Computational analysis of the evolution of glutathione peroxidase 6 (GPX6) activation free energy

Nayanika Das

Department of Biosciences, Universitat de Vic - Universitat Central de Catalunya, 08500 Vic

September 25, 2024

1 EVB Model

One of the major challenges inherent to all QM/MM approaches is the high computational cost associated with the QM part of the calculation [2]. This is important when it comes to QM/MM free energies where extensive conformational sampling is required. When one tries to deploy extensive DFT calculations to the QM region, this makes the cost of the computation much higher [2]. So, we come to a hybrid QM/MM method using the empirical valence bond (EVB), which has proven to be very useful for calculating thermodynamic activation parameters for chemical reactions [**?**]. Certainly, EVB can capture the changes in the environment around the protein with a simple force field description and not a large DFT or QM/MM cluster, giving an idea of how the enzyme proceeds to work under certain environmental conditions [**?**]. If the activation free energy is correctly predicted at a given temperature, this just means that the sum of ΔH and $-T\Delta S$ is correctly predicted. A recent change that has been made is for cases where the reference reaction would be rather complex to parameterize the EVB potential energy surface directly on DFT calculations [**?**]. To understand the origin of the thermodynamic activation parameters in the EVB model, it is necessary to understand the basis of EVB. The EVB approach constructs potential energy surfaces. This method elegantly transforms a force field-based description of individual electronic ground state (GS) species into a quantum chemical framework using valence bond theory, giving an empirically-based QM/MM description of chemical reactivity [2]. Any number of such states can in principle be used, but often a simple two-state model is considered for an elementary chemical reaction step. The system is then represented by an $n \times n$ EVB Hamiltonian [2].

$$\begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$$

Here, H_{11} and H_{22} are the energies of the two valence states calculated using classical force fields. The off-diagonal matrix elements represent the coupling between the two states. The value of the coupling term H_{11} needs to be calibrated on a reference reaction, and there is also a second parameter that must be calibrated and is phase-independent, the alpha shift [?], which corresponds to the constant difference in free energy between the reacting values in the two states [?]. These reference values (EVB parameters) are either devised from experimental data, commonly in aqueous solution, or directly by QM/MM calculations on the enzyme. Catalytic residues are largely conserved in enzymes as they lower the activation energy of reactions and thereby can increase enzymatic turnover [?]. Mutations in these active sites typically reduce catalytic activity [?]. In this study, we focus on the selenoprotein Glutathione Peroxidase 6 (GPX6) while exploring the replacement in several mammalian lineages of the rare amino acid selenocysteine (Sec) for cysteine (Cys) [?]. GPXSec activity classically reduces hydroperoxides, particularly hydrogen and lipid peroxides, with glutathione (GSH) as a cofactor [?]. GPXCys-containing proteins act on alternative substrates for peroxidation and may have additional functions, including signaling and oxidative protein folding. Thus, all GPX proteins may protect cells from oxidative stress [?]. The goal of this article was to calculate the activation free energy barrier for the Glutathione Peroxidase 6 active site cysteine-containing mouse and selenocysteine-containing human enzyme using Empirical Valence Bond simulation to determine the significant barrier difference with the change in the active site residue from Cys/Sec and the variants obtained from the previous phylogenetic study [?]. Furthermore, we aim to show analysis on the effect of the catalytic activity of both due to variants imposing environmental changes inside the protein.

2 Computational Methodology

The first step in preparing a simulation has missing force field parameters along with standard parameters and library files. In this case, it is Selenium (U). After equilibration, we continue with the FEP simulations which constitute the core part of the EVB methodology. The FEP protocol involves a gradual change of the (mapping) potential energy by the coupling parameter lambda. This change of potential will drive the system from the region of configurational space corresponding to the reactant state (RS) to that corresponding to the product state (PS). The MD simulation required for a free energy calculation often proceeds in multiple stages. The initial stage runs at a very low temperature with strong coupling to the temperature bath (energy minimization) to relax strain in the initial structure. Following this is a stepwise heating of the simulated system and equilibration at the target temperature. For perturbation simulations, this phase is composed of a series of simulations using intermediate potentials defined by different sets of weight coefficients for the FEP states.

2.1 Computational Model Preparation

The initial structure was taken from a snapshot of MD simulation done with OpenMM [**?**]. For creating the selenium parameters, FFLD in Maestro was used. We used the charges provided by Maestro. Protonation states were predicted using PROPKA [**?**], where the hydrogen and solvent were added using the Q program [**?**]. For the transition states, FFLD in Maestro and QM/MM parameters were used. The TIP3P water model was used in combination with the protein parameters mentioned above.

2.2 MD/EVB Simulations

Spherical boundary conditions [**?**] were applied to the system using the Q program [**?**], with 50 Å diameter water covering the protein. After equilibration of the system, two separate MD/EVB free energy perturbation (FEP) calculations were done: for step 1 (until formation of the selenolate ion) and for step 2 (to the formation of selenenic acid), which is the product of the reaction. The FEP protocol involves a gradual change of the (mapping) potential energy by the coupling parameter. For each of the FEP calculations, 51 discrete windows were considered, each window of 10 ps at 2 fs time step, giving a total of 1.02 ns of sampling of each free energy profile. These calculations were replicated 10 times starting from the minimized structure. The EVB gas phase shift α and the off-diagonal H_{ij} were determined iteratively to obtain average reaction free energies and barrier heights of all four systems in study. The four barriers were determined according to the QM/MM calculations that have been previously done (TS1 - 16.4 kcal/mol and Product - 18.8 kcal/mol).

References

[1] Schrödinger Release 2024-2: Maestro, Schrödinger, LLC, New York, NY, 2024.

[2] Carvalho, A. T. P., Barrozo, A., Doron, D., Vardi Kilshtain, A., Major, D. T., & Kamerlin, S. C. L. (2014). Challenges in computational studies of enzyme structure, function, and dynamics. *Journal of Molecular Graphics and Modelling*, 54, 62–79. https://doi.org/10.1016/j.jmgm.2014.09.003