



# БАЛТИЙСКИЙ НАУЧНО-ИНЖЕНЕРНЫЙ КОНКУРС

Санкт-Петербург, 30 января - 2 февраля 2017 года

## Разработка программного инструмента для предсказания комплекса белков

Реутский Даниил Александрович и Ухатов Александр Александрович.

г. Санкт-Петербург, лицей «Физико-техническая Школа», 11а класс.

Руководитель: Яковлев Павел Андреевич, директор департамента вычислительной биологии компании BIOCAD

Жизнедеятельность организмов во многом регулируется белок-белковыми взаимодействиями. Таким образом передаются сигналы в клетках, обеспечивается транспорт микро- и макроэлементов, работа иммунной системы. Последняя иногда даёт сбой, поэтому необходимо создавать новые молекулы для нейтрализации конкретных враждебных белков. Для этого требуется понимать механизм взаимодействия между ними. Эти задачи можно решать не только экспериментально, но и с помощью математического моделирования, тем самым сократив финансовые и временные затраты. Одна из таких задач состоит в предсказании оптимальной позиции двух белковых молекул друг относительно друга (иными словами, в предсказании строения комплекса белков). Цель данной работы — разработка *программного инструмента* для предсказания комплекса белков. Инструмент должен обладать минимальным временем работы и высокой предсказательной способностью. Всё исследование было разделено на несколько подзадач: 1) поддержка чтения файлов со структурами молекул; 2) преобразование структуры молекулы до трёхмерной кубической сетки (квантование); 3) поиск оптимального взаимного расположения двух молекул; 4) проверка выбранных положений на основании электростатического взаимодействия молекул друг с другом.

Для решения первой подзадачи требовалось изучить документацию формата файлов PDB. Этот формат является стандартом для представления трёхмерных структур белков и используется одной из крупнейших баз данных — Protein Data Bank ([www.rcsb.org](http://www.rcsb.org)). Для решения второй задачи была использована структура данных Range-tree, а в последствии был разработан более быстрый способ, основанный на округлении чисел до целых. Решение третьей подзадачи использует Быстрое прямое и обратное Преобразования Фурье (готовая реализация Jtransforms). Идею использовать его мы позаимствовали из статьи «Molecular surface recognition: Determination of geometric fit between proteins and their ligands by correlation techniques» (Proc. Natl. Acad. Sci. USA, vol. 89 no. 6, 2195–2199). В четвёртой подзадаче каждая молекула рассмотрена, как система точечных зарядов, подчиняющихся закону Кулона. Отсюда делается вывод о том, будут ли молекулы отталкиваться или притягиваться. Дополнительно мы реализовали визуализацию структур итогового белкового комплекса и каждой молекулы в отдельности (использована библиотека Java3D)

В результате был разработан программный инструмент, который моделирует оптимальную с точки зрения геометрии и электростатики стыковку двух белковых молекул. Рассмотрены, реализованы, оптимизированы методы и алгоритмы, позволяющие уменьшить время работы. Инструмент был протестирован на уже известных комплексах, результаты согласуются с экспериментальными данными. Этот результат был достигнут не сразу, а после подбора некоторых параметров (размер сетки, эффективные радиусы атомов).

В дальнейшем можно уточнить физическую модель взаимодействия, приняв к рассмотрению Ван-дер-Ваальсовы взаимодействия, энергии вращения и влияние растворителя. С помощью параллельных вычислений можно ускорить отдельные части программы. А применения реализованных методов с небольшими изменениями позволяют решать некоторые задачи из совершенно других областей, например, из области распознавания образов.