

Министерство образования и науки РФ  
федеральное государственное автономное образовательное учреждение высшего  
образования «Московский физико-технический институт (национальный  
исследовательский университет)»  
Факультет общей и прикладной физики  
Образовательная программа "Вычислительная физика конденсированного состояния  
и живых систем"

# **Влияние конформации остатка пролина на структурно-динамические характеристики трансмембранного домена рецептора инсулина**

Диплом на соискание степени бакалавра

**Выполнила:**

студентка 626 группы  
Щекотихина Дарья Денисовна

**Научный руководитель:**

доктор физико-математических наук  
Ефремов Роман Гербертович

Москва 2020

## Аннотация

Рецептор инсулина (IR) - один из важнейших рецепторов в организме человека: он отвечает за регуляцию гомеостаза глюкозы. Нарушения в его работе приводят к таким болезням, как диабет. Несмотря на важность, молекулярная структура рецептора инсулина до конца не расшифрована, т.к. представлено сравнительно мало данных о трансмембранным домене. По этой причине трансмембранный домен IR представляет особый интерес для исследования компьютерными методами. В данной работе проводилась оценка роли конформации пептидной связи Gly-Pro в трансмембранным домене рецептора инсулина в структуре этого домена с помощью метода молекулярной динамики. Были получены структуры с цис-конформацией пептидной связи Gly-Pro и предложен механизм активации, предполагающий цис-транс изомеризацию пептидов на нескольких этапах.

# Содержание

<b>1 Введение</b>	<b>4</b>
<b>2 Обзор литературы</b>	<b>5</b>
2.1 Рецептор инсулина . . . . .	5
2.1.1 Модели активации IR . . . . .	6
2.2 Трансмембранный домен IR . . . . .	7
2.2.1 Исследование трансмембранного домена . . . . .	7
2.2.2 Экспериментальные данные . . . . .	7
2.2.3 Данные моделирования . . . . .	8
2.3 Пролин в трансмембранных $\alpha$ -спиралях . . . . .	9
2.3.1 Влияние пролина на свойства спирали . . . . .	9
2.4 Компьютерное моделирование биологических молекул . . . . .	9
2.4.1 Молекулярная динамика . . . . .	10
<b>3 Материалы и методы</b>	<b>11</b>
3.1 Структуры ТМ пептидов . . . . .	11
3.2 Получение траектории МД . . . . .	12
3.3 Разработанные алгоритмы обработки траектории МД . . . . .	14
3.3.1 Получение информации о геометрии трансмембранной $\alpha$ -спирале . . . . .	14
3.3.2 Поиск пересечений двух $\alpha$ -спиралей . . . . .	14
<b>4 Результаты и обсуждение</b>	<b>16</b>
4.1 Мономеры с цис- связью . . . . .	16
4.1.1 Исследование геометрии пептидов . . . . .	17
4.2 Предсказание структуры димеров с цис- связью . . . . .	19
4.3 Обсуждение результатов . . . . .	23
<b>5 Выводы</b>	<b>25</b>
<b>Список литературы</b>	<b>26</b>

# 1 Введение

Компьютерное моделирование - это построение модели объектов реального мира с помощью компьютера. При наличии достаточно точной модели, можно делать предположения о свойствах объекта на основе этой модели. Это позволяет рассчитывать характеристики в том числе тех объектов, которые по какой-либо причине невозможно или сложно исследовать экспериментально.

Методы компьютерного моделирования находят применение как в процессе создания новых материалов, так и в процессе исследования биологических систем (так называемая "сухая биология"). Численное моделирование, работая в совокупности с экспериментальными методами, способно значительно ускорить и удешевить исследование. Структурно-динамическая информация, полученная численно, ложится в основу эксперимента, снижая затраты на поиск необходимых структур.

Рецептор инсулина (IR) - один из ключевых рецепторов в организме человека. Он отвечает за такие важные процессы, как метаболическая регуляция гомеостаза глюкозы и рост и развитие плода человека. Нарушения в работе IR приводят к трудноизлечимым болезням, таким как диабет. Поэтому исследование процесса активации IR - это актуальная задача современной биологии.

Исследование структур, расположенных в липидной мемbrane оказывается трудоемкой задачей для экспериментальных методов, таких как спектроскопия ядерного магнитного резонанса и рентгеноструктурный анализ. Тем не менее, трансмембранный (ТМ) домен IR играет важную роль в процессе активации IR. Отсутствие достаточного количества данных о ТМ домене IR затрудняет понимание процесса активации рецептора. Поэтому перспективным методом исследования ТМ домена IR является компьютерное моделирование. В настоящей работе проводятся исследования структуры трансмембранного домена IR с помощью методов компьютерного моделирования.

Существуют экспериментальные данные, указывающие на то, что ТМ домен IR может существовать в различных конформациях, отличающихся геометрией пептидов с N конца. Есть основания полагать, что эти геометрические изменения связаны с изменением конформации пептидной связи Gly-Pro. Целью этой работы является исследование структурно-динамических свойств ТМ домена IR в различных конформациях.

## 2 Обзор литературы

### 2.1 Рецептор инсулина

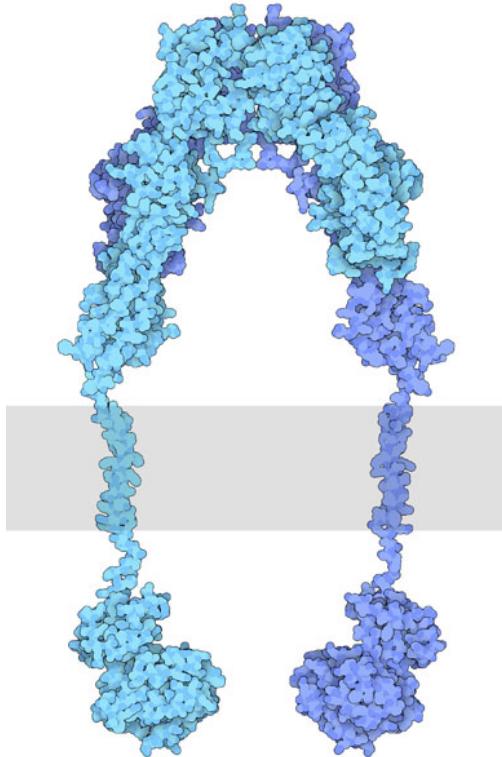


Рис. 1: Схематичное изображение пространственной структуры IR. Внеклеточный домен изображен сверху; внутриклеточный (киназный домен) изображен снизу; мембрана схематично показана серым цветом. <https://pdb101.rcsb.org/motm/182>

Инсулиновый рецептор (IR) - это рецептор, активируемый инсулином и инсулино-подобными факторами роста I и II (IGF-I и IGF-II). Он принадлежит обширному классу рецепторных тирозинкиназ и представляет собой интегральный мембранный белок. Этот белок является димером, состоящим из двух одинаковых молекул, пронизывающих мембрану.

В рецепторе инсулина выделяют внеклеточный, трансмембранный и внутриклеточный (киназный) домены (см рис. 1). Внеклеточный домен связывается с лигандом, трансмембранный домен передает сигнал от внеклеточного домена к киназному, киназный домен осуществляет реакцию фосфорилирования остатков тирозина в сигнальном белке, что приводит к каскаду реакций внутри клетки.

Одна из ролей IR в организме заключается в регуляции гомеостаза глюкозы. Этот процесс можно кратко описать следующим образом: пища поступает в организм, вследствие расщепления углеводов происходит повышение уровня глюкозы в крови; в ответ на повышение уровня глюкозы вырабатывается гормон инсулин; инсулин связывается

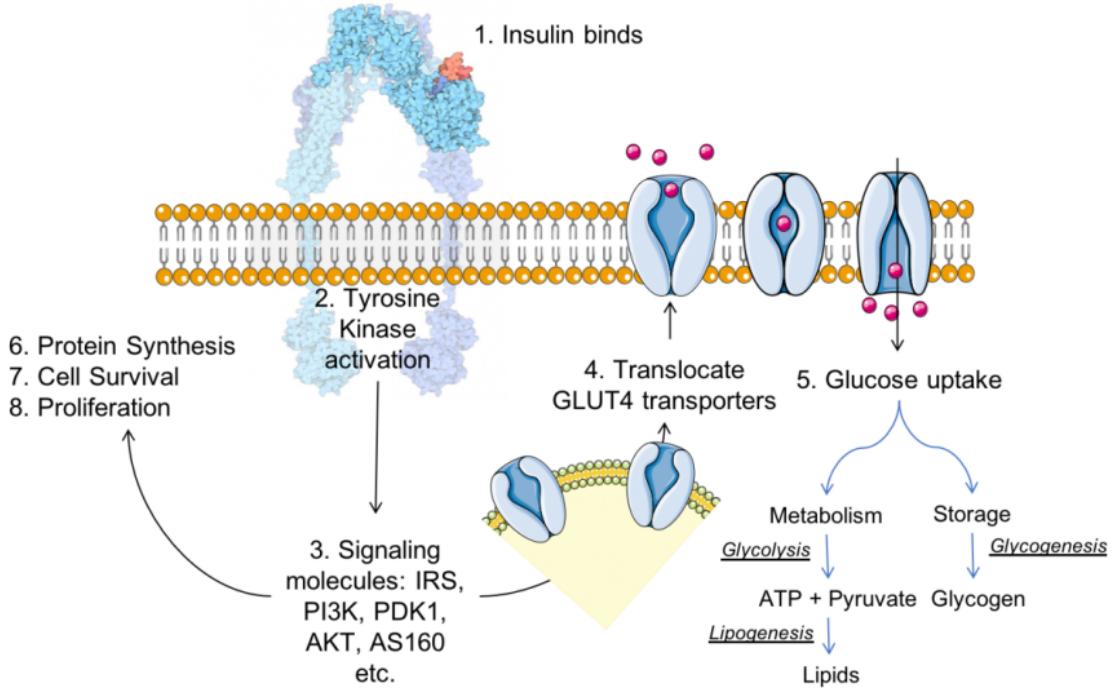


Рис. 2: Схема каскада реакций, запускаемых рецептором инсулина (синий) в результате связывания с инсулином (красный). <https://pdb101.rcsb.org/global-health/diabetes-mellitus/drugs/insulin/insulin-receptor>

с внеклеточным доменом рецептора; происходит передача сигнала внутрь клетки - к внутриклеточному домену; киназный домен запускает каскад реакций, приводящий к увеличению завата глюкозы клеткой. Вследствие этого уровень глюкозы в крови понижается.

Это справедливо для одной из изоформ IR. В организме человека рецептор инсулина существует в двух изоформах: IR-A и IR-B. Они различаются по аффинности к IGF-I и IGF-II. Так, установлено, что IR-A отвечает за внутриутробный рост и развитие, а IR-B отвечает за метаболическую регуляцию [2]. В контексте данной работы нас интересует активация IR инсулином, поэтому мы опускаем рассмотрение IR-A.

Поскольку рецептор отвечает за столь важные процессы в организме человека, нарушения в его работе приводят к различным клиническим проявлениям, в том числе к диабету. Поэтому то, как происходит процесс активации IR является актуальным вопросом современной биохимии. Знание этого процесса на молекулярном уровне может послужить базой для создания лекарств на основе структуры белка.

### 2.1.1 Модели активации IR

На данный момент общепринятой моделью активации IR, подтверждающейся экспериментальными данными, является предложенная Lee [5] (см рис. 3С). В данной модели после активации инсулином две части внеклеточного домена "сходятся тем самым

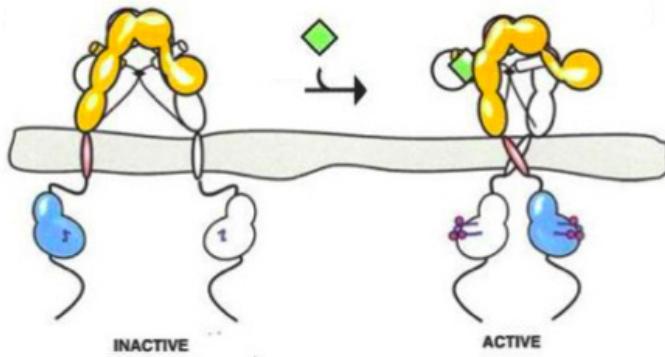


Рис. 3: Модель активации инсулина, основанная на димеризации ТМ домена. [4]

сближая  $\alpha$ -спирали ТМ домена. ТМ домен димеризуется и далее, из-за такого изменения конформации ТМ участка, киназный домен переходит в активированное состояние.

## 2.2 Трансмембранный домен IR

### 2.2.1 Исследование трансмембранного домена

В то время как для внеклеточного и киназного доменов возможно подробное исследование структуры экспериментальными методами, в том числе с помощью криоэлектронной микроскопии [3], трансмембранный домен представляет особую сложность, т.к. он может существовать только в липидном окружении. Таким образом, в изучении ТМ домена особый интерес представляют методы компьютерного моделирования. Они позволяют на базе неполных структурных данных проверять теоретические предположения на основе оценки стабильности структур, а так же получать представление о динамике полученных структур в модельном окружении (см. раздел 2.4).

### 2.2.2 Экспериментальные данные

С помощью метода двумерной ЯМР спектроскопии была расшифрована структура ТМ домена IR в мицеллах [6]. Структура для последовательности остатков 940–988 была получена в мицеллах ДФХ (додецилфосфохолина). Несмотря на то, что в работе удалось получить только мономеры белка, авторы указывают на возможность существования структур в форме димеров в мицеллах. У мономеров внемембранные участки являются подвижными и располагаются в водном окружении. Мембранный участок имеет структуру  $\alpha$ -спирали. При этом был отмечен кинк в этой спирали, созданный связью Gly960-Pro961, потенциально важный для функции домена.

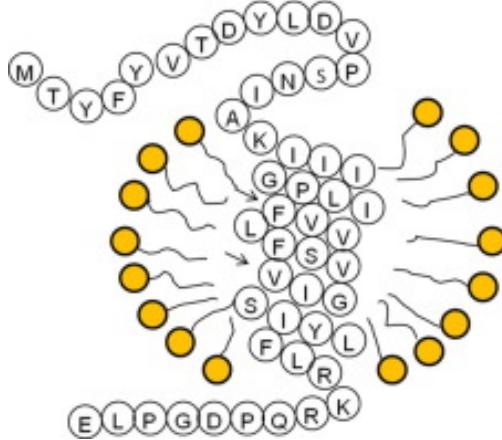


Рис. 4: Схематическое изображение ТМ домена IR в мицелле детергента (желтым указаны головки липидов, буквы обозначают аминокислотный остаток.) [6].

В лаборатории биомолекулярной ЯМР-спектроскопии ИБХ РАН был проведен аналогичный эксперимент с целью получения димерных форм в мицеллах ДФХ. Был выделен участок 944-973 и созданы образцы мономерных и димерных форм в мицеллах. Спектры показали необычные результаты в отношении кинка Gly-Pro. Спектр мономера двоится в N-концевых остатках, начиная с Pro961. Это говорит об изменении локального окружения этих остатков, что позволяет судить о наличии двух форм мономера, различающихся по форме с N-конца.

В спектре димеров начиная с пролина сигнал троится. То есть, аналогичным образом, существует три конформации димера ТМ домена IR.

TO-DO добавить график вторичной структуры и спектр (если это возможно)

Итак, целью данной работы является предсказание структур, которые были бы согласованы с результатом вышеописанного эксперимента, путем компьютерного моделирования.

### 2.2.3 Данные моделирования

Ранее в лаборатории моделирования биомолекулярных систем ИБХ РАН были предсказаны наиболее вероятные структуры димеров для ТМ доменов IR [7]. Для получения димеров на основе последовательности мономера был использован программный пакет PREDDIMER [8].

Структуры димеров, полученные с помощью PREDDIMER, затем были помещены в мембранные окружение из молекул пальмитоилолеилфосфатидилхолина (POPC), проведена минимизация энергии, релаксация и расчет траектории молекулярной динамики (3.2). Затем были предложены 3 наиболее стабильные димерные структуры. Все они имеют характерную водородную связь Ser969-Ser969.

TO-DO добавить картинку этих структур из диплома Мифтаха (если это возможно)

## 2.3 Пролин в трансмембранных $\alpha$ -спиралях

### 2.3.1 Влияние пролина на свойства спиралей

Принято считать, что остаток пролина в трансмембранных  $\alpha$ -спиралях играет функциональную роль. На это указывает сравнительно высокая встречаемость пролина в трансмембранных спиралах по сравнению со спиралями водорастворимых белков [9], [10]. Кроме того, было показано, что внесение мутаций в Pro в ТМ домене транспортных белков приводит к их деактивации [11]. Все это дает основание полагать, что пролин может быть функционально важен трансмембранных доменах большого количества белков.

Как компьютерное моделирование, так и экспериментальные данные показывают, что пролин вносит конформационные изменения в геометрию ТМ  $\alpha$ -спиралей [13], [12]. Цис-конформация пептидной связи между пролином и предшествующим ему остатком создает излом в  $\alpha$ -спирале, что может изменять ее положение в мембране и, как следствие, меняет функциональные свойства пептида.

Также было замечено, что пролин в трансмембранных спиралах имеет достаточно консервативное положение: он чаще всего оказывается "вытащен" в область полярных головок молекул липидов, составляющих мембрану [14]. Такое специфическое положение в свою очередь имеет связь с процессом димеризации ТМ спиралей. Так, с помощью поиска мотивов, образующих водородные связи в димерах, было исследовано, что пролин, находящийся в гидрофильной области ТМД (трансмембранныго домена) индуцирует димеризацию мономеров в мембране [15]. Этот факт вызывает особый интерес в контексте изучения механизма работы ТМД в IR.

Еще одним фактом в пользу того, что пролин функционально важен в механизме работы рецептора инсулина является то, что пролин консервативен во всем семействе IR. В него входят receptor инсулина, receptor инсулиноподобного фактора роста 1 (IGF-1R) и кецептор, подобный receptorу инсулина, (IRR). В ТМ доменах всех трех белков присутствует пролин в гидрофобной части.

## 2.4 Компьютерное моделирование биологических молекул

Атомистическое моделирование (AM) - это численный метод исследования вещества. AM рассматривает вещество как совокупность атомов, взаимодействие между которыми определяет физические свойства вещества.

## 2.4.1 Молекулярная динамика

Молекулярная динамика (МД) - один из методов атомистического моделирования. Он заключается в решении уравнений движения системы  $N$  взаимодействующих атомов:

$$m_i \frac{\partial r_i}{\partial t^2} = F_i, \quad i = 1..N; \quad F_i = \frac{-\partial V}{\partial r_i} \quad (1)$$

Потенциал  $V(r_1, \dots, r_N)$  определяется параметрами силового поля<sup>1</sup>.

Полученная система уравнений решается через некоторый малый промежуток времени  $\Delta t$  на некотором отрезке  $[t_0, t_1]$  при заданных начальных условиях  $(r_1(t_0), \dots, r_N(t_0))$ . В результате, совокупность точек  $(r_1, \dots, r_N)$  в каждый момент времени  $t_0, t_0 + \Delta t, t_0 + 2\Delta t, \dots, t_1$  составляет траекторию MD. Эта траектория определяет эволюцию системы в течение промежутка времени  $[t_0, t_1]$ .

При расчете макросистем ( $10^5$  атомов) аналитическое решение уравнений (1) на большом промежутке времени не представляется возможным для современной техники. Поэтому в MD уравнения решаются численно с использованием выбранной разностной схемы.

Критерием расчета MD является стабильность системы предмет + окружение. Система должна прийти к уравновешенному состоянию, в котором флуктуации остаются постоянными на некотором достаточно большом временном отрезке  $[t_S, t_2] \gg \Delta t$ . В качестве оценочной функции используется среднеквадратичное отклонение положения атомов от изначального или равновесного положения, усредненное либо по атомам (СКО), либо по временному промежутку (СКФ). Усреднение по атомам позволяет оценить стабильность структуры в целом, а усреднение по времени позволяет отслеживать участки повышенной подвижности структуры, которые часто играют большое значение в функции молекулы.

Описанный выше функционал реализуется открытыми программными пакетами, такими как NAMD и Gromacs.

---

<sup>1</sup>Какое-либо взаимодействие атомов можно описать в рамках соответствующей модели (к примеру, валентную связь между атомами можно описать как идеально жесткую связь или в модели гармонического осциллятора). Модель дает аналитическое выражение для потенциала взаимодействия. Силовое поле - это совокупность потенциалов разных взаимодействий с коэффициентами, подобранными эмпирически, чтобы модель соответствовала предмету в реальной жизни.

### 3 Материалы и методы

#### 3.1 Структуры ТМ пептидов

Изначальная структура мономера ТМ домена IR с транс-конформацией связи Gly957-Pro958 взята из [6]. Пептид модифицируется по длине свободных остатков и по конформации связи Gly957-Pro958.

Таблица 1: Последовательности пептидов

Название пептида	Аминокислотная последовательность
длинный	VTDYLDVPSNIAKIIGPLIFVFLFSVVIGSIYLFLRKQRQPDG
короткий	SNIAKIIGPLIFVFLFSVVIGSIYLFLRKKR

Для получения структуры мономера с цис-конформацией пептидной связи Gly-Pro эта пептидная связь была повернута на 180°с помощью программы Pymol. Т.к. пептидная связь является идеально жесткой в поле GROMOS 43a2x, после искусственного изменения угла пептидной связи этот угол сохранится в процессе рассчета траектории МД.

Таблица 2: Описание структур пептидов

Обозначение пептида	Описание пептида
ТРАНС-Д	Длинный пептид с транс-конформацией связи Gly957-Pro958
ТРАНС-К	Короткий пептид с транс-конформацией связи Gly957-Pro958
ЦИС-Д	Длинный пептид с цис-конформацией связи Gly957-Pro958
ЦИС-К	Короткий пептид с цис-конформацией связи Gly957-Pro958

В качестве основы для исследования влияния цис-конформации связи Gly957-Pro958 на структурно-динамические параметры димеров ТМ домена IR были использованы структуры димеров ТМ домена IR, полученные в [7] и отобранные, как наиболее стабильные. В таблице 3 представлены интерфейс взаимодействия  $\alpha$ -спиралей в димере и их угол скрещивания.

Для получения структур димеров с цис-конформацией связи Gly-Pro с помощью программного пакета Gromacs было проведено наложение структуры длинного цис-мономера на соответствующую часть димера по совпадающей части (Ple960-Lys978). Далее были выявлены пересечения двух спиралей димеров между собой методом, описанным в разделе 3.3.2. В результате были найдены номера остатков, в которых происходит пересечение  $\alpha$ -спиралей пептидов.

Таблица 3: Димеры ТМ домена IR

Номер структуры	Интерфейс взаимодействия	Угол скрещивания, $^{\circ}$
1	-----VF-SV-IG-YL-----	-55
2	-S-IA-II-L-F-S-G-YL-	-5,3
6	-S-K-I-L-F-----	40

### 3.2 Получение траектории МД

Основные стадии вычислительного эксперимента для мембранных белка следующие:

- Соотнесение параметров силового поля для белков и липидов.
- Сборка модельной системы: погружение белка в мембранное окружение.
- Минимизация энергии составленной системы.
- Релаксация системы: приведение системы к нужным значениям температуры и давления.
- Рассчет траектории MD.
- Оценка стабильности.
- Обработка полученных результатов.

Изначально, белок и мембрана<sup>2</sup> представлены в виде отдельных файлов, содержащих информацию о координатах каждого атома и принадлежности к определенной группе. Для объединения белка и липидов в единую систему используется подход InflateGRO [17].

На следующем шаге происходит минимизация энергии алгоритмом градиентного спуска. После этого с помощью алгоритмов термостата V-rescale и баростата Berendsen в коротком прогоне молекулярной динамики происходит релаксация системы к определенной температуре и давлению. Отрелаксированная система готова к расчету траектории MD.

После получения итоговой траектории вычисляются среднеквадратичное отклонение атомов главной цепи от изначального положения (СКО) и среднеквадратичная

<sup>2</sup>В связи с достаточно однородной структурой строения мембраны как липидного бислоя, расчет мембранных белков всегда проводится в периодических граничных условиях. Таким образом, файл хранит небольшой участок бислоя, который самостабилизирован так, как если бы бислон был бесконечен. Это является преимуществом исследования мембранных белков с помощью моделирования, в отличие от экспериментальных методов, в которых используются мицеллы, кубическая липидная фаза и т.д.

флуктуация атмов главной цепи (СКФ), по которым можно судить о стабильности белка в мембране.

В настоящей работе, в качестве мембраны используется заранее уравновешенный бислой из молекул пальмитоилолеилфосфатидилхолина (ПОФХ). В качестве программного пакета, реализующего алгоритмы МД, а также для оценки стабильности полученной траектории, используется пакет Gromacs 2019. Детали вычислительного эксперимента приведены в таблице 4.

Таблица 4: Параметры расчета траектории МД

Программный пакет	Gromacs 2019
Число молекул в системе	1-2 молекулы белка, ≈90 липидов, ≈4500 молекул воды
Молекулы мембранны	ПОФХ
Силовое поле	GROMOS 43a2x
Радиус для построения списка ближайших соседей	12A
Учет электростатических взаимодействий	Суммирование по Эвальду
Радиус отсечки электростатических взаимодействий	10A
Учет взаимодействий Ван-дер-Ваальса	Потенциал Леннарда-Джонса с отсечкой
Радиус отсечки взаимодействий Ван-дер-Ваальса	12A
Термостат	V-rescale
Баростат	Berendsen
<b>Минимизация энергии</b>	
<b>методом градиентного спуска</b>	
Максимальное число шагов	100000
Максимальный размер шага	0.01 нм
<b>Релаксация липидного окружения</b>	
Шаг интегрирования по времени	0.002 пс
Число шагов	2 500 000 (всего 5 нс)
Нагрев системы	от 5 К до 315 К за первые 500 пс
<b>Молекулярная динамика</b>	
Шаг интегрирования по времени	0.002 пс
Число шагов	100 000 000 (всего 200 нс)
Температура	315 K

### 3.3 Разработанные алгоритмы обработки траектории МД

#### 3.3.1 Получение информации о геометрии трансмембранный $\alpha$ -спиралей

Необходимым аппаратом в исследовании трансмембранных  $\alpha$ -спиралей является обработка информации о геометрических характеристиках спирали в мемbrane: наклон спирали относительно перпендикуляру к мемbrane, изломы спирали, ориентация и т.д. Для решения подобных задач был разработан специальный алгоритм.

Идеальный виток  $\alpha$ -спиралей, ориентированный по оси Z (ось Z перпендикулярна плоскости мембраны), методом наименьших квадратов фитится на соответствующий участок трансмембранный спиралей. Результатом фита является матрица поворота витка в пространстве. Данная матрица содержит всю информацию о геометрии данного участка спиралей. Получая такие матрицы для витков по всей длине спиралей, можно установить геометрические характеристики всего пептида.

В данной работе этот алгоритм применяется для определения угла наклона спиралей относительно перпендикуляра к плоскости мембраны и для отслеживания изломов спиралей. Для определения угла наклона, единичный вектор, направленный вдоль оси Z, умножается на матрицу поворота, тогда проекция полученного вектора на ось Z определяет косинус угла наклона спиралей (см. рис. 5a). Для отслеживания излома спиралей, каждый полученный вектор вычитается из предыдущего. В итоге, модуль этой разности показывает наличие излома на данном участке спиралей (см. рис. 5b).

Естественным образом, из вышеописанного модуля разности можно также рассчитать угол излома  $\varphi$ :

$$\varphi = \arccos\left(\frac{|\overline{\Delta R}|}{2} - 1\right).$$

Для реализации всех алгоритмов была использована вспомогательная библиотека MDAnalysis (<https://www.mdanalysis.org>).

#### 3.3.2 Поиск пересечений двух $\alpha$ -спиралей

С помощью последовательного наложения витка  $\alpha$ -спиралей на всю длину пептида также можно определить возможные пересечения мономеров в димерах, полученных методом наложения. Для этого для всех возможных пар  $\overline{R}_1, \overline{R}_2$  высчитывается разность. Если модуль этой разницы меньше некоторого  $\varepsilon$ , то на этом участке регистрируется пересечение (см. рис. 6).

Величина  $\varepsilon$  взята равной среднему диаметру  $\alpha$ -спиралей: 1.5 нм.

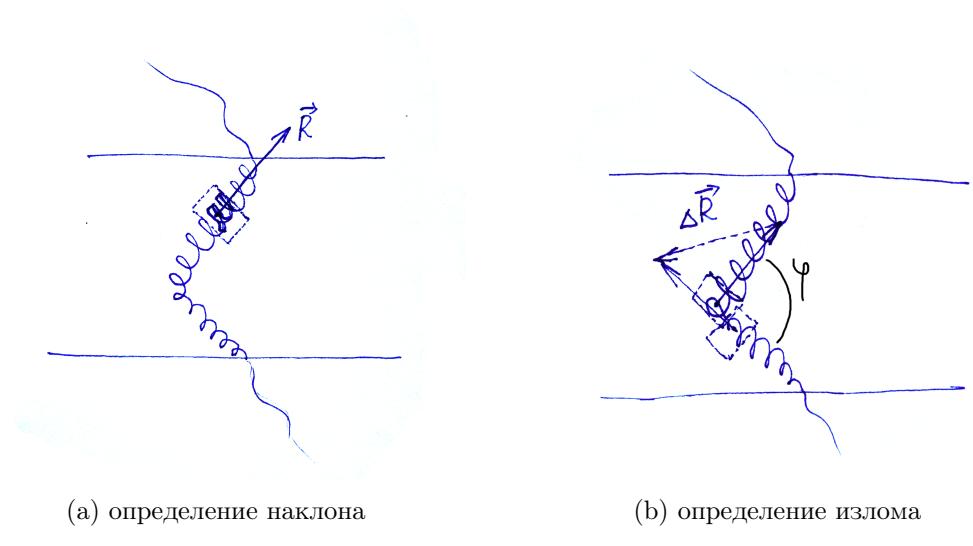


Рис. 5: Схема обработки геометрических данных трансмембранный  $\alpha$ -спиралей

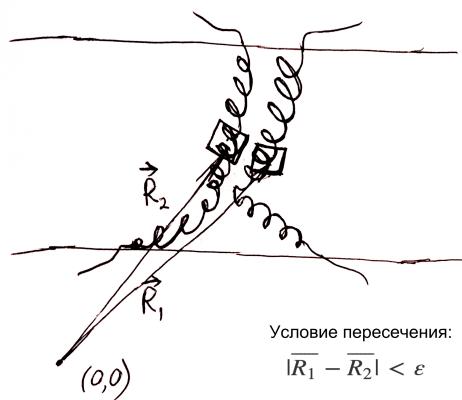


Рис. 6: Схема определения пересечений между мономерами в модельном димере.

## 4 Результаты и обсуждение

### 4.1 Мономеры с цис- связью

В работе были получены структуры мономеров с цис и транс-конформацией пептидной связи Gly957-Pro958. В процессе расчетов было замечено, что длина внешних свободных участков сильно влияет на структуру пептида. Поэтому расчет проводился как для коротких пептидов с обрезанными свободными участками цепи с N и C концов, так и для длинных (см. таблицу 2). Полученные структуры показаны на рис. 7.

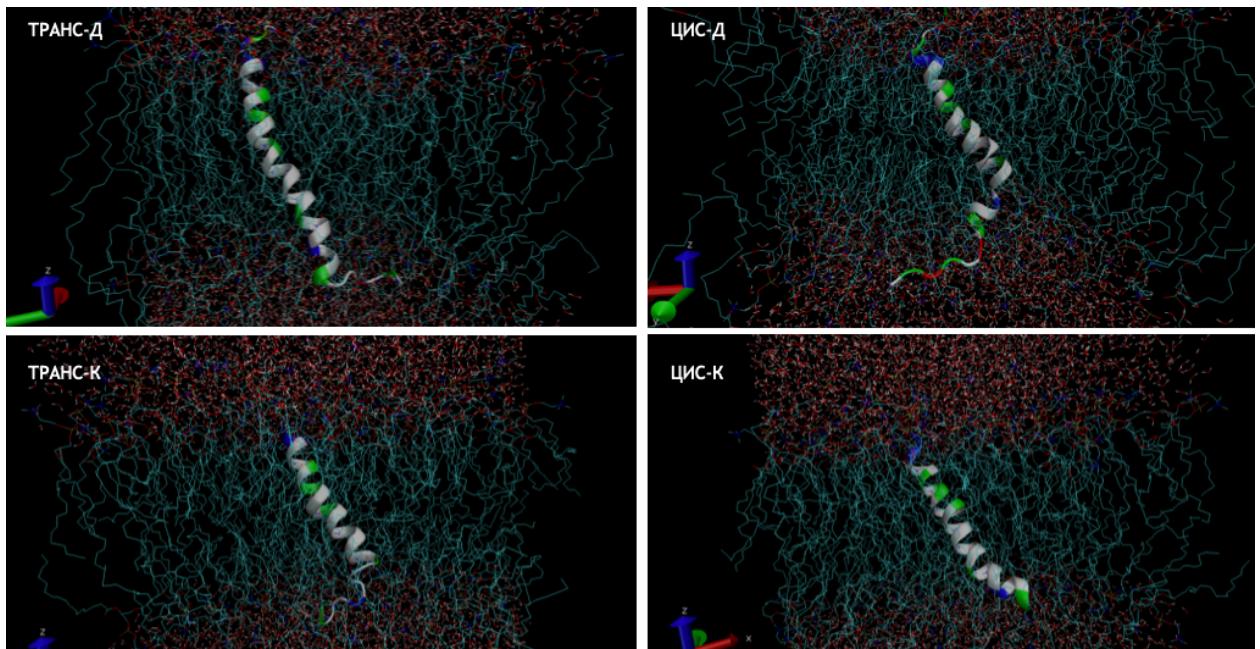


Рис. 7: Стабильные структуры мономеров: ТРАНС-Д, ЦИС-Д, ТРАНС-К, ЦИС-К. **ТО-DO сделать более красивую картинку, отметить Pro.**

На рис. 8 представлены графики СКО и СКФ для всех четырех пептидов. По графикам СКО можно видеть, что исследованные пептиды стабильны в мембранном окружении, т.к. устанавливается стабильное значение СКО. По графикам СКФ можно сравнить подвижность соответствующих остатков основной цепи в случае цис и транс конформации. Интересно заметить, что в случае короткого транс пептида наблюдается высокая подвижность в области остатков Gly957-Pro958. Это значит, что, эта область является нестабильной. В длинном транс пептиде, из-за дополнительного уравновешивания длинным внешним участком, такого эффекта не наблюдается.

Также стоит отметить, что у обоих пептидов с цис-связью пролин располагается в области полярных головок фосфолипидов мембраны (см. рис. 7).

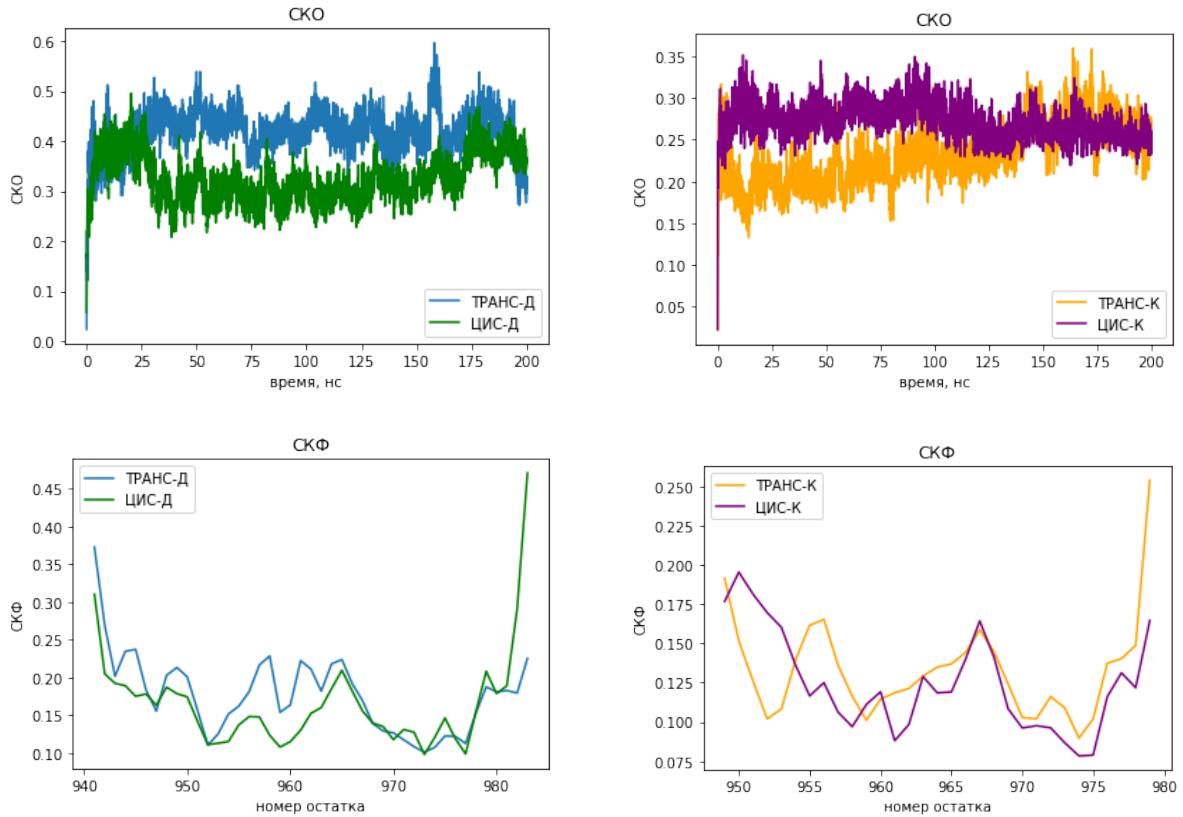


Рис. 8: Графики СКО и СКФ четырех пептидов: синий - ТРАНС-Д, зеленый - ЦИС-Д, желтый - ТРАНС-К, фиолетовый - ЦИС-К.

#### 4.1.1 Исследование геометрии пептидов

С помощью алгоритма, описанного в разделе 3.3.1, была проведена обработка результатов траектории МД с целью выявления закономерностей в геометрии разных конформаций пептида (см. рис. 9, 10).

Можно видеть, что пептид ТРАНС-Д является прямым и стоит вертикально в мембране. Пептиды ЦИС-Д и ТРАНС-К имеют наклон относительно оси Z (ЦИС-Д  $\approx 45^\circ$ , ТРАНС-К  $\approx 65^\circ$ .) Также, по красному графику СКО наложения витка можно видеть, что в области Pro958 пептида ЦИС-Д предположительно есть перелом - виток не наложился на искривленный участок. У пептида ЦИС-К можно видеть развернутую область с N конца после остатка пролина.

По графикам угла излома можно видеть, что у пептида ЦИС-Д самый большой излом присутствует в области остатков Gly957-Pro958. Аналогичный излом можно видеть и у ЦИС-К на неразвернутом участке.

Далее была дополнительна получена зависимость угла излома  $\alpha$ -спирали пептидов в области связи Gly957-Pro958 от времени симуляции (рис. 11). Можно видеть, что у транс-пептидов нет излома (угол  $\approx 180^\circ$ ), а у цис-пептидов есть излом (угол  $130 - 140^\circ$ ),

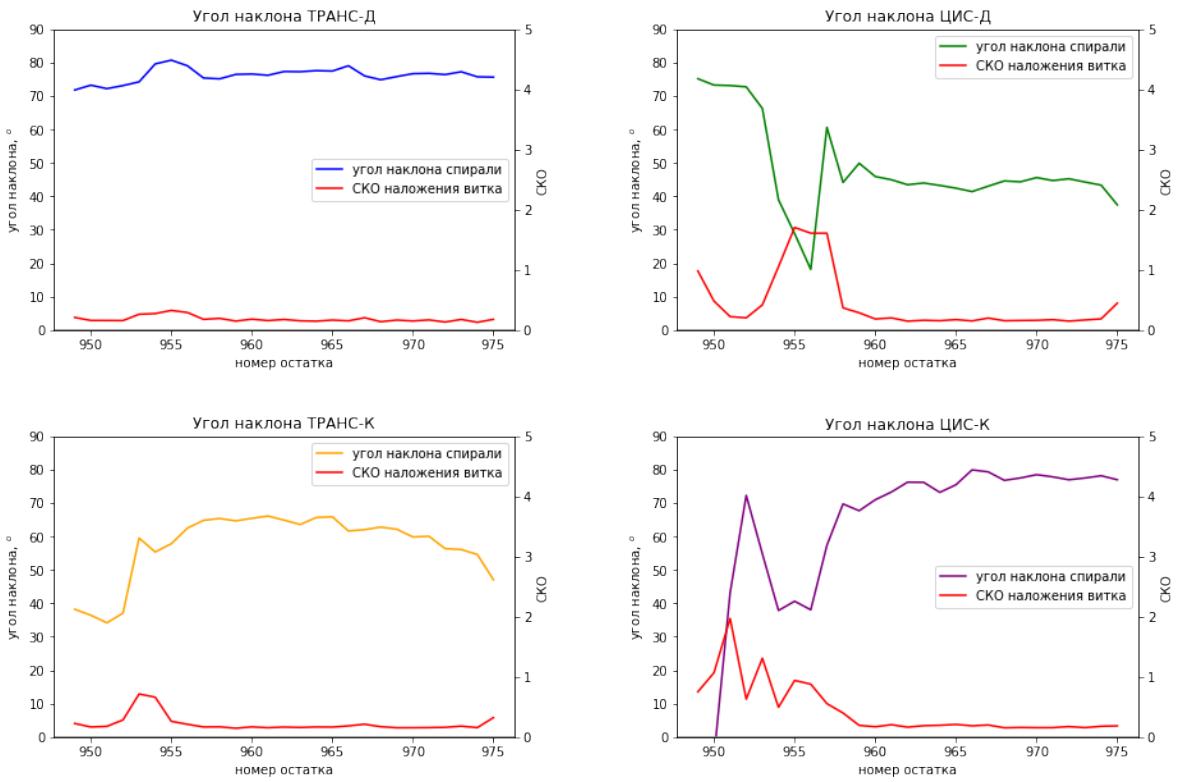


Рис. 9: Графики наклона  $\alpha$ -спирали в зависимости от остатка (180°- спираль расположена вертикально).

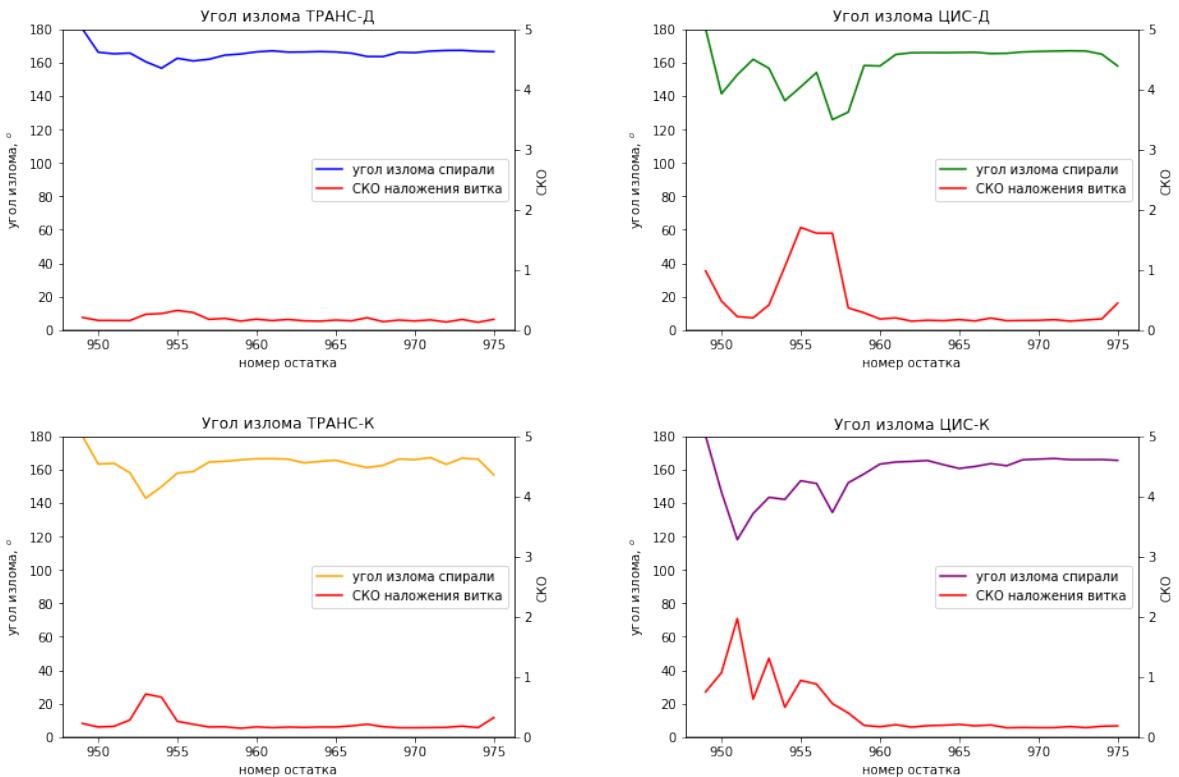


Рис. 10: Графики излома  $\alpha$ -спирали в зависимости от остатка (180°- прямая спираль.)

и он сохраняется неизменным на протяжении всего времени симуляции. Из этого можно сделать вывод, что структура цис-пептидов является жесткой в области Gly957-Pro958, что является важным фактом для дальнейшего поиска структур димеров с цис-связью.

Из полученных данных можно сделать вывод, что цис-конформация пептидной связи Gly957-Pro958 создает излом в трансмембранный  $\alpha$ -спирале пептида ТМ домена IR. При этом, наличие или отсутствие примембранных участков также влияет на структуру пептида, что согласуется с предположением о том, что конформационные изменения в ТМ домене IR индуцируются через внешние участки, соединяющие ТМ домен с внеклеточным и киназным доменами.

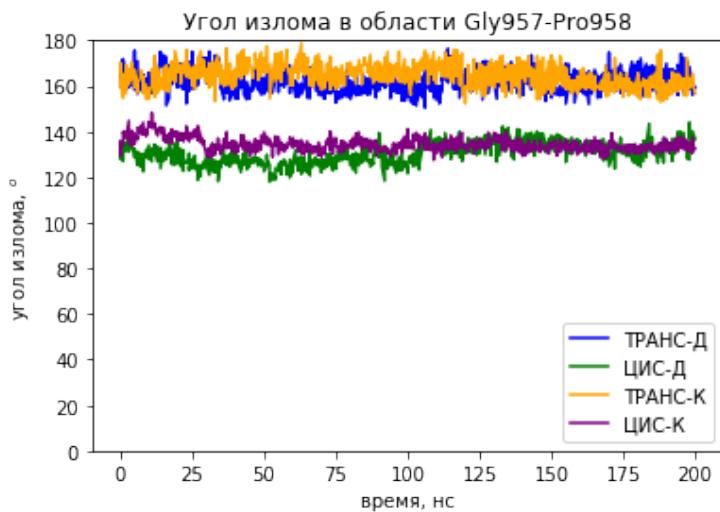


Рис. 11: График угла излома  $\alpha$ -спиралы в зависимости от времени в траектории МД.

## 4.2 Предсказание структуры димеров с цис- связью

В результате анализа структур мономеров с цис-связью было обнаружено, что структура в области Gly957-Pro958 является жесткой. Это накладывает геометрическое ограничение на возможные структуры димеров с цис-связью. Так, если соединять мономеры с неподходящим углом скрещивания  $\alpha$ -спиралей, то мономеры в димере могут пересекаться, как схематически отображено на рис. 12

С целью отбора возможных структур димеров с цис-транс, транс-цис и цис-цис пептидной связью Gly-Pro был использован алгоритм 3.3.2. За основу были взяты димеры, полученные в работе [7]. Результаты работы алгоритма представлены в таблице 5. В таблице показаны найденные пересечения по номерам остатков.

Очевидно, что структуры, содержащие пересечения, геометрически невозможны в силу жесткости мономера. Остальные структуры подходят для дальнейшего рассчета

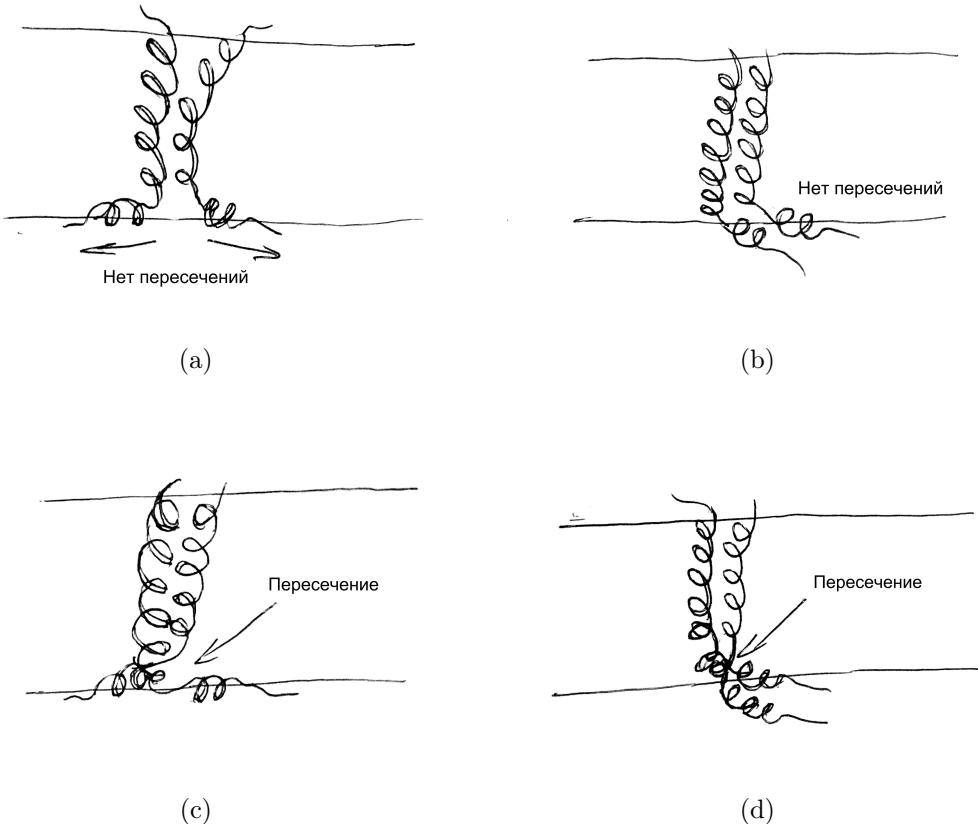


Рис. 12: Схема пересечений жестких мономеров в димерах. Димеры а, б - без пересечений, димеры с, д - есть пересечения. **ТО-ДО:** поменять на реальные картинки **после наложения - ?**

траектории МД.

В силу предположения о том, что димеризация происходит для цис-мономеров (см. 2.3.1), далее мы будем исследовать только цис-цис димеры. Как видно из таблицы 5, это цис-цис димер, полученный из структуры с номером 2.

Таким образом, была получена единственная структура цис-цис димера, возможная из геометрических соображений. Она является стабильной в мембране (см. рис. 13) и представляет собой симметричную структуру без пересечений (см. 12а).

Далее представлен график СКФ (рис. 14). Для сравнения рядом приведен аналогичный график для транс-транс димера, на базе которого был получен данный цис-цис димер. По графику СКФ можно сделать вывод, что ЦИС-ЦИС димер более стабилен, чем аналогичный ТРАНС-ТРАНС димер с таким же профилем димеризации.

Таблица 5: Пересечения мономеров в димерах с ЦИС конформацией

Номер димера	Пересечения (номера остатков)		
	ЦИС-ТРАНС	ЦИС-ЦИС	ТРАНС-ЦИС
1	-	954	-
2	-	-	-
6	959	952	-

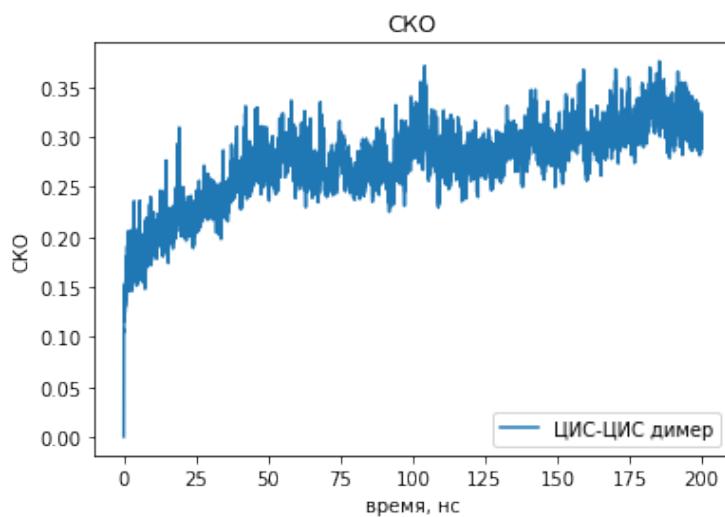


Рис. 13: График СКО для димера с цис конформацией обоих мономеров.

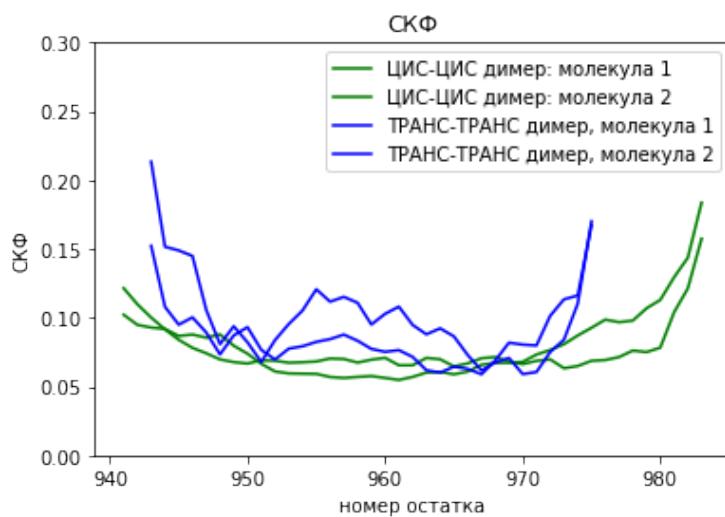
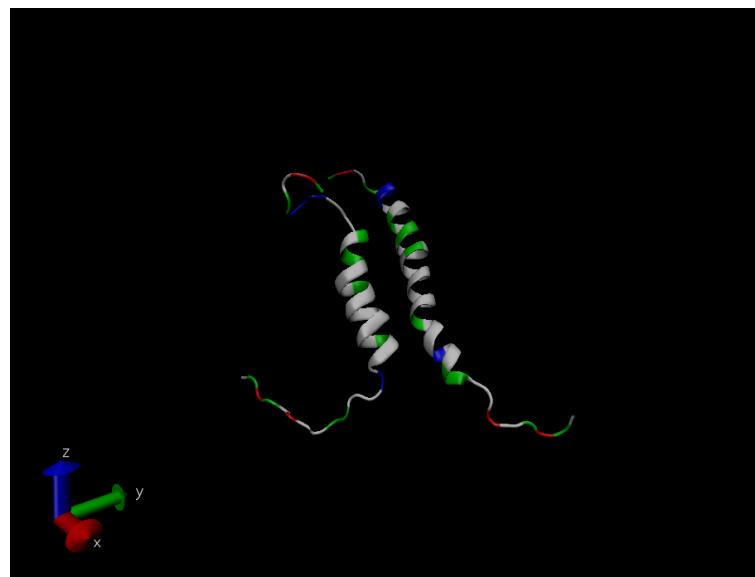
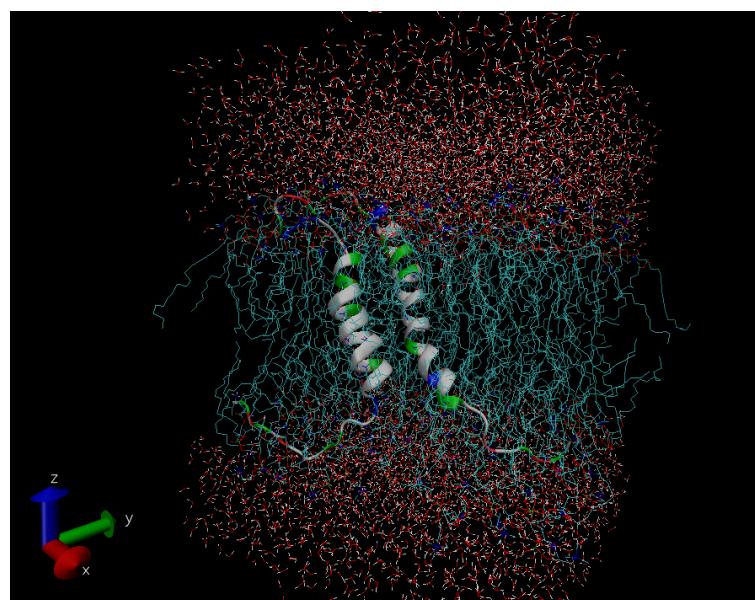


Рис. 14: Графики СКФ димеров с цис-цис и транс-транс конформациями мономеров (димер номер 2).



(a) мембрана скрыта



(b) мембрана показана

Рис. 15: Стабильный димер с цис-конформацией обоих мономеров.

### 4.3 Обсуждение результатов

При внесении цис-конформации пептидной связи Gly957-Pro958 наблюдаются следующие эффекты: пролин оказывается на краю мембраны и в области пептидной связи образуется резкий излом. Это находится в соответствии с предшествующими работами [13], [12].

Отдельно из анализа структурно-динамических свойств мономеров с цис и транс конформацией можно выделить замечание, что длина цепочки внешних остатков с N и C конца влияет на стабильность определенной конфигурации в области Gly957-Pro958. Из этого можно выдвинуть предположение, что внешние участки, которые соединяют ТМ домен с внеклеточным и внутриклеточным доменами, могут индуцировать цис-транс переход в мономерах ТМ домена.

Учитывая, что димеризации может способствовать наличие пролина в гидрофильной части мембраны [15], предложена гипотеза о том, что димеризация происходит, когда мономеры принимают цис-конформацию. Таким образом, в неактивной форме IR (см. рис. 3) за счет изменений во внеклеточном домене через примембранные участки индуцируется транс-цис переход в мономерах ТМ домена. После этого, вследствие повышенной способности к димеризации (см. [15]), мономеры димеризуются и переходят в состояние цис-цис димера (см. рис. 15). Димеризация оказывается предпочтительной в том числе потому, что из-за геометрического излома ТМ участки оказываются приближены друг к другу (см. рис. 16). Далее происходит переход из цис-цис димера в транс-транс димер - активное состояние IR (вероятно, переход происходит за счет приближения внешних участков с N конца).

Описанная выше модель активации основана на предположении о том, что в процессе активации ТМ домен работает симметрично: мономеры не образуют цис-транс и транс-цис димеры. Тем не менее, это предположение не имеет строгих обоснований, и поэтому нельзя отрицать возможности асимметрии в процессе активации IR. Поэтому модель с цис-транс изомеризацией может быть в дальнейшем дополнена с учетом несимметричных структур.

Также стоит отметить, что практически все полученные структуры для мономеров и димеров ТМ домена IR оказались в той или иной степени наклоненными в мемbrane. Это может говорить о том, что толщина мембраны, в которой производились расчеты (ПОФХ) является недостаточной для ТМ домена IR. IR и другие рецепторы семейства IR обычно расположены в липидных рафтах - утолщенных участках мембраны, укрепленных гликосфинголипидами и холестерином [18]. Из этого следует, что реальные свойства ТМ домена в жизни могут несколько отличаться от полученных и учет специфических свойств мембраны может уточнить результаты, полученные в настоя-

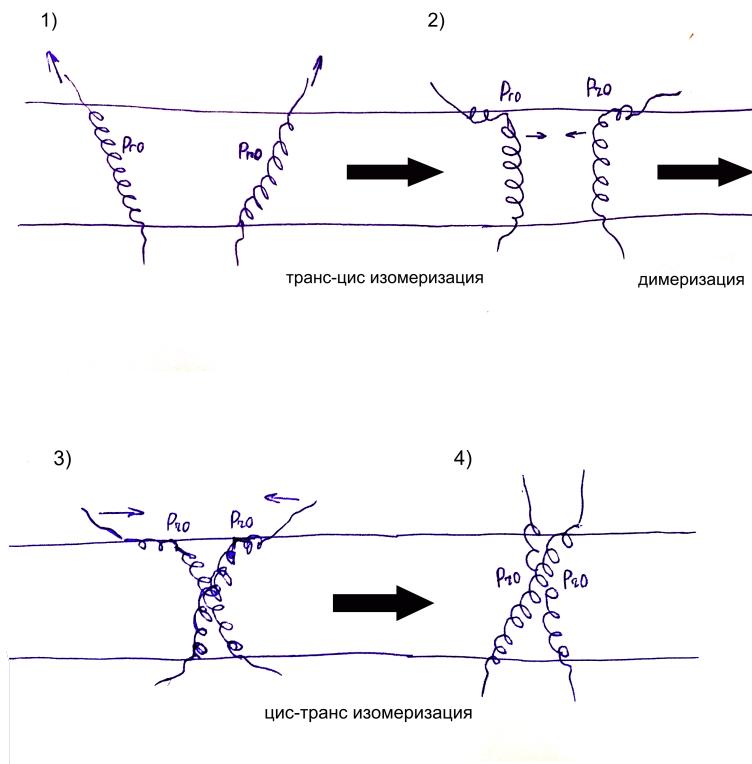


Рис. 16: Схема активации IR с цис-транс изомеризацией связи Gly-Pro в ТМ домене. 1. Внеклеточным доменом (не показан) индуцируется транс-цис-переход. 2. Цис-мономеры димеризуются в цис-цис димер. 3. Внеклеточным доменом индуцируется цис-транс переход - димер переходит в транс-транс конфигурацию. 4. ТМ домен в активной форме транс-транс димера.

щей работе.

## 5 Выводы

По результатам работы можно сделать следующие выводы:

- Мономеры ТМ домена IR с цис-конформацией пептидной связи Gly957-Pro958 стабильны в мембранном окружении и, следовательно, могут существовать в природе. Цис-конформация пептида меняет его геометрические свойства: в области пептидной связи образуется излом и пролин оказывается вытянут к областью полярных головок липидов.
- Подвижность мономеров в области связи Gly957-Pro958 с цис и транс конфигурациями пептидной связи зависит от длины внешних участков с N и C конца: для длинного пептида транс-конформация стабильна, для короткого пептида - нестабильна.
- Пептид с цис- конформацией пептидной связи Gly957-Pro958 является жестким в области этой связи.
- Существует стабильная модель возможного димерного состояния ТМ домена IR, допускающая цис-конфигурацию обоих мономеров. Этот димер имеет симметричную структуру с удаленными друг от друга N концами, лежащими на границе мембранны в области полярных головок фосфолипидов.

Существование димера с цис-цис конфигурацией позволяет предположить возможность механизма активации IR, предполагающего цис-транс измеризацию связи Gly-Pro ТМ домене. Полученные структуры мономеров и димера могут быть использованы для дальнейшего исследования ТМ домена IR экспериментальными методами.

## Список литературы

- [1] Staker B. L., Buchko G. W., Myler P. J. Recent contributions of structure-based drug design to the development of antibacterial compounds //Current opinion in microbiology. – 2015. – Т. 27. – С. 133-138.
- [2] Belfiore A, Malaguarnera R, Vella V, et al. : Insulin Receptor Isoforms in Physiology and Disease: An Updated View. Endocr Rev. 2017;38(5):379–431. 10.1210/er.2017-00073
- [3] Uchikawa E, Choi E, Shang G, Yu H, Bai XC. Activation mechanism of the insulin receptor revealed by cryo-EM structure of the fully liganded receptor-ligand complex Elife. 2019; 2019 Aug 22.
- [4] De Meyts P. The insulin receptor and its signal transduction network. In: De Groot LJ, Chrousos G, Dungan K, Feingold KR, Grossman A, Hershman JM, Koch C, Korbonits M, McLachlan R, New M, Purnell J, Rebar R, Singer F, Vinik A, editors. Endotext [Internet]. South Dartmouth, MA: MDText.com, Inc. (2000).
- [5] Lee J, Miyazaki M, Romeo GR and Shoelson SE. Insulin receptor activation with transmembrane domain ligands. J Biol Chem 289: 19769-19777 (2014)
- [6] Q. Li, et al. Solution structure of the transmembrane domain of the insulin receptor in detergent micelles Biochim. Biophys. Acta, 1838 (2014), pp. 1313-1321
- [7] Zamaletdinov M.F., Kuznetsov A.S., Maurice P ., Efremov R.G. Putative role of transmembrane domain dimerization in activation of proteins from the Insulin receptor family J Bioenerg Biomembr. 2018. V. 50, no. 6, p. 515.
- [8] Polyansky A. A. et al. PREDDIMER: a web server for prediction of transmembrane helical dimers //Bioinformatics. – 2014. – Т. 30. – №. 6. – С. 889-890.
- [9] Woolfson DN, Mortishire-Smith RJ, Williams DH. Conserved positioning of proline residues in membrane-spanning helices of ion-channel proteins. Biochem Biophys Res Commun 1991; 175: 733–737
- [10] Alex Peralvarez-Marin, Jose-Luis Bourdelande, Enric Querol, Esteve Padros (2006) The role of proline residues in the dynamics of transmembrane helices: the case of bacteriorhodopsin, Molecular Membrane Biology, 23:2, 127-135
- [11] Slepkov ER, Chow S, Lemieux MJ, Fliegel L. Proline residues in transmembrane segment IV are critical for activity, expression and targeting of the Na<sup>+</sup>?/H<sup>+</sup> exchanger isoform 1. Biochem J 2004; 379: 31–38
- [12] Cordes FS, Bright JN, Sansom MSP. Proline-induced distortions of transmembrane helices. J Mol Biol 2002; 323: 951–960

- [13] L.K. Iyer, S. Vishveshwara A model for transmembrane helix with a cis-proline in the middle FEBS Lett., 374 (1995), pp. 21-24
- [14] D.N. Woolfson, R.J. Mortishire-Smith, D.H. Williams Conserved positioning of proline residues in membrane-spanning helices of ion-channel proteins Biochem. Biophys. Res. Commun., 175 (1991), pp. 733-737
- [15] N. Sal-Man, D. Gerber, Y. Shai Proline localized to the interaction interface can mediate self-association of transmembrane domains Biochim. Biophys. Acta - Biomembr., 1838 (2014), pp. 2313-2318
- [16] Lummis, S., Beene, D., Lee, L. et al. Cis-trans isomerization at a proline opens the pore of a neurotransmitter-gated ion channel. Nature 438, 248–252 (2005)
- [17] C. Kandt, W.L. Ash, D.P. Tieleman Setting up and running molecular dynamics simulations of membrane proteins Methods, 41 (4) (2007), pp. 475-488
- [18] Delle Bovi R. J. et al. Sterol structure dependence of insulin receptor and insulin-like growth factor 1 receptor activation //Biochimica et Biophysica Acta (BBA)-Biomembranes. – 2019. – Т. 1861. – №. 4. – С. 819-826.