An Integrative Strategy to Model Complex Biological Assemblies

Ivaylo Ivanov, Georgia State University

Associated NERSC Project: Modeling Assemblies and Interactions in Eukaryotic Clamp Loading (m1254)

NISE Award: 960,000 Hours
Award Date: June 2011

Conventional structural biology methods have provided valuable but fragmentary insight into the architecture of the ternary assemblies of DNA with the proteins FEN1 and PCNA. These assemblies play a crucial role in replication - the common mechanism through which all organisms duplicate their genetic material. Our computational work would take advantage of state-of-the-art molecular modeling methods to elucidate the structure and dynamics of these important replication complexes.

DNA replication is accomplished by a dynamic protein assembly termed the replisome. The central goal of our work is to develop novel multiscale computational protocols, which would allow integrative modeling of protein-nucleic acid complexes within the replisome. Understanding the inner workings of this complex molecular machine is undeniably among the great challenges in the biological sciences. Breakthroughs in this field are expected to yield great advances in biomedicine specifically in the diagnosis, prognosis and treatment of cancer and genetic disorders. Furthermore, the fundamental biological mechanisms and the dynamic protein associations in DNA synthesis and repair are conserved across bacterial, archaeal and eukaryotic organisms and govern many metabolic processes occurring in response to environmental stress as cells strive to regain homeostasis. In this respect, our work also parallels ongoing effort at Lawrence Berkeley National Laboratory (LBNL) in microbial genomics aimed at unraveling the functions of microorganisms that are environmentally important or related to DOE’s bioenergy initiatives. Microbial genomes and metabolic responses to environmental stresses such as ultraviolet radiation, ionizing radiation and oxidative stress are of direct relevance to the mission of the DOE. Our project is also well aligned with the stated goals of the DOE Biological Systems Science division “to develop the computational capabilities and systems needed to predictively design and model complex biological systems” and “to develop and support DOE national user facilities for use in fundamental structural biology”.

Many individual components of the cell’s replication machinery have already been solved by protein crystallography. How these components come together to form a functioning molecular machine has, however, remained elusive. Furthermore, detailed structural characterization of large protein and nucleic acid assemblies is often impossible by any single experimental or computational method. Therefore, advancement of this field would necessarily involve integration of knowledge from a variety of sources – biochemical and biophysical experiments as well as molecular simulation and structural bioinformatics. For instance, electron microscopy (EM), small angle X-ray scattering (SAXS), fluorescence (FRET), footprinting, chemical cross-linking are among a number of methods, which could collectively be used to overcome the limitations inherent in static crystallographic structures and probe the assembly and conformational dynamics of macromolecular complexes. Each of these techniques comes with its own limits of accuracy, applicability and spatial/temporal resolution. Achieving a consistent molecular view will, therefore, require advanced computational methods in order to integrate experimental data from vastly disparate sources. A central goal of our work is to develop such novel multiscale computational protocols and apply them to replisomal complexes.

Within the replisome, replicative polymerases rapidly and faithfully duplicate the cell’s genetic material prior to cell division. Proliferating Cell Nuclear Antigen (PCNA) serves as an accessory protein whose role is to topologically tether these polymerases to DNA and ensure processive replication. Beyond involvement in replication, PCNA is also a recognized master coordinator of cellular responses to DNA damage and interacts with many DNA repair proteins and cell cycle checkpoint inhibitors. In this capacity, PCNA serves not only as a mobile platform for the attachment of these proteins to DNA but, importantly, plays an active role in the recruitment and release of these crucial participants at the replication fork. These are key processes in PCNA biology, which are incompletely understood from a structural point of view. This proposal puts forward an integrative strategy to model the ternary assemblies of Flap-endonuclease 1 (FEN1) with its double flap DNA substrate and with PCNA (or with the alternative heterotrimeric sliding clamp Rad9-Rad1-Hus1 (9-1-1)). While structural snapshots are available for the individual components of these assemblies and, in some cases, for the binary complexes (PCNA-DNA, PCNA, FEN1-DNA, 9-1-1 complex), the larger assemblies present extreme challenges to molecular crystallography (MX) and will require integration of data from a variety of sources (EM, SAXS, MX) through advanced computational methods. A modular multi-scale approach will be adopted wherein (i) large conformational ensembles for the complexes are generated at first through advanced sampling methods and (ii) these ensembles are systematically reduced to eliminate conformations incompatible with existing experimental constraints. The final refined models will be subject to experimental validation and guide future experimental efforts in this exciting research area. The exceptional computing resources available at NERSC will create opportunities for this work on the cutting edge between computation and structural biology that will lead to deeper understanding of the inner workings of biological assemblies engaged in replication.

TABLE OF CONTENTS