

PRIMoRDiA Meets EasyHybrid3: Unlocking the Exploration of the Electronic Structure of Biomolecules Through Hybrid QC/MM Simulations On the Fly.

Igor Barden Grillo,[†] José Fernando Ruggiero Bachega,[‡] Martin J. Field,[¶] Larissa Pereira,[†] and Gerd Bruno Rocha[†]

[†]*Departament of Chemistry of Federal University of Paraíba, João Pessoa - PB, Brazil*

[‡]*Department of Pharmacosciences, Federal University of Health Sciences of Porto Alegre, Centro Histórico, Porto Alegre- RS,90050-170, Brazil.*

[¶]*Laboratoire de Chimie et Biologie des Métaux, UMR5249, Université Grenoble I, CEA, CNRS, 17 avenue des Martyrs, 38054 Grenoble Cedex 9, France*

[§]*Theory Group, Institut Laue-Langevin, 71 avenue des Martyrs CS 20156, 38042 Grenoble Cedex 9, France*

E-mail:

Abstract

Hybrid simulations are powerful computational chemistry tools to obtain structural and thermodynamics information about enzymatic catalysis reactions. In addition to such relevant applications, these calculations also provide the resolved electronic structure, which is required to estimate the electronic energy in the SCF process. Recently, our group is using such information to obtain several quantum chemical descriptors to profile the steering forces underlying biological processes. In our group's recent main recent publications we show: how semiempirical Hamiltonian could be used to obtain the Fukui functions from polypeptides with modified descriptors, applying them to ligand-binding process; to profile enzymatic catalysis reaction coordinates; to study the main interactions on the active site of main protease SARS-COV-II, comparing with from the SARS-COV-I. The main goal of this work is to present our software PRIMoRDiA, which is a new application written to deal efficiently with large outputs of quantum chemical programs and to calculate more than thirty global and local quantum chemical descriptors. The new version of the program, is planned to work directly with the electronic structure obtained with pDynamo software through EasyHybrid 3.0 interface, allowing such theoretical analysis on the fly for hybrid QC/MM simulations.

Introduction

The developing of a computational protocols such the quantum chemical/quantum mechanical hybrid schemes, allowed the study of complex process that occurs in large biomolecular systems. This turned the Hybrid QC/MM methods the state of art and it comes to the energetic and structural studies of enzymatic catalysis. Although, in the great majority of those simulations the resolved electronic structure is throw away, i.e. the wave function information is not utilized even though the computational cost to obtain it already was paid.

PRIMoRDiA software is a new application released to the community prepared to deal with large output files from quantum chemical packages, transforming the electronic structure

information in quantum chemical descriptors, global and local, that describes the reactivity propensity and/or provide some electronic properties that can be used to train correlation models.

We already used PRIMoRDiA in some studies for profile and characterize reaction paths of enzymatic reactions, from trajectories generated with QC/MM simulations.

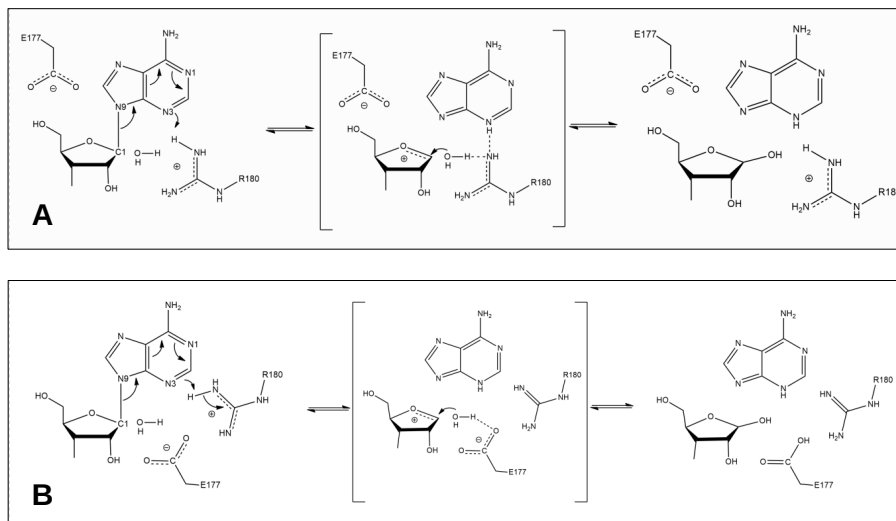


Figure 1

Methods

the system was prepared for QC/MM simulations with the following steps: i) parametrization of the cofactor NADP; ii) parametrization of the substrate; ii) solvation of the complex, ions addition to neutralize the total charge and protein parameters generation with AMBER force field; iv) GROMACS energy minimization of the entire system. After these steps, the force field parameters and coordinates files were loaded in the pDynamo library, where the system were prepared, QC/MM atoms, and fixed atoms were defined and the further simulations were performed.

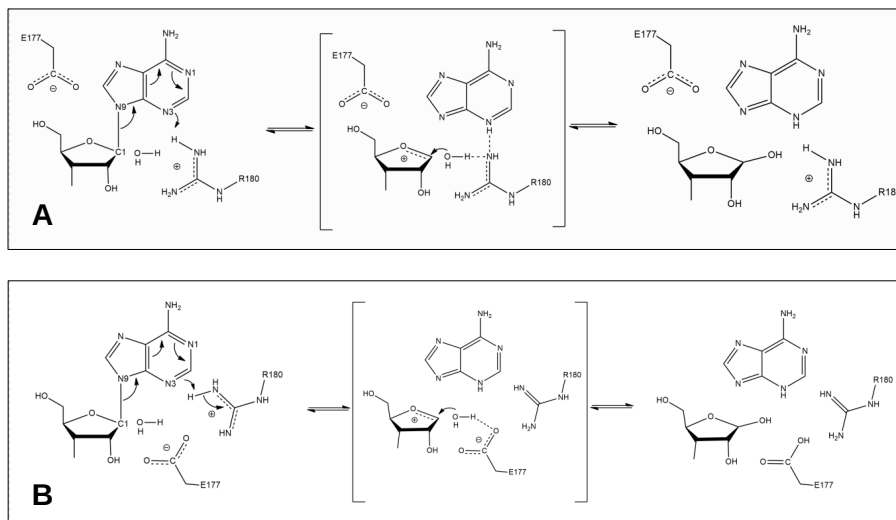


Figure 2

Structure Preparation and Parametrization

the library files for tleap were generated with parmchk2 from ambertools. The glucose-6-phosphate parameters were generated with am1-bcc charge scheme using antechamber program considering -2 of formal charge. We used the amberff99SB force field for the protein atoms, tip3p model to build a solvation cubic box respecting the 12Å between the surface of protein complex until the box end. Also, seven sodium ions were added to neutralize the total system charge.

We build the complex parameters and coordinates using tleap program, the parameters files were converted to GROMACS format using the Parmd library for Python 3, for energy minimization. We performed this minimization using steepest descent algorithm, periodic boundary conditions in all dimensions, 50000 maximum number of steps and 1000 kJ/mol/nm force tolerance and 0.01 of energy step.

QC/MM Simulations Set Up

The subsequent coordinates and the force field parameters were loaded and treated by the library classes and functions of pDynamo¹ and EasyHybrid². Firstly, the box was set to a sphere, taking the hydride (H1) as the center and including all atoms of the residues with up to 30 Å of distance. The atoms of the residues within 22 Å from the center was considered as mobile atoms and beyond as fixed, i.e., they only interact with the other atoms but without to get their coordinates updated. This avoids unnecessary gradient calculations without losing the effect of this outer layer. The distances were chosen to consider the NADP molecule present in the other binding pocket, as well to provide a minimum water molecules layer. After that, a geometry optimization was performed using pDynamo conjugated gradient minimizer with a $0.1 \text{ kJ.mol}^{-1}.\text{Å}^{-1}$ for root mean square tolerance.

The most critical definition is the atoms that will be treated with the quantum chemistry potential. An unnecessary inclusions of atoms, in respect to the description of the reactions, could lead to unpractical simulations times for energy calculations employing DFT functionals or further free energy evaluations. Thus, we included only the most important moieties to the proposed catalysis mechanism, which includes: the nicotinamide ring atoms from NADP; all glucose-6-phosphate atoms; the imidazole ring atoms of the catalytic histidine-337 and the near histidine-175; amine group of the lysine-145 and the carbon bind with the nitrogen, which is a residue very near the oxygen O1 of the substrate; and finally the carboxylic oxygens and carbon of the aspartate-174, that could activate the HIS237. The pruned system and division set up is shown in the ??.

Reactivity Descriptors

Results

QC/MM Molecular Dynamics Sampling

Reactivity Descriptors

Statistical Analysis

Structural and Reactivity Analysis

Conclusions

References

- (1) Field, M. J. The pDynamo program for molecular simulations using hybrid quantum chemical and molecular mechanical potentials. *Journal of Chemical Theory and Computation* **2008**, *4*, 1151–1161.
- (2) Bachega, J. F. R.; Timmers, L. F. S.; Assirati, L.; Bachega, L. R.; Field, M. J.; Wymore, T. GTKDynamo: A PyMOL plug-in for QC/MM hybrid potential simulations. 2013.