

Multi-scale Modeling of Enzyme Catalysis, Biomolecular Recognition and Protein Folding

Pratul K. Agarwal¹, Al Geist²

Keywords: multi-scale protein modeling, protein dynamics, enzymes, biomolecular recognition, folding, cyclophilin A, methyltransferase

1 Introduction.

Computational molecular biology continues to make significant impact on the understanding of biomolecular structure and function. Protein structure, folding, dynamics and function span multiple scales of time (see Figure 1). Experimental techniques, including NMR, X-ray crystallography, neutron scattering, and biochemical and mutational analysis, can access only a select range of time-scales. Collective information over multiple time-scales is required for gaining better understanding of many cellular processes. Obtaining much needed collective information at multiple time-scales, however, still remains a challenge.

Multi-scale modeling of protein functions and protein folding can provide insights, which complement and extend information available from experiments. However, novel theoretical and computational techniques are needed, since the present techniques and software are incapable of modeling the length and time-scales relevant for protein folding and most protein functions. We are developing new methodologies to enable multi-scale modeling in the following areas:

I. Enzyme catalysis: There is considerable interest in modeling enzyme catalysis as most cellular process involves biochemical reactions in presence of enzymes. Proteins enzymes are highly efficient catalysts. For more than a century, biochemists have known the importance of enzyme structure. However, increasing evidence now continues to reveal that proteins are dynamically active assemblies with a link between structure and dynamics [1-3]. Multi-scale modeling of enzyme catalysis will provide insights into the role of dynamical events in the enzyme complex on the reaction kinetics.

II. Biomolecular Recognition: Many cellular processes (including transcription, DNA replication/repair) involve proteins specifically recognizing DNA sequences or other proteins. Multi-scale modeling will investigate the biophysical factors which influence biomolecular recognition.

III. Protein folding: Events in protein folding range from picoseconds to seconds (and longer). The fundamental mechanism by which an extended protein folds into the unique 3-dimensional structure still remains an unsolved problem. Multi-scale modeling of the key events in this process will make an impact on predicting structure of a protein from its primary sequence.

2 Methods.

A variety of theoretical and computational techniques have been used to perform multi-scale modeling of protein complexes. The techniques include multiple sequence and protein structure alignment, molecular modeling (molecular dynamics), dynamic cross-correlation, normal mode analysis and quasi-harmonic analysis. Further, we have developed a novel algorithm to simulate the process of protein folding on next generation of supercomputers.

¹ Computational Biology Institute and Computer Science and Mathematics Division, Oak Ridge National Laboratory, Bethel Valley Road, Oak Ridge, TN USA. E-mail: agarwalpk@ornl.gov

² Computational Biology Institute and Computer Science and Mathematics Division, Oak Ridge National Laboratory, Bethel Valley Road, Oak Ridge, TN USA. E-mail: gst@ornl.gov

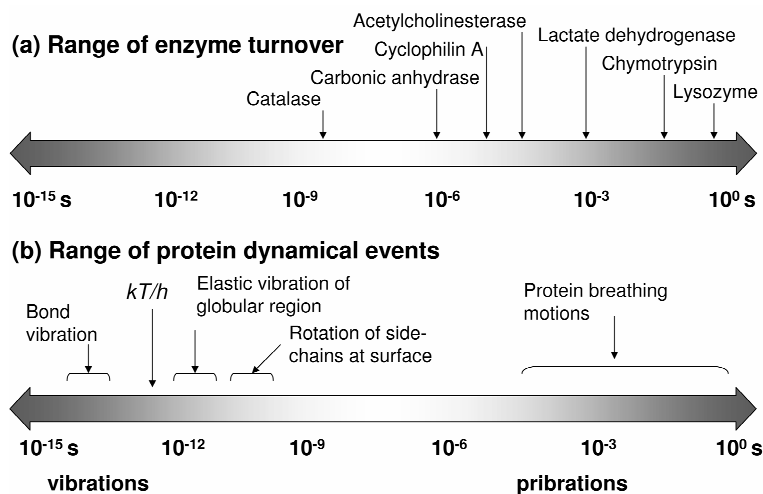


Figure 1: Protein function (such as enzyme catalysis) and protein dynamical events involve multiple time-scales.

3 Results.

Multi-scale modeling of protein functions including enzyme catalysis, biomolecular recognition and protein folding has provided novel insights into the link between protein structure, dynamics and function. Investigations of the peptidyl-prolyl cis/trans isomerization catalyzed by enzyme cyclophilin A reveals the presence of a network of protein vibrations, which extends from surface loop residues of the enzyme all the way to the active site [1-2]. The dynamical events in this network occur on the time-scales ranging from picosecond to microsecond-millisecond (time scale of the reaction). Genomic and structural analyses indicated that the residues and hydrogen-bonds forming crucial points in the network are conserved in several cyclophilin structures from species ranging from bacteria to human. This network is a conserved part of the enzyme structure and promotes catalysis by altering the transition state barrier recrossing [3]. Investigations of biomolecular recognition of DNA with methylated cytosine by m5c-methyltransferase *M. HhaI* indicates that dynamical events in protein-DNA complexes may play a role in recognition in protein complexes [4]. Further, the insights into dynamical events associated with a protein in solution have been used to design a novel algorithm to simulate the process of protein folding. This algorithm is based on overcoming local energy barriers and provides insights into the mechanism of protein folding a naturally occurring protein, Ubiquitin [5].

References

- [1] Agarwal, P. K., Geist, A and Gorin A. 2004. Protein Dynamics and Enzymatic Catalysis: Investigating the Peptidyl-Prolyl cis/trans Isomerization Activity of Cyclophilin A, *Biochemistry* **43**, 10605-10618.
- [2] Agarwal, P. K. 2004. Computational studies of the mechanism of cis/trans isomerization in HIV-1 catalyzed by cyclophilin A, *Proteins: Structure Function and Bioinformatics* **56**, 449-463.
- [3] Agarwal, P.K. 2005. Protein Dynamics Plays a Role in Rate Enhancement by Enzymes. *Submitted to Proceedings of the National. Academy of Sciences USA*.
- [4] Agarwal, P. K., Geist A., Yalla S.K. The role of dynamics in biomolecular recognition in protein-DNA complexes. *Manuscript under preparation*.
- [5] Agarwal, P.K. and Geist A., 2005. Designing a novel algorithm for simulating protein folding on next generation supercomputers, *Submitted to Journal of Parallel and Distributed Computing*