

## Задача №...

### «Молекулярная динамика мембранных белков на примере адренорецептора»

асс. В.Н.Новоселецкий, проф. Шайтан К. В.

#### **Аннотация задачи**

Вычислительные методы приобретают всё большую значимость во всех сферах науки, в том числе в исследовании биохимических процессов, протекающих в живой клетке, и в фармацевтической отрасли. В разработке новых лекарственных соединений используют алгоритмы компьютерного моделирования межмолекулярных взаимодействий в биологически значимых комплексах белков-рецепторов с низкомолекулярными соединениями – их лигандами. Такие взаимодействия лежат в основе большинства биохимических процессов, отвечающих за передачу сигналов, межклеточное распознавание, рецепцию и множества других. Нарушение работы биомолекулярных рецепторных комплексов – причина очень многих заболеваний, и белки-рецепторы часто используют в качестве мишени для лекарственного воздействия.

Прямое численное моделирование молекулярной динамики биологических макромолекул и их систем стало возможным благодаря достижениям современной вычислительной техники. В настоящее время метод молекулярной динамики – это один из важнейших способов исследования биологических нанообъектов, в основе которого лежит представление моделируемой молекулярной системы в виде совокупности взаимодействующих материальных точек, движение которых описывается классическими уравнениями Ньютона. Численное интегрирование этих уравнений позволяет рассчитать траектории движения всех точек, а математическая обработка этих траекторий дает возможность получать структурные и динамические характеристики моделируемой системы и отслеживать, таким образом, эволюцию системы во времени.

**Цели задачи:** освоение методики проведения молекулярно-динамических расчетов структур мембранных белков в программе Gromacs на примере бета2-адренэргического рецептора и анализ его конформационных изменений.

**Объекты исследования:** пространственные структуры молекулы бета2-адренэргического рецептора (АР) и липидного бислоя, состоящего из 288 молекул пальмитоилолеилфосфатидилхолина (ПОФХ).

**Методы:** молекулярно-динамический расчет (МД) траектории движения атомов.

### **План работы:**

1. Изучение статической структуры адренорецептора, взятой из базы данных белковых структур PDB. Рассмотрение положения лиганда в сайте связывания АР с помощью программы VMD.
2. Создание топологии молекулы лиганда для силового поля Gromacs с использованием веб-интерфейса программы PRODRG. Проверка правильности полученной топологии путем запуска пробного расчета МД лиганда в водной среде.
3. Встраивание структуры комплекса рецептор-лиганд в липидный бислой вручную с использованием традиционного протокола и автоматически с использованием программ пакета Gromacs.
4. Запуск расчета МД полученной белок-липидной системы.
5. Анализ полученной МД траектории средствами пакета Gromacs.
6. Формулирование выводов о проделанной работе.

### **Предполагаемые результаты и навыки:**

- Приобретение навыков работы с базой данных белковых структур PDB, программой визуализации структур биомолекул VMD;
- Приобретение навыков составления файлов с описанием топологии низкомолекулярных соединений с помощью программы PRODRG;
- Приобретение навыков выполнения молекулярно-динамических расчетов и проведения анализа их результатов в программном пакете Gromacs.

Полученные результаты продемонстрируют стабильность белковой структуры в липидном окружении. При выполнении работы студенты ознакомятся с теорией и практикой проведения молекулярно-динамических расчетов биомолекул, а также получат навыки работы с базами данных и программами, снабженными веб-интерфейсом.

## **Оценка итогов проведенного практикума:**

Основной формой отчетности студентов является письменный отчет по результатам исследования, который должен включать:

1. краткое теоретическое введение с описанием исследуемых объектов и методов;
2. четко сформулированные цели и задачи исследования;
3. корректно обработанные и представленные в удобной для восприятия форме результаты проведенных экспериментов и обсуждение полученных результатов;
4. четкие выводы по полученным результатам;
5. список использованной дополнительной литературы.

По представленному отчету преподаватель проводит со студентами устную беседу, в ходе которой студенты отвечают на вопросы по теоретическим основам выполняемой задачи, ходу выполнения работы, результатам и сделанным выводам. Отчет оценивается преподавателем и ответственным за проведение практикума.

## **Организация практикума:**

Работа выполняется группой студентов до 6 человек под руководством преподавателя. Местом проведения являются лаборатории кафедры биофизики биологического факультета МГУ. В ходе выполнения работы используется высокопроизводительный Линукс-кластер Лаборатории молекулярной динамики (кафедра Биоинженерии Биологического факультета МГУ им. М.В. Ломоносова) на процессорах 2 x Dual Core AMD Opteron 285 (2.6 GHz) и терминалы на базе ноутбуков ASUS A6M.

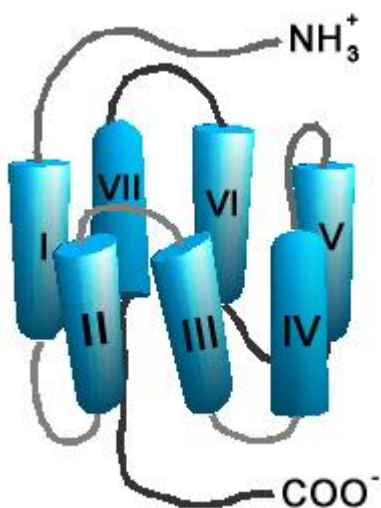
## **Описание задачи**

### **Введение**

Метод молекулярной динамики (МД) является одним из наиболее востребованных инструментов теоретического изучения структуры и динамики биологических объектов различного масштаба в пространстве и времени: от функциональных групп низкомолекулярных соединений до макромолекулярных систем, и от пикосекунд до сотен микросекунд. В основе метода МД лежит расчёт классических траекторий движения макромолекулы в фазовом пространстве координат и импульсов её атомов. В настоящее

время метод МД широко используется для исследования структурных, динамических и термодинамических характеристик биологических молекул и их комплексов. К задачам, решаемым с помощью этого метода, относятся изучение стабильности белков и процессов ионного транспорта, исследование конформационных превращений систем, сворачивания белков, докинга лигандов.

Рецепторы, чье действие опосредовано G-белками, (рецепторы семейства GPCR (G protein-coupled receptors, GPCRs)) – это важнейший класс мембранных белков (МБ), отвечающий за трансмембранную (ТМ) передачу сигналов в клетке, реакцию на свет, генерацию ТМ электростатического потенциала и т.д. В настоящее время рецепторы GPCR являются мишенями для 50% вновь создаваемых лекарственных препаратов, так как с их дисфункцией связывают многие заболевания - аллергия, шизофрения, гипертоническая болезнь, астма, психозы и другие. Геном человека кодирует около 1000 различных рецепторов этого семейства, из которых примерно половина является обонятельными, а остальные рецепторы разделяют на классы по типу лиганда, с которым они взаимодействуют. Характерной особенностью рецепторов семейства GPCR является уникальная консервативность пространственной укладки: все они состоят из ТМ домена, образованного семью фрагментами (предположительно,  $\alpha$ -спиралями, примерно ортогональными плоскости мембраны), и соединяющих их петель (при этом N-конец белка находится во внеклеточной области, а С-конец — в цитоплазматической) (Рис. 1).



**Рис. 1.** Схематичная структура рецепторов семейства GPCR. Цилиндрами показаны трансмембранные  $\alpha$ -спирали.

Трансмембранный домен рецепторов всех семейств имеет примерно один и тот же размер, в то время как внеклеточные домены (особенно внеклеточный) могут существенно варьировать в размере для разных подсемейств GPCR. Этот тип ТМ укладки, изначально гипотетический, был подтвержден с появлением в 2000 году кристаллографической структуры зрительного родопсина из сетчатки быка (pdb-код 1U19 и др.), которая до недавнего времени являлась единственной структурой высокого разрешения GPCR-рецептора. Однако благодаря существенному прогрессу в экспрессии и выделении значительных количеств белка для ряда рецепторов, в последние годы были получены структуры еще целого ряда рецепторов семейства GPCR –  $\beta_2$ -адреноцептора человека (pdb-код 2RH1 и др.), зрительного родопсина кальмара,  $\beta_1$ -адреноцептора индейки (pdb-код 2VT4) и аденозинового рецептора человека. Все они также имеют семиспиральную укладку ТМ домена и пространственно очень схожи как со зрительным родопсином быка, так и между собой. Природными лигандами адренэргических рецепторов являются адреналин и норадреналин (Рис. 2). В то же время наряду с прояснением вопросов, касающихся строения рецепторов, получение их структуры породило много новых вопросов, касающихся их функционирования. Получение ответа на такие вопросы подразумевает, прежде всего, детальное изучение функционирования этих и других рецепторов GPCR.



Рис. 2. Природные агонисты адренэргических рецепторов – адреналин (слева) и норадреналин

## Материалы и методы

### Описание метода МД

В основе методов молекулярной динамики лежит модельное представление о многоатомной молекулярной системе, в которой все атомы представляют собой материальные точки, а поведение отдельного атома описывается классическими уравнениями движения:

$$m_i \frac{d^2 \vec{r}_i(t)}{dt^2} = \vec{F}_i \quad (1)$$

где  $i$  – номер атома ( $1 \leq i \leq n$ ,  $n$  – полное число атомов в системе),  $m_i$  – масса атома,  $\vec{r}_i$  – радиус-вектор атома,  $\vec{F}_i$  – равнодействующая сил, действующих на атом.

Численное интегрирование уравнения (1) дает траекторию системы с известными в каждый момент времени скоростями и координатами каждого атома, а также значения энергии системы.

Равнодействующая сила складывается из двух составляющих:

$$\vec{F}_i = -\frac{\partial U(\vec{r}_1, \dots, \vec{r}_n)}{\partial \vec{r}_i} + \vec{F}_i^{ex} \quad (2)$$

Это, во-первых, градиент потенциальной энергии системы  $U$ , включающей в себя сумму всех взаимодействий атомов (3), а во-вторых, дополнительная сила  $\vec{F}_i^{ex}$ , определяемая взаимодействиями с молекулами среды (внешнее воздействие).

Первая составляющая – сила, действующая на данный атом со стороны всех остальных атомов, она рассчитывается как производная потенциальной энергии. Потенциальную энергию системы можно представить в виде суммы вкладов от различных типов взаимодействий между атомами:

$$U = U_b + U_v + U_\varphi + U_f + U_{qq} + U_{vdw} + U_{Hb} \quad (3)$$

где  $U_b$  – потенциальная энергия валентных связей (4),  $U_v$  – валентных углов (5),  $U_\varphi$  – торсионных углов,  $U_f$  – плоских групп и псевдоторсионных углов (6),  $U_{qq}$  – кулоновских сил (7),  $U_{vdw}$  – взаимодействий Ван-дер-Ваальса (8),  $U_{Hb}$  – водородных связей (9).

Энергия валентных связей описывается параболическим потенциалом (4):

$$U_b(r) = \frac{1}{2} \sum_{i=1}^{N_b} K_{b,i} (r_i - r_{0,i})^2 \quad (4)$$

где  $K_{b,i}$  – эффективная жёсткость валентной связи,  $i$  – номер связи в молекуле,  $N_b$  – полное число валентных связей,  $r_i$  – длина связи,  $r_{0,i}$  – равновесная длина связи.

Замена реального потенциала, описывающего валентные взаимодействия, на параболический оправдана тем, что при комнатных температурах колебания валентных связей малы. В то же время параболическое приближение делает возможным вести расчёт при высоких температурах без разрыва связи, который был бы неизбежен при более аккуратном приближении (Рис. 3).

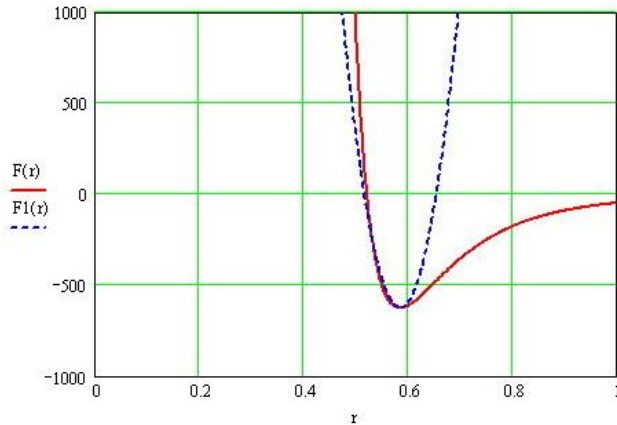


Рис. 3. Сравнение реального (показан сплошной красной линией) и параболического (показан пунктирной синей линией) потенциалов для валентной связи. Числовые параметры взяты произвольными.

Потенциальная энергия валентных углов также представляется гармоническим потенциалом:

$$U_v(\alpha) = \frac{1}{2} \sum_{i=1}^{N_v} K_{v,i} (\alpha_i - \alpha_{0,i})^2 \quad (5)$$

где  $K_{v,i}$  – эффективная упругость валентного угла,  $i$  – номер валентного угла,  $N_v$  – полное число валентных углов,  $\alpha_i$  – значение валентного угла,  $\alpha_{0,i}$  – его равновесное значение.

Потенциальная энергия для торсионных углов, плоских групп и псевдоторсионных углов задается общим выражением (6), представляющим собой ряд Фурье. Было установлено, что во всех случаях достаточно оставлять не более четырёх членов ряда (включая нулевой).

$$U_\varphi(\alpha) = \frac{1}{2} \sum_{\varphi=1}^{\phi} \sum_{l=1}^L K_{\varphi,l} \left[ 1 + g_{\varphi,l} \cos(n_{\varphi,l} \alpha_\varphi) \right] \quad (6)$$

где  $K_{\varphi,l}$  – константа,  $\varphi$  – номер торсионного угла,  $l$  – номер гармоники,  $g_{\varphi,l}$  – вклад гармоники в потенциал торсионного угла ( $-1 < g_{\varphi,l} < 1$ ),  $n_{\varphi,l}$  – кратность гармоники. Потенциалы  $U_f$  и  $U_\varphi$  отличаются константами.

Потенциальная энергия взаимодействия заряженных атомов характеризуется кулоновским потенциалом:

$$U_{qq}(\vec{r}_i, \vec{r}_j) = \sum_{i \neq j} \frac{q_i q_j}{\varepsilon r_{i,j}} \quad (7)$$

где  $\vec{r}_i, \vec{r}_j$  – радиус-векторы взаимодействующих атомов,  $q_i, q_j$  – их парциальные заряды,  $\varepsilon$  – диэлектрическая проницаемость среды (для вакуума  $\varepsilon = 1$ ),  $r_{i,j} = |\vec{r}_i - \vec{r}_j|$  – расстояние между атомами.

Взаимодействие между атомами, не связанными валентной связью, представляет собой силы отталкивания и притяжения электродинамической природы (Ван-дер-ваальсов потенциал). Общая энергия взаимодействия приближенно описывается с помощью потенциала Леннарда-Джонса (8) или потенциала для водородной связи (9).

$$U_{vw}(\vec{r}_i, \vec{r}_j) = \sum_{i \neq j} \left[ \frac{A}{r_{i,j}^{12}} - \frac{B}{r_{i,j}^6} \right] \quad (8)$$

$$U_{Hb}(\vec{r}_i, \vec{r}_j) = \sum_{i \neq j} \left[ \frac{A'}{r_{i,j}^{12}} - \frac{B'}{r_{i,j}^{10}} \right] \quad (9)$$

где  $A$  и  $B$ ,  $A'$  и  $B'$  – константы, определяющие глубину потенциальной ямы и расположение её минимума,  $\vec{r}_i, \vec{r}_j$  – радиус-векторы взаимодействующих атомов,  $r_{i,j} = |\vec{r}_i - \vec{r}_j|$  – расстояние между атомами.

Силовым полем называется файл, содержащий в себе все указанные выше параметры межатомных взаимодействий. Существует большое число силовых полей, параметризованных для различных типов молекул и уровней детализации, и исследователь может выбрать силовое поле в соответствии с типом решаемой задачи.



## Описание файлов

Для проведения молекулярно-динамических расчетов с помощью программы Gromacs требуется несколько основных файлов.

ffG43a2.atp, ffG43a2bon.itp	файлы с перечнем типов атомов и перечнем типов связей, валентных и двугранных углов в силовом поле ffG43a2 соответственно
test.top	файл с топологией системы для пробного запуска МД лиганда
em_test.sh, heat_test.sh, md_test.sh	файлы со скриптами для запуска минимизации энергии пробной системы, её нагревания и расчета МД соответственно
em.sh, heat.sh, mdpr.sh, md.sh	файлы со скриптами для запуска минимизации энергии мембранной системы, её нагревания, релаксации липидов и расчета МД соответственно
ions.mdp, em. mdp, heat. mdp, mdpr. mdp, md. mdp	файлы со параметрами для добавления ионов и выполнения расчетов по минимизации энергии мембранной системы, её нагреванию, релаксации липидов и расчета МД соответственно
porc288_pore19.pdb, porc.itp	файлы с координатами липидного бислоя из молекул ПОФХ с вырезанной порой радиусом 19А и топология такой молекулы
obrez.awk	скрипт для удаления внутримембранной воды
system.top, start.top	Файлы с описанием топологии мембранной системы

## Порядок выполнения работы

Предполагается, что работа выполняется в ОС Linux или аналогичной ей, а у студентов есть навыки работы с подобными ОС.

### 1. Подготовка структурных файлов

На сайте банка данных о белковых структурах PDB (Protein Databank <http://www.rcsb.org/pdb/home/home.do>) найдите страницу, посвященную структуре бета2-адренэргического рецептора человека с PDB ID 2RH1. Изучите эту страницу и скопируйте структурный файл в свою рабочую директорию.

Запустите программу визуализации VMD, набрав `vmd` в командной строке терминала. Ознакомьтесь с особенностями программы, используя учебные материалы, представленные на сайте <http://www.ks.uiuc.edu/Research/vmd/current/docs.html>. Визуализуйте структуру AP, изобразив белковую цепь в ленточном представлении (Cartoon или Ribbons), а молекулу лиганда в стержневом (Licorice). Проследите ход главной цепи молекулы, обратите внимание на расположение небелковых молекул в данной структуре.

Откройте структурный файл в текстовом редакторе и изучите его содержимое, руководствуясь при этом официальной документацией по формату \*.pdb (<http://www.wwpdb.org/documentation/format32/v3.2.html>). Найдите и удалите строки, соответствующие остаткам с 1002-го по 1161-й. Этих остатков нет в природном рецепторе и они были введены в рецептор исключительно с целью улучшения склонности его молекул к кристаллизации. Также удалите все строки, не относящиеся к координатному блоку, и строки с координатами небелковых молекул, кроме молекулы лиганда (каразолола - CAU). Сохраните измененный структурный файл под именем 2RH1\_A.pdb. Закройте текстовый редактор.

Откройте файл 2RH1\_A.pdb с помощью подходящей программы (например, Maestro), ориентируйте его по оси OZ и сохраните под именем 2RH1\_A\_OZ.pdb. В текстовом редакторе разделите этот файл на две части – рецептор и лиганд – и сохраните их под именами 2RH1\_rec\_OZ.pdb и 2RH1\_lig\_OZ.pdb соответственно.

Файл со структурой фосфолипидного бислоя подготовлен заранее. Ознакомьтесь с ним с помощью текстового и молекулярного редакторов, особое внимание уделите строению одиночного фосфолипида – пальмитоилолеилфосфатидилхолина (ПОФХ, POPC) (Рис. 4).

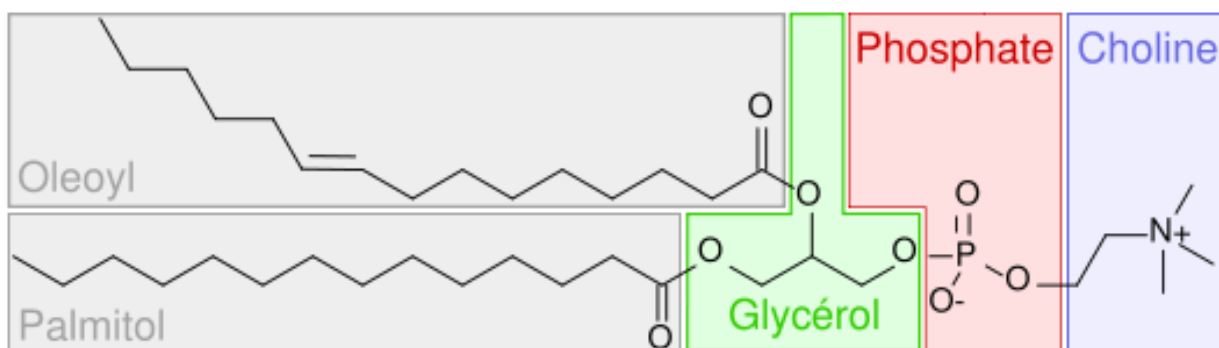


Рис. 4. Схема строения пальмитоилолеилфосфатидилхолина (ПОФХ, POPC).

## 2. Создание топологии лиганда

Программами пакета Gromacs предусмотрен анализ белковых структур, состоящих из стандартных аминокислот, и автоматическое создание файлов с их топологией – т.е. перечнем атомов, связей, валентных и двугранных углов с присвоенными параметрами силового поля. Однако для соединений, для которых нет традиционной номенклатуры атомов, автоматическое присвоение параметров силового поля представляется затруднительным и в настоящее время топологию таких соединений приходится описывать вручную. Существенную помощь при этом может оказать веб-сервис PRODRG (<http://davapc1.bioch.dundee.ac.uk/prodrg/>), в котором реализована возможность создания «полуфабриката» файла топологии по структурному файлу. Создайте такой полуфабрикат для молекулы каразалола и сохраните его в рабочей директории под названием 2RH1\_lig\_OZ.top. Среди строк этого файла содержатся, например, такие:

```
[ atoms ]
; nr      type  resnr resid  atom  cgnr   charge   mass
  1       CH3   1  CAU    C21    1     0.001  15.0350
  2       CH1   1  CAU    C20    1     0.006  13.0190
  ...
  14      CR61  1  CAU    C12    3    -0.035  13.0190
  ...
[ bonds ]
; ai  aj  fu    c0, c1, ...
  1   2   1    0.153  334720.0    0.153  334720.0 ;  C21  C20
  ...
```

Обратите внимание на типы атомов в соответствующем блоке. Изучение файла ffG43a2.atp показывает, что типы атомов CH3 и CH1 в силовом поле ffG43a2 есть, а тип CR61 отсутствует и должен быть заменен на CR1. Аналогично проверьте остальные типы

атомов. Далее, руководствуясь файлом ffG43a2bon.itp присвойте правильные типы связей, а также валентных и двугранных углов. Так, поскольку атомы C21 и C20 относятся к типу CHn и имеют межатомное расстояние 0,153 нм, то связь между ними имеет тип gb\_26 и вышеприведенная строка из блока [ bonds ] будет выглядеть следующим образом:

```
1 2 1 gb_26 ; C21 C20
```

Аналогично отредактируйте остальные строки в файле с топологией лиганда. Сохраните его под именем 2RH1\_lig\_OZ.itp.

### 3. Создание двухкомпонентной расчетной ячейки и проведение для неё расчетов МД

Создайте на дисковом пространстве вычислительного кластера еще одну рабочую директорию и скопируйте в неё с ПК файлы 2RH1\_A\_OZ.pdb, 2RH1\_lig\_OZ.pdb и 2RH1\_lig\_OZ.itp.

Для проверки правильности топологии лиганда можно запустить пробный расчет МД лиганда в водной среде. Для этого сначала поместите центр масс лиганда в начало координат с помощью команды

```
editconf -f 2RH1_lig_OZ.pdb -o 2RH1_lig_OZ-c.pdb -box 5.0 5.0 5.0 -c
```

(обратите внимание на размеры ячейки 5.0 x 5.0 x 5.0 нм, которые в данном случае определяются параметрами используемого алгоритма и не могут быть меньше указанных размеров).

Затем перенумеруйте и переименуйте атомы, а также отсортируйте строки в 2RH1\_lig\_OZ-c.pdb согласно 2RH1\_lig\_OZ.itp (атомы водорода, содержащиеся в 2RH1\_lig\_OZ.itp, должны быть и в 2RH1\_lig\_OZ-c\_renum.pdb).

Создайте расчетную ячейку, поместив молекулу лиганда в кубик из молекул воды:

```
genbox -cp 2RH1_lig_OZ-c_renum.pdb -cs -o test.gro
```

Исправьте файл топологии вашей системы test.top, указав в нем правильные имена упоминаемых файлов и количество молекул воды.

Запустите скрипт с командами минимизации энергии полученной системы em\_test.sh, затем нагревания heat\_test.sh и собственно молекулярной динамики md\_test.sh.

По окончании расчета визуализуйте систему с помощью VMD и рассмотрите изменение конформации молекулы лиганда в процессе МД. Если она остается стабильной, то параметры силового поля в топологии лиганда отнесены правильно и можно приступить к расчету МД комплекса рецептор-лиганд в бислое.

#### 4. Создание четырехкомпонентной расчетной ячейки и проведение для неё расчетов МД

Аналогично предыдущему заданию поместите центр масс комплекса рецептор-лиганд в начало координат:

```
editconf -f 2RH1_A_OZ.pdb -o 2RH1_A_OZ-c.pdb -box 9.62 9.62 8.00 -c
```

(в данном случае размеры ячейки 9.62 x 9.62 x 8.00 нм определяются типом используемого бислоя (ПОФХ\_288) и размерами встраиваемого комплекса – он должен убираться с запасом в 3 – 5 Å);

Вырежьте в текстовом редакторе лиганд из файла 2RH1\_A\_OZ-c.pdb в файл 2RH1\_lig\_OZ-c.pdb, рецептор сохраните как 2RH1\_rec\_OZ-c.pdb.

Аналогично предыдущему заданию перенумеруйте и переименуйте атомы, а также отсортируйте строки в 2RH1\_lig\_OZ-c.pdb согласно 2RH1\_lig\_OZ.itp (использование созданного ранее 2RH1\_lig\_OZ-c\_renum.pdb недопустимо – они в разных системах координат!!).

Создайте файл с топологией рецептора:

```
pdb2gmx -f 2RH1_rec_OZ-c.pdb -p 2RH1_rec_OZ-c -o 2RH1_rec_OZ-c_m.pdb -ighn
```

(На запрос о силовом поле выберите ffG43a2 и обратите внимание на не скомпенсированный заряд системы, который будет отображен после исполнения этой команды!).

Переименуйте файл 2RH1\_rec\_OZ-c.top в 2RH1\_rec\_OZ-c.itp и отредактируйте последний (нужно удалить строку #include “ffG43a2.itp” в начале файла и всего строки после первого #endif в конце файла).

Визуализуйте с помощью VMD совместно структуру комплекса (2RH1\_A\_OZ-c.pdb) и бислоя (porc288\_pore19.pdb). Удалите некоторые липидов во избежание

наталкиваний и протыканий с молекулой комплекса (скорректированный файл с бислоем удобно сохранить как `popc288_pore19m.pdb`).

Объедините комплекс рецептор-лиганд с бислоем для создания расчетной ячейки:

```
cat 2RH1_rec_OZ-c_m.pdb 2RH1_lig_OZ-c_renum.pdb
popc288_pore19m.pdb > system.pdb
```

и удалите некоординатные строки (TER, ENDMDL, ...) из середины `system.pdb`.

Закончите создание расчетной ячейки, заполнив свободное пространство молекулами воды:

```
genbox -cp system.pdb -cs -o system.gro
```

Поскольку алгоритм заполняет все свободное пространство, куда могла бы уместиться хотя бы одна молекула воды, а присутствие таких молекул в гидрофобной части бислоя нежелательно, удалите их оттуда с помощью скрипта:

```
awk -f obrez.awk system.gro > system1.gro
```

Подсчитайте число оставшихся молекул воды (SOL) и липидов (POP) в файле `system1.gro`:

```
awk 'BEGIN{i=0}($0~"SOL"){i++}END{print i/3}' system1.gro
awk 'BEGIN{i=0}($0~"POP"){i++}END{print i/52}' system1.gro
```

Исправьте соответствующие значения в файле топологии системы `system.top` – должны быть правильно указаны `2RH1_rec_OZ-c.itp`, `2RH1_lig_OZ-c.itp` и количества молекул воды и липидов.

При наличии не скомпенсированного заряда необходимо добавить противоионы:

```
grompp -p system.top -c system1.gro -f tmp.mdp -o system1.tpr
genion -s system3.tpr -o start.gro -nn 3      (добавлено 3 иона Cl)
```

И исправить файл `start.top`, правильно указав в нем число всевозможных молекул и ионов.

Запустите скрипт с командами минимизации энергии полученной системы `em.sh`, затем нагревания `heat.sh`, релаксации липидов `mdpr.sh` и собственно молекулярной динамики `md.sh`.

По окончании расчета визуализуйте систему с помощью VMD (Рис. 5) и рассмотрите изменение конформации молекул рецептора и лиганда в процессе МД.

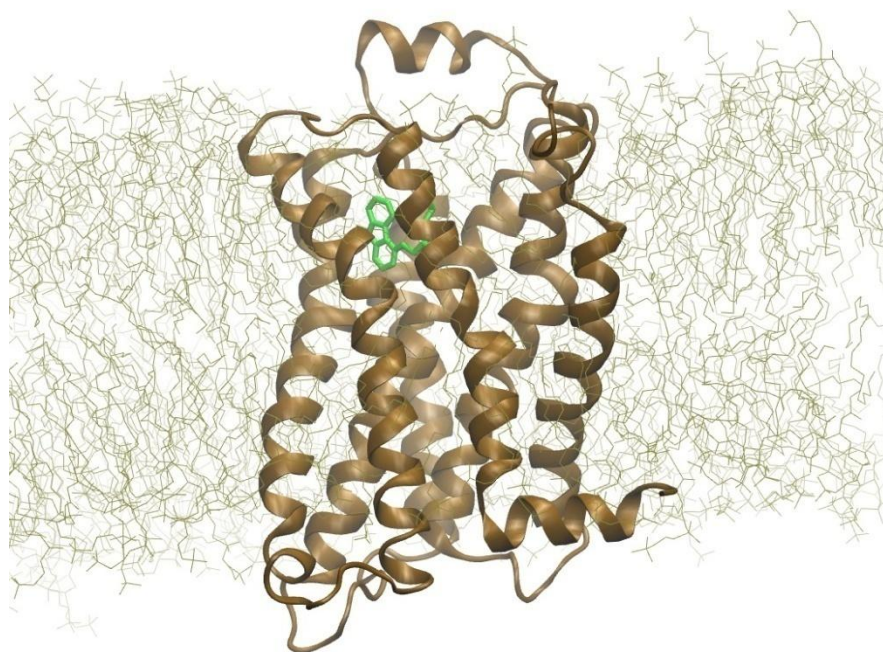


Рис. 5. Расположение бета2-адренэргического рецептора (показан в ленточном представлении) в комплексе с каразололом (показан зеленым) относительно липидного бислоя.

Начиная с версии 4.5 в пакете Gromacs доступна функция `g_membed`, выполняющая встраивание белка в липидный бислой. Освойте её самостоятельно.

## 5. Изучение влияния начальных условий на результаты расчетов

Как и в любом физическом эксперименте, в расчетах МД правильный выбор начальных условий оказывает решающее влияние на достоверность результатов. В данной работе предлагается проварьировать следующие начальные условия:

- a) температуру расчета;
- b) ионную силу раствора.
- c) ...

*Влияние температуры.* Температура – это мера тепловой энергии, заключенной в теле или материи, которая, в свою очередь, является результатом хаотического движения

составляющих тело частиц. Из молекулярно-кинетической теории следует, что кинетическая энергия частиц и температура тела связаны соотношением

$$E = \sum_{i=1}^N \frac{m_i v_i^2}{2} = \frac{3}{2} NkT, \quad (10)$$

где  $E$  – суммарная кинетическая энергия частиц,  $N$  – число частиц,  $k$  – постоянная Больцмана, а  $T$  – температура.

В ходе расчета траектории в методе по причине влияния погрешностей вычисления температура системы может изменяться. Для поддержания температуры системы на заданном уровне используются специальные алгоритмы – термостаты. Одним из наиболее ранних способов является использование термостата Берендсена, который широко используется и в настоящее время, однако следует иметь в виду, что его применимость ограничена системами с большим количеством частиц.

Температура, при которой мы желаем произвести вычисления задается в файлах с параметрами (heat. mdp, mdpr. mdp, md. mdp). Для рассматриваемой в настоящей работе системы белка в липидном бислое наиболее подходящей представляется температура 300К, как близкая к физиологической, но в то же время поучительно рассмотреть эволюцию системы и при несколько отличающихся температурах, например, 290К или 310К. Для проведения расчетов с указанными температурами внесите соответствующие изменения файлы с параметрами (\*.mdp).

*Влияние ионной силы раствора.* Ионная сила раствора — это мера интенсивности электрического поля, создаваемого ионами в растворе. Вычисляется по формуле, впервые предложенной Льюисом, и составляет полусумму произведений концентрации каждого иона в растворе на квадрат его заряда:

$$I_c = \frac{1}{2} \sum_{B=1}^n c_B z_B^2, \quad (11)$$

где  $c_B$  — молярные концентрации отдельных ионов (моль/л),  $z_B$  заряды ионов.

Например, для раствора NaCl с концентрацией 0,001 моль/л, в котором присутствуют два вида однозарядных ионов  $\text{Na}^+$  и  $\text{Cl}^-$  с концентрациями также равными 0,001 моль/л, ионная сила будет вычисляться следующим образом:



$I(\text{NaCl}) = 0,5(z^2(\text{Na}^+) \cdot c(\text{Na}^+) + z^2(\text{Cl}^-) \cdot c(\text{Cl}^-)) = 0,5(1^2 \cdot c(\text{NaCl}) + (-1)^2 \cdot c(\text{NaCl})) = c(\text{NaCl})$ ,  
то есть ионная сила в данном случае будет равна концентрации раствора. Для электролитов, в которых присутствуют многозарядные ионы, ионная сила обычно превышает молярность раствора.

Ионная сила раствора имеет большое значения для подбора условий растворимости и осаждения белков. При низкой ионной силе ( $< 0.2\text{M}$ ) растворимость белка увеличивается при повышении концентрации соли, так как экранирование притяжения противоположно заряженных групп приводит к разрыхлению структуры белка. При высокой ионной силе ( $> 0.2\text{M}$ ) - растворимость белка понижается из-за высаливания и обезвоживания. Мембранные белки менее подвержены влиянию ионной силы водного раствора на структуру, но их внеклеточные домены обогащены полярными и заряженными остатками, поэтому изучение влияния концентрации соли на эволюцию системы представляется интересным.

Изотонический раствор хлорида натрия в воде содержит его 0.91% по массе или 0.28% в мольном отношении. Таким образом, если система содержит около 10 000 молекул воды, то 28 из них нужно заменить на ионы  $\text{Na}^+$  и столько же на ионы  $\text{Cl}^-$ . Выполните строгий расчет количества ионов для своей системы. Выполните расчеты с изотонической концентрацией  $\text{NaCl}$ , вдвое больше изотонической и вдвое меньше изотонической ().

## 6. Анализ полученной траектории

По окончании вычислений можно изучить стабильность и подвижность различных компонентов системы в процессе МД. Нас, прежде всего, будут интересовать рецептор и лиганд.

Кристаллографические структуры белков получают в специфических условиях, достаточно несхожих с физиологическими. Поэтому представляется вполне очевидным, что в процессе МД конформация белка изменится. Для прояснения вопроса о том, насколько сильно это изменение и каких участков белка оно касается, можно вычислить среднеквадратичное отклонение (СКО) различных типов атомов белка от их начального состояния, а также среднеквадратичные флуктуации (СКФ) атомов в зависимости от номера остатка, которому они принадлежат.

$$CKO = \sqrt{\frac{\sum_i \left( (x_i - x_i^0)^2 + (y_i - y_i^0)^2 + (z_i - z_i^0)^2 \right)}{n}},$$

где сумма берется по всем  $n$  атомам молекулы,  $x_i, y_i, z_i$  и  $x_i^0, y_i^0, z_i^0$  – координаты атома  $i$  в рассматриваемой и референсной структурах, соответственно.

Традиционно оба эти показателя, СКО и СКФ, рассчитываются для  $C\alpha$ -атомов белковой цепи, однако можно рассчитывать их и для всех атомов основной цепи белка или даже вообще всех тяжелых атомов белка. Атомы водорода не принимаются во внимание, поскольку определение их координат недоступно методу рентгено-структурного анализа (РСА) и поэтому они отсутствуют в кристаллографической структуре. Для расчета СКО выполните следующую команду:

```
g_rms -s mdpr.gro -f md.trr
```

В появившемся приглашении укажите номер набора атомов, по которому будет выполнено совмещение структур перед расчетом СКО, а затем номер набора, для которого требуется выполнить сам расчет. Получившийся файл `rmsd.xvg` можно визуализировать в программе `xmgrace`:

```
xmgrace rmsd.xvg &
```

Получившийся график будет иметь следующий вид (Рис. 6). Хорошо видно, что график СКО выходит на плато уже через 500 пс, что говорит о том, что система стабилизировалась.

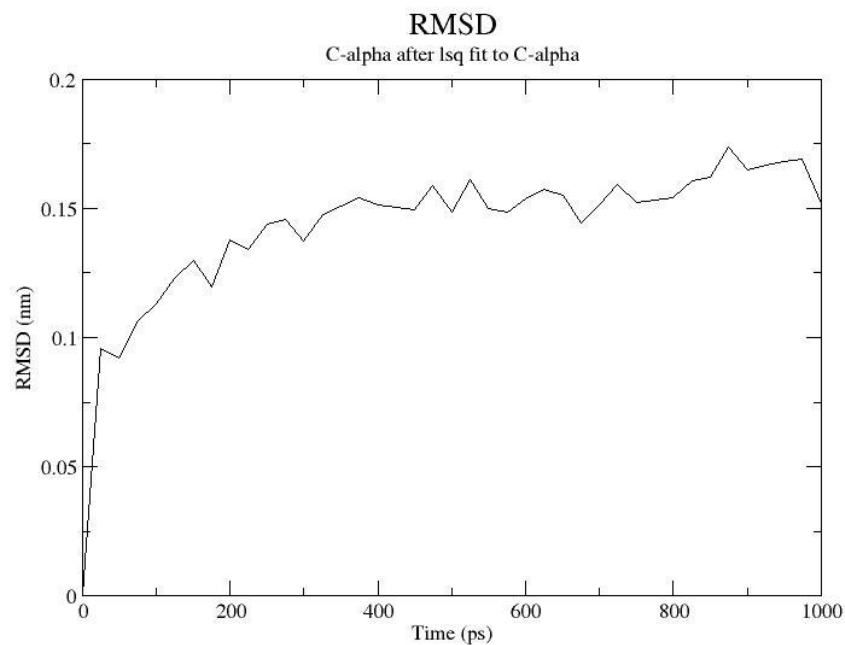


Рис. 6. Среднеквадратичное отклонение C $\alpha$ -атомов структуры бета2-адренэргического рецептора от начальной конформации в процессе молекулярной динамики. По оси абсцисс показано время в пикосекундах, по оси ординат отклонение в нанометрах.

Для расчета флуктуаций выполните следующую команду:

```
g_rmsf -s mdpr.gro -f md.trr
```

В появившемся приглашении укажите номер набора, для которого требуется выполнить сам расчет. Получившийся файл `rmsf.xvg` можно визуализовать аналогично предыдущему (Рис. 7). На получившемся рисунке прекрасно видна неравномерная подвижность атомов в процессе МД. Данный факт вам предлагается объяснить самостоятельно.

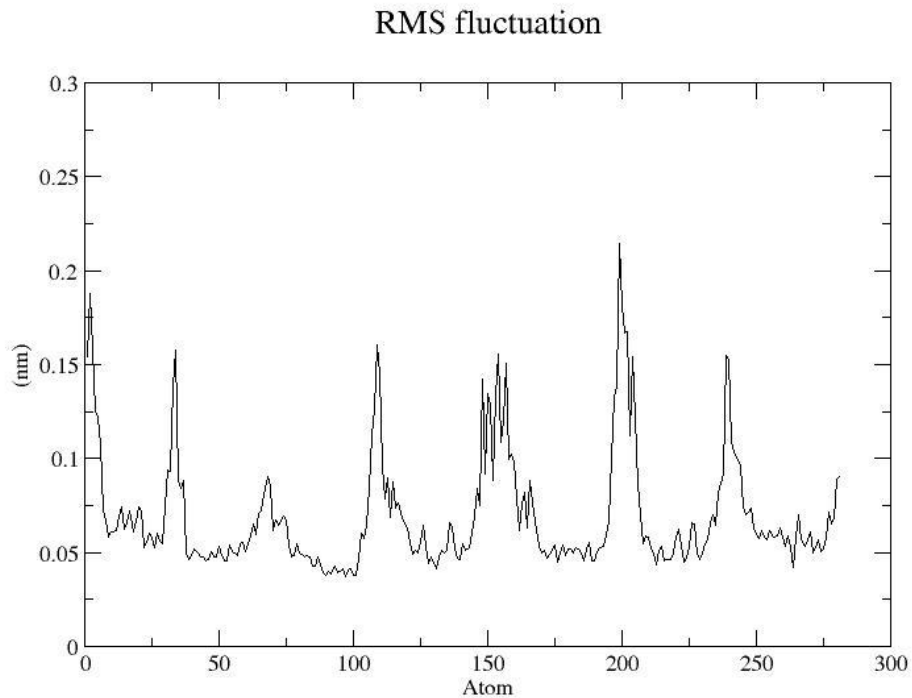


Рис. 7. Среднеквадратичные флуктуации Ca-атомов структуры бета2-адренэргического рецептора от их положения в начальной конформации в процессе молекулярной динамики. По оси абсцисс показан номер остатка, по оси ординат отклонение в нанометрах.

Еще одним показателем стабильности структуры служит величина общей энергии системы. Для её расчета выполните команду:

```
g_energy -f md.edr
```

На получившемся рисунке (Рис. 8) видно, что энергия системы, претерпевая в первые моменты времени резкое уменьшение, в дальнейшем стабилизируется на постоянном уровне аналогично СКО-Sa.

## Gromacs Energies

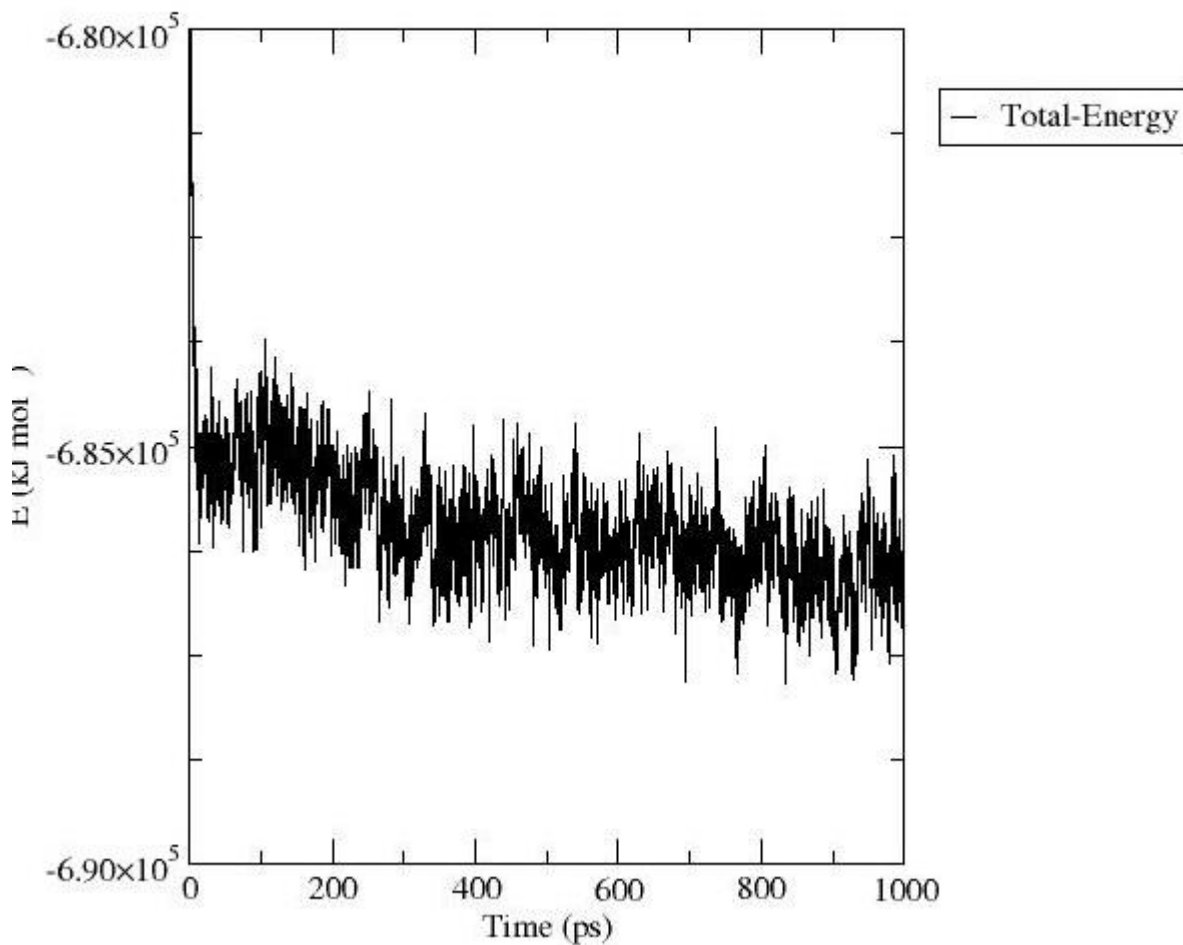


Рис. 8. Изменение общей энергии системы в процессе МД. По оси абсцисс показано время в пикосекундах, по оси ординат энергия в кКал/моль.

При необходимости МД траектория может быть конвертирована в формат \*.pdb следующей командой:

```
trjconv -f md.trr -s md.tpr -o md/md_receptor_dt200.pdb -dt 200  
-fit rot+trans -sep -pbc nojump
```

*(директорию md нужно создать предварительно; -dt 200 – интервал нарезания фреймов в пикосекундах; -fit rot+trans – совмещение фреймов с нулевым при помощи указанных преобразований; -sep – запись каждого фрейма в свой файл; -pbc*

nojump – *исправление возможного перескока рецептора или лиганда в соседнюю ячейку*);

```
trjconv -f md.trr -s md.tpr -o md/md_ligand_dt200.pdb -dt 200 -  
fit rot+trans -sep -pbc nojump
```

Такое разбиение фреймов на рецептор и лиганд может быть полезно, например, если предполагается последующее изучение различных взаимодействий между ними.

## Вопросы

1. Каково строение бета2-адренэргического рецептора и какую функциональную роль он выполняет? Приведите примеры его лигандов.
2. Каковы физическая природа и функциональный вид межмолекулярных взаимодействий?
3. Какие данные можно извлечь из базы данных белковых структур PDB? Каково строение файла в формате pdb?
4. Как устроена мембрана животной клетки? Что такое фосфолипиды? Приведите примеры заряженных и незаряженных фосфолипидов.

## ЛИТЕРАТУРА

1. К.В. Шайтан, С.С. Сарайкин, Метод молекулярной динамики, 1999.
2. Описание истории и методов МД  
[http://en.wikipedia.org/wiki/Molecular\\_dynamics](http://en.wikipedia.org/wiki/Molecular_dynamics)
3. Описание программного пакета Gromacs для выполнения расчетов МД  
<http://www.gromacs.org/Documentation/Manual>
4. Описание программы VMD для визуализации молекулярных структур  
<http://www.ks.uiuc.edu/Research/vmd/>
5. Описание бета2-адренэргического рецептора  
[http://ru.wikipedia.org/wiki/Бета-2\\_адренорецептор](http://ru.wikipedia.org/wiki/Бета-2_адренорецептор)  
[http://en.wikipedia.org/wiki/Adrenergic\\_receptor,\\_beta\\_2](http://en.wikipedia.org/wiki/Adrenergic_receptor,_beta_2)
6. Описание фосфолипидов  
<http://ru.wikipedia.org/wiki/Фосфолипиды>  
<http://en.wikipedia.org/wiki/Phospholipid>