

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

Nayanika Das,[†] Other Authors,[‡] and Jordi Villà-Freixa^{*,†,¶}

[†]Computational Biochemistry and Biophysics Lab, Research Group on Bioinformatics and Bioimaging (BP²), Department of Biosciences, Universitat de Vic - Universitat Central de Catalunya, 08500 Vic, Spain

[‡]Here is were the other authors work

¶IRIS-CC

E-mail: jordi.villa@uvic.cat

Abstract

Introduction

Selenium (Se), in the form of selenocysteine (Sec, U—the 21st amino acid) occurs in 25 proteins in the human proteome. Insertion of Sec into a protein is much more complicated than the other 20 amino acids because a UGA stop codon must be recoded as a sense codon for Sec.¹ The complexity of this process signifies that Sec must fulfill a chemical function that exerts biological pressure on the genome to maintain the Sec-insertion machinery^{1,2}. The view of Sec as a sophisticated innovation implies that for each occurrence of Sec in an

enzyme there is a unique and specific reason for the use of Se to enhance the enzymatic reaction relative to that of S.¹ The view of Sec as a sophisticated innovation implies that for each occurrence of Sec in an enzyme there is a unique and specific reason for the use of Se to enhance the enzymatic reaction relative to that of S.¹ This view also implies that since Se “speeds reactions” Sec should have widely substituted for Cys in enzymes, which clearly has not occurred. Specific reasons for the usage of Sec might include the enhanced nucleophilic character of Se relative to S, or another might be the much lower pKa of a selenol relative to that of a thiol.¹ Although the catalytic triad of glutathione peroxidase (GPX) has been well recognized, there has been little evidence for the relevance of the interactions among the triad amino acid, i.e. selenocysteine (U), glutamine (Q), and tryptophan (W). Hence, the mechanism of has been studied in various aspects. GPXSec activity classically reduces hydroperoxides, particularly hydrogen and lipid peroxides, with glutathione (GSH) as a cofactor.³ GPXCys containing proteins act on alternative substrates for peroxidation and may have additional functions, including signalling and oxidative protein folding. Thus, all GPX proteins may protect cells from oxidative stress.³

Mechanism of GPX

A catalytic mechanism proposed for GPX3 by Morokuma et al. based on DFT calculations has the resting state of Sec as selenol as shown in the above figure. In the first part of this reaction, hydrogen peroxide coordinates to the active site of the enzyme and the proton transfer is happening to the Gln83 residue^{4, 5} Morokuma⁴ did an ONIOM(QM:MM) method to evaluate the quantitative effect of the protein surroundings on the energetics of an enzymatic reaction. In the first step, the formation of the selenolate anion Se^- occurs via the proton transfer from the Se through the oxygen (O1) atom of hydrogen peroxide to the neighboring Gln83, leading to the intermediate (III).⁴ The computed barrier for the creation of the selenolate anion is 16.4 kcal/mol.⁴ In the second step of the stepwise mechanism, the O1-O2 bond of H_2O_2 is cleaved. During this process, one hydroxyl fragment (O1H) is trans-

ferred to the selenolate anion Se^- to form selenenic acid (R-SeOH), while simultaneously the second hydroxyl fragment (O₂H) accepts the previously transferred proton from Gln83 to form a water molecule (IV). Overall barrier (from II to IV) for the formation of selenenic acid (E-Se-OH) becomes 18.0 kcal/mol.⁴

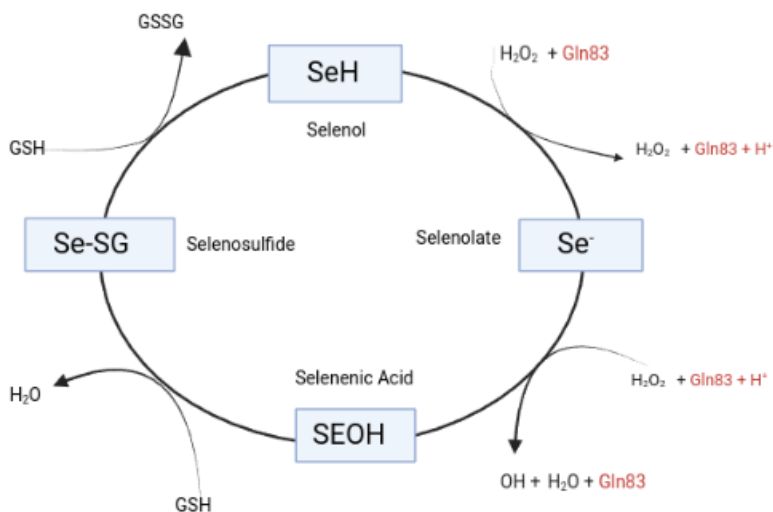


Figure 1: Catalytic Cycle of GPX as given by Morokuma et al. where the resting state of selenium is Selenol

Flohe et al⁶ explored the mechanism in a different way wherein the selenocysteine proton moves via water to the indol nitrogen of the Trp residue and then the hydrogen peroxide and with the result of formation of the products are instantly. They show essentials of

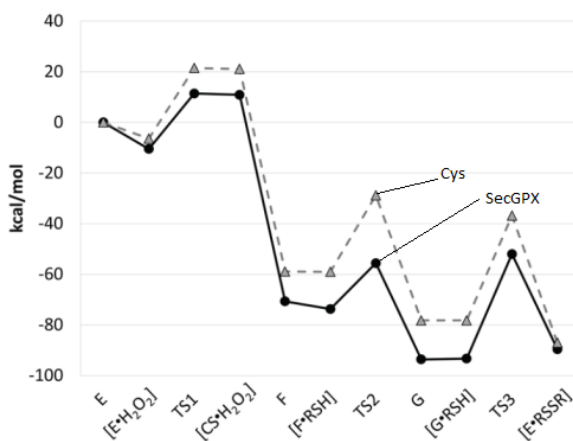


Figure 2: Energetic profile of the cycle for a SecGPX and the corresponding Cys, as DFT-calculated

the DFT-calculated oxidation of SecGPx by H_2O_2 . The calculation was performed with

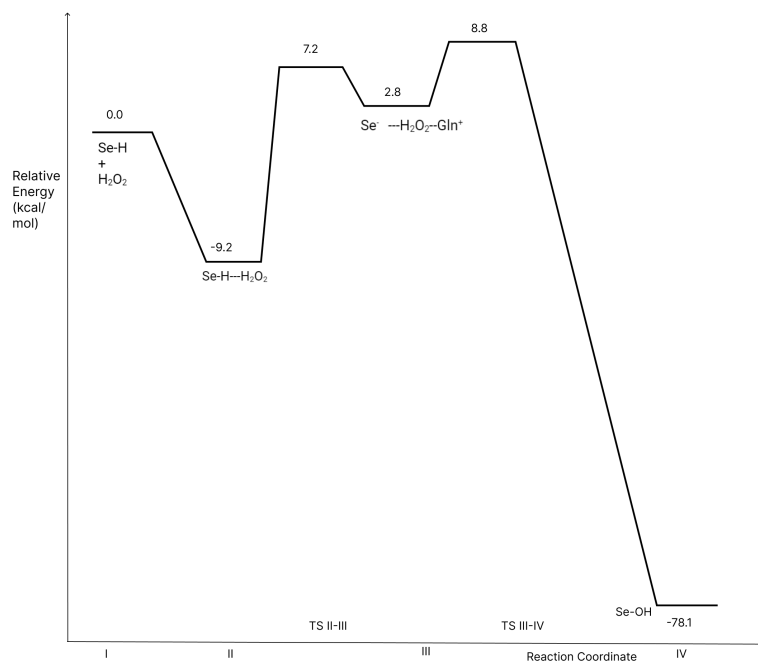


Figure 3: Potential Energy Diagram for the hydrogen peroxide reduction mechanism by GPX by Morokuma et al.

E, as shown in the figure, but with a single water molecule and H_2O_2 bound to the reaction center.⁶ The transition state, leading to the charge-separated form CS, involves three concurrent steps. CS evolves to the selenenic form F and a water molecule.⁶ Formation of the charge separated form does not depend on binding of H_2O_2 but just on the presence of water. Binding of H_2O_2 to CS then generates the identical unstable $CS \cdot H_2O_2$ complex which decays as described.⁶ In the transition state (TS1), the selenol proton moves via H_2O_2 and water to the Trp nitrogen, thus creating a charge-separated species. The resulting complex $[CS \cdot H_2O_2]$ is so unstable that it decays without any energy barrier. In G the selenocysteine is thiylated by a thiol substrate.⁶

Empirical Valence Bond Model based on the QM/MM calculations

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.⁷ Certainly EVB can capture the changes in the environment around the protein, with the 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.⁷ 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.⁷ To understand the origin of the thermodynamic activation parameters in the EVB model, it is necessary to understand the basis of EVB. EVB approach is construction of potential energy surfaces. 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 a $n \times n$ EVB Hamiltonian.⁷

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

Here, H_{11} and H_{22} are the energies of the two valence states which are calculated using classical force fields. The off-diagonal matrix elements represent the coupling between the two states. The value of the coupling term H_{12} , needs to be calibrated on a reference reaction and there is also a second parameter that must be calibrated and it is phase independent,⁸ the alpha shift, which corresponds to the constant difference in free energy between the reacting values in the two states.⁸ These reference values (EVB parameters) are either devised from experimental data, commonly in aqueous solution, or directly by QM/MM calculations on the enzyme.

The Concept of Epistasis and Protein Sequence

As proteins evolve, they follow trajectories through sequence space, so this topology also determines how mutation, drift, selection, and other forces can drive genetic and functional

evolution. Catalytic residues are largely conserved in enzymes as they lower the activation energy of reactions and thereby can increase enzymatic turnover.³ Epistasis can cause a mutation that confers or improves a function in one protein to have no effect or even be strongly deleterious in a related protein, attempts to give rise to natural sequence variation.⁹ So, now the question is how prevalent is epistasis within proteins.⁹ One way to gain insights into this in the context of protein evolution is to compare the effects of some mutation when it is introduced into different proteins related by evolutionary homologs.⁹ This is because protein homologs point to both strong and pervasive effects of epistasis that cause the functional effects of mutations to differ between related proteins.⁹ Epistatic interactions between mutant sites can help in explaining why evolution follows certain pathways.¹⁰ Hence, in this study we focus on selenoprotein Glutathione Peroxidase 6 (GPX6) homologs, essentially mouse and human, exploring the replacement for the rare amino acid Selenocysteine (Sec) to Cysteine (Cys) in Human and Cysteine (Cys) to Selenocysteine (Sec) in Mouse. The goal of this article was to calculate the activation free energy barrier for the Glutathione Peroxidase 6 homologs using Empirical Valence Bond simulation, in order to determine the significant barrier difference with the change in the active site residue from cys/sec. And further more showing analysis on the affect of the catalytic activity of both due to variants imposing environmental changes inside the protein.

Active site of GPX contains SEC49 in human GPX6 and CYS49 in mouse GPX6. While they are surrounded by residues GLN83, TRP157, LEU51, TYR55, GLY50. These residues are highly conserved in all GPX isoforms and take an important part in the mechanism. The most feasible mechanism that we followed in this article is the SeH-Gln83 by using the implementation of the DFT calculated reference reaction previously done by Morokuma et al.⁴ to set up our EVB calculations.

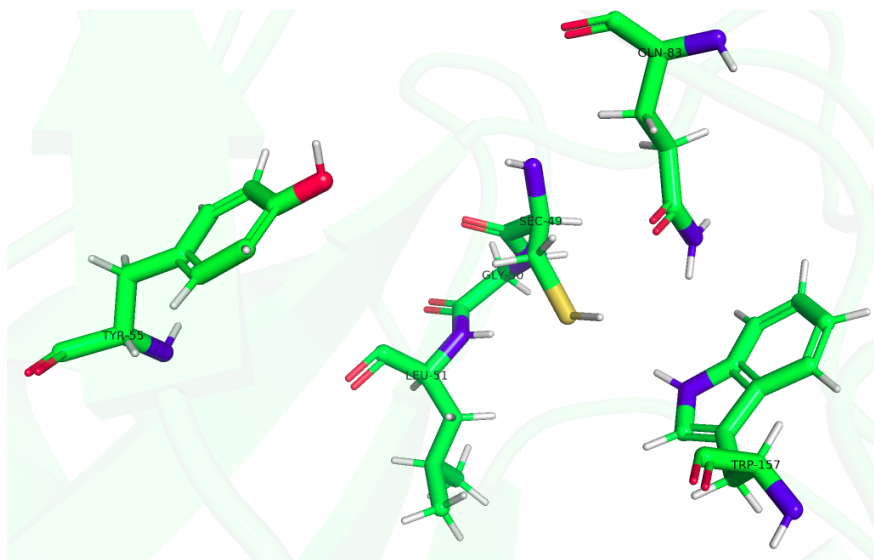


Figure 4: Active Site of GPX6 Human

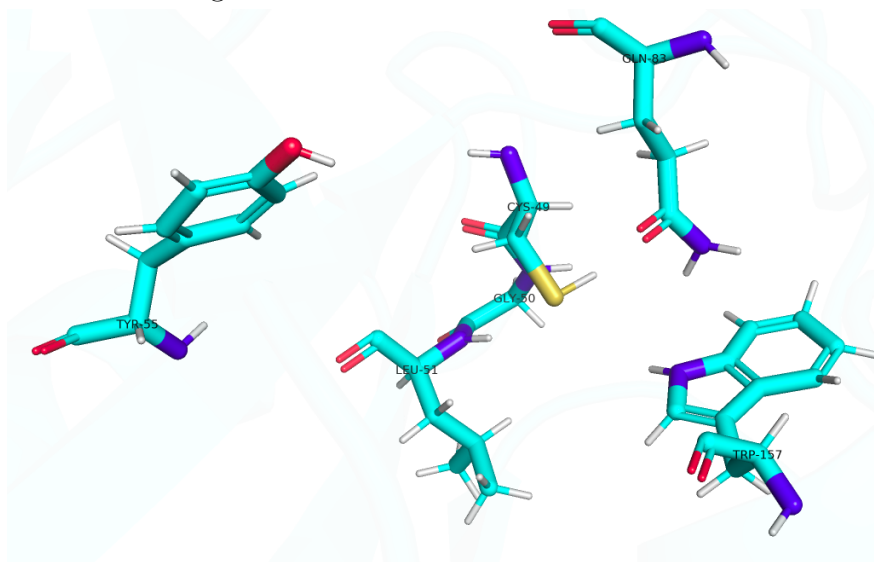


Figure 5: Active Site of GPX6 Mouse

Computational Methodology

The first step in preparing a simulation that has missing force field parameters along with standard parameters and library files. In this case it is Selenium (U). After equilibration, continuing 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 λ . This change of potential will drive the system from the region of configurational space corresponding to the reactant state (RS) to that corresponding to product state (PS). The MD simulation required for a free energy calculation often proceeds in multiple stages. The initial stage is running at a very low temperature with strong coupling to the temperature bath (energy minimisation) to relax strain in the initial structure. Following 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.

Computational Model Preparation

The initial structure was taken from snapshot of MD simulation done with openMM.¹¹ For creating the selenium parameters, FFLD in maestro was used. We used the charges provided by Maestro. Protonation states were predicted using PROPKA,¹² where the hydrogen and solvent was added using Q program¹³ For the transition states, FFLD in maestro and QM/MM parameters was used. The TIP3P water model was used in combination with the protein parameters mentioned above.

MD/EVB simulations

Spherical boundary conditions¹⁴ were applied to the system using Q program¹³ using 50 Å diameter water covering the protein.

After equilibration of the system, two separate MD/EVB free energy perturbation (FEP)

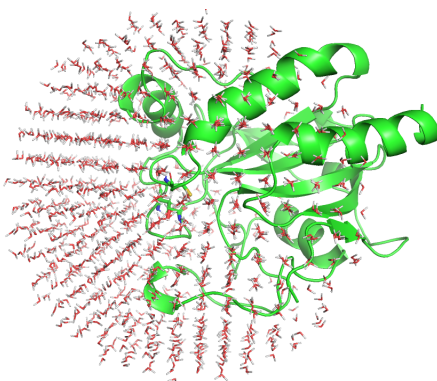


Figure 6: Sphere of water molecules around the protein

calculations were done for the step 1 (until formation of selenolate ion) and the 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 decrease windows were considered, each window of 10 ps at 2 fs time step, that gave 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 in order to have an average reaction free energies and the barrier heights of all the four systems in study. The four barriers we got were essentially determined according to the QM/MM calculations that have been previously done.

References

- (1) Hondal, R. J.; Ruggles, E. L. Differing views of the role of selenium in thioredoxin reductase. 2011.
- (2) Cardey, B.; Enescu, M. Selenocysteine versus cysteine reactivity: A theoretical study of their oxidation by hydrogen peroxide. *Journal of Physical Chemistry A* **2007**, *111*, 673–678.

- (3) Rees, J.; Sarangi, G.; Cheng, Q.; Floor, M.; Andrés, A. M.; Miguel, B. O.; Villà-Freixa, J.; Arnér, E. S.; Castellano, S. Ancient loss of catalytic selenocysteine spurred convergent adaptation in a mammalian oxidoreductase. *Genome Biology and Evolution* **2024**,
- (4) Prabhakar, R.; Vreven, T.; Frisch, M. J.; Morokuma, K.; Musaev, D. G. Is the protein surrounding the active site critical for hydrogen peroxide reduction by selenoprotein glutathione peroxidase? An ONIOM study. *Journal of Physical Chemistry B* **2006**, *110*, 13608–13613.
- (5) Prabhakar, R.; Vreven, T.; Morokuma, K.; Musaev, D. G. Elucidation of the mechanism of selenoprotein glutathione peroxidase (GPx)-catalyzed hydrogen peroxide reduction by two glutathione molecules: A density functional study. *Biochemistry* **2005**, *44*, 11864–11871.
- (6) Orian, L.; Mauri, P.; Roveri, A.; Toppo, S.; Benazzi, L.; Bosello-Travain, V.; Palma, A. D.; Maiorino, M.; Miotto, G.; Zaccarin, M.; Polimeno, A.; Flohé, L.; Ursini, F. Selenocysteine oxidation in glutathione peroxidase catalysis: An MS-supported quantum mechanics study. *Free Radical Biology and Medicine* **2015**, *87*, 1–14.
- (7) Oanca, G.; Åqvist, J. Why Do Empirical Valence Bond Simulations Yield Accurate Arrhenius Plots? *Journal of Chemical Theory and Computation* **2024**, *20*, 2582–2591.
- (8) Oanca, G.; van der Ent, F.; Åqvist, J. Efficient Empirical Valence Bond Simulations with GROMACS. *Journal of Chemical Theory and Computation* **2023**, *19*, 6037–6045.
- (9) Starr, T. N.; Thornton, J. W. Epistasis in protein evolution. 2016.
- (10) Storz, J. F. Compensatory mutations and epistasis for protein function. 2018.

- (11) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L. P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Computational Biology* **2017**, *13*.
- (12) Søndergaard, C. R.; Olsson, M. H.; Rostkowski, M.; Jensen, J. H. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of p K a values. *Journal of Chemical Theory and Computation* **2011**, *7*, 2284–2295.
- (13) Marelius, J.; Kolmodin, K.; Feierberg, I.; Åqvist, J. Q: A molecular dynamics program for free energy calculations and empirical valence bond simulations in biomolecular systems. 1999.
- (14) King, G.; Warshel, A. A surface constrained all-atom solvent model for effective simulations of polar solutions. *The Journal of Chemical Physics* **1989**, *91*, 3647–3661.