

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

Данная работа посвящена исследованию роли конформации пептидной связи пролина в трансмембранном домене рецептора инсулина в контексте механизма активации рецептора.

Содержание

1	Введение	4
2	Обзор литературы	5
2.1	Рецептор инсулина	5
2.1.1	Модели активации IR	6
2.2	Трансмембранный домен IR	7
2.2.1	Экспериментальные данные	7
2.2.2	Данные моделирования	9
2.3	Пролин в трансмембранных α -спиралях	9
2.3.1	Влияние пролина на свойства спирали	9
2.3.2	Цис-транс изомеризация	10
2.4	Компьютерное моделирование биологических молекул	10
2.4.1	Молекулярная динамика	11
2.4.2	Классический метод расчета MD для трансмембранного белка . .	12
2.4.3	PREDDIMER и метод потенциала средней силы	13
3	Материалы и методы	14
4	Результаты и обсуждение	15
5	Выводы	16
	Список литературы	17

1 Введение

TO-DO

- про компьютерные методы исследования - про сухую биологию
- про рецептор инсулина, какой он важный - про медицину, как важно этот рецептор изучать - почему нам нужны тут комп. методы
- про эту работу в контексте общего исследования, к чему это приведет

2 Обзор литературы

2.1 Рецептор инсулина

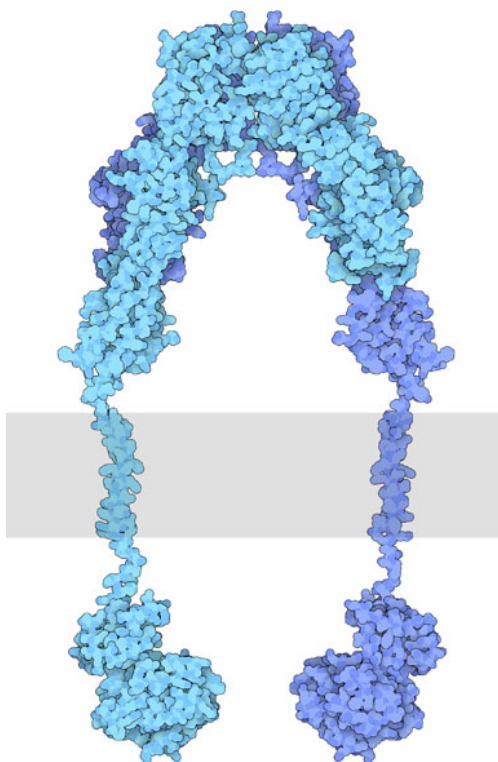


Рис. 1: ??? (Какое?) изображение IR. Внеклеточный домен изображен сверху; внутриклеточный (киназный домен) изображен снизу; мембрана схематично показана серым цветом. <https://pdb101.rcsb.org/motm/182>

Инсулиновый рецептор (IR) - это рецептор, связывающийся с молекулой инсулина. Он принадлежит обширному классу тирозинкиназных рецепторов и представляет собой трансмембранный белок, представленный в виде димера.

В рецепторе инсулина выделяют внеклеточный, трансмембранный и внутриклеточный (киназный) домены (см рис. 1). Внеклеточный домен связывается с инсулином, трансмембранный домен передает сигнал от внеклеточного домена к киназному, киназный домен запускает каскад реакций внутри клетки.

Роль 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 с внеклеточным доменом рецептора

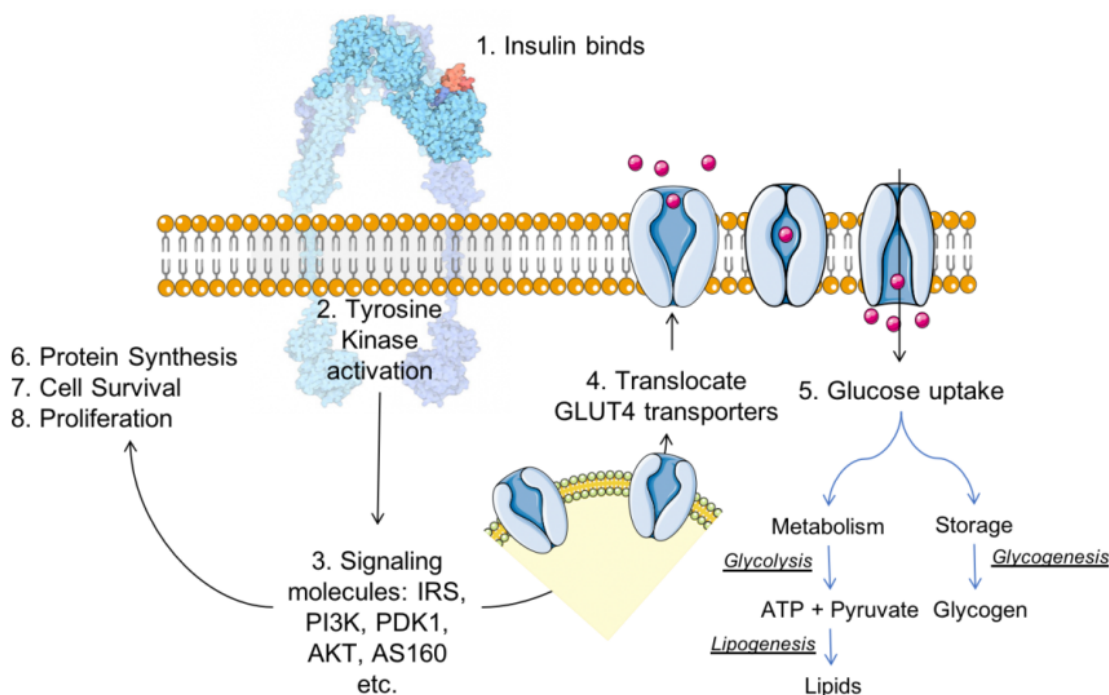


Рис. 2: Схема каскада реакций, запускаемых рецептором инсулина (синий) в результате связывания с инсулином (красный). <https://pdb101.rcsb.org/global-health/diabetes-mellitus/drugs/insulin/insulin-receptor>

- Происходит передача сигнала внутрь клетки - к внутриклеточному домену.
- Киназный домен запускает каскад реакций, приводящий к увеличению захвата глюкозы клеткой. Вследствие этого уровень глюкозы в крови понижается.

Поскольку рецептор отвечает за столь важный процесс в организме человека, нарушения в его работе приводят к различным клиническим проявлениям, в том числе к диабету. Поэтому процесс активации IR - это актуальный вопрос современной биохимии. Знание этого процесса на молекулярном уровне может послужить базой для создания лекарств на основе структуры белка (Structure-based drug design (TO-DO add reference to some article??)).

2.1.1 Модели активации IR

В то время как для внеклеточного и киназного доменов возможно подробное исследование экспериментальными методами, в том числе с помощью сгусто ЕМ [1], трансмембранный домен представляет особую сложность из-за необходимости создания липидного окружения. Поэтому были предложены разные механизмы активации IR [2] (см рис. 3).

На данный момент общепринятой моделью, подтверждающейся экспериментальными данными, является предложенная Lee [3] (см рис. 3С). В данной модели ТМ домены

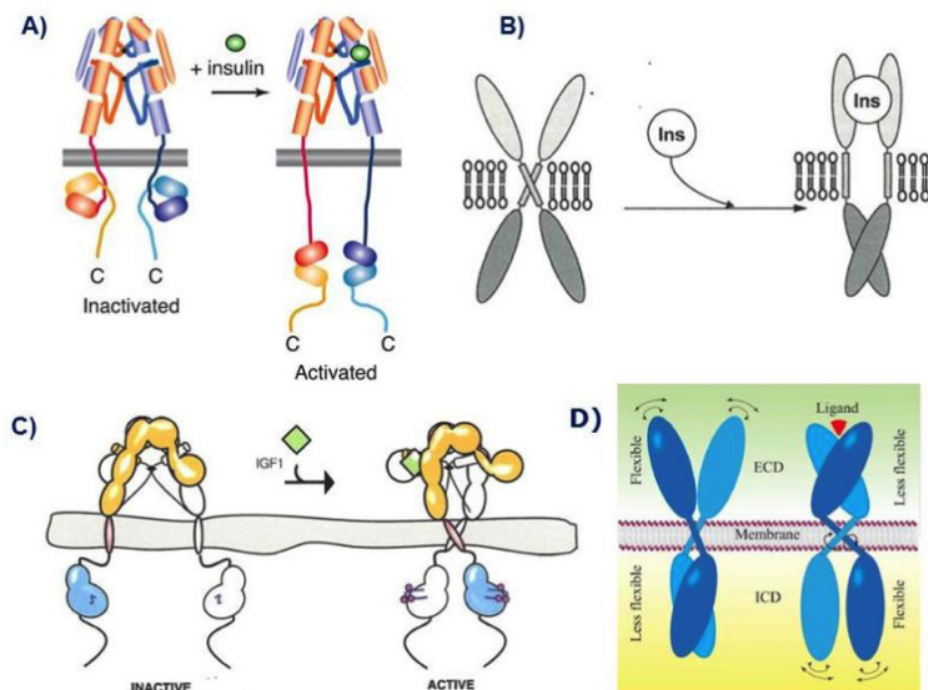


Рис. 3: Models of the mechanism of ligand activation of the insulin receptor tyrosine kinase. A: The "yo-yo" model. B: The TM domain ligand-induced separation model. C: The TM domain ligand-induced approximation model. D: The "rotation model". [2]

оказываются достаточно свободными и подвижными в мембране, чтобы иметь возможность самодимеризации. Далее, из-за такого изменения конформации ТМ участка, киназный домен переходит в активированное состояние.

Таким образом, ввиду особой сложности для эксперимента, 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 Трансмембранный домен IR

2.2.1 Экспериментальные данные

ТО-ДО добавить инфу про двумерную ЯМР спектроскопию (краткое описание метода)

С помощью вышеописанного метода была расшифрована структура ТМ домена IR в мицеллах [4]. Был выделен участок рецептора 940–988, содержащий трансмембранный домен и получен спектр (?? ТО-ДО как его назвать??). По спектру была расшифрована

структура TMD.

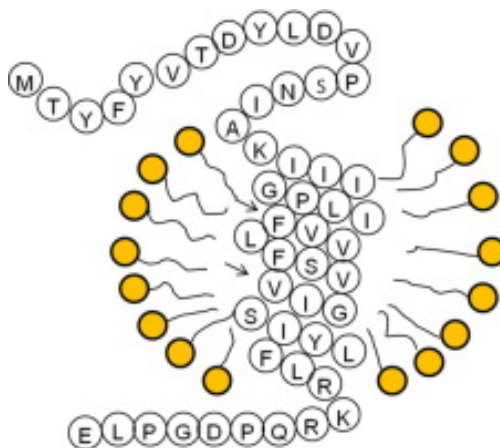


Рис. 4: Схематическое изображение ТМ домена IR (остатки 940–988) в мицелле [4].

В статье также содержится ряд замечаний по поводу полученной структуры.

- Несмотря на то, что в работе удалось получить только мономеры белка, авторы указывают на возможность существования структур в форме димеров в мицеллах.
- Внемембранные участки являются подвижными и располагаются в водном окружении.
- Мембранный участок имеет структуру α -спирали. При этом был отмечен кинк в этой спирали, созданный связью Gly960-Pro961, потенциально важный для функции домена.

В лаборатории биомолекулярной ЯМР-спектроскопии ИБХ РАН был проведен аналогичный эксперимент с целью получения димерных форм в мицеллах DPC. Был выделен участок 944-973 и созданы образцы мономерных и димерных форм в мицеллах. Спектры показали необычные результаты в отношении кинка Gly-Pro. Конкретно, спектр мономера двоятся в N-концевых остатках, начиная с остатка Pro961, 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 TMD IR.

ТО-ДО добавить график вторичной структуры и спектр (если это возможно)

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

2.2.2 Данные моделирования

Ранее в лаборатории моделирования биомолекулярных систем ИБХ РАН были предсказаны наиболее вероятные структуры димеров для TMD IR [5]. Для получения димеров на основе последовательности мономера был использован программный пакет PREDDIMER (см. раздел 2.4.3).

Структуры димеров, полученные с помощью PREDDIMER, затем были помещены в мембранное окружение из молекул пальмитоилолеилфосфатидилхолина (POPC). Далее была проведена минимизация энергии, релаксация и расчет 50 нс траектории молекулярной динамики (2.4.2). Затем методом потенциала средней силы (2.4.3) были определены 3 наиболее стабильные димерные структуры. Все они имеют характерную водородную связь Ser969-Ser969. Эти модели выбраны для исследования структуры димеров в контексте цис-транс конфигураций в настоящей работе.

ТО-DO добавить картинку этих структур из диплома Мифтаха (если это возможно) или сделать свои картинки

2.3 Пролин в трансмембранных α -спиралях

2.3.1 Влияние пролина на свойства спирали

Принято считать, что остаток пролина в трансмембранных спиралях играет функциональную роль. На это указывает сравнительно высокая встречаемость пролина в трансмембранных спиралях по сравнению со спиралями водорастворимых белков [6], [7]. Кроме того, было показано, что внесение мутаций в Pro в ТМ домене транспортных белков приводит к деактивации их механизма действия [8]. Все это дает основание полагать, что пролин играет ключевую роль в механизме действия трансмембранных доменов большого количества белков.

Как компьютерное моделирование, так и экспериментальные данные показывают, что пролин вносит конформационные изменения в геометрию ТМ α -спирали [10], [9]. Цис- конформация пептидной связи между XXX-Pro создает так называемый "кинк" в α -спирали, что может изменять ее ориентацию в мембране и, как следствие, меняет функциональные свойства пептида. Структуры с кинками отслеживаются в эксперименте наряду с обычными прямыми спиралями, что позволяет говорить о существовании обеих конформаций.

Также было замечено, что пролин в трансмембранных спиралях имеет достаточно консервативное положение: он чаще всего оказывается "вытащен" в область гидрофильных головок [11]. 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 (трансмембранного домена) индуцирует димеризацию мономеров в мембране [12]. Этот факт вызывает особый интерес в контексте изучения механизма работы ТМД в инсулиновом рецепторе.

2.3.2 Цис-транс изомеризация

В исследовании цис-транс конформаций α -спиралей в мембранах особняком стоит вопрос о цис-транс изомеризации. В эксперименте тяжело зафиксировать конформационный переход, меняющий геометрию пептида в связи с необходимостью накопления данных от множества пептидов. Моделирование транс-цис перехода вычислительными методами также является трудоемкой задачей, требующей введения более сложных силовых полей. Это приводит к недостатку данных о возможности, а также роли цис-транс изомеризации в работе ТМД.

Так же сложность в вопрос конформационного изменения вносит энергия перехода. Если в мембранной спирали происходит подобное изменение с образованием кинка посередине, то дополнительно к энергии изомеризации пептидной связи добавится энергия поворота участка α -спирали в мембране. Это наличие избыточной энергии может являться причиной того, что пролины преимущественно оказываются в гидрофильной области, т.к. в таком случае исчезает проблема поворота участка спирали в мембране.

Несмотря на сложности в получении данных, известны рецепторы, в которых цис-транс изомеризация, приводящая к конформационным изменениям, является механизмом передачи сигнала, к примеру рецептор родопсин, в котором поглощение кванта света дает энергию изомеризации. Также подобные изменения в ТМ спиральях показаны в ионных каналах с помощью мутагенеза пролина [13].

2.4 Компьютерное моделирование биологических молекул

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

Методы АМ 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, колоссально ускоряет и удешевляет исследование. Структурно-динамическая информация, полученная численно, ложится в основу эксперимента, снижая затраты на поиск необходимых структур.

2.4.1 Молекулярная динамика

Молекулярная динамика (MD) - 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 поля¹.

Полученная система уравнений решается через некоторый малый промежуток времени Δt 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 удачного расчета MD является стабильность системы предмет + окружение. Систем 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, усредненное либо по атомам (RMSD), либо по временному промежутку (RMSF). 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. В настоящей работе используется пакет Gromacs.

¹Какое-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.

2.4.2 Классический метод расчета MD для трансмембранного белка

Стандартный расчетный эксперимент для трансмембранного белка состоит из пяти основных частей:

- Погружение белка в мембранное окружение.
- Минимизация энергии составленной системы.
- Релаксация системы.
- Расчет траектории MD.
- Обработка полученных результатов.

Изначально, белок и мембрана² представлены в виде отдельных файлов, содержащих информацию о координатах каждого атома и принадлежности к определенной группе. Поскольку ТМ белок не существует без мембранного окружения, а мембрана является сложной стабилизированной системой, возникает проблема погружения белка в мембрану. Для решения этой проблемы используется подход InflateGRO [14].

Мембрана растягивается в ширину в 4 раза и совмещается с белком. Те липиды, которые попали в пересечение с молекулой белка, удаляются. После этого мембрана постепенно сжимается с коэффициентом 0.95 с минимизацией энергии на каждом шаге, пока не достигнет своей стандартной плотности.

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

После получения итоговой траектории проводится ее визуальная оценка с помощью программ визуализации. Затем вычисляется RMSD и RMSF, по которым окончательно можно судить о стабильности белка в мембране.

²В 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 бислоем 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.

2.4.3 PREDDIMER и метод потенциала средней силы

In vivo формирование трансмембранных димеров происходит при соблюдении соответствия свойств поверхностей. Поэтому, по карте гидрофобности мономеров можно предсказать, как будет формироваться димерная форма. Программа PREDDIMER реализует такой подход. Она позволяет предсказывать димеры по последовательности для α -спиральных ТМ пептидов.

Сначала по аминокислотной последовательности формируются идеальные спирали. Затем происходит вычисление их поверхности и расчет ее гидрофобных свойств. Затем полученные двумерные карты сопоставляются и с помощью варьирования угла между этими картами ищется интерфейс димеризации. В итоге для наиболее вероятных вариантов восстанавливается конечная структура димера.

После уравнивания структуры методом MD, возникает задача вычисления энергии взаимодействия мономеров в димере, чтобы установить наиболее стабильные. Для этого используется метод потенциала средней силы.

TO-DO добавить инфу про метод потенциала средней силы

3 Материалы и методы

4 Результаты и обсуждение

5 Выводы

Список литературы

- [1] 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.
- [2] 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).
- [3] Lee J, Miyazaki M, Romeo GR and Shoelson SE. Insulin receptor activation with transmembrane domain ligands. *J Biol Chem* 289: 19769-19777 (2014)
- [4] 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
- [5] 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.
- [6] 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
- [7] 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
- [8] Slepko ER, Chow S, Lemieux MJ, Fliegel L. Proline residues in transmembrane segment IV are critical for activity, expression and targeting of the Na⁺/H⁺ exchanger isoform 1. *Biochem J* 2004; 379: 31–38
- [9] Cordes FS, Bright JN, Sansom MSP. Proline-induced distortions of transmembrane helices. *J Mol Biol* 2002; 323: 951–960
- [10] L.K. Iyer, S. Vishveshwara A model for transmembrane helix with a cis-proline in the middle *FEBS Lett.*, 374 (1995), pp. 21-24
- [11] 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

- [12] 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
- [13] 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)
- [14] 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