova O., Zenkin N., Leon F., Irschik H., Jansen R., Severinov K., Darst S.A. // EMBO J. 2005. V. 24. № 4. P. 674–682.
- Esyunina D., Klimuk E., Severinov K., Kulbachinskiy A. // Proc. Natl. Acad. Sci. USA. 2015. V. 112. № 7. P. 2017–2022.
- Sousa R. // Cell. 2008. V. 135. № 2. P. 205–207.
- Darst S.A. // Trends Biochem. Sci. 2004. V. 29. № 4. P. 159–162.
- Monod J., Changeux J.P., Jacob F. // J. Mol. Biol. 1963. V. 6. № 4. P. 306–329.
- Monod J., Wyman J., Changeux J.P. // J. Mol. Biol. 1965. V. 12. № 1. P. 88–118.
- Perutz M.F., Rossmann M.G., Cullis A.F., Muirhead H., Will G. // Nature. 1960. V. 185. P. 416–422.
- Perutz M.F., Wilkinson A.J., Paoli M., Dodson G.G. // Annu. Rev. Biophys. Biomol. Struct. 1998. V. 27. № 1. P. 1–34.
- Eaton W.A., Henry E.R., Hofrichter J., Mozzarelli A. // Nat. Struct. Mol. Biol. 1999. V. 6. № 4. P. 351–358.

18. Koshland D.E., Nemethy G., Filmer D. // *Biochemistry*. 1966. V. 5. № 1. P. 365–385.
19. Conway A., Koshland D.E. // *Biochemistry*. 1968. V. 7. № 11. P. 4011–4023.
20. Makshakova O.N., Semenyuk P.I., Kuravsky M.L., Ermakova E.A., Zuev Y.F., Muronetz V.I. // *J. Struct. Biol.* 2015. V. 190. № 2. P. 224–235.
21. Eigen M. // *Nobel Symp.* 1967. V. 5. P. 333–369.
22. Arkin M.R., Wells J.A. // *Nat. Rev. Drug Discov.* 2004. V. 3. № 4. P. 301–317.
23. Schirmer T., Evans P.R. // *Nature*. 1990. V. 343. № 6254. P. 140–145.
24. Poorman R.A., Randolph A., Kemp R.G., Henrikson R.L. // *Nature*. 1984. V. 309. № 5967. P. 467–469.
25. Santamaría B., Estévez A.M., Martínez-Costa O.H., Aragón J.J. // *J. Biol. Chem.* 2002. V. 277. № 2. P. 1210–1216.
26. Ikeda Y., Taniguchi N., Noguchi T. // *J. Biol. Chem.* 2000. V. 275. № 13. P. 9150–9156.
27. Ikeda Y., Tanaka T., Noguchi T. // *J. Biol. Chem.* 1997. V. 272. № 33. P. 20495–20501.
28. Frauenfelder H., McMahon B.H., Austin R.H., Chu K., Groves J.T. // *Proc. Natl. Acad. Sci. USA*. 2001. V. 98. № 5. P. 2370–2374.
29. Hilser V.J., Thompson E.B. // *Proc. Natl. Acad. Sci. USA*. 2007. V. 104. № 20. P. 8311–8315.
30. Chennubhotla C., Yang Z., Bahar I. // *Mol. Biosyst.* 2008. V. 4. № 4. P. 287–292.
31. Henrich S., Salo-Ahen O.M., Huang B., Rippmann F.F., Cruciani G., Wade R.C. // *J. Mol. Recognit.* 2010. V. 23. № 2. P. 209–219.
32. Weisel M., Proschak E., Schneider G. // *Chem. Cent J.* 2007. V. 1. № 7. P. 1–17.
33. Yu J., Zhou Y., Tanaka I., Yao M. // *Bioinformatics*. 2010. V. 26. № 1. P. 46–52.
34. Yaffe E., Fishelovitch D., Wolfson H.J., Halperin D., Nussinov R. // *Nucl. Acids Res.* 2008. V. 36. Suppl 2. P. W210–W215.
35. Le Guilloux V., Schmidtke P., Tuffery P. // *BMC Bioinformatics*. 2009. V. 10. № 1. P. 168.
36. Volkamer A., Kuhn D., Grombacher T., Rippmann F., Rarey M. // *J. Chem. Information Modeling*. 2012. V. 52. № 2. P. 360–372.
37. Laurie A.T.R., Jackson R.M. // *Bioinformatics*. 2005. V. 21. № 9. P. 1908–1916.
38. Hernandez M., Ghersi D., Sanchez R. // *Nucl. Acids Res.* 2009. V. 37. Suppl 2. P. W413–W416.
39. Suplatov D., Kirilin E., Arbatsky M., Takhaviev V., Švedas V. // *Nucl. Acids Res.* 2014. V. 42. № W1. P. W344–W349.
40. Todd A.E., Orengo C.A., Thornton J.M. // *J. Mol. Biol.* 2001. V. 307. № 4. P. 1113–1143.
41. Huang B., Schroeder M. // *BMC Struct. Biol.* 2006. V. 6. № 1. P. 19.
42. Glaser F., Morris R.J., Najmanovich R.J., Laskowski R.A., Thornton J.M. // *Proteins*. 2006. V. 62. № 2. P. 479–488.
43. Kalinina O.V., Gelfand M.S., Russell R.B. // *BMC Bioinformatics*. 2009. V. 10. № 1. P. 174.
44. Varfolomeev S.D., Gurevich K.G., Poroykov V.V., Sobolev B.N., Fomenko A.E. // *Dokl. Biochem. Biophys.* 2001. V. 379. № 1. P. 252–254.
45. Халудлин И.Г., Суплатов Д.А., Шалаева Д.Н., Оцука М., Асано Я., Швядас В.К. // *Acta Naturae*. 2010. T. 2. № 2. C. 70–74.
46. Суплатов Д.А., Аржаник В.К., Швядас В.К. // *Acta Naturae*. 2011. T. 3. № 1. C. 99–105.
47. Yang J.S., Seo S.W., Jang S., Jung G.Y., Kim S. // *PLoS Comput. Biol.* 2012. V. 8. № 7. P. e1002612–e1002612.
48. Kalinina O.V., Mironov A.A., Gelfand M.S., Rakhmaninova A.B. // *Protein Sci.* 2004. V. 13. № 2. P. 443–456.
49. Suplatov D., Shalaeva D., Kirilin E., Arzhanik V., Švedas V. // *J. Biomol. Struct. Dyn.* 2014. V. 32. № 1. P. 75–87.
50. Campbell E.A., Korzheva N., Mustaev A., Murakami K., Nair S., Goldfarb A., Darst S.A. // *Cell*. 2001. V. 104. № 6. P. 901–912.
51. Porter C.T., Bartlett G.J., Thornton J.M. // *Nucl. Acids Res.* 2004. V. 32. Suppl 1. P. D129–D133.
52. Huang Z., Zhu L., Cao Y., Wu G., Liu X., Chen Y., Wang Q., Shi T., Zhao Y., Wang Y., et al. // *Nucl. Acids Res.* 2011. V. 39. Suppl 1. P. D663–D669.
53. Seifert A., Tatzel S., Schmid R.D., Pleiss J. // *Proteins*. 2006. V. 64. № 1. P. 147–155.
54. Suplatov D., Panin N., Kirilin E., Shcherbakova T., Kudryavtsev P., Švedas V. // *PloS One*. 2014. V. 9. № 6. P. e100643.
55. Ichiye T., Karplus M. // *Proteins*. 1991. V. 11. № 3. P. 205–217.
56. Liu J., Nussinov R. // *Proc. Natl. Acad. Sci. USA*. 2008. V. 105. № 3. P. 901–906.
57. Dima R.I., Thirumalai D. // *Protein Sci.* 2006. V. 15. № 2. P. 258–268.
58. Reynolds K.A., McLaughlin R.N., Ranganathan R. // *Cell*. 2011. V. 147. № 7. P. 1564–1575.
59. Ferguson A.D., Amezcua C.A., Halabi N.M., Chelliah Y., Rosen M.K., Ranganathan R., Deisenhofer J. // *Proc. Natl. Acad. Sci. USA*. 2007. V. 104. № 2. P. 513–518.
60. Bornscheuer U.T., Huisman G.W., Kazlauskas R.J., Lutz S., Moore J.C., Robins K. // *Nature*. 2012. V. 485. № 7397. P. 185–194.
61. Imming P., Sinning C., Meyer A. // *Nat. Rev. Drug Discov.* 2006. V. 5. № 10. P. 821–834.
62. Overington J.P., Al-Lazikani B., Hopkins A.L. // *Nat. Rev. Drug Discov.* 2006. V. 5. № 12. P. 993–996.
63. Arnold F.H. // *Acc. Chem. Res.* 1998. V. 31. № 3. P. 125–131.
64. Reetz M.T. // *Enzyme Catalysis in Organic Synthesis*. / Third Ed. 2012. P. 119–190.
65. Reetz M.T., Carballeira J. D. // *Nat. Protocols*. 2007. V. 2. № 4. P. 891–903.
66. Kazlauskas R.J., Bornscheuer U.T. // *Nat. Chem. Biol.* 2009. V. 5. № 8. P. 526–529.
67. Knowles J., Gromo G. // *Nat. Rev. Drug Discov.* 2003. V. 2. № 1. P. 63–69.
68. Roses A.D., Burns D.K., Chissoe S., Middleton L., Jean P.S. // *Drug Discov. Today*. 2005. V. 10. № 3. P. 177–189.
69. Payne D.J., Gwynn M.N., Holmes D.J., Pompliano D.L. // *Nat. Rev. Drug Discov.* 2007. V. 6. № 1. P. 29–40.
70. Chan P.F., Holmes D.J., Payne D.J. // *Drug Discov. Today: Therapeutic Strategies*. 2004. V. 1. № 4. P. 519–527.
71. Reetz M.T., Bocola M., Carballeira J.D., Zha D., Vogel A. // *Angew. Chem. Int. Ed. Engl.* 2005. V. 44. № 27. P. 4192–4196.
72. Нилов Д.К., Прохорова Е.А., Швядас В.К. // *Acta Naturae*. 2015. T. 7. № 2. C. 62–68.
73. Kuravsky M.L., Barinova K.V., Asryants R.A., Schmalhauzen E.V., Muronetz V.I. // *Biochimie*. 2015. V. 115. P. 28–34.
74. Pargellis C., Tong L., Churchill L., Cirillo P.F., Gilmore T., Graham A.G., Grob P.M., Hickey E.R., Moss N., Pav S., et al. // *Nat. Struct. Mol. Biol.* 2002. V. 9. № 4. P. 268–272.
75. Esnouf R., Ren J., Ross C., Jones Y., Stammers D., Stuart D. // *Nat. Struct. Mol. Biol.* 1995. V. 2. № 4. P. 303–308.
76. Conn P.J., Christopoulos A., Lindsley C.W. // *Nat. Rev. Drug Discov.* 2009. V. 8. № 1. P. 41–54.
77. Vázquez-Figueroa E., Chaparro-Riggers J., Bommarius A.S. // *ChemBioChem*. 2007. V. 8. № 18. P. 2295–2301.
78. Jochens H., Aerts D., Bornscheuer U.T. // *Protein Eng. Des. Sel.* 2010. V. 23. № 12. P. 903–909.
79. Suplatov D.A., Besenmatter W., Švedas V.K., Svendsen A. // *Protein Eng. Des. Sel.* 2012. V. 25. № 11. P. 689–697.
80. Pleiss J. // *Curr. Opin. Biotechnol.* 2011. V. 22. № 5. P. 611–617.

81. Damborsky J., Brezovsky J. // *Curr. Opin. Chem. Biol.* 2014. V. 19. P. 8–16.
82. Suplatov D., Voevodin V., Švedas V. // *Biotechnology J.* 2015. V. 10. № 3. P. 344–355.
83. Cole S., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E., et al. // *Nature*. 1998. V. 393. № 6685. P. 537–544.
84. Galperin M.Y., Koonin E.V. // *Curr. Opin. Biotechnol.* 1999. V. 10. № 6. P. 571–578.
85. Moir D.T., Shaw K.J., Hare R.S., Vovis G.F. // *Antimicrob. Agents Chemother.* 1999. V. 43. № 3. P. 439–446.
86. Krissinel E., Henrick K. // *Acta Crystallogr. D Biol. Crystallogr.* 2004. V. 60. № 12. P. 2256–2268.
87. Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. // *Nucl. Acids Res.* 1997. V. 25. № 17. P. 3389–3402.
88. Fischer J.D., Mayer C.E., Söding J. // *Bioinformatics*. 2008. V. 24. № 5. P. 613–620.
89. Menke M., Berger B., Cowen L. // *PLoS Comput Biol.* 2008. V. 4. № 1. P. e10.
90. Notredame C., Higgins D.G., Heringa J. // *J. Mol. Biol.* 2000. V. 302. № 1. P. 205–217.
91. Suplatov D., Kirilin E., Takhaveev V., Švedas V. // *J. Biomol. Struct. Dyn.* 2014. V. 32. № 11. P. 1752–1758.