Structural and Functional Decomposition of Universal Stress Protein A from *M. luteus*

Tae Y. Lim[1,2], and Mehmet Şen[2]
Department of Computer Science[1], and Biology and Biochemistry[2]

UNIVERSITY of HOUSTON

Background

The aim of this study was to determine the conserved structure of UspA 712 molecule from *Micrococcus Luteus* and its potential post-translational modification sites, the oldest known dormant bacteria. When structurally compared to other 32 UspA structures, two major structural ensembles are identified: a single-lobed fold, which homodimerizes and a double-lobed fold. Interestingly, both Usp structural families share an invariant classic core structure yet have significant divergence at specific looping regions. Specifically, these regions contain variety of post-translational modifications, attributing diverse functions to each protein. Furthermore, the correspondence of flexibility may associate with functionality.

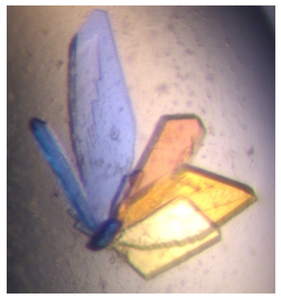

Figure 1A: UspA 712 crystal

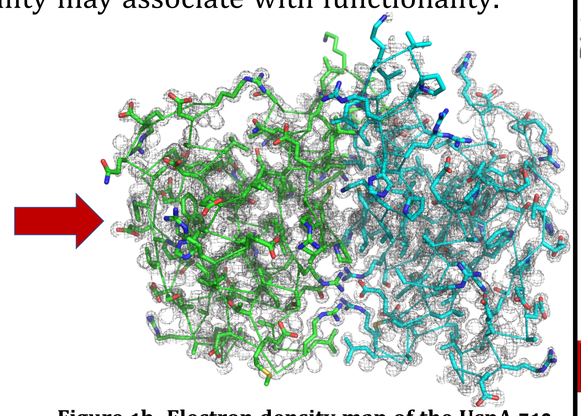

Figure 1b: Electron density map of the UspA 712

Methodology

- Due to missing regions of amino acids, molecular replacement was essential. Missing regions in UspA 712's structure were computationally built using Coot and refined using Phenix.
- All 33 Usp structures from the RCSB Databank were minimized utilizing AMBER/SANDER Molecular Dynamics to reduce the clashing of amino acids and the proteins' energy system.
- Multiple protein structural alignment algorithm was then implemented to retrieve the corresponding superposition coordinates of the 33 Usp structures.
- In PCA, the invariant atomic core positions were determined to provide confirmation of highly conserved regions in the cα-backbone amongst the structures. Furthermore, the structures were clustered into 3 main groups (red, blue, and green).
- In eNMA, the vibrational modes and protein flexibility are calculated by modeling the atoms as point masses connected by springs, which represent the interatomic force fields. The fluctuations are in terms of Root Mean Inner Squared Product.

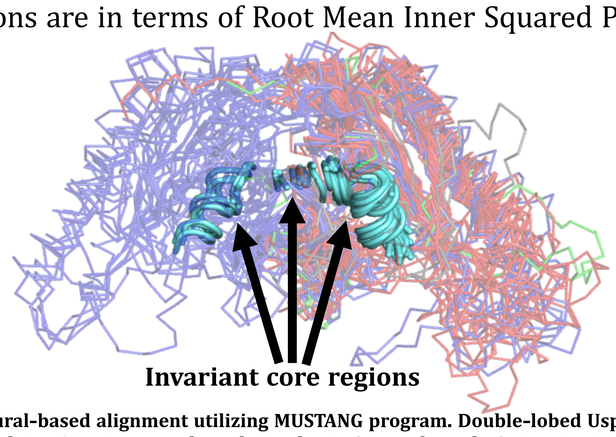

Figure 2: Structural-based alignment utilizing MUSTANG program. Double-lobed Usp structures (blue), single-lobed Usp Structures (red), and Mycobacterium Tuberculosis H37Rv structure (green)

Principal Component Analysis

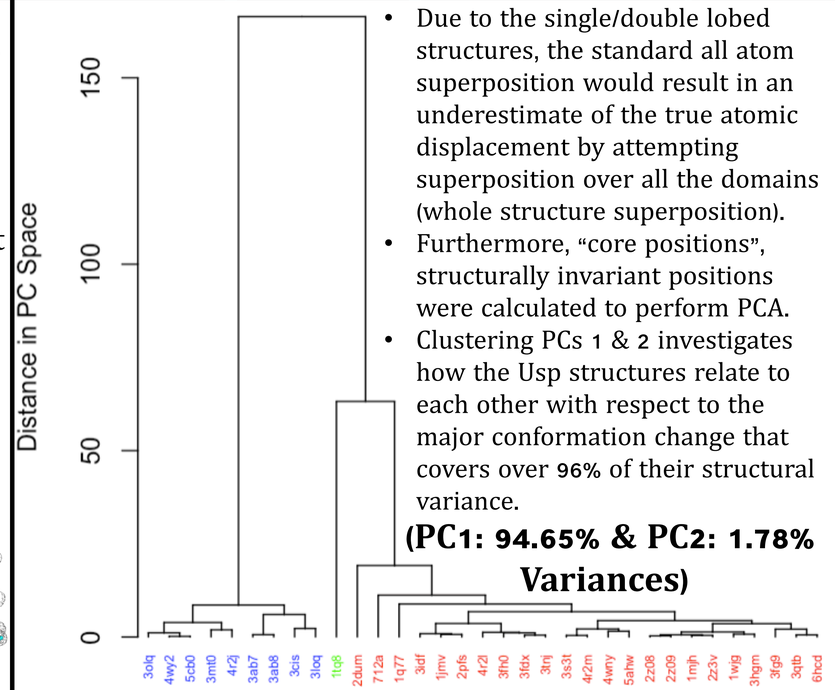

Figure 4: PC1 vs. PC2 Dendrogram showing distance in PC Space from each structure

- Due to the single/double lobed structures, the standard all atom superposition would result in an underestimate of the true atomic displacement by attempting superposition over all the domains (whole structure superposition).
- Furthermore, "core positions", structurally invariant positions were calculated to perform PCA.
- Clustering PCs 1 & 2 investigates how the Usp structures relate to each other with respect to the major conformation change that covers over 96% of their structural variance.

(PC1: 94.65% & PC2: 1.78% Variances)

Ensemble Normal Modes Analysis

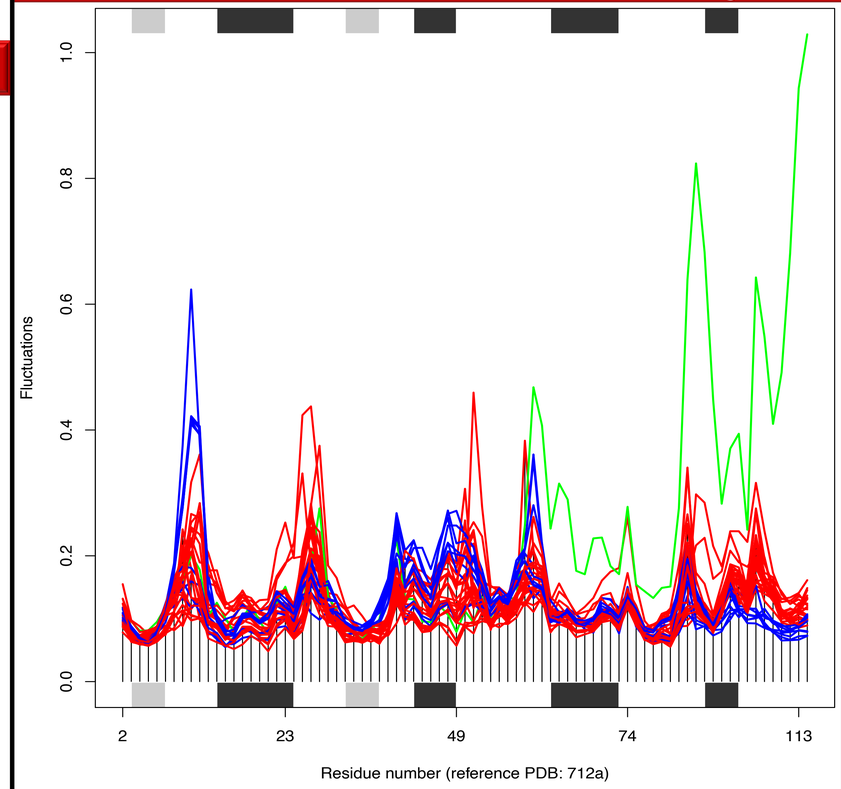

Figure 5: Ensemble Normal Modes Fluctuations for each Usp structure (colors according to respective clusters from PCA); Grey shows the β-sheets and Black shows the α- helices

- According to NMA, there exists a highly flexible region ~90th residue in the loop prior to the folding of the α-helix.
- Remarkably, there exists an expected Lysine acetylation site (Lys93), for post-translational modification, insisting responsiblity for biofilm formation in UspA from *M. luteus*.
- Single and double-lobed Usp structures share an invariant fold. Contrastingly, the highly flexible regions directs there are diverse functions for each corresponding structure.

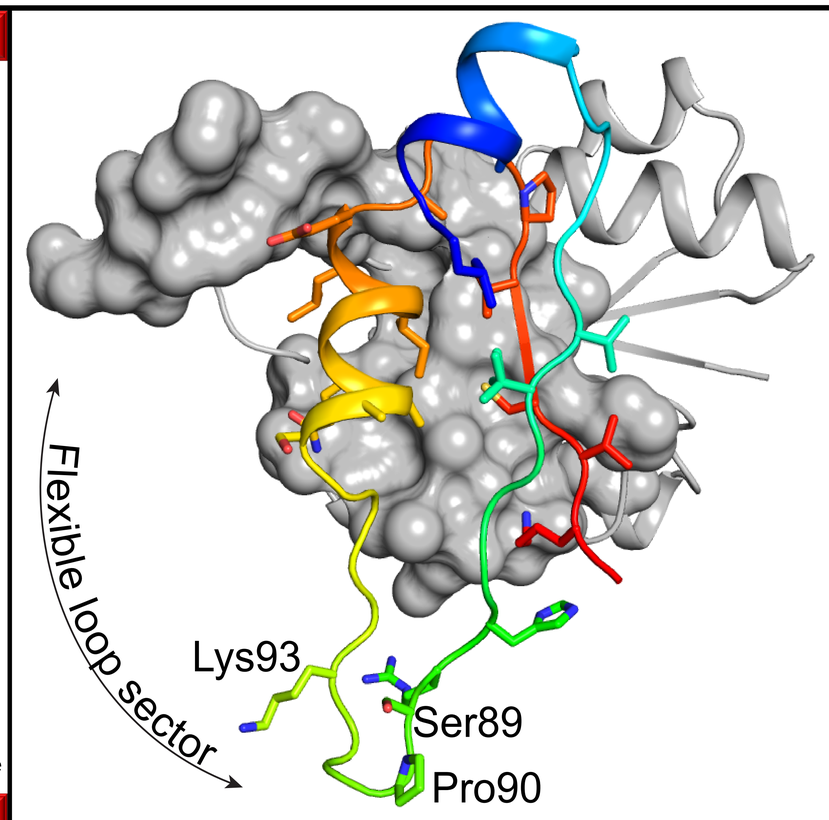

Figure 6: Structural units tuning UspA structure and function are mapped onto our crystal structure. Residues that undergo potential translational modification on the highly flexible loop are shown in stick. Co-evolving residues that are not located on the UspA core are colored in rainbow.

Discussion

- PCA demonstrates hierarchical clustering of Usp proteins based on their 3D-sequence alignment, which are conspicuously categorized into two major classes: single-lobed and homodimeric double-lobed structures.
- One highly variant section is highly flexible (NMA) and consists of post-translational modification motifs that could be pivotal in the function of each Usp protein.
- The emerging evidence displays that Usp molecules are not solely expressed under stress conditions, and their functions could be vital to bacterial growth.
- Usp molecules are emerging drug targets because they are not found in metazoans, but bacteria, fungi, and archaea.
- Usp molecules comprise the same structural scaffold, yet limited changes appear to effectively alter their function. This observation could potentially be utilized in designing new proteins with new functions, on existing protein scaffolds.

References

- Tkaczuk, Karolina L et al. "Structural and functional insight into the universal stress protein family." *Evolutionary applications* vol. 6,3 (2013): 434-49. doi:10.1111/eva.12057
- Xin-Qiu Yao, Lars Skjærven, and Barry J. Grant *The Journal of Physical Chemistry B* **2016** *120* (33), 8276-8288 DOI: 10.1021/acs.jpcb.6b01991
- Mali, Sujina, et al. "A Proteomic Signature of Dormancy in an Actinobacterium: Micrococcus luteus." *Journal of Bacteriology* (2017): JB-00206.
- Samanta, Sintu et al. "A universal stress protein in *Mycobacterium smegmatis* sequesters the cAMP-regulated lysine acyltransferase and is essential for biofilm formation." *The Journal of biological chemistry* vol. 295,6 (2020): 1500-1516. doi:10.1074/jbc.RA119.011373

Acknowledgements

This study was supported by the University of Houston:
Summer Undergraduate Research Fellowship

Contact Information: daniel.taeyounglim@gmail.com