Bilal Walid Hamdanieh & Marwa Ghraizi

Structural Bioinformatics

With Dr. Joseph Rebehmed

Structural Quality Assessment

This report will briefly portray the overall structure of Spectrin Pleckstrin Homology (PH) Domain bound with an Inositol Phosphate and will analyze the Surface Features at the Binding Site. Majorly, the project will include a Structural Quality Assessment of the relative model of PDBID: IBTN.

A- Overall Functional-Structural Features of a Pleckstrin Homology (PH) Domain:

Phosphoinositides (PIs) are lipid components of cellular membranes that function as signaling molecules. The inositol headgroups of Phosphoinositides are differentially phosphorylated, and selectively bound by a variety of protein modules, including the family of Pleckstrin Homology (PH) domain proteins.

PH domains were the first phosphoinositide binding domain identified and serve important roles in kinase signaling and cytoskeletal organization. PH domains consist of 100–120 amino acids that form a six-stranded β-sandwich with a C-terminal-helix *(Figure 1)*.

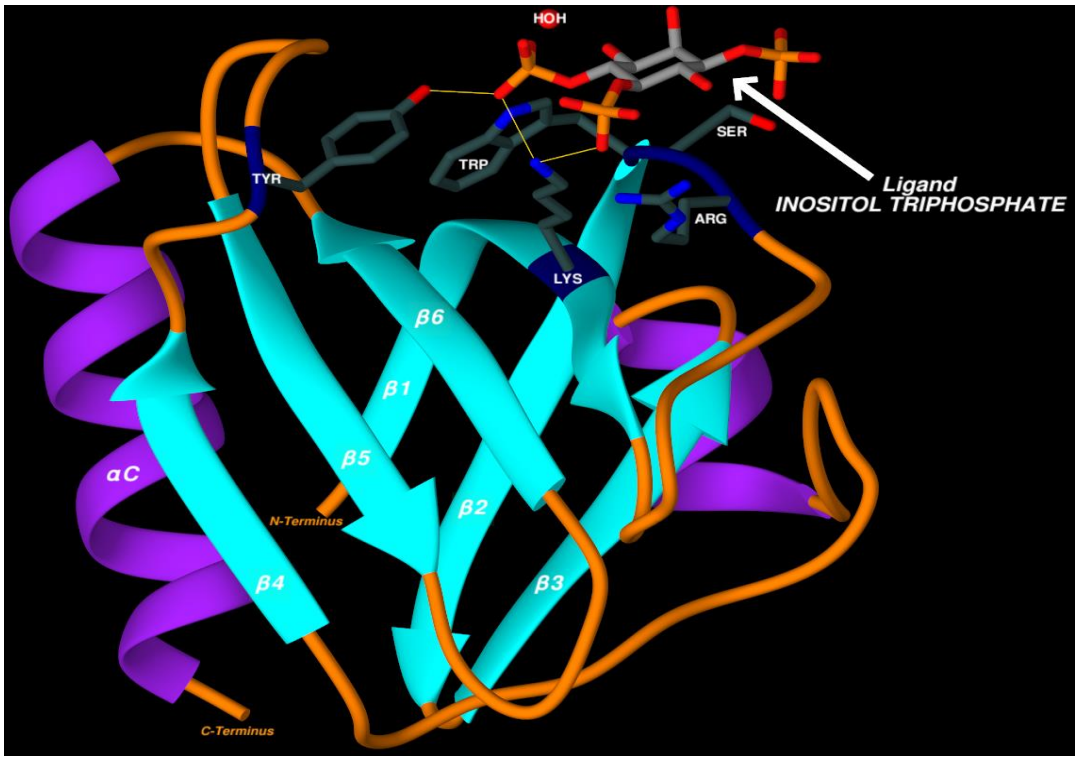


Figure 1: An overall ribbon view of the structure of the PH domain with the binding site for Inositol Phosphates generated via Chimera.

B- Surface Analysis of the Binding Site for PI in the PH Domain:

Physicochemical properties play a key role in stabilizing protein-ligand complexes notably electrostatic attractions. The negatively charged ligand in question here is stabilized by several positively charged residues: Tryptophane, Lysine and Arginine whose clouds are represented in blue in the figure below. In addition to Tyrosine that can create a positive and negative cloud charge around its hydroxyl group OH *(Figure 2)*

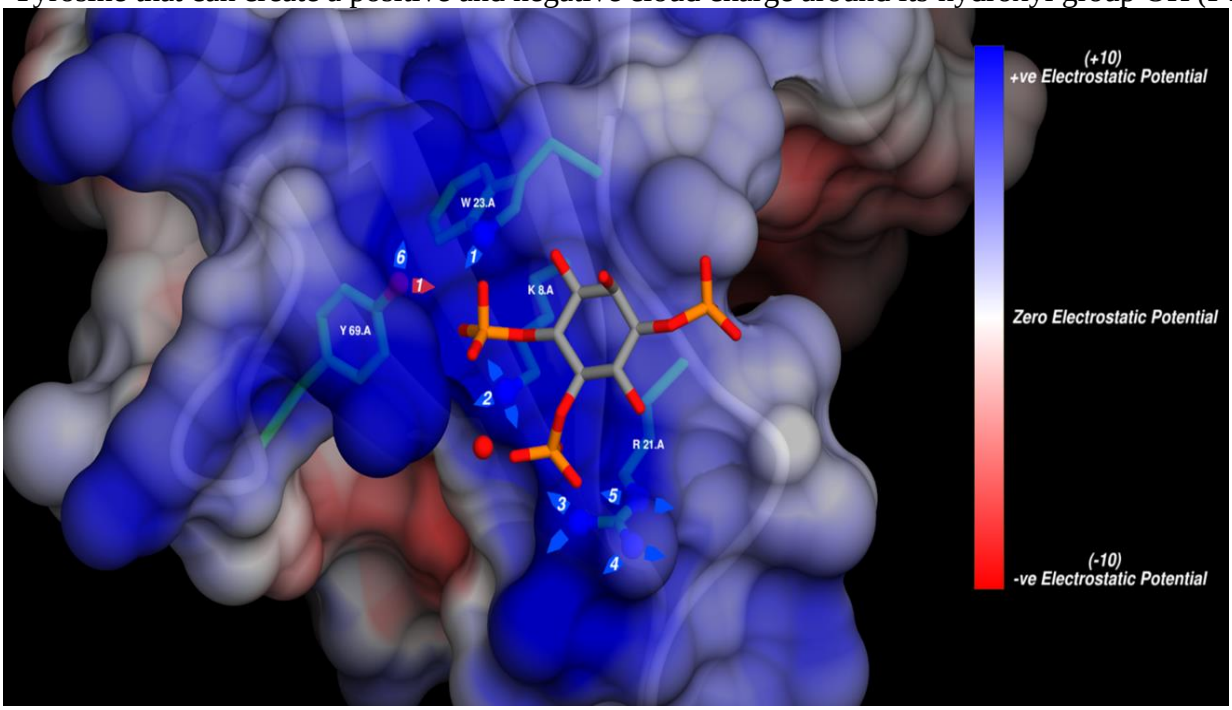


Figure 2: Representation of the electrostatic potential map of the surface of the binding site including key annotated residues that interact with the ligand via electrostatic attraction. The binding site contains 6 positively charged residues and 1 negatively charged therefore making the overall charge positive which allows the binding to the negatively charged ligand.

C- Structural Quality Assessment

Global Parameters:

1. **Resolution:** the model resolution documented by the depositor ranges between the lowest resolution power of 8Å and the highest resolution power of 2.00Å. However, Electron Density Server (EDS) showed the lowest resolution power to be 24.40Å and the highest resolution power to be 2.03Å. Due to the usual inconsistency in resolution, and it being overstated, the difference between the documented resolutions is insignificant, and the model can be considered of fair resolution.
2. **R-Factor**: The R-Factor is documented by the model depositor to be 0.205. The latter shows that the model reflects a true agreement between the observed X-ray diffraction patterns and the computed patterns after model composition.
3. **Atomic B-Factor:** An approximation of the average atomic B-factors using Wilson B-factor revealed a value of 25.5($Å^2$). The latter shows that on average, the position of the atoms is reliable (low uncertainty). Upon testing in Chimera, all residues showed an average B-factor less than 40$Å^2$ which further validates the reliability of the atomic coordinates' certainty. On further analysis, we deduced that Lysine70 has the highest average B-Factor of approximately 35.5$Å^2$, where the Zeta-Nitrogen atom (NZ) has the highest B-Factor value of approximately 46.5$Å^2$. However, Lysine is located in **loop/turn** which explains its higher B-Factor being the ability to flexibly move in space therefore implying lower certainty in its position upon model construction (Figure 3).

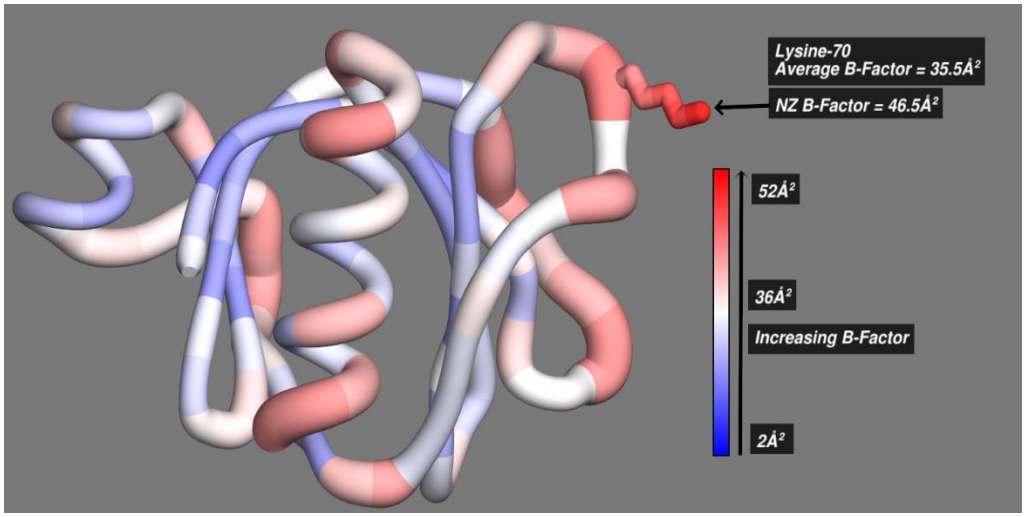
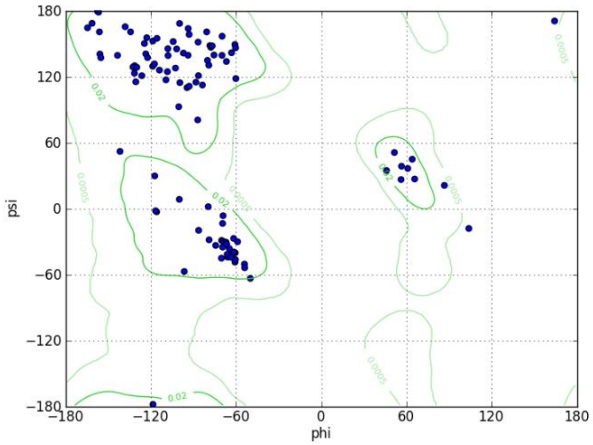


Figure 3 Putty/Sausage view of the protein generated by Chimera depending on the B-factor

Stereochemical Parameters:

1. **Ramachandran's Plot**: Upon validation using PDBsum, PROSESS, and MOLprobity, and after generating the Ramachandran's plot in Chimera, we have concluded that the model lacks any outliers, where:



⇨ 88.8% of the residues belong to the most favoured regions of ramachandran's plot (delimited by a green line)

⇨ 11.2% of the residues belong to the additional allowed regions of ramachandran's plot.

For a resolution of 2.00Å, the percentage of outliers must not exceed 10%. For the model in study, an exceptional 0% outliers is seen, indicating that the model passes the quality check.

Figure 4: Ramachandran Plot generated Via Chimera

2. **Bad Contacts (Clashes):** are unfavorable interactions where atoms are too close together; the distance between two atoms is less than the sum of their radii. For the model under study, 22 clashes are observed and validated through PDBsum, full PDB file report, PROSESS, and MOLprobity. These clashes could have occurred due to several reasons: improper refinement, chain mistracing, improper residues geometry parameters, etc. One of the significant clashes in the model under study is the clash between Lys42 NZ and Ser46 OG with an overlap of 0.572Å (Chimera; Figure 5).
   As shown in the figure below, the sum of the radii (3.25Å) is greater than the distance between the two atoms, thus these two atoms are clashing.

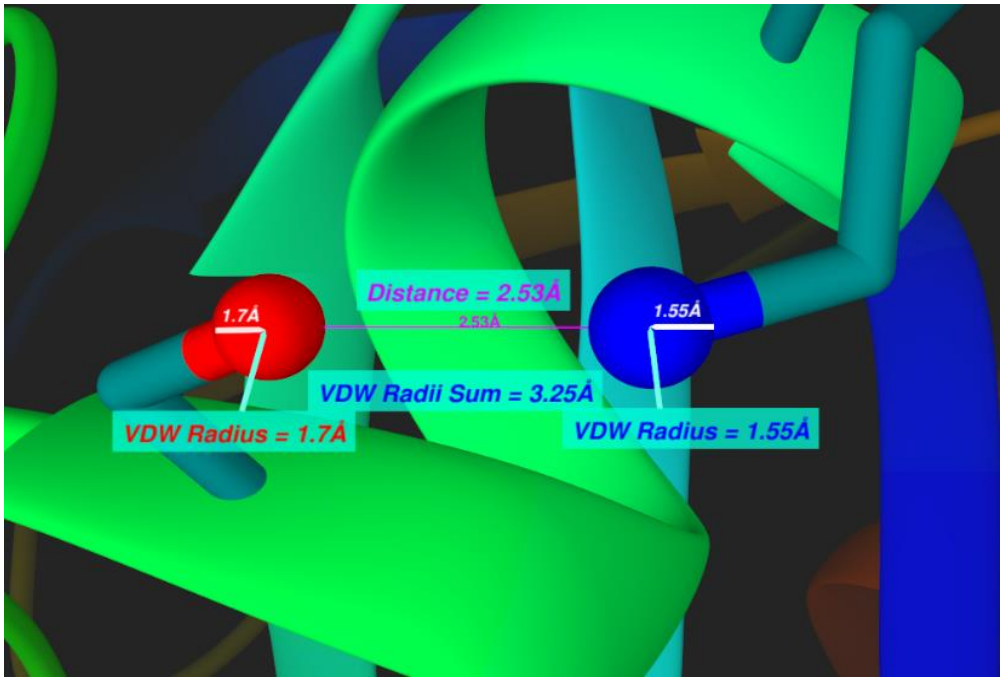


Figure 5: Illustration of the atomic clash between Lys42 and Ser46 generated by Chimera

Conclusion

The overall quality assessment of Pleckstrin Homology (PH) Domain bound with Inositol Phosphate demonstrates a model of good quality characterized by a high resolution, absence of Ramachandran outliers and reliable R-factor and B-factor. However, 22 atomic clashes are present including clashes between residues like Lysine42 and Serine46 along with a relatively high B-factor for Lysine70 justified by its presence in a geometrically flexible loop.

Reference

Ceccarelli, D. F., Blasutig, I. M., Goudreault, M., Li, Z., Ruston, J., Pawson, T., & Sicheri, F. (2007). Non-canonical interaction of phosphoinositides with pleckstrin homology domains of Tiam1 and ArhGAP9. *Journal of Biological Chemistry*, *282*(18), 13864-13874.