Roll number: 1068

CBSB3 Mini-Project Proposal

Project 7

Word count: 793

Brief Background & Motivation

Helicobacter pylori infects around 50% of people all over the world and can cause diseases like gastritis and gastric cancer. The expression of protein CagA has been proven to correlate with *H. pylori* virulence. When attached to gastric epithelial cells, the bacteria inject CagA into the cells, while the protein undergoes tyrosine phosphorylation and induces a growth factor-like response, leading to cellular morphological changes. SHP-2 is a tyrosine phosphatase in the plasma comprising two SRC homology 2 (SH2) domains. It plays an important role in the regulation of mitosis, spreading, migration, and adhesion of cells. Phosphorylated CagA can bind with SH2 domains within SHP-2 and stimulate phosphatase activity, therefore, deregulation of SHP-2 by CagA has the potential to transform gastric epithelial cells into cancer cells. Glu-Pro-Ile-Tyr-Ala (EPIYA) sequences that are classified as four variants (A, B, C, and D) are phosphorylation sites through CagA. EPIYA-A and EPIYA-B are conservative motifs presented in CagA, while EPIYA-C and EPIYA-D are specific motifs carried separately by Western and East Asian isolates of *H. pylori*. According to the article by Takeru Hayashi, the binding affinity of CagA and SHP-2 positively correlates with the number of phosphorylated EPIYA-C and EPIYA-D, and the latter in the East Asian CagA has a larger effect on the affinity. It is also proved that phosphorylated EPIYA-D binds more tightly to the N-SH2 domain than phosphorylated EPIYA-C, which might be a reason why gastric carcinoma is prevalent in East Asia. The affinity difference is due to the diversity of the amino acid residue at pY + 5 positio. N-SH2 and C-SH2 domains of SHP-2 are both able to bind with EPIYA-C and D, however, only the N-SH2 domain is confirmed to have a higher affinity to EPIYA-D. In this case, we want to have a deeper understanding of SHP-2-CagA complex by analyzing the interaction intensity between the C-SH2 domain and EPIYA-C and D.

Aims/Hypotheses

In this mini-project, I plan to validate that the N-SH2 domain of SHP-2 binds to both phosphorylated EPIYA-C and D segments and has higher affinity with EPIYA-D than with EPIYA-C; test the affinities of the C-domain of SHP-2 interacting with phosphorylated EPIYA-C and D; furthermore, compare the binding of EPIYA with the N-domain and with the C-domain.

Oriented peptide array libraries and protein docking are two ways to study protein-protein interactions, nevertheless, they are either expensive or poorly compatible. Thus, the molecular dynamics (MD) approach is chosen to demonstrate that EPIpYA-D binds to either the N-SH2 domain or C-SH2 domain of SHP-2 and is more stable than EPIpYA-C does in the MD simulation.

Brief Summary of Methods

In molecular dynamics simulations, the atoms and molecules are treated as point masses connected by bonds, angles, and dihedrals. The non-bonded interactions, such as van der Waals and electrostatic forces, are also considered. The simulation proceeds by integrating the equations of motion to update the positions and velocities of the particles at each step. MD simulations allow me to investigate the behavior of SHP-2 and CagA and their interactions between SH2 domains and EPIYAs over time, providing insights that may be difficult to obtain through experimental techniques alone. By simulating the motion of atoms based on classical mechanics, MD simulations can capture the conformational changes, flexibility, and stability of proteins and their complexes. By doing this. I can explore the binding interface, identify key residues involved in the interaction, and estimate the strength of the binding.

Work Plan & Feasibility Assessment

There are 5 weeks plus 5 days in total from April 8th to May 17th which is the deadline for the mini-project report. I plan to use 5 weeks to work on the project itself and 5 days to finish the report.

Week 1: Do a literature review and find the best approach and software for the project.

Week 2: Obtain the protein structures of SHP-2 and CagA from Protein Data Bank (PDB). Dock the EPIYA-C or D motif to either the N-SH2 or C-SH2 domain if needed. EPIpYA-C or D binding with C or D-SH2 domain are experimental groups. The EPIYA segments without phosphorylation are set to be negative control groups since they have very low binding strength with the SH2 domain.

Week 3: Use software like CHARMM to modify the structure and assign force field parameters to it, minimize the overall energy of the system, and equilibrate it at the appropriate temperature and pressure.

Week 4: Run the MD simulation with software like GROMACS to capture the dynamics of the interaction between the SH2 domains and the EPIYA motifs.

Week 5: Plot the affinity results of the simulation, analyze and interpret.

If something doesn’t go as intended, I will discuss it with my tutor and try to solve the problem or reduce some validation work.