

# Молекулярное моделирование рецепторов физиологически активных веществ для целей медицинской химии †

И.И.Баскин, В.А.Палюлин, Н.С.Зефирова

Московский государственный университет им. М.В.Ломоносова, Химический факультет  
119991 Москва, Ленинские горы, факс (495)939–0290

Рассмотрена общая методология молекулярного моделирования биологических рецепторов и их взаимодействия с лигандами. Обсуждены методики построения пространственных моделей белков, молекулярного докинга, оценки качества моделей, определения свободной энергии связывания белков с лигандами. Представлены методы молекулярного дизайна новых лекарственных препаратов, основанные на молекулярных моделях биологических мишеней: виртуальный скрининг и дизайн *de novo*. Приведены примеры перечисленных выше подходов для моделирования ряда фармакологически значимых рецепторов, анализа взаимодействий рецептор–лиганд и дизайна новых биологически активных органических соединений.

Библиография — 314 ссылок.

## Оглавление

I. Введение	539
II. Методы молекулярного моделирования белков	540
III. Молекулярное моделирование различных типов биологических рецепторов	547

## I. Введение

Рецепторами физиологически активных веществ (ФАВ) являются молекулы, как правило, частично гликозилированных белков, обычно находящихся в клеточных мембранах (либо в мембранах клеточных органоидов). Биологическая функция таких белков состоит в передаче внешнего сигнала клетке (либо клеточному органоиду). Передача осуществляется путем связывания молекулы природного агониста (нейромедиатора или гормона), что приводит к определенным конформационным изменениям рецептора. Следствием этого является передача сигнала внутри клетки (либо клеточного органоида) посредством выделения вторичных мессенджеров (в случае метаболитных рецепторов) или иницирования трансмембранных ионных токов (в случае ионотроп-

ных рецепторов). Именно с участием рецепторов осуществляется регулирование основных биологических функций организма на клеточном уровне.

Вводимые в организм органические молекулы также могут связываться с белковыми макромолекулами рецепторов и вызывать в них конформационные изменения (в случае полных или частичных агонистов) либо, наоборот, блокировать их (в случае антагонистов), либо изменять интенсивность внутриклеточного ответа (в случае модуляторов). Органические молекулы регулируют биологическую функцию организма, воздействуя на передачу внутриклеточных сигналов. Этим и обусловлена возможность лечения различных заболеваний с помощью соответствующих лекарственных препаратов. Рецепторы ФАВ наряду с ферментами являются основными типами биомишеней, на которые направлено действие лекарств. Поэтому знание пространственного строения рецепторов ФАВ и особенно их сайтов связывания с органическими лигандами важно для разработки новых лекарственных веществ.

Основным методом изучения пространственного строения биологических макромолекул является рентгеноструктурный анализ. Возможно, что уже в недалеком будущем подобные исследования будут проводить также и при помощи спектроскопии ЯМР, однако в настоящее время сфера применения этого метода ограничена изучением молекул небольших пептидов, а также сегментов либо доменов в составе сложных белковых молекул. Основное ограничение

**И.И.Баскин.** Кандидат химических наук, ведущий научный сотрудник лаборатории органического синтеза кафедры органической химии химического факультета МГУ.

Телефон: (495)939–3557, e-mail: baskin@org.chem.msu.ru

**В.А.Палюлин.** Кандидат химических наук, ведущий научный сотрудник той же лаборатории. Телефон: (495)939–3557, e-mail: vap@org.chem.msu.ru

**Н.С.Зефирова.** Академик, заведующий той же кафедрой.

Телефон: (495)939–1620, e-mail: zefirov@org.chem.msu.ru

Область научных интересов авторов: органическая химия, поиск количественных соотношений структура – свойство и структура – активность органических соединений, искусственные нейронные сети.

Дата поступления 29 января 2009 г.

† Данная статья завершает серию публикаций обзорных работ, отражающих основные направления научных исследований в области медицинской химии (см. *Успехи химии* 78 (5) (2009)).

рентгеноструктурного анализа связано с необходимостью иметь кристаллы анализируемых объектов. Относительно легко можно получать кристаллы небольших водорастворимых белков, но крайне трудно — белков и белковых комплексов, интегрированных в клеточную мембрану. К последнему типу относится большинство локализованных в клеточной мембране рецепторов ФАВ. Экспериментально изучено пространственное строение лишь очень небольшого числа таких рецепторов. Именно поэтому возможность теоретического предсказания пространственного строения рецепторов ФАВ и изучения их взаимодействия с молекулами биологически активных органических соединений методами компьютерного моделирования — одна из первоочередных задач в медицинской химии и фармакологии.

## II. Методы молекулярного моделирования белков

Молекулярное моделирование рецептора ФАВ обычно включает построение грубой модели пространственной структуры молекулы, ее уточнение и изучение динамических свойств макромолекулы белкового рецептора при помощи методов молекулярной динамики, а также теоретическое изучение его взаимодействия с лигандами.

### 1. Построение пространственной структуры молекулы белка методами сравнительного моделирования

В настоящее время построение начальной структуры белка в большинстве случаев основано на применении методов сравнительного моделирования, называемых также методами моделирования по гомологии. При помощи таких методов строят пространственную модель макромолекулы белка, исходя из его аминокислотной последовательности, путем переноса экспериментально полученной информации об исследованных белках-шаблонах на моделируемые белки. Подобный перенос может осуществляться на глобальном и локальном уровнях.

На глобальном уровне переносят информацию о типе основных цепей (задаваемом положениями  $C_{\alpha}$ -атомов) макромолекулы (как правило, но необязательно, в рамках таких структурно-консервативных элементов вторичной структуры белка, как  $\alpha$ -спирали,  $\beta$ -слои и др.) белка-шаблона на моделируемый белок. В основе такого подхода лежит представление о том, что все нативные белки возникли в результате длительной молекулярной эволюции, которую на генетическом уровне можно описать как последовательность мутаций (с некоторыми усложнениями, связанными с процессами обмена генетическим материалом, например, при помощи кроссинговера). Это означает, что молекулярная эволюция белка включала последовательные процессы точечной замены аминокислот и, возможно, вставок или удалений небольших участков аминокислотной последовательности. Для осуществления биологических функций в живом организме очень важно сохранение определенных черт пространственного строения молекулы белка.

Для выживания организма важно правильное функционирование всех его белков, значит, пространственное строение эволюционно-связанных нативных белков, по крайней мере на глобальном уровне, должно быть сходным. Следовательно, информация на макроуровне о пространственном строении нативных белков может быть перенесена с белка-шаблона (или с нескольких белков-шаблонов) с известным из эксперимента пространственным строением на эволю-

ционно-связанные с ними нативные белки, для которых требуется построить молекулярные модели. Чем ближе первичная структура моделируемого белка к первичной структуре белка-шаблона, тем теснее между ними эволюционная связь и будет точнее построенная пространственная модель. Белки с близкими последовательностями аминокислот всегда эволюционно связаны и всегда обладают схожим пространственным строением. Обратное, однако, неверно — иногда белки с сильно отличающимися аминокислотными последовательностями обладают схожим пространственным строением. В качестве примера можно привести семейство рецепторов, сопряженных с G-белками, что свидетельствует об очень далекой эволюционной связи между ними, поэтому надо применять другие критерии для ее установления.

Перенос информации на локальном уровне применяют при построении боковых цепей всех аминокислот, а также при моделировании. В этих случаях используют библиотеку ротамеров боковых цепей и базу данных белковых петель. Такие массивы информации формируются при анализе пространственного строения всех известных белков, а не только эволюционно-связанных.

Процесс сравнительного моделирования белка включает следующие стадии: 1) выбор белка-шаблона; 2) построение выравнивания аминокислотных последовательностей моделируемого и шаблонного белков; 3) построение пространственной модели моделируемого белка; 4) оценка качества построенной модели.<sup>1</sup> Как правило, первая и вторая стадии объединяются в одну.

#### а. Выбор шаблона и построение выравнивания аминокислотных последовательностей

Первая стадия сравнительного моделирования заключается в поиске для моделируемого белка эволюционно-связанного с ним белка-шаблона с известным пространственным строением. Для этого можно, например, сравнить последовательности аминокислот моделируемого белка с последовательностями всех белков, для которых установлено пространственное строение, и выбрать в качестве шаблонов те, для которых мера близости последовательностей превышает определенный порог. В этом случае для сравнения первичных структур белков применяют процедуру выравнивания: производится поиск взаимного расположения аминокислотных последовательностей одной под другой в таблице, при котором в каждом столбце располагаются друг под другом по возможности одинаковые либо близкие по свойствам аминокислоты. Если для такого построения требуется внести разрывы в последовательности, то это обычно делается, особенно если они соответствуют аминокислотам, не входящим в структурно-консервативные элементы вторичной структуры, в частности  $\alpha$ -спирали и  $\beta$ -слои. Различают глобальное выравнивание (сравниваются последовательности близких размеров) и локальное (сравнивается целиком одна последовательность с участком внутри другой последовательности большего размера). Обычно для построения этих видов выравнивания соответственно применяют основанные на методах динамического программирования<sup>2,3</sup> алгоритмы Нидлмана–Вунша<sup>4</sup> и Смита–Ватермана.<sup>5</sup> В обоих случаях осуществляется комбинаторная минимизация общей штрафной функции за различие находящихся друг под другом аминокислот. Эту функцию рассчитывают, используя заранее выбранные эмпирические матрицы замен,<sup>6</sup>

которые определяют штраф за различие типов аминокислот. Из матриц замен наиболее распространены матрицы семейств PAM<sup>7</sup> и BLOSUM,<sup>8</sup> причем последние более предпочтительны при моделировании по гомологии на основании структур эволюционно-удаленных белков.<sup>9</sup> Выравнивание последовательностей аминокислот может быть как парным, так и множественным (с участием трех и большего числа последовательностей). Как правило, множественное выравнивание приводит к построению более адекватных молекулярных моделей, даже если в него входит последовательность аминокислот только одного белка-шаблона.

Операции выбора шаблона и построения выравнивания часто сопряжены с поиском в очень больших базах данных, содержащих все известные аминокислотные последовательности. В качестве инструментов поиска обычно используют такие программы, как FASTA<sup>10</sup> и BLAST<sup>11,12</sup> (для парного выравнивания), а также PSI-BLAST<sup>12,13</sup> (для множественного выравнивания). Однако следует отметить, что ввиду ориентации программ FASTA и BLAST на максимальную производительность, необходимую для работы с большими объемами информации, качество выравниваний не всегда оказывается высоким. Поэтому перед моделированием имеет смысл перестроить выравнивания с помощью программ, ориентированных на более тщательное построение выравниваний, например программ семейств CLUSTAL<sup>14–16</sup> и COFFEE.<sup>17,18</sup>

Альтернативным способом поиска белка-шаблона и построения выравниваний является метод «протягивания»,<sup>19,20</sup> который представляет собой эффективный вычислительный подход к распознаванию типа пространственной укладки белка<sup>21</sup> и предсказанию его структуры.<sup>22,23</sup> В основе этого подхода лежит представление о том, что молекулярная эволюция происходит преимущественно путем внесения мутаций без изменения типа пространственной укладки (ТПУ). Образование нового ТПУ — очень редкое событие, поэтому накоплено относительно небольшое число ТПУ (по оценочным данным, 1500–1600), которыми обусловлено огромное разнообразие возникших в ходе эволюции белковых структур. Судя по всему, появление существенно большего числа ТПУ невозможно. Экспериментально изученные к настоящему времени методами рентгеноструктурного анализа и спектроскопии ЯМР белки образуют 1100 ТПУ. Следовательно, велика вероятность (70–80%) того, что произвольно выбранный белок имеет один из ТПУ, для которого имеется хотя бы один представитель с известной структурой, а значит, пространственное строение выбранного белка может быть предсказано методами сравнительного моделирования.

В методе «протягивания» аминокислотная последовательность моделируемого белка при помощи методов динамического программирования<sup>2,3</sup> «протягивается» через каждый ТПУ из имеющихся в библиотеке. Затем с использованием специальной эмпирической функции энергии определяется тот ТПУ, который наилучшим образом совместим с данной последовательностью. Описание ТПУ белка включает характеристику его трехмерной структуры, поэтому эту процедуру называют также 1D–3D-выравниванием. Используемая при этом эмпирическая функция энергии имеет статистическую природу, т.е. связана соотношением Больцмана с вероятностями встречаемости аминокислот данного типа в некотором структурном контексте (например, снаружи или внутри белковой глобулы и внутри какого элемента вторичной структуры), а также частоты встречае-

мости аминокислот данных типов на определенном расстоянии друг от друга.<sup>24</sup>

## б. Построение моделей

Построение пространственной структуры молекулы белка методами сравнительного моделирования производится на основе информации, содержащейся в выравнивании аминокислотных последовательностей моделируемого белка и белков-шаблонов (их может быть один или несколько), в пространственных структурах белков-шаблонов, а также в статистике распределения локальных геометрических параметров по всему массиву белков с известным пространственным строением. В настоящее время различают три основных подхода к построению моделей:<sup>25</sup> 1) сборка фрагментов; 2) комбинирование сегментов; 3) перенос статистических распределений геометрических характеристик.

**Сборка фрагментов.** Метод сборки фрагментов — исторически первый подход к построению пространственных структур белков методом сравнительного моделирования. В его основе лежит следующая эмпирическая закономерность: в структурах белков, входящих в близкородственные семейства (например, в семейство сериновых протеаз млекопитающих), можно выделить два типа участков.<sup>26</sup> Пространственные структуры участков первого типа строго консервативны внутри семейства белков, поэтому их часто называют структурно-консервативными зонами. Как правило, они состоят из стабильных элементов вторичной структуры, в частности  $\alpha$ -спиралей и  $\beta$ -листов. Пространственные структуры участков второго типа, напротив, сильно варьируются внутри семейства белков. Обычно они соответствуют реальным петлям в белковых структурах, поэтому все участки этого типа условно (в контексте данного метода моделирования) стали называть петлями.

В методе сборки на первом этапе осуществляется построение фрагментов моделируемого белка внутри структурно-консервативных участков путем переноса координат атомов основной цепи белков-шаблонов на моделируемый белок. Далее для каждой петли этого белка проводится поиск структуры в базе данных по петлям. База формируется по результатам анализа всех белков с известным пространственным строением. Наиболее подходящие из найденных петель, т.е. наиболее удачно стыкующиеся с уже построенными фрагментами белка и наиболее близкие по гомологии с соответствующими аминокислотными последовательностями моделируемых петель, служат шаблонами для задания координат атомов основной цепи внутри петель моделируемого белка. Затем для всех аминокислот задаются координаты атомов боковых цепей. Для этого используют данные библиотеки ротамеров, содержащей для каждого типа аминокислот и для каждого типа вторичной структуры наиболее часто встречающиеся значения торсионных углов. Имеется множество модификаций данной процедуры, различающихся лишь в деталях.<sup>27</sup> Наиболее известная компьютерная программа, работающая в соответствии с этим подходом, — COMPOSER<sup>28</sup> входила в состав комплекса SYBYL.

**Комбинирование сегментов.** Сравнительное моделирование по методу комбинирования сегментов основано на разбиении моделируемого белка на небольшие сегменты, для моделирования каждого из которых используют в качестве шаблона отдельный сегмент белка с известным пространственным строением, выбираемый при помощи процедуры локального выравнивания аминокислотных последовательностей.<sup>29</sup>

**Перенос статистических распределений геометрических характеристик.** В настоящее время метод переноса статистических распределений геометрических характеристик наиболее часто используют при построении пространственных моделей белков.<sup>30</sup> Его можно рассматривать как расширенный метод дистанционной геометрии, позволяющий при наличии геометрических ограничений на межатомные расстояния генерировать удовлетворяющие им конформации молекул.<sup>31,32</sup> Принципиальное же отличие рассматриваемого подхода от метода дистанционной геометрии заключается в том, что в нем каждое геометрическое ограничение описывается не точными значениями минимума и максимума соответствующего геометрического параметра, а функцией плотности вероятности его распределения, параметризованной на массиве белков с известной геометрией.<sup>30</sup> В рамках этого подхода используют функции плотности вероятности распределения геометрических характеристик двух основных типов: расстояний между  $C_{\alpha}$ -атомами основной цепи и внутренних геометрических параметров (длин связей, межатомных расстояний, валентных и торсионных углов) молекулы белка. Распределения для геометрических характеристик первого типа аппроксимируются исходя из известной геометрии белков-шаблонов, а для второго типа — на большом массиве белков с известным пространственным строением.

В рамках данного подхода путем суммирования отрицательных логарифмов функций плотности вероятности для различных геометрических характеристик формируется молекулярная функция плотности вероятности, зависящая от координат атомов белка. Эта функция является псевдоэнергетической; ее значение равно сумме вкладов перечисленных выше геометрических характеристик, причем каждый из вкладов определяется тем, насколько часто значение соответствующей характеристики встречается в массиве информации о белках с известной геометрией, используемом при параметризации. Путем оптимизации молекулярной функции плотности вероятности методами нелинейного программирования находят координаты всех атомов моделируемого белка.<sup>30</sup> В соответствии с данной методологией работает, в частности, весьма распространенная программа MODELLER,<sup>33–38</sup> с помощью которой осуществлено сравнительное моделирование значительного числа рецепторов ФАВ. Сходным образом работает также программа Rosetta в режиме сравнительного моделирования.<sup>39–41</sup>

**Точность моделирования по гомологии.** Принято считать, что точность моделирования по гомологии зависит от степени идентичности аминокислотных последовательностей моделируемого белка и шаблона. По имеющимся в литературе оценкам,<sup>25</sup> при идентичности > 50% модели являются достаточно надежными, возможные ошибки обычно затрагивают лишь конформации боковых цепей, а среднеквадратичное отклонение от экспериментально определенных структур составляет  $\sim 1 \text{ \AA}$ . При идентичности первичных структур 30–50% ошибки уже касаются и хода основной цепи в петлях; при идентичности < 30% ошибочным может оказаться также тип укладки основной цепи белка, а среднеквадратичное отклонение от эксперимента > 3  $\text{ \AA}$ .<sup>25</sup> Однако такие оценки не следует воспринимать как догму, поскольку во многих случаях (например, при моделировании рецепторов, сопряженных с G-белками) удается построить модели удовлетворительного качества и при меньшей идентичности аминокислотных последовательностей.

## в. Оценка качества моделей

Процесс моделирования пространственной структуры белка неоднозначен, поскольку на него влияет большое число факторов: выбор белков-шаблонов, деталей построения выравнивания аминокислотных последовательностей, выбор случайной начальной точки для оптимизации геометрии и т.д. Поэтому большое значение имеет процедура оценки качества моделей белков, позволяющая выбирать как наилучшие из построенных моделей, так и выявлять в них ошибки, вносить необходимые коррективы в процесс моделирования. В настоящее время применяют пять основных подходов к оценке качества построенных пространственных моделей белковых макромолекул: сравнение с результатами биологических экспериментов, использование геометрических критериев, метод статистических потенциалов, физико-химические подходы и методы статистического обучения.

**Сравнение с результатами биологических экспериментов.** Наиболее эффективным и часто используемым подходом к верификации пространственных моделей белков является метод сайт-направленного мутагенеза, в котором применяют принципы генной инженерии.<sup>42,43</sup> С его помощью можно, в частности, изменить тип любой из аминокислот в молекуле белка, что дает возможность экспериментально изучать влияние такой замены на какую-либо из функций белка. В этом случае критерием правильности построенной пространственной модели является возможность с ее помощью объяснить наблюдаемое влияние замены аминокислоты на изучаемую функцию белка, например на внутриклеточный ответ под действием определенного агониста, локализованного в мембране клетки рецептора ФАВ. В последнем случае наиболее сильно должен измениться внутриклеточный ответ при замене аминокислот, находящихся в близком контакте с молекулой агониста внутри сайта связывания.

Таким образом, используя информацию, полученную в генно-инженерных экспериментах по сайт-направленному мутагенезу, можно проводить частичное или даже полное картирование молекулы рецептора. Это позволяет точно идентифицировать сайты связывания различных лигандов, и дает ценные сведения для выбора правильной пространственной модели. На таком принципе основаны, в частности, итерационные процедуры уточнения модели, включающие несколько последовательных этапов: 1) построение сравнительной модели белка; 2) составление, согласно построенной модели, списка аминокислот, образующих сайт связывания; 3) замена при помощи сайт-направленного мутагенеза этих аминокислот на, например, аланин; 4) измерение функций молекул-мутантов; 5) сравнение полученных данных с прогнозируемыми на основании пространственной модели белка; 6) в случае расхождений необходимо перестроение модели белка. С помощью подобной итерационной процедуры построена пространственная модель холецистокининового рецептора.<sup>44,45</sup>

Однако не следует переоценивать возможности метода сайт-направленного мутагенеза для проверки качества построенной пространственной модели, поскольку изменение биологического ответа вследствие мутации может происходить по множеству различных механизмов, которые могут быть и не связаны с расположением данного остатка в активном центре белка. В частности, точечные мутации

могут приводить к глобальным изменениям пространственного строения молекулы белка и влиять на возможность осуществления необходимых для его функционирования конформационных перестроек. Кроме того, мутации аминокислот, расположенных вдали от активного центра, могут влиять на доступ лиганда к активному центру, а также на встраивание белка в клеточные мембраны. Поэтому необходимо применять дополнительные критерии правильности построенных моделей.

Для оценки качества моделей белков могут быть полезны биологические эксперименты, позволяющие выявлять сайты гликозилирования. В этом случае критерием правильности модели является факт нахождения аминокислот, подвергающихся гликозилированию, на поверхности белковой глобулы, контактирующей с водной фазой. Этой же цели может служить выявление сайтов палмитилирования, которые также должны располагаться на поверхности белка, но при этом они должны быть локализованы у края мембраны.

**Геометрические критерии.** При проверке геометрических критериев в построенной пространственной модели белка вычисляют значения длин всех имеющихся связей и всех валентных углов. Очевидно, что отклонение хотя бы одного из найденных значений от стандартных указывает на наличие локальной ошибки в модели. На это же указывает и слишком малое расстояние между парой несвязанных атомов в молекуле. Локальные ошибки во многих случаях могут быть устранены в процессе последующего уточнения модели при оптимизации.

Необходимым этапом проверки модели является вычисление абсолютных стереохимических конфигураций  $C_{\alpha}$ -атома основной цепи белка, а также  $C_{\beta}$ -атома боковой цепи изолейцина. Очевидно, что наличие стереохимической конфигурации  $R$  хотя бы у одного центра из этих атомов означает грубую ошибку в модели. Следует также отметить, что проверку стереохимической конфигурации указанных выше атомов необходимо проводить многократно и на всех этапах моделирования, поскольку неправильная конфигурация может возникнуть даже на заключительных этапах при оптимизации геометрии молекулы при помощи методов молекулярной механики и молекулярной динамики. То же относится и к проверке *транс*-конфигураций всех амидных связей в основной цепи белка. Из геометрических критериев проверяется также отклонение атомов от плоскости во всех бензольных ядрах в остатках фенилаланина и тирозина, в гетероциклических ядрах в остатках триптофана и гистидина, в гуанидиниевой группе гистидина, во всех карбоксильных и карбамидных группах, а также вокруг всех амидных связей в основной цепи белка.

При проверке геометрических критериев следует провести анализ распределений значений торсионных углов  $\varphi$  и  $\psi$  в основной цепи молекулы белка. Для каждого типа аминокислот значения этих углов используют в качестве координат при определении расположения соответствующей аминокислоты на карте Рамачандрана. В таком случае критерием правильности построенной пространственной модели молекулы белка будет число аминокислот, попавших в запрещенные области карт Рамачандрана, — оно не должно существенно превышать аналогичное число в известных из эксперимента пространственных структурах белков-шаблонов, по которым проводилось сравнительное моделирование.

**Метод статистических потенциалов.** Правильность переноса информации на моделируемые белки на локальном уровне контролируют, проверяя рассмотренные выше гео-

метрические критерии; для оценки достоверности модели на глобальном уровне (см. выше) используют статистические потенциалы.<sup>24, 46–49</sup> Под статистическим потенциалом в данном случае понимается функция, показывающая относительную<sup>‡</sup> вероятность наличия данной пространственной структуры у белка. Вероятность выражена в энергетической шкале при помощи соотношения Больцмана. Такие функции формируются аддитивными по вкладам, вносимым всеми структурными признаками, которыми могут быть, например, наличие контакта между аминокислотами данных типов либо присутствие аминокислоты данного типа в определенном структурном контексте. Величину каждого из вкладов определяют путем статистического анализа относительной частоты встречаемости соответствующего структурного признака в базе данных, содержащей все известные пространственные структуры белков. Эмпирические функции, используемые в упомянутом выше методе «протягивания», также являются статистическими потенциалами.

Статистические потенциалы могут быть использованы при оценке качества моделей в двух режимах: глобальном и локальном. В глобальном режиме статистический потенциал вычисляют для всей молекулы белка и применяют для отбора лучшей модели из множества возможных. Именно для этого используется статистический потенциал DOPE<sup>50, 51</sup> в программе MODELLER.<sup>33–38</sup> В локальном режиме статистический потенциал вычисляется отдельно для каждого из аминокислотных остатков и применяется для построения специальных графиков — так называемых 3D-профилей,<sup>52</sup> показывающих зависимость парциального статистического потенциала от номера аминокислотного остатка. С помощью 3D-профилей можно обнаружить конкретные участки внутри белковой молекулы, которые смоделированы неправильно. Эта методология реализована в ряде программ (а также Web-сервисов), в частности Verify3D<sup>21, 52, 53</sup> и ProSA.<sup>46, 54</sup> Однако следует иметь в виду, что упомянутые выше методы параметризованы для глобулярных белков, поэтому при построении моделей рецепторов, связанных с мембранами, такие методы правомерно использовать только для проверки правильности моделирования внеклеточных и внутриклеточных доменов, находящихся в водном окружении.

**Физико-химические подходы.** Физико-химические подходы к оценке качества модели основаны на вычислении при помощи методов молекулярной механики общей энергии молекулы белка с использованием силовых полей, обязательно учитывающих в неявном виде сольватацию. В качестве примера можно привести специально разработанное для этой цели силовое поле EFF,<sup>55, 56</sup> которое позволяет различать нативную конформацию молекулы белка и неправильно построенные ее молекулярные модели.

**Методы статистического обучения.** Ввиду разнообразия методов оценки качества построенных пространственных моделей молекул белка актуальны задачи выбора наиболее релевантных среди них и нахождения такого способа их комбинирования, который обеспечивает более высокое качество распознавания правильных молекулярных моделей по сравнению с применением индивидуальных подходов. В настоящее время для решения этой задачи обычно используют современные методы статистического обучения, такие как искусственные нейронные сети<sup>57–59</sup> и метод опорных векторов.<sup>60–62</sup> В одном из первых исследований искусственные

<sup>‡</sup> Относительно реперного состояния со случайно размещенными в пространстве аминокислотами.

нейронные сети были использованы для улучшения работы программы GenTHREADER, которое достигалось за счет комбинирования статистического потенциала, используемого в методе «протягивания», оценочной функции, описывающей качество выравнивания аминокислотных последовательностей, длины моделируемой аминокислотной последовательности и некоторых других характеристик.<sup>63</sup>

В статье<sup>51</sup> описано применение метода опорных векторов для комбинирования статистического потенциала DOPE,<sup>50, 51</sup> трех статистических потенциалов из системы MODPIPE<sup>§</sup> и двух оценочных функций PSIPRED/DSSP<sup>68–70</sup> с целью получения новой оценочной функции, позволяющей наилучшим образом распознавать правильные пространственные модели молекул белка.

## 2. Построение модели структуры молекулы белка методами *ab initio*

Очевидным недостатком методов сравнительного моделирования является то, что с их помощью нельзя построить пространственную структуру белка, для которого отсутствует подходящий белок-шаблон.<sup>71–75</sup> В таких случаях для построения молекулярных моделей начали применять методы моделирования *ab initio*, для которых не требуется наличия шаблонов. Наиболее известной программой, реализующей данную методологию, является Rosetta,<sup>76–80</sup> в рамках которой при помощи стохастической процедуры минимизации эмпирической функции энергии (по сути близкой к рассмотренным выше статистическим потенциалам) осуществляется построение пространственной модели белка путем ее сборки из микрофрагментов. Аналогично работает и программа SimFold,<sup>81</sup> в которой при построении молекулярной модели производится минимизация одноименной функции энергии, однако эта функция имеет физико-химическую природу. Следует иметь в виду, что при моделировании больших молекул белка с использованием методов *ab initio* получаются менее точные модели, чем при использовании методов сравнительного моделирования при наличии подходящих шаблонов.

## 3. Уточнение модели и изучение динамических свойств макромолекул рецепторов

Вслед за первоначальным построением пространственной модели молекулы рецептора методом сравнительного моделирования обязательно следует стадия уточнения путем оптимизации геометрии молекулы методом молекулярной механики<sup>82, 83</sup> с использованием эмпирических силовых полей.<sup>84</sup> Во многих исследованиях оптимизированную геометрию молекулы белка используют в качестве исходной для проведения моделирования методом молекулярной динамики.<sup>83, 85</sup> В таких случаях молекулу водорастворимого белка (либо находящегося в водной среде домена белка) обычно подвергают сольватации — помещают в куб (в некоторых случаях в октаэдр или сферу), заполненный молекулами воды, для достижения суммарного нулевого электрического заряда в него добавляют противоионы. Для

учета сольватации используют также расчеты, в которых влияние воды аппроксимировано при помощи континуальных моделей сольватации (например, модели Борна). При моделировании доменов рецепторов ФАВ, встроенных в мембрану клеток, модель белка встраивают в предварительно подготовленную молекулярную модель липидного бислоя (пример см. в статье<sup>86</sup>).

Стадия моделирования с помощью метода молекулярной динамики обычно проводится для достижения следующих целей. Во-первых, в процессе непродолжительного моделирования можно уточнить геометрию молекулы, поскольку система может выйти из мелких локальных энергетических минимумов, а при более длительном моделировании можно проследить влияние сольватирующей среды (например, мембраны) на строение белка. Во-вторых, при проведении молекулярно-динамического моделирования в течение определенного времени после достижения равновесия получают усредненные значения геометрических и энергетических характеристик, которые можно сопоставлять с экспериментальными данными. Особенно это относится к изучению комплексов белок–лиганд. В-третьих, при проведении длительного (> 5–10 нс) молекулярно-динамического моделирования возможно изучение динамических свойств макромолекул. В применении к моделированию рецепторов ФАВ это означает возможность изучения конформационных перестроек в белковой макромолекуле, приводящих к активации рецептора, что необходимо для поиска потенциальных лекарственных веществ, влияющих на функционирование рецептора.

При моделировании рецепторов ФАВ и их комплексов с лигандами методами молекулярной механики и молекулярной динамики наиболее часто используют следующие программы и программные комплексы: AMBER,<sup>87, 88</sup> CHARMM,<sup>89</sup> NAMD,<sup>90</sup> GROMOS,<sup>91</sup> GROMACS.<sup>92</sup>

## 4. Критическая оценка методов предсказания структуры белков

Для объективной оценки того, какие именно подходы и какие компьютерные программы и расчетные серверы дают возможность построить наиболее точные пространственные модели белков, регулярно, начиная с 1994 г., проводится эксперимент по прогнозированию пространственной структуры белков, исходя из аминокислотных последовательностей. В ходе этого эксперимента, названного «критической оценкой методов предсказания структуры белков» (Critical Assessment of Techniques for Protein Structure Prediction (CASP)), исследовательские группы, участвуют в своего рода соревновании, в ходе которого они могут проверить предсказательную способность своих подходов. Каждый из участников такого соревнования получает аминокислотные последовательности белков, пространственное строение которых им неизвестно, но известно организаторам соревнования. Задача участника — построить пространственные модели этих белков. Результаты сравнения построенных моделей с пространственными структурами, определенными из эксперимента, регулярно (один раз в два года) выкладываются организаторами соревнования на сервер, который доступен в сети Интернет по адресу <http://predictioncenter.org>. Кроме того, подробные отчеты об эксперименте публикуются в специальных номерах журнала «Proteins», также доступных на этом сервере.

Таким образом, изучая публикуемые материалы, можно получить объективную картину, показывающую возможности методов прогнозирования пространственной струк-

§ Автоматическая система сравнительного моделирования молекул белка, основанная на использовании программы PSIBLAST<sup>12, 13</sup> для построения выравнивания аминокислотных последовательностей и программы MODELLER<sup>33–38</sup> для построения пространственных моделей.<sup>64–67</sup>

туры белков и помогающую любому пользователю грамотно подойти к выбору средств моделирования интересующих его белков.

## 5. Моделирование взаимодействия рецепторов с лигандами

Последним и наиболее важным для целей конструирования лекарственных веществ этапом молекулярного моделирования рецепторов ФАВ является изучение их взаимодействия с лигандами. Анализ взаимодействия рецепторов с лигандами обычно преследует две основные цели: подробно описать способ связывания молекулы лиганда с сайтом связывания рецептора и использовать полученную при этом информацию для поиска новых лигандов для данного рецептора.

### а. Молекулярный докинг

Основным средством предсказания способа связывания биологической макромолекулы с лигандом является процедура молекулярного докинга,<sup>93–98</sup> позволяющая отыскать наиболее энергетически выгодное расположение лиганда в сайте связывания биологической макромолекулы. В настоящее время процедура молекулярного докинга и анализ его результатов являются ключевыми при поиске новых лекарственных препаратов,<sup>99–104</sup> и поэтому она имеет большое значение для медицинской химии. Следует подчеркнуть, что также важны результаты докинга в биологические мишени не только органических молекул малого размера, но и небольших пептидов и даже целых белков (в последнем случае говорят о белок-белковом докинге<sup>105–109</sup>). Действительно, подробное знание структуры комплекса рецептора с пептидным лигандом дает ценную информацию для дизайна пептидомиметиков (небольших органических молекул, имитирующих действие пептидов),<sup>110–115</sup> а знание строения белок-белковых комплексов облегчает задачу создания лекарственных препаратов, биологическое действие которых обусловлено ингибированием белок-белковых взаимодействий небольшими органическими молекулами.<sup>116, 117</sup>

Первоначально процедуры молекулярного докинга предназначались для поиска оптимального расположения молекулы лиганда относительно молекулы биомишени (главным образом, белка) без рассмотрения в явном виде возможности изменения геометрии каждого из них при связывании (докинг с «жесткой» биомишенью и «жестким» лигандом). При этом используют процедуру помещения «ключа» (лиганда) в «замок» (в белковую биомишень).<sup>118</sup> Наиболее эффективный способ реализации такой процедуры состоит в описании форм «замка» и «ключа» (например, при помощи молекулярных поверхностей либо сферических гармоник Фурье<sup>119–121</sup>) и поиске такого взаимного расположения в пространстве этих форм, при котором они окажутся комплементарными друг другу.<sup>122</sup>

Альтернативный и более традиционный способ докинга состоит в поиске при помощи стохастических методов оптимизации (таких как генетические алгоритмы и алгоритмы имитации отжига) наиболее энергетически выгодного положения «ключа» внутри «замка».<sup>123</sup> В последнем случае с целью повышения эффективности вычислений энергию взаимодействия обычно рассчитывают путем интерполяции значений, предварительно рассчитанных в узлах решетки.<sup>124</sup> В настоящее время практически все программы докинга дополнены алгоритмами, дающими возможность учитывать конформационную гибкость небольших молекул ли-

гандов, а последние версии программ, как правило, обладают еще и ограниченной возможностью учета гибкости белковой макромолекулы, по крайней мере путем подстройки конформаций боковых цепей аминокислот в сайте связывания.<sup>125</sup> В отдельных публикациях (см., например,<sup>126–128</sup>) описаны способы учета гибкости основной цепи белка при докинге. Как правило, главным фактором, препятствующим учету гибкости молекулы белка при докинге, является пока еще очень резкое снижение скорости работы программ.

В настоящее время наиболее распространенными программами докинга, ориентированными на лиганды небольшого размера, являются DOCK,<sup>129–132</sup> AUTODOCK,<sup>123, 133–135</sup> GOLD,<sup>136, 137</sup> Glide,<sup>138–140</sup> FlexX.<sup>141–144</sup> В ряде публикаций проведено сравнение между ними (см., например,<sup>145</sup>). Из программ белок-белкового докинга наиболее распространены RosettaDock,<sup>146, 147</sup> ZDOCK,<sup>148, 149</sup> а также PatchDock.<sup>150</sup>

### б. Оценка свободной энергии связывания и ранжирование лигандов по средству к рецептору

Для целей медицинской химии большое значение имеют количественные оценки свободной энергии взаимодействия белок – лиганд (фактически констант связывания), поскольку именно она определяет биологическую активность лиганда. Среди используемых для этого подходов можно выделить неэмпирические, полуэмпирические и эмпирические.

Под неэмпирическими подходами к оценке свободной энергии взаимодействия белок – лиганд обычно понимают основанную на строгом применении аппарата статистической термодинамики группу близких методов, различающихся лишь техникой интегрирования: метод возмущения свободной энергии, метод термодинамического интегрирования и метод плавного роста.<sup>151–153</sup> В основе этих методов лежит процедура построения термодинамического цикла, включающего «алхимические» стадии постепенного преобразования («мутации») одной химической структуры в другую. Каждое такое преобразование проводится через большое число промежуточных «нефизических» состояний; для каждого из состояний путем тщательного молекулярно-динамического моделирования находится среднее по траектории значение внутренней энергии системы, которое используется при интегрировании, проводимом с целью вычисления свободной энергии этого преобразования. Такая процедура является строгой и поэтому приводит к результатам, хорошо согласующимся с экспериментальными данными. Тем не менее ее практическое применение требует гигантских объемов вычислений ввиду необходимости моделировать при помощи молекулярной динамики большое число (по крайней мере, несколько десятков) «нефизических» состояний. Вследствие этого практическое применение неэмпирических методов расчета свободной энергии для целей медицинской химии ограничено изучением взаимодействия лишь одного или нескольких наиболее перспективных органических лигандов с макромолекулой биомишени.

В полуэмпирических методах, в отличие от неэмпирических, не проводят молекулярно-динамических расчетов большого числа «нефизических» состояний, поэтому они требуют гораздо (в десятки раз) меньших затрат машинного времени. Одним из наиболее распространенных полуэмпирических методов является ММ-PBSA (ММ-GBSA),<sup>154–157</sup> который сочетает в себе расчет средней энергии взаимодействия белок – лиганд при помощи молекулярной динамики, оценку

конформационной энтропии комплекса путем расчета нормальных колебаний, вычисление электростатической составляющей энергии сольватации по методу Пуассона–Больцмана<sup>158</sup> либо по обобщенному методу Борна,<sup>159</sup> а также оценку гидрофобной составляющей энергии сольватации при помощи вычисления площади поверхности, доступной для сольватации.<sup>160</sup>

Альтернативным рассмотренным методом является более простой (хотя и менее точный) метод линейной энергии взаимодействия (Linear Interaction Energy (LIE)),<sup>161</sup> основанный на оценке свободной энергии образования комплекса белок–лиганд при помощи эмпирической формулы с подстроеными параметрами, учитывающими изменения электростатической и леннард-джонсовской составляющих внутренней энергии при комплексообразовании. Необходимые компоненты энергии вычисляют при помощи молекулярно-динамического моделирования.

Полуэмпирические методы вполне могут быть применены к небольшим выборкам, включающим несколько десятков молекул; их успешно применяют в исследованиях, направленных на поиск новых лекарственных препаратов. Тем не менее из-за необходимости проведения молекулярно-динамических расчетов полуэмпирические методы все еще остаются достаточно трудоемкими, а поэтому непригодны для проведения скрининга больших баз данных по свойствам органических соединений.

Для очень быстрой полуколичественной оценки свободной энергии взаимодействия рецептор–лиганд либо просто для ранжирования лигандов по средству к рецептору используют эмпирические оценочные функции.<sup>162–165</sup> Для их вычисления необходимо знать только пространственное строение комплекса белок–лиганд, которое может быть найдено в результате докинга. Поэтому оценочные функции обычно рассчитывают сразу же по завершении процедуры докинга. В настоящее время различают три категории оценочных функций: молекулярно-механические, физико-химические и основанные на знаниях.

Молекулярно-механические оценочные функции могут быть использованы как для нахождения в процессе докинга правильного положения лиганда внутри полости рецептора, так и для приближенной оценки свободной энергии комплексообразования. Они представляют собой молекулярно-механические силовые поля, дополненные средствами оценки энергии сольватации и конформационной энтропии. Примером могут служить оценочные функции программ AUTODOCK,<sup>123</sup> GOLD (функция GoldScore)<sup>136, 137</sup> и программного комплекса Sybyl (функция D-Score) (www.tripos.com).

Физико-химические (эмпирические) оценочные функции состоят из вкладов, имеющих четкую природу, — вклады, ответственные за гидрофобное взаимодействие, образование водородных связей, солевых мостиков, замораживание вращений вокруг связей при образовании комплексов и др. Коэффициенты при этих вкладах обычно подбирают при помощи множественного линейного регрессионного анализа таким образом, чтобы воспроизводились экспериментальные данные по свободной энергии взаимодействия для комплексов белок–лиганд с известной пространственной структурой. Примерами могут служить оценочные функции Score1,<sup>166</sup> Score2,<sup>167</sup> Score,<sup>168</sup> X-Score,<sup>169, 170</sup> F-Score из программного комплекса Sybyl, а также функции ChemScore,<sup>171, 172</sup> LigScore,<sup>173</sup> PLP и LUDI из программного комплекса Cerius<sup>2</sup> (<http://accelrys.com/products/cerius2/>), HINT.<sup>174</sup>

Оценочные функции, основанные на знаниях, базируются на тех же принципах, что и упомянутые выше статистические потенциалы. В их основе лежит подсчет числа контактов между атомами лиганда и белка, причем вклад каждого типа контакта вычисляется по относительной частоте его встречаемости в базе данных, содержащей известные пространственные структуры комплексов белок–лиганд, после перевода в энергетическую шкалу при помощи соотношения Больцмана. Полученные значения чисел контактов могут быть использованы для ранжирования лигандов по средству к белку, а также в семействах близких белков скоррелированы со значениями свободной энергии связывания белок–лиганд. Примерами могут служить оценочные функции PMF,<sup>175, 176</sup> DrugScore,<sup>177</sup> SmoG,<sup>178–180</sup> VHB,<sup>181</sup> ITScore.<sup>182, 183</sup>

Несмотря на большое число разработанных оценочных функций, ни одна из них не зарекомендовала себя как надежный инструмент для оценки констант связывания белок–лиганд или хотя бы для ранжирования лигандов по средству к белку. Поэтому в последнее время все большее внимание привлекает уже многократно доказавший свою эффективность подход, основанный на поиске «консенсуса» предсказаний, которые дают различные оценочные функции.<sup>184–187</sup> Разработано несколько стратегий поиска «консенсуса» для оценочных функций.<sup>187</sup> Кроме того, продемонстрирован значительный эффект, достигаемый при комбинировании оценочных функций при помощи методов статистического обучения с использованием экспериментальных данных.<sup>188–190</sup> Подобные идеи легли и в основу ряда гибридных оценочных функций, таких как CSCORE<sup>185</sup> и GFScore.<sup>191</sup>

## 6. Молекулярный дизайн новых лекарственных препаратов

Ценность построенных моделей рецепторов ФАВ для медицинской химии заключается в том, что с их помощью можно проводить молекулярный дизайн новых лекарственных препаратов, основываясь на знании пространственных структур биомишеней.<sup>104, 192–202</sup> Для этого проводят докинг лигандов в построенную молекулярную модель рецептора и отбирают лиганды, для которых предсказано наибольшее сродство к рецептору. В этом случае в зависимости от источника лигандов говорят либо о виртуальном скрининге, либо о дизайне *de novo*.

### а. Виртуальный скрининг

В практике медицинской химии при виртуальном скрининге проводится «просеивание» больших баз данных органических соединений через набор вычислительных «сит» с целью отбора потенциальных лекарственных препаратов.<sup>203–214</sup> Обычно используемые базы данных содержат либо каталоги коммерчески доступных органических соединений, либо виртуальные библиотеки соединений, которые можно легко синтезировать, например, методами комбинаторной химии.<sup>215</sup> В случае виртуального скрининга,<sup>204, 205, 208</sup> основанного на знании пространственного строения биомишеней, в качестве основного «сита» используют процедуру докинга в молекулярную модель рецептора в сочетании с применением оценочных функций для отбора лигандов с наибольшим сродством к биомишени. Как правило, при этом используют и дополнительные «сита», позволяющие на максимально ранней стадии отбросить соединения с неудовлетворительными фармакокинетическими свойствами и потенциальные



токсиканты, а также структуры, не содержащие фармакофоров, необходимых для связывания с нужной биомшенью.

#### б. Дизайн *de novo*

В отличие от виртуального скрининга, при котором используют базы данных известных лигандов, с помощью методов дизайна *de novo*<sup>216–219</sup> генерируют новые структуры лигандов и оценивают их свойства в момент генерирования. Построение новых структур лигандов осуществляется внутри молекулярной модели сайта связывания рецептора.<sup>220</sup> В большинстве случаев вычислительная процедура методов дизайна *de novo* основана на применении генетического алгоритма<sup>220</sup> (либо метода «роя частиц»<sup>221</sup>), в рамках которого осуществляется комбинаторная оптимизация структур органических соединений, приводящая к таким наборам лигандов, для которых оценочные функции прогнозируют наибольшее сродство к рецептору.

Очевидным недостатком виртуального скрининга при разработке новых лекарственных препаратов по сравнению с методом дизайна *de novo* является то, что используемые при скрининге библиотеки химических соединений составляют очень малую часть всего разнообразия химических структур. Поэтому с помощью методов дизайна *de novo* в принципе можно построить отсутствующие в базах данных структуры лигандов со значительно лучшими целевыми характеристиками. Существенным недостатком дизайна *de novo* является то, что генерируемые в рамках данного подхода химические структуры могут оказаться синтетически труднодоступными. Очевидно, что его дальнейшее развитие должно быть связано с интеграцией с методами компьютерного синтеза и методами автоматической оценки синтетической доступности органических соединений.<sup>222</sup> Наиболее известными компьютерными программами, с помощью которых можно проводить дизайн *de novo* новых лекарственных препаратов, являются LigBuilder,<sup>223</sup> LUDI,<sup>224–226</sup> модуль LeapFrog в составе программного комплекса Sybyl, ADAPT,<sup>220</sup> SPROUT (в сочетании с программами HIPPO для анализа сайта связывания рецептора и CAESA для оценки синтетической доступности генерируемых лигандов),<sup>227–229</sup> PRO\_LIGAND,<sup>230, 231</sup> AlleGrow (<http://bostondenovo.com/Allegrow.htm>).

### III. Молекулярное моделирование различных типов биологических рецепторов

Рецепторы ФАВ — сложные структуры, иногда состоящие из нескольких доменов и нескольких субъединиц, которые к тому же претерпевают глобальные конформационные переходы между несколькими пространственными формами, соответствующими состояниям покоя, активации, а в некоторых случаях десенсибилизации. Это лишь одна из многих причин (можно упомянуть также множественность сайтов связывания, аллостерической модуляции и др.), из-за которых типы воздействия небольших органических молекул на функционирование рецепторов отличаются большим разнообразием. В частности, молекулы лекарственных веществ могут выступать в качестве агонистов (полных, частичных, нейтральных и обратных), коагонистов, антагонистов (конкурентных и неконкурентных), модуляторов (отрицательных и положительных) и т.д. Вследствие этого молекулярное моделирование разных типов биологических рецепторов и их комплексов с органическими лигандами имеет свои особенности, краткое изложение которых и составляет основное

содержание данной части обзора. Мы остановимся лишь на некоторых типах рецепторов ФАВ, в практическом осуществлении моделирования которых у авторов данной публикации имеется опыт.

#### 1. Рецепторы, сопряженные с G-белками

Рецепторы, сопряженные с G-белками, являются биологическими мишенями для ~40% современных лекарственных препаратов.<sup>232</sup> В функциональном плане этот тип рецепторов объединяет следующее: все они являются белками, встроенными в клеточную мембрану, которые передают сигнал внутрь клетки за счет того, что связываются с молекулами агонистов, «выловленными» из межклеточной среды, претерпевают при связывании с ними конформационную перестройку, приводящую к взаимодействию с молекулой G-белка и формированию внутриклеточного ответа. Несмотря на чрезвычайно высокую фармакологическую важность этого типа рецепторов, лишь в 2007–2008 гг. удалось при помощи рентгеноструктурного анализа изучить пространственное строение всего лишь двух из более тысячи его представителей —  $\beta_2$ -адренергического рецептора<sup>233–235</sup> и  $A_{2A}$ -аденозинового рецептора в комплексе с антагонистом.<sup>236</sup> Именно поэтому методы компьютерного моделирования до сих пор играли и в обозримом будущем будут играть ведущую роль как в изучении пространственного строения и функционирования рецепторов, сопряженных с G-белками, так и в дизайне новых лекарственных препаратов, ориентированных на эти биомишени.<sup>237–239</sup>

В структурном плане все рецепторы, сопряженные с G-белками, роднит то, что их трансмембранные части устроены одинаково — представляют собой пучок из пронизывающих мембрану семи  $\alpha$ -спиралей, соединенных между собой тремя внутриклеточными и тремя внеклеточными петлями. Вследствие этого авторы ранних работ по сравнительному моделированию таких рецепторов использовали в качестве шаблона молекулу бактериородопсина, пространственная структура которого была первоначально установлена методом дифракции электронов,<sup>240, 241</sup> а затем рентгеноструктурного анализа.<sup>242</sup> Следует отметить, что первоначально моделированию подвергались только трансмембранные  $\alpha$ -спирали без соединяющих их петель.

В 2000 г. была расшифрована кристаллическая структура родопсина быка<sup>243</sup> — сопряженного с G-белком рецептора, управляемого светом. Строение и функционирование рецепторов, сопряженных с G-белками, значительно ближе к строению родопсина млекопитающих, чем к строению бактериородопсина, поэтому именно структуру родопсина стали использовать в качестве шаблона при сравнительном моделировании. Этим обусловлены существенно более высокая точность молекулярных моделей рецепторов, а также возможность моделирования большинства петель, что и определило использование построенных моделей для направленного поиска новых лекарственных препаратов. В 2007 г. благодаря расшифровке кристаллической структуры  $\beta_2$ -адренергического рецептора<sup>233–235</sup> удалось проверить точность построенных таким образом моделей рецепторов. Оказалось, что трансмембранная область, составляющая большую часть сайта связывания лигандов, моделируется практически идеально, однако имеются существенные ошибки в моделировании второй внеклеточной петли, которая прикрывает сайт связывания лигандов.

Несмотря на общность пространственного строения трансмембранной части рецепторов, сопряженных с G-бел-

ками, имеются существенные различия между шестью классами рецепторов данного типа. Различия касаются прежде всего аминокислотных последовательностей, способов функционирования рецепторов и строения внеклеточных доменов. Вследствие этого каждый из этих классов требует отдельного подхода к моделированию.

#### а. Первый класс рецепторов, сопряженных с G-белками

Первый класс рецепторов, сопряженных с G-белками, составляют белки, подобные родопсину. Их объединяет общий механизм активации и передачи сигнала.<sup>244</sup> Благодаря этому каждый трансмембранный сегмент рецепторов данного класса содержит один или несколько полностью консервативных внутри класса аминокислотных остатков, наличие которых необходимо для реализации механизма. Такие остатки могут быть использованы для построения выравнивания аминокислотных последовательностей. Сочетание построенного выравнивания с использованием структуры бычьего родопсина в качестве шаблона лежит в основе большинства последних работ по молекулярному моделированию первого класса рецепторов, сопряженных с G-белками. В качестве примеров можно привести построение пространственных моделей аденозиновых (рис. 1),<sup>86, 245 – 249</sup> пуриновых<sup>250, 251</sup> и мелатониновых<sup>252, 253</sup> рецепторов, а также их комплексов с лигандами.

Моделирование рецепторов, подобных родопсину, с использованием структуры последнего в качестве шаблона имеет много ограничений. Прежде всего, имеющиеся в кристалле молекулы родопсина находятся в состоянии покоя (т.е. не взаимодействуют с G-белками и не инициируют тем самым внутриклеточный ответ), поэтому в результате сравнительного моделирования получается конформация молекулы рецептора, также находящегося в состоянии покоя. Построенная таким образом молекулярная модель соответ-

ствует конформации рецептора, с которой стабилизирующие обратные агонисты образуют устойчивые комплексы. По кривой доза – ответ обратные агонисты обычно идентифицируются как конкурентные антагонисты, поэтому построенные модели могут быть использованы только для изучения связывания рецептора с антагонистами, и на основе этих моделей можно проводить направленный поиск только антагонистов к данному рецептору.

Если же в задачу исследователя входит разработка агонистов (которые образуют стабильные комплексы с активированной формой рецептора), то обычно применяют один из следующих подходов. Во-первых, можно не рассматривать в явном виде конформационную подвижность рецептора и моделировать связывание агонистов и антагонистов с одной и той же пространственной моделью, фиксируя различия между агонистами и антагонистами только по тому, с какими участками в сайте связывания они контактируют. На практике, несмотря на ошибочность лежащего в основе этого подхода предположения о наличии единой конформации рецептора для связывания со всеми типами лигандов, он неплохо работает. Для объяснения данного факта часто приводится модель, согласно которой агонисты сначала связываются с неактивной формой рецептора и лишь потом, взаимодействуя с определенными участками сайта связывания, инициируют конформационную перестройку молекулы рецептора, ведущую к его активации. Еще одним объяснением может служить то, что при активации рецептора структура сайта связывания лигандов меняется относительно слабо, поэтому должна существовать сильная корреляция между сродством лиганда к обеим формам рецептора. И тем не менее для решения ряда задач такой подход оказался неприемлемым,<sup>254</sup> поскольку с его помощью трудно различить потенциальные агонисты в ходе автоматического виртуального скрининга либо дизайна *de novo*.

Второй, более новый подход состоит в построении отдельных пространственных моделей для рецептора в состоянии покоя и для его активированной формы. Поскольку до недавнего времени о детальном механизме активации рецепторов, подобных родопсину, было известно очень мало, то исследователи получали пространственную модель агонист-связывающей активированной формы рецептора, исходя из его основной формы, путем вращения и сдвига некоторых трансмембранных  $\alpha$ -спиралей с учетом представлений о возможном механизме активации, при этом добивались согласования с определенными косвенными экспериментальными данными. В частности, таким образом были построены<sup>255</sup> модели активной формы рецепторов, подобных родопсину, которые были использованы для осуществления докинга агонистов в процессе виртуального скрининга<sup>255</sup> и при реализации процедуры «обратного» скрининга, позволяющего идентифицировать возможные биологические мишени для лигандов.<sup>256</sup>

Третий подход, характерный для последних работ, состоит в построении молекулярной модели метародопсина II (активированной формы родопсина), которую можно использовать в качестве шаблона для сравнительного моделирования агонист-связывающей активированной формы рецепторов, подобных родопсину. Впервые это было осуществлено в работах Никифоровича с соавт.<sup>257–259</sup> путем поиска оптимального пространственного расположения трансмембранных сегментов, соответствующего экспериментальным данным. Эта же цель была достигнута в работе Тихоновой с соавт.<sup>260</sup> путем компьютерного моделирования процесса активации родопсина при помощи специальных

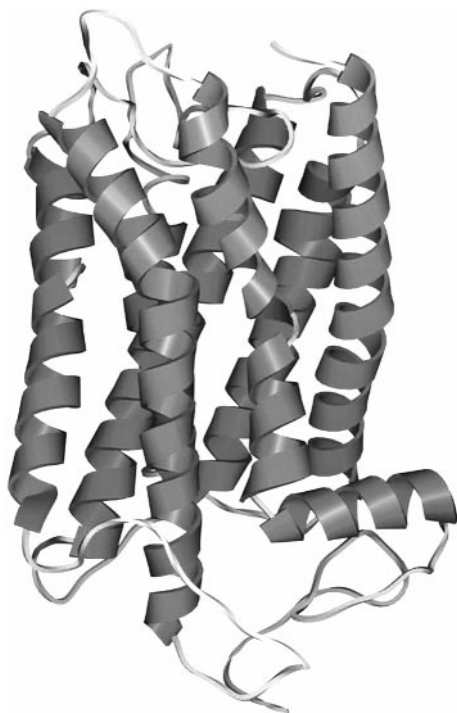


Рис. 1. Молекулярная модель A<sub>2B</sub>-аденозинового рецептора.

модификаций метода молекулярной динамики в сочетании с использованием экспериментальных данных по определенным межатомным расстояниям. Разновидностью данного подхода является получение модели активированной формы рецептора из неактивной при помощи моделирования процесса активации рецептора с использованием метода направленной молекулярной динамики.<sup>261</sup>

Наконец, четвертый подход мог бы состоять в использовании в качестве шаблона пространственной структуры метародопсина II, установленной сначала при помощи спектроскопии ЯМР,<sup>262</sup> а затем методом рентгеноструктурного анализа.<sup>263</sup> Однако модели получаются недостаточно точными.

Серьезной проблемой в моделировании рецепторов, подобных родопсину, является прогнозирование пространственного строения внеклеточных петель, непосредственно взаимодействующих с объемными пептидными лигандами. Приемлемые молекулярные модели удается построить только в сочетании с многочисленными экспериментальными данными по сайт-направленному мутагенезу и с данными спектроскопии ЯМР (см., например, работу<sup>45</sup>).

#### б. Второй класс рецепторов, сопряженных с G-белками

В отличие от первого класса рецепторов, сопряженных с G-белками, для которых лигандами являются молекулы относительно небольших размеров, с сайтом связывания внутри трансмембранной части, лигандами второго класса рецепторов являются пептидные гормоны большого размера, к типичным представителям которых относятся секретин и паратиреоидный гормон.<sup>264</sup> Кроме трансмембранной части рецепторы этого класса содержат еще и N-концевой домен, также принимающий участие в связывании лигандов. Пространственные модели трансмембранной части этих рецепторов были формально построены методом сравнительного моделирования с использованием родопсина быка в качестве шаблона,<sup>256</sup> однако корректность этих моделей не подтверждена экспериментальными данными. Аминокислотная последовательность N-концевого домена секретинного рецептора не похожа ни на одну последовательность никакого другого белка с известным пространственным строением, поэтому его молекулярная модель была построена<sup>265</sup> методом *ab initio* с привлечением экспериментальных данных.

#### в. Третий класс рецепторов, сопряженных с G-белками

К третьему классу рецепторов, сопряженных с G-белками, относятся метаболитные глутаматные рецепторы,<sup>266, 267</sup> а также рецептор ГАМК<sub>B</sub> (ГАМК –  $\gamma$ -аминомасляная кислота). Они представляют собой димеры, каждый мономер которых состоит из трех доменов: N-концевого, трансмембранного и «богатого цистеином». Все агонисты и конкурентные антагонисты связываются с N-концевым доменом (который иногда уподобляют растению росичке), поскольку он состоит из двух половинок глобулы, захлопывающихся, как только в пространство между ними попадает молекула агониста. В свою очередь это вызывает глобальную конформационную перестройку всего димера, ведущую к передаче сигнала внутрь клетки. Неконкурентные антагонисты связываются с трансмембранным доменом.

**Моделирование N-концевого домена.** Первые работы по теоретическому изучению пространственного строения открытой и закрытой форм N-концевого домена подтипа

mGluR<sub>1</sub> и его взаимодействия с лигандами проводили методом сравнительного моделирования с использованием в качестве шаблонов бактериальных периплазматических белков.<sup>268, 269</sup> На основе результатов моделирования была сформулирована гипотеза о том, что агонисты при связывании с белком лежат плашмя, инициируя захлопывание половинок глобулы, а антагонисты становятся поперек, препятствуя этому. В дальнейшем модель открытой формы (рис. 2) была существенно уточнена благодаря применению техники множественного выравнивания аминокислотных последовательностей,<sup>270</sup> а модель закрытой формы — благодаря использованию метода «протягивания».<sup>271</sup> На основе построенной модели открытой формы N-концевого домена в статье<sup>270</sup> была выдвинута гипотеза о важности остатка Arg78 для связывания лигандов в подтипе mGluR<sub>1</sub>. Эта гипотеза сразу же была подтверждена экспериментально.<sup>272</sup> Вместе с тем с учетом этой же модели была поставлена под сомнение упомянутая выше гипотеза о способе связывания антагонистов.<sup>273</sup> Основные результаты моделирования в дальнейшем подтверждены результатами рентгеноструктурного анализа N-концевого домена mGluR<sub>1</sub>.<sup>274</sup> Далее, используя модель mGluR<sub>1</sub> в качестве шаблона, были построены модели N-концевых доменов для всех подтипов метаболитных глутаматных рецепторов,<sup>275</sup> что дало возможность провести сравнительный анализ их сайтов связывания.

**Моделирование трансмембранного домена.** Лиганды трансмембранного домена метаболитного глутаматного рецептора являются его очень активными и селективными неконкурентными антагонистами. Строение этого домена определено при помощи сравнительного моделирования с использованием пространственной структуры родопсина в качестве шаблона.<sup>276–278</sup> Построенные модели применяли для поиска новых неконкурентных антагонистов при помощи процедуры виртуального скрининга.<sup>277</sup>

**Моделирование домена, «богатого цистеином».** Пространственные модели домена, «богатого цистеином», метаболитных глутаматных рецепторов впервые были построены авторами работы<sup>279</sup> при помощи сравнительного моделирования с использованием выравнивания аминокислотных

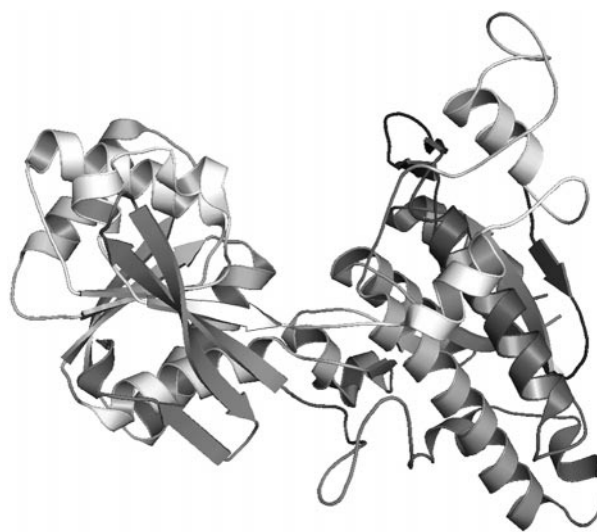


Рис. 2. Модель N-концевого домена mGluR<sub>1</sub> в открытой форме.

последовательностей, полученного по методу «протягивания».

### г. Frizzled-рецепторы

Frizzled-рецепторы (называемые также FZD-рецепторами) в последнее время стали относить (наряду со Smoothend-рецепторами) к шестому классу рецепторов,<sup>¶</sup> сопряженных с G-белками.<sup>280, 281</sup> На взаимодействии FZD-рецепторов с Wnt-белками, выступающими в роли их лигандов, основан сигнальный путь Wnt – FZD, играющий важную роль в процессах эмбрионального развития и функционирования организма. В то же время нарушения в этом сигнальном пути характерны при появлении злокачественных опухолей.<sup>282</sup> Поэтому сайт взаимодействия между FZD-рецепторами и Wnt-белками — одна из мишеней, на которую может быть направлено действие препаратов для лечения онкологических заболеваний.

В структуре FZD-рецепторов, кроме состоящего из семи  $\alpha$ -спиралей трансмембранного домена, содержатся также внеклеточный CRD-домен, «богатый цистеином», который и участвует во взаимодействии с Wnt-белками.<sup>283</sup> Пространственное строение этого домена изучено при помощи рентгеноструктурного анализа.<sup>283</sup> Из-за низкой растворимости Wnt-белков, взаимодействия Wnt – FZD экспериментально не изучены. Поэтому было предпринято молекулярное моделирование Wnt-белков и их комплексов с FZD-рецепторами.<sup>284–286</sup> Поскольку оказалось невозможным найти подходящий шаблонный белок для сравнительного моделирования, пространственная структура Wnt-белка была построена методом *ab initio* с использованием программы Rosetta.<sup>76–80</sup> Для осуществления белок-белкового докинга использовали сервер PatchDock.<sup>150</sup> Полученную молекулярную модель комплекса Wnt – FZD применяли как для изучения взаимодействия данных белков,<sup>284</sup> так и для дизайна непептидных ингибиторов этого взаимодействия.<sup>285</sup>

## 2. Ионотропные глутаматные рецепторы

Ионотропные глутаматные рецепторы делятся на три типа по названию избирательных агонистов, которыми являются 2-амино-3-(3-гидрокси-5-метилизоксазол-4-ил)пропионовая кислота (AMPA-рецепторы), *N*-метил-D-аспарагиновая кислота (NMDA-рецепторы) и каиновая кислота (каинатные рецепторы).<sup>4</sup> Эти рецепторы представляют собой сложные супрамолекулярные комплексы, состоящие из нескольких субъединиц, которые образуют ионный канал.<sup>266, 287, 288</sup> Прохождение ионного тока через этот канал и приводит к внутриклеточному ответу. Каждая субъединица глутаматных рецепторов состоит из нескольких доменов, каждый из которых обладает одним или несколькими центрами связывания фармакологически важных лигандов.<sup>266, 287, 288</sup>

### а. Моделирование внеклеточных агонист(коагонист)-связывающих доменов

Внеклеточные агонист(коагонист)-связывающие домены содержат сайты связывания природного агониста (глутамата) либо (в случае NMDA-рецептора) природного коагониста (глицина или D-серина).<sup>266, 287, 288</sup> Глутамат- и глицин-

связывающие сайты NMDA-рецептора имеют сходное пространственное строение и расположены на разных субъединицах рецептора. При одновременном связывании агонистов и коагонистов происходят конформационные превращения, приводящие к активации рецептора, открытию ионного канала и проникновению ионов металла внутрь клетки. Конкурентные антагонисты, связываясь с этими же сайтами, наоборот, препятствуют активации рецептора.

Первые представления о пространственном строении и механизме активации внеклеточных агонист-связывающих доменов ионотропных глутаматных рецепторов были получены на основании обнаруженной гомологии с последовательностями аминокислот некоторых бактериальных периплазматических связывающих белков.<sup>289</sup> Первая попытка построения молекулярных моделей глутаматного и глицинового сайтов NMDA-рецептора при помощи сравнительного моделирования с использованием периплазматического белка LAOBP в качестве шаблона предпринята авторами работы.<sup>290</sup> В 2000 г. были опубликованы результаты рентгеноструктурного анализа искусственного белка, сконструированного на основе агонист-связывающего домена AMPA-рецептора в комплексах с некоторыми его лигандами.<sup>291</sup> Это дало возможность, с учетом относительно высокой гомологии между аминокислотными последовательностями AMPA- и NMDA-рецепторов в рассматриваемом домене, построить существенно более качественные молекулярные модели глутаматного (рис. 3)<sup>292, 293</sup> и глицинового (рис. 4)<sup>294, 295</sup> сайтов связывания NMDA-рецептора. Полученные модели были использованы для объяснения известных закономерностей структура – активность для лигандов глутаматного сайта,<sup>293</sup> при построении количественных моделей связывания лигандов с глутаматным сайтом AMPA-рецептора<sup>296</sup> и с глициновым сайтом NMDA-рецептора,<sup>295</sup> для разработки концепции «полей селективности»,<sup>297</sup> а также для проведения виртуального скрининга и создания сфокусированных библиотек потенциальных лигандов NMDA- и AMPA-рецепторов.<sup>298</sup>

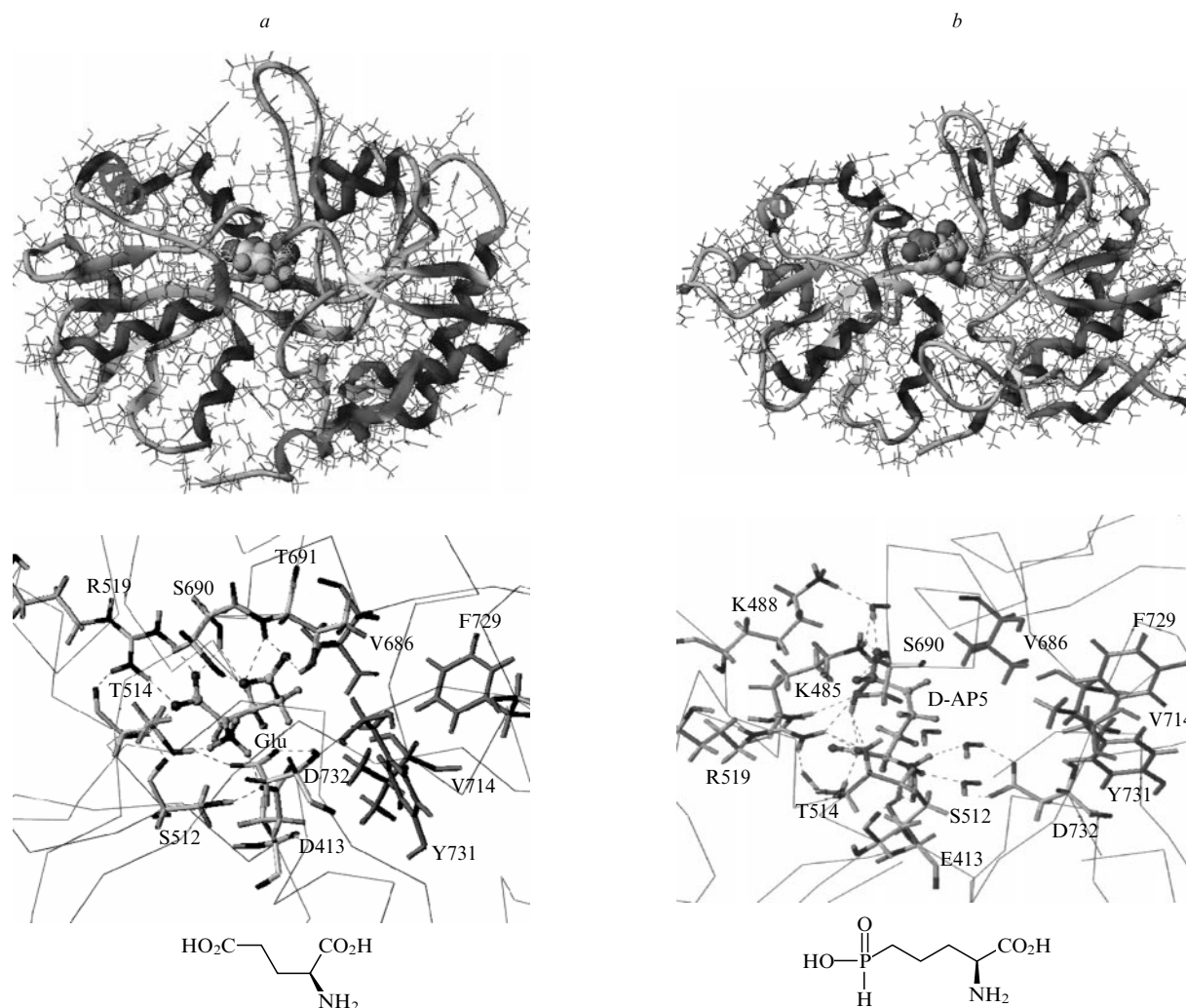
### б. Моделирование ионного канала NMDA-рецептора

Молекулярное моделирование ионного канала ионотропных глутаматных рецепторов основано на общности их архитектуры с калиевыми ионными каналами бактерий.<sup>299</sup> Построена пространственная модель ионного канала NMDA-рецептора (рис. 5)<sup>300</sup> по методу сравнительного моделирования с использованием структуры калиевого канала KcsA<sup>301</sup> в качестве шаблона. Полученную модель канала использовали для изучения взаимодействия с неконкурентным антагонистом (блокатором канала) МК-801 и для дизайна его гибких аналогов.<sup>300</sup> В дальнейшем эта модель была уточнена и применена для количественного и качественного моделирования связывания блокаторов канала.<sup>302</sup>

### в. Моделирование N-концевых модуляторных доменов

Особенность ионотропных глутаматных рецепторов заключается в наличии множества сайтов связывания в случае молекул, выполняющих модуляторную функцию (как положительную, так и отрицательную) по отношению к NMDA-рецептору. Для значительной части фармакологически важных органических соединений, действующих на центральную нервную систему, сайты связывания расположены в N-концевых доменах NMDA-рецептора. Именно этим обусловлена актуальность построения пространственных моделей доме-

¶ Литературные данные о моделях рецепторов четвертого и пятого классов пока отсутствуют.



**Рис. 3.** Комплекс агониста (глутамата) с закрытой формой (а) и антагониста (AP5) с открытой формой (b) глутамат-связывающего домена NMDA-рецептора.

Приведены вид глобулы с лигандом (вверху) и сайт связывания (внизу).

нов; такие модели могли бы способствовать разработке новых лекарственных препаратов для лечения нейродегенеративных заболеваний. Первые молекулярные модели доменов построены в работе<sup>303</sup>. Был применен метод сравнительного моделирования с использованием пространственной структуры N-концевого домена метаботропного глутаматного рецептора mGluR<sub>1</sub>,<sup>274</sup> для предсказаний элементов вторичной структуры целевого белка использовали программу PSI-Pred,<sup>69</sup> а метод «протягивания» — для выравнивания аминокислотных последовательностей. На основе полученных моделей доменов были выдвинуты предположения о способах связывания модуляторов NMDA-рецептора, а также построены количественные модели связывания аналогов ифенпродила.<sup>303</sup>

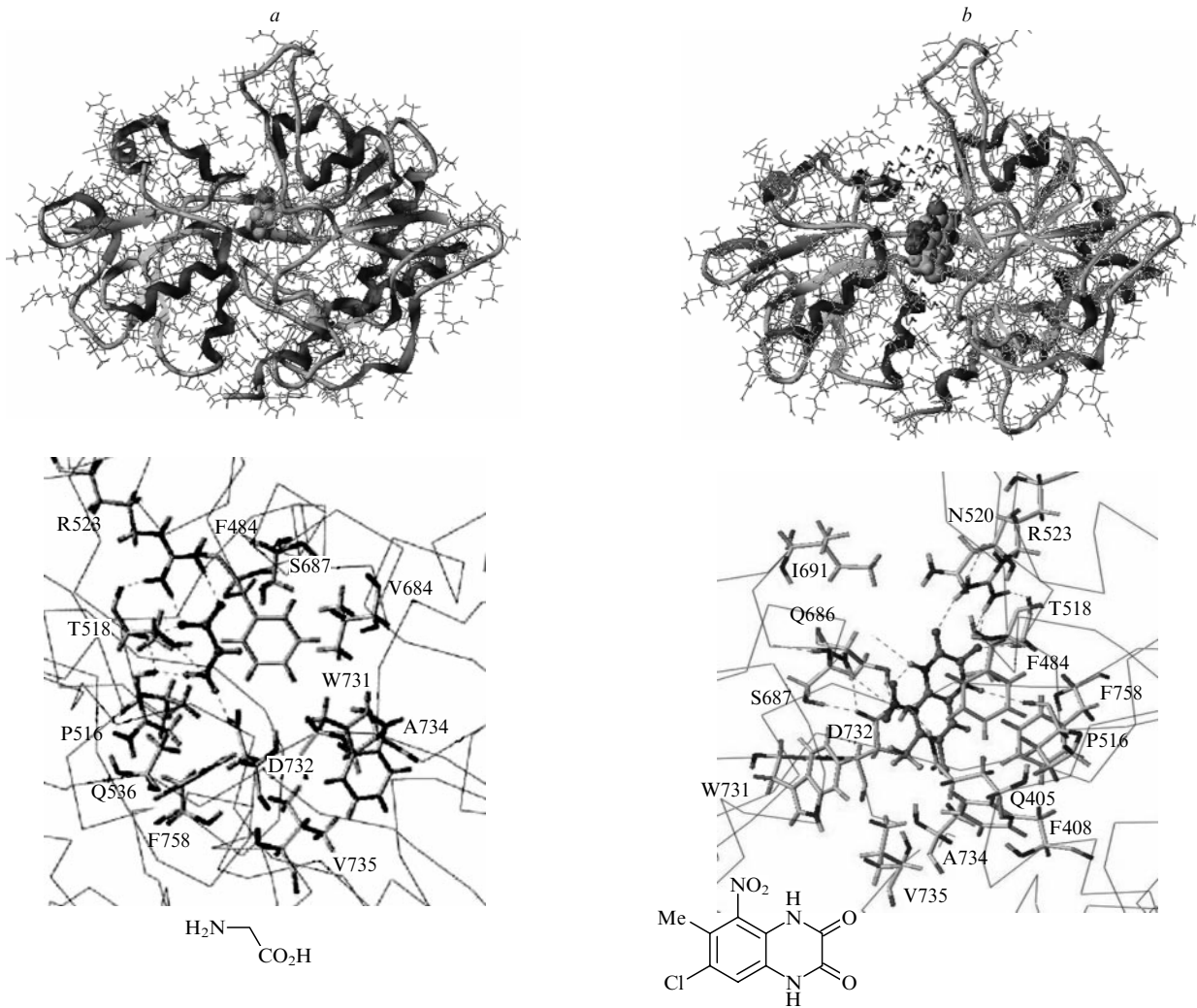
#### г. Моделирование связывания положительных аллостерических модуляторов с AMPA-рецептором

В последние годы при разработке лекарственных препаратов особое внимание уделяется положительным аллостерическим модуляторам AMPA-рецептора.<sup>304,305</sup> Методом рентгеноструктурного анализа было установлено, что сайт их связывания расположен между двумя субъединицами

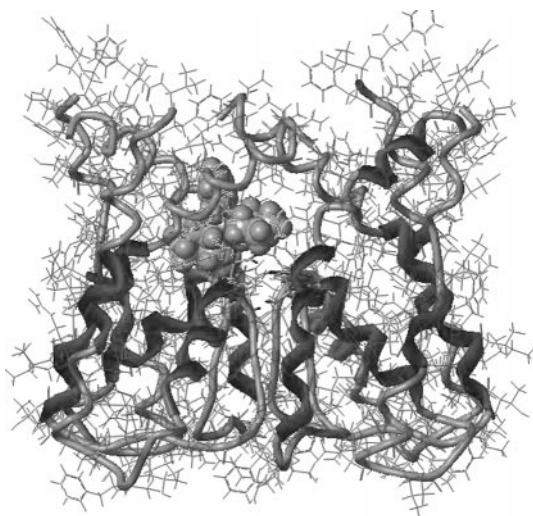
AMPA-рецептора. Механизм действия положительных модуляторов, по-видимому, заключается в том, что они затормаживают взаимное вращение субъединиц, необходимое для десенситилизации рецептора (для закрытия ионного канала, предварительно открытого после связывания глутамата) и тем самым могут увеличивать величину и продолжительность ионного потока через канал.<sup>306</sup> Найденный сайт связывания был использован для моделирования взаимодействия положительных модуляторов с AMPA-рецептором и для дизайна новых препаратов.<sup>307,308</sup>

### 3. Рецепторы, сходные по пространственному строению с никотиновым рецептором

Никотиновый рецептор и подтипы А и С рецептора  $\gamma$ -аминоназойной кислоты (ГАМК<sub>A</sub> и ГАМК<sub>C</sub>), третий подтип серотониновых рецепторов (5-HT<sub>3</sub>), а также глициновый рецептор относятся к классу рецепторов, сходных по пространственному строению с никотиновым рецептором; иногда их называют никотиноидными или цистеин-петельными. Они представляют собой интегрированные в клеточную мембрану пентамерные супрамолекулярные структуры, формирующие ионный канал. Взаимодействие с молекулами



**Рис. 4.** Комплекс агониста (глицина) с закрытой формой (а) и антагониста (6-метил-5-нитро-7-хлорхиноксалин-2,3-диона) с открытой формой (б) глицин-связывающего домена NMDA-рецептора. Приведены вид глобулы с лигандом (вверху) и сайт связывания (внизу).



**Рис. 5.** Модель ионного канала NMDA-рецептора с блокатором фенциклидином.



**Рис. 6.** Модель ГАМК<sub>A</sub>-рецептора в открытой форме.

агонистов приводит к изменению конформации рецептора и открытию ионного канала. Большое значение для исследования пространственной структуры представителей такого семейства рецепторов имеют результаты рентгеноструктурного анализа ацетилхолин-связывающего белка<sup>309</sup> и его комплексов с лигандами.<sup>310,311</sup> С использованием этих структур в качестве шаблонов для методов сравнительного моделирования построены молекулярные модели рецепторов ГАМК<sub>A</sub> (см.<sup>312</sup>) (рис. 6) и ГАМК<sub>C</sub>,<sup>313,314</sup> осуществлен теоретический анализ взаимодействия рецепторов с молекулами агонистов и конкурентных антагонистов, а также предложен механизм активации рецепторов. Полученные модели могут быть использованы для поиска новых потенциально активных лигандов при помощи виртуального скрининга.<sup>313</sup>

\* \* \*

В настоящее время методы молекулярного моделирования играют важную роль в рациональном дизайне новых лекарственных препаратов. Особенно велико их значение при выборе в качестве биологической мишени для действия лекарственных средств рецепторов ФАВ. Учитывая усовершенствования методологии молекулярного моделирования белков и взаимодействий белок–лиганд, значительное увеличение вычислительных мощностей и быстрое накопление огромного количества экспериментальных данных, необходимых для совершенствования вычислительных подходов, можно ожидать, что роль молекулярного моделирования в медицинской химии будет возрастать. Именно вычислительные методы, интегрирующие молекулярное моделирование, биоинформатику, хемоинформатику и хемоинформатику, будут определять лицо медицинской химии будущего.

## Литература

- M.A.Marti-Renom, A.C.Stuart, A.Fiser, R.Sanchez, F.Melo, A.Sali. *Annu. Rev. Biophys. Biomol. Struct.*, **29**, 291 (2000)
- R.Giegerich, C.Meyer, P.Steffen. *Sci. Comput. Progr.*, **51**, 215 (2004)
- S.R.Eddy. *Nat. Biotechnol.*, **22**, 909 (2004)
- S.B.Needleman, C.D.Wunsch. *J. Mol. Biol.*, **48**, 443 (1970)
- T.F.Smith, M.S.Waterman. *J. Mol. Biol.*, **147**, 195 (1981)
- S.F.Altschul. *J. Mol. Biol.*, **219**, 555 (1991)
- M.O.Dayhoff, R.Schwartz, B.C.Orcutt. *Atlas of Protein Sequence and Structure. Vol. 5 (Suppl. 3)*. (Ed. M.O.Dayhoff). National Biomedical Research Foundation, Washington, DC, 1978. P. 345
- S.Henikoff, J.G.Henikoff. *Proc. Natl. Acad. Sci. USA*, **89**, 10915 (1992)
- S.R.Eddy. *Nat. Biotechnol.*, **22**, 1035 (2004)
- D.J.Lipman, W.R.Pearson. *Science*, **227**, 1435 (1985)
- S.F.Altschul, W.Gish, W.Miller, E.W.Myers, D.J.Lipman. *J. Mol. Biol.*, **215**, 403 (1990)
- S.F.Altschul, T.L.Madden, A.A.Schaffer, J.Zhang, Z.Zhang, W.Miller, D.J.Lipman. *Nucleic Acids Res.*, **25**, 3389 (1997)
- S.F.Altschul, D.J.Lipman. *Proc. Natl. Acad. Sci. USA*, **87**, 5509 (1990)
- D.G.Higgins. *Methods Mol. Biol.*, **25**, 307 (1994)
- J.D.Thompson, D.G.Higgins, T.J.Gibson. *Nucl. Acid. Res.*, **22**, 4673 (1994)
- M.A.Larkin, G.Blackshields, N.P.Brown, R.Chenna, P.A.McGettigan, H.McWilliam, F.Valentin, I.M.Wallace, A.Wilm, R.Lopez, J.D.Thompson, T.J.Gibson, D.G.Higgins. *Bioinformatics*, **23**, 2947 (2007)
- C.Notredame, L.Holm, D.G.Higgins. *Bioinformatics*, **14**, 407 (1998)
- C.Notredame, D.G.Higgins, J.Heringa. *J. Mol. Biol.*, **302**, 205 (2000)
- D.T.Jones, W.R.Taylor, J.M.Thornton. *Nature (London)*, **358**, 86 (1992)
- D.T.Jones, C.Hadley. *Bioinformatics: Sequence, Structure and Databanks*. (Eds D.G.Higgins, W.R.Taylor). Springer-Verlag, Heidelberg, 2000. P. 3
- J.U.Bowie, R.Luthy, D.Eisenberg. *Science*, **253**, 164 (1991)
- D.T.Jones. *Methods Mol. Biol.*, **143**, 131 (2000)
- D.T.Jones. *Curr. Opin. Struct. Biol.*, **10**, 371 (2000)
- S.H.Bryant, C.E.Lawrence. *Proteins*, **16**, 92 (1993)
- D.Baker, A.Sali. *Science*, **294**, 93 (2001)
- J.Greer. *J. Mol. Biol.*, **153**, 1027 (1981)
- B.Wallner, A.Elofsson. *Protein Sci.*, **14**, 1315 (2005)
- C.M.Topham, P.Thomas, J.P.Overington, M.S.Johnson, F.Eisenmenger, T.L.Blundell. *Biochem. Soc. Symp.*, **57**, 1 (1990)
- M.Levitt. *J. Mol. Biol.*, **226**, 507 (1992)
- A.Sali, T.L.Blundell. *J. Mol. Biol.*, **234**, 779 (1993)
- T.F.Havel, G.M.Crippen, I.D.Kuntz, J.M.Blaney. *J. Theor. Biol.*, **104**, 383 (1983)
- T.F.Havel, I.D.Kuntz, G.M.Crippen. *J. Theor. Biol.*, **104**, 359 (1983)
- A.Sali, L.Potterton, F.Yuan, H.van Vlijmen, M.Karplus. *Proteins*, **23**, 318 (1995)
- R.Sanchez, A.Sali. *Proteins*, (Suppl. 1), 50 (1997)
- R.Sanchez, A.Sali. *Methods Mol. Biol.*, **143**, 97 (2000)
- A.Fiser, A.Sali. *Methods Enzymol.*, **374**, 461 (2003)
- N.Eswar, B.Webb, M.A.Marti-Renom, M.S.Madhusudhan, D.Eramian, M.Y.Shen, U.Pieper, A.Sali. *Curr. Protoc. Protein Sci., Ch. 2, Unit 2.9* (2007)
- N.Eswar, D.Eramian, B.Webb, M.Y.Shen, A.Sali. *Methods Mol. Biol.*, **426**, 145 (2008)
- D.Chivian, D.E.Kim, L.Malmstrom, P.Bradley, T.Robertson, P.Murphy, C.E.Strauss, R.Bonneau, C.A.Rohl, D.Baker. *Proteins*, **53** (Suppl. 6), 524 (2003)
- C.A.Rohl, C.E.Strauss, D.Chivian, D.Baker. *Proteins*, **55**, 656 (2004)
- D.Chivian, D.Baker. *Nucleic Acids Res.*, **34**, e112 (2006)
- E.P.Garvey, C.R.Matthews. *Biotechnology*, **14**, 37 (1990)
- P.N.Bryan. *Methods Mol. Biol.*, **40**, 271 (1995)
- E.Archer, B.Maigret, C.Escricut, L.Pradayrol, D.Fourmy. *Trends Pharmacol. Sci.*, **24**, 36 (2003)
- E.Archer-Lahlou, I.Tikhonova, C.Escricut, M.Dufresne, C.Seva, L.Pradayrol, L.Moroder, B.Maigret, D.Fourmy. *J. Med. Chem.*, **48**, 180 (2005)
- M.J.Sippl. *Proteins*, **17**, 355 (1993)
- M.J.Sippl. *J. Comput. Aided Mol. Des.*, **7**, 473 (1993)
- K.Park, M.Vendruscolo, E.Domany. *Proteins*, **40**, 237 (2000)
- P.Narang, K.Bhushan, S.Bose, B.Jayaram. *J. Biomol. Struct. Dyn.*, **23**, 385 (2006)
- M.Y.Shen, A.Sali. *Protein Sci.*, **15**, 2507 (2006)
- D.Eramian, M.Y.Shen, D.Devos, F.Melo, A.Sali, M.A.Marti-Renom. *Protein Sci.*, **15**, 1653 (2006)
- D.Eisenberg, J.U.Bowie, R.Luthy, S.Choe. *Faraday Discuss.*, **25** (1992)
- R.Luthy, J.U.Bowie, D.Eisenberg. *Nature (London)*, **356**, 83 (1992)
- M.Wiederstein, M.J.Sippl. *Nucleic Acids Res.*, **35**, W407 (2007)
- T.Lazaridis, M.Karplus. *J. Mol. Biol.*, **288**, 477 (1999)
- T.Lazaridis, M.Karplus. *Proteins*, **35**, 133 (1999)
- J.Gasteiger, J.Zupan. *Angew. Chem., Int. Ed. Engl.*, **105**, 503 (1993)
- Н.М.Гальберштам, И.И.Баскин, В.А.Палюлин, Н.С.Зефилов. *Успехи химии*, **72**, 706 (2003)
- И.И.Баскин, В.А.Палюлин, Н.С.Зефилов. *Рос. хим. журн.*, **50**, 86 (2006)
- V.Vapnik. *Statistical Learning Theory*. Wiley-Interscience, 1998
- V.N.Vapnik. *IEEE Trans. Neural Netw.*, **10**, 988 (1999)

62. I.Ovidiu. In *Reviews in Computational Chemistry*. (Eds T.R.C.Kenny, B.Lipkowitz). 2007. P. 291
63. L.J.McGuffin, D.T.Jones. *Bioinformatics*, **19**, 874 (2003)
64. U.Pieper, N.Eswar, H.Braberg, M.S.Madhusudhan, F.P.Davis, A.C.Stuart, N.Mirkovic, A.Rossi, M.A.Marti-Renom, A.Fiser, B.Webb, D.Greenblatt, C.C.Huang, T.E.Ferrin, A.Sali. *Nucleic Acids Res.*, **32**, D217 (2004)
65. U.Pieper, N.Eswar, F.P.Davis, H.Braberg, M.S.Madhusudhan, A.Rossi, M.Marti-Renom, R.Karchin, B.M.Webb, D.Eramian, M.Y.Shen, L.Kelly, F.Melo, A.Sali. *Nucleic Acids Res.*, **34**, D291 (2006)
66. U.Pieper, N.Eswar, A.C.Stuart, V.A.Ilyin, A.Sali. *Nucleic Acids Res.*, **30**, 255 (2002)
67. R.Sanchez, U.Pieper, N.Mirkovic, P.I.de Bakker, E.Wittenstein, A.Sali. *Nucleic Acids Res.*, **28**, 250 (2000)
68. W.Kabsch, C.Sander. *Biopolymers*, **22**, 2577 (1983)
69. L.J.McGuffin, K.Bryson, D.T.Jones. *Bioinformatics*, **16**, 404 (2000)
70. N.Colloc'h, C.Etchebest, E.Thoreau, B.Henrissat, J.P.Mornon. *Protein Eng.*, **6**, 377 (1993)
71. A.Kolinski, J.Skolnick. *Proteins*, **18**, 338 (1994)
72. J.U.Bowie, D.Eisenberg. *Proc. Natl. Acad. Sci. USA*, **91**, 4436 (1994)
73. K.Yue, K.A.Dill. *Protein Sci.*, **5**, 254 (1996)
74. R.Srinivasan, G.D.Rose. *Proteins*, **22**, 81 (1995)
75. D.Chivian, T.Robertson, R.Bonneau, D.Baker. *Methods Biochem. Anal.*, **44**, 547 (2003)
76. K.T.Simons, C.Kooperberg, E.Huang, D.Baker. *J. Mol. Biol.*, **268**, 209 (1997)
77. K.T.Simons, R.Bonneau, I.Ruczinski, D.Baker. *Proteins*, (Suppl. 3), 171 (1999)
78. K.T.Simons, I.Ruczinski, C.Kooperberg, B.A.Fox, C.Bystroff, D.Baker. *Proteins*, **34**, 82 (1999)
79. R.Bonneau, J.Tsai, I.Ruczinski, D.Chivian, C.Rohl, C.E.Strauss, D.Baker. *Proteins*, (Suppl. 5), 119 (2001)
80. R.Bonneau, C.E.Strauss, C.A.Rohl, D.Chivian, P.Bradley, L.Malmstrom, T.Robertson, D.Baker. *J. Mol. Biol.*, **322**, 65 (2002)
81. Y.Fujitsuka, G.Chikenji, S.Takada. *Proteins*, **62**, 381 (2006)
82. У.Буркерт, Н.Л.Эллингджер. *Молекулярная механика*. Мир, Москва, 1986. С. 364
83. T.Schlick. *Molecular Modeling and Simulation: An Interdisciplinary Guide*. Springer-Verlag, New York, 2002
84. A.D.Mackerell Jr. *J. Comput. Chem.*, **25**, 1584 (2004)
85. J.M.Haile. *Molecular Dynamics Simulation*. Wiley, New York, 1997
86. A.A.Ivanov, I.I.Baskin, V.A.Palyulin, L.Piccagli, P.G.Baraldi, N.S.Zefirov. *J. Med. Chem.*, **48**, 6813 (2005)
87. D.A.Case, T.E.Cheatham III, T.Darden, H.Gohlke, R.Luo, K.M.Merz Jr., A.Onufriev, C.Simmerling, B.Wang, R.J.Woods. *J. Comput. Chem.*, **26**, 1668 (2005)
88. J.W.Ponder, D.A.Case. *Adv. Protein Chem.*, **66**, 27 (2003)
89. B.R.Brooks, R.E.Brucoleri, B.D.Olafson, D.J.States, S.Swaminathan, M.Karplus. *J. Comput. Chem.*, **4**, 187 (1983)
90. J.C.Phillips, R.Braun, W.Wang, J.Gumbart, E.Tajkhorshid, E.Villa, C.Chipot, R.D.Skeel, L.Kale, K.Schulten. *J. Comput. Chem.*, **26**, 1781 (2005)
91. M.Christen, P.H.Hunenberger, D.Bakowies, R.Baron, R.Burgi, D.P.Geerke, T.N.Heinz, M.A.Kastenholz, V.Krautler, C.Oostenbrink, C.Peter, D.Trzesniak, W.F.van Gunsteren. *J. Comput. Chem.*, **26**, 1719 (2005)
92. D.Van Der Spoel, E.Lindahl, B.Hess, G.Groenhof, A.E.Mark, H.J.Berendsen. *J. Comput. Chem.*, **26**, 1701 (2005)
93. T.Lengauer, M.Rarey. *Curr. Opin. Struct. Biol.*, **6**, 402 (1996)
94. I.Halperin, B.Ma, H.Wolfson, R.Nussinov. *Proteins*, **47**, 409 (2002)
95. R.D.Taylor, P.J.Jewsbury, J.W.Essex. *J. Comput. Aided Mol. Des.*, **16**, 151 (2002)
96. R.T.Kroemer. *Biochem. Soc. Trans.*, **31**, 980 (2003)
97. V.Mohan, A.C.Gibbs, M.D.Cummings, E.P.Jaeger, R.L.DesJarlais. *Curr. Pharm. Des.*, **11**, 323 (2005)
98. S.F.Sousa, P.A.Fernandes, M.J.Ramos. *Proteins*, **65**, 15 (2006)
99. D.A.Gschwend, A.C.Good, I.D.Kuntz. *J. Mol. Recognit.*, **9**, 175 (1996)
100. R.Abagyan, M.Totrov. *Curr. Opin. Chem. Biol.*, **5**, 375 (2001)
101. B.K.Shoichet, S.L.McGovern, B.Wei, J.J.Irwin. *Curr. Opin. Chem. Biol.*, **6**, 439 (2002)
102. J.C.Alvarez. *Curr. Opin. Chem. Biol.*, **8**, 365 (2004)
103. D.B.Kitchen, H.Decornez, J.R.Furr, J.Bajorath. *Nat. Rev. Drug Discov.*, **3**, 935 (2004)
104. R.T.Kroemer. *Curr. Protein Pept. Sci.*, **8**, 312 (2007)
105. M.J.Sternberg, H.A.Gabb, R.M.Jackson. *Curr. Opin. Struct. Biol.*, **8**, 250 (1998)
106. G.R.Smith, M.J.Sternberg. *Curr. Opin. Struct. Biol.*, **12**, 28 (2002)
107. J.J.Gray. *Curr. Opin. Struct. Biol.*, **16**, 183 (2006)
108. A.M.Bonvin. *Curr. Opin. Struct. Biol.*, **16**, 194 (2006)
109. D.W.Ritchie. *Curr. Protein Pept. Sci.*, **9**, 1 (2008)
110. T.Kieber-Emmons, R.Murali, M.I.Greene. *Curr. Opin. Biotechnol.*, **8**, 435 (1997)
111. M.G.Bursavich, D.H.Rich. *J. Med. Chem.*, **45**, 541 (2002)
112. J.M.Ahn, N.A.Boyle, M.T.MacDonald, K.D.Janda. *Mini Rev. Med. Chem.*, **2**, 463 (2002)
113. J.J.Perez, F.Corcho, O.Llorens. *Curr. Med. Chem.*, **9**, 2209 (2002)
114. B.M.Rajesh, J.Iqbal. *Curr. Pharm. Biotechnol.*, **7**, 247 (2006)
115. J.Vagner, H.Qu, V.J.Hruby. *Curr. Opin. Chem. Biol.*, **12**, 292 (2008)
116. S.Betzi, A.Restouin, S.Opi, S.T.Arold, I.Parrot, F.Guerlesquin, X.Morelli, Y.Collette. *Proc. Natl. Acad. Sci. USA*, **104**, 19256 (2007)
117. S.R.Stauffer. *Curr. Top. Med. Chem.*, **7**, 961 (2007)
118. W.L.Jorgensen. *Science*, **254**, 954 (1991)
119. W.Cai, X.Shao, B.Maigret. *J. Mol. Graph. Model.*, **20**, 313 (2002)
120. R.J.Morris, R.J.Najmanovich, A.Kahraman, J.M.Thornton. *Bioinformatics*, **21**, 2347 (2005)
121. A.Kahraman, R.J.Morris, R.A.Laskowski, J.M.Thornton. *J. Mol. Biol.*, **368**, 283 (2007)
122. B.K.Shoichet, I.D.Kuntz, D.L.Bodian. *J. Comput. Chem.*, **13**, 380 (1992)
123. G.M.Morris, D.S.Goodsell, R.S.Halliday, R.Huey, W.E.Hart, R.K.Belew, A.J.Olson. *J. Comput. Chem.*, **19**, 1639 (1998)
124. E.C.Meng, B.K.Shoichet, I.D.Kuntz. *J. Comput. Chem.*, **13**, 505 (1992)
125. Y.Zhao, M.F.Sanner. *J. Comput. Aided Mol. Des.*, **22**, 673 (2008)
126. S.Chaudhury, J.J.Gray. *J. Mol. Biol.*, **381**, 1068 (2008)
127. N.Kamiya, Y.Yonezawa, H.Nakamura, J.Higo. *Proteins*, **70**, 41 (2008)
128. N.Andrusier, E.Mashiach, R.Nussinov, H.J.Wolfson. *Proteins*, **73**, 271 (2008)
129. I.D.Kuntz. *Science*, **257**, 1078 (1992)
130. I.D.Kuntz, E.C.Meng, B.K.Shoichet. *Acc. Chem. Res.*, **27**, 117 (1994)
131. E.C.Meng, I.D.Kuntz. *Drug Inform. J.*, **28**, 733 (1994)
132. D.T.Moustakas, P.T.Lang, S.Pegg, E.Pettersen, I.D.Kuntz, N.Broojmans, R.C.Rizzo. *J. Comput. Aided Mol. Des.*, **20**, 601 (2006)
133. D.S.Goodsell, A.J.Olson. *Proteins*, **8**, 195 (1990)
134. D.S.Goodsell, G.M.Morris, A.J.Olson. *J. Mol. Recognit.*, **9**, 1 (1996)
135. G.M.Morris, D.S.Goodsell, R.Huey, A.J.Olson. *J. Comput. Aided Mol. Des.*, **10**, 293 (1996)
136. M.L.Verdonk, J.C.Cole, M.J.Hartshorn, C.W.Murray, R.D.Taylor. *Proteins*, **52**, 609 (2003)
137. J.C.Cole, J.W.M.Nissink, R.Taylor. *Virtual Screening in Drug Discovery*. (Eds B.Shoichet, J.Alvarez). Taylor Francis; CRC Press, Boca Raton, 2005



138. R.A.Friesner, J.L.Banks, R.B.Murphy, T.A.Halgren, J.J.Klicic, D.T.Mainz, M.P.Repasky, E.H.Knoll, M.Shelley, J.K.Perry, D.E.Shaw, P.Francis, P.S.Shenkin. *J. Med. Chem.*, **47**, 1739 (2004)
139. T.A.Halgren, R.B.Murphy, R.A.Friesner, H.S.Beard, L.L.Frye, W.T.Pollard, J.L.Banks. *J. Med. Chem.*, **47**, 1750 (2004)
140. M.P.Repasky, M.Shelley, R.A.Friesner. *Curr. Protoc. Bioinformatics*, (Ch. 8), Unit 8.12 (2007)
141. H.Claussen, M.Gastreich, V.Apelt, J.Greene, S.A.Hindle, C.Lemmen. *Curr. Drug. Discov. Technol.*, **1**, 49 (2004)
142. S.S.Cross. *J. Chem. Inf. Model.*, **45**, 993 (2005)
143. B.Kramer, M.Rarey, T.Lengauer. *Proteins*, **37**, 228 (1999)
144. I.Schellhammer, M.Rarey. *Proteins*, **57**, 504 (2004)
145. J.C.Cole, C.W.Murray, J.W.Nissink, R.D.Taylor, R.Taylor. *Proteins*, **60**, 325 (2005)
146. S.Lyskov, J.J.Gray. *Nucl. Acids Res.*, **36**, W233 (2008)
147. C.Wang, O.Schueler-Furman, I.Andre, N.London, S.J.Fleishman, P.Bradley, B.Qian, D.Baker. *Proteins*, **69**, 758 (2007)
148. R.Chen, L.Li, Z.Weng. *Proteins*, **52**, 80 (2003)
149. K.Wiehe, B.Pierce, W.W.Tong, H.Hwang, J.Mintseris, Z.Weng. *Proteins*, **69**, 719 (2007)
150. D.Schneidman-Duhovny, Y.Inbar, V.Polak, M.Shatsky, I.Halperin, H.Benyamini, A.Barzilai, O.Dror, N.Haspel, R.Nussinov, H.J.Wolfson. *Proteins*, **52**, 107 (2003)
151. M.A.McCarrick, P.Kollman. *Methods Enzymol.*, **241**, 370 (1994)
152. W.Wang, O.Donini, C.M.Reyes, P.A.Kollman. *Annu. Rev. Biophys. Biomol. Struct.*, **30**, 211 (2001)
153. H.J.Woo, B.Roux. *Proc. Natl. Acad. Sci. USA*, **102**, 6825 (2005)
154. J.Srinivasan, T.E.Cheatham III, P.Cieplak, P.A.Kollman, D.A.Case. *J. Am. Chem. Soc.*, **120**, 9401 (1998)
155. I.Massova, P.A.Kollman. *J. Am. Chem. Soc.*, **121**, 8133 (1999)
156. A.Weis, K.Katebzadeh, P.Soderhjelm, I.Nilsson, U.Ryde. *J. Med. Chem.*, **49**, 6596 (2006)
157. F.Fogolari, A.Brigo, H.Molinari. *Biophys. J.*, **85**, 159 (2003)
158. B.Honig, A.Nicholls. *Science*, **268**, 1144 (1995)
159. B.Jayaram, D.Sprou, D.L.Beveridge. *J. Phys. Chem. B*, **102**, 9571 (1998)
160. D.Sitkoff, K.A.Sharp, B.Honig. *J. Phys. Chem.*, **98**, 1978 (1994)
161. J.Aqvist, J.Marelius. *Combin. Chem. High Throughput Screen.*, **4**, 613 (2001)
162. R.Wang, Y.Lu, S.Wang. *J. Med. Chem.*, **46**, 2287 (2003)
163. R.Wang, Y.Lu, X.Fang, S.Wang. *J. Chem. Inf. Comput. Sci.*, **44**, 2114 (2004)
164. X.D.Li, T.J.Hou, X.J.Xu. *Acta Phys. Chim. Sinica*, **21**, 504 (2005)
165. A.N.Jain. *Curr. Protein Pept. Sci.*, **7**, 407 (2006)
166. H.J.Böhm. *J. Comput. Aided Mol. Des.*, **8**, 243 (1994)
167. H.J.Böhm. *J. Comput. Aided Mol. Des.*, **12**, 309 (1998)
168. R.Wang, L.Liu, L.Lai, Y.Tang. *J. Mol. Model.*, **4**, 379 (1998)
169. R.Wang, L.Lai, S.Wang. *J. Comput. Aided Mol. Des.*, **16**, 11 (2002)
170. C.Obiol-Pardo, J.Rubio-Martinez. *J. Chem. Inf. Model.*, **47**, 134 (2007)
171. M.D.Eldridge, C.W.Murray, T.R.Auton, G.V.Paolini, R.P.Mee. *J. Comput. Aided Mol. Des.*, **11**, 425 (1997)
172. C.W.Murray, T.R.Auton, M.D.Eldridge. *J. Comput. Aided Mol. Des.*, **12**, 503 (1998)
173. A.Krammer, P.D.Kirchhoff, X.Jiang, C.M.Venkatachalam, M.Waldman. *J. Mol. Graph. Model.*, **23**, 395 (2005)
174. P.Cozzini, M.Fornabaio, A.Marabotti, D.J.Abraham, G.E.Kellogg, A.Mozzarelli. *J. Med. Chem.*, **45**, 2469 (2002)
175. I.Muegge, Y.C.Martin, P.J.Hajduk, S.W.Fesik. *J. Med. Chem.*, **42**, 2498 (1999)
176. I.Muegge. *J. Med. Chem.*, **49**, 5895 (2006)
177. H.F.Velec, H.Gohlke, G.Klebe. *J. Med. Chem.*, **48**, 6296 (2005)
178. R.S.DeWitte, A.V.Ishchenko, E.I.Shakhnovich. *J. Am. Chem. Soc.*, **119**, 4608 (1997)
179. R.S.DeWitte, E.I.Shakhnovich. *J. Am. Chem. Soc.*, **118**, 11733 (1996)
180. A.V.Ishchenko, E.I.Shakhnovich. *J. Med. Chem.*, **45**, 2770 (2002)
181. M.Feher, E.Dereety, S.Roy. *J. Chem. Inf. Comput. Sci.*, **43**, 1316 (2003)
182. S.Y.Huang, X.Zou. *J. Comput. Chem.*, **27**, 1866 (2006)
183. S.Y.Huang, X.Zou. *J. Comput. Chem.*, **27**, 1876 (2006)
184. P.S.Charifson, J.J.Corkery, M.A.Murcko, W.P.Walters. *J. Med. Chem.*, **42**, 5100 (1999)
185. R.D.Clark, A.Strizhev, J.M.Leonard, J.F.Blake, J.B.Matthew. *J. Mol. Graph. Model.*, **20**, 281 (2002)
186. M.Feher. *Drug Discov. Today*, **11**, 421 (2006)
187. A.Oda, K.Tsuchida, T.Takakura, N.Yamaotsu, S.Hirono. *J. Chem. Inf. Model.*, **46**, 380 (2006)
188. A.E.Klon, M.Glick, J.W.Davies. *J. Med. Chem.*, **47**, 4356 (2004)
189. R.Teramoto, H.Fukunishi. *J. Chem. Inf. Model.*, **47**, 526 (2007)
190. S.Costanzi, I.G.Tikhonova, M.Ohno, E.J.Roh, B.V.Joshi, A.O.Colson, D.Houston, S.Maddileti, T.K.Harden, K.A.Jacobson. *J. Med. Chem.*, **50**, 3229 (2007)
191. S.Betzi, K.Suhre, B.Chetrit, F.Guerlesquin, X.Morelli. *J. Chem. Inf. Model.*, **46**, 1704 (2006)
192. J.Greer, J.W.Erickson, J.J.Baldwin, M.D.Varney. *J. Med. Chem.*, **37**, 1035 (1994)
193. P.M.Colman. *Curr. Opin. Struct. Biol.*, **4**, 868 (1994)
194. C.L.M.J.Verlinde, W.G.J.Hol. *Structure*, **2**, 577 (1994)
195. P.J.Whittle, T.L.Blundell. *Annu. Rev. Biophys. Biomol. Struct.*, **23**, 349 (1994)
196. T.L.Blundell. *Nature (London)*, **384**, 23 (1996)
197. T.J.Marrone, J.M.Briggs, J.A.McCammon. *Annu. Rev. Pharmacol. Toxicol.*, **37**, 71 (1997)
198. K.Gubernator, H.J.Böhm. *Structure-Based Ligand Design, Methods and Principles in Medicinal Chemistry*. Wiley-VCH, Weinheim, 1998
199. M.Isiguro. *Tanpakushitsu Kakusan Koso. (Protein, Nucleic Acid, Enzyme)*, **45**, 880 (2000)
200. G.Klebe. *J. Mol. Med.*, **78**, 269 (2000)
201. A.C.Anderson. *Chem. Biol.*, **10**, 787 (2003)
202. M.MDaile, C.Mait, P.A.Holt, J.M.Maguire, J.B.Meier, M.C.Miller, L.Petraccone, J.O.Trent. *Exp. Mol. Pathol.*, **86**, 141 (2009)
203. X.Barril, R.E.Hubbard, S.D.Morley. *Mini Rev. Med. Chem.*, **4**, 779 (2004)
204. S.Ghosh, A.Nie, J.An, Z.Huang. *Curr. Opin. Chem. Biol.*, **10**, 194 (2006)
205. A.Good. *Curr. Drug Discov. Develop.*, **4**, 301 (2001)
206. A.N.Jain. *Curr. Drug Discov. Develop.*, **7**, 396 (2004)
207. T.Langer, R.D.Hoffmann. *Curr. Pharm. Des.*, **7**, 509 (2001)
208. P.D.Lyne. *Drug Discov. Today*, **7**, 1047 (2002)
209. C.McInnes. *Curr. Opin. Chem. Biol.*, **11**, 494 (2007)
210. I.Muegge, S.Oloff. *Drug Discov. Today: Technolog.*, **3**, 405 (2006)
211. W.P.Walters, M.T.Stahl, M.A.Murcko. *Drug Discov. Today*, **3**, 160 (1998)
212. A.S.Reddy, S.P.Pati, P.P.Kumar, H.N.Pradeep, G.N.Sastry. *Curr. Protein Pept. Sci.*, **8**, 329 (2007)
213. G.Schneider, H.J.Böhm. *Drug Discov. Today*, **7**, 64 (2002)
214. B.Waszkowycz. *Curr. Drug Discov. Develop.*, **5**, 407 (2002)
215. D.V.S.Green. *Prog. Med. Chem.*, **41**, 61 (2003)
216. P.M.Dean. *Mol. Biotechnol.*, **37**, 237 (2007)
217. T.Honma. *Med. Res. Rev.*, **23**, 606 (2003)
218. H.Mausser, W.Guba. *Curr. Drug Discov. Develop.*, **11**, 365 (2008)
219. G.Schneider, U.Fechner. *Nat. Rev. Drug Discov.*, **4**, 649 (2005)
220. S.C.H.Pegg, J.J.Haresco, I.D.Kuntz. *J. Comput. Aided Mol. Des.*, **15**, 911 (2001)
221. M.Hartenfeller, E.Proschak, A.Schuller, G.Schneider. *Chem. Biol. Drug Des.*, **72**, 16 (2008)
222. J.Gasteiger. *J. Comput. Aided Mol. Des.*, **21**, 307 (2007)
223. R.Wang, Y.Gao, L.Lai. *J. Mol. Model.*, **6**, 498 (2000)

224. H.J.Bohm. *J. Comput. Aided Mol. Des.*, **6**, 593 (1992)
225. H.J.Bohm. *J. Comput. Aided Mol. Des.*, **6**, 61 (1992)
226. H.J.Bohm. *J. Comput. Aided Mol. Des.*, **8**, 623 (1994)
227. V.Gillet, A.P.Johnson, P.Mata, S.Sike, P.Williams. *J. Comput. Aided Mol. Des.*, **7**, 127 (1993)
228. V.J.Gillet, W.Newell, P.Mata, G.Myatt, S.Sike, Z.Zsoldos, A.P.Johnson. *J. Chem. Inf. Comput. Sci.*, **34**, 207 (1994)
229. V.J.Gillet, G.Myatt, Z.Zsoldos, A.P.Johnson. *Persp. Drug Discov. Des.*, **3**, 34 (1995)
230. D.E.Clark, D.Frenkel, S.A.Levy, J.Li, C.W.Murray, B.Robson, B.Waszkowycz, D.R.Westhead. *J. Comput. Aided Mol. Des.*, **9**, 13 (1995)
231. D.R.Westhead, D.E.Clark, D.Frenkel, J.Li, C.W.Murray, B.Robson, B.Waszkowycz. *J. Comput. Aided Mol. Des.*, **9**, 139 (1995)
232. D.Filmore. *Mod. Drug Discov.*, **7**, 24 (2004)
233. V.Cherezov, D.M.Rosenbaum, M.A.Hanson, S.G.Rasmussen, F.S.Thian, T.S.Kobilka, H.J.Choi, P.Kuhn, W.I.Weis, B.K.Kobilka, R.C.Stevens. *Science*, **318**, 1258 (2007)
234. S.G.Rasmussen, H.J.Choi, D.M.Rosenbaum, T.S.Kobilka, F.S.Thian, P.C.Edwards, M.Burghammer, V.R.Ratnala, R.Sanishvili, R.F.Fischetti, G.F.Schertler, W.I.Weis, B.K.Kobilka. *Nature (London)*, **450**, 383 (2007)
235. D.M.Rosenbaum, V.Cherezov, M.A.Hanson, S.G.Rasmussen, F.S.Thian, T.S.Kobilka, H.J.Choi, X.J.Yao, W.I.Weis, R.C.Stevens, B.K.Kobilka. *Science*, **318**, 1266 (2007)
236. V.P.Jaakola, M.T.Griffith, M.A.Hanson, V.Cherezov, E.Y.Chien, J.R.Lane, A.P.Ijzerman, R.C.Stevens. *Science*, **322**, 1211 (2008)
237. M.Ashton, M.H.Charlton, M.K.Schwarz, R.J.Thomas, M.Whittaker. *Combin. Chem. High Throughput Screen.*, **7**, 441 (2004)
238. P.Jimonet, R.Jäger. *Curr. Drug Discov. Develop.*, **7**, 325 (2004)
239. S.Radestock, T.Weil, S.Renner. *J. Chem. Inf. Model.*, **48**, 1104 (2008)
240. N.Grigorieff, T.A.Ceska, K.H.Downing, J.M.Baldwin, R.Henderson. *J. Mol. Biol.*, **259**, 393 (1996)
241. Y.Kimura, D.G.Vassilyev, A.Miyazawa, A.Kidera, M.Matsushima, K.Mitsuoka, K.Murata, T.Hirai, Y.Fujiyoshi. *Nature (London)*, **389**, 206 (1997)
242. E.Pebay-Peyroula, G.Rummel, J.P.Rosenbusch, E.M.Landau. *Science*, **277**, 1676 (1997)
243. K.Palczewski, T.Kumasaka, T.Hori, C.A.Behnke, H.Motoshima, B.A.Fox, I.Le Trong, D.C.Teller, T.Okada, R.E.Stenkamp, M.Yamamoto, M.Miyano. *Science*, **289**, 739 (2000)
244. L.A.Rubenstein, R.G.Lanzara. *J. Mol. Struct. (THEOCHEM)*, **430**, 57 (1998)
245. А.А.Иванов, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. *Докл. АН*, **386**, 548 (2002)
246. А.А.Иванов, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. *Вестн. МГУ. Сер. 2. Химия*, **43**, 231 (2002)
247. А.А.Иванов, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. *Докл. АН*, **289**, 404 (2003)
248. А.А.Ivanov, V.A.Palyulin, N.S.Zefirov. *J. Mol. Graph. Model.*, **25**, 740 (2007)
249. А.А.Ivanov, V.Wang, A.M.Klutz, V.L.Chen, Z.G.Gao, K.A.Jacobson. *J. Med. Chem.*, **51**, 2088 (2008)
250. А.А.Ivanov, H.Ко, L.Cosyn, S.Maddileti, P.Besada, I.Fricks, S.Costanzi, T.K.Harden, S.V.Calenbergh, K.A.Jacobson. *J. Med. Chem.*, **50**, 1166 (2007)
251. А.А.Ivanov, S.Costanzi, K.A.Jacobson. *J. Comput. Aided Mol. Des.*, **20**, 417 (2006)
252. А.А.Иванов, А.Э.Воронков, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. *Докл. АН*, **394**, 822 (2004)
253. А.Э.Воронков, А.А.Иванов, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. *Докл. АН*, **403**, 409 (2005)
254. C.Bissantz. *J. Recept. Signal Transduct. Res.*, **23**, 123 (2003)
255. C.Bissantz, P.Bernard, M.Hibert, D.Rognan. *Proteins*, **50**, 5 (2003)
256. C.Bissantz, A.Logean, D.Rognan. *J. Chem. Inf. Comput. Sci.*, **44**, 1162 (2004)
257. G.V.Nikiforovich, G.R.Marshall. *Biochemistry*, **42**, 9110 (2003)
258. G.V.Nikiforovich, G.R.Marshall. *Biochem. Biophys. Res. Commun.*, **345**, 430 (2006)
259. G.V.Nikiforovich, C.M.Taylor, G.R.Marshall. *Biochemistry*, **46**, 4734 (2007)
260. I.G.Tikhonova, R.B.Best, S.Engel, M.C.Gershengorn, G.Hummer, S.Costanzi. *J. Am. Chem. Soc.*, **130**, 10141 (2008)
261. E.Marco, M.Foucaud, I.Langer, C.Escrieut, I.G.Tikhonova, D.Fourmy. *J. Biol. Chem.*, **282**, 28779 (2007)
262. G.Choi, J.Landin, J.F.Galan, R.R.Birge, A.D.Albert, P.L.Yeagle. *Biochemistry*, **41**, 7318 (2002)
263. D.Salom, D.T.Lodowski, R.E.Stenkamp, I.Le Trong, M.Golczak, B.Jastrzebska, T.Harris, J.A.Ballesteros, K.Palczewski. *Proc. Natl. Acad. Sci. USA*, **103**, 16123 (2006)
264. A.J.Harmar. *Genome Biol.*, **2**, R3013 (2001)
265. W.R.Taylor, R.E.Munro, K.Petersen, R.P.Bywater. *Comput. Biol. Chem.*, **27**, 103 (2003)
266. H.Brauner-Osborne, J.Egebjerg, E.O.Nielsen, U.Madsen, P.Krogsgaard-Larsen. *J. Med. Chem.*, **43**, 2609 (2000)
267. H.Brauner-Osborne, P.Wellendorph, A.A.Jensen. *Curr. Drug Targ.*, **8**, 169 (2007)
268. G.Costantino, R.Pellicciari. *J. Med. Chem.*, **39**, 3998 (1996)
269. G.Costantino, A.Macchiarulo, R.Pellicciari. *J. Med. Chem.*, **42**, 5390 (1999)
270. И.И.Баскин, М.С.Беленикин, Г.Костантино, В.А.Палюлин, Р.Пелличари, Н.С.Зефи́ров. *Докл. АН*, **374**, 347 (2000)
271. М.С.Беленикин, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. *Докл. АН*, **383**, 643 (2002)
272. А.А.Jensen, P.O.Sheppard, P.J.O'Hara, P.Krogsgaard-Larsen, H.Brauner-Osborne. *Eur. J. Pharmacol.*, **397**, 247 (2000)
273. М.С.Беленикин, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. *Докл. АН*, **384**, 113 (2002)
274. N.Kunishima, Y.Shimada, Y.Tsuji, T.Sato, M.Yamamoto, T.Kumasaka, S.Nakanishi, H.Jingami, K.Morikawa. *Nature (London)*, **407**, 971 (2000)
275. М.С.Беленикин, И.И.Баскин, Г.Костантино, В.А.Палюлин, Р.Пелличари, Н.С.Зефи́ров. *Докл. АН*, **386**, 261 (2002)
276. М.С.Беленикин, Г.Костантино, В.А.Палюлин, Р.Пелличари, Н.С.Зефи́ров. *Докл. АН*, **393**, 827 (2003)
277. T.Noese, A.Jirgensons, I.Starchenkova, S.Renner, I.Jaunzeme, D.Trifanova, M.Hechenberger, T.Bauer, V.Kauss, C.G.Parsons, G.Schneider, T.Weil. *ChemMedChem*, **2**, 1763 (2007)
278. M.Vanejvs, C.Jatzke, S.Renner, S.Muller, M.Hechenberger, T.Bauer, A.Klochkova, I.Pyatkin, D.Kazyulkin, E.Aksenova, S.Shulepin, O.Timonina, A.Naasis, A.Gutcaits, C.G.Parsons, V.Kauss, T.Weil. *J. Med. Chem.*, **51**, 634 (2008)
279. М.С.Беленикин, В.А.Палюлин, Н.С.Зефи́ров. *Докл. АН*, **394**, 259 (2004)
280. Н.С.Huang, P.S.Klein. *Genom. Biol.*, **5**, 234 (2004)
281. C.C.Malbon. *Front. Biosci.*, **9**, 1048 (2004)
282. X.He. *Dev. Cell*, **4**, 791 (2003)
283. C.E.Dann, J.C.Hsieh, A.Rattner, D.Sharma, J.Nathans, D.J.Leahy. *Nature (London)*, **412**, 86 (2001)
284. А.Э.Воронков, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. *Докл. АН*, **412**, 262 (2007)
285. А.Е.Voronkov, I.I.Baskin, V.A.Palyulin, N.S.Zefirov. *J. Mol. Graph. Model.*, **26**, 1179 (2008)
286. А.Э.Воронков, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. *Докл. АН*, **419**, 696 (2008)
287. R.Dingledine, K.Borges, D.Bowie, S.F.Traynelis. *Pharmacol. Rev.*, **51**, 7 (1999)
288. P.E.Chen, D.J.Wyllie. *Br. J. Pharmacol.*, **147**, 839 (2006)

289. P.J.O'Hara, P.O.Sheppard, H.Thogersen, D.Venezia, V.A.Haldeman, V.McGrane, K.M.Houamed, C.Thomsen, T.L.Gilbert, E.R.Mulvihill. *Neuron*, **11**, 41 (1993)
290. V.Laube, H.Hirai, M.Sturgess, H.Betz, J.Kuhse. *Neuron*, **18**, 493 (1997)
291. N.Armstrong, E.Gouaux. *Neuron*, **28**, 165 (2000)
292. И.Г.Тихонова, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. Докл. АН, **382**, 550 (2002)
293. I.G.Tikhonova, I.I.Baskin, V.A.Palyulin, N.S.Zefirov, S.O.Bachurin. *J. Med. Chem.*, **45**, 3836 (2002)
294. И.Г.Тихонова, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. Докл. АН, **382**, 840 (2002)
295. I.G.Tikhonova, I.I.Baskin, V.A.Palyulin, N.S.Zefirov. *J. Med. Chem.*, **46**, 1609 (2003)
296. И.Г.Тихонова, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. Докл. АН, **389**, 123 (2003)
297. I.I.Baskin, I.G.Tikhonova, V.A.Palyulin, N.S.Zefirov. *J. Med. Chem.*, **46**, 4063 (2003)
298. И.Г.Тихонова, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. Изв. АН. Сер. хим., 1282 (2004)
299. T.Kuner, P.H.Seeburg, H.R.Guy. *Trends Neurosci.*, **26**, 27 (2003)
300. S.Bachurin, S.Tkachenko, I.Baskin, N.Lermontova, T.Mukhina, L.Petrova, A.Ustinov, A.Proshin, V.Grigoriev, N.Lukoyanov, V.Palyulin, N.Zefirov. *Ann. NY Acad. Sci.*, **939**, 219 (2001)
301. D.A.Doyle, J.M.Cabral, R.A.Pfuetzner, A.Kuo, J.M.Gulbis, S.L.Cohen, B.T.Chait, R.MacKinnon. *Science*, **280**, 69 (1998)
302. И.Г.Тихонова, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. Докл. АН, **396**, 551 (2004)
303. И.Г.Тихонова, И.И.Баскин, В.А.Палюлин, Н.С.Зефи́ров. Докл. АН, **397**, 549 (2004)
304. A.Arai, M.Kessler, P.Xiao, J.Ambros-Ingerson, G.Rogers, G.Lynch. *Brain Res.*, **638**, 343 (1994)
305. K.A.Yamada. *Neurosci. Lett.*, **249**, 119 (1998)
306. Y.Sun, R.Olson, M.Horning, N.Armstrong, M.Mayer, E.Gouaux. *Nature (London)*, **417**, 245 (2002)
307. И.Г.Тихонова, М.И.Лавров, В.А.Палюлин, Н.С.Зефи́ров. Докл. АН, **399**, 268 (2004)
308. Пат. 2333211 РФ (2008)
309. K.Brejč, W.J.van Dijk, R.V.Klaassen, M.Schuermans, J.van Der Oost, A.B.Smit, T.K.Sixma. *Nature (London)*, **411**, 269 (2001)
310. P.H.Celie, S.E.van Rossum-Fikkert, W.J.van Dijk, K.Brejč, A.B.Smit, T.K.Sixma. *Neuron*, **41**, 907 (2004)
311. Y.Bourne, T.T.Talley, S.B.Hansen, P.Taylor, P.Marchot. *EMBO J.*, **24**, 1512 (2005)
312. В.И.Чупахин, В.А.Палюлин, Н.С.Зефи́ров. Докл. АН, **408**, 693 (2006)
313. Д.И.Осолодкин, В.И.Чупахин, В.А.Палюлин, Н.С.Зефи́ров. Докл. АН, **412**, 404 (2007)
314. D.I.Osolodkin, V.I.Chupakhin, V.A.Palyulin, N.S.Zefirov. *J. Mol. Graph. Model.*, **27**, 813 (2009)

#### MOLECULAR SIMULATION OF RECEPTORS OF PHYSIOLOGICALLY ACTIVE COMPOUNDS FOR PURPOSES OF MEDICAL CHEMISTRY

I.I.Baskin, V.A.Palyulin, N.S.Zefirov

Department of Chemistry, M.V.Lomonosov Moscow State University  
Leninskie Gory, 119991 Moscow, Russian Federation, Fax +7(495)939-0290

The general strategy of the molecular simulation of biological receptors and their interaction with ligands is considered. The procedures for construction of 3D protein models, molecular docking, evaluation of model quality, determination of the free energy of protein binding with ligands are discussed. The methods of molecular design of new medicaments based on molecular models of biological targets: virtual screening and *de novo* design, are presented. Examples of the above-listed approaches for the simulation of a number of pharmacologically significant receptors, analysis of receptor–ligand interactions and design of new biologically active organic compounds are given.

Bibliography — 314 references.

Received 29th January 2009