Principal Component Analysis to Explore Transition Pathway of the Conformational Change in *E. Faecalis* Prolyl-tRNA Synthetase upon Substrate Binding

Samuel Fehling, Alexander Strom, Sudeep Bhattacharyya, and Sanchita Hati
Department of Chemistry, University of Wisconsin-Eau Claire, WI 54702

Abstract

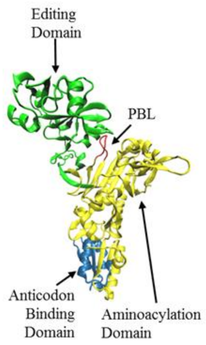
Aminoacyl-tRNA synthetases (AARSs) catalyze the esterification of transfer RNAs with their cognate amino acids. These multi-domain enzymes undergo conformational changes upon substrate binding. To understand the molecular mechanism of the population-shift from substrate–free conformations to bound conformation(s), we analyzed the substrate–bound and substrate–free conformations of *E. Faecalis* Prolyl-tRNA Synthetase (ProRS) by performing Principal Component Analysis (PCA) of the molecular dynamic simulation trajectories. In the present study, three approaches are taken for analysis: 1) PCA of individual all–atom molecular dynamics simulation trajectories for substrate–bound and substrate–free ProRS (PDB Code: 2J3M) systems, 2) PCA on the combined MD trajectories of substrate–bound and substrate–free ProRS, and 3) PCA of active site MD trajectories for each system. Our analyses of individual molecular dynamic simulations indicate substrate–bound and substrate–free principal components are intrinsically different; no conformational overlap was identified for these two systems. PCA of MD trajectories demonstrated that active site dynamics are fundamentally distinct between substrate–bound and substrate–free systems.

Aminoacyl-tRNA Synthetases

- Aminoacyl-tRNA synthetases are multi-domain proteins, which are responsible for catalyzing the aminoacylation of tRNA in a two-step reaction:

$$AA + ATP + ARS \rightarrow AA\text{-}AMP.ARS + PP_i \quad (1)$$
$$AA\text{-}AMP.ARS + tRNA \rightarrow AA\text{-}tRNA + AMP + ARS \quad (2)$$

- Many aminoacyl-tRNA synthetases are considered allosteric because tRNA binding at the anticodon domain modulates the catalytic activity of some of these enzymes.



- The anti-codon binding domain interacts with the anticodon region of the tRNA to ensure covalent attachment between the tRNA and the cognate amino acid.
- Aminoacylation domain is the location used to catalyze the aminoacylation of tRNA.
- Proline Binding Loop (PBL) prevents the spontaneous hydrolysis of adenylate intermediate.
- The editing domain hydrolyses incorrectly paired aminoacyl-tRNA molecules.

- **Two proposed models for long-range allosteric communications**
  - **The "induced-fit" model:** substrate-induced conformational change propagated through a single residue-residue interaction pathway.
  - **The "population-shift" model:** a perturbation at a distant site that alters the conformational equilibrium through "pre-existing" multiple pathways of residue-residue interactions.
  - **Pre-Existing pathways:** Communication pathways are inherent to the protein and new pathways cannot be generated. There are multiple allosteric pathways within a protein and information can only be sent through these pathways.

Hypothesis

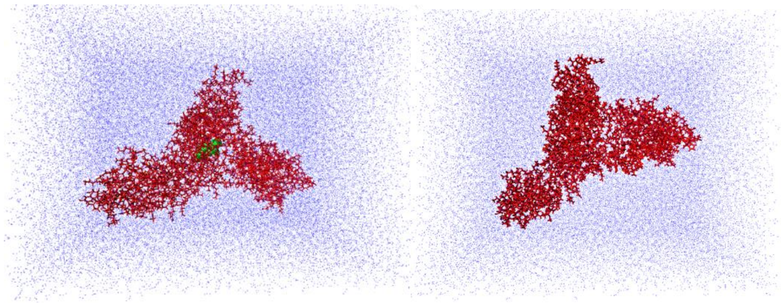
- The "population–shift" model is the shift of conformational ensemble of a protein due to binding or release of substrate/product, which drives the ensemble to altered energy conformations.
- Conformational transitions take place through well-defined residue-residue interaction networks.

Methods

Molecular Dynamic (MD) Simulation

- Figure a) illustrates substrate–bound ProRS submerged in a water box. Red represents the ProRS enzyme, blue represents the water box, and green is the bound substrate, pyridoxal phosphate (PLP).
- Figure b) illustrates substrate–free ProRS submerged in a water box. Red represents the ProRS enzyme and blue represents the water box.
- Molecular dynamic simulations were performed for 30 ns for both substrate–bound and substrate–free ProRS systems.

a) *b)*



Principal Component Analysis (PCA)

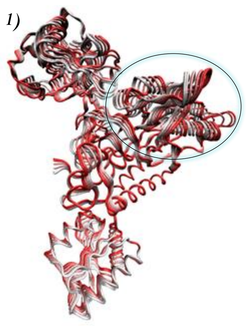
- PCA converts a set of potentially correlated variables into a set of uncorrelated variables called principal components.
- The first principal component includes the highest amount of variance.
- Variance contains data of global protein movement, the slow movement of the protein backbone.
- ProRS MD trajectories were analyzed individually and collectively.

Results

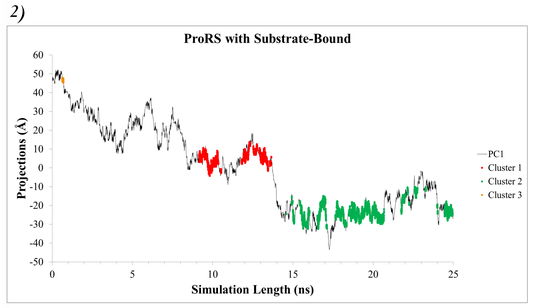
PCA of Individual All-Atom Molecular Dynamic Simulation Trajectories

- Figure 1) illustrates the substrate–bound ProRS time-lapse trajectory of PC1. Starting conformation in red and end in white. PLP is not shown.
- Figure 2) displays Projection (Å) vs. Simulation Length (ns).
- Figure 3) illustrates the substrate–free ProRS time-lapse trajectory of PC1. Starting conformation in red and end in white.
- Figure 4) displays Projection (Å) vs. Simulation Length (ns).

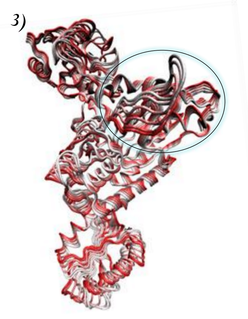
1)



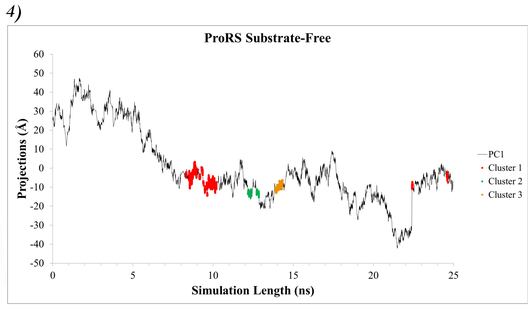
2)



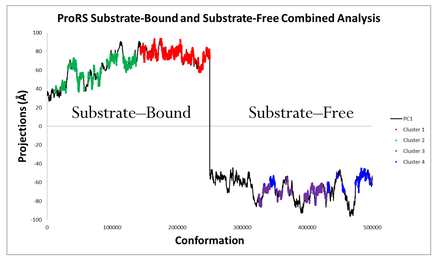
3)



4)

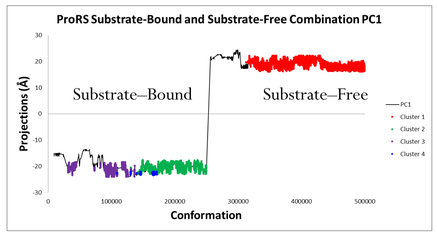


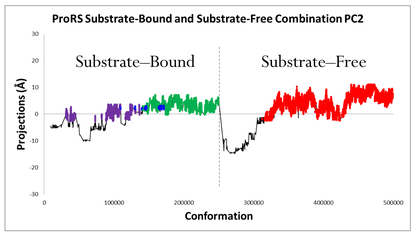
PCA of Combined MD Trajectories

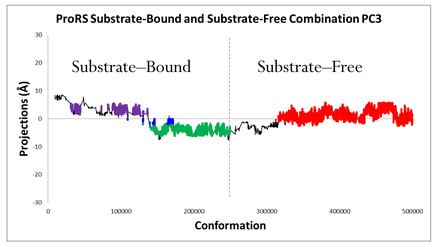


- Slower motion of substrate–bound and substrate–free systems remain inherently different along PC1 projected path.
- Clusters designate groups of conformations with similar magnitude of motion and direction from the beginning structure (conformation).

PCA of Combined Active Site MD Trajectories







- PC1 represents global protein movement - the slower fluctuations of the protein backbone.
- PC2 and PC3 represent local protein movement - the fast fluctuations of the protein backbone.

Conclusions

- PCA of individual and combined MD trajectories indicate substrate–bound and substrate–free principal components are intrinsically different; no conformational overlap was identified regarding these two systems.
- Close scrutiny of the PCA of combined MD trajectories also demonstrated that active site dynamics are fundamentally distinct between substrate–bound and substrate–free systems.

Future Directions

- Perform molecular dynamic simulations for extended period of time (100 ns) for both substrate–bound and substrate–free ProRS systems.
- Investigate transition intermediate conformations that share a bridge between substrate–bound and substrate–free systems in search of similar intermediate conformations.
- Experimental results indicate that mutations along the pre-existing pathway, N305A, F412A, and E234A, display a large impact in enzyme catalysis.
- *In silico* mutations to probe how mutations along pre-existing pathways alter substrate–bound and substrate–free conformations.

References

1. Glykos, N.M. (2006J. Comput. Chem., 27, 1765-1768
2. Glykos, Nicholas. Carma Documentation, N M Glykos Group. N.p., n.d. Web. 06 Dec. 2013.
3. Humphrey, W., Dalke, A. and Schulten, K., J. Molec. Graphics, 1996, vol. 14, pp. 33-38.
4. James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kale, and Klaus Schulten. Scalable molecular dynamics with NAMD. Journal of Computational Chemistry, 26:1781-1802, 2005. abstract, journal
5. Koukos, P.I. & Glykos, N.M. (2013), J. Comput. Chem., 34, 2310-2312
6. Ma, B., S. Kumar, C.-J. Tsai, and R. Nussinov. Protein Engineering Design and Selection 12.9 (1999)
7. Palmer, Michael. Principal Components Analysis. N.p., n.d. Web. 06 Dec. 2013.
8. Alexander M. Strom, Samuel C. Fehling, Stephanie N. Tadayon, Yer Yang, Dr. Sudeep Bhattacharyya, and Dr. Sanchita Hati., The Journal of Molecular Modeling, 2014

Acknowledgements:
- TeraGrid Resources
- Office of Research & Sponsored Programs at UWEC
- Computational resources were provided in part by the MERCURY supercomputer consortium (http://mercuryconsortium.org/) under NSF grant CHE-1229354