



<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *Journal of Physical Chemistry B*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Lundberg, M., Morokuma, K. (2007)

Protein environment facilitates O₂ binding in non-heme iron enzyme: An insight from ONIOM calculations on isopenicillin N synthase (IPNS).

Journal of Physical Chemistry B, 111(31): 9380-9389

<http://dx.doi.org/10.1021/jp071878g>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-145466>

Protein Environment Facilitates O₂ Binding in Non-Heme Iron Enzyme. An Insight from ONIOM Calculations on Isopenicillin N Synthase (IPNS).

Marcus Lundberg¹, and Keiji Morokuma^{1,2}.

¹ *Fukui Institute for Fundamental Chemistry, Kyoto University, 34-4 Takano Nishihiraki-cho, Sakyo, Kyoto 606-8103, Japan*

² *Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, GA 30322, USA*

Abstract. Binding of dioxygen to a non-heme enzyme has been modeled using the ONIOM combined quantum mechanical/molecular mechanical (QM/MM) method. For the present system, isopenicillin N synthase (IPNS), binding of dioxygen is stabilized by 8-10 kcal/mol for a QM:MM (B3LYP:Amber) protein model compared to a quantum mechanical model of the active site only. In the protein system, the free energy change of O₂ binding is close to zero. Two major factors consistently stabilize O₂ binding. The first effect, evaluated at the QM level, originates from a change in coordination geometry of the iron center. The active-site model artificially favors the deoxy state (O₂ not bound), since it allows too large rearrangements of the five-coordinate iron site. This error is corrected when the protein is included. The corresponding effect on binding energies is 3-6 kcal/mol, depending on the coordination mode of O₂ (side-on or end-on). The second major factor that stabilizes O₂ binding is van der Waals interactions between dioxygen and the surrounding enzyme. These interactions, 3-4 kcal/mol at the MM level, are neglected in models that include only the active site. Polarization of the active site by surrounding amino acids does not have a significant effect on the binding energy in the present system.

Keywords. QM/MM, 2-histidine-1-carboxylate, density functional theory, binding energy

· Corresponding Author: E-mail: morokuma@fukui.kyoto-u.ac.jp, Phone:+81-75-711-7843

I. Introduction

Oxygen-activated non-heme iron enzymes perform a wide range of chemical transformations and play important roles in many metabolic pathwaysⁱ. An important common step in these enzymes is the binding of O₂ to the metal center, a step also shared by many heme and Cu-enzymes. Dioxygen binding activates the enzymes and triggers the catalytic reactions. The present contribution studies this important O₂ binding step in detail. More specifically, it is shown how the surrounding protein, included with a quantum mechanical/molecular mechanical (QM/MM) model, affects O₂ binding in the mononuclear non-heme iron enzyme isopenicillin N synthase (IPNS).

Both experimental and theoretical studies have provided insights into the catalytic mechanisms of the mononuclear non-heme iron enzymes. In theoretical studies, the requirement for a high-level quantum-mechanical description of the transition-metal center, typically by hybrid density-functional theory (DFT), has in most cases limited computational models to the active site only. Although this approach has had considerable successⁱⁱ, the disadvantage is that potential effects of the protein environment cannot be evaluated. In this aspect, O₂ binding is especially interesting since the active-site DFT approach apparently has problems in describing this reaction. For non-heme iron systems, active-site calculations using the B3LYP functional typically estimate O₂ binding to be endoergic by 10 kcal/mol or moreⁱⁱⁱ. Such a large endoergicity would leave little room for oxygen-activated chemistry, and is not consistent with the observed reactivity of these enzymes.

Incorrect estimations of O₂ binding energies can either be due to errors of the method or deficiencies in the model. Consistent accuracy for coordinatively unsaturated transition-metal species has not yet been reached^{iv-vi}, but hybrid functionals tend to give reasonable energies for many saturated (or nearly saturated) transition-metal systems relevant for biological applications^{vii,viii}. However, for many reactions the results are highly sensitive to the amount of Hartree-Fock (HF) exchange included in the functional. For biologically relevant iron systems, hybrid functionals have in many cases proved to be good choices for calculations of relative energies^{ix-xii}. B3LYP performs relatively well for O₂ binding to non-heme iron when compared to high-level methods like CCSD(T) and CASPT2^{xiii}. However, there are suggestions that the optimum amount of HF exchange could be 10-15 %^{xiv-xvi}, compared to 20 % in B3LYP.

A decrease of the amount of HF exchange in B3LYP to 15 % increases O₂ binding by a several kcal/mol^{vii}. However, even with this correction, O₂ still does not bind to non-heme iron in active-site models.

In previous reports, extensions of the quantum mechanical (QM) treatment by including the protein in a molecular mechanics (MM) description (QM/MM models)^{xvii-xxii} have shown significant effects on the non-heme Fe-O₂ interactions^{xxiii-xxv}. It is therefore reasonable to assume that deficiencies in the active-site model may contribute to the problems in describing O₂ binding in non-heme iron systems. There exists a number of QM/MM studies of O₂ binding to heme systems, but they will not be further discussed since the electronic and geometric features of the O₂-bound state are significantly different between the different enzyme families. **Heme systems are characterized by strong ligand fields whereas non-heme systems have relatively weak ligand fields. Further, the four-dentate heme group puts geometrical restrictions on the geometry of the iron center that are significantly different from those of the non-heme systems.**

To this point, only one study has quantified the energetic effect of the protein on O₂ binding in a non-heme enzyme. Adding the surrounding protein to an active-site model of the dinuclear non-heme iron site in hemerythrin stabilizes the oxy state (O₂ bound) by 10 kcal/mol^{xxiii}. The final B3LYP/OPLS-AA binding energy agrees well with the fact that hemerythrin binds dioxygen reversibly. The major protein contributions comes from; (1) van der Waals interactions between O₂ and protein atoms and (2) polarization of the active site by a neighboring carboxylate side-chain, not included in the quantum mechanical model. Other QM/MM calculations also suggest significant effects of the protein on dioxygen binding. In the dinuclear iron center of methane monooxygenase, the Fe-O₂ distance decreases from 3.22 Å in an active-site model to 2.16 Å in an ONIOM (B3LYP:Amber) model^{xxiv}. It should also be noted that a CASSCF/Amber treatment of 2-oxoglutarate, a mononuclear non-heme iron enzyme, gives different geometries, electronic structure and spin-state ordering compared to a DFT active-site model^{xxv}; however, from this study it is not possible to separate the effect of the change in method from the effect of the protein.

Recently, two new studies of the O₂-bound (or peroxide) state in non-heme enzymes have been reported^{xxvi,xxvii}. As the number of experimentally accessible states increases, the interest to model O₂ binding and the O₂-bound state will also increase. To improve the understanding of dioxygen interactions in protein systems, and to outline the important effects in protein modeling, the present contribution investigates O₂ binding in the mononuclear non-heme enzyme isopenicillin N synthase (IPNS). The effect of the protein is

obtained by comparing two computational approaches, the active-site DFT model and an ONIOM DFT:MM model that includes the entire protein. Since the high-level QM treatment is the same in the two approaches, all changes in geometries, electronic structures and energies can be attributed to the surrounding protein.

Isopenicillin N synthase uses the full four-electron oxidative power of O₂ to transform its three-peptide substrate δ -(L- α -aminoadipoyl)-L-cysteinyl-d-valine (ACV) to isopenicillin N (see Fig. 1)^{xxviii}. This is a key step in synthesis of the important β -lactam antibiotics penicillins and cephalosporins^{xxix}. Since biosynthesis is the most efficient way to synthesize the β -lactam core, the enzymatic mechanism is also of industrial interest^{xxx}.

--- Figure 1 ---

Computational investigations of IPNS are facilitated by the existence of several X-ray structures, representing different states of the enzymatic reaction^{xxxi-xxxiii}. In all structures, iron is coordinated by a facial triad consisting of two histidines and one carboxylate. IPNS shares this 2-His-1-carboxylate motif with many other oxygen-activated enzymes, e.g., extradiol dioxygenase, α -ketoglutarate-dependent oxygenases, rieske dioxygenases and pterin-dependent hydroxylases^{xxxiv}. In IPNS, the initial step of the catalytic cycle is binding of the ACV substrate. This gives a five-coordinate iron site where iron is ligated by His214, Asp216, His270, the thiolate group of the substrate and one water molecule^{xxxii} (see Fig. 2). In the following sections, the X-ray structure of this state will be called the five-coordinate X-ray structure. At this point, iron is in the high-spin Fe(II)-state (S=2)^{xxxv,xxxvi}. This is the most common spin-state for this family of enzymes and reflects the weak ligand field of the coordinating amino acids^{xxxvii}.

--- Figure 2 ---

Substrate binding to the active site increases the enzymes affinity for oxygen, but the proposed O₂-bound state has not been isolated and no experimental data is therefore available. However, supplying NO instead of O₂ makes it possible to obtain an X-ray structure of an NO-bound state^{xxxii}. Binding of NO does not lead to any major structural rearrangements of the active site, but Fe becomes six-coordinate with octahedral coordination. NO binds in the end-on mode (only one substrate atom, in this case nitrogen, coordinates to iron) in a position trans to Asp216. In the following sections, this X-ray structure will be called the six-coordinate X-ray structure. It is commonly assumed that O₂ binds in the same position as NO (see Fig. 2).

Despite the similarities of their active sites, enzymes in the 2-His-1-carboxylate family seem rather flexible when it comes to O₂ binding. **Until recently, the X-ray structure of naphthalene dioxygenase^{xxxviii} represented the only case where a dioxygen molecule had actually been trapped. In this structure, oxygen binds in a side-on conformation (both oxygen coordinating to iron) essentially trans to a bidentate carboxylate ligand. In the recent X-ray structure of extradiol dioxygenase, oxygen appears to bind in a side-on conformation, trans to a monodentate glutamate ligand^{xxvi}.** Other X-ray structures of 2-His-1-carboxylate enzymes, with NO analogues, show that extradiol dioxygenase BphC probably binds O₂ in a similar position as IPNS^{xxxix}, while clavamate synthase probably binds O₂ trans to one of the histidine ligands^{xl}. In the present investigation of IPNS, it is assumed that O₂ occupies the same position as NO and only binding trans to Asp216 is considered. However, both side-on and end-on binding are taken into account.

2. Computational Methods

QM/MM calculations have been performed using the two-layer ONIOM scheme^{xxi,xli,xlii}. In this approach the system is divided in two parts. The “model” system includes the reactive part of the enzyme and this system is treated by an accurate (high-level) computational method that can describe bond breaking and formation. The “real” system includes all protein atoms and is treated with a low-level (fast) computational method. Bonds between atoms in the selected “model” system and the atoms only included in the “real” system are capped with hydrogen link atoms. Interactions between the two layers have mainly been treated at the classical level (mechanical embedding or ME), but electrostatic interactions have also been accounted for in a semi-classical way by incorporating the MM charges into the QM Hamiltonian (electronic embedding or EE)^{xviii,xxi}.

In the present investigation, the high-level method is the hybrid DFT functional B3LYP^{xliii} and the low-level method is the Amber force field^{xliv}. Calculations are performed using Gaussian03^{xlv}. The applicability of the ONIOM (B3LYP:Amber) method for describing structures and energies of enzymatic systems has recently been illustrated for several systems, e.g., soluble and particulate methane monooxygenase, methylmalonyl-CoA mutase and glutathione peroxidase^{xxiv,xlvi-xlviii}.

Results from the ONIOM protein calculations are compared to results from a DFT model that includes only the enzyme active site. To facilitate comparisons between protein

and active-site systems, the “model” part of the ONIOM system matches the active-site model, concerning both the atom selection and to the choice of method. The ONIOM “model” system (and the active-site model) includes Fe, the three amino acids His214, Asp216 and His270 (*Aspergillus nidulans* numbering) that coordinate Fe, a water ligand and selected parts of the substrate (see Fig. 3). **The truncation of the N-C bond in the substrate leads to a false polar N-H bond. However, the reaction proceeds at a considerable distance from this bond, and this truncation should therefore not affect the present results.** Initial coordinates are taken from the X-ray structure where Fe is six-coordinate (1BLZ, 1.45 Å resolution)^{xxxii}, and NO is replaced by O₂. Including dioxygen, the model system contains 65 atoms. Geometries are optimized at the B3LYP/6-31G(d) level and final energies evaluated using the 6-311+G(d,p) basis set. Increasing the basis set to 6-311+G(2df,2pd) for the active-site model had a limited effect (<0.3 kcal/mol). In the active-site model Cartesian coordinates of link atoms were frozen to keep the geometry closer to the X-ray structure (see Fig. 3). One atom in the substrate carboxylate is also kept frozen to avoid the formation of artificial hydrogen bonds. All calculations are performed with the unrestricted formalism. Reported energies do not contain corrections for spin-contamination. For the active-site model, corrections using the Heisenberg Hamiltonian formalism^{xlix} would be ≤0.7 kcal/mol.

--- Figure 3 ---

The “real” ONIOM system consists of 5368 atoms. It includes all protein atoms, and also all crystallographic water molecules within 20 Å of the iron center. No additional waters, e.g. a solvation shell, were added to the model. During the first geometry optimization, some surface residues change their orientation, but the overall structure is retained. In the present static optimization approach, the benefit of a solvation sphere of water molecules would be small, since the important dielectric effect of water comes from reorientation of water dipoles. For the present reaction, with limited solvation effects, neglect of water solvation should not affect the results.

Amino acids are assigned standard Amber parameters. Glu and Asp residues are assigned as negative while Lys and Arg residues are assigned as positive. Histidine protonation states are assigned based on a visual inspection of their local environment. As a result, histidines 47, 62, 82, 114, 124, 135, 259 are assigned as HIE (N_ε protonated) while histidines 174, 214, 270 are assigned as HID (N_δ protonated).

The ONIOM system is fully optimized for all stationary points. The calculation of ONIOM energies and gradients require two separate MM calculations, one of the model

system and one of the real system. To perform these calculations, Amber parameters have to be assigned to all bonds. However, the energy and gradient terms from these bonded parameters (except those for the link atoms) effectively cancel in the ONIOM formalism, since they are identical for the real and the model system. In the present study, non-standard bonded terms were extracted from a Hessian calculation using the program XYZviewer (unpublished). An alternative solution would be to set all the bonded terms to zero. This would give identical results in an ONIOM QM:MM calculation, but not in a pure MM optimization.

However, non-bonded interactions between “model” and “real” system are important parts of the total ONIOM energy. It turns out that the van der Waals interactions between dioxygen and surrounding protein have significant effects on the binding energies. These effects are calculated using Amber parameters taken from oxygen in amides and anionic acids ($R=1.6612 \text{ \AA}$, $\epsilon=0.21 \text{ kcal/mol}$)^{xliv}. Fe parameters are taken from heme, but the values have no effect on binding energies. Atoms in the model system, including amino acids, are assigned point charges from a B3LYP/6-31G(d) ESP (Merz-Kollman) calculation of the optimized system without O₂. In mechanical embedding, these charges are not updated during the reaction. Dioxygen charges are artificially set to 0.0. The use of zero charge for O₂ is suboptimal, but is considered preferable relative to a non-consistent change in the MM parameterization during a reaction step. Instead of reparameterizing all charges when dioxygen is introduced, electrostatic effects directly involving O₂ are evaluated using the electronic embedding scheme. In these calculations, atoms in the “model” system are assigned point charges from a Mulliken population analysis. In the electronic embedding scheme, these charges are self-consistently updated in each QM:MM micro iteration.

Binding energy calculations have also been performed starting from an X-ray structure representing the state where iron is five-coordinate (1BKO, 1.3 Å resolution)^{xxxii}. The computational model for these calculations was prepared in an identical way as described above.

All calculations use enzyme + free O₂ as reference state. To obtain a reasonable estimate of the O₂ binding energies, the electronic binding energies calculated with the active-site model are corrected by adding zero-point, thermal and solvent corrections. Hessian contributions are calculated from a freely optimized active-site model where the terminal methyl groups are removed from the ligands and the substrate model is replaced by methylthiolate.

Solvent corrections for active-site models are calculated using IEFPCM ($\epsilon=4$). The reference state of dioxygen is aqueous solution and experimental solubility data give a solvation energy for O₂ of +3.9 kcal/mol. This solvation effect thus favors the O₂ bound state.

Protein effects are quantified by comparing potential energies from the ONIOM protein system and the active-site model. In this comparison, zero-point, thermal and solvent corrections are not included, which implicitly assumes that these contributions are the same for the two systems. The major parts of the thermal and solvation corrections originate from the use of free O₂ in the reactant state, and should be the same in all calculations.

Solvent effects are difficult to handle in QM/MM calculations. At present, it is not possible to evaluate the solvent effects in the ONIOM protein system using continuum solvation methods. As an approximation, the effect of continuum solvation is taken from evaluations of the active-site model. Since the solvation radius is significantly smaller in the active-site model than in the protein, this procedure leads to an overestimation of the solvation effect in the ONIOM system. However, the difference in solvation energy between the protein's deoxy and the oxy states is < 1 kcal/mol in the active-site model. Even if this effect is overestimated in the protein system, errors should be significantly smaller than 1 kcal/mol and therefore not critical in the present calculations.

Assuming similar free-energy corrections for active site and protein models also neglects protein contributions to free energy that have been argued to be important in enzyme catalysis¹. With the static optimization approach used in the present study, it is not possible to sample the conformational space and obtain protein free energies. However, since O₂ is a small molecule, it can be argued that protein conformation and flexibility should remain essentially the same before and after binding. The total binding entropy should therefore be relatively well described by the harmonic approximation. If this is correct, the remaining concern with the static optimization approach is that observed differences in protein energy should be directly coupled to the reaction coordinate, i.e., to avoid contributions from spurious changes in MM minima. A conservative solution is to use similar protein configurations for reactant and product structures. This was achieved by iterating between reactant and product structures until energies converge. In the present study this required three to four iterations.

3. Results and Discussions

To facilitate the discussion of protein effects on O₂ binding, the present section starts with a short description of how these effects will be analyzed. The term “protein effect” is used for the change in relative energy of the oxy state (O₂ bound), compared to the deoxy state (O₂ not bound), when going from an active-site model to an ONIOM model that includes the entire protein. In the ONIOM QM/MM description, the protein effect can be divided into two parts, a high-level (QM) effect, evaluated using B3LYP, and a low-level (MM) effect, evaluated using the Amber force field. To clarify what is included in these terms, the expression for the ONIOM energy is repeated:

$$E(\text{ONIOM}) = E(\text{High : Model}) + E(\text{Low : Real}) - E(\text{Low : Model})$$

High and Low represent the computational methods while Model and Real represent the selected systems (see Computational Details). The high-level effect is obtained by comparing binding energies from the E(High:Model) term with binding energies from the active-site model. This effect represents changes in geometry of the active site, and indirectly in electronic structure, caused by interactions with the surrounding protein. The low-level effects are obtained from direct calculations of [E(Low:Real)-E(Real:Model)] for reactant (deoxy) and product (oxy) states. Low-level effects include non-bonded interactions between atoms in the model system and the surrounding protein, as well as all interactions between atoms present only in the real system (i.e. the surrounding protein).

Results for the active-site model are presented first since they constitute the basis for further investigations. This is followed by the results from the ONIOM calculations and a discussion of the origin of the protein effects.

A. Active-Site Model Calculations. Prior to O₂ binding, Fe is in a high-spin Fe(II)-state (S=2)^{xxxv,xxxvi}. Binding of triplet dioxygen (S=1) gives an Fe-O₂ system where the highest spin multiplicity is a septet (S=3). In the active-site study, all possible spin multiplicities, septet, quintet, triplet and singlet, are investigated. In addition, both end-on and side-on binding modes (see Fig. 2) are taken into account. This gives a total of eight different states. However, preliminary investigations indicated that singlet states are unstable by 15-20 kcal/mol (compared to the septet state at the B3LYP 6/31G(d) level). They were therefore excluded from more detailed investigations. In addition, no true minima could be found for a side-on triplet state. Remaining five states are listed in Table 1.

---- Table 1 ---

In the B3LYP description, septet and quintet states can be seen as Fe(III)-superoxo radical species. With the present method, the best indication of the iron oxidation state is the spin population. For high-spin iron complexes, spin populations of 3.7-3.8 approximately correspond to Fe(II), of 4.0-4.2 to Fe(III), and of 3.0-3.3 to Fe(IV). The difference between the calculated spin populations and the formal number of unpaired d-electrons is due to spin delocalization^{li}. However, the amount of delocalization is almost the same for all iron complexes of the same oxidation state. Spin populations are therefore good indicators of the oxidation states of high-spin iron systems.

When dioxygen binds, an electron is transferred from Fe(II) to an anti-bonding 3π orbital in O_2 . This is reflected in the Fe Mulliken spin population of 4.1 (see Table 1), which is characteristic of an Fe(III) state (see above). For the septet state, the spin on the superoxo radical is aligned parallel with the spins on Fe(III), while for the quintet state the alignment is anti-parallel. The absolute value of total spin of the superoxo radical is ≥ 1 in the septet state and ≤ 1 in the quintet state, since electrons formally assigned to ligands are substantially backdonated. Consequently, spin formally assigned to d-orbitals on Fe is delocalized to all the ligands. A population analysis for the septet state with side-on binding of O_2 is available in the Supporting information. The electron transfer reduces the formal O-O bond order and increases the bond length from 1.26 Å in O_2 to 1.30-1.31 Å in the superoxo radical.

In the triplet state, the electronic structure is less clear. O-O bond distance and spin populations (see Table 1) indicate a substantial component of an Fe(II)- O_2 state with four spins on Fe(II) and two spins on O_2 antiferromagnetically coupled.

As seen above, in the active-site calculations the side-on septet state is the most stable state, but energy differences are relatively small. The ordering of the states is in agreement with similar calculations on a biomimetic non-heme iron complex^{liii}. Still, the limited accuracy of the current computational method does not allow a final determination of the preferred electronic state. However, the most important result is that even in the most stable state, O_2 is unbound by 10.6 kcal/mol, mainly due to the large entropic cost of binding free O_2 .

The entropy effects are calculated using the Sackur-Tetrode equation, which takes into account the loss of translational entropy when free dioxygen becomes trapped in the active site. However, it can be argued that the substrate has lower entropy in solution than in gas phase^{liii}. The present treatment could therefore overestimate the entropy costs of O_2 binding. Experimental data for a non-heme complex show entropy effects of 11 kcal/mol for

the initial oxygen-activated reaction step^{liv}. If it is assumed that these entropy effects are mainly due to the association of O₂, the presently calculated values (e.g. 12 kcal/mol for the septet side-on state) are not unreasonable, although they should still be treated with some caution.

As indicated in the introduction, the binding energy is expected to highly dependent on the choice of functional. To briefly check the stability of the results, electronic energies were recalculated with a limited number of functionals, using the B3LYP geometries. PBE1PBE (hybrid) and B1B95 (meta-hybrid) gave results within 1 kcal/mol of the B3LYP result, while non-hybrid functionals significantly stabilizes the oxy state. Relatively to B3LYP the effect is 12.3 kcal/mol for B86 (GGA), 7.6 kcal/mol for BLYP (GGA), and 19.6 kcal/mol for BB95 (meta-GGA). The effect of the exchange component can also be seen by changing the amount of HF exchange in B3LYP. A decrease from 20 to 15 % stabilizes the oxy state by 4.4 kcal/mol, and the effect is almost linear down to 0 %. If 15 % HF exchange is used, this gives an endoergicity of O₂ binding of 6.3 kcal/mol. As discussed in the introduction, a large endoergicity of dioxygen binding does not give a catalytically active system. In isopenicillin N synthase the rate-limiting barrier is 16.8 kcal/mol^{liv}, which leaves 10 kcal/mol for the first reaction barrier. The enzymatic mechanism of isopenicillin N synthase does not seem to be compatible with such a significant endoergicity of O₂ binding^{lvi}.

To further investigate the results for O₂ binding, efforts are made to investigate the possible limitations of the active-site model. For this purpose, the two lowest-energy structures, the side-on and end-on septet structures are chosen for further investigations with an extended ONIOM protein model. Including both structures makes it possible to see how the protein affects the relative energy of the two conformations.

B. ONIOM Protein System Calculations. In all the following comparisons, electronic binding energies of the active-site models (ΔE) are directly compared to ONIOM binding energies for the protein system.

The first set of ONIOM protein calculations start from the X-ray structure where Fe is six-coordinate, after replacing NO with O₂. Following each structure optimization, O₂ is iteratively removed and added until the binding energy converges. Relative energies, selected spin populations and structural parameters for all ONIOM calculations are listed in Table 2.

--- Table 2 ---

For the side-on septet structure the relative energy of the oxy state is calculated to be -5.4 kcal/mol, see Table 2. Compared to +2.7 kcal/mol for the active-site model, the protein effect is -8.1 kcal/mol, making the O₂ binding more favorable. For the end-on septet the relative energy changes from +5.2 kcal/mol to -5.2 kcal/mol when the protein environment is added. In this case the protein effect is -10.3 kcal/mol. Dividing the protein effects for side-on binding into its high and low-level components gives a high-level (QM) effect of -3.2 kcal/mol and a low-level (MM) effect of -4.9 kcal/mol. Corresponding numbers for end-on binding are -5.8 kcal/mol and -4.5 kcal/mol. These effects are significant and change the free energies of O₂ binding from highly endoergic to close to thermoneutral (see Fig. 4). As described earlier, a decrease in the amount of HF exchange to 15 % favors O₂ binding by 4.4 kcal/mol, and would result in an overall favorable energy for O₂ binding.

--- Figure 4 ---

The origin of the high-level effect will be analyzed first, followed by a dissection of the MM contributions. The end-on structure will be taken as example since the total protein effect is larger for that system. In the present ONIOM-ME calculations, high-level effects reflect changes in the active-site geometry. Calculations on methane monooxygenase indicated large protein effects on the geometry of the O₂ bound species, which could possibly explain the effect on the binding energy^{xxiv}. However, a comparison of the geometries between active site and ONIOM models in isopenicillin N synthase does not show any major changes in the Fe-O₂ interactions (see Fig. 5). Looking directly at dioxygen, the O-O distance increases by 0.01 Å (from 1.30 to 1.31 Å), while the smallest Fe-O distance change by 0.01 Å (from 2.06 Å to 2.05 Å) (see Table 2). There is change in the Fe-O-O angle from 104.2° to 115.5°, seen as an increase in the distance between Fe and the non-coordinating oxygen (Fe-O2 in Table 2). The Mulliken spin population of iron is constant at 4.11, while the spins on oxygen goes from 0.63 and 0.71 in the active-site model to 0.58 and 0.67 in the ONIOM model. The structure of the active site is also similar between the two models, e.g. no new hydrogen bonds to O₂ are formed in the protein structure. Based only on these observations, there is no clear explanation for the 5.7 kcal/mol QM effect on the reaction energy.

--- Figure 5 ---

In calculations of relative energies, the reactant deserves the same attention as the product. A comparison of the geometries of the deoxy structures between ONIOM and active-site models shows some important differences (see Fig. 5). The most striking observations are that in the active-site model, His270 and the substrate thiolate relax significantly from their positions in the protein system. As a quantitative measure, the root

mean square deviation (RMSD) between the active-site and the ONIOM model can be calculated. Looking at iron and the five atoms coordinating Fe in both structures, the RMSD is 0.25 Å in the reactant deoxy structure. This can be compared to 0.06 Å in the product structure.

In many systems, active-site models allow rotation of metal ligands out of their positions in the protein. Since these displacements usually are the same during a reaction sequence, they do not significantly affect the reaction energy. However, in the present reaction there is a transition from a five to a six-coordinate metal site and the relaxation of substrate and His270 becomes larger in the five-coordinate reactant, as illustrated in Fig. 5. Including the protein in a QM:MM description seems to give a similar description of the geometry for both structures, and consequently the protein correction is larger for the deoxy structure than for the O₂ bound state. Energetically, the protein therefore destabilizes the deoxy structure relative to the oxy state, which in turn gives a more exothermic reaction. It should be noted that the large flexibility of the active-site model is realized even though the present active-site model uses frozen Cartesian coordinates from the X-ray structure (see Fig 3). A stricter scheme for freezing amino-acid residues could reduce the artificial rotation in the active-site model, but then possibly at the expense of a less accurate treatment of other parts of the reaction.

The total relaxation energy, i.e. the energy difference between the QM part in active-site and QM:MM models^{lvii} is in the range of 31 kcal/mol for the deoxy structure. Although far from all this energy is due to *artificial* ligand relaxation in the active-site model, the magnitude of these energies illustrate the potential importance of using accurate active-site geometries.

The high-level protein effect can appear in two different situations. Either the active-site model is too flexible, which gives an artificially stable reactant state, or the ONIOM model is too rigid, artificially stabilizing a state similar to the original X-ray structure (in the present case the product state where iron is six-coordinate). To test these two different possibilities, identical O₂ binding calculations were performed using a crystal structure representing the reactant state (where iron is five-coordinate). If the ONIOM model would be too rigid, these calculations would artificially stabilize the reactant and the high-level effect should change sign. However, this is not what was found. Starting from the five-coordinate X-ray structure, the relative energy of the oxy state becomes -5.3 kcal/mol, very similar to the value for the six-coordinate X-ray structure (see Table 2). The high-level effect is -5.5

kcal/mol, close to -5.8 kcal/mol in the preceding calculation. The geometries of the O₂ bound states are also very similar for the two calculations as can be seen in Fig 6.

--- Figure 6 ---

These observations indicate that the ONIOM model is flexible enough to give a good description of the active site, no matter which starting structure that is used. The high-level protein effect should therefore be essentially correct. The reason why the high-level effect is slightly larger for the end-on structure, compared to the side-on structure, is probably that the latter requires structural rearrangements to form a seven-coordinate iron geometry. This structure also seems to be artificially stabilized by the active-site model, although less than the deoxy state.

Now turning the attention to the low-level effects, they are around -5 kcal/mol in all calculations (see Table 2). The low-level effects include non-bonded interactions between atoms in the model system and the surrounding protein, as well as all interactions between atoms present only in the surrounding protein. Looking again at end-on binding of O₂, calculated using the X-ray structure where Fe is six-coordinated, the total MM effect is -4.5 kcal/mol. This effect can easily be separated into components, and Fig. 7 shows the individual MM contributions to the reaction energy. These contributions are also listed in Table 3, together with the contributions for all other calculations.

--- Figure 7 ---

--- Table 3 ---

The total low-level effect of the protein is between -(4.5~5.0) kcal/mol, but the contributions from individual MM terms differ slightly. In general, the largest MM contribution comes from van der Waals interactions, and they always stabilize O₂ binding. This is similar to the findings for hemerythrin, where van der Waals interactions favored O₂ binding by as much as 6 kcal/mol^{xxiii}. In the present case, the effect is between -2.3 and -4.7 kcal/mol, depending on the method (see Table 3). Direct van der Waals interactions between protein and O₂ are -(3.2~3.5) kcal/mol, evaluated by replacing the original dioxygen atom type with a new type that lacks van der Waals interactions.

The sum of van der Waals interactions in the deoxy state are attractive, but in the oxy state this sum extends over two more atoms (the dioxygen atoms). These new van der Waals interactions stabilize the oxy state relative to the deoxy state, All atoms outside the model system have attractive van der Waals interaction with oxygen and the closest amino acids (Val272, Leu223 and Phe211) should therefore be most important. Since O₂ does not have

any interactions with the protein in the reactant, any interactions in the product directly affects the binding energy.

It should be kept in mind that the calculated effect depends on the assigned van der Waals parameters for dioxygen. In the present calculation, these parameters are simply transferred from the most similar atom types in the Amber force field (see Computational details). However, even if the oxygen parameters would change by 20 %, the change in binding energy would be ≤ 0.6 kcal/mol.

The effect of van der Waals interactions in the QM:MM calculations does not show that the protein can “use” van der Waals interactions to preferentially stabilize the oxy state. The sum of van der Waals interactions between O₂ and its surroundings is expected to be similar in solution (reactant) and protein (product). However, the results show that these interactions are significant and cannot be neglected in calculations of binding energies since they are indirectly included in the aqueous reference state. The situation is different for the proceeding steps of the enzymatic reaction. The van der Waals interactions between O₂ and protein should then be more or less constant and the neglect of these interactions in active-site models should not affect calculated reaction energies.

In contrast to the van der Waals interactions, the Coulomb contributions to O₂ binding in the mechanical embedding calculations do not come from interactions with dioxygen since the point charges of the O₂ molecule are artificially set to zero (see Computational details). To get a more accurate estimate of the electrostatic interactions between dioxygen and surrounding protein, end-on and side-on structures originally optimized using the ONIOM-ME method are reoptimized using the ONIOM-EE scheme.

First looking at the end-on structure, the protein effect calculated with EE is -10.8 kcal/mol, compared to -10.3 kcal/mol for ME (see Table 2). For EE calculations, one part of the electrostatic effect is evaluated quantum mechanically and the other classically, and a division between high and low-level effects is not straightforward. However, it is easy to remove bonded and van der Waals terms and count classical Coulomb effects together with quantum ones. If this is done, the “high-level+Coulomb” effect is -5.8 kcal/mol for EE compared to -6.9 kcal/mol for ME.

There are no major changes in geometry when going from ME to EE, as can be seen in Fig. 8. Mulliken charge populations are also rather stable, with the charge on Fe goes from 1.19 in ME to 1.28 in EE, but a similar increase is observed also in the reactant. **Note that the charge population of iron is a poor indication of the oxidation state^{lviii}.** The total oxygen charge remains constant at -0.53. The observed changes when going from ME to EE are thus

relatively small. It therefore seems that polarization of the active site by surrounding amino acids, an effect that can only be taken into account with EE, has a limited role in determining the binding energy, at least in the present system. In the hemerythrin study, polarization of the active site was found to be important^{xxiii}. That effect was attributed to a carboxylate group that accepted a hydrogen bond from one of the histidine ligands coordinating iron. In isopenicillin N synthase, the histidines form hydrogen bonds with backbone carbonyls and the expected effect of polarization should therefore be smaller.

--- Figure 8 ---

Finally looking at the low-level bonded contributions, they are in general relatively small and no detailed investigation of these effects has been made. However, for the calculations using the five-coordinate X-ray structure, there are several large non-bonded MM contributions, together with very significant Coulomb contributions (see Table 3). These contributions reflect the fact that reactant and product structures represent different MM minima. When using the five-coordinate X-ray structure, it is necessary to move Val272 away from the binding site in order to fit O₂ into the optimized reactant structure. This gives a difference in orientation of Val272 between reactant and product, which is the main cause of the large individual MM components in these calculations. Still, the relative energies of different Val272 orientations are very similar (within 0.2 kcal/mol). The total MM effect is therefore similar to the six-coordinate X-ray structure.

4. Conclusions

In the present study, inclusion of the entire protein in an ONIOM QM/MM protein model has a significant effect on O₂ binding energies in isopenicillin N synthase. Compared to an active-site model, O₂ binding is stabilized by 8-10 kcal/mol, depending on the coordination mode. It is concluded that the active-site model allows too large rearrangements of the iron coordination sphere, which artificially stabilizes the reactant structure where iron is five-coordinate. The magnitude of this effect is almost 6 kcal/mol for end-on binding of dioxygen. The effect for side-on binding is smaller since the flexibility of the active-site model also stabilizes the seven-coordinate side-on complex. Protein effects evaluated at the molecular mechanics level further stabilize O₂ binding by 4-5 kcal/mol. Van der Waals interactions between O₂ and the surrounding enzyme contribute 3-4 kcal/mol, but this value

of course depends on the choice of van der Waals parameters. In general, individual low-level contributions show fluctuations in the 1 kcal/mol range.

Compared to the study of O₂ binding in hemerythrin^{xxiii}, the total protein effect is similar (approximately 10 kcal/mol). In both cases, van der Waals interactions with O₂ give important contributions. However, polarization of the active site is less important in isopenicillin N synthase compared to hemerythrin.

During the preparation of the present manuscript we became aware of a study of the properties of NO and O₂ bound to IPNS^{lix}. In these calculations, O₂ binding is close to thermoneutral in the active-site model, which is different from our results. The main cause of the difference is the different amount of HF exchange used in the two studies (B3LYP (20 % HF exchange) compared to BP86+10 % HF exchange).

Calculations of O₂ binding energies in non-heme iron systems using active-site DFT models are problematic. The present results suggest that the main reason for these problems is the limited model size, rather than significant problems with the method itself. A decrease in the amount of HF exchange in the B3LYP functional to 15 % is still required to achieve favorable O₂ binding, and since no experimental estimates of binding energies are available, no solid claims can be made regarding the optimum functional.

The present study illustrates the benefits of a QM/MM treatment of protein systems. Dioxygen binding is a reaction step where large protein effects can be expected since all modeled protein interactions with O₂ directly influence the reaction energy. Still, including the surrounding protein in an MM description can be generally used to improve optimized structures with limited additional computational cost.

Acknowledgements. We thank Drs. Rajeev Prabhakar, Thom Vreven and Tomasz Borowski for helpful suggestions, and Momotaro Takeda for help with calculations. One of the authors (ML) acknowledges a Fukui Institute for Fundamental Chemistry Fellowship. The present work was in part supported by a CREST (Core Research for Evolutional Science and Technology) grant in the Area of High Performance Computing for Multi-scale and Multi-physics Phenomena from the Japan Science and Technology Agency (JST). The use of computational resources at the Fukui Institute for Fundamental Chemistry is also acknowledged.

Supporting Information. List of all optimized structures with their energies (in Hartrees) and the energy components for ONIOM when applicable. **Structure files with coordinates of all the optimized structures (in XYZ format). Detailed population analysis of the oxy state (septet, side-on binding).** This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- ⁱ Costas, M.; Mehn, M. P.; Jensen, M. P.; Que, L., Jr. *Chem. Rev.* **2004**, *104*, 939-986.
- ⁱⁱ Bassan, A.; Blomberg, M. R. A.; Borowski, T.; Siegbahn, P. E. M. *J. Inorg. Biochem.* **2006**, *100*, 727-743.
- ⁱⁱⁱ Bassan, A.; Blomberg, M. R. A.; Siegbahn, P. E. M. *Chem. Eur. J.* **2003**, *9*, 106-115.
- ^{iv} Furche, F.; Perdew, J. P. *J. Chem. Phys.* **2006**, *124*, 044103.
- ^v Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *124*, 224105.
- ^{vi} Jensen, K. P.; Roos, B. O.; Ryde, U. *J. Chem. Phys.* **2007**, *126*, 014103.
- ^{vii} Siegbahn, P. E. M. *J. Inorg. Biochem.* **2006**, *11*, 316-323.
- ^{viii} Lundberg, M.; Siegbahn, P. E. M. *J. Comp. Chem.* **2005**, *26*, 661-667.
- ^{ix} Quiñonero, D.; Musaev, D. G.; Morokuma, M. *Inorg. Chem.* **2003**, *42*, 8449-8455.
- ^x Swart, M.; Groenhof, A. R.; Ehlers, A. W.; Lammertsma, K. *J. Phys. Chem. A* **2004**, *108*, 5479-5483.
- ^{xi} Rong, C.; Lian, S.; Yin, D.; Shen, B.; Zhong, A.; Bartolotti, L.; Liu, S. *J. Chem. Phys.* **2006**, *125*, 174102.
- ^{xii} Strickland, N.; Harvey, J. N. *J. Phys. Chem. B* **2007**, *111*, 841-852.
- ^{xiii} Bassan, A.; Borowski, T.; Siegbahn, P. E. M. *Dalton Trans.* **2004**, *20*, 3153-3162.
- ^{xiv} Reiher, M.; Salomon, O.; Hess, B. A. *Theor. Chem. Acc.* **2001**, *107*, 48-55.
- ^{xv} Reiher, M. *Inorg. Chem.* **2002**, *41*, 6928-6935.
- ^{xvi} Schenk, G.; Pau, M. Y. M.; Solomon, E. I., *J. Am. Chem. Soc.* **2004**, *126*, 505-515.
- ^{xvii} Field, M. J.; Bash, P. A.; Karplus, M. *J. Comp. Chem.* **1990**, *11*, 700-733.
- ^{xviii} Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580-10594.
- ^{xix} Murphy, R. B.; Philipp, D. M.; Friesner, R. A. *J. Comp. Chem.* **2000**, *21*, 1442-1457.
- ^{xx} Kollman, P. A.; Kuhn, B.; Peräkylä, M. *J. Phys. Chem. B* **2002**, *106*, 1537-1542.
- ^{xxi} Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery, J. A., Jr.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comp.* **2006**, *2*, 815-826.
- ^{xxii} Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. *Chem. Rev.* **2006**, *106*, 3210-3235.
- ^{xxiii} Wirstam, M.; Lippard, S. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2003**, *125*, 3980-3987.
- ^{xxiv} Hoffman, M.; Khavrutskii, I. V.; Musaev, D. G.; Morokuma, K. *Int. J. Quant. Chem.* **2004**, *99*, 972-980.
- ^{xxv} Nemukhin, A. V.; Grigorenko, B. L.; Topol, I. A.; Burt, S. K. *Int. J. Quant. Chem.* **2006**, *106*, 2184-2190.
- ^{xxvi} Kovaleva, E. G.; Lipscomb, J. D. *Science* **2007**, *316*, 453-457.
- ^{xxvii} Katona, G.; Carpentier, P.; Nivière, V.; Amara, P. Adam, V.; Ohana, J.; Tsanov, N.; Bourgeois, D. *Science* **2007**, *316*, 449-453.
- ^{xxviii} Baldwin, J. E.; Bradley, M. *Chem. Rev.* **1990**, *90*, 1079-1088.
- ^{xxix} Schenk, W. A. *Angew. Chem. Int. Ed.* **2000**, *39*, 3409-3411.
- ^{xxx} Andersson, I.; Terwisscha van Scheltinga, A. C.; Valegård, K. *Cell. Mol. Life Sci.* **2001**, *58*, 1897-1906.
- ^{xxxi} Roach, P. L.; Clifton, I. J.; Fülöp, V.; Harlos, K.; Barton, G. J.; Hajdu, J.; Andersson, I.; Schofield, C. J.; Baldwin, J. E. *Nature* **1995**, *375*, 700-704.
- ^{xxxii} Roach, P. L.; Clifton, I. J.; Hensgens, C. M. H.; Shibata, N.; Schofield, C. J.; Baldwin, J. E. *Nature* **1997**, *387*, 827-830.
- ^{xxxiii} Burzlaff, N. I.; Rutledge, P. J.; Clifton, I. J.; Hensgens, C. M. H.; Pickford, M.; Adlington, R. M.; Roach, P. L.; Baldwin, J. E. *Nature* **1999**, *401*, 721-724.

- ^{xxxiv} Koehntop, K. D.; Emerson, J. P.; Que, L., Jr. *J. Biol. Inor. Chem.* **2005**, *10*, 87-93.
- ^{xxxv} Chen, V. J.; Orville, A. M.; Harpe, M. R.; Frolik, C. A.; Surerus, K. K.; Münck, E.; Lipscomb, J. D. *J. Biol. Chem.* **1989**, *264*, 21677-21681.
- ^{xxxvi} Scott, R. A.; Wang, S. K.; Eidsness, M. K.; Kriauciunas, A.; Frolik, C. A.; Chen, V. J. *Biochemistry* **1992**, *31*, 4596-4601.
- ^{xxxvii} Solomon, E. I.; Brunold, T. C.; Davis, M. I.; Kemsley, J. N.; Lee, S.-K.; Lehnert, N.; Neese, F.; Skulan, A. J.; Yang, Y. S.; Zhou, J. *Chem. Rev.* **2000**, *100*, 235-349.
- ^{xxxviii} Karlsson, A.; Parales, J. V.; Parales, R. E.; Gibson, D. T.; Eklund, H.; Ramaswamy, S. *Science* **2003**, *299*, 1039-1042.
- ^{xxxix} Sato, N.; Uragami, Y.; Nishizaki, T.; Takahashi, Y.; Sasaki, G.; Sugimoto, K.; Nonaka, T.; Masai, E.; Fukuda, M.; Senda, T. *J. Mol. Biol.* **2002**, *321*, 621-636.
- ^{xl} Zhang, Z.; Ren, J.; Harlos, K.; McKinnon, C. H.; Clifton, I. J.; Schofield, C. J. *FEBS Lett.* **2002**, *517*, 7-12.
- ^{xli} Maseras, F.; Morokuma, K. *J. Comp. Chem.* **1995**, *16*, 1170-1179.
- ^{xlii} Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *J. Mol. Str. (Theochem)* **1999**, *461*, 1-23.
- ^{xliii} Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648-5652.
- ^{xliv} Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.
- ^{xlvi} Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, 2004.
- ^{xlvi} Yoshizawa, K.; Shiota, Y. *J. Am. Chem. Soc.* **2006**, *128*, 9873-9881.
- ^{xlvii} Kwiecien, R. A.; Khavrutskii, I. V.; Musaev, D. G.; Morokuma, K.; Banerjee, K. R.; Paneth, P. *J. Am. Chem. Soc.* **2006**, *128*, 1287-1292.
- ^{xlviii} Prabhakar, R.; Vreven, T.; Frisch, M. J.; Morokuma, K.; Musaev, D. G. *J. Phys. Chem. B* **2006**, *110*, 13608-13613.
- ^{xlvi} Noodleman, L.; Case, D. A. *Adv. Inorg. Chem.* **1992**, *38*, 423-470.
- ^l Klähn, M.; Braun-Sand, S.; Rosta, E.; Warshel, A. *J. Phys. Chem. B* **2005**, *109*, 15645-15650.
- ^{li} Blomberg, M. R. A.; Siegbahn, P. E. M. *Theor. Chem. Acc.* **1997**, *97*, 72-80.
- ^{lii} Borowski, T.; Bassan, A.; Siegbahn, P. E. M. *Inor. Chem.* **2004**, *43*, 3277-3291.
- ^{liii} Amzel, M. P. *Proteins* **1997**, *28*, 144-149.
- ^{liv} Mehn, M. P.; Fujisawa, K.; Hegg, E. L.; Que, L., Jr. *J. Am. Chem. Soc.* **2003**, *125*, 7828-7842.
- ^{lv} Kriauciunas, A.; Frolik, C. A.; Hassell, T. C.; Skatrud, P. L.; Johnson, M. G.; Holbrok, N. I.; Chen, V. J. *J. Biol. Chem.* **1991**, *266*, 11779-11788.
- ^{lvi} Lundberg, M.; Morokuma, K. unpublished.

^{lvii} Torrent, M.; Vreven, T.; Musaev, D. G.; Morokuma, K.; Farkas, Ö. ; Schlegel, H. B. *J. Am. Chem. Soc.* **2002**, 124, 192-193.

^{lviii} Siegbahn, P. E. M.; Crabtree R. H. *J. Am. Chem. Soc.* **1997**, 119, 3103-3113.

^{lix} Brown, C. D.; Neidig, M. L.; Neibergall, M. B.; Lipscomb, J. B.; Solomon E. I. *J. Am. Chem. Soc. ASAP article*, <http://dx.doi.org/10.1021/ja071364v>.

Table 1. Stationary structures and energies for O₂ binding to Fe in the active site of IPNS, calculated using the active-site model. Structures are listed in order of stability. Binding energies are calculated relative to free O₂ and a high-spin Fe(II) complex (quintet). Atom labels are given in Fig. 3.

Spin /Coordination	Mulliken spin density			Distance (Å)			ΔE	ΔH _{corr}	-TΔS ^a	Solv. ^b	ΔG
	Fe	O1	O2	Fe-O1	Fe-O2	O1-O2					
Septet/side-on	4.09	0.73	0.64	2.09	2.13	1.31	2.7	0.8	11.7	-4.7	10.6
Septet/end-on	4.11	0.63	0.71	2.06	2.69	1.30	5.2	0.2	12.0	-6.0	11.3
Quintet/end-on	4.09	-0.12	-0.50	1.91	2.95	1.31	5.8	0.2	13.1	-5.7	13.4
Quintet/side-on	4.11	-0.39	-0.28	2.08	2.16	1.31	7.2	0.8	11.1	-5.5	13.6
Triplet/end-on	3.30	-0.76	-0.55	1.94	2.75	1.28	3.6	1.7	14.8	-5.5	14.7

^a The entropy cost of trapping free O₂ is 14.5 kcal/mol. ^b -3.9 kcal/mol comes from the solvation of O₂.

Table 2. Structures and energies for O₂ binding to Fe in the active site of IPNS, calculated using different ONIOM setups. Results for the corresponding active-site models are included to facilitate comparisons. 6-coord indicates that calculations use the X-ray structure where Fe is originally six-coordinate, while 5-coord indicates that calculations use the X-ray structure where Fe is originally five-coordinate. Atom labels are shown in Fig. 3.

X-ray structure.	Method ^a	Mulliken spin			Distance (Å)				Protein effect (kcal/mol)		ΔE (kcal/mol)	
		Fe	O1	O2	Fe- O1	Fe- O2	O1- O2	Q M	MM	Total	Active- site	ONIOM
Side-on septet structures												
	Act. site	4.09	0.73	0.64	2.09	2.13	1.31	-	-	-	2.7	-
6-coord	ME	4.08	0.68	0.63	2.07	2.13	1.32	- 3.2	-4.9	-8.1	2.7	-5.4
5-coord	ME	4.09	0.67	0.66	2.07	2.15	1.31	- 2.9	-4.8	-7.7	2.7	-5.0
6-coord	EE	4.14	0.71	0.66	2.09	2.17	1.31	NA	NA	-7.6	2.7	-4.9
End-on septet structures												
	Act. site	4.11	0.63	0.71	2.06	2.69	1.30	-	-	-	5.2	-
6-coord	ME	4.11	0.58	0.67	2.05	2.87	1.31	- 5.8	-4.5	-10.3	5.2	-5.2
5-coord	ME	4.13	0.57	0.68	2.04	2.79	1.31	-	-5.0	-10.5	5.2	-5.3

								5.5				
6-coord	EE	4.19	0.60	0.67	2.08	2.91	1.31	NA	NA	-10.8	5.2	-5.6

^a ME: Mechanical embedding. EE: Electronic embedding (see text for details).

Table 3. Low-level protein contributions to the reaction energy for O₂ binding, calculated using ONIOM. For ONIOM-EE calculations, it is not straightforward to quantify low-level electrostatic effects since Coulomb interactions are calculated both classically and semi-classically.

		MM contributions (kcal/mol)							
X-ray structure	Embedding ^a	Coulomb	Van der Waals				Out-of-plane	Total MM	
			Stretch	Bend	Torsion				
Side-on septet structures									
6-coord	ME	-1.6	-2.3	0.2	-0.1	-1.2	0.1	-4.9	
5-coord	ME	-5.0	-2.3	0.2	1.7	0.8	-0.2	-4.8	
6-coord	EE	NA	-4.4	-0.0	-0.2	-1.0	0.1	NA	
End-on septet structures									
6-coord	ME	-1.2	-2.9	0.2	0.1	-0.8	0.1	-4.5	
5-coord	ME	-3.4	-2.6	0.2	1.3	-0.1	-0.3	-5.0	
6-coord	EE	NA	-4.7	0.1	-0.2	-0.3	0.1	NA	

^a ME: Mechanical embedding. EE: Electronic embedding (see text for details).

Figure captions

Figure 1. Enzymatic reaction catalyzed by isopenicillin N synthase (IPNS). O₂ is used to transform the substrate δ -(L- α -aminoadipoyl)-L-cysteinyl-d-valine (ACV) to isopenicillin N.

Figure 2. Schematic illustrations of the active site in IPNS. Before O₂ binding iron is five-coordinated with a square pyramidal geometry (left). Dioxygen can bind in end-on configuration (only one oxygen coordinating iron) and this creates an octahedral iron site (middle). Side-on binding of O₂, where both oxygens coordinate to iron, is also considered (right).

Figure 3. Active-site model used in the present studies. The same atom selection also makes up the “model” system in the ONIOM calculations. The ONIOM protein system is fully optimized but in the active-site model, atoms marked with **X** have frozen Cartesian coordinates from X-ray structure.

Figure 4. Free energies for binding of O₂ to Fe in the active site of IPNS, including protein effects calculated using the 6-coord. X-ray structure. All calculations use enzyme + free O₂ as reference state and results include zero point, thermal and solvent corrections from an active-site model.

Figure 5. Geometry of the ONIOM model system (blue) compared to the geometry of the active-site model (silver) for deoxy (left) and oxy (right) structures. Note the different relaxation of substrate and His270 in the two states. ONIOM calculations are performed with ME and use the X-ray structure where Fe is six-coordinate. Structural views are designed by first aligning all heavy atoms, followed by a transformation so that the iron atoms overlap. Selected distances and spin populations are listed in Table 2.

Figure 6. Comparison of ONIOM model geometries for the oxy state calculated using two different X-ray structures, one where iron is originally six-coordinate (blue) and one where iron is originally five-coordinated (white). The structural view is designed by first aligning all heavy atoms, followed by a transformation so that the iron atoms overlap. The RMSD for iron and its coordinating atoms (O₂ not included) is 0.06 Å. Selected distances and spin populations are listed in Table 2.

Figure 7. Low-level protein effects on to the reaction energy for O₂ binding, calculated using ONIOM-ME and the 6-coordinate X-ray structure. The total effect is shown together with the individual MM components.

Figure 8. Comparison of geometries from an optimization using ONIOM-ME (blue) and an optimization using ONIOM-EE (bronze). Structural views are designed by first aligning all heavy atoms, followed by a transformation so that the iron atoms overlap. The RMSD for iron and its coordinating atoms (O₂ not included) is 0.04 Å. Selected distances and spin populations are listed in Table 2.

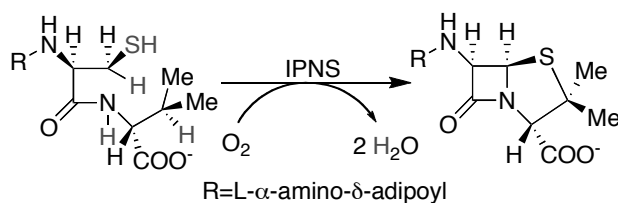


Figure 1. Enzymatic reaction catalyzed by isopenicillin N synthase (IPNS). O_2 is used to transform the substrate δ -(L- α -aminoadipoyl)-L-cysteinyl-D-valine (ACV) to isopenicillin N.

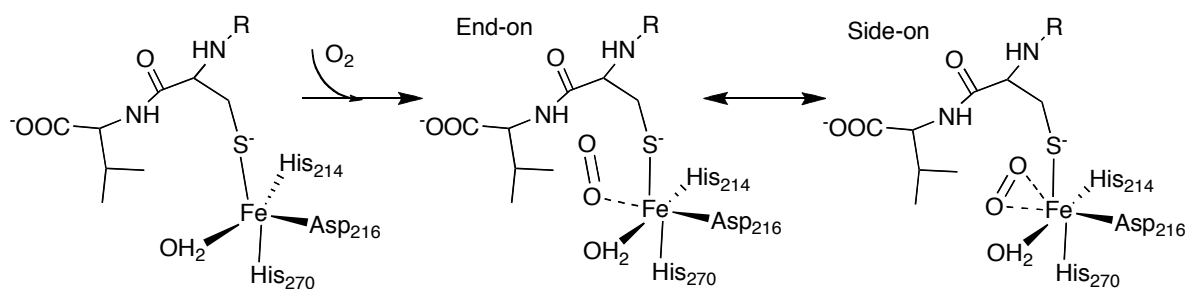


Figure 2. Schematic illustrations of the active site in IPNS. Before O_2 binding iron is five-coordinated with a square pyramidal geometry (left). Dioxygen can bind in end-on configuration (only one oxygen coordinating iron) and this creates an octahedral iron site (middle). Side-on binding of O_2 , where both oxygens coordinate to iron, is also considered (right).

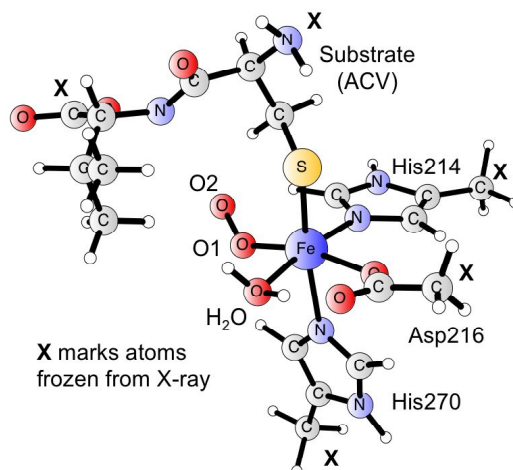


Figure 3. Active-site model used in the present studies. The same atom selection also makes up the “model” system in the ONIOM calculations. The ONIOM protein system is fully optimized but in the active-site model, atoms marked with X have frozen Cartesian coordinates from X-ray structure.

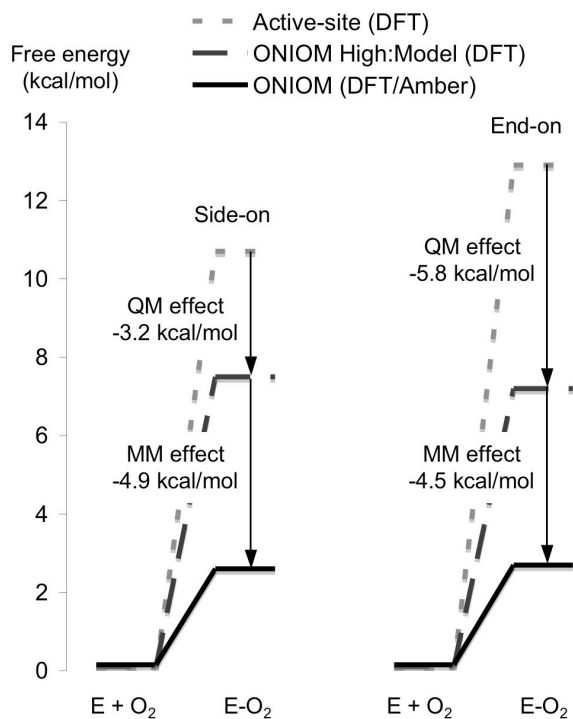


Figure 4. Free energies for binding of O_2 to Fe in the active site of IPNS, including protein effects calculated using the 6-coord. X-ray structure. All calculations use enzyme + free O_2 as reference state and results include zero point, thermal and solvent corrections from an active-site model.

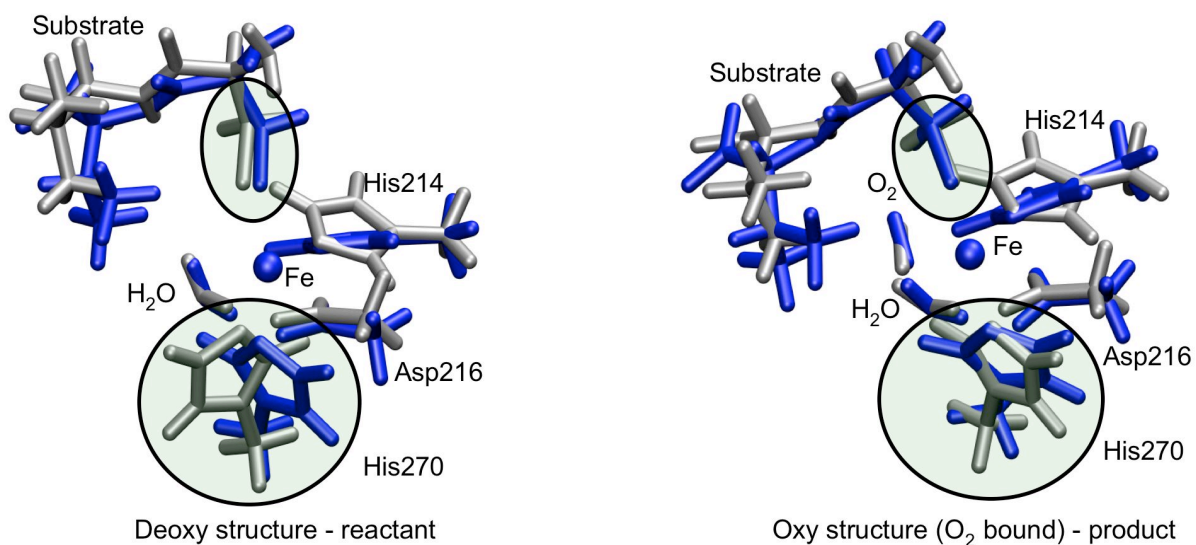


Figure 5. Geometry of the ONIOM model system (blue) compared to the geometry of the active-site model (silver) for deoxy and oxy structures. Note the different relaxation of substrate and His270 in the two states. ONIOM calculations are performed with ME and use the X-ray structure where Fe is six-coordinate. Structural views are designed by first aligning

all heavy atoms, followed by a transformation so that the iron atoms overlap. Selected distances and spin populations are listed in Table 2.

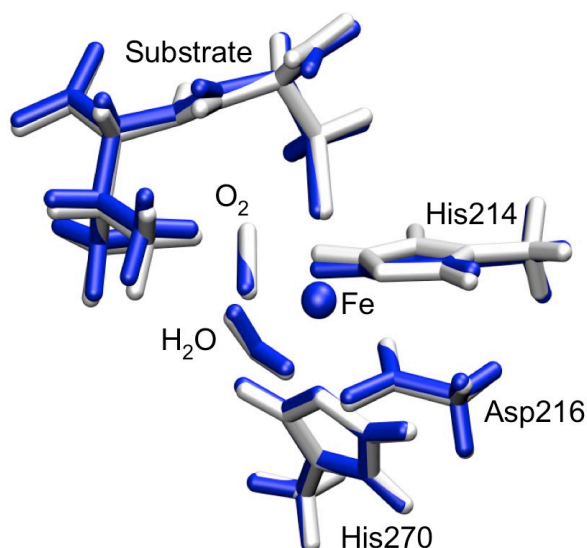


Figure 6. Comparison of ONIOM model geometries for the oxy state calculated using two different X-ray structures, one where iron is originally six-coordinate (blue) and one where iron is originally five-coordinated (white). The structural view is designed by first aligning all heavy atoms, followed by a transformation so that the iron atoms overlap. The RMSD for iron and its coordinating atoms (O_2 not included) is 0.06 \AA . Selected distances and spin populations are listed in Table 2.

Figure 7. Low-level protein effects on to the reaction energy for O_2 binding, calculated using ONIOM-ME and the 6-coordinate X-ray structure. The total effect is shown together with the individual MM components.

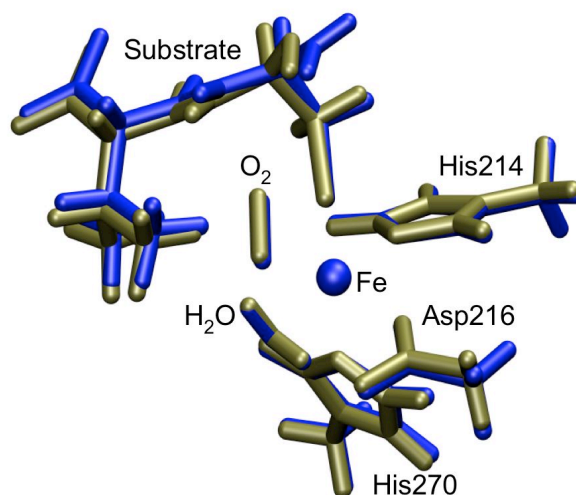


Figure 8. Comparison of geometries from an optimization using ONIOM-ME (blue) and an optimization using ONIOM-EE (bronze). Structural views are designed by first aligning all heavy atoms, followed by a transformation so that the iron atoms overlap. The RMSD for iron and its coordinating atoms (O_2 not included) is 0.04 Å. Selected distances and spin populations are listed in Table 2.