# $O_2$ activation in salicylate 1,2-dioxygenase

## A QM/MM study reveals the role of His162

Dong, Geng; Ryde, Ulf

*Published in:*
Inorganic Chemistry

*DOI:*
10.1021/acs.inorgchem.6b01732

2016

*Document Version:*
Peer reviewed version (aka post-print)

# $O_2$ Activation in Salicylate 1,2-Dioxygenase: A QM/MM Study Reveals the Role of His162

*Geng Dong* and Ulf Ryde*

Department of Theoretical Chemistry, Lund University, Chemical Centre, P. O. Box 124, SE-221 00 Lund, Sweden

*Note that unfortunately the results in ref. 15 are incorrectly described: Mutants of Arg127 were inactive in the experiments, not active. We are sorry for this mistake.*

**KEYWORDS**

$O_2$ activation, nonheme iron enzymes, QM/MM, big-QM calculations, QM/MM free-energy perturbation, QTCP, density functional theory

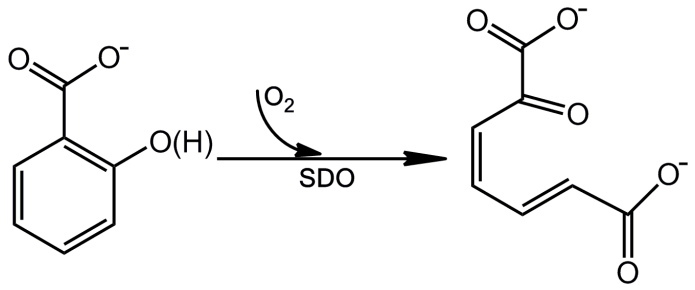
**ABSTRACT**

Nonheme iron enzymes play an important role in the aerobic degradation of aromatic ring systems. Most enzymes can only cleave substrates with electron-rich substituents, e.g. with two hydroxyl groups. However, salicylate 1,2-dioxygenase (SDO) can cleave rings with only a single hydroxyl group. We have investigated the oxygen-activation mechanism of the ring fission of salicylate by SDO by computational methods, using combined quantum mechanical and molecular mechanical (QM/MM) geometry optimisations, large QM calculations with 493 atoms, and QM/MM free-energy perturbations. Our results demonstrate that the reactive Fe–$O_2$ species is best described as a Fe(III)−$O_2^{\bullet-}$ state, which is triplet $O_2$ binding to quintet Fe(II), leading to a one-electron transfer from Fe(II) to $O_2$. Subsequently, the $O_2^{\bullet-}$ group of this species attacks the aromatic ring of substrate to form an alkylperoxo intermediate. Mutation studies have suggested that His162 is essential for the catalysis. Our calculations indicate that His162 plays a role as an acid−base catalyst, providing a proton to the substrate.

## INTRODUCTION

The cleavage of aromatic rings by nonheme iron enzymes is an important biological process in aerobic degradation. Three main classes of ring-fission dioxygenases have been discovered so far. The class I and II enzymes (also called intradiol and extradiol ring-cleaving dioxygenases) require two hydroxyl groups in ortho positions on the substrate (e.g. catechol or protocatechuate).[1-4] The class III enzymes can cleave aromatic compounds with only a single hydroxyl group. Salicylate 1,2-dioxygenase (SDO) belongs to the family of homo-tetrameric class III ring fission dioxygenases, together with 3-hydroxyanthranilate-3,4-dioxygenase, 1-hydroxy-2-naphthoate dioxygenase, gentisate dioxygenase, and homogentisate dioxygenase (HGDO).[5-13] The SDO enzyme utilizes $O_2$ to cleave the aromatic ring of the substrate salicylate (Sal) and inserts two oxygen atoms into the product (Scheme 1).

**Scheme 1.** Overall reaction catalysed by SDO




There are some similarities between class III and II enzymes. For instance, they both have a Fe(II) ion in their active sites, whereas a Fe(III) ion is utilized in the intradiol dioxygenases.[1]

However, the coordination in extradiol dioxygenases and SDO is different. The former contains a 2-His–1-carboxylate motif, whereas three His residues bind to the Fe(II) ion in SDO. In extradiol dioxygenases, there is a conserved His residue that can act as an acid−base catalyst. In SDO, residues Arg127 and His162 may play a similar role.[5] Mutation studies[14, 15] have shown that His162 is essential for catalysis: The His162Ala SDO mutant lost almost all activities when reacting with salicylate. However, when Arg127 was mutated to Ile, Asn, or Lys, no significant loss of activity was observed. *This is incorrect: Mutants of Arg127 were inactive according to ref. 15.*

Dioxygen activation by mononuclear nonheme iron enzymes has been extensively studied by experimental methods.[3, 4, 16-19] However, many details of the mechanisms are still unclear. Therefore, many groups have employed quantum mechanical (QM), as well as combined QM and molecular mechanics (QM/MM) methods to study these enzymes.[20-22] The $O_2$ adduct (Fe–$O_2$) has been proposed to be the reactive oxygen species in homoprotocatechuate 2,3-dioxygenase (HPCD, one of the extradiol dioxygenases), using different substrates.[23-26] It was first described as Fe(II)−$O_2^{\bullet-}$–substrate$^{\bullet}$ state by Siegbahn and Haeffner.[25] Neese and coworkers instead suggested a Fe(III)−$O_2^{\bullet-}$ species using a larger QM model,[26] whereas a Fe(III)−$O_2^{\bullet-}$/Fe(II)−$O_2^{\bullet-}$–substrate$^{\bullet}$ hybrid state was found by QM/MM calculations.[23] For the intradiol dioxygenases studies, a Fe(III)−$O_2^{\bullet-}$ state was proposed as reactive species,[28-30] whereas a Fe(II)−$O_2^{\bullet-}$–substrate$^{\bullet}$ species was found in the aminophenol dioxygenase.[31] One of the class III enzymes, HGDO, has been studied by Siegbahn and coworkers, and they found the reactive Fe–$O_2$ species to be in the Fe(II)−$O_2^{\bullet-}$–substrate$^{\bullet}$ state.[32] Recently, HGDO was reinvestigated by Lai and coworkers using QM/MM methods.[33] They considered both protonated and deprotonated states of substrate and located both the Fe(II)−$O_2^{\bullet-}$–substrate$^{\bullet}$ and Fe(III)–$O_2^{\bullet-}$ states. On the

other hand, oxygen activation by cysteine dioxygenase follows a different mechanism.[27, 34] The ground state of the Fe–$O_2$ adduct is the singlet Fe(II)−$O_2^{\bullet-}$ state, and the attack of the superoxo group on the substrate involves a spin-state crossing from singlet to quintet. In addition, in some other nonheme enzymes, Fe(III)−$O_2^{\bullet-}$ was suggested as the reactive species, viz. $\alpha$-ketoacid-dependent oxygenases,[35-39] isopenicillin N synthase,[40, 41] hydroxyethylphosphonate dioxygenase,[42] and (S)-2-hydroxypropyphosphonate epoxidase.[3, 4, 16-19, 43]

For SDO, the oxygen-activation mechanism has been investigated using QM/MM calculations and the reactive oxygen species was deemed to be in between the Fe(II)−$O_2$ and Fe(II)−$O_2^{\bullet-}$–salicylate$^{\bullet}$ states, with a partial electron transfer from the substrate to $O_2$.[44] However, other possible states were not considered and they did not include the essential His162 in their QM region. In this article, we investigated the oxygen activation in SDO by QM/MM, QM-cluster, and big-QM calculations, as well as QM/MM free-energy perturbations. In particular, we study the role of the active-site His162 and Arg127 residues.

## METHODS

### *The protein*

All calculations were based on the 2.1-Å crystal structure of salicylate 1,2-dioxygenase with the substrate salicylate from the bacterium *Pseudaminobacter salicylatoxidans* (PDB code 3NJZ), which contains only a single subunit.[6] $O_2$ was added to the iron centre, giving a Fe(II) ion coordinated by three His, $O_2$, and the substrate, as is shown in Figure 1.

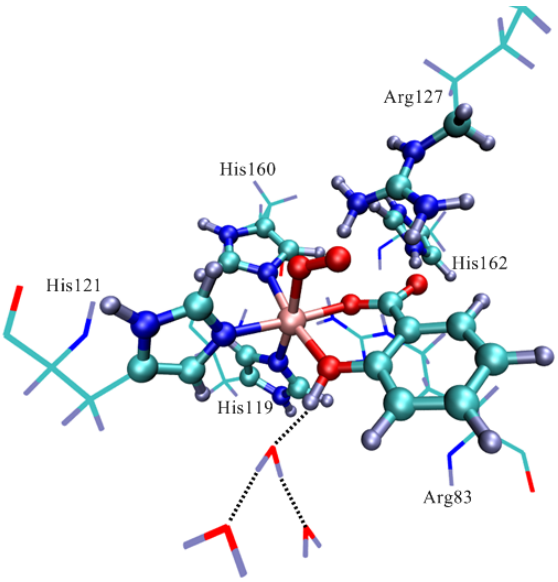



**Figure 1.** The active site of SDO. The atoms shown in ball-and-stick mode were included in the QM region in the QM/MM calculations. The $O_2$ molecule and the three water molecules were not present in the crystal structure.

The enzyme was set up in the same way as in our previous studies.[45] Residues Glu4 to Ile48 and 42 water molecules close to these residues were removed, because they are more than 37 Å away from active site. The protonation state of all the residues were determined by using PROPKA,[46] a study of the hydrogen-bond pattern around the His residues, the solvent accessibility, and the possible formation of ionic pairs. All Arg, Lys, Asp, and Glu residues were assumed to be charged. His119, His121, and His160 were protonated on the ND1 atom, His248 was assumed to be protonated on the NE2 atom, whereas the other His residues were assumed to be doubly protonated.

His162 forms a hydrogen bond to the carboxyl group of substrate and to the hydroxyl group of Thr137. Therefore, it may transfer protons to substrate. The hydroxyl group of salicylate has a p$K_a$ value of 13.8,[47] indicating that it is protonated in water solution. However,

when it binds to Fe, the p$K_a$ value is reduced and the proton may be transferred to solution through a chain of water molecules (cf. Figure 1). Therefore, we studied both the protonated and deprotonated states of the substrate. In the previous study, only the deprotonated state of the substrate was considered.[44] We found that if the substrate is deprotonated, the proton on the doubly protonated His162 transferred to carboxyl group of substrate in the QM/MM optimisations. Therefore, to obtain a deprotonated substrate, a singly protonated His162 had to be considered. Consequently, three models were used in this, viz. the substrate protonated on the hydroxyl group and doubly protonated His162, the substrate protonated on the carboxylate group and His162 protonated on the ND1 atom, and the deprotonated substrate with His162 protonated on the NE2 atom. The net charge of the simulated systems was neutral, –1 and –2 in the three simulations, respectively. No counter ions were added for any of the systems.

The protein was solvated with water molecules forming a sphere with a radius of 40 Å around the geometric centre, using the leap module of the Amber software package (6240 water molecules and 23 640 atoms in total).[48] The added protons and water molecules were optimised by a 120-ps simulated annealing calculation, followed by a minimisation, keeping the other atoms fixed at the crystal-structure positions.

*QM/MM calculations*

The QM/MM calculations were performed with the ComQum software.[49, 50] In this approach,[51, 52] the protein and solvent are split into two subsystems: System 1 (the QM region) was relaxed by QM methods. It consisted of the iron ion, the first-coordination-sphere of Fe (His119, His121, His160, salicylate and $O_2$), as well as His162 and Arg127. Histidines were modelled as an imidazole and arginine as methylguanidinium (ball-and-stick model in Figure 1). System 2

contained the remaining part of the protein and the solvent. It was kept fixed at the original (crystallographic) coordinates in the QM/MM calculations (but it was relaxed in the free-energy perturbations). We also tried calculations in which all atoms in any residue within 6 Å of any atom in system 1 were relaxed by a full MM minimisation in each step of the QM/MM geometry optimisation. They gave similar results, as can be seen from the geometries and energies in Table S2 in the supporting information (SI).

In the QM calculations, system 1 was represented by a wavefunction, whereas all the other atoms were represented by an array of partial point charges, one for each atom, taken from MM libraries. Thereby, the polarisation of the QM region by the surroundings is included in a self-consistent manner (electrostatic embedding). When there is a bond between systems 1 and 2 (a junction), the hydrogen link-atom approach was employed: The QM region 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.[49, 53] All atoms were included in the point-charge model, except the CL atoms.[54]

The total QM/MM energy in ComQum was calculated as[49, 50]

$$E_{\mathrm{QM/MM}} = E^{\mathrm{HL}}_{\mathrm{QM1+ptch2}} + E^{\mathrm{CL}}_{\mathrm{MM12},q_1=0} - E^{\mathrm{HL}}_{\mathrm{MM1},q_1=0} \qquad (1)$$

where $E^{\mathrm{HL}}_{\mathrm{QM1+ptch2}}$ is the QM energy of the QM region truncated by HL atoms and embedded in the set of point charges modelling system 2 (but excluding the self-energy of the point charges). $E^{\mathrm{HL}}_{\mathrm{MM1,q1=0}}$ is the MM energy of the QM region, still truncated by HL atoms, but without any electrostatic interactions. Finally, $E^{\mathrm{CL}}_{\mathrm{MM12,q1=0}}$ is the classical energy of all atoms in the system with CL atoms and with the charges of the QM region set to zero (to avoid double counting of the electrostatic interactions). By this approach, which is similar to the one used in the ONIOM method,[55] errors caused by the truncation of the QM region should cancel.

The geometry optimisations were continued until the energy change between two iterations was less than 2.6 J/mol ($10^{-6}$ a.u.) and the maximum norm of the Cartesian gradients was below $10^{-3}$ a.u. The QM calculations were carried out using Turbomole 7.0 software.[3] The geometry optimisations were performed using the B3LYP[56-58] functional in combination with def2-SV(P)[59] basis set, including empirical dispersion corrections with the DFT-D3 approach, as implemented in Turbomole.[60] The MM calculations were performed with the Amber software,[48] using the Amber ff14SB force field.[61] The energies were improved by single-point calculations using the larger def2-TZVP[62] and def2-QZVP[62, 63] basis set for all the atoms. The discussed spin densities (based on Mulliken populations) were obtained at the B3LYP-D3/def2-SV(P) level.

*QM-cluster calculations*

QM-cluster calculations were performed with Turbomole 7.0 software.[3] Geometries were obtained at the B3LYP-D3/def2-SV(P) level in gas phase without any restraints. Geometries were also optimised at the same level in a conductor-like screening model (COSMO) continuum solvent with a dielectric constant of 4 or 80, employing optimised radii (and 2.0 Å for Fe) and default parameters for a water-like solvent.[64-66] Single-point energy calculations were performed at B3LYP-D3/def2-TZVPD level in both gas phase and COSMO models with the dielectric constant of 4 or 80. Acidity constants (p$K_a$ values) were estimated using the solvated energies and using a value of 1131.00 kJ/mol for the solvation and translation free energy of the proton.[67]

Ten models were employed in this paper (Figure 2). They were used to estimate relative stability of the various protonation states and to study the roles of His162 and Arg127. In Models 1, 4, and 8 the substrate was deprotonated, whereas in the other seven models, the substrate was protonated either on the non-coordinating carboxylic oxygen atom (models 2, 3, 5, and 6) or on

the phenolic oxygen atom (models 7 and 9–10). Models 4–6 included a model of Arg127, whereas models 3, 6, 8, 9, and 10 included a model of His 162. The latter group was doubly protonated in model 10 and singly protonated in the other models, donating a hydrogen bond to the non-coordinating carboxylic oxygen atom of the substrate in models 8 and 9, but receiving a hydrogen bond from the protonated non-coordinating carboxylic oxygen atom of the substrate in models 3 and 6.

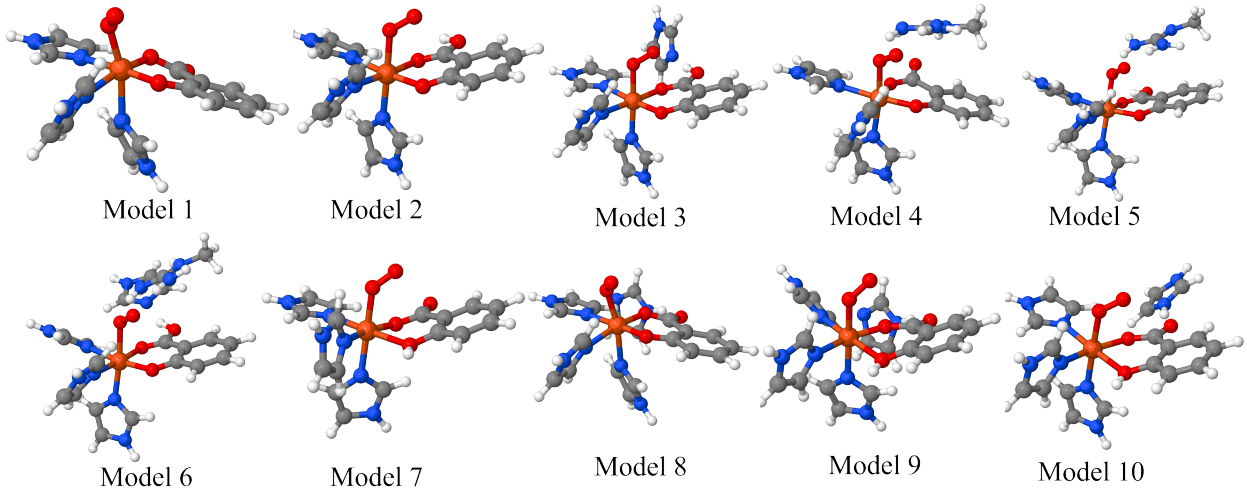



**Figure 2.** The models used in the QM-cluster calculations

*Big-QM calculations*

To obtain stable and converged energies, we employed the big-QM approach:[68] The minimal QM region, consisting of the Fe ion and its first-sphere ligands, was extended with all residues with at least one atom within 4.5 Å of the minimal QM region (including eight water molecules) and junctions were moved at least two residues away. In addition, all charged groups, buried inside the protein were included (Arg83 and Asp174). This gave a QM region of 493 atoms, shown in Figure 3. The big-QM calculations were performed on coordinates from the QM/MM

calculations and with a point-charge model of the surroundings, because this gave the fastest calculations in our previous test.[69] To this big-QM energy, we added the DFT-D3 dispersion correction, calculated for the same big QM region with Becke–Johnson damping,[70] third-order terms, and default parameters for the B3LYP functional using dftd3 program.[60] We also included a MM correction

$$E_{\mathrm{MM}} = E^{\mathrm{CL}}_{\mathrm{MM12},q_1=0} - E^{\mathrm{CL}}_{\mathrm{MM1},q_1=0} \qquad (2)$$

yielding a standard QM/MM energy, but with the big-QM system as the QM region.

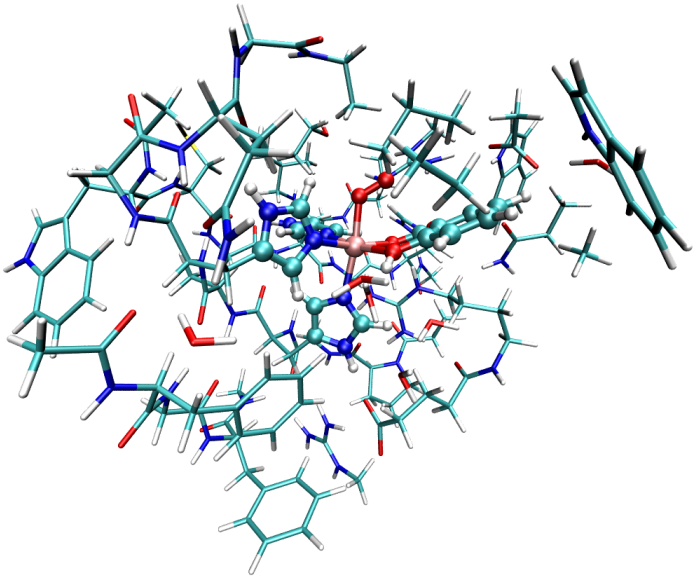


**Figure 3.** Atoms included in the big-QM calculations

*QTCP calculations*

The QTCP approach (QM/MM thermodynamic cycle perturbation) is a method to calculate the free energy difference between two states, A and B, with a high-level QM/MM method, using sampling at only the MM level.[71-73] It employs the thermodynamic cycle in Figure 4, showing that the free-energy difference is obtained from three terms:

$$\Delta G_{\mathrm{QTCP}}(\mathrm{A} \rightarrow \mathrm{B}) = -\Delta G_{\mathrm{MM}\rightarrow\mathrm{QM/MM}}(\mathrm{A}) + \Delta G_{\mathrm{MM}}(\mathrm{A} \rightarrow \mathrm{B}) + \Delta G_{\mathrm{MM}\rightarrow\mathrm{QM/MM}}(\mathrm{B}) \qquad (3)$$

The QTCP calculations were performed as has been described before:[72-74] First, each state of interest was optimised by QM/MM, keeping system 2 fixed at the crystal structure. Then, the protein was further solvated in an octahedral box of TIP3P water molecules,[75] extending at least 9 Å from the QM/MM system. For one of the states, the system was first subjected to a 1000-step minimisation, keeping the atom in the QM region fixed and restraining all heavy atoms in the crystal structure with a force constant of 418 kJ/mol/Å$^2$. Then, two 20-ps MD simulations were run with the heavy atoms still restrained. The first was run with a constant volume and the second with a constant pressure. Next, the size of the periodic box was equilibrated by a 100-ps MD simulation with a constant pressure and only the heavy atoms in QM region restrained to the QM/MM structure. Finally, for all states, an equilibration of 200 ps and a production simulation of 400 ps were run with a constant volume for each state. During the production run, snapshots were collected every 2 ps.

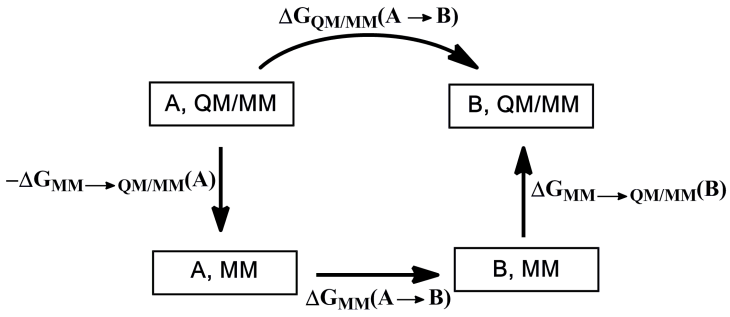



**Figure 4.** The thermodynamic cycle employed in the QTCP calculations

Based on these 200 snapshots, three sets of free-energy perturbations (FEPs) were performed, as shown in Figure 4. First, FEPs were performed at MM level in the forward and reverse directions along the reaction coordinate by changing the charges and coordinates of the QM region to those of the QM/MM calculations,[74] The charges were first modified in nine steps, keeping the coordinates at those of the A state. Then, the coordinates were modified in five steps

to those of the B state (with the charges of the B state). Second, MM → QM/MM FEPs were performed for both the A and B states, keeping the QM regions fixed, as has been described before.[72, 73] 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.

Reported energies are the big-QM/MM energies including dispersion and the MM correction ($E^{\mathrm{SV(P)}}_{\mathrm{bigQM/MM}}$). This energy was extrapolated to the def2-QZVP level using QM calculations with the normal QM region and including the point-charge model ($E^{\mathrm{QZVP}}_{\mathrm{QM}} - E^{\mathrm{SV(P)}}_{\mathrm{QM}}$):

$$E_{\mathrm{bigQM/MM}} = E^{\mathrm{SV(P)}}_{\mathrm{bigQM/MM}} + E^{\mathrm{QZVP}}_{\mathrm{QM}} - E^{\mathrm{SV(P)}}_{\mathrm{QM}} \qquad (4)$$

The big-QM/MM energies were used to compare the energy difference of similar structures in different spin states. To the big-QM/MM energies, the QTCP energy correction (i.e. the difference between the QTCP and QM/MM energies, $E^{\mathrm{SV(P)}}_{\mathrm{QTCP}} - E^{\mathrm{SV(P)}}_{\mathrm{QM/MM}}$) was added:

$$E_{\mathrm{tot}} = E_{\mathrm{bigQM/MM}} + E^{\mathrm{SV(P)}}_{\mathrm{QTCP}} - E^{\mathrm{SV(P)}}_{\mathrm{QM/MM}} \qquad (5)$$

The final energy $E_{\mathrm{tot}}$ was used to describe the reaction energy barrier. The various energy components are collected in Table S1 in the supporting information.

## RESULTS AND DISCUSSION

The binding of triplet $O_2$ to quintet Fe(II) can result in four spin states, viz. the septet, quintet, triplet, and singlet state. In this paper, we examined all these spin state for Fe–$O_2$ adduct. However, the singlet state was more than 40 kJ/mol higher in energy than the most stable state, so it will not be further discussed.

### *The Fe–$O_2$ adduct with a protonated substrate*

First, we studied the reaction of a protonated substrate with SDO, i.e. salicylate with a deprotonated carboxylate group but with a protonated hydroxyl group and His162 doubly protonated. In the septet and triplet states, $O_2$ bound only weakly to the iron ion with Fe–$O_1$ distances of 2.59 and 2.47 Å, respectively ($O_1$ and $O_2$ are the oxygen atoms of $O_2$ with $O_1$ coordinating to Fe, as can be seen in Figure 5). However, in the quintet state ($^5$**A1**), a much tighter complex was obtained with Fe–$O_1$ and $O_1$–$O_2$ bond lengths of 2.04 and 1.25 Å, respectively, and a Fe–$O_1$–$O_2$ angle of 125˚. In the previous QM/MM study (using the similar B3LYP/def2-SVP method for the geometries, but with a deprotonated substrate and without any His162 model in the QM region),[44] the Fe–$O_1$ bond length in the quintet state was 1.98 Å and the Fe–$O_1$–$O_2$ angle was 124˚. For HPCD, a side-on binding conformation of $O_2$ was trapped in both experimental[76] and computational studies.[23, 24] We tried to find such a side-on species also for SDO, but without any success. It seems likely that the side-on $O_2$ conformation in HPCD is stabilised by interactions with a second-sphere His residue that is not present in SDO.

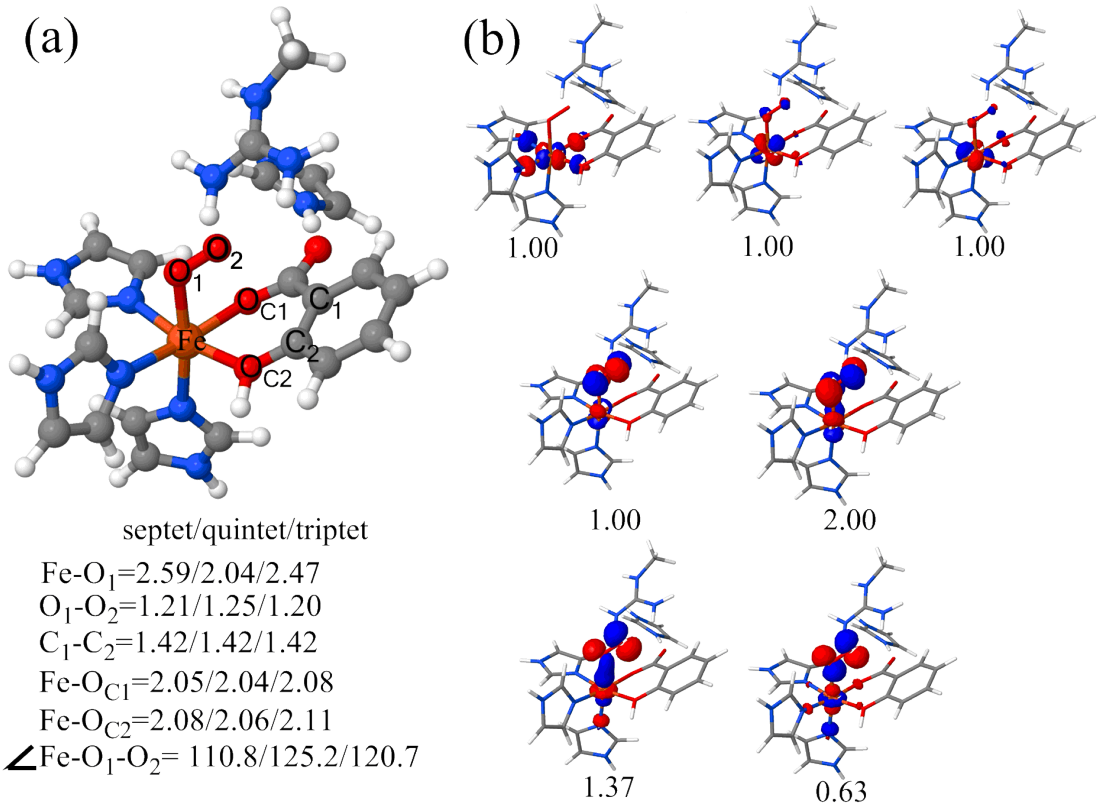



**Figure 5.** (a) Optimised structure of the Fe–$O_2$ adduct with a protonated substrate in various spin states. All distances are in Å and angles are in degrees. (b) Fe 3*d* and $O_2$ π* natural orbitals in the quintet state together with the occupation numbers.

Next, we focus on the electronic structure of $^5$**A1**. It can be seen from the natural orbitals shown in Figure 5 that there are three singly-occupied, almost pure, Fe 3*d* orbitals. In addition, there are four orbitals that are mixtures of Fe 3*d* and $O_2$ π* orbitals. Among the two orbitals involving the $\pi^*_{op}$ orbitals (out of the $FeO_2$ plane), one is singly occupied and the other is doubly occupied. Since the Fe 3*d* and $O_2$ π* orbitals have the almost equal contributions, we can assign 1.5 electrons each to Fe and $O_2$. On the other hand, the two orbitals involving the $O_2$ $\pi^*_{ip}$ orbitals (in the $FeO_2$ plane) have occupation number of 0.63 and 1.37. Again, the Fe 3*d* and $O_2$ π*

orbitals have almost equal contributions, so we can assign one electron each to Fe and $O_2$. This would give 5.5 electrons on Fe and 2.5 electron on $O_2$, indicating a partial electron transfer from the Fe(II) $3d_{yz}$ orbital to the $\pi^*_{op}$ orbital when $O_2$ binds to Fe, i.e. the electronic structure is in between Fe(II)–$O_2$ and Fe(III)–$O_2^{\bullet-}$. This is supported by the spin populations on Fe, $O_2$ and the substrate, which are 4.03, –0.28, and 0.07. They show an antiferromagnetic coupling between Fe and $O_2$, and indicate somewhat lower populations on Fe and $O_2$ than expected from the orbitals. However, it is always observed that high-spin Fe(III) delocalises some of its spin on all first-sphere ligands,[23] which would decrease the spin both on Fe and on the antiferromagnetically coupled $O_2$. In particular, the spin population on substrate is 0.07, which is similar to what is found on the His ligands (0.04–0.05). Thus, no significant radical character on the substrate was detected.

In addition, another quintet Fe–$O_2$ species ($^5$**A1'**) was also located. The geometry of $^5$**A1'** is similar to that of $^5$**A1**, but they have different electronic structures (Figure S1). In $^5$**A1**, half of an electron is transferred from the Fe(II) $3d_{yz}$ orbital to the $O_2$ $\pi^*_{op}$ orbital. However, in $^5$**A1'**, this partial electron is instead transferred from the Fe(II) $3d_{xz}$ orbital to the $O_2$ $\pi^*_{ip}$ orbital, whereas the $O_2$ $\pi^*_{op}$ orbital is singly occupied. This species is 14 kJ/mol ($E_{\text{bigQM/MM}}$) less stable than $^5$**A1**.

*The Fe–$O_2$ adduct with the substrate protonated on the carboxylate group*

Next, we studied the Fe–$O_2$ adduct with a deprotonated substrate (i.e. salicylate with a double negative charge, deprotonated on both the carboxylate and hydroxyl groups) and His162 still doubly protonated. Interestingly, when the substrate was deprotonated, a proton transfer was always observed from His162 to the non-coordinating carboxylate oxygen atom of the substrate during the QM/MM optimisation, so that the latter still was a monoanion (Figure 6). This is

consistent with the rather high second p$K_a$ value for the salicylate (13.8). In the previous QM/MM study of SDO,[44] His162 was not included in the QM region, so they could not observe such a proton transfer. We found that the O–O bond length in the quintet state ($^5$**A2**) is 0.04 and 0.08 Å longer than in septet ($^7$**A2**) and triplet states ($^3$**A2**), and the Fe–$O_1$ bond is 0.27 and 0.33 Å shorter, respectively. This means that interaction between Fe and dioxygen is stronger in the quintet state than in the septet and triplet states, as was also observed for the protonated substrate. Compared to $^5$**A1**, the Fe–$O_1$ distance is 0.03 Å shorter in $^5$**A2** and the O–O bond is 0.04 Å longer. Moreover, the Fe–$O_{C2}$ bond is shorter (because $O_{C2}$ is deprotonated) and the Fe–$O_{C1}$ bond is longer because the carboxyl group is protonated. Finally, the hydrogen bond between Arg127 and $O_2$ is 0.27 Å shorter.

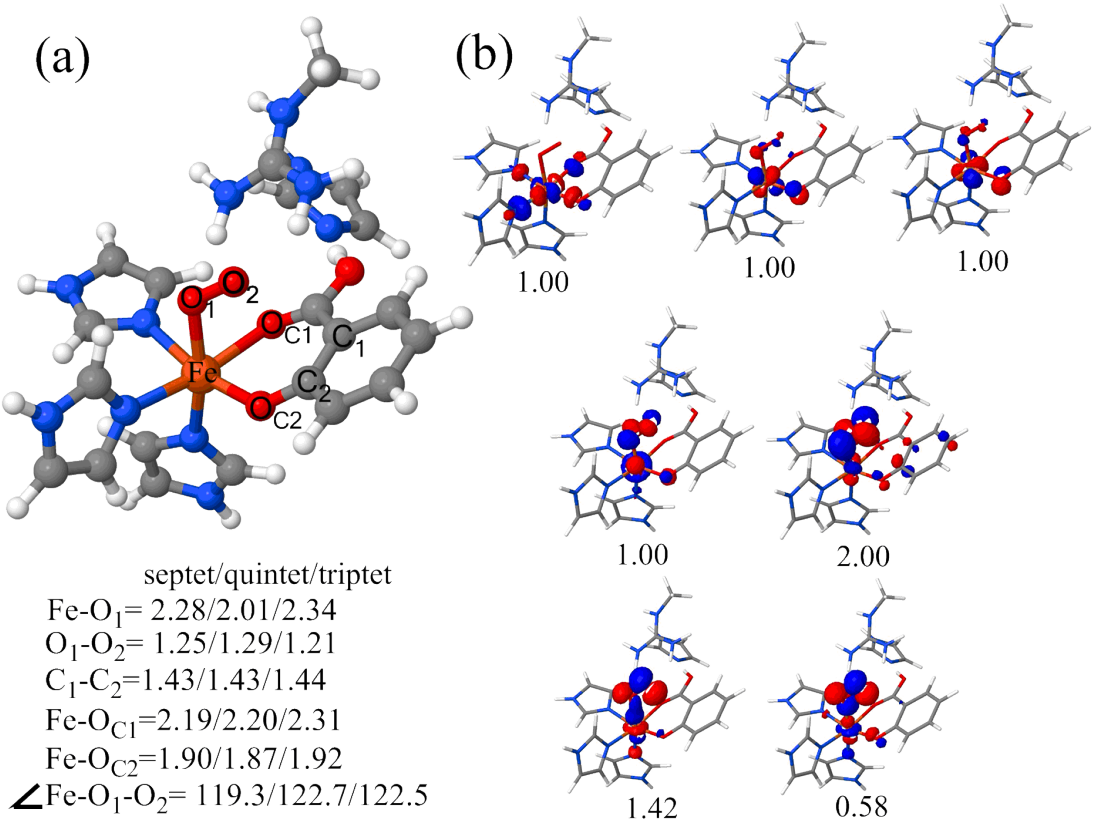


**Figure 6.** (a) Optimised structure of the Fe–$O_2$ adduct with the substrate protonated on the carboxylate group (by a proton transfer from His162). All distances are in Å and angles are in degrees. (b) Fe 3*d* and $O_2$ π* natural orbitals in the quintet state together with the occupation numbers.

The septet and triplet states involve the triplet $O_2$ ferromagnetic or antiferromagnetic coupled with a quintet Fe(II). The orbitals of the $^5$**A2** state (shown in Figure 6b) are similar to those of the $^5$**A1** state, except those involving the $O_2$ $\pi^*_{op}$ orbitals. The latter two orbitals (middle row in Figure 6b), now have uneven contributions: The one that is singly occupied is mainly Fe 3*d*, whereas the doubly occupied one is dominated by $O_2$ π*. This indicates that there are more electrons on $O_2$, which is also confirmed by the spin populations on Fe and $O_2$, which are 4.10 and –0.60, i.e. showing a significant radical character on $O_2$. Thus, $^5$**A2** is close to a Fe(III)−$O_2^{\bullet-}$ state (with the electronic configuration $\pi^2_{op}\pi^1_{ip}$), with an antiferromagnetic coupling between Fe and $O_2$. In the previous QM/MM study,[44] the spin population on $O_2$ was only −0.27, which was interpreted as a Fe(II)−$O_2$ state. Our results show that His162 has strong influence on the electronic structure of SDO and therefore should be included in the QM region.

The spin density on the substrate is 0.32 in the $^5$**A2** state. This is significantly larger than for the $^5$**A1** state, but it is still only a minor radical character. Most of the spin is found on the $O_{C2}$ atom that binds to Fe (0.23). In the high-spin Fe(II) complex without $O_2$, the spin on the salicylate substrate is 0.15.

We also found another quintet state species in which the Fe centre is intermediate spin, ferromagnetically coupled to $O_2$ (the spin populations are 2.68 on Fe and 1.28 on $O_2$). However, this state was 66 kJ/mol less stable than $^5$**A2** at the B3LYP/def2-TZVP level.

Previous B3LYP calculations on nonheme iron enzymes have predicted a septet ground

state.[23, 24, 32, 44] However, in this work the quintet state is predicted to be ground state, based on the big-QM energies (Eq. 4), but the septet ($^7$**A2**) and triplet ($^3$**A2**) states are only 5 and 1 kJ/mol less stable than $^5$**A2**, respectively. This shows that all three states are close in energy. However, for the subsequent steps in reaction mechanism, only the quintet state was reactive: For the septet state, no transition states or products were found. For the triplet state, the transition state (at B3LYP/def2-TZVP level) lay 139 kJ/mol above the transition state in quintet state, and the product was 73 kJ/mol less stable than the product in quintet.

The **A2** model differs from the **A1** model in that it contains one proton less. Moreover, the substrate is protonated on the non-coordinating carboxylic oxygen atom, rather than on the phenolic oxygen atom. The latter may seem unexpected, because a phenolic group is appreciably less acidic than a carboxylate group (the p$K_a$ values of benzoic acid and phenol are 4.2 and 10.0, respectively). However, the coordination to the metal ion changes this preference: the QM-cluster models 3 and 9 in Figure 2 contain the same atoms, but in model 3, the substrate is protonated on the non-coordinating carboxylate oxygen atom (and this proton is involved in a hydrogen bond to the His162 model), whereas in model 9, the substrate is instead protonated on the phenolic oxygen atom (and the non-coordinating carboxylic oxygen atom receives a hydrogen bond from the His162 model). QM calculations show that model 3 is actually 82, 59, and 44 kJ/mol more stable than model 9 in gas phase or in a COSMO continuum solvent with dielectric constants of 4 (a protein-like environment) or 80 (water), respectively. Likewise, model 2 (with the substrate protonated on the non-coordinating carboxylate oxygen atom, but no His model) was 72, 46, and 27 kJ/mol more stable than model 7 (protonated on the phenolic oxygen atom) in gas phase or in a COSMO continuum solvent with dielectric constants of 4 or 80, respectively. For this reason, we did not perform any QM/MM calculations on the tautomer

with the phenolic oxygen atom of the substrate protonated.

*The Fe–$O_2$ adduct with a deprotonated substrate*

In order to obtain a structure of the Fe–$O_2$ adduct with a fully deprotonated substrate (with a double negative charge, as in the previous QM/MM study of SDO[44]), we changed the protonation state of His162 from doubly to singly protonated with the proton on the NE2 atom that forms a hydrogen bond to the substrate. Such a state of His162 is possible, because it is connected to the solvent by a proton-transfer network involving Leu153, Thr154, a crystal-water molecule, and Thr137 (shown in Figure 7). The corresponding quintet state with the fully deprotonated substrate ($^5$**A3**) has spin populations on Fe, $O_2$, and substrate of 4.14, –0.73, and 0.44, respectively. Thus, $^5$**A3** is in the Fe(III)−$O_2^{\bullet-}$ state and has the same electronic configuration as $^5$**A2**. The Fe−$O_1$ bond length is 0.04 Å longer, the $O_1$−$O_2$ and Fe−$O_{C2}$ bonds are 0.01 Å longer, the $C_1$−$C_2$ bond length is same, whereas the Fe–$O_{C1}$ bond length is 0.19 Å shorter, compared to the $^5$**A2** state.

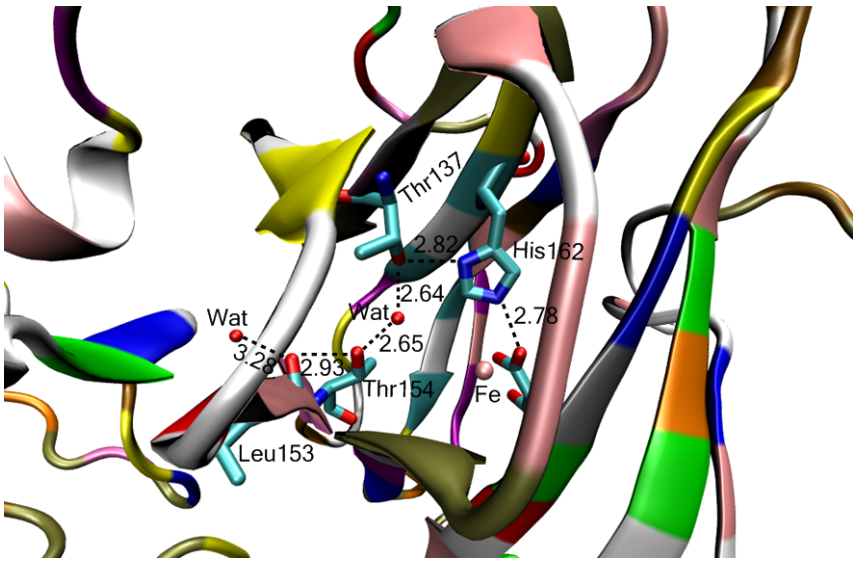



**Figure 7.** The hydrogen-bond network around the His162 residue in crystal structure

*Fe–$O_2$ attack on the substrate*

Several Fe–$O_2$ adducts in different spin states were located in this study using various protonation states of substrate. To understand which one is the reactive species, we studied the first step in the reaction mechanism, i.e. the attack of the Fe–$O_2$ adduct on substrate to form an alkylperoxo-bridged species ($^5$**B**). First, we studied the mechanism of protonated substrate reacting with enzyme, i.e. the attack of $O_2$ on $C_1$ in $^5$**A1** (Figure 8). The reaction energies were estimated by QTCP (Eq. 5). We found an activation energy barrier of 91 kJ/mol and the reaction was endothermic by 64 kJ/mol. The spin populations, which are 3.81 and 0.02 on the Fe and $O_2$, indicate that the product $^5$**B1** is a high-spin Fe(II)–alkylperoxo species, i.e. the delocalised electron has returned to Fe.

Next, we investigated the same reaction starting from $^5$**A2** species instead. In this case, the energy barrier was 77 kJ/mol, which is 14 kJ/mol lower than for $^5$**TS1**. The product $^5$**B2** was 72 kJ/mol higher in energy than $^5$**A2**. Finally, the reaction starting from deprotonated substrate $^5$**A3**

was investigated. The energy barrier was 71 kJ/mol, and the $^5$**B2** product was 54 kJ/mol higher in energy than $^5$**A3**. Thus, the activation barrier for $^5$**A3** was 5 kJ/mol lower than for $^5$**A2**.

In this discussion and in Figure 8, we have implicitly assumed that the three reactant states $^5$**A1**, $^5$**A2**, and $^5$**A3** are isoenergetic, to facilitate the comparison of the reaction barriers. Of course, the assumption is not true. The three states differ in the number of protons ($^5$**A1** has one proton more and $^5$**A3** one proton less than $^5$**A2**) and they can therefore be related by two p$K_a$ values. Unfortunately, it is much harder to estimate p$K_a$ values by QM methods than reaction and activation energies, because the net charge changes in the reaction, leading to very effects of the surroundings. Still, we have employed the QM cluster models 10, 8, and 3 to estimate the relative p$K_a$ values of states $^5$**A1**, $^5$**A2**, and $^5$**A3**. The results in Table 1 suggest that at pH 7, model 3 (corresponding to the $^5$**A2** state) is most stable. However, the results are sensitive to the dielectric constant employed in the COSMO continuum solvation model. With a dielectric constant of 4, often assumed to be similar to a protein environment, model 8 is only 2 p$K_a$ unit (11 kJ/mol) less stable, whereas with a water-like dielectric constant of 80, instead model 10 is only 1 p$K_a$ units (7 kJ/mol) less stable.

The experimental rate constant for salicylate reacting with wild-type SDO is 0.9 $s^{-1}$,[14] which corresponds to ~68 kJ/mol. This is in reasonable agreement with the barriers calculated for the $^5$**A3** and $^5$**A2** states. Clearly, the substrate is protonated before it binds to Fe. Upon binding, the proton transfers to solution and this is coupled with electron transfer from Fe to $O_2$. During this process, a proton on His162 will transfer to substrate if His162 is doubly protonated.

**Table 1.** Calculated p*K*a values of QM cluster models 3, 8, and 10, calculated at the B3LYP-D3/def2-TZVPD level using the COSMO continuum solvation model with two different dielectric constants (ε).

| States | ε | |
|---|---|---|
| | 4 | 80 |
| Model 10 to 3 | -14.0 | 5.8 |
| Model 3 to 8 | 8.9 | 12.5 |

We also examined the mechanism of $O_2$ attacking the $C_2$ atom of substrate, but we failed to find any transition state or product.

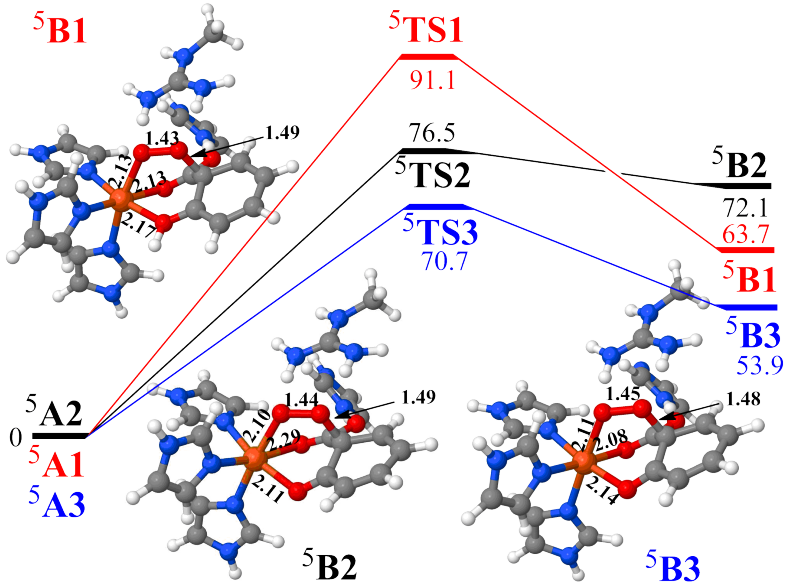



**Figure 8**. Reaction mechanism of Fe–$O_2$ attacking the substrate for the three protonation states of the substrate and His162, A1, A2, and A3. Energies are $E_{tot}$ ( Eq. 5) in kJ/mol.

The final energies are composed of several energy contributions as is shown in Eq. 5. These energies are listed in Table S1 in the Supporting Information. The final energies are based on the big-QM calculations at B3LYP/def2-SV(P) level. To this energy, a standard QM/MM term is added, as well as DFT-D3 energy corrections, including third-order terms and Becke–Johnson damping. The MM terms are quite small and typically negative (0–5 kJ/mol), owing to the large big-QM region. The dispersion corrections are 3–11 kJ/mol with a varying

sign. The basis-set corrections (to def2-QZVP) are 0–13 kJ/mol. The QTCP correction (difference between the QTCP and QM/MM results) is negative and up to 20 kJ/mol. It can be seen from Table S1 that the results for $^5$**A1** from each level of theory always gives the highest activation barrier. For the $^5$**A2** and $^5$**A3**, the energies are rather close (1–5 kJ/mol).

*Exploring the roles of Arg127 and His162*

Mutation experiments[15] have shown that His162 is essential for catalysis, because the His162Ala mutant was almost inactive. However, when Arg127 was mutated to Ile, Asn, or Lys, no significant loss of activity was observed. *This is incorrect: Mutants of Arg127 were inactive according to ref. 15.* We have used QM-cluster calculations to reveal the roles of the two residues using models 1–6 in Figure 2. We also studied the relationship between the geometric and electronic structure using QM/MM method.

For each of the six models, we optimised the geometry of the $O_2$ adduct. The resulting structures are shown in Figure 2 and all key bond distances, angles, and spin densities are listed in Table 2. Four of models (1 and 4–6) give a strong radical character on $O_2$ (0.7–0.8 spin density) and also a relatively long $O_1$–$O_2$ bond 1.29–1.31 Å, indicating a superoxide character. The Fe–$O_{C1}$ bond distance is long if the substrate is protonated (Models 2, 3, 5, and 6). From Model 1, we can see that when the $O_2$ binds to Fe(II) and the substrate is deprotonated, one electron is transferred from Fe to $O_2$. If the substrate is protonated (Models 2 and 3), the electron cannot transfer. As we have seen, His162 plays the role of acid–base catalyst, protonating the substrate if it is doubly protonated. This would keep the electron on Fe(II) in the absence of other residues. However, comparing Models 2 and 3 with 3 and 6, we see that the electron can transfer if a model of Arg127 included, forming a hydrogen bond with $O_2$. This indicates that Arg127

also has a role, promoting the electron transfer from Fe to $O_2$. The experiments showed that the enzyme reactivates did not be change when Arg127 was mutated to Ile, Asn, or Lys. *This is incorrect: Mutants of Arg127 were inactive according to ref. 15.* Asn and Lys can also form a hydrogen bond to $O_2$ and may play a similar role.

**Table 2.** The bond distance and angle for the six models in QM-cluster calculations

| | Bond distance (Å) | | | | Angle (°) | Spin density | | |
|---|---|---|---|---|---|---|---|---|
| | $O_1$–$O_2$ | Fe–$O_1$ | Fe–$O_{C1}$ | Fe–$O_{C2}$ | ∠Fe–$O_1$–$O_2$ | Fe | $O_2$ | Sal[a] |
| Model 1 | 1.29 | 1.99 | 1.93 | 1.91 | 124.9 | 4.12 | -0.67 | 0.42 (0.34) |
| Model 2 | 1.24 | 2.04 | 2.12 | 1.94 | 125.0 | 3.90 | -0.15 | 0.13 |
| Model 3 | 1.25 | 2.01 | 2.07 | 1.93 | 124.6 | 3.94 | -0.24 | 0.20 |
| Model 4 | 1.31 | 2.06 | 1.97 | 1.90 | 122.6 | 4.17 | -0.82 | 0.48 (0.36) |
| Model 5 | 1.30 | 2.07 | 2.15 | 1.89 | 117.0 | 4.21 | -0.82 | 0.36 (0.28) |
| Model 6 | 1.30 | 2.06 | 2.11 | 1.88 | 120.4 | 4.18 | -0.83 | 0.41 (0.31) |

[a] Values in brackets is the sum of the spin density on $O_{C1}$ and $O_{C2}$

Therefore, we used QM/MM calculations to investigate if the Arg127 affects the electron transfer in the protein. We mutated the Arg127 to Ile and put it into MM system for the $^5$**A2** and $^5$**A3** states. The spin populations on Fe and $O_2$ were 3.83 and 0.02 in $^5$**A2** and 3.90 and –0.12 in $^5$**A3**, respectively. This indicates that no electron is transferred from Fe(II) to $O_2$. Thus, our results indicate that a Arg127Ile mutation would change the electronic state of the Fe–$O_2$ adduct. However, it is possible that a water molecule may replace the hydrogen bond to the substrate in the mutant.

## CONCLUSION

In this paper, the oxygen-activation mechanism of salicylate 1,2-dioxygenase has been investigated by QM/MM methods, and we have also explored the roles of the active-site residues His162 and Arg127. Our calculations demonstrate that the reactive oxygen species is Fe(III)−$O_2^{\bullet-}$, indicating that $O_2$ attracts one electron when binding to Fe(II). Three protonation states of His162 were used to investigate the oxygen-activation mechanism and our results show similar reactivity in the reaction starting from $^5$**A2** and $^5$**A3**. The proton of His162 in $^5$**A2** will transfer to the carboxyl atom of substrate. Thus, His162 plays the role of an acid–base catalyst.

## ASSOCIATED CONTENTS

### Supporting Information

Energy components in Eq. 5; Fe 3*d* and $O_2$ $\pi^*$ orbitals in $^5$**A1'**; results of calculations in which all atoms within 6 Å of the QM system were relaxed; and the Cartesian coordinates of key intermediates. This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

### Corresponding Author

*E-mail: Geng.Dong@teokem.lu.se

### Author Contributions

GD did all the calculations. GD and UR wrote the manuscript together.

### Notes

Both authors have given approval to the final version of the manuscript. The authors declare no competing financial interests.

## ACKNOWLEDGMENT

This investigation has been supported by grants from the Swedish research council (project 2014-5540), the China Scholarship Council, and COST through Action CM1305 (ECOSTBio). The computations were performed on computer resources provided by the Swedish National Infrastructure for Computing (SNIC) at Lunarc at Lund University.

**REFERENCES**

(1) Que, L.; Ho, R. Y. N. Dioxygen activation by enzymes with mononuclear non-heme iron active sites. *Chem. Rev.* **1996,** *96*, 2607-2624.

(2) Bugg, T. D. H.; Ramaswamy, S. Non-heme iron-dependent dioxygenases: unravelling catalytic mechanisms for complex enzymatic oxidations. *Curr. Opin. Chem. Biol.* **2008,** *12*, 134-140.

(3) Bruijnincx, P. C. A.; van Koten, G.; Gebbink, R. J. M. K. Mononuclear non-heme iron enzymes with the 2-His-1-carboxylate facial triad: recent developments in enzymology and modeling studies. *Chem. Soc. Rev.* **2008,** *37*, 2716-2744.

(4) Costas, M.; Mehn, M. P.; Jensen, M. P.; Que, L. Dioxygen activation at mononuclear nonheme iron active sites: enzymes, models, and intermediates. *Chem. Rev.* **2004,** *104*, 939-986.

(5) Matera, I.; Ferraroni, M.; Burger, S.; Scozzafava, A.; Stolz, A.; Briganti, F. Salicylate 1,2-dioxygenase from Pseudaminobacter salicylatoxidans: crystal structure of a peculiar ring-cleaving dioxygenase. *J. Mol. Biol.* **2008,** *380*, 856-868.

(6) Ferraroni, M.; Matera, I.; Steimer, L.; Burger, S.; Scozzafava, A.; Stolz, A.; Briganti, F. Crystal structures of salicylate 1,2-dioxygenase-substrates adducts: a step towards the comprehension of the structural basis for substrate selection in class III ring cleaving dioxygenases. *J. Struct. Biol.* **2012,** *177*, 431-438.

(7) Zhang, Y.; Colabroy, K. L.; Begley, T. P.; Ealick, S. E. Structural studies on 3-hydroxyanthranilate-3,4-dioxygenase: the catalytic mechanism of a complex oxidation involved in NAD biosynthesis. *Biochemistry* **2005,** *44*, 7632-7643.

(8) Colabroy, K. L.; Zhai, H. L.; Li, T. F.; Ge, Y.; Zhang, Y.; Liu, A. M.; Ealick, S. E.; McLafferty, F. W.; Begley, T. P. The mechanism of inactivation of 3-hydroxyanthranilate-3,4-dioxygenase by 4-chloro-3-hydroxyanthranilate. *Biochemistry* **2005,** *44*, 7623-7631.

(9) Adachi, K.; Iwabuchi, T.; Sano, H.; Harayama, S. Structure of the ring cleavage product of 1-hydroxy-2-naphthoate, an intermediate of the phenanthrene-degradative pathway of Nocardioides sp strain KP7. *J. Bacteriol.* **1999,** *181*, 757-763.

(10) Adams, M. A.; Singh, V. K.; Keller, B. O.; Jia, Z. C. Structural and biochemical characterization of gentisate 1,2-dioxygenase from Escherichia coli O157 : H7. *Mol. Microbiol.* **2006,** *61*, 1469-1484.

(11) Chen, J.; Li, W.; Wang, M. Z.; Zhu, G. Y.; Liu, D. Q.; Sun, F.; Hao, N.; Li, X. M.; Rao, Z. H.; Zhang, X. C. Crystal structure and mutagenic analysis of GDOsp, a gentisate 1,2-dioxygenase from silicibacter pomeroyi. *Protein Sci.* **2008,** *17*, 1362-1373.

(12) Harpel, M. R.; Lipscomb, J. D. Gentisate 1,2-dioxygenase from pseudomonas. substrate coordination to active-site $Fe^{2+}$ and mechanism of turnover. *J. Biol. Chem.* **1990,** *265*, 22187-22196.

(13) Jeoung, J. H.; Bommer, M.; Lin, T. Y.; Dobbek, H. Visualizing the substrate-, superoxo-, alkylperoxo-, and product-bound states at the nonheme Fe(II) site of homogentisate dioxygenase. *PNAS* **2013,** *110*, 12625-12630.

(14) Ferraroni, M.; Steimer, L.; Matera, I.; Burger, S.; Scozzafava, A.; Stolz, A.; Briganti, F. The generation of a 1-hydroxy-2-naphthoate 1,2-dioxygenase by single point mutations of salicylate

1,2-dioxygenase--rational design of mutants and the crystal structures of the A85H and W104Y variants. *J. Struct. Biol.* **2012,** *180*, 563-571.

(15) Eppinger, E.; Ferraroni, M.; Burger, S.; Steimer, L.; Peng, G.; Briganti, F.; Stolz, A. Function of different amino acid residues in the reaction mechanism of gentisate 1,2-dioxygenases deduced from the analysis of mutants of the salicylate 1,2-dioxygenase from pseudaminobacter salicylatoxidans. *Biochim. Biophys. Acta Proteins Proteomics* **2015,** *1854*, 1425-1437.

(16) Abu-Omar, M. M.; Loaiza, A.; Hontzeas, N. Reaction mechanisms of mononuclear non-heme iron oxygenases. *Chem. Rev.* **2005,** *105*, 2227-2252.

(17) Rohde, J. U.; Bukowski, M. R.; Que, L. Functional models for mononuclear nonheme iron enzymes. *Curr. Opin. Chem. Biol.* **2003,** *7*, 674-682.

(18) Solomon, E. I.; Wong, S. D.; Liu, L. V.; Decker, A.; Chow, M. S. Peroxo and oxo intermediates in mononuclear nonheme iron enzymes and related active sites. *Curr. Opin. Chem. Biol.* **2009,** *13*, 99-113.

(19) Kovaleva, E. G.; Neibergall, M. B.; Chakrabarty, S.; Lipscomb, J. D. Finding intermediates in the $O_2$ activation pathways of non-heme iron oxygenases. *Acc. Chem. Res.* **2007,** *40*, 475-483.

(20) Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R. Z.; Siegbahn, P. E. M. Quantum chemical studies of mechanisms for metalloenzymes. *Chem. Rev.* **2014,** *114*, 3601-3658.

(21) Quesne, M. G.; Borowski, T.; de Visser, S. P. Quantum mechanics/molecular mechanics modeling of enzymatic processes: caveats and breakthroughs. *Chem. Eur. J.* **2016,** *22*, 2562-2581.

(22) Bassan, A.; Borowski, T.; Siegbahn, P. E. M. Quantum chemical studies of dioxygen activation by mononuclear non-heme iron enzymes with the 2-His-1-carboxylate facial triad. *Dalton Trans.* **2004**, 3153-3162.

(23) Dong, G.; Shaik, S.; Lai, W. Z. Oxygen activation by homoprotocatechuate 2,3-dioxygenase: a QM/MM study reveals the key intermediates in the activation cycle. *Chem. Sci.* **2013,** *4*, 3624-3635.

(24) Dong, G.; Lai, W. Z. Reaction mechanism of homoprotocatechuate 2,3-dioxygenase with 4-nitrocatechol: implications for the role of substrate. *J. Phys. Chem. B* **2014,** *118*, 1791-1798.

(25) Siegbahn, P. E. M.; Haeffner, F. Mechanism for catechol ring-cleavage by non-heme iron extradiol dioxygenases. *J. Am. Chem. Soc.* **2004,** *126*, 8919-8932.

(26) Christian, G. J.; Ye, S. F.; Neese, F. Oxygen activation in extradiol catecholate dioxygenases - a density functional study. *Chem. Sci.* **2012,** *3*, 1600-1611.

(27) Kumar, D.; Thiel, W.; de Visser, S. P. Theoretical study on the mechanism of the oxygen activation process in cysteine dioxygenase enzymes. *J. Am. Chem. Soc.* **2011,** *133*, 3869-3882.

(28) Borowski, T.; Siegbahn, P. E. M. Mechanism for catechol ring cleavage by non-heme iron intradiol dioxygenases: a hybrid DFT study. *J. Am. Chem. Soc.* **2006,** *128*, 12941-12953.

(29) Georgiev, V.; Noack, H.; Borowski, T.; Blomberg, M. R. A.; Siegbahn, P. E. M. DFT study on the catalytic reactivity of a functional model complex for intradiol-cleaving dioxygenases. *J. Phys. Chem. B* **2010,** *114*, 5878-5885.

(30) Jastrzebski, R.; Quesne, M. G.; Weckhuysen, B. M.; de Visser, S. P.; Bruijnincx, P. C. A. Experimental and computational evidence for the mechanism of intradiol catechol dioxygenation by non-heme iron(III) complexes. *Chem. Eur. J.* **2014,** *20*, 15686-15691.

(31) Dong, G.; Lu, J. R.; Lai, W. Z. Insights into the mechanism of aromatic ring cleavage of noncatecholic compound 2-aminophenol by aminophenol dioxygenase: a quantum mechanics/molecular mechanics study. *Acs Catal* **2016,** *6*, 3796-3803.

(32) Borowski, T.; Georgiev, V.; Siegbahn, P. E. M. Catalytic reaction mechanism of homogentisate dioxygenase: a hybrid DFT study. *J. Am. Chem. Soc.* **2005,** *127*, 17303-17314.

(33) Qi, Y.; Lu, J.; Lai, W. Insights into the reaction mechanism of aromatic ring cleavage by homogentisate dioxygenase: a quantum mechanical/molecular mechanical Study. *J. Phys. Chem. B* **2016,** *120*, 4579-4590.

(34) de Visser, S. P. Elucidating enzyme mechanism and intrinsic chemical properties of short-lived intermediates in the catalytic cycles of cysteine dioxygenase and taurine/α-ketoglutarate dioxygenase. *Coord. Chem. Rev.* **2009,** *253*, 754-768.

(35) Borowski, T.; Bassan, A.; Siegbahn, P. E. M. Mechanism of dioxygen activation in 2-oxoglutarate-dependent enzymes: a hybrid DFT study. *Chem. Eur. J.* **2004,** *10*, 1031-1041.

(36) Borowski, T.; Bassan, A.; Siegbahn, P. E. M. 4-hydroxyphenylpyruvate dioxygenase: a hybrid density functional study of the catalytic reaction mechanism. *Biochemistry* **2004,** *43*, 12331-12342.

(37) de Visser, S. P. Can the peroxosuccinate complex in the catalytic cycle of taurine/α-ketoglutarate dioxygenase (TauD) act as an alternative oxidant? *Chem. Commun.* **2007**, 171-173.

(38) Diebold, A. R.; Brown-Marshall, C. D.; Neidig, M. L.; Brownlee, J. M.; Moran, G. R.; Solomon, E. I. Activation of α-Keto acid-dependent dioxygenases: application of an $\{FeNO\}^7/\{FeO_2\}^8$ methodology for characterizing the initial steps of $O_2$ activation. *J. Am. Chem. Soc.* **2011,** *133*, 18148-18160.

(39) Ye, S. F.; Riplinger, C.; Hansen, A.; Krebs, C.; Bollinger, J. M.; Neese, F. Electronic structure analysis of the oxygen-activation mechanism by $Fe^{II}$- and $\alpha$-ketoglutarate ($\alpha$KG)-dependent dioxygenases. *Chem. Eur. J.* **2012,** *18*, 6555-6567.

(40) Brown, C. D.; Neidig, M. L.; Neibergall, M. B.; Lipscomb, J. D.; Solomon, E. I. VTVH-MCD and DFT studies of thiolate bonding to $\{FeNO\}^7/\{FeO_2\}^8$ complexes of isopenicillin N synthase: substrate determination of oxidase versus oxygenase activity in nonheme Fe enzymes. *J. Am. Chem. Soc.* **2007,** *129*, 7427-7438.

(41) Lundberg, M.; Morokuma, K. Protein environment facilitates $O_2$ binding in non-heme iron enzyme. An insight from ONIOM calculations on isopenicillin N synthase (IPNS). *J. Phys. Chem. B* **2007,** *111*, 9380-9389.

(42) Hirao, H.; Morokuma, K. Ferric superoxide and ferric hydroxide are used in the catalytic mechanism of hydroxyethylphosphonate dioxygenase: a density functional theory investigation. *J. Am. Chem. Soc.* **2010,** *132*, 17901-17909.

(43) Milaczewska, A.; Broclawik, E.; Borowski, T. On the catalytic mechanism of (S)-2-hydroxypropylphosphonic acid epoxidase (HppE): a hybrid DFT study. *Chem. Eur. J.* **2013,** *19*, 770-780.

(44) Roy, S.; Kästner, J. Synergistic substrate and oxygen activation in salicylate dioxygenase revealed by QM/MM simulations. *Angew. Chem. Int. Edit.* **2016,** *55*, 1168-1172.

(45) Dong, G.; Ryde, U. Protonation states of intermediates in the reaction mechanism of [NiFe] hydrogenase studied by computational methods. *J. Biol. Inorg. Chem.* **2016,** *21*, 383-394.

(46) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *J. Chem. Theor. Comput.* **2011,** *7*, 525-537.

(47) Mock, W. L.; Morsch, L. A. Low barrier hydrogen bonds within salicylate mono-anions. *Tetrahedron* **2001,** *57*, 2957-2964.

(48) Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E. III; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Monard, G.; Needham, P.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Salomon-Ferrer, R.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; York, D. M.; Kollman, P. A. (2014), AMBER 14, University of California, San Francisco.

(49) Ryde, U. The coordination of the catalytic zinc ion in alcohol dehydrogenase studied by combined quantum-chemical and molecular mechanics calculations. *J. Comput. Aided Mol. Des.* **1996,** *10*, 153-164.

(50) Ryde, U.; Olsson, M. H. M. Structure, strain, and reorganization energy of blue copper models in the protein. *Int. J. Quantum Chem.* **2001,** *81*, 335-347.

(51) Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem. Int. Edit.* **2009,** *48*, 1198-1229.

(52) Ryde, U. QM/MM calculations on proteins. *Meth. Enz.* **2016,** *577*, 119-158.

(53) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. Frontier bonds in QM/MM methods: a comparison of different approaches. *J. Phys. Chem. A* **2000,** *104*, 1720-1735.

(54) Hu, L.; Söederhjelm, P.; Ryde, U. On the convergence of QM/MM energies. *J. Chem. Theor. Comput.* **2011,** *7*, 761-777.

(55) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. ONIOM: a multilayered integrated MO+MM method for geometry optimizations and single

point energy predictions. A test for Diels-Alder reactions and $Pt(P(t\text{-}Bu)_3)_2+H_2$ oxidative addition. *J. Phys. Chem.* **1996,** *100*, 19357-19363.

(56) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic-behavior. *Phys. Rev. A* **1988,** *38*, 3098-3100.

(57) Lee, C. T.; Yang, W. T.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron-density. *Phys. Rev. B* **1988,** *37*, 785-789.

(58) Becke, A. D. Density-fuctional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993,** *98*, 5648-5652.

(59) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully optimized contracted Gaussian-basis sets for atoms Li to Kr. *J. Chem. Phys.* **1992,** *97*, 2571-2577.

(60) dftd3 software http://www.thch.uni-bonn.de/tc/index.php?section=downloads&subsection=DFT-D3&lang=english.

(61) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theor. Comput.* **2015,** *11*, 3696-3713.

(62) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005,** *7*, 3297-3305.

(63) Rappoport, D.; Furche, F. Property-optimized Gaussian basis sets for molecular response calculations. *J. Chem. Phys.* **2010,** *133*, 134105.

(64) Klamt, A.; Schuurmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perk. T 2* **1993**, 799-805.

(65) Klamt, A.; Jonas, V.; Burger, T.; Lohrenz, J. C. W. Refinement and parametrization of COSMO-RS. *J. Phys. Chem. A* **1998,** *102*, 5074-5085.

(66) Schäfer, A.; Klamt, A.; Sattel, D.; Lohrenz, J. C. W.; Eckert, F. COSMO implementation in TURBOMOLE: extension of an efficient quantum chemical code towards liquid systems. *Phys. Chem. Chem. Phys.* **2000,** *2*, 2187-2193.

(67) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Aqueous solvation free energies of ions and ion-water clusters based on an accurate value for the absolute aqueous solvation free energy of the proton. *J. Phys. Chem. B* **2006,** *110*, 16066-16081.

(68) Sumner, S.; Söderhjelm, P.; Ryde, U. Effect of geometry optimizations on QM-cluster and QM/MM studies of reaction energies in proteins. *J. Chem. Theor. Comput.* **2013,** *9*, 4205-4214.

(69) Hu, L.; Söderhjelm, P.; Ryde, U. Accurate reaction energies in proteins obtained by combining QM/MM and large QM calculations. *J. Chem. Theor. Comput.* **2013,** *9*, 640-649.

(70) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011,** *32*, 1456-1465.

(71) Luzhkov, V.; Warshel, A. Microscopic models for quantum-mechanical calculations of chemical processes in solutions - LD/AMPAC and SCAAS/AMPAC calculations of solvation energies. *J. Comput. Chem.* **1992,** *13*, 199-213.

(72) Rod, T. H.; Ryde, U. Accurate QM/MM free energy calculations of enzyme reactions: methylation by catechol O-methyltransferase. *J. Chem. Theor. Comput.* **2005,** *1*, 1240-1251.

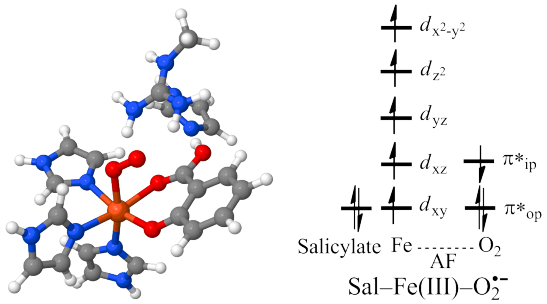
(73) Rod, T. H.; Ryde, U. Quantum mechanical free energy barrier for an enzymatic reaction. *Phys. Rev. Lett.* **2005,** *94*, 138302.

(74) Heimdal, J.; Kaukonen, M.; Srnec, M.; Rulisek, L.; Ryde, U. Reduction potentials and acidity constants of Mn superoxide dismutase calculated by QM/MM free-energy methods. *Chemphyschem* **2011,** *12*, 3337-3347.

(75) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983,** *79*, 926-935.

(76) Kovaleva, E. G.; Lipscomb, J. D. Crystal structures of $Fe^{2+}$ dioxygenase superoxo, alkylperoxo, and bound product intermediates. *Science* **2007,** *316*, 453-457.

For Table of Contents Only:




The reactive Fe–$O_2$ species in salicylate 1,2-dioxygenase is best described as a Fe(III)–$O_2^{\bullet-}$ state. It is a high spin of Fe(III) ion antiferromagnetically coupled to a superoxide anion. This species attacks the substrate to form the alkylperoxo species. His162 plays the role of an acid–base catalyst, providing a proton to the substrate.