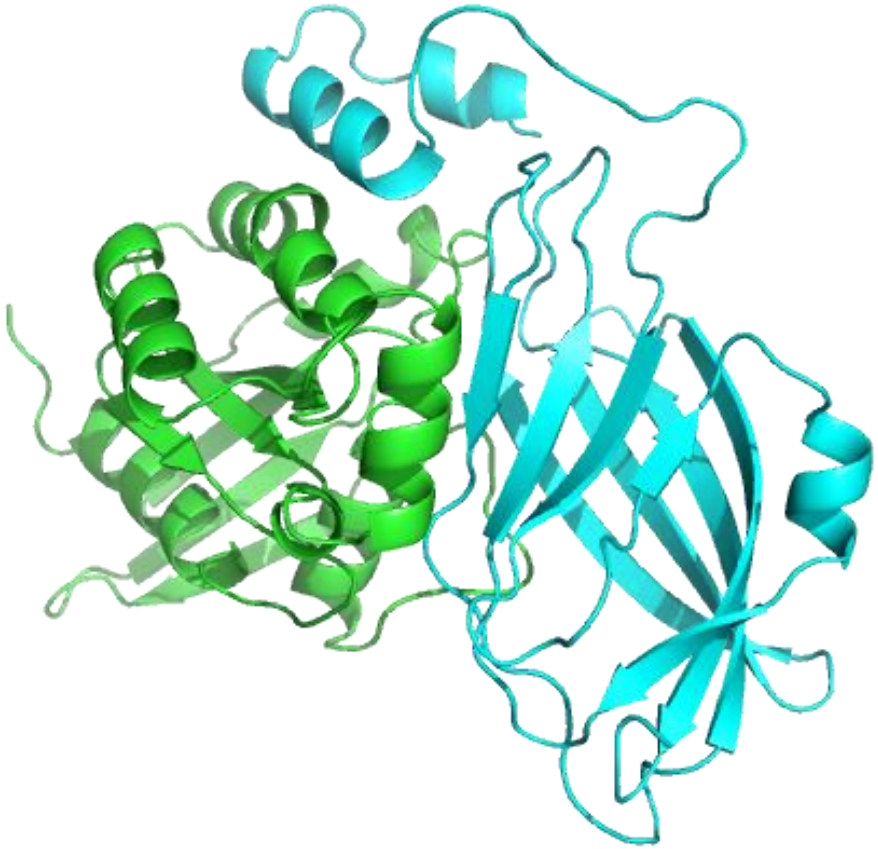
Project Proposal

Deep Learning Guided Complex Refinement

Presented by

Mohimenul Karim

Protein Complex



Background

- Understand protein-protein interaction to understand biological function
- Know protein complex structure to understand protein-protein interaction
- Different experimental methods to determine complex structure
  - X-ray crystallography
  - Nuclear magnetic resonance
- Problems
  - Time consuming
  - Expensive

Background

- Rigid body docking approach
- Underlying assumption
  - No conformational change upon binding
- Problem
  - Inaccuracy in the interface contact
  - Atomic clashes
- Hence, refinement is needed

Existing Complex Refinement Methods

- Some of the existing complex refinement methods
  - RDOCK
  - HADDOCK
  - GalaxyRefineComplex
- Dependent on
  - Energy minimization
  - Molecular dynamics simulations

Our Hypothesis

- Leverage deep learning prediction to guide protein complex refinement
- Hypothesis
  - Our deep learning prediction will perform well which will guide the complex refinement and improve the quality of the protein complexes

Our Approach

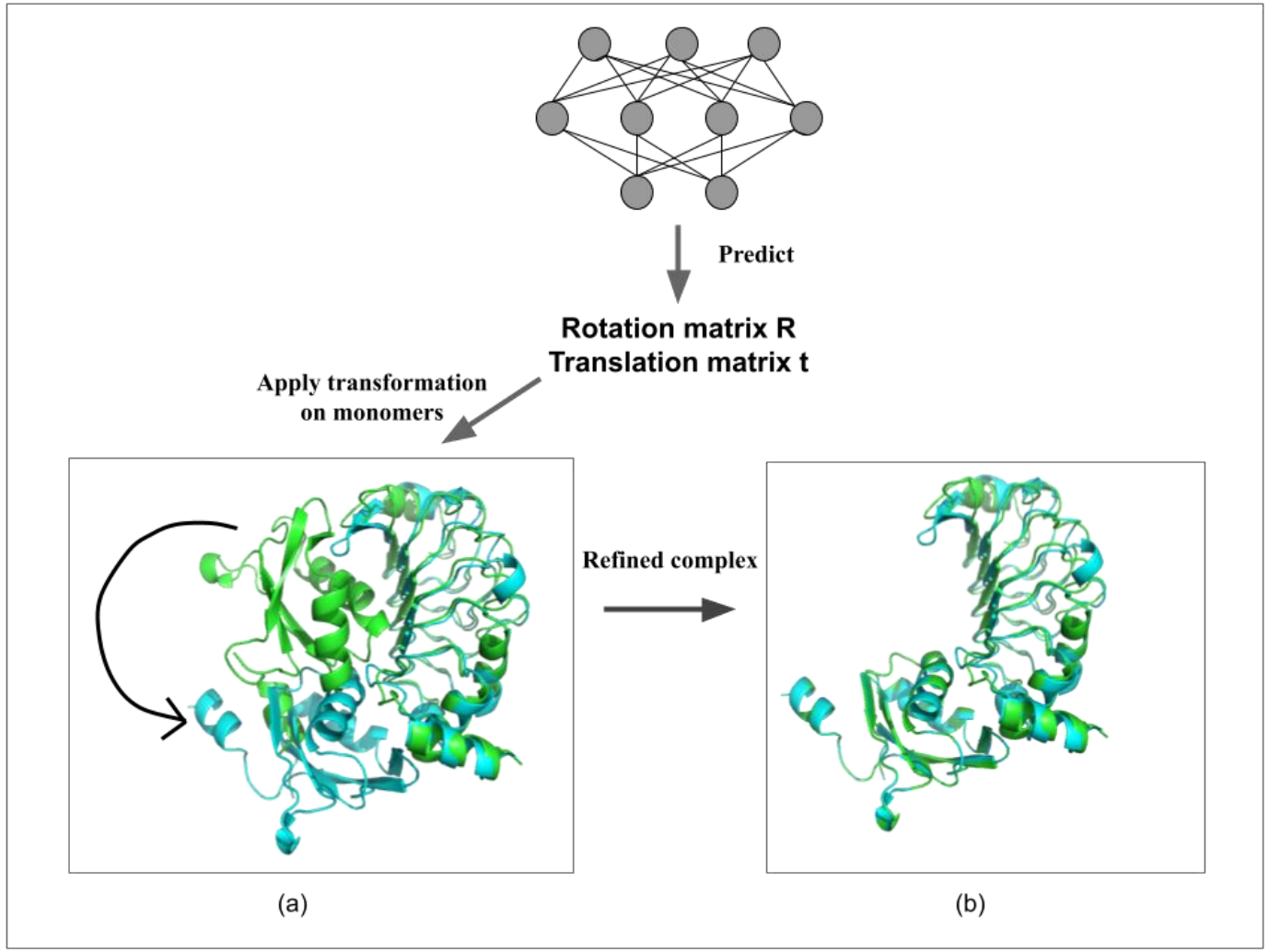
- Develop a protein complex refinement method powered by deep learning
- Predict rotation and translation matrix for individual monomers of the near native complex
- The metrices can be used to orient the individual monomers of the complex in a correct formation

Basic Mathematical Idea

- The positions of the atoms in an individual monomer of the near native complex: X
- The positions of the atoms in the corresponding individual monomer of the native complex: X′
- The rotation and translation matrices required to bring the individual monomers closer to the native: R (rotation), t (translation)
- Predict R, t and apply to X so that

$$X' = RX + t$$

Overview of our method

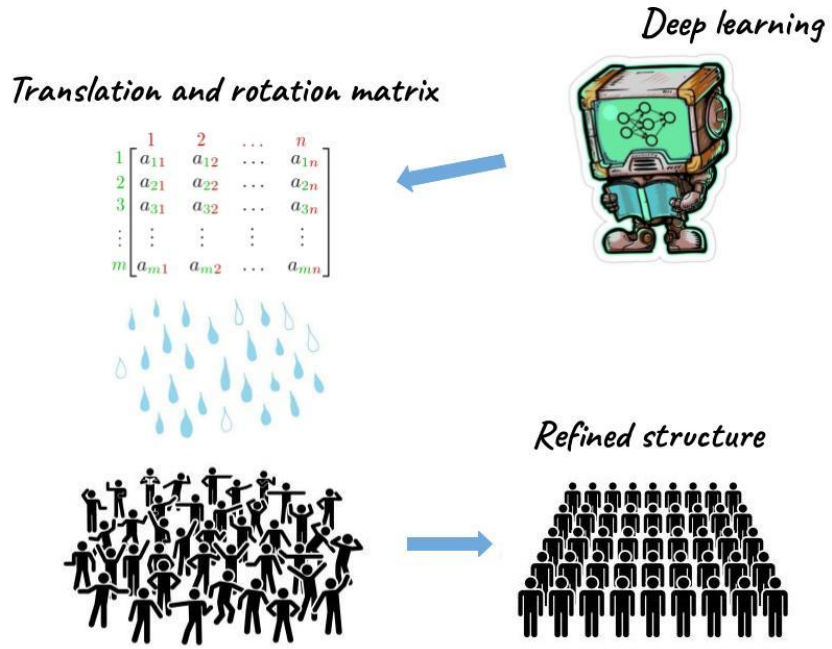


Evaluation Metrics

- Mean squared error (MSE), as our prediction is going to be a regression problem
- TM-score
- Global Distance Test – Total Score (GDT_TS)
- Interface RMSD (I-RMSD)

Project Milestones

List of Tasks	Timeframe
Collect dataset	March 05, 2022
Calculate transformation matrix for a few complexes and check the validity of the proposed approach by applying the matrix on the complexes	March 06, 2022
Automate the calculation of transformation matrix for all complexes	March 14, 2022
Generate features from the dataset	March 21, 2022
Prepare deep learning model for prediction	March 28, 2022
Train the model and predict the transformation matrices	March 29, 2022 – April 10, 2022



The images above were taken from https://www.google.com/imghp and modified for presentation purpose