VIP

Electrostatic Perturbations in the Substrate-Binding Pocket of Taurine/α-Ketoglutarate Dioxygenase Determine its Selectivity

Hafiz Saqib Ali[a, b] and Sam P. de Visser*[a, c]

Abstract: Taurine/α-ketoglutarate dioxygenase is an important enzyme that takes part in the cysteine catabolism process in the human body and selectively hydroxylates taurine at the C^1-position. Recent computational studies showed that in the gas-phase the C^2–H bond of taurine is substantially weaker than the C^1–H bond, yet no evidence exists of 2-hydroxytaurine products. To this end, a detailed computational study on the selectivity patterns in TauD was performed. The calculations show that the second-coordination sphere and the protonation states of residues play a major role in guiding the enzyme to the right selectivity. Specifically, a single proton on an active site histidine residue can change the regioselectivity of the reaction through its electrostatic perturbations in the active site and effectively changes the C^1–H and C^2–H bond strengths of taurine. This is further emphasized by many polar and hydrogen bonding interactions of the protein cage in TauD with the substrate and the oxidant that weaken the *pro-R* C^1–H bond and triggers a chemoselective reaction process. The large cluster models reproduce the experimental free energy of activation excellently.

Introduction

The nonheme iron dioxygenases belong to the large and most diverse class of mononuclear metalloenzymes that utilize molecular oxygen.[1] In general, there are two types of oxygenases, namely those that utilize both atoms of molecular oxygen (O_2) in substrate activation, i.e. the dioxygenases, and those that use only one oxygen atom of O_2 to activate the substrate while the other oxygen atom is reduced to a water molecule, i.e. the monooxygenases.[2] Some of these oxygenases, particularly the dioxygenases, often require a co-substrate, such as α-ketoglutarate (α-KG) to generate the active species in the catalytic cycle that reacts with substrate.[1,3] These so-called α-KG-dependent dioxygenases generally contain a ferrous iron and react with α-KG through oxidative decarboxylation to form succinate and with release of carbon dioxide produce a high-valent iron(IV)-oxo species. Many biomimetic iron(IV)-oxo models have been created and studied and it was shown that they are highly reactive and often react with substrates through oxygen atom transfer.[1d,g,h,4] The most common mechanism in nonheme iron dioxygenases is the activation of an aliphatic C—H bond that is then converted into an alcohol.[1–5] The α-KG-dependent nonheme iron dioxygenases are an important class of metalloenzymes which are present in almost all forms of life and carry out a range of important biochemical transformations. For instance, they are involved in biosynthesis reactions in the human body, including that of 4-hydroxyproline, which is an essential component in collagen formation.[6] In bacteria, the α-KG-dependent nonheme iron dioxygenases have been implicated with the biosynthesis of antibiotics,[7] while in plants hormones such as flavonols are synthesized by a series of enzymes that includes several nonheme iron dioxygenases.[5a]

Taurine/α-KG-dependent dioxygenase (TauD) is one of the best examples of an α-KG-dependent dioxygenase and has been extensively studied.[1,3,5] The human TauD isozyme takes part in the cysteine metabolism pathway, where it hydroxylates taurine selectively at the *pro-R* C^1-position. In a subsequent reaction step hydroxytaurine splits off sulfate to trigger the sulfur recycling mechanism in the body.[8] Figure 1 shows the extract of the active site of TauD alongside its general hydroxylation reaction mechanism. The TauD active site structure shown in Figure 1 was taken from the crystal structure coordinates deposited under the 1OS7 protein data bank (pdb) file.[9] The central iron (Fe) atom in TauD is connected to the protein via covalent bonds with two histidine (His_{99} and His_{255}) and the carboxylate group of an aspartic acid (Asp_{101}) residue. In addition, the iron(II) binds α-KG to give the iron a pentacoordinated ligand structure. Substrate taurine is bound
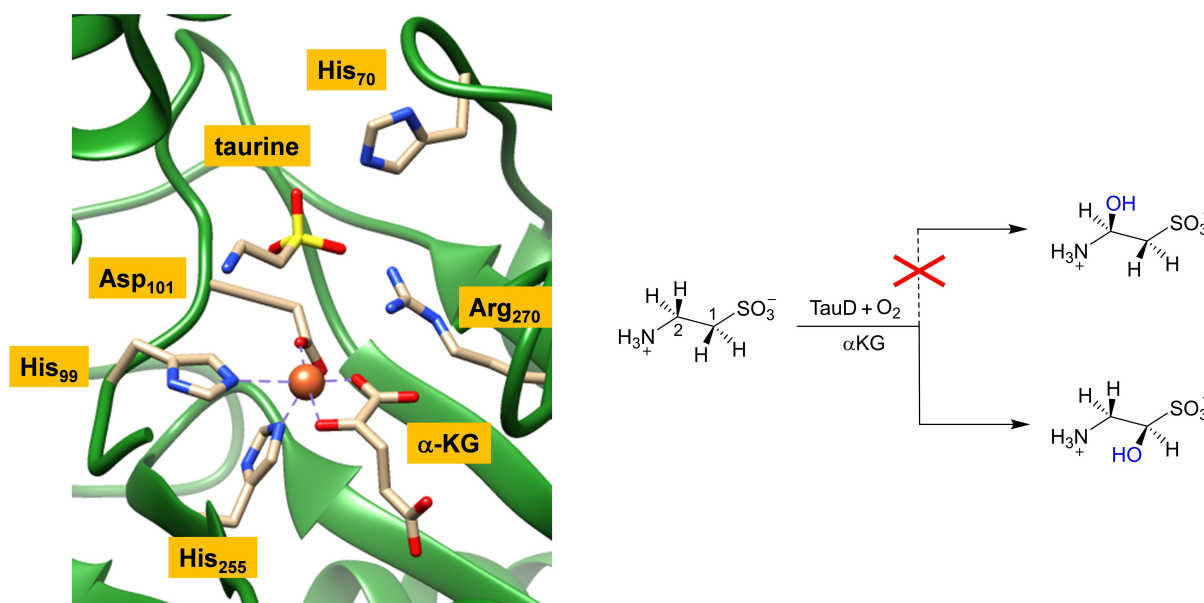
[a] Dr. H. S. Ali, Dr. S. P. de Visser
Manchester Institute of Biotechnology
The University of Manchester
131 Princess Street, Manchester M1 7DN (UK)
E-mail: sam.devisser@manchester.ac.uk

[b] Dr. H. S. Ali
Department of Chemistry
The University of Manchester
Oxford Road, Manchester M13 9PL (UK)

[c] Dr. S. P. de Visser
Department of Chemical Engineering and Analytical Science
The University of Manchester
Oxford Road, Manchester M13 9PL (UK)

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202104167

Chemistry Europe
European Chemical
Societies Publishing

Figure 1. *Left*: Active site structure of TauD enzyme as taken from the 1OS7 pdb file. *Right*: General reaction mechanism of hydroxylation of taurine by TauD and products obtained.

close to the iron center and forms hydrogen bonding interactions with the side chains of His$_{70}$ and Arg$_{270}$. Experimental studies showed that TauD uses α-KG and dioxygen to react on the iron(II) center to form an iron(IV)-oxo active species.[10] It was characterized for TauD using UV-Vis absorption, Mössbauer spectroscopic, electron paramagnetic resonance and extended X-ray absorption fine structure (EXAFS) studies.[10] Moreover, kinetic isotope effect studies and the measurements of rate constants for substrate activation identified the iron(IV)-oxo species as the active species of TauD.[5c,10a] In particular, the iron(IV)-oxo species was found to abstract a hydrogen atom from taurine at a rate of $k_{exp} = 13$ s^{-1} at 5 °C.
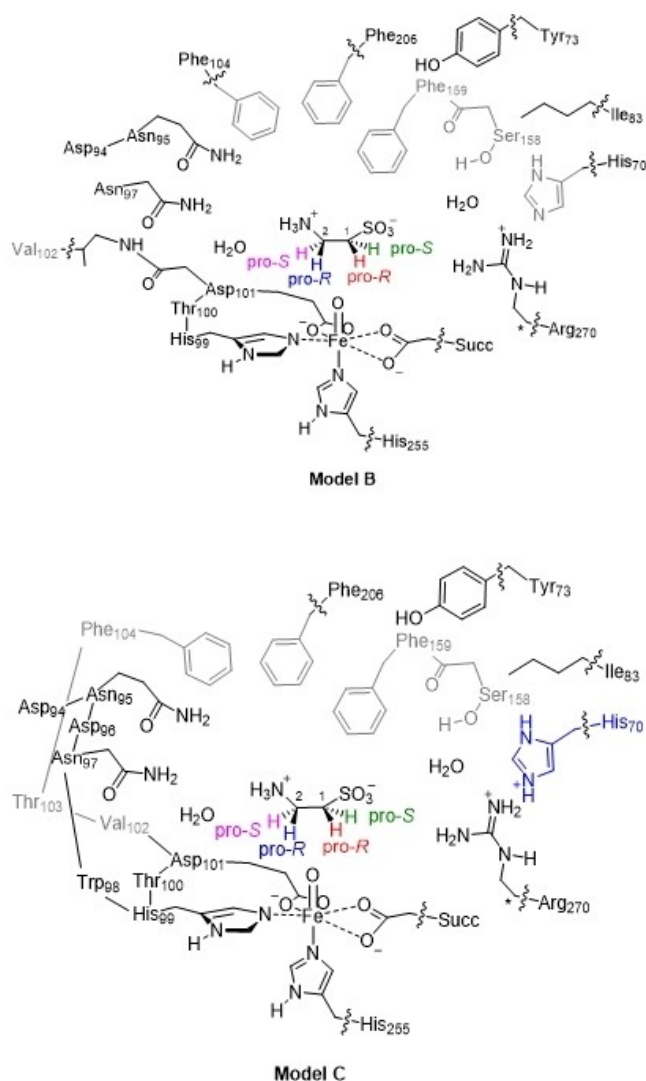
Many computational studies have subsequently focused on the catalytic cycle of TauD and the C^1–H substrate activation pathways using density functional theory (DFT) cluster models as well as quantum mechanics/molecular mechanics methods (QM/MM).[11] All of these studies, however, only considered *pro-R* C^1–H activation of taurine, even though taurine has four aliphatic C–H bonds in close proximity to the iron(IV)-oxo active species that are chemically different. A recent report showed that in the gas-phase the C^2–H bond of taurine is considerably weaker than the C^1–H bond.[12] Based on these bond strengths, TauD should react through C^2-hydroxylation rather than C^1-hydroxylation, which it does not do. Therefore, TauD reacts through negative catalysis, whereby a thermodynamically unfavorable channel is catalyzed over one that has a larger driving force.[12,13] The question is how TauD manages this negative catalysis and how the thermodynamically favorable C^2–H activation channel is blocked.

To find out why and how TauD reacts regio- and stereo-selectively on the *pro-R* C^1-position of taurine, and whether alternative pathways are feasible, we decided to do a computational study on the substrate hydroxylation of taurine at the

pro-R C^1–H, *pro-S* C^1–H, *pro-R* C^2–H and *pro-S* C^2–H positions using a combination of density functional theory (DFT) on cluster models and QM/MM on complete enzymatic structures. Furthermore, we investigate how electrostatic perturbations of, for example, the protein, can trigger a selective reaction process through studying residues in different protonation states. Our work shows that electrostatic perturbations of charged residues in the second-coordination sphere of the protein guide an otherwise unselective reaction process to a regio- and stereo-selective reaction mechanism. In particular, the protonated histidine group in the substrate binding pocket is vital to direct the local dipole moment in such a way to guide the reaction to C^1-hydroxylation.

Results and Discussion

To gain mechanistic insight into the substrate selectivity process of taurine hydroxylation by TauD enzymes, we performed a series of DFT and QM/MM calculations. As shown by kinetic isotope effect measurements on the enzyme, the taurine activation is performed by an iron(IV)-oxo species.[10a] Therefore, we created the corresponding iron(IV)-oxo species (**Re**) using quantum chemical cluster models but also performed QM/MM on the full structure. The models were based on the 1OS7 pdb file[9] and after initial set-up that included a molecular dynamics (MD) simulation (see details in the Methods section and Supporting Information), we established that the protein structure and substrate placement are highly rigid. We initially created three cluster models based on the last step of the MD simulation, namely a minimal model **A** and two larger cluster models **B** and **C** (Figure 2). Data on the small cluster model **A** can be found in the Supporting Information, while we focus on

Chem. Eur. J. **2022**, *28*, e202104167 (2 of 15)

© 2021 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202104167
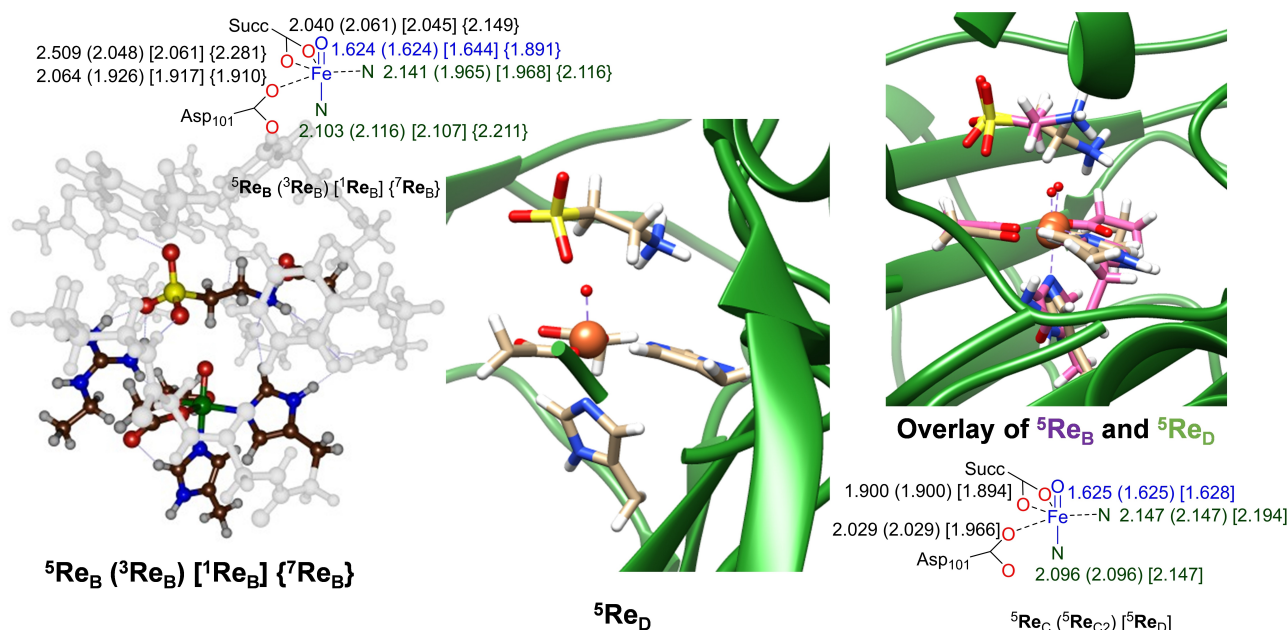
Chemistry
Europe
European Chemical
Societies Publishing

Figure 2. DFT cluster models **B** and **C** used in this work. Wiggly lines show where bonds were cut and link-atoms were added. Atoms labelled with a star were kept fixed during the geometry optimizations.

the larger cluster models only. The large cluster models had 244 (model **B**) and 279 (model **C**) atoms in total and contain the iron(IV)-oxo and substrate with their first- and second coordination sphere included. In particular, several protein chains were part of the cluster models, whereby the amino acid residues of those chains that are pointing into the active site were included in the model, while all other residues from the chosen chains were truncated to Gly residues. The cluster models, therefore, include the hydrogen bonding network around the substrate and oxidant that positions these groups in the active site. In model **B** the His$_{70}$ side chain is singly protonated while it is doubly protonated in model **C**. For optimal comparison, we did some test calculations, whereby we took the model **C** structures and removed the proton from His$_{70}$: Model **C2**.

Subsequently, a geometry optimization using four DFT cluster models (**Re$_A$**, **Re$_B$**, **Re$_C$** and **Re$_{C2}$**) and one QM/MM model **Re$_D$** was performed. A comparison of the optimized geometries

of **Re$_B$**, **Re$_C$**, **Re$_{C2}$** and **Re$_D$** are given in Figure 3. The cluster models **Re$_B$** were calculated in multiple spin states. The quintet spin state ^{5}**Re$_B$** is the ground state with the triplet spin state (^{3}**Re$_B$**) higher in energy by 10.4 kcal mol^{-1} at UB3LYP-D3/BS2// UB3LYP/BS1 + ZPE + E$_{solv}$ level of theory with ZPE = zero-point energy and E$_{solv}$ = solvation energy, while the singlet and septet structures are higher in energy by 19.2 and 10.9 kcal mol^{-1}. Our calculated electronic configuration of the iron(IV)-oxo species, therefore, is in agreement with electron paramagnetic resonance (EPR) studies that characterized the iron(IV)-oxo species as a quintet spin state.[10a] Geometrically, the ^{5}**Re$_B$** structure has the carboxylate group of succinate bound as a bidentate ligand with Fe–O distances of 2.040 and 2.509 Å. Similar distances are seen for the interactions of the carboxylate group of the Asp$_{101}$ ligand with iron, while the two nitrogen ligands (of His$_{99}$ and His$_{255}$) are at an Fe–N distance of 2.141 and 2.103 Å. The iron(IV)-oxo bond is short at 1.624 Å. This value is in excellent agreement with experimental EXAFS studies that identified an Fe–O distance of 1.62 Å.[10c] The singlet and triplet geometries similarly to the quintet spin state have short Fe–O distances of 1.644 and 1.624 Å, while the septet spin state has the distance elongated to 1.891 Å. The iron first-coordination sphere distances in models **C** and **C2** are virtually identical; hence addition of a second-coordination sphere proton does not affect the electronic configuration and geometry of the iron center. Using model **C** and QM/MM we find short Fe–O distances of 1.625 and 1.628 Å, while the Fe–N distances are 2.147/2.194 (equatorial) and 2.096/2.147 (axial) Å, respectively. As such, structures ^{5}**Re$_B$**, ^{5}**Re$_C$** and ^{5}**Re$_D$** have a similar first-coordination sphere. Indeed, an overlay of the optimized geometries of ^{5}**Re$_B$** and ^{5}**Re$_D$** (right-hand-side of Figure 3) puts most atoms in a similar position and shows that the first-coordination sphere of the two models as well as the protein chains are matching well. Our optimized geometries also compare well to previous DFT and QM/MM calculations on the iron(IV)-oxo species of TauD as well as analogous αKG-dependent nonheme iron dioxygenases.[11,14]

The resonance Raman difference spectrum generated of TauD in a reaction with ^{16}O$_2$ gave a fingerprint for the iron(IV)-oxo stretch vibration at 821 cm^{-1} that dropped to 787 cm^{-1} when the heavy isotope ^{18}O$_2$ was used.[10b,15] A geometry optimization of ^{5}**Re$_B$** using a modest LANL2DZ basis set with core potential on iron and 6-31G on the rest of the atoms reproduces the experimental vibrational spectrum well and gives an Fe–O bond of 1.655 Å and an Fe–O stretch vibration of 817 cm^{-1}. At the UB3LYP/BS1 level of theory the calculated vibrational frequencies of ^{5}**Re$_B$** for Fe=^{16}O versus Fe=^{18}O gives a drop of the iron(IV)-oxo stretch vibration by 38 cm^{-1} using the recommended scaling factor of 0.95 of Scott and Radom.[16] The calculated vibrational frequency difference is in good agreement with the difference spectrum measured by Proshlyakov et al.[15] As such, we decided to continue our studies with UB3LYP/BS1 as it matches the experimental structure the best. In particular, our calculated optimized geometry, vibrational frequencies and electronic configuration of the quintet spin state excellently reproduces experimental characterization.
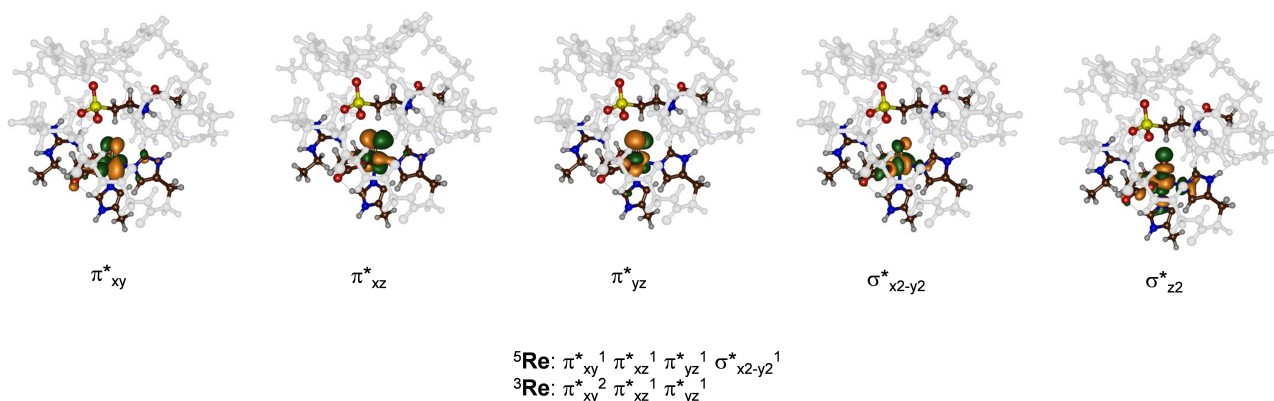
Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202104167

Chemistry
Europe
European Chemical
Societies Publishing

Figure 3. Optimized geometries of iron(IV)-oxo complexes calculated in this work. Left hand-side: DFT cluster results of ^{1,3,5,7}**Re**$_B$. Middle and bottom-right: Optimized geometries of ^{5}**Re**$_C$, ^{5}**Re**$_{C2}$ and ^{5}**Re**$_D$. Right-hand-side: overlay of the ^{5}**Re**$_B$ and ^{5}**Re**$_D$ optimized geometries. Bond lengths are given in angstroms.

Iron(IV)-oxo complexes have close-lying spin states as a result of their orbital occupation, which is dependent on their ligand features. Thus, nonheme iron(IV)-oxo species in octahedral ligand environment typically have a triplet spin ground state with the quintet spin higher in energy.[17] By contrast, trigonal bipyramidal coordinated iron(IV)-oxo is often found in a quintet spin ground state. Furthermore, the effect of solvation and external perturbations has been shown to affect spin-state ordering and reactivities dramatically.[18] The ordering of the spin states in the reactants and local minima along the reaction mechanism may affect the kinetics and as such all structures were calculated in multiple spin states.[19]

Figure 4 displays the high-lying occupied and virtual molecular orbitals that determine the spin state ordering in the reactants complex. All contain a metal 3d-contribution and hence are labelled with the 3d-type. These orbitals are the π^*_{xy}, π^*_{xz}, π^*_{yz}, σ^*_{x2-y2} and σ^*_{z2}. The lowest in energy are the three π^* orbitals for the antibonding interactions of the metal 3d orbitals with the $2p_x$ and $2p_y$ atomic orbitals on the oxo group along the z-axis (π^*_{xz} and π^*_{yz}), while the π^*_{xy} orbital is in the equatorial plane. In addition, there is the σ^*_{x2-y2} orbital for the antibonding interactions of the $3d_{x2-y2}$ orbital on iron with σ-orbitals in the xy-plane from the nitrogen of His$_{99}$ and the oxygen atoms of Asp$_{101}$ and succinate. Finally, the σ^*_{z2} orbital along the z-axis is based on the interactions of the $3d_{z2}$ on iron with the $2p_z$ on the oxo group and the $2p_z$ on the nitrogen atom of the axial histidine ligand (His$_{255}$). In the quintet spin ground state the structure has electronic configuration of $\pi^*_{xy}{}^1$ $\pi^*_{xz}{}^1$ $\pi^*_{yz}{}^1$ $\sigma^*_{x2-y2}{}^1$, while the σ^*_{z2} orbital is virtual. By contrast, the triplet spin state has configuration $\pi^*_{xy}{}^2$ $\pi^*_{xz}{}^1$ $\pi^*_{yz}{}^1$. The septet

^{5}**Re**: $\pi^*_{xy}{}^1$ $\pi^*_{xz}{}^1$ $\pi^*_{yz}{}^1$ $\sigma^*_{x2-y2}{}^1$
^{3}**Re**: $\pi^*_{xy}{}^2$ $\pi^*_{xz}{}^1$ $\pi^*_{yz}{}^1$

Figure 4. Molecular valence orbitals of the iron(IV)-oxo species of TauD.

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202104167

Chemistry Europe
European Chemical
Societies Publishing

spin state can be reached from the quintet spin state by a promotion of an electron from the doubly occupied π_O orbital on the oxo group to σ^*_{z2}. For our TauD model the optimized singlet and septet spin state structures are high in energy and hence the singlet and septet spin states were not considered further.

DFT cluster model B calculations on the stereoselectivity of taurine hydroxylation

Next, we considered taurine hydroxylation at the C^1 and C^2 positions of the substrate which we investigated for the cluster model **B**. Thus, we studied the activation of all four C–H bonds of taurine, namely the *pro-R* as well as the *pro-S* sites of the substrate. These C–H bonds are defined as C1R, C1S, C2R and C2S and give hydroxytaurine with the corresponding stereochemistry. We calculated a stepwise mechanism with an initial hydrogen atom abstraction leading to a radical intermediate (**IM1**) via a transition state **TS1**$_{HA}$. In a subsequent OH rebound step via transition state **TS2**$_{reb}$ the hydroxytaurine product complexes (**P$_H$**) are reached. Details of the mechanism and the nomenclature of the structures are given in Scheme S1 (Supporting Information). The potential energy landscape for taurine hydroxylation leading to *R*-1-hydroxytaurine, *S*-1-hydroxytaurine, *R*-2-hydroxytaurine and *S*-2-hydroxytaurine for model ^{5}**Re$_B$** is shown in Figure 5. In agreement with previous

computational studies and experimental observation,[10a] the rate-determining step is the hydrogen atom abstraction via ^{5}**TS1**$_{HA}$. Interestingly, we find all four hydrogen atom abstraction barriers within a window of 3.0 kcal mol^{-1}. In particular, the lowest energy barrier is for the *pro-S* C^2–H hydrogen atom abstraction (^{5}**TS1**$_{HA,B,C2S,\sigma}$) with a magnitude of 9.3 kcal mol^{-1}. About 1.6 kcal mol^{-1} higher in energy, is the barrier passing ^{5}**TS1**$_{HA,B,C2R,\sigma}$, which leads to the same radical intermediate ^{5}**IM1**$_{C2,B}$ as the transition state ^{5}**TS1**$_{HA,B,C2S,\sigma}$. The hydrogen atom abstraction steps from the C^2-position to reach ^{5}**IM1**$_{C2,B}$ is exothermic by $\Delta E + ZPE = -8.1$ kcal mol^{-1} with respect to ^{5}**Re$_B$**, while the radical intermediate at C^1 (^{5}**IM1**$_{C1,B}$) is -8.8 kcal mol^{-1} more stable. The lowest energy C^1–H hydrogen atom abstraction barrier is from the *pro-S* C^1–H group and has a barrier of 11.8 kcal mol^{-1}, while the *pro-R* hydrogen abstraction is at 12.2 kcal mol^{-1} with respect to ^{5}**Re$_B$**. As such the calculations predict dominant C^2-hydroxylation of taurine, although some small amount of C^1-hydroxylation could be possible. The results of the calculations on model **B**, therefore, are in disagreement with the natural system as only C^1-hydroxylation products are observed. Therefore, alternative models and methods were applied that investigated this discrepancy.

The optimized transition state structures are given in Figure 5 as well and show the typical features of hydrogen atom abstraction transition states by nonheme iron(IV)-oxo complexes.[17,20] In all structures the Fe–O bond has elongated to 1.78 Å in the transition state and the substrate approaches
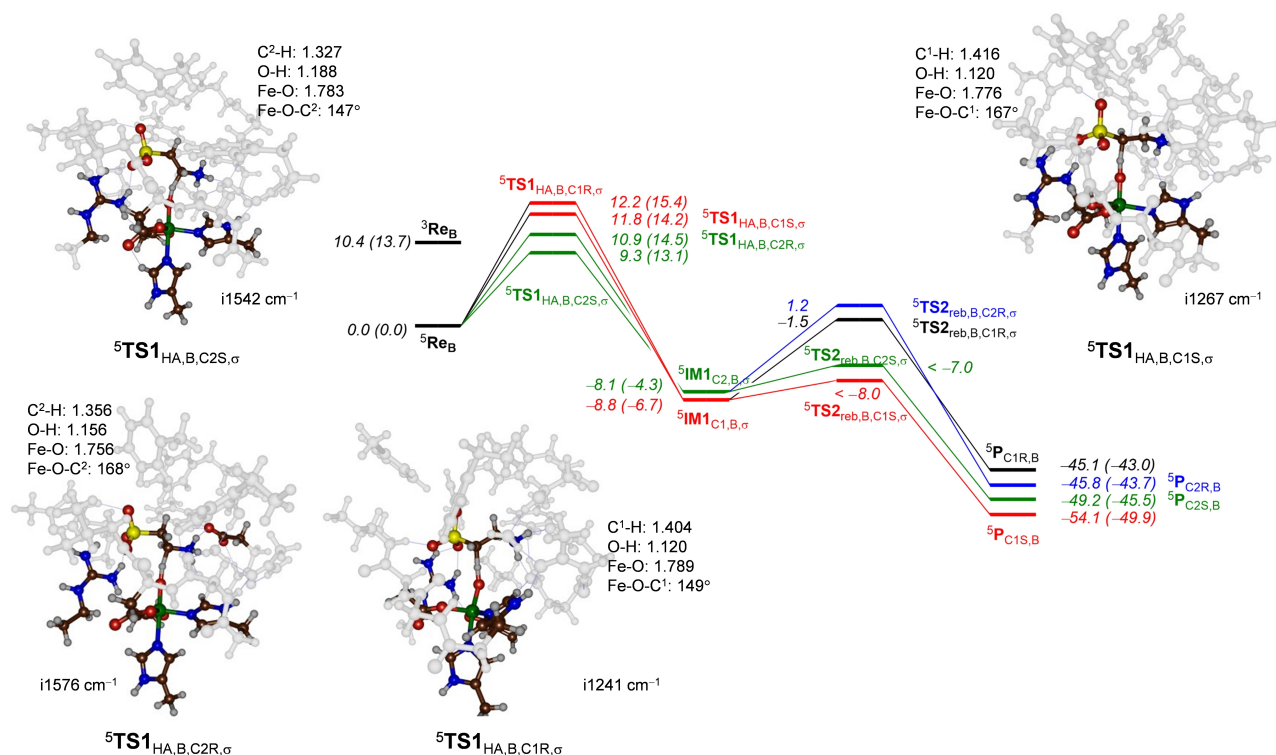
Figure 5. UB3LYP-D3/BS2//UB3LYP/BS1 calculated potential energy landscape for taurine hydroxylation at the *pro-R* and *pro-S* C^1–H and C^2–H positions of taurine in ^{5}**Re$_B$** as calculated in Gaussian-09. Energies (in kcal mol^{-1}) contain solvent, dispersion and ZPE corrections, while free energies are given in parenthesis. Optimized geometries of the hydrogen atom abstraction transition states give bond lengths in angstroms, angles in degrees in the imaginary frequency in cm^{-1}.

Chem. Eur. J. **2022**, 28, e202104167 (5 of 15)

© 2021 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202104167

Chemistry
Europe
European Chemical
Societies Publishing

the oxidant under a large Fe—O—C angle that ranges from 147° for $^5\text{TS1}_{\text{HA,B,C2S},\sigma}$ to 167° for $^5\text{TS1}_{\text{HA,B,C1S},\sigma}$. In general, all structures are product-like with a short O—H and long C—H distances. For all four hydrogen atom abstraction pathways, a radical intermediate with electronic configuration $\pi^*_{xy}{}^1 \pi^*_{xz}{}^1 \pi^*_{yz}{}^1 \sigma^*_{x2-y2}{}^1 \sigma^*_{z2}{}^1 \phi_{\text{Sub}}{}^1$ is formed, whereby the substrate orbital (ϕ_{Sub}) contains a down-spin electron, while the other orbitals have a single up-spin electron. In previous studies this pathway was called the $^5\sigma$-pathway and it often shows attack of substrate from the top as the σ^*_{z2} orbital is being filled.[21] Structures of hydrogen atom abstraction transition states, radical intermediates and rebound transition states with this electronic configuration have the subscript σ added to the label.

The alternative electron transfer during the hydrogen atom abstraction leads to radical intermediates $\textbf{IM1}_{\text{C1,B},\pi}$ and $\textbf{IM1}_{\text{C2,B},\pi}$ with configuration $\pi^*_{xy}{}^2 \pi^*_{xz}{}^1 \pi^*_{yz}{}^1 \sigma^*_{x2-y2}{}^1 \sigma^*_{z2}{}^0 \phi_{\text{Sub}}{}^1$, whereby the substrate radical is up-spin and ferromagnetically coupled to the metal 3d unpaired electrons. The latter configuration is called the $^5\pi$-pathway configuration and usually has a transition state structure with more bent Fe—O—C angle (typically around 120°).[21] Although we were unable to characterize any $^5\pi$-type hydrogen atom abstraction transition states, we were able to find the radical intermediates $^5\textbf{IM1}_{\text{C1,B},\pi}$ and $^5\textbf{IM1}_{\text{C2,B},\pi}$ that have a $^5\pi$-type electronic configuration. These two radical intermediates are well higher in energy than the $^5\textbf{IM1}_{\text{C1,B},\sigma}$ and $^5\textbf{IM1}_{\text{C2,B},\sigma}$ structures shown in Figure 5 by 16.2 and 15.6 kcal mol^{-1}, respectively. Therefore, the $^5\pi$-pathway is a high energy pathway that is inaccessible for our TauD model studied here. These pathways also have substantial rebound barriers of well over 12 kcal mol^{-1} with respect to the reactants complex (Supporting Information, Table S6) and hence the $^5\pi$-pathway will not be able to compete with the $^5\sigma$-pathway for this oxidant.

Despite the fact that the lowest energy barriers correspond to C2S and C2R hydrogen atom abstraction pathways, the lowest energy intermediate is the C^1-radical although by only a small margin ($\Delta E + ZPE = 0.7$ kcal mol^{-1}). Clearly, the substrate binding pocket and the protein environment bind and position the substrate in a specific orientation to guide the reaction towards a specific regio- and stereoselectivity. After the radical

intermediates a small rebound barrier for the C1S and C2S pathways leads to hydroxytaurine product complexes, while the C1R and C2R pathways have considerable rebound barriers of 7.3 and 9.3 kcal mol^{-1} as estimated from constraint geometry scans. The rebound steps have large exothermicity and generate products within a small margin of energy with the most stable product complex to be the S-1-hydroxytaurine product. All hydrogen atom abstraction transition states calculated for model **B** have a large imaginary frequency (> i1000 cm^{-1}) and hence the reaction is expected to proceed with a large amount of quantum chemical tunneling and will be affected by the substitution of a hydrogen atom by deuterium.

We repeated several of the geometry optimizations for the hydrogen atom abstraction step with different density functional theory methods and basis sets, however, all these calculations (see Supporting Information) predicted the same product distributions and did not change the order of the transition states.

DFT cluster model C calculations on the regioselectivity of taurine hydroxylation

As the DFT cluster calculations on Model **B** predict the wrong product distributions as compared to experiment, we decided to expand the cluster structure to model **C** that has most of the protein chains around substrate and oxidant included and therefore restricts substrate approach. Moreover, the model has the active site His$_{70}$ residue in its doubly protonated form. Figure 6 gives the hydrogen atom abstraction energy landscape for the pathways for activation of the C1R, C1S, C2S and C2R positions of taurine as calculated with model **C**. As can be seen the C^1—H hydrogen atom abstraction barriers are by far the lowest at $\Delta E + ZPE = 14.5$ and 15.8 kcal mol^{-1}, while the two transition states for C^2—H hydrogen atom abstraction are well higher at 21.2 and 23.1 kcal mol^{-1}. As such, the model **C** calculations predict dominant C^1-hydroxylation of taurine in agreement with experimental observation. As the margins between the C^1—H and C^2—H hydrogen atom abstraction
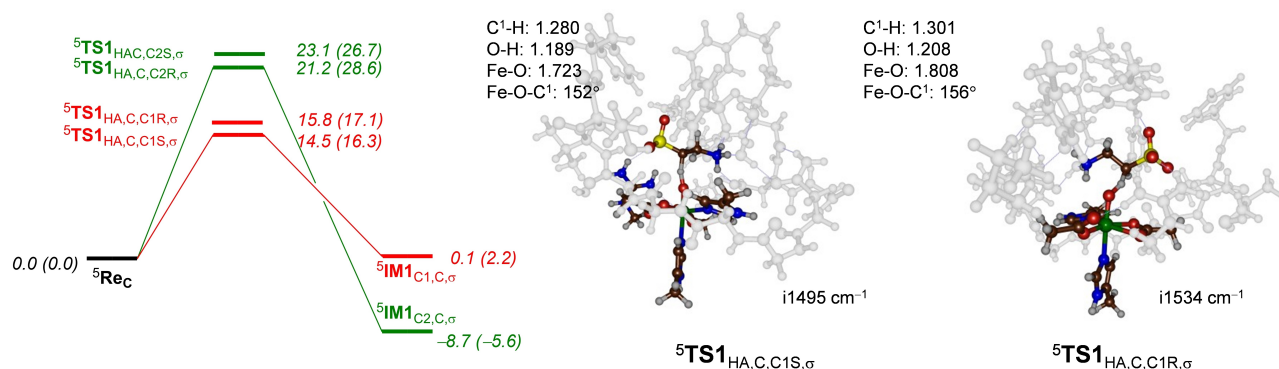
Figure 6. UB3LYP-D3/BS2//UB3LYP/BS1 calculated potential energy landscape for taurine hydroxylation at the *pro-R* and *pro-S* C^1—H and C^2—H positions of taurine in $^5\text{Re}_C$ as calculated in Gaussian-09. Energies (in kcal mol^{-1}) contain solvent, dispersion and ZPE corrections, while free energies are given in parenthesis. Optimized geometries of the hydrogen atom abstraction transition state give bond lengths in angstroms, angles in degrees in the imaginary frequency in cm^{-1}.

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202104167

**Chemistry
Europe**
European Chemical
Societies Publishing

barriers are large, the probability of obtaining C^2-hydroxylated products appears negligible. Interestingly, the radical intermediate $^5IM1_{C2,C,\sigma}$ is more stable than $^5IM1_{C1,C,\sigma}$ by almost 9 kcal mol^{-1}. Therefore, the order of the radical intermediates and transition states is opposite and TauD appears to react by negative catalysis, where a thermodynamically less favorable pathway is kinetically favored.

The optimized geometry of the $^5TS1_{HA,C,C1S,\sigma}$ transition state is shown on the right-hand-side of Figure 6. It has a large imaginary frequency of i1495 cm^{-1} that indicates the reaction will proceed with a significant amount of quantum tunneling and will have a large kinetic isotope effect. The geometry has a relatively central hydrogen atom at a distance of the donor carbon atom of 1.280 Å and at a distance of 1.189 Å to the acceptor oxo group. The C^1–H–O angle is 152°. The *pro-R* hydrogen atom abstraction transition state also has a large imaginary frequency of i1534 cm^{-1} and consequently, will experience a large kinetic isotope effect. The structure is relatively central with similar O–H and C–H distances, i.e. 1.208 and 1.301 Å, respectively. After the formation of the C^1-radical ($^5IM1_{C1,C,\sigma}$) a favorable rebound pathway with a negligible barrier leads to *R*-1-hydroxytaurine products selectively in agreement with experimental observation. It appears, therefore, that large cluster models that incorporate a significant number of atoms of the second coordination sphere in TauD and particularly a charged His$_{70}$ residue give a good representation of the actual enzyme. To test the effect of the proton on the His$_{70}$ residue, we created model **C2** that is model **C** without the proton on His$_{70}$. Subsequently, the reactants and hydrogen atom abstraction transition states were reoptimized. Indeed, we see a change in regioselectivity in model **C2** as compared to model **C**. In particular, the lowest hydrogen atom abstraction barriers for model **C2** are the C^2–H hydrogen atom abstraction barriers by more than 3 kcal mol^{-1} over the C^1–H barriers. Therefore, the choice of the model in TauD and particularly the inclusion of protonated groups in the substrate binding pocket such as His$_{70}$ are essential for the proper description of the long-range electrostatic effects of the active site and influence barrier heights and selectivities.

QM/MM calculations on the stereoselectivity of taurine hydroxylation

In addition to the DFT cluster calculations, we also decided to study the structure and reactivity of TauD with QM/MM methods on the full protein. Figure 7 gives the hydrogen atom abstraction transition state structures for the C1R, C1S, C2S and C2R pathways of TauD as calculated with QM/MM using model **D** and minimal QM region **A**. In agreement with experimental observation the QM/MM calculated reaction mechanisms give the $^5TS1_{HA,DA,C1R,\sigma}$ barrier as the lowest in energy. In particular, the $^5TS1_{HA,DA,C1R,\sigma}$ and $^5TS1_{HA,DA,C1S,\sigma}$ barriers are at $\Delta E + ZPE = 8.3$ and 8.5 kcal mol^{-1} with respect to the iron(IV)-oxo reactant complex and are considerably lower in energy than the ones for the C^2–H hydrogen atom abstractions, namely $\Delta E + ZPE = 24.6$ kcal mol^{-1} for $^5TS1_{HA,DA,C2S,\sigma}$ and $\Delta E + ZPE = 23.8$ kcal mol^{-1} for $^5TS1_{HA,DA,C2R,\sigma}$. These energetic differences implicate that the reaction will be highly selective and virtually no C^2-hydroxylation will happen by the enzyme. As such, the QM/MM calculations predict the correct selectivity as determined experimentally and give a reversal of selectivity as compared to the cluster models mentioned above. A previous QM/MM study on TauD on a considerably smaller enzyme model reported a
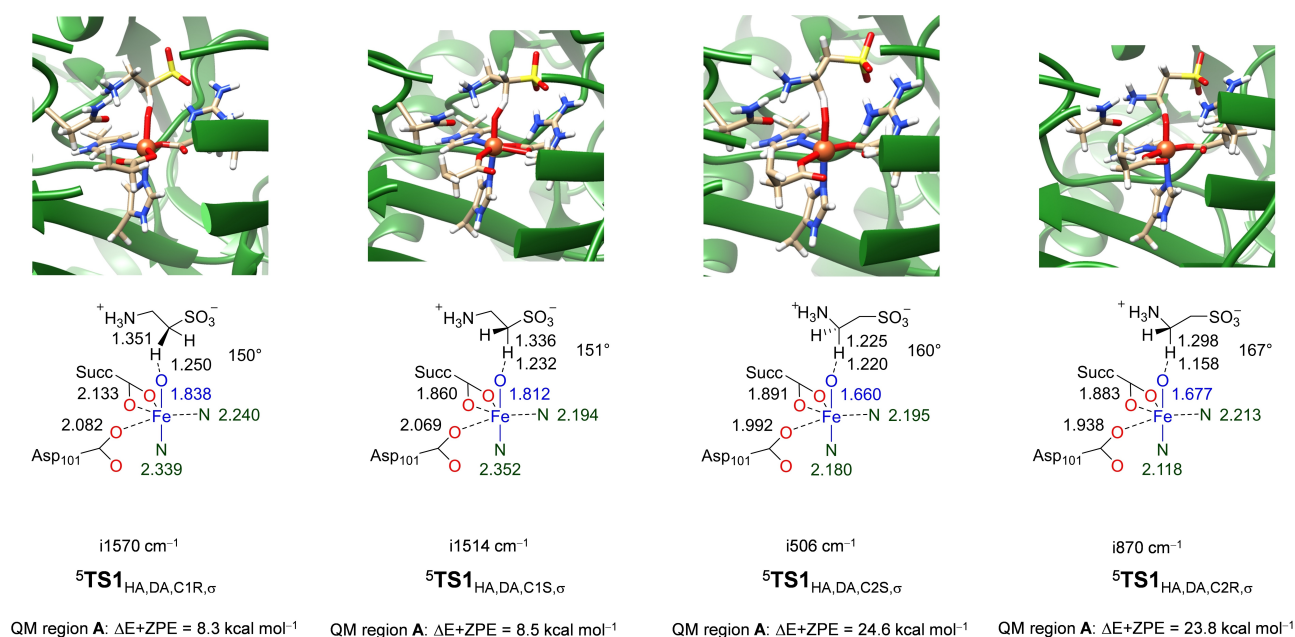
Figure 7. QM/MM optimized transition state geometries using QM region **A** calculated at ONIOM(UB3LYP-D3/BS1:Amber) level of theory. Relative energies include ZPE and were obtained with basis set BS2 relative to ^5Re_D. Bond lengths are in angstroms and the C–H–O angles in degrees.

Chem. Eur. J. **2022**, 28, e202104167 (7 of 15)

© 2021 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202104167

Chemistry Europe
European Chemical
Societies Publishing

hydrogen atom abstraction barrier of $\Delta E + ZPE = 6.7$ kcal mol^{-1} for the pro-R-C^1–H abstraction reaction based on constraint geometry scans, which is a result that is in good agreement with what we find here.[11e] To find out if the choice of the QM region affects the energetics of the calculation, we expanded the QM region to the same atoms as described in cluster model **B** and did a single point QM/MM calculation at ONIOM(UB3LYP-D3/BS1:Amber) level of theory. However, the choice of the QM region has a dramatic effect on the energetics and reverses the ordering of the transition states and predicts selective C^2-hydroxylation rather than C^1-hydroxylation. This is opposite of experimental observation and therefore the choice of the QM region and the protonation state of the second-coordination sphere residues is essential for the correct description of the transition state energies. Moreover, the same selectivity is obtained with QM/MM with QM region **DA** and the cluster model **A**, which both treat the same atoms with a QM method.

The ^{5}**TS1**$_{HA,DA,C2S,\sigma}$ and ^{5}**TS1**$_{HA,DA,C2R,\sigma}$ transition states have short Fe–O distances of 1.660 and 1.677 Å, respectively, and also have a relatively small imaginary frequency of i506 and i870 cm^{-1}. These unusual distances and frequency values point to large constraints on the chemical structure of the transition states. Indeed the O–H–C^2 angle is relatively large, namely 160 and 167°. By contrast, the ^{5}**TS1**$_{HA,DA,C1S,\sigma}$ and ^{5}**TS1**$_{HA,DA,C1R,\sigma}$ structures have a significantly longer Fe–O bond of 1.812 and 1.838 Å due to electron transfer from the substrate into the antibonding σ^*_{z2} orbital, while the electron transfer is later in the C^2 transition states. Thus, the ^{5}**TS1**$_{HA,DA,C1R,\sigma}$ and ^{5}**TS1**$_{HA,DA,C1S,\sigma}$ have a spin density of −0.60 and −0.58 on the substrate, which is typical for $^5\sigma$-type transition states and the transfer of an up-spin electron from substrate into the σ^*_{z2} orbital. The ^{5}**TS1**$_{HA,DA,C1R,\sigma}$ and ^{5}**TS1**$_{HA,DA,C1S,\sigma}$ structures are product-like with a shorter O–H distances of 1.250/1.232 Å than the C^1–H distance of 1.351/1.336 Å. Interestingly, both structures have a bent O–H–C^1 orientation of 150–151°, while typically values around 180° are expected for this pathway.

Comparison of DFT and QM/MM results and link to experiment

Experimentally a rate constant $k_{exp} = 13$ s^{-1} at 5 °C was determined for hydrogen atom abstraction by an iron(IV)-oxo species leading to predominant R-1-hydroxytaurine products.[10a] Using transition state theory, this rate constant corresponds to a free energy of activation at 5 °C of $\Delta G^+_{exp,5°C} = 14.83$ kcal mol^{-1}. To find out how our various models compare to experiment, we summarize the various transition state energies obtained with QM cluster and QM/MM methods described in this work in Table 1. Using cluster model **B** a $\Delta G^+_{HA,B,C1R,\sigma} = 15.4$ kcal mol^{-1} is obtained (Figure 5, Table 1), which is in excellent agreement with the experimentally obtained result. Unfortunately, we located lower energy barriers for hydrogen atom abstraction from the C^2 and pro-S C^1 positions of taurine. Therefore, cluster model **B** predicts the correct kinetics for the pro-R C^1–H abstraction pathway but for this model the wrong kinetics for

Table 1. DFT calculated free energies of activation (ΔG^+) for the pro-R C^1–H, pro-S C^1–H, pro-S C^2–H and pro-R C^2–H hydrogen atom abstraction pathways from taurine by TauD models.[a,b]

Method/Model	TS1$_{C1R,\sigma}$	TS1$_{C1S,\sigma}$	TS1$_{C2S,\sigma}$	TS1$_{C2R,\sigma}$
DFT model **A**	4.2	18.1	13.8	15.0
DFT model **B**	15.4	14.2	13.1	14.5
DFT model **B2**	27.8	26.8	24.4	30.5
DFT model **C**	17.1	16.3	26.6	28.6
DFT model **C2**	30.9	25.7	24.9	28.6
QM/MM model **DA**	6.7	8.3	26.4	24.9
QM/MM model **DB**	22.6	21.8	13.3	9.4
Experiment[c]	14.83			

[a] ΔG_{BS2} values contain energies calculated at UB3LYP/BS2 with solvent model included, while geometries are optimized at UB3LYP/BS1. [b] Values in kcal mol^{-1}. [c] Rate constant from Ref. [10a] converted into a free energy using transition state theory.

the alternative hydrogen atom abstraction pathways is obtained and hence it gives an overall wrong selectivity.

We then expanded the cluster model to model **C** and obtained a $\Delta G^+_{HA,C,C1R,\sigma} = 17.1$ kcal mol^{-1}, while the corresponding C^2–H hydrogen atom abstraction barriers are well higher in energy. Moreover, both ^{5}**TS1**$_{HA,C,C1R,\sigma}$ and ^{5}**TS1**$_{HA,C,C1S,\sigma}$ transition states lead to the same radical intermediate that preferentially reacts to form R-1-hydroxytaurine. Consequently, the slightly enlarged model **C** predicts the correct kinetics and regio- and enantioselectivity of the TauD enzyme well. Finally, with QM/MM a free energy of activation of $\Delta G^+_{HA,DA,C1R,\sigma} = 6.7$ kcal mol^{-1} is found that is well lower in energy than the other C–H abstraction barriers. Therefore, QM/MM with QM region **A** gives the correct selectivity but its barriers are well underestimated with respect to experiment. Most probably this is the result of the choice of a small QM region for the QM/MM calculations as it matches the result of the minimal cluster model **A** the best. Finally, to establish the effect of the His$_{70}$ protonation state on the reactivity and selectivity of TauD enzymes, we took model **C** and removed one of the His$_{70}$ protons to obtain model **C2**. We then reoptimized all four hydrogen atom abstraction transition states and find the C1S and C2S barriers to be close in free energy, namely $\Delta G = 25.7$ kcal mol^{-1} for ^{5}**TS1**$_{HA,C2,C1S,\sigma}$ and $\Delta G = 24.9$ kcal mol^{-1} for ^{5}**TS1**$_{HA,C2,C2S,\sigma}$. Therefore, the His$_{70}$ protonation state is essential for TauD to obtain the correct selectivity and guide the reaction to C^1-hydroxylation. The absolute barriers of the model **C2** transition states; however, are well higher than those for model **B** by more than $\Delta G = 10$ kcal mol^{-1} (see Table 1). Therefore, we analyzed the structural differences between model **B** and **C** further to find the origin of the lowering of the transition state barriers.

To understand the second-coordination sphere better, we created a geometric overlay of the ^{5}**Re**$_B$ and ^{5}**Re**$_D$ structures. As discussed above in Figure 2 an overlay of the first-coordination sphere of the iron center gives negligible differences between the two structures that would not explain reactivity differences. Therefore, long-range perturbations of the second-coordination sphere must affect substrate binding and positioning and its thermochemical properties. We then focused on the second-coordination sphere and particularly the taurine binding

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202104167

Chemistry Europe
European Chemical
Societies Publishing

position, see Figure 8. Although the C^1H_2–SO_3^- group of taurine is virtually in the same position in ^5Re_B, ^5Re_C and ^5Re_D, actually the amine group of the substrate has shifted slightly and in $^5Re_C/^5Re_D$ it forms a hydrogen bond with the oxo group. As a consequence of this new hydrogen bond there is a shift in the hydrogen bonding network around the taurine amine group and also changes in the positions of the Tyr_{73}, Asp_{94}, and Asn_{95} side chains in the substrate binding pocket are observed. Most probably this hydrogen bonding interaction from the taurine amine group to the oxo triggers a hydrogen atom from the C1S position and makes hydrogen atom abstraction from the other three positions more difficult. Thus, a hydrogen bond of 1.961 Å between the taurine amine group and oxo is formed, while it also interacts with the oxygen atom of Asn_{95} at a distance of 2.009 Å.

The amine group of taurine interacts further with the carboxylate of Asp_{94} and through a crystal water molecule with the alcohol group of Ser_{158} in a tight hydrogen bonding network. A similar hydrogen bonding network is seen in cluster model ^5Re_C. Our calculations, therefore, show the importance of the hydrogen bonding network around taurine involving the side-chains of Asp_{94}, Asn_{95} and Ser_{158}. Indeed, active site mutations where these residues were replaced by alternative (smaller) side-chains led to reduced activity of TauD and much smaller rate constants.[22]

As this hydrogen bonding network has shifted in the DFT cluster model ^5Re_B with respect to that in $^5Re_C/^5Re_D$, we decided to explore the effect of this hydrogen bonding network further and considered an alternative cluster model $^5Re_{B2}$ that is based on ^5Re_B and the corresponding hydrogen atom transition states but has the groups Asp_{94}, Asn_{95}, Ser_{158} and Phe_{159} removed. The calculations on Model **B2** raise the barrier of $^5TS1_{HA,B2,C1S,\sigma}$ to

$\Delta G^\ddagger = 26.8$ kcal mol^{-1} as compared to the model **B** transition state. By contrast, the $^5TS1_{HA,B2,C1R,\sigma}$, $^5TS1_{HA,B2,C2S,\sigma}$ and $^5TS1_{HA,B2,C2R,\sigma}$ barriers are raised even higher to $\Delta G^\ddagger = 27.8$, 24.4 and 30.5 kcal mol^{-1}, respectively. As such, the hydrogen bonding network lowers the C^1–H and C^2–H abstraction barriers dramatically and affects the rate of the reaction by positioning the substrate in a better orientation. Despite the fact that the hydrogen bonding network of Asp_{94}, Asn_{95} and Ser_{158} influences the kinetics of the reaction dramatically, actually it does not influence the regioselectivity dramatically and the same ordering of the transition states is obtained with the lowest barrier for $^5TS1_{HA,B2,C2S,\sigma}$ for models **B** and **B2**. Our calculations, therefore, match experimental data on TauD variants with Asp94Ala, Asn95Ala, Asn95Asp, Ser158Ala, Phe159Leu and Phe159Val mutations that showed a considerable reduction of the rate constant of taurine hydroxylation as compared to the wild-type enzyme.[22,23] It was suggested that these mutants will bring additional crystal water molecules into the substrate binding pocket that disrupt the hydrogen bonding network and make the substrate more mobile.

To understand how the hydrogen bonding network and particularly the close approach of the protonated amine group of substrate to the oxidant affects the charge and spin distributions we analyzed the spin densities in detail. Figure 9 shows the optimized hydrogen atom abstraction transition states and the spin density obtained on the FeO and taurine groups. In the QM/MM optimized transition states for C^1–H abstraction, a large amount of spin is found on the substrate moiety (-0.60 in $^5TS1_{HA,DA,C1R,\sigma}$ and -0.58 in $^5TS1_{HA,DA,C1S,\sigma}$). By contrast, in the corresponding model **B** transition states only -0.44 spin is found on the substrate. Therefore, the hydrogen bonding interaction between the amine group of substrate and the iron(IV)-oxo species polarizes the C^1–H bond stronger and causes an earlier electron transfer from substrate to oxidant. In the QM/MM structure, the C^2–H transition states are late with little electron transfer in the transition states: $^5TS1_{HA,D,C2S,\sigma}$ has $\rho_{taurine} = 0.16$ and $^5TS1_{HA,D,C2R,\sigma}$ has $\rho_{taurine} = -0.10$. By contrast, these transition states have significant radical character in the cluster model **B** optimized structures with values of -0.43 and -0.50, respectively. As a consequence, a local electric field and dipole moment perturbs the electronic configuration of the C^2–H hydrogen atom abstraction transition states dramatically and results in a change of ordering from the large cluster model to the QM/MM system. A comparison of the spin densities for $^5TS1_{HA,B,C1R,\sigma}$ and $^5TS1_{HA,C,C1R,\sigma}$ gives very similar values, which is as expected as their free energy of activation was very much alike. By contrast, the spin density values for the two pro-S and pro-R C^2–H hydrogen atom abstraction transition states for model **B** and **C** are very different as expected from the dramatic differences in barrier height.

Environmental effects on C–H bond strengths

Previously, we showed that external perturbations such as hydrogen bonding interactions, electric dipole moments and electric field effects can influence C–H bond strengths and
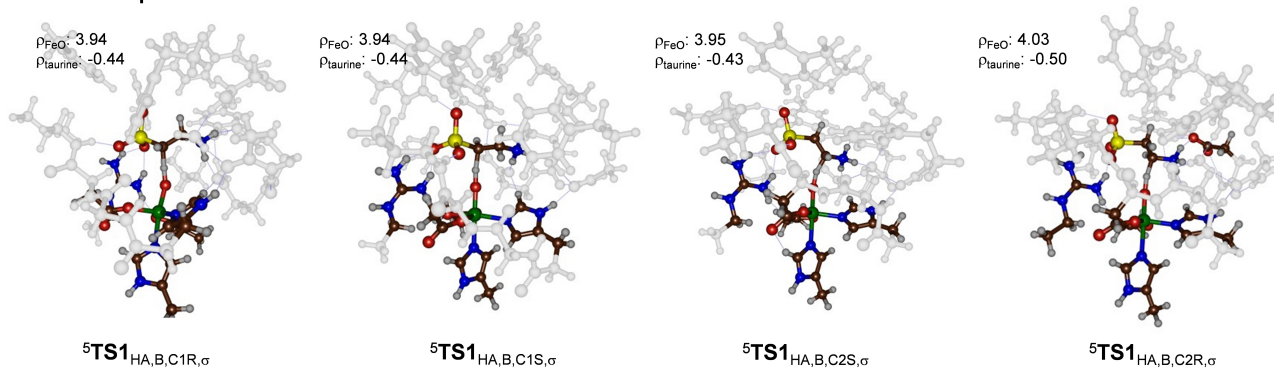

Figure 8. Hydrogen bonding network surrounding taurine substrate in ^5Re_C and ^5Re_D and an overlay of ^5Re_B (in purple) and ^5Re_D. Bond lengths are given in angstroms. The changes in position of protein residues between model **B** and **D** are identified with orange arrows.

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202104167

Chemistry Europe
European Chemical
Societies Publishing

Model B spin densities:

ρ_{FeO}: 3.94
$\rho_{taurine}$: -0.44

ρ_{FeO}: 3.94
$\rho_{taurine}$: -0.44

ρ_{FeO}: 3.95
$\rho_{taurine}$: -0.43

ρ_{FeO}: 4.03
$\rho_{taurine}$: -0.50

$^5\mathbf{TS1}_{HA,B,C1R,\sigma}$ $\quad$ $^5\mathbf{TS1}_{HA,B,C1S,\sigma}$ $\quad$ $^5\mathbf{TS1}_{HA,B,C2S,\sigma}$ $\quad$ $^5\mathbf{TS1}_{HA,B,C2R,\sigma}$

Model C spin densities:

ρ_{FeO}: 3.87
$\rho_{taurine}$: -0.42

ρ_{FeO}: 3.87
$\rho_{taurine}$: -0.32

ρ_{FeO}: 3.50
$\rho_{taurine}$: 0.29

ρ_{FeO}: 3.92
$\rho_{taurine}$: -0.36

Model D spin densities:

ρ_{FeO}: 4.13
$\rho_{taurine}$: -0.60

ρ_{FeO}: 4.09
$\rho_{taurine}$: -0.58

ρ_{FeO}: 3.68
$\rho_{taurine}$: 0.16

ρ_{FeO}: 3.83
$\rho_{taurine}$: -0.10

Figure 9. Group spin densities of the hydrogen atom abstraction transition states for the DFT model **B** and **C** and the QM/MM model **D** with QM region **A**.

selectivities.[24] In particular, electric field effect calculations by various groups identified changes in reaction rates, as well as reaction mechanism with external electric field perturbations included.[25] Therefore, we initially calculated the substrate C–H bond dissociation free energy (BDFE) values of taurine for all four C–H bonds from the homolytic cleavage of the C–H bond by calculating the substrate, a hydrogen atom and the substrate with a hydrogen atom removed. The gas-phase calculated BDFE values have been taken from Ref. [12]. The data obtained at the UB3LYP/BS2 level of theory are shown in Figure 10. As can be seen from Figure 10, with a solvent model included in the calculations the four C–H bond strengths are in a narrow window of 2.0 kcal mol⁻¹ with the C^1–H bond the weakest of those bonds. Nevertheless, the small differences in substrate C–H bond strength would imply products originating from all four possible hydrogen atom abstraction channels.

We also calculated the BDFE values in the gas-phase (without solvent model) and found $BDFE_{C1,no\ solvent} =$
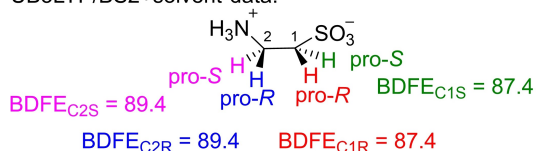
$85.8\ \text{kcal mol}^{-1}$ and $BDFE_{C2,no\ solvent} = 74.7\ \text{kcal mol}^{-1}$.[12] Consequently, the addition of a solvent model changes the ordering of the BDFE values for the C^1–H and C^2–H bond strengths. As such the C–H bond strengths will be sensitive to external perturbations and polar groups. As the BDFE values of taurine appear to be sensitive to external polarization effects, we decided to calculate them using electric field effect perturbations included.

Experimentally, only *pro-R* C^1–H hydrogen atom abstraction takes place, which implies that the protein environment through external electrostatic perturbations must affect the regioselectivity of the reaction and maybe even the C–H bond strengths. In particular, in recent work of our group we showed that external electric fields along a specific direction can change C–H bond strengths and in enzymes guide a regioselectivity pathway.[24a,c] Therefore, we performed electric field based calculations on an isolated taurine substrate using fields along the positive and negative x-, y- and z-axis of the molecules. Figure 11 shows the evolvement of the BDFE values for C1R and C2R under electric field perturbations along the molecular x-, y- and z-axis of taurine. As can be seen a field along the molecular skeleton of taurine (z-axis) has little effect on the relative BDFE values for *pro-R* C^1–H and C^2–H bond strengths and along the whole series the C^1–H is the weakest of the two. By contrast, a field along the molecular y-axis, i.e. along the *pro-R* C^1–H bond, either pushes electron density towards the carbon or pulls it away. As a consequence with a positive field along the y-axis the C^1–H bond is weakest, while with a field in the opposite direction the C^2–H bond becomes the weakest of the two bonds. Finally, with a field along the x-direction the two BDFE values become almost equal, which should result in a mixture of products originating from C^1–H and C^2–H hydrogen atom

UB3LYP/BS2+solvent data:

H_3N^+ $\quad$ SO_3^-

pro-*S* H $\quad$ H pro-*S*
pro-*R* $\quad$ pro-*R*

$BDFE_{C2S} = 89.4$ $\qquad$ $BDFE_{C1S} = 87.4$

$BDFE_{C2R} = 89.4$ $\qquad$ $BDFE_{C1R} = 87.4$
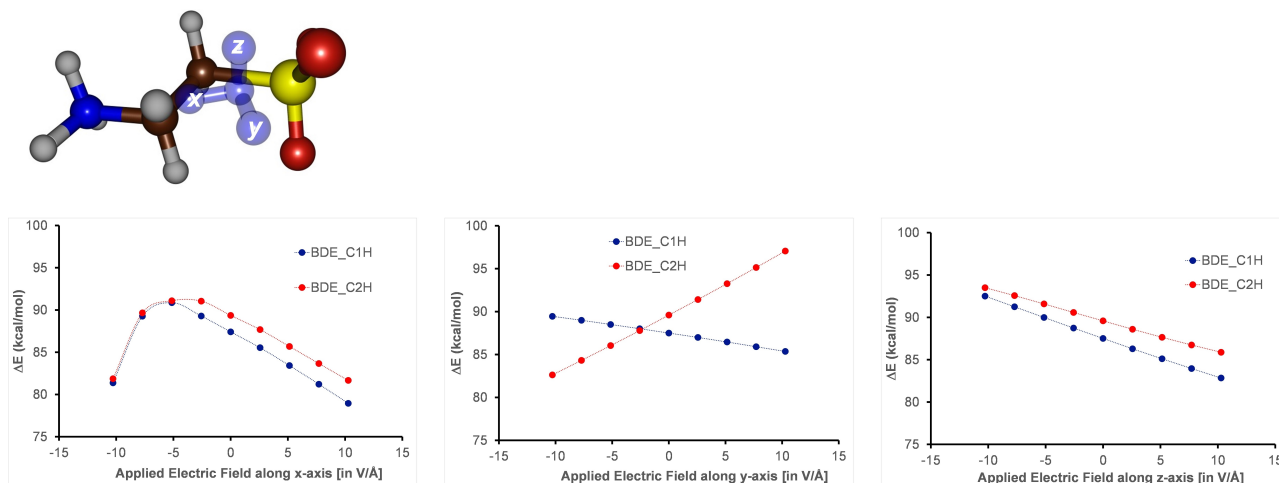
UB3LYP/BS2, gas-phase:

$BDFE_{C2} = 74.7$ $\qquad$ $BDFE_{C1} = 85.8$

Figure 10. Calculated BDFE values for the *pro-R* and *pro-S* C–H bonds of an isolated taurine molecule as calculated at UB3LYP/BS2 level of theory: Data obtained at ΔG_{solv} (top) or ΔG_{gas} level of theory (in kcal mol⁻¹).

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202104167

**Chemistry
Europe**
European Chemical
Societies Publishing

Figure 11. Electric field effect perturbations on the C^1–H (in blue) and C^2–H (in red) bond strengths (BDFE) of taurine as calculated at UB3LYP/6-311 + G* in the gas-phase. The x-, y-, and z-axis are defined through the centre of mass of taurine with the positive direction as defined in Gaussian identified with the arrows in blue. The field values are reported in V/Å. The electric field perturbation was added to the solvent-corrected BDFE values for *pro-R* C^1–H and C^2–H from Figure 10.

abstraction processes. These electric field results show that the electrostatic perturbations of the protein can have a major effect on the regioselectivity of the reaction and push the mechanism into a specific direction.

Figure 12 shows the electric dipole moment vector in complexes ^5Re_B and ^5Re_C. Thus, in ^5Re_C the vector points along the C^2–H bond of the substrate and according to the electric field effect perturbations from Figure 11 that perturbs along the molecular y-axis of the substrate and strengthens these C^2–H bonds. Indeed, model **C** shows high energy barriers for hydrogen atom abstraction from the C^2–H bonds of taurine. By contrast to the electric dipole vector in model ^5Re_C, the vector in ^5Re_B points in a very different direction that corresponds to the x- and z-direction electric fields of Figure 11. As shown in Figure 11, fields along the molecular x- and z-axis of the substrate have little effect on the relative bond strengths of the C^1–H and C^2–H bonds. As a consequence, model **B** has an

electric dipole vector that little affects the C–H bonds strengths at the C^1–H and C^2–H bonds of taurine. Indeed, using model **B** we find all hydrogen atom abstraction barriers within a small window of about 3 kcal mol^{-1} that would predict an unselective reaction with multiple products. A further analysis of the 1OS7 pdb file (right-hand-side of Figure 12) focuses on the charged residues in the protein structure. The TauD structure contains 19 glutamate and 14 aspartate amino acids with a negatively charged side-chain, while there are 19 arginine and 9 lysine amino acids with a positively charged side-chain. Therefore, the protein structure is overall negatively charged and these polar amino acids induce an electric dipole moment and a local electric field on the reaction center. The dipole moment shown for large cluster model **C** indeed points in the same direction as the favorable electric field direction for selective C^1-hydroxylation. By contrast, the dipole moment for model **B** points in the wrong direction as compared to the favorable electric field
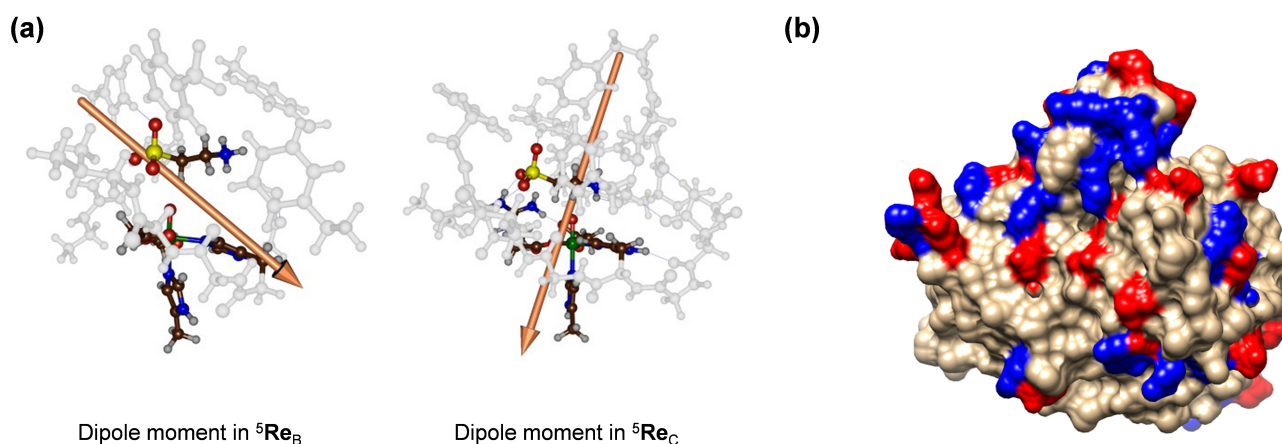
Figure 12. (a) Electric dipole moment vector of the reactant complexes ^5Re_B and ^5Re_C. (b) Charge distribution in TauD as analysed from the 1OS7 pdb file: positively charged amino acids in blue and negatively charged ones in red.

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202104167

Chemistry Europe
European Chemical
Societies Publishing

direction for selective C^1-hydroxylation and indeed predicts favorable C^2-hydroxylation instead. Consequently, large cluster models like model **C** reproduce the structure and reactivity of TauD well thanks to the long-large interactions and charge and dipole moments of the second-coordination sphere that influence substrate positioning and its reactivity.

Conclusion

A combination of density functional theory cluster calculations and quantum mechanics/molecular mechanics studies are presented on taurine/α-ketoglutarate-dependent dioxygenase. In particular, pathways leading to *R*- and *S*-1-hydroxytaurine and *R*- and *S*-2-hydroxytaurine have been performed. The reactions are stepwise via a radical intermediate and create an intermediate of σ-type with a high-spin iron(III) antiferromagnetically coupled to a substrate radical. A number of large cluster models as well as QM/MM with various size QM region have been studied. These cluster models range from a minimal model that contains the first-coordination sphere and substrate only (model **A**) to models with expanded second-coordination sphere environments surrounding the oxidant and substrate. We show that the choice of the model is crucial for the calculations and particularly there is an important dipole moment in the active site structure that guides the reaction to the correct selectivity. This dipole moment is accomplished through a protonated His residue in the vicinity of the substrate (His_{70}), which needs to be in its protonated form in the catalytic cycle. Models without a proton on this His_{70} residue predict the wrong selectivity and/or kinetics. In addition, the calculations highlight the hydrogen bonding network of Asp_{94}, Asn_{95} and Ser_{158} that position the amine group of taurine substrate and position it in an ideal orientation for optimal kinetics. Our calculations using large cluster models that include key hydrogen bonding interactions in the substrate binding pocket and His_{70} protonated reproduce experimental rates and selectivity excellently and give a dominant C^1-hydroxylation channel leading to *R*-1-hydroxytaurine products. This is triggered by charged active site residues including a protonated His_{70} group.

Experimental Section

Set up of the system: Our model set-up starts from the crystal structure coordinates of substrate- and α-KG-bound TauD as deposited by the 1OS7 pdb file, which is an enzymatic monomer.[9] Hydrogen atoms were added to the structure based on pH 7 conditions with the H++ webserver.[26] The MCPB.py tool as implemented in the AMBER-18 software package was used to generate the force field parameters for a penta-coordinated iron(II) centre, while for the description of the protein atoms the ff14SB4 force field parameters were used.[27,28] Substrate taurine and α-KG were described with the second generation General Amber Force Field (GAFF2) using AM1-BCC charges.[29] These GAFF2 parameters were generated with the antechamber and parmchk2 modules of the AMBER-18 software package. Finally, the substrate-bound enzyme structure was solvated in a cubic box of TIP3P water molecules and subsequently neutralized by adding Na^+ ions.

Thereafter, a molecular dynamics (MD) simulation was carried out using the Particle Mesh Ewald Molecular Dynamics (PMEMD) module as implemented in the AMBER-18 software package.[30] The system was minimized using 2000 steps of steepest decent minimization with a starting restraint potential of 2.0 kcal mol^{-1} on all heavy atoms of the protein, which was gradually released. After that the system was heated for 100 ps from a temperature of 0 K to 298.15 K under NVT ensemble conditions with the Langevin thermostat and subsequently equilibrated for 1 ns under NPT conditions using the Berendsen barostate without constraints to the energy and the structure.[31] Thereafter, a 20 ns MD simulation under NPT conditions was carried out in 2 fs time-steps, using the SHAKE protocol on hydrogen atoms and a 10 Å non-bonded cut-off with periodic boundary conditions. The MD simulation (Supporting Information, Figures S1 and S2) shows that the system is highly rigid and little changes with respect to the crystal structure coordinates are seen.

QM/MM approaches: QM/MM calculations follow procedures as reported previously.[32] In general, the QM/MM calculations were carried out using ONIOM scheme as implemented in the Gaussian software package.[33,34] The QM/MM calculations were setup using molUP, a VMD plugin, from the equilibrated final structure of the MD simulation.[35] The QM/MM model **D** consist of the complete protein, substrate and water molecules within 15 Å of the protein and has a total of 9,551 atoms and has His_{70} in its singly protonated form. The QM region consist of the first-coordination sphere around the iron(IV)-oxo group and substrate and the link-atom procedure was used to bridge the QM and MM regions.[36] QM region **A** has 53 atoms, while QM region **B** contains the atoms as shown in Figure 2a. The QM region is treated using DFT with the unrestricted B3LYP hybrid functional and a basis set with LANL2DZ (with electron core potential) on iron and 6-31G* on the rest of the atoms: basis set BS1,[37,38] while the MM region was treated with the Amber force field.[28,30] We initially ran QM/MM geometry optimizations using QM region **A** and followed these up with a set of single point calculations whereby the QM region was expanded to the atoms shown in Figure 2 under model **B**. The electrostatic interactions between the QM and MM regions were described with the electronic embedding scheme. The geometry optimizations were performed with UB3LYP/BS1 level of DFT and followed by an analytical frequency calculation at 298.15 K and 1 atm. The transition states were located using potential energy scan calculations followed by full geometry optimization with the Berny algorithm. The energy values were corrected through single point calculations at the ONIOM(UB3LYP/BS2:Amber) level of theory with BS2 a basis set representing LACV3P+ (with electron core potential) on iron and 6-311+G* on the rest of the atoms.

The M06 density functional method was also tested for the QM region in the QM/MM calculations,[39] however this gave essentially similar energetics than those found with B3LYP, see Supporting Information.

DFT cluster models: Cluster models focus on the substrate and oxidant binding environment and the second-coordination sphere and consider all atoms in the model with a quantum chemical approach.[40] As the MD simulation gives little variety between the active site orientation of the various snapshots, we created five cluster models (**A**, **B**, **B2**, **C** and **C2**) based on the last step of the MD simulation, see Figure 2. We truncated the active site model by selecting the residues of amino acids and second coordination sphere groups, which determine substrate and oxidant binding and positioning and particularly include all polar (hydrogen bonded and π-stacking interactions) close to the metal and substrate. In the active site model we replaced the iron(II) ion by iron(IV)-oxo species, while α-KG was replaced by succinate. A truncated model (minimal) cluster model was considered, namely a minimal cluster

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202104167

**Chemistry
Europe**
European Chemical
Societies Publishing

model **A** that contains only the first coordination sphere of residues to iron and the substrate and had 72 atoms. Thereafter, larger cluster models **B** and **C** were created that incorporate the environments around substrate and oxidant and their hydrogen bonding interactions. The active site cluster model **B** incorporates the metal and its first-coordination sphere as well as the substrate and a number of second-coordination sphere residues that determine the size and shape of the substrate binding pocket and incur hydrogen bonding interactions. Thus, the model includes a short protein chain of amino acids that links to the equatorial ligands of the iron, namely His_{99}–Thr_{100}–Asp_{101}–Val_{102}. In addition, two short chains for Ser_{158}–Phe_{159} and Asp_{94}–Asn_{95} were included. The axial histidine ligand (His_{255}) of iron was abbreviated to methylimidazole and α-KG replaced by acetate. Finally, the side chains of the residues His_{70}, Tyr_{73}, Ile_{83}, Asn_{97}, Phe_{104}, Phe_{206} and Arg_{270} as well as substrate taurine were included. Overall our cluster model contains 244 atoms and has a neutral charge. To prevent the model from changing dramatically in structure during the geometry optimizations we placed some constraints on some of the C^{α}-protein atoms as identified with a star in Figure 2. Cluster Model **C** contains 279 atoms and is based on the QM/MM optimized geometry and had the atoms of cluster model **B** expanded with the full peptide chain Asp_{94}–Asn_{95}–Asp_{96}–Asn_{97}–Trp_{98}–His_{99}–Thr_{100}–Asp_{101}–Val_{102}–Thr_{103}–Phe_{104} with Asp_{96}, Trp_{98} and Thr_{100} truncated to a Gly residue. In addition, model **C** has the His_{70} residue doubly protonated. To test the protein environment on the kinetics and selectivity of the reactions, we created two further models based on models **B** and **C**, namely **B2** and **C2**. Model **C2** is similar to model **C** but has His_{70} singly protonated, while model **B2** is model **B** with the chains of Asp_{94}, Asn_{95}, Ser_{158} and Phe_{159} removed.

DFT procedures: The Gaussian-09 software package was used for all quantum chemical calculations discussed here.[33] Following previous experience with cluster models of nonheme iron complexes,[41] we utilized the unrestricted B3LYP density functional method in combination with a LANL2DZ (with electron core potential) on iron and 6-31G* on the rest of the atoms: basis set BS1.[37,38] This method was shown to reproduce experimental spin-state assignments and rate-constants well.[42] All geometry optimizations were performed with these basis sets in which all local minima were verified by the absence of negative eigenvalues in the vibrational frequency analysis while all the transition state structures were found using the Berny algorithm, and verified by vibrational analysis and visualized by animating the imaginary frequency. For key transition states also intrinsic reaction coordinate calculations were done, which confirmed the transition states to connect with the local minima as identified. In order to improve the energetics, single point energies on the optimized geometries were calculated at UB3LYP/BS2, whereby BS2 represents the LACV3P+ (with electron core potential) on iron and 6-311+G* on the rest of the atoms. The latter set of calculations included a continuum polarized conductor model (CPCM) with a dielectric constant mimicking chlorobenzene and the dispersion correction (D3) developed by Grimme.[43] For a selection of structures, we ran full geometry optimizations at UB3LYP-D3/BS1 level of theory, but these studies gave the same trends as those obtained at UB3LYP/BS1, see Supporting Information for details. To obtain the free energies at 298.15 K and 1 atm, the zero-point energy (ZPE), thermal corrections and entropy contribution evaluated from the unscaled vibrational frequencies at the UB3LYP/BS1 level of theory were then added to the electronic energies calculated from the same level of DFT.

The primary kinetic isotopic effect (KIE's) were calculated using the classical equations due to Eyring (Eq. (1)) and with tunnelling corrections included due to Wigner (Eq. (2) and 3).[44] In the Eyring KIE, the activation free energy ($\Delta G^{\ddagger}$) of hydrogen and deuterium

substituted reaction was considered in the gas phase at room temperature (T = 298.15 K) with R being the gas constant. By contrast, the Wigner tunnelling factor incorporates a factor that considers the change in the magnitude of the imaginary vibrational frequency in the transition state for the hydrogen and deuterium substituted systems. In Equation (3), h is Planck's constant and k_B is the Boltzmann constant, The KIE values for the various hydrogen atom abstraction pathways were calculated by replacing the both *pro-R* and *pro-S* hydrogen atoms at the C^1 and C^2 positions of taurine substrate.

$$KIE_{Eyring} = \exp[(\Delta G^{\ddagger}_D - \Delta G^{\ddagger}_H)/RT] \tag{1}$$

$$KIE_{Wigner} = KIE_{Eyring} \times Q_{t,H}/Q_{t,D} \tag{2}$$

$$Q_t = 1 + (h\nu/k_BT))^2/24 \tag{3}$$

Finally, we performed the electric field effect (EFE) calculations as implemented in Gaussian software package using the "field" keyword.[33] The EFE is calculated by performing a single point energy calculation on various optimized geometries with an electric field perturbation along the molecular x, y or z-axis with the positive direction as defined in Gaussian. These electric field perturbations were done with various field magnitudes in either positive or negative field directions along each principal axis.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

[1] a) E. I. Solomon, T. C. Brunold, M. I. Davis, J. N. Kemsley, S. K. Lee, N. Lehnert, F. Neese, A. J. Skulan, Y. S. Yang, J. Zhou, *Chem. Rev.* **2000**, *100*, 235–349; b) T. D. H. Bugg, *Curr. Opin. Chem. Biol.* **2001**, *5*, 550–555; c) M. J. Ryle, R. P. Hausinger, *Curr. Opin. Chem. Biol.* **2002**, *6*, 193–201; d) M. Costas, M. P. Mehn, M. P. Jensen, L. Que Jr, *Chem. Rev.* **2004**, *104*, 939–986; e) M. M. Abu-Omar, A. Loaiza, N. Hontzeas, *Chem. Rev.* **2005**, *105*, 2227–2252; f) C. Krebs, D. G. Fujimori, C. T. Walsh, J. M. Bollinger Jr, *Acc. Chem. Res.* **2007**, *40*, 484–492; g) S. P. de Visser, D. Kumar (Eds.), *Iron-containing enzymes: Versatile catalysts of hydroxylation reactions in nature*, Royal Society of Chemistry Publishing, Cambridge (UK) **2011**; h) A. R. McDonald, L. Que Jr, *Coord. Chem. Rev.* **2013**, *257*, 414–428.
[2] a) A. G. Jarvis, *Curr. Opin. Chem. Biol.* **2020**, *58*, 63–71; b) A. Das, C. Hessin, Y. Ren, M. Desage-El Murr, *Chem. Soc. Rev.* **2020**, *49*, 8840–8867; c) A. C. Ghosh, C. Duboc, M. Gennari, *Coord. Chem. Rev.* **2021**, *428*, 213606.

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202104167

Chemistry Europe
European Chemical
Societies Publishing

[3] R. P. Hausinger, *2-Oxoglutarate-dependent oxygenases*, Royal Society of Chemistry, Cambridge **2015**, Chapter 1, pp. 1–58.

[4] a) P. C. A. Bruijnincx, G. van Koten, R. J. M. Klein Gebbink, *Chem. Soc. Rev.* **2008**, *37*, 2716–2744; b) P. Comba, M. Kerscher, *Coord. Chem. Rev.* **2009**, *253*, 564–574; c) W. Nam, Y.-M. Lee, S. Fukuzumi, *Acc. Chem. Res.* **2014**, *47*, 1146–1154; d) K. Ray, F. F. Pfaff, B. Wang, W. Nam, *J. Am. Chem. Soc.* **2014**, *136*, 13942–13958; e) G. Mukherjee, J. K. Satpathy, U. K. Bagha, M. Q. E. Mubarak, C. V. Sastri, S. P. de Visser, *ACS Catal.* **2021**, *11*, 9761–9797.

[5] a) M. D. White, E. Flashman, *Curr. Opin. Chem. Biol.* **2016**, *31*, 126–135; b) L. Vicens, G. Olivo, M. Costas, *ACS Catal.* **2020**, *10*, 8611–8631; c) J. M. Bollinger Jr, J. C. Price, L. M. Hoffart, E. W. Barr, C. Krebs, *Eur. J. Inorg. Chem.* **2005**, *21*, 4245–4254; d) E. G. Kovaleva, J. D. Lipscomb, *Nat. Chem. Biol.* **2008**, *4*, 186–193.

[6] a) M. A. McDonough, V. Li, E. Flashman, R. Chowdhury, C. Mohr, B. M. Lienard, J. Zondlo, N. J. Oldham, I. J. Clifton, J. Lewis, L. A. McNeill, R. J. Kurzeja, K. S. Hewitson, E. Yang, S. Jordan, R. S. Syed, C. J. Schofield, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9814–9819; b) K. Gorres, R. T. Raines, *Crit. Rev. Biochem. Mol. Biol.* **2010**, *45*, 106–124; c) J. D. Vasta, A. Choudhary, K. H. Jensen, N. A. McGrath, R. T. Raines, *Biochemistry* **2017**, *56*, 219–227.

[7] a) J. E. Baldwin, M. Bradley, *Chem. Rev.* **1990**, *90*, 1079–1088; b) M. G. Thomas, Y. A. Chan, S. G. Ozanick, *Agents Chemotherap.* **2003**, *47*, 2823–2830; c) X. Yin, T. M. Zabriskie, *ChemBioChem.* **2004**, *5*, 1274–1277; d) F. Kudo, A. Miyanaga, T. Eguchi, *Nat. Prod. Rep.* **2014**, *31*, 1056–1073.

[8] a) G. D. Straganz, B. Nidetzky, *ChemBioChem* **2006**, *7*, 1536–1548; b) C. A. Joseph, M. J. Maroney, *Chem. Commun.* **2007**, 3338–3349; c) S. Aluri, S. P. de Visser, *J. Am. Chem. Soc.* **2007**, *129*, 14846–14847; d) D. Buongiorno, G. D. Straganz, *Coord. Chem. Rev.* **2013**, *257*, 541–563; e) E. P. Tchesnokov, M. Fellner, E. Siakkou, T. Kleffmann, L. W. Martin, S. Aloi, I. L. Lamont, S. M. Wilbanks, G. N. L. Jameson, *J. Biol. Chem.* **2015**, *290*, 24424–24437; f) C.-C. G. Yeh, C. Pierides, G. N. L. Jameson, S. P. de Visser, *Chem. Eur. J.* **2021**, *27*, 13793–13806.

[9] a) J. R. O'Brien, D. J. Schuller, V. S. Yang, B. D. Dillard, W. N. Lanzilotta, *Biochemistry* **2003**, *42*, 5547–5554; b) H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235–242.

[10] a) J. C. Price, E. W. Barr, B. Tirupati, J. M. Bollinger Jr, C. Krebs, *Biochemistry* **2003**, *42*, 7497–7508; b) D. A. Proshlyakov, T. F. Henshaw, G. R. Monterosso, M. J. Ryle, R. P. Hausinger, *J. Am. Chem. Soc.* **2004**, *126*, 1022–1024; c) P. J. Riggs-Gelasco, J. C. Price, R. B. Guyer, J. H. Brehm, E. W. Barr, J. M. Bollinger Jr, C. Krebs, *J. Am. Chem. Soc.* **2004**, *126*, 8108–8109; d) M. L. Neidig, C. D. Brown, K. M. Light, D. Galonić Fujimori, E. M. Nolan, J. C. Price, E. W. Barr, J. M. Bollinger Jr, C. Krebs, C. T. Walsh, E. I. Solomon, *J. Am. Chem. Soc.* **2007**, *129*, 14224–14231.

[11] a) T. Borowski, A. Bassan, P. E. M. Siegbahn, *Chem. Eur. J.* **2004**, *10*, 1031–1041; b) S. P. de Visser, *J. Am. Chem. Soc.* **2006**, *128*, 9813–9824; c) A. V. Nemukhin, I. A. Topol, R. E. Cachau, S. K. Burt, *Theor. Chem. Acc.* **2006**, *115*, 348–353; d) S. Sinnecker, N. Svensen, E. W. Barr, S. Ye, J. M. Bollinger Jr, F. Neese, C. Krebs, *J. Am. Chem. Soc.* **2007**, *129*, 6168–6179; e) E. Godfrey, C. S. Porro, S. P. de Visser, *J. Phys. Chem. A* **2008**, *112*, 2464–2468; f) H. Chen, W. Lai, J. Yao, S. Shaik, *J. Chem. Theory Comput.* **2011**, *7*, 3049–3053; g) E. A. C. Bushnell, G. B. Fortowsky, J. W. Gauld, *Inorg. Chem.* **2012**, *51*, 13351–13356; h) H. J. Kulik, C. L. Drennan, *J. Biol. Chem.* **2013**, *288*, 11233–11241; i) A. Wójcik, M. Radoń, T. Borowski, *J. Phys. Chem. A* **2016**, *120*, 1261–1274; j) S. Álvarez-Barcia, J. Kästner, *J. Phys. Chem. B* **2017**, *121*, 5347–5354; k) R. N. Manna, T. Malakar, B. Jana, A. Paul, *ACS Catal.* **2018**, *8*, 10043–10050; l) J. E. M. N. Klein, G. Knizia, *Angew. Chem. Int. Ed.* **2018**, *57*, 11913–11917; *Angew. Chem.* **2018**, *130*, 12089–12093; m) S. S. Chaturvedi, R. Ramanan, N. Lehnert, C. J. Schofield, T. G. Karabencheva-Christova, C. Z. Christov, *ACS Catal.* **2020**, *10*, 1195–1209; n) M. Srnec, S. R. Iyer, L. M. K. Dassama, K. Park, S. D. Wong, K. D. Sutherlin, Y. Yoda, Y. Kobayashi, M. Kurokuzu, M. Saito, M. Seto, C. Krebs, J. M. Bollinger Jr, E. I. Solomon, *J. Am. Chem. Soc.* **2020**, *142*, 18886–18895; o) H. S. Ali, R. H. Henchman, J. Warwicker, S. P. de Visser, *J. Phys. Chem. A* **2021**, *125*, 1720–1737.

[12] S. P. de Visser, Y.-T. Lin, H. S. Ali, U. K. Bagha, G. Mukherjee, C. V. Sastri, *Coord. Chem. Rev.* **2021**, *439*, 213914.

[13] a) J. Rétey, *Angew. Chem. Int. Ed.* **1990**, *29*, 355–361; *Angew. Chem.* **1990**, *102*, 373–379; b) B. Vögeli, T. J. Erb, *Curr. Opin. Chem. Biol.* **2018**, *47*, 94–100.

[14] a) K. M. Light, J. A. Hangasky, M. J. Knapp, E. I. Solomon, *J. Am. Chem. Soc.* **2013**, *135*, 9665–9674; b) M. G. Quesne, R. Latifi, L. E. Gonzalez-Ovalle, D. Kumar, S. P. de Visser, *Chem. Eur. J.* **2014**, *20*, 435–446; c) A. Timmins, M. Saint-André, S. P. de Visser, *J. Am. Chem. Soc.* **2017**, *139*,

9855–9866; d) J. Xue, J. Lu, W. Lai, *Phys. Chem. Chem. Phys.* **2019**, *21*, 9957–9968.

[15] P. K. Grzyska, E. H. Appelman, R. P. Hausinger, D. A. Proshlyakov, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 3982–3987.

[16] A. P. Scott, L. Radom, *J. Phys. Chem.* **1996**, *100*, 16502–16513.

[17] a) H. Hirao, D. Kumar, L. Que Jr, S. Shaik, *J. Am. Chem. Soc.* **2006**, *128*, 8590–8606; b) I. Prat, J. S. Mathieson, M. Güell, X. Ribas, J. M. Luis, L. Cronin, M. Costas, *Nat. Chem.* **2011**, *3*, 788–793; c) R. Latifi, M. A. Sainna, E. V. Rybak-Akimova, S. P. de Visser, *Chem. Eur. J.* **2013**, *19*, 4058–4068; d) A. Ansari, A. Kaushik, G. Rajaraman, *J. Am. Chem. Soc.* **2013**, *135*, 4235–4249; e) T. Z. H. Gani, H. J. Kulik, *ACS Catal.* **2018**, *8*, 975–986; f) V. Dantignana, J. Serrano-Plana, A. Draksharapu, C. Magallón, S. Banerjee, R. Fan, I. Gamba, Y. Guo, L. Que Jr, M. Costas, A. Company, *J. Am. Chem. Soc.* **2019**, *141*, 15078–15091.

[18] a) S. P. de Visser, F. Ogliaro, P. K. Sharma, S. Shaik, *Angew. Chem. Int. Ed.* **2002**, *41*, 1947–1951; *Angew. Chem.* **2002**, *114*, 2027–2031; b) L. Bernasconi, E. J. Baerends, *J. Am. Chem. Soc.* **2013**, *135*, 8857–8867; c) M. Lundberg, T. Borowski, *Coord. Chem. Rev.* **2013**, *257*, 277–289; d) O. Cussó, I. Garcia-Bosch, X. Ribas, J. Lloret-Fillol, M. Costas, *J. Am. Chem. Soc.* **2013**, *135*, 14871–14878; e) D. Mandal, D. Mallick, S. Shaik, *Acc. Chem. Res.* **2018**, *51*, 107–117; f) D. Sahoo, M. G. Quesne, S. P. de Visser, S. P. Rath, *Angew. Chem. Int. Ed.* **2015**, *54*, 4796–4800; *Angew. Chem.* **2015**, *127*, 4878–4882.

[19] a) S. Shaik, H. Chen, D. Janardanan, *Nat. Chem.* **2011**, *3*, 19–27; b) C. T. Saouma, J. M. Mayer, *Chem. Sci.* **2014**, *5*, 21–31.

[20] a) S. P. de Visser, *J. Am. Chem. Soc.* **2010**, *132*, 1087–1097; b) H. Li, Y. Liu, *ACS Catal.* **2020**, *10*, 2942–2957; c) R. Kumar, B. Pandey, A. Sen, M. Ansari, S. Sharma, G. Rajaraman, *Coord. Chem. Rev.* **2020**, *419*, 213397.

[21] a) S. P. de Visser, *J. Am. Chem. Soc.* **2006**, *128*, 15809–15818; b) A. Decker, J.-U. Rohde, E. J. Klinker, S. D. Wong, L. Que Jr, E. I. Solomon, *J. Am. Chem. Soc.* **2007**, *129*, 15983–15996; c) S. Yeh, F. Neese, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1228–1233.

[22] P. K. Grzyska, M. J. Ryle, G. R. Monterosso, J. Liu, D. P. Ballou, R. P. Hausinger, *Biochemistry* **2005**, *44*, 3845–3855.

[23] K. P. McCusker, J. P. Klinman, *J. Am. Chem. Soc.* **2010**, *132*, 5114–5120.

[24] a) S. Ghafoor, A. Mansha, S. P. de Visser, *J. Am. Chem. Soc.* **2019**, *141*, 20278–20292; b) Y.-T. Lin, A. Stańczak, Y. Manchev, G. D. Straganz, S. P. de Visser, *Chem. Eur. J.* **2020**, *26*, 2233–2242; c) H. S. Ali, R. H. Henchman, S. P. de Visser, *Chem. Eur. J.* **2021**, *27*, 1795–1809; d) Y.-T. Lin, H. S. Ali, S. P. de Visser, *Chem. Eur. J.* **2021**, *27*, 8851–8864; e) H. S. Ali, R. H. Henchman, S. P. de Visser, *ChemCatChem* **2021**, *13*, 3054–3066.

[25] a) S. Shaik, S. P. de Visser, D. Kumar, *J. Am. Chem. Soc.* **2004**, *126*, 11746–11749; b) N. S. Hill, M. L. Coote, *J. Am. Chem. Soc.* **2018**, *140*, 17800–17804; c) S. Shaik, D. Danovich, J. Joy, Z. Wang, T. Stuyver, *J. Am. Chem. Soc.* **2020**, *142*, 12551–12562; d) Z. Ma, N. Nakatani, H. Fujii, M. Hada, *Bull. Chem. Soc. Jpn.* **2020**, *93*, 187–193; e) O. Kirshenboim, A. Frenklah, S. Kozuch, *Chem. Sci.* **2021**, *12*, 3179–3187; f) S. Yu, P. Vermeeren, T. A. Hamlin, F. M. Bickelhaupt, *Chem. Eur. J.* **2021**, *27*, 5683–5693; g) P. Besalú-Sala, M. Solà, J. M. Luis, M. Torrent-Sucarrat, *ACS Catal.* **2021**, *11*, 14467–14479.

[26] J. C. Gordon, J. B. Myers, T. Folta, V. Shoja, L. S. Heath, A. Onufriev, *Nucleic Acids Res.* **2005**, *33*, 368–371.

[27] P. Li, K. M. Merz, *J. Chem. Inf. Model.* **2016**, *56*, 599–604.

[28] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

[29] J. Träg, D. Zahn, *J. Mol. Model.* **2019**, *25*, 39–48.

[30] D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, D. Ghoreishi, M. K. Gilson, H. Gohlke, A. W. Goetz, D. Greene, R. Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. J. Mermelstein, K. M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R. C. Walker, J. Wang, H. Wei, R. M. Wolf, X. Wu, L. Xiao, D. M. York, P. A. Kollman, *AMBER-2018*, University of California, San Francisco **2018**.

[31] a) R. J. Loncharich, B. R. Brooks, R. W. Pastor, *Biopolymers* **1992**, *32*, 523–535; b) Y. Lin, D. Pan, J. Li, L. Zhang, X. Shao, *J. Chem. Phys.* **2017**, *146*, 124108.

[32] a) M. G. Quesne, T. Borowski, S. P. de Visser, *Chem. Eur. J.* **2016**, *22*, 2562–2581; b) T. S. Hofer, S. P. de Visser, *Front. Chem.* **2018**, *6*, 357.

[33] *Gaussian*-09, Revision D.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada,

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202104167

Chemistry Europe
European Chemical
Societies Publishing

M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian, Inc., Wallingford CT **2013**.

[34] T. Vreven, K. S. Byun, I. Komáromi, S. Dapprich, J. A. Montgomery Jr, K. Morokuma, M. J. Frisch, *J. Chem. Theory Comput.* **2006**, *2*, 815–826.

[35] H. S. Fernandes, M. J. Ramos, N. M. F. S. A. Cerqueira, *J. Comput. Chem.* **2018**, *39*, 1344–1353.

[36] D. Bakowies, W. Thiel, *J. Phys. Chem.* **1996**, *100*, 10580–10594.

[37] a) A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652; b) C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.

[38] a) P. J. Hay, W. R. Wadt, *J. Chem. Phys.* **1985**, *82*, 270–283; b) W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **1972**, *56*, 2257–2261; c) M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees, J. A. Pople, *J. Chem. Phys.* **1982**, *77*, 3654–3662.

[39] Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **2008**, *120*, 215–241.

[40] a) S. Shaik, D. Kumar, S. P. de Visser, A. Altun, W. Thiel, *Chem. Rev.* **2005**, *105*, 2279–2328; b) M. R. A. Blomberg, T. Borowski, F. Himo, R.-Z. Liao, P. E. M. Siegbahn, *Chem. Rev.* **2014**, *114*, 3601–3658.

[41] a) G. Mukherjee, A. Alili, P. Barman, D. Kumar, C. V. Sastri, S. P. de Visser, *Chem. Eur. J.* **2019**, *25*, 5086–5098; b) S. Louka, S. M. Barry, D. J. Heyes, M. Q. E. Mubarak, H. S. Ali, L. M. Alkhalaf, A. W. Munro, N. S. Scrutton, G. L. Challis, S. P. de Visser, *J. Am. Chem. Soc.* **2020**, *142*, 15764–15779; c) S. B. Han, H. S. Ali, S. P. de Visser, *Inorg. Chem.* **2021**, *60*, 4800–4815.

[42] a) F. G. Cantú Reinhard, A. S. Faponle, S. P. de Visser, *J. Phys. Chem. A* **2016**, *120*, 9805–9814; b) X.-X. Li, V. Postils, W. Sun, A. S. Faponle, M. Solà, Y. Wang, W. Nam, S. P. de Visser, *Chem. Eur. J.* **2017**, *23*, 6406–6418.

[43] a) J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.* **2005**, *105*, 2999–3009; b) S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.

[44] J. Bigeleisen, M. Wolfsberg, *Advances in Chemical Physics*, I. Prigogine (Ed.), John Wiley & Sons, Inc. **1958**, Volume l, pp 15–76.