

# Імітаційне моделювання синтезу білків

С.В. Беседіна  
кафедра інформаційних технологій  
Черкаський національний університет  
імені Богдана Хмельницького  
Черкаси, Україна  
bes\_sv@ukr.net

А.Ю. Кришталь  
кафедра інформаційних технологій  
Черкаський національний університет  
імені Богдана Хмельницького  
Черкаси, Україна  
kristal7@i.ua

## Protein synthesis simulation modelling

S. Byesyedina  
Department of Information Technology  
Cherkassy National University  
By Bohdan Khmelnytsky  
Cherkassy, Ukraine  
bes\_sv@ukr.net

A. Kryshstal  
Department of Information Technology  
Cherkassy National University  
By Bohdan Khmelnytsky  
Cherkassy, Ukraine  
kristal7@i.ua

**Анотація**—Обґрунтовано застосування методу QCP при емуляції синтезу білків, що дозволяє значно розширити знання просторової організації структури молекул білка. Запропоновано використання алгоритму QCP для модифікації наявних структур та отримання нових білків із новими властивостями.

**Abstract**—QCP method used to emulating the synthesis of proteins. It is increasing knowledge base of protein structure. QCP method can be used for modification exist and generation new structures.

**Ключові слова** — імітаційне моделювання; синтез білків; алгоритм QCP; протеїн.

**Keywords** — simulation; protein synthesis; QCP algorithm; protein.

### I. Вступ

Значний розвиток комп'ютерів та суперкомп'ютерів, які сьогодні стали самостійним інструментом пізнання в сучасній молекулярній біології та у результаті розшифрування структури геномів вірусів, мікроорганізмів та геному людини з'явилися величезні обсяги інформації щодо послідовностей нуклеїнових кислот та білків. На основі накопиченої інформації процес моделювання потребує наявності можливості бачити як моделюючи так і експериментальну структуру білка [1].

При розробці нових лікарських препаратів та оцінці впливу довкілля на структуру білків (протеїнів) виникає проблема імітаційного моделювання синтезу білків за умови зовнішніх збурень. На сьогодні, ефективність дослідження цих процесів визначається кількістю ресурсів, які на них витрачаються, та кількістю корисної інформації, яку здобули в результаті дослідження. У випадку використання вартісних програмних продуктів

(Rosetta@Home, Protein Data Bank, WebProAnalyst і т.ін.) ефективність знижується із зростанням їх вартості.

Тому постає задача експериментального дослідження результатів роботи існуючих методів для визначення допустимого зниження вартості моделювання синтезу білків.

### II. ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕНЬ

Метою цієї роботи є розробка системи для імітаційного моделювання синтезу білків із застосуванням методу Quaternion Characteristic Polynomial (QCP), для побудови трьохмірної структури білка, що дозволить відобразити структуру в процесі поєднання різних білків у нормальних умовах, з метою максимального наближення їх до тієї, яка відтворюється у живій клітині.

Поставлене завдання вирішується на основі застосування існуючих методів та розробки нового алгоритму моделювання синтезу білків, що дозволяє значно розширити їх можливості. У практиці до цього часу широке застосування методів [4-6] обмежувалось алгоритмічною складністю та високими коштами побудови систем для їх моделювання.

### III. ОСНОВИ МОДЕЛЮВАННЯ СИНТЕЗУ БІЛКІВ

Емпірично встановлено, що якщо послідовності двох білків ідентичні один одному більш ніж на 30%, то вони майже напевно є «родичами» і ступінь еволюційної дивергенції ще не настільки велика, щоб їх структури втратили спільність. Ці спостереження і є основою методики передбачення просторової структури, названою моделюванням на підставі гомології [3]. Цей процес включає кілька кроків, головними з яких є пошук структурного шаблону і побудова амінокислотного

рівняння. Вирішальним фактором, що визначає якість одержуваних моделей, є ступінь гомології послідовностей модельованого білка і шаблону.

Процеси фолдинга і міжмолекулярних взаємодій білків можуть бути змодельовані з використанням молекулярної механіки зокрема, молекулярної динаміки і методу Монте-Карло, які все частіше використовують переваги паралельних і розподілених обчислень або молекулярного моделювання [1].

В основу моделювання покладено алгоритм QCP, який є найбільш швидким з відомих алгоритмів для розв'язання задачі мінімізації середньоквадратичного відхилення (СКВ) між двома множинами точок. Алгоритм визначає матрицю повороту 3x3, на яку треба повернути другу множину точок, щоб мінімізувати СКВ [2]. Сьогодні, найбільш поширеними є такі типи моделювання білків: моделювання укладання «з перших принципів» (ab initio) [4], передбачення на основі гомології (homology modeling) [5], трідинг (threading) [6].

Оскільки, система моделювання повинна представляти оцінку коректності своїх результатів, вхідними даними модуля є дві складні просторові структури білка, які представляються у вигляді координат атомів у трьохвимірній системі координат. Блок-схема роботи розробленої системи подана на рис. 1.

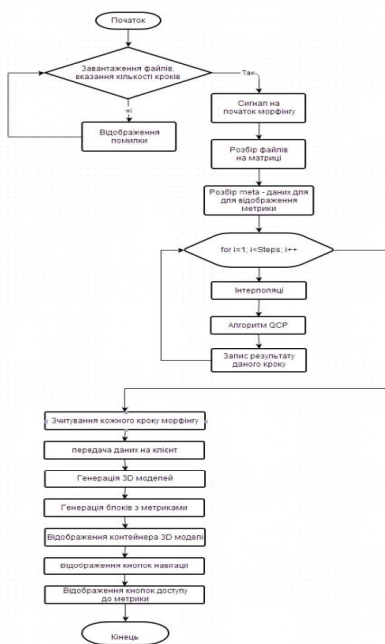


Рис. 1. Блок-схема діяльності розробленої системи

У результаті модельного експерименту вдалось виявити, що конкатенація вхідних даних призводить до удосконалення в проектуванні білка, який у свою чергу може безпосередньо застосовувати при створенні нових ферментів, вакцин, тощо. Процедура аналізу результатів повинна бути направлена на глибокий аналіз структурної схожості білкових молекул. Порівняння змодельованих експериментальних білкових структур виконується у три підходи: 1) Візуальне порівняння. У людини, яка досліджує

є можливість наглядно порівняти результат моделювання та експериментальних даних, а також провести операцію обертання й масштабування для візуальної оцінки коректності параметрів використовуваної моделі. 2) Оцінка СКВ, яка в молекулярному моделюванні використовується в якості просторової міри близькості двох моделей, де низьке СКВ означає близькість двох структур. 3) Порівняння структур як вузлів. Проекція трьохрівневих структур білка на площині представляє собою складний вузол, тому теорія вузлів дозволяє представити заплутані структури у вигляді полінома.

Результат проведеного моделювання синтезу білка на різних етапах зі зміною координат подано на рис. 2.

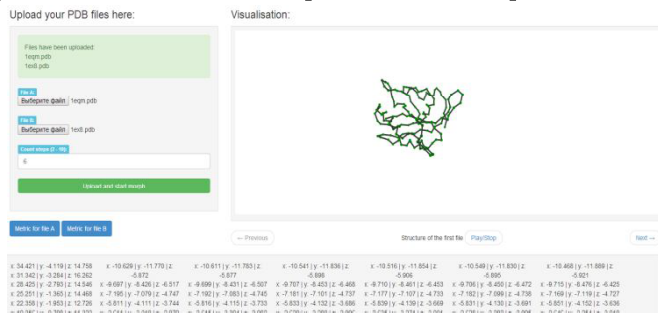


Рис. 2. Вікно для моделювання синтезу білків

## ВИСНОВКИ

Дані про проміжні структури, що представляють зміну білка з однієї конформації в іншу, надають багато інформації про його функції. Експериментальні методи є недостатніми для вирішення цього завдання за рахунок практичних і технологічних обмежень. З цієї причини, структурна біологія гостро потребує алгоритмів, які можуть точно передбачити проміжні структури, що відображають процес конформаційних змін білка.

Встановлено, чим менше відхилення від еталонного структурного коду, тим взаємодія білків в межах одного препарату підсилюється. Це дозволяє забезпечити параметричну оптимізацію процесу синтезу нових елементів. Ефективність нового запропонованого методу моделювання підвищується за рахунок того, що він безкоштовний та дозволяє отримувати аналогічні результати.

## ЛІТЕРАТУРА REFERENCES

- [1] К.О. Одинець, С.С. Івахно, Д.Б.Ковальський, Б.Т. Токоєнко, О.І. Корнелюк, Структурна біоінформатика в постгенному еру, Київ, 2004.
- [2] Liu P, Agrafiotis DK, Theobald DL (2011) [Online]. Available: <http://theobald.brandeis.edu/qcp>.
- [3] Hillisch A., Pineda L.F., Hilgenfeld R. Utility of homology models in the drug discovery process. Drug Discov. (2004) Today 15, 659–669.
- [4] M. S. Shell, 2009. Ab Initio methods [Online]. Available: <http://www.ieee.org/documents/ieeecitationref.pdf>.
- [5] Elmar Krieger, Sander B. Nabuurs and Gert Vriend, Homology Modeling 2005.
- [6] Jinbo Xu, Feng Jiao, Libo Yu, Protein Structure Prediction Using Threading, 2008, pp. 91-121.