# Dalton Transactions

## Accepted Manuscript

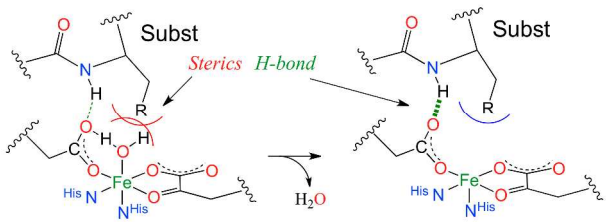

This is an *Accepted Manuscript*, which has been through the RSC Publishing peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, which is prior to technical editing, formatting and proof reading. This free service from RSC Publishing allows authors to make their results available to the community, in citable form, before publication of the edited article. This *Accepted Manuscript* will be replaced by the edited and formatted *Advance Article* as soon as this is available.

To cite this manuscript please use its permanent Digital Object Identifier (DOI®), which is identical for all formats of publication.

More information about *Accepted Manuscripts* can be found in the **Information for Authors**.

Please note that technical editing may introduce minor changes to the text and/or graphics contained in the manuscript submitted by the author(s) which may alter content, and that the standard **Terms & Conditions** and the **ethical guidelines** that apply to the journal are still applicable. In no event shall the RSC be held responsible for any errors or omissions in these *Accepted Manuscript* manuscripts or any consequences arising from the use of any information contained in them.




Factors that lead to coordinative unsaturation in nonheme-Fe(II) enzymes include sterics, facial triad carboxylate H-bonding, and strong cosubstrate donor ligation.

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

# First- and second-sphere contributions to Fe(II) site activation by cosubstrate binding in nonheme Fe enzymes

**Kenneth M. Light,[a] John A. Hangasky,[b] Michael J. Knapp[b] and Edward I. Solomon*[a]**

***Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX***

**DOI: 10.1039/b000000x**

**Nonheme Fe(II) enzymes exhibit a general mechanistic strategy where binding all cosubstrates opens a coordination site on the Fe(II) for $O_2$ activation. This study shows that strong-donor ligands, steric interactions with the substrate and second-sphere H-bonding to the facial triad carboxylate allow for five-coordinate site formation in this enzyme superfamily.**

Non-heme Fe(II) (NHFe(II)) enzymes generally utilize a redox-active cosubstrate and an Fe(II) center to activate $O_2$ for reaction with organic species.[1] These enzymes can be broadly divided into five classes based on the cosubstrate employed as an electron source for $O_2$ activation. Members of the α-ketoglutarate (αKG)-dependent dioxygenase class utilize a bidentate αKG ligand as a source of two electrons needed for the reaction.[1a] For extradiol dioxygenases, the substrates themselves are bidentate catecholate ligands that provide the two electrons for $O_2$ activation.[1b] Members of the pterin-dependent and Rieske dioxygenase classes use non-ligated pterins and $Fe_2S_2$ Rieske clusters, respectively, as electron sources.[1c] A fifth class of NH Fe(II) enzymes acts on redox-inactive substrates which bind directly to the Fe(II) center and in some cases require an additional cosubstrate to supply electrons. This eclectic class includes isopenicillin *N* synthase (IPNS),[1d] 1-aminocyclopropane-1-carboxylic acid oxidase (ACCO)[1e] and (*S*)-2-hydroxypropylphosphonic acid epoxidase (HppE),[1f] among others. Members of this superfamily bind to Fe(II) through a 2-His, 1-Asp/Glu facial triad of protein-derived ligands and additional $H_2O$'s to maintain a 6-coordinate (6C) Fe(II) center. This loses a $H_2O$ ligand to become 5-coordinate (5C) and reactive towards $O_2$ only when all cosubstrates are present, though what induces this water ligand to dissociate has not been well-explored. In this study we use computational methods on an enzyme system where various possible contributions to ligand dissociation can be evaluated to understand how NHFe(II) enzymes become active once all necessary components are in place.

Hypoxia-inducible factor 1α (HIF-1α) is the master controller of the human cellular hypoxic response.[2] HIF-asparginyl hydroxylase (identified formerly as FIH-1) is an αKG-dependent NHFe(II) enzyme that inactivates HIF-1α by hydroxylating Asn803 in the C-terminal transaction domain (CAD) of HIF-1α, preventing the HIF-αβ homodimer from binding to transcriptional co-activator p300 and activating genes involved with increasing

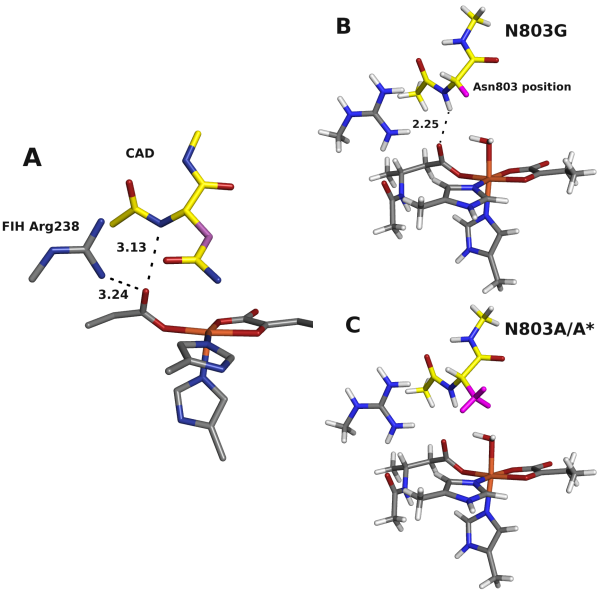



Figure 1. A.) View of the FIH-1 active site showing H-bonds to the facial triad carboxylate. CAD Fragment is in yellow with hydroxylation target in magenta. B.) The N803G model for analyzing H-bonding effects only. C.) The N803A/A* model for including steric effects. Atoms in magenta replace the side chain of N803.

cellular $O_2$ levels.[3] The active site of αKG/CAD-bound FIH-1 as determined by crystallography is shown in Figure 1A where notable second-sphere interactions are the hydrogen bonds (H-bonds) from the backbone amide of Asn803 of CAD and from the sidechain of FIH Arg238 to the non-ligated O of the facial triad carboxylate.[3] It has been previously shown in a related enzyme system taurine/αKG dioxygenase (TauD)[4] that H-bonding between this carboxylate O and the ligated $H_2O$ stabilizes the Fe-$OH_2$ bond, which is destabilized by the strong electron-donating character of the αKG ligand. From DFT calculations[4] the ΔG for binding $H_2O$ to an αKG-bound NHFe(II) site is ≈+8 kcal/mole in the absence of this H-bond, but is ≈-1 kcal/mole if this H-bond is present. Magnetic circular dichroism (MCD) spectra of Fe(II)- and Fe(II)/αKG-bound forms of TauD[4] and FIH[5a] show an increased splitting of the $e_g$ d orbitals upon binding of αKG, which has been ascribed to the weakened Fe-$OH_2$ bond. Upon binding substrate, H-bonds from CAD Asn803 and FIH-1 Arg238 may disrupt the carboxylate–$H_2O$ H-bond and therefore promote water loss upon CAD binding as is observed in the MCD

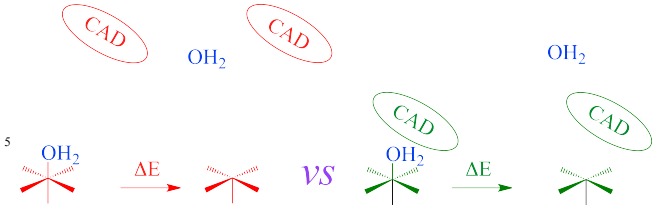



Figure 2. Computational strategy for determining contributions to $H_2O$ loss.

spectrum of Fe(II)/αKG/CAD-bound FIH, which showed a mixture of 5C and 6C forms.[5a] The β-carbon of CAD Asn803 (colored magenta in Figure 1) is the target for hydroxylation and is poised directly above the Fe(II) active site. This group will potentially sterically clash with a coordinated $H_2O$ to further promote loss of this ligand. Due to the presence of steric, H-bonding, and strong electron-donor effects from equatorially bidentate bound αKG, FIH was deemed a good NHFe(II) system for exploring the different possible contributions to the 6C→5C conversion in activation of NHFe(II) to react with $O_2$.

To explore the relative contributions of sterics and second-sphere H-bonding to water loss, a computational approach to correlate the thermodynamics of water elimination with various effects was undertaken as outlined in Figure 2. DFT calculations were performed on both $H_2O$-coordinated 6C and $H_2O$-dissociated 5C geometries of FIH wherein the CAD substrate was either present in its crystallographic location (Figure 2, right) or *ca.* 40 Å away (Figure 2, left) where it could not interact with the FIH active site. Initial calculations, which included the full Asn803 side chain, indicated that while the process of coordinated water loss for the CAD-unbound form has a ΔG≈0 kcal/mole, water loss from the CAD-bound form is favorable with a ΔG≈-7 kcal/mole. (See Supporting) The ΔE for removal of water for the CAD-bound form was also ≈5 kcal/mole more favorable than that of the CAD-unbound form. (See Supporting)

In order to separate the contributions of substrate sterics and H-bonding, three CAD models were evaluated. The side chain of CAD-Asn803 was changed to a Gly residue as shown in Figure 1B, (N803G) for the inclusion of H-bonding effects with minimal steric interaction. Replacement of CAD-Asn803 with an Ala residue (N803A) retains the H-bonding while introducing the steric interaction between the coordinated water and the methylene group of Asn803 (Figure 1C). The N803A model still allows for flexibility of the methyl group, so to further model the anchoring of CAD Asn803 (by H-bonding with FIH Arg238 and Gln239) a third model, N803A*, was used in which the methyl group is constrained in its position during geometry optimization. The changes in energy (E), enthalpy (H), and Gibbs free energy (G) upon loss of the water ligand are given in the Supporting Information. For all three models of the CAD substrate ΔG for $H_2O$ loss is ≈ 0 when CAD is unbound and is from ≈ -4 to -9 kcal/mole when CAD is bound. This reproduces the behavior of the CAD model with the full Asn803 side chain and correctly predicts that coordinated water will dissociate upon CAD binding. Figure 3 compares the relative energies of the 6C ($H_2O$ bound, bottom) and 5C ($H_2O$ lost, top) CAD-bound (green) and unbound (red) forms for the three models, and decouples the effects of H-bonding from those of sterics on the stability of these

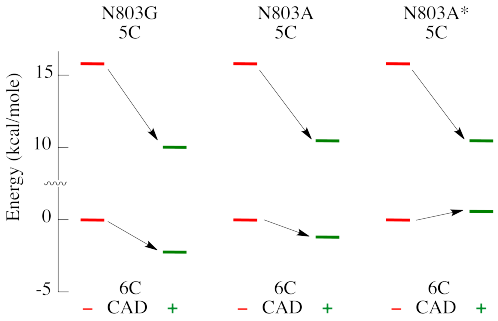



Figure 3. DFT Thermodynamic study of water loss. Forms listed in red are for CAD-unbound models whereas those in green are for CAD-bound models. All energies are in kcal/mole.

forms. For each group of models the energy of the 6C CAD-unbound form is set to 0 kcal/mole. For the N803G (H bond) model, binding CAD stabilizes both the 6C and 5C forms, but stabilizes the latter by 5.8 kcal/mole *vs* 2.6 kcal/mole for the 6C form. This result indicates that the H bond between the substrate amide NH and the carboxylate O stabilizes the 5C form relative to the 6C form (by ≈3 kcal/mole *in silico*), shifting the equilibrium towards uncoordinated water. This increased stabilization is depicted in Figure 4 by a weaker amide-carboxylate H bond for the 6C form (left) relative to the strong H bond in the 5C form (right). This contribution solely reflects the H-bonding interaction, as steric effects between the substrate and water are minimal.

While the stabilization of 5C FIH upon CAD binding is effectively constant for all the models in Figure 3, top, the stabilization of 6C FIH decreases from N803G to N803A, until binding CAD to 6C N803A* results in destabilization (Figure 3, bottom). This trend reflects the increase in steric clash between the CAD sidechain and the coordinated water, also depicted in Figure 4, left (light–red curves). Thus, the energetic effects of the H bond between the CAD backbone amide NH and the non-ligated O of the facial triad carboxylate and the steric interaction of the substrate with the coordinated $H_2O$ are comparable in promoting loss of $H_2O$ upon substrate binding, and are similar in magnitude to the contribution of the strong donor ligand αKG to $H_2O$ release (i.e. ≈9 kcal/mole).[4]

From this and prior studies, three contributions to $H_2O$ loss in FIH-1 have been identified: Steric interactions, which destabilize the 6C form, H-bonding to the non-ligated facial triad carboxylate O, which stabilizes the 5C form, and the strong donor properties of the αKG ligand, which also stabilize the 5C form. These findings can be extended to the other members of the αKG-dependent class, where all three contributions are present. While for many αKG -dependent enzymes the substrate cannot form an H bond to the facial triad carboxylate, this H bond is provided by a protein residue from the enzyme itself which stabilizes the 5C Fe(II) site, such as Tyr299 in clavaminate synthase[6] and Thr239 in anthocyanidin snythase.[7]

The three contributions to water dissociation discussed above are also present in the other classes of NHFe(II) enzymes. In addition to αKG -dependent enzymes, classes that involve cosubstrate binding to Fe(II) include the extradiol dioxygenases[1b] and the non-redox-active bound substrate class[1d-f] as described

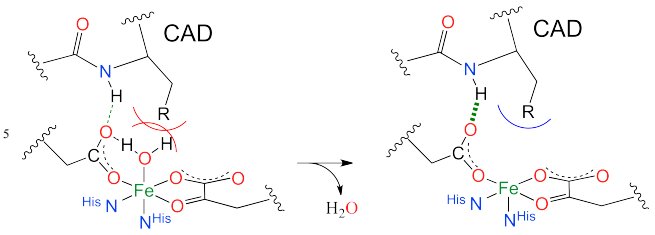



Figure 4. Diagram showing the H-bonding and steric contributions to 5C site formation in FIH-1 upon CAD binding.

above. Like the αKG-dependent enzymes, members of these two classes possess good electron-donating ligands: catecholates for the extradiol dioxygenases,[1b] the δ-(L-α-aminoadipoyl)-L-cysteinyl-D-valine (ACV) substrate thiolate group for IPNS,[1d] amine and carboxylate ligation from ACC for ACCO,[1e] and hydroxyl and phosphonate ligation from HPP for HppE.[1f] Extradiol dioxygenase active sites bind substrate such that the remaining Fe(II) coordination site is *trans* to the facial triad carboxylate, and with no H-bond, the $H_2O$ in this position dissociates to form a 5C site.[1b] In IPNS the substrate ACV, in addition to binding to Fe(II) as a strong donor, sterically clashes with a coordination site to help promote water loss, and a second-sphere enzyme residue H-bonding partner for the facial triad carboxylate is also present in the form of Thr221.[1c] In the case of HppE, HPP does not sterically clash with the remaining coordination site for $H_2O$, and HppE lacks a second-sphere H-bonding partner for the facial triad carboxylate. However, HPP can stabilize a 5C site through its donation of negative charge, and the short distance (≈ 2.8 Å) between the HPP hydroxyl O and facial triad carboxylate O implies the presence of an H-bond, which would prevent the carboxylate from H-bonding to a potential $H_2O$ ligand.[1e] The case of ACCO is more difficult to examine due to the lack of crystallography on the ACC-bound form. However, the resting form of the enzyme does show the presence of a second-sphere H-bonding partner for the facial triad carboxylate in residue Asn216, and the ACC substrate, which binds through an amine and carboxylate group, would donate significant electron density to the Fe(II) to promote a 5C active site.[8]

Pterin-dependent hydroxylases and Rieske dioxygenases, which use non-ligated pterins and $Fe_2S_2$ Rieske clusters respectively as electron sources,[1c] do not generally possess a second-sphere enzyme residue for H-bonding, nor do they possess a cosubstrate that binds to Fe(II). Steric destabilization of 6C Fe(II) must therefore play a significant role in opening up the active site. However, it should be noted that enzymes in these two classes generally have the facial triad carboxylate bind in a bidentate mode once all cosubstrates are present, and this bidentate ligand may provide additional charge density for stabilization of the 5C site.[9]

Spectroscopic studies have shown that NHFe(II) sites favour a 5C geometry only when all cosubstrates are present.[10] A 5C active site plays an important role in catalytic activity as shown by the 200-fold higher rate of Fe-$O_2$ bonding of (tyrosine+pterin)-bound tryptophan hydroxylase relative to the pterin-only-bound form of the enzyme,[11] the ≈ 5000-fold higher rate of reaction for substrate-bound *vs* substrate-free halogenase SyrB2,[12] and the ≈ 17500-fold higher rate of Rieske site oxidation of substrate-bound over substrate-free naphthalene dioxygenase.[13] All of these enzymes have been observed from MCD to show the 6C→5C conversion with substrate binding in combination with binding of cosubstrates (if any) or a reduced Rieske cluster for Rieske-dependent dioxygenases. It is important that the Fe site remain 6C until the cosubstrates have been appropriately bound, as uncoupled reaction with $O_2$ can lead to unwanted side reactions including production of $H_2O_2$[14] and enzyme self-hydroxylation.[15]

In summary, three major contributions to stabilization of the experimentally observed 5C Fe(II) center in FIH-1 upon substrate binding have been identified: Steric interactions between substrate and coordinated $H_2O$ which destabilizes the 6C form, H-bonding between second-sphere enzyme residues and the facial triad carboxylate which stabilizes the 5C form, and the strong electron-donating character of the bound αKG which also stabilizes the 5C form. These contributions are found in all the 5 classes of NHFe(II) enzymes listed above, and although in some instances a given contribution is absent, the other contributions appear to compensate. Nature has produced a flexible system for inducing the opening of the coordination sphere of the Fe(II) active site for $O_2$ reactivity that can be tailored to the properties of the cosubstrates involved in catalysis.

This work was supported by the U.S. National Institute of Health (GM 40392 to E.I.S. and GM 077413 to M.J.K.). K.M.L. and J.A.H. were supported by the Althouse Stanford Graduate Fellowship and the CBI Training Grant (NIH T32-GM008515), respectively.

## Notes and references

[a] *Department of Chemistry, Stanford Univeristy, Stanford, USA. Fax: 650 725 0259 Tel: 650 723 9104; E-mail: Edward.solomon@stanford.edu*
[b] *Department of Chemistry, University of Massachussetts, Amherst, Amherst, USA. Fax: 413 545 4490; Tel: 413 545 4001; E-mail: mknapp@chem.umass.edu*

† Electronic Supplementary Information (ESI) available: Computational modeling details, table of calculated thermodynamic parameters and coordinates for all geometry-optimized structures. See DOI: 10.1039/b000000x/

1 (*a*) Hausinger R., *Crit. Rev. Biochem. Mol.*, 2004, **39**, 21–68; (*b*) E. G. Kovaleva, M. B. Neibergall, S. Chakrabarty and J. D. Lipscomb, *Acc. Chem. Res.*, 2007, **40**, 475–483; (*c*) E. Olsson, K. Teigen, A. Martinez and V. R. Jensen, *Adv. Inor. Chem.*, 2010, **62**, 437–500; (*d*) N. I. Burzlaff, P. J. Rutledge, I. J. Clifton, C. M. H. Hensgens, M. Pickford, R. M. Adlington, P. L. Roach and J. E. Baldwin, *Nature*, 1999, **401**, 721–724; (*e*) A. M. Rocklin, D. L. Tierney, V. Kofman, N. M. W. Brunhuber, B. M. Hoffman, R. E. Christoffersen, N. O. Reich, J. D. Lipscomb and L. Que, Jr., *Proc. Natl Acad. Sci. USA*, 1999, **96**, 7905–7909; (*f*) L. J. Higgins, F. Yan, P. Liu, H. W. Liu and C. L. Drennan, *Nature*, 2005, **437**, 838–844.
2 R. J. Hampton–Smith and D. J. Peet, *Ann. N. Y. Acad. Sci.*, 2009, **1177**, 19–29.
3 J. M. Elkins, K. S. Hewitson, L. A. McNeill, J. F. Seibel, I. Schlemminger, C. W. Pugh, P. J. Ratcliffe and C. J. Schofield, *J. Biol. Chem.*, 2003, **278**, 1802–1806.
4 M. L. Neidig, C. D. Brown, K. M. Light, D. G. Fujimori, E. M. Nolan, J. C. Price, E. W. Barr, J. M. Bollinger, C. Krebs, Jr., C. T. Walsh and E. I. Solomon, *J. Am. Chem. Soc.*, 2007, **129**, 14224–14231.
5 (*a*) K. M. Light, J. A. Hangasky, M. J. Knapp and E. I. Solomon, *J. Am. Chem. Soc.*, 2013, **135**, 9665–9674; (*b*) J. A. Hangasky, C. Y. Taabazuing, M. A. Vallliere and M. J. Knapp, *Metallomics*, 2013, **5**, 287–301.

6 Z. Zhang, J. Ren, D. K. Stammers, J. E. Baldwin, K. Harlos, and C. J. Schofield, *Nature Struct. Biol.*, 2000, **7**, 127–133.

7 R. C. Wilmouth, J. J. Turnbull, R. W. D. Welford, I. J. Clifton, A. G. Prescott, and C. J. Schofield, *Structure*, 2002, **10**, 93–103.

8 Z. Zhang, J.–S. Ren, I. J. Clifton and C. J. Schofield, *Chem. Biol.* 2004, **11**, 1383–1394.

9 (*a*) O. A. Andersen, A. J. Stokka, T. Flatmark and E. Hough, *J. Mol. Biol.*, 2003, **333**, 747–757; (*b*) A. Karlsson, J. V. Parales, R. E. Parales, D. T. Gibson, H. Eklund and S. Ramaswamy, *Science* 2003, **299**, 1039–1042.

10 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 and J. Zhou, *Chem. Rev.*, 2000, **100**, 235–349.

11 M. S. Chow, B. E. Eser, S. A. Wilson, K. O. Hodgson, B. Hedman, P. F. Fitzpatrick and E. I. Solomon, *J. Am. Chem. Soc.*, 2009, **131**, 7685–7698.

12 M. L. Matthews, C. M. Krest, E. W. Barr, F. H. Vaillancourt, C. T. Walsh, M. T. Green, C. Krebs and J. M. Bollinger, Jr., *Biochem.*, 2009, **48**, 4331–4343.

13 M. D. Wolfe, J. V. Parales, D. T. Gibson and J. D. Lipscomb, *J. Biol. Chem.*, 2001, **276**, 1945–1953.

14 J. N. Kemsley, E. C. Wasinger, S. Datta, N. Mitić, T. Acharya, B. Hedman, J. P. Caradonna, K. O. Hodgson and E. I. Solomon, *J. Am. Chem. Soc.*, 2003, **125**, 5677–5686.

15 M. J. Ryle, A. Liu, R. B. Muthukumaran, R. Y. N. Ho, K. D. Koehntop, J. McCracken, L. Que, Jr. and R. P. Hausinger, *Biochem.*, 2003, **42**, 1854–1862.