What determines the selectivity of arginine dihydroxylation by the nonheme iron enzyme OrfP?

Document Version
Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

OPEN ACCESS

WILEY-VCH

What determines the selectivity of arginine dihydroxylation by the nonheme iron enzyme OrfP?

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

[a]   Mr H. S. Ali, Dr R. H. Henchman, Dr S. P. de Visser
        Manchester Institute of Biotechnology, The University of Manchester, 131 Princess Street, Manchester, M1 7DN, United Kingdom
        E-mail: sam.devisser@manchester.ac.uk
[b]   Mr H. S. Ali, Dr R. H. Henchman
        Department of Chemistry, The University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom
[c]   Dr S. P. de Visser
        Department of Chemical Engineering and Analytical Science, The University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom

        Supporting information for this article is given via a link at the end of the document.

Abstract: The nonheme iron enzyme OrfP reacts with L-Arg selectively to form the $3R,4R$-dihydroxyarginine product, which in mammals can inhibit the nitric oxide synthase enzymes involved in blood pressure control. To understand the mechanisms of dioxygen activation of L-Arg by OrfP and how it enables two sequential oxidation cycles on the same substrate, we performed a density functional theory study on a large active site cluster model. We show that substrate binding and positioning in the active site guides a highly selective reaction through C^3–H hydrogen atom abstraction. This happens despite the fact that the C^3–H and C^4–H bond strengths of L-Arg are very similar. Electronic differences in the two hydrogen atom abstraction pathways drive the reaction with an initial C^3–H activation to a low-energy $^5\sigma$-pathway, while substrate positioning destabilizes the C^4–H abstraction and sends it over the higher-lying $^5\pi$-pathway. We show that substrate and monohydroxylated products are strongly bound in the substrate binding pocket and hence product release is difficult and consequently its lifetime will be long enough to trigger a second oxygenation cycle.

Introduction

Natural products including antibiotics often are build up from sugar and amino acid components.[1] Nature utilizes a range of enzymes to modify amino acids to give the natural products its structure and function. These biosynthesis processes often require a high regio- and chemoselectivity for the reaction, which is difficult to achieve in chemical catalysis. Understanding of how the natural product synthesis by enzymes is achieved and how this high selectivity is obtained is important for biotechnology and could enable the biosynthesis of useful products with high regio-, enantio- and stereoselectivity and minimize waste products.

A class of metalloenzymes involved in the biosynthesis of many natural products are the nonheme iron dioxygenases, which are found in most organisms.[2] For instance, in humans the biosynthesis of the amino acid analogue $4R$-hydroxyproline is performed by proline-4-hydroxylase enzymes in an enantio- and stereoselective reaction mechanism, which is an essential component of collagen and gives it its functional and structural properties.[3] Many nonheme iron dioxygenases react via a highly selective reaction mechanism, where the substrate is tightly bound in the substrate binding pocket near the nonheme iron cofactor.

Interestingly, several nonheme iron dioxygenases activate an arginine residue as part of a natural product synthesis reaction such as an antibiotics biosynthesis reaction in bacteria.[1ab,4] Thus, VioC hydroxylates a free arginine molecule selectively at the C^3-position as a precursor to the biosynthesis of viomycin,[5] while NapI desaturates an arginine substrate at the C^3–C^4 bond through two sequential hydrogen atom abstraction reactions.[6] These enzymes were studied spectroscopically and kinetically and their substrate specificity tested. Recent studies on the nonheme iron enzyme GetI showed it to hydroxylate L-Arg at the C^4-position, although it was not clear whether arginine is its natural substrate.[7] A similar situation was found in the ethylene forming enzyme that apart from catabolizing succinate to three CO_2 molecules, also appears to bind and activate an arginine residue but hydroxylate it at the C^5-position.[8] Therefore, several nonheme iron dioxygenases appear to activate an L-Arg substrate differently to give either C^3-hydroxylation, C^4-hydroxylation, C^5-hydroxylation or C^3–C^4 desaturation through the use of O_2 and α-ketoglutarate (αKG) on an iron center.

A recently discovered and characterized nonheme iron dioxygenase from *Streptomyces*, named OrfP, was shown to activate a free arginine amino acid and produce yet another product in a reaction with dioxygen, namely the $3R,4R$-dihydroxyarginine.[9] Subsequently, the $3R,4R$-dihydroxoarginine is converted into streptolidine and attached to a glucosamine sugar scaffold to form the streptothricin product. OrfP; therefore, performs the dihydroxylation of L-Arg as a precursor reaction to the biosynthesis of the streptothricin antibiotics. Isolation of OrfP and studies of its product distributions showed that the major component was $3R,4R$-dihydroxyarginine; however, small amounts of $3R$-hydroxyarginine, $3S$-hydroxyarginine and $4S$-hydroxyarginine, are also observed (Figure 1), although their origins remain unclear. Note that $3R,4R$-dihydroxyarginine was found to have inhibitory effects on inducible nitric oxide synthase enzymes in mammals; an enzyme involved in blood pressure control and inflammation.[10] Hence, its selective biosynthesis pathways are important in medicinal and biotechnology work and may have important applications.
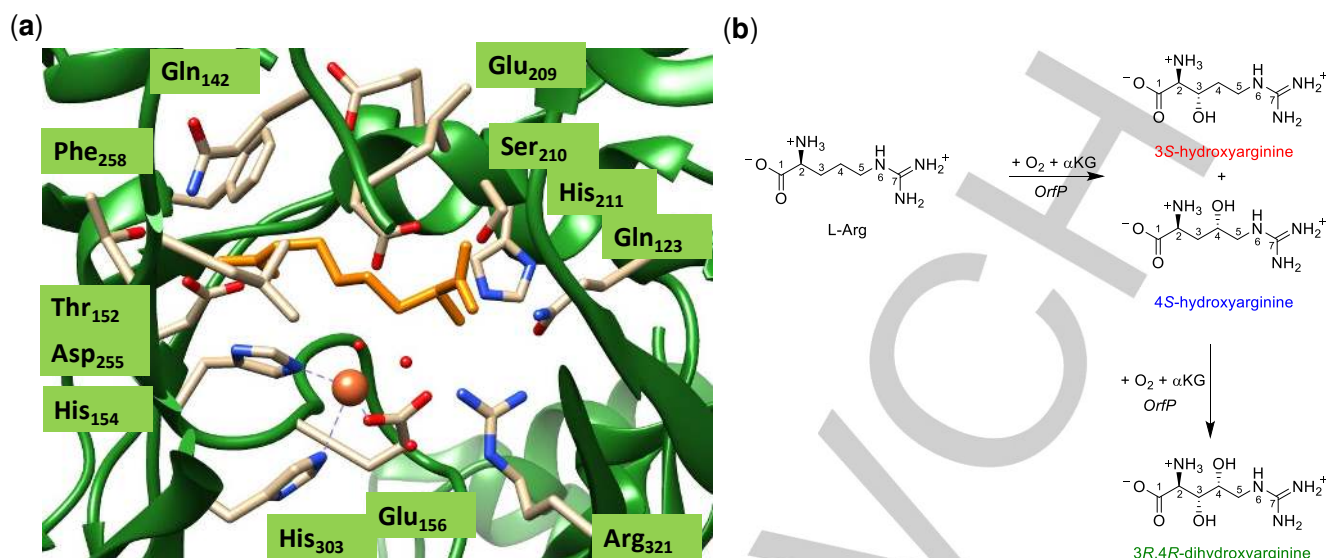
Figure 1. a) Extract of the crystal structure coordinates (4M2E pdb file) of L-homo-arginine (amber) bound OrfP with key residues and iron (light brown) highlighted; b) Products and by-products formed in OrfP activation of L-Arg.

An extract of the active site of OrfP is given in Figure 1 as taken from the 4M2E protein databank (pdb) file.[11,12] The iron is linked to the protein through interactions with the side chains of His$_{154}$, Glu$_{156}$ and His$_{303}$ group in a typical facial 2-His/1-Glu triad, which is common for many nonheme iron dioxygenases.[13] In the 4M2E pdb,[11] the other ligand sites of the metal are occupied by water molecules. It is known that OrfP utilizes α-ketoglutarate (αKG or also called 2-oxoglutarate) and dioxygen and binds these on its iron center and presumably forms an iron(IV)-oxo active species. For several other nonheme iron dioxygenases the iron(IV)-oxo species has been characterized with UV-Vis absorption, electroparamagnetic resonance and Mössbauer spectroscopic studies and shown to be the active species that reacts with substrate.[14,15]

How OrfP can perform a double dioxygenation reaction by preventing release of hydroxyarginine in favor of a second catalytic cycle is unknown. Although crystal structure coordinates of OrfP have been obtained, they do not give a clear understanding on the product release and substrate oxidation selectivity, which warrants a detailed computational study. In the OrfP pdb structure, the substrate analogue L-homo-arginine (Figure 1) binds tightly into the substrate-binding pocket and is surrounded by polar groups. In particular, its guanidinium group forms a salt bridge with Asp$_{255}$ and experiences hydrogen bonding interactions from the alcohol group of Thr$_{152}$ and the oxo group of Gln$_{142}$. The carboxylate group of L-homo-arginine forms a salt bridge with Arg$_{321}$, which is also in hydrogen bonding distance to the iron ligand Glu$_{156}$. The amine group of the substrate is positioned with hydrogen bonding and salt-bridge interactions to the side chains of the Gln$_{123}$, Asp$_{208}$, Ser$_{210}$ and His$_{211}$ residues, which will influence the reactivity of the substrate with the cofactor. Hence we reasoned that a large cluster model will be needed to describe the OrfP structure and reactivity well. In this work, a large cluster model incorporating the substrate and cofactor and their direct environment is studied to explore the reaction pathways of L-Arg activation leading to 3S-hydroxyarginine, 3R-

hydroxyarginine, 4S-hydroxyarginine and 3R,4R-dihydroxyarginine products. Moreover, we focus on how the hydroxyarginine release is prevented in favor of a second oxygenation cycle.

Results

First oxygenation cycle of L-Arg by OrfP.
We created an OrfP cluster model, see Scheme 1, from the crystal structure coordinates of the OrfP enzyme deposited in the protein databank under the 4M2E pdb file.[11,12] This structure represents an OrfP tetramer with L-homo-arginine as substrate mimic bound. We took chain A of the protein and replaced L-homo-arginine by L-arginine manually. The structure has iron coordinated to the side chains of two histidine groups (His$_{154}$ and His$_{303}$) and the carboxylate group of Glu$_{126}$. As α-ketoglutarate (αKG) is missing from the pdb file, it was docked in position with the SwissDock web server bound to iron in the same plane as the side chains of His$_{154}$ and Glu$_{156}$.[16] A 20ns molecular dynamics simulation shows the structure to be highly rigid and little changes to the active site and second-coordination sphere structure of the iron center was found. Next, a model containing the first and second coordination sphere of the metal center was created that included the oxidant with its direct ligands and the main components of the substrate (L-Arg) and succinate binding pocket: see Scheme 1.

Subsequently, we explored the first hydroxylation pathway of L-Arg substrate using the model described in Scheme 1. OrfP enzymes perform two consecutive hydroxylation reactions on L-Arg and each of these cycles uses one molecule of O$_2$ and a molecule of αKG to form an iron(IV)-oxo species and succinate. Spectroscopic studies showed this iron(IV)-oxo to react with L-Arg.[11] We will start with the substrate activation steps in the first catalytic cycle, where L-Arg is bound to an iron(IV)-oxo(succinate) complex and follow the reaction mechanism of substrate hydroxylation leading to 3S-hydroxyarginine, 3R-hydroxyarginine

and 4S-hydroxyarginine products using the cluster model described in Scheme 1 (Model **A**). Thereafter, the second hydroxylation cycle was investigated from these mono-hydroxylated arginine product complexes by replacing the iron(II) site by an iron(IV)-oxo species (Model **B**). Model **A** and **B**; therefore, have the same overall charge. The model type is given in subscript after the label of the structure.
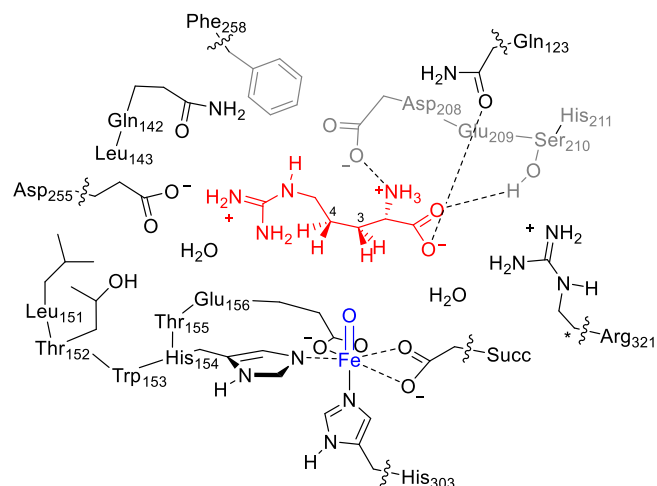

Scheme 1. DFT cluster model investigated in this work with substrate highlighted in red and iron(IV)-oxo in blue. The wiggly lines identify where the protein chain was cut and a hydrogen atom inserted.

Our cluster model of the iron(IV)-oxo species with L-Arg bound was geometry optimized in the low-lying triplet and quintet spin states (^{3,5}**Re**$_A$) using UB3LYP/BS1, while the quintet spin structure was also minimized with a larger basis set at UB3LYP/BS3 (^{5}**Re**$_{A,BS3}$). Figure 2 gives optimized geometries of the reactant complexes ^{3,5}**Re**$_A$ as obtained with DFT using cluster model **A**. The Fe–O distance in the quintet spin state is short and typical of a double bond, while it is somewhat larger in the triplet spin state. The axial histidine ligand is at a distance of 2.078 Å in the quintet spin state and at 2.035 Å in the triplet spin state, while a geometry optimization with a larger basis set, i.e. BS3 basis set, elongates it slightly to 2.157 Å. These structures of the nonheme iron(IV)-oxo species match previous calculations on similar species well.[17] Moreover, experimental studies on analogous nonheme iron enzymes, such as TauD and P4H established an Fe–O distance of about 1.62 Å and characterized them as quintet spin ground states.[14,15] Therefore, the calculations match previously reported experimental structures of enzymatic nonheme iron(IV)-oxo complexes well. Furthermore, an overlay of the ^{5}**Re**$_A$ structure with the crystal structure coordinates shows that little changes to the structure have occurred during the geometry optimization and the model still has all features of the protein (Supporting Information, Figure S4).

The nearest hydrogen atom from the oxo group is the pro-S hydrogen atom at the C^3-position and its distance is about 2.667 Å in the quintet spin state, while the pro-R hydrogen atom is located at 4.061 Å and the pro-S hydrogen atom on the C^4-position at 5.008 Å. Based on the substrate positioning in OrfP; therefore, it appears that the C^3-position is located closest to the metal center and hence should be the preferred site of activation.

The quintet spin state is the ground state and is well separated from the triplet spin state by ΔE+ZPE = 7 kcal mol^{-1}. As such, the triplet spin state plays little role in the reaction mechanism and the reaction takes place through single-state-reactivity of the quintet spin state surface only. This is as expected for an iron(IV)-oxo species in trigonal bipyramidal configuration that usually gives a high-spin ground state.[18] The quintet spin state is characterized with a molecular orbital occupation of $\pi^*_{xy}{}^1 \pi^*_{xz}{}^1 \pi^*_{yz}{}^1 \sigma^*_{x2-y2}{}^1$, while in the triplet spin state it is $\pi^*_{xy}{}^2 \pi^*_{xz}{}^1 \pi^*_{yz}{}^1$. The spin-state-ordering, spin-state energies and orbital configurations match previous work on nonheme iron enzymes and biomimetic models well and show that OrfP has the usual features of the first-coordination sphere.[17,19]
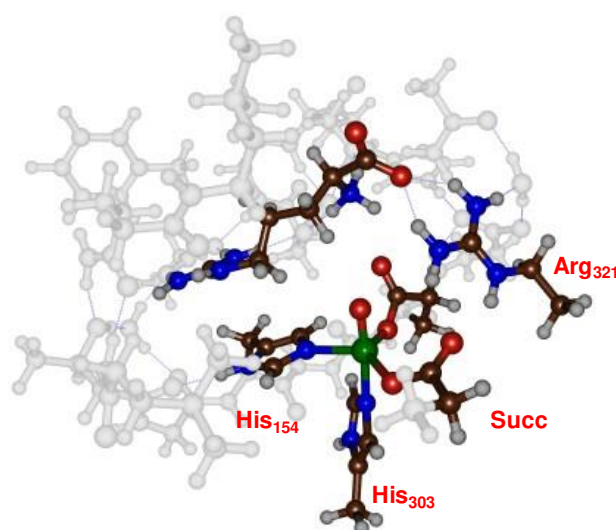
C^{3}H$_{proS}$-O: 2.667 (2.376) [2.532]
C^{3}H$_{proR}$-O: 4.061 (4.125) [4.017]
C^{4}H$_{proR}$-O: 5.008 (3.110) [4.831]

Fe-O: 1.650 (1.616) [1.756]
Fe-N(His$_{303}$): 2.078 (2.157) [2.035]

ΔE+ZPE = 0.0 [7.0]
ΔG = 0.0 [7.9]

^{5}**Re**$_A$ (^{5}**Re**$_{A,BS3}$) [^{3}**Re**$_A$]

Figure 2. UB3LYP/BS1 optimized geometries of ^{5,3}**Re**$_A$ with bond lengths in angstroms and the relative (free) energies in kcal mol^{-1}. UB3LYP/BS3 optimized geometry ^{5}**Re**$_{A,BS3}$ data are given in parenthesis.

Next we considered substrate hydroxylation by the reactant complexes ^{3,5}**Re**$_A$ at the C^3 and C^4 positions to produce 3S-hydroxyarginine, 3R-hydroxyarginine and 4S-hydroxyarginine. The overall reaction scheme that was investigated with the definition of the various intermediate and transition state structures is given in Scheme 2. To this end we calculated the hydrogen atom abstraction from the two different hydrogen atoms on the C^3-position, designated the C3R (or pro-R) and C3S (or pro-S) hydrogen atoms, and the nearest hydrogen from the C^4 group, i.e. the C4S hydrogen atom. Transition states (**TS1**$_{HA}$) for all positions were located and lead to a radical intermediate (**IM1**) representing an iron(III)-hydroxo species with a nearby substrate radical on either the C^3 or C^4 position of the substrate.
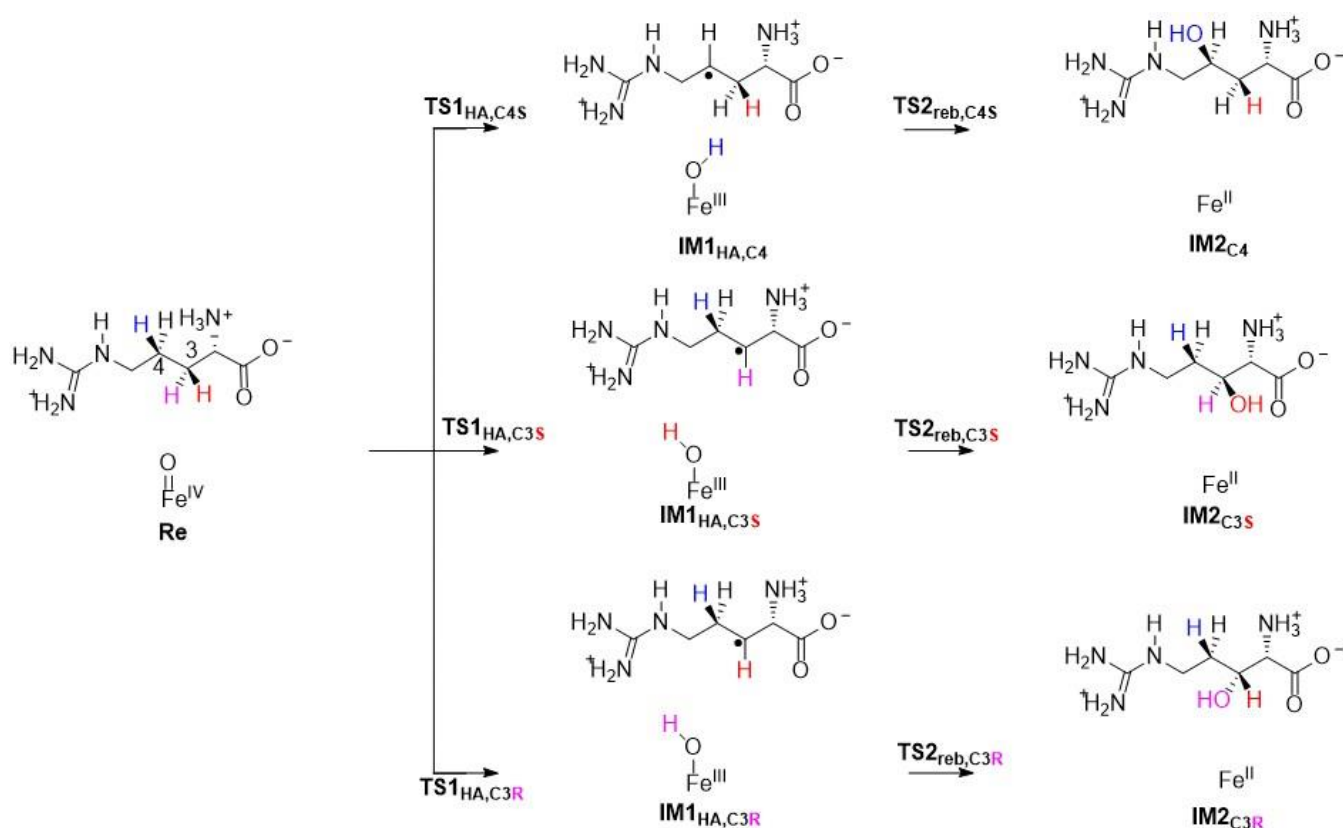
Scheme 2. Reaction mechanisms of arginine hydroxylation by the iron(IV)-oxo species of OrfP as studied in this work.

A subsequent OH rebound to the substrate (via transition state **TS2**$_{reb}$) leads to the mono-hydroxylated products (**IM2**) that are bound to an iron(II) species. The C–H bond that is activated in each step in the first substrate hydroxylation mechanism is identified with C3R, C3S and C4S in subscript after the label of the structure for the pathways leading to 3R-hydroxyarginine, 3S-hydroxyarginine and 4S-hydroxyarginine.

The reaction mechanism as described in Scheme 2 was calculated for the large cluster model **A** of OrfP for substrate hydroxylation leading to 3S-hydroxyarginine, 3R-hydroxyarginine and 4S-hydroxyarginine products. The potential energy landscape with key optimized geometries is shown in Figure 3. As can be seen from Figure 3 the lowest enthalpy of activation ($\Delta E^{\ddagger}$+ZPE) for hydrogen atom abstraction is from the pro-S C³–H position via ⁵**TS1**$_{HA,C3S}$ with a value of $\Delta E^{\ddagger}$+ZPE = 12.1 kcal mol⁻¹. Close in enthalpy of activation is the pro-R C³–H hydrogen atom abstraction barrier via ⁵**TS1**$_{HA,C3R}$ at $\Delta E^{\ddagger}$+ZPE = 15.5 kcal mol⁻¹. The hydrogen atom abstraction barrier from the C⁴–H position is much higher in energy than the one from the C³–H positions, i.e. $\Delta E^{\ddagger}$+ZPE = 23.2 kcal mol⁻¹. Based on these hydrogen atom abstraction barriers, a selective hydrogen atom abstraction from the C³–H position is predicted with the pro-S channel dominant, although small amounts of pro-R cannot be excluded. The calculations therefore match experimental observation[9] that singly hydroxylated arginine at the C³-position is formed. We also calculated the ⁵**TS1**$_{HA,C3S}$ and ⁵**TS1**$_{HA,C4S}$ structures at UB3LYP/BS3 level of theory and find $\Delta E^{\ddagger}$+ZPE energies at

UB3LYP-D3/BS2//UB3LYP/BS3 of 7.9 and 23.4 kcal mol⁻¹, respectively. As such these barriers predict the same trends as seen with UB3LYP/BS1 optimized structures and hence we continued with this method only.

Interestingly, when entropy and thermal corrections are added to the enthalpy, the pro-R hydrogen atom abstraction barrier becomes the lowest energy pathway: $\Delta G^{\ddagger}$ = 11.6 kcal mol⁻¹ for the pro-R C³–H pathway, while for the pro-S C³–H pathway $\Delta G^{\ddagger}$ = 13.5 kcal mol⁻¹ is found. Consequently, both hydrogen atoms on the C³ atom of L-arginine can be abstracted by the iron(IV)-oxo species and this should give a mixture of 3R-hydroxyarginine and 3S-hydroxyarginine products. These products have indeed been observed experimentally.[9] To understand why entropy and thermal corrections reverse the ordering of ⁵**TS1**$_{HA,C3S}$ and ⁵**TS1**$_{HA,C3R}$, we analyzed the structures in more detail. The transition state geometries (⁵**TS1**$_{HA,C3S}$, ⁵**TS1**$_{HA,C3R}$ and ⁵**TS1**$_{HA,C4S}$) are shown in Figure 3. The two C³–H hydrogen atom abstraction barriers are relatively central with similar C³–H and O–H distances. In particular, the ⁵**TS1**$_{HA,C3R}$ has C³–H and O–H distances of 1.265 and 1.272 Å, respectively, while those distances are 1.302 and 1.233 Å for ⁵**TS1**$_{HA,C3S}$. By contrast, the C⁴–H hydrogen atom abstraction barrier is more product-like with considerably longer C⁴–H distance than O–H distance: 1.358 versus 1.187 Å. The interesting difference between the three structures relates to the Fe–O–C angle in the transition states. Thus, ⁵**TS1**$_{HA,C3S}$ has a large Fe–O–C³ angle of 167°, whereas in the ⁵**TS1**$_{HA,C3R}$ structure it is 139° and in ⁵**TS1**$_{HA,C4S}$ the angle is 127°.
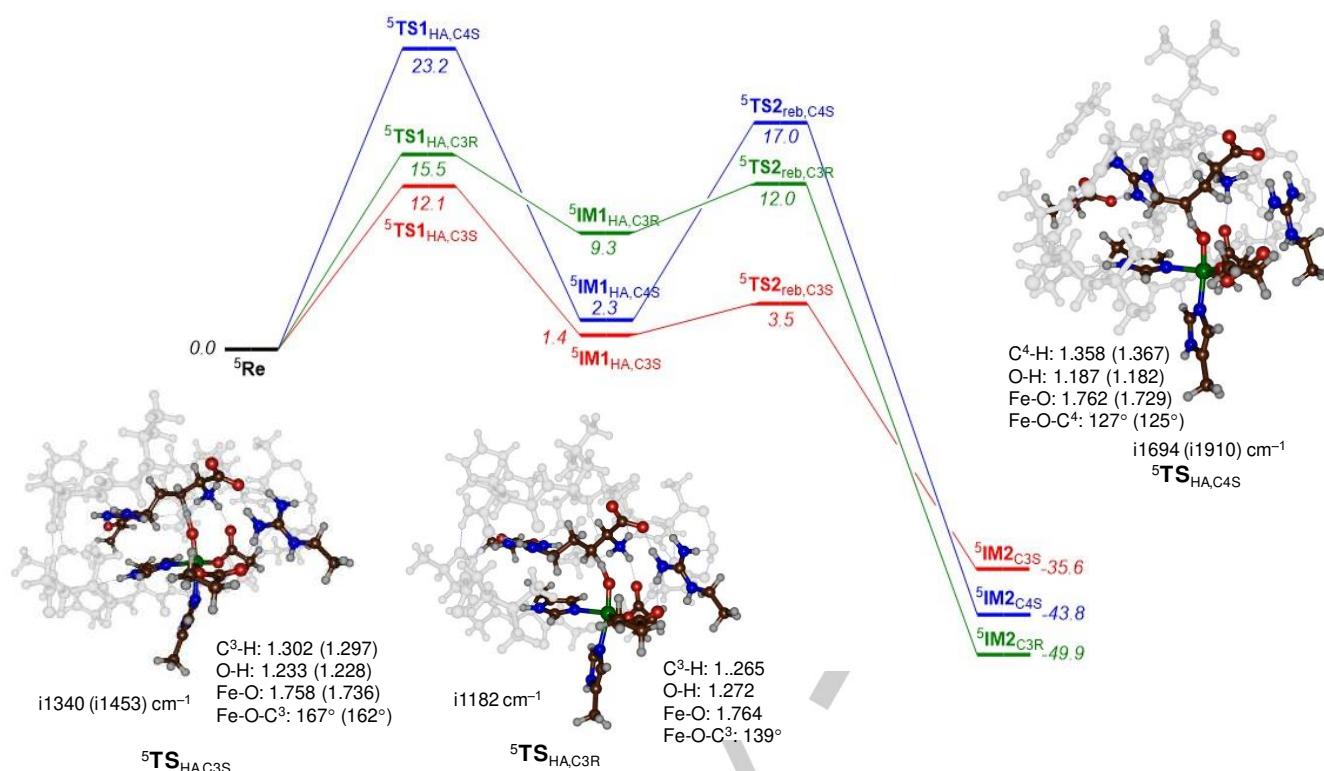
Figure 3. Potential energy landscape for L-arginine hydroxylation at the C^3 and C^4 positions with energies (in kcal mol^{-1}) relative to ^{5}**Re**$_A$ calculated at UB3LYP-D3/BS2//UB3LYP/BS1. Optimized geometries of the transition states are given with bond lengths in angstroms, the bond angle in degrees and the imaginary frequency in cm^{-1}. UB3LYP/BS3 optimized structures are given in parenthesis.

We also optimized the ^{5}**TS1**$_{HA,C3S}$ and ^{5}**TS1**$_{HA,C4S}$ structures with UB3LYP/BS3; however, very little changes in the optimized geometries with respect to UB3LYP/BS1 are seen. For both structures the C–H and O–H distances are within 0.01Å and only the Fe–O is shortened by up to 0.04Å. As such the basis set has little influence on the optimized structures and BS1 and BS3 give qualitative similar results.

Typically in hydrogen atom abstraction transition states by nonheme iron(IV)-oxo species the quintet spin state pathways shows approach of the substrate from the top and an almost linear Fe–O–C bond angle.[20] Despite this large difference in oxidant approach between ^{5}**TS1**$_{HA,C3S}$ and ^{5}**TS1**$_{HA,C3R}$ based on the Fe–O–C^3 angle; actually an overlay of the two optimized geometries does not show major differences (Supporting Information, Figure S5). However, there are major differences in vibrational entropy between the two structures, whereby ^{5}**TS1**$_{HA,C3S}$ has a 12.1 cal mol^{-1} K^{-1} larger vibrational entropy than ^{5}**TS1**$_{HA,C3R}$. Moreover, when all small vibrations with magnitude smaller than 50 cm^{-1} are removed from the equation, the entropy difference is reversed and ^{5}**TS1**$_{HA,C3R}$ has a larger vibrational entropy by 16.0 cal mol^{-1} K^{-1}. This corresponds to a free energy stabilization of 4.8 kcal mol^{-1} and reverses the individual barrier heights. The large vibrational entropy contribution of ^{5}**TS1**$_{HA,C3R}$ is unrealistic and probably the result of a gas-phase model, where the vibrational contributions are overestimated. Thus, previously we did a comparative study on experimental and computational enthalpy and free energies of activation for oxygen atom transfer

reactions and found the entropy to be overestimated by as much as 50% in model.[21] Moreover, the enthalpy values predict the experimental product distributions correctly and appear to be more realistic. As such, we will focus on ΔE+ZPE values only.

To test whether replacement of the transferring hydrogen atom by deuterium would have an effect on the barrier heights, we re-evaluated the vibrational frequencies by replacing the pro-R and pro-S hydrogen atom on the C^3-position by deuterium. Thus, with a deuterium atom on the pro-R position and a hydrogen atom at the pro-S position the ^{5}**TS1**$_{HA,C3R}$ and ^{5}**TS1**$_{HA,C3S}$ barriers change to ΔG‡ = 12.8 and 13.5 kcal mol^{-1}, respectively. Consequently, the order of the hydrogen atom abstraction barriers does not change, but the energy gap narrows to within 1 kcal mol^{-1}. The reverse situation with a deuterium atom on the pro-S position and a hydrogen atom at the pro-R position gives free energies of activation of ΔG‡ = 11.7 and 14.7 kcal mol^{-1} and widens the energy gap.

As the hydrogen atom abstraction step is rate-determining for the first hydroxylation cycle of L-Arg by OrfP, we calculated the kinetic isotope effects (KIEs) using the Eyring and Wigner methods. Both C3R and C3S pathways give a KIE$_{Eyring}$ of about 8, which rises to 14 – 15 when quantum mechanical tunneling corrections are added. These are typical values of hydrogen atom abstraction kinetic isotope effects that are commonly seen in hydrogen atom abstraction reactions by nonheme iron(IV)-oxo complexes in enzymatic and biomimetic model complexes.[22]
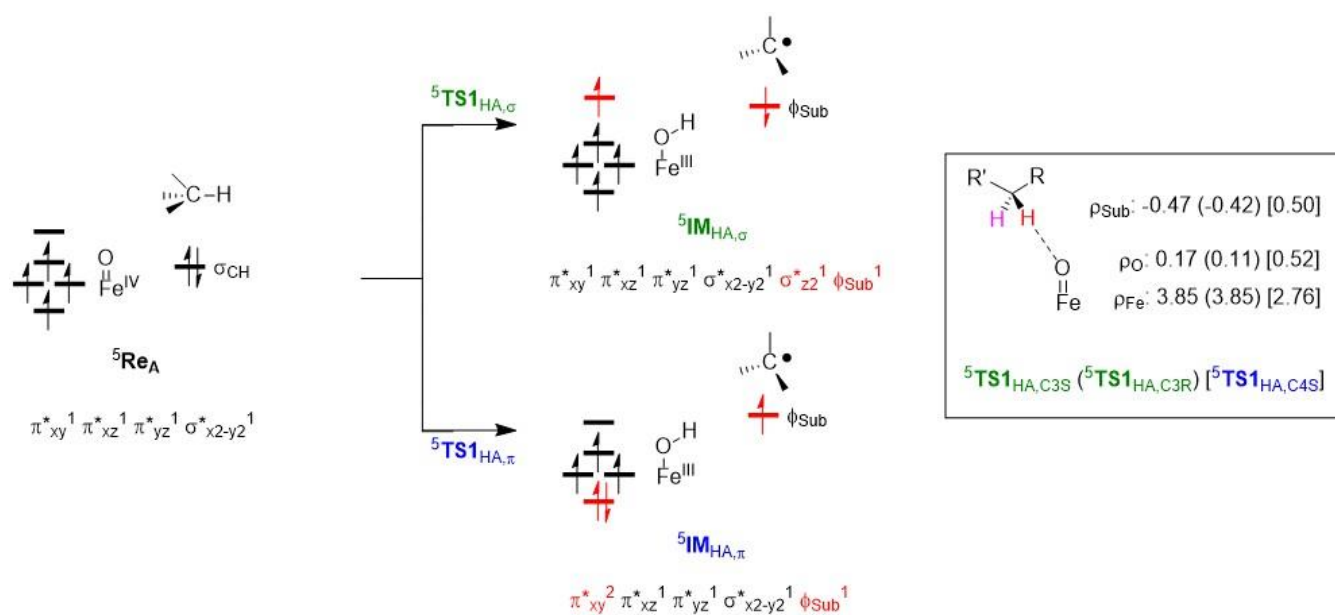
Figure 4. Electron transfer processes for hydrogen atom abstraction and OH rebound via $^5\sigma$ (top) and $^5\pi$ (bottom) pathways and group spin densities (ρ) as obtained from the optimized transition state geometries.

After the hydrogen atom abstraction transition states all structures relax to a radical intermediate **IM1**; however, the two structures with a radical on C^3 (^{5}**IM1**$_{C3S}$ and ^{5}**IM1**$_{C3R}$) are very wide apart with ^{5}**IM1**$_{C3S}$ lower in energy than ^{5}**IM1**$_{C3R}$ by 7.9 kcal mol^{-1}. By contrast, the ^{5}**IM1**$_{C4S}$ is within 1.1 kcal mol^{-1} of ^{5}**IM1**$_{C3S}$. Therefore, the ordering of the radical intermediates ^{5}**IM1** are different from those of the transition states and probably is the result of the tight substrate binding and positioning that affect the kinetics dramatically.

Next, an OH rebound step takes place from the **IM1** intermediates. The pathway from ^{5}**IM1**$_{C3}$ gives small rebound transition states of less than 3 kcal mol^{-1} to give alcohol products with large exothermicity. This small rebound barrier will imply that ^{5}**IM1**$_{C3}$ has a short lifetime and collapse to products without rearrangement and stereochemical scrambling of products. On the other hand, the rebound barrier ^{5}**TS2**$_{reb,C4S}$ is 14.7 kcal mol^{-1} above ^{5}**IM1**$_{C4S}$ and hence the radical intermediate ^{5}**IM1**$_{C4S}$ will have a finite lifetime. Interestingly, the alcohol product complexes (**IM2**) have different ordering than the **IM1** states, whereby the most stable structure is ^{5}**IM2**$_{C3R}$. Overall, the DFT modelling on cycle 1 of OrfP shows that a mixture of 3S-hydroxyarginine and 3R-hydroxyarginine may be expected as those pathways have competing reaction barriers and rate constants.

To understand the different substrate to oxidant angles (Fe−O−C angle), we show the possible electron transfer pathways in the hydrogen atom abstraction and OH rebound steps in Figure 4. As discussed above the reactant has $\pi^*_{xy}{}^1$ $\pi^*_{xz}{}^1$ $\pi^*_{yz}{}^1$ $\sigma^*_{x2-y2}{}^1$ configuration, while the substrate C−H bond is occupied with two electrons. Upon hydrogen atom transfer, the C−H bond cleaves homolytically and a radical is left on the substrate in orbital ϕ_{Sub}. The electron of the hydrogen atom moves into the metal 3d-block and generally there are two possibilities called the $^5\sigma$ and $^5\pi$ pathways.[23] In the $^5\sigma$ pathway an electron transfer from the substrate into the virtual σ^*_{z2} orbital takes place to give a radical intermediate ^{5}**IM1**$_{HA,\sigma}$ with configuration $\pi^*_{xy}{}^1$ $\pi^*_{xz}{}^1$ $\pi^*_{yz}{}^1$ $\sigma^*_{x2-y2}{}^1$

$\sigma^*_{z2}{}^1$ $\phi_{Sub}{}^1$. As the σ^*_{z2} orbital is located along the Fe−O axis the substrate will approach from the top and an almost linear Fe−O−C angle is obtained in the transition state. In the radical intermediate the five metal-type orbitals are antiferromagnetically coupled to a radical on the substrate and are exchange stabilized. The alternative mechanism for substrate hydroxylation is via the $^5\pi$-pathway, where the initial electron transfer from the substrate into the singly occupied π^*_{xy} orbital takes place. This then gives an intermediate with electronic configuration $\pi^*_{xy}{}^2$ $\pi^*_{xz}{}^1$ $\pi^*_{yz}{}^1$ $\sigma^*_{x2-y2}{}^1$ $\phi_{Sub}{}^1$, where now all unpaired electrons are ferromagnetically coupled. The radical rebound pathways brings the $^5\sigma$ and $^5\pi$ pathways together into the same alcohol product complex ^{5}**IM2** with configuration $\pi^*_{xy}{}^2$ $\pi^*_{xz}{}^1$ $\pi^*_{yz}{}^1$ $\sigma^*_{x2-y2}{}^1$ $\sigma^*_{z2}{}^1$. Electron transfer into the π^*_{xy} orbital gives a side-on approach and a more bent structure (typically around 120°) is found. Based on the electron-transfer processes these pathways are called the $^5\sigma$ and $^5\pi$-pathways. To find out, whether our hydrogen atom abstraction pathways belong to the $^5\sigma$ or $^5\pi$-type, we analyzed the molecular orbitals and unpaired spin density of the three hydrogen atom abstraction transition states ^{5}**TS1**$_{HA}$.

Group spin densities of ^{5}**TS1**$_{HA,C3S}$, ^{5}**TS1**$_{HA,C3R}$ and ^{5}**TS1**$_{HA,C4S}$ transition states are given in Figure 4. The spin density for the C^3−H hydrogen atom abstraction is large on the iron ($\rho_{Fe} = 3.85$ for both), while negative spin density is accumulating on the substrate. Consequently, these spin densities characterize both C^3−H hydrogen atom abstraction barriers as $^5\sigma$ pathway structures. By contrast, the ^{5}**TS1**$_{HA,C4S}$ transition state has positive spin density on the substrate ($\rho_{Sub} = 0.50$), while the spin on iron is only $\rho_{Fe} = 2.76$. Therefore, the ^{5}**TS1**$_{HA,C4S}$ transition state is of the $^5\pi$-pathway rather than of $^5\sigma$.

In an attempt to find the $^5\sigma$-pathway structures, some molecular orbitals of the $^5\pi$-pathway transition states were swapped; however, during the SCF convergence this electronic configuration was not stable and relaxed to the $^5\pi$ electronic configuration and geometry instead.
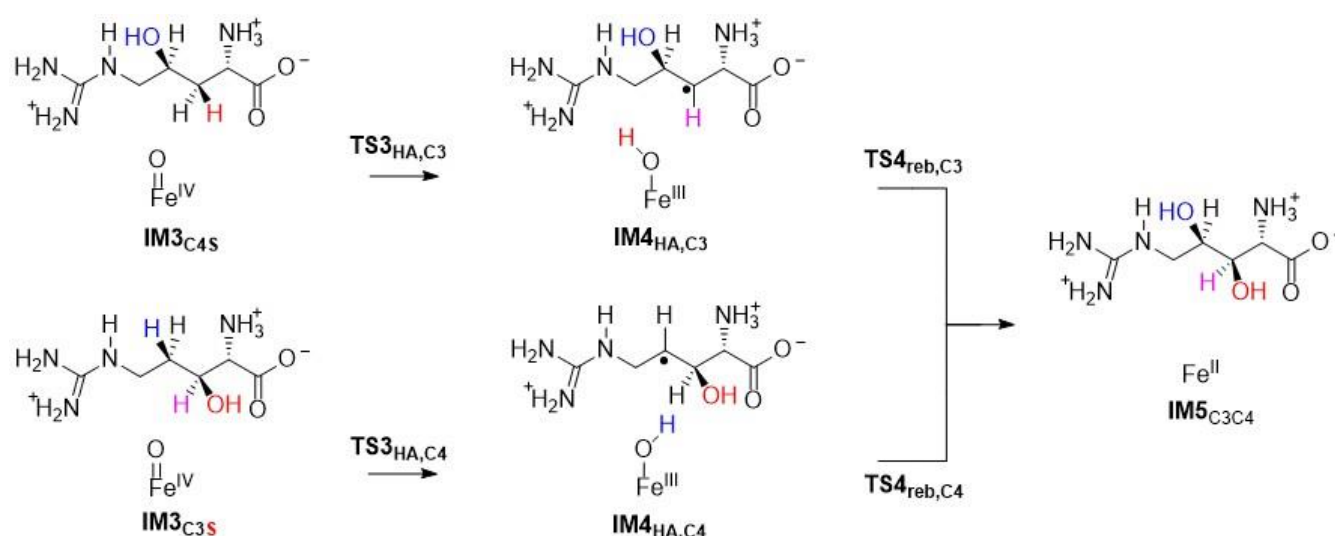
Scheme 3. Reaction mechanisms investigated in this work for hydroxyarginine hydroxylation by an iron(IV)-oxo species of OrfP.

Therefore, the C^3–H and C^4–H hydrogen atom abstraction pathways give differences in electronic configuration and electron transfer processes that is forced upon the system through substrate binding and positioning. These electronic and stereochemical effects push the reaction to C^3-hydroxylation of L-Arg selectively.

Second oxygenation cycle of L-Arg by OrfP.
In the next set of calculations we explored the second hydroxylation cycle of arginine. Experimental studies show that OrfP enzymes are able to incorporate two hydroxyl groups into an arginine molecule.[1f,9,11] We hypothesized that after the first hydroxylation step of the substrate is completed, the hydroxyarginine is not released from the enzyme but only succinate. Subsequently, another catalytic cycle starts with binding of a new molecule of αKG and O_2 to the iron center that then react to form a second iron(IV)-oxo and succinate species. To this end, we took the optimized geometry of the 3S-hydroxyarginine, 3R-hydroxyarginine and 4S-hydroxyarginine bound iron(II) complexes, i.e. ^{5}**IM2**$_{C3S}$, ^{5}**IM2**$_{C3R}$ and ^{5}**IM2**$_{C4S}$, and replaced the iron(II) group by iron(IV)-oxo to form ^{5}**IM3**$_{C3S}$, ^{5}**IM3**$_{C3R}$ and ^{5}**IM3**$_{C4S}$, which kept the overall charge the same. Thereafter, the hydroxylation of the pro-R C^4–H group of 3S-hydroxyarginine in ^{5}**IM3**$_{C3S}$ and the hydroxylation of the pro-R C^3–H group in 4S-hydroxyarginine was studied to form the 3R,4R-dihydroxyarginine product **IM5**$_{C3C4}$ (Scheme 3). Note that the dihydroxylation changes the stereochemistry on atom C^3 from 3S in the mono-hydroxylated species to 3R in the dihydroxylated arginine. The second hydroxylation step starts with a hydrogen atom abstraction via ^{5}**TS3**$_{HA}$ to form a radical intermediate ^{5}**IM4**$_{HA}$. A radical rebound via transition state ^{5}**TS4**$_{reb}$ gives the dihydroxylated arginine product complex ^{5}**IM5**$_{C3C4}$. Figure 5 shows the potential energy landscape for the second hydroxylation cycle starting from these ^{5}**IM3**$_B$. In addition to these two pathways we also investigated the hydrogen atom abstraction from the C^3–H group in 3R-hydroxyarginine.

The iron(IV)-oxo species for the second reaction cycle for 3S-hydroxyarginine and 3R-hydroxyarginine bound complexes (^{5}**IM3**$_{B,C3S}$ and ^{5}**IM3**$_{B,C3R}$) are shown in Figure 5. They have similar features as the reactant species of cycle 1 with an Fe–O distance of 1.656 and 1.650 Å, respectively, and the same electronic configuration of $\pi^*_{xy}{}^1 \pi^*_{xz}{}^1 \pi^*_{yz}{}^1 \sigma^*_{x2-y2}{}^1$. The only difference is a hydrogen bonding interaction between the alcohol group of 3S-hydroxyarginine and the oxo group, which may have elongated the Fe–O bond slightly. Nevertheless, the C^4–H group is positioned close to the oxo group at a distance of 3.159 Å ^{5}**IM3**$_{B,C3S}$. In contrast, in ^{5}**IM3**$_{B,C3R}$ the C^3–H group is much further away and positioned at a distance of 5.019 Å. Therefore, we decided to explore hydrogen atom abstraction from the C^3–H position instead: the C^3H–O distance is 3.985 Å in ^{5}**IM3**$_{B,C3R}$.

Thereafter, the hydrogen atom abstraction by the iron(IV)-oxo species ^{5}**IM3**$_{B,C3S}$ from the C^4–H position was calculated and a barrier of $\Delta E^{\ddagger}$+ZPE = 11.6 kcal mol^{-1} is found. This is much lower in energy than the hydrogen atom abstraction from the C^4–H position obtained for L-arginine, i.e. 23.2 kcal mol^{-1} see Figure 3. An analysis of the group spin densities of ^{5}**TS3**$_{HA,C4}$ gives a spin of 3.93 on iron and a spin of −0.42 on the C^4 atom of the substrate and an orbital occupation of $\pi^*_{xy}{}^1 \pi^*_{xz}{}^1 \pi^*_{yz}{}^1 \sigma^*_{x2-y2}{}^1 \sigma^*_{z2}{}^1 \phi_{Sub}{}^1$. Consequently, ^{5}**TS3**$_{HA,C4}$ has a $^5\sigma$ electronic configuration, while ^{5}**TS1**$_{HA,C4}$ had a $^5\pi$ configuration. These differences are probably the result of the hydrogen bonding interaction from the hydroxo group of substrate to the oxo group that constraints the substrate approach and withdraws electron density. Indeed in previous work we showed that hydrogen bonding interactions to an iron(IV)-oxo species influence reaction barriers.[18] Geometrically, the transition state ^{5}**TS3**$_{HA,C4}$ is relatively central with almost equal C–H and O–H distances, i.e. 1.245 Å (C^4–H bond) and 1.296 Å (O–H bond). The C^4–O–Fe angle, however, is somewhat bent (140°) while for the first hydrogen atom abstraction transition states (**TS1**$_{HA}$ above) have an angle close to 180°. Consequently, there is considerable constraint in the structure probably due to hydrogen bonding interactions in the substrate binding pocket.
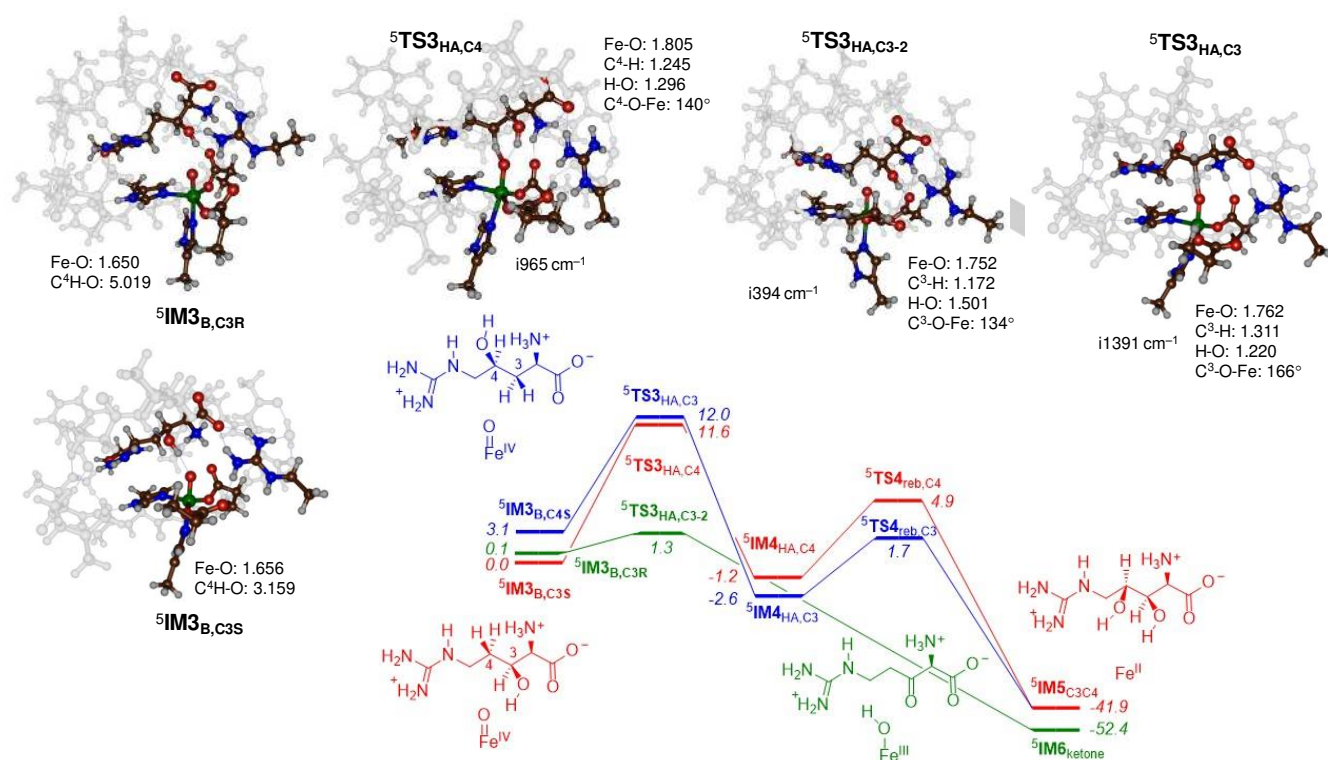
Figure 5. Potential energy landscape for 3-hydroxyarginine and 4-hydroxyarginine hydroxylation to form the dihydroxoarginine and β-ketoarginine products. Energies calculated at UB3LYP-D3/BS2//UB3LYP/BS1. Outside parenthesis are ΔE+ZPE values, while free energies are in parenthesis. Optimized geometries of the transition states are given with bond lengths in angstroms, angles in degrees and the imaginary frequency in cm^{-1}.

After the transition state, the system relaxes to a radical intermediate via an almost thermoneutral process (ΔE+ZPE = −1.2 kcal mol^{-1} with respect to ^{5}**IM3**$_{B,C3S}$). A subsequent rebound barrier of 6.1 kcal mol^{-1} above ^{5}**IM4**$_{HA,C4}$ leads to the dihydroxylated arginine products ^{5}**IM5**$_{C3C4}$ with large exothermicity. The alternative pathway starting from ^{5}**IM3**$_{B,C4S}$ to form ^{5}**IM5**$_{C3C4}$ was also studied. Thus, a transition state of ΔE‡+ZPE = 12.0 kcal mol^{-1} leads to a radical intermediate that is 2.6 kcal mol^{-1} more stable than ^{5}**IM3**$_{B,C3S}$. A small OH rebound barrier of 4.3 kcal mol^{-1} above ^{5}**IM4**$_{HA,C3}$ then leads to the 3R,4R-dihydroxyarginine product. The transition state geometry of ^{5}**TS3**$_{HA,C3}$ is given in Figure 5 and has the transferring the hydrogen atom almost midway between the donor and acceptor groups: the C^3–H distance is 1.311 Å and the H–O distance is 1.220 Å. The C^3–O–Fe angle is 166° and therefore much larger than the corresponding angle in ^{5}**TS3**$_{HA,C4}$ of 140°.

These reaction barriers are low in energy and of the same order of magnitude as the first hydrogen atom abstraction from the pro-S position of the C^3–H group in Figure 3 above. Therefore, the dihydroxylated product can be formed from either 3S-hydroxyarginine or 4S-hydroxyarginine with similar rate constants and reaction barriers. However, as shown in Figure 3 above, the first hydroxylation step is highly selective and will give dominant 3S-hydroxyarginine through the first oxygen atom transfer cycle. Finally, we explored activation of 3R-hydroxyarginine in a second reaction cycle. As the C^4–H group is located far away, due to the positioning of 3R-hydroxyarginine in the substrate binding pocket we instead studied the C^3–H abstraction pathway via transition

state ^{5}**TS3**$_{HA,C3-2}$. A barrier of only 1.3 kcal mol^{-1} with respect to ^{5}**IM3**$_{B,C3S}$ is found, which is not surprising as a weak tertiary C–H bond is broken. This transition state is early with a short C^3–H distance of 1.172 Å and a long O–H distance of 1.501 Å. The barrier has a small imaginary frequency of i394 cm^{-1} with a dominant C^3–H–O stretch vibration. However, the transition states also show movement for the OH group of the substrate in the direction of the carboxylate group of Glu$_{156}$. Indeed, after the hydrogen atom abstraction barrier the system does not relax to a radical intermediate but a rapid second hydrogen atom transfer takes place from the substrate OH group to the Glu$_{156}$ group, which desaturates the C^3–O bond of substrate and form the β-keto-arginine product (^{5}**IM6**$_{ketone}$). This product cannot be formed from 3S-hydroxyarginine and 4S-hydroxyarginine as the tertiary C–H bond points away from the iron-oxo group and hence these groups are not accessible for the iron(IV)-oxo species. It is evident from our calculations, therefore, that 3R-hydroxyarginine most likely leads to desaturation of the substrate through a very efficient and low-energy reaction process in the first reaction cycle and consequently, substrate binding and positioning should block this potential pathway. As such, substrate positioning is crucial in OrfP enzymes and the substrate binding pocket is evolved to maximize the yield of products and minimize the amount of by-products. In particular, the tertiary C–H bonds of the mono-hydroxylated arginine should point away from the metal center so that no desaturation pathways become accessible.
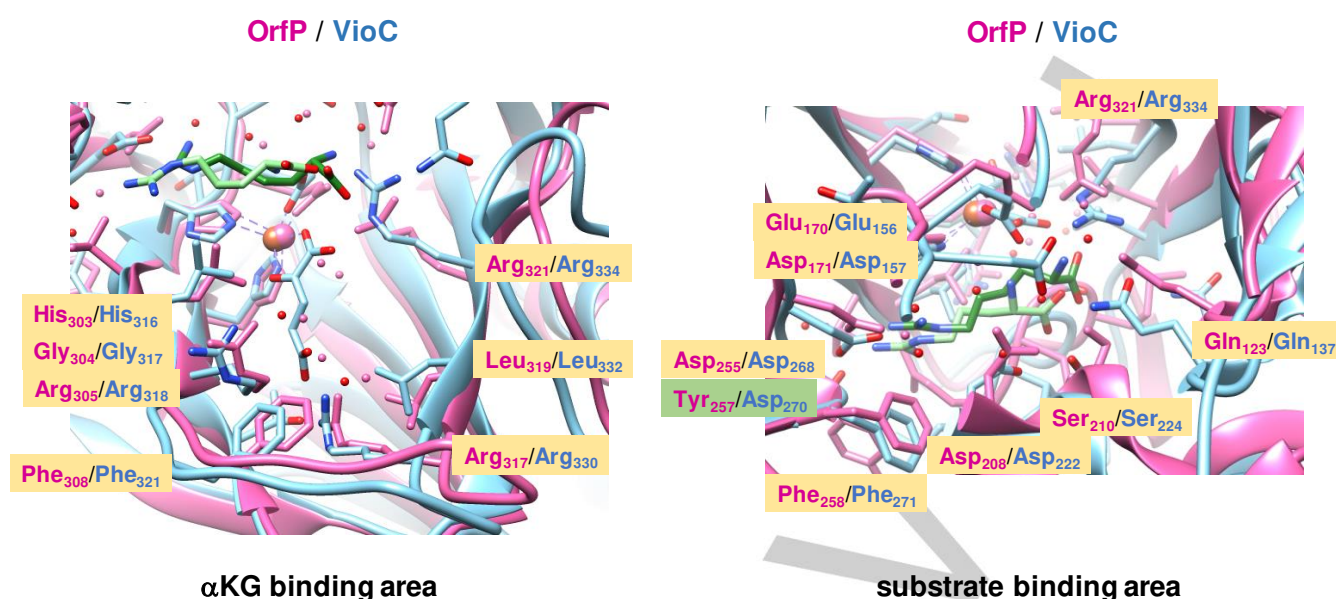
Figure 6. Overlay of the pdb files of OrfP[9] and VioC[5d] with the focus on the αKG binding pocket (left) and L-Arg binding pocket (right).

Discussion

To understand the details of the dihydroxylation mechanism of L-Arg by OrfP enzymes, we analyzed the thermochemical properties of the oxidant and substrate in more detail and made a comparison with analogous enzymes. We compared the structures of two arginine activating nonheme iron dioxygenases, namely OrfP and VioC. These two nonheme iron dioxygenases react differently with L-arginine as a substrate, whereby VioC selectively hydroxylates it at the C^3-position and OrfP performs the dihydroxylation to form $3R,4R$-dihydroxyarginine. A structural comparison between the OrfP and VioC crystal structure coordinates (4M2E versus 6ALM pdb files)[5d,9] is given in Figure 6 as an overlay of the active site regions of the two enzymes. Both enzymes utilize αKG, dioxygen and L-Arg on a nonheme iron center that is bound to the protein through a 2-His/1-Glu facial triad coordination. The overlay on the left focuses on the αKG binding area, while the one on the right zooms into the L-Arg binding area. As can be seen from Figure 6, the two enzymes have an almost identical αKG binding loop that starts from the axial histidine ligand (His$_{303}$ in OrfP and His$_{316}$ in VioC) and the peptide chain continues with HGRXXFQXRYDGXDRWLKR. Hence in this loop of 19 amino acids only four residues (labeled as X) do not match between the two proteins and all of those amino acids point away from the αKG binding pocket. Clearly, the αKG binding loop is highly conserved and may show a similar amino acid chain in most αKG dependent nonheme iron dioxygenases. Indeed, it has been reported that a HX$_{13}$R loop from the axial ligand reaches a conserved Arg residue that binds and positions αKG in the substrate binding pocket.[24] The overlay on the left-hand-side of Figure 6 indeed puts the residues of this loop in virtually the same position in both enzymes. In OrfP the Arg$_{317}$ forms a salt bridge with αKG, while the analogous residue

Arg$_{330}$ has that role in VioC. Interestingly, the next Arg in this chain (Arg$_{321}$ in OrfP/Arg$_{334}$ in VioC) locks the carboxylate group of the substrate in a salt bridge.

Another conserved region between the OrfP and VioC structures relates to the His and Glu iron ligands (His$_{168}$ and Glu$_{170}$ for OrfP and His$_{154}$ and Glu$_{156}$ for VioC) with a peptide region WHTEDAFXPY, whereby the Asp residue (Asp$_{171}$ for OrfP and Asp$_{157}$ for VioC) forms a hydrogen bonding interaction with the NH$_3^+$ group of L-Arg. By contrast the guanidinium group of L-Arg forms a salt bridge with Asp$_{255}$ (OrfP) and Asp$_{268}$ (VioC). In addition, the substrate binding pocket is aligned with the conserved residues Asp$_{208}$/Asp$_{222}$, Ser$_{210}$/Ser$_{224}$ and Phe$_{258}$/Phe$_{271}$, for OrfP and VioC, respectively. The only difference seen from the overlay of the structures appears to be the position of a substrate binding pocket Gln residue (Gln$_{123}$ in OrfP and Gln$_{137}$ in VioC) that has shifted inside in OrfP. In addition, the replacement of Tyr$_{257}$ (OrfP) with Asp in VioC on the edge of the substrate binding pocket is observed. The latter is highlighted in green in Figure 6 and implicates that VioC has an additional hydrogen bonding interaction to the substrate guanidium group that is missing in OrfP. The overlay of the pdb structures of VioC and OrfP shown in Figure 6; therefore, implies that their active site features, and αKG and substrate binding environments are highly alike. As such, the structures do not give a clear reason why different reaction products are obtained in the two reaction mechanisms. The analysis and comparison of the crystal structure coordinates of OrfP and VioC; however, implicates that both enzymes will preferentially activate L-Arg on the C^3-position as that group appears to be closest to the metal center. Indeed, our lowest energy reaction pathways for the first hydroxylation cycle in OrfP give the lowest barriers for hydrogen atom abstraction from the C^3–H position leading to $3S$-hydroxyarginine products. Consequently, both VioC and OrfP are C^3-activating nonheme iron enzymes of L-Arg substrate due to careful substrate positioning in the substrate binding pocket.

To gain further insight into the differences of the L-Arg activation mechanisms of VioC and OrfP, we created an active site cluster model of VioC with 3S-hydroxyarginine bound and optimized its geometry at UB3LYP/BS1: $^5\text{IM2}_{\text{C3S,VioC}}$. An overlay of the $^5\text{IM2}_{\text{C3S,VioC}}$ structure with the OrfP optimized geometry of 3S-hydroxyarginine, i.e. $^5\text{IM2}_{\text{C3S}}$, is shown in Figure 7. As can be seen from Figure 7, most of the protein residues are in a similar position between the two structures. In particular, the first coordination sphere ligands, i.e. the 2-His/1-Glu coordinated ligands and the succinate are in virtually the same position. Clearly, the differences in reactivity are not the result of differences in the coordination environment of the metal and hence come from the second coordination sphere and product release.
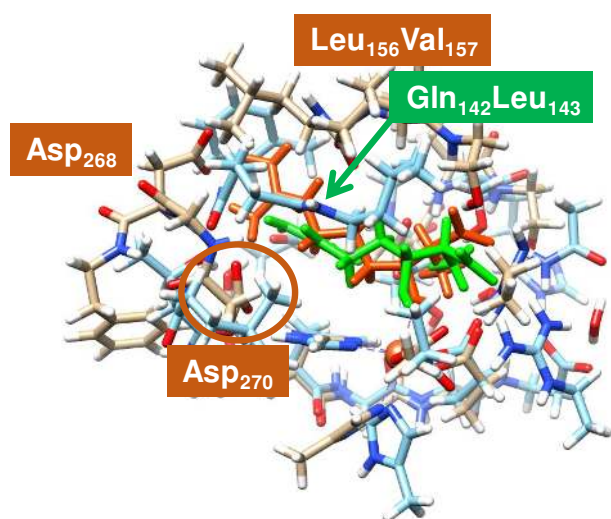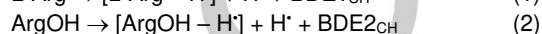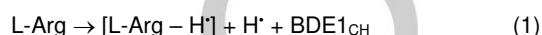

Figure 7. Overlay of UB3LYP/BS1 optimized geometries of 3S-hydroxyarginine bound iron(II) complexes of OrfP and VioC. Protein backbone of VioC in grey with 3S-hydroxyarginine in orange, while the protein backbone for OrfP is in light blue with the 3S-hydroxyarginine in green. Key differences between the two structures are highlighted in orange for VioC and green for OrfP.

As discussed in Figure 6 above, VioC has an extra carboxylic acid group in the substrate binding pocket, namely Asp_{270}, where OrfP has a Tyr residue. The optimized geometry of 3S-hydroxyarginine in the VioC model as a result is positioned differently and the guanidinium group is twisted with respect the OrfP structure. This points the hydroxo group more towards the iron atom and may prevent further dioxygen binding to the iron center. In contrast, in OrfP the guanidinium group of 3S-hydroxyarginine only forms a salt bridge with Asp_{255}, which points the hydroxyarginine tail down, while the other terminus is pointed slightly up. This means the hydroxyl group is slightly further from the metal center and a gap has appeared where O_2 can be inserted to trigger a new catalytic cycle. The tight substrate positioning in VioC may prevent this. In addition to the extra carboxylic acid group in VioC, the row of amino acids Leu_{156}-Val_{157} is located slightly higher in the substrate binding pocket than the corresponding residues (Gln_{142}-Leu_{143}) in OrfP. Therefore, the binding pocket in OrfP is tighter and smaller in OrfP than it is in VioC and consequently product release will be slowed down. These seemingly small differences between the two enzymes determine the product release mechanism and enable a second oxidation cycle in OrfP, which does not happen in VioC.

Next, we did a thermochemical analysis of the substrate, the intermediates and products and focused on the C–H bond strengths of the various aliphatic positions of L-Arg and the hydroxyarginine isomers in the structure in the protein. To this end, we calculated the bond dissociation energy (BDE) of various C–H bonds of L-Arg by calculating an isolated L-Arg molecule, a H-atom and the substrate with one hydrogen atom removed, Eq. 1.[25] The energy difference between these three structures is then the BDE1_{CH} for that particular position. In addition, the C–H bond dissociation energies of several C–H bonds in the hydroxyarginine structures from the geometries of the **IM2** intermediates were calculated (Eq 2): BDE2_{CH}.

$$\text{L-Arg} \rightarrow [\text{L-Arg} - \text{H}^\bullet] + \text{H}^\bullet + \text{BDE1}_{\text{CH}} \qquad (1)$$
$$\text{ArgOH} \rightarrow [\text{ArgOH} - \text{H}^\bullet] + \text{H}^\bullet + \text{BDE2}_{\text{CH}} \qquad (2)$$

The BDE1_{CH} and BDE2_{CH} for L-Arg, 3S-hydroxyarginine, 4S-hydroxylarginine and 3R-hydroxyarginine were calculated for various C–H bonds at the UB3LYP/6-311++G** level of theory with solvent corrections included and are summarized in Figure 8. As can be seen the highlighted three C–H bond strengths of L-Arg are within 3.3 kcal mol^{-1} from each other, whereby the $\text{BDE1}_{\text{C4S-H}}$ and $\text{BDE1}_{\text{C3R-H}}$ are the lowest in energy at 94.0/94.1 kcal mol^{-1}, while the $\text{BDE1}_{\text{C3S-H}}$ is slightly higher in energy at 97.3 kcal mol^{-1}, respectively.

Therefore, if the aliphatic C–H abstraction reaction from L-Arg by the enzyme is governed by the C–H bond strength of the substrate only, the reaction should proceed with dominant hydrogen atom abstraction from the pro-R C^3 and pro-S C^4 position of the substrate with minor amounts of pro-S C^3 hydroxylation. The DFT calculations shown above in Figure 3 on the enzymatic model, in contrast, show that the lowest barrier is obtained for C^3–H hydrogen atom abstraction from the pro-S C^3 position. Therefore, the 3S-hydroxyarginine is predicted to be the dominant product in the first reaction cycle based on the DFT cluster calculations even though the pro-S C^3–H bond is not the weakest C–H bond in the substrate. Thus, under ideal substrate approach, i.e. without substrate perturbation from the protein, the C^4-hydroxylation should be the dominant product. Clearly, substrate positioning and the tightness of the substrate binding pocket, guides the reaction to the pro-S C^3–H bond selectively.

The C^4–H pathway, by contrast, appears to be higher in energy in the calculations on the enzymatic system as a result of access to a higher energy potential energy landscape with $^5\pi$ configuration. In addition, electrostatic interactions from the protein may destabilize its pathway. These wide differences in hydrogen atom abstraction barriers must result from the substrate binding and positioning that affects the accessibility of the substrate by the oxidant.

Technically, the first hydrogen atom abstraction step in Figure 3 should correlate with the energy to break the C–H bond in the substrate minus the energy to form the O–H bond in the iron(III)-hydroxo intermediate.[25] Previously, for a small cluster model complex we calculated a $\text{BDE}_{\text{FeO-H}}$ value of 93.0 kcal mol^{-1}.[26] Based on the difference between BDE1_{CH} and $\text{BDE}_{\text{FeO-H}}$, we would predict a reaction enthalpy from reactants to **IM1$_{\text{HA}}$** of 4.3, 1.1 and 1.0 kcal mol^{-1} for the pro-S C^3, pro-S C^4 and pro-R C^3 pathways, respectively. These values are close to the **IM1$_{\text{HA,C3S}}$** and **IM1$_{\text{HA,C4S}}$** energies and shows that the optimized structures of the radical intermediates have limited disruption through the protein environment.
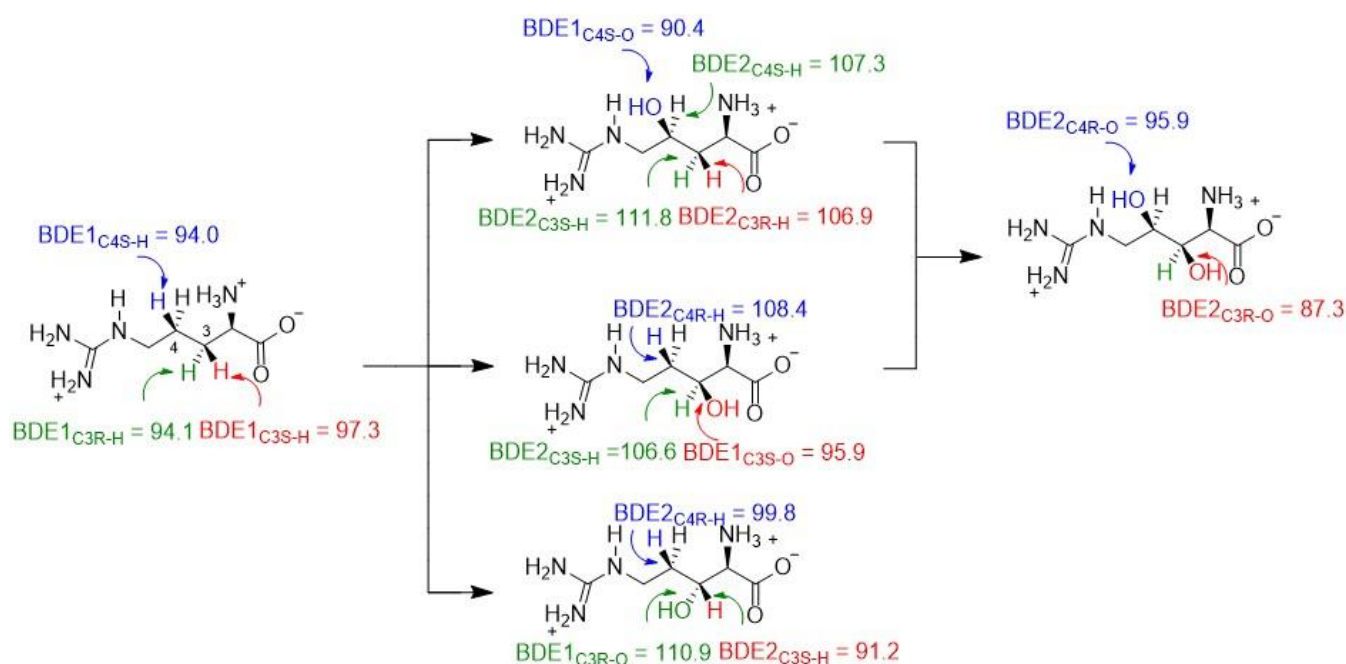
Figure 8. UB3LYP/6-311++G**+ZPE calculated C–H bond dissociation energies for various bonds in L-Arg. Values in kcal mol^{-1}.

However, the kinetics is strongly affected by the shape and size of the protein pocket and hence the transition states (**TS1**$_{HA}$) follow a different ordering than the radical intermediates (**IM1**$_{HA}$) and cover a wider energy range.

We also calculated the strength of various possible C–OH bonds that are formed, designated BDE1$_{C-O}$, via a similar procedure as the BDE1$_{CH}$ values, but where we keep the geometry of the hydroxyarginine as in the **IM2** structures. The three C–O bond strengths in 3*S*-hydroxyarginine, 3*R*-hydroxyarginine and 4*S*-hydroxyarginine are 95.9, 110.9 and 90.4 kcal mol^{-1}, respectively. Consequently, based on the thermodynamics of the reaction there are many C–H bonds in the substrate of almost equal strength. Under ideal conditions, where the substrate approach is unperturbed; therefore, a mixture of products will be formed. However, substrate positioning reduces the number of reaction products and guides the reaction to C^3–H activation selectively.

In a final set of calculations, we explored the C–H bond strengths of various C–H bonds in the hydroxyarginine complexes, where we keep the hydroxyarginine as in the geometry of the **IM2** structures. These BDE2$_{CH}$ values are given in Figure 8 as well. The C^3–H BDE2$_{CH}$ values in 4*S*-hydroxyarginine are 111.8 (pro-*S*) and 106.9 (pro-*R*) kcal mol^{-1}, whereas the tertiary C–H bond strength at the C^4-position is 107.3 kcal mol^{-1}. Thus, the weakest C–H bonds for 4*S*-hydroxyarginine are the pro-*R* C^3 and C^4 C–H bonds. As such it is important that the C^4–H bond points away from the reaction center as its abstraction will not lead to 3*R*,4*R*-dihydroxyarginine products. Indeed, our optimized geometry shows this bond to point upwards.

For 3*S*-hydroxyarginine we calculated C–H bond energies for the pro-*R* C^4 position of 108.4 kcal mol^{-1}, while the tertiary C^3–H bond has a strength of BDE2$_{CH}$ = 106.6 kcal mol^{-1}. Also for 3*S*-

hydroxyarginine it is crucial that the C^3–H bond points away from the reaction center and also in this case our optimized geometry shows it to point away from the iron atom. Finally, the 3*R*-hydroxyarginine structure (bottom structure in Figure 8) has a very weak tertiary pro-*S* C^3–H bond of BDE2$_{C3S-H}$ = 91.2 kcal mol^{-1} and a well stronger pro-*R* C^4–H bond of BDE2$_{C4R-H}$ = 99.8 kcal mol^{-1}. Clearly, the weakest bond will lead to the 3,3,-dihydroxylated product or give desaturation to form β-ketoarginine and hence substrate positioning needs to avoid bringing this hydrogen atom close to the active site. Consequently, positioning of 3*S*-hydroxyarginine and 4*S*-hydroxyarginine in the OrfP binding pocket gives orientations where the weak tertiary C–H bond is oriented away from the reaction center and it will be unlikely to be activated, while that is not possible for 3*R*-hydroxyarginine bound. Finally, we estimated the binding free energy (BFE) of 3*S*-hydroxyarginine, 3*R*-hydroxyarginine and 4*S*-hydroxyarginine in the substrate binding pockets of OrfP and VioC. To this end we took the optimized geometries ^{5}**IM2**$_{C3S}$, ^{5}**IM2**$_{C3R}$ and ^{5}**IM2**$_{C4S}$ and split the system into hydroxyarginine and protein and then did a single point frequency calculation of each of the fragments to estimate the binding free energy of the hydroxyarginine in the binding pocket. The largest binding free energy (BFE) for OrfP is found for 4*S*-hydroxyarginine with a value of 101 kcal mol^{-1}, while it is 96 kcal mol^{-1} for 3*R*-hydroxyarginine and 94 kcal mol^{-1} for 3S-hydroxyarginine. These binding free energies implicate that 4*S*-hydroxyarginine will be the strongest bound and it will be difficult to release it from the substrate binding pocket, whereas 3*S*- and 3*R*-hydroxyarginine form a slightly weaker interaction with the protein pocket. Therefore, the binding free energies of the singly hydroxylated products do not give an explanation for the dihydroxylation process in OrfP.
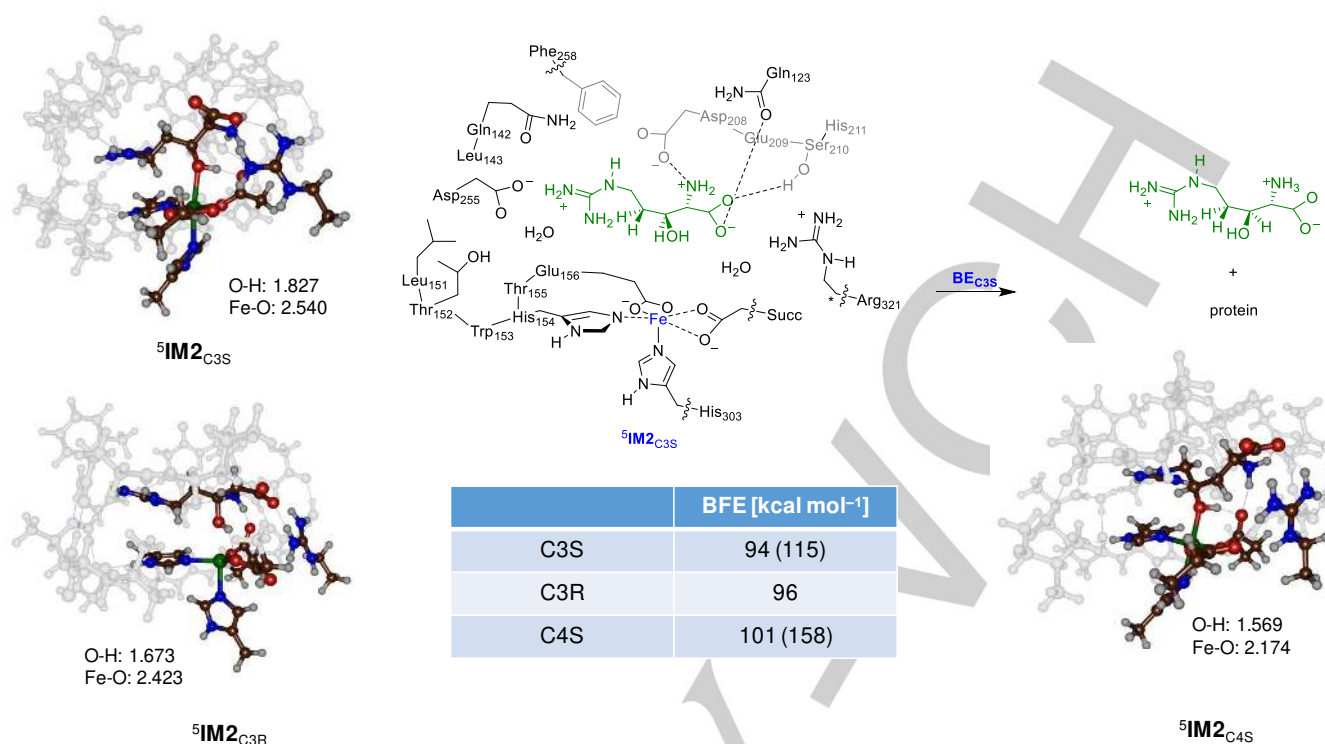
Figure 9. UB3LYP/BS1 optimized geometries of singly hydroxylated Arg complexes **IM2** with bond lengths in angstroms. Also given are calculated binding free energies (BFE) of singly hydroxylated Arg in the binding pocket of OrfP (VioL) in kcal mol⁻¹.

Optimized geometries of the ^{5}**IM2**$_{C3S}$, ^{5}**IM2**$_{C3R}$ and ^{5}**IM2**$_{C4S}$ structures are shown in Figure 9. In both ^{5}**IM2**$_{C3S}$ and ^{5}**IM2**$_{C3R}$ the Fe–O bond between the metal and the product is long (>2.4Å), which is much longer than that typically seen for a covalent bond and hence will be a weak intermolecular interaction. The structure with hydroxyarginine the strongest bound, i.e. ^{5}**IM2**$_{C4S}$, has a very short Fe–O interaction of only 2.174 Å. It also has the shortest O–H distance of the hydroxyl group of hydroxyarginine with a neighboring oxygen atom donor: in this case a distance of 1.569 Å to the carboxylate of Glu$_{156}$ is found. Thanks to these short distances the 4*S*-hydroxyarginine will bind stronger than 3*S*- and 3*R*-hydroxyarginine.

To understand substrate binding better we also calculated a VioC model with 3*S*-hydroxyarginine and 4*S*-hydroxyarginine bound. The corresponding BFE values are 115 and 158 kcal mol⁻¹, respectively. These values implicate the hydroxyarginine binds stronger in the VioC structure than in the OrfP structure, which would contradict the product release seen in those enzymes. Probably, the larger binding energy for VioC with respect to OrfP is due to the extra carboxylic acid group in the binding area, i.e. Asp$_{270}$ that forms a strong link with the substrate. However, the size of the empty substrate binding pocket we measure from the **IM2** optimized structures in OrfP are 1161 and 1187 Å^3 for ^{5}**IM2**$_{C3S}$ and ^{5}**IM2**$_{C4S}$, respectively, while the corresponding values for the VioC structures are 1322 and 1291 Å^3. Therefore, the product binding pocket in VioC is larger and gives the substrate and product more flexibility and mobility enables the hydroxyarginine to break free and escape, while the tight substrate/product binding pocket in OrfP locks the hydroxyarginine in and prevents it from escaping. It appears, therefore, that a single additional amino acid

in VioC, namely Asp$_{270}$, can pull the hydroxyarginine product away from the iron center and enable its release from the substrate binding pocket. This product-release mechanism appears to be missing in OrfP. It would be interesting to see whether the Tyr257Asp or Tyr257Glu mutations in OrfP would indeed enable release of monohydroxylated arginine from the substrate binding pocket and affect the reactivity.

Conclusion

The work presented here represents a computational study into the dihydroxylating nonheme iron dioxygenase OrfP. Using large active site cluster models we calculated the mechanism on all low-lying spin states. We show that the reaction proceeds by two consecutive hydroxylation reactions by an iron(IV)-oxo species. The first cycle has a rate-determining hydrogen atom abstraction from the C3S-position of substrate and is followed by a small rebound barrier to give 3*S*-hydroxyarginine with a small preference over 3*R*-hydroxyarginine, while the 4*S*-hydroxyarginine pathway is well higher in energy. The second cycle then binds αKG and oxygen to form another iron(IV)-oxo species. Also the second cycle has a rate-determining hydrogen atom abstraction step with similar barriers for the pathways starting from 3*S*-hydroxyarginine and 4*S*-hydroxyarginine to form the dihydroxylated product. Interestingly, the calculations show that 3*R*-hydroxyarginine in a second cycle would – via a small reaction barrier – be converted into β-ketoarginine through a desaturation step rather than lead to hydroxylation. Overall the calculations reveal that the reaction happens through negative

catalysis, where a low energy pathway, i.e. the breaking of the C^4–H bond, is avoided in favor of the breaking of a stronger bond. This selectivity is the result of substrate positioning in a very tight binding pocket that guides substrate and oxidant to the C^3–H group for substrate hydroxylation in cycle 1. Our calculations highlight the function of an active site Asp residue as a hinge to lift the monohydroxylated product out of the binding pocket of VioC, which is missing in OrfP.

Experimental Section

Model Set-Up.

A density functional theory (DFT) cluster model was created using procedures described previously.[27] The 4M2E pdb file was used,[11,12] whereby we selected chain A and replaced L-homo-Arg substrate manually by L-Arg and docked αKG in the structure using the SwissDock web server.[16] Subsequently, hydrogen atoms were added in Chimera at pH = 7,[28] and we made sure to check all polar residues: All Asp and Glu side-chains were deprotonated and all Arg and Lys side-chains protonated. The His residues in the protein were visually inspected for neighbouring hydrogen bonding donor and acceptor groups and based on the analysis all were chosen to be singly protonated on either the N_δ or N_ϵ atom. In the active site model we replaced the iron(II) ion by an iron(IV)-oxo species with a starting distance of 1.62 Å and the oxo trans to His303, while αKG was replaced by succinate to create the active oxidant species in the enzyme. These changes kept the overall charge of the model the same, which was thereafter solvated with water in Chimera. A 20 ns molecular dynamics simulation was run on this structure (Figure S1 and S2, Supporting Information). The MD simulation gave very little changes to the protein structure and kept all features of the active site in tact during the full run. As such, OrfP is a rigid protein, where the active site and substrate binding pocket show little changes over time.

Next, we analysed the first and second coordination sphere of the iron and substrate environment in detail and included in the model all groups that determine the shape and constraints of the substrate and cofactor binding environments and particularly included key hydrogen bonding interactions of charged residues and salt-bridges. Thus, the cluster model includes two long protein chains that circumvent the cofactor and substrate-binding pocket, namely the peptide chains Leu151-Thr152-Trp153-His154-Thr155-Glu156 and Asp208-Glu209-Ser210-His211. Amino acid residues pointing away from the active site where shortened to Gly residues in the model, i.e. Trp153, Thr155 and Glu209, but their protein backbone was kept in the model. The side chains of the residues Gln142-Leu143, Asp255, Phe258, and Arg321 were included in the model. A hydrogen atom was added to the positions where a bond was cut to restore the valencies of the atoms. The axial histidine ligand (His303) of iron was shortened to methylimidazole. Finally, based on the solvated structure we decided to include two water molecules in the model: one positioned near the carboxylate group of the substrate and the other near the guanidinium group. Overall, our DFT cluster model consists of 275 atoms, has overall charge of –1 and was calculated with multiplicity singlet, triplet and quintet spin. As the model contains many internal hydrogen bonding interactions, no constraints on the system were used. An overlay of the optimized reactant structure and the original pdb file (Supporting Information Figure S4) indeed shows little differences on the overall shape and structure of the chemical system.

Procedures.

The Gaussian-09 software package was used for all quantum chemical calculations.[29] Following previous experience with cluster models of nonheme iron dioxygenases,[30] we utilized the unrestricted B3LYP density functional method for geometry optimizations, constraint geometry scans and frequency calculations. The basis set used was a LANL2DZ (with electron core potential, ECP) on the iron and 6-31G on the rest of the atoms (H, C, N, O) designated basis set BS1.[31,32] Test calculations with alternative density functional methods were performed on analogous systems and showed little change in spin-state-ordering, optimized geometries and overall reaction mechanism;[33] hence B3LYP/BS1 was used for geometry optimizations. All local minima were verified by the presence or absence of negative eigenvalues in the vibrational frequency analysis while all the transition state structures were found using the Berny algorithm,[34] and confirmed by vibrational analysis to have one imaginary mode, which was animated and shown to correspond to the reaction coordinate. In order to correct the energetics and account for the effect of solvent, single-point energies of the optimized geometries were evaluated with the UB3LYP-D3/BS2 level of theory,[31,32,35] whereby BS2 is 6-311+G* on H, C, O, N and LACV3P+ (with electron core potential) on iron. The latter set of calculations included a conductor-like polarizable continuum model (CPCM) with a dielectric constant of ε = 5.7 mimicking chlorobenzene,[36] which has been shown to be a good representative of an enzyme active site.[27a] For several structures the geometries were reoptimized at the UB3LYP/BS3 level of theory (basis set BS3 has LACV3P+ with core potential on iron and 6-31G* on the other atoms), but similar geometries and reaction barriers were obtained. Therefore, most of the work was done using the UB3LYP/BS2//UB3LYP/BS1 approach.

Since the zero-point energies (ZPE) and vibrational entropy contributions are sometimes affected by small real frequencies for internal motions, we excluded vibrations with values below 50 cm⁻¹ and recalculated the ZPE, thermal and entropy corrections. Free energies are calculated at 298.15 K and 1 atm, and include the thermal corrections evaluated from the unscaled vibrational frequencies at the UB3LYP/BS1 level of theory, with solvent, dispersion and entropy contributions.

Kinetic isotope effects (KIEs) were estimated using the classical Eyring equation Eq. 3 based on differences in free energy of activation ($\Delta G^\ddagger$) of the hydrogen and deuterium-substituted systems as described previously.[37] Tunneling corrections (Q_t) to the KIE were estimated using the Wigner correction as described in Eq. 4 and is based on the change in imaginary frequency of the transition state (ν), see Eqs. 4 and 5. In Eqs. 3 – 5, R is the gas constant, T is the temperature (298 K), h is Planck's constant and k_B is Boltzmann's constant.

$$KIE_{Eyring} = \exp\{(\Delta G^\ddagger_D - \Delta G^\ddagger_H)/RT\} \tag{3}$$

$$KIE_{Wigner} = KIE_{Eyring} \times Q_{tH}/Q_{tD} \tag{4}$$

$$Q_t = 1 + (h\nu/k_B T)^2/24 \tag{5}$$

The substrate binding free energy (BFE) was calculated by taking the optimized geometries of the three hydroxyarginine complexes (**IM2**C4S, **IM2**C3S and **IM2**C3R). A single point frequency calculation at UB3LYP/BS1 with CPCM included was done on all structures. Thereafter, the structures were split into two components, namely the hydroxyarginine and the protein, i.e. **IM2** minus the hydroxyarginine. We then did a frequency at the same level of theory on both the hydroxyarginine and the protein structures. The binding free energy is evaluated from the difference in free energy of the three individual components.

Acknowledgements

[1]   For recent reviews in the field, see, e.g., a) F. Kudo, A. Miyanaga, T. Eguchi, *Nat. Prod. Rep.* **2014**, *31*, 1056–1073; b) L.-F. Wu, S. Meng, G.-

L. Tang, *Biochem. Biophys. Acta* **2016**, *1864*, 453–470; c) R. G. S. Berlinck, A. F. Bertonha, M. Takaki, J. P. G. Rodriguez, *Nat. Prod. Rep.* **2017**, *34*, 1264–1301; d) S.-S. Gao, N. Naowarojna, R. Cheng, X. Liu, P. Liu, *Nat. Prod. Rep.* **2018**, *35*, 792–837; e) M. Hellwig, *Angew. Chem. Int. Ed.* **2019**, *58*, 16742–16763; f) J. B. Hedges, K. S. Ryan, *Chem. Rev.* **2020**, *120*, 3161–3209.

[2]   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; i) M. D. White, E. Flashman, *Curr. Opin. Chem. Biol.* **2016**, *31*, 126–135.

[3]   a) C. J. Schofield, Z. Zhang, *Curr. Opin. Struc. Biol.* **1999**, *9*, 722–731; b) K. Gorres, R. T. Raines, *Crit. Rev. Biochem. Mol. Biol.* **2010**, *45*, 106–124; c) 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; d) S. P. de Visser, *Chem. Record* **2018**, *18*, 1501–1516.

[4]   D. Baud, P.-L. Saaidi, A. Monfleur, M. Harari, J. Cuccaro, A. Fossey, M. Besnard, A. Debard, A. Mariage, V. Pellouin, J.-L. Petit, M. Salanoubat, J. Weissenbach, V. de Berardinis, A. Zaparucha, *ChemCatChem.* **2014**, *6*, 3012–3017.

[5]   a) J. H. Carter II, R. H. Du Bus, J. R. Dyer, J. C. Floyd, K. C. Rice, P. D. Shaw, *Biochemistry* **1974**, *13*, 1227–1233; b) X. Yin, T. M. Zabriskie, *ChemBioChem.* **2004**, *5*, 1274–1277; c) V. Helmetag, S. A. Samel, M. G. Thomas, M. A. Marahiel, L.-O. Essen, *FEBS J.* **2009**, *276*, 3669–3682; d) A. J. Mitchell, N. P. Dunham, R. J. Martinie, J. A. Bergman, C. J. Pollock, K. Hu, B. D. Allen, W.-c. Chang, A. Silakov, J. M. Bollinger Jr, C. Krebs, A. K. Boal, *J. Am. Chem. Soc.* **2017**, *139*, 13830–13836; e) N. P. Dunham, A. J. Mitchell, J. M. Del Río Pantoja, C. Krebs, J. M. Bollinger Jr, A. K. Boal, *Biochemistry* **2018**, *57*, 6479–6488.

[6]   N. P. Dunham, W.-c. Chang, A. J. Mitchell, R. J. Martinie, B. Zhang, J. A. Bergman, L. J. Rajakovich, B. Wang, A. Silakov, C. Krebs, A. K. Boal, J. M. Bollinger Jr, *J. Am. Chem. Soc.* **2018**, *140*, 7116–7126.

[7]   C. R. Zwick III, M. B. Sosa, H. Renata, *Angew. Chem. Int. Ed.* **2019**, *58*, 18854–18858.

[8]   a) S. Martinez, M. Fellner, C. Q Herr, A. Ritchie, J. Hu, R. P. Hausinger, *J. Am. Chem. Soc.* **2017**, *139*, 11980–11988; b) Z. Zhang, T. J. Smart, H. Choi, F. Hardy, C. T. Lohans, M. I. Abboud, M. S. W. Richardson, R. S. Paton, M. A. McDonough, C. J. Schofield, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 4667–4672; c) M. Li, S. Martinez, R. P. Hausinger, J. P. Emerson, *Biochemistry* **2018**, *57*, 5696–5705.

[9]   Z. Guo, J. Li, H. Qin, M. Wang, X. Lv, X. Li, Y. Chen, *Angew. Chem. Int. Ed.* **2015**, *54*, 5175–5178.

[10]   Y. Masuda, C. Maruyama, K. Kawabata, Y. Hamano, T. Doi, *Tetrahedron* **2016**, *72*, 5602–5611.

[11]   C.-Y. Chang, S.-Y. Lyu, Y.-C.; Liu, N.-S. Hsu, C.-C. Wu, C.-F. Tang, K.-H. Lin, J.-Y. Ho, C.-J. Wu, M.-D. Tsai, T.-L. Li, *Angew. Chem. Int. Ed.* **2014**, *53*, 1943–1948.

[12]   H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucl. Acids Res.* **2000**, *28*, 235–242.

[13]   a) P. C. A. Bruijnincx, G. van Koten, R. J. M. Klein Gebbink, *Chem. Soc. Rev.* **2008**, *37*, 2716–2744; b) S. Kal, L. Que Jr, *J. Biol. Inorg. Chem.* **2017**, *22*, 339–365; c) S. R. Iyer, V. D. Chaplin, M. J. Knapp, E. I. Solomon, *J. Am. Chem. Soc.* **2018**, *140*, 11777–11783.

[14]   a) D. A. Proshlyakov, T. F. Henshaw, G. R. Monterosso, M. J. Ryle, R. P. Hausinger, *J. Am. Chem. Soc.* **2004**, *126*, 1022–1023; b) J. M. Bollinger Jr, J. C. Price, L. M. Hoffart, E. W. Barr, C. Krebs, *Eur. J. Inorg. Chem.* **2005**, 4245–4254.

[15]   a) J. C. Price, E. W. Barr, B. Tirupati, J. M. Bollinger Jr, C. Krebs, *Biochemistry* **2003**, *42*, 7497–7508; b) 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; c) W. Yan, H. Song, F. Song, Y. Guo, C.-H. Wu, A. S. Her, Y. Pu, S. Wang, N. Naowarojna, A. Weitz, M. P. Hendrich, C. E. Costello, L. Zhang, P. Liu, Y. J. Zhang, *Nature* **2015**, *527*, 539–543; d) D. Galonić Fujimori, E. W. Barr, M. L. Matthews, G. M. Koch, J. R. Yonce, C. T. Walsh, J. M. Bollinger Jr, C. Krebs, P. J. Riggs-Gelasco, *J. Am. Chem. Soc.* **2007**, *129*, 13408–13409.

[16]   A. Grosdidier, C. Zoete, O. Michielin, *Nucl. Acids Res.* **2011**, *39*, 270–277.

[17]   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) 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; d) E. A. C. Bushnell, G. B. Fortowsky, J. W. Gauld, *Inorg. Chem.* **2012**, *51*, 13351–13356; e) H. J. Kulik, C. L. Drennan, *J. Biol. Chem.* **2013**, *288*, 11233–11241; f) A. Wójcik, M. Radoń, T. Borowski, *J. Phys. Chem. A* **2016**, *120*, 1261–1274; g) X. Song, J. Lu, W. Lai, *Phys. Chem. Chem. Phys.* **2017**, *19*, 20188–20197; h) H. Su, X. Sheng, W. Zhu, G. Ma, Y. Liu, *ACS Catal.* **2017**, *7*, 5534–5543; i) R. N. Manna, T. Malakar, B. Jana, A. Paul, *ACS Catal.* **2018**, *8*, 10043–10050; j) Y.-T. Lin, A. Stańczak, Y. Manchev, G. D. Straganz, S. P. de Visser, *Chem. Eur. J.* **2020**, *26*, 2233–2242.

[18]   a) R. Latifi, M. A. Sainna, E. V. Rybak-Akimova, S. P. de Visser, *Chem. Eur. J.* **2013**, *19*, 4058–4068; b) D. Sahoo, M. G. Quesne, S. P. de Visser, S. P. Rath, *Angew. Chem. Int. Ed.* **2015**, *54*, 4796–4800.

[19]   a) E. Godfrey, C. S. Porro, S. P. de Visser, *J. Phys. Chem. A* **2008**, *112*, 2464–2468; b) A. R. Diebold, C. D. Brown-Marshall, M. L. Neidig, J. M. Brownlee, G. R. Moran, E. I. Solomon, *J. Am. Chem. Soc.* **2011**, *133*, 18148–18160; c) C. R. Wick, H. Lanig, C. M. Jäger, N. Burzlaff, T. Clark, *Eur. J. Inorg. Chem.* **2012**, 4973–4985; d) G. Dong, S. Shaik, W. Lai, *Chem. Sci.* **2013**, *4*, 3624–3635; e) A. Wójcik, M. Radoń, T. Borowski, *J. Phys. Chem. A* **2016**, *120*, 1261–1275; f) X. Song, J. Lu, W. Lai, *Phys. Chem. Chem. Phys.* **2017**, *19*, 20188–20197; g) A. S. Faponle, F. P. Seebeck, S. P. de Visser, *J. Am. Chem. Soc.* **2017**, *139*, 9259–9270; h) S. S. Chaturvedi, R. Ramanan, N.i Lehnert, C. J. Schofield, T. G. Karabencheva-Christova, C. Z. Christov, *ACS Catal.* **2020**, *10*, 1195–1209.

[20]   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.

[21]   F. G. Cantú Reinhard, A. S. Faponle, S. P. de Visser, *J. Phys. Chem. A* **2016**, *120*, 9805–9814.

[22]   a) S. Shaik, D. Kumar, S. P. de Visser, *J. Am. Chem. Soc.* **2008**, *130*, 10128–10140; 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; c) P. Barman, F. G. Cantú Reinhard, U. K. Bagha, D. Kumar, C. V. Sastri, S. P. de Visser, *Angew. Chem. Int. Ed.* **2019**, *58*, 10639–10643.

[23]   a) D. Kumar, H. Hirao, L. Que Jr, S. Shaik, *J. Am. Chem. Soc.* **2005**, *127*, 8026–8027; b) S. P. de Visser, *Angew. Chem. Int. Ed.* **2006**, *45*, 1790–1793; c) C. Geng, S. Ye, F. Neese, *Angew. Chem. Int. Ed.* **2010**, *49*, 5717–5720; d) S. Ye, F. Neese, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1228–1233; e) L. Bernasconi, E. J. Baerends, *J. Am. Chem. Soc.* **2013**, *135*, 8857–8867; f) A. Timmins, M. G. Quesne, T. Borowski, S. P. de Visser, *ACS Catal.* **2018**, *8*, 8685–8698.

[24]   S. P. de Visser, *Chem. Eur. J.* **2020**, *26*, 5308–5327.

[25]   a) F. G. Bordwell, J.-P. Cheng, *J. Am. Chem. Soc.* **1991**, *113*, 1736–1743; b) J. M. Mayer, *Annu. Rev. Phys. Chem.* **2004**, *55*, 363–390; c) S. P. de Visser, *J. Am. Chem. Soc.* **2010**, *132*, 1087–1097.

[26]   S. Ghafoor, A. Mansha, S. P. de Visser, *J. Am. Chem. Soc.* **2019**, *141*, 20278–20292.

[27]   a) M. R. A. Blomberg, T. Borowski, F. Himo, R.-Z. Liao, P. E. M. Siegbahn, *Chem. Rev.* **2014**, *114*, 3601–3658; b) S. P. de Visser, M. G. Quesne, B. Martin, P. Comba, U. Ryde, *Chem. Commun.* **2014**, *50*, 262–282; c) M. G. Quesne, T. Borowski, S. P. de Visser, *Chem. Eur. J.* **2016**, *22*, 2562–2581; d) M. Pickl, S. Kurakin, F. G. Cantú Reinhard, P. Schmid, A. Pöcheim, C. K. Winkler, W. Kroutil, S. P. de Visser, K. Faber, *ACS Catal.* **2019**, *9*, 565–577.

14

[28] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.

[29] *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, 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, 2010.

[30] a) M. G. Quesne, R. Latifi, L. E. Gonzalez-Ovalle, D. Kumar, S. P. de Visser, *Chem. Eur. J.* **2014**, *20*, 435–446; b) A. Timmins, M. Saint-André, S. P. de Visser, *J. Am. Chem. Soc.* **2017**, *139*, 9855–9866.

[31] 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.

[32] 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.

[33] M. A. Kaczmarek, A. Malhotra, G. A. Balan, A. Timmins, S. P. de Visser, *Chem. Eur. J.* **2018**, *24*, 5293–5302.

[34] C. Peng, H. B. Schlegel, *Isr. J. Chem.* **1993**, *33*, 449–454.

[35] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.

[36] M. Cossi, G. Scalmani, N. Rega, V. Barone, *J. Chem. Phys.* **2002**, *117*, 43–54.

[37] a) P. Barman, P. Upadhyay, A. S. Faponle, J. Kumar, S. S. Nag, D. Kumar, C. V. Sastri, S. P. de Visser, *Angew. Chem. Int. Ed.* **2016**, *55*, 11091–11095; b) G. Mukherjee, C. W. Z. Lee, S. S. Nag, F. G. Cantú Reinhard, D. Kumar, C. V. Sastri, S. P. de Visser, *Dalton Trans.* **2018**, *47*, 14945–14957.

Entry for the Table of Contents
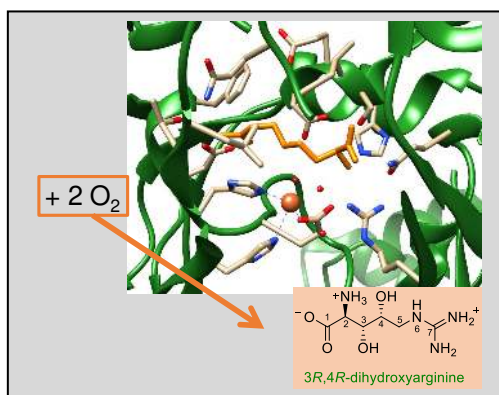
Text for Table of Content:

Density functional theory calculations on the arginine dihydroxylating dioxygenase OrfP focus on the mechanism of arginine activation. We find a selective initial C^3–H oxygenation pathway followed by a second oxygenation cycle on the C^4–H group. A comparison with other arginine activating enzymes highlight a tight substrate binding pocket that locks the product and enables two oxygenation cycles.