QM/MM study of the stereospecific proton exchange of glutathiohydroxyacetone by glyoxalase I

Jafari, Sonia; Ryde, Ulf; Irani, Mehdi

Published in: Results in Chemistry

DOI: 10.1016/j.rechem.2019.100011

2019

Document Version: Publisher’s PDF, also known as Version of record

Link to publication

Citation for published version (APA):

Creative Commons License:
CC BY

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
QM/MM study of the stereospecific proton exchange of glutathiohydroxyacetone by glyoxalase I

Sonia Jafari a,b, Ulf Ryde b, Mehdi Irani a,*

a Department of Chemistry, University of Kurdistan, P.O. Box 66175-416, Sanandaj, Iran
b Department of Theoretical Chemistry, Lund University, P. O. Box 124, SE-221 00 Lund, Sweden

ARTICLE INFO

Article history:
Received 23 July 2019
Accepted 10 October 2019

Keywords:
Glyoxalase I
Enantioselectivity
Reaction mechanism
QM/MM
Free energy perturbation
DFT

ABSTRACT

We have performed quantum mechanics (QM), molecular mechanics (MM) and hybrid QM/MM calculations to study the stereospecific proton exchange of glutathiohydroxyacetone (HOC-SG) by glyoxalase I (GlxI). We did the QM/MM calculations with a large QM system (246 atoms) to investigate the proton-exchange mechanism. Moreover, single-point big-QM energies with 1303 atoms in the big QM system and 22,412 atoms in the MM system were used to compare the energy difference of the stationary structures. GlxI catalyzes the exchange of the pro-S, but not the pro-R hydroxymethyl proton of HOC-SG with a deuterium from the D2O solvent. Classical molecular dynamics simulations with different protonation states of Glu99, Glu172 and HOC-SG led to the determination of most stable species (Glu-172 is protonated and the alcoholic oxygen of HOC-SG is deprotonated). The QM/MM results showed that before binding of HOC-SG, both active-site glutamates are charged, whereas HOC-SG is protonated. When HOC-SG binds, its alcoholic proton (H0) can point toward either Glu-99 or Glu-172. However, if the substrate binds so that H0 is directed toward Glu-99, it is not transferred, whereas if it is directed toward Glu-172, the latter abstracts H0. The results showed that transferring H0 to the glutamates from the reactant states is the key step to make the proton exchange reaction possible. Our calculations show that order of basicity of the glutamates and HOC-SG inside the enzyme is: Glu-172 > HOC-SG > Glu-99. The calculations allow us to propose a reaction mechanism for the stereospecific proton exchange of HOC-SG by GlxI with an overall barrier of 14.1 kcal/mol.

© 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Enzymes are biological catalysts. Unlike other catalysts, they are usually highly selective and specific for particular chemical reactions. They generally catalyze only one type of reactions; for example, acetylene hydratase catalyzes the non-redox hydration of acetylene in the anaerobic unsaturated hydrocarbon metabolism [1]. Some enzymes are both chemoselective and enantioselective; for example, lipase B from Candida antarctica favors the R enantiomer of 1-phenylethanol [2]. However, only a few enzymes accept both enantiomers of a chiral substrate but convert them to only one enantiomer of the product [3,4]. The typical example is glyoxalase I (EC 4.4.1.5, lactoylglutathione lyase; GlxI), which converts both enantiomers of its substrate into the same enantiomer of its product [3].

Classically, chemoselectivity and enantioselectivity of enzymes can be related to their large, folded, twisted and entangled tertiary and quaternary structures, which give them unique three-dimensional structures. The catalytic activity of enzymes takes place in a relatively small region, the active site. The remaining three-dimensional structure provides a skeleton and modulator for the active site, either by imposing a special shape or electron density to the active-site pocket. A simple understanding of the enantioselectivity of enzymes can be gained from a hollow mold model. A left hand cannot be fitted into a hollow mold of a right hand and likewise an S enantiomer of a substrate cannot bind to the active site of an enantioselective enzyme that is designed to bind the R enantiomer. Of course, this gives a very rough description of enzyme selectivity, neglecting contributions from dynamics and electrostatics effects from the enzymes tertiary and quaternary structures. In addition, it cannot explain selectivity of enzymes with a symmetric active site (e.g. GlxI, the case of this study). On the other hand, computational enzymatic methods can give a detailed and accurate description of enzyme catalysis and selectivity. Such methods have been used to investigate selectivity of some enzymes [5–16]. Herein, we apply quantum mechanics (QM), molecular mechanics (MM) and hybrid QM/MM methods to study the unusual stereospecificity of GlxI.

In the last two decades, different aspects of the catalytic mechanism of GlxI have been studied [3,7,10,17–28]. As shown in Scheme 1, the enzyme accepts both enantiomers of its chiral substrate (hemithioacetals of methylglyoxal with a variety of aromatic and aliphatic α-ketoaldehydes) but gives only one enantiomer of its product (S-D-
Interestingly, crystal structures of the enzyme [22] show a symmetric active site (cf. Fig. 1) with a Zn ion coordinated by Glu-33, His-126, Glu-99 and Glu-172, as well as either two water molecules (Fig. 1a) or an inhibitor (Fig. 1b) [22]. This symmetry can explain why GlxI accepts and converts both S- and R-enantiomers of the substrate. However, it does not explain why the enzyme produces only one enantiomer of the product. On the contrary, we would expect from a symmetric active site to produce both enantiomers of the product and not only one of them. In fact, QM-cluster calculations using rather small models could not explain the unusual stereospecificity of GlxI [10]. This implies that something more complicated takes place inside the enzyme that allows it to produce only one enantiomer of the product. A deeper look at the active site reveals an important asymmetry. Fig. 1b shows that in the crystal structure of an S-[N-hydroxy-N-(p-iodophenyl)carbamoyl]glutathione (HIC-SG) inhibited GlxI, Glu-172 is displaced from the Zn ion, whereas Glu-99 remained coordinated to the ion (the shortest Zn-O distance for the two residues is 3.26 and 1.90 Å, respectively).

Two experimental facts confirm the asymmetry of the active-site glutamate residues: i) Glu-172 or Glu-99 initiate the catalytic reaction of GlxI by abstracting the H1 atom from the S- or R-substrate, respectively (cf. Scheme 2 for the naming of atoms) [22,25–27]. However, experiments showed that isomerization of R- and S-glutathiolactaldehydes by GlxI is characterized by a nonstereospecific proton abstraction followed by a partially shielded proton transfer to the si face of the cis-enediol intermediate [3]. ii) GlxI exclusively catalyzes the exchange of the pro-S (H3) but not the pro-R (H3) hydroxymethyl proton of a product analogue (glutathiohydroxyacetone; HOC-SG) with deuterium from the D2O solvent (Scheme 3). In addition, the enzyme mediates the exchange of the pro-S, but not pro-R hydroxymethyl deuterium of [2H]-HOC-SG with a proton from the H2O solvent [20].

HOC-SG is not the normal substrate of GlxI, but based on experimental observations, Creighton and Hamilton proposed that a substrate analogue is processed in the same way as the normal substrate [27]. Considering this fact, we recently studied the proton-exchange reaction of HOC-SG by GlxI (the reaction shown in Scheme 3) [7]. We employed the QM-cluster approach [29] and classical molecular dynamics (MD) simulations. The results indicated that the unusual stereospecificity of GlxI is caused by the more flexible environment of Glu-172 compared to that of Glu-99 [7]. We used the higher flexibility of Glu-172 to justify the higher basicity of it, which was proposed by Cameron et al. [22] i.e. the higher flexibility of Glu-172 allows it to dissociate from the active-site Zn ion and then attain a higher basicity. The higher flexibility and basicity of Glu-172 allow it to abstract and transfer protons much easier than Glu-99 in the catalytic reaction [7].

Several lines of evidence indicate that Glu-172 is protonated in the crystal structure and in the HOC-SG bounded enzyme: i) In the HIC-SG...
(this molecule has a similar structure to HOC-SG) bounded crystal structure of GlxI [22], the inhibitor is directly coordinated to the Zn ion, displacing Glu-172. It is well known that metals lower the \( pK_a \) of their ligands. For example, the \( pK_a \) value of water in aqueous metal complexes is reduced from 15.7 for bulk water to 12.8 and 2.2 for \( \text{Ca}^{2+} \) and \( \text{Fe}^{3+} \), respectively [30]. Therefore, it is possible that HIC-SG is deprotonated and a proton may actually reside on the dissociated Glu-172. ii) QM-cluster calculations starting from the 1QIN crystal structure and with the normal substrate of GlxI show that Glu-172 is protonated but not Glu-99 in reactant states [7,10]. iii) Calculations starting from a crystal structure with a rather symmetric active site [21] show that if the two glutamate ligands are free to move during the optimization, the protonated Glu-172 will dissociate from the zinc ion [25].

Despite all studies, there still are no quantitative calculations of the amount of basicity of the glutamate residues. In this paper, we extrapolate these hypotheses to a HOC-SG bounded GlxI and use QM/MM methods to investigate if HOC-SG is negatively charged or neutral in the enzyme and if its protonation state depends on the conformation of its alcoholic group. Then, we compare computationally the basicity of the glutamate residues and the bound substrate. Moreover, we also determine a reaction path for the proton-exchange reaction and study how it is affected by the conformation of HOC-SG in the enzyme.

2. Methods

2.1. The protein

All calculations are based on the 2.0 Å crystal structure of human GlxI (PDB code 1QIN) [22], which is composed of two identical and entangled subunits each with 183 residues, two Zn ions and two HIC-SG molecules. This structure was selected because it contains the intermediate analogue (HIC-SG) bound to the active site, which resembles binding of HOC-SG (cf. Fig. 2). The experimental data come from yeast GlxI, but there is no crystal structure for GlxI of this organism. However, Aronsson et al. showed that GlxI from both human and yeast contains Glu-172. The experimental data come from yeast GlxI, but there is no crystal structure for GlxI of this organism. However, Aronsson et al. showed that GlxI from both human and yeast contains Glu-172.

The calculated force constants and charges are given in Tables S1–S3. The stereospeciﬁc protonation states. The MD simulations were performed using the CPU-accelerated pmemd code [35–37] of AMBER 16 [38]. The protein and HOC-SG were described with the Amber ff14SB [39] and GAFF [40] force fields, respectively. Water molecules were described explicitly using the TIP3P model [41]. The Zn sites were treated by a non-bonded model with restraints between the metal and the ligands [42], implemented as NMR bond restraints (nmropt = 1 option and &rst name lists; a restraints file is given in the Supporting Information). The force restraints, used the metal-ligand distances observed in the crystal structure (averaged over the two subunits). In addition, for the metal sites, RESP charges were employed [43], fitted to electrostatic potentials calculated and sampled with the Merz–Kollman scheme [44]. The force constants and RESP charges were obtained from the optimized models in Fig. S1 in the Supporting Information. These calculations were performed at B3LYP-D3/def2-SV(P) [45–50] level of theory, using the Turbomole software (version 7.1) [51]. The force constants were calculated by the method of Seminario [52,53] and averaged over interactions of the same types. The calculated force constants and charges are given in Tables S1–S3 in the Supporting Information. The Amber ff14SB force field employs charges, calculated with the Hartree–Fock method, but it is well-known that metal-sites are poorly described at the Hartree–Fock level of theory. Therefore, we prefer to obtain the charges with a DFT level of theory with significant amount of Hartree–Fock exchange [54].

The setup of the MD simulations is similar to that in our recent works [7,55]. The enzyme was solvated in a periodic truncated octahedral box of TIP3P water molecules, extending at least 14 Å from the solute using the

![Scheme 2](image)

Scheme 2. The proposed initial step of the GlxI reaction, showing that (a) Glu-172 abstracts H1 from the S-substrate and (b) Glu-99 abstracts H1 from the R-substrate.

![Scheme 3](image)

Scheme 3. The stereospecific proton exchange of HOC-SG, catalyzed by GlxI.

![Fig. 2](image)

Fig. 2. Superposition of HOC-SG and HIC-SG from the crystal structure, showing the similarity in size and shape of the two molecules. HOC-SG and HIC-SG are shown in balls and tubes, respectively. Hydrogen atoms are omitted for clarity.
turbulent program in the Amber suite [38]. No counter ions were added to any of the systems. The final system contained ~44,680 atoms. After the solvation, the systems were subjected to 1000 cycles of minimization, with the heavy atoms of the protein restrained toward the starting structure with a force constant of 100 kcal/mol/Å². This was followed by a 20 ps constant-volume and a 20 ps constant-pressure equilibration with the same restraints, but a force constant of 50 kcal/mol/Å². Finally, the systems were equilibrated for 1 ns without any restraints, followed by a 300 ns production simulation, during which coordinates were sampled every 10 ps. The temperature was kept constant at 300 K using Berendsen’s weak coupling isotropic algorithm with a collision frequency of 2 ps⁻¹ [57]. Long-range electrostatics were handled by particle-mesh Ewald summation [58] with a fourth-order B spline interpolation and a tolerance of 10⁻⁶. The cut-off radius for Lennard-Jones interactions was set to 8 Å. All bonds involving hydrogen atoms were constrained to their equilibrium values using the SHAKE algorithm [59], allowing for a time step of 2 fs during the simulations. The root-mean-square deviation (RMSD) was calculated with the AMBER module [60], analyzing trajectories with coordinates which were saved every 10 ps and using the crystal structure as the reference. The reported values are averages over 30,000 sets of coordinates.

For the QM/MM calculations, the protein was instead solvated in a sphere with a radius of 40 Å around the geometrical center of the protein. The final system contained ~23,700 atoms. The positions of the added protons and water molecules, as well as all atoms of HOC-SG were optimized in 100 steps of steepest descent minimization, keeping all other heavy atoms fixed at the crystal-structure positions. The allowed substrate to relax in the active site (it was inserted manually into the enzyme). Then, the added protons and water molecules were optimized by a 240 ps simulated annealing calculation (up to 370 K), followed by a 1000 steps minimization, keeping the other atoms fixed at the crystal-structure positions. No SHAKE restraints were employed and the time step was 0.5 fs.

2.3. QM/MM calculations

The QM/MM calculations were performed with the ComQum software [61,62]. In this approach [63,64], the protein and solvent are split into three subsystems: System 1 (the QM region) was treated by QM methods. System 2 consisted of all residues or water molecules within 6 Å of any atom in system 1. In some calculations (called protein-fixed), system 2 was kept fixed at the original (crystallographic) coordinates. Such calculations were run to ensure that the surroundings residues are in the same local minimum in all calculations, making the energies stable, although all outer-sphere reorganization is ignored. In other calculations (called protein-free), all residues and solvent molecules in system 2 were allowed to relax by a full MM minimization in each cycle of the QM optimization. Note also that the full protein is relaxed in the QTCP calculations (see below). Finally, system 3 contained the remaining part of the protein and the solvent. It was kept fixed at the original coordinates (equilibrated crystal structure).

In the QM calculations, system 1 was represented by a wave function, whereas all the other atoms were represented by an array of partial point charges, one for each atom, taken from MM libraries. Thereby, the polarization of the QM system by the surroundings is included in a self-consistent manner. When there is a bond between systems 1 and 2 (a junction), the hydrogen link-atom approach was employed. In this approach, the QM system was capped with hydrogen atoms (hydrogen link atoms, HL), the positions of which are linearly related to the corresponding carbon atoms (carbon link atoms, CL) in the full system [61,65]. All atoms were included in the point-charge model, except the CL atoms [66]. The point charges do not necessarily sum up to an integer, because the Amber force field does not employ charge groups [67]. The total QM/MM energy in ComQum was calculated as [61,62],

$$E_{\text{QM/MM}} = E_{\text{QM1}}^{\text{HL}} + E_{\text{MM123}}^{\text{CL}} + E_{\text{MM123}}^{\text{q1}} + E_{\text{MM1}}^{\text{q1}} + E_{\text{MM1}}^{\text{q1}}$$

where $E_{\text{QM1}}^{\text{HL}}$ is the QM energy of system 1, truncated by HL atoms and embedded in the set of point charges representing systems 2 and 3 (but excluding the self-energy of the point charges) $E_{\text{MM1}}^{\text{q1}}$ is the MM energy of the QM system, still truncated by HL atoms, but without any electrostatic interactions. Finally, $E_{\text{MM123}}^{\text{q1}}$ is the classical energy of all atoms in the system with CL atoms and with the charges of the QM system set to zero (to avoid double counting of the electrostatic interactions). By using this approach, which is similar to the one used in the Oniom method [68], errors caused by the truncation of the quantum

---

**Fig. 3.** The QM system used in the QM/MM calculations. HOC-SG, the Zn ion and HOH-404 are shown in a ball-and-stick representation and the protein residues are shown as tubes.
system should partly cancel out. Thus, ComQum employs a subtractive scheme with electrostatic embedding and van der Waals link-atom corrections [69].

The QM/MM geometry optimizations were performed using the Turbomole software [51] at the TPSS/def2-SV(P) level of theory [50,70]. The calculations were sped up by expanding the Coulomb interactions in an auxiliary basis set, the resolution-of-identity approximation [71,72] and empirical dispersion corrections were included with the DFT-D3 approach [48] and Becke–Johnson damping [49] as implemented in Turbomole. The MM calculations were performed with the Amber software [67], using the Amber ff14SB [39] and GAFF [40] force fields for the protein and HOC-SG, respectively. Water molecules were described by the TIP3P model [41].

All first-order stationary states (reactants, products and intermediates) were fully optimized without any restraints. To go from one first-order stationary point to another, we moved a proton step by step, applying either an H-O or a H-C restraint with a force constant of 1 a.u. and changing the corresponding target value in steps of 0.2 Å (0.1 Å around the maximum in energy). For example, going from IM1 to IM2, we scanned the H$_\text{O}$-O1 distance in 11 steps. In each step, all other degrees of freedom were optimized. The transition states were approximated as the highest point on the potential energy surface along the reaction coordinates. Scans were performed in both directions, i.e. both from the reactant and from the product, to ensure that the paths are connected and that the hysteresis is minor. Forward and backward reaction profiles are compared in Fig. S2 in the Supporting Information.

The QM system consisted of the Zn ion, models of Gln-33 and His-126 (the glutamine was represented by propanamide and histidine by methyl-imidazole), and all atoms of Thr-97, Leu-98, Glu-99, Leu-100, Thr-101, Trp-170, Ile-171, Glu-172, Ile-173, Leu-174, HOC-SG and HOH-404 (cf. Fig. 3), giving a total of 246 atoms. The QM/MM calculations were performed when system 2 was fixed at the crystallographic coordinates.

We have also performed a large amount of preliminary calculations with 10 other QM systems. Those QM/MM calculations were performed when system 2 was either fixed or free to relax. The calculations led us to choose a proper QM system. The results are collected in the Supporting Information.

2.4. Big-QM calculation

Previous studies have shown that QM/MM energies strongly depend on the size of the studied QM system [66,73]. Neutral residues have a significant effect on energies only when they are within 4.5 Å of the active site [66] but all buried charges in the protein significantly affect the energies [66,73]. Moreover, QM/MM calculations are sensitive to locations of the junctions between the MM and QM system [64,66]. Therefore, we have developed the big-QM method which moves the junctions 2–3 residues away from the QM system and includes all charged residues and all residues with at least one atom within 4.5 Å of any atom in the original QM system [74,75].

In this work, our big-QM system consisted of all chemically reasonable groups (i.e. keeping functional groups, as well as conjugated and aromatic systems intact) with at least one atom within 4.5 Å of any atom in the QM system shown in Fig. 3. Moreover, all junctions were moved at least three residues away from the QM system. There is only one buried charge group in GlxI that is neither in the QM systems nor is coordinated to a metal ion (Arg-37). This group was also included in the big-QM system. This gave a system of 1303 atoms, shown in Fig. 4. The big QM energy calculations were performed on the QM/MM optimized structures. The calculations were performed as described above, but they employed also the multiple-accelerated resolution-of-identity $f$ approach (marij keyword) [76]. The calculations were run with the parallel version of Turbomole [77]. To the big-QM energies, we added the DFT-D3 dispersion correction and a MM correction (E$\text{MM1}_{\text{IM123,q1'=0}}$ – E$\text{MM1}_{\text{IM1,q1'=0}}$), yielding a standard QM/MM energy but with the big-QM system as the QM region.

2.5. QM/MM-PBSA calculations

The QM/MM-PBSA method [78,79] can be viewed as a post-processing of QM/MM calculations to obtain more stable energies, including a solvation term obtained by continuum methods. It is an adaptation of the widely used MM/PBSA approach [80,81] for QM/MM calculations. Two variants of QM/MM-PBSA were used. They differ in how the QM energy and the charges on the QM atoms are calculated [33,78,82]. In the first approach, denoted by Vac in the following, the QM energy and the charges were calculated from a vacuum wave function. In the second approach, denoted Pitch below, the QM wave function is obtained with a point-charge model of the MM system. The calculations were automatized and performed by Linux shell scripts, which are available from the authors on request. Further details of the QM/MM-PBSA calculations can be found in the Supporting Information and in http://www.teokem.lu.se/~ulf/Methods/qmmm_pbsa.html.

![Fig. 4. Atoms included in the big-QM system.](image)

![Scheme 4. The thermodynamic cycle employed in the QTCP calculations.](image)
First, each state of interest was optimized by QM/MM, keeping QM part was subjected to a 1000-step minimization, keeping the atoms in the at least 9 Å from the QM/MM system. For the reactant state, the system was solvated in an octahedral box of TIP3P water molecules, extending the coordinates at those of the reactant state. Then, the coordinates were collected every 2 ps. 200 snapshots were run for each state. During the production run, 200 snapshots were run with a constant pressure equilibrated the size with a constant volume, the second with a constant pressure. Next, a 100-ps MD simulation with a constant pressure equilibrated the size of the periodic box and only the heavy atoms in the QM system restrained to the QM/MM structure. The final structure of this simulation was used as the starting structure also for the other state. Finally, an equilibration of 200 ps and a production of 400 ps were run with a constant volume for each state. During the production run, 200 snapshots were collected every 2 ps.

Based on these 200 snapshots, three sets of FEPs were performed as has been described before [83–85]. First, each state of interest was optimized by QM/MM, keeping QM/MM system at MM level and two FEP calculations in method space from MM to QM/MM, one each for the I and the J states.

$$\Delta G_{QMM}^{\text{MM}} (I\rightarrow J) = \Delta G_{\text{MM}} (I\rightarrow J) - \Delta G_{\text{QMM}} (I) + \Delta G_{\text{QMM}} (J)$$

The QTCP calculations were performed as has been described before [33,85]. First, each state of interest was optimized by QM/MM, keeping systems 2 and 3 fixed at the crystal structure. Then, the protein was further solvated in an octahedral box of TIP3P water molecules, extending at least 9 Å from the QM/MM system. For the reactant state, the system was subjected to a 1000-step minimization, keeping the atoms in the QM part fixed and restraining all heavy atoms from the crystal structure with a force constant of 100 kcal/mol Å². Then, two 20 ps MD simulations were run with the heavy atoms still restrained. The first was run with a constant volume, the second with a constant pressure. Next, a 100-ps MD simulation with a constant pressure equilibrated the size of the periodic box and only the heavy atoms in the QM system restrained to the QM/MM structure. The final structure of this simulation was used as the starting structure also for the other state. Finally, an equilibration of 200 ps and a production of 400 ps were run with a constant volume for each state. During the production run, 200 snapshots were collected every 2 ps.

Based on these 200 snapshots, three sets of FEPs were performed as is shown in Scheme 4. First, FEPs were performed at the MM level in the forward and reverse direction along the reaction coordinate by changing the charges and coordinates of the QM system to those of the QM/MM calculations [33]. Charges were first modified in nine steps, keeping the coordinates at those of the reactant state. Then, the coordinates were modified in five steps to those of the product state (with the charges in the product state). Finally, MM → QM/MM FEPs were performed for both the reactant and product states, keeping the QM systems fixed, as has been described before [84,85]. All FEP calculations were performed with the local software calcqtcp. Further details of the QTCP calculations can be found in http://signe.teokem.lu.se/~ulf/Methods/qtcp.html.

### 3. Results and discussion

The results are presented in three subsections. First, we describe our tests made to decide a proper protonation state and conformation of the alcoholic oxygen of HOC-SG (O1). Glu-99 and Glu-172 in the active site. For the starting crystal structure, we get a deprotonated substrate and a protonated Glu-99 and Glu-172 (Scheme 5c and a, respectively). Additionally, the H3 atom may be transferred to the carboxylate groups of either of the glutamates, giving a deprotonated substrate and a protonated Glu-99 and Glu-172 (Scheme 5d and a, respectively). Thus, four conformations are possible and we call them A, B, C and D states as is summarized in Table 1 and shown in Scheme 5.

We have performed 300 ns MD simulations for each of these four protonation states. Then, we calculated mass weighted RMSDs from the starting crystal structure for the heavy atoms of Glu-99 and Glu-172, as has been done before for His residues in three proteins [86].

### Table 1

Protonation states of the glutamates and HOC-SG in the MD simulations.

<table>
<thead>
<tr>
<th></th>
<th>Glu-99 protonation state</th>
<th>Glu-172 protonation state</th>
<th>HOC-SG protonation state</th>
<th>HOC-SG conformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Deprotonated</td>
<td>Protonated</td>
<td>Deprotonated</td>
<td>-</td>
</tr>
<tr>
<td>B</td>
<td>Deprotonated</td>
<td>Deprotonated</td>
<td>Protonated</td>
<td>H3 to 172</td>
</tr>
<tr>
<td>C</td>
<td>Deprotonated</td>
<td>Deprotonated</td>
<td>Protonated</td>
<td>H3 to 99</td>
</tr>
<tr>
<td>D</td>
<td>Protonated</td>
<td>Deprotonated</td>
<td>Deprotonated</td>
<td>-</td>
</tr>
</tbody>
</table>

### Table 2

Mass-weighted RMSD values (Å) of the two active-site glutamate residues for the four MD simulations differing in the protonation state (cf. Table 1). Glu-99A and Glu-172A belong to active site 1 and Glu-99B and Glu-172B belong to active site 2. Plots of the time dependence of the RMSDs are shown in Figs. S7–10 in the Supporting Information.

<table>
<thead>
<tr>
<th></th>
<th>Active site 1</th>
<th>Active site 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Glu-99A</td>
<td>Glu-172A</td>
</tr>
<tr>
<td>A</td>
<td>0.49</td>
<td>0.43</td>
</tr>
<tr>
<td>B</td>
<td>0.33</td>
<td>0.91</td>
</tr>
<tr>
<td>C</td>
<td>0.37</td>
<td>0.89</td>
</tr>
<tr>
<td>D</td>
<td>0.28</td>
<td>1.03</td>
</tr>
</tbody>
</table>
and for homocitrate and nearby residues in nitrogenase [55]. The RMSD of heavy atoms from the crystal structure gives information about which protonation state is more realistic. The philosophy is that if an incorrect protonation state is employed, the atoms in that residue or in nearby residues will move to release steric or electrostatic clashes or to form new favorable interactions [86]. Therefore, the RMSD will be higher for incorrect protonation states (provided that the reference structure is representative for the state studied).

The calculated RMSDs are summarized in Table 2. It can be seen that the RMSD of Glu-99 does not vary significantly in the four simulations and two subunits (0.28–0.49 Å; cf. Table 2). However, RMSDs of the Glu-172 residue are much lower in the A state simulation than in the others (0.43 and 0.58 compared to 0.89–1.06 Å). Thus, the MD results indicate that Glu-172 is probably protonated and HOC-SG is deprotonated (owing to the lower RMSD of this residue in the A state simulation) in the HOC-SG bounded enzyme. The differences between the two active sites give an impression of the precision of the results. They are 0.02–0.03 Å for most of the protonation states, but up to 0.06 Å for the B state and 0.09–0.15 Å for the A state.

To obtain further support for the preferred protonation state of the glutamate residues and HOC-SG, we performed also QM/MM, big-QM, QM/MM-PBSA and QTCP calculations for the four states shown in Scheme 5. The calculated relative energies are collected in Table 3. All methods suggest that A is the most stable protonation state, whereas the D state is very unstable (7–9 kcal/mol less stable than the A state). Moreover, we could not find any QM/MM local minima for the B and D states (the energies of these states in Table 3 were calculated when the distance between H₂ and the carboxylate group of the corresponding Glu residue was constrained to 1.0 Å).

Thus, our results suggest the following scenario for the binding of HOC-SG to the GlxI: Before the binding, both active-site Glu residues are deprotonated and therefore negatively charged, whereas HOC-SG is protonated. When HOC-SG binds, the proton can point toward either Glu-99 or Glu-172, i.e. the H₂to99 or H₂to172 conformations to the active site. However, our calculations show that in the H₂to172 conformation, the alcoholic proton is abstracted by Glu-172 resulting in the A state, whereas in the H₂to99 conformation, the proton stays on HOC-SG (state C).

### 3.2. Reaction paths

Next, we performed QM/MM calculations with a large QM system (246 atoms) to investigate the proton-exchange mechanism of HOC-SG by GlxI. This method includes electrostatic effects of the surroundings in a self-consistent manner in the wave function, which is more realistic than the implicit surrounding in QM-cluster calculations. Moreover, single-point big-QM energies with 1303 atoms in the QM system were used to compare the energy difference of the stationary structures along the reaction path. These would give a more reliable reaction path than that of the previous QM-cluster calculations [7]. The energies discussed in Sections 3.2.1 and 3.2.2 and the reaction energy profile in Fig. 5 are based on the big-QM energies. The corresponding QM/MM and QTCP results are shown in Fig. S11 in the Supporting Information (they give a similar reaction profile).

Among the protons of the substrate analogue, H₂ can be exchanged noncatalytically by a deuterium (it is an alcoholic proton and can be exchanged by a deuterium from D₂O in the reaction medium). Therefore, we suppose that H₂ has already been exchanged by a deuterium atom before entering the enzyme active site. In the following, we examine how the protein manages to exchange this proton with Hg, but not with H₈ (atom names are shown in Scheme 5).

#### 3.2.1. Reaction paths from the H₂to99 conformation

As described in the previous section, the most stable state of the H₂to99 conformation is C, in which H₂ is on O1 of HOC-SG and is directed toward Glu-99. The schematic structure of the C state is shown in Scheme 5c. Starting from the C state, there are three possible proton...
transfers (H$_2$O and H$_2$ to Glu-99 and H$_2$ to Glu-172; these are represented by arrows in Scheme 5c). Our calculations showed that H$_2$O could not move to Glu-99 in the first step for the H$_2$O-to99 conformation (the proton transfer shown by a red arrow in Scheme 5c). This would produce the D state, which is unstable and can only be obtained when the H$_2$O and Glu-99 distance be constrained. The second proton transfer (H$_2$ to Glu-99) resulted in the structure that is shown in Scheme 6a. From the resulting state shown in Scheme 6a, there are two possible proton transfers (H$_2$ to Glu172 and H$_2$ to C1). The former transfer is unlikely because it would lead to a two-coordinated C1 atom which energetically is not feasible (the transfer shown by the red arrow in Scheme 6a) and the later will go back to the C state (the transfer shown by the green arrow in Scheme 6a). Therefore, this path was not tested.

The third proton transfer from the C state (H$_2$ to Glu-172) produced the structure shown in Scheme 6b. The barrier and reaction energy for this step was 5.3 and 3.9 kcal/mol, respectively. As can be seen from Scheme 6b there is only one reasonable proton transfer from this structure (H$_2$ to Glu-99; the transfer shown by a red arrow in Scheme 6b). This proton transfer was unsuccessful (H$_2$O returned to the starting point on O1). In conclusion, there is no path for the H$_2$O-to99 conformation to exchange H$_2$ by a deuterium.

### 3.2.2. Reaction paths from the H$_2$O-to172 conformation

For the H$_2$O-to172 conformation, we started the calculations from the A protonation state in Scheme 5a, as the results in section 3.1 showed that when HOC-SG binds in the H$_2$O-to172 conformation to the enzyme, its alcoholic proton is abstracted by Glu-172, resulting in the A state. Starting from the A state, there are two reasonable proton transfers (H$_2$O to O1 and H$_2$O to Glu-99; these are represented by arrows in Scheme 5a). Our calculations showed that H$_2$O could not move to O1 in the first step (the proton transfer shown by a red arrow in Scheme 5a); this would produce the B state, which is 1.2 kcal/mol less stable than A and can only be obtained if the H$_2$O and O1 distance is constrained. The second proton transfer (H$_2$O to Glu-99) would result in the structure that is shown in Scheme 7a (IM1). The energy barrier for this step is 14.1 kcal/mol and IM1 is 0.4 kcal/mol more stable than the A state (cf. Fig. 5 for the energies). The barrier for this proton abstraction is 8.8 kcal/mol higher than the corresponding proton abstraction from the C state (H$_2$ from C1 by Glu-172 in the C state; 14.1 vs 5.3 kcal/mol). This shows that Glu-172 is a stronger nucleophile than Glu-99 (the nucleophiles in the C and A states are Glu-172 and Glu-99, respectively).

From IM1, there are two reasonable proton transfers (H$_2$O to O1 and H$_2$O to C1; the transfers are shown by arrows in Scheme 7a). The results showed that the barrier of the former transfer is 13.3 kcal/mol and reaches to a second intermediate (IM2; the structure is shown in Scheme 7b) which is 8.4 kcal/mol higher than IM1. From IM2 there is only one reasonable proton transfer (H$_2$O to C1; shown by an arrow in Scheme 7b). This transfer leaves a product in which H$_2$O is exchanged by H$_3$O$_+$ (Pr; shown in Scheme 7c), in agreement with experimental findings (we supposed that H$_2$O was previously exchange by a deuterium before entering HOC-SG to the active site). This step has a low barrier (1.6 kcal/mol). The Pr state is a variant of the C state in Scheme 5c in which H$_2$O, H$_2$O, and H$_2$O are replaced by H$_3$O$_+$, H$_2$O, and H$_2$O respectively (compare the structures of Scheme 5c and Scheme 7c). The two states are almost degenerate (C is only 0.8 kcal/mol higher than Pr).

Our results showed that the other transfer from the IM1 state (H$_2$O to C1; shown by a bold green arrow in Scheme 7a) directly reaches to the Pr state (H$_2$O simultaneously moves to O1 while moving H$_2$O to C1). However, this proton transfer has a quite high barrier of 18.4 kcal/mol.

According to the results, there is no path from the H$_2$O-to99 conformation of HOC-SG to exchange its H$_2$O by a deuterium. On the other hand, there are two alternative paths from the H$_2$O-to172 conformation to a proton exchanged product (B $\rightarrow$ A $\rightarrow$ IM1 $\rightarrow$ IM2 $\rightarrow$ Pr and B $\rightarrow$ A $\rightarrow$ IM1 $\rightarrow$ Pr). The former has a lower overall barrier (14.1 vs. 18.0 kcal/mol; cf. Fig. 5), which would make it ~5000 times faster than the later (using the Arrhenius equation at 300 K). Thus, the second path is unlikely. The calculated overall energy barriers are in reasonable agreement with the observed solvent exchange rate constant [20] ($k = 1$ s$^{-1}$, corresponding to an activation barrier of ~17 kcal/mol).

Based on the results, we propose a mechanism for the stereospecific proton exchange of HOC-SG by GlxI as is shown in Scheme 8. First, HOC-SG exchanges non-catalytically its alcoholic proton with a deuterium from the solvent, before entering to the active site. Next, the substrate binds to the enzyme, giving either the B or C state. The later does not accomplish any path to reach the proton exchanged product. However, the former passes the deuterium to Glu-172, resulting the Pr state. Then, H$_2$O is abstracted by Glu-99, resulting IM1 (this is the rate determining step with a barrier of 14.1 kcal/mol). After that, H$_2$O moves to O1, resulting IM2. Then, the deuterium is transferred from Glu-172 to C1, leading to the proton exchanged product (the Pr state). Finally, Pr dissociates from the active site, allowing the enzyme to

![Diagram](image-url)
start a new catalytic cycle. According to this mechanism, the product will always have the deuterium in the pro-S position, in accordance with the experimental findings [20]. The proposed mechanism in Scheme 8 is slightly different from the mechanism proposed based on QM-cluster calculations [7]. They are the same in all states. However, in the IM1 state of the QM-cluster based mechanism, H_o is on O1 instead of being on O6.

3.3. Flexibility of the glutamates in the reaction path

We have previously suggested that the origin of the unusual stereospecificity of GlxI is the higher flexibility of Glu-172 compared to Glu-99 (our molecular dynamics simulations showed that RMSD and RMSF values of Glu-172 are higher than those of Glu-99 and that the flexible loops inside the enzyme are closer to Glu-172) [7]. Here, we have further investigated this hypothesis for the reaction path shown in Scheme 8. We calculated RMSD of heavy atoms of the two active-site glutamate residues in the consecutive QM/MM stationary states (i.e. going from the crystal structure to B, from B to A, from A to IM1, from IM1 to IM2, and from IM2 to Pr) and collected them in Table 4. We can see that in all steps, except the third one (A → IM1), the RMSD values of Glu-172 are greater than those of Glu-99. The average RMSDs of Glu-172 is larger than that of Glu-99 along the reaction path (0.35 vs. 0.22 Å; cf. the last column of Table 4). This shows that Glu-172 is more flexible not only in the crystal structure but also along the reaction path.

Table 4

<table>
<thead>
<tr>
<th>Step</th>
<th>RMSD Glu-99 (Å)</th>
<th>RMSD Glu-172 (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crystal → B</td>
<td>0.21</td>
<td>0.80</td>
</tr>
<tr>
<td>B → A</td>
<td>0.08</td>
<td>0.11</td>
</tr>
<tr>
<td>A → IM1</td>
<td>0.25</td>
<td>0.16</td>
</tr>
<tr>
<td>IM1 → IM2</td>
<td>0.40</td>
<td>0.44</td>
</tr>
<tr>
<td>IM2 → Pr</td>
<td>0.15</td>
<td>0.24</td>
</tr>
<tr>
<td>Average</td>
<td>0.22</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Scheme 8. Our proposed mechanism for the proton exchange reaction of HOC-SG by GlxI, based on the QM/MM calculations.
4. Conclusions

In this investigation, we have applied classical MD simulations and QM/MM optimizations, together with Big-QM, QM/MM-PBSA and QTCP calculations to study the stereospecific proton exchange of HOCSG by Gxl. MD simulations with different protonation states of Glu-99, Glu-172 and HOCSG led to the determination of most stable species. QM/MM results showed that if the substrate binds in the H2to172 conformation, H2 is abstracted by Glu-172, whereas if it binds in H2to99 conformation, H2 remains on O1 and is directed toward Glu-99.

Our QM/MM results showed that there is no reaction path from the H2to99 conformation led to any proton exchange (the key proton transfers were unsuccessful e.g. H2 to Glu-99). However, for the H2to172 conformation we found a feasible reaction path (Scheme 8). The results showed that transferring H2 to the glutamates from the reactant states is the key step to make the proton exchange reaction possible. This proton transfer is spontaneous in the H2to172 conformation (resulting A from B), but is not possible for the H2to99 conformation. Our calculations show that the order of basicity of the substrate and the glutamates is Glu-172 > HOCSG > Glu-99 (cf. the results in Table 3).

Finally, we also investigated the hypothesis of the higher flexibility of Glu-172 within the reaction path suggested by the QM/MM calculations. The results showed that in the reaction path, the RMSDs of Glu-172 are greater than those of Glu-99. This shows that Glu-172 is more flexible than Glu-99, not only in the crystal structure, but also along the suggested reaction path.

In summary, this and our recent studies on the catalytic mechanism of Gxl [7,10] give the following view of the proton exchange reaction of HOCSG by Gxl: The active site Zn ion is symmetrically coordinated by Gln-33, His-126, the two active site glutamates and two water molecules before binding any substrate. The substrate analogue binds to the active site displacing the two coordinated water molecules. However, binding HOCSG in its H2to172 conformation dissociates Glu-172 from the Zn ion. This takes place by abstracting the aliphatic proton of HOCSG by Glu-172. On the other hand, for HOCSG bound in the H2to99 conformation, nothing happens. This is because Glu-99 is a weaker base than the aliphatic oxygen of HOCSG and therefore it cannot abstract H2 and pass it to C1. On the other hand, HOCSG binds in the H2to172 conformation and exchanges H2 by a deuterium from the medium as is shown in Scheme 8.

Acknowledgements

This investigation has been supported by grants from the Swedish National Infrastructure for Computing (SNIC) at Lunarc at Lund University and at HPC2N at Umeå University. The calculations were performed on computer resources provided by the Swedish National Infrastructure for Computing (SNIC) at Lunarc at Lund University and at HPC2N at Umeå University.

Appendix A. Supplementary data

Protonation states of the protein residues in the MD simulations and QM/MM calculations, details of the MD simulations, a restraints file for the Zn–ligand bond restraints in the MD simulations, optimized models used to calculate force constants and charges for the MM force field, MM charges, force constants and restraints used for the Zn clusters, comparison of the forward and backward reaction profiles, details of the QM/MM-PBSA calculations, QM/MM results of ten other QM systems, plots of time dependence of RMSDs of the glutamates in the MD simulations, comparison of the QM/MM, big-QM and QTCP reaction profiles, QTCP free energy profiles for several of the QM systems. Supplementary data to this article can be found online at doi: https://doi.org/10.1016/j.jrecchem.2019.100011.


