Journal of Molecular Graphics and Modelling 60 (2015) 15–23

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Qgui: A high-throughput interface for automated setup and analysis of free energy calculations and empirical valence bond simulations in biological systems

Geir Villy Isaksen[a], Tor Arne Heim Andberg[a], Johan Åqvist[b], Bjørn Olav Brandsdal[a,*]

[a] The Centre for Theoretical and Computational Chemistry, Department of Chemistry, University of Tromsø, N9037 Tromsø, Norway
[b] Department of Cell and Molecular Biology, Biomedical Center, Uppsala University, Box 596, S-75124 Uppsala, Sweden

ARTICLE INFO

Article history:
Received 12 February 2015
Received in revised form 13 May 2015
Accepted 14 May 2015
Available online 21 May 2015

Keywords:
Thermodynamic activation parameters
Free energies
Q
Empirical valence bond
Linear interaction energy
Molecular dynamics
Free energy perturbation

ABSTRACT

Structural information and activity data has increased rapidly for many protein targets during the last decades. In this paper, we present a high-throughput interface (Qgui) for automated free energy and empirical valence bond (EVB) calculations that use molecular dynamics (MD) simulations for conformational sampling. Applications to ligand binding using both the linear interaction energy (LIE) method and the free energy perturbation (FEP) technique are given using the estrogen receptor (ERα) as a model system. Examples of free energy profiles obtained using the EVB method for the rate-limiting step of the enzymatic reaction catalyzed by trypsin are also shown. In addition, we present calculation of high-precision Arrhenius plots to obtain the thermodynamic activation enthalpy and entropy with Qgui from running a large number of EVB simulations.

1. Introduction

Computer modeling and simulations have now become important tools to study complex systems in many areas of chemistry and biology. With the tremendous increase in speed and availability of high performance computers it is nowadays possible to perform quantitative calculations on highly complex systems such as the ribosome [1,2] and virus capsids [3–5]. Biochemical topics where computational modeling is frequently used to obtain novel insights include among others ligand binding and design, folding of proteins, enzyme catalysis and protein-membrane interactions. The ability to accurately calculate free energies enables us to characterize the structure and energetics of molecular complexes, and is often the key to understand many biological functions. In this respect, it is especially worth emphasizing that energetics often provides the most important and useful link between structure and function of biomolecular systems.

Molecular dynamics simulations represent an efficient way to sample the thermally accessible region of conformational space with microscopic models of molecular systems. Together with the potential energy (given by the force field), the ensemble of sampled structures form the basis for further calculations of, for example, free energies. Thermodynamic and kinetic experiments provide valuable data in terms of binding free energies, solvation energies and activation free energies, which are directly comparable to results obtained with simulations techniques. It thus becomes possible not only to verify and validate calculated properties with experiments, but also to make predictions that can be examined with experiments.

Historically, classical MD simulations to calculate free energies have been a low-throughput technique. This picture is now changing mainly due to advancements in algorithms and the scaling performance of software together with the increase in computer power witnessed during the last decade. Several different methods to compute free energies that rely on MD for conformational sampling are presently available [6–11]. These methods are generally more complicated to apply when compared to for example simplified ligand docking and scoring approaches, especially when the ligands are structurally dissimilar [12,13]. Computational tools that simplify the use of conformational sampling techniques and make the entire process of computing free energies more efficient are therefore desirable. Graphical user interfaces have the potential not only to lower the barrier for new users, but also of making the process of calculating free energies more efficient.

* Corresponding author. Tel.: +47 77 64 40 57; fax: +47 77 64 47 65.
E-mail address: bjorn-olav.brandsdal@uit.no (B.O. Brandsdal).

http://dx.doi.org/10.1016/j.jmgm.2015.05.007
1093-3263/

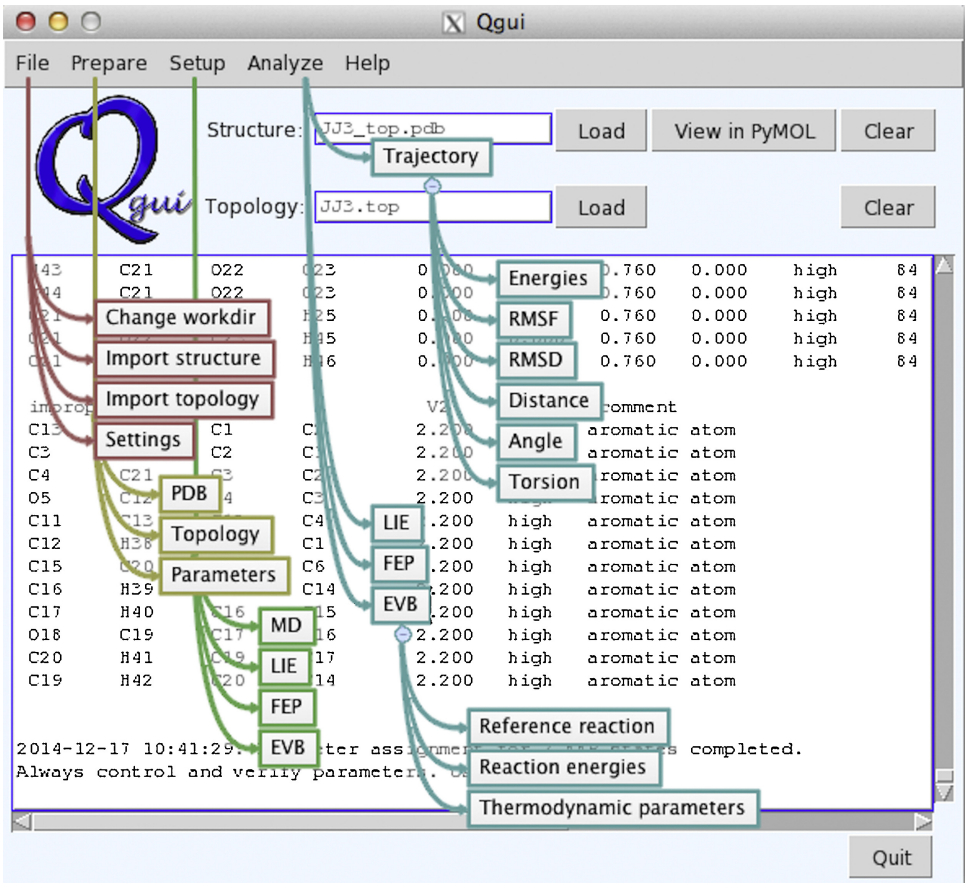

Fig. 1. The main window of the graphical user interface to Q (Qgui) with illustrated overview of dropdown menus and their contents.

Another issue related to computations of free energies is how to sample phase space most efficiently, and to ensure that the relevant parts are really captured by the simulations. It has become standard practice to conduct several replicas of the simulations to ensure that statistically significant results with adequate sampling of the relevant parts of phase space are achieved. This is typically done by repeating the simulations with different initial conditions, which may be done by assigning different initial velocities or different starting coordinates when running the MD simulations. The main challenge with this procedure is management and handling of the enormous amount of data generated.

Historically, most studies have focused on how to predict and calculate the Gibbs free energy (ΔG), as direct calculation of enthalpies and entropies involves large energy contributions. While the free energy typically converges within reasonable simulation time, obtaining converged enthalpic and entropic contributions pose a major computational challenge. Nonetheless, knowledge about the enthalpic and entropic contributions is desirable and provides important chemical information about molecular systems. We have recently shown that not only the activation free energy, but also the activation enthalpy and entropy can be calculated for enzymatic reactions using the EVB approach [14,15]. Of particular relevance here is determination of high precision computational Arrhenius plots in order to obtain thermodynamic activation parameters [16,17]. This was recently illustrated by Isaksen et al. [17] with more than 3000 independent simulations yielding a total simulation time of ~2 μs. When high precision is needed, sufficient sampling becomes critical. The development of a graphical user interface was initially driven by our need for an efficient way of setting up free energy simulations and to analyze large amounts of data produced using the MD program Q [18]. Obtaining reliable and converged reaction free energies typically involves averaging over several hundreds of MD trajectories. Furthermore, in order to obtain reliable thermodynamic activation parameters it is necessary to run the simulations at several temperatures (typically 6–10) over a suitable range for the enzyme of interest. This procedure is of course not limited to enzymatic reactions, and can easily be applied to van't Hoff plots and ligand binding.

Here we present an overview of the newly developed graphical user interface, Qgui. Examples of how it is used to build molecular topologies for MD simulations and to compute ligand binding free energies with the LIE [7,19] and FEP [6,20] method are given. We also show how Qgui enables efficient studies of chemical reactions and to obtain thermodynamics activation parameters for enzymatic reactions. It should be noted that Qgui is not limited to Q, and can in principle be extended to function with other MD software packages.

2. Methods

2.1. Qgui—Graphical user interface to Q

Qgui has been written in the Python language using object oriented design principles based on the TkInter GUI library (https://docs.python.org/2/library/tkinter.html). The Qgui consists of a main window, which provides feedback to the user, and four dropdown menus, as illustrated in Fig. 1. The Qgui main window (Fig. 1) is simplistic and mainly functions as a feedback/logging window for all the different tools available. PDB files can be imported directly from the RCSB Protein Data Bank (http://www.pdb.org). To visualize and manipulate structures, Qgui communicates with any