Обзорная статья

## МОЛЕКУЛЯРНЫЙ ДОКИНГ: РОЛЬ НЕВАЛЕНТНЫХ ВЗАИМОДЕЙСТВИЙ В ОБРАЗОВАНИИ КОМПЛЕКСОВ БЕЛКОВ С НУКЛЕОТИДАМИ И ПЕПТИДАМИ

© 2010 г. Т. В. Пырков<sup>\*, \*\*#</sup>, И. В. Озеров<sup>\*, \*\*\*</sup>, Е. Д. Балицкая<sup>\*, \*\*\*</sup>, Р. Г. Ефремов<sup>\*</sup>

\* Учреждение Российской академии наук Институт биоорганической химии им. акад. М.М. Шемякина и Ю.А. Овчинникова РАН, 117997 ГСП, Москва, В-437, ул. Миклухо-Маклая, 16/10;

\*\* Московский физико-технический институт (государственный университет), Москва;

\*\*\*Московский государственный университет им. М.В. Ломоносова, Москва

Поступила в редакцию 23.11.2009 г. Принята к печати 29.01.2010 г.

Знание пространственной структуры комплексов клеточных белков и мембранных рецепторов с лигандами является важным шагом на пути к пониманию механизмов их функционирования. Рациональный поиск и дизайн новых лекарственных соединений также нуждается в структурной информации о взаимодействиях прототипов лекарств с белком-мишенью. В данной работе предложен краткий обзор вычислительных методов молекулярного докинга (от англ. docking - стыковка), позволяющих предсказывать конформацию лиганда внутри активного центра белка. На примере недавних разработок, проводившихся в лаборатории моделирования биомолекулярных систем ИБХ РАН, продемонстрированы подходы к увеличению точности и эффективности современных алгоритмов докинга. Особое внимание уделено гидрофобным / гидрофильным и стэкингвзаимодействиям, ответственным за молекулярное узнавание определенных фрагментов лиганда. Эти типы контактов не всегда адекватно отражены в алгоритмах оценки энергии межмолекулярных взаимодействий существующих программ для докинга (оценочных функциях), что в конечном итоге приводит к ошибочным предсказаниям трехмерных структур комплексов. Именно поэтому всесторонний учет таких взаимодействий является одной из наиболее актуальных задач молекулярного моделирования.

**Ключевые слова:** оценочная функция докинга; водородные связи; гидрофобные взаимодействия; стэкинг.

## Введение

В настоящее время методы компьютерного молекулярного моделирования становятся неотъемлемой частью фундаментальных исследований, направленных на изучение молекулярных механизмов функционирования белков, а также и прикладных проектов, связанных с рациональным дизайном новых лекарственных соединений.

Метод молекулярного моделирования, целью которого является поиск наиболее достоверной ориентации и конформации лиганда в центре связывания белка-мишени, называется молекулярным докингом. Молекулярный докинг позволяет предсказывать

Сокращения: ОФ – оценочная функция; МГП – молекулярный гидрофобный потенциал; NAD – никотинамидадениндинуклеотид; FAD – флавинадениндинуклеотид; PDB – база данных трехмерных структур белков (Protein Data Bank).

<sup>&</sup>lt;sup>#</sup>Автор для связи (тел.: (495) 336-20-00; эл. почта: pyrkov@nmr.ru).

пространственную структуру комплекса рецептор–лиганд и свободную энергию его образования, исходя из данных о пространственной структуре рецептора, известной с разрешением в несколько ангстрем (например, полученной с помощью рентгеноструктурного анализа), и химической структуре лиганда. Достоинства этого метода очевидны: установление ключевых аминокислотных остатков в активном центре белка, позволяющее изучать структурно-динамические основы ферментативных реакций на атомном уровне; рациональный дизайн лигандов и/или рецепторов с заранее заданными селективностью, кинетическими свойствами и т.д. Докинг также используют в процессе виртуального высокопроизводительного скрининга (сканирования) баз данных (англ. virtual high throughput screening – vHTS), который значительно снижает затраты проектов, направленных на поиск новых эффективных и селективных лигандов. Вследствие этого, развитие и оптимизация алгоритмов докинга является на данный момент областью активных научных разработок.

В течение последних лет в литературе появилось большое число исследований, ориентированных на сравнение различных алгоритмов докинга и используемых ими оценочных функций между собой [1–10]. Такие исследования призваны помочь выбрать из многочисленных программ докинга наиболее подходящую для определенного класса соединений или для биомолекулярных систем со специфическими свойствами. Рассмотрим проблемы современных алгоритмов докинга и подходы, позволяющие повысить их точность.

В общем случае здесь можно выделить две независимые (хотя и взаимосвязанные) составляющие процедуры докинга – алгоритм конформационного поиска и оценочную функцию (ОФ).

## Алгоритм конформационного поиска

В классическом варианте молекулярного докинга задача алгоритма конформационного поиска сводится к перебору конформационного пространства комплекса за счет варьирования торсионных углов лиганда и его перемещения как целого относительно неподвижной структуры белка-мишени. Современные алгоритмы конформационного поиска в подавляющем большинстве случаев находят конформации, близкие к экспериментальным, за сравнительно короткое время. Тем не менее есть факторы, также влияющие на успешность докинга, которые часто не учитываются в стандартных алгоритмах. Один из таких факторов – конформационная подвижность белкамишени, в большинстве случаев сопровождающая связывание лиганда. Диапазон подвижности может быть разным – начиная с небольшой «подстройки» боковых цепей и заканчивая масштабными доменными движениями [11]. Эти движения играют большую роль: если взять структуру белка, оптимальную для связывания данного лиганда, то

результат докинга наверняка будет точнее, чем если взять любую другую (например, апоформу). На первый взгляд самым логичным решением этой проблемы является учет подвижности белка в программе докинга. К сожалению, современные вычислительные средства не позволяют проводить такое моделирование за приемлемое время, так как молекула белка очень велика, и учет подвижности по всем степеням свободы может привести к так называемому «комбинаторному взрыву» (астрономическому увеличению числа возможных вариантов). Лишь в некоторых программах предусмотрена ограниченная подвижность сайтов связывания белка (как правило, на уровне небольшой адаптации конформаций боковых цепей остатков активного центра). Другой подход к этой проблеме состоит в докинге в несколько *разных* конформаций одного и того же белка с последующим выбором лучших решений из каждого запуска докинга. Третий подход – найти некоторую универсальную структуру белка-мишени, с участием которой докинг давал бы достаточно хорошие результаты для различных классов лигандов. При этом уменьшается число «пропущенных» (но правильных) решений, однако сильно также возрастает и число неверных вариантов.

Более подробное описание упомянутых подходов к проблеме конформационной подвижности рецептора можно найти в обзорах [12, 13]. В дальнейшем под молекулярным докингом мы будем подразумевать его классический вариант с использованием неподвижной структуры белка мишени.

## Оценочная функция

Оценочные функции (ОФ), используемые в процессе докинга, служат для вычисления примерной энергии комплексов и ранжирования различных предполагаемых конформаций лиганда в сайте связывания на каждом шаге конформационного поиска (рис. 1). Но ОФ не всегда хорошо справляются с этой задачей, внося ошибки при моделировании структуры комплекса рецептор–лиганд там, где они должны указывать верное решение. Заранее предполагая неточность в ОФ, обычно рассматривают не единственную структуру комплекса, а целый набор возможных вариантов, рассчитывая на то, что среди них окажется правильный или хотя бы близкий к нему. Из этого набора можно выбрать наиболее правдоподобный вариант, основываясь, например, на известных экспериментальных данных о роли тех или иных аминокислотных остатков активного центра рассматриваемого белка в связывании лигандов. Фактически, в данном случае используется дополнительный фильтр, оптимизирующий положение лиганда для данного конкретного белка-мишени [11–20].

К решению проблемы ранжирования результатов докинга можно подойти и с другой стороны – использовать лигандспецифичные ОФ, повышающие вероятность найти

правдоподобное решение в случае достаточно узкого класса химических веществ, – например, нуклеотидов, пептидов и др. Такие ОФ учитывают важные для распознавания этих соединений взаимодействия с рецептором, – например, стэкинг в случае нуклеотидов [21, 22], водородные связи пептидов с основной цепью белка-мишени [23] и другие типы контактов. В этом обзоре будут изложены результаты исследований, проводимых в лаборатории моделирования биомолекулярных систем ИБХ РАН, направленных на разработку лигандспецифичных ОФ. При этом, кроме указанных типов межмолекулярных взаимодействий, особое внимание будет также уделено исследованию роли гидрофобных / гидрофильных контактов, которые важны при распознавании белком амфифильных лигандов.

ОФ показывает, какая из ориентаций лиганда в сайте связывания рецептора наиболее правдоподобна или, если сравниваются несколько разных лигандов, – какой из них обладает наибольшим сродством к белку-мишени. Обычно энергию связывания лиганда с рецептором раскладывают на отдельные независимые слагаемые – термы, отражающие различные физические взаимодействия. Линейная комбинация этих термов и представляет собой ОФ. Все ОФ, применяемые в современных алгоритмах докинга, можно условно разделить на три типа: 1) основанные на силовых полях, 2) эмпирические и 3) статистические ОФ. Подробное описание этих типов ОФ можно найти в обзорах [24–27]; здесь же мы лишь кратко остановимся на их основных особенностях.

Оценка энергии взаимодействия рецептор–лиганд на основе силового поля кажется наиболее естественным подходом, поскольку термы силового поля напрямую соответствуют тем или иным взаимодействиям, определяющим молекулярное распознавание. Примером ОФ, основанной на эмпирическом силовом поле, может служить функция, включенная в широко известный пакет для докинга DOCK [28]. Ввиду того, что лиганды, как правило, ковалентно не связаны с рецептором, ОФ на основе молекулярно-механических силовых полей включают термы, характеризующие невалентные взаимодействия: в основном, электростатические и Ван-дер-Ваальсовы силы. Главным ограничением ОФ на основе силовых полей является то, что они рассчитаны на оптимизацию структуры молекулы, учитывая изменение лишь энтальпийной составляющей энергии взаимодействия. Между тем, связывание лиганда с рецептором сопровождается эффектом десольватации, а также изменением энтропии [25], не учитываемыми в расчетах молекулярной механики. Эта особенность существенно ограничивает применение такого типа ОФ в молекулярном докинге.

С этой точки зрения существенным преимуществом обладают так называемые «эмпирические» ОФ. В отличие от ОФ, основанных на силовых полях, они включают термы, описывающие межмолекулярные контакты зачастую более примитивным

способом, – без проведения прямых аналогий с парными межмолекулярными физическими взаимодействиями. Предсказательная способность ОФ зависит не только от конкретного вида термов, описывающих те или иные взаимодействия, но и от весовых коэффициентов при термах, определяемых исходя из параметризации с использованием обучающих наборов экспериментальных данных о структурах комплексов, – как в функции Chemscore, разработанной в 1997 г. [289] (уравнение 1).

В данном случае межмолекулярные взаимодействия представлены в виде линейной комбинации термов, описывающих различные виды контактов: водородные связи ( $C_{hb}$ ), гидрофобные взаимодействия ( $C_{lip}$ ), взаимодействия с ионами металлов ( $C_{met}$ ); кроме того, учитывается число «замороженных» при связывании торсионных углов  $H_{rotb}$ , отражающих изменение энтропии ( $C_{rot}$ ). Благодаря упрощенному виду термов можно существенно сократить расчеты и сэкономить вычислительные ресурсы. Например, координационные связи с ионами металлов  $M(r_{lr})$  или гидрофобные контакты  $L(r_{lr})$  могут быть описаны с помощью расстояний  $r_{lr}$  между соответствующими атомами лиганда и рецептора, хотя такое приближение и не является физически корректным. Водородные связи описываются эмпирическими геометрическими параметрами (расстояние между донором и акцептором  $r_{AD}$  и угол между ними и атомом водорода  $\alpha_{ADH}$ ), а не их энергетическими характеристиками.

Эффективность эмпирических ОФ сильно зависит от используемых при их параметризации обучающих наборов. Последнее, конечно, накладывает определенные ограничения на область применения этих ОФ – надежность получаемых результатов не может быть гарантирована в том случае, если молекулы белка-мишени и/или лиганда существенно отличаются по характеру межмолекулярных взаимодействий от обучающей выборки. Например, ОФ, которая наиболее эффективна при рассмотрении комплексов, где доминируют полярные контакты (водородные связи, солевые мостики и т.д.), может выдавать ошибочные результаты в случаях, когда гидрофобные контакты играют ключевую роль в распознавании лиганда рецептором, и наоборот.

# $\Delta G_{bind} = C_0 + C_{lip} \sum L(r_{lr}) + C_{hb} \sum H(r_{AD}, \alpha_{ADH}) + C_{met} \sum M(r_{lm}) + C_{rot} N_{rot}.$ (1)

Третий тип ОФ – статистические, основанные на кривых радиального распределения атомов в структурах комплексов лиганд–рецептор, полученных экспериментально. В дальнейшем эти кривые могут быть преобразованы в статистические потенциалы, которые, однако, в большей степени предназначены для предсказания ориентации лиганда в активном сайте, чем для оценки свободной энергии связывания. Наиболее известные из таких ОФ – PMF [30] и ASP [31], где парные потенциалы параметризованы для каждого

типа атомов и определяют среднюю объемную плотность атомов одного типа и их радиальную функцию распределения относительно каждого атома другого типа.

Статистические ОФ неявным образом характеризуют те особенности взаимодействия лиганда с рецептором, которые сложно описать в явном виде, – например, взаимодействия ароматических групп. В то же время данный подход сильно зависит от разделения атомов на типы (как правило, число типов атомов в таких схемах существенно превышает число химических элементов) и, как и эмпирические ОФ, от состава обучающей выборки. Кроме того, интерпретация статистических ОФ в общепринятых термах различных взаимодействий (электростатических, гидрофобных и т.д.) является затруднительной.

В последнее время эмпирические ОФ нашли широкое применение в алгоритмах докинга. Популярность использования эмпирических ОФ обусловлена тем, что с накоплением все большего количества экспериментальных данных о строении и биохимии молекулярных комплексов для них становится достаточно просто составлять обучающую выборку и конструировать системоспецифичные или настраиваемые ОФ. При создании системоспецифических ОФ возможны два подхода. Первый заключается в использовании экспериментальных данных о связывании различных лигандов с изучаемым рецептором [16, 32, 33]. Полученная белок-специфичная ОФ, учитывающая структурнофункциональные особенности сайта, позволяет эффективно проводить поиск новых лигандов для этого белка по базам данных низкомолекулярных соединений. Второй подход основан на создании лигандспецифичных критериев. Он менее распространен, так как требует большого количества экспериментальных данных о взаимодействиях отдельного класса лигандов с различными белками. Тем не менее на данный момент уже существуют несколько успешных примеров создания таких специфичных ОФ для пептидов [23], углеводов [34, 35], АТР [36, 37].

Одной из важных проблем, возникающих при выборе наиболее подходящих критериев оценки для исследования конкретной системы белок–лиганд, является учет различных (наиболее значимых) межмолекулярных взаимодействий в рассматриваемом комплексе. Остановимся подробнее на методах учета некоторых из этих взаимодействий.

## Водородные связи

Водородные связи часто играют определяющую роль в формировании структуры биологических макромолекул – белков и нуклеиновых кислот, – а также в образовании их комплексов с низкомолекулярными веществами. В эмпирических ОФ наличие в молекуле водородных связей определяется по простым геометрическим критериям – расстоянию между донором и акцептором и углу между ними и атомом водорода.

Нередко для определенного типа лигандов характерно образование определенных мотивов (комбинаций) водородных связей с белком. Эти комбинации водородных связей вносят свой вклад в специфическое узнавание лигандов в активном центре белка-мишени. Поиск и учет таких мотивов в ОФ может существенно улучшить предсказательную способность докинга на том спектре мишеней / лигандов, для которых характерен этот мотив. С этой целью представляют интерес исследования связывания пептидных и нуклеотидных лигандов с белками.

В частности, на основе анализа экспериментально установленных пространственных структур комплексов белок–пептид из баз данных Brookhaven Protein Data Bank (PDB) [38] и PDBbind [39] было показано, что атомы азота и кислорода основной цепи пептидных лигандов статистически достоверно образуют водородные связи преимущественно с главной цепью белка-рецептора и в меньшей степени – с атомами боковых групп аминокислотных остатков (рис. 2). Объясняется это тем, что часто в таких белках участок β-тяжа экспонирован в активный центр и предоставляет пептидному лиганду или его аналогу возможность образовать большое число именно таких связей, также приняв вытянутую форму β-тяжа, – как, например, в матриксных металлопротеиназах [40].

Для нуклеотидных лигандов, т.е. таких, в состав которых входят азотистые основания (ATP, ADP, GTP, NAD, FAD и др.), также характерны определенные мотивы водородных связей с сайтом связывания в белке-рецепторе. Причем в данном случае они во многом напоминают сетку водородных связей в парах нуклеотидов в двойной спирали ДНК (рис. 3) [41]. Мотивы водородных связей могут быть учтены в ОФ в виде суммы вкладов отдельных связей, составляющих мотив.

## Гидрофобные взаимодействия

Неполярные молекулы обладают свойством объединяться друг с другом в полярном растворителе, стремясь, таким образом, минимизировать площадь контакта с ним. На макроскопическом уровне это наблюдается в виде так называемых гидрофобных (липофильных) взаимодействий, которые играют важнейшую роль в формировании пространственной структуры биологических макромолекул и целых систем, таких, как, например, липидная бислойная мембрана [42]. Также, в ряде случаев они ответственны и за специфичное узнавание рецептор–лиганд [43, 44].

Несмотря на широкую распространенность и фундаментальное значение для структуры биомакромолекул и их комплексов, природа физических сил, лежащих в основе гидрофобного эффекта, не определена достаточно точно для их количественного описания. Это объясняется тем, что гидрофобный эффект складывается из множества разных межмолекулярных взаимодействий: электростатического, ван-дер-ваальсова, энтропийного

эффекта и др. В связи с этим до сих пор не предложено универсального метода численной оценки свободной энергии гидрофобных взаимодействий.

Тем не менее осознание ключевой роли гидрофобных взаимодействий в организации биомолекулярных систем дало толчок развитию эмпирических подходов к оценке гидрофобности. Самый простой способ, используемый, в частности, в функции Chemscore [28], – разделить все атомы на гидрофобные (углерод, сера, галогены) и гидрофильные (остальные атомы). Число контактов между гидрофобными атомами может служить приблизительной оценкой величины гидрофобных взаимодействий лиганда с рецептором. Однако этот подход имеет серьезные ограничения и с его помощью не всегда можно добиться корректного предсказания. Так, гетероциклический фрагмент в составе лиганда (например, аденин в АТР) будет идентифицирован как гидрофильный, хотя известно, что аденин, входящий в состав АТР, предпочтительно связывается в гидрофобных карманах активных сайтов [21, 22, 45–47] и, следовательно, обладает гидрофобными свойствами.

Более точные методы, учитывающие при расчете свойств отдельных частей молекулы не только тип химического элемента, но и влияние окружения, основаны на экспериментальных данных по распределению органических соединений между полярной и неполярной фазами (обычно вода/октанол). Логарифм коэффициента распределения вещества между двумя средами – *logP* – традиционно служит количественной мерой его гидрофобности. Для составления ОФ часто используется линейное приближение, согласно которому молекулярный коэффициент *logP* может быть представлен в виде линейной комбинации атомных констант гидрофобности [48, 49]. В системе параметризации [49] каждый атом молекулы относится к одному из 120 топологических типов, включающих явно заданные атомы водорода. Также известны другие системы параметризации – без атомов водорода [50], фрагментальные [51, 52] и др.

Эмпирические атомные константы гидрофобности используют при изучении пространственного распределения гидрофобных/гидрофильных свойств, например, на поверхности молекулы. Согласно концепции молекулярного гидрофобного потенциала (МГП), гидрофобные свойства рассчитывают на поверхности молекулы или в любой другой точке пространства, используя функцию, зависящую от расстояния (рис. 4, уравнение 2):

$$\mathbf{M}\Gamma\Pi_{j} = \sum_{i} f_{i} \times g(r_{ij}) \qquad (2)$$

где МГП<sub>*j*</sub> – величина МГП в точке пространства *j*,  $f_i$  – атомная константа гидрофобности,  $g(r_{ij})$  – функция затухания, зависящая от расстояния между атомом *i* и точкой *j*. Необходимо отметить, что, как уже было сказано выше, строгого описания гидрофобных сил не существует, поэтому используемые функции  $g(r_{ij})$  носят, по сути, интуитивный и

эмпирический характер. При этом в некоторых работах было показано, что одни из них лучше описывают свойства низкомолекулярных соединений, а другие – макромолекул, таких, как белки [53].

Для характеристики вклада гидрофобных взаимодействий при образовании комплексов лиганд–рецептор предложены различные критерии, оценивающие соответствие их гидрофобных/гидрофильных свойств. Существуют методы, основанные на сравнении величины и знака констант гидрофобности контактирующих атомов [54–56] или значений МГП двух молекул на поверхности интерфейса [44]. Другие методы оценивают гидрофобные взаимодействия просто по числу контактов между гидрофобными атомами или химическими группами [28], либо по доле гидрофобной поверхности, экранированной от растворителя [57], либо по комплементарности гидрофобных областей на поверхности двух молекул (образованию энергетически выгодного контакта между этими областями) [58]. Для количественной оценки комплементарности гидрофобных свойств систем типа ATP-белок мы использовали параметр, учитывающий долю заглубленной гидрофобной поверхности лиганда [36].

Кроме того, мы показали, что небольшое смещение МГП молекулы лиганда в гидрофобную область (+0.03 в единицах *logP*) дает более правдоподобную картину распределения свойств для низкомолекулярных соединений, особенно для нуклеозидов и нуклеотидов. Возможность варьировать смещение шкалы атомных констант гидрофобности не реализована в программных пакетах для молекулярного моделирования, поэтому мы создали для этого собственное приложение, доступное на веб-сервере PLATINUM (http://model.nmr.ru/platinum/) [59].

Другим примером успешного применения разработанного нами подхода является докинг лигандов циклооксигеназы-1 и 2 и тромбина. В этих случаях также наблюдается хорошая комплементарность гидрофобных свойств лиганда и активного центра белка [60]. Кроме того, гидрофобный контакт играет определяющую роль при связывании антагонистов β<sub>2</sub>-адренорецептора в его активном сайте [61–63]. Использование ОФ, учитывающей МГП, позволяет эффективно предсказывать энергию связывания агонистов и антагонистов этого рецептора [64].

## Стэкинг

Среди различных типов межмолекулярных контактов особого внимания заслуживает стэкинг ароматических колец, также играющий важную роль в молекулярном узнавании. Хотя многие лиганды, в том числе и лекарственные соединения, содержат ароматические фрагменты и группы, стэкинг часто не учитывается явно при составлении ОФ.

Стэкинг-взаимодействия наблюдаются между двумя ароматическими группами, в результате чего они принимают определенную ориентацию друг относительно друга в

пространстве. Самый известный пример – двойная спираль ДНК, где азотистые основания в результате стэкинга располагаются параллельно друг другу [65, 66]. Также возможно и перпендикулярное («Т-образное») взаимное расположение ароматических колец, что показано, например, для бензола [67]. Кроме того, ароматические соединения имеют тенденцию участвовать в *π*-катионном взаимодействии, при котором образуется контакт между положительно заряженными группами и электронным облаком кольца [68–70].

Ароматические контакты, так же как и водородные связи, могут быть описаны с помощью геометрических критериев. В наших исследованиях взаимодействий АТР с различными белками было показано, что аденину свойственно образование стэкинга с боковыми цепями аминокислотных остатков белка, преимущественно с фенилаланином [36], что согласуется и с результатами других работ [21, 22, 45, 46, 71, 72]. Для того чтобы выявить параметры контактов, образованных в результате стэкинга, мы предлагаем анализировать взаимную ориентацию двух ароматических фрагментов соединений в терминах расстояния между центрами циклов и угла между их плоскостями (рис. 5).

Диапазон значений параметров, определяющих наличие или отсутствие стэкинга в расчетных алгоритмах, до сих пор остается не выясненным и в оценочных критериях выбирается достаточно произвольно [73, 74]. Картину усугубляет еще и тот факт, что многие ароматические соединения стремятся расположиться не только параллельно, но и перпендикулярно друг к другу, как это было показано для аминокислот в белках [74, 75] и в моделях, состоящих из простых углеводородов, – бензола, нафталина и т.д. [16, 76–79]. Уточнение взаимного расположения ароматических колец для конкретных систем повышает эффективность оценки качества и достоверности структур комплексов белоклиганд, предсказываемых методами молекулярного моделирования. Например, в результате анализа пространственных структур комплексов различных белков с лигандами, содержащими в своей структуре гуанин, было установлено, что для таких лигандов характерно образование «параллельного» стэкинга с расстоянием от плоскости кольца гуанина до центра ароматического фрагмента аминокислотного остатка (h), равным 3-4 Å, и иногда – «Т-образных» контактов с h = 4.5 - 5.5 Å, как в случае с тирозином (рис. 6) [41]. Стоит заметить, что гуанидиновая группа аргинина, так же как и ароматические аминокислоты, склонна образовывать параллельные контакты с аденином.

## Консенсусный подход

Результатом докинга, как правило, является не одна структура комплекса белоклиганд, а целый набор наиболее вероятных (с лучшими значениями ОФ) ориентаций лиганда в сайте связывания. Поэтому даже в случае, если недостаточная точность используемой в процессе докинга ОФ не позволяет выбрать из полученного набора «правильный» вариант, соответствующий нативной ориентации лиганда, всегда есть

возможность переранжировать этот набор по более эффективному критерию. Такой метод получил название консенсусного докинга. Надо отметить, что многие программные пакеты для молекулярного докинга используют несколько различных ОФ, что существенно облегчает реализацию такого подхода на практике. Кроме того, было замечено, что среди различных ОФ одни более эффективны в процессе поиска возможных конформаций и ориентаций лиганда в сайте, а другие – в финальном ранжировании этих вариантов по значениям энергии взаимодействия белок–лиганд, как это, например, реализовано в программе LeadFinder [80].

В качестве примера консенсусного докинга может служить исследование, посвященное сравнению двух ОФ, реализованных в программе докинга GOLD. Одна из них, Goldscore [81], использует термы силовых полей – ван-дер-ваальсовы взаимодействия и водородные связи. Другая – эмпирическая ОФ Chemscore [28] – учитывает водородные связи, гидрофобные взаимодействия и координационные связи с ионами металлов. Различные ориентации лиганда в сайте связывания, сгенерированные с помощью Goldscore, затем были отранжированы по значениям Chemscore, что существенно повысило долю верно предсказанных структур комплексов белок–лиганд. Обратная операция – докинг с применением Chemscore и последующее ранжирование решений докинга по Goldscore – также позволила улучшить результаты по сравнению с применением каждой из ОФ в отдельности. В литературе встречается также и другой термин – консенсусное ранжирование. Он означает несколько иной подход: здесь ОФ комбинируют не за счет последовательного их применения, а производят ранжирование результатов докинга по критерию, представляющему собой взвешенную комбинацию различных ОФ [1–8, 82].

Системоспецифичные (настраиваемые) ОФ также могут быть использованы в «консенсусном докинге». Входящие в уравнение весовые коэффициенты различных термов взаимодействия могут быть оптимизированы для частных случаев, – например, для конкретного белка-мишени или определенного класса лигандов [11–20]. Так, например, известно, что необходимым условием эффективного ингибирования тромбина является водородная связь ингибитора с боковой цепью остатка Asp189 в активном сайте тромбина [83], а для циклинзависимой киназы 2 ингибиторы должны образовывать водородную связь с атомом азота основной цепи остатка Leu83 [11]. Такого рода данные могут быть использованы и в стандартных ОФ в качестве дополнительных фильтров или добавочных термов наряду с уже существующими.

Основываясь на положительных результатах разработки ATP-специфичной ОФ, мы поставили задачу найти универсальную лигандспецифичную функцию для более широкого класса лигандов. Используя простую модель, представляющую собой комбинацию всего двух термов, описывающих водородные связи и гидрофобные взаимодействия, мы выявили

некоторые общие закономерности, отражающие характерные межмолекулярные контакты различных типов лигандов с белками-рецепторами. Так, для средних по размеру (10–25 тяжелых атомов) или амфифильных соединений гидрофобные контакты играют бо́льшую роль, чем для маленьких, или, наоборот, более крупных лигандов, а также пептидов [60].

С целью реализации данного подхода на практике нами была создана программа PLATINUM (Protein-Ligand ATtractions Investigation NUMerically), доступная в виде вебсервиса на сайте лаборатории моделирования биомолекулярных систем ИБХ РАН: http://model.nmr.ru/platinum/ [59]. PLATINUM позволяет количественно оценивать распределение гидрофобных/гидрофильных свойств молекул, учитывать гидрофобные и стэкинг-взаимодействия, водородные связи. Также веб-сервер PLATINUM позволяет анализировать взаимодействия и ранжировать структуры комплексов белок–лиганд, полученные с помощью докинга, и может быть использован в консенсусном подходе при ранжировании результатов молекулярного докинга (рис. 7).

### Заключение и перспективы

В данном обзоре освещены основные недостатки современных методов молекулярного докинга и возможные пути их преодоления путем оптимизации функций оценки энергии межмолекулярных взаимодействий, применяемых в процессе докинга. Особое внимание уделено методам учета гидрофобных взаимодействий, а также стэкинга и водородных связей, - в частности, определенных мотивов таких связей, ответственных за специфичное узнавание пептидных и нуклеотидных лигандов их рецепторами. Все представленные методы проиллюстрированы конкретными примерами их применения. Рассмотренный «консенсусный» подход позволяет эффективно комбинировать описанные здесь новые критерии с уже существующими для повышения точности теоретического предсказания пространственной структуры комплексов белок-лиганд. По мере накопления новых экспериментальных данных о межмолекулярных взаимодействиях в таких комплексах этот подход будет получать все большее распространение благодаря развитию системоспецифичных критериев. Созданный с этой целью веб-сервер PLATINUM в будущем будет дополнен возможностью автоматической настройки оценочных функций, а также расширен за счет готовых оценочных критериев для некоторых наиболее широко распространенных классов лигандов, – таких, как нуклеотиды и пептиды.

Помимо оценочных функций на результаты докинга может также влиять конформационная подвижность белка-рецептора, которая не учитывается в стандартных алгоритмах. Дело в том, что при рассмотрении большого числа таких связей, в белкерецепторе возникает проблема «комбинаторного взрыва». Большинство современных методов докинга достигли высокой эффективности в воспроизведении ориентации и конформации лиганда в сайте связывания, откуда он предварительно был удален. В то же

время допущение о конформационной неподвижности молекулы рецептора в процессе связывания верно лишь для части исследуемых белков – по некоторым оценкам, примерно 50% [11]. При этом специальные исследования свидетельствуют о необходимости настройки оценочных функций для модели «подвижного рецептора» [84, 85].

## БЛАГОДАРНОСТИ

Работа выполнена при финансовой поддержке РФФИ (гранты № 07-04-01514-а, 09-04-13813-офи\_ц), программ РАН «Молекулярная и клеточная биология» и «Основы фундаментальных исследований нанотехнологий и наноматериалов» и Федерального агентства по науке и инновациям в поддержку ведущих научных школ (грант НШ-4728.2006.4).

СПИСОК ЛИТЕРАТУРЫ

- Bursulaya B.D., Totrov M., Abagyan R., Brooks C.L. // J. Comput. Aided Mol. Des. 2003.
   V. 17. P. 755–763.
- 2. Schulz-Gasch T., Stahl M. // J. Mol. Model. 2003. V. 9. P. 47-57.
- Ferrara P., Gohlke H., Price D.J., Klebe G., Brooks C.L. // J. Med. Chem. 2004. V. 47.
   P. 3032–3047.
- 4. *Kellenberger E., Rodrigo J., Muller P., Rognan D.* // PROTEINS. 2004. V. 57. P. 225–242.
- 5. Kontoyianni M., McClellan L.M., Sokol G.S. // J. Med. Chem. 2004. V. 47. P. 558–565.
- 6. Perola E., Walters W.P., Charifson P.S. // PROTEINS. 2004. V. 56. P. 235-249.
- Wang R., Lu Y., Fang X., Wang S. // J. Chem. Inf. Comput. Sci. 2004. V. 44. P. 2114– 2125.
- Zhou Z., Felts A.K., Friesner R.A., Levy R.M. // J. Chem. Inf. Model. 2007. V. 47. P. 1599– 1608.
- Cross J.B., Thompson D.C., Rai B.K., Baber J.C., Fan K.Y., Hu Y., Humblet C. // J. Chem. Inf. Model. 2009. V. 49. P. 1455–1474.
- 10. Cheng T., Li X., Li Y. Liu Z., Wang R. // J. Chem. Inf. Model. 2009. V. 49. P. 1079–1093.
- 11. Betts M.J., Sternberg M.J. // Protein Eng. 1999. V. 12. P. 271-283.
- 12. Zhong H., Tran L.M., Stang J.L. // J. Mol. Graph. Model. 2009. V. 28. P. 558-575.
- 13. Kokh D.B., Wenzel W. // J. Med. Chem. 2008. V. 51. P. 5919–5931.
- Ferrara P., Curioni A., Vangrevelinghe E., Meyer T., Mordasini T., Andreoni W., Acklin P., Jacoby E. // J. Chem. Inf. Model. 2006. V. 46. P. 254–263.
- 15. Thomas M.P., McInnes C., Fischer P.M. // J. Med. Chem. 2006. V. 49. P. 92–104.
- 16. Jansen J.M., Martin E.J. // Curr. Opin. Chem. Biol. 2004. V. 8. P. 359-364.
- 17. Morley D.S., Afshar M. // J. Comput. Aided Mol. Des. 2004. V. 18. P. 189–208.
- Park M.S., Dessal A.L., Smrcka A.V., Stern H.A. // J. Chem. Inf. Model. 2009. V. 49. P. 437–443.
- Hevener K.E., Zhao W., Ball D.M., Babaoglu K., Qi J., White S.W., Lee R.E. // J. Chem. Inf. Model. 2009. V. 49. P. 444–460.
- 20. Deng W., Verlinde C.L.M.J. // J. Chem. Inf. Model. 2008. V. 48. P. 2010–2020.
- 21. Mao L., Wang Y., Liu Y., Hu X. // J. Mol. Biol. 2004. V. 336. P. 787-807.
- 22. Saito M., Go M., Shirai T. // Protein Eng. Des. Sel. 2006. V. 19. P. 67-75.
- 23. Rognan D., Lauemoller S.L., Holm A., Buus S., Tschinke V. // J. Med. Chem. 1999. V. 42.
  P. 4650–4658.

- 24. Kitchen D.B., Decornez H., Furr J.R., Bajorath J. // Nat. Rev. Drug. Discov. 2004. V. 3. P. 935–949.
- 25. Moitessier N., Englebienne P., Lee D., Lawandi J., Corbeil C.R. // Br. J. Pharmacol. 2008.
  V. 153. P. S7–S26.
- 26. Tuccinardi T. // Comb. Chem. High Throughput Screen. 2009. V. 12. P. 303-314.
- 27. Zoete V., Grosdidier A., Michielin O. // J. Cell. Mol. Med. 2009. V. 13. P. 238-248.
- 28. *Ewing T.J., Makino S., Skillman G.A.* // J. Comput. Aided Mol. Des. 2001. V. 15. P. 411–428.
- 29. Eldridge M.D., Murray C.W., Auton T.R., Paolini G.V., Mee R.P. // J. Comput. Aided Mol. Des. 1997. V. 11. P. 425–445.
- 30. Muegge I. // J. Med. Chem. 2006. V. 49. P. 5895–5902.
- 31. Mooij W.T.M., Verdonk M.L. // PROTEINS. 2005. V. 61. P. 272–287.
- 32. Amini A., Shrimpton P.J., Muggleton S.H., Sternberg M.J. // PROTEINS. 2007. V. 69. P. 823–831.
- 33. Catana C., Stouten P.F.W. // J. Chem. Inf. Model. 2007. V. 47. P. 85-91.
- 34. Laederach A., Reilly P.J. // PROTEINS. 2005. V. 60. P. 591–597.
- Kerzmann A., Fuhrmann J., Kohlbacher O., Neumann D. // J. Chem. Inf. Model. 2008. V.
   48. P. 1616–1625.
- 36. Pyrkov T.V., Kosinsky Y.A., Arseniev A.S., Priestle J.P., Jacoby E., Efremov R.G. // PROTEINS. 2007. V. 66. P. 388–398.
- 37. Pyrkov T.V., Efremov R.G. // Int. J. Mol. Sci. 2007. V. 8. P. 1083–1094.
- Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., Bourne P.E. // Nucleic Acid. Res. 2000. V. 28. P. 235–242.
- 39. Wang R., Fang X., Lu Y., Yang C.Y., Wang S. // J. Med. Chem. 2005. V. 48. P. 4111-4119.
- 40. Whittaker M., Floyd C.D., Brown P., Gearing A.J.H. // Chem. Rev. 1999. V. 99. P. 2735–2776.
- 41. Pyrkov T.V., Pyrkova D.V., Balitskaya E.D., Efremov R.G. // Acta Naturae. 2009. V. 1. P. 124–127.
- 42. Efremov R.G., Chugunov A.O., Pyrkov T.V., Priestle J.P., Arseniev A.S., Jacoby E. // Curr. Med. Chem. 2007. V. 14. P. 393–415.
- 43. *Gaillard P., Carrupt P.A., Testa B., Boudon A.* // J. Comput. Aided. Mol. Des. 1994. V. 8. P. 83–96.
- 44. Testa B., Carrupt P.A., Gaillard P., Billois F., Weber P. // Pharm. Res. 1996. V. 13. P. 335–343.
- 45. Denissiouk K.A., Rantanen V.V., Johnson M.S. // PROTEINS. 2001. V. 44. P. 282–291.
- 46. Cappello V., Tramontano A., Kock U. // PROTEINS. 2002. V. 47. P. 106-115.

- 47. Kuttner Y.Y., Sobolev V., Raskind A., Edelman M. // PROTEINS. 2003. V. 52. P. 400-411.
- 48. Viswanadhan V.N., Ghose A.K., Revankar G.R., Robins R.K. // J. Chem. Inf. Comput. Sci. 1989. V. 29. P. 163–172.
- 49. *Ghose A.K., Viswanadhan V.N., Wendoloski J.J.* // J. Phys. Chem. 1998. V. 102. P. 3762–3772.
- 50. Wang R., Gao I., Lay L. // Perspect. Drug Discov. Des. 2000. V. 19. P. 47-66.
- 51. *Klopman G., Li J.-Y., Wang S., Dimayuga M.* // J. Chem. Inf. Comput. Sci. 1994. V. 34. P. 752–781.
- 52. Meylan W.M., Howard P.H. // J. Pharm. Sci. 1995. V. 84. P. 83-92.
- 53. Heiden W., Moeckel G., Brickmann J. // J. Comput. Aided. Mol. Des. 1993. V. 7. P. 503– 514.
- 54. Wang R., Luhua L., Shaomeng W. // J. Comp. Aided. Mol. Des. 2002. V. 16. P. 11-26.
- 55. Meng E.C., Kuntz I.D., Abraham D.J., Kellogg G.E. // J. Comput. Aided. Mol. Des. 1994.
  V. 8. P. 299–306.
- 56. Efremov R.G., Alix A.J.P. // J. Biomol. Struct. Dyn. 1993. V. 11. P. 483-507.
- 57. Bohm H.J. // J. Comput. Aided Mol. Des. 1994. V. 8. P. 243-256.
- 58. Exner T.E., Keil M., Brickmann J. // J. Comput. Chem. 2002. V. 23. P. 1176–1187.
- Pyrkov T.V., Chugunov A.O., Krylov N.A., Nolde D.E., Efremov R.G. // Bioinformatics.
   2009. V. 25. P. 1201–1202.
- 60. *Pyrkov T.V., Priestle J.P., Jacoby E., Efremov R.G. //* SAR QSAR Environ. Res. 2008. V. 19. P. 91–99.
- 61. Bree F., Tayar N., Van de Waterbeemd H., Testa B., Tillement J.P. // J. Recept. Res. 1986.
  V. 6. P. 381–409.
- 62. Contreras M.L., Wolfe B.B., Molinoff P.B. // J. Pharmacol. Exp. Ther. 1986. V. 237. P. 154–164.
- 63. Contreras M.L., Wolfe B.B., Molinoff P.B. // J. Pharmacol. Exp. Ther. 1986. V. 237. P. 165–172.
- 64. Novoseletsky V.N., Pyrkov T.V., Efremov R.G. // SAR QSAR Environ. Res. 2010. (в печати)
- 65. Gotoh O. // Adv. Biophys. 1983. V. 16. P. 1-52.
- 66. Sponer J., Leszczynski J., Hobza P. // J. Biomol. Struct. Dynam. 1996. V. 14. P. 117–135.
- 67. Jorgensen W.L., Severance D.L. // J. Am. Chem. Soc. 1990. V. 112. P. 4768-4774.
- 68. Waters M.L. // Curr. Opin. Chem. Biol. 2002. V. 6. P. 736-741.
- 69. Meyer E.A., Castellano R.K., Diederich F. // Angew. Chem. Int. Ed. 2003. V. 42. P. 1210– 1250.
- 70. Tewari A.K., Dubey R. // Bioorg. Med. Chem. 2008. V. 16. P. 126-143.

- 71. Traxler P., Furet P. // Pharmacol. Ther. 1999. V. 82. P. 195–206.
- 72. Sigel H., Griesser R. // Chem. Soc. Rev. 2005. V. 34. P. 875–900.
- 73. Deng Z., Chuaqui C., Singh J. // J. Med. Chem. 2004. V. 47. P. 337–344.
- 74. Chakrabarti P., Bhattacharyya R. // Prog. Biophys. Mol. Biol. 2007. V. 95. P. 83-137.
- Chelli R., Gervasio F.L., Procacci P., Schettino V. // J. Am. Chem. Soc. 2002. V. 124. P.
   6133–6143.
- 76. *Tsuzuki S., Honda K., Uchimaru T., Mikami M., Tanabe K. //* J. Am. Chem. Soc. 2001. V. 124. P. 104–112.
- 77. Small D., Zaitsev V., Jung Y., Rosokha S.V., Head-Gordon M., Kochi J.K. // J. Am. Chem. Soc. 2004. V. 126. P. 13850–13858.
- 78. Sato T., Tsuneda T., Hirao K. // J. Chem. Phys. 2005. V. 123. P. 104307.
- 79. Malathy Sony S.M., Ponnuswamy M.N. // Crystal Growth & Design. 2006. V. 6. P. 736– 742.
- Stroganov O.V., Novikov F.N., Stroylov V.S., Kulkov V., Chilov G.G. // J. Chem. Inf. Model. 2008. V. 48. P. 2371–2385.
- 81. Jones G., Willett P., Glen R.C., Leach A.R., Taylor R.D. // J. Mol. Biol. 1997. V. 267. P. 727–748.
- Renner S., Derksen S., Radestock S., Morchen F. // J. Chem. Inf. Model. 2008. V. 48. P. 319–332.
- Blomberg D., Fex T., Xue Y., Brickmann K., Kihlberg J. // Org. Biomol. Chem. 2007. V.
   5(16). P. 2599–2605.
- 84. Pyrkov T.V., Kosinsky Y.A., Arseniev A.S., Priestle J.P., Jacoby E., Efremov R.G. // J. Chem. Inf. Model. 2007. V. 47. P. 1171–1181.
- 85. Bindewald E., Skolnick J. // J. Comput. Chem. 2005. V. 26. P. 374–383.

## Подписи к рисункам Пыркова Т.В. и др.

**Рис. 1.** Общая схема процесса молекулярного докинга. Начальный этап – генерация различных конформеров лиганда. Затем происходит их ранжирование в соответствии с качеством межмолекулярных контактов, образованных между лигандом и активным центром рецептора. Дальнейшая оптимизация конформеров будет происходить до тех пор, пока не наступят условия, завершающие цикл (см. текст).

**Рис. 2.** Соотношение числа водородных связей пептидных групп лигандов с атомами основной и боковых цепей аминокислотных остатков белка-рецептора.

**Рис. 3.** Пример образования мотива водородных связей в комплексах GTP с молибдоптерин-гуанин-динуклеотидсинтазой (PDB-код 1FRW), аденилосукцинатсинтазой (PDB-код 1LOO) и ксантин-гуанин-фосфорибозилтрансферазой (PDB-код 1A97).

Рис. 4. Применение концепции МГП для учета гидрофобной комплементарности. (*a*) Комплементарность гидрофобных/гидрофильных свойств в комплексе ATP– Ca-ATP-аза (на рисунке показан только ATP-связывающий домен Ca-ATP-азы). Отчетливо видна комплементарность гидрофобных свойств (выделено черным цветом) аденина лиганда и его «гидрофобного кармана» в белке. Рисунок получен с помощью веб-сервера PLATINUM (http://model.nmr.ru/platinum/) по методу МГП, основанному на атомных константах гидрофобности. (*б*) МГП<sub>*j*</sub> – значение МГП в точке пространства *j*, *f<sub>i</sub>* – атомная константа гидрофобности, *g*(*r<sub>ij</sub>*) – функция затухания.

**Рис. 5.** Геометрические параметры, описывающие стэкинг-взаимодействие между двумя ароматическими циклами: расстояния между центрами циклов в плоскости одного из них (смещение, *d*) и вдоль перпендикуляра к ней (высота, *h*); угол между плоскостями циклов (*α*).

**Рис. 6.** Распределение геометрических параметров взаимного расположения двух ароматических циклов при стэкинг-взаимодействии на примере фенольного кольца остатка тирозина белка и аденина различных лигандов, образующих комплекс с белкамирецепторами. Понятия высоты, смещения и угла  $\alpha$  между двумя взаимодействующими циклами приведены на рис. 5. Черные и серые точки отражают, соответственно, параллельное (cos2 $\alpha$  = 0.6 ÷ 1.0) и перпендикулярное (T-образное) (cos2 $\alpha$  = 0.0 ÷ 0.4) взаимное расположение колец.

**Рис. 7.** Пример расчета термов межмолекулярных взаимодействий на веб-сервере PLATINUM (http://model.nmr.ru/platinum/) для различных ориентаций ATP в активном центре Ca-ATP-азы (PDB код 1VFP), полученных с помощью докинга. Все результаты докинга сравнивали с нативной (выделено темным цветом) структурой лиганда по

значению среднеквадратичного отклонения по координатам атомов (rmsd). Термы, вычисленные с помощью веб-сервера PLATINUM, включают площадь липофильного S<sub>LL</sub> и гидрофильного S<sub>HH</sub> контактов, а также площадь поверхности интерфейса лиганда с рецептором S<sub>buried</sub> (для сравнения дана общая площадь поверхности лиганда S<sub>total</sub>); Match<sup>1</sup> и Match<sup>2</sup> отражают степень комплементарности гидрофобных/гидрофильных и гидрофобных свойств лиганда и рецептора, соответственно; H-bonds, Stack, Stack(Gua- $\pi$ ) – число водородных связей, ароматического стэкинга и стэкинга с положительно заряженными гуанидиновыми группами, соответственно (дробные числа – результат вычислений весовых функций геометрических параметров этих контактов – см. раздел «Стэкинг»).

# Пырков



# Пырков



# Пырков



Пырков



Пырков



# Пырков



# Пырков

Рис. 7

Results Receptor name: 1VFP\_prt.mol2 Reference ligand: 1VFP\_atp.pdb.

Ligand name	r.m.s.d.	ГГр	# H-bonds	$S_{L/L}$	$S_{H/H}$	Sburied	Stotal	Match <sup>1</sup>	Match <sup>2</sup>	# Stack.	# Stack. Gua–π
1VFP_atp 🛛 🚽 🗬	0.00	1	5.36	0.95	248.23	290.06	319.68	0.7794	0.0424	1.933	0
gold_soln_atp_m1_2 🛃 🗬	1.94	0.714	5.55	5.25	248.80	289.11	322.97	0.7866	0.208	0.804	0
gold_soln_atp_m1_5 🛃 🗬	8.16	0.208	2.04	0.00	234.87	287.97	335.25	0.7006	0	0	0.121
gold_soln_atp_m1_4 🛃 🗬	1.19	0.632	4.27	0.25	244.18	280.63	313.54	0.7796	0.0128	1.891	0
gold_soln_atp_m1_1 🛃 🗬	1.21	0.684	4.16	0.06	242.53	279.18	309.37	0.7842	0.0034	1.742	0
gold_soln_atp_m1_3 🛃 🗬	3.89	0.55	1.87	1.77	229.30	252.59	311.52	0.7418	0.1083	1.659	0

# MOLECULAR DOCKING: ROLE OF INTERMOLECULAR CONTACTS IN FORMATION OF COMPLEXES OF PROTEINS WITH NUCLEOTIDES AND PEPTIDES

T. V. Pyrkov<sup>\*, \*\*#</sup>, I. V. Ozerov<sup>\*, \*\*\*</sup>, E. D. Balitskaya<sup>\*, \*\*\*</sup>, R. G. Efremov<sup>\*</sup>

<sup>#</sup>*Phone/fax:* +7(495) 336-20-00; *e-mail:* pyrkov@nmr.ru

\* Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry Russian Academy of Sciences, Ul. Miklukho-Maklaya, 16/10, 117997 GSP, Moscow V-437, Russia
\*\* Moscow Institute of Physics and Technology (State University)
\*\*\* Lomonosov Moscow State University

Knowledge of 3D-structure of protein-ligand complex is a major prerequisite for understanding the functioning mechanism of cellular proteins and membrane receptors. This is also of a great help in rational drug design projects. In the present paper we briefly review the molecular docking approaches used to predict possible orientation of a ligand in the protein binding site. The recent trends to improve the accuracy and efficiency of docking algorithms are demonstrated with the results obtained in Laboratory of Biomolecular Modeling. Particular attention is paid to proteinligand hydrophobic and stacking interactions responsible for molecular recognition of ligand fragments. Such type of interactions are not always adequately represented in scoring criteria of docking applications that leads to mismatch in 3D-structure complexes predictions. That is why further inquiry of methods to account for these interactions is now the area of active research.

Key words: scoring function; hydrogen bonds; hydrophobic interactions; stacking

# Для переводчика: Пырков Т.В., Биоорган. химия №4, 2010

рационального дизайна новых	rational drug design
лекарственных соединений	
Докинг	docking
Ранжирование	scoring
мотивы водородных связей	hydrogen-bonding motifs
гидрофобные взаимодействия	hydrophobic interactions
Стэкинг	stacking
Оценочная функция	scoring function
Лиганд	ligand
конформационный поиск	conformation search
белок-мишень	protein-target
основная цепь [белка]	[protein] backbone
Сродство	affinity
Терм	term
Ван-дер-Ваальс	van der Waals
потенциал Леннарда-Джонса	Lennard-Jones potential
наталкивания между атомами	atom clashes
функция-ступенька	step function
число замороженных при связывании	number of rotational bonds frozen upon
лиганда торсионных углов	binding
кривая радиального распределения	radial distribution function
[активный] сайт	[active] site
матриксная металлопротеиназа	matrix metalloproteinase
Узнавание	recognition
низкомолекулярные соединения	small molecules
Заглубленный	embedded
двойная спираль ДНК	double-stranded DNA
Т-образный	T-shaped
π-катионном взаимодействии	$\pi$ -cation interaction
гуанидиновая группа	guanidine group
консенсусный подход	consensus approach
Пересортировать	re-score
система-специфичные (настраиваемые)	system-specific (adapted) scoring functions
оценочные функции	
конформационная подвижность	conformational flexibility