GA

BW

EXPLORING THE STRUCTURE AND DYNAMICS OF CONVERGED ENSEMBLES OF DNA AND RNA THROUGH MOLECULAR DYNAMICS SIMULATIONS

Allocation: NSF PRAC/12,000 Knh
PI: Thomas E. Cheatham III[1]
Co-PIs: Adrian Roitberg[2], Carlos Simmerling[3], David Case[4]
Collaborators: Rodrigo Galindo-Murillo[1], Christina Bergonzo[5], Daniel R. Roe[6]

[1]University of Utah
[2]University of Florida
[3]Stonybrook University
[4]Rutgers University
[5]Institute for Bioscience and Biotechnology
[6]National Heart, Lung, and Blood Institute

EXECUTIVE SUMMARY

Over the past two years of using Blue Waters and taking full advantage of its computational power, we have made extensive progress in the realm of biomolecular simulation methodologies, specializing in nucleic acid structure, dynamics, and ligand–protein binding. The AMBER package for biomolecular simulation and its GPU (graphics processing unit) code has proven to be high-performance and reliable, taking full advantage of Blue Waters. As our group continues to develop specialized software and methodologies to analyze the vast amount of sampling information, we further increase our understanding of relevant biological DNA, RNA, and protein structures.

RESEARCH CHALLENGE

Molecular Dynamics (MD) simulations have been one of the most important tools in the computational chemist's toolbox for the last 25 years. Useful as it is, this technique has considerable limitations, mainly in two areas: force field validation and conformational sampling. These two problems are deeply related: As more sampling time is achieved, force field discrepancies are found. The force field is temporarily fixed; however, as computational power grows, more sampling time is achieved and new force field issues are found. Blue Waters has enabled our group to achieve enough sampling time to rigorously validate and assess protein, DNA, and RNA force fields in order to expose existing drawbacks.

METHODS & CODES

Blue Waters provides the necessary computing power to perform tests and benchmarks of several simulation methodologies. This allows us to increase the amount of space sampled for a particular biomolecular system and to create new enhanced sampling techniques in order to obtain a converged ensemble. Our group is focusing on two areas: multi-dimensional replica exchange (M-REMD) and ensemble simulations. In the case of replica exchange, we have explored multiple small RNA systems (tetranucleotides, hairpins, loops, etc.) and we have achieved a converged ensemble that generated insight into successes and failures of force fields, ion models, water models, and modeling procedures. This information helps to pinpoint problem areas in the models used and to guide the next steps of research.

The other methodology we have used involves multiple independent copies of a particular biomolecular system, or ensemble simulations. This lets us explore increasing sampling space without the introduction of any biasing or enhanced mechanism. These simulations allow us to study the process of the DNA–ligand binding mechanism in fully atomistic ways, which provide further insight in order to design novel small molecules to increase biological activity. Of major interest is the study of a family of planar copper-compounds (with general formula $[Cu(N–N)(N–O)]NO^3$ and $[Cu(N–N)(O–O)]NO^3$; where the N–N ligand denotes either 2,2'-bipyridine or 1,10-phenanthroline (the aromatic ligand); N–O represents an essential amino acid or peptides; and the O–O represents a nonaromatic ligand—either acetylacetonate or salicylaldehydate) that experimentally show higher biological activity at a lower dosage with respect to the hallmark of transition-metal drugs: cisplatinum. Using Blue Waters and several milliseconds of sampling time, we observed five principal binding sites (Fig. 1). Binding modes a) through e) represent the result of unbiased interactions of the ligands with the Drew Dickerson dodecamer, with the sequence GCGCAATTGCGC, using the bsc0 force field for nucleic acids. Binding mode a) represents stacking on the edges of the DNA, modes b) and d) are where the ligand binds into the minor groove of the helix. Mode c) starts in the minor groove and then the ligand moves into an AT base pair, pushing the AT bases toward the major groove as the ligand slides into the resulting cavity. Mode e) is an intercalated mode as a result of the terminal base pairs fraying, allowing the ligand to stack into the exposed bases.

RESULTS & IMPACT

Key results are described in detail in our publications (below).

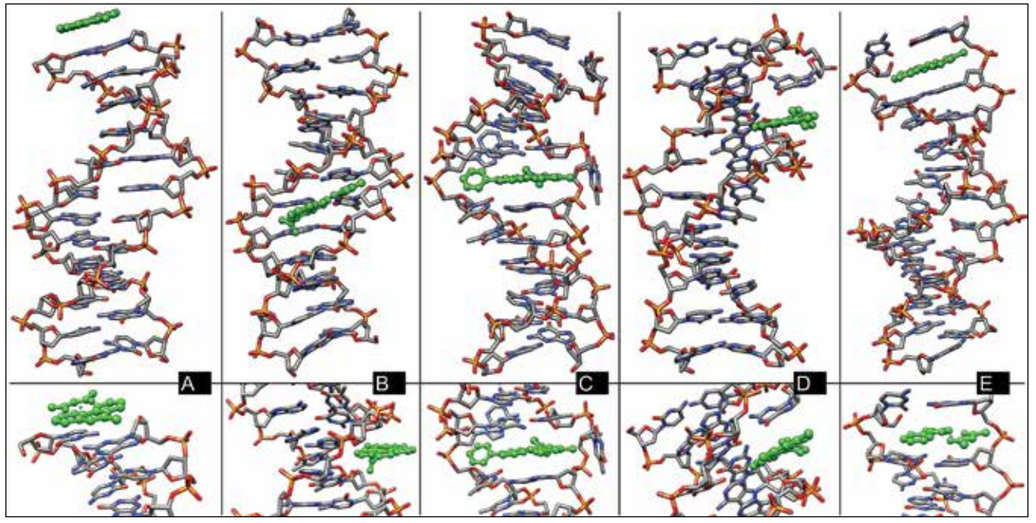

Figure 1: Modes of binding between a 12-mer double-stranded DNA and a planar copper complex—a) stacking at the end of the DNA; b) and d) minor groove interaction; c) base-pair eversion mechanism; and e) intercalation at the terminal base pairs.

WHY BLUE WATERS

Because Blue Waters is optimized for GPU-centered calculations, as well as massive CPU parallelization, it has been a key tool for our group and allowed us to independently run hundreds of biomolecular simulations of myriad systems of DNA and RNA. In some cases, we have reached well over millisecond timescales of combined sampling time. In addition, the highly trained Blue Waters staff have been in close collaboration to develop custom and specialized code capable of performing the analysis of terabytes of simulation data.

PUBLICATIONS AND DATA SETS

http://amber.utah.edu

Galindo-Murillo, R., et al., Intercalation processes of copper complexes in DNA. *Nuc. Acids Res.*, 43 (2015), pp. 5364–5376.

Galindo-Murillo, R., et al., Assessing the current state of Amber force field modifications for DNA. *J. Chem. Theory Comp.*, 12 (2016), pp. 4114–4127, DOI: 10.1021/acs.jctc.6b00186.

Heidari, Z., et al., Using Wavelet Analysis to Assist in Identification of Significant Events in Molecular Dynamics Simulations. *J. Chem. Inf. Model*, 56 (2016), pp. 1282–1291.

Bergonzo, C., and T. E. Cheatham, III, Mg^{2+} binding promotes SLV as a scaffold in Varkud Satellite Ribozyme SLI-SLV kissing loop junction. *Biophys. J.*, 113 (2017), pp. 313–320, DOI: 10.1016/j.bpj.2017.06.008.