



HAL
open science

Theoretical characterization of the shikimate 5-dehydrogenase reaction from *Mycobacterium tuberculosis* by hybrid QC/MM simulations and quantum chemical descriptors

Igor Barden Grillo, José Fernando Ruggiero Bachega, Luís Fernando S. M. Timmers, Rafael A. Caceres, Osmar Norberto de Souza, Martin J. Field, Gerd Bruno Rocha

► To cite this version:

Igor Barden Grillo, José Fernando Ruggiero Bachega, Luís Fernando S. M. Timmers, Rafael A. Caceres, Osmar Norberto de Souza, et al.. Theoretical characterization of the shikimate 5-dehydrogenase reaction from *Mycobacterium tuberculosis* by hybrid QC/MM simulations and quantum chemical descriptors. *Journal of Molecular Modeling*, 2020, 26 (11), pp.297. 10.1007/s00894-020-04536-9. hal-02987221

HAL Id: hal-02987221

<https://hal.science/hal-02987221v1>

Submitted on 24 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Theoretical Characterization of the Shikimate 5-dehydrogenase Reaction from *Mycobacterium Tuberculosis* by Hybrid QC/MM Simulations and Quantum Chemical Descriptors

Igor Barden Grillo · José Fernando
Ruggiero Bachega · Luis Fernando S.M.
Timmers · Rafael A. Caceres · Osmar
Norberto de Souza · Martin J. Field ·
Gerd Bruno Rocha

Received: date / Accepted: date

Igor Barden Grillo

Departamento de Química, Centro de Ciências Exatas e da Natureza Universidade Federal da Paraíba, João Pessoa, Brazil
E-mail: barden.igor@gmail.com

José Fernando Ruggiero Bachega

Laboratório de Bioinformática, Modelagem e Simulação de Biosistemas (LABIO), Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) Programa de Pós-Graduação em Biologia Celular e Molecular, Av. Ipiranga 6681, 90619-900, Porto Alegre, RS, Brazil
Present address: Departamento de Farmacociências, Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSA)

Luis Fernando S.M. Timmers

Laboratório de Bioinformática, Modelagem e Simulação de Biosistemas (LABIO), Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) Programa de Pós-Graduação em Biologia Celular e Molecular, Av. Ipiranga 6681, 90619-900, Porto Alegre, RS, Brazil
Present address: Programa de Pós-Graduação em Biotecnologia (PPGBiotec), Universidade do Vale do Taquari - Univates, Rua Avelino Talini, 171 - Bairro Universitário, Lajeado, RS, Brazil

Rafael A. Caceres

Departamento de Farmacociências, Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSA)
Programa de Pós-Graduação em Ciências da Saúde (PPGCS) e Programa de Pós-Graduação em Biociências (PPGBio), Porto Alegre - Brasil

Osmar Norberto de Souza

Laboratório de Bioinformática, Modelagem e Simulação de Biosistemas (LABIO), Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) Programa de Pós-Graduação em Biologia Celular e Molecular, Av. Ipiranga 6681, 90619-900, Porto Alegre, RS, Brazil
Programa de Pós-Graduação em Biologia Celular e Molecular, Av. Ipiranga 6681, 90619-900, Porto Alegre, RS, Brazil

Martin J. Field

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

Abstract In this study, we have investigated the enzyme shikimate 5-dehydrogenase from the causative agent of tuberculosis, *Mycobacterium tuberculosis*. We have employed a mixture of computational techniques, including molecular dynamics, hybrid quantum chemical/molecular mechanical potentials, relaxed surface scans, fast quantum chemical descriptors and free-energy simulations, to elucidate the enzyme’s reaction pathway. Overall, we find a concerted mechanism, with a single transition state, that proceeds by an energetically uphill hydride transfer, followed by an energetically downhill proton transfer. Our mechanism and calculated free energy barrier for the reaction, 64.9 kJ mol^{-1} , are in good agreement with those predicted from experiment. An analysis of quantum chemical descriptors along the reaction pathway indicated a possibly important, yet currently unreported, role of the active site threonine residue, Thr65.

Keywords *Mycobacterium tuberculosis* · free-energy profiles · fast quantum chemical descriptors · quantum chemical/molecular mechanical potentials · reaction mechanism · shikimate 5-dehydrogenase

1 Introduction

Enzymes are the most important catalysts in life chemistry. These proteins can increase reaction rates by a factor of as much as 10^8 in some cases, thereby tuning the chemical reactions in metabolic pathways. As a result, enzymes have substantial relevance in many areas of biochemistry, including biotechnology, nanoscience and medicinal chemistry. The selection of a suitable protein to target is a crucial step in the drug development process of bioactive molecules, such as drugs and herbicides. The protein should either be in the metabolic pathway that is present in the pathogen but absent in the hosts, following the rule of selective toxicity, or be sufficiently different that it can be specifically targeted.

The shikimate (SKM) pathway [1] is a seven-step enzymatic pathway that converts erythrose 4-phosphate and phosphoenolpyruvate to chorismate, a biosynthetic precursor of aromatic amino acids, vitamins B9 and K1, ubiquinone, and salicylate. This route is present in plants, fungi, bacteria, and apicomplexan parasites, but not in metazoans, such as humans beings, which makes the enzymes from this enzymatic pathway promising targets for non-toxic antimicrobials and herbicides. The widely used herbicide known as glyphosate is an effective inhibitor of 5-enolpyruvyl-shikimate-3-phosphate (EPSP) synthase, the sixth step on the SKM pathway.

An example of an organism in which the SKM pathway has been targeted is *Mycobacterium tuberculosis*. This bacterium is the causative agent of tuberculosis which, according to the WHO, is the most fatal infectious disease

Gerd Bruno Rocha

Departamento de Química, Centro de Ciências Exatas e da Natureza Universidade Federal da Paraíba, João Pessoa, Brazil

with over one million deaths per year [2]. Gene disruption experiments have demonstrated that the SKM pathway is essential for the viability of *M. tuberculosis* [3] and, as a result, substantial efforts have been made to develop inhibitors against its constituent enzymes, including 3-dehydroquinate dehydratase [4], shikimate kinase [5, 6, 7] and EPSP synthase [8].

Shikimate 5-dehydrogenase (SDH), encoded by the gene *aroE*, catalyzes the fourth step of the SKM pathway. It converts dehydroshikimate (DHK) to SKM via a NADPH-dependent reduction, a reaction shown in Figure 1. A concerted catalytic mechanism for the SDH reaction was first proposed by Fonseca and coworkers [9], in which the transfers of the hydride from NADPH (atom H42 in the figure) and the proton from the solvent occur simultaneously. Subsequently, Gan and coworkers proposed that the SDH-catalysed reaction in *Aquifex aeolicus* followed the same mechanism [10]. More recently, Peek and coworkers highlighted the importance of the conserved lysine and aspartate residues in the SDH reaction [11].

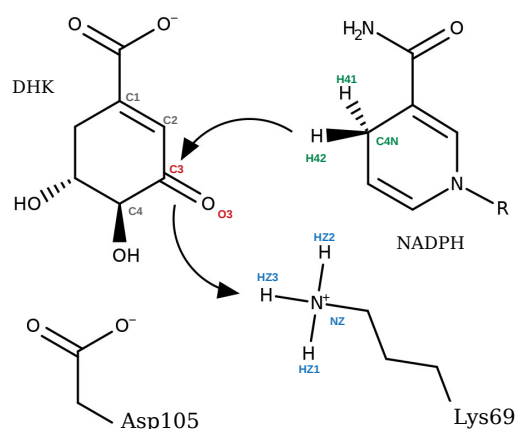


Fig. 1 Proposed mechanism of the shikimate 5-dehydrogenase catalytic reaction.

Although experimental studies have supplied valuable insights into the SDH reaction in various organisms, it is difficult to fully elucidate a mechanism via *in vitro* experiments alone. Computational approaches can act as a complement to experiment by providing information that is difficult to obtain experimentally, such as a knowledge of transition state (TS) structures. Here, we employ molecular simulation methods to investigate the reaction of the shikimate 5-dehydrogenase from *M. tuberculosis* (MtSDH). We first used classical molecular dynamics (MD) simulations along with molecular mechanical (MM) potentials to determine the most populated conformations of the MtSDH enzyme complexed with its substrates prior to enzyme catalysis. Subsequently we

performed MD simulations with hybrid quantum chemical (QC)/MM potentials to study the reduction catalyzed by the enzyme. Finally, we analyzed in detail the reactivity of the reacting atoms using reactivity descriptors (RDs), which are theoretical quantities that correspond to well established chemical concepts, such as softness and hardness. RDs have already been applied in the context of biomolecules and have been shown to be able to reliably reproduce reactivity information for polypeptides and for ligand-protein interactions [12], and also to theoretically characterize enzymatic catalytic mechanisms [13]. Together our simulation results and RD analysis, in combination with previous experimental data, provide a detailed picture of the reaction mechanism of MtSDH.

2 Methods

In this section we present the setup and details of the computational simulations performed in this study. We applied three different levels of simulation protocol, namely: molecular mechanical/molecular dynamics (MM/MD), in which we sampled representative reactive structures of the system; quantum chemical/molecular mechanical (QC/MM) potentials for scanning possible reaction paths, searching for transition states, and calculating reaction free energies; and reactivity descriptors (RDs) for characterizing atom reactivity along the reaction path so as to elucidate some of the stabilizing interactions that provide the enzyme with its catalytic power. A summary of our simulation protocol is given in Figure 2.

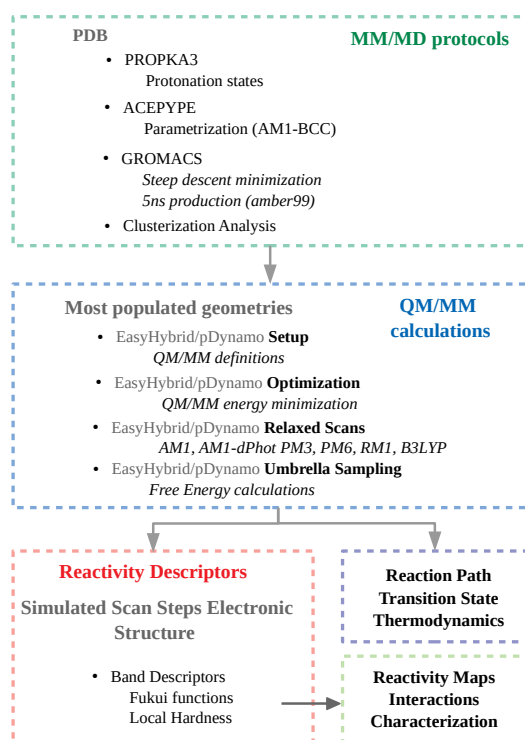


Fig. 2 Flowchart of the simulation protocol used in this study.

2.1 MM/MD simulations

The coordinates of the enzyme MtSDH were obtained from the Protein Data Bank (pdb id: 4P4G) [14]. The original shikimate molecule present in the structure was edited and converted to 3-dehydroshikimate using the Avogadro molecular editor [15], and the NADPH molecule was introduced into the MtSDH active site via the superposition of the crystal structure of SDH from *Thermus thermophilus* (pdb id: 2EV9) [16].

MD simulations of MtSDH associated with its substrates were carried out using the Amber99SB force field implemented in GROMACS 2016.3 [17]. The substrate's force field was parameterized using ACPYPE [18] and its atom partial charges assigned by AM1-bcc. NADPH parameters were obtained from the Bryce group Amber Database [19,20]. Protonation states of the protein were assigned at pH 7.5 using the PROPKA3 [21] web server. The protein-substrate complex was fully solvated in a cubic box. Periodic bound conditions were employed, with van der Waals interactions computed using a 12 Å cut-off, and electrostatic contributions calculated via a particle mesh Ewald method using a grid with 1.2 Å spacing.

Simulations were performed in the isothermal-isobaric ensemble. The temperature was kept constant (298.15 K) using a V-rescale [22] thermostat with a coupling time of 0.1 ps, and the pressure was maintained at 1 bar by the Berendsen barostat [23]. The system evolved in time steps of 2 fs, since a LINCS [24] algorithm was used to restrain the lengths of all bonds involving hydrogen atoms. Prior to MD, the system was energy minimized with a steepest-descent energy algorithm. Altogether, MD simulations were performed for a duration of 5 ns after equilibration, with snapshots of the system being saved every 20 ps, giving 2500 snapshots for subsequent analysis.

2.2 QC/MM simulations

Snapshots from the MM/MD simulations were used as starting points for the QC/MM potential calculations. All QC/MM simulations were performed with pDynamo 1.9.0 [25, 26] coupled to ORCA 4.0 [27]. The EasyHybrid/GTKDynamo interface [28] was used to setup the system. To reduce the computational cost of the QC/MM calculations we employed a subset of the MM/MD system that consisted of the 7432 atoms within 25 Å of the H42(NADPH) atom. To preserve the integrity of the system, the positions of the atoms at the boundary of this subset were frozen. The QC region had 66 atoms with a total electronic charge of zero, consisting of the side chain of Lys69, DHK, and the nicotinamide and ribose portions of the NADPH (see Figure S1). pDynamo’s standard hydrogen link atom approximation was employed to saturate the broken bonds at the boundary between QC and MM regions.

The reaction was studied using a relaxed surface scan procedure. This was performed with five of the semiempirical QC methods implemented in pDynamo, namely AM1, AM1/d-PhoT, PM3, PM6 and RM1, together with B3LYP/3-21G and B3LYP/6-31+G* density functional theory (DFT) methods available in ORCA. These scans were supplemented by calculations to determine free energy profiles as a function of a reaction coordinate using umbrella sampling and the weighted histogram analysis method (WHAM), as implemented in pDynamo. For each profile, 40 umbrella sampling windows were employed. For each window, 10 ps of equilibration dynamics followed by 50 ps of production dynamics were performed, using distance-dependent harmonic biasing potentials with force constants of 200 kJ mol⁻¹Å⁻². Finally, as a consistency check on the quality of the scan reaction paths, we carried out refinements using the nudged elastic band (NEB) algorithm [29] implemented in pDynamo.

2.3 Reactivity descriptor calculations

Conceptual DFT (CDFT) provides mathematical definitions of commonly-used chemical concepts. These definitions are drawn from the fundamental differential equation of DFT, Equation 1 [30], either by taking its derivatives

or manipulating its Taylor expansions, so as to provide quantitative estimates of the changes in electronic structure as the number of electrons varies.

$$dE = \mu dN + \int \rho(r) d\nu(r) dr \quad (1)$$

The derivatives of energy with respect to the number of electrons produce global quantities. The first derivative is the electronic chemical potential, μ , which is a measure of the capacity of the system to donate electrons (Equation 2). The second derivative is the hardness, η , which indicates the resistance of the system to donate electrons, and whose inverse is the softness S (Equation 3).

$$\mu = \left(\frac{\partial E}{\partial N} \right)_{\nu} \quad (2)$$

$$\eta = \frac{1}{S} = \left(\frac{\partial^2 E}{\partial N^2} \right)_{\nu} \quad (3)$$

The derivatives of electron density with respect to the number of electrons result in local RDs that assign a number to each point in three-dimensional space. The first derivative of electron density with respect to the number of electrons is the Fukui function (Equation 4) [31]. The electron density in regions where this derivative reaches its maximum has the greatest propensity to increase as the number of electrons vary. Due to the discontinuity in the number of electrons, this derivative is determined using finite differences where the densities are calculated by performing DFT calculations at the N -electron geometry but with $N - 1$, N and $N + 1$ electrons [32]. This leads to two local RDs, namely the electrophilic attack susceptibility (EAS) (Equation 5), and the nucleophilic attack susceptibility (NAS) (Equation 6). The latter can also be calculated using frontier molecular orbitals, with the EAS and NAS approximated by the density of HOMO (Equation 7) and LUMO (Equation 8), respectively.

$$f(r) = \left(\frac{\partial \rho(r)}{\partial N} \right)_{\nu} \quad (4)$$

$$EAS(r) = f^{-}(r) = \rho(r)_N - \rho(r)_{N-1} \quad (5)$$

$$NAS(r) = f^{+}(r) = \rho(r)_{N+1} - \rho(r)_N \quad (6)$$

$$EAS(r) = |\phi_{\text{HOMO}}(r)|^2 \quad (7)$$

$$NAS(r) = |\phi_{\text{LUMO}}(r)|^2 \quad (8)$$

The local RDs can be condensed to atoms, thereby assigning values to atomic centers. Equation 9 shows the generic condensation method for a Fukui

function whose value is integrated over a region of space that encloses an atom k . In the case of finite difference Fukui functions, the condensation is done using the partial gross charges of each atom derived from the requisite $N - 1$, N and $N + 1$ -electron calculations.

$$f_k^a = \int_{\Omega_k} f(r)_k^a dr \quad (9)$$

A local softness function can be defined in terms of the Fukui function by multiplication by the global softness, S . Like the Fukui function, soft-soft interactions are related to orbital superposition, as opposed to the hard-hard interactions that arise from electrostatic and weak forces. By contrast, however, there is no unambiguous definition of a local version of hardness, although the mathematical development of CDFT has led to several solutions for its approximation. Grillo and co-workers [13] showed that the local hardness was crucial for the depiction of enzymatic reaction profiles, and that reasonable values for an atom k could be obtained using the molecular electrostatic potential definition in Equation 10. In this equation, N is the total number of electrons, and the sum runs over all atoms, l , other than k .

$$\frac{1}{2N} \sum_{l \neq k} \frac{f^-(r_l)}{|r_k - r_l|} \quad (10)$$

The calculation of RDs faces specific issues when employed for large biomolecules, mainly due to the size of the system. For finite-difference calculations, the standard perturbation of removing and adding one electron to estimate the derivative is often insufficiently precise due to the large number of electrons compared to the small number present in the molecules typically studied by organic chemists. The solution proposed by Khandonghin and coworkers [33] is to use a larger charge perturbation of at least five electrons, but this can lead to difficulties in converging the QC calculation. By contrast, frontier orbital methods are cheaper as only one QC calculation is required, rather than the three needed for the finite-difference methods. Nevertheless, problems arise in large biological polymers because of degeneracies among the molecular orbitals near the HOMO-LUMO gap. To counter this, Fukushima and coworkers [34] showed that the reactivity of enzymes can be related to orbital energy levels in windows as far as 1-5 eV from the HOMO and LUMO energies, suggesting that these should also be included in the definitions of the RDs. Because of this, in this study, we have used an energy-weighted approach for the calculation of the Fukui functions, shown for the EAS and the NAS in Equation 11 and Equation 12, respectively. In these equations the sums run over all orbitals whose energy lies within a given energy band, b , of the HOMO (EAS) or LUMO (NAS).

$$f_{EW}^- = \sum_{i=b}^{HOMO} e^{-|E_i - E_{HOMO}|} |\psi_i|^2 \quad (11)$$

$$f_{EW}^+ = \sum_{i=LUMO}^b e^{-|E_i - E_{LUMO}|} |\psi_i|^2 \quad (12)$$

3 Results and Discussions

3.1 Structures from the MM/MD simulations

We carried out MM/MD simulations to determine the range of conformations that were accessible to the protein-substrate system. For the reduction reaction, two pertinent distances that characterize it are those between the hydride, H42(NADPH), and C3(DHK), and between the proton, HZ3(Lys69), and O3(DHK), respectively. We show in Figure 3A scatter plots of these two distances obtained from each of the MD snapshots, from which we retrieved the most representative structure obtained from the simulation. This structure is given in Figure 3B, with H42-C3 and HZ3-O3 distances of 2.90 Å and 2.85 Å, respectively. We use this as the starting point for the hybrid QC/MM calculations.

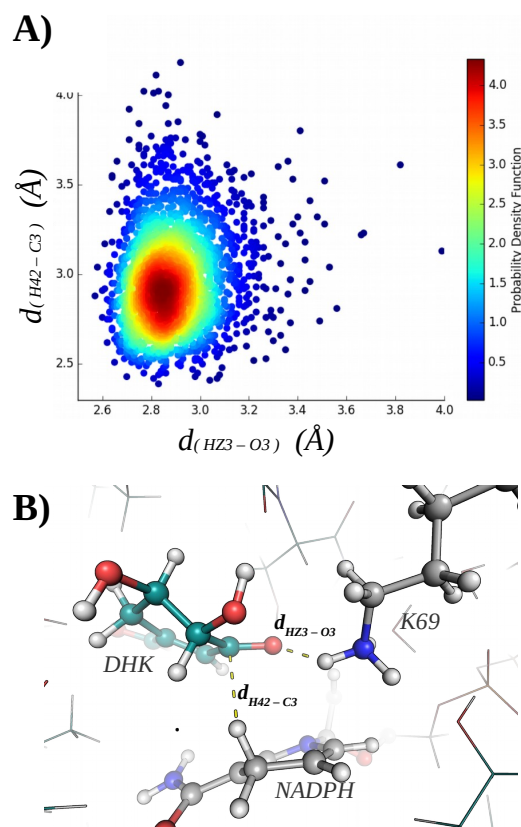


Fig. 3 Analysis of structures from the MM/MD simulations. (A) Scatter plots of the hydride transfer H42-C4N and proton transfer HZ3-O3 distances for each of the MD snapshots; (B) The MM/MD structure used for the QC/MM calculations. The tertiary structure is represented as ribbons, and the residues involved in the reaction as balls and sticks with CPK colors. The carbon atoms of substrate (DHK) are colored in green. Image generated with PyMOL [35].

3.2 Description of the SDH catalyzed reaction

We performed relaxed surface scans of the reduction reaction as a function of two coordinates. The first considered, which characterizes the hydride transfer, is the difference between the C4N-H42(NADPH) and H42-C3(DHK) distances, whereas that for the proton transfer was the difference between NZ-HZ3(Lys69) and HZ3-O3(DHK) distances. From these simulations, we calculated PESs as a function of both coordinates, displayed in Figure S2, and one-dimensional potential energy curves determined by simulating only the hydride transfer coordinate, depicted in Figure 4.

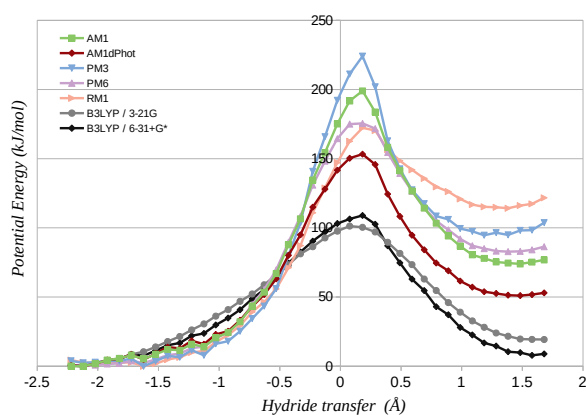


Fig. 4 Relaxed surface scan potential energy profiles for the SDH reaction from *M. tuberculosis* using different QC/MM methods as a function of the hydride transfer reaction coordinate. All structures with the semiempirical and B3LYP/3-21G QC/MM methods were fully geometry optimized as a function of the reaction coordinate. By contrast, the B3LYP/6-31+G* values are single points on top of the B3LYP/3-21G structures.

The surfaces in Figure S2 show that the reaction proceeds with a single energy barrier, corresponding to a TS structure in which the hydride is almost completely transferred, but the proton transfer has yet to occur. After the TS there is an increase in the C3-O3 bond length of SKM and a spontaneous reorientation of the general base, Lys69 towards O3. Proton transfer itself occurs after hydride transfer has finished. The geometries obtained with the different QC/MM potentials agree well, although the TS energies differ substantially. This can be seen in the one-dimensional energy profiles in Figure 4, with the B3LYP/3-21G QC/MM potential presenting the lowest energy barrier, of 101.2 kJ mol⁻¹. This is in agreement with kinetic isotope effect measurements that suggest a concerted mechanism with a single TS [9].

Figure 5 shows the critical structures along the pathway arising from the 2-D scans with the AM1/d-PhoT QC/MM potential. In addition, it gives the PES obtained by performing single point calculations with a B3LYP/6-31+G* DFT QC/MM potential at the B3LYP/3-21G optimized geometries. The DFT/MM surface gives the same mechanism as the semiempirical ones, but its energies are much lower, with a TS energy of only 101.2 kJ mol⁻¹. The same is observed for the equivalently calculated 1-D DFT/MM profile for the hydride transfer (Figure 4). NEB calculations also corroborate the presence of a single TS and a concerted mechanism. These NEB calculations, including those fully optimized with a B3LYP/6-31+G*/D3 DFT QC/MM potential, gave very similar profiles and energy barriers to the ones obtained by scans, and so will not be discussed further here.

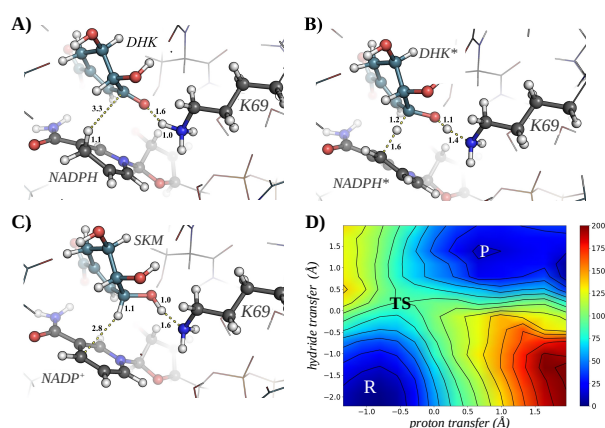


Fig. 5 Two-dimensional PES of the SDH reaction from *M. tuberculosis*. (A) Enzyme-substrate (ES) complex at the beginning of the reaction; (B) transition state (TS) of the reaction; (C) enzyme-product (EP) complex obtained at the end of the reaction. All distances are given in angstroms. (D) two-dimensional PES for the transfer of hydride from NADPH to the C3 carbon of DHK and for the proton from lysine to the O3 oxygen of DHK. Reactants, transition state, and products are highlighted as R, TS and P, respectively. The energies are in kJ mol^{-1} and were obtained by performing B3LYP/6-31+G*/MM single point calculations at the optimized geometries from the AM1d-PhoT/MM PES. The residues involved in the reaction are represented as balls and sticks with CPK colors. The carbon atoms of substrate (DHK) and product (SKM) are colored in green. Image generated with PyMOL [35].

Finally, to obtain an estimate of the free energy of the reaction, we carried out umbrella sampling simulations (Figure 6) for the rate-limiting hydride transfer reaction using the B3LYP/3-21G QC/MM potential. The resulting barrier is $116.9 \text{ kJ mol}^{-1}$, which is about 52 kJ mol^{-1} higher than the experimental one. However, we obtain an adjusted activation free energy barrier of 64.9 kJ mol^{-1} , which is in much better agreement with the experimental barrier of 61.9 kJ mol^{-1} , if we employ a correction equivalent to the difference of the semi-empirical/MM and DFT/MM TS potential energies. This assumes that the dynamical effects contributing to the free energy are well reproduced by the umbrella sampling simulations, but that the static electronic effects are more accurately described by the DFT/MM potential.

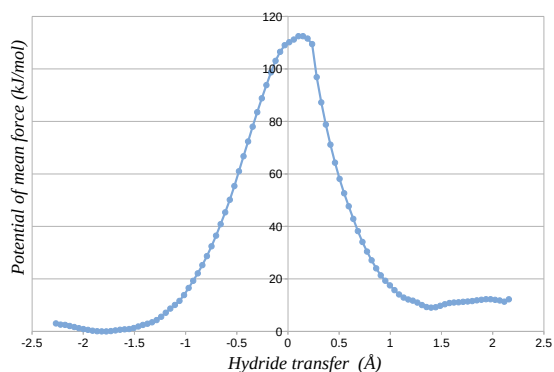


Fig. 6 Reaction coordinate of the MtSDH described by umbrella sampling approach. The initial geometries were retrieved from the relaxed scan calculated using the QC(B3LYP/3-21G)/MM potential. Image generated with Grace [36].

3.3 Reactivity Descriptors

We calculated Fukui functions, electrophilic and nucleophilic attack susceptibilities, and local hardness using the electronic structure of each step of the 2-D coordinate simulated reaction with the semiempirical AM1 Hamiltonian. In Figure 7, we show contour plots of the local hardness and EAS values for the transferred hydride atom, and the EAS for the O3 of DHK which is protonated by Lys69. The transfer of high charge density groups, such as hydrides and protons, is expected to be governed mainly by Coulombic forces and are labelled as hard-hard processes in reactivity theories. In Figure 7A, the local hardness of the hydride atom increases as it is transferred, showing that the electrostatic forces that drive this reaction increases as the hydride atom starts bonding with the C3 atom from DHK. The hardness reaches its maximum value at the TS structure (point P2 of Figure 7A) but is smaller in the parts of the PES where the lysine proton has already been transferred. The propensity of H42(NADPH) to react as a hydride anion, carrying electron density to the DHK from the nicotimide ring of NADPH, is indicated in the EAS contour plot of Figure 7B. The EAS is only high in the TS structure and decreases rapidly as the hydride transfer completes. Finally, Figure 7C displays the EAS contour plot for the O3 oxygen of DHK, which shows that it activates only after a significant displacement in the hydride transfer reaction coordinate or, in other words, that it will tend to abstract the proton from the lysine only after hydride transfer.

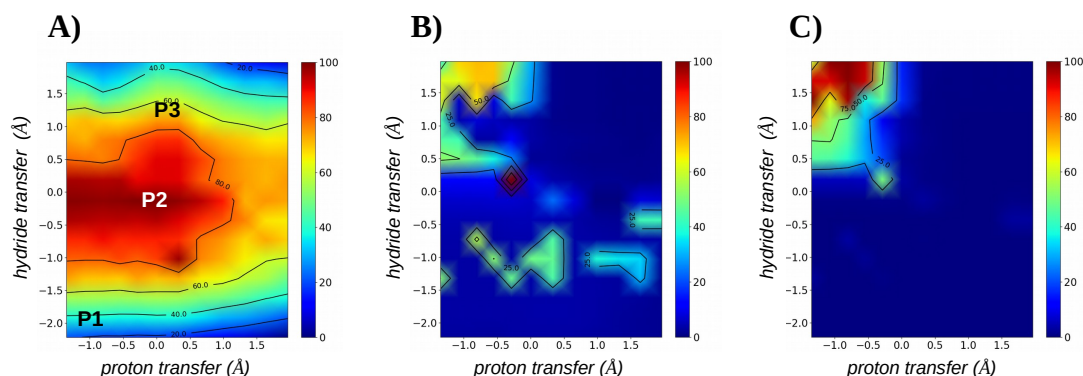


Fig. 7 RD contour plots for the simulated 2-D relaxed surface scan of the MtSDH catalyzed reaction. A) Local hardness for H42(NADPH): P1 — initial structure; P2 — TS structure; P3 — structure where the hydride transfer is almost completed. B) EAS for H42(NADPH). C) EAS for O3(DHK).

These RD results support the observation that the mechanism of the MtSDH reaction occurs in two-steps, with an initial hydride transfer followed by abstraction of the lysine proton by DHK. Figure 8 focuses on the RDs for two key structures from the reaction path. Figure 8A shows the net propensity of attack susceptibility, or dual descriptor, for the initial structure (P1 in Figure 7A). This is calculated as the difference of the EAS and NAS, with positive values (red) indicating regions where an atom is more prone to receive a nucleophilic attack and accommodate electron density, and negative values (blue) regions where it is more prone to suffer an electrophilic attack and donate electron density. In this initial geometry, the nicotinamide ring of NADPH and the aspartic acid, Asp105, show positive values. This can be interpreted as the NADPH not having a propensity to donate charge, and Asp105 near Lys69 being prone to receive electron density thereby increasing the propensity of Lys69 to lose one of its protons. In Figure 8B the partial charges of the initial structure are depicted and show negative partial charges on the nicotinamide ring atoms, but positive charges on the hydrogens, including the one that is expected to be transferred as a hydride. After hydride transfer it can be observed in Figure 8C that NADP has no net reactivity, the DHK atoms have reactivity indicative of a nucleophilic species, and Asp105 continues to present itself as an electrophile. The partial charges for this structure after the hydride transfer are illustrated in Figure 8D, showing the high negative partial charge on the O3(DHK) atom that abstracts the lysine proton.

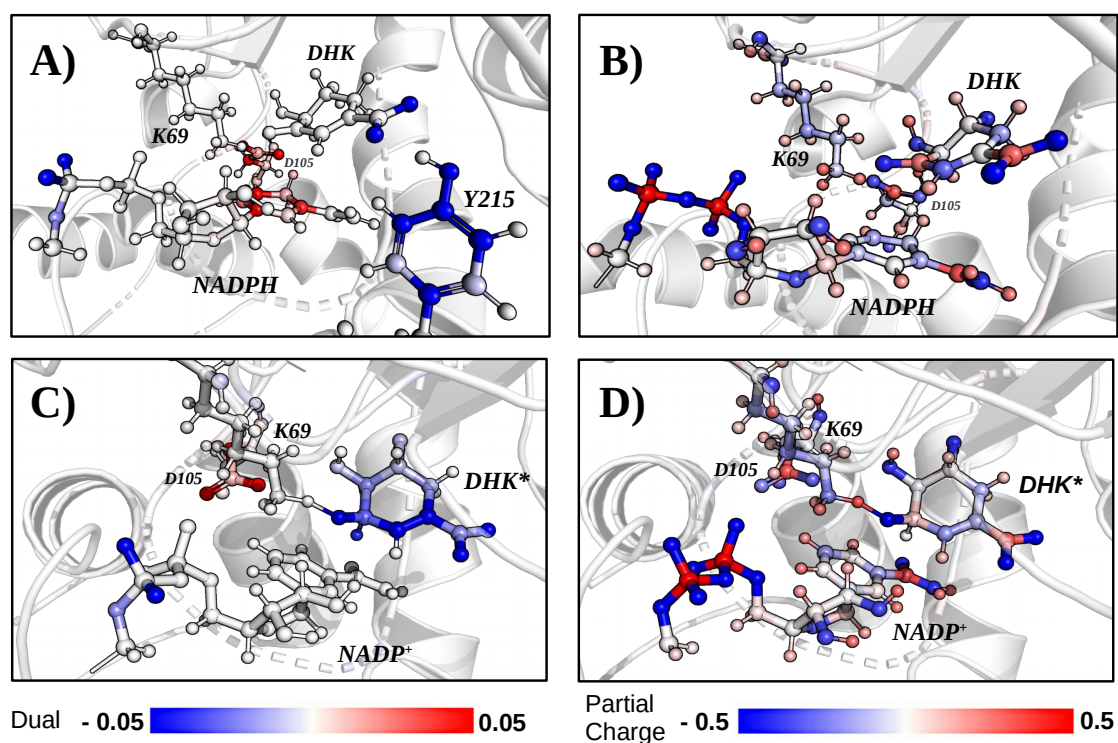


Fig. 8 Normalized reactivity maps for the key structures in the reaction path. A) Net propensity of attack susceptibility (or dual descriptor) for the initial structure from the 2-D surface scan for each atom. B) Partial charges for each atom in the initial structure from the 2-D surface scan. C) Net propensity of attack susceptibility for the structure after the hydride transfer. D) Partial charge for the structure after the hydride transfer for each atom.

The reactivity maps of Figure 8 support the hypothesis that proton transfer is only activated after hydride transfer, and also that charge delocalization from the NADPH to DHK is expected, but only after a certain displacement in the hydride transfer reaction coordinate has occurred. As has already been stated, electrostatic interactions appear to be the main driving force for the hydride transfer reaction, thereby explaining why a charge transfer, a soft-soft interaction driven process, is not predictive of the hydride transfer.

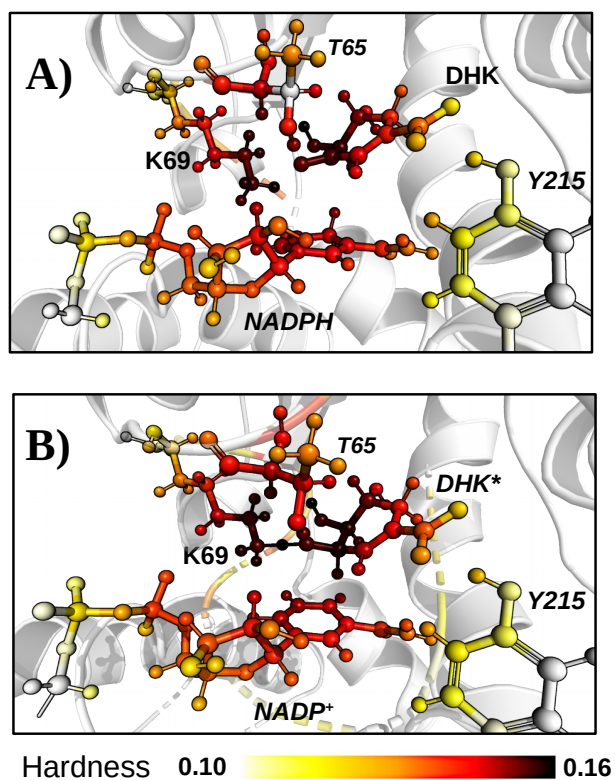


Fig. 9 Local hardness reactivity maps. A) initial structure (P1). B) TS structure (P2).

Figure 9 displays local hardness reactivity maps for the initial (P1) and TS (P2) structures, which highlights the roles of some nearby amino acid residues, such as the threonine, Thr65. Comparing the local hardness for Thr65 between the initial (Figure 9A) and TS (Figure 9B) structures, it is evident that there is an increase in the values of the threonine side chain atoms, which thereby stabilize the TS through electrostatic interactions.

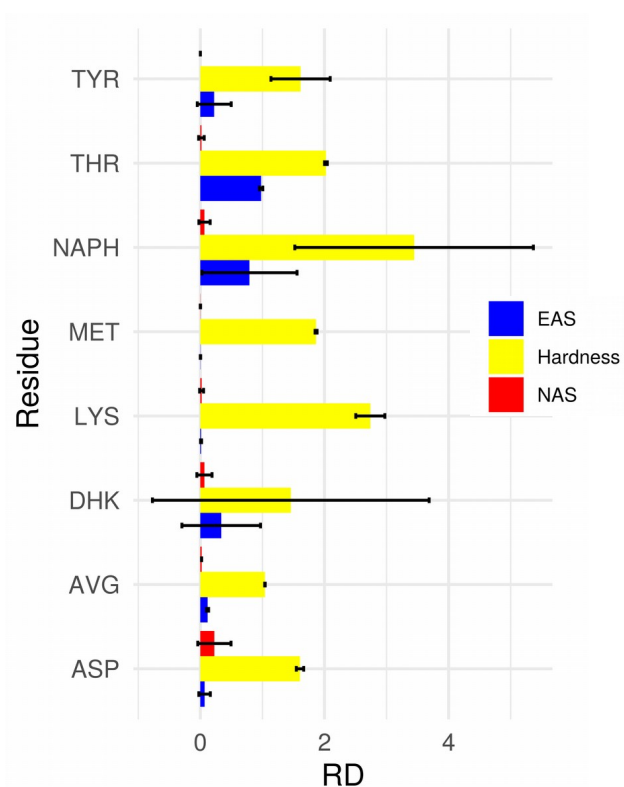


Fig. 10 Average EAS, hardness and NAS values for selected active site amino-acids, substrates and cofactors calculated for all structures of the 2-D relaxed surface scan. AVG denotes the average for all residues and black bars are the standard deviations of the respective averages.

In Figure 10 the average EAS, hardness and NAS values for selected active site amino-acids and substrates are plotted for all structures of the 2-D relaxed surface scan. The averages were obtained by averaging over all atom values in the respective residues. The figure also shows the average RD values over all residues for each structure (denoted by AVG), together with the standard deviation in the structure averages. It can be seen that Thr65 has significant local hardness in all geometries sampled by the surface scan procedure, with low variability, indicating that it is an important residue in TS stabilization. It also has a high EAS value, meaning that it has a propensity to donate electron density. As expected for reactant species, the hardnesses of the substrate, DHK, and the cofactor, NADPH, have high variability. Likewise, Asp105 has significant NAS values for the entire sampled set of geometries, supporting its hypothesized role in increasing the activity of Lys69.

4 Conclusions

In this work we have investigated the reaction mechanism of the shikimate 5-dehydrogenase from *Mycobacterium tuberculosis* using a variety of hybrid QC/MM potentials in combination with relaxed surface scans, free-energy simulations and an analysis of a variety of reactivity descriptors. To our knowledge, this is the first computational study of the reaction in this enzyme. Overall, the methods predict a consensus mechanism in which there is no stable intermediate. Instead the reaction proceeds by an energetically uphill hydride transfer, followed by an energetically downhill reorientation and then transfer of the proton towards the O3 oxygen of DHK. The TS for the reaction corresponds to a structure in which there is almost complete transfer of the hydride. Our mechanism is consistent with double kinetic isotope effect measurements that suggest a single TS and a concerted mechanism, and our DFT-corrected free energy barrier for the reaction is also in good agreement with that estimated from experiment. Overall, our results confirm those obtained experimentally, but they also provide additional information, including insights into the driving force underlying the reaction, the important roles of some peripheral amino acids in the active site, and the structure of the TS structure, which should be valuable for developing more selective inhibitors for this enzyme.

Acknowledgements This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior - Brasil (CAPES) - Finance Code 001. LFSMT acknowledges a postdoctoral scholarship awarded by CAPES, and JFRB acknowledges a postdoctoral scholarship (DOCFIX) awarded by CAPES and by FAPERGS.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Herrmann, K. M., and Weaver, L. M. (1999). The Shikimate Pathway. Annual Review of Plant Physiology and Plant Molecular Biology, 50, 473-503.
2. pass
3. Parish, T., Stoker, N. G. (2002). The common aromatic amino acid biosynthesis pathway is essential in *Mycobacterium tuberculosis*. Microbiology, 148, 3069-3077.
4. Petersen, G. O., Saxena, S., Renuka, J., Soni, V., Yogeewari, P., Santos, D. S., Bizarro, C. V., Sriram, D. (2015). Structure-based virtual screening as a tool for the identification of novel inhibitors against *Mycobacterium tuberculosis* 3-dehydroquinate dehydratase. Journal of Molecular Graphics and Modeling, 60,124-31.
5. Alturki, M. S., Fuanta, N. R., Jarrard, M. A., Hobrath, J. V., Goodwin, D. C., Rantsó, T. A., Calderón, A. I. (2017). A multifaceted approach to identify non-specific enzyme inhibition: Application to *Mycobacterium tuberculosis* shikimate kinase. Bioorganic and Medicinal Chemistry Letters,28, 802-808.
6. Mehra, R., Rajput, V. S., Gupta, M., Chib, R., Kumar, A., Wazir, P., Khan, I. A., Nargotra, A. (2016). Benzothiazole Derivative as a Novel *Mycobacterium tuberculosis*

- Shikimate Kinase Inhibitor: Identification and Elucidation of Its Allosteric Mode of Inhibition. *Journal of Chemical Information and Modeling*, 56, 930-940.
7. Yao, J., Wang, X., Luo, H., Gu, P. (2017). Understanding the Catalytic Mechanism and the Nature of the Transition State of an Attractive Drug-Target Enzyme (Shikimate Kinase) by Quantum Mechanical/Molecular Mechanical (QM/MM) Studies. *Chemistry*, 23, 16380-16387.
 8. Timmers, L. F. S. M., Neto, A. M. S., Montalvão, R. W., Basso, L. A., Santos, D. S., Norberto de Souza, O. (2017). EPSP synthase flexibility is determinant to its function: computational molecular dynamics and metadynamics studies. *Journal of Molecular Modeling*, 23, 197.
 9. Fonseca, I. O., Silva, R. G., Fernandes, C. L., Norberto de Souza, O., Basso, L. A., Santos, D. S. (2007). Kinetic and chemical mechanisms of shikimate dehydrogenase from *Mycobacterium tuberculosis*. *Archives in Biochemistry Biophysics*, 457, 123-133.
 10. Gan, J., Wu, Y., Prabakaran, P., Gu, Y., Li, Y., Andrykovitch, M., Liu, H., Gong, Y., Yan, H., Ji, X. (2007). Structural and Biochemical Analyses of Shikimate Dehydrogenase AroE from *Aquifex aeolicus*: Implications for the Catalytic Mechanism. *Biochemistry*, 46, 9513-9522.
 11. Peek, J., Lee, J., Hu, S., Senisterra, G., Christendat, D. (2011). Structural and mechanistic analysis of a novel class of shikimate dehydrogenases: evidence for a conserved catalytic mechanism in the shikimate dehydrogenase family. *Biochemistry*, 50(40), 8616-8627.
 12. Grillo, I. B., Urquiza-Carvalho, G. A., Chaves, E. J. F., and Rocha, G. B. (2020). Semiempirical methods do Fukui functions: Unlocking a modeling framework for biosystems. *Journal of Computational Chemistry*, 41(9), 862-873.
 13. Grillo, I. B., Urquiza-Carvalho, G., Bachega, J. F. R., and Rocha, G. B. (2020). Elucidating Enzymatic Catalysis using Fast Quantum Chemical Descriptors. *Journal of Chemical Information and Modeling*.
 14. pass
 15. pass
 16. pass
 17. van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., Berendsen, H. J. C. (2005). GROMACS: fast, flexible, and free. *Journal of Computational Chemistry*, 26, 1701-1718.
 18. da Silva, A. W., and Vranken, W. F. (2012). ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Research Notes*. 5, 367.
 19. <http://research.bmh.manchester.ac.uk/bryce/amber/>
 20. Holmberg, Niklas, Ulf Ryde, and Leif Bülow. Redesign of the coenzyme specificity in L-lactate dehydrogenase from *Bacillus stearothermophilus* using site-directed mutagenesis and media engineering. *Protein engineering* 12.10 (1999): 851-856
 21. Olsson, M. H. M., Søndergaard, C. R., Rostkowski, M., Jensen, J.H. (2011). PROPKA3: consistent treatment of internal and surface residues in empirical pK_a predictions. *Journal of Chemical Theory and Computation*, 7, 525-537.
 22. Bussi, G., Donadio, D., Parrinello, M. (2007). Canonical sampling through velocity rescaling. *The journal of chemical physics*, 126, 014101.
 23. Berendsen, H. J. C., Postma, J. P. M., DiNola, A., Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The journal of chemical physics*, 81, 3684-3690.
 24. Hess, B., Bekker, H., Berendsen, H. J. C., Fraaije, J. G. E. M. (1997). LINCS: a linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18, 1463-1472.
 25. Field, M. J. (2007). *A Practical Introduction to the Simulation of Molecular System*. Cambridge, UK: Cambridge University Press.
 26. Field, M. J. (2008). The pDynamo Program for Molecular Simulations using Hybrid Quantum Chemical and Molecular Mechanical Potentials. *Journal of Chemical Theory and Computation*, 4, 1151-1161.
 27. Neese, F. (2011). The ORCA program system. *WIREs Computational Molecular Science*, 2, 73-78.

-
28. Bachega, J. F. R., Timmers, L. F. S. M., Assirati, L., Bachega, L. B., Field, M. J., Wymore, T. (2013). GTKDynamo: a PyMOL plug-in for QC/MM hybrid potential simulations. *Journal of Computational Chemistry*, 34, 2190-2196.
 29. Aleksandrov, A., Field, M. J. (2012). A Hybrid Elastic Band String Algorithm for Studies of Enzymatic Reactions. *Phys. Chem. Chem. Phys.*, 14, 12544-12553.
 30. Parr, R. G., Donnelly, R. A., Levy, M., and Palke, W. E. (1978). Electronegativity: the density functional viewpoint. *The Journal of Chemical Physics*, 68(8), 3801-3807.
 31. pass
 32. pass
 33. pass
 34. pass
 35. Schrödinger, LLC (2010) The PyMOL molecular graphics system, version 1.3r1. Schrödinger, LLC, New York
 36. Turner, P. J. (2005). XMGRACE, Version 5.1.19. Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, OR.