

На правах рукописи

БУСЛАЕВ ПАВЕЛ ИЛЬИЧ

**Применение метода главных компонент для анализа конформаций
молекул липидов, полученных методом молекулярной динамики**

Специальность 03.01.02 – Биофизика

АВТОРЕФЕРАТ

диссертации на соискание учёной степени
кандидата физико-математических наук

Москва – 2018

Работа выполнена в Федеральном государственном автономном образовательном учреждении высшего образования «Московский физико-технический институт (государственный университет)» (МФТИ).

Научный руководитель:

кандидат физико-математических наук

Гущин Иван Юрьевич

Официальные оппоненты:

Шайтан Константин Вольдемарович, доктор физико-математических наук профессор, Федеральное государственное бюджетное образовательное учреждение высшего образования "Московский государственный университет имени М.В.Ломоносова", биологический факультет, профессор.

Гуртовенко Андрей Алексеевич, доктор физико-математических наук профессор РАН, Федеральное государственное бюджетное учреждение науки «Институт высокомолекулярных соединений Российской академии наук», ведущий научный сотрудник.

Ведущая организация:

Федеральное государственное бюджетное учреждение науки Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова Российской академии наук.

Защита состоится 20 сентября 2018 года в 16:00 на заседании диссертационного совета Д 212.156.06 на базе Московского физико-технического института (государственного университета) по адресу: 141701, Московская область, г. Долгопрудный, Институтский переулок, д.9.

С диссертацией можно ознакомиться в библиотеке МФТИ, а также на интернет-сайте диссертационного совета <https://mipt.ru/education/post-graduate/D212-156-06/>.

Автореферат диссертации разослан «__» июня 2018 г.

Учёный секретарь
диссертационного совета

Гец Артем Викторович

Общая характеристика работы

Актуальность работы

Биологические мембраны являются одними из важнейших частей живых клеток, выполняя большое количество активных и пассивных функций. Входящие в их состав мембранные белки являются мишенями для более чем 50% представленных на рынке лекарств. При этом биологические мембраны являются также и одними из наиболее сложных для изучения как экспериментальными, так и вычислительными методами систем.

Вычислительные методы в последнее время играют все большую роль в исследовании биологических объектов, что было отмечено Нобелевской премией по химии 2013 года (Карплус, Левитт и Варшель). Одной из основных задач в усовершенствовании вычислительных методов является повышение точности и скорости получения результатов расчёта. При решении данных задач полезными оказываются методы обработки больших данных и машинного обучения.

В настоящее время моделирование методом молекулярной динамики широко используется в научной среде, позволяя получить информацию, дополняющую экспериментальные подходы. Моделирование используется для изучения большого количества систем таких как мицеллы, липидные бислои и везикулы, липидные кубические и гексагональные фазы. Такие важные физиологические процессы, как образование пор в мембране, слияние и разделение мембран и др. активно изучаются при помощи моделирования методом молекулярной динамики. Также данный метод позволяет изучать влияние добавок, начиная от мелких молекул и вплоть до мембранных белков, на мембранную систему.

Анализ конформаций молекул липидов методом главных компонент, являющийся темой данного исследования – это новейший способ анализа больших данных, получаемых при моделировании мембранных систем методом молекулярной динамики. Разработанный метод позволит уточнить используемые для моделирования силовые поля, количественно изучать поведение разнообразных мембранных систем, тем самым упрощая и ускоряя, в том числе, процесс разработки лекарств, действующих на биологические мембраны и их компоненты.

Степень разработанности

Моделирование биологических мембран затруднено необходимостью описывать поведение не только водного растворителя, окружающего интересующие нас объекты, но и липидов, составляющих основу мембраны. Один из важных факторов успешного моделирования – применение правильного набора параметров, описывающих взаимодействие между атомами системы (силовое поле). Несмотря на то, что на данный момент разработано большое количество силовых полей, отсутствует количественная оценка влияния тех или иных параметров непосредственно на конформации липидных молекул.

Для проверки липидных силовых полей используют величины, которые можно измерить экспериментально: плотности, поверхностное натяжение, модули упругости, потенциалы диполей, данные о диффузии, данные рентгеновской дифракции, параметры порядка ядер дейтерия, полученные методом ЯМР, а также времена спин-релаксации. Все эти характеристики являются макроскопическими и практически не содержат информации о конформациях отдельных молекул липидов. Поэтому при исследовании мембранных систем при помощи моделирования методом молекулярной динамики в основном рассчитываются лишь макроскопические параметры: площадь, приходящаяся на молекулу липида, распределения электронных плотностей для различных химических групп, диффузия молекулы липида как целого. Другие рассчитываемые величины, как например параметры порядка, описывают определенные движения молекулы, но не движения молекулы целиком.

Наконец, в последнее время все больше внимания уделяется сравнению наборов параметров для мембранных систем, используемых для расчёта взаимодействия между атомами при моделировании методом молекулярной динамики. Однако в этих работах внимание уделяется лишь усреднённым макропараметрам, а не локальным движениям молекул липидов. Алгоритмы для количественного сравнения конформаций липидов, особенно полученных при моделировании с использованием силовых полей различного разрешения, на данный момент отсутствуют.

Цель работы

Целями данной работы являются:

1. Проведение моделирования мембранных систем методом молекулярной динамики с использованием различных силовых полей для последующего анализа и сравнения получаемых результатов
2. Разработка подхода к анализу конформаций и динамики молекул липидов с помощью метода главных компонент
3. Разработка подхода к прямому количественному сравнению конформаций и динамики молекул липидов, получаемых в различных моделированиях, с помощью метода главных компонент
4. Изучение влияния выбранного силового поля на получаемые конформации и динамику молекул липидов
5. Изучение влияния уровня разрешения силовых полей на получаемые конформации и динамику молекул липидов

Основные задачи исследования

Для проверки возможности применения метода главных компонент для анализа конформаций и динамики молекул липидов необходимо получить траектории липидных бислоев, рассчитанные с помощью различных силовых полей. В частности, необходимо моделирование бислойных систем с использованием различных современных силовых полей.

Следующая задача – применение метода главных компонент для анализа конформаций молекул липидов, получаемых при моделировании. На этом этапе необходимо разработать алгоритмы, позволяющие количественно описать конформационное пространство, доступное молекулам липидов при моделировании, и алгоритмы, которые позволят оценивать динамику процесса обхода доступных конформаций. Далее, необходимо разработать алгоритмы, позволяющие напрямую сравнивать конформации и динамику молекул липидов, получаемые при моделировании с использованием различных силовых полей.

Разработанные алгоритмы анализа и сравнения далее необходимо применить для рассчитанных траекторий. На этом этапе необходимо провести

тщательную проверку и обработку разработанных алгоритмов и получить финальные результаты. Интерпретация полученных результатов поможет найти слабые места тестируемых силовых полей.

Один из возможных способов ускорения расчётов при моделировании методом молекулярной динамики – это использование крупнозернистых силовых полей. При таком подходе несколько атомов молекулы объединяются в одну частицу-«псевдоатом», и рассматривается только взаимодействие этих частиц. Ускорение получается за счёт заметного уменьшения количества взаимодействующих частиц, однако, при этом очевидно теряются некоторые детали. Заключительной задачей данной работы, является изучение влияния уровня разрешения силового поля на конформации и динамику молекул липидов.

Научная новизна

В данной работе впервые представлен метод анализа конформаций молекул липидов с помощью метода главных компонент, единственный из известных на сегодняшний день методов анализа молекул липидов, позволяющий количественно описывать конформационную динамику отдельных молекул липидов. Результаты анализа показывают принципиальные различия между силовыми полями, применяемыми для моделирования молекул липидов, а также позволяют оценить характерные временные масштабы процессов, происходящих с отдельными молекулами липидов. Также, анализ позволил напрямую сравнить конформации липидов, промоделированных с помощью силовых полей различного разрешения.

Практическая значимость

Практическая ценность результатов диссертации заключается в том, что:

1. Численно охарактеризовано пространство конформаций, доступное молекуле липида. Показано, какое минимальное число степеней свободы необходимо для эффективного описания конформаций молекул липидов. Показано, что основной степенью свободы для всех изученных

липидных силовых полей является движение углеводородных хвостов навстречу друг другу в плоскости молекулы липида.

2. Численно определены характерные временные масштабы, необходимые для эффективного обхода пространства доступных конформаций. Сформулированы рекомендации по выбору минимальной длительности моделирования.
3. Показано, как численно и с сохранением всей информации сравнить конформации и динамику молекул липидов, промоделированных с использованием различных силовых полей; как численно измерить влияние параметров системы, таких как температура, кривизна, добавки, на поведение липидов.
4. Показано, что разработанный подход позволяет выявлять критические различия между силовыми полями. Показаны различия между современными широко используемыми силовыми полями.
5. Устранение обнаруженных различий может быть использовано для дальнейшего уточнения и улучшения существующих липидных силовых полей, в том числе силовых полей низкого разрешения.
6. Показано, что различия между силовыми полями последнего поколения существенно меньше, чем между предыдущими версиями силовых полей. Выделены силовые поля, дающие сильно отличающиеся и, скорее всего, недостоверные результаты.
7. Показано, что крупнозернистое силовое поле сохраняет от 75 до 100% наиболее коллективных конформационных изменений, которые описывают более 90% вариации конформационных изменений молекулы липида.
8. Показано, что ускорение динамики при использовании крупнозернистых силовых полей происходит неравномерно: для коллективных движений ускорение не так велико, как для локальных конформационных изменений молекулы липида.

Методы исследования

В работе были использованы следующие методы и технологии:

1. Моделирование методом молекулярной динамики
2. Структурное выравнивание методом Кабша

3. Методы линейной алгебры для расчёта матриц ковариации и главных компонент
4. Статистические методы для анализа результатов
5. Программирование на языках Python, Mathematica, Tcl и C
6. Визуализация молекулярных структур при помощи программного обеспечения PyMOL

Положения, выносимые на защиту

1. Показано, что метод главных компонент может быть использован для изучения конформаций и динамики молекул липидов. Определены основные коллективные движения (степени свободы) атомов липидных молекул и соответствующие им характерные временные масштабы.
2. Показано, что различия между липидными силовыми полями последнего поколения существенно меньше, чем между предыдущими версиями силовых полей. Выделены силовые поля, дающие сильно отличающиеся и, скорее всего, недостоверные результаты. Показано, что крупнозернистое силовое поле сохраняет от 75 до 100% наиболее коллективных конформационных изменений, которые описывают более 90% вариации конформационных изменений молекулы липида.
3. Показано, что на уровне отдельных липидов внесение ненасыщенных связей в углеводородные хвосты молекул липидов приводит к их разупорядочению, тогда как динамика остается практически неизменной; увеличение температуры, наоборот, незначительно влияет на конформации молекул липидов, но существенно ускоряет обход пространства доступных состояний.

Апробация результатов работы

Основные результаты работы докладывались на следующих научных конференциях и семинарах:

1. 2017, 60 конференция МФТИ, Разработка программного обеспечения для автоматического анализа конформаций молекул липидов методом главных компонент

2. 2017, 60 конференция МФТИ, Применение метода главных компонент для исследования динамики липидов в мембранных системах
3. 2017, 60 конференция МФТИ, Моделирование фазовых переходов в липидных системах
4. 2017, The future of biomembrane simulations: hidden pitfalls and future challenges, Principal Components Analysis of Lipid Molecule Conformations in Atomistic and CG simulations”
5. 2017, MARTINI Workshop 2017, “Principal Components Analysis of Lipid Molecule Conformations in Atomistic and CG Simulations”
6. 2016, 59 конференция МФТИ, Статистический анализ динамики липидов в биологической мембране: развитие и перспектива метода главных компонент
7. 2016, Biomembranes2016, Principal components analysis of lipid molecule motions in molecular dynamics simulation
8. 2016, Gordon Research Conference – Computational Chemistry, Principal components analysis of lipid molecule motions in molecular dynamics simulation
9. 2016, Gordon Research Seminar – Computational Chemistry, Principal components analysis of lipid molecule motions in molecular dynamics simulation
10. 2015, 58 конференция МФТИ, Статистический анализ динамики липидов в мембране
11. 2014, Biomembranes2014, Principal components analysis of lipid molecule motions in molecular dynamics simulation
12. 2014, From Computational Biophysics to Systems Biology, Principal components analysis of lipid molecule motions in molecular dynamics simulation
13. 2013, 56 конференция МФТИ, Статистический анализ динамики липидов в мембране

По материалам диссертации опубликовано 3 статьи в рецензируемых журналах, входящих в список, утверждённый Высшей аттестационной комиссией.

Структура и объём работы

По своей структуре диссертационная работа состоит из введения, трёх глав, списка литературы. Объём диссертации составляют 191 страница текста, включая 71 рисунок и 4 таблицы. Список использованных источников содержит 471 наименование.

Основное содержание работы

Диссертационная работа состоит из введения, трёх глав и заключения.

Во введении обоснована актуальность темы диссертационной работы, сформулированы цели, задачи и методы исследования, отражены научная новизна, практическая значимость работы, описана апробация полученных результатов.

Первая глава представляет литературный обзор современного состояния проблем исследования биологических мембран, использования моделирования методом молекулярной динамики и применения метода главных компонент для анализа конформаций биологических молекул.

Вторая глава включает в себя описание материалов и методов, использованных при выполнении диссертационной работы. Описаны процедуры получения начальных структур для проведения моделирования; описаны условия проведения моделирования; описаны алгоритмы, использованные для обработки данных и анализа полученных результатов.

Третья глава состоит из пяти подразделов. В **первом подразделе третьей главы** показано, как метод главных компонент может быть использован для анализа конформаций и динамики молекул липидов, полученных при моделировании, и для сравнения конформаций и динамики молекул липидов, полученных при моделировании с использованием различных силовых полей.

Метод главных компонент применялся к координатам тяжёлых атомов молекул липидов, получаемых при моделировании методом молекулярной динамики, предварительно выровненных относительно общей для всех изучаемых систем опорной (референсной) структуры (рисунок 1). В первую очередь вычислялась матрица ковариации смещений атомов относительно среднего положения, её собственные значения и вектора. Особенность метода

главных компонент заключается в том, что вариация смещений атомов вдоль собственного вектора с наибольшим собственным значением максимальна. Так, оказалось, что для описания 90% конформационных изменений молекулы липида при моделировании методом главных компонент достаточно 12-14 наиболее значимых собственных векторов. При этом первая главная компонента представляет из себя движения гидрофобных хвостов молекулы липида навстречу друг другу, аналогично движению концов ножниц. Остальные движения оказываются менее коллективными и включают разнообразные скручивания и изгибы молекулы (рисунок 1).

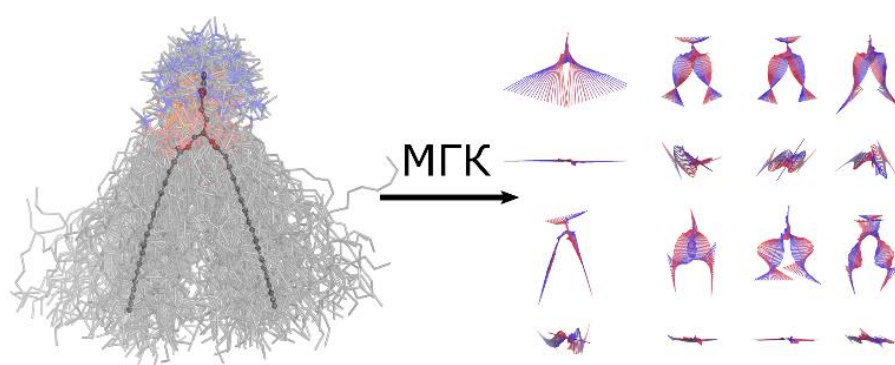


Рисунок 1 – Применение метода главных компонент к конформациям молекул липидов. Первая главная компонента соответствует движению гидрофобных хвостов в плоскости молекулы липида (движение похожее на движения концов ножниц). Следующие движения являются более сложными и включают разнообразные изгибы и скручивания

При проведении моделирования методом молекулярной динамики важную роль играют параметры, описывающие взаимодействия между атомами (силовые поля). Для изучения влияния этих параметров на конформации молекул липидов были проведены моделирования с использованием 8 различных широко используемых силовых полей. Оказалось, что первая главная компонента одинакова для всех моделирований, в то время как остальные главные компоненты несколько отличались.

Метод главных компонент может также быть использован для количественного описания сходств и различий пространств конформаций, доступных молекулам липидов, в моделированиях с использованием

различных силовых полей. Во-первых, можно непосредственно сравнить матрицы ковариации смещений атомов, получаемые в различных моделированиях. Анализ показал, что в целом матрицы ковариаций, получаемые для различных силовых полей, можно сгруппировать в кластеры. Так, например, системы, промоделированные с использованием силовых полей последнего поколения (SLipids, Lipid14 и CHARMM36), позволяют получить более похожие между собой ансамбли конформаций, чем силовые поля предыдущего поколения (рисунок 2). Силовые поля, использующие модель объединённых атомов (атомы водорода объединяются с тяжёлыми атомами, с которыми они ковалентно связаны, в «псевдоатомы») (Berger, GROMOS43A1-S3, GROMOS54A7), также образуют свой отдельный кластер (рисунок 2). Следующий подход, позволяющий сравнить характер движений вдоль главных компонент – это сравнение распределений плотности вероятности проекций координат атомов на главные компоненты. Здесь важно отметить, что такое сравнение возможно только в общем для сравниваемых траекторий базисе. Для всех моделирований оказалось, что распределение вдоль первой главной компоненты сильно отличается от нормального; распределения вдоль других компонент близки к нормальным (рисунок 2).

Последующий анализ показал, что три пика, видимых на распределениях вдоль первой главной компоненты, соответствуют конформациям молекулы липида, в которых гидрофобные хвосты находятся в контакте, разделены одним, или разделены двумя гидрофобными хвостами других липидов, соответственно. Распределения вдоль первой главной компоненты позволяют выявить существенные различия между моделированиями, однако, как и в случае матриц ковариации, видно, что силовые поля последнего поколения дают более схожие результаты, чем силовые поля предыдущего поколения.

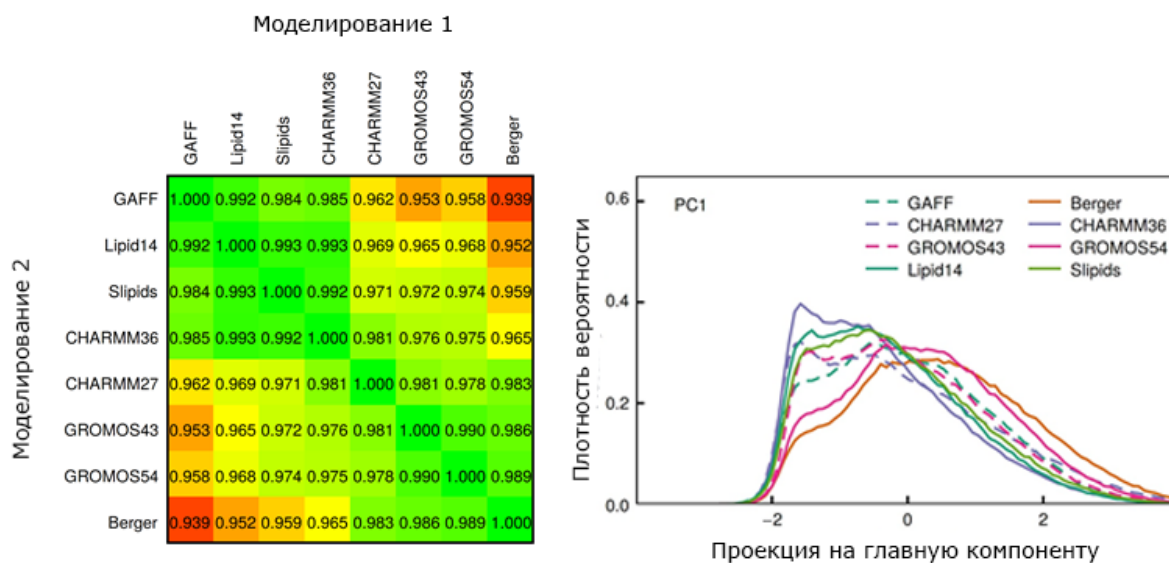


Рисунок 2 – Количественное сравнение пространств доступных конформаций молекул липидов с помощью метода главных компонент. Слева показаны значения коэффициента корреляции Пирсона для пар матриц ковариации, полученных в различных моделированиях. Справа показаны распределения плотности вероятности вдоль первой главной компоненты для липидов, промоделированных с помощью различных силовых полей

Помимо численного описания и сравнения конформаций, метод главных компонент позволяет также изучать динамику молекул липидов. Для этого были использованы два подхода. Во-первых, рассчитывались времена схождения распределений. Для этого, распределения, полученные для полной и частичной (за определённый отрезок времени) траекторий сравнивались между собой с помощью статистического теста Колмогорова-Смирнова. Чем ближе распределения полной и частичной траектории, тем меньше мера Колмогорова-Смирнова. Характерные временные масштабы схождения распределений вдоль главной компоненты вычислялись как время, за которое мера Колмогорова-Смирнова упадёт в e^2 раз относительно начального уровня (рисунок 3). Во-вторых, рассчитывались автокорреляции значений проекций координат атомов на главные компоненты (рисунок 3). Результаты, полученные этими двумя способами, оказались связаны линейным соотношением (время падения автокорреляций приблизительно равно 0,28 времени схождения распределений). При этом, характерные времена

схождения распределений позволяют оценить скорость обхода липидом пространства доступных конформаций и, тем самым, предоставляют релевантную информацию о скорости процесса эквilibрации отдельного липида.

Сравнение характерных временных масштабов движений молекул липидов, промоделированных с помощью различных силовых полей, как и сравнение конформаций, показывает большее сходство между силовыми полями последнего поколения. При этом, даже для силового поля SLipids, для которого молекула липида уравнивается быстрее всего, время уравнивания для первой главной компоненты составляет порядка 100 нс. Поэтому, можно сделать вывод, что минимальная необходимая длительность моделирования мембранных систем составляет по крайней мере 100 нс.

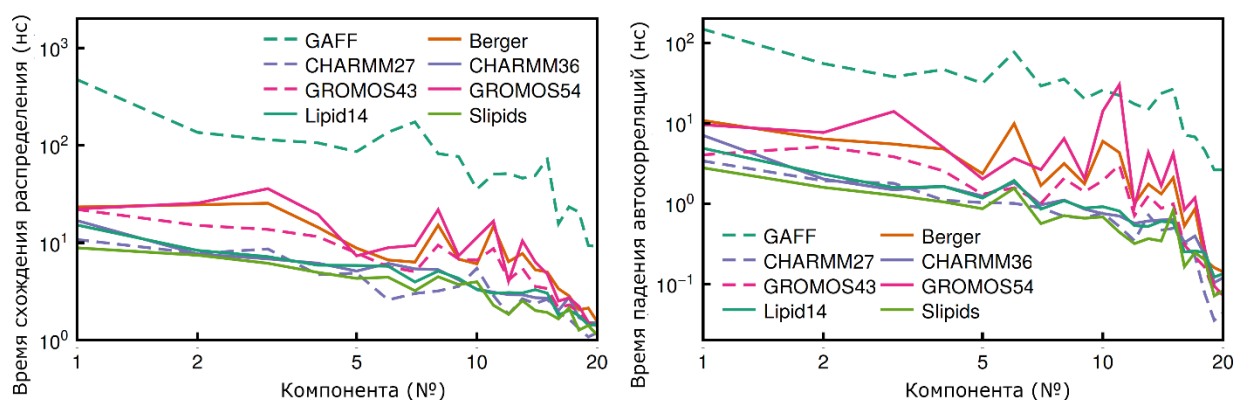


Рисунок 3 - Характерные времена схождения распределений (слева) и падения автокорреляций (справа) для моделирований с использованием разных силовых полей в общем базисе

Во втором подразделе третьей главы показано, как уровень разрешения силового поля влияет на конформации и динамику молекул липидов. В первом разделе рассматривались силовые поля высокого разрешения, для которых каждый тяжёлый атом рассматривается отдельно. При моделировании методом молекулярной динамики число межатомных взаимодействий, а как следствие и вычислительная сложность, растёт как N^2 . Для моделирования систем больших размеров и больших временных масштабов были разработаны силовые поля низкого разрешения, так называемые крупнозернистые силовые поля, в которых несколько атомов

объединялись в группы, которые далее рассматриваются как отдельные взаимодействующие частицы – «псевдоатомы» (рисунок 4). Применение метода главных компонент позволяет сравнить конформацию и динамику молекул липидов, получаемые при моделировании с использованием силовых полей высокого и низкого разрешений.

Оказалось, что первая главная компонента для липидов, промоделированных с помощью силового поля низкого разрешения, является такой же, как и для липидов, промоделированных с помощью силового поля высокого разрешения. Матрицы ковариации также оказались достаточно похожими. Однако, так как число степеней свободы молекулы липида в крупнозернистом представлении уменьшено, какая-то информация о конформациях всё же теряется. Для количественного описания степени сохранения информации о коллективных движениях вдоль главных компонент при трансформации к крупнозернистому представлению была предложена величина

$$\alpha = \frac{\text{RMSD}^{cg}}{\text{RMSD}^{aa}} = \frac{\sqrt{\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n \omega_{ij} (e_{ji}^{aa_{cg}})^2}}{\sqrt{\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n \omega_{ij} (e_j^{aa})^2}} = \frac{\sqrt{\frac{1}{m} \sum_{i=1}^m (e_i^{cg})^2}}{\sqrt{\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n \omega_{ij} (e_j^{aa})^2}} \leq 1 \quad (1)$$

где e_i^{aa} – это n -мерный вектор смещений атомов в полноатомном представлении, $e_{ji}^{aa_{cg}}$ – n -мерный вектор смещений атомов, где смещения каждого атома j совпадает со смещением соответствующего объединённого атома i крупнозернистого представления, e_i^{cg} – m -мерный вектор смещений объединённых атомов в крупнозернистом представлении после трансформации, ω_{ij} – матрица преобразования к крупнозернистому представлению.

Оказалось, что конформации молекул липидов при движении вдоль первых 15 наиболее коллективных собственных компонент сохраняются на 75-100% при трансформации от полноатомного представления CHARMM36 к крупнозернистому представлению Martini (рисунок 4). Остальные, более локальные главные компоненты сохраняются гораздо хуже (рисунок 4).

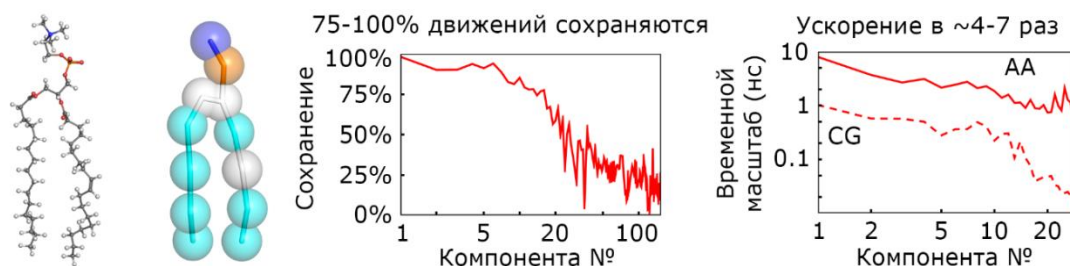


Рисунок 4 – Влияние уровня разрешения на конформации и динамику молекул липидов. Слева показаны модели липидов при полноатомном и крупнозернистом моделировании. В центре – сохранение информации об основных движениях при переходе к крупнозернистому представлению. Справа – изменение характерных временных масштабов при переходе к крупнозернистому представлению

При этом так как первые 12-14 главных компонент позволяют описать более 90% структурного разнообразия молекул липидов, и эти конформации хорошо сохраняются при трансформации к крупнозернистому базису, можно сделать вывод, что крупнозернистое силовое поле Martini позволяет с достаточной степенью точности воспроизводить пространство доступных конформаций молекулы липида. Однако, были выявлены и конформации, достигаемые в полноатомном представлении, но не достигаемые в крупнозернистом, и наоборот.

Динамика молекул липидов при моделировании с помощью крупнозернистых силовых полей ускоряется примерно в 4-7 раз (рисунок 4). При этом интересно, что для коллективных движений это ускорение практически одинаково, а для локализованных движений это ускорение значительно больше. Таким образом, ускорение динамики у молекул липидов происходит неравномерно.

В третьем подразделе третьей главы показано, как насыщенность углеводородных цепей влияет на конформации и динамику молекул липидов, промоделированных с помощью полноатомных и крупнозернистых силовых полей. Для этого сравнивались конформации четырёх фосфатидилхолинов с различными сочетаниями олеоильного и стеароильного гидрофобных хвостов. Оказалось, что позиции олеоильных хвостов в липидах DOPC, SOPC и SOPC совпадают, а позиции стеароильных углеводородных хвостов совпадают в

липидах DSPC, SOPC и OSPC. Сравнения матриц ковариаций, собственных векторов и распределений проекций координат атомов также показывают одинаковые различия в парах DOPC и SOPC, DOPC и OSPC, SOPC и DSPC, и OSPC и DSPC. Различия в парах DOPC и DSPC, и SOPC и OSPC также одинаковы. Эти наблюдения подтверждают гипотезу об аддитивном влиянии гидрофобных углеродных хвостов на конформации молекул липидов. Стоит отметить, что анализ конформаций с помощью метода главных компонент не позволил выявить различий между липидами SOPC и OSPC, хотя, например, экспериментально измеренные температуры фазовых переходов для этих липидов несколько отличаются.

Характерные временные масштабы для исследованных липидов оказались близки, однако наблюдалась и общая закономерность: чем менее насыщены углеводородные цепи липида, тем медленнее он уравнивается.

В четвёртом подразделе третьей главы проведен анализ влияния температуры на конформации и динамику молекул липидов. Показано, что температура практически не влияет на конформации молекул липидов: матрицы ковариации, рассчитанные для моделирований, проведённых при различных температурах, практически совпадают. С понижением температуры, липиды становятся несколько более упорядоченными, распределения проекций координат атомов на главные компоненты сужаются, но никаких других значимых отличий не наблюдается. В то же время, динамика молекул липидов с понижением температуры замедляется. Интересно, что температура одинаково влияет на динамику движений вдоль всех главных компонент.

В пятом подразделе третьей главы проанализированы структурные различия между основными микробными родопсинами. Показано, что размер полости в окрестности ретиналя и Шиффова основания в белках, которые прокачивают ионы, составляет ~ 12 Å. В белках, прокачивающих протон или передающих сигнал, размер этой полости меньше и составляет ~ 10 Å. Возможно, именно это отличие является ключевым и лежит в основе селективности тех и других насосов.

В заключении диссертационной работы представлены основные выводы и перспективы работы, а также приведен библиографический список цитируемой литературы.

Основные выводы и перспективы работы

1. Проведено моделирование липидных бислоев методом молекулярной динамики с использованием силовых полей General Amber Force Field, Berger, CHARMM27, CHARMM36, GROMOS43A1-S3, GROMOS54A7, Slipids, Lipid14, Martini 2.2.
2. Разработан подход к анализу конформаций и динамики молекул липидов на основе метода главных компонент, позволяющий количественно описать доступное молекулам липидов конформационное пространство для молекул липидов, выявить наиболее значимые коллективные движения и рассчитать характерные временные масштабы, на которых происходят структурные изменения.
3. Разработан подход к сравнению конформаций и динамики молекул липидов, получаемых в различных моделированиях, позволяющий эффективно находить различия и сходства в доступных конформационных пространствах молекул липидов в общем базисе и динамики обхода этих конформационных пространств.
4. С помощью разработанного подхода проведено сравнение конформаций и динамики молекул липидов, полученных при моделировании липидных бислойных систем с использованием наиболее применяемых липидных силовых полей. Выявлены сходства и различия между силовыми полями. Показано, что силовые поля последней версии позволяют получать сходные конформации и динамику, тогда как для силовых полей предыдущего поколения конформации различались заметно сильнее.
5. С помощью разработанного подхода проведено сравнение конформаций и динамики молекул липидов, полученных при моделировании липидных бислойных систем с использованием силовых полей различного разрешения. Показано, что основные коллективные движения при моделировании с помощью крупнозернистого силового поля сохраняют от 75 до 100% информации, получаемой при полноатомном моделировании. Найдены конформации, которые достигаются в полноатомном моделировании и не достигаются в крупнозернистом, и наоборот. Показано, что обход конформационного

пространства молекулой липида при крупнозернистом моделировании происходит в 4-7 раз быстрее, чем при полноатомном. При этом показано, что различные коллективные движения ускоряются по-разному.

6. Проведено сравнение доступных конформаций и динамики молекул липидов DOPC, SOPC, OSPC и DSPC. Показано, что пространства доступных конформаций для всех четырёх липидов в значительной мере перекрываются, однако, насыщенные липиды более компактны. Также показано, что при одной и той же температуре динамика различных липидов практически неотличима.
7. Проведено сравнение доступных конформаций и динамики молекул липида DOPC при различных температурах. Показано, что при понижении температуры динамика всех коллективных движений замедляется одинаково.
8. В будущем разработанный подход может быть использован для уточнения существующих липидных силовых полей, для изучения влияния молекул и белков на конформации и динамику липидов, и в конечном счете для разработки лекарств.

Опубликованные работы по теме диссертации

Материалы диссертации опубликованы автором достаточно полно в следующих работах:

1. Principal component analysis of lipid molecule conformational changes in molecular dynamics simulations / P. Buslaev, V. Gordeliy, S. Grudinin, I. Gushchin // *J. Chem. Theory Comput.* – 2016. – Vol. 12, no. 3 – P. 1019-1028
2. Effects of Coarse Graining and Saturation of Hydrocarbon Chains on Structure and Dynamics of Simulated Lipid Molecules / P. Buslaev, V. Gushchin // *Scientific reports* – 2017. – Vol. 7, no. 1 – P. 11476
3. Structure of the light-driven sodium pump KR2 and its implications for optogenetics / I. Gushchin, V. Shevchenko, V. Polovinkin, V. Borshchevskiy, P. Buslaev, E. Bamberg, V. Gordeliy // *The FEBS Journal* – 2016. – Vol. 283 – P. 1232-1238

Буслаев Павел Ильич

**Применение метода главных компонент для анализа конформаций
молекул липидов, полученных методом молекулярной динамики**

**АВТОРЕФЕРАТ ДИССЕРТАЦИИ НА СОИСКАНИЕ УЧЁНОЙ СТЕПЕНИ
КАНДИДАТА ФИЗИКО-МАТЕМАТИЧЕСКИХ НАУК**

Подписано в печать . Формат 60×84 1/16. Усл. печ. л. 1,2. Тираж 100 экз. Заказ № 156.

Федеральное государственное автономное образовательное учреждение высшего образования «Московский физико-технический институт (государственный университет)». Отдел оперативной полиграфии «Физтех-Полиграф»
141701, Московская обл., г. Долгопрудный, Институтский пер. д. 9.