Bioinformatics, 35(18), 2019, 3508–3509
doi: 10.1093/bioinformatics/btz107
Advance Access Publication Date: 13 February 2019
Applications Note

Structural bioinformatics

Biomolecular Reaction and Interaction Dynamics Global Environment (BRIDGE)

Tharindu Senapathi[1], Simon Bray[2], Christopher B. Barnett[1], Björn Grüning[2,3,*] and Kevin J. Naidoo [1,4,*]

[1]Scientific Computing Research Unit and Department of Chemistry, University of Cape Town, Rondebosch 7701, South Africa, [2]Department of Computer Science, University of Freiburg, 79110 Freiburg, Germany, [3]Center for Biological Systems Analysis (ZBSA), University of Freiburg, 79104 Freiburg, Germany and [4]Institute for Infections Disease and Molecular Medicine, Faculty of Health Science, University of Cape Town, Rondebosch 7701, South Africa

*To whom correspondence should be addressed.
Associate Editor: Alfonso Valencia

Received on October 8, 2018; revised on January 14, 2019; editorial decision on February 7, 2019; accepted on February 11, 2019

Abstract

Motivation: The pathway from genomics through proteomics and onto a molecular description of biochemical processes makes the discovery of drugs and biomaterials possible. A research framework common to genomics and proteomics is needed to conduct biomolecular simulations that will connect biological data to the dynamic molecular mechanisms of enzymes and proteins. Novice biomolecular modelers are faced with the daunting task of complex setups and a myriad of possible choices preventing their use of molecular simulations and their ability to conduct reliable and reproducible computations that can be shared with collaborators and verified for procedural accuracy.

Results: We present the foundations of Biomolecular Reaction and Interaction Dynamics Global Environment (BRIDGE) developed on the Galaxy platform that makes possible fundamental molecular dynamics of proteins through workflows and pipelines via commonly used packages, such as NAMD, GROMACS and CHARMM. BRIDGE can be used to set up and simulate biological macromolecules, perform conformational analysis from trajectory data and conduct data analytics of large scale protein motions using statistical rigor. We illustrate the basic BRIDGE simulation and analytics capabilities on a previously reported CBH1 protein simulation.

Availability and implementation: Publicly available at https://github.com/scientificomputing/BRIDGE and https://usegalaxy.eu

Contact: kevin.naidoo@uct.ac.za or bjoern.gruening@gmail.com

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Molecular mechanisms underlie biological phenomena. Consequently, locating molecular modeling tools within a genomics research platform consolidates the three components of the bioinformatics ecosystem that enables a seamless progression from (i) DNA, RNA or protein sequence analysis to (ii) gene expression profiling, metabolic and functional pathway analysis to (iii) molecular structural analysis of proteins—identification of therapeutic targets, development of biomarkers and examination of protein alterations. It is this intention to link genomic analytics to molecular simulations that motivates our development, reported here, of the Biomolecular Reaction and Interaction Dynamics Global Environment (BRIDGE) within the Galaxy platform (Afgan *et al.*, 2018; Goecks *et al.*, 2010). BRIDGE is a web server, based on the Galaxy framework, to perform molecular dynamics (MD) simulations of biomolecules and conduct statistical analyses on the trajectory data produced.

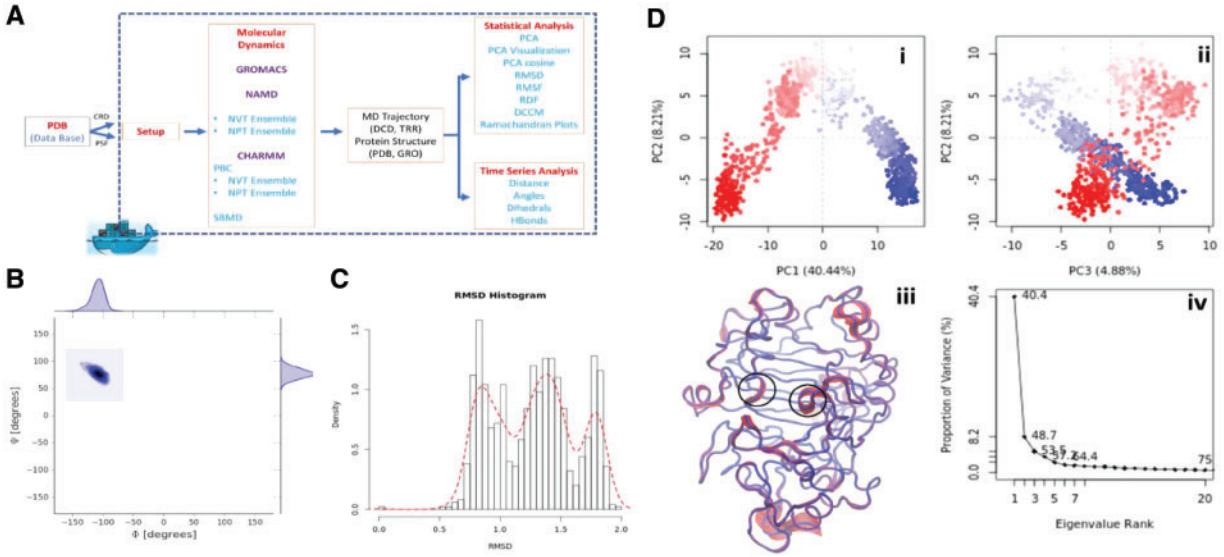

Fig. 1. (**A**) Overview of the tools to run MD and analyze in the BRIDGE MD platform (https://github.com/scientificomputing/BRIDGE). (**B**) Ramachandran plot of the cellulose ligand glycosidic bond that is targeted for hydrolysis. (**C**) Histogram of the RMSD of Cα atoms of the protein. (**D**) PCA of the Cα atoms of the protein backbone motions

Protein functions are carried out by atomistic scale binding interactions of their residues in catalytic/binding domains in the case of enzymes/lectins/antibodies and conformational changes of their backbones and loops that regulate transport, folding, etc. (Fersht, 1999). It is now universally accepted that probing the molecular mechanisms of reaction and interactions are best made using MD simulations.

The Galaxy software platform is an open-source platform that has historically focused on sequencing analysis of all kinds (Goecks *et al.*, 2010), but now is a general framework for data analysis, functioning beyond life sciences. Here we add molecular modeling functionalities to Galaxy, specifically the capability to perform (i) MD simulations, (ii) statistical mechanics analyses (PDFs, time correlation functions, etc.) on dynamics trajectories and (iii) statistical analysis and big data analytics on individual biomolecules or families of biomolecular structures and configurations.

A set of Galaxy tool wrappers (Fig. 1A) have been developed to set up and run classical MD simulations using either one (or all) of the CHARMM (Brooks *et al.*, 2009), NAMD (Phillips *et al.*, 2005) and GROMACS (Lindahl *et al.*, 2001), MD engines. Following this, wrappers were developed to include the features of MDAnalysis (Gowers *et al.*, 2016) for subsequent statistical mechanics analysis and finally wrappers for the Bio3D (Skjærven *et al.*, 2014) package were written to make the statistical analysis of structural/conformational biomolecular motions produced from trajectories possible.

2 Demonstration and conclusion

We illustrate some of the analytical tools able to investigate conformational changes by analysis of a typical short protein simulation such as for CBH1 (see Supplementary Material for details). The Ramachandran style dihedral angle plot of a key glycosidic linkage of the oligosaccharide ligand is computed using the Ramachandran Plots tool (Fig. 1B). The protein motion is analyzed using the root mean square deviation (RMSD) tool. Three distinct conformations around RMSD of 0.8, 1.2 and 1.8 Angstrom can be seen from the RMSD histogram (Fig. 1C).

Protein conformational changes can be investigated in greater detail using tools in the statistical analyses module. Here PCA was used to discover the statistically meaningful conformations in the 5 *ns* CBH1 trajectory (Fig. 1D). The principal motions within the trajectory and the vital motions needed for conformational changes were identified. Two distinct groupings along the dominant PC1 plane (Fig. 1Di and iv) indicating a non-periodic conformational change are identified. The groupings along the PC2 and PC3 planes (Fig. 1Dii) do not completely cluster separately implying that these global motions are periodic. The PC1 is linked to an active site motion (Fig. 1Diii) that limits the motion to a key glycosidic bond (Fig. 1B).

Funding

This work was supported by the South African Research Chairs Initiative (SARChI) of the Department of Science and Technology (DST) and National Research Foundation (NRF) [grant 48103].

Conflict of Interest: none declared.

References

Afgan,E. *et al.* (2018) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.*, **46**, W537–W544.

Brooks,B.R. *et al.* (2009) CHARMM: the biomolecular simulation program. *J. Comput. Chem.*, **30**, 1545–1614.

Fersht,A. (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*. W.H. Freeman, New York.

Goecks,J. *et al.* (2010) Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.*, **11**, R86.

Michaud-Agrawal,N. *et al.* (2011) MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **32**, 2319–2327.

Lindahl,E. *et al.* (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. *Mol. Model. Annu.*, **7**, 306–317.

Phillips,J.C. *et al.* (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, **26**, 1781–1802.

Skjærven,L. *et al.* (2014) Integrating protein structural dynamics and evolutionary analysis with Bio3D. *BMC Bioinformatics*, **15**, 399.