

# CBSB3 mini-project proposal of project3:

## modelling mutational effects on protein structure

roll number: 0022

word count: 956

### Background & motivation

The study of protein structure is critical for understanding protein function, and one way to investigate the impact of amino acid mutations on protein stability is through site-directed mutagenesis. The bacteriophage T4 lysozyme (T4L) protein has been extensively studied and is often used as a model protein to examine stability, ligand binding, and conformational changes (Wang, Papaleo and Lindorff-Larsen, 2016). Early studies by Matsumura and colleagues involved mutating multiple residues to cysteines to investigate the impact of disulfide bonds on protein stability, led to the creation of the pseudo wildtype (\*wt) T4L by eliminating potential disulfide bond-forming residues (Matsumura, Signor and Matthews, 1989). Further point mutations such as the A42S mutation (206L, Fig1) in the hydrophobic interior of T4L, which introduces a polar hydroxyl group (Blaber *et al.*, 1993), and the L99A mutation (3DMV, Fig1) that creates a cavity within the protein, have provided insight into the effects of polarity and cavity formation on protein stability and ligand binding (Liu, Baase and Matthews, 2009). The 2CL9 mutant, which introduces G113A and R119P mutations, led to the reversal of the stability of the two conformations of the T4L protein (Bouvignies *et al.*, 2011), a topic that was further studied using molecular dynamics simulations by Wang *et al.* in 2016. Overall, these site-directed mutagenesis studies on T4L have contributed greatly to our understanding of protein structure and function.

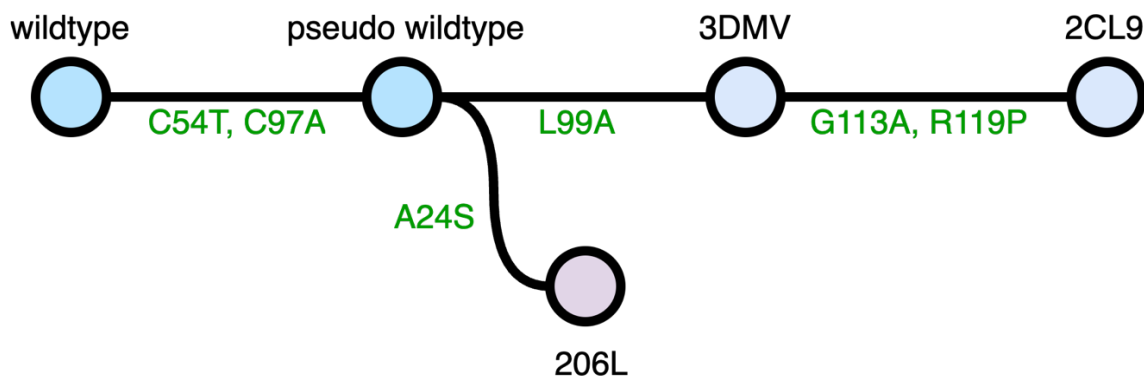


Figure1. The relationship between different T4L mutants. Each cycle, together with the black ID represents a protein mutant, green notations represent the mutation happen from the previous protein.

### Aims & hypotheses

This project aims to conduct further research on the L99A T4L (3DMV) protein. This protein possesses a number of fascinating properties that make it an excellent model system for understanding protein structure and dynamics. Unlike other normal T4L mutant proteins, it has a cavity structure, allowing it

to be used to study protein-ligand interactions. Additionally, it has two relatively stable conformations, and it has been demonstrated that the transition between these conformations is related to ligand binding (Kamenik *et al.*, 2021). Thus, this project proposes the following mutation studies on L99A T4L: (1) choosing an amino acid residue near its internal cavity and mutating it to a polar residue to examine the change in protein folding free energy before and after the mutation. Previous research has shown that introducing polar amino acid residues inside the protein can greatly destabilize the protein (Blaber *et al.*, 1993). Investigating whether the destabilizing effect changes when the polar group is introduced around the cavity can deepen our understanding of the destabilizing effect of polar groups inside proteins. (2) Using the Rosetta structure-based design calculation to determine the substituents that predictably stabilize the excited state relative to the ground state. Mutations similar to G113A and R119P will be selected, and their effects on the stability of the two conformations will be predicted. Based on previous research, it is assumed that these mutations will decrease or even reverse the free energy difference between the ground and excited state conformations (Bouvignies *et al.*, 2011). (3) Conducting ligand docking analysis on the mutated proteins generated in the previous steps, with particular attention to the binding ability of the ligand in different conformations. It is hypothesized that the mutation variant with the open cavity conformation, which has a larger cavity and more solvent-exposed regions, will have a greater binding affinity for larger and polar ligands, while the opposite is true for the closed cavity conformation (Kamenik *et al.*, 2021). This proposed research will contribute to our understanding of protein-ligand interactions and protein dynamics.

## Summary of methods

This project requires multiple analysis on different aspects including protein stability, conformational changes, and ligand binding. Structural design calculations based on rosetta can be used to determine the best substituents that can stabilize the excited state relative to the ground state. To investigate the effect of introducing point mutation, the FoldX software can be used to calculate the change in folding free energy before and after the mutation. The effects of mutations on protein-ligand interactions can be evaluated using ligand docking software such as Autodock combined with Chimera. To compare and fit the simulated data to experimental data, various analysis tools can be used. For example, molecular dynamics trajectories can be analyzed using software such as VMD or PyMOL to visualize protein conformational changes and interactions with ligands. Free energy calculations can be analyzed using tools such as the Python-based PyEMMA software, which can calculate free energy landscapes and provide insights into protein dynamics.

## Work plan & feasibility assessment

Week 1: Literature review on L99A T4L protein and its properties. Identify the amino acid residue to be mutated and perform initial structure analysis.

Week 2: Conduct Rosetta structure-based design calculations to determine the substituents that can stabilize the excited state relative to the ground state.

Week 3-4: Mutate the selected amino acid residue and the identified mutations similar to G113A and R119P. Analyze the effects of the mutations on protein stability using molecular dynamics simulations.

Week 5: Conduct ligand docking analysis on the mutated proteins generated in the previous steps to investigate the binding ability of the ligand in different conformations. Start writing the final report.

Week 6: Analyze the results obtained from the previous steps and draw conclusions. Finish writing the final report.

Overall, the proposed project appears to be feasible, given that the necessary software and tools are readily available. However, some specific considerations should be taken into account to ensure the successful completion of the project. First, the availability and quality of the necessary data and structures must be verified. The quality of the data of L99A T4L (3DMV) protein structure from PDB should be carefully assessed before use. Secondly, the computational demands of the Rosetta design calculation can vary greatly depending on the size and complexity of the protein system, requiring access to suitable computational resources such as high-performance computing clusters or cloud computing services.

## Reference

Blaber, M. *et al.* (1993) 'Energetic cost and structural consequences of burying a hydroxyl group within the core of a protein determined from Ala → Ser and Val → Thr substitutions in T4 lysozyme', *Biochemistry*, 32(42), pp. 11363–11373. Available at: <https://doi.org/10.1021/bi00093a013>.

Bouvignies, G. *et al.* (2011) 'Solution structure of a minor and transiently formed state of a T4 lysozyme mutant', *Nature*, 477(7362), pp. 111–114. Available at: <https://doi.org/10.1038/nature10349>.

Kamenik, A.S. *et al.* (2021) 'Energy penalties enhance flexible receptor docking in a model cavity', *Proceedings of the National Academy of Sciences*, 118(36), p. e2106195118. Available at: <https://doi.org/10.1073/pnas.2106195118>.

Liu, L., Baase, W.A. and Matthews, B.W. (2009) 'Halogenated Benzenes Bound within a Non-polar Cavity in T4 Lysozyme Provide Examples of I...S and I...Se Halogen-bonding', *Journal of Molecular Biology*, 385(2), pp. 595–605. Available at: <https://doi.org/10.1016/j.jmb.2008.10.086>.

Matsumura, M., Signor, G. and Matthews, B.W. (1989) 'Substantial increase of protein stability by multiple disulphide bonds', *Nature*, 342(6247), pp. 291–293. Available at: <https://doi.org/10.1038/342291a0>.

Wang, Y., Papaleo, E. and Lindorff-Larsen, K. (2016) 'Mapping transiently formed and sparsely populated conformations on a complex energy landscape', *eLife*. Edited by Y. Shan, 5, p. e17505. Available at: <https://doi.org/10.7554/eLife.17505>.