

Structure and assembly of the diiron cofactor in the heme-oxygenase-like domain of the *N*-nitrosourea-producing enzyme SznF

Molly J. McBride,^a Sarah R. Pope,^b Kai Hu,^{b,l} Jeffrey W. Slater,^a C. Denise Okafor,^{b*} Emily P. Balskus,^{c*} J. Martin Bollinger, Jr.,^{a,b*} Amie K. Boal^{a,b*}

^aDepartment of Chemistry and ^bDepartment of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania 16802, United States

^cDepartment of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, United States

^lPresent address: Department of Molecular, Cell, and Cancer Biology, University of Massachusetts Medical School, Worcester, MA 01605

*To whom correspondence may be addressed: cdo5093@psu.edu, balskus@chemistry.harvard.edu, jmb21@psu.edu, akb20@psu.edu

Classification: Biological Sciences – Biochemistry, Physical Sciences – Biophysics and Computational Biology

Keywords: *N*-oxygenase, x-ray crystallography, iron binding, bioinformatics, peroxo, streptozotocin

Abstract

In biosynthesis of the pancreatic cancer drug streptozotocin, the tri-domain nonheme-iron oxygenase, SznF, hydroxylates N^δ and $N^{\omega'}$ of N^ω -methyl-L-arginine before oxidatively rearranging the triply modified guanidine to the *N*-methyl-*N*-nitrosourea pharmacophore. A previously published structure visualized the mono-iron cofactor in the enzyme's C-terminal cupin domain, which effects the final rearrangement, but exhibited disorder and minimal metal occupancy in the site of the proposed diiron cofactor in the *N*-hydroxylating heme-oxygenase-like (HO-like) central domain. Here we leverage our recent report of an intensely absorbing μ -peroxodiiron(III/III) intermediate formed from the $\text{Fe}_2(\text{II/II})$ complex and O_2 to understand assembly of the diiron cofactor in the HO-like domain and to obtain structures with both SznF iron cofactors bound. Tight binding at one diiron subsite is associated with a conformational change, which is followed by weak binding at the second subsite and rapid capture of O_2 by the $\text{Fe}_2(\text{II/II})$ complex. Differences between iron-deficient and iron-replete structures reveal both the conformational change required to form the O_2 -reactive $\text{Fe}_2(\text{II/II})$ complex and the structural basis for cofactor instability, showing that a ligand-harboring core helix dynamically refolds during metal acquisition and release. The cofactor also coordinates an unanticipated Glu ligand contributed by an auxiliary helix implicated in substrate binding by docking and molecular dynamics simulation. The additional ligand is conserved in another experimentally validated HO-like *N*-oxygenase but not in two known HO-like diiron desaturases. Among ~9600 sequences identified bioinformatically as belonging to the emerging HO-like diiron protein (HDO) superfamily, ~25% have this carboxylate residue and are thus tentatively assigned as *N*-oxygenases.

Significance statement

The enzyme SznF assembles the *N*-nitrosoourea pharmacophore of the drug streptozotocin. Its central *N*-oxygenase domain resembles heme-oxygenase (HO) and belongs to an emerging superfamily of HO-like diiron enzymes (HDOs) with unstable metallocofactors that have resisted structural characterization. We investigated assembly of the O₂-reactive diiron complex from metal-free SznF and Fe(II) and leveraged this insight to obtain the first structure of a functionally assigned HDO with intact cofactor. Conformational changes accompanying cofactor acquisition explain its instability, and the observation of an unanticipated glutamate ligand that is conserved in only a subset of the HDO sequences provides a potential basis for top-level assignment of enzymatic function. Our results thus provide a roadmap for structural and functional characterization of novel HDOs.

Enzymes with histidine- and carboxylate-coordinated dinuclear iron cofactors catalyze biologically, biomedically, and environmentally important reactions in organisms from bacteria to humans (1, 2). Most of these enzymes use dioxygen as a co-substrate and cleave strong C–H, O–H, or N–H bonds in their reactions. Extensively studied members of this functional class share the conserved four-helix-bundle architecture initially recognized in the iron-storage protein, ferritin (Fig. S1) (3). Examples of ferritin-like diiron oxidases and oxygenases (FDOs) include the β subunit of class Ia ribonucleotide reductase (4, 5), soluble methane monooxygenase (sMMO) (6-10), stearoyl acyl carrier Δ^9 desaturase (Δ^9 D) (11, 12), toluene/*o*-xylene monooxygenase (ToMO) (13-16), and arylamine *N*-oxygenases from the biosynthetic pathways that build the antibiotics aureothin (AurF) (17-20) and chloramphenicol (CmlI) (21-25). Each of these proteins reduces O₂ at its Fe₂(II/II) cofactor to form a μ -peroxo-Fe₂(III/III) intermediate (1). In some cases (e.g., sMMO) additional proteins (MMOB) must bind to the FDO component (MMOH) to activate its cofactor for efficient O₂ capture (8), and, in others (e.g., Δ^9 D) (11), binding of the substrate triggers this step. Early studies emphasized the conversion of the relatively inert initial peroxide-level intermediates to more reactive high-valent [Fe(IV)-containing] complexes by reductive or redox-neutral O–O bond cleavage (4, 5, 26-29), but later studies, including those on the *N*-oxygenases, suggested reaction of substrates directly with the initial mid-valent peroxide adducts (18-21, 23). In these cases, activating protonation and/or isomerization of the canonical μ -1,2-peroxide core was generally posited. Regardless of the pathway, the reactions of these proteins most commonly generate μ -(hydr)oxo-Fe₂(III/III) product clusters, which must be reduced in situ by electron transfer proteins (e.g., ferredoxins or dedicated reductase components) for subsequent turnovers (1, 2). Thus, the diiron clusters of the FDOs act as stable cofactors, cycling through multiple redox states in discrete oxidative and reductive half-reactions.

In the last decade, O₂-reactive, His/carboxylate-coordinated diiron cofactors have been identified in several other protein architectures, including the HD domain (30), the metallo- β -

lactamase fold (31), the TIM-barrel fold (32, 33), and a HEAT-repeat domain (34, 35). Another new structural type, the seven-helix architecture first recognized in heme oxygenase (HO) (Fig. S1) (2, 36), has emerged recently as a second privileged diiron scaffold that may rival the ferritin-like fold in the versatility of transformations it supports. At least four HO-like diiron enzymes (HDOs) – a pair of *N*-oxygenases (37-40) and a pair of olefin-forming lyase/oxidase (hereafter, desaturase) enzymes (41-44) – have been described to date, and several others have been identified within biosynthetic gene clusters (Fig. S2) (45-48). The presumptive enzymatic activity of the first structurally characterized protein in this new superfamily, *Chlamydia* protein associated with death domains (CADD), is still unknown (49, 50). The first HDO to be assigned an enzymatic function, UndA, oxidatively decarboxylates dodecanoic (lauric) acid to undecene (41). An x-ray crystal structure of UndA revealed only a single iron ion with His₂/carboxylate protein coordination within the relatively open HDO core, and it was initially suggested that UndA uses the observed monoiron(II) center for O₂ activation (41). A pair of later studies showed, by contrast, that the iron cofactor is dinuclear (42, 43). One of these studies characterized a μ -peroxo-Fe₂(III/III) complex on the reaction pathway (43). In that work, concerted efforts to solve a structure with both cofactor subsites fully occupied were unsuccessful, and regions of the protein implicated in providing ligands for the second iron site (core α 3) gave weak electron density, indicative of structural disorder. In roughly concurrent studies, the aforementioned pair of HO-like *N*-oxygenases, involved in biosynthesis of streptozotocin (SznF) (37, 38) and azomycin (RohS) (40), also failed to retain their predicted diiron cofactors. For the case of SznF, multiple attempts at cofactor reconstitution, either prior or subsequent to crystallization, did not yield a high-quality structure of the holoenzyme, although partially metallated forms were successfully characterized (Fig. S3) (37). The apparent conservation of cofactor instability among the first several functionally assigned proteins in the emerging HDO superfamily (37, 38, 40, 41, 43) was taken as evidence that they might function by a novel *modus operandi*, perhaps utilizing iron more as a labile co-substrate than as a stably bound, in-situ-recycled cofactor (39). Regardless of the reason, no

structure of a catalytically relevant (diferrous or diferric) holoenzyme HDO complex for a functionally assigned member of this emerging superfamily has yet been reported.

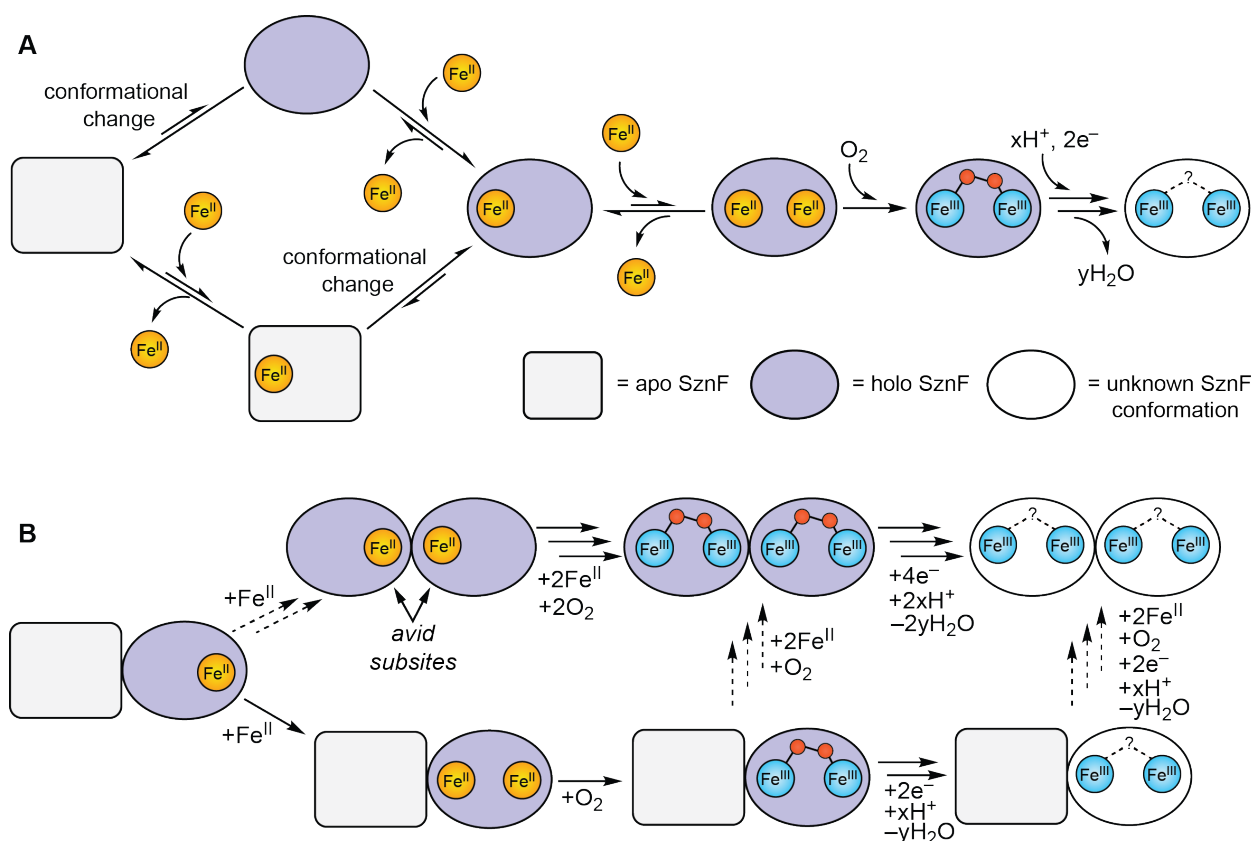
SznF has, in addition to its HDO central domain, an N-terminal dimerization domain and a C-terminal mono-iron cupin domain (Fig. S4). The HDO domain sequentially hydroxylates each of the two unmethylated guanidino nitrogen atoms of N^{ω} -methyl-L-arginine (L-NMA), with N^{δ} being targeted first, before the cupin domain rearranges the triply modified L-Arg to N^{δ} -hydroxy- N^{ω} -methyl- N^{ω} -nitroso-L-citrulline (Fig. S2), the proposed source of the *N*-nitroso-urea pharmacophore of streptozotocin (37). The first SznF structures had a single iron ion in the cupin domain coordinated (in a mode potentially similar to that required for the final rearrangement step) by the N^{δ} -hydroxy- N^{ω} -methyl-L-arginine (L-HMA) product of the first HDO-mediated hydroxylation, but, as previously noted, did not reveal the structure of the HDO diiron cofactor (Fig. S3). In a later study, we showed that, far from being incompetent for iron binding, the HDO domain of the protein can readily acquire Fe(II) from solution and use it to capture O₂ in a μ -peroxo-Fe₂(III/III) complex (39). Acceleration of decay of this complex by mixing with either HDO substrate (L-NMA or L-HMA) showed that this complex either is the *N*-oxygenating complex for both steps or undergoes a substrate-promoted, rapid, reversible, and disfavored conversion to the oxygenating intermediate, which is then efficiently captured by the substrate. Here, we used the spectroscopic signature of this intermediate (absorption feature centered at 629 nm with molar absorptivity of $\sim 3 \text{ mM}^{-1}\text{cm}^{-1}$) to report on assembly of O₂ reactive Fe₂(II/II) complex from the apo protein and Fe(II) and kinetically dissect the assembly process. We then further leveraged this understanding to solve the structure of SznF with its diiron cofactor intact, thus providing the first look at a mechanistically relevant reactant complex of any member of this emerging enzyme superfamily. Comparison of an improved structure of SznF with its HDO domain in its metal-free (apo) form to that of the Fe₂(II/II) complex affords a detailed structural description of the process of Fe(II) acquisition and the major protein conformational changes accompanying it. Together, the results

provide a structural rationale for the cofactor instability seen in each of the initially described HDO systems.

Results and Discussion

Pathway to the μ -Peroxo-Fe₂(III/III) Intermediate in SznF from Apo Protein, Fe(II), and O₂.

We recently demonstrated accumulation of a μ -peroxo-Fe₂(III/III) intermediate to a high level (~ 0.65 enzyme equivalents) upon exposure of a solution containing apo SznF and 3 equiv Fe(II) to excess O₂ (39). The intermediate exhibits intense long-wavelength absorption ($\lambda_{\text{max}} = 629 \text{ nm}$, $\epsilon_{629} \sim 3 \text{ mM}^{-1}\text{cm}^{-1}$), which presumably arises from a peroxide-to-iron charge-transfer (LMCT) transition. In this study, we used development of this feature as the readout to dissect cofactor assembly and intermediate formation. By varying the concentrations, the molar ratios, and the order of mixing of the apo protein, Fe(II) and O₂ reactants, we resolved four steps in this process (Scheme 1A). The apo protein binds Fe(II) tightly at one diiron subsite, and this first, tight-binding step requires a relatively slow ($< 40 \text{ s}^{-1}$ at 5°C) protein conformational change [either before or after Fe(II) association] that partly rate-limits formation of the μ -peroxo-Fe₂(III/III) intermediate from the apo protein and free Fe(II). The second iron subsite has modest affinity and is not saturated even with free Fe(II) concentrations in the 10^{-4} – 10^{-3} -M range. Fe(II) binding at the less avid subsite is also relatively sluggish (~ 1.5 – $3.0 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$) and is followed by more rapid capture of O₂ (~ 1.5 – $5 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$). The data also suggest that the associated monomers of the homodimer interact anti-cooperatively to engender partial "half-of-sites" reactivity (Scheme 1B), which manifests differently in initiation of the reaction by either (i) mixing of the pre-formed Fe(II)•SznF complex with O₂ or (ii) simultaneous mixing of the apo protein with Fe(II) and O₂. We summarize the data revealing this complex cofactor-assembly mechanism and likely anti-cooperativity in the following sections before taking up the structural basis for the effects.



Scheme 1. Assembly and reaction of the SznF diiron cofactor. (A) A conformational change accompanies Fe(II) binding to the avid subsite in the HDO domain of SznF. (B) Rationale for the divergent kinetic behavior of reactions initiated from pre-formed Fe(II)-SznF complexes (*top pathway*) and the apo protein (*bottom pathway*).

Effect of Fe(II):Protein Ratio on the Kinetics of the Reaction of the Fe(II)•SznF(-Δcupin)

Complex with O₂. The kinetics of formation and decay of the μ -peroxo-Fe₂(III/III) intermediate in reaction of the pre-formed Fe(II)•SznF complex with O₂ depend in a complex way on the metal-to-protein ratio in the anoxic reactant solution (Fig. 1A). Several informative trends can be seen in these stopped-flow absorption traces. The reaction with Fe(II):SznF = 1 fails to accumulate the μ -peroxo-Fe₂(III/III) intermediate (Fig. 1A, *blue*). Accumulation of the intermediate increases with Fe(II):SznF ratio in the range of 2-4 (Fig. 1A, *red, orange, green, cyan, purple*), and "saturation" sets in as the ratio reaches 6.0 (Fig. 1A, *black*). The data imply that the O₂-reactive dinuclear complex is not populated when the apo protein is allowed to reach equilibrium with equimolar Fe(II). Mixing of the anoxic Fe(II)•SznF reactant solution with an anoxic solution of the colorimetric

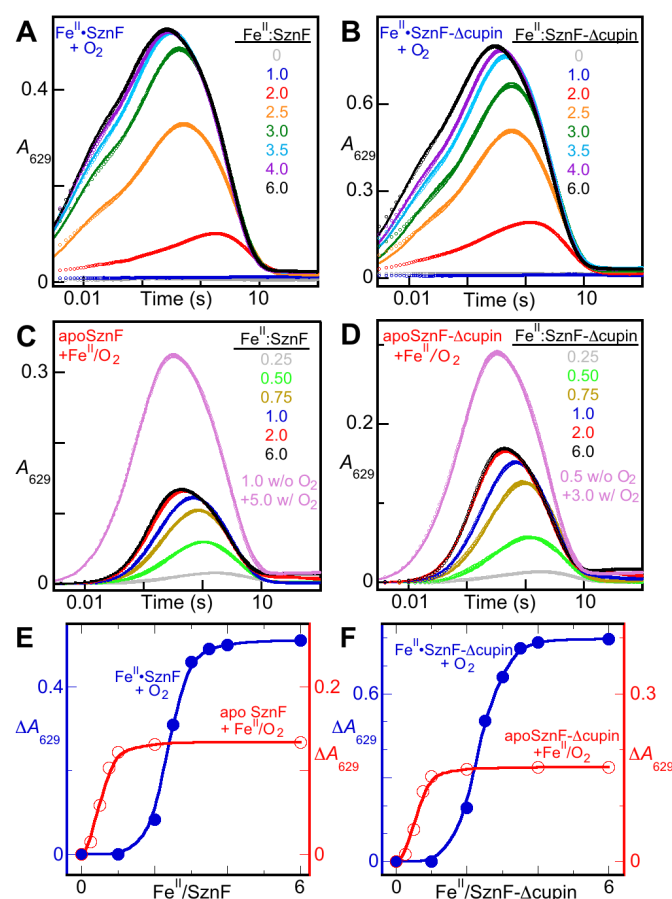


Fig. 1. Stopped-flow absorption experiments monitoring assembly of the O_2 -reactive $\text{Fe}_2(\text{II/II})$ cofactor in the HDO domain of SznF(- Δ cupin) by its conversion to the intensely absorbing μ -peroxo- $\text{Fe}_2(\text{III/III})$ intermediate. (A) Absorbance at 629 nm ($A_{629\text{nm}}$) versus time after an anoxic solution containing wild-type SznF and Fe(II) was mixed at 5 °C with an equal volume of O_2 -saturated buffer. (B) Traces from corresponding reactions of the Δ cupin variant. The solid lines in A and B are fits of the equation (see *SI Appendix*) appropriate for the model of Scheme S1 to the data. (C, D) Traces from reactions of a solution of SznF(- Δ cupin) containing 0-1 equiv Fe(II) with (additional) Fe(II) and O_2 . The solid lines are fits of equation 1 (*SI Appendix*) to the data. (E, F) Differences between maximum and minimum $A_{629\text{nm}}$ for the traces in A-D plotted versus Fe(II):SznF(- Δ cupin). The solid lines are interpolation fits to the data. When no Fe(II) was present in the protein solution, it was saturated with O_2 , and the Fe(II) solution was anoxic and in buffer. When Fe(II) was present in the protein solution, it was anoxic and the Fe(II) solution was both O_2 -saturated and in 1 mM H_2SO_4 rather than buffer.

Fe(II) indicator ferrozine ($\lambda_{\text{max}} = 562 \text{ nm}$, $\epsilon_{562} = 27.9 \text{ mM}^{-1}\text{cm}^{-1}$ (51, 52), rather than with O_2 -saturated buffer, revealed that, with $\text{Fe(II):SznF} = 1$, essentially all the Fe(II) is bound (Fig. S5). Free Fe(II) is chelated rapidly ($k_{\text{obs}} > 50 \text{ s}^{-1}$) by ferrozine under the conditions of this experiment,

engendering a fast phase of development in the A_{562} kinetic trace. There is no such rapid phase with an Fe(II):SznF ratio of 1.0 (Fig. S5A-B, *blue*), but increased ratios give steadily increasing amplitude attributable to free Fe(II) (Fig. S5A-B, *red, green, cyan, purple, black*). In principle, tight binding in the cupin domain of the wild-type enzyme could fully sequester one equiv Fe(II) to prevent its binding in the HDO domain, thus preventing formation of the peroxo-Fe₂(III/III) complex upon exposure of the Fe(II)•SznF reactant to O₂. However, the variant protein with all three His ligands (407/409/448) in the cupin domain replaced by Ala (SznF-Δcupin) behaves similarly, exhibiting neither accumulation of the μ -peroxo-Fe₂(III/III) complex (upon mixing with O₂) nor a fast phase of chelation by ferrozine when the anoxic protein reactant solution contains equimolar Fe(II) and protein (Fig. 1B and Fig. S5C-D). It thus appears that binding of Fe(II) in the cupin domain does not contribute to the observed complex behavior, and we hereafter denote Fe(II)•SznF complexes in terms of only the two sites contributed by the HDO domain. These observations suggest that pre-incubation of apo SznF(-Δcupin) with one equiv Fe(II) leads to sequestration of the metal in an unreactive mononuclear complex *with the metal bound in one of the two cofactor subsites of the HDO domain*. This situation would arise, most simply, from substantially different intrinsic affinities of the two subsites and only modest (or no) cooperativity between them.

A second insight that is obvious from inspection of the ΔA_{629} traces, particularly those from the reactions with Fe(II):SznF = 2.5 - 6, is that formation of the intermediate is kinetically biphasic. The faster phase of development, complete within ~ 0.04 s, acquires increasing amplitude with increasing Fe(II):SznF, coming to represent ~ half of the total rise phase at the highest Fe(II):SznF ratio examined (Fig. S6A). The apparent first-order rate constant associated with this phase, extracted by fitting equation 1 (appropriate for Scheme S1; see below) to the data (solid lines in Fig. 1A), varies by less than 10% from the mean of $1.5 \times 10^2 \text{ s}^{-1}$ for all Fe(II):SznF ratios tested from 2.5 to 6 (Fig. S6B). In this range, free Fe(II) – reflected by the fast phase of chelation in the

anoxic mix with ferrozine – increases proportionally to the Fe(II):protein ratio (Fig. S5B, triangles). The kinetic insensitivity of the fast phase of intermediate formation to the concentration of free Fe(II) suggests that it is associated with trapping of O₂ by the dinuclear Fe₂(II/II)•SznF complex already formed in the anoxic pre-incubation. The slower phase in the A₆₂₉-versus-time traces becomes faster and loses amplitude as the Fe(II):SznF ratio increases from 3 to 6 (Fig. S6A). Its apparent rate constant from the fitting analysis (k_1') nearly doubles (from 6 to 12 s⁻¹) across this range (Fig. S6B, triangles), implying that intermediate formation in the slow phase requires a bimolecular step involving free Fe(II). Succinctly, the more rapid phase of intermediate formation reflects protein with its O₂-reactive Fe₂(II/II) cofactor fully assembled, and the slower phase reflects protein with only the more avid subsite occupied (Scheme 1). The fraction of the protein in the mononuclear form decreases as Fe(II):SznF increases (because the less avid subsite is increasingly occupied), and the amplitude associated with the slower, [Fe(II)]-dependent phase of intermediate formation thus decreases. The other effect of increasing Fe(II):SznF is that any mononuclear complex that still remains can proceed more rapidly through the two-step sequence of (i) Fe(II) binding at the less avid subsite and (ii) O₂ addition, because the concentration of free Fe(II) rises proportionally (as shown by the ferrozine experiment) and the effective first-order rate constant for association at the less avid site increases accordingly.

Kinetics of the Reaction of Oxygenated Apo SznF(-Δcupin) with Fe(II). The kinetic behavior of the reaction initiated by mixing oxygenated apo SznF(-Δcupin) with Fe(II) is drastically different from that seen upon initiation from the pre-formed complex (Fig. 1C, F). First, there is a pronounced lag phase in formation of the intermediate, regardless of the Fe(II):protein ratio. The lag phase is not seen in the reaction of the pre-formed complex with O₂ which implies that it is associated with a slow step in acquisition of Fe(II) by the apo protein. The fact that, even at very low Fe(II):SznF(-Δcupin) ratios (0-1), which do not support accumulation of intermediate in reaction of the pre-formed complex with O₂, give considerable intermediate accumulation when

starting from apo protein implies that the slow step is associated with sequestration of Fe(II) in the tight site. In other words, with time given for this slow step to reach completion in the absence of O₂, equimolar Fe(II) ends up sequestered in only the more avid subsite, but, with O₂ present from the start, filling of the avid subsite is followed by a two-step sequence of Fe(II) acquisition by the weaker site and addition of O₂, and this sequence is faster than the initial binding in the avid subsite. One outcome of this situation is an efficient use of even sub-stoichiometric Fe(II) for still relatively rapid (despite the lag phase) intermediate formation when apo protein is mixed with Fe(II). A second, remarkable outcome is that an Fe(II):SznF(-Δcupin) ratio of 2 is essentially "saturating" in mixing apo protein with Fe(II) and O₂; an increase to a ratio of 6 results in a very modest effect on the kinetics of the intermediate, in marked contrast to the behavior of the reaction of the pre-formed complex with O₂. Even a mere 1 equiv Fe(II), which supports essentially no intermediate accumulation if allowed to bind to the protein in advance of O₂ exposure, supports accumulation of > 90 % of the maximum quantity possible if it is mixed with apo protein in the presence of O₂. In essence, "saturation" – as reported by the kinetics of the intermediate – requires much less Fe(II) if it is not allowed to reach equilibrium in advance of O₂ exposure.

A second notable and (as yet) incompletely understood characteristic of the reaction as initiated from the apo protein is a markedly diminished amplitude in the A₆₂₉-versus-time traces. Regression analysis of the traces according to Eq. 1 returned rate constants of ~ 40 s⁻¹ for the lag phase and 2-8 s⁻¹ [varying with Fe(II):SznF(-Δcupin)] for the formation phase, with a maximum amplitude of 0.15 (0.19 for SznF-Δcupin). This amplitude is only ~ one-fourth the value seen in the traces of Fig. 1A-B. Put simply, two steps occurring with these best-fit rate constants would not limit accumulation to the drastic extent seen in this comparison. A potential explanation lies in the dimeric nature of SznF and the precedent from the ferritin-like diiron enzymes of "half-of-sites reactivity." Prior filling of the more avid subsites in both paired monomers of the dimer may allow for both monomers to form the intermediate rapidly (in either the fast or slow phase seen in Fig.

1A-B), whereas the cascade of Fe(II) binding at the less avid subsite and O₂ addition to the dinuclear cluster that ensues after the apo protein dimer first binds Fe(II) in the more avid subsite may, by conformational coupling, block reaction of the other, paired monomer. In principle, the two phases of intermediate formation seen in reaction of the pre-formed Fe(II)•SznF(-Δcupin) complex with O₂ could similarly represent reactions of pair monomers, the first having both subsites filled and the second requiring Fe(II) binding to the less avid subsite prior to O₂ addition. In this scenario, conformational coupling across the dimer could engender a binding anti-cooperativity that serves to suppress formation of fully loaded Fe(II)₄•SznF dimer (Scheme S1A), as seen before in the ferritin-like class Ia ribonucleotide reductase β subunit from *E. coli* (53, 54). The fact that the amplitude for the Δcupin variant is greater than that for the wild-type protein may also relate to a perturbation of intermonomer coupling caused by replacement of all three His ligands with Ala, although an increased molar absorptivity of the intermediate in the variant scaffold is also theoretically possible.

A Rate-Limiting Conformational Change Associated with Fe(II) Binding in the Avid Subsite.

The results of a hybrid experiment, in which SznF(-Δcupin) was first incubated with ≤ 1 equiv Fe(II) in the absence of O₂ and the resultant complex was then mixed simultaneously with additional Fe(II) and O₂, confirm that binding in the avid subsite drives or locks in a conformational change that is responsible for the lag phase in intermediate formation in the reaction of apo protein with Fe(II) and O₂. The quantity of Fe(II) in the anoxic pre-incubation in this experiment does not, by itself, support intermediate accumulation upon subsequent exposure to O₂ (Fig. 1A-B, *blue*, and Fig. S7). Furthermore, the larger quantity of Fe(II) mixed (along with O₂) with the resultant complex would, if added directly to the apo protein, support only the drastically diminished accumulation noted previously (Fig. 1C-D, *black*, and Fig. S7). It is therefore informative that addition of Fe(II) in *both* stages removes the pronounced lag in intermediate formation and allows ~ twice as much to accumulate as in the reaction of apo protein directly with the same quantity of

Fe(II) (Fig. 1 C,F, *pink*, and Fig. S7). Prior binding just to the avid subsite in the absence of O₂ accomplishes the conformational change in advance of mixing with additional Fe(II) and O₂, while leaving the second, less avid subsite still unoccupied. Mixing with Fe(II) and O₂ then allows intermediate formation to proceed with the effective first-order rate constant characteristic of Fe(II) binding to the less avid subsite ($\sim 12 \text{ s}^{-1}$) and only a much shorter lag phase attributable to fast O₂ addition ($1.8 - 4.5 \times 10^2 \text{ s}^{-1}$ with $\sim 0.9 \text{ mM O}_2$, Fig. S8).

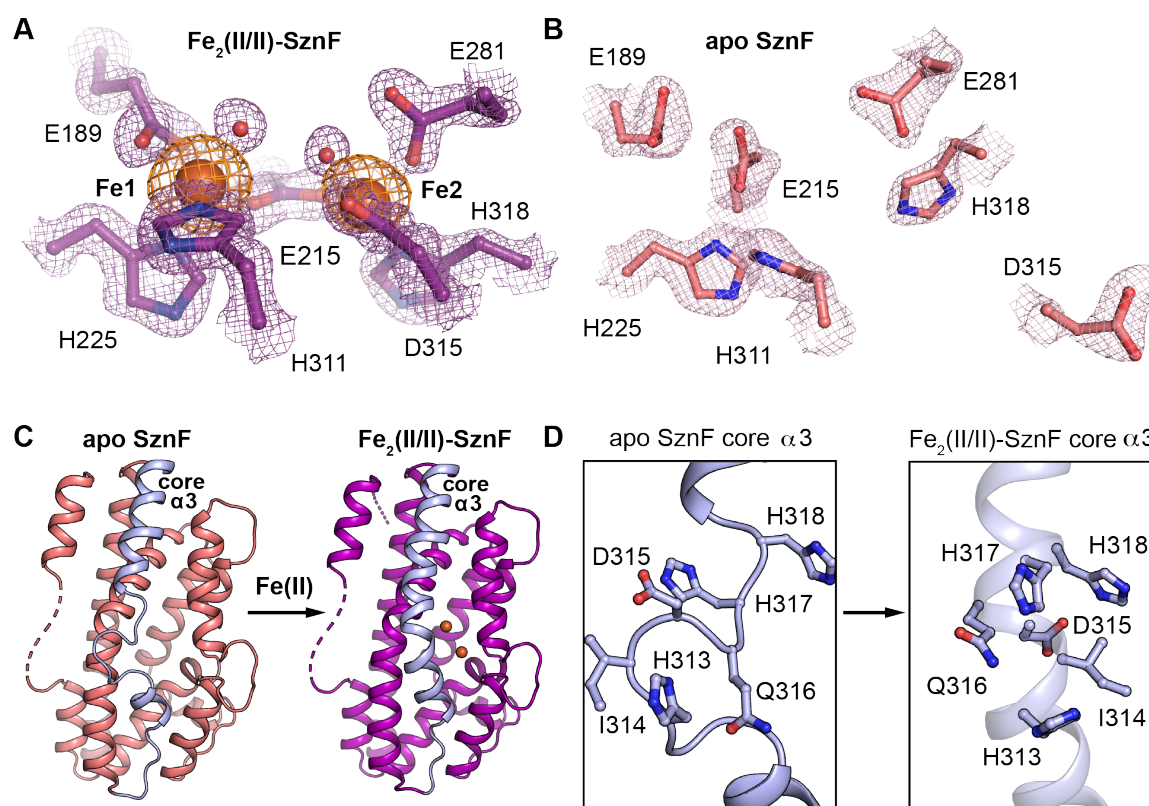


Fig. 2. X-ray crystal structures of Fe₂(II/II)-SznF (A) and apo SznF (B) reveal a conformational change in core $\alpha 3$ (C, D) upon metal binding. $2F_o - F_c$ maps (purple or pink mesh) are shown contoured at 1.5σ (Fe₂(II/II)-SznF) or 1.0σ (apo SznF). An iron anomalous difference map (orange mesh) is shown contoured at 3.0σ . Selected amino acid side chains are shown in stick format. Fe(II) ions and water molecules are shown as orange and red spheres, respectively.

X-ray Crystal Structure of SznF with an Intact Diron Cofactor Reveals a Seventh Cofactor Ligand in the HDO Domain. In previously reported structures of SznF (PDB accession codes 6M9S, 6M9R), the HDO domain was largely devoid of iron (Fig. S3), despite concerted efforts to

isolate the protein with an intact oxidized cofactor or reconstitute with Fe(II) and oxygen (37). Our recognition in a recent study that the Fe₂(III/III) complex is unstable (39), and our elucidation here of the complex Fe₂(II/II) cofactor assembly dynamics, suggested that prior crystallization of the apo protein followed by loading of Fe(II) *in crystallo* might allow for structural characterization of the holoenzyme complex. As recently reported, we produced the protein in *E. coli* cultures supplemented with Mn(II) to limit Fe(II) uptake during over-expression and purification and chelated the Mn(II) to produce the naive [not previously reacted with Fe(II) and O₂] apo protein. We then crystallized the protein anoxically and briefly soaked crystals in a mildly acidic solution of Fe(II) for 5-10 min before freezing them for collection of x-ray diffraction datasets (*SI Appendix*, Table S1). This approach yielded a 1.66-Å resolution crystallographic model with high occupancy (67-100%) of all four Fe(II) binding sites in the two HDO domains of the SznF dimer. The identity of the Fe₂(II/II) cofactor was validated by collection of anomalous diffraction datasets at the Fe and Mn x-ray *K*-edge absorption energy peaks. (Fig. 2A and Fig. S4). Strong peaks in the anomalous difference map were observed only in the datasets acquired at the Fe absorption energy. The structure confirms that the diiron cofactor is held within core helices 1-3 by the conserved 3-His/3-carboxylate ligand set (Fig. S1) predicted from alignment of its sequence with those of other characterized HDOs (Fig. 3).

In addition to verifying the six that were predicted, the new structure identifies an additional iron ligand, Glu189. CADD and UndA – two other structurally characterized HDOs – lack a coordinating residue at this position (Fig. S9). The SznF Glu189 carboxylate binds in bidentate mode to Fe1 (Fig. 2A and Fig. S10). Interestingly, FDO *N*-oxygenases, such as AurF and CmlI, have an analogous extra Fe1 ligand relative to their C/O-H-cleaving counterparts (Figs. S1, S11) (17, 22). However, the identity and origin of the additional ligand is different in the two scaffolds. In the FDOs, the extra ligand is a His provided by the same core helix (core α4) that harbors the Glu and His ligands in the second of the tandem pair of EXXH motifs that identifies the scaffold

(Fig. S1). In SznF, Glu189 is instead contributed by one of the auxiliary helices (aux α 2) found in the HDO scaffold (Figs. S1, S11). . Despite these differences, the extra ligand serves to differen-

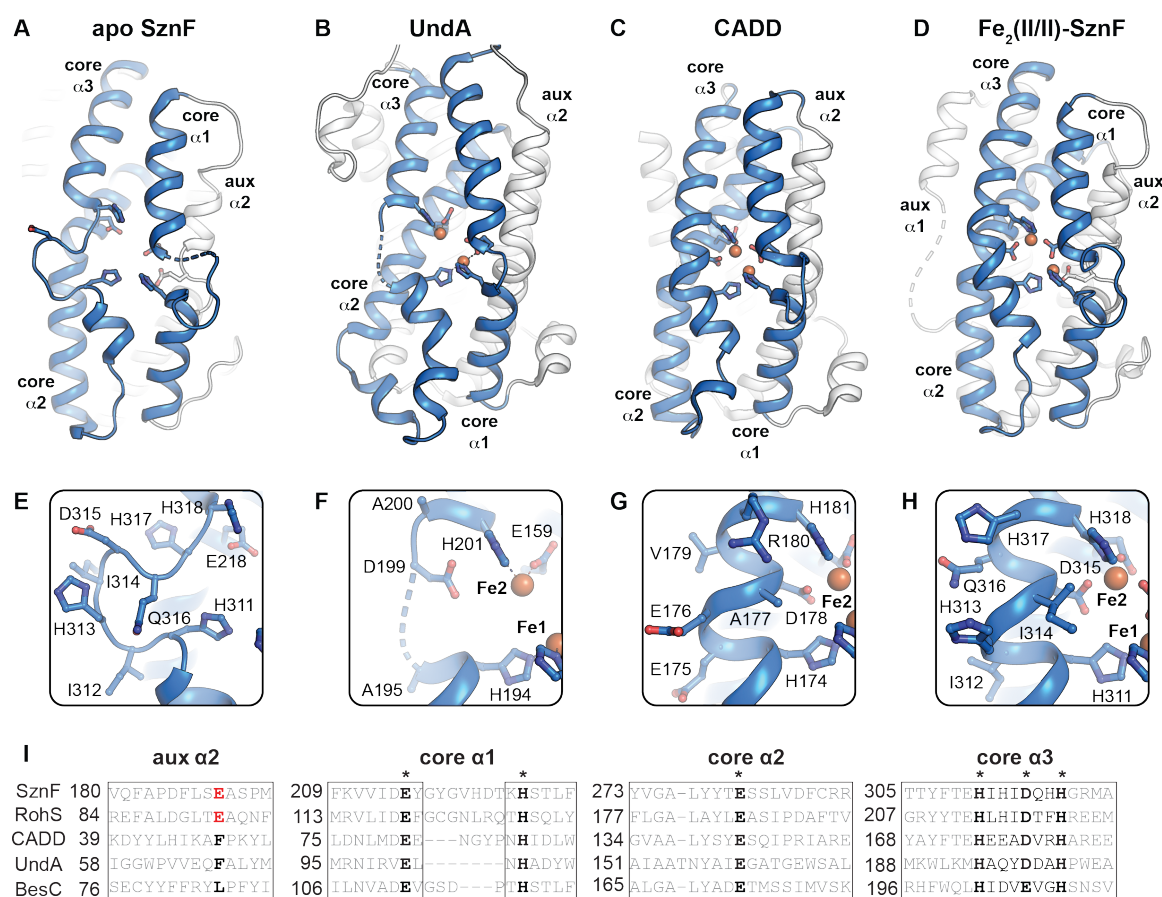


Fig. 3. Structural comparison and sequence alignment of functionally diverse HDO proteins. Apo SznF (A, E) (PDB accession code 6XCV) and UndA (B, F) (PDB accession code 6P5Q) x-ray structures lack full-occupancy cofactors and exhibit disrupted or disordered regions in core α 3. CADD (C, G) (PDB accession code 1RCW) and the iron-bound form of SznF (D, H) (PDB accession code 6VZY) contain full-occupancy diiron cofactors and maintain a canonical α -helix structure throughout core α 3. A sequence alignment (I) of characterized *N*-oxygenases (SznF, RohS), oxidative desaturases (BesC, UndA), and protein of unknown function (CADD) shows that the newly identified Glu ligand in aux α 2 of SznF is conserved in the *N*-oxygenases but absent in CADD and the desaturases. However, all enzymes conserve branched and polar amino acids in the metal-binding region of core α 3, suggesting that all HDOs characterized to date are equipped to employ the conformationally-gated metal-binding mechanism described here for SznF.

tiate the cofactor subsites in both scaffolds, giving a higher coordination number for Fe1 than for Fe2. It has been suggested for AurF and CmlI that saturation of the Fe1 coordination sphere by

the extra His disfavors the η^2 interaction of the dioxygen unit with this iron that would be required for O–O-bond cleavage to generate a high-valent state, thus stabilizing the peroxide-level complex for its direct oxidization of the substrate nitrogen (22). This hypothesis seems reasonable also for the case of SznF.

In the Fe₂(II/II) SznF cofactor, the six-coordinate Fe1 has a distorted-octahedral geometry, whereas the five-coordinate Fe2 has roughly square-pyramidal symmetry. Each site additionally coordinates a water molecule trans to a His ligand (Fig. 2A, Fig. S10). Interestingly, these presumably labile waters are separated by a distance of only 2.8 Å, and they project toward one another into a solvent-lined cavity that could potentially harbor the L-NMA or L-HMA substrate (see below). As such, their locations might approximate the positions of the O-atoms of the bridging peroxide in the detected intermediate. However, formation of the μ -peroxo-Fe₂(III/III) intermediate would likely require significant contraction along the Fe–Fe vector: the \sim 5.0 Å separation observed in the structure is much greater than those found in most structurally characterized μ -peroxo-Fe₂(III/III) complexes with similar ligand sets (2.5 - 3.4 Å) (1). This long intermetallic separation in reduced SznF is enforced by both local and global structural features. The cluster has a single bridging ligand, the carboxylate of Glu215 from core α 2, which coordinates in a μ -1,3 mode. By contrast, reduced complexes of FDOs generally have two or three bridging carboxylates and intermetallic distances of less than 4 Å (1). In some cases, a μ - η^1, η^2 mode of one of these carboxylates is associated with contraction along the Fe–Fe axis, and this geometry is proposed to be diagnostic of the O₂-reactive state (55). No such bridge is present in reduced SznF (Fig. S11).

The long Fe–Fe vector in SznF may also arise, in part, from irregularities within the helices that provide the ligands. Of the four helices that provide ligands, two – aux α 2 and core α 1 – contain significant loop/turn interruptions, and the other two – core α 2 and core α 3 – exhibit more subtle deviations from canonical backbone H-bonding that result in curvature of the helical axis.

These irregularities would, in general, be expected to increase flexibility in the vicinity of the cofactor (Fig. S12). Interestingly, 3 of the 4 ligand-providing helices of the HDO scaffold face the solvent. Their surface exposure, coupled with the aforementioned structural distortions, likely underpins the unusual iron-binding properties of SznF. Furthermore, the large separation of the Fe(II) ions and single carboxylate bridge connecting them could explain why the two cluster subsites behave essentially as one tighter (cofactor-like) subsite and one weaker (substrate-like) subsite, with no obvious cooperativity between them.

Visualization of a Conformational Change in Core Helix 3 Accompanying Cofactor

Assembly. Comparison of the occupancies of the two cofactor subsites – 90-100% for subsite 1 versus 60-80% for subsite 2 – implies that subsite 1 has the greater affinity. This assignment, which would appear to correlate with the greater coordination number of Fe1, would implicate Fe(II) binding to subsite 1 in driving the conformational change. To elucidate the nature of this change, we also solved the structure of the naive apo protein, omitting the Fe(II) soak before freezing the crystals for data collection. The resulting diffraction datasets, collected to a resolution of 1.8 Å, yielded more well-defined electron density for the protein than did the datasets collected in the prior study (Fig. 2B). Most notably, the entire segment spanning core helix $\alpha 3$ – both backbone and side chains – could be definitively modeled, whereas it had previously exhibited partial disorder (Fig. S3). In the new apo structure, the segment spanning residues 311-318, with three cofactor ligands (His311, Asp315 and His318), adopts an ordered, non-helical conformation. By contrast, in the iron-bound structure, this same segment is essentially α -helical (although three canonical backbone interactions are absent), and core $\alpha 3$ is uninterrupted. Remodeling of secondary structure in the residues spanning core $\alpha 3$ thus appears to be a key part of the conformational change seen in the kinetics of cofactor assembly in SznF (Fig. 2C, D).

In addition to this $\alpha 3$ backbone remodeling, multiple, substantial movements of cofactor-coordinating side chains are associated with iron binding. The histidine and carboxylate ligands

contributed by aux $\alpha 2$, core $\alpha 1$, and $\alpha 2$ rotate. The ligands in core $\alpha 3$ change position even more drastically. Asp315 undergoes the greatest shift, moving ~ 7 Å from the surface of the protein and folding into the cavity that harbors the cofactor. Interestingly, this side chain is a ligand only to Fe2, in the subsite of lesser affinity. Repositioning of Asp315 is likely a crucial part of the conformational change, as it should properly configure the second subsite. The data paint a detailed picture of the cofactor assembly process, in which loading of subsite 1 initiates refolding of core $\alpha 3$ to bring subsite 2 ligand Asp315 into position to coordinate the second Fe(II).

Role of Core $\alpha 3$ Dynamics in Cofactor Instability of HDOs. Comparison of the structure of core $\alpha 3$ in the HDO domain of SznF to those of its counterparts in UndA and CADD (Fig. 3A-H) reveals correlations that may be functionally significant. CADD appears to be an outlier in the emerging superfamily (at least at this early stage) in having a stable cluster that allows it to be isolated with bound metal (49). As in Fe₂(II/II)-SznF, core $\alpha 3$ of CADD, which we presume to have been characterized in its oxidized Fe₂(III/III) state, is fully α -helical in the published structure. By contrast, UndA shares with SznF the property of cofactor instability and still has not been crystallographically characterized with its diiron cluster intact, despite concerted efforts toward this goal (43). Core $\alpha 3$ in the UndA structure is partially disordered, as in the previously reported structures of SznF (37), and the Asp315 equivalent, Asp198, resides in or near this unmodeled stretch (Fig. S9). However, the fact that substitution of this residue abolished activity in UndA was taken to imply that it is a cofactor ligand, which would, in turn, suggest that it must also be brought into position during assembly of the O₂-reactive complex (43). Here, an additional event, fatty acid (substrate) binding, is necessary to produce the reactive state. In UndA, it is not known whether substrate binding precedes, interrupts, or follows binding of the two Fe(II) ions. It seems likely that metal binding in subsite 2 with coordination of the mobile core $\alpha 3$ carboxylate – brought into place by a conformational change driven by binding of Fe(II) (to subsite 1) and/or substrate – could be the final step, as in SznF. Moreover, the structural features associated with weak Fe(II) binding in

subsite 2 may be conserved across the larger superfamily. Core $\alpha 3$ is completely surface exposed in all three structurally characterized HDOs. Analysis of the conformationally mobile segment in SznF core $\alpha 3$ (HIHIDQHH) shows that it contains entirely polar or β -branched side chains. Burial of such a sequence in the interior of an α -helix might be disfavored in the absence of compensating metal-ligand interactions. Alignment of the corresponding segments of the functionally assigned HDOs (Fig. 3I) shows that all are rich in polar or branched (or both) amino acids.

Identification of the Substrate Binding Site in the SznF HDO Domain by Computational Modeling Guided by Structural Comparisons. Because initial efforts to obtain an experimental structure of SznF in complex with one of its substrates (L-NMA or L-HMA) were unsuccessful, we used computational docking and molecular dynamics to assess how L-NMA might bind and, importantly, how the regiochemistry of hydroxylation might switch. Cavity-mapping analysis (56) with the coordinates of the $\text{Fe}_2(\text{II/II})$ -SznF structure revealed a pocket near the cofactor with sufficient volume ($>1000 \text{ \AA}^3$) to accommodate the substrate (Fig. S13). The pocket resides on the same face of the metal cluster as the solvent ligands postulated to approximate the positions of the peroxide oxygen atoms in the intermediate. Substrate bound within this pocket would be ideally poised for direct nucleophilic attack on the peroxide, one mechanism proposed for the FDO *N*-oxygenases (18, 21). The SznF cavity is directly connected to the protein surface, terminating near a disordered six-residue sequence in aux $\alpha 1$. This opening provides a plausible access route for substrate to enter the HDO active site. FDO *N*-oxygenases are also purported to undergo a closed-to-open transition to gate substrate access to the active site, although the proposed access route is distinct (22).

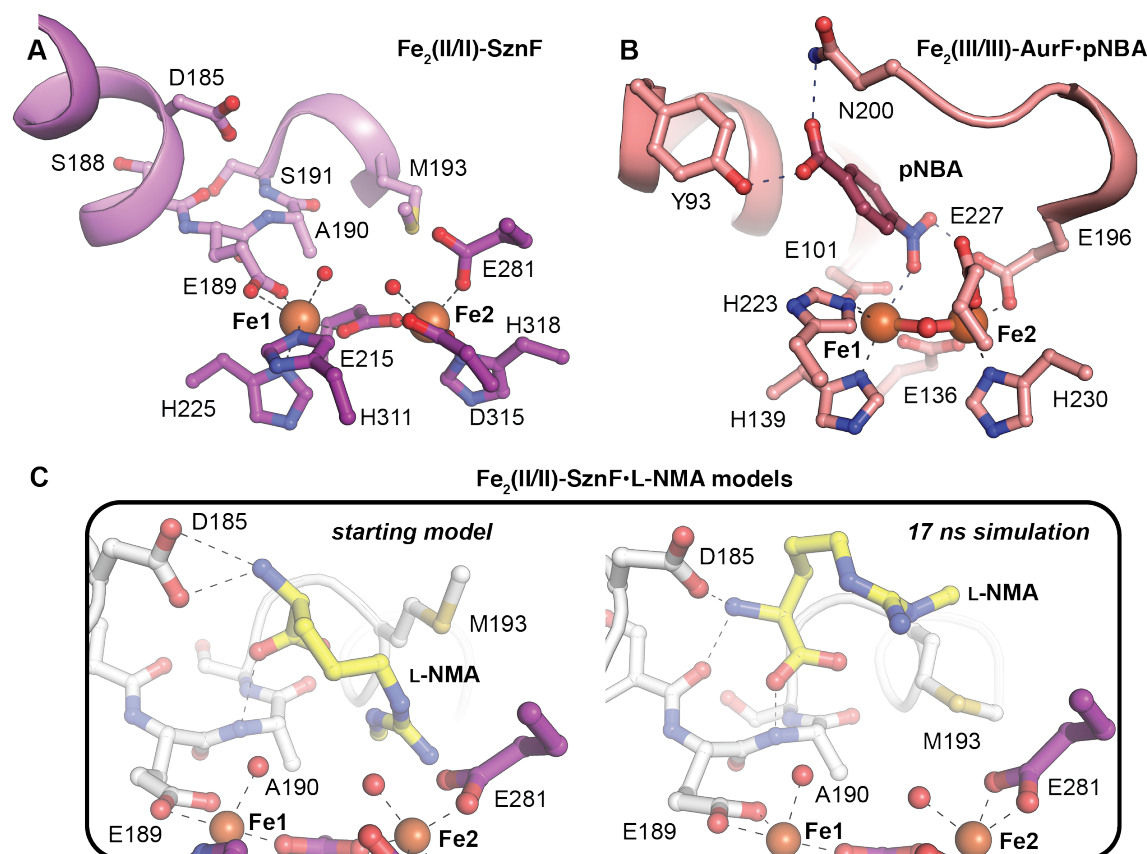


Fig. 4. Aux $\alpha 2$ (violet in panel A, white in panel C) contributes an unanticipated seventh ligand and predicted substrate-binding determinants. A comparison of the metal coordination environment and putative substrate or product binding pockets in SznF (PDB accession code 6VZY) (A) and AurF (PDB accession code 3CHT) (B) reveals similarities between the HDO and FDO *N*-oxygenases. An energy minimized docking model of the SznF substrate, L-NMA, before (*left view*) and after (*right view*) molecular dynamics simulation (C) shows that predicted substrate H-bonding interactions in aux $\alpha 2$ can be accommodated. Selected amino acids and substrates are shown in stick format. Fe^{II} ions and water molecules are shown as orange and red spheres, respectively.

Despite the ionic character of both SznF HDO substrates, this hypothetical binding site is largely hydrophobic. Only a single charged side chain, that of Asp185, projects into the pocket. As noted, it resides on aux $\alpha 2$, which also contains the newly identified Glu189 ligand. Aux $\alpha 2$ is interrupted near these side chains by a type II β -turn insertion (Fig. 4A). Glu189 is located at the tip of this β -turn. The motif both enables a closer approach of the coordinating carboxylates and exposes additional backbone polar functional groups that could interact with the substrate. This

interruption could at least partly complement the charges of the substrates via helical dipoles. Specifically, the bisection of aux $\alpha 2$ by the β -turn exposes negative and positive dipoles that could orient near the α -ammonium and α -carboxylate groups, respectively. The turn is flanked at its C-terminus by Met193. Its thioether side chain projects into the active site. Along with other hydrophobic groups in the substrate pocket, the Met193 side chain could contribute to selectivity for the *N*-methylated L-Arg over the more prevalent, unmethylated form. Intriguingly, the next residue, Met194, projects away from the active site toward its counterpart in the opposite monomer, where it could play a role in the proposed cross-dimer anti-cooperativity in Fe(II) binding (Fig. S14). Thus, aux $\alpha 2$ not only delivers a ligand that was not anticipated from sequence comparisons with other HDOs but may also (i) coordinate intermonomer communication relevant to cofactor assembly and (ii) provide multiple substrate binding determinants.

To test the idea that the aux $\alpha 2$ functional groups contribute to binding and proper orientation of L-NMA, we generated a docking model of the Fe₂(II/II)-SznF•L-NMA complex. We reasoned that the α -amine and α -carboxylate groups might interact with the side chain of Asp185 and the backbone amide of Ala190, respectively. We subjected an energy-minimized docking model of this complex to molecular dynamics (MD) simulations (Fig. 4C and Fig. S15). The simulations show that the substrate can be accommodated in the open cavity above the cofactor while maintaining the predicted contacts with the Asp185 carboxylate and Ala190 amide nitrogen within the β -turn. While these interactions persist throughout the simulation, the L-NMA side chain remains dynamic, populating different rotamers and occupying different positions within the active site (Fig. S15). In general, the pair of guanidino nitrogen atoms that are hydroxylated in the sequential steps remain close to the cofactor; their relatively high mobility may enable the switch in regiochemistry (37, 38).

The modeled structure provides a basis to compare the location and nature of the substrate-binding pocket in SznF with those in other HDOs and FDOs. Some published structures

of UndA include bound lauric acid (Figs. S9 and S16) (41, 43) and, although none of these have high occupancies of iron subsite 2, they do reveal direct coordination of the substrate carboxylate by Fe1, in approximately the same location as the Glu189 ligand in SznF. Correspondingly, aux $\alpha 2$ of UndA lacks the β -turn insertion that harbors the extra carboxylate ligand and is instead uninterrupted in all published structures (41, 43). The carboxylate of Glu189 in SznF and the substrate carboxylate in UndA might thus be, in a sense, functionally analogous. If this carboxylate-Fe1 interaction is necessary to set the geometry and tune the electronic structure of the cofactor for O₂ capture, the difference between the UndA and SznF structures would explain why the O₂-addition step is markedly accelerated by substrate binding in the former HDO (43) but not in the latter (39). Despite this difference, aux $\alpha 2$ of UndA, like the corresponding helix in SznF, contributes groups that interact with the substrate: the hydrophobic side chains of Val65 and Phe68 both contact the alkyl chain of the bound lauric acid (41, 43). The comparison suggests that analysis of aux $\alpha 2$ sequences in HDOs might aid in identifying their substrates and perhaps engineering their specificities and outcomes.

Although the FDO and HDO scaffolds have completely different topologies, the *N*-oxygenases in the two superfamilies appear to use a conserved coordination sphere (3-His/4-carboxylate) that, in both cases, includes an increased coordination number of their Fe1 site in comparison to their C/O-H activating counterparts (Figs. S1, S11, S16). However, in the FDOs, the extra ligand resides on a core helix that, across the entire superfamily, contributes two other cofactor ligands (17), whereas, in SznF, it extends from an auxiliary helix that contributes only this ligand, and only in a subset of the superfamily (Fig. S1). AurF is the only FDO *N*-oxygenase crystallographically characterized to date with its product bound (17). The two systems share a similar relative disposition of the substrate/product and cofactor (Fig. 4A and B). In AurF, the 4-nitrobenzoic acid (pNBA) product is held over Fe1 by polar contacts between the product carboxylate and side chains of Tyr93 and Asn200 from core helices $\alpha 1$ and $\alpha 3$, respectively. In

the $\text{Fe}_2(\text{II/II})\text{-SznF}\cdot\text{L-NMA}$ model, the approach and number/type of contacts are similar. Our model predicts that the substrate is held over Fe1 by the aforementioned interactions with its α -ammonium and α -carboxylate groups. The FDO *N*-oxygenases have also resisted characterization in their substrate-bound forms, perhaps due to relatively low substrate affinities. In both the FDO and HDO systems, the failure of the substrate to impact the kinetics of O_2 addition may indicate that this step precedes substrate binding in the preferred pathway (18, 21, 39). In this scenario, the $\text{Fe}_2(\text{II/II})\text{-SznF}\cdot\text{L-NMA}$ complex would be out of sequence and challenging to capture in the crystal.

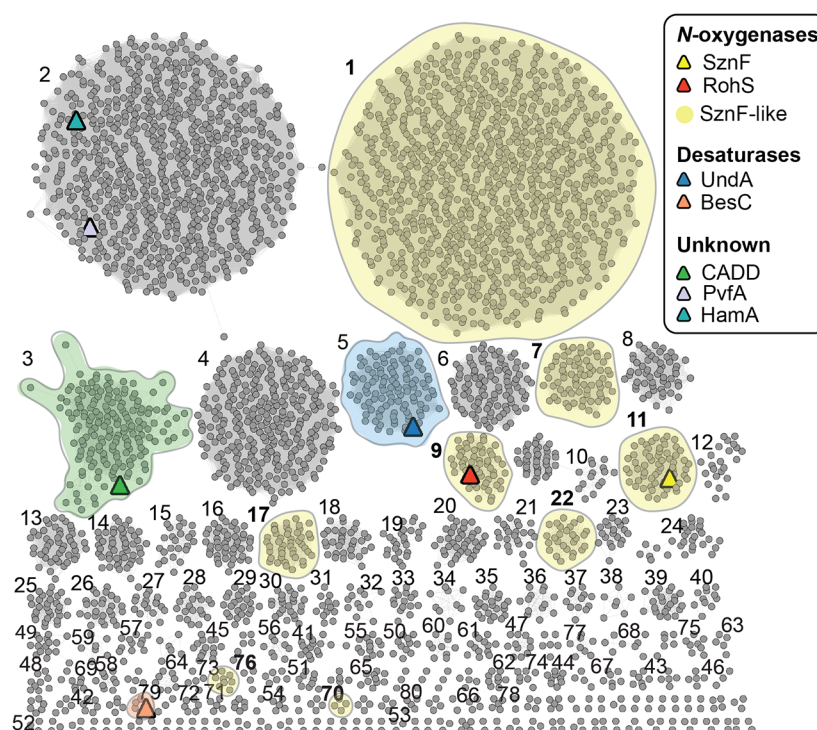


Fig. 5. Bioinformatic analysis of the HDO superfamily. A sequence similarity network of 9600 predicted HDO sequences shows that many sequence families remain uncharacterized at the structural and/or functional level. Homologs of known *N*-oxygenases (red, yellow), oxidative desaturases (blue, peach), or structurally characterized proteins of unknown function (green, lavender, teal) are indicated as a colored triangle. Clusters designated SznF-like are shaded yellow, indicating conservation of the extra Glu ligand in aux $\alpha 2$.

Bioinformatic analysis of the HDO superfamily. Comparison of the known structural and functional characteristics of SznF and UndA highlights potentially predictive correlations. The presence of the carboxylate ligand from aux $\alpha 2$ may (i) obviate direct substrate coordination to the cofactor and thus its triggering of O₂ addition and (ii) stabilize the diiron-O₂ adduct with respect to O-O-bond cleavage, favoring oxidation only of substrates that can attack as nucleophiles. Conversely, the absence of the extra carboxylate should leave this site vacant for substrate coordination, triggering of O₂ addition, and (more speculatively) progression to a more reactive complex capable of H• abstraction from an aliphatic carbon (e.g., C3 of lauric acid, as proposed in recent studies of UndA). We searched sequence databases to segregate potential HDOs into groups that conserve the aux $\alpha 2$ β -turn insertion harboring the extra Glu from those that don't. Use of an alignment score of 51 in generating a sequence similarity network (SSN) from the ~ 9600 predicted HDO proteins that we identified (Fig. 5) separated the reference sequences [SznF and RohS (validated *N*-oxygenases), UndA and BesC (validated desaturases), and CADD (unknown function)] into different clusters, consistent with their distinct activities and targets. Within the top 80 clusters in the network (ranked by number of sequences), we searched for conservation of an aux $\alpha 2$ β -turn and coordinating carboxylate by manual alignment to SznF and other reference sequences. Predicted protein sequences in eight of these clusters (clusters 1, 7, 9, 11, 17, 22, 70, and 76) conserve features of the insertion/extra-ligand motif (Fig S17, Table S2). Clusters 9 and 11 contain the known *N*-oxygenases, RohS and SznF, respectively. We suspect that the other six clusters might also contain *N*-oxygenases with properties similar to SznF. One of these groups, cluster 1, is the largest of the SSN, containing 1780 sequences, roughly 20% of the HDO sequence pool evaluated here. Cluster 1 is enriched in sequences from non-pathogenic *Streptomyces* and *Mycobacterium* organisms (Fig. S17). Each of the remaining four clusters contains 50-200 sequences, and these are similarly dominated by predicted proteins from environmental aquatic and soil microbes, as opposed to pathogens or commensal bacteria.

Apart from cluster 11, in which all sequences are sufficiently long (Fig. S18) to suggest fusion to a separate (e.g., cupin) domain, the remaining predicted *N*-oxygenases are uniformly shorter, indicating that they likely act as simple *N*-oxygenases. Testing of these top-level functional predictions is likely to vastly expand the known repertoire of the emerging HDO superfamily.

Methods

Detailed materials and methods are provided in the *SI Appendix*.

Stopped-flow absorption (SF-Abs) spectroscopy. Stopped-flow absorption experiments were performed at 5 °C on an Applied Photophysics Ltd. (Leatherhead, UK) SX20 stopped-flow spectrophotometer housed in an MBraun anoxic chamber and equipped with either a photodiode-array (PDA) or photomultiplier tube (PMT) detector. Details of reaction conditions and data analysis are provided in the figure legends and *SI Appendix*.

X-ray crystallographic characterization of apo and Fe₂(II/II)-SznF. Apo N-terminally His₆-tagged SznF was overexpressed and purified as detailed in the *SI Appendix*. SznF was crystallized under anoxic conditions in a stringently apo form in 16% (w/v) PEG 4000, 0.1 M MgCl₂, and 0.1 M Tris pH 8.5. To obtain the Fe(II)-bound form of the enzyme, crystals were soaked for 5-10 min in precipitant solution supplemented with 10 mM ferrous ammonium sulfate. Data collection and refinement procedures are described in the *SI Appendix*.

Bioinformatic analysis of HDO enzymes. An initial pool of 42,581 unique sequences annotated as either heme-oxygenase or heme-oxygenase-like proteins was extracted from Interpro (Nov. 2019 release). Using length filters and a custom Python script, this pool was analyzed for conservation of the HDO core helix diiron ligands (Fig. S1) to obtain a final set of ~ 9600 candidate HDO enzymes as input for SSN generation with the Enzyme Function Initiative Enzyme Similarity Tool (57). Using manual and automated sequence alignments and homology model generation, eight of the top 80 (ranked by number of sequences) sequence clusters were annotated as likely

N-oxygenases on the basis of conservation of the additional carboxylate ligand and β -turn motif in aux $\alpha 2$.

Acknowledgements

We gratefully acknowledge use of the resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817). GM/CA@APS has been funded in whole or in part with Federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). The Eiger 16M detector at GM/CA-XSD was funded by NIH grant S10 OD012289. This work was additionally supported by National Institutes of Health grants GM119707 to A.K.B. and GM138580 to J.M.B.

REFERENCES

1. Jasniewski AJ & Que L, Jr. (2018) Dioxygen activation by nonheme diiron enzymes: diverse dioxygen adducts, high-valent intermediates, and related model complexes. *Chem Rev* 118(5):2554-2592.
2. Rajakovich LJ, *et al.* (2020) Emerging structural and functional diversity in proteins with dioxygen reactive dinuclear transition metal cofactors. *Comprehensive Natural Products III*, ed Liu H-w, Begley, T. P. (Elsevier, Amsterdam) DOI:10.1016/B978-0-12-409547-2.14864-4.
3. Theil EC (1987) Ferritin: Structure, gene, regulation and cellular function in animals, plants, and microorganisms. *Annu Rev Biochem* 56(1):289-315.
4. Bollinger JM, Jr., *et al.* (1991) Mechanism of assembly of the tyrosyl radical-dinuclear iron cluster cofactor of ribonucleotide reductase. *Science* 253(5017):292-298.
5. Bollinger JM, Jr., *et al.* (1994) Mechanism of assembly of the tyrosyl radical-diiron(III) cofactor of *E. coli* ribonucleotide reductase. 2. kinetics of the excess Fe^{2+} reaction by optical, EPR, and Mössbauer spectroscopies. *J Am Chem Soc* 116(18):8015-8023.
6. Banerjee R, Jones JC, & Lipscomb JD (2019) Soluble methane monooxygenase. *Annu Rev Biochem* 88(1):409-431.
7. Tinberg CE & Lippard SJ (2011) Dioxygen activation in soluble methane monooxygenase. *Acc Chem Res* 44(4):280-288.
8. Merckx M, *et al.* (2001) Dioxygen activation and methane hydroxylation by soluble methane monooxygenase: A tale of two irons and three proteins. *Angew Chem Int* 40(15):2782-2807.
9. Wallar BJ & Lipscomb JD (1996) Dioxygen activation by enzymes containing binuclear non-heme iron clusters. *Chem Rev* 96(7):2625-2658.
10. Tinberg CE & Lippard SJ (2009) Revisiting the mechanism of dioxygen activation in soluble methane monooxygenase from *M. capsulatus* (Bath): evidence for a multi-step, proton-dependent reaction pathway. *Biochemistry* 48(51):12145-12158.
11. Fox BG, Lyle KS, & Rogge CE (2004) Reactions of the diiron enzyme stearyl-acyl carrier protein desaturase. *Acc Chem Res* 37(7):421-429.
12. Broadwater JA, Ai J, Loehr TM, Sanders-Loehr J, & Fox BG (1998) Peroxodiferic intermediate of stearyl-acyl carrier protein delta 9 desaturase: oxidase reactivity during single turnover and implications for the mechanism of desaturation. *Biochemistry* 37(42):14664-14671.
13. Murray LJ, García-Serres R, Naik S, Huynh BH, & Lippard SJ (2006) Dioxygen activation at non-heme diiron centers: characterization of intermediates in a mutant form of toluene/o-xylene monooxygenase hydroxylase. *J Am Chem Soc* 128(23):7458-7459.
14. Murray LJ, *et al.* (2007) Characterization of the arene-oxidizing intermediate in ToMOH as a diiron(III) species. *J Am Chem Soc* 129(46):14500-14510.
15. Song WJ & Lippard SJ (2011) Mechanistic studies of reactions of peroxodiiron(III) intermediates in T201 variants of toluene/o-xylene monooxygenase hydroxylase. *Biochemistry* 50(23):5391-5399.
16. Acheson JF, Bailey LJ, Brunold TC, & Fox BG (2017) In-crystal reaction cycle of a toluene-bound diiron hydroxylase. *Nature* 544(7649):191-195.
17. Choi YS, Zhang H, Brunzelle JS, Nair SK, & Zhao H (2008) In vitro reconstitution and crystal structure of *p*-aminobenzoate *N*-oxygenase (AurF) involved in aureothin biosynthesis. *Proc Natl Acad Sci U S A* 105(19):6858-6863.

18. Korboukh VK, Li N, Barr EW, Bollinger JM, Jr., & Krebs C (2009) A long-lived, substrate-hydroxylating peroxodiiron(III/III) intermediate in the amine oxygenase, AurF, from *Streptomyces thioluteus*. *J Am Chem Soc* 131(38):13608-13609.
19. Li N, Korboukh VK, Krebs C, & Bollinger JM, Jr. (2010) Four-electron oxidation of *p*-hydroxylaminobenzoate to *p*-nitrobenzoate by a peroxodiferric complex in AurF from *Streptomyces thioluteus*. *Proc Natl Acad Sci U S A* 107(36):15722-15727.
20. Park K, *et al.* (2017) Peroxide activation for electrophilic reactivity by the binuclear non-heme iron enzyme AurF. *J Am Chem Soc* 139(20):7062-7070.
21. Makris TM, *et al.* (2015) An unusual peroxo intermediate of the arylamine oxygenase of the chloramphenicol biosynthetic pathway. *J Am Chem Soc* 137(4):1608-1617.
22. Knoot CJ, Kovaleva EG, & Lipscomb JD (2016) Crystal structure of CmlI, the arylamine oxygenase from the chloramphenicol biosynthetic pathway. *J Biol Inorg Chem* 21(5):589-603.
23. Jasniewski AJ, Komor AJ, Lipscomb JD, & Que L, Jr. (2017) Unprecedented (*m*-1,1-peroxo)diferric structure for the ambiphilic orange peroxo intermediate of the nonheme *N*-oxygenase CmlI. *J Am Chem Soc* 139(30):10472-10485.
24. Komor AJ, *et al.* (2017) CmlI *N*-oxygenase catalyzes the final three steps in chloramphenicol biosynthesis without dissociation of intermediates. *Biochemistry* 56(37):4940-4950.
25. Komor AJ, Jasniewski AJ, Que L, & Lipscomb JD (2018) Diiron monooxygenases in natural product biosynthesis. *Nat Prod Rep*.
26. Lee SK, Nesheim JC, & Lipscomb JD (1993) Transient intermediates of the methane monooxygenase catalytic cycle. *J Biol Chem* 268(29):21569-21577.
27. Lee SK, Fox BG, Froland WA, Lipscomb JD, & Münck E (1993) A transient intermediate of the methane monooxygenase catalytic cycle containing an Fe^{IV}Fe^{IV} cluster. *J Am Chem Soc* 115(14):6450-6451.
28. Nesheim JC & Lipscomb JD (1996) Large kinetic isotope effects in methane oxidation catalyzed by methane monooxygenase: evidence for C-H bond cleavage in a reaction cycle intermediate. *Biochemistry* 35(31):10240-10247.
29. Brazeau BJ & Lipscomb JD (2000) Kinetics and activation thermodynamics of methane monooxygenase compound Q formation and reaction with substrates. *Biochemistry* 39(44):13503-13515.
30. Aravind L & Koonin EV (1998) The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem Sci* 23(12):469-472.
31. Daiyasu H, Osaka K, Ishino Y, & Toh H (2001) Expansion of the zinc metallo-hydrolase family of the beta-lactamase fold. *FEBS Lett* 503(1):1-6.
32. Kenney GE, *et al.* (2018) The biosynthesis of methanobactin. *Science* 359(6382):1411.
33. Dong L-B, *et al.* (2019) Characterization and crystal structure of a non-heme diiron monooxygenase involved in platensimycin and platencin biosynthesis. *J Am Chem Soc* 141(31):12406-12412.
34. Park J-H, *et al.* (2006) Molecular cloning, expression, and structural prediction of deoxyhypusine hydroxylase: A HEAT-repeat-containing metalloenzyme. *Proc Natl Acad Sci U S A* 103(1):51.
35. Han Z, *et al.* (2015) Crystal structure of the peroxo-diiron(III) Intermediate of deoxyhypusine hydroxylase, an oxygenase involved in hypusination. *Structure* 23(5):882-892.
36. Poulos TL (2014) Heme enzyme structure and function. *Chem Rev* 114(7):3919-3962.
37. Ng TL, Rohac R, Mitchell AJ, Boal AK, & Balskus EP (2019) An *N*-nitrosating metalloenzyme constructs the pharmacophore of streptozotocin. *Nature* 566(7742):94-99.

38. He HY, Henderson AC, Du YL, & Ryan KS (2019) Two-enzyme pathway links L-arginine to nitric oxide in *N*-nitroso biosynthesis. *J Am Chem Soc* 141(9):4026-4033.
39. McBride MJ, *et al.* (2020) A peroxodiiron(III/III) intermediate mediating both *N*-hydroxylation steps in biosynthesis of the *N*-nitroso urea pharmacophore of streptozotocin by the multi-domain metalloenzyme SznF. *J Am Chem Soc* 142(27):11818-11828.
40. Hedges JB & Ryan KS (2019) In vitro reconstitution of the biosynthetic pathway to the nitroimidazole antibiotic azomycin. *Angew Chem Int Ed Engl* 58(34):11647-11651.
41. Rui Z, *et al.* (2014) Microbial biosynthesis of medium-chain 1-alkenes by a nonheme iron oxidase. *Proc Natl Acad Sci U S A* 111(51):18237-18242.
42. Manley OM, Fan R, Guo Y, & Makris TM (2019) Oxidative decarboxylase UndA utilizes a dinuclear iron cofactor. *J Am Chem Soc* 141(22):8684-8688.
43. Zhang B, *et al.* (2019) Substrate-triggered formation of a peroxo-Fe₂(III/III) intermediate during fatty acid decarboxylation by UndA. *J Am Chem Soc* 141(37):14510-14514.
44. Marchand JA, *et al.* (2019) Discovery of a pathway for terminal-alkyne amino acid biosynthesis. *Nature* 567(7748):420-424.
45. Hermenau R, *et al.* (2018) Gramibactin is a bacterial siderophore with a diazeniumdiolate ligand system. *Nat Chem Biol* 14(9):841-843.
46. Hermenau R, *et al.* (2019) Genomics-driven discovery of NO-donating diazeniumdiolate siderophores in diverse plant-associated bacteria. *Angew Chem Int Ed Engl* 58(37):13024-13029.
47. Jenul C, *et al.* (2018) Biosynthesis of fragin is controlled by a novel quorum sensing signal. *Nat Commun* 9(1):1297.
48. Kretsch AM, *et al.* (2018) Discovery of (dihydro)pyrazine *N*-oxides via genome mining in *Pseudomonas*. *Org Lett* 20(16):4791-4795.
49. Schwarzenbacher R, *et al.* (2004) Structure of the *Chlamydia* protein CADD reveals a redox enzyme that modulates host cell apoptosis. *J Biol Chem* 279(28):29320-29324.
50. Adams NE, *et al.* (2014) Promiscuous and adaptable enzymes fill "holes" in the tetrahydrofolate pathway in *Chlamydia* species. *MBio* 5(4):e01378-01314.
51. Umback NJ & Norton JR (2002) Indirect determination of the rate of iron uptake into the apoprotein of the ribonucleotide reductase of *E. coli*. *Biochemistry* 41(12):3984-3990.
52. Stookey LL (1970) Ferrozine - a new spectrophotometric reagent for iron. *Anal Chem* 42(7):779-&.
53. Pierce BS, Elgren TE, & Hendrich MP (2003) Mechanistic implications for the formation of the diiron cluster in ribonucleotide reductase provided by quantitative EPR spectroscopy. *J Am Chem Soc* 125(29):8748-8759.
54. Pierce BS & Hendrich MP (2005) Local and global effects of metal binding within the small subunit of ribonucleotide reductase. *J Am Chem Soc* 127(10):3613-3623.
55. Voegtli WC, *et al.* (2003) Variable coordination geometries at the diiron(II) active site of ribonucleotide reductase R2. *J Am Chem Soc* 125(51):15822-15830.
56. Ho BK & Gruswitz F (2008) HOLLOW: generating accurate representations of channel and interior surfaces in molecular structures. *BMC Struct Biol* 8(1):49.
57. Zallot R, Oberg N, & Gerlt JA (2019) The EFI web resource for genomic enzymology tools: leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. *Biochemistry* 58(41):4169-4182.