

*На правах рукописи*



Новоселецкий Валерий Николаевич

**УЧЁТ ГИДРОФОБНЫХ ВЗАИМОДЕЙСТВИЙ  
ПРИ ОЦЕНКЕ ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ МЕМБРАННЫХ  
БЕЛКОВ И РАЗРАБОТКЕ ДЕЙСТВУЮЩИХ НА НИХ ЛИГАНДОВ  
(НА ПРИМЕРЕ АНТАГОНИСТОВ  $\beta$ 2AR-РЕЦЕПТОРА)**

Специальность: 03.01.02 – Биофизика

**АВТОРЕФЕРАТ**

диссертации на соискание ученой степени  
кандидата физико-математических наук

Москва - 2010

Работа выполнена в Лаборатории моделирования биомолекулярных систем Учреждения Российской академии наук Института биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова РАН, на кафедре физико-химической биологии и биотехнологии Московского физико-технического института.

Научный руководитель:

Доктор физико-математических наук, профессор  
Ефремов Роман Гербертович

Официальные оппоненты:

доктор физико-математических наук,  
профессор  
Ризниченко Галина Юрьевна

доктор биологических наук,  
профессор  
Коротков Евгений Вадимович

Ведущая организация:

Институт биомедицинской химии им. В.Н. Ореховича РАН

Защита диссертации состоится « 25 » ноября 2010 г. в 14 ч. 00 мин. на заседании диссертационного совета Д 501.001.96 Биологического факультета Московского Государственного университета им. М.В. Ломоносова по адресу: 119992, Москва, Воробьевы горы, МГУ, Биологический факультет, кафедра Биофизики, «Новая» аудитория.

С диссертацией можно ознакомиться в библиотеке Биологического факультета Московского Государственного университета им. М.В. Ломоносова.

Автореферат разослан « 15 » октября 2010 г.

Ученый секретарь  
Диссертационного совета,  
кандидат биологических наук



М.Г. Страховская

## **1. Общая характеристика работы**

### **1.1. Актуальность проблемы**

Мембранные белки (МБ) широко участвуют в таких важных биологических процессах, как передача сигналов, транспорт веществ, создание и поддержание трансмембранных (ТМ) потенциалов и т.д., составляя при этом около 30% от общего числа белков, синтезируемых клеткой. Наиболее обширным классом МБ являются рецепторы, чье действие опосредовано G-белками (G-protein Coupled Receptors - GPCR), которые ответственны за клеточную реакцию на гормоны и нейротрансмиттеры, а также за зрение, обоняние и вкус. Неправильное функционирование этих рецепторов порождает большое число заболеваний (Schoneberg et al., 2004), поэтому неудивительно, что в настоящее время действие приблизительно 30% представленных на рынке лекарственных препаратов направлено именно на рецепторы семейства GPCR. Особенно полезным для конструирования лекарств является наличие пространственной структуры соответствующего рецептора, однако возможности экспериментальных методов в этой области ещё ограничены, хотя за последние годы прогресс стал особенно заметен. Так, до недавнего времени была известна структура высокого разрешения лишь для одного белка из семейства GPCR – зрительного родопсина, выделенного из сетчатки быка (Palczewski et al, 2000). Теперь же достижения в технике кристаллизации МБ позволили получить структуры бета2-адренэргического рецептора человека (Cherezov et al., 2007), бета1-адренэргического рецептора индейки (Warne T. et al., 2008) и альфа2-аденозинового рецептора человека (Jaakola et al., 2008). Все эти структуры оказались весьма схожими в области ТМ доменов и заметно отличающимися в области внемембранных петель, что указывает на сохранившуюся нетривиальность в моделировании структур других рецепторов семейства GPCR.

Для понимания механизмов функционирования белков на молекулярном уровне, а также для рационального конструирования новых лекарственных препаратов, требуется знание пространственной структуры белков в комплексах с лигандами – природными субстратами, ингибиторами и пр. Однако определение структуры комплекса МБ с лигандами всё ещё сопряжено со значительными техническими трудностями. В связи с этим, а также благодаря росту вычислительных мощностей современных компьютеров, все более широкое распространение получают методы молекулярного моделирования структуры комплекса. Одним из таких методов является молекулярный докинг –

процедура предсказания структуры комплекса белок-лиганд (Kitchen et al., 2004). Полученные с помощью докинга данные о расположении лиганда в активном центре белка позволяют детально изучать взаимодействия между лигандом и рецептором, что дает возможность судить о биологической роли лиганда. Таким образом, численная оценка процессов распознавания является важным шагом на пути к пониманию функциональной роли этих процессов и предсказанию аффинности и селективности лигандов. Одним из явлений, оказывающих заметное влияние на связывание лигандов, является гидрофобный эффект. В частности, в литературе подробно описано влияние гидрофобности бета-блокаторов на их фармакокинетические, фармакодинамические и клинические свойства (Mannhold, 2005). Для рассмотрения вопроса о гидрофобных свойствах можно воспользоваться концепцией молекулярного гидрофобного потенциала (МГП), который основывается на эмпирических константах атомной гидрофобности, рассчитанных из экспериментальных значений коэффициентов распределения в среде октанол-вода ( $\log P$ ) для органических соединений. Ранее в многочисленных работах уже было показано, что МГП может быть использован, в том числе, и при построении 3D-QSAR моделей, позволяя улучшить результаты предсказания констант связывания и биологической активности (Testa et al., 1996).

Рассмотрение как структур МБ, так и комплексов белок-лиганд с целью выявления лучших из них требует применения различных оценочных функций. В настоящей работе описывается новый метод оценки качества упаковки полипептидных цепей в ТМ доменах мембранных белков. Кроме того, предложен интегрированный подход, объединивший в себе традиционный анализ решений молекулярного докинга с количественным описанием гидрофобной/ гидрофильной организации комплексов рецептора с его лигандами. Данный подход стал возможным благодаря появлению кристаллографической структуры бета2-адренэргического рецептора и позволил уточнить наши представления о роли гидрофобных взаимодействий в связывании бета-блокаторов с этим рецептором.

## **1.2. Цели настоящей работы**

- Разработка новых эффективных функций для оценки качества упаковки альфа-спиральных доменов МБ, в т.ч. в крупнозернистом представлении, на основе молекулярного гидрофобного потенциала;
- Разработка и тестирование метода учета индуцированного соответствия рецептора и лиганда на примере комплексов адренэргического рецептора с антагонистами, для которых известны пространственные структуры высокого разрешения;

- Разработка и тестирование фрагментарной функции оценки гидрофобной комплементарности лиганда и сайта связывания на примере антагонистов бета2-адренэргического рецептора.

### **1.3. Основные результаты, выносимые на защиту**

1. На основе анализа представительного набора пространственных структур МБ высокого разрешения предложен метод количественной оценки качества упаковки их ТМ доменов, основанный на использовании молекулярного гидрофобного потенциала. Показана применимость метода при рассмотрении больших наборов теоретических моделей для выявления тех из них, которые наиболее близки к экспериментальной структуре.

2. На основе анализа кристаллографических структур бета2-адренэргического рецептора человека и данных о связывании с ним ряда антагонистов создан метод предсказания константы связывания аналогичных соединений, использующий фрагментарную гидрофобную комплементарность лиганда и рецептора. Предложен алгоритм модификации существующих антагонистов бета2-адренэргического рецептора человека для повышения их активности.

### **1.4. Научное значение и новизна работы**

Все представленные результаты получены впервые. Разработанный метод качества упаковки ТМ альфа-спиральных доменов является новым подходом к моделированию МБ и, в частности, рецепторов семейства GPCR и химерных рецепторов на их основе. Ключевым свойством нашего метода, которое отличает его от аналогичных методов, разработанных ранее, является его ориентация на крупнозернистое представление белковых структур. Такое нововведение позволяет, с одной стороны, избавиться от излишней чувствительности к конформации боковых цепей, а с другой существенно ускоряет расчеты, позволяя проводить оценку большого числа конформаций белковой структуры в разумные сроки.

Предложенный метод предсказания константы связывания антагонистов с бета2-адренэргическим рецептором человека исключительно на основе гидрофобных характеристик взаимодействия лиганда с рецептором также оригинален. Другой особенностью метода является фрагментарный подход, позволяющий избирательно оценивать комплементарность различных групп, что позволило добиться высокой корреляции предсказанных констант связывания с экспериментальными данными.

## 1.5. Практическое значение работы

Предложенные в работе методы могут найти применение в построении и оптимизации теоретических моделей МБ, в частности, таких важных классов как GPCR или каналов. Метод фрагментарной оценки гидрофобных взаимодействий белок-лиганд позволит с большей аккуратностью идентифицировать активные соединения и предсказывать структуры комплексов, чем стандартные, широко используемые методы докинга.

Указанные подходы могут быть востребованы как в фундаментальных исследованиях структурно-динамических аспектов взаимодействия белков с лигандами, так и в фармацевтической области при рациональном конструировании прототипов новых лекарственных препаратов. Особенно актуальной представляется разработка высокоаффинных и селективных лигандов для рецепторов семейства GPCR, что требует создания точных пространственных моделей этих белков.

## 1.6. Публикации и апробация

Основные результаты работы изложены в 6 статьях, опубликованных в российских и международных реферируемых научных журналах. Материалы работы докладывались на I международной конференции «Математическая биология и биоинформатика» (Пушино, 2006), XIV конференции «Математика. Компьютер. Образование» (Пушино, 2007), XIX зимней молодежной научной школе «Перспективные направления физико-химической биологии и биотехнологии» (Москва, 2007), XV конференции «Математика. Компьютер. Образование» (Дубна, 2008), IV Российском симпозиуме «Белки и пептиды» (Казань, 2009), Пятом международном симпозиуме по вычислительным методам в токсикологии и фармакологии с использованием ресурсов интернет (СМТРИ-2009) (Стамбул, 2009), XVII конференции «Математика. Компьютер. Образование» (Дубна, 2010).

## 1.7. Структура и объем диссертации

Диссертационная работа имеет следующую структуру. В *Главе 1 (Введение)* сформулированы основные цели и задачи исследования, и обоснована его практическая важность. *Глава 2* представляет собой обзор литературы. Он посвящен современным методам моделирования структур мембранных белков вообще и рецепторов GPCR в особенности, а также методам предсказания биологической активности низкомолекулярных соединений. Полученные результаты и их обсуждение приведены в *Главе 3*, состоящей из двух разделов. В *разделе 3.1* описана разработка и тестирование оценивающих функций для структур мембранных белков на основе анализа доступных экспериментальных структурных данных. В *разделе 3.2* описан оригинальный подход к

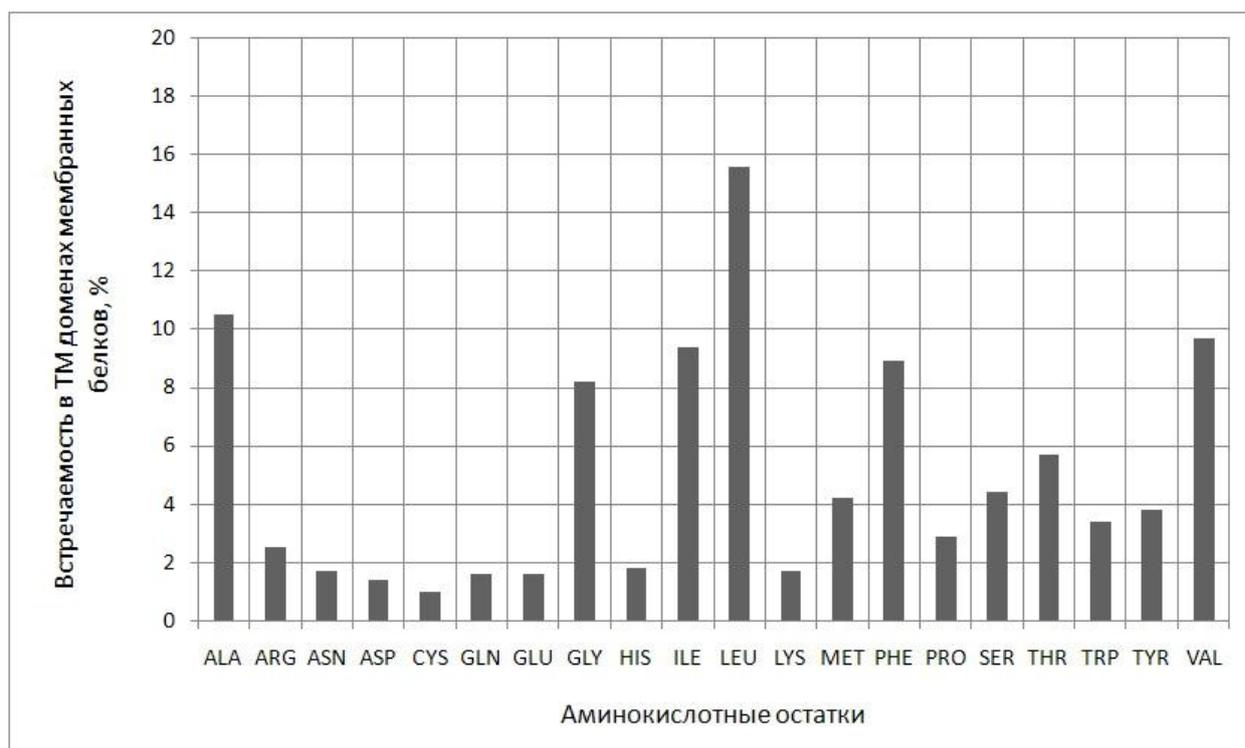
предсказанию констант связывания антагонистов бета2-адренэргического рецептора. *Глава 4 (Заключение)* составляет перечень основных результатов работы, обсуждение их научно-практического значения и дальнейших перспектив исследований в данной области. Описание методов и алгоритмов, использованных в работе, дано в *Главе 5*. Завершает диссертацию список литературы.

Диссертация изложена на 137 страницах машинописного текста, содержит 43 рисунка и 12 таблиц. Список литературы включает 188 источников.

## 2. Основные результаты и их обсуждение

### 2.1. Разработка и тестирование оценочных функций для мембранных белков

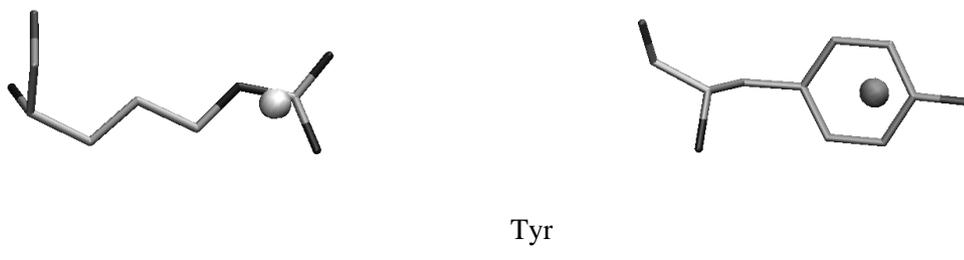
*Создание базы остатков.* Для проведения анализа из базы данных PDB был выбран представительный набор из 18 структур неродственных белков с разрешением  $\leq 3.5 \text{ \AA}$ , содержащих около 5 тысяч остатков в ТМ доменах. Как и следовало ожидать, преобладают в этой базе неполярные остатки (Leu, Ala, Val, Ile, ...) (Рис. 1).



**Рис. 1.** Доля различных аминокислотных остатков в ТМ доменах мембранных белков.

*Крупнозернистое представление.* Для моделирования систем биомолекул на больших временных и пространственных масштабах можно использовать крупнозернистое (КЗ) представление этих систем, суть которого заключается в том, что группа реальных атомов заменяется на один псевдоатом, что позволяет резко уменьшить число взаимодействующих частиц и время расчета их взаимодействий. Поскольку конкретная реализация такого представления не столь важна, то мы остановились на одночастичном представлении, когда один аминокислотный остаток представляется одним зерном (Рис. 2). Носителями свойств остатков выбирали геометрические центры их боковых цепей без учета атомов водорода. Кроме того, для остатков Leu, Ile, Met, Phe, Trp, His, Tyr, Gln, Glu, Lys, Arg не учитывали  $C_{\beta}$ -атомы, а для Lys и Arg – ещё и  $C_{\gamma}$ -атомы. Подобные смещения зёрен от геометрических центров остатков к их функциональным

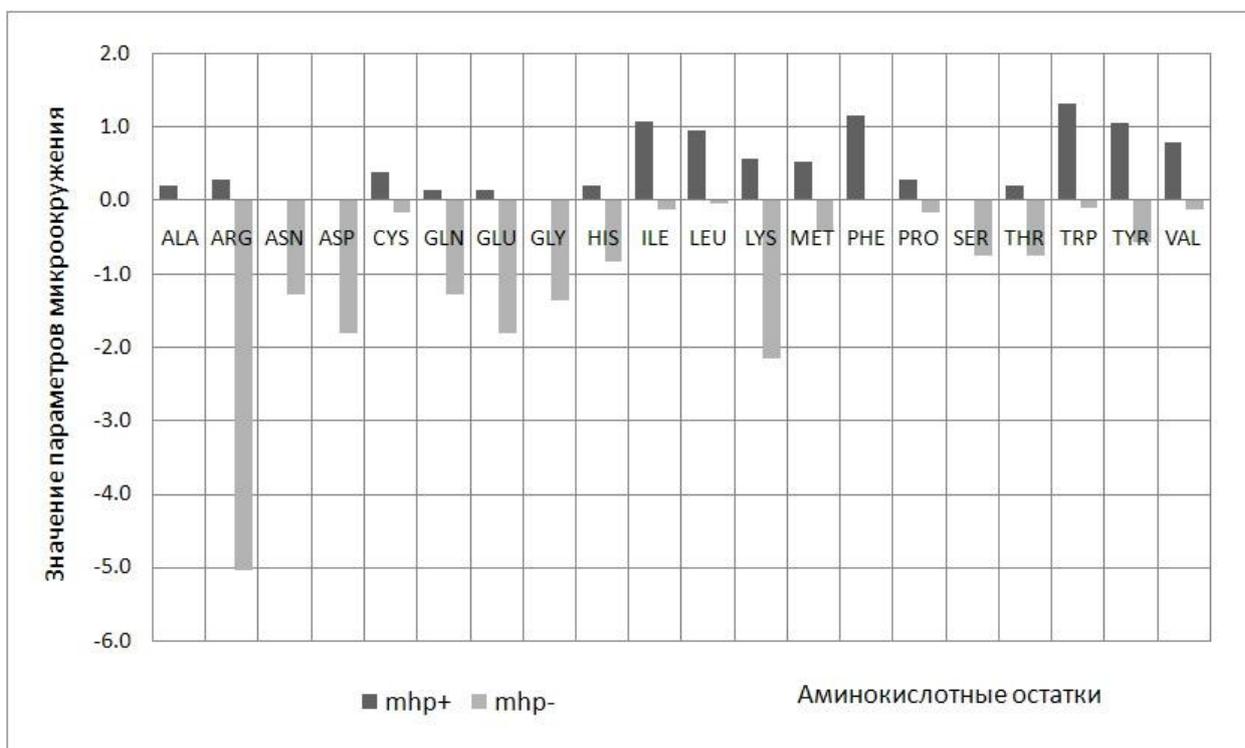
группам были использованы для лучшего соответствия координат зёрен координатам функциональных групп, свойства которых эти зёрна призваны представлять. При необходимости предлагаемый подход может быть адаптирован и к другим вариантам КЗ представлений.



**Рис. 2.** Полноатомное и крупнозернистое представление аргинина и тирозина.

Микроокружение остатков в структуре белка может быть описано с помощью различных параметров. Определив предпочтения остатков обучающего набора обладать теми или иными значениями выбранных параметров, можно проанализировать эти параметры в тестовой структуре и сделать вывод о близости микроокружения каждого остатка в тестовой структуре микроокружению остатков такого же типа в структурах обучающего набора. Как и прежде (Chugunov et al., 2007), было выбрано два параметра – гидрофильная и гидрофобная составляющие микроокружения. Будем называть их  $Fp$  и  $Fnp$ . Для расчета этих параметров был использован подход молекулярного гидрофобного потенциала – МГП (Gaillard et al., 1994; Efremov et al., 2007).

*Вычисление параметров микроокружения для крупнозернистого представления.* При создании крупнозернистой оценочной функции (КЗОФ) для каждого из 20 типов остатков были протабулированы значения двух параметров –  $mhp^+$  и  $mhp^-$ , которые вычисляются суммированием, соответственно, положительных и отрицательных значений МГП-параметров атомов этого остатка (Ghose et al., 1998), с учетом смещения шкалы МГП ( $MHP\text{-offset} = -0.06$  на атом), введенного для повышения точности оценки (Рис. 3).

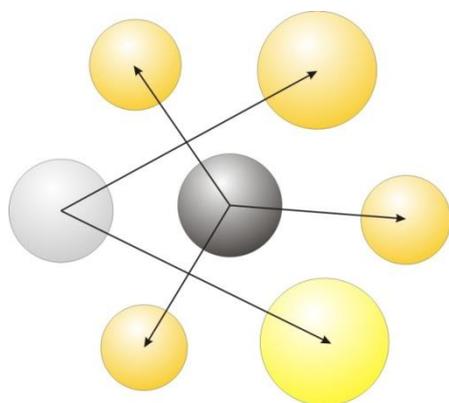


**Рис. 3.** Гидрофобные-гидрофильные свойства боковых цепей аминокислотных остатков.

Затем параметры микроокружения рассматриваемого остатка могут быть вычислены через параметры  $mhp^+$  и  $mhp^-$  остатков, составляющих это микроокружение, т.е. соседних с рассматриваемым. Для этого дополнительно были введены следующие параметры зёрен:

а) заглубленность  $\alpha = 1 - \frac{\left| \sum_i \vec{r}_i \right|}{\sum_i |\vec{r}_i|}$ , где  $\vec{r}_i$  – вектор от рассматриваемого зерна к соседнему с ним

(т.е. находящемуся ближе, чем 7 Å) (Рис. 4);

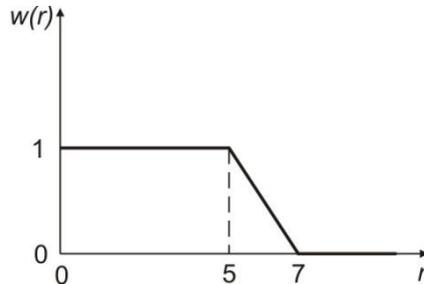


**Рис. 4.** Иллюстрация алгоритма вычисления заглубленности. Центральное зерно, очевидно, является заглубленным; для него  $\sum_i \vec{r}_i = 0$ . Крайние зерна являются экспонированными; для них  $\sum_i \vec{r}_i > 0$ . Для наглядности проведены не все радиус-векторы.

б) суммарная гидрофильность  $Np = \sum_i w(r_i) \cdot mhp_i^-$  и гидрофобность

$Nnp = \sum_i w(r_i) \cdot mhp_i^+$ , наведенная соседними зёрнами, а весовая функция  $w(r)$

представлена на рисунке (Рис. 5).



**Рис. 5.** Весовая функция  $w(r)$ , используемая для расчета наведенного МГП.

Все вышеуказанные суммы вычисляются по зёрнам, не принадлежащим той же спирали, что и анализируемое зерно.

Теперь параметры микроокружения могут быть вычислены по формулам

$$Fp = \alpha \sqrt{\frac{Np}{2(Np + Nnp)}}, \quad Fnp = \alpha \sqrt{\frac{Nnp}{2(Np + Nnp)}}. \quad (1)$$

Квадратный корень использован для придания распределениям остатков по параметрам  $Fp$  и  $Fnp$  более наглядного вида.

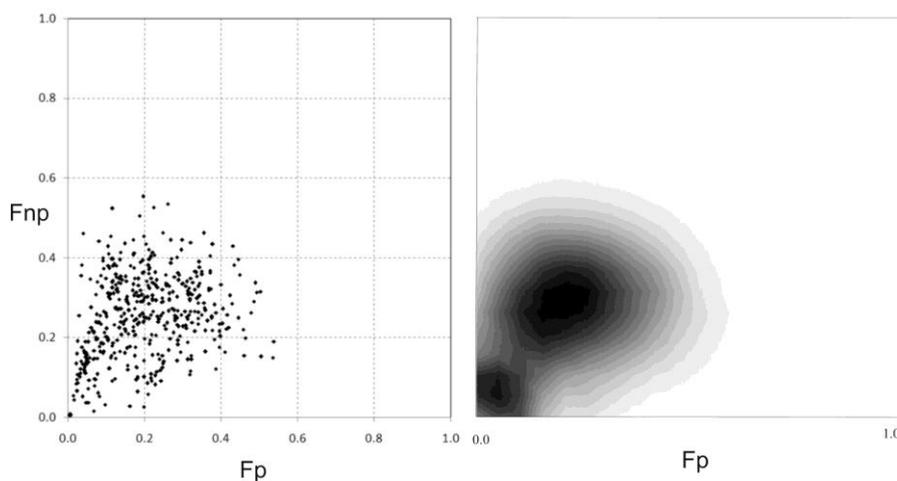
*Вычисление параметров микроокружения для полноатомного представления.* Сначала для  $\alpha$ -спирали, содержащей анализируемый остаток, строили молекулярную поверхность, для каждой точки которой вычисляли её гидрофобность и заглубленность с учетом всех атомов рассматриваемой структуры. При этом использовали весовую функцию, аналогичную приведенной на рисунке (Рис. 5), но интервал убывания составлял  $2 \div 4 \text{ \AA}$ . Затем для анализируемого остатка вычисляли общее число точек его поверхности ( $Nt$ ), включая число заглубленных гидрофобных ( $Nnp$ ) и заглубленных гидрофильных ( $Np$ ) точек. Тогда параметры  $Fp$  и  $Fnp$  могут быть вычислены по тем же формулам (1), но теперь заглубленность должна быть рассчитана как  $\alpha = \frac{Np + Nnp}{Nt}$ .

*Построение оценочных функций.* Имея в виду природу параметров  $Fp$  и  $Fnp$ , очевидно, что их значения принадлежат диапазону от 0 до 1. Разбиение этого диапазона на 20 интервалов для каждого из параметров дает 400 классов окружения. Теперь каждый остаток из обучающего набора может попасть в тот или иной класс окружения в

зависимости от значений  $Fp$  и  $Fnp$ , которыми он обладает. Соответствие каждого из 20 типов остатков с каждым классом окружения может быть выражено следующей функцией

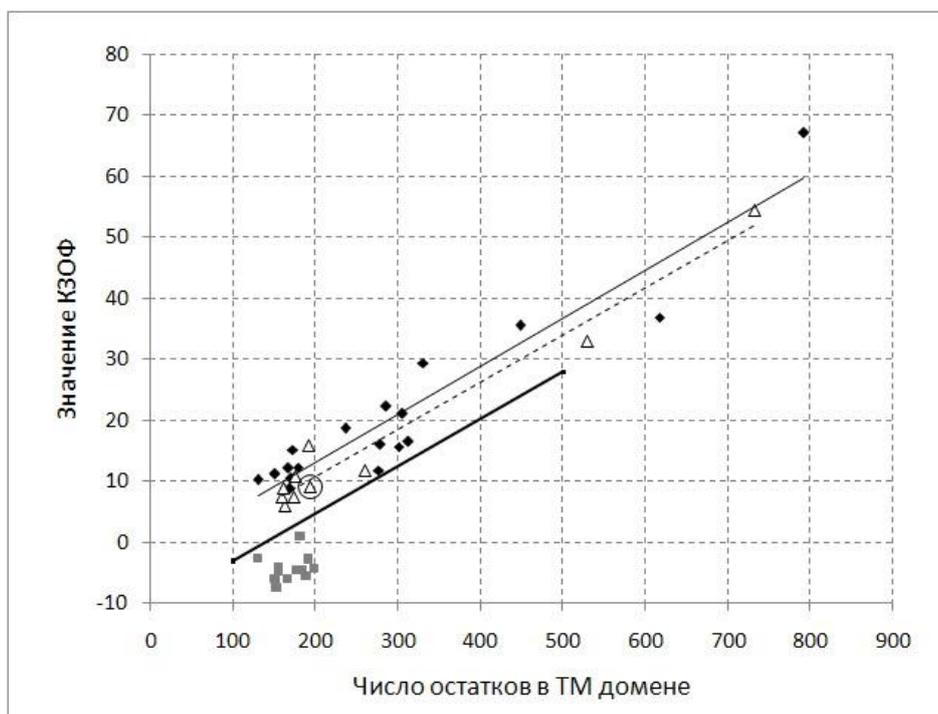
$$S_{ij} = \ln\left(\frac{P_{ij}}{P_j}\right)$$

Здесь  $P_{ij}$  – вероятность обнаружения остатка типа  $i$  в классе  $j$ , а  $P_j$  – вероятность обнаружения остатка произвольного типа в классе  $j$ . Таким образом, если какой-либо остаток имеет повышенное сродство к какому-либо классу окружения, то вероятность его обнаружения там будет выше средней и  $S_{ij} > 0$ . Поскольку не представлялось возможным собрать данные о встречаемости каждого типа остатков в каждом классе окружения, то была применена интерполяция (Рис. 6). Параметры интерполяции были подобраны для достижения самосогласованности метода.



**Рис. 6.** Слева – распределение остатков аланина по параметрам  $Fp$  и  $Fnp$ . Справа – сглаженный вид плотности этого распределения.  $Fp$  может принимать значения 0 до 1 по горизонтальной оси,  $Fnp$  аналогично – по вертикальной.

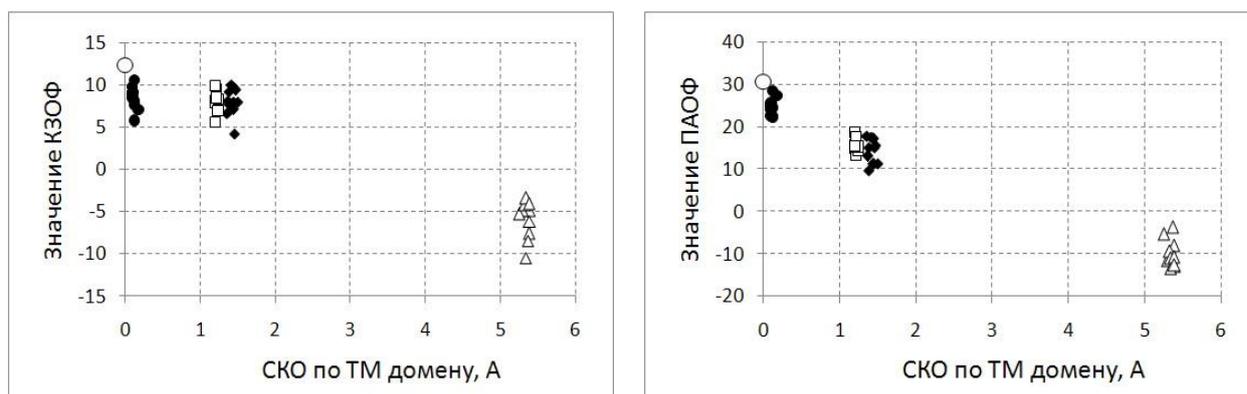
*Тестирование оценочных функций.* Для проверки применимости разработанных ОФ на практике они были использованы для оценки тестового набора МБ. Исходя из предположения, что с точки зрения микроокружения остатков белки обучающего и тестового наборов устроены сходным образом, можно сделать вывод, что структуры с близкой длиной ТМ сегментов должны иметь близкие значения оценочных функций. На рисунке (Рис. 7) показана зависимость значений КЗОФ от длины последовательности ТМ домена для обучающего и тестового наборов МБ.



**Рис. 7.** Зависимость значений КЗОФ от длины ТМ домена для обучающего набора МБ (чёрные ромбы, сплошная аппроксимирующая прямая), тестового набора (белые треугольники, прерывистая прямая) и набора теоретических моделей зрительного родопсина (серые квадраты). Кристаллографическая структура зрительного родопсина (1U19) обведена. Толстая чёрная линия – «прямая отклонения 3 Å» (пояснения в тексте).

Несмотря на то, что вариацией подгоночных параметров не удалось достичь полного совпадения линий тренда для обучающего и тестового наборов, для них обоих прослеживается четкая тенденция к практически равномерному увеличению значений КЗОФ с ростом длины ТМ домена. Кроме того, хорошо видно, что кристаллографическая структура (КС) зрительного родопсина получает заметно более высокое значение ОФ, чем его модели. Это наблюдение, а также ряд аналогичных (изложены ниже), позволяют выявить связь между отклонением модели от КС и соответствующим уменьшением значения ОФ: чем ближе структура модели к КС, тем большим значением ОФ она обладает. Таким образом, становится возможным сформулировать критерий корректности структуры МБ («прямая отклонения 3 Å», см. ниже). В случае с полноатомной оценочной функцией (ПАОФ) ситуация полностью аналогична (данные не приведены). Здесь удалось добиться лучшего соответствия линий тренда для обучающего и тестового наборов, особенно в диапазоне 150 – 300 а.о. Этот диапазон наиболее важен с точки зрения фармацевтических приложений, поскольку такую длину ТМ домена имеют большинство рецепторов (в частности, рецепторы GPCR). Так, длина ТМ домена родопсина составляет 194 а.о., бета2-адреноцептора – 167 а.о.

*Критерий корректности структуры МБ.* Если доступно несколько шаблонов при моделировании на основании гомологии, то проблема выбора наиболее подходящего из них достаточно актуальна. Рассмотрим построение моделей галородопсина (PDB-код 1E12) по гомологии с несколькими шаблонами различной степени родства, а также оценку полученных моделей предлагаемыми ОФ. В качестве шаблонов были выбраны КС бактериородопсина (PDB-код 1BRR) и зрительного родопсина (PDB-код 1U19), а также самого галородопсина.

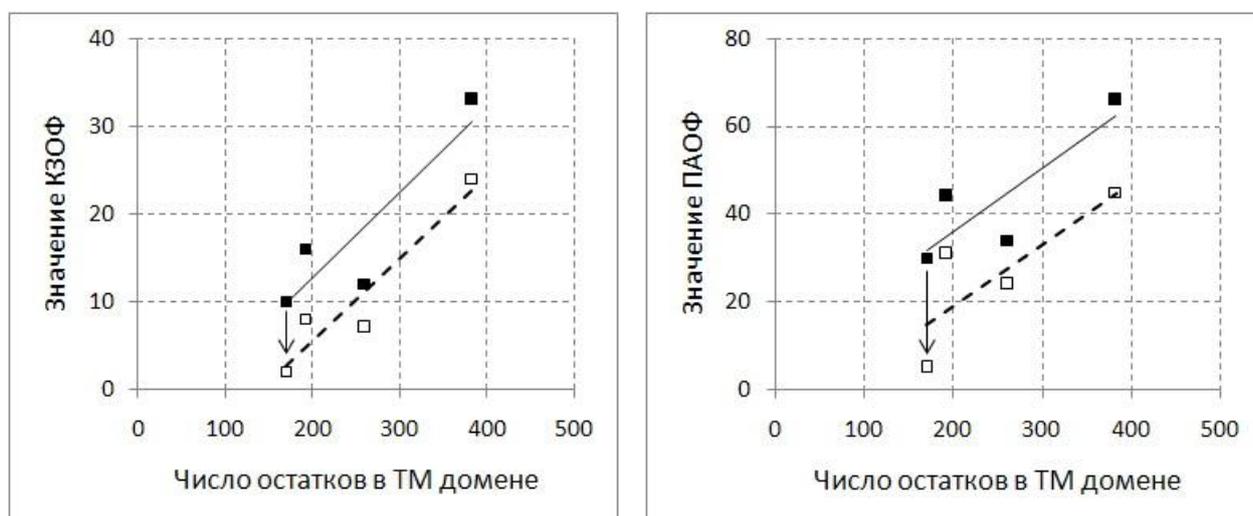


**Рис. 8.** Зависимость КЗОФ (слева) и ПАОФ (справа) от близости моделей к нативной структуре. Белый круг – кристаллографическая структура галородопсина. Модели построены по шаблонам (в порядке возрастания СКО): самого галородопсина (черные кружки), бактериородопсина (белые квадраты), сенсорного родопсина II (черные ромбы) и зрительного родопсина (белые треугольники).

На рисунке (Рис.) приведена зависимость значений ОФ для моделей галородопсина, построенных по гомологии с каждым упомянутым шаблоном, от близости моделей к нативной структуре, в терминах среднеквадратичного отклонения (СКО). Хорошо видно, что чем выше СКО от нативной структуры, тем ниже значение ОФ для соответствующей модели. Это наблюдение подтверждает, что применение метода оценки качества упаковки МБ при моделировании на основании гомологии позволяет идентифицировать наиболее близкие к нативной структуры. К сожалению, КЗОФ демонстрирует низкую чувствительность в диапазоне СКО  $0 \div 2 \text{ \AA}$ , что, вероятно, является следствием самой природы крупнозернистого представления и не может быть устранено.

Подобное моделирование может быть проведено ещё для нескольких МБ с различной длиной ТМ домена: кальциевой АТФ-азы, аммониевого транспортёра и цитохром С оксидазы. Как и в случае моделей галородопсина, для всех наборов моделей хорошо прослеживается убывание значений ОФ с ростом СКО, а также нечувствительность КЗОФ в диапазоне отклонений  $0 \div 2 \text{ \AA}$ . Если теперь отметить те значения ОФ, при которых

модели (будь у нас такие) имели бы отклонение от кристаллографической структуры, равное  $3 \text{ \AA}$ , то применительно к аммониевому транспортёру это будет  $\sim 2$  для КЗОФ и  $\sim 5$  для ПАОФ (Рис. 8). Результат моделирования и оценки моделей галородопсина и еще трех вышеупомянутых МБ удобно представить в виде диаграмм (Рис. 9).

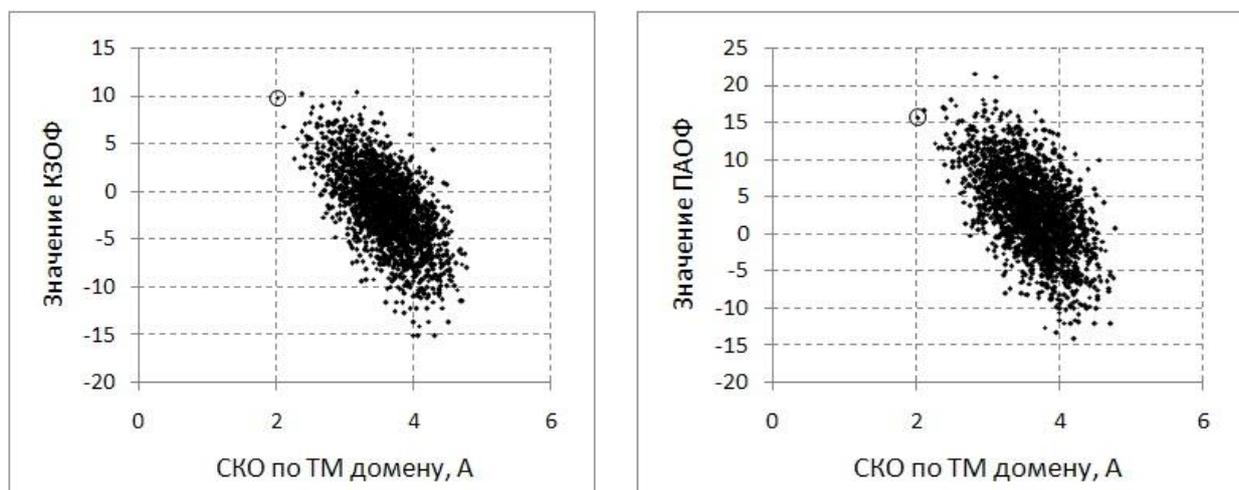


**Рис. 9.** Уменьшение значений (показано стрелкой) КЗОФ (слева) и ПАОФ (справа) при отклонении моделей от КС на  $3 \text{ \AA}$  для галородопсина (170 а.о. в ТМ домене), кальциевой АТФ-азы (192 а.о.), аммониевого транспортёра (261 а.о.) и цитохром С оксидазы (383 а.о.). Чёрные квадраты соответствуют КС мембранных белков, белые – их гипотетическим моделям с СКО  $3 \text{ \AA}$ .

Чётко видно, что регрессионные прямые, соответствующие оценкам КС и гипотетических моделей, практически параллельны. Это наблюдение позволяет предположить общность данного явления не только для четырёх смоделированных белков, но для всех остальных МБ, а также для обеих ОФ. Руководствуясь такими соображениями, можно провести «прямые отклонения на  $3 \text{ \AA}$ » и на диаграмме, где представлена зависимость значений ОФ от длины ТМ домена для обучающего и тестового наборов структур, а также теоретических моделей зрительного родопсина (Рис. 7). Хорошо видно, что все точки, соответствующие КС МБ, качество которых нами постулируется, как хорошее, лежат выше «прямой отклонения на  $3 \text{ \AA}$ ». Модели же родопсина, которые в сравнении с его КС мы можем оценить, как недостаточно хорошие, оказываются лежащими ниже этой прямой. Оба указанных обстоятельства позволяют использовать «прямые отклонения на  $3 \text{ \AA}$ » (и соответствующие им зависимости значений ОФ от длины ТМ домена) в качестве критерия правильности упаковки структуры МБ.

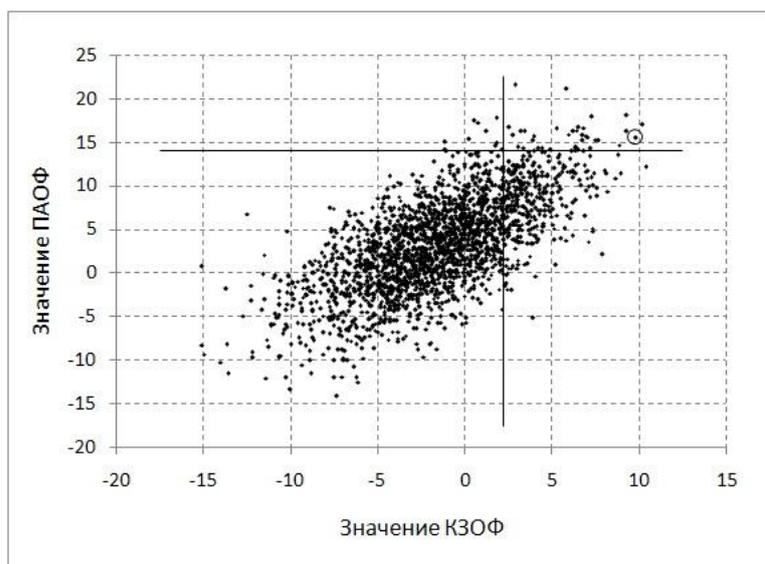
*Возможность выявления ошибок в выравнивании.* При моделировании на основании гомологии смещения в выравнивании приводят к существенному искажению получаемых

моделей по сравнению с оптимальными. Для изучения возможности идентифицировать такие искажения с помощью КЗОФ и ПАОФ был создан набор моделей  $\beta_2$ -адренорецептора по шаблону зрительного родопсина, но с систематическими ошибками в выравнивании ТМ домена. Ошибки состояли в независимом смещении для каждой из 7 спиралей последовательности модели относительно последовательности шаблона на  $0, \pm 1$  остаток по сравнению с оптимальным выравниванием; для каждого варианта выравнивания была построена одна модель. Таким образом, всего было получено  $3^7 = 2187$  моделей, из которых лишь одна обладает оптимальным выравниванием (Рис. 10). Разумеется, при достаточном уровне идентичности аминокислотных последовательностей правильное выравнивание может быть легко построено с помощью общепринятых методов, но в рассматриваемом случае этот уровень составляет менее 30% по ТМ домену, что является критическим при моделировании на основании гомологии.



**Рис. 10.** Значения КЗОФ (слева) и ПАОФ (справа) моделей адренорецептора, построенных по шаблону зрительного родопсина с различными вариантами выравнивания, и их СКО по ТМ  $\alpha$ -атомам от КС адренорецептора (2RH1). Точки, соответствующие модели с правильным выравниванием, обведены.

Хорошо видно, что модель, отвечающая правильному выравниванию, обладает не только минимальным значением СКО, но и достаточно высокими значениями ОФ, соответствующими их нахождению выше «прямой отклонения на  $3 \text{ \AA}$ » (Рис. 7). Если теперь не принимать во внимание данные о КС адренорецептора, то распределение его моделей по КЗОФ и ПАОФ будет иметь следующий вид (Рис. 11).



**Рис. 11.** Распределение моделей адренорецептора, построенных по шаблону зрительного родопсина при искаженном выравнивании, по значениям КЗОФ и ПАОФ. Точка, соответствующая модели с правильным выравниванием, обведена. Вертикальная линия – пороговое величина значения КЗОФ, соответствующее отклонению в  $3 \text{ \AA}$  для длины ТМ домена 167 а.о. (что наблюдается в случае с адренорецептором). Горизонтальная линия – аналогично для ПАОФ.

Легко видеть, что модель, отвечающая правильному выравниванию, является «хорошей» как по КЗОФ, так и по ПАОФ, что существенно сужает область поиска этой модели среди прочих в случае, если бы правильность её выравнивания не была нам известна заранее.

#### *Выводы (1)*

а) Предложен оригинальный метод «крупнозернистого» представления белковых структур.

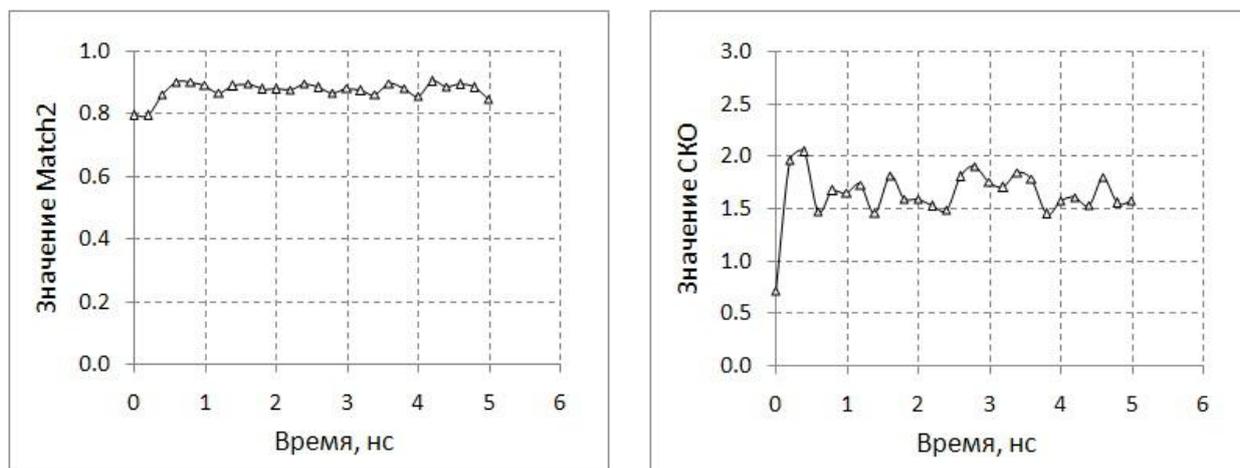
б) Разработаны функции оценки качества упаковки мембранных белков на основе «крупнозернистого» и полноатомного представления.

в) Предложен критерий, позволяющий разграничивать множество моделей на «плохие» и «хорошие» (по СКО от КС), основываясь лишь на значении оценочных функций.

г) Разработанные методы и функции реализованы в программе PROMETHEUS (PROtein MEmbrane THEoretical Ultimate Score), созданной в ходе выполнения диссертационной работы.

## 2.2. Гидрофобная комплементарность бета-блокаторов в сайте связывания $\beta$ 2-адренэргического рецептора и её использование для предсказания константы связывания

*Гидрофобная организация комплексов GPCR с лигандами.* Для анализа гидрофобной организации комплексов использовали интегральный параметр Match2 (Pyrkov et al, 2009), соответствующий комплементарности гидрофобных свойств поверхности лиганда и рецептора. Этот параметр может изменяться в диапазоне от 0, что соответствует отсутствию комплементарности, до 1, означающей полное соответствие. В качестве примера комплекса GPCR с лигандом рассмотрим зрительный родопсин. Его лиганд ретиналь имеет сопряженную полиеновую структуру и является гидрофобным. Естественно ожидать поэтому, что и находящиеся в контакте с ретиналем аминокислотные остатки также будут гидрофобными. С расшифровкой пространственной структуры родопсина стало возможным полное рассмотрение всех таких остатков, и, действительно, большей частью они оказываются либо гидрофобными, либо имеющими значительные гидрофобные фрагменты. В то же время, контактируют с ретиналем и несколько гидрофильных остатков. Все эти контакты вносят свой вклад в величину комплементарности гидрофобных свойств ретиналя, которая в кристаллографической структуре (PDB код 1U19) составляет 0,85. Для выяснения того, насколько стабильным является полученное значение комплементарности, был выполнен расчет МД молекулы зрительного родопсина с ретиналем, встроенной в гидратированный липидный бислой.

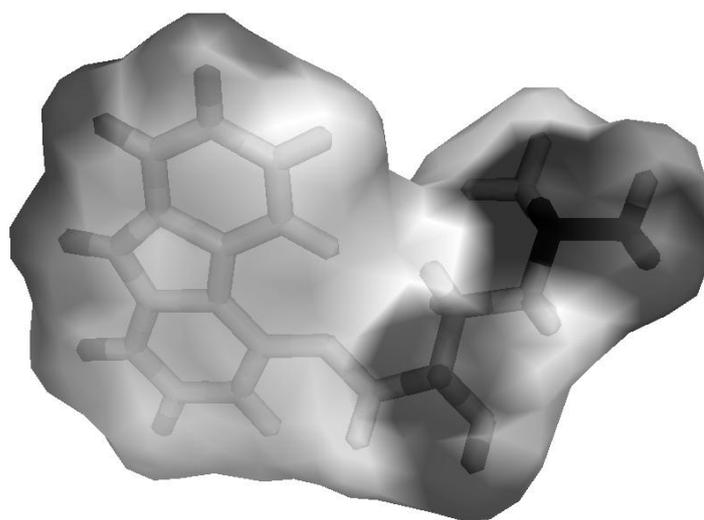


**Рис. 12.** Флуктуации значений комплементарности гидрофобных свойств ретиналя в активном центре (слева) и СКО атомов ретиналя (справа) в процессе МД.

Оказалось, что значения комплементарности подвержены лишь незначительным случайным колебаниям (Рис. 12, слева), в то время как СКО от стартовой структуры атомов как ретиналя (Рис. 12, справа), так и активного центра родопсина (данные не приведены) флуктуирует значительно сильнее. Таким образом, комплементарность лучше

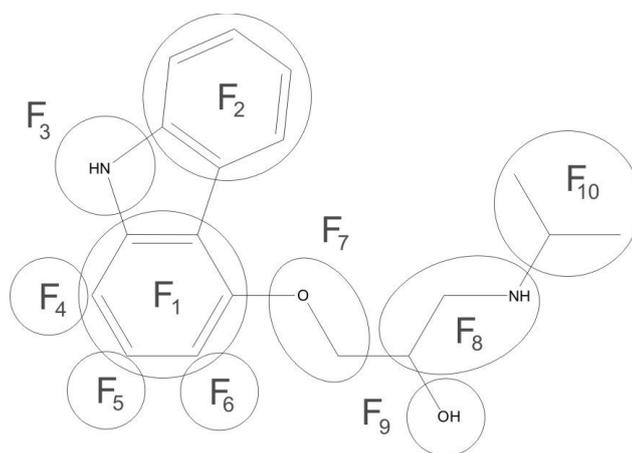
подходит для оценки правильного положения лиганда в активном центре, что делает её ценным инструментом, например, при ранжировании решений докинга.

*Комплементарность гидрофобных свойств в комплексах бета-блокаторов.* Появление кристаллографической структуры бета2-адренорецептора с различными лигандами и возможность их использования в качестве мишеней для молекулярного докинга позволяет детально изучить гидрофобную организацию комплексов антагонистов с рецептором. Для исследования взаимодействий лигандов с рецептором использовали веб-сервер PLATINUM (Pyrkov et al., 2009), а также программные пакеты GOLD и Glide. Были рассмотрены как оценивающие функции целиком, так и их отдельные термы, описывающие физически различные межмолекулярные взаимодействия, такие как гидрофобные контакты и водородные связи. Оказалось, что ни одна из функций не показала приемлемой корреляции с константами связывания и, следовательно, не может быть использована для надежных предсказаний (данные не приведены). Аналогично оказались неприменимыми и параметры собственной гидрофобности лиганда, т.е. гидрофобности, наведенной на поверхности лиганда его же собственными атомами, а также интегральная комплементарность гидрофобных и гидрофильных свойств лиганда и активного центра. Причина этих неудач, вероятно, кроется в том, для лиганда в целом комплементарность гидрофобных свойств между ним и рецептором достаточно низка (Рис. 13).



**Рис. 13.** Молекула каразолола из кристаллографической структуры бета2-АР (2RH1) и ее молекулярная поверхность, раскрашенная по значениям комплементарности гидрофобных свойств лиганда и рецептора. Благоприятные контакты (гидрофобно-гидрофобные и гидрофильно-гидрофильные) показаны светлым, неблагоприятные – темным.

Поскольку интегральные параметры бета-блокаторов не проявили достаточной корреляции с константами связывания, было решено обратиться к параметрам молекулярных групп. Для учета гидрофобных предпочтений различных групп в молекулах лигандов использовали разбиение молекулы антагонистов на 10 фрагментов, которое ниже проиллюстрировано на примере бета-блокатора каразолола (Рис. 14). Фрагмент  $F_1$  представляет собой ароматическую систему (как правило, бензольное кольцо). Фрагмент  $F_2$  в случае каразолола также представляет собой бензольное кольцо. Во фрагменте  $F_3$  всего два атома – азот и водород. Фрагменты  $F_4, F_5, F_6$  отсутствуют, так как располагающиеся на соответствующих местах атомы водорода отнесены к фрагменту  $F_1$ . Фрагментом  $F_7$  является группа  $-O-CH_2-$  (отсутствует у ряда бета-блокаторов). Фрагменты  $F_8$  и  $F_9$  одинаковы для всех лигандов и содержат группы  $-CH-CH_2-NH_2^+$  и  $-OH$ , соответственно. Фрагмент  $F_{10}$  в случае каразолола представляет собой изопропил, но сильно варьирует от лиганда к лиганду.



**Рис. 14.** Схема разбиения молекулы антагониста на фрагменты на примере каразолола (подробности см. в тексте).

Для каждого фрагмента в молекулах лигандов вычисляли комплементарность гидрофобных свойств Match2, для чего учитывали только те точки молекулярной поверхности, которые соответствовали атомам рассматриваемого фрагмента. Оказалось, что фрагмент  $F_1$  во всех случаях проявляет высокую комплементарность, а фрагмент  $F_9$  – отсутствие таковой. В самом деле,  $F_1$  представляет собой бензольное кольцо в контакте с гидрофобным окружением, а  $F_9$  является гидроксильной группой, которая, очевидно, не способна проявлять гидрофобную комплементарность из-за своей гидрофильной природы. Кроме того, чрезвычайно высокую комплементарность была выявлена у фрагмента  $F_2$  в некоторых лигандах, в то время как другие фрагменты продемонстрировали низкую комплементарность.

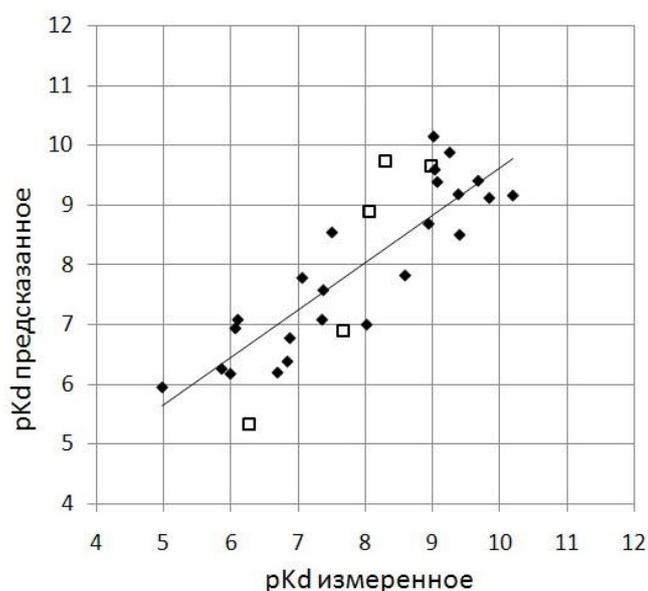
*Фрагментарная функция расчета аффинности.* После того, как определены основные благоприятные факторы связывания антагонистов, становится возможным составить оценочную функцию для предсказания их аффинности. Мы рассмотрели функцию следующего общего вида:

$$pKd_{pred} = A + B \cdot MMH + \sum_i C_i \cdot Match2(F_i),$$

где ММН – середина диапазона распределения гидрофобности (полусумма максимального и минимального значения собственного МГП на поверхности лиганда), Match2(F<sub>i</sub>) – гидрофобная комплементарность *i*-го фрагмента лиганда, а А, В, С<sub>і</sub> – весовые коэффициенты. Среднее значение и медиана гидрофобности лиганда также были опробованы вместо ММН, но оказалось, что именно ММН позволяет получить наиболее эффективную ОФ для предсказания аффинности. Определение наиболее существенных факторов и подгонка коэффициентов были выполнены с помощью множественной регрессии. Коэффициенты А, В, С<sub>і</sub> были усреднены по 8 редуцированным обучающим наборам в процессе кросс-валидации при 6 исключенных точках (leave-6-out cross-validation) и итоговое уравнение получило вид:

$$pKd_{pred} = 10.0 + 13.2 \cdot MMH + 2.4 \cdot Match2(F_2) - 1.8 \cdot Match2(F_4) + \\ + 2.1 \cdot Match2(F_6) + 3.1 \cdot Match2(F_8)$$

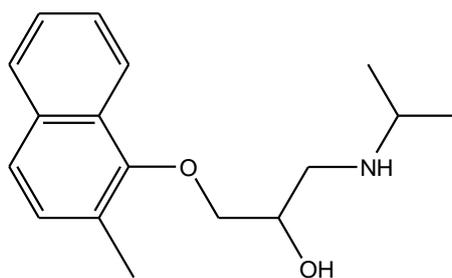
Соотношение предсказанных и экспериментальных значений аффинности для набора обратных агонистов и антагонистов с широким спектром значений pKd (от 5 до 11) показано на рисунке (Рис. 15) ( $R^2 = 0.77$ ;  $s = 0.76$ ;  $F_{5, 18} = 11.95$ ).



**Рис. 15.** Соотношение предсказанных и экспериментальных значений аффинности (pKd) для набора обратных агонистов и антагонистов бета2-АР ( $R^2 = 0.77$ ). Ромбы и квадраты относятся к обучающему набору (24 случая) и тестовому (5 случаев), соответственно.

Коэффициент корреляции для предсказанных значений ( $q^2$ ) составил 0.62, при том, что приемлемым значением в данном случае является  $q^2 > 0.3$ . Таким образом, наши результаты подтверждают возможность использования комплементарности гидрофобных свойств для предсказания аффинности соединений.

*Соединения с потенциально высокой аффинностью.* Принимая во внимание вышесказанное, можно попытаться модифицировать ингибитор с уже высокой аффинностью в сторону дальнейшего ее увеличения. Для этого необходимо улучшить гидрофобную комплементарность его отдельных фрагментов с окружающими этот фрагмент остатками рецептора. Из формулы для расчета значений константы связывания  $pKd_{пред}$  видно, что благодаря положительности коэффициентов  $C_2$  и  $C_6$  эффективным может быть введение гидрофобных заместителей в качестве фрагментов  $F_2$  и  $F_6$  (Рис.). Для проверки этого предположения с использованием программы комбинированного докинга CombiGlide был выполнен поиск оптимальных заместителей в качестве фрагмента  $F_6$  для пропранолола (фрагмент  $F_2$  в нем уже присутствует и состоит из двух атомов углерода и двух атомов водорода бензольного кольца). После выполнения всех вычислений, описанных выше (докинг с помощью CombiGlide, минимизация энергии полученного комплекса с наложением ограничений на лиганд и оценка с помощью фрагментарной ОФ) получили значение  $pKd_{пред} = 11.1$  в случае метильного заместителя (Рис. 16). Это значение заметно выше тех, которые получены для обучающего и тестового наборов, однако, несомненно, данное предсказание нуждается в экспериментальном подтверждении.



**Рис. 16.** Потенциальный бета-блокатор аналогичный пропранололу, но отличающийся от него наличием метильного заместителя у бензольного кольца.

#### *Выводы (2)*

- а) Предложен оригинальный метод разбиения молекулы лиганда на фрагменты для детального учета гидрофобной комплементарности.
- б) Разработана функция количественной оценки аффинности бета-блокаторов.
- в) Предложен ряд соединений, предположительно обладающих высоким сродством к бета2-адренэргическому рецептору.

## СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Чугунов А.О., Новоселецкий В.Н., Арсеньев А.С., Ефремов Р.Г. Новый метод оценки качества упаковки трансмембранных  $\alpha$ -спиральных доменов в белках. (2007). *Биохимия* Т. 72, № 3, 358-367.  
English version: Chugunov A.O., Novoseletsky V.N., Arseniev A.S., Efremov R.G. A novel method for packing quality assessment of transmembrane alpha-helical domains in proteins. (2007) *Biochemistry (Moscow)*. V. 72, N 3, 293-300.
2. Chugunov A.O., Novoseletsky V.N., Nolde D.E., Arseniev A.S., Efremov R.G. A method to assess packing quality of transmembrane  $\alpha$ -helices in proteins. I. Parameterization using structural data. (2007). *J. Chem. Inf. Mod.* v. 47, N 3, 1150-1162.
3. Chugunov A.O., Novoseletsky V.N., Nolde D.E., Arseniev A.S., Efremov R.G. A method to assess packing quality of transmembrane  $\alpha$ -helices in proteins. II. Validation by “correct vs. misleading” test. (2007). *J. Chem. Inf. Mod.* v. 47, N 3, 1163-1170.
4. Чугунов А.О., Новоселецкий В.Н., Арсеньев А.С., Ефремов Р.Г. Интегральные мембранные белки: подход к созданию реалистичных моделей *in silico*. (2007). *Мембраны, изд. ВИНТИ, Серия «Критические технологии», № 1 (33), С. 21-31.*
5. Новоселецкий В.Н., Чугунов А.О., Ефремов Р.Г. Метод распознавания некорректно свернутых моделей  $\alpha$ -спиральных трансмембранных доменов белков. В: *Математика. Компьютер. Образование.* Сб. трудов XIV международной конференции. Под общей редакцией Г.Ю. Ризниченко Ижевск: Научно-издательский центр "Регулярная и хаотическая динамика", 2007. Том 2, 392 стр. Стр. 313-318.
6. Novoseletsky V.N., Pyrkov T.V., Efremov R.G. Analysis of hydrophobic interactions of antagonists with the beta2-adrenergic receptor. (2010) *SAR and QSAR in Environmental Research* v. 21, N 1, 37-55.
7. Новоселецкий В.Н., Чугунов А.О., Ефремов Р.Г. (2006). Метод оценки упаковки трансмембранных  $\alpha$ -спиралей в моделях мембранных белков. *VIII чтения, посвященные памяти академика Ю.А. Овчинникова*, Москва, стр. 85.
8. Новоселецкий В.Н., Чугунов А.О., Ефремов Р.Г. «Метод распознавания некорректно свернутых моделей  $\alpha$ -спиральных трансмембранных доменов белков». *Современные проблемы фундаментальных и прикладных наук – общая и прикладная физика: Сборник трудов 49-й научной конференции МФТИ*, Долгопрудный, (2006).  
[http://bio.fizteh.ru/student/mipt\\_conference/conf\\_result.html](http://bio.fizteh.ru/student/mipt_conference/conf_result.html)
9. Chugunov A.O., Novoseletsky V.N., Efremov R.G. “A method to assess correct/misfolded structures of transmembrane domain of membrane proteins”. *Proceedings of the Fifth*

- International Conference on Bioinformatics of Genome Regulation and Structure*,  
Novosibirsk, Russia 16-22 July, (2006), **1**, 247-251.
10. Novoseletsky, V.N., Chugunov A.O., Efremov, R.G. (2006). Assessment of the packing quality of  $\alpha$ -helical transmembrane domains. *Proceedings of the 1<sup>st</sup> International Conference on Mathematical Biology and Bioinformatics*, Pushchino, Russia, 14-15. Изд-во "МАКС-ПРЕСС", Москва, 2006.
  11. Новоселецкий В.Н., Чугунов А.О., Ефремов Р.Г. Метод распознавания некорректно свернутых моделей  $\alpha$ -спиральных трансмембранных доменов белков. *XIV Международная конференция «Математика. Компьютер. Образование»*. 22-27 января 2007 г., Пущино, Сборник научных тезисов под ред. Г.Ю. Ризниченко, НИЦ «Регулярная и хаотическая динамика», г. Ижевск, 2007, С. 176.
  12. Новоселецкий В.Н., Чугунов А.О., Ефремов Р.Г. Оптимизация конформаций боковых цепей в моделях мембранных белков. *XIX зимняя молодежная научная школа «Перспективные направления физико-химической биологии и биотехнологии»*, 7-9 февраля 2007 г., Тезисы докладов и стендовых сообщений под ред. Т.В. Овчинниковой и Т.И. Соркиной, ИБХ РАН, Москва, 2007, С. 55.
  13. Novoseletsky V.N., Chugunov A.O., Efremov R.G. A method to recognize misfolded models of alpha-helical transmembrane protein domains. In: *Biophysics and the Challenges of Emerging Threats*. Proc. of the NATO Advanced Study Institute on Biophysics and the Challenges of Emerging Threats Erice, Sicily, Italy 19-30 June 2007. Series: NATO Science for Peace and Security Series. Subseries: NATO Science for Peace and Security Series B: Physics and Biophysics. Puglisi, Joseph D. (Ed.), 2009, Approx. 230 p., ISBN: 978-90-481-2366-7, p. 160.
  14. Chugunov A.O., Novoseletsky V.N., Efremov R.G. Quality of computer-built models of membrane proteins accessed via knowledge-based approach. *Forth International Symposium on Computational Methods in Toxicology and Pharmacology Integrating Internet Resources*. September 1-5, 2007, Moscow, Russia, p. 89.
  15. Новоселецкий В.Н., Пыркков Т.В., Ефремов Р.Г. Функция оценки качества структуры мембранных доменов в «крупнозернистом» представлении. *XV Международная конференция «Математика. Компьютер. Образование»*. 28 января – 2 февраля 2008 г., Дубна, Сборник научных тезисов под ред. Г.Ю. Ризниченко, НИЦ «Регулярная и хаотическая динамика», г. Ижевск, 2008, С. 200.
  16. Efremov R.G., Nolde D.E., Polyansky A.A., Novoseletsky V.N., Volynsky P.E. "Modern computational pharmacology: movement toward biomembranes". *Proceedings of the Sixth*

*International Conference on Bioinformatics of Genome Regulation and Structure,*

Novosibirsk, Russia 22-28 June, (2008), **1**, 66

17. Novoseletsky V.N., Pyrkov T.V., Efremov R.G. Analysis of hydrophobic organization of GPCR complexes with ligands. *Fifth International Symposium on Computational Methods in Toxicology and Pharmacology Integrating Internet Resources*, Abstract Book, p. 111, Istanbul, Turkey, 2009.
18. Новоселецкий В.Н., Пыркков Т.В., Ефремов Р.Г. Использование шаблона из идеальных альфа-спиралей для моделирования структуры комплекса адренэргического рецептора с антагонистом. *Тезисы докладов IV российского симпозиума «Белки и пептиды»*, стр. 175, Изд-во «ФизтехПресс», Казань, 2009.
19. Новоселецкий В.Н., Пыркков Т.В., Ефремов Р.Г. Анализ гидрофобной комплементарности сайта связывания бета2-адренэргического рецептора. *XVII Международная конференция «Математика. Компьютер. Образование»*. 25-30 января 2010 г., Дубна, Сборник научных тезисов под ред. Г.Ю. Ризниченко, стр. 255, НИЦ «Регулярная и хаотическая динамика», г. Ижевск, 2010.

## СПИСОК ЦИТИРУЕМОЙ ЛИТЕРАТУРЫ

1. Cherezov, V., Rosenbaum, D.M., Hanson, M.A., Rasmussen, S.G., Thian, F.S., Kobilka, T.S., Choi, H.J., Kuhn, P., Weis, W.I., Kobilka, B.K., Stevens, R.C. (2007). High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318**, 1258–1265
2. Chugunov, A.O., Novoseletsky, V.N., Arseniev, A.S., Efremov, R.G. (2007). A novel method for packing quality assessment of transmembrane alpha-helical domains in proteins. *Biochemistry (Mosc.)* **72**, 293–300.
3. Efremov, R.G., Chugunov, A.O., Pyrkov, T.V., Priestle, J.P., Arseniev, A.S., Jacoby, E. (2007). Molecular lipophilicity in protein modeling and drug design. *Curr. Med. Chem.* **14**, 393–415.
4. Gaillard, P., Carrupt, P.A., Testa, B., Boudon, A. (1994). Molecular lipophilicity potential, a tool in 3D QSAR: method and applications. *J. Comput. Aided. Mol. Des.* **8**, 83–96.
5. Ghose, A.K., Viswanadhan, V.N., Wendoloski, J.J. (1998). Prediction of hydrophobic (lipophilic) properties of small organic molecules using fragmental methods: an analysis of ALOGP and CLOGP methods. *J. Phys. Chem.* **102**, 3762–3772.
6. Jaakola, V.P., Griffith, M.T., Hanson, M.A., Cherezov, V., Chien, E.Y., Lane, J.R., Ijzerman, A.P., Stevens, R.C. (2008). The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **322**, 1211–1217.
7. Kitchen, D.B., Decornez, H., Furr, J.R., Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* **3**, 935–49.
8. Mannhold, R., (2005). The impact of lipophilicity in drug research: a case report on beta-blockers. *Mini Rev. Med. Chem.* **5**, 197–205.
9. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., Le Trong, I., Teller, D.C., Okada, T., Stenkamp, R.E., Yamamoto, M., Miyano, M. (2000). Crystal Structure of Rhodopsin: A G-Protein–Coupled Receptor. *Science* **289**, 739–745.
10. Pyrkov, T.V., Chugunov, A.O., Krylov, N.A., Nolde, D.E., Efremov, R.G. (2009). PLATINUM: a web tool for analysis of hydrophobic/hydrophilic organization of biomolecular complexes. *Bioinformatics* **25**, 1201–1202.
11. Schöneberg, T., Schulza, A., Biebermann, H., Hermsdorf, T., Römpler, H., Sangkuhl, K. (2004). Mutant G-protein–Coupled Receptors as a Cause of Human Diseases. *Pharmacol. Ther.* **104**, 173–206.
12. Testa, B., Carrupt, P.A., Gaillard, P., Billois, F., Weber, P. (1996). Lipophilicity in molecular modeling. *Pharm. Res.* **13**, 335–343.
13. Warne, A., Serrano-Vega, M.J., Baker, J.G., Moukhametzianov, R., Edwards, P.C., Henderson, R., Leslie, A.G.W., Tate, C.G., Schertler, G.F.X. (2008) Structure of the Beta1-Adrenergic G Protein-Coupled Receptor. *Nature* **454**, 486.