

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ВИСОКИХ ТЕХНОЛОГІЙ

Завідувач кафедри молекулярної біотехнології та біоінформатики

к.б.н, доц. Нипорко О. Ю.

Протокол №_____ засідання кафедри

від “_____” _____ 2023 р

**СПЕКТРАЛЬНИЙ ТЮНІНГ РОДОПСИНУ *BOS TAURUS* ДЛЯ
ДОСЯГНЕННЯ БАТОХРОМНОГО ЕФЕКТУ**

Випускна кваліфікаційна робота магістра
студентки спеціальності 091 Біологія
ОП «Біоінформатика і структурна біологія»
Нечепуренко Діани Віталіївни

Науковий керівник від кафедри
завідувач кафедри молекулярної біотехнології
та біоінформатики
к.б.н., доцент **Нипорко Олексій Юрійович**

Робота виконана у Computational drug discovery lab

Brain Science Institute of Korean Institute of Science and Technology

Під керівництвом доктора біологічних наук

Hahnbeom Park (박한범)

Оцінка захисту роботи

Київ – 2023 р.

Анотація

Нечепуренко Д.В. Спектральний тюнінг родопсину *Bos taurus* для досягнення батохромного ефекту.

Ця робота присвячена створенню молекулярної моделі родопсину з унікальними спектральними характеристиками для застосування в оптогенетичних дослідженнях. Родопсини є світлочутливими білками, які можуть бути використані для контролю активності клітин за допомогою світла. Дослідники використовують генетично модифіковані версії родопсинів, називані оптогенетичними інструментами, для специфічного контролю активності нейронів та інших клітин. Однак наразі відомо про успішні спектральні зсуви родопсинів до менших довжин хвиль. У цій роботі описується підхід до створення родопсину з батохромним ефектом, тобто зі здатністю активуватися під дією світла більшої довжини хвилі.

Дослідження, проведене на базі Korea Institute of Science and Technology, пропонує використання химерного білку, який поєднує бичачий родопсин з внутрішньоклітинним доменом дофамінового рецептору, для створення нових оптогенетичних інструментів зі специфічними властивостями. Результати дослідження вказують на можливість створення родопсину зі здатністю активуватися під дією світла 473 нм та запуском внутрішньоклітинного молекулярного каскаду дофамінового рецептора. На основі отриманих даних, була сформульована та поставлена задача змінити саме хромофорні характеристики бичачого родопсину з можливістю активації під дією довших довжин хвиль. Нові варіанти родопсинів з унікальними характеристиками піків поглинання у поєднанні з химерними внутрішньоклітинними доменами дадуть змогу одночасної специфічної

активації різних нейронних відповідей в рамках одного оптогенетичного експерименту.

Для досягнення поставленої мети ми запропонували детальний порівняльний аналіз, включаючи коеволюційний аналіз, аналіз послідовностей, докінг-аналіз та аналіз молекулярної динаміки. Ці підходи дозволяють раціонально вводити точкові мутації у молекулярну структуру родопсину, уникаючи значних структурних порушень. Крім того, важливим етапом є отримання негативного контролю, який використовується для моделювання молекулярної динаміки родопсину. Це здійснюється шляхом введення експериментально відомих мутацій, які порушують структуру родопсину.

Результати цієї роботи сприятимуть розвитку оптогенетики шляхом створення нових оптогенетичних інструментів з унікальними спектральними характеристиками. Здатність бичачого родопсину активуватися під дією світла більшої довжини хвиль розширить можливості контролю активності клітин та нейронних відповідей. Дані дослідження можуть мати велике значення для вивчення функціонування нервової системи та пошуку нових шляхів лікування розладів, пов'язаних з нею.

Ключові слова: родопсин, оптогенетика, світлочутливі білки, спектральні характеристики, молекулярне моделювання, мутагенез, докінг, молекулярна динаміка, батохромний ефект.

Зміст

Анотація	2
Вступ	5
Розділ 1. Характеристика спектральних особливостей опсинів	7
1.1. Функціональні особливості родопсину	11
1.1.1. Метародопсини	12
1.1.1. Дезактивація родопсину	15
1.2. Правило «п'яти сайтів» та «ОН-сайтів» для кольорового зору	16
1.2.1. Розвиток зору червоного та зеленого кольорів у ссавців і принцип, відомий як правило «п'яти сайтів»	16
1.2.2. Правило «ОН-сайтів»	18
Розділ 2. Використання бичачого родопсину в оптогенетиці	20
Розділ 3. Практична частина	23
3.1. Додаткова валідація правила «ОН-сайтів»	23
3.2. MSA та коеволюційний аналіз	27
3.3. Мутагенез в PyRosetta	32
3.4. Референсний докінг в Rosetta GALigandDock	36
3.5. Симуляція молекулярної динаміки та негативний контроль	39
Висновки	42
Джерела	44

Вступ

Родопсини - це світлочутливі білки, які застосовуються в оптогенетиці для контролю активності нейронів та інших клітин у відповідь на світло. Родопсини знаходяться в сітківці ока, де вони відповідають за виявлення світла та ініціювання зорової реакції. Дослідники використовують генетично модифіковані версії родопсинів, що можуть експресуватися в конкретних клітинах, наприклад, нейронах. Ці модифіковані родопсини називають оптогенетичними інструментами, і вони дозволяють дослідникам контролювати активність клітин за допомогою світла. Родопсини зазнають конформаційних змін, коли піддаються впливу світла певної довжини хвилі, що викликає активацію клітини.

Для оптогенетичних досліджень використовуються декілька типів родопсинів, включаючи каналородопсини, галородопсини та археродопсини. Бичачий родопсин є одним з них і часто використовується в оптогенетиці як шаблон для створення нових оптогенетичних інструментів зі специфічними властивостями та функціями.

Одне з недавніх досліджень, що було проведене на базі Korea Institute of Science and Technology (KIST) в лабораторії оптичного зондування та контролю динаміки білків показало успішне використання химерного білку бичачого родопсину з внутрішньоклітинним доменом дофамінового рецептору[1]. Такі результати нашої роботи на ідею створення варіацій химерних білків з активацією різних внутрішньоклітинних сигналів під дією світла різного спектру. Оскільки на даний момент відомо лише про успішні спектральні зсуви родопсину в сторону менших довжин хвиль, у дослідженні, описаному в даній

роботі описано підхід до створення батохромного ефекту (зсув до більших довжин хвиль).

Актуальністю даної роботи є необхідність побудови молекулярної моделі родопсину з унікальними спектральними характеристиками для використання в оптогенетичних дослідженнях. Родопсин, здатний активуватися під дією світла червоного спектру забезпечить можливість селективної одночасної активації різних нейронних відповідей.

Мета: дослідити за допомогою методів молекулярного моделювання, мутагенезу, докінгу та молекулярної динаміки варіанти бичачого родопсину, які можуть призвести до батохромного ефекту.

Завдання: за допомогою гомологічного моделювання побудувати структури білків опсинів з піками поглинання в спектрі довших довжин хвиль; провести докінг мутованих варіантів із лігандом (ретиналом); провести симуляції молекулярної динаміки; проаналізувати результати.

Розділ 1. Характеристика спектральних особливостей опсинів

Зорові опсини тварин — це група білків, які належать до суперкласу мембранних рецепторів, зв'язаних з G-білками (GPCR), і утворюють функціонально різноманітні світлочутливі фотопігменти [2]. Опсини є унікальними серед GPCR, оскільки кожен з них постійно зв'язує неактивну форму свого специфічного ліганду, надаючи майже миттєву відповідь на фотоактивацію [3]. Світлочутливість забезпечується зв'язаною молекулою або хромофором, цис-ретиналем (або спорідненим), який зазвичай походить від вітаміну А. Вільний хромофор сітківки максимально поглинає фотони ультрафіолетового (УФ) світла близько 380 нм, але його чутливість зміщується, щоб реагувати на більш високу енергію з меншою довжиною хвилі, коли він зв'язаний з опсином ультрафіолетового випромінювання з короткою довжиною хвилі (SW), і на нижчу енергію, довші довжини хвилі, коли прив'язані до синього середньохвильового (MW) або довгохвильового (LW) опсину.

Хромофор утворює ковалентний зв'язок із залишком лізину (K) у місці зв'язування ліганду опсину через протоновану основу Шиффа. Цей зв'язок зберігається у всіх опсинах тварин і стабілізується сусіднім негативно зарядженим залишком, протиіоном (рис. 1.1). Коли неактивна цис-форма хромофора поглинає фотон світла, вона змінюється на нову конформацію повністю транс-ретинальної. Це, у свою чергу, викликає конформаційні зміни в опсині, пов'язаному з білком Gt у хребетних або білком Gq у безхребетних, таким чином активуючи чіткі сигнальні каскади, що ведуть до сприйняття світла [4]. Оскільки комплекс хромофор-опсин (також званий родопсин) налаштований на певну довжину хвилі світла, максимальне поглинання опсину (λ_{max}) представляє фенотип, який можна безпосередньо виміряти. Хоча останні

дослідження виявили невізуальні ролі традиційних візуальних опсинів [5], такі як розрізнення температури і смакові функції, тісний зв'язок між генотипом і фенотипом спектральної чутливості робить опсини потужною моделлю для розуміння фенотипової адаптації.

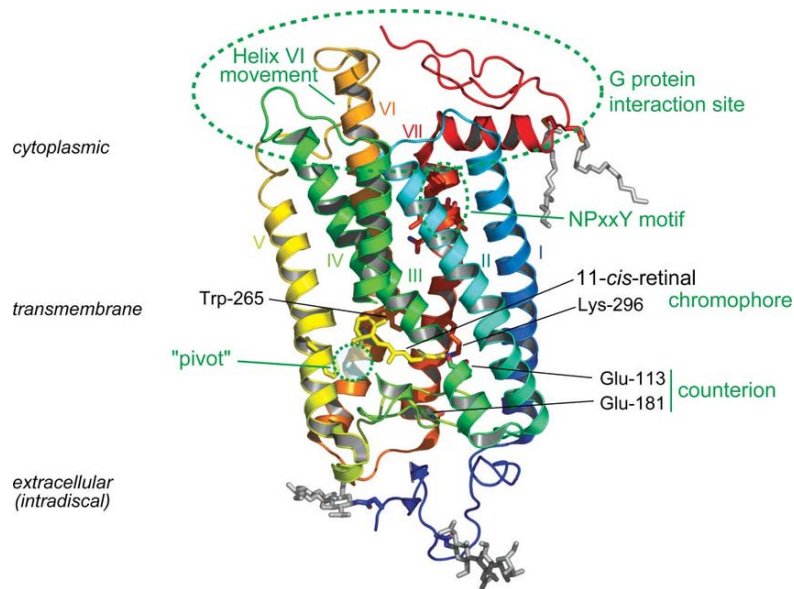


Рис. 1.1. Структура родопсину з визначеними ключовими залишками [6]. Показано структурну модель бичачого родопсину (код PDB 1U19).

Опсини виникли на початку еволюції тварин і дублювалися [7], щоб дати початок трьом основним групам генів: с-опсином і г-опсином, основним зоровим пігментам хребетних і членистоногих, відповідно, і менш відомим G-пов'язаним рецепторам сітківки/Go-опсином [8]. У той час як зорові пігменти хребетних є білками, пов'язаними з Gt, які активують каскад трансдуцину (Gt), що призводить до реакції гіперполяризації [9], опсини безхребетних пов'язані з білками Gq, які ініціюють каскад фосфоінозитолу, що призводить до реакції деполіаризації фоторецепторних клітин [10]. На додаток до активації різних шляхів фототрансдукції, кожна група містить різні гени опсинів, які дають початок зоровим пігментам, які еволюціонували від одного предка через

ідіосинкратичні дуплікації генів і максимально поглинають на різних довжинах хвиль.

Різноманіття предків зорових опсинів хребетних включає чотири сімейства генів: чутливі до короткохвильового випромінювання – SWS1 (УФ, λ_{max} 344–445 нм) і SWS2 (синій, λ_{max} 400–470 нм), а також середньо- та довгохвильово чутливі опсини – MWS (зелений, λ_{max} 480–530 нм) та LWS (червоний, λ_{max} 500–570 нм) відповідно [11], [12]. Візуальні γ -опсини членистоногих включають п'ять добре охарактеризованих сімейств, включаючи чутливі до коротких хвиль (SW), чутливі до середньої довжини хвилі (MW1 і MW2) і чутливі до довгих хвиль (LW1 і LW2) [13]. Предкові лінії комах, швидше за все, містили три типи опсинів: SW (або UV), MW (або синій) і LW (або зелений) опсин, з розширенням кількості та/або втратою генів, що пояснює дивовижне розмаїття опсинів, що спостерігається серед існуючих груп комах. Таким чином, родини опсинів хребетних і безхребетних є функціонально аналогічними, але філогенетично віддалено спорідненими, вони конвергувалися незалежно, щоб поглинати однакові діапазони світлового спектру.

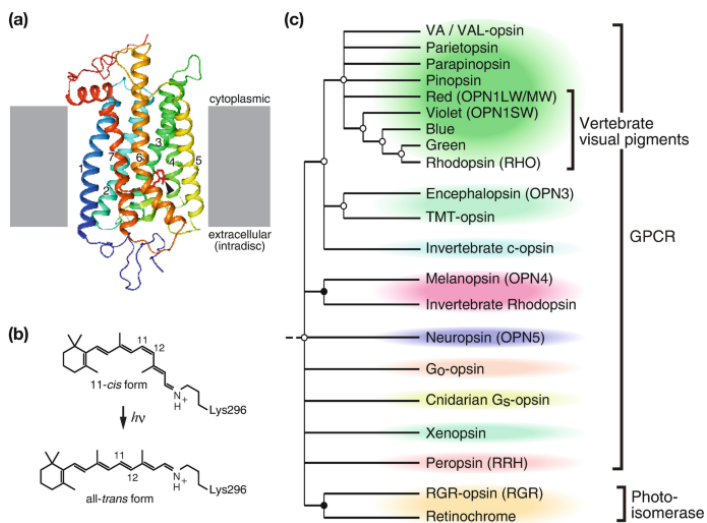


Рис. 1.2. Структура та різноманітність білків родини опсинів. (a) Кристалічна структура бичачого родопсину. (b) Хімічні структури 11-цис

(верхня) і повністю транс (нижня) форм ретиналю (ретиніліденова основа Шиффа). (с) Філогенетичний зв'язок між підродинами в родині опсинів. У дужках вказано номенклатуру генів людини. Вузли з замкнутим колом представляють дивергенцію видів, тоді як вузли з незамкненим колом представляють дублювання генів [14].

Основним фактором, що визначає спектральну чутливість фоторецептора, є спектр поглинання експресованого зорового пігменту. Оптичні та електрофізіологічні фактори, такі як скринінг та фільтрація пігментів, а також нейронна обробка інформації, можуть додатково зсунути пік чутливості фоторецептора до певної довжини хвилі світла. На додаток до ефектів тюнінгу бічної фільтрації, максимуми поглинання зорових пігментів можна зрушити шляхом зміни типу хромофора з поширеного ретиналю (A1) або гідроксиретиналю (A3) на дегідроретиналь (A2), що спонукає зоровий пігмент поглинати хвилі з більшою довжиною. Інший механізм включає зміну амінокислотних залишків опсину, які знаходяться всередині або поруч із сайтом зв'язування хромофору, змінюючи здатність хромофора поглинати певні довжини хвилі світла. Цей механізм пояснює велике розмаїття спектральних чутливостей, необхідних для кольорового зору, і є результатом динамічної історії диверсифікації та молекулярної еволюції опсинів [15].

Вражаюча фенотипова конвергенція в спектральній чутливості в основних групах опсинів хребетних і безхребетних не обов'язково означає конвергенцію механізмів тюнінгу на молекулярному рівні. Оскільки амінокислотні заміни ключових залишків, що оточують хромофор, відбувалися протягом еволюційної історії кожної лінії, різні механізми налаштування, ймовірно, мали модульовану спектральну чутливість різних груп опсинів [16], [17].

Механізми, що лежать в основі спектрального налаштування у зорових опсинах хребетних, були детально охарактеризовані, зокрема за допомогою спектроскопічного аналізу з рекомбінантними білками опсинів дикого типу та мутантних [18]. Навпаки, механізми налаштування опсинів безхребетних все ще в основному невідомі через труднощі отримання достатнього виходу очищених рекомбінантних пігментів для спектроскопічного аналізу, особливо для певних груп опсинів. Тим не менш, останніми роками спостерігається значний прогрес у доступності геномних ресурсів безхребетних, у вдосконаленні біоінформатики та інструментів моделювання, і головним чином у розробці надійних молекулярних аналізів для вивчення спектральної чутливості зорових пігментів безхребетних.

«Справжній родопсин» Rh1 хребетних [19] виявляє рівні слабкого освітлення («скотопічний зір») у щелепних хребетних, менш чутливий до різниць у кольорі та має високу чутливість до змін освітленого та темного середовища, форми та руху. Незважаючи на те, що Rh2 тісно пов'язаний з Rh1 [20], він експресується виключно в колбочкових фоторецепторах і є фотопсином за біохімічними властивостями. Колбочкові опсини SWS1, SWS2, LWS і Rh2 забезпечують «фотичний зір» в умовах високого рівня освітлення та відповідають за колірний зір. Виходячи з економності підтипів опсинів, присутніх у існуючих лініях, ранній предок хребетних, ймовірно, мав фототичний і скотопічний зір із с-опсинами SWS1, SWS2, LWS, Rh1 і Rh2 у їхньому геномі.

1.1. Функціональні особливості родопсину

Щоб активувати каскад трансдукції, а потім вимкнути його, необхідно (хоча недостатньо) активувати, а потім дезактивувати молекулу родопсину.

Нижче ми розглянемо молекулярні переходи, що лежать в основі активації та відключення родопсину [21].

1.1.1. Метародопсини

Ранні проміжні продукти. Давно відомо, що після майже миттєвої зміни під час фотоізомеризації молекула родопсину проходить серію переходів через різні проміжні стани, і рис. 1.3 показує поточний вигляд основних фотоінтермедіатів та їх переходи. У темному стані спокою хромофор родопсину, 11-цис-ретиналь, ковалентно зв'язаний через протоновану основу Шиффа з Lys296 на TM7 [22], [23]; протон у основі Шиффа стабілізується протионом у Glu113 на TM3.

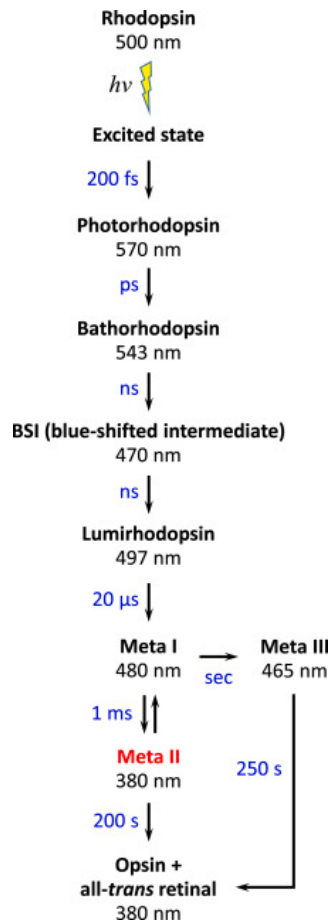


Рис. 1.3. Фотопродукти фотолізу родопсину [21].

У темряві 11-цис-ретинальний хромофор зв'язується з апопротеїном опсину за допомогою протонованої основи Шиффа, утворюючи таким чином темний стан спокою родопсину. При поглинанні кванта світла хромофор ізомеризується протягом 200 фс з 11-цис-конфігурації до повністю-транс-конфігурації [24], [25]. Потім білок термічно релаксує через низку проміжних продуктів, кожна з яких характеризується своїм спектром поглинання УФ/видимого світла, із зазначеною довжиною хвилі максимального поглинання (λ_{max}) і з приблизно вказаним часом для родопсину ссавців *in vivo*. Активний стан Meta II, який зв'язує та активує трансдуцин G-білка, показано червоним кольором. Гідроліз зв'язку ретиніліденової основи Шиффа генерує вільний опсин і транс-ретиналь.

Із надзвичайною швидкістю (на шкалі часу від фс до нс) ізомеризована молекула проходить через стани, які називаються фотородопсин, батородопсин і BSI (проміжний продукт із зміщенням в спектр синього кольору), перш ніж досягти стану, який називається люміродопсин (Lumi), який максимально поглинає при довжині хвилі 497 нм, дуже близькому до піку родопсину. З вивільненням напруги в ретиноїді енергія передається в білок, але активація залишається поблизу сайту зв'язування ретиналю. Цілком імовірно, що вже в Lumi більша частина енергії фотонів, поглиненої родопсином, була передана апопротеїну. У рентгенівській структурі Lumi [26] викривлення транс-ретиналю в батородопсині вже послаблено через дислокацію β -іононового кільця.

Метародопсини I і II. При температурі тіла Meta I перетворюється на Meta II приблизно за 1 мс, і в умовах *in vivo* рівноважне положення цього етапу наближається до Meta II. Цей перехід включає депротонування зв'язку Шиффової основи, який був протонований у стані спокою та у всіх фотоінтермедіатах до Meta I; в результаті спостерігається значний зсув

спектрального поглинання від видимого до ультрафіолетового до піку при 380 нм для Meta II [27]–[29]. Спочатку передбачалося, що Meta II складається з одного молекулярного стану, але згодом стало зрозуміло, що існують три різні молекулярні підстани, кожна з депротонованою основою Шиффа, і тому спектрально нерозрізнена. Відповідно, усі ці три підстани підпадають під загальний термін «метародопсин II»; вони називаються Meta IIa, Meta IIb і Meta IIbH⁺ [30], [31]. У першому з них (Meta IIa) відбувається лише незначна структурна зміна порівняно з Meta I, окрім транслокації протонів; у другому (Meta IIb) відбувається значне переміщення TM6 назовні; а в третьому (Meta IIbH⁺) протон поглинається з водного середовища для нейтралізації Glu134 у мотиві ERY. Саме цей останній стан, Meta IIbH⁺, є «активним» у тому сенсі, що він каталізує активацію G-білка, трансдуцину; тому це дорівнює терміну «R*» або «Rh*», який використовується в електрофізіології.

Треступенева активація. Дані переходи порушують два основних обмеження, які утримують основний стан родопсину неактивним, а саме: мережу водневих зв'язків Шиффової основи, яка забезпечує центральний іонний «замок», та іонний «замок» TM3–TM6. Під час 3-етапного переходу від Meta I до Meta IIbH⁺ ці два основні обмеження послідовно порушуються, і тепер молекула має високу ймовірність перебувати в активному стані. Кінцевим результатом фотоізомеризації, що супроводжується цими переходами, є те, що родопсин перетворюється протягом мс з майже повністю каталітично неактивного на потужно каталітично активний.

Метародопсин III. Хоча «нормальним» шляхом розпаду Meta II є вивільнення транс-ретиналю з утворенням опсину, також відомо, що не весь Meta II розпадається таким чином. Тимчасово з'являється помаранчевий продукт ($\lambda_{\text{max}} \approx 465$ нм), який повільно замінює прозорий продукт Meta II

($\lambda_{\text{max}} = 380 \text{ nm}$). Мета III утворюється за хвилини, і вона так само розпадається на опсин, але (за відсутності регуляторних факторів) ще повільніше. При високому рН до 40% Мета I/II проходить через Мета III [32]–[34].

1.1.1. Дезактивація родопсину

Активність Мета II^{H+} (R*) дуже швидко вимикається шляхом двоетапного процесу фосфорилювання та зв'язування арестину, як показано на рис. 1.4, де етапи 1 і 2 представляють етапи активації, узагальнені вище. Крок 3 показує фосфорилювання R* родопсинкіназою. Крок 4 показує зв'язування арестину. Етапи 3 і 4 відбуваються швидко, в результаті чого активність R* відключається протягом 100 мс.

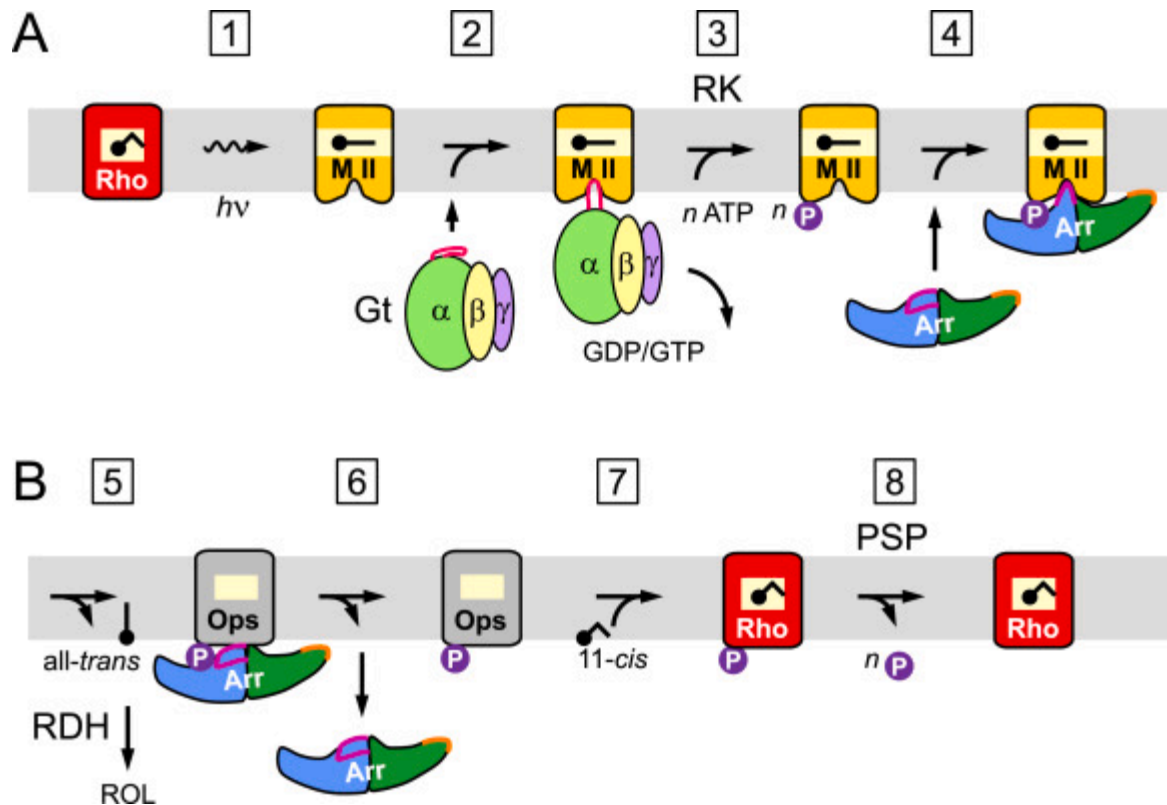


Рис. 1.4. Дезактивація родопсину та повернення до стану спокою [21].

1.2. Правило «п'яти сайтів» та «ОН-сайтів» для кольорового зору

Більшість ссавців мають дихроматичний кольоровий зір, але гоміноїди та мавпи Старого Світу мають трихроматичний зір, оскільки він допомагає їм виявляти жовті та червоні фрукти на тлі плямистого листя [35]. Хоча було досягнуто значного прогресу у вивченні структурно-функціональних зв'язків, відповідальних за зір у тьмяному світлі, розуміння кольорового зору на молекулярному рівні залишається складним питанням. Незважаючи на труднощі, експериментальні дослідження мутагенезу дали цінні відомості, необхідні для ідентифікації ключових молекулярних елементів, що опосередковують кольоровий зір [36].

1.2.1. Розвиток зору червоного та зеленого кольорів у ссавців і принцип, відомий як правило «п'яти сайтів»

Закономірності, описані нижче, були отримані переважно в результаті дослідження саме опсинів приматів [37], оскільки визначення патернів еволюції кольорового зору у найближчих до людини представників могли б відповісти на фундаментальні питання виникнення кольорового зору людини. Однак, було вирішено взяти до уваги будь-які підтвержені правила чи закони, що стосуються родини опсинів, оскільки їх у подальшому можна редукувати до хіміко-фізичних закономірностей, що в теорії, працюватимуть на будь-якій структурі опсинів.

Отже, згідно з першими роботами Йокоями, основну відмінність між чутливими до середньої довжини хвилі (MWS) або зеленими зоровими пігментами та чутливими до довгохвильових променів (LWS) або червоними зоровими пігментами у ссавців можна віднести до певних варіацій амінокислот. Зокрема, амінокислотний мотив AFA (позиції 180, 277 і 285) знаходиться в зеленому пігменті, тоді як мотив SYT присутній на відповідних червоного

пігменту. Це правило «трьох місць» було широко виявлено у приматів, що підтверджено дослідженнями, проведеними у 1990-их роках. Крім того, у людей додаткові незначні коригування червоного та зеленого пігментів вимагають відмінностей амінокислот у сайтах 116, 230, 233 та 309 [38]. Отже, широко визнано, що варіації в амінокислотах у трьох ділянках, зокрема в 277 і 285, головним чином відповідають за значну різницю в зеленому та червоному кольоровому зорі, що спостерігається у людей та інших приматів.

Сформульоване припущення було допрацьовано та вдосконалено тим же Йокоюмою, який справедливо вважається одним з першопрохідців у дослідженні (зокрема біоінформатичному) еволюції кольорового зору. У результаті дослідження, проведеного у 1988 році, було виявлено, що зміни в амінокислотах, зокрема S180A (S → A на ділянці 180), H197Y, Y277F, T285A та A308S, зміщують максимальну довжину хвилі поглинання (λ_{\max}) червоного та зеленого зорових пігментів до синього кінця спектру адитивним способом. Щоб ширше дослідити застосовність правила «п'яти сайтів», дослідники проаналізували часткові амінокислотні послідовності червоних і зелених пігментів п'яти груп ссавців: парнокопитних, хижих, зайцеподібних, непарнокопитних і гризунів. Результати показали, що пігменти kota (*Felis catus*), собаки (*Canis familiaris*) і кози (*Capra hircus*), усі з яких мають АНУТА у п'яти критичних позиціях, мали значення λ_{\max} близько 530 нм. Навпаки, пігмент щура (*Rattus norvegicus*) з АУУТS у цих місцях мав значення λ_{\max} приблизно 510 нм, що точно узгоджується з прогнозами, заснованими на правилі «п'яти сайтів». Проте спостережувані значення λ_{\max} відповідних пігментів у європейського кролика (*Oryctolagus cuniculus*), білохвостого оленя (*Odocoileus virginianus*), сірої білки (*Sciurus carolinensis*) і морської свинки (*Cavia procellus*) постійно перевищували прогнозовані значення більш ніж на 10 нм. Це свідчить про існування додаткових молекулярних механізмів, що

впливають на сприйняття червоного та зеленого кольорів. Аналіз передбачуваних амінокислотних послідовностей предкових організмів показує, що **сучасні червоні та зелені пігменти ссавців еволюціонували з одного предкового зелено-червоного гібридного пігменту шляхом цільових амінокислотних замін.**

1.2.2. Правило «ОН-сайтів»

Згідно з правилом «п'яти сайтів», зміни амінокислот у ділянках 164, 181, 261, 269 і 292 пояснюють 50-нм синій зсув між червоним (560 нм) і зеленим (~510–530 нм) зоровими пігментами. Очевидно, амінокислотні заміни стосуються заміни SHYTA композицією AYFAS (S164A, H181Y, Y261F, T269A, A292S). Більш простий спосіб поглянути на правило полягає в тому, що заміна серину на аланін (S→A), гістидину тирозином (H→Y), тирозину фенілаланіном (Y→F), треоніну аланіном (T→A) і аланіну серином (A→S) призведе до втрати або збільшення гідроксильної групи зоровим пігментом, що, у свою чергу, може спричинити зміну спостережуваного λ_{max} .

Згідно з правилом «ОН-сайтів» [36], якщо амінокислота містить ОН-групу і присутня поблизу частини β -іононового кільця ретиналю в опсині, вона стабілізує збуджений стан відносно основного стану хромофору і спричинить червоний зсув λ_{max} . Навпаки, якщо амінокислота містить ОН-групу та присутня поблизу основи Шиффа, вона стабілізує основний стан відносно збудженого стану та індукує синій зсув λ_{max} . Іншими словами, взаємодія ретиналю з диполярними залишками поблизу β -іононового кільця збільшить делокалізацію електронів і зменшить зміну довжини зв'язку (BLA), що призводить до червоних зсувів, тоді як взаємодія ретиналю з диполярними залишками поблизу Шиффової основи зменшить делокалізацію електронів і збільшить BLA, що призводить до синього зсуву. Це принцип був

продемонстрований [36], використовуючи амінокислотні заміни в сайтах 164, 261 і 269 (A164S, F261Y і A269T) у пігментах M/LWS мавп, у позиції 292 (A292S, A292T, A292Y) у пігменті бичачого RH1 і в позиції 111 (Y111) у пігменті RH1 кальмарів. Отримані результати дозволяють вважати, що **правило «ОН-сайтів» є специфічним для місця, а не для пігменту.**

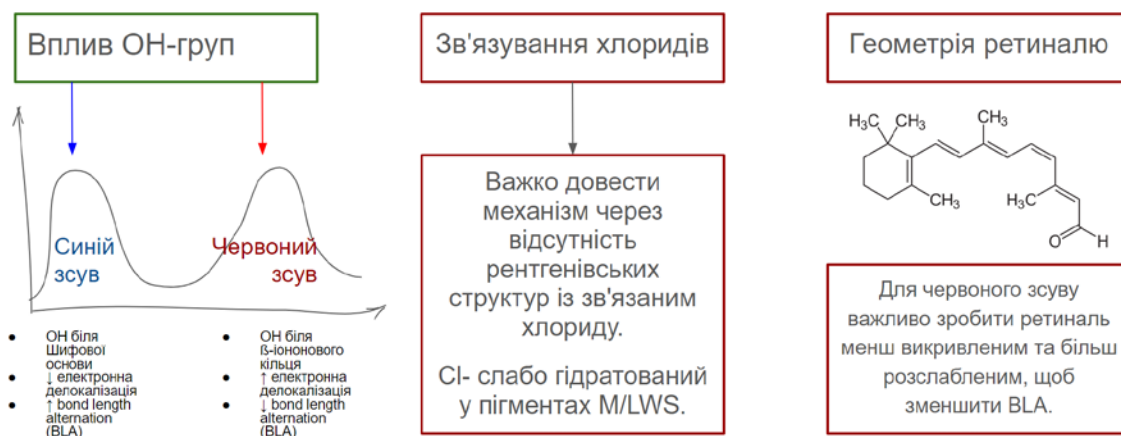


Рис.1.5. Узагальнене розуміння властивостей сайту зв'язування, що впливають на спектральні характеристики опсину.

Розділ 2. Використання бичачого родопсину в оптогенетиці

Галузь нейронаук отримала значну користь від оптогенетики, яка дозволяє точно контролювати нейронну активність. Одним із широко використовуваних інструментів в оптогенетиці є Каналродопсин (ChR), світлочутливий іонний канал, який може активувати нейрони при освітленні. Однак вибірково активувати специфічну нейротрансмісію в мозку за допомогою ChR складно. Щоб усунути це обмеження, була представлена нова оптогенетична техніка під назвою OptoXRs. OptoXR включає оптичну активацію специфічних G-білкових рецепторів (GPCR). Концепція спочатку була продемонстрована за допомогою OptoB2AR, рецепторної химери, що поєднує частини бичачого родопсину та B2AR [39]. Використовуючи синє світло, OptoB2AR активує специфічні для B2AR сигнальні події.

Відповідно до цієї стратегії було розроблено OptoDRD1 для дослідження ролі рецепторів DRD1 у функціях винагороди [40]. У нещодавньому дослідженні лабораторією Optical Sensing and Control of Protein Dynamics на базі Korea Institute of Science and Technology (KIST) було розроблено OptoDRD2 шляхом поєднання світлочутливої частини бичачого родопсину з внутрішньоклітинною частиною DRD2. Експерименти підтвердили, що стимуляція синім світлом клітин, що експресують OptoDRD2, індукує DRD2-специфічні шляхи. Потім було експресовано OptoDRD2 у збудливих нейронах латеральної блідої кулі (LGP) у мозку миші, щоб дослідити функцію DRD2 у популяціях глутаматергічних клітин ланцюга базальних гангліїв. Оптогенетична стимуляція OptoDRD2 у цих нейронах LGP може впливати на різні параметри руху, такі як частота та тривалість руху (рис. 2.1).

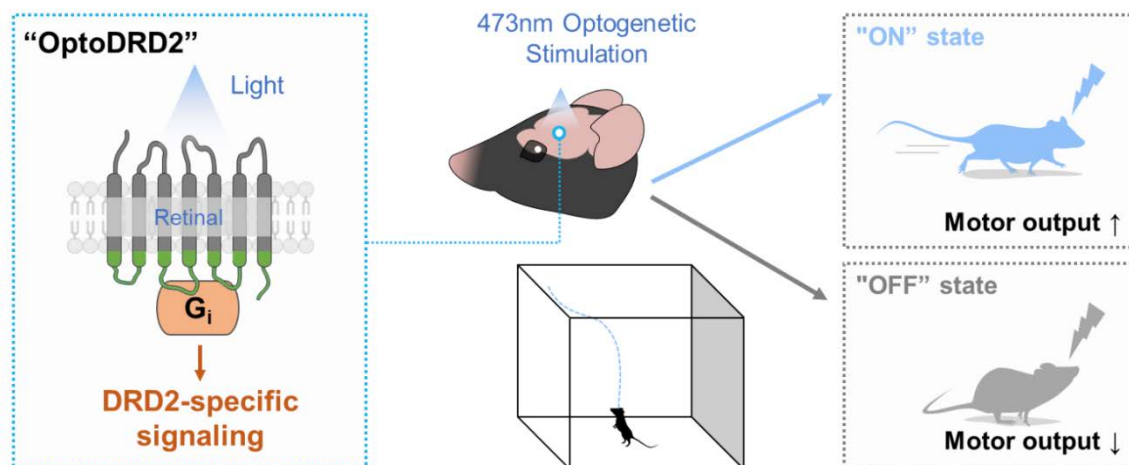


Рис. 2.1. Принцип дії OptoDRD2. Ладобраторією Optical Sensing and Control of Protein Dynamics було розроблено OptoDRD2, рецептор дофаміну типу 2 (DRD2), який реагує на синє світло. Під впливом синього світла OptoDRD2 активує специфічні сигнальні шляхи, пов'язані з DRD2. Завдяки експресії OptoDRD2 у бічній блідій кулі мишей спостерігали посилення моторних виходів. Цей OptoDRD2 можна ефективно використовувати для дослідження функцій DRD2 у певних областях мозку. Зображення надано професором Jihye Seong.

Успішна розробка химерного родопсину з внутрішньоклітинним доменом від іншого рецептору з родини GPCR підштовхнуло до ідеї створення й інших подібних білків, що під дією світла будуть активувати різні внутрішньоклітинні сигнали. Крім того, можливість контрольованої одночасної/поступової активації може значно розширити сучасні можливості методів оптогенетики. Таким чином, було вирішено розробити модифіковані бичачі родопсини, що могли б активуватися при інших довжинах хвиль (зокрема під дією червоного чи зеленого світла).

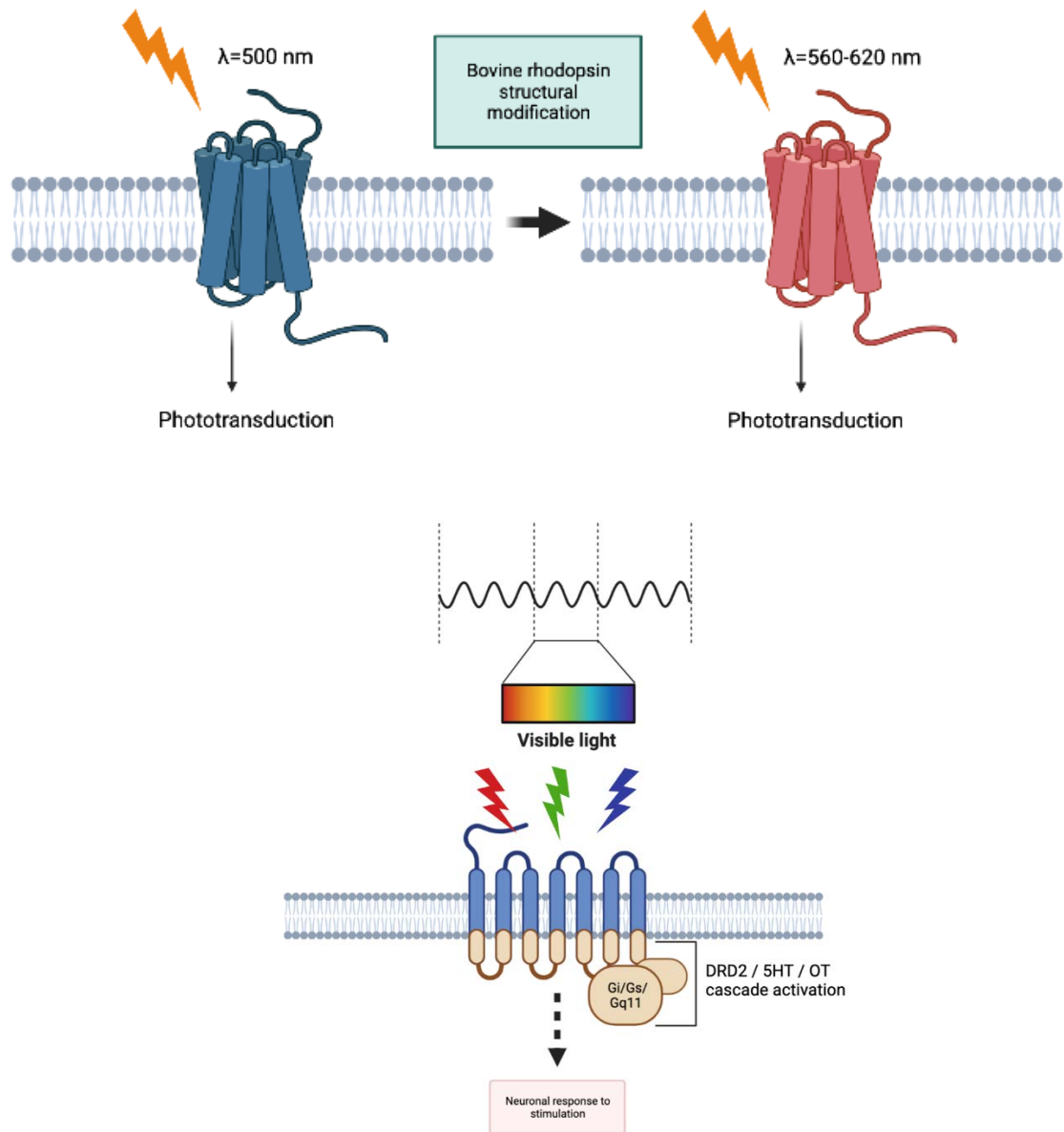


Рис. 2.2. Узагальнена ідея спектрального тюнінгу бичачого родопсину. Зліва показана довжина хвилі, при якій активується дикий тип бичачого родопсину, справа – бажаний результат батохромного зсуву. Знизу зображено потенційні варіанти внутрішньоклітинних доменів для створення химерних білків.

Розділ 3. Практична частина

Використання біоінформатичних методів для досягнення поставленої задачі обумовлено необхідністю в мінімізації хибних спроб введення точкових мутацій в бичачий родопсин *in vitro*.

В даній роботі використовувалися такі методи *in silico*:

1. Гомологічне моделювання.
2. Локальне вирівнювання послідовностей та аналіз множинного вирівнювання послідовностей (MSA).
3. Коеволюційний аналіз.
4. Мутагенез.
5. Референсний докінг.
6. Симуляція молекулярної динаміки.

Під час даного експерименту використовувались такі ресурси : Protein Data Bank (PDB), UniProt, NCBI, KIST Computational Cluster та OIST HPC Deigo. Та наступні програми/інструменти: Discovery Studio Visualization (DS Visualization) [41], AlphaFold2, ESMFold, PyMol [42], BLAST, GREMLIN, PyRosetta, Rosetta GALigandDock, CHARMM-GUI [43] та OpenMM.

3.1. Додаткова валідація правила «ОН-сайтів»

Зважаючи на попередньо сформульоване дослідниками припущення, що правила локалізації гідроксильних груп в сайті зв'язування не є видоспецифічними та поширюються на будь-які типи опсинів, було вирішено спершу додатково перевірити дане правило. Для цього було взято дані про

спектри поглинання опсинів деяких членистоногих, опубліковані в одному з нещодавніх досліджень присвяченому їх еволюції (рис 3.1).

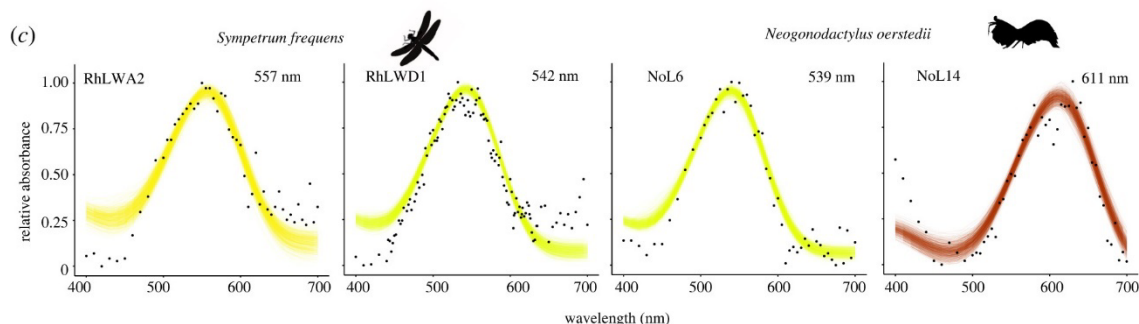


Рис. 3.1. Спектри поглинання в темряві були виміряні для специфічних опсинів довгохвильових одональних комах і стоматодних ракоподібних [2].

Було обрано варіант опсину довгої хвилі поглинання NoL14 *Neogonodactylus oerstedii* для аналізу та підтвердження/спростування правила. Структуру чутливого до зеленого опсину NoL14 LW Gq з стоматоподи *Neogonodactylus oerstedii* було змодельовано з використанням родопсину *Hasarius adansoni* (PDB 6i9k) як шаблону, оскільки не існує експериментально отриманої структури білка. Амінокислота послідовність опсину глибоководної креветки була взята з бази UniProt за кодом A0A6H0X1Q8, павука – B1B1U5. Для перевірки, чи можлива побудова структури білка креветки на основі структури павука-стрибунця, було проведено вирівнювання (рис. 3.2) та визначено відсоток ідентичності, який складає 47.97%. Даний відсоток співпадінь задовольняє умову для гомологічного моделювання.

Гомологічна модель була побудована в AlphaFold2 та візуалізована в PyMol (рис. 3.3).

tr B1B1U5 B1B1U5_9ARAC	MLPHAAKMAARVAGDHDGRNISIVDLLPEDMLPMIHEHWYKFPMPMETSMMHYILGMLIIVI	60
tr A0A6H0X1Q8 A0A6H0X1Q8_NE00E	-MSY-YDDSASAAAYTNPYGNHTVVDTPADMLHLIDSHWYGFPLNDMWMYGLAFFISVT	58
	: : . : * . * : * : * * : * * * : * . * * * * * : : * * : * * *	
tr B1B1U5 B1B1U5_9ARAC	GIISVSGNGVVMYLMMTVKNLRTPGNFLVNLALSDFGMLFFMPTMSINCFMFAETWVIGP	120
tr A0A6H0X1Q8 A0A6H0X1Q8_NE00E	GILAVGGNFVVIWVFMCTKALRSPSNYYVNLAISDFTLMACMCPVINSYRTWIFGP	118
	** : * . * * * * : * * * * * : * * * * * * * : : * * : * * : * * : * * *	
tr B1B1U5 B1B1U5_9ARAC	FMCELYGMIGSLFGSASIWVSLVMITLDRYVIVKGMAGKPLTKVALLRMLFVWISLGN	180
tr A0A6H0X1Q8 A0A6H0X1Q8_NE00E	TFCYVYAAIGSLTGCASIFMTCLISYDRYVIVKGIKKPLTTGKAMLMLLVWLTSAAW	178
	: * : * . * * * * * * * * : : * : * * * * * * : * * * * * * : * * *	
tr B1B1U5 B1B1U5_9ARAC	TIAPMYGWSRYVPEGSMTSCTIDYIDTAINPMSYLIAYAFVYFVPLFIIICYAFIVMQ	240
tr A0A6H0X1Q8 A0A6H0X1Q8_NE00E	TFAPFFGWSRYVPEGNMTACGDYLRGSILDQTYLWSYTTWCYFMTFVFIYVYCFIVAA	238
	* * * : *	
tr B1B1U5 B1B1U5_9ARAC	VAAHEKSLREQAKMNIKSLRSNEDNKKASAEFRLAKVAFMTICCFMAMTYPYLTLSFLG	300
tr A0A6H0X1Q8 A0A6H0X1Q8_NE00E	VRNHEKAMREQAKKMGVKSRLGDADAQKKSADCKLAKIALINVSLEWFMAMTYPYAIINIAG	298
	* * * * * : *	
tr B1B1U5 B1B1U5_9ARAC	IFSDRTWLTTPMTSVWGAIFAKASACYNPIVYGISHPKYRAALHDKFPCLKCGSDSPKGD	360
tr A0A6H0X1Q8 A0A6H0X1Q8_NE00E	LT-NKEIVTPLFSIWGSVFAKANTVYNPVYVYIASHPKYKALYQKLPWLQCAPDVPEDDS	357
	: : : * * : * * * : * * * * * * : * * * * * * * * * * * * * * * * * * *	
tr B1B1U5 B1B1U5_9ARAC	ASTVAESEKAGE---	372
tr A0A6H0X1Q8 A0A6H0X1Q8_NE00E	XSTASTTTTGGEKA	372
	** : : . * *	

Рис. 3.2. Результати вирівнювання послідовностей родопсину *Hasarius adansoni* (PDB 6i9k) та опсину довгої хвилі поглинання *Neogonodactylus oerstedii*.

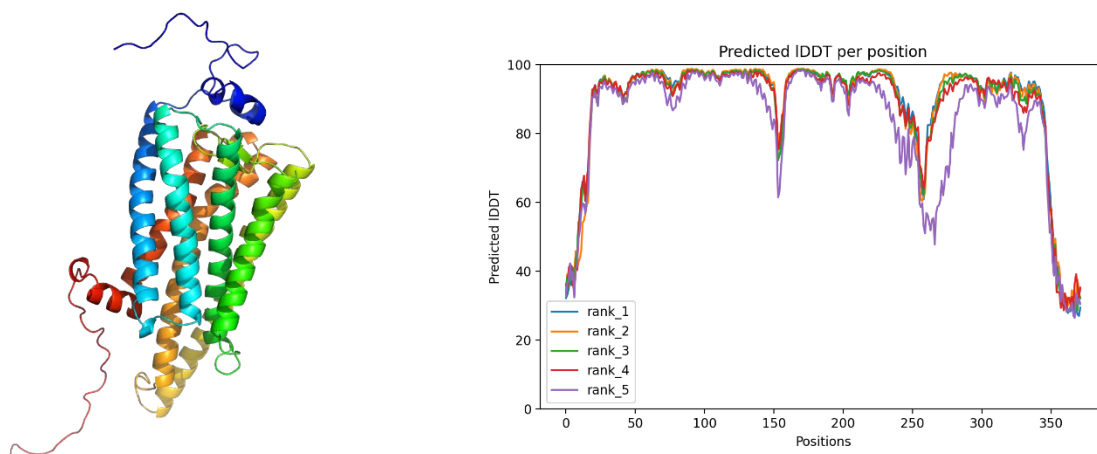


Рис. 3.3. Модель *N.oerstedii* (зліва) та тест різниці локальних відстаней LDDT (справа). LDDT — це метрика, яка використовується для оцінки точності та якості прогнозів структури білка, створених AlphaFold2. Він вимірює узгодженість між прогнозованою моделлю та експериментальною структурою на локальному рівні шляхом порівняння відстаней між

відповідними атомами в прогнозованій та експериментальній структурах. Оцінки LDDT варіюються від 0 до 100, причому вищі оцінки вказують на кращу узгодженість і вищу якість прогнозів.

Після структурного вирівнювання побудованої моделі *Neogonodactylus oerstedii* та експериментально отриманої структури бичачого родопсину (PDB 1U19) в PyMol та порівняльного аналізу ключових 10 амінокислот сайту зв'язування з ретиналем (рис. 3.4.), було отримано результати, представлені в Додатку 1. Цікавим є той факт, що 3 з 5 відмінних залишків у опсині довгої хвилі узгоджуються з попередньо сформульованим правилом «ОН-сайтів»: ILE189THR, MET207TYR, PHE208THR (позиції вказані відносно родопсину *Bos taurus*).

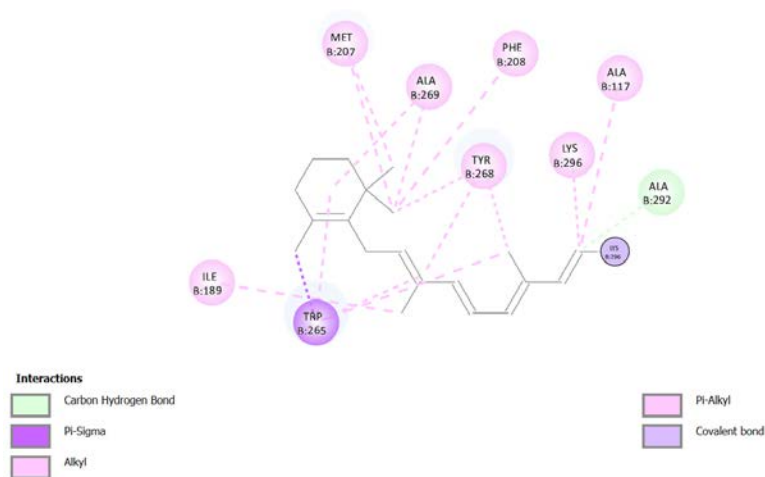


Рис. 3.4. 2D-діаграма зв'язків у сайті зв'язування з ретиналем в структурі бичачого родопсину. Отримано з використанням Discovery Studio Visualization.

Таким чином, дані спостереження додатково підтверджують, що наявність гідроксильних груп в певних позиціях сайту зв'язування також зустрічаються в природі серед тих видів, що мають зміщені піки поглинання в сторону довших довжин хвиль.

3.2. MSA та коеволюційний аналіз

MSA (Multiple Sequence Alignment) означає «множинне вирівнювання послідовностей». Це процес вирівнювання більш ніж двох послідовностей білків або нуклеотидів з метою виявлення подібностей, консервативних регіонів та структурних мотивів між ними. MSA широко використовується в біоінформатиці та геноміці для вивчення еволюційних відносин між різними послідовностями та виявлення консервативних регіонів, які можуть бути важливими для функцій білків або РНК.

Коеволюційний аналіз, відомий також як аналіз кореляційної еволюції, використовується для виявлення кореляційних зв'язків між мутаціями в парах амінокислот або нуклеотидів у білках або геномах. Ідея полягає в тому, що якщо дві амінокислоти або нуклеотиди часто змінюються разом під час еволюції, це може свідчити про їх взаємодію або функціональну залежність.

Коеволюційний аналіз зазвичай проводиться на основі MSA. Шляхи аналізу коеволюції можуть варіювати, включаючи методи, такі як статистика взаємоінформації (Mutual Information), моделі Маркова (Markov Models) та алгоритми машинного навчання. Ці методи дозволяють виявляти статистично значущі кореляції між мутаціями та ідентифікувати пари амінокислот або нуклеотидів, які можуть взаємодіяти або бути важливими для структури та функцій білків або РНК.

Внутрішньомолекулярний коеволюційний аналіз родопсину *Bos taurus* та LW опсину *Neogonodactylus oerstedii* був проведений з метою вирішення двох питань:

1. Чи існують такі амінокислотні залишки в області сайту зв'язування бичачого родопсину з ретиналем, які в процесі еволюції

коеволюціонували з іншими амінокислотними залишками та відіграють значну роль у підтриманні структурної цілісності родопсину?

2. Чи існують такі амінокислотні залишки в області сайту зв'язування опсину поглинання довгої хвилі глибоководної креветки, які також структурно коеволюціонували, і, відповідно, мають проєктуватися на структуру родопсину лише як мутація?

Для вирішення цієї задачі необхідно було отримати MSA файли відповідно для родопсину *Bos taurus* та опсину *Neogonodactylus oerstedii*. Розглянемо алгоритм на прикладі останнього.

Використовуючи інструмент BLAST, вводимо послідовність та отримуємо результат вирівнювання зі 100 найближчими сиквенсами (рис. 3.5).

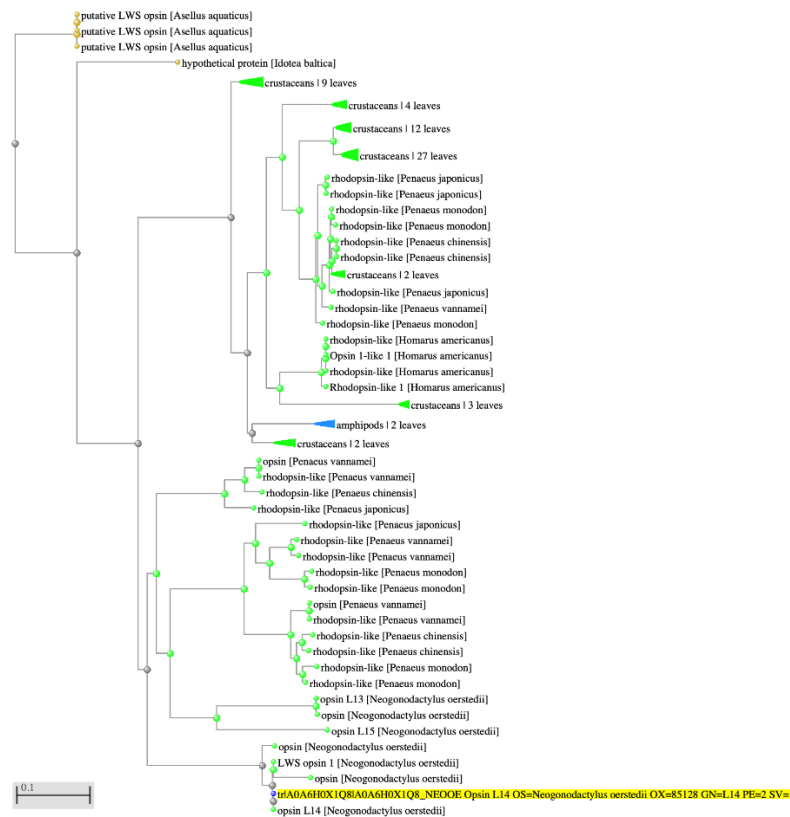


Рис. 3.5. Філогенетичне дерево, побудоване на основі отриманого MSA файлу. Жовтим виділено цільовий сиквенс.

Для подальшого коеволюційного аналізу використовувався онлайн-сервер, розроблений Baker's lab (тією ж лабораторією, що розробляла Rosetta). GREMLIN — це метод вивчення статистичної моделі сімейства білків, який фіксує як консервативні патерни, так і коеволюцію в заданому сімействі. Сила вимірної коеволюції доволі добре передбачає контакти між залишками в 3D-структурі білка.

Результати коеволюційного аналізу опсину *Neogonodactylus oerstedii* представлені та пояснені нижче.

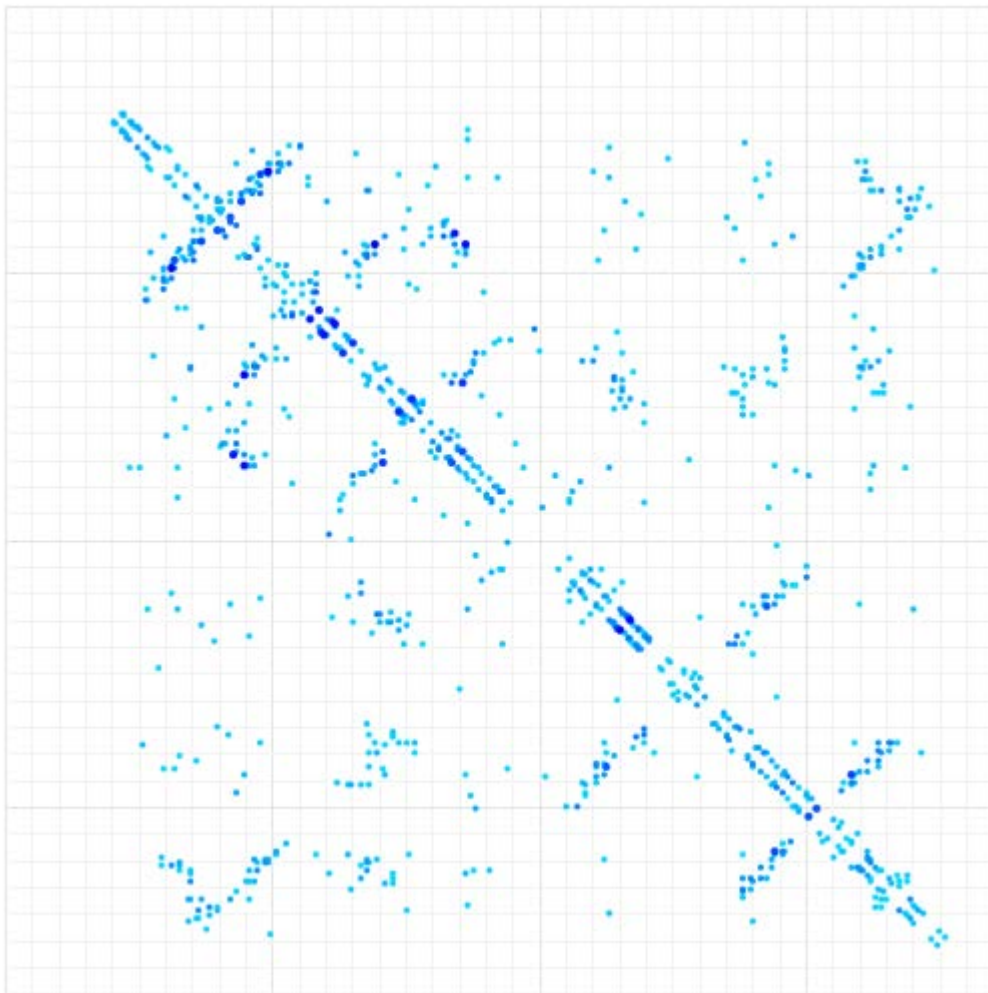


Рис. 3.6. Пари залишків, відсортовані за силою коеволюції. Чим темніші та більші сині крапки, тим вища сила коваріації.

Для аналізу власне амінокислотних пар було написано та використано код мовою Python:

```

coevol_analysis

import pandas as pd
from itertools import chain

def search(data,num):

    ll = []
    for j in num:
        rows = []
        for n in range(len(data)):
            if j == data.loc[n][0] or j == data.loc[n][1]:
                rows.append(data.loc[n])

        ll.append(rows)

    return ll

#----- Main -----

df = pd.read_csv("/Users/kistintern4/Documents/coevolution/opsin_or_rhodopsin.txt", sep="\t")
nums = [296]

lists = search(df,nums)

# ind = 0
# print("Number is : ",nums[ind])
# print(lists[ind])

flatten_list = list(chain.from_iterable(lists))
print(pd.DataFrame(flatten_list))

df1=pd.DataFrame(flatten_list)

df1.to_excel('/Users/kistintern4/Documents/coevolution/coevol_results.xls')
```

Рис. 3.7. Програма для екстракції необхідних даних зі згенерованих вихідних даних GREMLIN.

Код імпортує необхідні бібліотеки: `pandas` для обробки та аналізу даних і `itertools` для зведення вкладених списків.

Визначається функція пошуку, яка приймає два аргументи: дані та число. Вона шукає рядки серед даних, де або значення в першому стовпці, або значення в другому стовпці відповідає будь-якому значенню в списку `num`. Функція збирає всі відповідні рядки для кожного значення в `num` і повертає список списків (`ll`).

Код зчитує вхідний файл CSV у `DataFrame` під назвою `df` за допомогою `pd.read_csv`. Роздільник для файлу вказано як `"\t"` (розділений табуляцією). У даному варіанті список `nums` визначається одним елементом, яким є число 296 (позиція амінокислоти), проте, в залежності від задачі, кількість позицій та самі позиції можуть легко змінюватися. Наприклад, аби отримати повну таблицю коеволюції всіх залишків з сайту зв'язування, є можливість вказати їх всі одразу.

Функція пошуку викликається з `df` і `nums` як аргументами, а результат зберігається в змінній `lists`. Змінна `lists` зведена за допомогою `itertools.chain.from_iterable` для створення єдиного списку під назвою `flatten_list`. `Flatten_list` використовується для створення нового `DataFrame` під назвою `df1`.

`Df1 DataFrame` зберігається як файл Excel під назвою «`coevol_results.xls`» за допомогою `df1.to_excel`.

Було виявлено, що у LW опсині *Neogonodactylus oerstedii* амінокислоти 189T та 118S (позиції вказані з проєкцією на бичачий родопсин) –

коеволюціонують та мають структурну залежність. Отже, при введенні мутації I189T необхідно також вводити заміну T118S в структурі родопсину.

3.3. Мутагенез в PyRosetta

Rosetta — це набір обчислювальних інструментів і алгоритмів, розроблених, зокрема, дослідниками Вашингтонського та Каліфорнійського університетів. Він призначений для моделювання та прогнозування структури, стабільності та функції білків.

Білкова інженерія — це процес модифікації або проектування білків для покращення їхніх властивостей або створення нових функцій. Rosetta може допомогти в цьому процесі, надаючи різні обчислювальні методи для прогнозування структури білка, стикування білка з білком, дизайну білка та аналізу взаємодії білок-ліганд.

Rosetta використовує комбінацію моделей на основі фізики, статистичних потенціалів і алгоритмів машинного навчання для дослідження та прогнозування структур і взаємодій білків. Він може генерувати моделі білкових структур на основі експериментальних даних, таких як рентгенівська кристалографія або спектроскопія ядерного магнітного резонансу (ЯМР). Крім того, Rosetta може оптимізувати білкові послідовності для покращення стабільності, афінності зв'язування або інших бажаних властивостей.

PyRosetta - це Python-інтерфейс до програмного забезпечення Rosetta для білкової інженерії. Він надає можливість взаємодіяти з функціональністю Rosetta за допомогою Python, що спрощує програмування і автоматизацію багатьох завдань у галузі мутагенезу білків.

Використання PyRosetta має кілька переваг для мутагенезу білків:

1. **Простота програмування:** Python є популярною та простою у використанні мовою програмування, що полегшує розробку скриптів для мутагенезу білків з використанням PyRosetta.
2. **Багатий набір функцій:** PyRosetta надає доступ до широкого спектра функціональності Rosetta, включаючи методи для створення, оптимізації та аналізу білкових структур.
3. **Потужність Rosetta:** PyRosetta базується на фундаментальних алгоритмах та моделях Rosetta, які дозволяють проводити якісну оцінку впливу мутацій на структуру та стабільність білків. Rosetta використовує енергетичні функції та статистичні потенціали для оцінки енергії та властивостей білків, що дозволяє передбачати ефект мутацій та вибирати оптимальні варіанти для певних дизайнерських цілей.

Варто взяти до уваги, що Rosetta та PyRosetta найефективніше працює на базі ОС Linux. Весь процес взаємодії з програмою та запуски скриптів відбувались на ssh сервері Lab of Computational Drug Discovery (Park's Lab) Korea Institute of Science and Technology (KIST).

Для введення точкових мутацій та релаксації структури було сформульовано такий алгоритм дій:

1. Створити робочу директорію для мутантного білка.
2. Підготувати інпути – файли параметрів для Rosetta та структуру холорецептора бичачого родопсину (разом зі зв'язаним лігандом).

А) Підготовка параметрів ліганду:

- протонування (починаючи з .mol2 файлу)

```
> /path/to/obabel input.mol2 -o mol2 -O input.h.mol2 -p 7
```

- протонування (починаючи з .pdb файлу)

```
> /path/to/obabel input.pdb -o mol2 -O input.h.mol2 -p 7
```

- додати частковий заряд:

```
> /path/to/obabel -i mol2 input.h.mol2 -o mol2 -O input.qh.mol2 --
partialcharge MMFF94
```

- конвертувати в формат Rosetta:

```
>/path-
```

```
to/Rosetta/source/scripts/python/public/generic_potential/mol2genparams.py
-s input.qh.mol2
```

```
> mv input.qh.params LG.params
```

Б) Підготувати apo.pdb (структуру рецептора без ліганду) в PyMol

В) Створити holo.pdb, де ліганд розміщується в приблизно навколо сайту зв'язування

```
> python place_ligand.py apo.pdb input.qh_0001.pdb [res1] [res2] [res3] ...
```

Де [res1] [res2] [res3] ... - номери залишків у apo.pdb, які мають контактувати з лігандом; переміститься до центру мас атомів [res1-n].

3. Створити .txt з позиціями мутацій:

```
> cat mutation.txt
```

M207Y

F208T

I189T

T118S

Ctrl+D

4. Запуск PyRosetta скрипту:

```
> source /path/to/pyrosetta/setup.sh
```

> python3.8 mutate_and_relax.py holo.pdb mutation.txt

В Додатку 2 представлено код програми, яка вводить точкові мутації в структуру білка та одразу проводить її релаксацію.

Загалом, було підготовано 5 мутантних варіантів родопсину *Bos taurus*:

Робоча назва	Мутовані позиції	Обґрунтування вибору	Додаткові характеристики
Mut1	F261Y/A269T	Експериментально підтверджені зсуви в чевроний спектр [44]	Найслабше зв'язування ретиналю, найбільший зсув (20 нм)
Mut2	A164S/A269T		Найсильніше зв'язування ретиналю, середній зсув (14 нм)
Mut3	A164S/F261Y		Середня сила зв'язування ретиналю, найменший зсув (12 нм)
Mut4	A164S/F261Y/A269T		Відсутність зв'язування з ретиналем (було взято для перевірки експерименту шляхом докінгу)
MutOH	M207Y/F208T/I189T/T118S	Варіант мутацій, підібраний на основі правила «ОН-сайтів», еволюційного та коеволюційного аналізів	Унікальний варіант мутації, який ще не був випробуваний в жодному експерименті

Таблиця 3.1. Варіанти мутацій, обрані для подальшого докінгу та перевірки біоінформатичним методами.

Окрім методу введення точкових мутацій в Rosetta, додатково було згенеровані гомологічні моделі мutowаних білків, побудовані внаслідок заміни амінокислот у самому сиквенсі з використанням AlphaFold2 та ESMFold.

3.4. Референсний докінг в Rosetta GALigandDock

Референсний докінг (англ. reference docking) - це метод в комп'ютерному моделюванні молекулярного докінгу, який використовується для передбачення взаємодії між біомолекулою (наприклад, білком) і малим молекулярним з'єднанням (наприклад, лікарським засобом). У референсному докінгу відбувається пошук оптимальної орієнтації та місцеположення малих молекул в активному центрі білка, використовуючи вже відомі структури комплексів білок-ліганд.

GALigandDock є одним з популярних методів референсного докінгу. Він має кілька переваг:

1. Використання генетичного алгоритму для генерації та оптимізації нових молекулярних з'єднань. Це дозволяє ефективно здійснювати пошук у просторі конформацій молекул, що може привести до знаходження більш точних інтеракційних конформацій.
2. Врахування гнучкості білка та ліганду під час докінгу. Це особливо важливо, оскільки багато білків можуть зазнавати конформаційних змін під час взаємодії з лігандом. Врахування гнучкості дозволяє отримати більш точні результати.

3. Можливість використання експериментальних даних, такі як інформація про взаємодію амінокислотних залишків або енергетичні дані, для покращення точності докінгу.

Цей метод було використано для докінгу мутованих варіантів білків родопсину, оскільки у нас вже були наявні експериментальні дані (структура) зв'язування ретиналю з білком дикого типу. Таким чином, можна було одразу скоригувати докінг та отримати точніші результати.

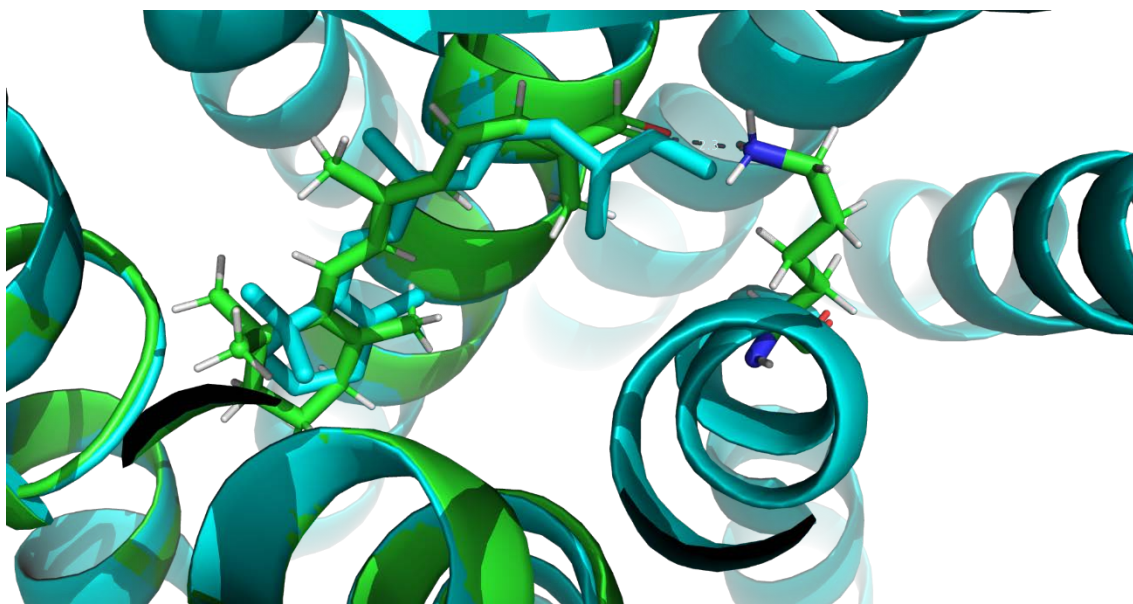


Рис. 3.8. Результати референсного докінгу молекули ретиналю з Mut4. Структура мутованого родопсину та 11-цис-ретиналю (зелений колір) була накладена на структуру родопсину дикого типу (бірюзовий колір) для наочності. Відстань очікуваного ковалентного зв'язку ліганду з Lys-296 показана пунктирною лінією та становить 2.3 Å.

Отримані результати є хибно позитивними, оскільки було експериментально підтверджено, що в триплетному мутанті не відбувається зв'язування з хромофором. Тому було вирішено також провести аналогічний референсний докінг з моделями, отриманими в AlphaFold2 та ESMFold.

Додаткова перевірка з цими моделями також дала хибно позитивний результат (рис. 3.9). Також з'явився хибно негативний результат у випадку докінгу моделі Mut1, побудованої в ESMFold (рис. 3.10). Незважаючи на те, що в інших випадках докінгу не спостерігалось таких абнормальностей, залишалось відкритим питання, чи це методи молекулярного моделювання та докінгу є недостатньо специфічними для аналізу поставленої задачі, й вони не можуть виявити дуже тонкі структурні зміни, які в реальності відбуваються при введенні певних мутацій, чи це порушення стосується насправді надмолекулярного рівня (порушення інкорпорації в мембрану тощо).

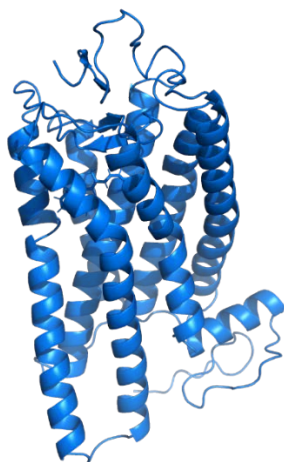


Рис. 3.9. Результати референсного докінгу з моделлю, побудованою в **AlphaFold2**. Молекула ретиналю знаходиться у сайті зв'язування.

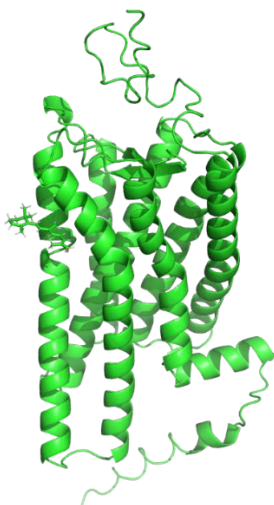


Рис. 3.10. Результати референсного докінгу Mut1 з моделлю, побудованою в ESMFold. Молекула ретиналю не зайшла в кишеню сайту зв'язування.

3.5. Симуляція молекулярної динаміки та негативний контроль

Щоб спробувати побачити тонкі молекулярні зміни, що можуть відбутися внаслідок введення даних мутацій та зрозуміти, як себе поведуть білки у динаміці, було вирішено застосувати метод симуляції молекулярної динаміки. Молекулярна динаміка (англ. molecular dynamics) - це комп'ютерний метод для моделювання руху атомів та молекул у системі впродовж певного часу. Вона використовує математичні рівняння та чисельні методи для відтворення руху частинок у системі, враховуючи їх взаємодію та зовнішні умови.

OpenMM є однією з бібліотек, яка може бути використана для моделювання молекулярної динаміки. OpenMM є відкритим програмним забезпеченням, розробленим для проведення симуляцій молекулярної динаміки на рівні атомів та частинок. OpenMM надає потужні інструменти для розрахунку енергій взаємодії між атомами, обчислення сил та переміщень частинок. Його ми використовували для проведення симуляцій.

Оскільки родопсин є трансмембранним білком, потрібно особливими чином підготувати файли-інпути для симуцій МД. **CHARMM-GUI** (Chemistry at HARvard Molecular Mechanics Graphical User Interface) - це онлайн-інструмент, розроблений для підготовки систем мембрани та мембранных білків для моделювання МД з використанням пакету CHARMM (Chemistry at HARvard Macromolecular Mechanics). CHARMM-GUI надає користувачам зручний інтерфейс, що дозволяє автоматизувати багато кроків підготовки системи перед МД симуляціями.

CHARMM-GUI має спеціалізовані модулі для підготовки мембрани та мембранных білків, які ми використовували в даному порядку:

1. Вибір типу мембрани.
2. Встановлення мембранної системи.
3. Вставка білка в мембрану.
4. Додання іонів, видалення води, налаштування початкових швидкостей частинок.

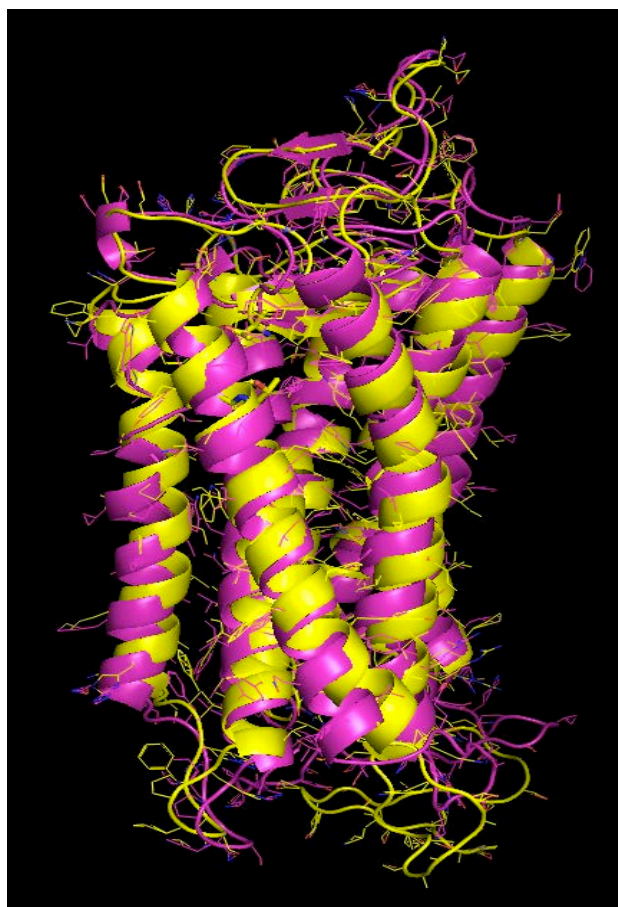




Рис. 3.11. Результати МД Mut4 (зліва) та MutOH (справа). Спостерігається зміщення спіралі V у Mut4 (бірюзовий колір) відносно дикого типу (фіолетовий колір). Водночас, MutOH, в якому наявні 4 точкові мутації, обрані на основі еволюційного аналізу опсинів, аналізу коеволюції LW опсину *Neogonodactylus oerstedii* та врахування правила «ОН-сайтів», не відображає таких значних структурних зміщень спіралей.

Аналізуючи результати симуляцій МД вказаних мутантних варіантів бичачого родопсину, та враховуючи той факт, що експериментально підтверджено, що в Mut4 триплеті не відбувається зв'язування з ретиналем, виникає припущення, що можливо саме дане зміщення V спіралі і є причиною, через яку ліганд не може зайти в кишеню сайту зв'язування. Водночас, відсутність значних динамічних змін MutOH, можливо, вказує на те, що запропонований мутант є більш стабільним за структурою.

Даних спостережень недостатньо, щоб зробити такі висновки. Тому, подальші кроки нашого дослідження такі:

1. Перевірка інших відомих точкових мутацій родопсинів, які спричиняють структурні порушення та відповідно, порушують зв'язування з ретиналем. Дані мутації [45] стануть своєрідним

негативним контролем для МД, щоб підтвердити надійність даного методу у детекції незначних структурних порушень, що призводять до незв'язування з хромофором.

2. Проведення МД в GROMACS, аналіз RMSD, RMSF, Н-зв'язків, FES, FEL та інших характеристик, що дозволять нам оцінити поведінку MutOH, а саме, його стабільність, вільну енергію, конформаційні стани тощо. Дані дослідження будуть проводитись на базі Okinawa Institute of Science and Technology (OIST) на потужностях HPC кластеру Deigo.

Висновки

В результаті проведеної роботи можна зробити декілька важливих висновків:

1. Здатність контролювати активацію родопсину за допомогою різних довжин хвиль світла розширює можливості оптогенетичних методів та допомагає вивчати нейронні ланцюги і розуміти роботу мозку.
2. Правило «п'яти сайтів» та правило «ОН-сайтів» пояснюють спостережуване зміщення спектральних характеристик зорових пігментів та вплив гідроксильних груп на стабілізацію збудженого або основного стану в зоровій системі. Аналіз опсину NoL14 у *Neogonodactylus oerstedii* підтверджує правило «ОН-сайтів» і підкреслює зв'язок гідроксильних груп зі спектральними характеристиками поглинання світла.

3. Внутрішньомолекулярний коеволюційний аналіз родопсину *Bos taurus* і довгохвильового опсину *Neogonodactylus oerstedii* виявив структурну залежність між певними амінокислотами, що підтверджує необхідність взаємозалежних мутацій у родопсині.
4. Отримано п'ять мутантних варіантів родопсину *Bos taurus*, включаючи унікальний варіант MutOH, який був обраний на основі правил «ОН-сайтів», еволюційного та коеволюційного аналізів.
5. Симуляція молекулярної динаміки (МД) показало зміщення спіралі V в Mut4, що може перешкоджати зв'язуванню ліганду, тоді як MutOH демонструє більшу структурну стабільність.

Джерела

- [1] H. Kim, G. Park, J. Kim, and J. Seong, “Development of OptoDRD2 to precisely control the function of DRD2 by light,” *IBRO Rep*, vol. 6, p. S522, Sep. 2019, doi: 10.1016/J.IBROR.2019.07.1626.
- [2] M. A. Liénard, W. A. Valencia-Montoya, and N. E. Pierce, “Molecular advances to study the function, evolution and spectral tuning of arthropod visual opsins,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 377, no. 1862. Royal Society Publishing, Oct. 24, 2022. doi: 10.1098/rstb.2021.0279.
- [3] N. Varma *et al.*, “Crystal structure of jumping spider rhodopsin-1 as a light sensitive GPCR,” *Proceedings of the National Academy of Sciences*, vol. 116, no. 29, pp. 14547–14556, Jul. 2019, doi: 10.1073/pnas.1902192116.
- [4] A. Terakita, E. Kawano-Yamashita, and M. Koyanagi, “Evolution and diversity of opsins,” *Wiley Interdiscip Rev Membr Transp Signal*, vol. 1, no. 1, pp. 104–111, Jan. 2012, doi: 10.1002/wmts.6.
- [5] N. Y. Leung and C. Montell, “Unconventional Roles of Opsins,” *Annu Rev Cell Dev Biol*, vol. 33, no. 1, pp. 241–264, Oct. 2017, doi: 10.1146/annurev-cellbio-100616-060432.
- [6] H. Tsukamoto and A. Terakita, “Diversity and functional properties of bistable pigments,” *Photochemical and Photobiological Sciences*, vol. 9, no. 11, pp. 1435–1443, 2010, doi: 10.1039/c0pp00168f.

- [7] M. Ramirez *et al.*, “The last common ancestor of most bilaterian animals possessed at least 9 opsins,” *Genome Biol Evol*, p. evw248, Oct. 2016, doi: 10.1093/gbe/evw248.
- [8] T. Nagata, M. Koyanagi, R. Lucas, and A. Terakita, “An all-trans-retinal-binding opsin peropsin as a potential dark-active and light-inactivated G protein-coupled receptor,” *Sci Rep*, vol. 8, no. 1, Dec. 2018, doi: 10.1038/s41598-018-21946-1.
- [9] E. Kawano-Yamashita, M. Koyanagi, S. Wada, H. Tsukamoto, T. Nagata, and A. Terakita, “Activation of Transducin by Bistable Pigment Parapinopsin in the Pineal Organ of Lower Vertebrates,” *PLoS One*, vol. 10, no. 10, p. e0141280, Oct. 2015, doi: 10.1371/journal.pone.0141280.
- [10] M. Koyanagi and A. Terakita, “Diversity of animal opsin-based pigments and their optogenetic potential,” *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, vol. 1837, no. 5, pp. 710–716, May 2014, doi: 10.1016/j.bbabi.2013.09.003.
- [11] G. H. Jacobs, “Evolution of colour vision in mammals,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1531, pp. 2957–2967, Oct. 2009, doi: 10.1098/rstb.2009.0039.
- [12] F. E. Hauser and B. S. Chang, “Insights into visual pigment adaptation and diversity from model ecological and evolutionary systems,” *Curr Opin Genet Dev*, vol. 47, pp. 110–120, Dec. 2017, doi: 10.1016/j.gde.2017.09.005.
- [13] C. J. van der Kooi, D. G. Stavenga, K. Arikawa, G. Belušič, and A. Kelber, “Evolution of Insect Color Vision: From Spectral Sensitivity to Visual Ecology,” *Annu Rev Entomol*, vol. 66, no. 1, pp. 435–461, Jan. 2021, doi: 10.1146/annurev-ento-061720-071644.

- [14] D. Kojima and Y. Fukada, “Spectroscopic Analysis of Wavelength Sensitivities of Opsin-Type Photoreceptor Proteins,” in *Circadian Clocks*, 2022, pp. 169–185. doi: 10.1007/978-1-0716-2577-4_8.
- [15] A. Terakita and T. Nagata, “Functional Properties of Opsins and their Contribution to Light-Sensing Physiology,” *Zoolog Sci*, vol. 31, no. 10, pp. 653–659, Oct. 2014, doi: 10.2108/zs140094.
- [16] E. Salcedo, D. M. Farrell, L. Zheng, M. Phistry, E. E. Bagg, and S. G. Britt, “The Green-absorbing *Drosophila* Rh6 Visual Pigment Contains a Blue-shifting Amino Acid Substitution That Is Conserved in Vertebrates,” *Journal of Biological Chemistry*, vol. 284, no. 9, pp. 5717–5722, Feb. 2009, doi: 10.1074/jbc.M807368200.
- [17] T. Saito, M. Koyanagi, T. Sugihara, T. Nagata, K. Arikawa, and A. Terakita, “Spectral tuning mediated by helix III in butterfly long wavelength-sensitive visual opsins revealed by heterologous action spectroscopy,” *Zoological Lett*, vol. 5, no. 1, p. 35, Dec. 2019, doi: 10.1186/s40851-019-0150-2.
- [18] F. E. Hauser, I. van Hazel, and B. S. W. Chang, “Spectral tuning in vertebrate short wavelength-sensitive 1 (SWS1) visual pigments: Can wavelength sensitivity be inferred from sequence data?,” *J Exp Zool B Mol Dev Evol*, vol. 322, no. 7, pp. 529–539, Nov. 2014, doi: 10.1002/jez.b.22576.
- [19] T. D. Lamb, E. N. Pugh, and S. P. Collin, “The Origin of the Vertebrate Eye,” *Evolution: Education and Outreach*, vol. 1, no. 4, pp. 415–426, Oct. 2008, doi: 10.1007/s12052-008-0091-2.
- [20] A. K. Mishra, C. Fritsch, R. Voutev, R. S. Mann, and S. G. Sprecher, “Homothorax controls a binary Rhodopsin switch in *Drosophila* ocelli,” *PLoS Genet*, vol. 17, no. 7, Jul. 2021, doi: 10.1371/journal.pgen.1009460.

- [21] K. P. Hofmann and T. D. Lamb, “Rhodopsin, light-sensor of vision,” *Prog Retin Eye Res*, vol. 93, p. 101116, Mar. 2023, doi: 10.1016/j.preteyeres.2022.101116.
- [22] P. W. Hildebrand *et al.*, “A Ligand Channel through the G Protein Coupled Receptor Opsin,” *PLoS One*, vol. 4, no. 2, p. e4382, Feb. 2009, doi: 10.1371/journal.pone.0004382.
- [23] T. Okada, O. P. Ernst, K. Palczewski, and K. P. Hofmann, “Activation of rhodopsin: new insights from structural and biochemical studies,” *Trends Biochem Sci*, vol. 26, no. 5, pp. 318–324, May 2001, doi: 10.1016/S0968-0004(01)01799-6.
- [24] M. Ben-Nun, F. Molnar, H. Lu, J. C. Phillips, T. J. Martínez, and K. Schulten, “Quantum dynamics of the femtosecond photoisomerization of retinal in bacteriorhodopsin,” *Faraday Discuss*, vol. 110, pp. 447–462, 1998, doi: 10.1039/a801310a.
- [25] R. R. Birge, “Nature of the primary photochemical events in rhodopsin and bacteriorhodopsin,” *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, vol. 1016, no. 3, pp. 293–327, Apr. 1990, doi: 10.1016/0005-2728(90)90163-X.
- [26] H. Nakamichi and T. Okada, “Local peptide movement in the photoreaction intermediate of rhodopsin,” *Proceedings of the National Academy of Sciences*, vol. 103, no. 34, pp. 12729–12734, Aug. 2006, doi: 10.1073/pnas.0601765103.
- [27] E. Ritter, K. Zimmermann, M. Heck, K. P. Hofmann, and F. J. Bartl, “Transition of Rhodopsin into the Active Metarhodopsin II State Opens a New Light-induced Pathway Linked to Schiff Base Isomerization,” *Journal of Biological Chemistry*, vol. 279, no. 46, pp. 48102–48111, Nov. 2004, doi: 10.1074/jbc.M406857200.

- [28] R. G. Matthews, R. Hubbard, P. K. Brown, and G. Wald, "Tautomeric Forms of Metarhodopsin," *Journal of General Physiology*, vol. 47, no. 2, pp. 215–240, Nov. 1963, doi: 10.1085/jgp.47.2.215.
- [29] F. Siebert, "Fourier-transform infrared spectroscopy applied to rhodopsin Non-invasive glucose measurement View project," 2005. [Online]. Available: <https://www.researchgate.net/publication/229632360>
- [30] M. Heck *et al.*, "Signaling States of Rhodopsin," *Journal of Biological Chemistry*, vol. 278, no. 5, pp. 3162–3169, Jan. 2003, doi: 10.1074/jbc.M209675200.
- [31] E. Crocker *et al.*, "Location of Trp265 in Metarhodopsin II: Implications for the Activation Mechanism of the Visual Receptor Rhodopsin," *J Mol Biol*, vol. 357, no. 1, pp. 163–172, Mar. 2006, doi: 10.1016/j.jmb.2005.12.046.
- [32] W. J. Degrip and K. J. Rothschild, "Chapter 1 Structure and mechanism of vertebrate visual pigments," 2000, pp. 1–54. doi: 10.1016/S1383-8121(00)80004-4.
- [33] M. Heck *et al.*, "Signaling States of Rhodopsin," *Journal of Biological Chemistry*, vol. 278, no. 5, pp. 3162–3169, Jan. 2003, doi: 10.1074/jbc.M209675200.
- [34] F. J. Bartl and R. Vogel, "Structural and functional properties of metarhodopsin III: Recent spectroscopic studies on deactivation pathways of rhodopsin," *Physical Chemistry Chemical Physics*, vol. 9, no. 14, p. 1648, 2007, doi: 10.1039/b616365c.
- [35] M. Nei, J. Zhang, and S. Yokoyama, "Color vision of ancestral organisms of higher primates," *Mol Biol Evol*, vol. 14, no. 6, pp. 611–618, Jun. 1997, doi: 10.1093/oxfordjournals.molbev.a025800.

- [36] S. Sekharan, K. Katayama, H. Kandori, and K. Morokuma, “Color Vision: ‘OH-Site’ Rule for Seeing Red and Green,” *J Am Chem Soc*, vol. 134, no. 25, pp. 10706–10712, Jun. 2012, doi: 10.1021/ja304820p.
- [37] S. Yokoyama and F. B. Radlwimmer, “The ‘five-sites’ rule and the evolution of red and green color vision in mammals,” *Mol Biol Evol*, vol. 15, no. 5, pp. 560–567, May 1998, doi: 10.1093/oxfordjournals.molbev.a025956.
- [38] A. B. Asenjo, J. Rim, and D. D. Oprian, “Molecular determinants of human red/green color discrimination,” *Neuron*, vol. 12, no. 5, pp. 1131–1138, May 1994, doi: 10.1016/0896-6273(94)90320-4.
- [39] R. D. Airan, K. R. Thompson, L. E. Fenno, H. Bernstein, and K. Deisseroth, “Temporally precise in vivo control of intracellular signalling,” *Nature*, vol. 458, no. 7241, pp. 1025–1029, Apr. 2009, doi: 10.1038/nature07926.
- [40] L. A. Gunaydin *et al.*, “Natural Neural Projection Dynamics Underlying Social Behavior,” *Cell*, vol. 157, no. 7, pp. 1535–1551, Jun. 2014, doi: 10.1016/j.cell.2014.05.017.
- [41] D. S. Biovia, “Discovery Studio Visualizer.” San Diego, 2019.
- [42] L. , & D. W. Schrödinger, “PyMol.” 2020.
- [43] S. Jo, T. Kim, V. G. Iyer, and W. Im, “CHARMM-GUI: A web-based graphical user interface for CHARMM,” *J Comput Chem*, vol. 29, no. 11, pp. 1859–1865, Aug. 2008, doi: 10.1002/jcc.20945.
- [44] T. Chan, M. Lee, and T. P. Sakmar, “Introduction of hydroxyl-bearing amino acids causes bathochromic spectral shifts in rhodopsin. Amino acid substitutions responsible for red-green color pigment spectral tuning,” *Journal of Biological*

Chemistry, vol. 267, no. 14, pp. 9478–9480, 1992, doi: 10.1016/s0021-9258(19)50115-6.

- [45] E. P. Rakoczy, C. Kiel, R. McKeone, F. Stricher, and L. Serrano, “Analysis of Disease-Linked Rhodopsin Mutations Based on Structure, Function, and Protein Stability Calculations,” *J Mol Biol*, vol. 405, no. 2, pp. 584–606, Jan. 2011, doi: 10.1016/j.jmb.2010.11.003.

Додаток 1

Залишки в сайті зв'язування зв'язування бичачого родопсину	<i>N. oerstedii</i> залишки в сайті зв'язування	Коментар
ALA 117	GLY 128	В літературі не знайдено пояснень.
ILE 189	THR 200	Це може призвести до меншого викривлення ретиналю.
MET 207	TYR 216	У відповідності до правил «ОН-сайтів»
PHE 208	THR 217	У відповідності до правил «ОН-сайтів»
TRP 265	TRP 288	-
TYR 268	TYR 291	-
ALA 269	ALA 292	-
ALA 292	SER 314	Вважається, що заміна SER/ALA зміщує поглинання в бік коротшої довжини хвилі
ALA 295	ALA 317	-
LYS 296	LYS 318	-

Додаток 2

```
import sys

import numpy as np

from pyrosetta import *

RELAXSCRIPT = '/home/hpark/scripts/cart1.script'

def aa1toaa3(aa1):

    mapper = {'A': 'ALA', 'C': 'CYS', 'D': 'ASP', 'E': 'GLU', 'F': 'PHE', 'G': 'GLY', 'H': 'HIS',
              'I': 'ILE', 'K': 'LYS', 'L': 'LEU', 'M': 'MET', 'N': 'ASN', 'P': 'PRO', 'Q': 'GLN',
              'R': 'ARG', 'S': 'SER', 'T': 'THR', 'V': 'VAL', 'W': 'TRP', 'Y': 'TYR'}

    return mapper[aa1]

def read_mutation(f):

    # syntax: A 100 G

    mutations = {}

    for l in open(f):

        words = l[:-1].split()

        res = int(words[1])

        tomut = aa1toaa3(words[2])

        mutations[res] = tomut

    return mutations

def find_residues_neightto_mut(pose, mutlist):

    # store CBxyz

    xyz_CB = []
```

```

for res in range(1, pose.size()): # skip ligand
    if pose.residue(res).has('CB'):
        xyz_CB.append(pose.residue(res).xyz('CB'))
    elif pose.residue(res).has('CA'):
        xyz_CB.append(pose.residue(res).xyz('CA'))

xyz_CB = np.array(xyz_CB)
mutlist = np.array(mutlist)
mutlist -= 1
xyz_mut = xyz_CB[mutlist]

D = np.array([[np.linalg.norm(xyz1 - xyz2) for xyz1 in xyz_mut] for xyz2 in xyz_CB])
neighs = np.where(D < 8.0)[0] + 1
return neighs

def main(pdb, mutationf, allow_backbone=False):
    params = 'LG.params' # hard-coded
    # setup pyrosetta

    init(f'-beta_cart -crystal_refine -extra_res_fa {params} -no_autogen_cart_improper -
constrain_relax_to_start_coords')

    pose = pose_from_file(pdb)

    mutations = read_mutation(mutationf)

    neighres = find_residues_neighto_mut(pose, list(mutations.keys()))

    print(neighres)

    # find neighs to mutating res

```

```
sfxn = create_score_function("beta_cart.wts")

# mutate

for res in mutations:

    mutator = protocols.simple_moves.MutateResidue(res, mutations[res])

    mutator.apply(pose)

pose.dump_pdb("mut1.pdb")

# set relax

mm = MoveMap()

mm.set_bb(False)

mm.set_jump(True)

for res in neighres:

    mm.set_chi(res, True)

ligres = pose.size()

mm.set_chi(ligres, True)

relax = protocols.relax.FastRelax(sfxn, RELAXSCRIPT)

relax.set_movemap(mm)

# run

relax.apply(pose)

pose.dump_pdb('mutrlx1.pdb')

if __name__ == "__main__":

    pdbref = sys.argv[1]

    mutationf = sys.argv[2]

    main(pdbref, mutationf)
```