Mag. Sarah Pratter

Mononuclear Nonheme Iron and
Oxygen Dependent Tayloring Enzymes from
Antibiotic Biosynthetic Pathways

_____

DISSERTATION

vorgelegt an der Technischen Universität Graz

zur Erlangung des akademischen Grades einer

Doktorin der Naturwissenschaften

Betreuer: Dipl.-Ing. Dr.techn. Grit Straganz und

Univ.-Prof. Dipl.-Ing. Dr.techn. Bernd Nidetzky,

Institut für Biotechnologie und Bioprozesstechnik

Graz, im Juni 2013

Deutsche Fassung:
Beschluss der Curricula-Kommission für Bachelor-, Master- und Diplomstudien vom 10.11.2008
Genehmigung des Senates am 1.12.2008

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am …………………………   …………………………………………..
                                                                   (Unterschrift)

Englische Fassung:

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

…………………………   …………………………………………..
         date   (signature)

Danksagung

Table of Contents

Abstract

Mononuclear nonheme iron enzymes catalyze particularly complex, highly specific reactions in the aerobic metabolism of organisms from all domains of life. In bacterial lipopeptide biosynthetic pathways they modify pharmaceutically relevant metabolites during or after their biosynthesis via a range of chemically challenging, O_2 dependent oxidations that are often still not well understood on a molecular basis. Within this work the putative alpha-ketoglutarate (α-KG) dependent halogenase HctB from the hectochlorin biosynthetic pathway in *Lyngbya majuscula* is characterized. This first principle study of a fatty acyl-halogenase reveals the regioselective dichlorination and oxygenation of its native hexanoyl-substrate. Additionally, unprecedented introduction of a vinyl-chloride functionality into the non-activated acyl-group is observed. HctB shows the potential for bromination and displays relaxed specificity towards middle- and long-chain fatty acyl-substrates. This suggests that HctB can bestow considerable diversity to a suitable target molecule and thereby potentially enables the synthesis of substances with altered properties. Supported by spectroscopic and stopped flow kinetic studies it is furthermore demonstrated that chloride plays a pivotal role in the activation of O_2 by the halogenase: the halide is crucial to generate the catalytically active, 5-coordinate square pyramidal Fe(II) center and accelerates its reactivity towards O_2 > 200 fold. Complementary MD simulations find a chloride-induced rearrangement of the resting enzyme structure, that suggests that chloride impacts the metal center geometry both directly, by its coordination, and indirectly, by facilitating the substrate's entering into the active site. In a second project the molecular basis of chiral hydroxylation in (*S*)-hydroxymandelate synthase is elucidated. Inversion of the enzyme's enantioselectivity is achieved by a rational design approach and accompanied by a triplication of enzyme efficiency. The crystal structure of the newly designed (*R*)-mandelate synthase and MD-simulation analyses give further insights into the structural basis of enantio-selective catalysis.

Zusammenfassung

Mononukleäre Nicht-Häm-Eisen (MNHE) Enzyme sind allgegenwärtig in der Natur und man findet sie dort im aeroben Metabolismus verschiedenster Spezies. In Bakterien modifizieren diese Proteine pharmazeutisch relevante Lipopeptide, während oder nach deren Biosynthese durch äußerst komplexe und hoch spezifische O_2 abhängigen Oxidationen. In vielen Fällen sind die molekularen Mechanismen dieser chemisch anspruchsvollen Umsetzungen noch nicht vollständig aufgeklärt. Im ersten Teil dieser Arbeit wird die alpha-Ketoglutarat (α-KG) abhängige Halogenase HctB, die einen Schritt in der Herstellung des Naturstoffes Hectochlorin in *Lyngbya majuscula* katalysiert, biochemisch charakterisiert. Die Studie, welche zum Ziel hatte, die Grundprinzipien der MNHE abhängigen Halogenierung von Fettsäurebausteinen aufzuklären, zeigt, dass das Enzym neben der regioselektiven Dichlorierung und Oxygenierung auch eine bislang unbekannte, kombinierte Desaturierung/ Halogenierung katalysiert, die zum Einbau einer Vinylchlorid-Funktion in die C6-Fettsäurekette des Substrates führt. HctB besitzt ein breites Substratspektrum und kann Brom als alternatives Halogen einbauen. Diese offenbare katalytische Vielfalt macht HctB zu einem vielversprechenden Biokatalysator für die biosynthetische Herstellung von neuen, bioaktiven Lipopeptidvarianten. In einer weiterführenden Studie wird mit Hilfe von Stopped-Flow und CD/MCD Analysen demonstriert, dass Chlor eine Schlüsselrolle in der Aktivierung von O_2 dieser Halogenase spielt: Das Halogen induziert die katalytisch aktive, quadratisch-pyramidale Geometrie am Eisenzentrum und erhöht dessen Reaktivität gegenüber O_2 um mehr als das 200-fache. In MD-Simulationen wird gezeigt, dass die Koordination des Chlors zum Eisen-Kofaktor, neben dem experimentell belegten direkten Einfluss auf das Metallzentrum, eine Konformationsänderung im Enzym bewirkt, welche dem Substrat erst den Zugang zum aktiven Zentrum eröffnet. Im dritten Teil dieser Arbeit wird die molekulare Grundlage der chiralen Hydroxylierung in der O_2- und 2-Ketosäure-abhängigen MNHE Dioxygenase (*S*)-Hydroxymandelatsynthase aufgeklärt. Durch rationelles Enzymdesign wurde die Enantioselektivität der Hydroxylase umgekehrt und ihre Effizienz um das Dreifache gesteigert. Die Lösung der Kristallstruktur und MD-Simulationen der neuartigen (*R*)-Mandelatsynthase eröffnen Einblicke in die Struktur-Funktions-Beziehung der enantio-selektiven Katalyse am MNHE-Zentrum.

More than just a halogenase: modification of fatty acyl-moieties by a trifunctional metal center

Sarah M. Pratter[a], Jakov Ivkovic[b], Ruth Birner-Gruenberger[c], Rolf Breinbauer[b], Klaus Zangger[d] and Grit D. Straganz[a]*
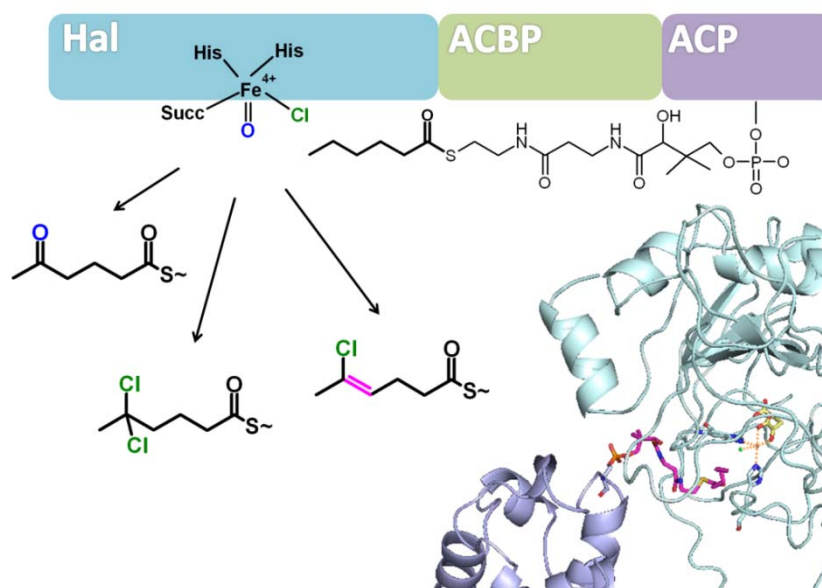
[a] Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12, A-8010 Graz, Austria.

[b] Institute of Organic Chemistry, Graz University of Technology, Stremayrgasse 9, A-8010 Graz, Austria.

[c] Institute of Pathology and Center of Medical Research, Medical University of Graz, Stiftingtalstraße 24, A-8010 Graz, Austria.

[d] Institute of Chemistry, University of Graz, Heinrichstraße 28, A-8010 Graz, Austria.

corresponding author, e-mail: grit.straganz@tugraz.at, Tel: +43-316-8738414 Fax: +43-316-8738434

Abstract

The highly selective oxidative halogenations by nonheme iron and α-ketoglutarate dependent enzymes are key reactions in the biosynthesis of lipopeptides and often bestow the metabolite's valuable bioactivity. Here we present the first biochemical characterization of an exponent from the group of fatty acyl-halogenases, HctB, from the hectochlorin biosynthetic pathway in *Lyngbya majuscula*. Its unprecedented three domain structure, which comprises an acyl-carrier protein domain, allows self-sufficient conversion of the covalently tethered hexanoyl-substrate. Structural analysis of the native product via ^{13}C-NMR reveals high regioselectivity but considerable catalytic promiscuity, which challenges the classification of HctB as a primary halogenase: Along with the proposed dichlorination, HctB performs oxygenation and an unprecedented introduction of a vinyl-chloride functionality into the non-activated carbon chain. Enzymatic activity is switched on by high chloride concentrations and single turnover kinetic measurements suggest that chloride triggers the primary step of O_2 activation in the reaction mechanism. The enzyme shows activity towards covalently tethered C6-C14 acyl-moieties and the relaxed substrate specificity is discussed based on a structural model of the enzyme substrate complex. Results suggest that fatty acyl-halogenation can introduce considerable structural diversity into the biosynthesis of lipopeptides.

Introduction

Nonribosomal lipopeptides from bacterial, fungal or botanical sources with antimicrobial, anti-infective, and anticancer properties have been used as potent pharmaceuticals for decades. Due to their diversity, the exploration of these complex natural products is a popular and promising route for the discovery of new drugs. Many lipopeptides bear halogenated aliphatic moieties - modifications that strongly influence the compounds' bioactivities.[1] The molecular basis of this chemically challenging catalysis is a focus of current research.

Mononuclear nonheme iron (MNH) and α-ketoglutaric acid (α-KG) dependent halogenases are arguably the most potent halogenating enzymes in nature, due to their unique ability to selectively transform carbon centers in non-activated lipopeptide building blocks. MNH dependent halogenases are generally found in polyketide synthase (PKS) / non-ribosomal peptide synthetase (NRPS) pathways and act on substrates that are tethered to the thiolation domain of an appendant acyl- or peptidyl-carrier protein (ACP/PCP) via a phosphopantetheinyl (PPT) moietiy.[1] Since the first report of an α-KG dependent MNH halogenase[2] only a handful of representatives have been biochemically characterized.[1,3-5] Analysis of the structure of *Pseudomonas syringae* threonyl halogenase SyrB2 has established the generally accepted view that α-KG dependent halogenases provoke halogenations due to a structural variation that sets them apart from the highly homologous keto-acid dependent oxygenases: instead of the prototypical metal binding facial triad of oxygenases, which comprises 2 histidines and 1 aspartate/glutamate, halogenases display a small hydrophobic amino acid that substitutes the carboxylic acid moiety and provides space for a chloride atom to coordinate to the iron cofactor.[6] During the

reaction mechanism[1] O_2 is activated and decarboxylates α-KG. The resulting high valent Fe(IV)=O species abstracts a hydrogen atom from the substrate, and a Cl-Fe(III)-OH species forms, in which the chloride successfully competes with the hydroxy ligand for transfer to the substrate radical.[1] This competition strongly depends on the positioning of substrate relative to the Cl-Fe(III)-OH species.[7]

Biochemical characterizations of α-KG dependent halogenases have mainly targeted mono-domain enzymes that modify non-activated amino acyl-precursors. Another type of α-KG dependent halogenases has been reported, which is embedded in a PKS/NRPS multi-domain protein and acts on activated carbon atoms: The JamE-halogenase domain (JamE-Hal) in the jamaicamide pathway chlorinates the gamma-position of a 3-hydroxy-3-(4-pentynyl)-glutaryl substrate, followed by a dehydratase and a decarboxylase, which then transform the α-hydroxy halogen to a terminal vinyl-chloride moiety. The closely related halogenase domain, CurA-Hal, from the curacin A pathway, which is involved in a cryptic chlorination reaction and the synthesis of a cyclopropyl moiety, shows similar enzymatic activity.[3] Yet, several bioactive natural products with long halogenated saturated alkyl chains are known (Figure 1): Hectochlorin and the related compounds dolabellin and lyngbyabellin A and B all harbour 5,5-dichloro-hexanoyl moieties,[8] which suggests the dihalogenation of a fatty acyl-building block. This implies a so far uncharacterized halogenase subtype, one that chlorinates unactivated carbons in non-amino acyl-substrates. Genetic data of the hypothetical fatty acyl-halogenase HctB from the hectochlorin pathway in *Lyngbya majuscula* indicate an unusual three-domain structure.[8] The N-terminus bears the catalytic halogenase domain, which is followed by an acyl-Coenzyme A binding protein (ACBP) of yet unknown function and the substrate-anchoring ACP. This distinct three-modular organization may be a common feature of fatty acyl-

6

halogenases, but so far only one other gene of similar organization has been reported[9] and more data will be needed to substantiate this notion.
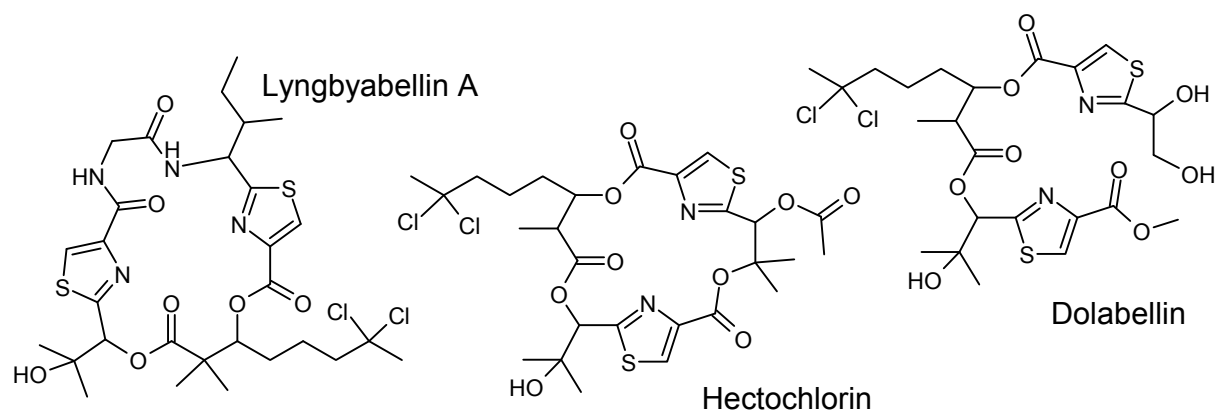
Figure 1. Examples of natural products with saturated halo fatty acyl-moieties.

Results and Discussion

The aim of this study was to explore the catalytic potential of the subfamily of MNH and α-KG dependent fatty acyl-halogenases that modify unactivated medium-chain fatty acyl-moieties. Native amino acyl- and keto acyl-substrates of so far characterized halogenases were converted with high regio- and chemoselectivity. Yet, the catalytic metal center of an α-KG dependent halogenase is in principle bifunctional and the strong preference of chlorination over oxygenation in halogenases is believed to be due to the substrate's precise positioning.[7] This appears to be at odds with the nature of unbranched, hydrophobic fatty acyl-moieties, which are more flexible than the small amino acid building blocks or the 3-hydroxy-3-acyl-glutaryl moieties. We were therefore particularly interested in the chemo- and

7

regioselectivities of fatty acyl-halogenases. The first exponent of these, for which the encoding gene cluster as well as the natural product are known, is HctB,[8] the target enzyme of this work.

Production and spectrometric characterization of functional HctB.

As the functional expression and reconstitution of multi-domain proteins presents special challenges, the purified enzyme was characterized regarding its secondary structure, its competence to bind divalent metal ions and the α-KG cofactor, as well as the efficiency of substrate loading. N- and C-terminal *Strep*-tag II full length versions of HctB were recombinantly expressed in *E. coli*. Soluble protein was obtained under an optimized procedure only, whereby host fermentation at 18°C and cell disruption with a Na_2CO_3 containing buffer were performed. SDS PAGE analysis confirmed HctB's purity (> 95%) and its theoretical molecular weight of 56.8 kDa (Figure S1). CD spectroscopy indicated properly folded protein with 24% α-helices, 25% β-sheets, 20% turns and 30% unordered regions (Figure S2). Due to higher protein yields (~ 40 mg L^{-1} cell culture), the C-terminally tagged form was used for all following experiments. Spectrophotometric iron determination revealed 4-8% Fe(II) in the isolated protein. The ability of HctB to bind divalent metals was probed via fluorescence titration of the enzyme with Cu(II) as described previously.[10] Substantial fluorescence quenching (~ 70%) and a typical saturation behavior were observed (Figure S3). The obtained K_d of 1.4 ± 0.7 µM confirmed the presence of a functional high affinity metal binding site in the protein preparation.

Enzymes from PKS/NRPS metabolic pathways typically require substrates that are covalently bound to a conserved serine residue, which is embedded in the thiolation domain of an ACP/PCP, via a PPT moiety. In the absence of an available native transferase we chose Sfp from *Bacillus subtili*s,[11] which is known to transfer a range of acyl-PPT moieties from respectively substituted Coenzyme A (CoA) analogues and covalently attach them to ACP domains that come from a variety of organisms. The PPT binding area of the ACP domain in HctB (YALGSAE) shows significant deviations from the consensus motif, (L/V)(G/L)(G/A/F/Y)(D/H/K/E)S(L/Q)(D/A/G).[12] Thus, in a first step, the principle ability of Sfp to transfer an acyl-PPT moiety onto HctB was assessed, using a fluorescence based method: Incubation of ApoHctB with Sfp and Bodipy-CoA, which Sfp can competently transfer, and subsequent analysis via SDS PAGE analysis demonstrated fluorescence labeling of HctB (Figure S4). Subsequently, the efficient modification of HctB with the PPT-form of the native hexanoyl-substrate was verified via MALDI-TOF MS analysis: The 450 Da mass shift from 56785 Da to 57235 Da after the Sfp reaction is in line with the theoretical mass shift (438 Da) that is expected upon addition of the hexanoyl-PPT moiety to ApoHctB (56778 Da) in order to give hexanoyl-PPTylated HoloHctB (57216 Da). No residual signal was found at the m/z of 56785 Da, which demonstrates an efficient transfer of the hexanoyl-PPT moiety to HctB (Figure 2).
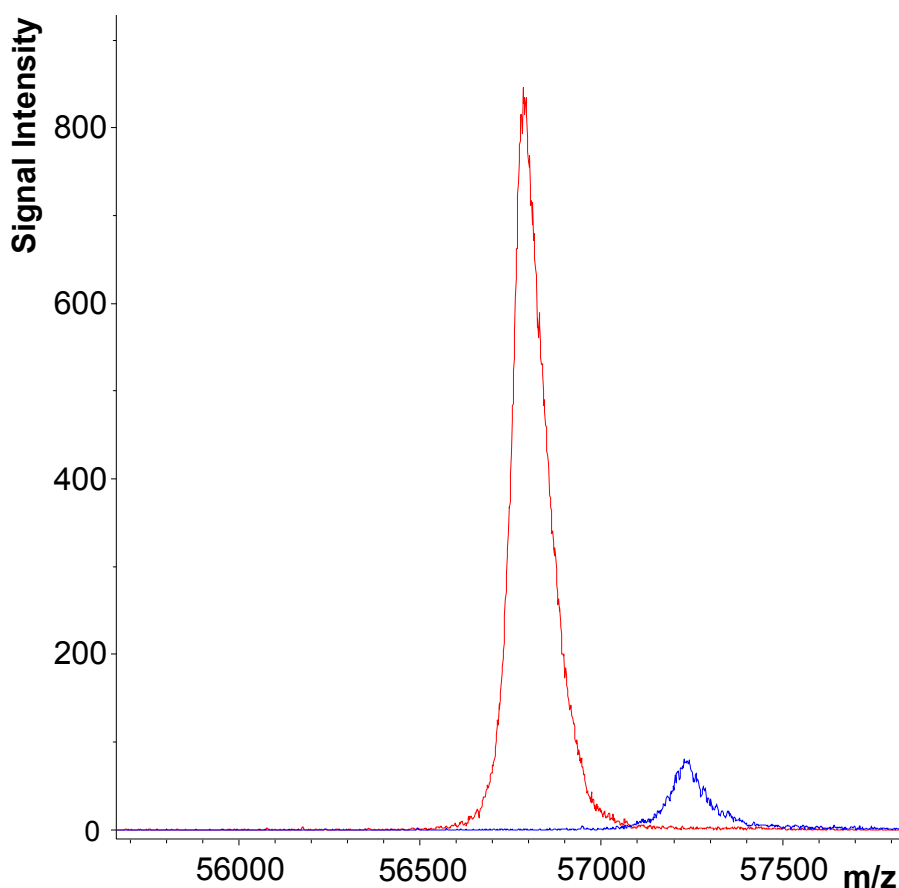
Figure 2. MADLI-TOF MS analysis of hexanoyl-PPT transfer onto HctB. Overlay and close-up of the mass spectra of ApoHctB (red) and HoloHctB after its reconstitution via Sfp and hexanoyl-CoA (blue).

Bidentate binding of the α-ketoacid cofactor to the Fe(II) site is a prerequisite for O_2 activation in α-KG dependent enzymes and can be characterized by the typical metal to ligand charge transfer (MLCT) band, which occurs at ~ 530 nm.[13] α-KG binding to the Fe(II) core in HoloHctB was affirmed by anaerobic UV/VIS spectrophotometry. The addition of 150 µM α-KG to 300 µM Fe(II)-loaded HoloHctB resulted in a broad absorbance band that was centered at ~ 500 nm. At equimolar amounts of α-KG and Fe(II)-HoloHctB an extinction coefficient of ε_{500} ~ 150 M^{-1} cm^{-1} was observed, which did not further increase upon addition of more α-KG. The band was present both in

the absence (Figure 3) and presence of NaCl (data not shown). This implies that the enzymatic Fe(II)-α-KG complex does not depend on the presence of chloride.
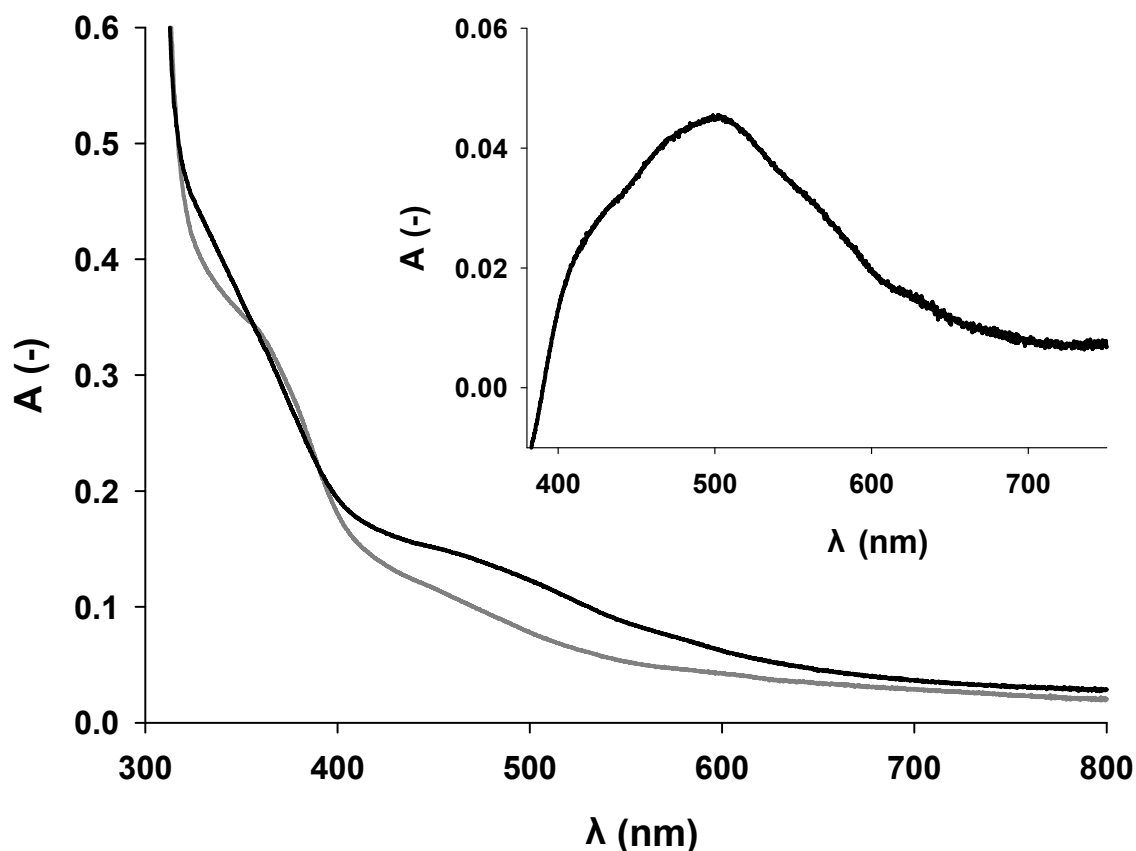
Figure 3. α-KG binding to HoloHctB. Spectra of 300 µM anaerobic *Fe(II)-HoloHctB* in the absence (gray) and presence (black) of 300 µM α-KG. Inset: The resulting difference spectrum depicts the MLCT band.

Characterization of HctB activity.

Halogenase activity was determined by monitoring O_2 consumption in air saturated enzyme preparations (260 µM O_2) over time. In the absence of Fe(II) or α-KG cofactors no O_2 depletion was detectable. In the presence of stoichiometric amounts of Fe(II) and α-KG, HoloHctB displayed elevated O_2 consumption rates that were

widely independent from further increased α-KG concentrations (Figure S5). Depending on the concentration of chloride, specific rates of O_2 depletion ranged from 0.42 ± 0.21 min^{-1} (< 0.1 mM NaCl) to 12.1 ± 4.1 min^{-1} (≥ 1 M NaCl) (Figure 4 a). Note that the specific rate determined under Fe(II), α-KG and chloride saturating conditions corresponds to the apparent first order rate constant of O_2 depletion of the single turnover trace ($k_{O2}{}^{app}$). The stoichiometry of O_2 consumption was characterized by determining the ratio of depleted O_2 and enzyme concentration. Under chloride saturating conditions α-KG addition and O_2 consumption were stoichiometric up to concentrations that corresponded to 1.9 ± 0.3 equivalents of HoloHctB and then reached a plateau (Figure S6). This is in line with a double oxidation of the substrate, whereby the employed HoloHctB is fully functional and O_2 consumption and substrate conversion are strongly coupled. Interestingly, full reaction rate and turnover were only achieved at NaCl concentrations > 500 mM, which indicates that high chloride concentrations, as present in sea water,[14] are needed to quantitatively activate HoloHctB.

In halogenases, the presence of substrate has been found to be crucial for O_2 activation.[15] To test for the necessity of substrate on HctB activity, the rates of O_2 consumption of ApoHctB were recorded and compared to those of HoloHctB. In the absence of chloride, ApoHctB showed strongly reduced activity (0.08 ± 0.06 min^{-1}). This basal rate was increased ~ 11-fold in the presence of ≥ 1 M chloride (0.87 ± 0.28 min^{-1}). By comparison, the presence of covalently linked substrate led only to a ~ 5-fold increase in O_2 consumption rates. When comparing Apo- and Holo-enzyme under NaCl saturation, covalently tethered substrate enhances the halogenase's activity ~ 14 times. By contrast, the addition of hexanoyl-CoA to ApoHctB did not significantly elevate O_2 consumption rates (1.15 ± 0.08 min^{-1}), indicating that covalent

linkage of the substrate to HctB is required for activity. The strong impact of the chloride concentration on single turnover O_2 consumption rates (Figure 4) suggests that not only substrate but also chloride is required for efficient primary O_2 activation in HctB. This is the first time that an impact of halide on the primary O_2 reduction rate has been reported and current investigations aim at clarifying by what molecular mechanism chloride triggers O_2 activity in halogenases.
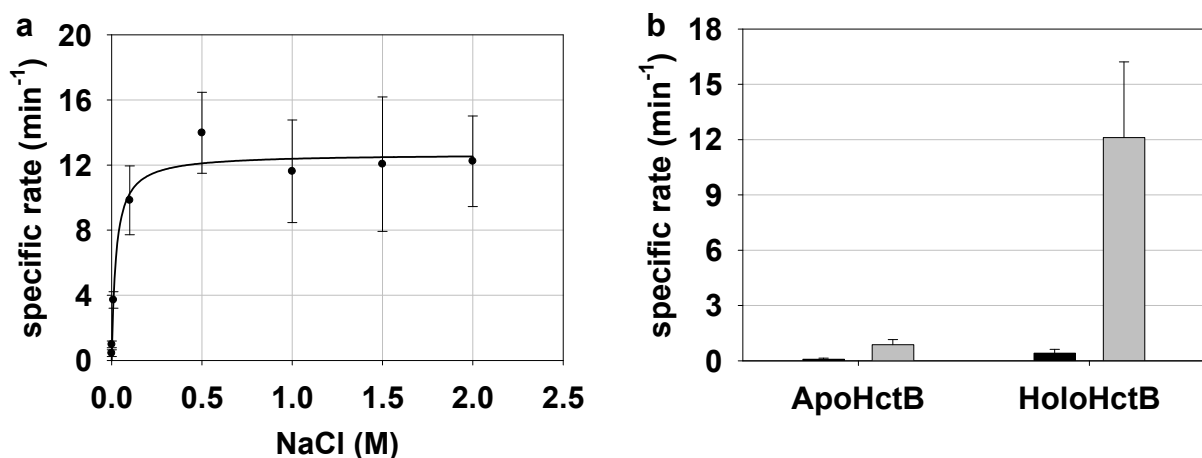
Figure 4. Dependency of HctB activity on the substrate and chloride concentrations under Fe(II) and α-KG saturated conditions. (a) Correlation of HoloHctB's specific rate on the NaCl concentration and (b) specific rates of ApoHctB and HoloHctB in the absence (≤ 1 mM, black) and presence (≥ 1 M, gray) of NaCl.

The halogenase and the ACP domain in HctB are separated by an ACBP of unknown function. To homologous ACBPs a role as an intracellular acyl-CoA transporter and pool former has been attributed and an involvement in the intermediary metabolism and gene regulation of various organisms has been proposed.[16] In order to test a putative role of this domain during catalysis, HoloHctB activity in the presence of a range of acyl-CoAs was investigated. However, addition of stoichiometric amounts of hexanoyl-CoA, octanoyl-CoA, decanoyl-CoA, lauryl-CoA, or myristoyl-CoA did not

significantly impact O_2 consumption characteristics of the enzyme (Note that MALDI-TOF MS data ruled out the presence of a putative tightly bound acyl-CoA ligand in HoloHctB preparations, which might have masked an impact of acyl-CoAs on the enzymatic activity in our assay).

Catalytic promiscuity.

MALDI-TOF MS analysis of HoloHctB revealed shifts of m/z ~ 70 and m/z ~ 160 after conversions with chloride and bromide, respectively, when compared to an unconverted HoloHctB control sample (Figure S7). Broadness of the obtained peaks suggested a mixture of indiscernible structures.

In order to qualitatively identify the native reaction products, HoloHctBs' acyl-moieties were hydrolyzed from their enzyme-bound PPT-arm after conversion, using the thioesterase TycF[17], and tested for modifications using a coupled HPLC-MS technique. Product solutions of NaCl-containing reactions, NaCl-free reactions, and unconverted Fe(II)- and α-KG- free controls were scanned for the m/z ratios of particular species. Samples from NaCl saturated HoloHctB conversions showed ions with m/z ratios of 129, 147 and 183, which correspond to oxo-hexanoate, monochloro-hexenoate and dichloro-hexanoate. No significant amount of hexanoate, which was the only species present in unconverted HoloHctB preparations, was found. In a conversion that lacked NaCl, hexanoate and oxo-hexanoate were detected. Surprisingly, also traces of monochloro-hexanoate, monochloro-hexenoate and dichloro-hexanoate were found, possibly due to traces of chloride in the enzyme preparations. Additionally a captured ion at 8.8 minutes in both NaCl free and NaCl saturated samples was consistent with a monohydroxy-hexanoate species (m/z = 131), but could not be unambiguously identified due to a lack of available standards.

Signals of the detected ions are presented in Figure S8. The presence of chloride in the reaction products was verified by its prototypical isotope pattern. The statistically expected molecular mass ratio for mono-chlorinated substances of 3.1:1 (^{35}Cl:^{37}Cl) and for di-chlorinated substances of 9.8:6.3:1 ($^{35}Cl_2$:$^{35}Cl^{37}Cl$:$^{37}Cl_2$) was found for the elution peaks at 33.8 min and 37.7 min and this confirmed their tentative identification as monochloro-hexenoate and dichloro-hexanoate, respectively (Figure S9).

To gain more structural information on the reaction products, acyl-moieties with a fully ^{13}C-labeled hexanoyl-backbone were employed. $^{13}C_6$-hexanoyl-CoA was obtained by oxidation of $^{13}C_6$-1-hexanol and subsequent, enhanced S-acylation to CoA by a known methodology[18] (Scheme S1) in a total yield of 36%. $^{13}C_6$-labeled products were subjected to NMR analysis in two ways, namely (i) still bound to HctB and, (ii) hydrolyzed and separated from the enzyme. A ^1H-^{13}C HSQC of the unconverted ^{13}C-labeled hexanoyl-moiety attached to HctB showed the expected broad NMR signals due to the sample's high molecular weight (Figure S10). After the enzymatic reaction a number of additional peaks was found, which could be assigned after cleavage from the enzyme using one-dimensional ^{13}C as well as two-dimensional ^1H-^{13}C HSQC and INADEQUATE spectra. The assignment was corroborated by comparison with ^{13}C and ^1H-^{13}C HSQC spectra of hexanoic acid, 5-oxo-hexanoic acid, 6-chloro-hexanoic acid and literature values for 5,5-dichlorohexanal.[19] For a semi-quantitative analysis of the relative amounts of the individual free products, ^{13}C spectra with a long relaxation delay (5s) were recorded and intensity comparisons were done on signals of the same multiplicity only. According to the obtained signals the amount of hexanoyl-substrate was ~ 10%, which indicates quantitative conversion. Identified reaction products confirmed results from HPLC-MS analysis and were 5-oxo-hexanoic acid (~ 50%), 5,5-

dichlorohexanoic acid (~ 25%) and 5-chloro-4-hexenoic acid (~ 25%). No signals for monohydroxy-hexanoate were found, which indicates that it is only a minor product component. A representative ^{13}C spectrum of free products is shown in Figure 5 a. Note that the observed carbon-carbon couplings due to the use of uniformly ^{13}C-labeled hexanoic acid precursors made the product peaks easily distinguishable from the strong singlets of the HEPES buffer. Products identified in the free form could also be found bound to the protein, where they show markedly increased line widths due to the significantly increased molecular weight of the enzyme. An overlay of an HSQC of free and HctB bound products is shown in Figure 5 b.

In summary, product analysis shows remarkable catalytic promiscuity. Intriguingly, beside dihalogenation and oxygenation, an unprecedented introduction of a vinyl-chloride functionality into the non-activated fatty acyl-chain is observed, whereby regioselectivity is pertained. This incorporation of a non-terminal vinyl-chloride moiety into a non-activated carbon chain by one enzyme contrasts the previously described MNH dependent methylene-chloride formation in jamaicamides,[3] which is accomplished by three enzymes and is restricted to activated substrates and the synthesis of terminal vinyl-chlorides.

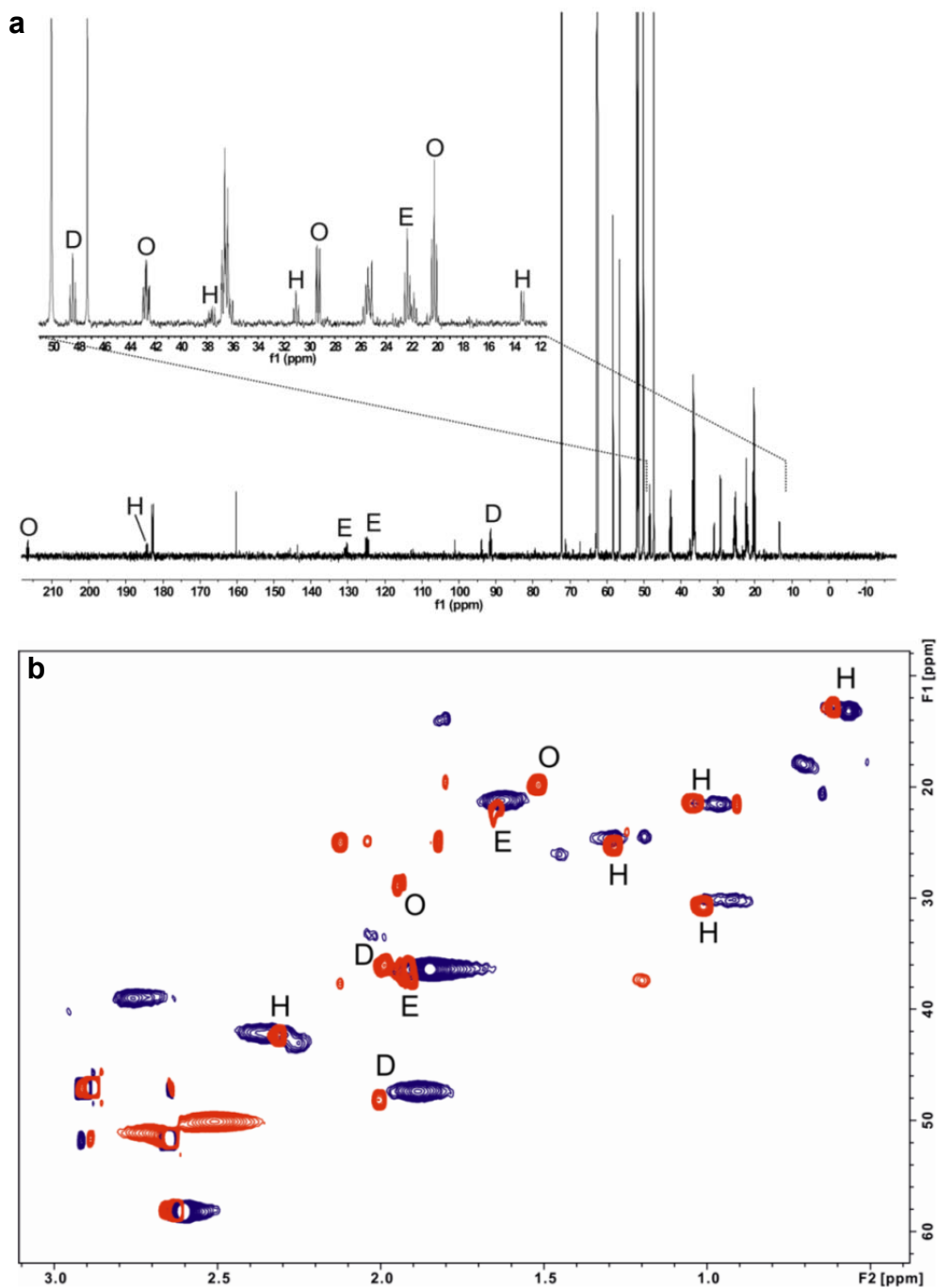*Figure 5. ^{13}C-NMR spectra of the HctB reaction products.* Panel a shows an one-dimensional ^{13}C NMR spectrum of free hexanoyl-products. Non overlapped signals of the individual products are labeled by H (hexanoic acid), O (5-oxo-hexanoic acid), E (5-chloro-4-hexenoic acid) and D (5,5-dichlorohexanoic acid). Panel b shows an overlay of ^{1}H-^{13}C HSQC spectra of free products (red) and products bound to HctB (blue).

Substrate and halogen promiscuity.

HctB displayed a remarkable promiscuity regarding alternative fatty acyl-substrates. HoloHctB variants that carried an octanoyl-, decanoyl-, lauryl- or myristoyl-group instead of the native hexanoyl-substrate showed considerable O_2 consumption rates. However, rates decreased with growing chain length, resulting in a ~ 5-fold drop for the myristoyl moiety (2.4 ± 1.0 min^{-1}), compared to the native substrate (Figure 6 a). O_2 consumption stoichiometry did not change significantly (Figure 6 b). The principle potential of HctB for bromination in the presence of 1 M KBr was confirmed in analogy to amino acyl-halogenases,[20] whereby rates were ~ 4 times slower (3.0 ± 1.1 min^{-1}) than for the respective chlorination reaction and O_2 consumption stoichiometry suggested < 2 conversions (Figure S11).
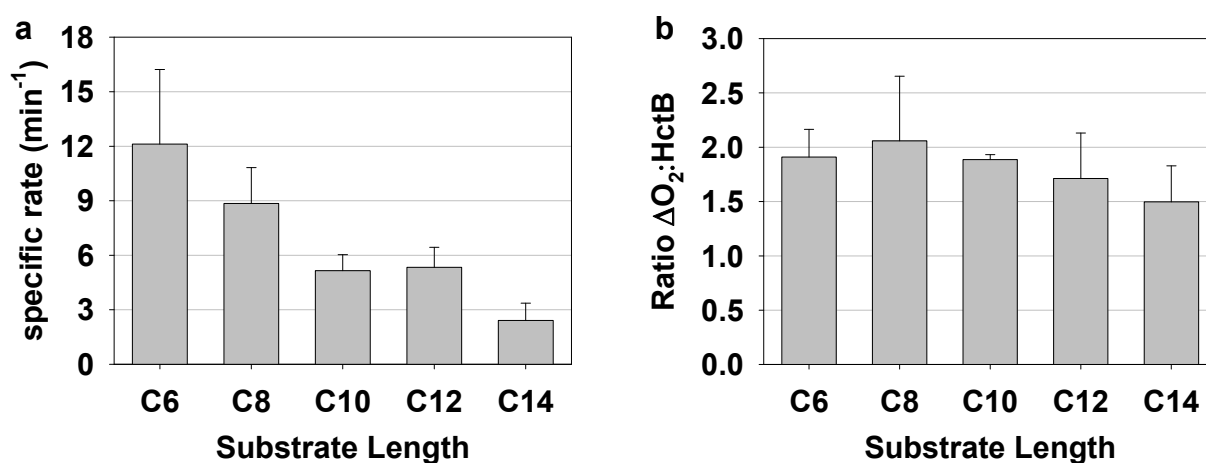

Figure 6. Activity of HctB towards putative alternative fatty acyl-substrates. (a) Specific rates and (b) O_2 consumption stoichiometry of hexanoyl- (C6), octanoyl- (C8), decanoyl- (C10), lauryl- (C12), and myristoyl- (C14) HoloHctB using cofactor saturating conditions are shown.

In order to rationalize the surprisingly relaxed substrate spectrum, a structural model of the HctB halogenase domain was generated and subjected to *in silico* docking of acyl-PPT moieties. The C4-C6 carbon atoms of the native substrate were rather close to the chloride ion (3.5-4.5 Å) and the iron cofactor (6-7 Å). Similar low energy positioning in vicinity to the catalytic metal center was found for all docked non-native fatty acyl-moieties (Figure 7). This is in good agreement with experimental data, which suggest that C8-C14 acyl-chains induce primary O_2 activation. In halogenases substrate binding has been found to remove the sixth ligand from the iron, presumably a water molecule that is displaced by the vicinity of substrate to the metal ion. This leads to a 5-coordinate iron center, which is competent to reduce O_2. In a separate approach docking of the ACP domain to the halogenase moiety resulted in a low energy form, where the conserved Ser 427 of the thiolation domain was located in vicinity (< 6 Å) to the PPT phosphate moieties. The resulting two-domain model contradicts a previously proposed role of the ACBP domain in substrate binding during catalysis.[9]
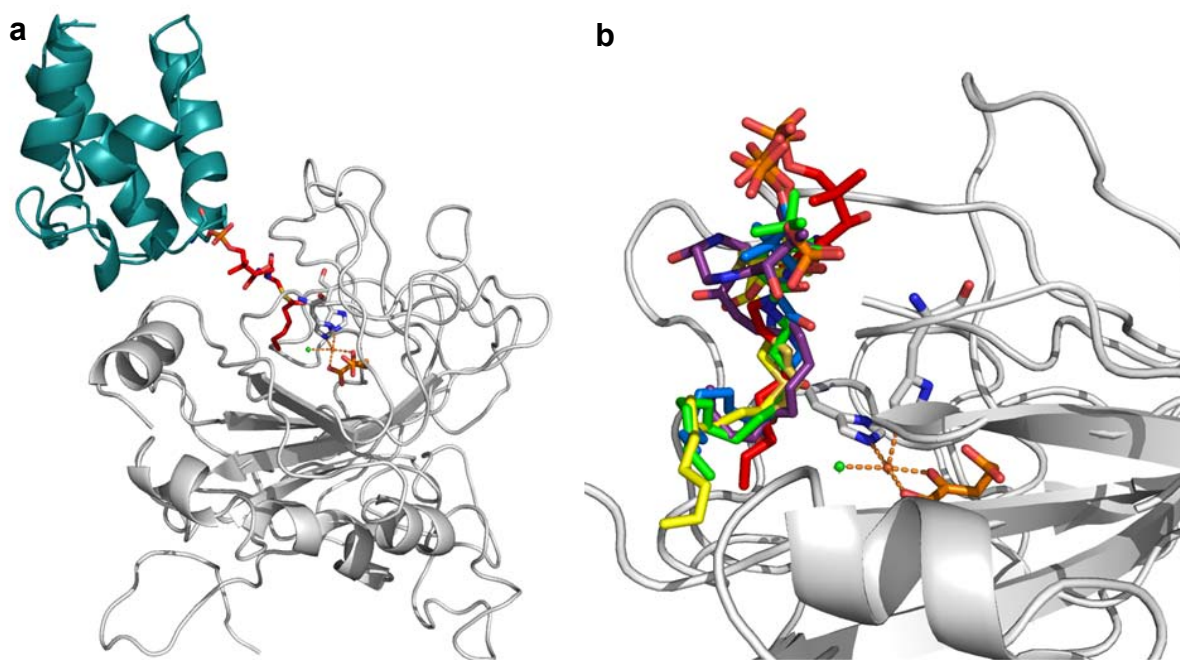
Figure 7. Structural model of the HctB halogenases-domain-ACP-substrate complex and truncated HctB halogenases-domain-fatty-acyl-complexes. Panel a shows the overall structure of a halogenases- (gray) and ACP- (teal) domain model with the native C6-substrate after covalent linkage of the hexanoyl-PPT (red) and ACP domain and its relaxation by MD simulation. Panel b: C6 (red), C8 (marine), C10 (green), C12 (yellow), and C14 (violet) PPT-acyl-substrates were docked into a HctB halogenase-domain model containing the metal coordinating histidines (gray sticks), the iron cofactor (orange ball), a chloride atom (green ball) and the α-KG cofactor (orange sticks). A close-up of the docking results is shown, whereby residues 1-14 and 223-226 were omitted for clarity.

Conclusion

Multifunctionality, including desaturase activity, has already been reported for α-KG dependent MNH enzymes, such as for clavaminate synthase,[21] but it has so far not been shown for α-KG dependent halogenases, which according to previous reports chemo- and regioselectively perform halogenation of their native substrates. By contrast, our study implies that HctB inserts considerable structural variety into the primary building block of hectochlorin by dichlorination, oxygenation and vinyl-chloride formation, whereby the presumed C6 specificity is probably afforded by the preceding acyl-ACP-synthetase, HctA. The introduction of a vinyl-chloride moiety into a non-activated carbon chain is to our knowledge unprecedented. The product pattern of HctB under chloride saturating conditions displays chloride and oxo-moieties exclusively in the C5 position and the double bond in position C4-C5. Also, α-KG consumption and O_2 depletion are strongly coupled, but it does not exceed two turnovers. Based on these findings we propose a reaction mechanism where a primary transfer of either a metal bound hydroxy or chloride species at C5 takes place. In a second oxidation cycle the remaining hydrogen at the C5 is abstracted. The comparably stable radical may then undergo a second chlorination/hydroxylation round or, alternatively, the intermediate Cl-Fe(III)-OH moiety may abstract a hydrogen atom from C4, resulting in the three observed reaction products (Scheme 1).

In summary, the catalytic properties of HctB infer that the product spectrum of the hectochlorin pathway in *L. majuscula* comprises structural variants with dichloro-, oxo-, and vinyl-chloride alkyl chains and, by extension, that exponents of the fatty acyl-halogenase group may generally display all three functionalities in comparable

measure. This is in line with the observation that lipopeptide pathways in nature often employ promiscuous enzymes that generate a range of reactions products, which are then assembled to a variety of structures. The resulting diversity of the secondary metabolites increases the chances of producing a bioactive variant that bestows evolutionary fitness to the producer.[22]
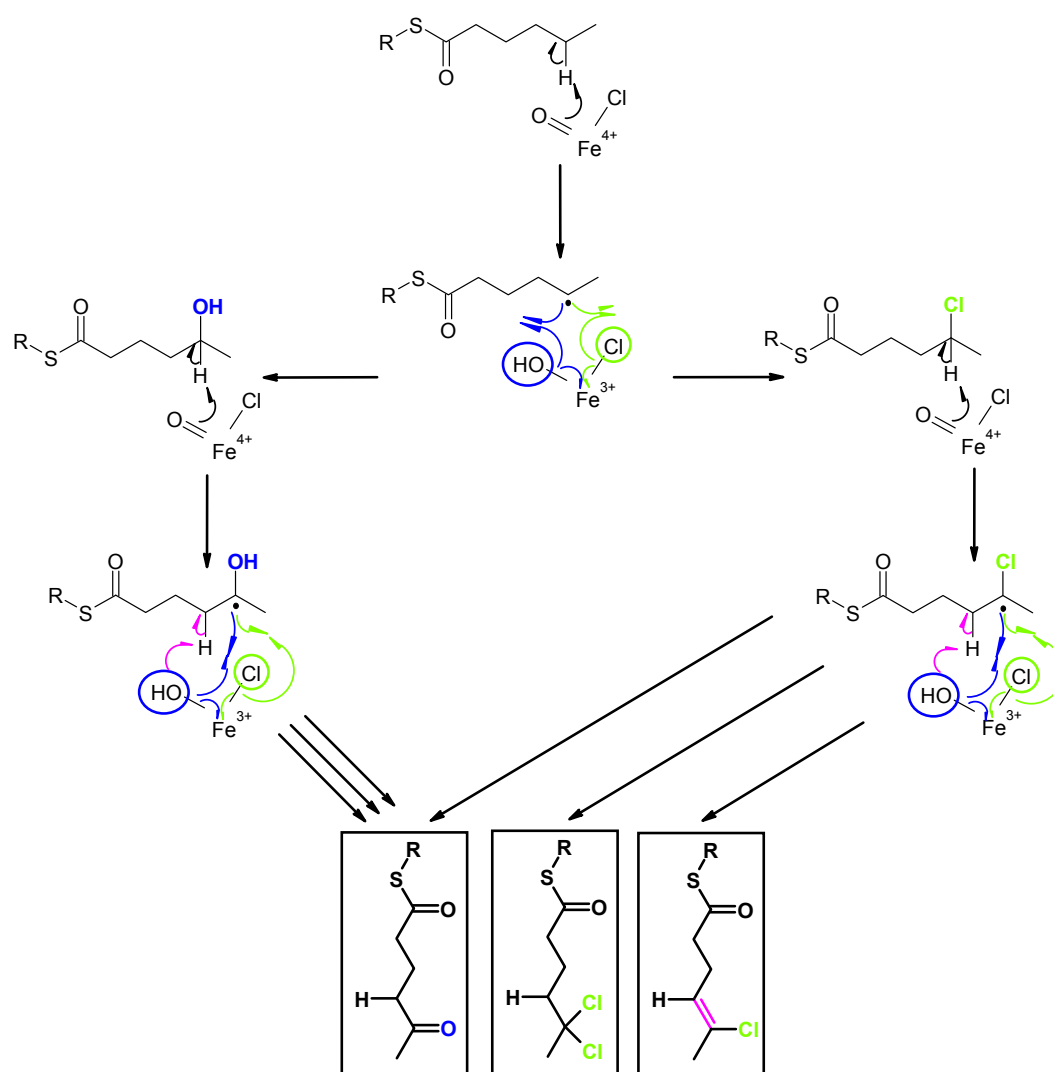
Scheme 1: Proposed mechanism of product formation in HctB. In the first oxidation cycle, the hexanoyl-substrate is chlorinated or hydroxylated at C5. In the second round of modification, again, a radical at the C5 position is generated, which either rebinds the chloride or the hydroxyl moiety to yield dichloro- or oxo-hexanoate, or abstracts a proton from the C4 position, which results in the vinyl-chloride product.

Materials and Methods

Chemicals. If not otherwise stated, chemicals were purchased from Sigma Aldrich or Carl Roth at the highest purity available. Phusion DNA Polymerase, *pKYB*1 vector and T4 DNA ligase were from New England BioLabs and restriction enzymes and buffers, and alkaline phosphatase from Fermentas.

Cloning. The HctB encoding DNA sequence (NCBI database AAY42394), which was C-terminally *Strep*-tag II tagged and codon optimized for expression in *E. coli*, and the C-terminally 6xHis tagged sequence of the PPT-transferase Sfp from *Bacillus subtilis*[11] (NCBI database AEK64474.1) were both purchased from Mr. Gene GmbH and cloned into the *pKYB*1 vector. The N-terminally *Strep*-tag II tagged *hctB* construct was prepared by cloning into the pASK-IBA7plus vector (IBA BioTAGnology). Sequence correctness of the constructs was confirmed by Eurofins MWG Operon.

Protein expression. Plasmids were introduced into *E. coli* BL21 Gold (DE3) cells by electroporation. HctB expression was performed in a medium that had previously been optimized for the expression of a MNH dioxygenase.[23] 300 ml main cultures were inoculated with an over-night culture to give an OD_{600} of 0.1-0.2. Cells were grown at 37°C and 110 rpm to an OD_{600} of ~ 0.7. Protein expression was induced at 18°C by addition of 100 µg ml^{-1} IPTG. After incubation at 18°C for 22 hours cells were harvested and resuspended in AP2 buffer (50 mM Tris, 150 mM NaCl, 100 mM Na_2CO_3, 5% glycerol, pH 8.5). Sfp and TycF were produced in 2xYT broth. Sfp was expressed according to the HctB protocol but at 25°C and using TMGD buffer[11] at a pH of 6.9. TycF was expressed as previously described.[17]

Protein purification. Host cells were disrupted by two passages of french pressing at approximately 24,000 psi. After removal of cell debris (4°C, 78000 rcf, 35 min), soluble *Strep-tag* II HctB protein was purified on a *Strep*-Tactin Superflow column (IBA BioTAGnology) according to the manufacturers' instructions and buffer exchanged into 20 mM HEPES pH 7.5. His-tagged Sfp and TycF proteins were purified via affinity chromatography using a Cu-chelat column and a stepwise imidazole gradient (50 mM Tris, 300 mM NaCl, pH 7.5, 0-400 mM imidazole). Protein samples were buffer exchanged into 20 mM HEPES pH 7.5 (TycF) or TMGD buffer pH 6.9 (Sfp). Cell disruption, affinity chromatography, buffer exchange and concentration of protein samples (in VivaSpin 20 concentrators from Sartorius) were performed at 4°C. The protein concentration of HctB was determined according to Edelhoch and Pace, as described previously[24]. Concentrations of Sfp and TycF proteins were determined by the Bio-Rad Protein Assay. The purity of the isolated enzymes and their molecular weights were checked via SDS-PAGE (GE Healthcare; Figures S1 and S12).

Fe(II) content. The quantity of Fe(II) in HctB preparations was determined spectrophotometrically at 25°C in 20 mM Tris/HCl pH 7.5 according to Hennessy *et al*[25] using $\geq$ 350 fold molar excess of FereneS. The detachment of iron from the native protein and concomitant formation of the colored Fe(II)-FereneS complex (ε = 35.5 mM^{-1} cm^{-1}) was monitored at 592 nm over 15-18 hours. Blanks measured in the absence of protein were subtracted.

Circular dichroism (CD) spectroscopy. Far-UV CD spectra were recorded with a Jasco J-715 spectropolarimeter at room temperature using a 0.02 cm pathlength cylindrical cell and instrument parameters as described previously.[26] Enzymes dissolved in 20 mM potassium phosphate buffer (pH 7.5) to a concentration of ~ 0.4

mg ml^{-1} were analysed. Three spectra obtained in the wavelength range of 300 to 185 nm were averaged and corrected by a blank spectrum lacking the enzyme, before the CD signal was converted into mean residue ellipticity. Data were processed using the DICHROWEB program.[27]

Fluorescence titration. Experiments were performed with 1 µM enzyme dissolved in 20 mM Tris/HCl buffer pH 7.5. CuCl$_2$ stock solutions were added in incremental amounts to give final concentrations of 0.5-300 µM Cu(II). Fluorescence data was collected at 25 °C in a Hitachi F-4500 fluorescence spectrophotometer (Hitachi High-Technologies). Samples were excited at 290 nm and emission was recorded between 310 and 500 nm at a scan speed of 1200 nm min^{-1}. For each CuCl$_2$ concentration the emission intensity at 340 nm (maximum fluorescence) of three replicates was averaged. Quenching of the signal was calculated by the following formula:

$$Q = \frac{(I_0 - I * f)}{I_0} \tag{1}$$

I$_0$ and I are the fluorescence signals in the absence and presence of Cu(II) and f is the dilution factor of the enzyme due to addition of titrand. Data from individual titration experiments were fit with the program SigmaPlot 9.0 (Systat Software Inc.) according to the following equation:

$$y = \frac{Q_{max} * x}{(K_d + x)} \tag{2}$$

K_d values of eight titration experiments were averaged.

Reconstitution of HoloHctB. Aliquots of 200-500 nmol ApoHctB were mixed with 0.1 equivalents of Sfp, 2 equivalents of the respective fatty acyl-CoA (C6-C14) and transfer buffer (10 mM $MgCl_2$, 75 mM Tris/Cl pH 7.5 final concentrations), resulting in a concentration of 200 to 500 µM HctB. Reaction samples were incubated at 24°C and 350 rpm for 1-1.5 hours on a thermo-shaker. Precipitated protein was removed by centrifugation prior to buffer exchange to either 20 mM Tris/H_2SO_4, pH 7.5 or 20 mM HEPES/NaOH, pH 7.5. HoloHctB samples were concentrated ~ 2-5 fold and protein concentrations were reassessed.

Bodipy-CoA transfer assay. Bodipy-CoA was prepared according to La Clair *et al*[28] and was then transferred as described for the fatty acyl-substrates. Samples were subjected to SDS PAGE analysis. The protein attached fluorescent dye was excited at 532 nm and recorded at 555 nm using a long pass filter on a Bio-Rad scanner.

Mass spectrometry. Protein solutions were desalted by repeated cycles of concentration and dilution in 0.1% trifluoroacetic acid using 10 kDa cut off ultrafiltration tubes (Amicon), diluted 1:10 (v:v) in 50% acetonitrile and 0.1% trifluoroacetic acid and mixed 1:1 (v:v) with a suspension of 5 mg super-DHB matrix (9:1 mixture of 2,5-dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid) in 50µL acetonitrile and 50 µL 0.1% trifluoroacetic acid. 1 µL was spotted onto an 800 µm MTP Anchorchip (Bruker) and left to dry. MALDI-TOF-MS measurements were performed on a Bruker Ultraflextreme instrument. Full MS spectra were obtained in positive linear mode after external calibration using an appropriate protein standard (10-70 kDa) in the mass/charge range from 8000 to 77000.

α-KG binding. Anaerobic HoloHctB samples (300 µM) in the absence or presence of 1 M NaCl were prepared in 20 mM HEPES/NaOH buffer, pH 7.5, in an air-tight Wheaton®-vial by purging the vial with N_2. Residual O_2 was consumed by adding an equimolar amount of acetylacetone, and depleting it with Dke1 (4-5 µM).[26] N_2-flushed, O_2-free Fe(II)- and α-KG cofactors were added as stock solutions in microliter quantities in a N_2-purged glove box to yield 300 µM Fe(II) and 0 µM / 150 µM / 300 µM / 600 µM α-KG. Spectra were recorded from 250 to 800 nm on a CARY 50 Bio UV-Visible Spectrophotometer (Varian). Three replicates were performed at a scan rate of 1200 nm min^{-1} for each condition and averaged for analysis.

HctB activity assay. Apo- and HoloHctB activities were recorded by monitoring O_2 consumption over time with a luminescence based Oxygen Microoptode (Presens) in an air-tight cell. Assays containing 100 µM enzyme were performed under stirring at 23°C in a total volume of 500 µl in air-saturated 20 mM Tris/H_2SO_4 or 20 mM HEPES/NaOH buffers pH 7.5, supplemented with 0-2 M NaCl or 1 M KBr. α-KG was added as an air-saturated aqueous stock solution (5 – 100 mM) to a final concentration of 50-1000 µM, while Fe(II) was added as an anaerobic aqueous 15 mM stock solution (150 µM final concentration). Apparent first order rate constants (k_{O2}^{app}) of single turnover kinetics at 260 µM O_2 were calculated based on fitting the data to a single **(3)** or double **(4)** exponential decay using Sigma Plot 9.0.

$$f = y_0 + a * e^{(-b*x)} \qquad \textbf{(3)}$$

$$f = y_0 + a * e^{(-b*x)} + c * e^{(-d*x)} \qquad \textbf{(4)}$$

y_0 is the final O_2 concentration after conversion, a and c are the amplitudes of the curve and b and d are the respective k_{O2}^{app} values. The determined rates were validated by using the initial reaction rate divided by the concentration of HctB active

sites. In the same way specific rates were calculated when reaction rates were too slow to allow the observation of complete signal decay and, consequently, fitting of the traces to equations **(3)** or **(4)**.

Isolation of reaction products. HoloHctB samples from activity assays were pooled, buffer exchanged into 20 mM HEPES/NaOH pH 7.5 and concentrated to ~ 1 mM. Acyl-moiety cleavage from the enzyme's ACP domain was accomplished by addition of the thioesterase TycF[17] (15-30 µM) and subsequent incubation at 30°C and 350 rpm overnight. Precipitated protein was removed by centrifugation. Supernatants were subjected to a VivaSpin 500 centrifugal filter (Sartorius) and the filtrate, which contained the cleaved carbonic acid, was collected. Samples for ^{13}C-NMR analysis were supplemented with 10% deuterium oxide.

HPLC-MS analysis. Reaction products, which had been cleaved from the enzyme, as well as standard solutions of hexanoic acid, 5-oxo-hexanoic acid and 6-chloro hexanoic acid were acidified with 1% formic acid (FA) (v/v). HPLC analysis was carried out using an Agilent Technologies system, equipped with a LiChroCART Purospher STAR column, at 30°C at a flow rate of 0.8 ml min^{-1} using 1% FA in water (solvent A) and 1% FA in acetonitrile (solvent B) according to the following protocol: 10 min isocratic flow of 10% B, a 30 min gradient from 10 % to 50% B, a 1 min gradient from 50 % to 90% B, 3 min isocratic flow of 90% B, a 1 min gradient of 90 % to 10% B and 10 min isocratic flow of 10% B. Separated components were screed for particular m/z ratios of the substrate and potentially formed reaction products using ESI-MS single ion monitoring (SIM) in a negative ion detection mode.

Synthesis of $^{13}C_6$-labeled acyl-CoA. The synthesis of $^{13}C_6$-hexanoyl-CoA was started from commercially available $^{13}C_6$-1-hexanol. The microscale model synthesis was developed first with non-labeled 1-hexanol. $^{13}C_6$-hexanoyl-CoA was obtained in total yield of 36% by oxidation of $^{13}C_6$-1-hexanol[29] and subsequent, enhanced *S*-acylation to CoA by a known methodology[18] using PyBOP as a coupling agent (Scheme S1).

NMR analysis. ^{13}C-labeled enzyme-product complexes and hydrolyzed reaction products were analysed on a Bruker Avance III 700 MHz NMR spectrometer equipped with a 5 mm TCI cryo probe at 298 K. For the one-dimensional ^{13}C NMR spectra 16384 scans were acquired in 8.5 hours and the ^1H-^{13}C HSQCs were recorded with 16 scans for each of the 128 increments. For the INADEQUATE spectra 512 scans were accumulated for each of the 64 transients.

HctB structural modeling and in silico docking. A structural model of the halogenase domain of HctB with (i) the ACP domain and (ii) a range of truncated fatty acyl-substrates was generated by a combination of *in silico* docking and MD simulations. Details are given in the Supporting Information.

Acknowledgements

References

1. Neumann, C.S., Fujimori, D.G., and Walsh, C.T. (2008) Halogenation strategies in natural product biosynthesis, *Chem. Biol. 15*, 99-109.

2. Vaillancourt, F.H., Yeh, E., Vosburg, D.A., O'Connor, S.E., and Walsh, C.T. (2005) Cryptic chlorination by a non-haem iron enzyme during cyclopropyl amino acid biosynthesis, *Nature 436*, 1191-1194.

3. Gu, L., Wang, B., Kulkarni, A., Geders, T.W., Grindberg, R.V., Gerwick, L., Håkansson, K., Wipf, P., Smith, J.L., Gerwick, W.H., and Sherman, D.H. (2009) Metamorphic enzyme assembly in polyketide diversification, *Nature 459*, 731-735.

4. Jiang, W., Heemstra, J.R., Forseth, R.R., Neumann, C.S., Manaviazar, S., Schroeder, F.C., Hale, K.J., and Walsh, C.T. (2011) Biosynthetic chlorination of the piperazate residue in kutzneride biosynthesis by KthP, *Biochemistry 50*, 6063-6072.

5. Neumann, C.S. and Walsh, C.T. (2008) Biosynthesis of (-)-(1*S*,2*R*)-allocoronamic acyl-thioester by an Fe(II)-dependent halogenase and a cyclopropane-forming flavoprotein, *J. Am. Chem. Soc. 130*, 14022-14023.

6. Blasiak, L.C., Vaillancourt, F.H., Walsh, C.T., and Drennan, C.L. (2006) Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis, *Nature 440*, 368-371.

7. Matthews, M.L., Neumann, C.S., Miles, L.A., Grove, T.L., Booker, S.J., Krebs, C., Walsh, C.T., and Bollinger, J.M. (2009) Substrate positioning controls the partition between halogenation and hydroxylation in the aliphatic halogenase, SyrB2, *Proc. Natl. Acad. Sci. USA. 106*, 17723-17728.

8. Ramaswamy, A.V., Sorrels, C.M., and Gerwick, W.H. (2007) Cloning and biochemical characterization of the hectochlorin biosynthetic gene cluster from the marine cyanobacterium *Lyngbya majuscula*, *J. Nat. Prod. 70*, 1977-1986.

9.   Nishizawa, A., Arshad, A.B., Nishizawa, T., Asayama, M., Fujii, K., Nakano, T., Harada, K., and Shirai, M. (2007) Cloning and characterization of a new hetero-gene cluster of nonribosomal peptide synthetase and polyketide synthase from the cyanobacterium *Microcystis aeruginosa* K-139, *J. Gen. Appl. Microbiol. 53*, 17-27.

10.   Straganz, G.D., and Nidetzky, B. (2005) Reaction coordinate analysis for β-diketone cleavage by the non-heme Fe^{2+}-dependent dioxygenase Dke1, *J. Am. Chem. Soc. 127*, 12306-12314.

11.   Quadri, L.E.N., Weinreb, P.H., Lei, M., Nakano, M.M., Zuber, P., and Walsh, C.T. (1998) Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases, *Biochemistry 37*, 1585-1595.

12.   Lambalot, R.H., Gehring, A.M., Flugel, R.S., Zuber, P., LaCelle, M., Marahiel, M.A., Reid, R., Kohsla, C., and Walsh, C.T. (1996) A new enzyme superfamily - the phosphopantetheinyl transferases, *Chem. Biol. 3*, 923-936.

13.   Ryle, M.J., Padmakumar, R., and Hausinger, R.P. (1999) Stopped-flow kinetic analysis of *Escherichia coli* taurine/alpha-ketoglutarate dioxygenase: interactions with alpha-ketoglutarate, taurine, and oxygen, *Biochemistry 38*, 15278-15286.

14.   Lefort-Tran, M., Pouphile, M., Spath, S., and Packer, L. (1988) Cytoplasmic membrane changes during adaptation of the fresh water cyanobacterium *Synechococcus* 6311 to salinity, *Plant Physiol. 87*, 767-775.

15.   Neidig, M.L., Brown, C.D., Light, K.M., Fujimori, D.G., Nolan, E.M., Price, J.C., Barr, E.W., Bollinger, J.M., Krebs, C., Walsh, C.T., and Solomon, E.I. (2007) CD and MCD of CytC3 and taurine dioxygenase: role of the facial triad in α-KG-dependent oxygenases, *J. Am. Chem. Soc. 129*, 14224-14231.

16.   Knudsen, J., Mandrup, S., Rasmussen, J.T., Andreasen, P.H., Poulsen, F., and Kristiansen, K. (1993) The function of acyl-CoA-binding protein (ACBP)/diazepam binding inhibitor (DBI), *Mol. Cell. Biochem. 123*, 129-138.

17.	Yeh, E., Kohli, R.M., Bruner, S.D., and Walsh, C.T. (2004) Type II thioesterase restores activity of a NRPS module stalled with an aminoacyl-S-enzyme that cannot be elongated, *ChemBioChem 5*, 1290-1293.

18.	Zhang, W., Bolla, M.L., Kahne, D., and Walsh C.T. (2010) A three enzyme pathway for 2-amino-3-hydroxycyclopent-2-enone formation and incorporation in natural product biosynthesis, *J. Am. Chem. Soc. 132*, 6402-6411.

19.	Sone, H., Kondo, T., Kiryu, M., Ishiwata, H., Ojika, M., and Yamada, K.J. (1995) Dolabellin, a cytotoxic bisthiazole metabolite from the sea hare *Dolabella auricularia*: structural determination and synthesis, *Org. Chem. 60*, 4774-4781.

20.	Vaillancourt, F.H., Vosburg, D.A., and Walsh, C.T. (2006) Dichlorination and bromination of a threonyl-S-carrier protein by the non-heme FeII halogenase SyrB2, *ChemBioChem 7*, 748-752.

21.	Salowe, S.P., Krol, W.J., Iwata-Reuyl, D., and Townsend, C.A. (1991) Elucidation of the order of oxidations and identification of an intermediate in the multistep clavaminate synthase reaction, *Biochemistry 30*, 2281-2292.

22.	Firn, R.D. and Jones, C.G. (2003) Natural products - a simple model to explain chemical diversity, *Nat. Prod. Rep. 20*, 382-391.

23.	Straganz, G.D., Slavica, A., Hofer, H., Mandl, U., Steiner, W., and Nidetzky, B. (2005) Integrated approach for production of recombinant acetylacetone dioxygenase from *Acinetobacter johnsonii*, *Biocatal. Biotransform. 23*, 261–269.

24.	Straganz, G.D., Diebold, A.R., Egger, S., Nidetzky, B., and Solomon, E.I. (2010) Kinetic and CD/MCD spectroscopic studies of the atypical, 3-His ligated, non-heme Fe^{2+} center in diketone dioxygenase: the role of hydrophilic outer shell residues in catalysis, *Biochemistry 49*, 996–1004.

25.	Hennessy, D.J.S., Reid, G.R., Smith, F.E., and Thompson, S.L. (1984) Ferene – a new spectrophotometric reagent for iron, *Can. J. Chem. 62*, 721-724.

26. Leitgeb, S., Straganz, G.D., and Nidetzky, B. (2009) Biochemical characterization and mutational analysis of the mononuclear non-haem Fe^{2+} site in Dke1, a cupin-type dioxygenase from *Acinetobacter johnsonii*, *Biochem. J. 418*, 403-411.

27. Whitmore, L. and Wallace, B.A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, *Nucleic Acids Res. 32*, W668–W673.

28. La Clair, J.J., Foley, T.L., Schegg, T.R., Regan, C.M., and Burkart, M.D. (2004) Manipulation of carrier proteins in antibiotic biosynthesis, *Chem. Biol. 11*, 195-201.

29. Zhao, M., Li, J., Song, Z., Desmond, R., Tschaen, D.M., Grabowski, E.J.J., and Reider, P.J. (1998) A novel chromium trioxide catalyzed oxidation of primary alcohols to the carboxylic acids, *Tetrahedron Lett. 39*, 5323-5326.

Supporting Information:

More than just a halogenase: modification of fatty acyl-moieties by a trifunctional metal center

Sarah M. Pratter[a], Jakov Ivkovic[b], Ruth Birner-Grünberger[c], Rolf Breinbauer[b], Klaus Zangger[d] and Grit D. Straganz[a]*

[a] Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12, A-8010 Graz, Austria

[b] Institute of Organic Chemistry, Graz University of Technology, Stremayrgasse 9, A-8010 Graz, Austria

[c] Institute of Pathology - Core Facility Mass Spectrometry, Medical University of Graz, Stiftingtalstraße 24, A-8010 Graz, Austria

[d] Institute of Chemistry, University of Graz, Heinrichstraße 28, A-8010 Graz, Austria

corresponding author, e-mail: grit.straganz@tugraz.at, Tel: +43-316-8738414 Fax: +43-316-8738434
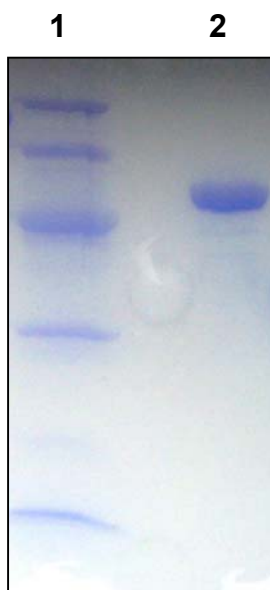
Figure S1. SDS-PAGE of purified ApoHctB. Lane 1: Low Molecular Weight (LMW) protein standard (GE Healthcare) (99 kDa, 66 kDa, 45 kDa, 30 kDa, 20.1 kDa, 14.4 kDa), lane 2: C-terminally tagged *Strep*-tag II HctB (56.8 kDa).
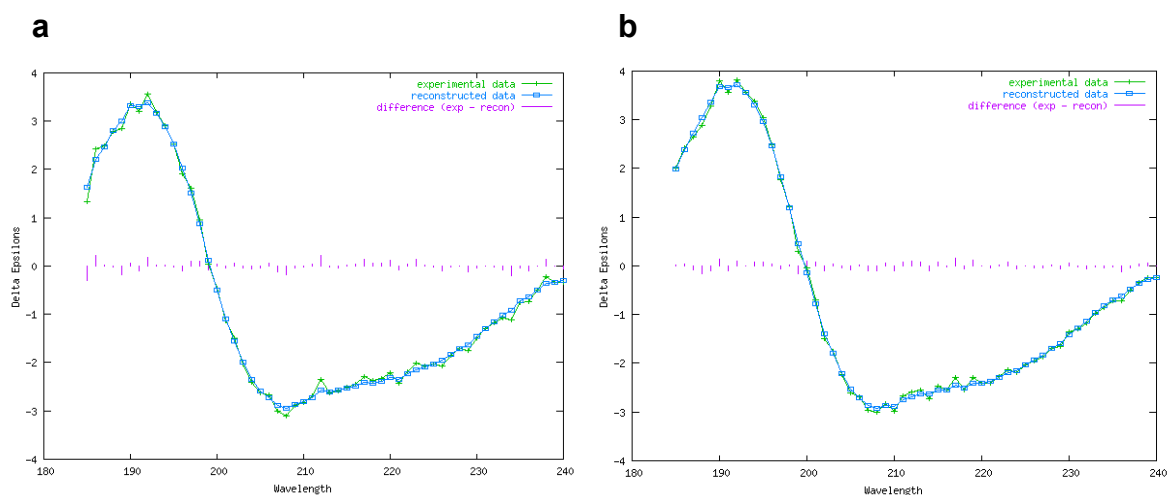
Figure S2. CD-Spectra of ApoHctB. The CD spectra of (a) C-terminally and (b) N-terminally *Strep*-tag II versions of HctB were recorded at ambient temperature. Spectra analysis using the DICHROWEB program indicated the same fraction of α-helices, β-sheets, turns, and unordered regions in both variants.

Figure S3. Representative fluorescence titration curve of ApoHctB and Cu(II). Data were collected at 25°C and pH 7.5 and a plot of relative fluorescence quenching at 340 nm gave a K_d of 1.4 ± 0.7.
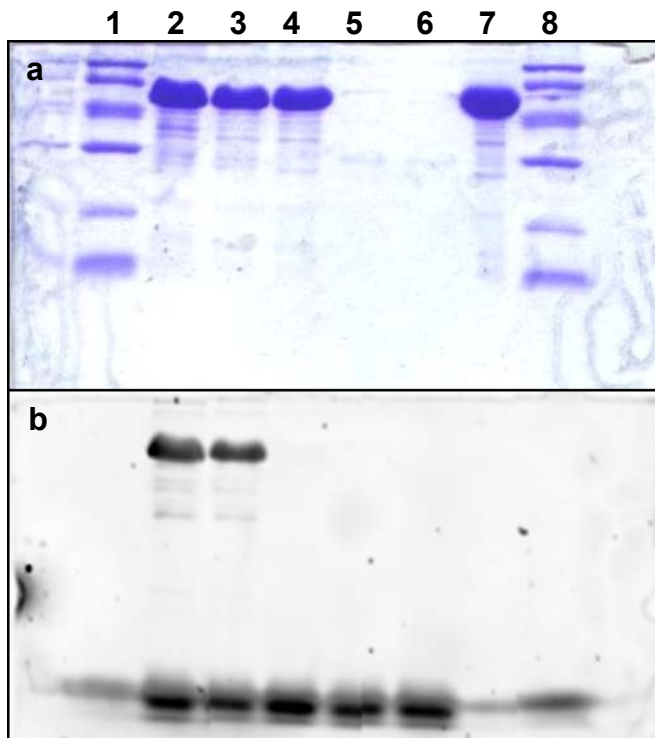
Figure S4. Bodipy-PPT transfer onto HctB by Sfp, SDS gel. Panel a shows the Coomassie blue stain of the proteins and panel b gives the signals of Bodipy that were detected at 555 nm. Lane 1 and 8: LMW protein standard (99 kDa, 66 kDa, 45 kDa, 30 kDa, 20.1 kDa, 14.4 kDa), Lane 2: HctB + Bodipy-CoA + prurified Sfp, Lane 3: HctB + Bodipy-CoA + Sfp c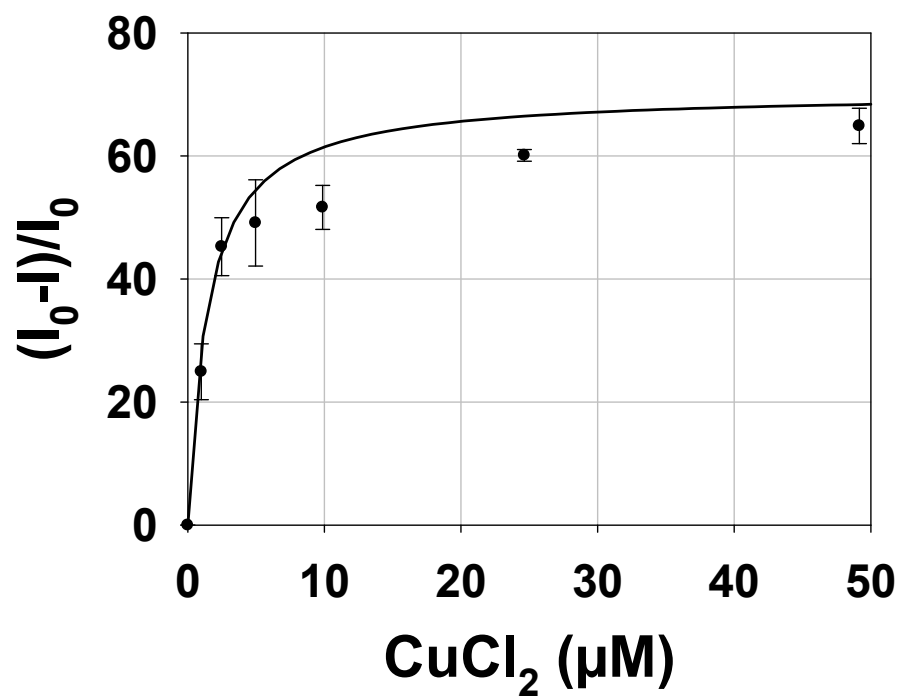rude cell extract, Lane 4: HctB + Bodipy-CoA, Lane 5: purified Sfp + Bodipy-CoA, Lane 6: Sfp crude cell extract + Bodipy-CoA, Lane 7: HctB only.
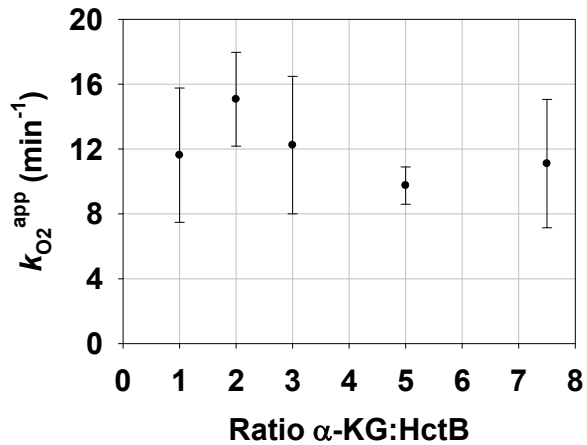
Figure S5. HoloHctB activity in air saturated Tris or HEPES buffers (20 mM, pH 7.5). Correlation of k_{O2}^{app} and α-KG concentration under NaCl saturating conditions (≥ 1 M) is shown.
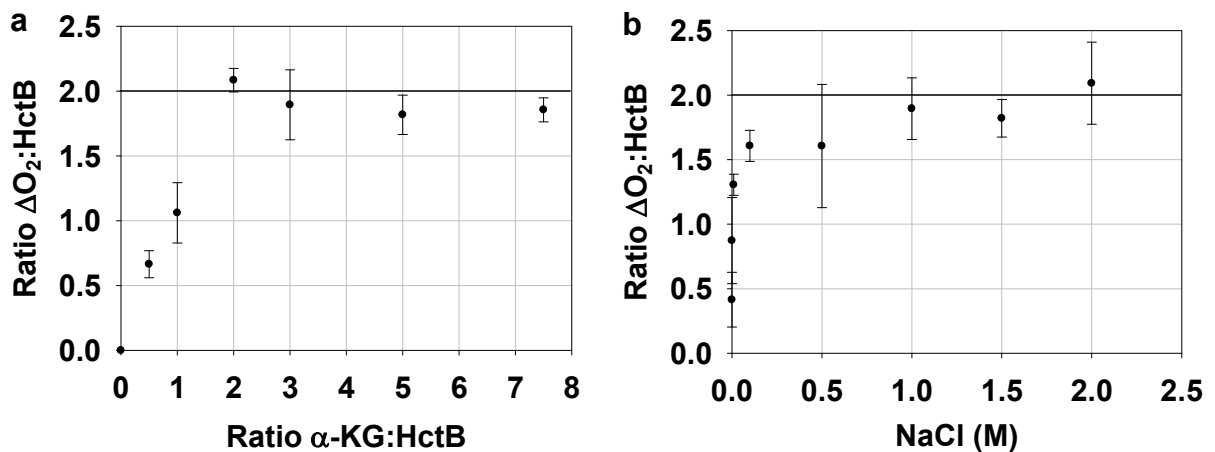
Figure S6. HoloHctB activity. Correlation of O_2 consumption stoichiometry and (a) α-KG concentration under NaCl saturating conditions (≥ 1 M) and (b) NaCl concentration under α-KG saturating conditions (300 µM) is shown.
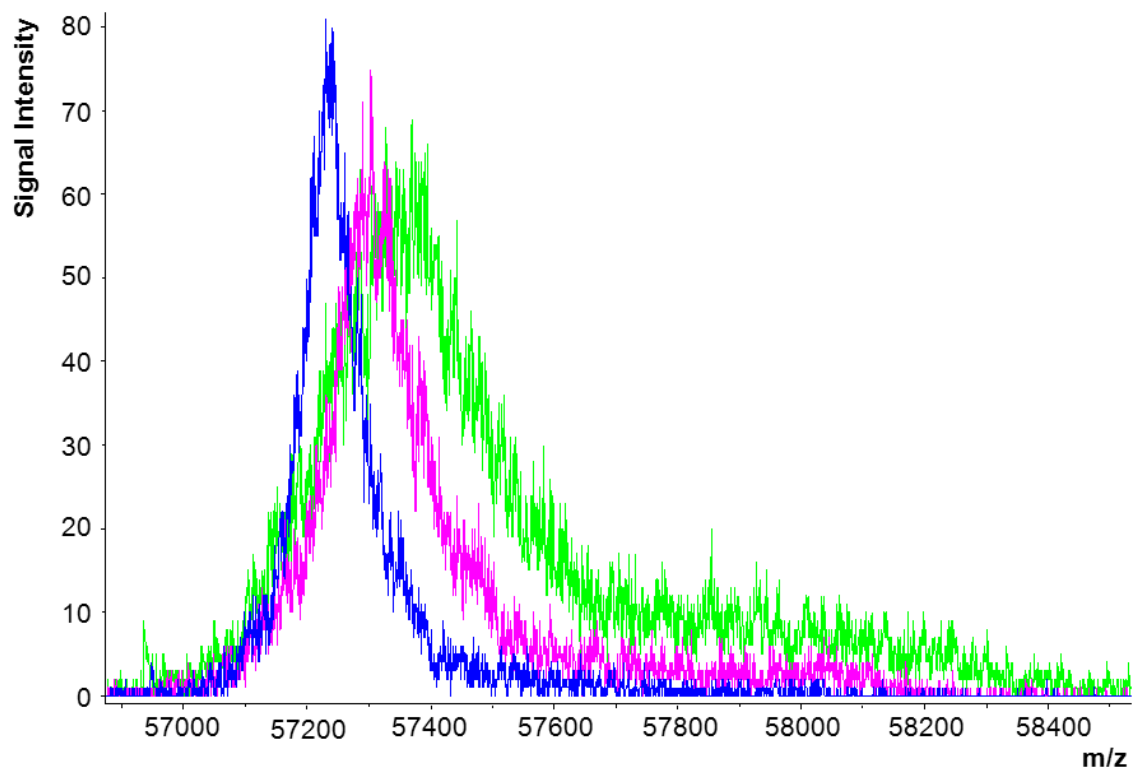
Figure S7. MALDI-TOF MS analysis of HoloHctB. Mass spectra before (blue) and after conversion with 1 M NaCl (pink) and 1 M KBr (green) are shown.
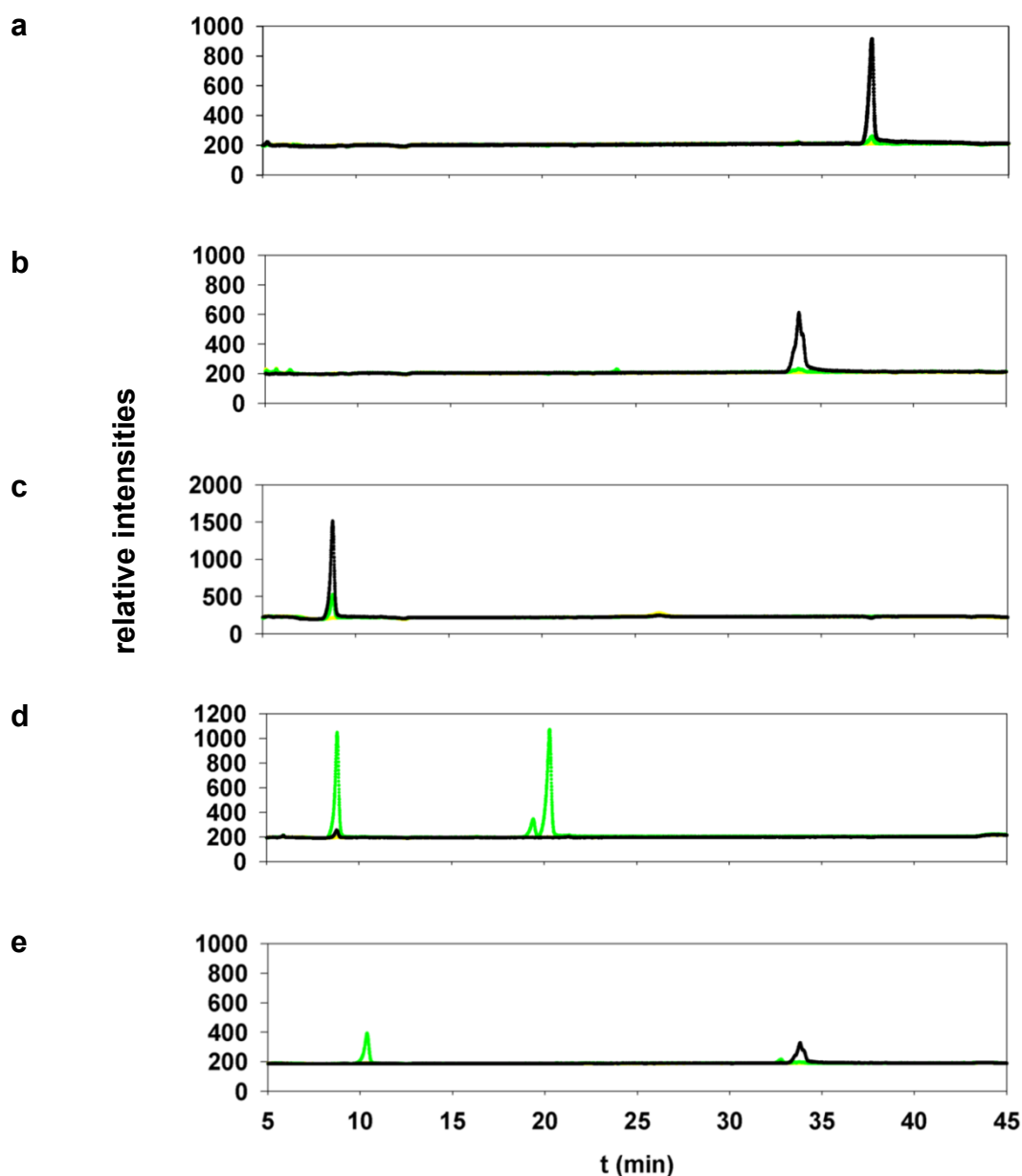
Figure S8. Qualitative determination of the HoloHctB reaction products via HPLC-MS analyses. Plots show scans for (a) di-^{35}Cl-hexanoate (m/z = 183), (b) ^{35}Cl-hexenoate or dihydroxy-hexanoate (m/z = 147), (c) oxo-hexanoate (m/z = 129), (d) hydroxy-hexanoate (m/z = 131), and (e) ^{35}Cl-hexanoate or ^{37}Cl-hexenoate (m/z = 149) in a HoloHctB control sample (yellow), a HoloHctB reaction excluding NaCl (green) and a HoloHctB reaction in the presence of 1 M NaCl (black). Note that HPLC MS is not a quantitative technique and no monohydroxy-species was found in NMR analysis, indicating that this is a minor reaction product.
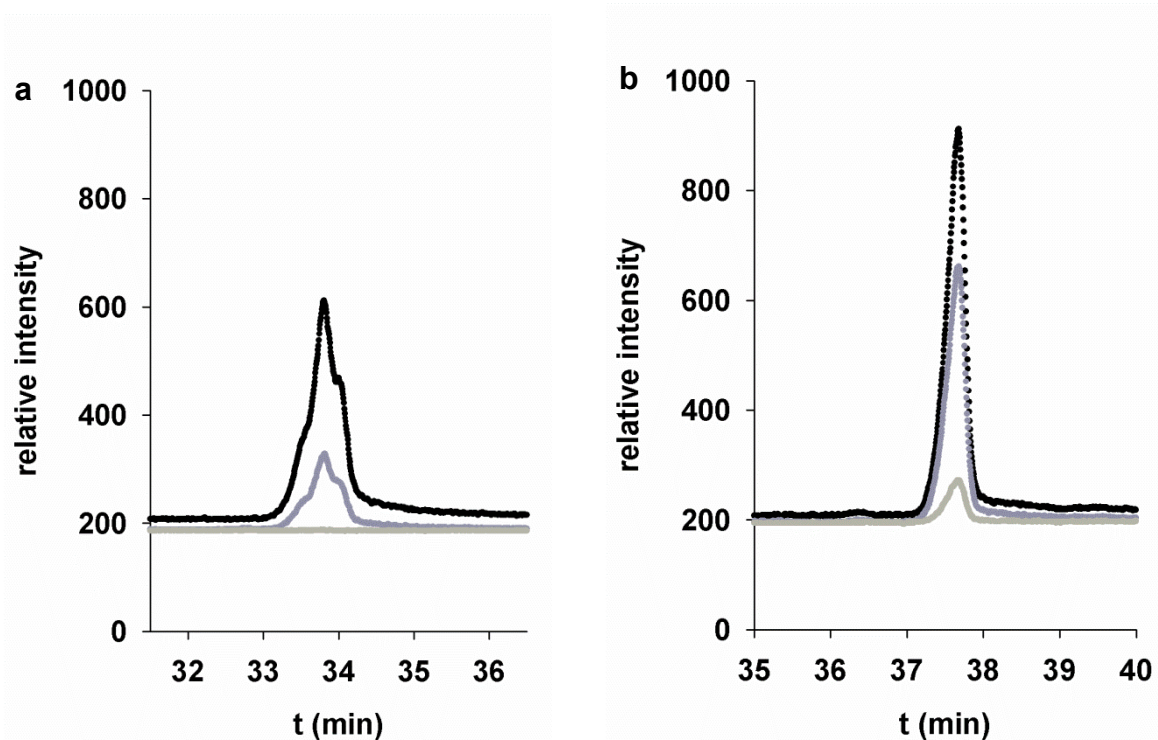
Figure S9. Verification of chloride incorporation by ^{35}Cl:^{37}Cl isotope dependent m/z ratio assessment. (a) Integrated signals of m/z = 147 (black) and m/z = 149 (dark gray) showed a ratio of 2.9:1. (b) Integrated signals for m/z = 183 (black), m/z = 185 (dark gray) and m/z = 187 (bright gray) show a ratio of 9.4:6.8:1. The spectra give evidence for (a) one and (b) two chloride atoms in the respective structures.
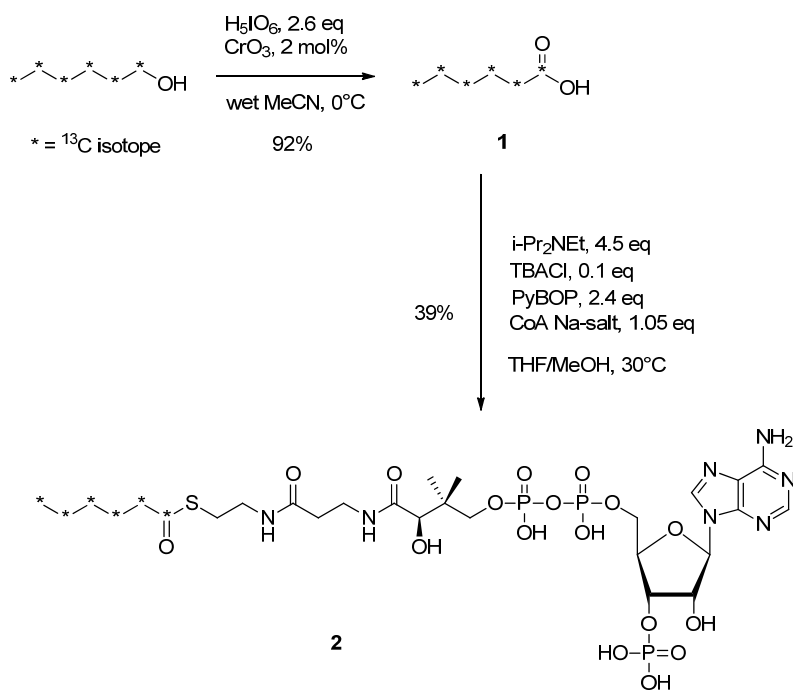
Scheme S1. Synthesis of $^{13}C_6$-hexanoyl-CoA.
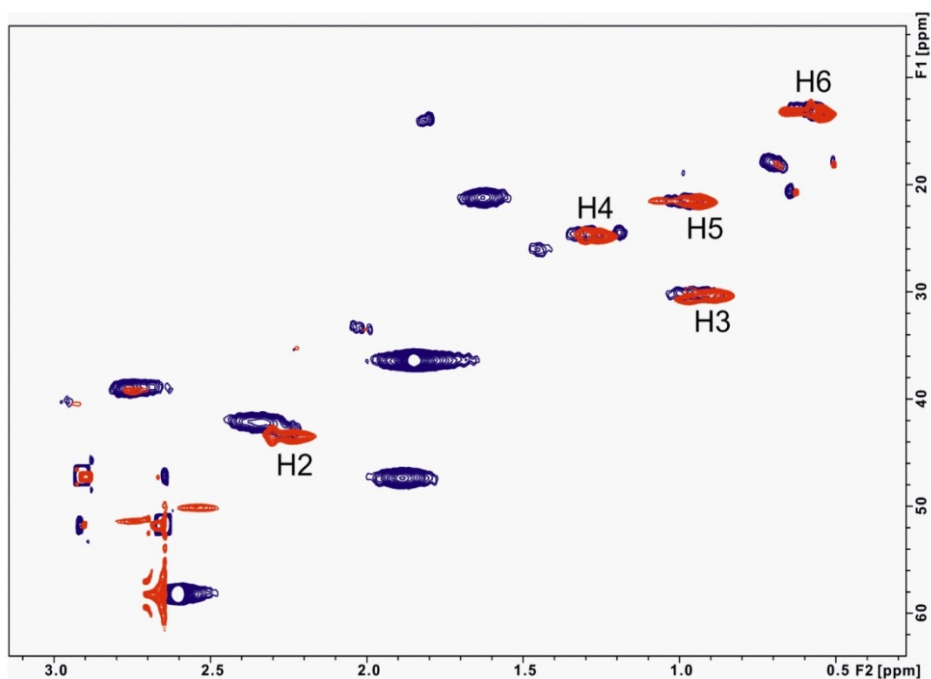
Figure S10. Overlay of ^{1}H-^{13}C HSQC spectra of the hexanoyl-substrate bound to HctB before (red) and after (blue) the enzymatic reaction. The assignment of the hexanoyl-group is indicated (H2-H6).
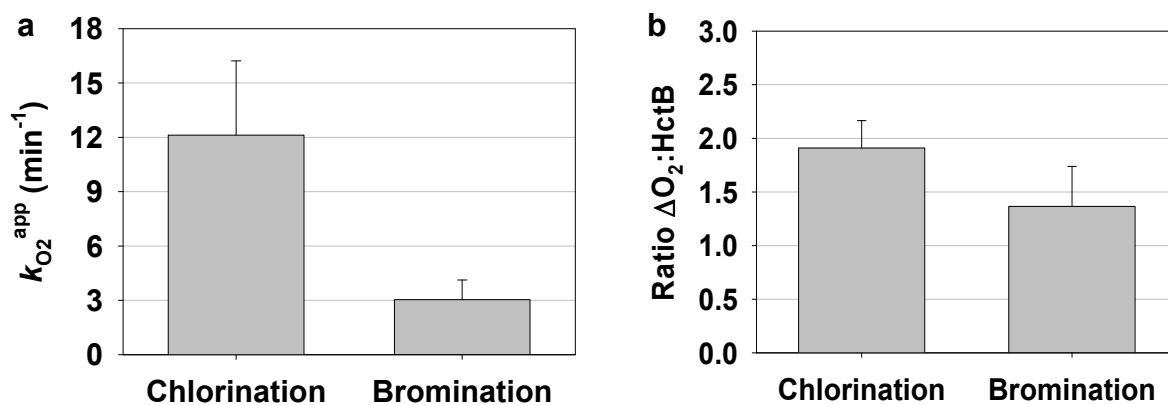
Figure S11. Substrate conversion in the presence of chloride vs. bromide. Plots show (a) the $k_{O2}{}^{app}$ and (b) the O_2 consumption stoichiometry of HoloHctB in the presence of 1 M NaCl and 1 M KBr, respectively.
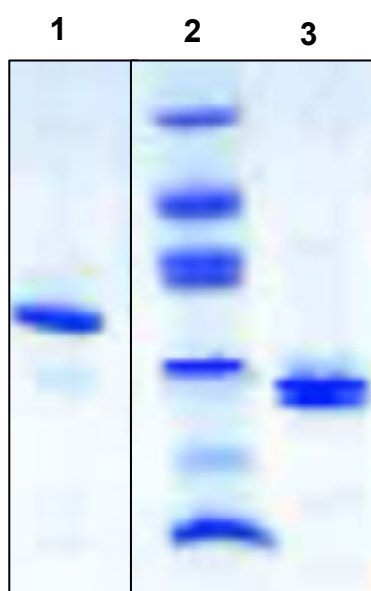
Figure S12. SDS-PAGE of purified accessory proteins. Lane 1: TycF (33.4 kDa); lane 2: LMW protein standard (99 kDa, 66 kDa, 45 kDa, 30 kDa, 20.1 kDa, 14.4 kDa); lane 3: Sfp (26.9 kDa).

HctB structural modeling and in silico docking.

The three conserved regions of HctB were identified via BLAST search.[S1] Based on these results, the 2-293 amino acid fragment (halogenase domain) and the 390-466 amino acid fragment (ACP domain) were modeled using the SWISS MODEL server[S2] and validated via Procheck.[S3] Global docking of the two domains was performed via the GraMM-X platform.[S4] The lowest energy structure was submitted to the Rosetta Docking server[S5] for local docking. In a complementary approach the C6-C14 acyl-PPT substrate moieties were *in silico* docked using the Autodock suite,[S6] which is integrated in the Yasara docking suite.[S7] The truncated ligand's terminal group was the phosphate moiety that is generally tethered to the ACP domain via covalent linkage to a conserved serine in the PPT-binding motif. To account for structural flexibility, five receptor structures were generated from the HctB domain model structure, each with alternative high-scoring side-chain rotamers. For *in silico* docking of the fully flexible ligands in a docking cell that surrounds the whole HctB domain at 298 K, the Amber03 force field[S8] was used. The resulting 120 models per ligand were subjected to cluster analysis. For generation of the halogenase and ACP-hexanoyl-PPT containing HctB model, the lowest energy hexanoyl-PPT conformation was superimposed with the previously *in silico* docked HctB-ACP complex. After covalent linkage of the phosphate and serine moieties and 'freezing' of the halogenase domain, the fatty acyl-moieties, the metal ion, and coordinating ligands, the system was energy minimized using a steepest descent minimization to remove conformational stress, followed by a simulated annealing minimization until convergence (< 0.05 kJ mol^{-1} per 200 steps). Integration time steps were set to 1.33 and 4 fs for intra- and intermolecular forces, respectively.

References

S1.   Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res. 25*, 3389-3402.

S2.   Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling, *Bioinformatics, 22*, 195-201.

S3.   Laskowski, R.A., Macarthur, M.W., Moss, D.S., and Thornton, J.M. (1993) PROCHECK: A program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283–291.

S4.   Tovchigrechko, A. and Vakser, I.A. (2006) GRAMM-X public web server for protein-protein docking, *Nucleic Acids Res. 34*, W310-W314.

S5.   Lyskov, S. and Gray, J.J. (2008) The RosettaDock server for local protein-protein docking*, Nucleic Acids Res. 36*, W233-W238.

S6.   Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., and Olson, A.J. (1998) Automated docking using a lamarckian genetic algorithm and and empirical binding free energy function, *J.Comput.Chem. 19*, 1639-1662.

S7.   YASARA (9.11.9 9.11.9) Vienna, Austria: YASARA Biosciences GmbH; 2012.

S8.   Duan, Y., Wu, C., Chowdhury, S., Lee, M.C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., and Lee, T. (2003) A point-charge force field for molecular mechanics simulations of proteins, *J. Comput. Chem. 24*, 1999-2012.

The role of chloride in the mechanism of O_2 activation at the mononuclear nonheme Fe(II) center of the halogenase HctB

Sarah M. Pratter[a], Kenneth M. Light[b], Edward I. Solomon[b] and Grit D. Straganz[a]*

[a] Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12, A-8010 Graz, Austria.

[b] Department of Chemistry, Stanford University, Stanford, California 94305, USA.

* *corresponding author*, e-mail: grit.straganz@tugraz.at, Tel: +43-316-8738414 Fax: +43-316-8738434

Abstract
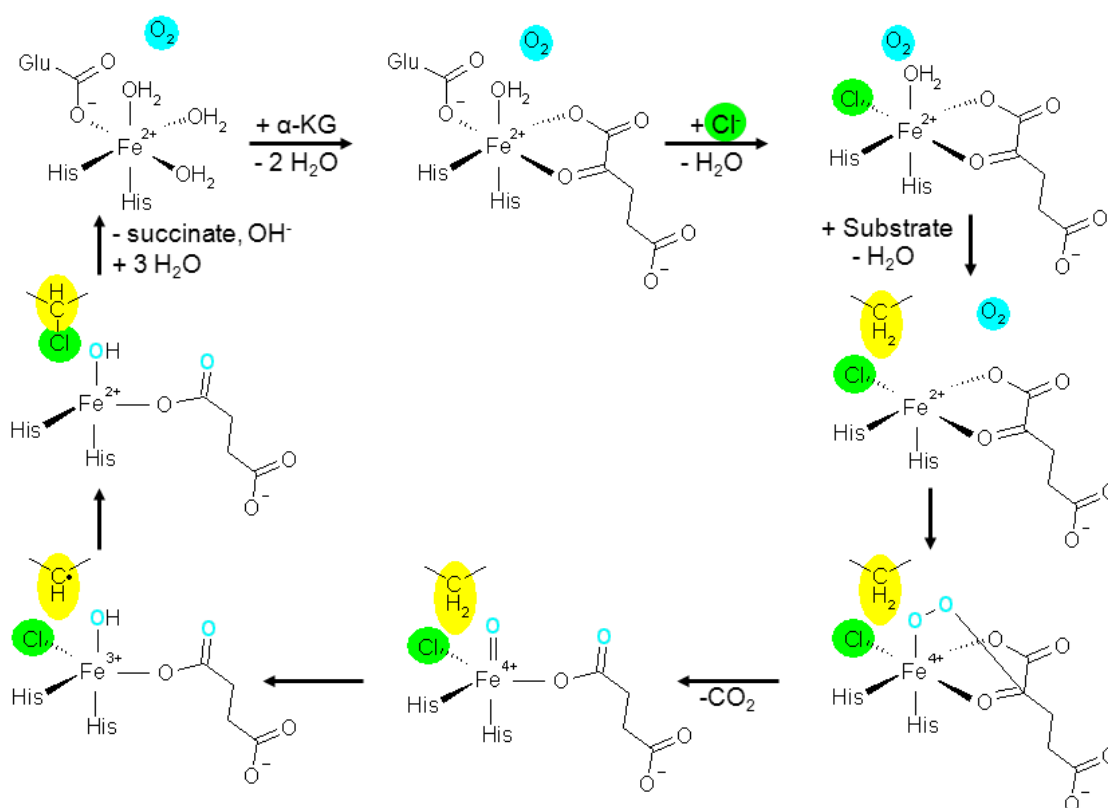
Mononuclear nonheme Fe(II) (MNH) and alpha-ketoglutarate (α-KG) dependent halogenases activate O_2 in order to perform oxidative halogenations of activated and non-activated carbon centers. While the mechanism of halide incorporation is well understood, the mechanism by which halogenases prevent oxidations in the absence of chloride is still obscure. Here, we characterize the impact of chloride on the metal center geometry and reactivity of the fatty acyl-halogenase HctB. Stopped flow kinetic studies show that the oxidative transformation of the Fe(II)-α-KG-enzyme complex is > 200-fold accelerated by saturating concentrations of chloride in both, the absence and presence of covalently bound substrate. By contrast, the presence of substrate, which generally brings about O_2 activation at enzymatic MNH centers, only has a ~ 10 fold effect in the absence of chloride. Circular dichroism (CD) and magnetic CD studies demonstrate that chloride binding triggers changes in the metal center geometry: it induces the partial formation of a species with square pyramidal 5-coordinate geometry, which is a prerequisite for O_2 activation at enzymatic MNH centers. Molecular dynamics simulations suggest that the binding of chloride to the metal center of HctB leads to a conformational change in the enzyme that makes the active site more easily accessible to substrate and thus facilitates the formation of the catalytically competent enzyme-substrate complex. Results are discussed in relation to other MNH dependent halogenases.

Introduction

In nature, the selective halogenation of non-activated carbon atoms is performed by mononuclear nonheme iron (MNH), O_2 and alpha-ketoglutaric acid (α-KG) dependent halogenases. The structural rational of this chemically challenging reaction has been elucidated rather recently[1-3] and to this date only a handful of MNH halogenases have been biochemically characterized.[4-11] Their molecular mechanism follows, in principle, that of the closely related α-KG dependent hydroxylases. There, the Fe(II) center of the resting enzyme is coordinated by a 2-His 1-carboxylate motif and three water ligands. Upon coordination of the α-KG cofactor, two water molecules are replaced whereby the metal ion's octahedral geometry is pertained. Only after substrate binding in the outer coordination sphere is the remaining water ligand displaced from the Fe(II) center and the resulting 5-coordinate species is prone to oxidation by O_2.[12] The α-keto cofactor is decarboxylated by the thus activated dioxygen species, which leads to the formation of a high valent Fe(IV)=O intermediate that abstracts a hydrogen atom from the substrate.[13] While in MNH dependent hydroxylases the reaction proceeds via transfer of the hydroxy group from the Fe(III)-OH intermediate onto the substrate radical, in halogenases an iron-coordinated chloride successfully competes with the hydroxy group and yields the halogenated product (Scheme 1). This is accomplished by an alteration of the halogenases' metal binding motif: There, a small, nonpolar amino acid replaces the acidic amino acid of the 2-His 1-carboxylate triad, and this allows a chloride atom to coordinate to the Fe(II) ion.

Several reasons have been invoked for the strong preference of halogenation over hydroxylation that has generally been observed in halogenases,[4-6,10] namely (i) the

48

lower redox potential of Cl• compared to OH•[14]; (ii) a possible bicarbonate formation between the metal bound hydroxy group and the α-KG derived CO_2 which prevents the transfer of the hydroxy group[15]; (iii) a possible protonation of the hydroxy group by a nearby Glu-Arg proton donor that results in the formation of water, which makes hydroxylation unfavorable[16]; (iv) the difference in binding strengths of the chloride and hydroxy group to the metal ion and the resulting energetic barrier for hydroxylation[17]; or (v) the positioning of the substrate radical relative to the Cl-Fe(III)-OH center.[18,19]
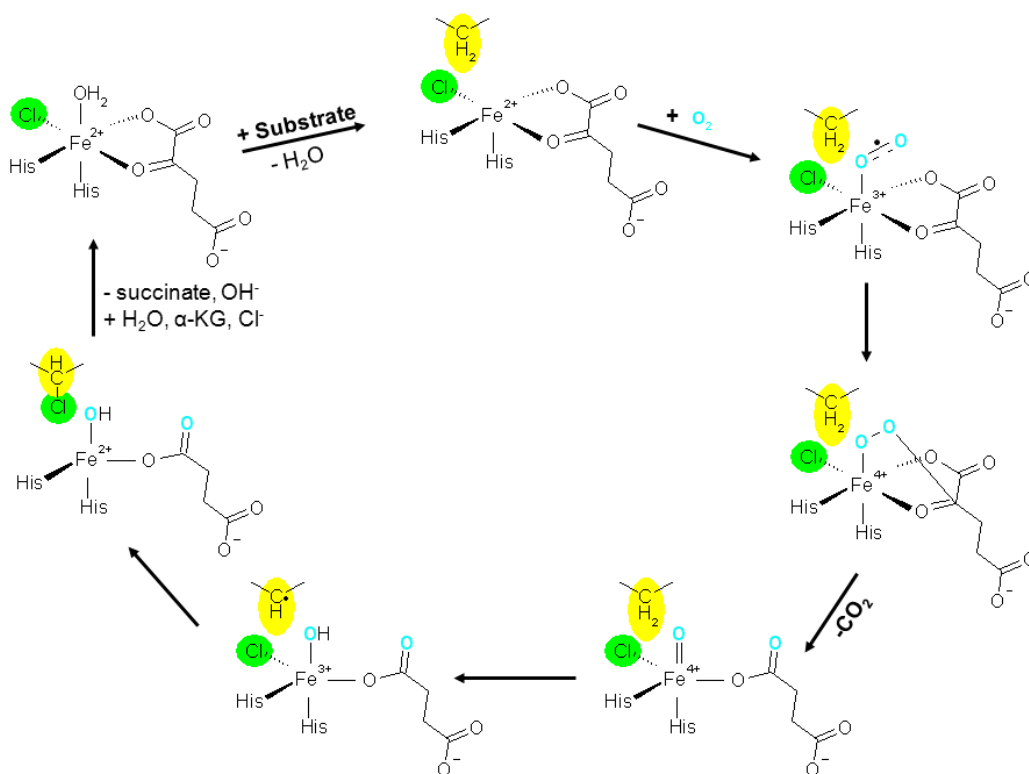
Scheme 1. Proposed molecular mechanism of Fe(II)- and α-KG dependent halogenation.[13] In the presence of α-KG and substrate, O_2 binds to the 5-coordinate Fe(II) center and decarboxylates the α-KG cofactor. This yields a highly reactive Fe(IV)=O intermediate that abstracts a proton from the substrate. Chloride rather than the hydroxy moiety is rebound by the substrate radical, resulting in a chlorinated reaction product.

While substantial efforts have been made to investigate the second half of the halogenases' reaction in order to rationalize the preference of substrate chlorination over hydroxylation after proton abstraction, not much attention has been paid to the first half of the catalytic cycle. Chloride bound crystal structures of the substrate free MNH halogenases SyrB2,[1] and CurA-Hal[3] show that the halide can coordinate to the iron even before substrate binds. On the other hand, in CytC3 no halide is present in the obtained crystal structures that contain either Fe(II) only or Fe(II)-α-KG, despite rather high chloride concentrations in the mother liquor (80 mM).[2] This suggests that halides are not necessarily bound a priori. To date no study exists that addresses the question how halogenases prevent putative alternative oxidation reactions in the absence of halide.

Recently, a fatty acyl-halogenase, HctB from *L. majuscula,* has been characterized in our lab.[11] The enzyme that modifies middle-chain fatty acyl-moieties displays an unprecedented three domain organization that may be prototypical for fatty acyl-halogenases and sets it apart from the mono-domain amino acyl-halogenases and the multi-domain ketide halogenases: An acyl-Coenzyme A binding protein and an acyl-carrier protein with an inherent thiolation site, where the substrate covalently binds via a phosphopantetheinyl bridge, are fused to the halogenase domain. This composition makes HctB self-sufficient in regard to a substrate binding entity. The trifunctional enzyme introduces 5-oxo-, 5,5-dichloro- and 5-chloro-4-vinyl moieties into the hexanoyl-substrate under chloride saturating conditions. In the course of the enzyme's biochemical characerization we observed that O_2 reduction was not only triggered by the presence of substrate but also by chloride under single turnover conditions. [11]

Based on these observations, a suspected role of chloride in primary O_2 activation in HctB is investigated in this study. A combination of circular dichroism (CD) and magnetic CD (MCD) spectroscopies, that directly observe the geometric and electronic structure of the enzymatic Fe(II) ligand field and of stopped flow kinetics reveals that the metal center is constituted in the absence of a halide ion but remains inert towards O_2 and that the presence of chloride is essential for triggering O_2 reduction at the metal center. Molecular dynamics (MD) simulations of the HctB halogenase domain are employed to gain insights into the role of the protein structure in chloride dependent O_2 activation.

Results

Stopped flow analysis.

In a previous study single turnover O_2 consumption rates of HctB had been found to be 5- and 11-fold elevated in the presence of substrate and chloride, respectively.[11] As the observed O_2 depletion rate may underestimate the actual rate of O_2 reduction during catalysis, the oxidation of the Fe(II)-α-KG complex was monitored spectrophotometrically via single turnover stopped flow measurements by recording the decay of its prototypical metal to ligand charge transfer (MLCT) transition (ε_{500nm} = 150 M^{-1})[11]. Therefore, anaerobic enzyme preparations of substrate-bound HctB (HoloHctB), preloaded with ~ 0.95 molar equivalents Fe(II) and α-KG, were mixed with O_2-enriched (650-1400 µM) and, optionally, 1M NaCl containing buffer. In the absence of NaCl the *α-KG-Fe(II)-HoloHctB* complex decayed slowly with a specific rate of $1.1*10^{-3}$ s^{-1} (c_{O2} = 700 µM). By contrast, under analogous conditions but in the

presence of chloride, the absorbance band disappeared within ~ 20 seconds. The trace displayed three distinct phases: A lag phase of ~ 60 ms (k_{SC1} = 35.1 s^{-1}) was followed by a phase of fast signal decay, which gave an apparent first order rate constant k_{SC2} of 6.74 s^{-1} and accounted for approximately one third of the total amplitude. The third phase of the reaction, a slower absorbance decrease (k_{SC3} = 0.22 s^{-1}), equaled the O_2 depletion rate (~ 0.20 s^{-1}) that had previously been determined using the O_2 sensor under similar conditions (Figure 1). Rates were not significantly dependent on the O_2 concentration (Table S1, Figure S1). The total amplitude of signal decay corresponded to 97 % of the applied α-KG-Fe(II) concentration, confirming a quantitative conversion of the complex. It is noteworthy that in previous studies it was observed that the Fe(IV)=O intermediate had a similar extinction coefficient at 520 nm as the Fe(II)-α-KG complex but additionally showed a significant absorbance increase at 318 nm (ε ~ 1500 M^{-1} cm^{-1}).[13] An analogous signal increase was not detectable during substrate conversion by HoloHctB. Instead, a slow signal increase at 320 nm in the absence (Figure S2 b) and presence (Figure S2 d) of chloride was obtained, indicating some Fe(II) oxidation as a side reaction.

When substrate-free HctB (ApoHctB) was subjected to stopped flow analysis, the chloride-free complex displayed an almost stable signal at 500 nm over the measured time range with a specific rate of 0.09*10^{-3} s^{-1} (Figure S2 a). By contrast, in the presence of chloride *α-KG-Fe(II)-ApoHctB* gave an absorbance increase over the whole recorded wavelength range. The fastest increase was observed at 320 nm. The absorbance trace at 500 nm had a lag phase, which indicated the formation of a secondary species. When traces were fit via formula **(1)** (see Material and Methods section) apparent first order rate constants of 0.15 s^{-1} (320 nm) and 0.03 s^{-1} (500 nm) together with estimated extinction coefficients k_c and k_c' of 4.3 mM^{-1} cm^{-1} (320 nm)

and 7.3 mM^{-1} cm^{-1} (500 nm) were obtained (Figure S2 c). Consequently, a putative

formation of the 320 nm species was assessed to account for < 3% of the total iron in

the samples of HoloHctB and choride free ApoHctB. Regarding the intensely

absorbing band at 500 nm, the stoichiometric decrease of the *α-KG-Fe(II)-HoloHctB*

complex according to ΔA_{500} and the relative stability of the signals in the absence of

NaCl implied that in the respective samples no significant buildup of this species

occurred (Figures S2 a, b, d).
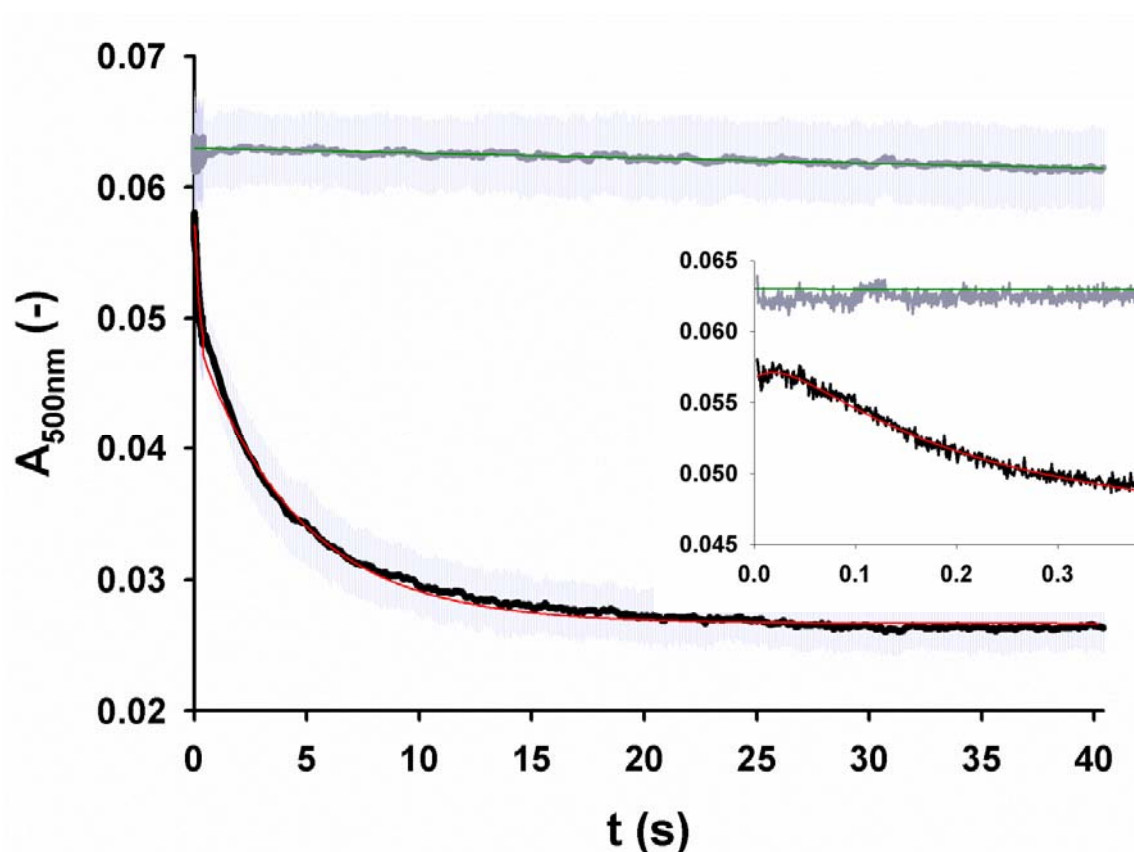
Figure 1. Stopped flow absorption kinetic traces of the α-KG-Fe(II)-HoloHctB complex decay. The conversion of the chromophoric α-KG-Fe(II) pair was monitored in the absence (gray) and presence (black) of 0.5 M NaCl at 500 nm and average traces were fit with a linear regression (green) and via a three-phasic model using the Pro-Kineticist software (Applied Photophysics, see Materials and Methods section for details), respectively. The insert shows the initial 400 ms phase of the reaction.

Summarizing, according to stopped flow measurements the decay of the *α-KG-Fe(II)-HoloHctB* complex is accelerated ≥ 200-fold by chloride, while the presence of covalently bound substrate in the absence of chloride enhanced the rate only by an order of magnitude.

CD and MCD spectroscopy.

In order to gain insights into the metal center geometries, combined CD/MCD spectroscopic measurements of the enzymatic Fe(II) centers of Apo- and HoloHctB in the presence and absence of saturating NaCl concentrations were applied. The methodology allows the determination of an Fe(II) center geometry, based on the energies of the respective d-d transitions: Free high spin Fe(II) exhibits five energetically degenerate d-orbitals. Upon the influence of an octahedral ligand field the d-orbitals split into a triply degenerate $^5T_{2g}$ ground state and a doubly degenerate ^5E_g excited state that are each further split in the distorted octahedral geometry of a protein environment: The ^5E_g state, which is located at ~ 10000 cm^{-1}, will split by ~ 2000 cm^{-1}. Removal of an axial ligand and concomitant formation of a square pyramidal 5-coordinated species results in the splitting of the ^5E_g state by ~ 5000 cm^{-1}, yielding transitions at ~ 10000 cm^{-1} and ~ 5000 cm^{-1}.[20]

<u>CD spectroscopy:</u> Anaerobic titrations of Apo- (≥ 3.5 mM) and HoloHctB (> 1.4 mM) with Fe(II) and α-KG were monitored via CD spectroscopy to investigate the conditions required for cofactor saturation of the enzyme. Upon addition of 0.9 molar equivalents of Fe(II), transitions in the region of 7000-12000 cm^{-1} were observed. Addition of another 0.9 equivalents Fe(II) did not significantly increase the signals (data for HoloHctB in Figure 2 a), for which an extinction coefficient Δε of ~ 0.03 cm^{-1} mM^{-1} was deduced. These data also implied, that the addition of 0.9 equivalents

Fe(II) was sufficient to load Apo- and Holo-enzyme quantitatively in the absence of α-KG and NaCl. Addition of a > 3-fold molar excess of α-KG, which ensure saturating conditions,[11] to Fe(II)-loaded enzymes led to a shift towards higher energy, indicating α-KG coordination to the metal (data for HoloHctB in Figure 2 b). Consequently 0.9 equivalents of Fe(II) and 3.4 equivalents of α-KG were chosen for subsequent experiments.
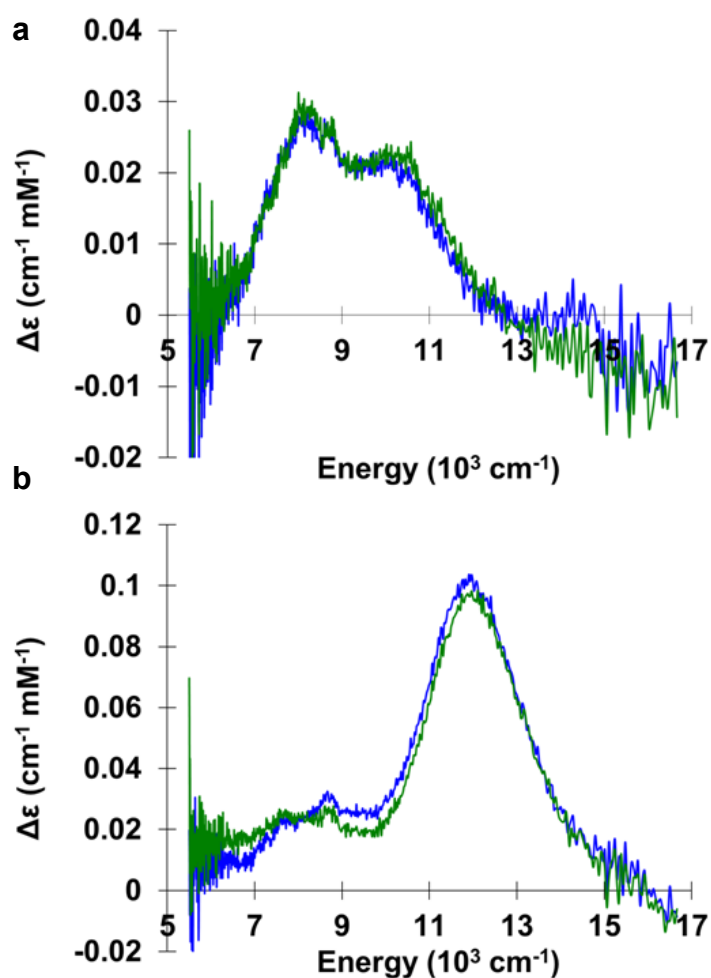

Figure 2. Fe(II)- and α-KG titration of HoloHctB (CD). A chloride-free enzyme solution (1.45 mM) was loaded with (a) 0.9 (blue) and 1.8 (green) molar equivalents of Fe(II) and additionally, with (b) 3.4 (blue) and 6.8 (green) molar equivalents of α-KG. CD spectra demonstrated saturation at 0.9 (Fe(II)) and 3.4 (α-KG) equivalents of the cofactors.

<u>MCD spectroscopy:</u> In the absence of NaCl, the MCD spectra of *Fe(II)-Apo- and Fe(II)-HoloHctB* showed two overlapping features with maxima at ~ 9200 and ~ 11200 cm^{-1} (Figure 3 a and c). The energy region of the transitions together with their splitting by ~ 2000 cm^{-1} indicated a 6-coordinate octahedral metal ligand field. In both, the *α-KG-Fe(II)-ApoHctB* and the *α-KG-Fe(II)-HoloHctB* complex bands shifted to ~ 8700 cm^{-1} and ~ 11700 cm^{-1} (Figure 3 b and d) and a shoulder in the high energy region, which started at ~ 15000 cm^{-1}, was observed. The latter corresponds to the MLCT band that appears upon α-KG coordination to the metal. The more pronounced splitting by ~ 3000 cm^{-1} is in line with a weakened axial ligand due to the binding of the strong electron-donor α-KG, which was also observed for the iron center of CytC3. [21]

When NaCl saturating conditions were applied, MCD features of both the *Fe(II)-ApoHctB* as well as the *Fe(II)-HoloHctB* samples, appeared at ~ 9200 and ~ 11200 cm^{-1} (Figure 4 a and c), resembling the respective chloride-free complexes. By contrast, the addition of chloride to the *α-KG-Fe(II)-ApoHctB* sample resulted in a dramatic change of spectral features compared to the respective NaCl-free sample. Two transitions at ~ 7700 cm^{-1} and ~ 12200 cm^{-1} were observed (splitting ~ 4500 cm^{-1}). Fitting of the 5 K spectrum furthermore suggested that an additional peak in the region of 9000 cm^{-1} was present and a further small absorbance peak in the ~ 13000 cm^{-1} range, which would indicate a second octahedral species, could not be ruled out. Given that no low energy transition at ~ 5000 cm^{-1} is present (Figure 4 b), the spectra clearly do not mirror a square pyramidal, 5-coordinate Fe(II) species, which is the generally accepted geometry that accomplishes O_2 activation. However, the partial formation of a 5-coordinate, trigonal-bipyramidal species, which would be in line with the feature at ~ 9000 cm^{-1}, cannot be excluded. The four features at ≤ 5000

cm^{-1}, ~ 7500 cm^{-1}, ~ 9800 cm^{-1}, and in the region of 12000 cm^{-1} in the spectra of *α-KG-Fe(II)-HoloHctB* (Figure 4 d) are in line with the presence of two different Fe(II) ligand geometries. In analogy to the *α-KG-Fe(II)-ApoHctB* center, features at ~ 7500 cm^{-1} and ~ 12000 cm^{-1} may reflect a residual 6-coordinate Fe(II) species. The transition at ≤ 5000 cm^{-1} together with the peak at ~ 9800 cm^{-1} (splitting > 4800 cm^{-1}) are prototypical for a square pyramidal, 5-coordinate Fe(II) geometry. In summary this implies that the presence of all cofactors, including α-KG, substrate and NaCl are stringent to turn the system 5-coordinate and thus prime the metal center for oxygen binding and catalytic activity.
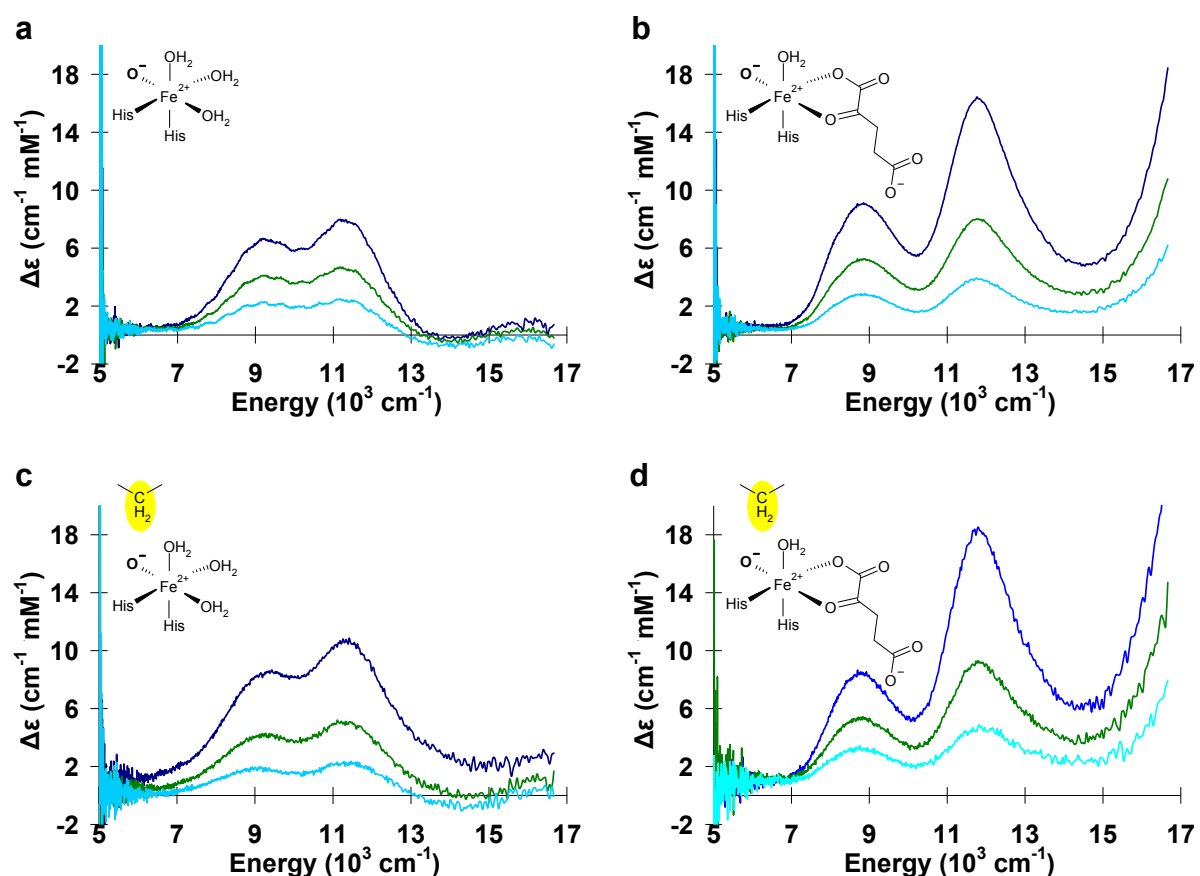
Figure 3. Low-temperature, 7 T MCD spectra of Fe(II)-HctB in the absence of NaCl. 5 K (dark blue), 20 K (green) and 40 K (light blue) spectra of *Fe(II)-ApoHctB* in the (a) absence and (b) presence of α-KG and spectra of *Fe(II)-HoloHctB* in the (c) absence and (d) presence of α-KG are shown.
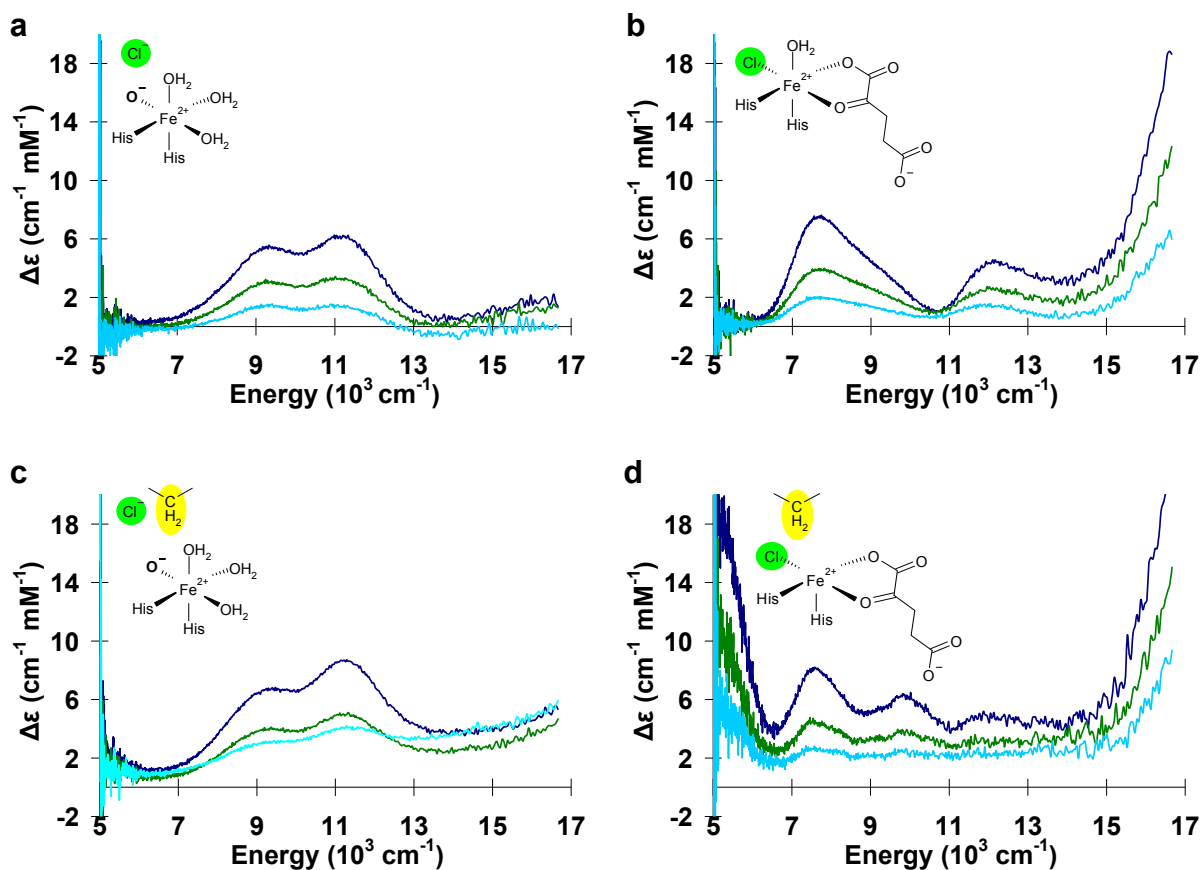
Figure 4. Low-temperature, 7 T MCD spectra of Fe(II)-HctB in the presence of NaCl. 5 K (dark blue), 20 K (green) and 40 K (light blue) spectra of *Fe(II)-ApoHctB* in the (a) absence and (b) presence of α-KG and spectra of *Fe(II)-HoloHctB* in the (c) absence and (d) presence of α-KG are shown.

To find out whether the maintenance of the 6-coordinate Fe(II) site in the chloride-free *α-KG-Fe(II)-HoloHctB* complex was due to a missing negative charge at the halide binding position, we recorded a spectrum of the respective sample at a pD of 9.1. However, the pD shift by 1.6 units, which previously brought about water deprotonation in a non-carboxylate coordinated metal center,[22] did not significantly alter the MCD spectrum (Figure S3).

MD simulations.

To elucidate the structural basis of chloride mediated O_2 activation, a model of the HctB halogenase domain was constructed based on the crystal structure of SyrB2 as described previously[11] and subjected to MD simulations in the presence and absence of 1M NaCl in the simulation cell. In order to ensure geometrically appropriate metal center geometry, an octahedral iron force field was defined and applied to the ligands, namely His112, His228, α-KG, and - where appropriate - chloride, as outlined in the Supplementary Information. 100 ns MD simulations of the chloride containing 2-His 1-chloro α-KG ligated HctB model and its chloride-free analogue showed that the over-all structures were stable over the simulated time. 90 ns snapshots of the simulations are shown in Figure 5. Notably, in the halide free simulation a major rearrangement at the metal center was observed: Residue Glu224, which pointed away from the metal center in the starting structures, reoriented and coordinated to the Fe(II) ion. The carboxylate moiety occupied the halide's position within 10 picoseconds and remained there (Figure 6 a). When the simulation of the chloride free HctB structure was repeated with a water molecule kept strongly bound to the metal ion at the position of the chloride ion, Glu224 adopted a similar orientation, whereby the residue H-bonded to the metal bound water molecule (data not shown). By contrast, in the chloride containing structure the carboxylate moiety of Glu224 remained pointing away from the HctB active site (Figure 6 a), presumably due to electrostatic interactions. The reorientation of Glu224 in the chloride free HctB structures led to major structural changes: The residue's flipping resulted in a H-bonding interaction between its backbone oxygen and the backbone nitrogen of Val114 (Figure 6 c) and drew the respective strands closer together (Figure 6 b). Furthermore, the side chains of Tyr115 and Arg225 reduced their distance from an average of 10 Å to 5-6 Å (Figure 6 c) regarding the O-7 of Tyr

and the N-ω of Arg225. Also, the backbone nitrogen of Glu224 formed a H-bond with the oxo group of the Asn140 side chain in the halide free structure (Figure 6 c).
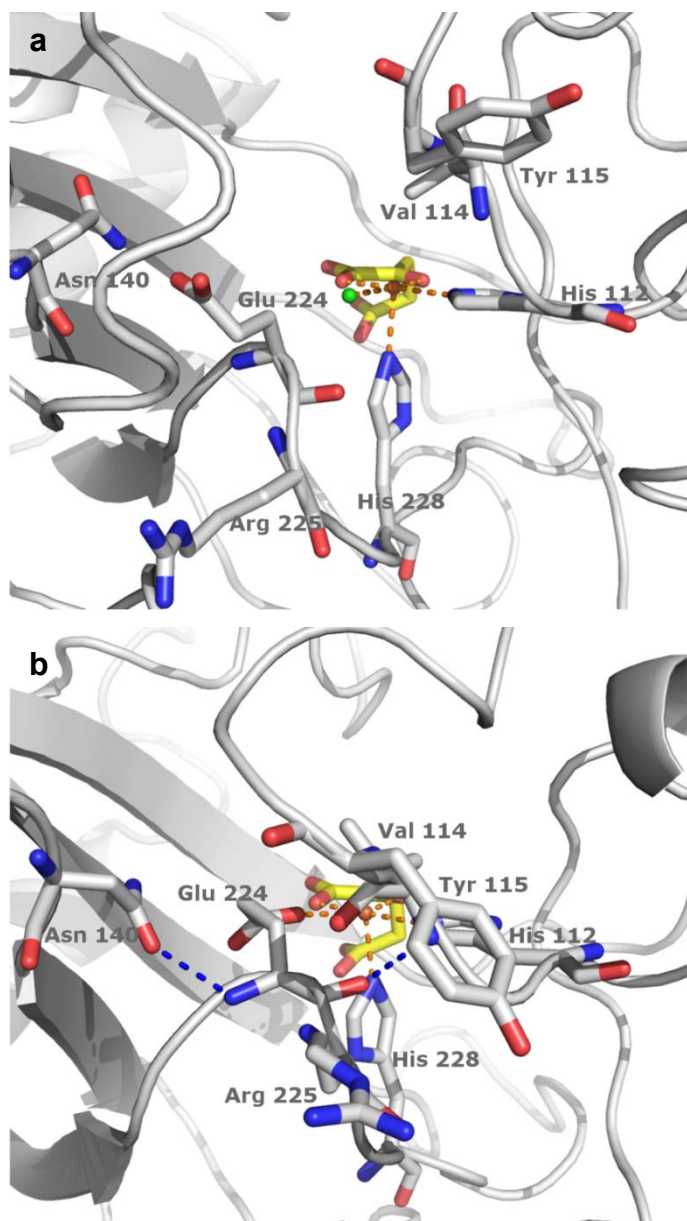
Figure 5. 90 ns snapshots of the HctB halogenase domains derived from MD simulations of the modeled structures. While in the chloride containing structure (a) the halide (green ball) coordinates to the octahedrally coordinated Fe(II) (orange ball), in the chloride free structure (b) Glu224 takes the chloride's position and this results in a rearrangement of the neighboring residues Val114, Tyr115, Asn140 and Arg225. The axially bound water molecules at the metal center are not displayed.
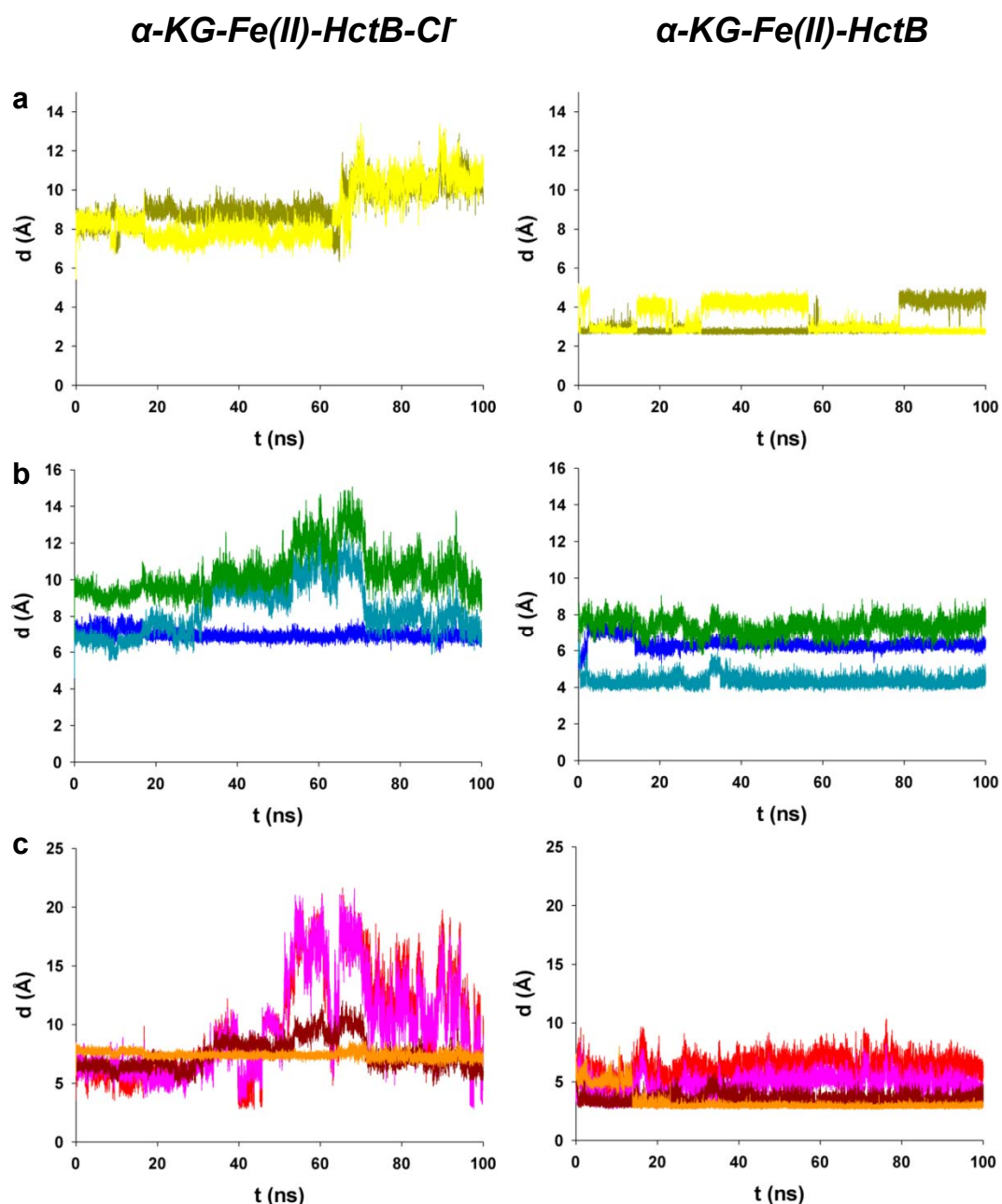
Figure 6. 100 ns MD simulations of residues involved in the chloride-triggered structural rearrangement in HctB in the presence (left panels) and absence (right panels) of 1 M chloride. Panel a shows distances of O-5 and O-5' of Glu224 to Fe(II). Panel b displays backbone distances from C-2 of Asn140 to N-2 of Glu224 (blue), N-2 of Val114 to C-2 of Glu224 (teal) and C-2 of Tyr115 to C-2 of Arg225 (green). Panel c gives the distances of the putative H-bonding atoms O-7 of Tyr115 and N-ω of Arg225 (red), O-7 of Tyr115 and N- ω' of Arg225 (pink), N-2 of 114 and O-1 of Glu224 (dark red), and O-4 of Asn140 and N-2 of Glu224 to (orange).

61

Discussion

The impact of chloride binding on the metal center geometry and the O_2 reduction activity of HctB.

The presented study demonstrates that although an enzymatic α-KG bound Fe(II) center is constituted in the absence of a halide, the latter impacts the geometric structure of the metal center and thereby triggers O_2 activation in HctB. CD/MCD spectroscopic studies of the metal centers in *Fe(II)-ApoHctB* and *Fe(II)-HoloHctB* show stable 6-coordinate geometry in the absence of any cofactor. This sets HctB apart from the related α-KG dependent halogenase CytC3, where α-KG was a prerequisite for Fe(II) binding to the enzyme.[21] HctB samples with and without chloride show similar spectral features. However, upon α-KG coordination the presence of the chloride ion has a major impact on the metal center geometry and reactivity. This suggests that HctB has a functional, self-sufficient metal binding site, but α-KG binding is required in order to allow chloride to coordinate to the metal ion.

In the *α-KG-Fe(II)-HoloHctB* complex, which should in principle contain all components that are necessary for O_2 activation, the presence of chloride is required to induce the loss of a ligand from the octahedrally coordinated iron center and this results in a square pyramidal metal species (Figure S5). The vacated coordination site allows binding of O_2 and results in an active site that is competent for catalysis. This is supported by stopped flow analyses, which demonstrate a chloride induced ≥ 200-fold increase in the rate of *α-KG-Fe(II)-HoloHctB* complex decay.

Also, in *α-KG-Fe(II)-ApoHctB* the octahedral geometry is disturbed in the presence of chloride: Features at ~ 7700 cm^{-1}, ~12200 cm^{-1} and ~ 9000 cm^{-1} indicate the presence of two species, which do, however, not display square pyramidal geometry. Yet, the iron-complex is clearly prone to oxidation by O_2: Appearance of an intensely absorbing species at 320 nm in the presence of O_2 is observed, which is not significantly formed in the other studied samples. Whether this is due to cofactor oxidation to Fe(III) or, alternatively, an Fe(IV)=O intermediate is not clear. Notably, Matthews *et al.* showed, that a SyrB2 reaction with amino acyl-free PPant-SyrB1 cofactor results in a remarkable stability of the Fe(IV)=O intermediate of at least 1000 s.[24] By analogy, in *α-KG-Fe(II)-ApoHctB* the Fe(IV)=O degradation may also be significantly inhibited by the lack of the proton-donating substrate.

A putative role of Glu224 in O_2 activation of HctB.

MD simulations of the HctB structures suggest major structural differences between the chloride bound and chloride free halogenase domains. In the absence of chloride, the carboxylate moiety of Glu224 is turned towards the iron center. This results in H-bonding interactions between the backbones of Glu224 and Val114 as well as the backbone of Glu224 and the side chain of Asn140 and reduces the distances between the Tyr115 and Arg225 sidechains. Notably, previous *in silico* docking studies have suggested that the fatty acyl-moiety accesses the active site between the strands that bear residues 114-115 and 224-225 (Figure S4).[11] As the respective residues converge in the chloride free model, an analogous active site access by the substrate's acyl-moiety could be severely hampered. Such a scenario can rationalize the substrate's surprising inability to affect MCD spectra and to markedly increase O_2 consumption rates in the absence of chloride. The apparent transient stability of the *α-KG-Fe(II)-HoloHctB* complex during single turnover kinetics, which is mirrored by

the ~ 60 ms lag phase, may thus reflect the reorganization of the enzyme upon chloride binding, involving the opening of the substrate channel, the substrate entrance to the active site and the formation of a 5-coordinate iron center before O_2 dependent catalysis can take place. Although it cannot be ruled out that in the absence of chloride a strongly bound water takes the halide's position adjacent to the metal center, direct coordination of Glu224 could help rationalize several observations in this study, such as the surprising stability of the iron cofactor in the absence of α-KG and the insensitivity of MCD spectra regarding a shift of the pD from 7.5 to 9.1. Furthermore, the apparent requirement of α-KG for the coordination of chloride might be explained by a weakening of the Glu224-Fe(II) bond upon α-KG coordination, which allows the replacement of Glu224 by the halide. Lastly, the presence of a carboxylate ligand in the absence of chloride may stabilize a sixths water ligand in analogy to other facial triad enzymes.[25]

A corresponding acidic amino acid residue is found in the other so far annotated fatty acyl-halogenase Psm3L, but it is not generally conserved in MNH halogenases. According to sequence alignments, the amino acyl-halogenases have a serine in the respective position and in the structure of SyrB2 this residue occupies a similar space as Glu224 in the HctB model. Therefore, the particular proposed molecular mechanism may be specific to the subtype of fatty acyl-halogenases. However, a principle role of chloride in the activation of halogenases by triggering changes in the protein structure may be a general feature of the MNH halogenase mechanism: Interestingly, in SyrB2 a major rearrangement of polypeptide strands was obtained upon binding of Fe(II) and chloride to an α-KG containing structure.[1] In CurA, concomitant coordination of α-KG and chloride to the iron center resulted in a structural reorganization.[3] By contrast, Apo-CytC3, which could not be crystallized

with a metal coordinated halide, did not change its open conformation upon simultaneous Fe(II) and α-KG binding. Even though structural data are scarce and the presence of charged buffer molecules may compromise their interpretation, it is noteworthy that in the available studies conformational rearrangements were always accompanied by a switch from a halide free to a halide coordinating metal center. Particularly for halogenases with low chloride affinity and those that perform multiple halogenations per substrate and therefore consume the bound halide, a chloride dependent activation makes sense as it prevents undesired hydroxylation after the primary halogenation step.

To elucidate by what precise structural interactions chloride binding activates the individual halogenases, additional and comparative studies of halogenase structures will be required. In the light of our findings, correct iron redox states and the absence of exogenous charged molecules in the active site will be crucial. This will allow to faithfully unveil the structural rearrangements upon chloride binding, which will likely be sensitive to the electrostatic interactions at work.

Materials and Methods

Recombinant ApoHctB was expressed, purified, and reconstituted to HoloHctB as described previously[11] except the final enzyme solution was exchanged into 20 mM Bis-Tris pH 7.5 for stopped flow analysis.

Stopped flow UV/VIS spectrophotometric analysis. The enzyme was made O_2 free by ~ 20 cycles of evacuation and N_2-flushing in an air-tight V-Vial (Wheaton, Millville/USA) capped with a screw-top septum (Supelco, Bellefonte/USA) and subsequently 0.95 equivalents of $Fe(II)SO_4*7H_2O$ and α-KG (Sigma Aldrich, St. Louis/USA) were added from stock solutions (10 mM) in a N_2-purged glove box. The enzyme preparations (440 μM – 550 μM active sites) were mixed at 10°C with equal volumes of 20 mM Bis-Tris buffer, pH 7.5 that contained ~ 1.4 mM O_2 if not stated otherwise and 1 M NaCl if required, using a SX20 Stopped-Flow UV/VIS spectrophotometer (Applied Photophysics, Leatherhead/UK), which was equipped with a polychromatic light source. Note that in a previous study the addition of one equivalent α-KG to HoloHctB (100 μM) yielded maximum initial O_2 consumption rates and thus, in the here performed enzyme kinetic measurements, which used ~ 95% cofactor-loaded active sites (> 100 μM), saturating conditions were ensured.[11] Spectroscopic analyses were carried out by using a SX20 photodiode array detector, whereby absorptions from 270 nm to 730 nm were recorded with 400 collected data points in the first 0.4 seconds and 1 data point per 100 ms in the remaining 40 or 60 seconds. Absorbance traces of ≥ 4 measurements per condition were averaged and fit using the following methods: (i) When rates of signal changes were so slow that they had linear characteristics over the monitored time, rates were determined based on the extinction coefficient of the *α-KG-Fe(II)-HctB* complex (ε_{500nm} = 0.15 mM^{-1})[11] or the theoretical extinction coefficient calculated from the chloride containing *α-KG-Fe(II)-ApoHctB* absorbance trace at 320 nm and 500 nm (this work). Rates were then related to the applied metal concentration to give specific rates. (ii) For traces that followed a monophasic exponential rise to a stable value, the following formula was applied:

$$f = y_0 + a(1 - e^{-k_c x})$$
(1)

(iii) In the case of apparently bi- or triphasic curves, the Pro-Kineticist 1.0.10 software (Applied Photophysics) was used and the trace at the respective constant wavelength was fit to a sequential model in the form of A→B→C and A→B→C→D, respectively.

CD/MCD spectroscopy. If not stated otherwise, for CD/MCD spectroscopy, protein samples were exchanged into deuterated 50 mM HEPES/NaOD buffer pD 7.5, containing 1 M NaCl if required, using 4 ml Amicon® Ultra Centrifugal Filters and concentrated to 1.5-3.5 mM. O_2 was removed from the samples by ~ 20 cycles of evacuation and argon flushing. Ferrous ammonium sulfate and α-KG were added to the enzyme preparations in microliter quantities from deuterated, anaerobic stock solutions in a N_2-purged glove box, where the final sample was filled into a CD or MCD cell. CD measurements were carried out on a Jasco J-730W spectropolarimeter at 283 K. Spectra were corrected by subtracting the respective cofactor-free spectrum. MCD samples were prepared from CD samples by adding glassing agent till saturation was reached at approximately 50% sucrose. Samples were injected into MCD cells, frozen and stored in liquid N_2 until use. MCD data were recorded on a Jasco J-730W spectropolarimeter equipped with an Oxford Instruments Spectromag 4000 superconducting magnet and a liquid N_2-cooled InSb detector. To affirm the authenticity of the MCD signals, field dependencies at -7, +3.5 and +7 T were recorded at 5 K. The C-term origin of the signals was proven by tracking temperature dependencies at 5, 20 and 40 K at +7 T (3 scans per condition were averaged).[20] The obtained data were corrected by subtraction of the zero-field spectrum at the respective temperature. MCD spectra were compared to their respective CD counterparts in order to verify that transitions had the same energies in CD and MCD.

MD simulations. A structural model of the HctB halogenase domain was created by homology modeling as described previously.[11] MD simulations were performed with the YASARA Structure suite, version 12.11.20 (YASARA Biosciences).[26] A periodic simulation cell, which comprised the whole enzyme and added 5 Å in each dimension was used with explicit solvent. The AMBER99 force field[27] was employed and long-range electrostatic potentials were calculated with the Particle Mesh Ewald (PME) method, with a cutoff of 7.864 Å.[28,29] Force field parameters for α-KG were generated with the AutoSMILES utility.[30] Fe(II) parameters were defined to reflect the principle octahedral geometry of the Fe(II) center. An equilibrium spring length that was derived from DFT calculations was used for all ligands[31] with a spring stretching force constant of 2000 N m^{-1}. Ligand-Fe-ligand angles of minimum energy were defined as 180° for opposite ligands and as 90° for all others, with an angle bending force constant of 1000 kJ mol^{-1} rad^2. In this way, force field parameters for (i) N-ε of His112, N-ε of His228 and α-KG coordinating carboxylate and keto-oxygen were defined for chloride free ApoHctB. (ii) For the chloride containing complex, analogous parameters were included for the iron coordinated chloride ion. Note that no force field parameters were defined for Glu224. (iii) In order to investigate the impact of a putative strongly bound water at the halide position, the chloride atom was substituted by a water molecule and respective parameters were adapted accordingly. The oxidation states of the iron cofactor in all complexes was adjusted to +2. The hydrogen bonding network was optimized by the method of Hooft and co-workers,[32] and YASARAs' pKa values at pH 7.5 were assigned.[33] The simulation cell was filled with water at a density of 0.997 g mL^{-1} using NaCl concentrations of 0.001 M and 1 M, respectively. After relaxation of the solvent the system was energy minimized, whereby a steepest descent minimization was applied to remove conformational stress, followed by a simulated annealing minimization until

convergence was reached (< 0.05 kJ mol^{-1} per 200 steps). Integration time steps were set to 1.33 and 4 fs for intra- and intermolecular forces, respectively. MD simulations at 298 K were initiated, whereby integration time steps for intramolecular and intermolecular forces were set to 1.25 fs and 2.5 fs, respectively.

Acknowledgements

References

(1)   Blasiak, L. C.; Vaillancourt, F. H.; Walsh, C. T.; Drennan, C. L. *Nature* **2006**, *440*, 368-371.

(2)   Wong, C.; Fujimori, G. D.; Walsh C. T.; Drennan, C. L. *J. Am. Chem. Soc.* **2009**, *131*, 4872-4879.

(3)   Khare, D.; Wang, B.; Gu, L.; Razelun, J.; Sherman D. H.; Gerwick W. H.; Hakansson, K.; Smith J. L. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14099-14104.

(4)   Vaillancourt, F. H.; Yeh E.; Vosburg D. A.; O'Connor S. E.; Walsh, C. T. *Nature* **2005**, *436*, 1191-1194.

(5)   Vaillancourt, F. H.; Yin, J.; Walsh, C. T. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 10111-10116.

(6)   Ueki, M.; Galonic, D. P.; Vaillancourt, F. H.; Garneau-Tsodikova, S.; Yeh, E.; Vosburg, D. A.; Schroeder, F. C.; Osada, H.; Walsh, C. T. *Chem. Biol.* **2006**, *13*, 1183-1191.

(7)   Galonic, D. P.; Vaillancourt, F. H.; Walsh, C.T. *J. Am. Chem. Soc.* **2006**, *128*, 3900-3901.

(8)   Neumann, C. S.; Walsh, C. T. *J. Am. Chem. Soc.* **2008**, *130*, 14022-14023.

(9)   Gu, L.; Wang, B.; Kulkarni, A.; Geders, T. W.; Grindberg, R. V.; Gerwick L.; Håkansson, K.; Wipf, P.; Smith, J. L.; Gerwick, W. H.; Sherman, D. H. *Nature* **2009**, *459*, 731-735.

(10)   Jiang, W.; Heemstra, J. R.; Forseth, R. R.; Neumann, C. S.; Manaviazar, S.; Schroeder, F.C.; Hale, K. J.; Walsh, C. T. *Biochemistry* **2011**, *50*, 6063-6072.

(11)   Pratter, S. M. *et al*, unpublished

(12)   Bruijnincx, P. C. A.; van Koten, G.; Klein Gebbink, R. J. M. *Chem. Soc. Rev.* **2008**, *37*, 2716–2744.

(13)   Galonic, D. P.; Barr, E. W.; Walsh, C. T.; Bollinger, J. M.; Krebs, C. *Nature Chem. Biol.* **2007**, *3*, 113-116.

(14)   Kojima, T.; Leising, R. A.; Yan, S.; Que, L., Jr. *J. Am. Chem. Soc.* **1993**, *115*, 11328-11335.

(15) de Visser, S. P.; Latifi, R. *J. Phys. Chem.* **2009**, *B 113*, 12-14.

(16) Pandian, S.; Vincent, M. A.; Hillier, I. H.; Burton, N. A. *Dalton Trans.* **2009**, *31*, 6201-6207.

(17) Kulik, H. J.; Blasiak, L. C.; Marzari, N.; Drennan, C. L. *J. Am. Chem. Soc.* **2009**, *131*, 14426-14433.

(18) Matthews, M. L.; Neumann, C. S.; Miles, L. A.; Grove, T. L.; Booker, S. J.; Krebs, C.; Walsh, C. T.; Bollinger, J. M. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 17723-17728.

(19) Kulik, H. J.; Drennan, C. L. *J. Biol. Chem.* **2013**, *288*, 11233-11241.

(20) 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.

(21) Neidig, M. L.; Brown, C. D.; Light, K. M.; Galonic´ Fujimori, D.; Nolan, E. M.; Price, J. C.; Barr, E. W.; Bollinger, J. M., Jr.; Krebs, C.; Walsh, C. T.; Solomon, E. I. *J. Am. Chem. Soc.* **2007**, *129*, 14224-14231.

(22) Straganz, G. D.; Diebold, A. R.; Egger, S.; Nidetzky, B.; Solomon, E. I. *Biochemistry* **2010**, *49*, 996-1004.

(23) Hamada, Y. Z.; Carlson, B. L.; Shank, J. T. *Syn. React. Inorg. Met.* **2003**, *33*, 1425–1440.

(24) Matthews, M. L.; Krest, C. M.; Barr, E. W.; Vaillancourt, F. H.; Walsh, C. T.; Green, M. T.; Krebs, C.; Bollinger, M. J., Jr. *Biochemistry* **2009**, *48*, 4331–4343.

(25) Zhou, J.; Kelly, W. L.; Bachmann, B. O.; Gunsior, M.; Townsend, C. A.; Solomon, E. I. *J. Am. Chem. Soc.* **2001**, *123*, 7388-7398.

(26) YASARA (9.11.9 9.11.9) Vienna, Austria: YASARA Biosciences GmbH; **2012**

(27) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem*. **2000**, *21*, 1049-1074 .

(28) Cornell, W.D.; et al. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.

(29) Essman, U.; et al. *J. Chem. Phys.* **1995**, *B 103*, 8577-8593.

(30) Jakalian, A.; Jack, D. B.; Bayly, C. *J. Comput. Chem.* **2002**, *23*, 1623-1641.

(31)   Pandian, S.; Vincent, M. A.; Hillier, I. H.; Burton, N. A. *Dalton Trans.* **2009**, 6201-6207.

(32)   Hooft, R. W.; Vriend, G.; Sander, C.; Abola, E.E. *Nature* **1996**, *381*, 272.

(33)   Krieger, E.; Nielsen, J. E.; Spronk, C.A.; Vriend, G. *J. Mol. Graph. Model.* **2006**, *25*, 481-486.

Supplementary Information:

The role of chloride in the mechanism of O_2 activation at the mononuclear nonheme Fe(II) center of the halogenase HctB

Sarah M. Pratter[a], Kenneth M. Light[b], Edward I. Solomon[b] and Grit D. Straganz[a]*

[a] Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12, A-8010 Graz, Austria

[b] Department of Chemistry, Stanford University, Stanford, California 94305, USA

corresponding author, e-mail: grit.straganz@tugraz.at, Tel: +43-316-8738414 Fax: +43-316-8738434

Table S1. Correlation of O_2 concentration and *α-KG-Fe(II)-HoloHctB* complex decay. First order rate constants in the presence of 0.5 M NaCl were determined at ~ 325 µM, ~ 550 µM, and ~ 700 µM O_2 at 500 nm.

O_2 concentration (µM)	k_{SC1} (s^{-1})	k_{SC2} (s^{-1})	k_{SC3} (s^{-1})
700	35 ± 5	6.7 ± 0.3	0.22 ± 0.01
550	(38)*	4.1 ± 0.2	0.19 ± 0.01
325	42 ± 8	5.2 ± 0.2	0.15 ± 0.01

* The determination of this value by fitting was not possible. Consequently, the value was kept constant in order to be able to fit k_{SC2} and k_{SC3}.
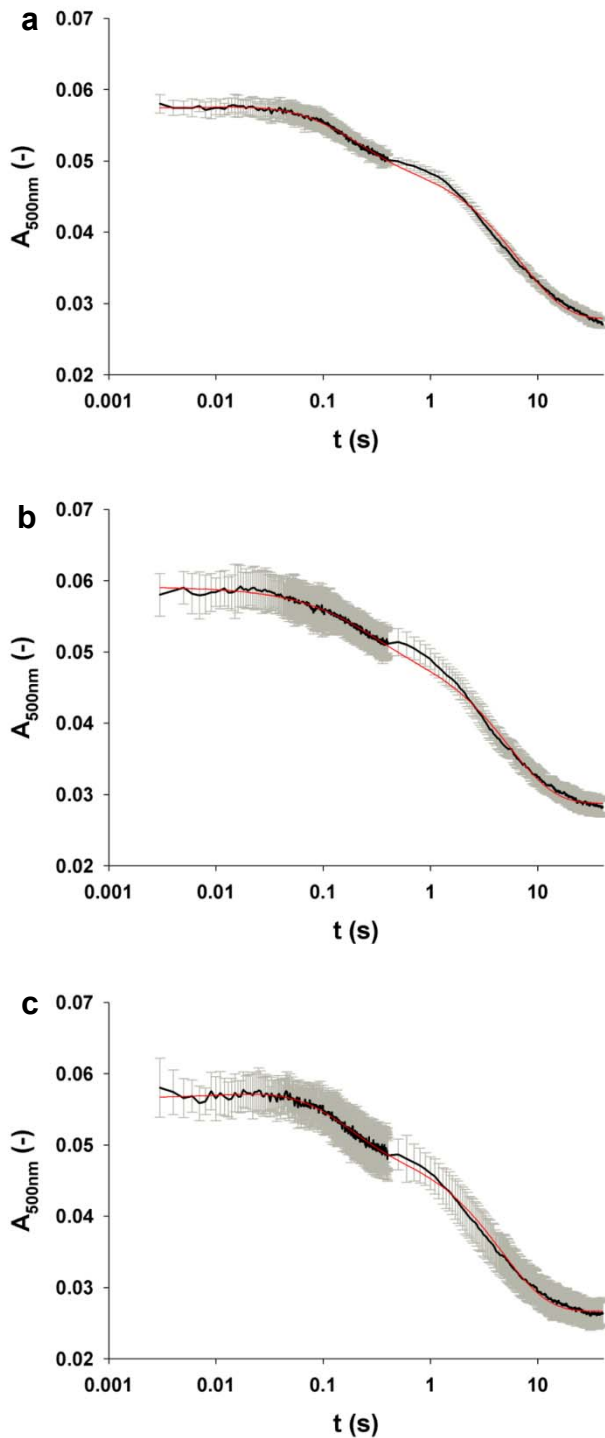
Figure S1. α-KG-Fe(II)-HoloHctB complex decay at varying O_2 concentrations. Average traces monitored at 500 nm in the presence of 0.5 M NaCl and (a) 325 µM, (b) 550 µM (c) and 700 µM O_2 (black) were fit (red) via the Pro-Kineticist 1.0.10 software according to the Materials and Methods section. Standard deviations are shown (gray).
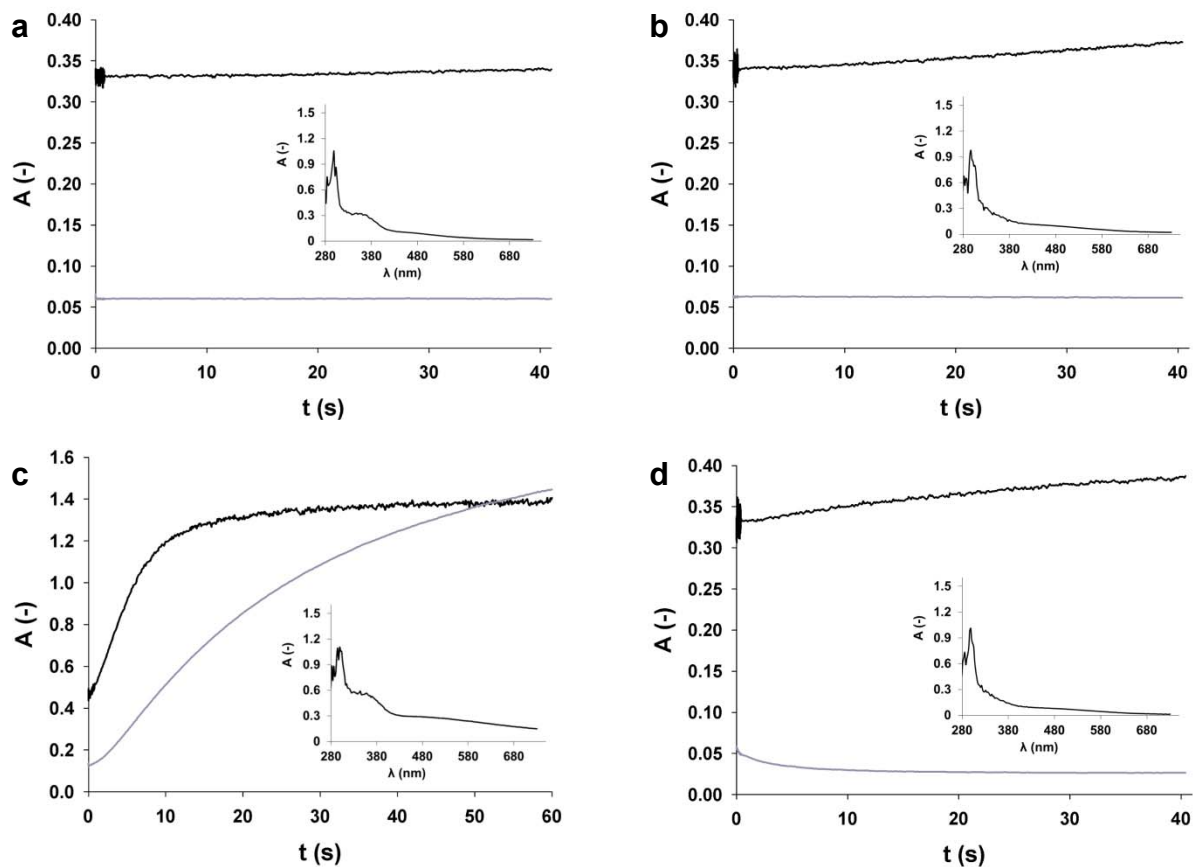
Figure S2. Stopped flow absorption kinetic traces of HctB. The reaction of *α-KG-Fe(II)-ApoHctB* with (a) O_2 and (c) O_2 and NaCl and the reaction of *α-KG-Fe(II)-HoloHctB* with (b) O_2 and (d) O_2 and NaCl was tracked at 320 nm (black) and 500 nm (gray). The inset shows the wavelength spectrum of the respective reaction at 0.05 s.
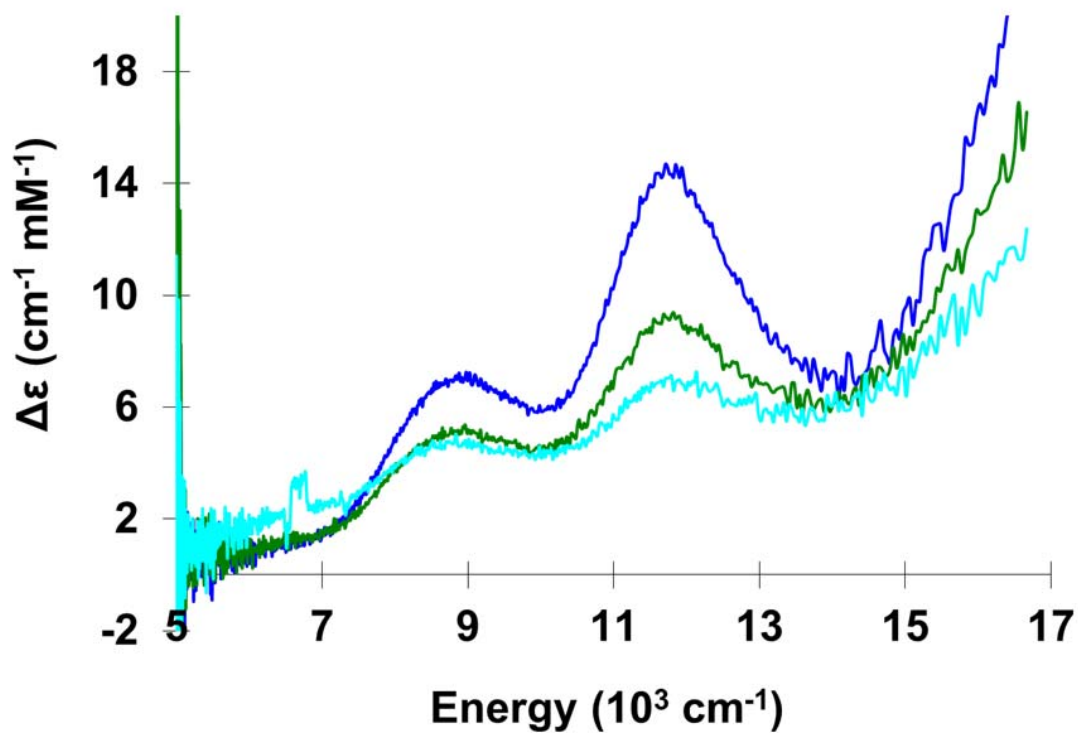
Figure S3. Low-temperature, 7 T MCD spectra of the α-KG-Fe(II)-HoloHctB complex in the absence of NaCl. 5 K (dark blue), 20 K (green) and 40 K (light blue) spectra were recorded at a pD of 9.1 in order to test for putative charge effects on the Fe(II) center geometry.
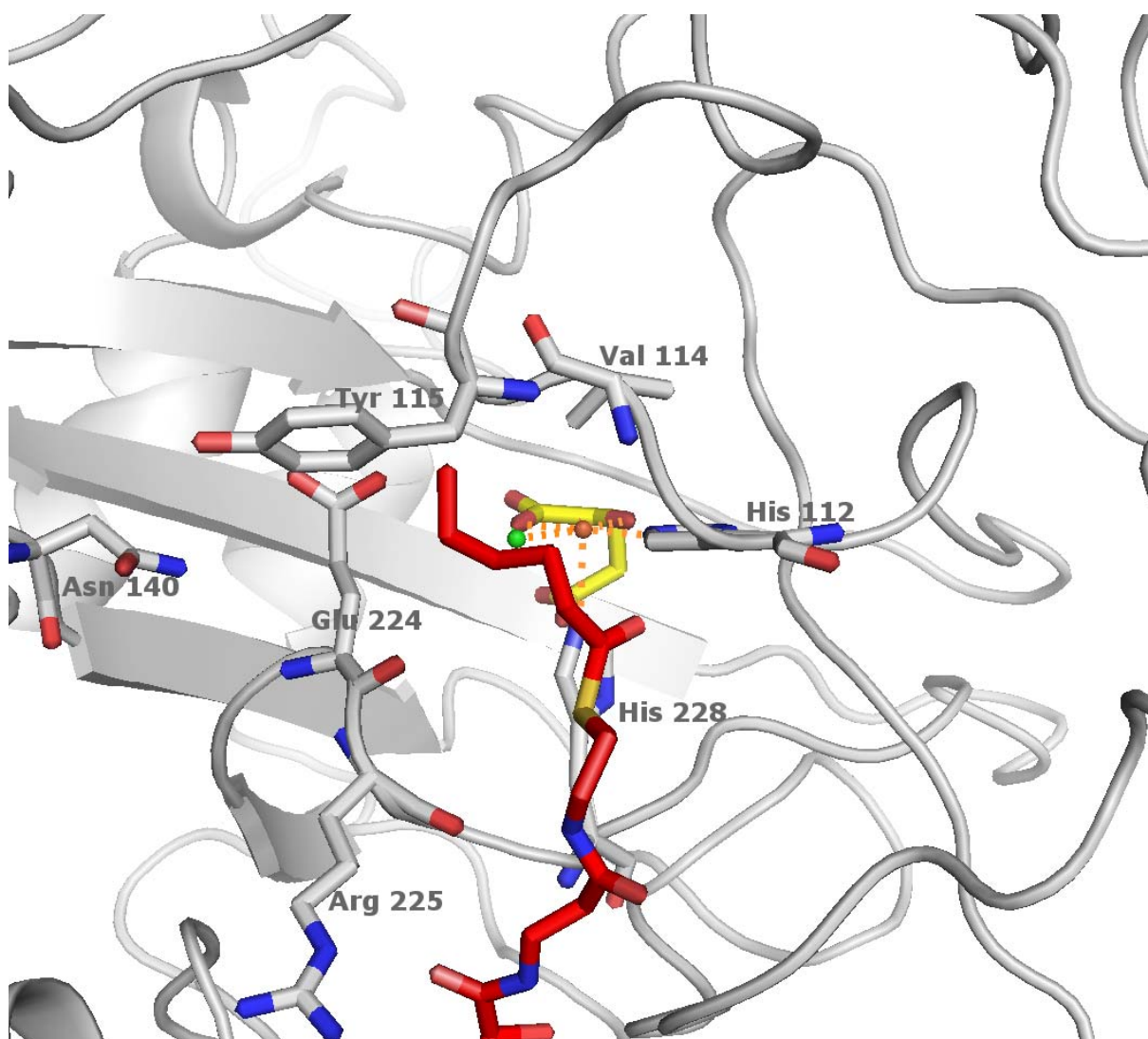
Figure S4. Model of the HctB halogenase domain with the hexanoyl-phosphopantetheinyl-substrate residue (red) docked into the active site.[11] The residues Val114, Tyr115, Glu224, and Arg225 line the substrate entrance tunnel to the active Fe(II) center (orange), which is ligated by His112 and His228 (gray), α-KG (yellow), and chloride (green). Note that the perspective shown corresponds to the one shown in Figure 5 a and b.
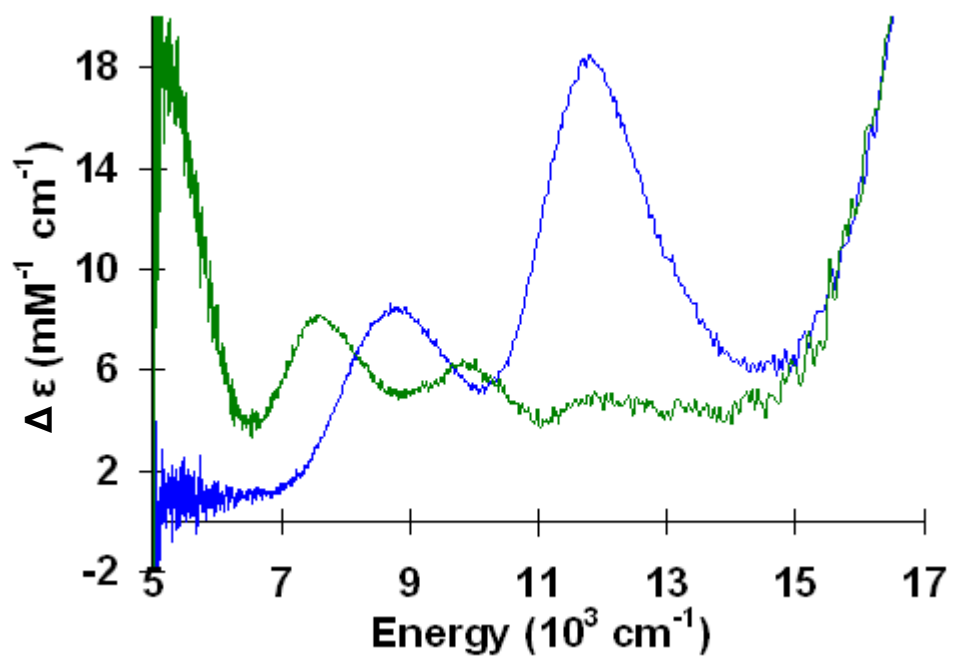
Figure S5. The role of chloride in metal center coordination in HoloHctB. Overlay of the MCD signals of the *α-KG-Fe(II)-HoloHctB* complex in the absence (blue) and presence (green) of NaCl.

Steric control of hydroxylation by tuning the interplay of metal center geometry and protein structure

Sarah M. Pratter[a], Cornelia Konstantinovics[a], Cristiana L. M. DiGiuro[a], Erich Leitner[b], Devesh Kumar[c], Sam P. de Visser[d], Gideon Grogan[e] and Grit D. Straganz[a*]

[a] Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12, A-8010 Graz, Austria.

[b] Institute of Analytical Chemistry and Food Chemistry, Graz University of Technology, Stremayrgasse 9, A-8010 Graz, Austria.

[c] Department of Applied Physics, School for Physical Sciences, Babasaheb Bhimrao Ambedkar University, Vidya Vihar, Rae Bareilly Road, Lucknow 226-025, India.

[d] Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, Manchester M1 7DN, United Kingdom.

[e] York Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York, YO10 5DD, United Kingdom.

corresponding author, e-mail: grit.straganz@tugraz.at, Tel: +43-316-8738414 Fax: +43-316-8738434

Abstract

O_2 dependent enzymatic iron centers can perform synthetically challenging stereoselective oxidations of aliphatic hydrocarbon moieties. The steric control of these reactions can open new synthetic routes for the biotechnological industry. We probe the molecular principles that control stereochemistry during catalysis at a mononuclear nonheme iron center. Based on the hypothesis of a strongly conserved trigonal bipyramidal metal center geometry throughout catalysis, we redesign (*S*)-*p*-hydroxymandelate synthase from *Streptomyces coelicolor* with the objective to obtain stereo-complementary reactivity. Phenylpyruvate oxidation to mandelate is switched from 99.6% (*S*)- to 97.4% (*R*)-selectivity in the tailored enzyme. The product-bound crystal structure of the created variant and molecular dynamics simulations demonstrate that the interplay of metal center geometry and protein scaffold leads to two distinct ligand binding modes that result in enantio-complementary pathways. The highly active (*R*)-mandelate synthase opens up a new route for the efficient biosynthesis of the pharmaceutical building block (*R*)-mandelate from feedstock.

Introduction

Natural stereoselective oxidations by molecular oxygen have long fascinated chemists and these chemically challenging, 'green' reactions have great potential in industrial processes.[1] However, the responsible enzymes in nature have a finite synthetic repertoire. A major goal is therefore to control their stereoselectivity and to produce a desired stereo-center within a certain target molecule. There have been some major successes recently in the design of stereo-controlled oxygenation reactions by cytochromes P450 (P450s),[2a-d] which have been established as a powerful platform for designed oxygenations.[3] By contrast, O_2 and mononuclear nonheme Fe^{II} (MNH) dependent enzymes are not generally used in enzyme design, even though they parallel and complement the catalytic potential of P450s with their ability to oxidize aliphatic carbon atoms. Their high regio-, chemo-, and stereoselectivities,[4] however, limit the applicability of an evolutionary approach. Rational methods, where target substitutions can be pinpointed based on the identification and quantitative description of the underlying structural factors of catalytic selectivity and efficiency may in future help to bridge this gap.

To investigate what features control the stereoselective hydroxylation by a MNH center, we chose the highly stereoselective model enzyme (*S*)-*p*-hydroxymandelate synthase (HMS),[5] with the goal to introduce (*R*)-selectivity into the scaffold. The reaction mechanism follows that of α-ketoacid dependent MNH enzymes and involves an oxidative decarboxylation of the *p*-hydroxyphenylpyruvate substrate's ketoacid moiety that leads to a high valent Fe^{IV}-oxo species.[6] A subsequent hydrogen atom abstraction from the α-carbon atom of the resulting reaction intermediate, *p*-hydroxyphenylacetate,[7] and rebound of the hydroxyl group gives the

chiral product (Figure 1).[6] The crystal structure of HMS from *Amycolatopsis orientalis* (*Ao*HMS) has been determined and shows the native reaction product coordinated to the metal center.[8] Recently, it was found that the high enantioselectivity can be compromised by substituents on the aromatic ring of phenylpyruvate (PP).[9] However, no structural features have been identified that cause the generally observed (*S*)-enantioselectivity of HMS.
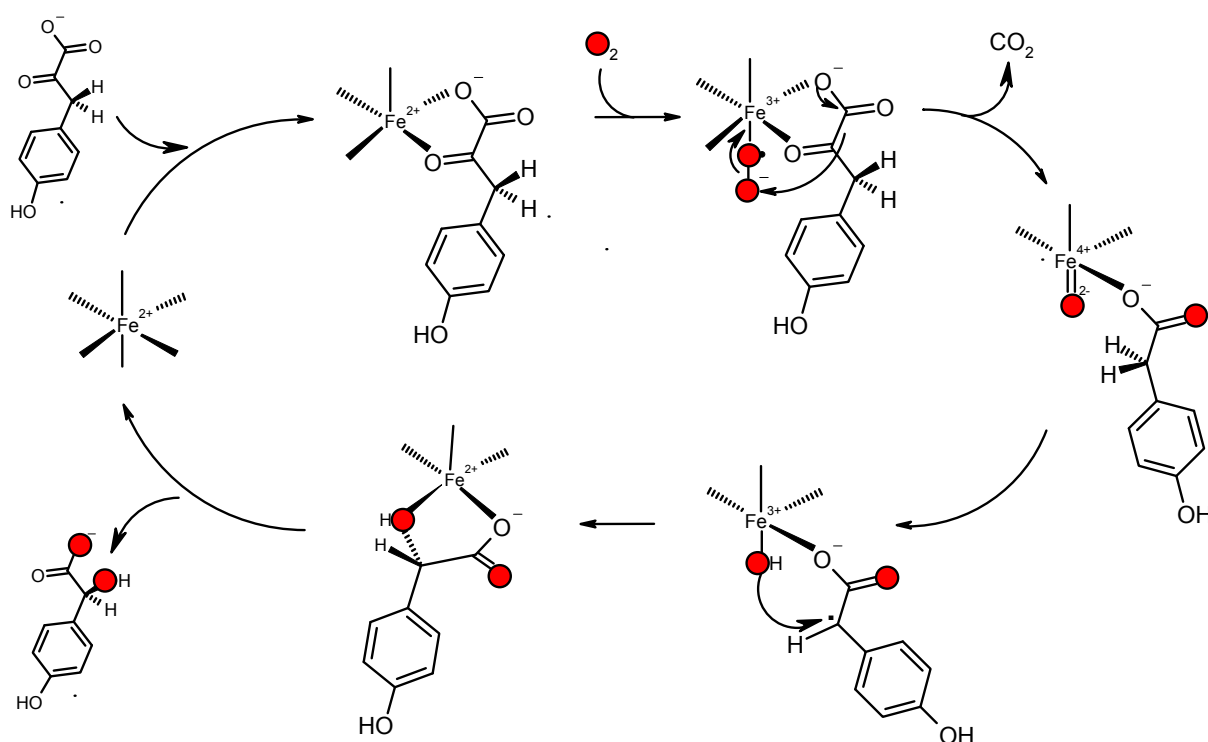
Figure 1. Generally accepted reaction mechanism of HMS enzymes.[6] The catalytic cycle is initiated by substrate binding, which primes the iron center for O_2 activation. O_2 then attacks the substrate and leads to the decarboxylation of the α-ketoacid and the formation of a high valent $Fe^{IV}=O$ species. The latter abstracts a benzylic hydrogen atom. In a rebound reaction, the resulting Fe^{III}-OH species transfers a hydroxyl radical to the benzylic position resulting in the mandelate product, which leaves the active site.

Results

Structural features that determine HMS enantioselectivity.

Upon close inspection of the metal center geometry in *in silico* docked complexes of HMS from *Streptomyces coelicolor* (*Sc*HMS) with mandelate (MA) analogs in (*S*)-confomation, the structural models revealed that the ligands were generally oriented in a slightly distorted trigonal bipyramidal geometry at the metal center. The MA's α-hydroxy group and the metal coordinating glutamate were in the axial position and two nitrogens from the coordinating histidines together with the ligand's carboxylate oxygen formed the trigonal basis of the pyramid. By contrast, the analogous (*R*)-MA-enzyme complex showed more pronounced deviations from the ideal symmetry (Figure S1).[9] Notably, density functional theory (DFT) calculations of the reactive $Fe^{IV}=O$ and $Fe^{III}-OH$ species in the catalytic cycle of α-ketoacid dependent enzymes found that a trigonal bipyramidal geometry was one of two probable organizations of these high valent iron centers.[10] The active site of the crystallized *Ao*HMS-(*S*)-hydroxy-MA complex displays a trigonal bipyramidal organization, and it has been proposed that the enzyme-product complex reflects the geometry of the final hydroxylation step.[8] This poses the question of what impact the ligand field symmetry may have on enantioselectivity during catalysis. In models of (*R*)- and (*S*)-MA bound *Sc*HMS, in which the trigonal bipyramidal geometry was constrained, the (*R*)-MA ligand did not occupy the hydrophobic pocket that accommodates the aromatic substituent of (*S*)-MA. In contrast to the non-restricted *in silico* docking results, the aromatic substituent pointed toward the presumable substrate entrance channel and thereby clashed with residue Tyr359 (Figure S1). *pro*-(*S*) and *pro*-(*R*) transition states of H-abstraction and rebound hydroxylation gave similar principle spatial orientations when docked into the trigonal bipyramidal metal center of

*Sc*HMS. Consequently we investigated how the aromatic ring's destabilization in the hydrophobic pocket and its stabilization in the alternative orientation influenced enantioselective catalysis. The steric clash with residue Tyr359 upon (*R*)-MA binding in our *in silico* model could be removed by its substitution with alanine. In order to obstruct the hydrophobic pocket of *Sc*HMS, potentially effective point mutations were identified using *in silico* site saturation mutagenesis with the following set of criteria: The variation (i) results in a low-energy apo-enzyme structure in order to prevent energetically driven reorganization of the active site; (ii) has low-energy amino acid conformers that all interfere with the hydrophobic binding pocket; and (iii) does not interact with the first coordination sphere of the metal ion and thus interfere with its electronic or geometric structure. Using the algorithm described in the Supplementary Information (Figure S2), six target single point substitutions (Ser221Met, Val223Phe, Val223Met, Val223Trp, Ile236Phe and Tyr359Ala) were identified and introduced into the *Sc*HMS scaffold *in vitro*.

Spectra of PP conversion products by the wild type and the resulting enzyme variants were subsequently analyzed via GC-MS and HPLC-MS: All but one variation had a significant impact on the enantioselectivity of the wild type that formed (*R*)-MA of only 0.4 % enantiomeric purity (Figure 2 a). The most pronounced, a ~ 1900-fold increase in (*R*)-selectivity was obtained by the mutation Ser221Met, which resulted in (*R*)-MA of 88.5 % enantiomeric purity, followed by the substitution Val223Phe with 65.1 % (*R*)-MA. While both of these mutations aimed at obstructing the hydrophobic pocket in the active site, the opening of space for (*R*)-MA binding by the Tyr359Ala mutation resulted in a ~ 47-fold increase of (*R*)-MA selectivity. Notably, Ser221, Val223 and Tyr359 are strongly conserved throughout all annotated putative *p*-hydroxyphenylpyruvate dioxygenases and HMS. Combinations of the most

successful substitutions further enhanced enantioselectivity: The double variants Val223Phe_Tyr359Ala and Ser221Met_Tyr359Ala produced (*R*)-MA with 93.1 % and 94.6 % enantiomeric purity, respectively. Triple variants Ser221Met_Val223Met_Tyr359Ala and Ser221Met_Val223Phe_Tyr359Ala further increased the ratio to 95.0 % and 97.4 % in favor of (*R*)-selective hydroxylation, which corresponds to a ~ 4700- and ~ 9300-fold enhancement compared to wild type *Sc*HMS. The latter translates to a difference in transition state energy $\Delta\Delta G\ddagger$ of 18.8 kJ mol^{-1}. No reaction side products that are indicative of C2-C3 cleavage or uncoupled decarboxylation[11] were formed in detectable amounts and product recovery for MA was > 95 % for all variants except for variant Ile236Phe, which showed substantial losses of ~ 80 % that were not accounted for. This may indicate the formation of alternative oxygenation products in this variant that escaped our analysis. Notably, ring oxidation products have been reported from mutational studies of the HMS' active sites.[12a,b]

Kinetic analyses.

All *Sc*HMS variants had enzymatic activities towards PP that were roughly in the same range of the wild type enzyme (k_{cat}^{app} = 0.62 s^{-1}, k_{cat}^{app}/K$_M$ = 2.1 mM^{-1} s^{-1}) (Figure 2 b). The point mutants Val223Phe and Ser221Met, which are believed to obstruct the hydrophobic pocket, increased catalytic efficiencies by an order of magnitude with k_{cat}^{app}/K$_M$ values of 40 mM^{-1} s^{-1} and 34 mM^{-1} s^{-1} respectively, while the active site 'opening' variant Tyr359Ala did not significantly alter wild type efficiency (k_{cat}^{app}/K$_M$ = 2.1 mM^{-1} s^{-1}). Remarkably, the combination of the single point mutations Ser221Met and Tyr359Ala led to a ~ 4- and ~ 5-fold increase of k_{cat}^{app} (2.4 s^{-1}) and k_{cat}^{app}/K$_M$ (10 mM^{-1} s^{-1}). The other double variant Val223Phe_Tyr359Ala, by contrast, showed a 42 % and 28 % decrease in the respective values (k_{cat}^{app} = 0.36 s^{-1}; k_{cat}^{app}/K$_M$ = 1.5 mM^{-1} s^{-1}). For the enzyme with highest (*R*)-enantioselectivity,

bearing the Ser221Met_Val223Phe_Tyr359Ala triplemutation, the turnover number was raised by 22% (k_{cat}^{app} = 0.76 s^{-1}). The apparent k_{cat}^{app}/K$_M$ of 6.9 mM^{-1} s^{-1} was > 3 times higher compared to the wild type enzyme. The less enantioselective Ser221Met_Val223Met_Tyr359Ala variant exhibited k_{cat}^{app} (2.0 s^{-1}) and k_{cat}^{app}/K$_M$ (18 mM^{-1} s^{-1}) values that were ~ 3- and > 8-fold greater than the wild types', respectively.
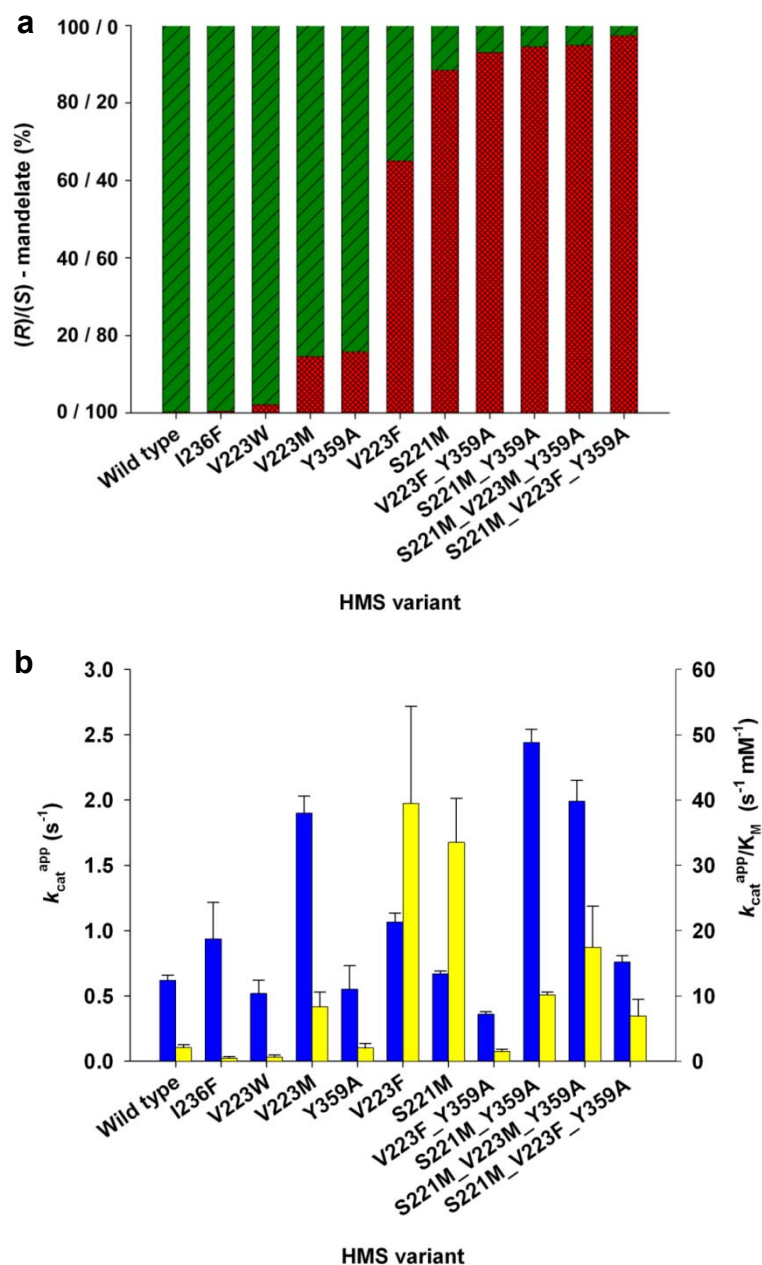

Figure 2. Enantioselectivities and activities of ScHMS wild type and variants. (a) The ratio of (R)- (red) and (S)- (green) MA products and (b) steady state kinetic constants k_{cat}^{app} (blue) and k_{cat}^{app}/K$_M$ (yellow) are shown.

Structural characterization of (R)-MA synthase.

The structure of the newly designed (R)-MA synthase, *Sc*HMS variant Ser221Met_Val223Phe_Tyr359Ala, in complex with (R)-MA and cobalt, which was used as a redox-inactive Fe^{II} mimic, was solved and refined to a resolution of 1.95 Å with R$_{cryst}$ and R$_{free}$ values of 17.5 % and 21.5 % respectively, using a monomer of the wild type structure of *Ao*HMS as the model. The asymmetric unit contained two molecules, each in complex with the non-native metal cofactor. The product's aromatic substituent occupied the free space at the beginning of the putative product exit channel,[9] which has been created by the Tyr359Ala mutation, while Met221 and Phe223 occupy space in the enzyme's hydrophobic binding pocket. The metal ion ligates (R)-MA in trigonal bipyramidal geometry and thus mirrors the product coordinated active site in the native structure of *Ao*HMS[11] (Figure 3).
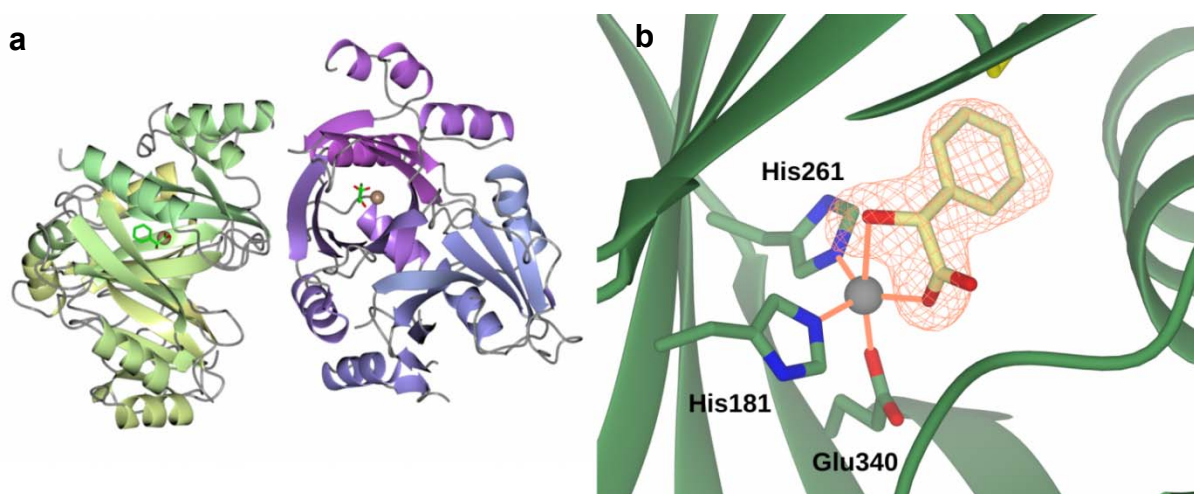
Figure 3. Crystal structure of the (R)-selective ScHMS triple variant. Panel a shows the overall structure of the observed dimer in complex with cobalt (brown) and the reaction product (R)-MA (green). Panel b shows a close-up of the active site: The oxo-groups of (R)-MA (yellow) and the side chains of His261, His181 and Glu340 form a trigonal bipyramidal geometry around the cobalt atom (grey). The ligand's electron density (coral) corresponds to the omit map (*Fo-Fc* map) contoured at a level of 3 σ obtained after the final rounds of refinement in the absence of the ligand.

Molecular dynamics (MD) simulations.

Computational MD studies of reaction intermediates in the scaffolds of wild type *Sc*HMS and its Ser221Met_Val223Phe_Tyr359Ala variant shed light on the impact of particular catalytic steps on the reaction's overall enantioselectivity. In each case intermediates and transition states from phenylacetate (PA) ligated Fe^{IV}=O centers up to the final products were analyzed, whereby the 'ideal' initial *pro*-(*S*) and *pro*-(*R*) geometries, which had been geometrically optimized by DFT calculations (Figures S3-S5), were the starting points of all simulations. The complexes generally showed two distinct localizations of the intermediates' phenyl rings that were prototypical for either the *pro*-(*S*) or the *pro*-(*R*) pathway (Figure 4 and Figures S6-S13). Inspections of ligand geometries and energies in the *Sc*HMS wild type protein suggest that product chirality is determined by the final hydroxylation step: Formation of the PA radical via the *pro*-(*S*) mechanism was energetically favorable. However, it displayed considerable out-of-plane flexibility and oscillated between a *pro*-(*S*) (43 %) and a *pro*-(*R*) (57 %) conformation, thus making both pathways equally feasible (Figure S10 and Table S4). The intermediate of the *pro*-(*S*) rebound transition state and the ultimate enzyme bound (*S*)-product were then both energetically favored compared to the (*R*)-selective pathway. By contrast, during MD simulations of the newly designed (*R*)-MA synthase the first, high valent Fe^{IV}=O complex of both *pro*-(*R*) and *pro*-(*S*) starting geometries adopted *pro*-(*R*) conformation. The subsequent PA radicals were entrapped in their respective *pro*-(*S*) (99 %) and *pro*-(*R*) positions (97 %) (Figure S10), while the *pro*-(*S*) pathway was energetically favored during hydroxylation. This suggests that the primary Fe^{IV}=O complex geometry co-determines enantioselectivity in the variant and apparently compensates the energetically favored (*S*)-trajectory during hydroxylation (Table S4). Overall, the computational methodology effectively rationalized the experimentally observed

enantioselectivities of the two enantio-complementary enzymes and gave a refined picture of the stereoselective hydroxylation process.
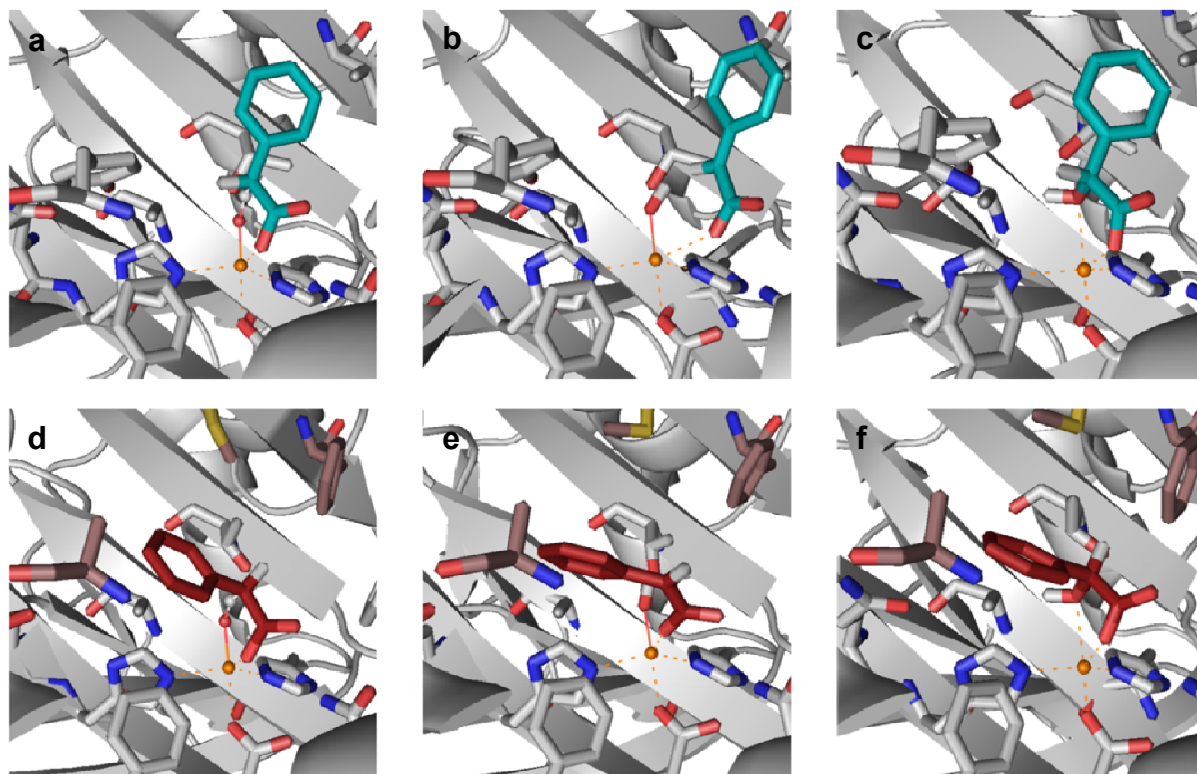
Figure 4. Structural representation of the catalytic steps that lead to chiral hydroxylation in wild type ScHMS (upper panels) and in the (R)-selective triple variant (lower panels). In panels a-e, the average structures of 10 ns MD simulations are shown, panel f depicts the variant crystal structure solved in this study. (Note that MD simulations gave similar active site geometries). The transition state of H-abstraction from PA by the Fe^{IV}=O center (a, d) results in an iron-bound hydroxy group and a substrate radical (b, e). The subsequent rebound step yields the chiral products (c, f). Residues within 5 Å of the metal (orange) are displayed, except Ile335, Leu358, and Ile236, which are omitted for clarity. Ligands of the (*S*)- and (*R*)-selective pathways are in teal and dark red, respectively. Mutations in the triple variant are shown in brown.

We demonstrate for the first time how the interplay of a rigid ligand field geometry and a protein structure control the stereoselective hydroxylation at an enzymatic metal center. The consideration of this principle has allowed the design of a highly (*R*)-selective MA synthase from an enantio-complementary HMS via the *in vitro* construction of a total of only ten enzyme variants. The resulting ~ 9300-fold enantioselectivity switch, the most pronounced change ever reported for an oxygenation, is accompanied by a triplication of enzyme efficiency. The designed chiral hydroxylase, which has no precedent in nature, opens up a new route for the efficient sustainable whole-cell (*R*)-MA biosynthesis from readily available renewable feedstock. Furthermore, this work and its underlying principles may be a first step to make enzymatic MNH centers accessible as a general platform for tailor-made biosynthetic oxygenations.

References

[1]   K. Sanderson, *Nature* **2011**, 469, 18-20.

[2]   a) S. Kille, F. E. Zilly, J. P. Acevedo, M. T. Reetz, *Nat. Chem.* **2011**, 3, 738–743; b) R. Agudo, G. D. Roiban, M. T. Reetz, *ChemBioChem* **2012**, 13, 1465-1473; c) W. L. Tang, Z. Li, H. Zhao, *Chem. Commun. (Camb)* **2010**, 46, 5461-5463; d) S. Q. Pham, G. Pompidor, J. Liu, X. D. Li, Z. Li, *Chem. Commun. (Camb)* **2012**, 48, 4618-4620.

[3]   V. B. Urlacher, M. Girhard, *Trends Biotechnol.* **2012**, 30, 26-36.

[4]   E. G. Kovaleva, J. D. Lipscomb, *Nat. Chem. Biol.* **2008**, 4, 186-193.

[5]   B. K. Hubbard, M. G. Thomas, C. T. Walsh, *Chem. Biol.* **2000**, 7, 931-942.

[6]   O. W. Choroba, D. H. Williams, J. B. Spencer, *J. Am. Chem. Soc.* **2000**, 122, 5389-5390.

[7]   D. D. Shah, J. A. Conrad, B. Heinz, J. M. Brownlee, G. R. Moran, *Biochemistry* **2011**, 50, 7694-7704.

[8]   J. Brownlee, P. He, G. R. Moran, D. H. Harrison, *Biochemistry* **2008**, 47, 2002-2013.

[9]   C. M. L. Di Giuro, C. Konstantinovics, U. Rinner, C. Nowikow, E. Leitner, Grit. D. Straganz, *PLOS ONE* **2013,** accepted

[10]   M. L. Neidig, A. Decker, O. W. Choroba, F. Huang, M. Kavana, G. R. Moran, J. B. Spencer, E. I. Solomon, *Proc. Natl. Acad. Sci. USA* **2006**, 103, 12966-12973.

[11]   T. K. Paine, H. Zheng, L. Que Jr., *Inorg. Chem.* **2005**, 44, 474-476.

[12]   a) M. Gunsior, J. Ravel, G. L. Challis, C. A. Townsend, *Biochemistry* **2004**, 43, 663-674; b) J. M. Brownlee, B. Heinz, J. Bates, G. R. Moran, *Biochemistry* **2010**, 49, 7218-7226.

Supplementary Information:

Steric control of hydroxylation by tuning the interplay of metal center geometry and protein structure

Sarah M. Pratter[a], Cornelia Konstantinovics[a], Cristiana L. M. DiGiuro[a], Erich Leitner[b], Devesh Kumar[c], Sam P. de Visser[d], Gideon Grogan[e] and Grit D. Straganz[a*]

[a] Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12, A-8010 Graz, Austria.

[b] Institute of Analytical Chemistry and Food Chemistry, Graz University of Technology, Stremayrgasse 9, A-8010 Graz, Austria.

[c] Department of Applied Physics, School for Physical Sciences, Babasaheb Bhimrao Ambedkar University, Vidya Vihar, Rae Bareilly Road, Lucknow 226-025, India.

[d] Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, Manchester M1 7DN, United Kingdom.

[e] York Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York, YO10 5DD, United Kingdom.

corresponding author, e-mail: grit.straganz@tugraz.at, Tel: +43-316-8738414 Fax: +43-316-8738434

A. *In silico* site saturated mutagenesis

A.1 Structural analysis of the product ligated active site.
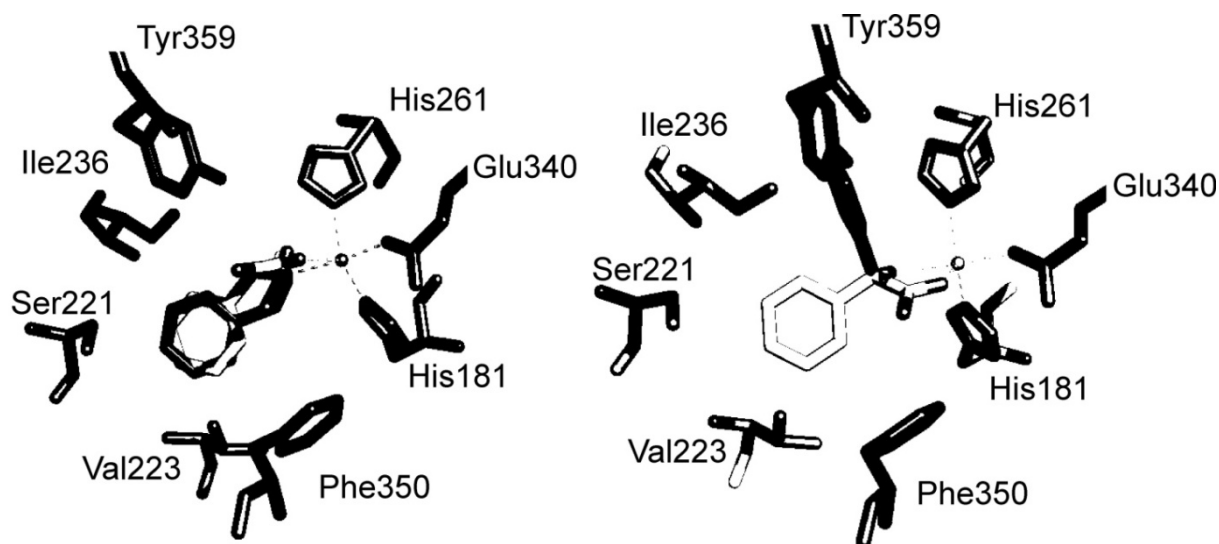
Figure S1. Active site models of wild type ScHMS, ligated with the putative products (S)-MA (white) and (R)-MA (black). Hydrogen bonds and amino acid residues in the ligand binding pocket are omitted for reasons of clarity. (a) Geometries of the (*R*)- and (*S*)-MA chelated metal center models, obtained by *in silico* docking, were analyzed. The (*S*)-MA ligated structure showed a distorted trigonal bipyramidal geometry that reflects product bound crystal structure data (PDB: 2R5V) with angle deviations < 2°. In the (*R*)-MA bound model, the positions of axial ligands, namely the carboxylate oxygen from Glu340 and the hydroxy oxygen from mandelate, were similarly preserved. However, the ligands that form the trigonal plane in the (*S*)-MA chelated structure showed major deviations in the (*R*)-MA bound model: Angles between the two metal-nitrogen bonds and the product's metal coordinated oxygen shifted from 110° to 139° (His 181) and from 124° to 100° (His261), respectively. The aromatic rings of the ligands in all structures occupied the same space. (b) When the coordination geometry of the crystal structure was preserved, the aromatic residue of (*R*)-MA was no longer situated in the substrate binding pocket but clashed with Tyr359. (Note that the cartoon from panel a is adapted from a previous study of our lab).[32]

A.2 Methods.

For *in silico* analysis, a structural model of *Sc*HMS, which was created using Swiss-model[1],[2] and validated using PROCHECK,[3] was applied.[32] The model was based on a sequence alignment with the crystal structure of *p*-(*S*)-hydroxy-MA ligated *Ao*HMS.[4] Total energies of the respective apo-enzyme systems were calculated *in vacuo* using the Gromos96 force field and Gromos96 43B1 parameters.[5] In a two-step approach *in silico* site saturation mutagenesis was performed. For each amino acid substitution the force field energies in a rigid protein structure were calculated for all rotamers using the penultimate rotamer library.[6] The respective structures were ranked in energy and variants with low energy topomers ($\Delta\Delta E$ < 10 kJ compared to the lowest energy) that did not intrude into the 2 Å sphere of the product's aromatic ring were excluded. Also, variants with low energy topomers ($\Delta\Delta E$ < 10 kJ) that interfered with metal ligating residues (2 Å sphere) and variants that only gave high energy protein structures (ΔE >10,000 kJ) were not considered. From the remaining mutations, those were excluded that had the potential to form a H-bond (cutoff 3.5 Å) with the $Fe^{IV}=O$ species that is formed during catalysis (Figure S2).

A.3 Results:
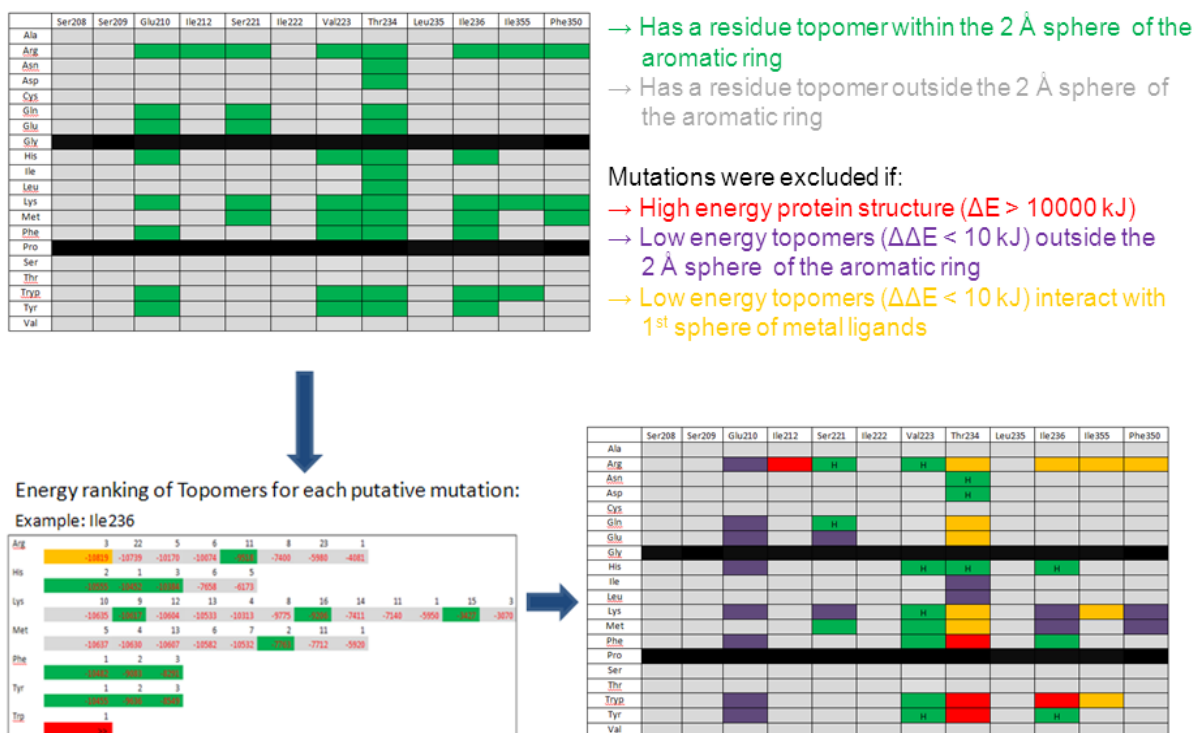
Figure S2. Flow chart that outlines the algorithm that was applied to an in silico site directed mutagenesis approach of ScHMS, in order to identify target residues that obstruct the binding pocket of the (S)-MA's aromatic ring. Residues which interfere with the aromatic binding pocket were identified (green, upper left panel) and variants were then subjected to force field calculations by the Gromos96 force field, whereby each rotamer of the library was considered. Species were then ranked in energy. The depicted sample table (left lower panel) outlines calculations for position Ile236: Only calculated rotamers that give potential energies < 0.0 J mol^{-1} for the total protein structure are shown. The rotamer number is given above the respective energy value. Cells are colored depending on the properties of the rotamer, whereby green indicates an orientation within 2 Å of the space occupied by the aromatic ring, grey marks an orientation outside of this 2 Å sphere. Orange indicates that the topomer enters the 2 Å sphere of the metal center, which comprises metal binding amino acid residues and the 2-hydroxyacetate moiety of the product ligand. Red indicates that the respective substitution does not give any conformation with an overall energy of < 0.0 J mol^{-1}. The lower right panel shows the final target variants (green), which were identified based on the following criteria: Variants with low energy topomers (ΔΔE < 10 kJ) that did not intrude into the 2 Å sphere of the product's aromatic ring (lilac) and variants with low energy topomers (ΔΔE <10 kJ) that interfere with metal ligating residues (2 Å sphere, yellow) were excluded. Substitutions that only give high energy protein structures (ΔE > 10,000 kJ) were also not considered. Furthermore, from the primarily identified variants, those were excluded that could form a H-bond (cutoff: 3.5 Å) with a putative Fe^{IV}=O intermediate (H in green squares).

B. HMS variant construction and expression and characterization of enzymatic activity and enantioselectivity

Chemicals were purchased from Sigma-Aldrich (St. Louis, USA) in the highest available purity. Enzymes and reagents for cloning were purchased from Promega (Madison, WI, USA). For the construction and production of enzyme variants, a p*KYB*1 vector, which bears the chromosomal *Sc*HMS gene fused to a C-terminal sequence that codes for a *Strep*-tag II streptavidin affinity tag, was subjected to site directed mutagenesis using the quick-change method.[7] Resulting DNA sequences were verified by sequence analyses at Eurofins MWG Operon (Ebersberg, Germany). *E. coli* BL 21 Gold (DE3) host cells (Stratagene, La Jolla, CA, USA), carrying the respective expression construct, were cultivated, harvested and purified as described previously.[8]

The activity of *Sc*HMS and variants was determined using the oxygen sensor 'Oxygen Microoptode' (Presens GmbH, Regensburg) and data were fit to the Michaelis-Menten equation to determine apparent turnover numbers (k_{cat}^{app}) and Michaelis Menten constants (K_M) at 260 µM O_2. Turnover numbers are related to the Fe(II) concentration as described previously.[32] After derivatization, reaction products were analyzed as MA esters via chiral GC-MS and unambiguously identified based on their retention time and mass-fingerprint using a method described previously.[32] Complementary chiral HPLC-MS chromatography was used to quantify the reaction products. [32]

C. Protein crystallization, data collection and processing, structure solution and refinement

C.1 Methods.

A solution of 10 mg mL^{-1} *Sc*HMS variant Ser221Met_Val223Phe_Tyr359Ala, which had been mixed previously with 1.25 mM CoCl$_2$ and 4 mM (*R*)-MA, was subjected to crystallization trials using a range of commercially available 96-well screens. The most promising hits were obtained using the Clear Strategy Screen[9] in conditions containing 0.1 M Tris/HCl buffer pH 8.5, 0.2 M MgCl$_2$ and 15% (w/v) PEG 4K. Larger crystals for diffraction analysis were prepared using the same enzyme-ligand solution with the hanging-drop (2 µl) vapour diffusion method in 24-well dishes. The best crystals were routinely obtained in drops containing 0.05 M Tris/Cl pH 9.0, 0.1 M MgCl$_2$, 7.5% (w/v) PEG 4K and 2.5% 2-Methyl-1,4-pentanediol. Crystals were flash-cooled in liquid nitrogen in the mother liquor and assessed for diffraction using a Rigaku Micromax-007HF x-ray source fitted with Osmic multilayer optics and a MARRESEARCH MAR345 imaging plate detector. Crystals that diffracted to more than 3 Å resolution were retained for detailed analysis. Datasets were collected on beamline io-4 at the Diamond Light Source Synchrotron in Oxford (data collection statistics are given in Table S1). The structure was solved with the program PHASER[10] using a monomer from the structure of *Ao*HMS (PDB 2R5V) as a model and built and refined using the programs COOT[11] and REFMAC[12] respectively. After building the protein and water molecules, a cobalt atom and the (*R*)-MA ligand were successfully modeled into the density of each subunit at the active site. The final metal- and product bound structure was validated using PROCHECK[3] and deposited in the Protein Databank with the accession code 3ZGJ.

C.2 Results.

Table S1. Data collection and refinement statistics for *Sc*HMS triple variant. Figures in brackets refer to the highest resolution shell.

	*Sc*HMS Ser221Met_Val223Phe_Tyr359Ala Complex with cobalt and (*R*)-MA
Beamline	Diamond I02
Wavelength (Å)	0.9795 Å
Space Group	*P*4
Unit cell (Å)	a = 122.9; = 122.9; c = 54.9
Resolution	61.44-1.95 (2.01-1.95)
Unique Reflections	57014 (4191)
Completeness (%)	100 (100)
R_{merge}	0.06 (0.49)
$R_{p.i.m.}$	0.04 (0.30)
Multiplicity	6.7 (6.9)
$<I/\sigma(I)>$	15.6 (3.9)
Protein atoms	5059
R_{cryst}/R_{free} (%)	17.5/21.5
r.m.s.d. 1-2 bonds (Å)	0.021
r.m.s.d. 1-3 bonds (°)	2.00
Average *B* factor (Å^2)	35

D. DFT modeling of (*R*)- and (*S*)-mandelate hydroxylation by a nonheme $Fe^{IV}=O$ active site model

D.1 Methods.

All calculations were performed using previously benchmarked and calibrated methods,[13],[14] and utilize DFT as implemented in *Gaussian*-09.[15] Our starting structure was taken from the *Sc*HMS model structures of this work, where we abbreviated the two histidine ligands (His181, His261) of iron with methylimidazole groups and the glutamic acid ligand (Glu340) with ethylacetate. In addition we added an oxo-group bound to iron and included the side chains of neighbouring glutamine (Gln325) and threonine (Thr234) as propanamide and ethanol, respectively, in our model that provide key stereochemical and hydrogen bonding interactions to the reaction substrate and/or products. The primary intermediate phenylacetate (PA), which can be viewed as the substrate of hydroxylation by the $Fe^{IV}=O$ center, is bound to the metal center with its carboxylate group. To prevent the model from changing too much with respect to the crystal structure coordinates we fixed the coordinates of the terminal methyl groups of the methylimidazole rings, the ethylacetate and ethanol. We investigated a substrate hydroxylation mechanism for PA leading to (*R*)- and (*S*)-MA products using an $Fe^{IV}=O$ intermediate. The reactions are stepwise via a radical intermediate (Int) and initiated by a hydrogen atom abstraction barrier (TS1). Subsequent rebound of the hydroxy group via a rebound transition state (TS2) gives hydroxylated products (MA). We investigated hydrogen atom abstraction leading to a *pro*-(*R*) and *pro*-(*S*) type intermediate (Int_R, Int_S) and their corresponding rebound barriers to (*R*)- and (*S*)-MA products. Since nonheme $Fe^{IV}=O$ intermediates are known to have close lying spin states,[16] we calculated all mechanisms on the lowest lying singlet, triplet and quintet spin states as identified with a superscript next to the label.

We used methods that were benchmarked against experimental free energies of activation for oxygen atom transfer reactions by nonheme iron model complexes.[17]-[19] The hybrid density functional UB3LYP was used in combination with an LACVP basis set (with core potential) on iron and 6-31G on the rest of the atoms for geometry optimizations (basis set B1), frequencies and geometry scans.[20]-[23] The latter were full geometry optimizations with one degree of freedom, the reaction coordinate, fixed. To obtain better quality energetics, we did single point calculations with a triple-zeta type basis set: LACV3P+* with core potential on iron and 6-311+G* on the rest of the atoms (basis set B2). In addition, single point calculations using the polarized continuum model were performed with a dielectric constant of $\varepsilon = 4.7$ and a probe radius of 2.6 Å. To test the effect of dispersion energy on the relative energies we did single point calculations using the MO6 density functional method.[24] The latter, however, had very little effect on the relative energies. Free energies contain zero-point vibrational, entropic and thermal corrections to the energy.

D.2 Results:

In agreement with previous nonheme $Fe^{IV}=O$ calculations[16] the quintet spin state is the ground state. Its separation from the triplet is much larger than commonly found. It appears therefore that the reaction proceeds on a dominant quintet spin state surface only and none of the other spin states is low enough to contribute significantly. The *pro*-(*R*) hydrogen atom abstraction has the lowest barrier (20.9 kcal mol^{-1} in the gas phase; 16.5 kcal mol^{-1} with solvent corrections included and 20.97 kcal mol^{-1} free energy of activation in solvent), whereas the *pro*-(*S*) barrier is 19.9 kcal mol^{-1} in the gas phase, 18.8 kcal mol^{-1} with solvent corrections included and 20.23 kcal mol^{-1} free energy of activation in solvent. Consequently, the DFT calculations show that the product ratios clearly depend on the environment.

Obviously, minor stereochemical interactions in the substrate binding pocket that stabilize one isomer over the other will have dramatic effects on the product distributions. Energies of all reaction intermediates are given in Tables S2 and S3, the respective optimized UB3LYP/B1 geometries are given in Figures S3-S5.
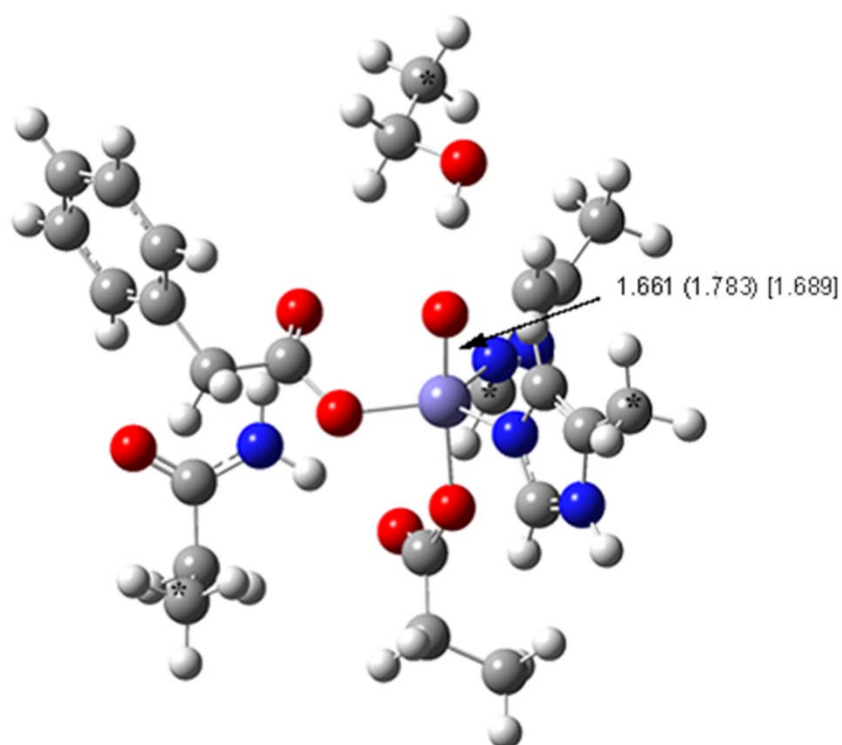
Figure S3. Optimized UB3LYP/B1 geometry of ^{5,3,1}*PA with bond lengths in angstroms.* Triplet spin data are presented in parenthesis and singlet spin data in square brackets.

Figure S4. Optimized UB3LYP/B1 geometry of (a) $^{5,3,1}TS1_R$ and (b) $^{5,3,1}TS2_S$ with bond lengths in angstroms. Triplet spin data are presented in parenthesis and singlet spin data in square brackets.

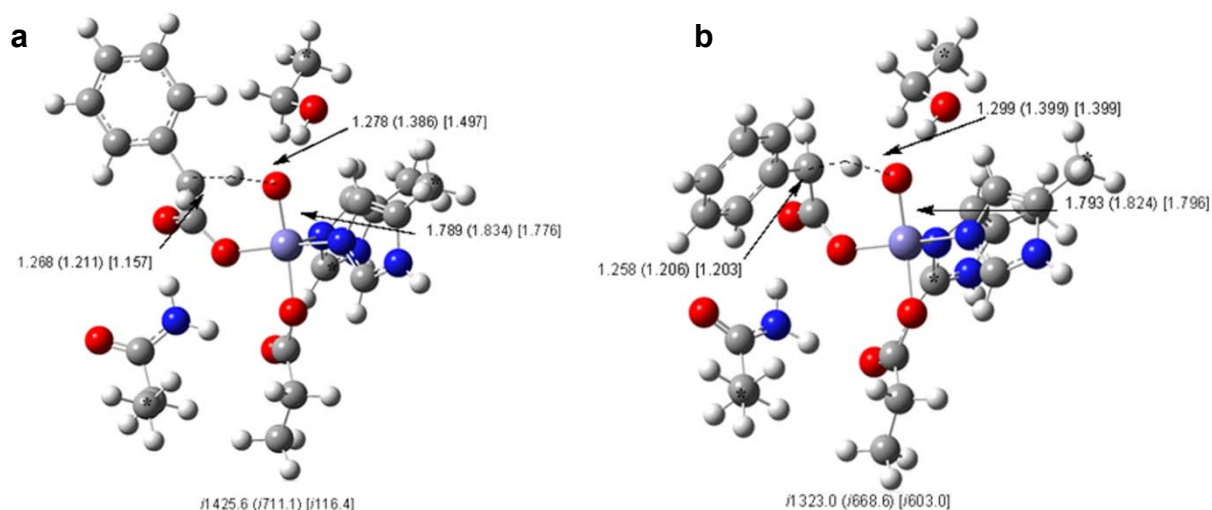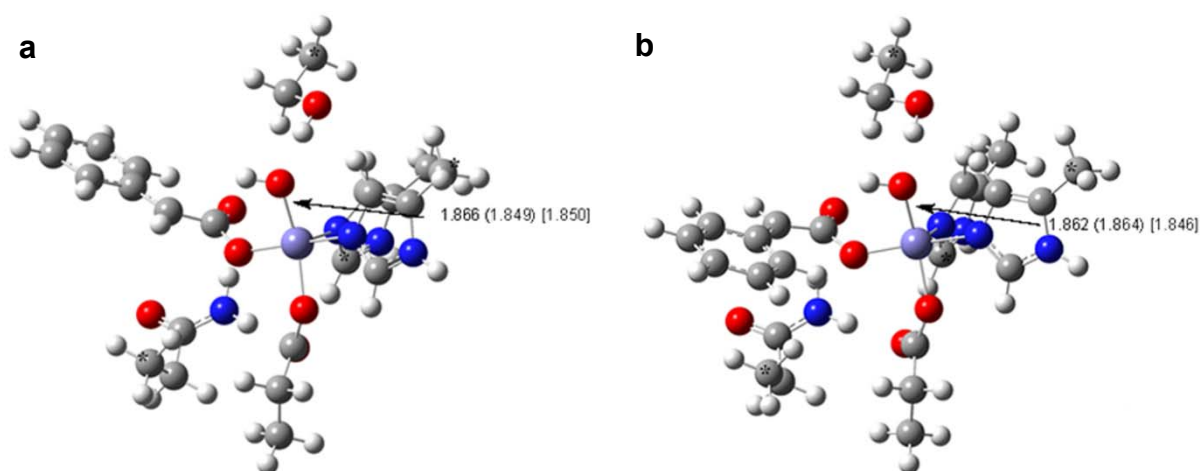*Figure S5. Optimized UB3LYP/B1 geometry of (a) $^{5,3,1}Int_R$ and (b) $^{5,3,1}Int_S$ with bond lengths in angstroms.* Triplet spin data are presented in parenthesis and singlet spin data in square brackets.

Table S2: Absolute energies (in au) of DFT calculated structures of PA hydroxylation by a nonheme Fe^{IV}=O model complex.

	E (B1, au)	ZPE	G (B1, au)	Esolv	E (B2, au)	E (M06, au)
^{1}PA	-1860.209117	0.610227	-1859.682290	-26.6168	-1861.145819	-1859.039926
^{3}PA	-1860.231685	0.609953	-1859.707924	-25.7427	-1861.168708	-1859.134146
^{5}PA	-1860.252190	0.610737	-1859.728642	-26.9741	-1861.195318	-1859.167912
R-form:						
^{1}TS1	-1860.201542	0.607568	-1859.682367	-31.1382	-1861.136175	-1859.081936
^{3}TS1	-1860.208731	0.604921	-1859.688030	-31.8295	-1861.145743	-1859.109196
^{5}TS1	-1860.216623	0.604093	-1859.692641	-31.3485	-1861.155363	-1859.124657
^{1}Int	-1860.244865	0.609725	-1859.718375	-29.8325	-1861.183217	-1859.122949
^{3}Int	-1860.256954	0.609037	-1859.735219	-26.8020	-1861.197147	-1859.156161
^{5}Int	-1860.257618	0.608710	-1859.734297	-29.9953	-1861.198415	-1859.163119
^{1}TS2	-1860.236524	0.609280	-1859.710865	-31.3207	-1861.174706	-1859.107210
^{3}TS2	-1860.233827	0.608549	-1859.709976	-26.3323	-1861.173990	-1859.140148
^{5}TS2	-1860.251104	0.608882	-1859.725872	-30.3899	-1861.192774	-1859.155073
^{1}MA	-1860.255922	0.611725	-1859.728042	-35.7180	-1861.200309	-1859.125782
^{3}MA	-1860.285883	0.612025	-1859.756153	-27.2513	-1861.221432	-1859.186726
^{5}MA	-1860.301432	0.610686	-1859.781447	-32.6302	-1861.246154	-1859.216180
S-form:						
^{1}TS1	-1860.200474	0.605874	-1859.678693	-30.8676	-1861.134949	-1859.075711
^{3}TS1	-1860.215979	0.605165	-1859.696840	-28.7075	-1861.153328	-1859.116342
^{5}TS1	-1860.220023	0.604531	-1859.700353	-28.0632	-1861.157462	-1859.126232
^{1}Int	-1860.246598	0.610539	-1859.722251	-28.5215	-1861.182449	-1859.126602
^{3}Int	-1860.254461	0.608498	-1859.729279	-29.6125	-1861.196052	-1859.160727
^{5}Int	-1860.256156	0.608420	-1859.731319	-29.4416	-1861.198631	-1859.162522
^{1}TS2	-1860.218913	0.609154	-1859.687477	-30.9179	-1861.157148	-1859.100500
^{3}TS2	-1860.231884	0.607788	-1859.706585	-30.5605	-1861.173070	-1859.138715
^{5}TS2	-1860.237749	0.607592	-1859.714655	-30.3186	-1861.178846	-1859.146513
^{1}MA	-1860.260087	0.611018	-1859.734497	-34.2311	-1861.205444	-1859.130890
^{3}MA	-1860.269313	0.611279	-1859.745116	-34.2866	-1861.215451	-1859.169027
^{5}MA	-1860.302640	0.610569	-1859.782731	-30.6793	-1861.250458	-1859.217872

Table S3: Relative energies (in kcal mol^{-1}) of DFT calculated structures of PA hydroxylation by a nonheme Fe^{IV}=O model complex.

	ΔE (B1)	ΔE+ZPE (B1)	ΔE (B2)	ΔE+ZPE (B2)	ΔE+ZPE+E$_s$ (B2)	ΔE+ZPE (M06)	ΔG (B1)	ΔG (B2)	ΔG+E$_{solv}$ (B2)
^{1}PA	27.03	26.71	31.06	30.74	31.10	79.99	29.09	33.12	33.48
^{3}PA	12.87	12.37	16.70	16.21	17.44	20.70	13.00	16.83	18.06
^{5}PA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R-form:									
^{1}TS1	31.78	29.79	37.11	35.12	30.96	51.96	29.04	34.37	30.20
^{3}TS1	27.27	23.62	31.11	27.46	22.60	33.20	25.48	29.32	24.47
^{5}TS1	22.32	18.15	25.07	20.90	16.53	22.97	22.59	25.34	20.97
^{1}Int	4.60	3.96	7.59	6.96	4.10	27.58	6.44	9.44	6.58
^{3}Int	-2.99	-4.06	-1.15	-2.21	-2.04	6.31	-4.13	-2.29	-2.11
^{5}Int	-3.41	-4.68	-1.94	-3.22	-6.24	1.74	-3.55	-2.09	-5.11
^{1}TS2	9.83	8.92	12.93	12.02	7.67	37.18	11.16	14.26	9.91
^{3}TS2	11.52	10.15	13.38	12.01	12.65	16.05	11.71	13.57	14.22
^{5}TS2	0.68	-0.48	1.60	0.43	-2.98	6.89	1.74	2.65	-0.76
^{1}MA	-2.34	-1.72	-3.13	-2.51	-11.26	27.06	0.38	-0.41	-9.16
^{3}MA	-21.14	-20.33	-16.39	-15.58	-15.86	-11.00	-17.26	-12.51	-12.78
^{5}MA	-30.90	-30.93	-31.90	-31.93	-37.59	-30.32	-33.14	-34.14	-39.79
S-form:									
^{1}TS1	32.45	29.40	37.88	34.83	30.94	54.81	31.34	36.77	32.88
^{3}TS1	22.72	19.23	26.35	22.85	21.12	28.86	19.96	23.58	21.85
^{5}TS1	20.19	16.29	23.76	19.86	18.77	22.26	17.75	21.32	20.23
^{1}Int	3.51	3.38	8.08	7.95	6.40	25.80	4.01	8.58	7.03
^{3}Int	-1.43	-2.83	-0.46	-1.87	-4.50	3.10	-0.40	0.56	-2.07
^{5}Int	-2.49	-3.94	-2.08	-3.53	-6.00	1.93	-1.68	-1.27	-3.74
^{1}TS2	20.88	19.89	23.95	22.96	19.01	41.31	25.83	28.90	24.96
^{3}TS2	12.74	10.89	13.96	12.11	8.52	16.47	13.84	15.06	11.47
^{5}TS2	9.06	7.09	10.34	8.36	5.02	11.45	8.78	10.05	6.71
^{1}MA	-4.96	-4.78	-6.35	-6.18	-13.43	23.41	-3.67	-5.07	-12.33
^{3}MA	-10.74	-10.40	-12.63	-12.29	-19.61	-0.36	-10.34	-12.23	-19.54
^{5}MA	-31.66	-31.76	-34.60	-34.71	-38.41	-31.46	-33.94	-36.88	-40.59

E. MD simulations

E.1 Methods.

All MD simulations were performed using the YASARA Structure suite version 12.11.20 (YASARA Biosciences).[25] A periodic simulation cell with dimensions of 81.25 Å, 66.67 Å, and 59.15 Å was employed with explicit solvent, using models of *Sc*HMS that had been built based on the crystal structure of *Ao*HMS (PDB 2R5V) and the *Sc*HMS triple mutant structure of this study (PDB 3ZGJ). The AMBER99[26] force field was used and long-range electrostatic potentials were calculated with the Particle Mesh Ewald (PME) method, with a cutoff of 7.864 Å.[27],[28] The ligand's force field parameters were generated with the AutoSMILES utility.[29] Iron parameters were defined and calibrated to reflect the trigonal bipyramidal geometry that is found in the crystal structures of *Ao*HMS[4] and *Sc*HMS from this study. Therefore, force field parameters were deduced, whereby equilibrium spring length for all ligands from DFT calculations were used (*vide supra*), and a spring stretching force constant of 2000 N m^{-1} was applied. Ligand-Fe-ligand angles of minimum energy were defined as 180° for the axial ligands - the Glu340 coordinating oxygen and the iron bound reactive oxygen species or the hydroxy group in the final product, 120° for the equatorial ligands Nε-His181, Nε-His261 and the coordinating oxygen from the respective reaction intermediate's or product's carboxylate residue, and 90° for all angles between one axial ligand, the metal center and one equatorial ligand, with all angle bending force constants set to of 1000 kJ mol^{-1} rad^2. For all MD simulations the metal center's trigonal bipyramidal geometry was thus constrained. For the MD simulations regarding transitions states of H-abstraction, TS1, and rebound hydroxylation, TS2, the transition state geometries of the atoms that partake in the reaction, namely the iron-oxo species, the reactive carbon center, and, regarding the

H-abstraction step, also the transferred hydrogen atom, were additionally fixed. The oxidation states of the iron cofactor were adjusted to +4 for the Fe^{IV}=O intermediate and to +3 for the transition states and the Fe^{III}-OH intermediate, respectively.

Structure pretreatment and simulation parameters: The hydrogen bonding network was optimized using the method of Hooft and coworkers,[30] and YASARAs' p*K*a values at pH 7.5 were assigned.[31] The simulation cell was filled with water at a density of 0.997 g mL^{-1} and neutralized with 0.05% NaCl. After relaxation of the solvent, the system was energy minimized using a steepest descent minimization to remove conformational stress, followed by a simulated annealing minimization until convergence (< 0.05 kJ mol^{-1} per 200 steps). Integration time steps were set to 1.33 and 4 fs for intra- and intermolecular forces, respectively. Temperature controlled MD simulations at 298 K were initiated, whereby integration time steps for intramolecular and intermolecular forces were set to 1.25 fs and 2.5 fs, respectively.
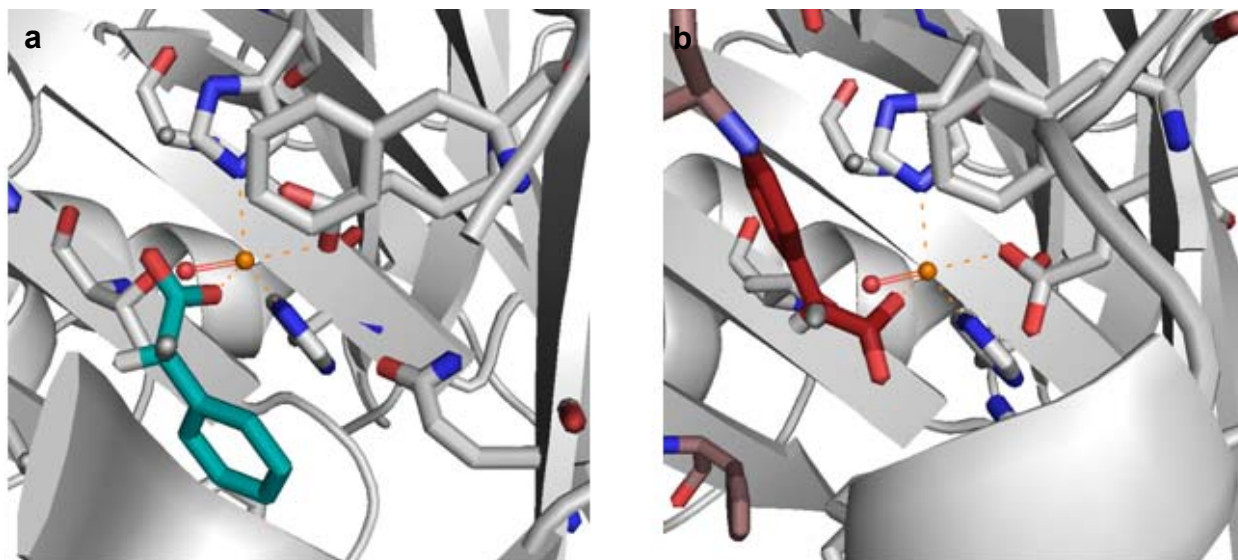
Figure S6. Fe^{IV}=O PA complexes in (a) a modeled ScHMS wild type structure with pro-(S) ligand conformation and (b) the ScHMS triple variant model with pro-(R) ligand conformation. Positioning of the ligand's phenyl ring is energetically enforced by the respective protein structure. Residues within 5 Å of the metal (orange) are displayed as gray sticks (except of Ile335, Leu358, and Ile236, which were omitted for clarity). Mutations in the ScHMS triple mutant are shown in brown.

Table S4: Average potential energies ΔG regarding the intermediate/product ligand and amino acid residues at positions 221, 223 and 359 in the (*R*)- versus (*S*)-specific complexes, as determined by MD simulations, are shown. Mean values describe the average values obtained during the 5 ns to 10 ns period, a time frame in which, generally, the systems had reached equilibrium. Standard deviations calculated from the mean energies of the respective five one-nanosecond segments of this 5-10 ns time frame are given.* Note that in our model systems the energy contributions of the ligands' respective stereo-complementary binding modes were generally overestimated by an order of magnitude. However, a system that gives realistically small contributions may complicate the identification of significant effects.

$\Delta G_{\text{wild type}}$ (kJ mol^{-1})								
	Ligand *pro(S)*	Ligand *pro(R)*	Ser221 *pro(S)*	Ser221 *pro(R)*	Val223 *pro(S)*	Val223 *pro(R)*	Tyr359 *pro(S)*	Tyr359 *pro(R)*
Fe^{IV}=O - center	-301±8	-296±5	-51±2	-58±3	-70±1	-70±1	-74±2	-72±3
TS1- H-abstraction	-101±8	-117±4	-53±6	-42±3	-73±2	-70±2	-68±2	-68±4
Fe^{III}-OH center	-122±2	-43±5	-50±2	-57±1	-64±1	-69±1	-71±2	-80±2
TS2-hydroxylation	-103±2	97±6	-49±1	-35±5	-61±2	-57±2	83±3	85±5
Fe^{II}-MA	202±12	225±5	-48±2	-43±7	-71±1	-69±2	-65±6	-77±3

$\Delta G_{\text{variant Ser221Met\_Val223Phe\_Tyr359Ala}}$ (kJ mol^{-1})								
	Ligand *pro(S)*	Ligand *pro(R)*	Met221 *pro(S)*	Met221 *pro(R)*	Phe223 *pro(S)*	Phe223 *pro(R)*	Ala359 *pro(S)*	Ala359 *pro(R)*
Fe^{IV}=O - center	n.d.**	293±8	n.d.**	-20±3	n.d.**	15±1	n.d.**	-6.5±1
TS1- H-abstraction	-68±13	-111±7	-17±3	-15±3	18±2	22±3	1.5±3.5	-0.5±1.5
Fe^{III}-OH center	-32±8	-35±3	-26±3	-24±4	17±3	22±2	3±1	-4±1
TS2-hydroxylation	-98±3	159±7	-15±3	-20±2	16±3	14±2	-3±2	-4±1
Fe^{II}-MA	233±3	247±3	13±2	19±4	-18±2	-20±3	-2±2	-9±2

* Note that when MD simulations for the wild type *pro*-(*S*) pathway were repeated using different random seed numbers, comparable mean energies of the 5-10 ns time frame were found and standard deviations were in a similar range as the standard deviations of the means of one-nanosecond segments from one simulation.

** In the triple mutant *pro*-(*S*) starting structure the PA ligand was repositioned after ~ 3 ns and adopted the *pro*-(*R*) conformation. Therefore, no energies are given for the *pro*-(*S*) conformation.
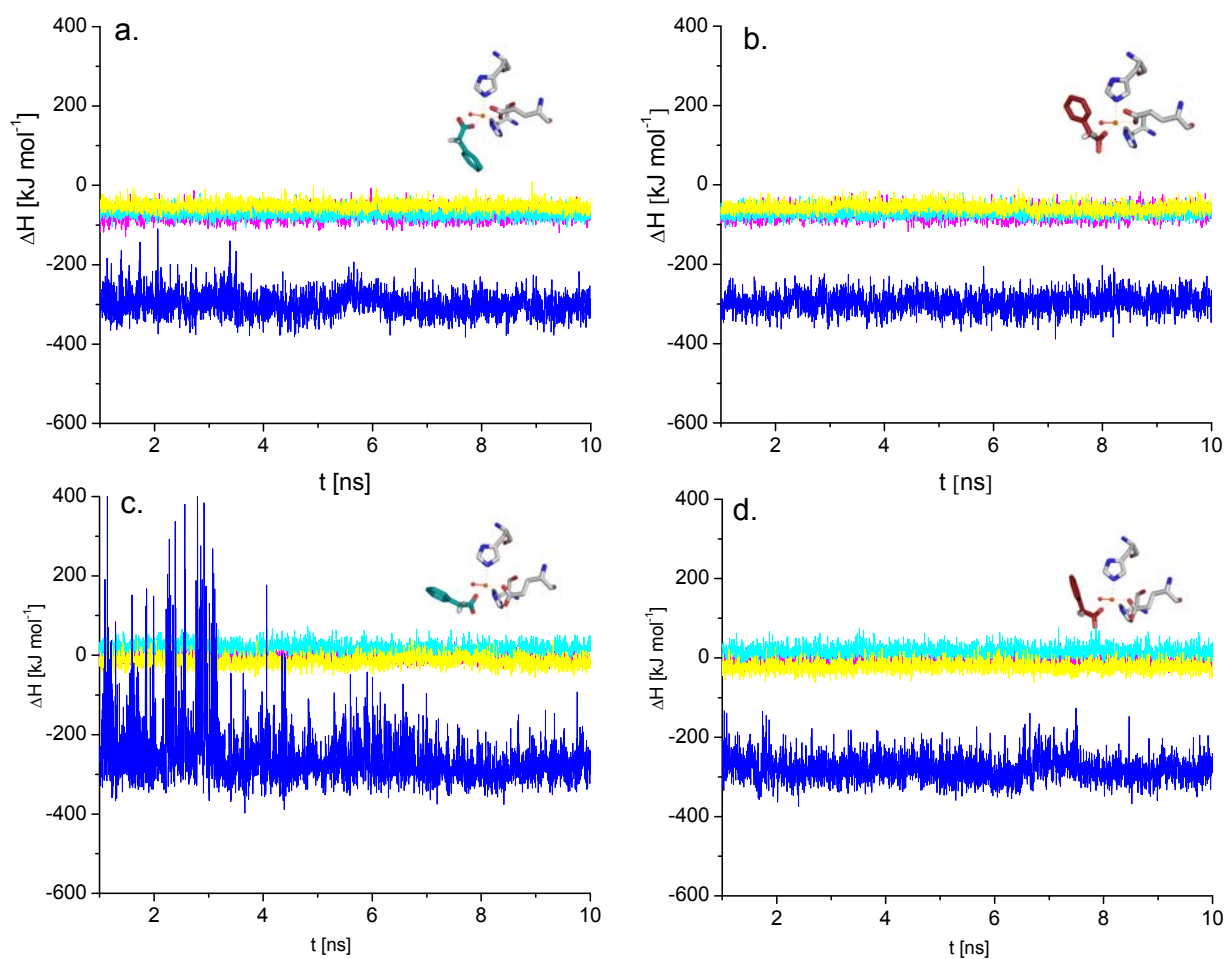
Figure S7. Potential energies of the intermediate PA ligand (blue) and the amino acid residues in positions 221, 223 and 359 (yellow, cyan, magenta), derived from MD-simulations of the reaction intermediates in complex with the respective protein structures. (a., b.) *Sc*HMS wild type and (c., d.) the *Sc*HMS triple variant Ser221Met_Val223Phe_Tyr359Ala structural models were equilibrated at 298 K, whereby the initial structures had the PA ligand positioned in optimal *pro*-(*S*)- (a., c.) and *pro*-(*R*)- (b., d.) H-abstraction geometry.
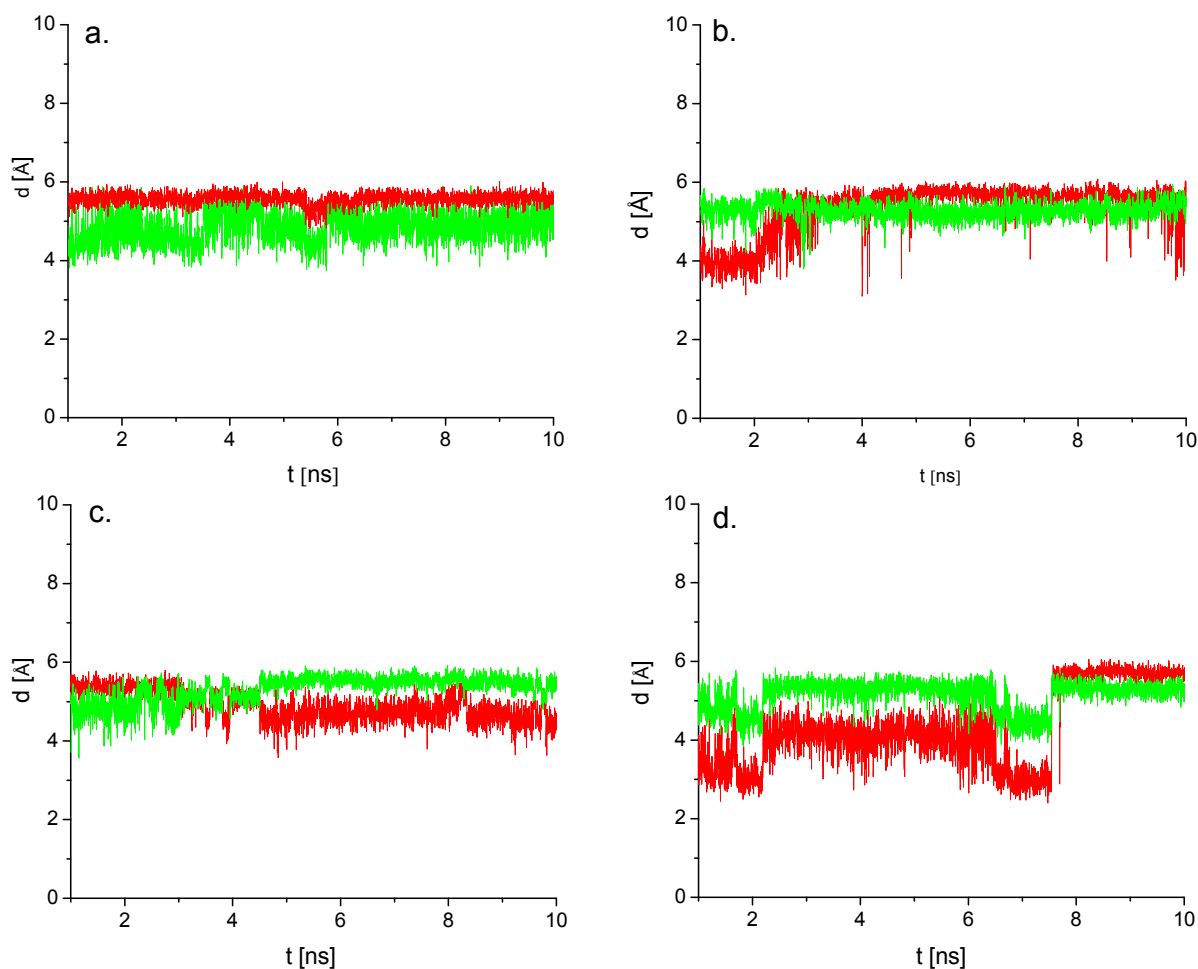
Figure S8. Distance of the Fe^{IV}=O oxygen and the target pro-(S)- (green) and pro-(R)- (red) hydrogen for abstraction from the PA intermediate species, derived from the respective MD-simulations of the reaction intermediate in complex with the respective protein structures (a., b.) ScHMS wild type and (c., d.) ScHMS triple variant Ser221Met_Val223Phe_Tyr359Ala as described in Figure S7. Results show that, due to ligand flexibility, even when the aromatic ring's respective in-pocket or out-of-pocket orientation is maintained, this does not correspond to a 'fixed' favorable positioning of either benzylic H-atom for abstraction (with the arguable exception of wild type enzyme in *pro*-(*S*)-abstraction geometry). However, each of the two principle aromatic ring positions will only allow the formation of either a *pro*-(*S*)- (in-pocket) or *pro*-(*R*)- (out of pocket) H-abstraction transition state at the trigonal bipyramidal iron center (see Figure S9).
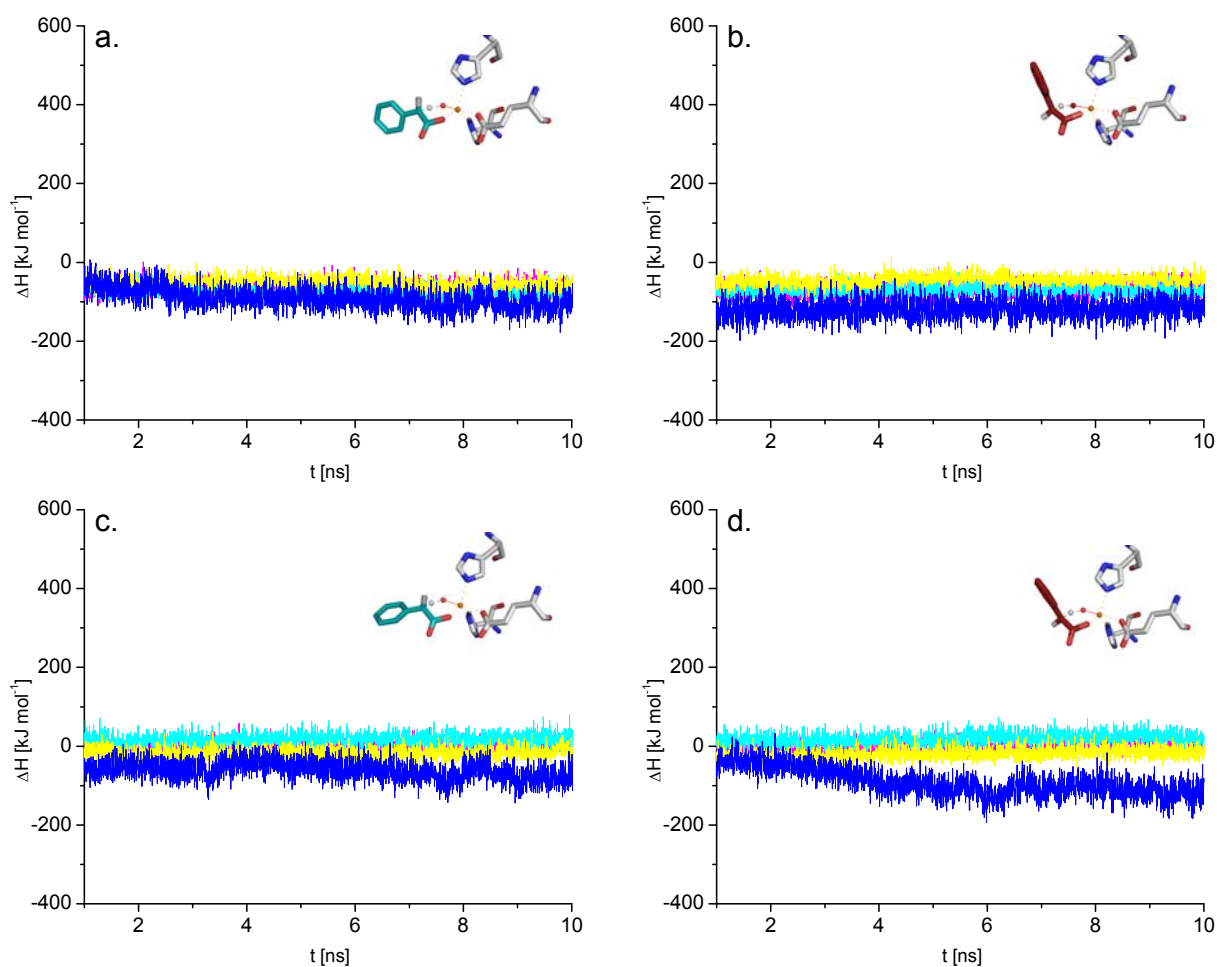
Figure S9. Potential energies of the ligand (blue) and the amino acid residues in positions 221, 223 and 359 (yellow, cyan, magenta) of the H-abstracting transition state, derived from MD-simulations of the reaction intermediates in complex with the respective protein structures. (a., b.) *Sc*HMS wild type and (c., d.) *Sc*HMS triple variant Ser221Met_Val223Phe_Tyr359Ala structural models were equilibrated at 298 K, whereby the initial ligand geometry reflected that of the *pro*-(*S*)- (a., c.) and *pro*-(*R*)- (b., d.) H-abstraction transition state derived from DFT-calculations. Potential energies of the fixed oxygen, hydrogen and carbon atoms that directly partake in catalysis are excluded from the calculated ligand energies.
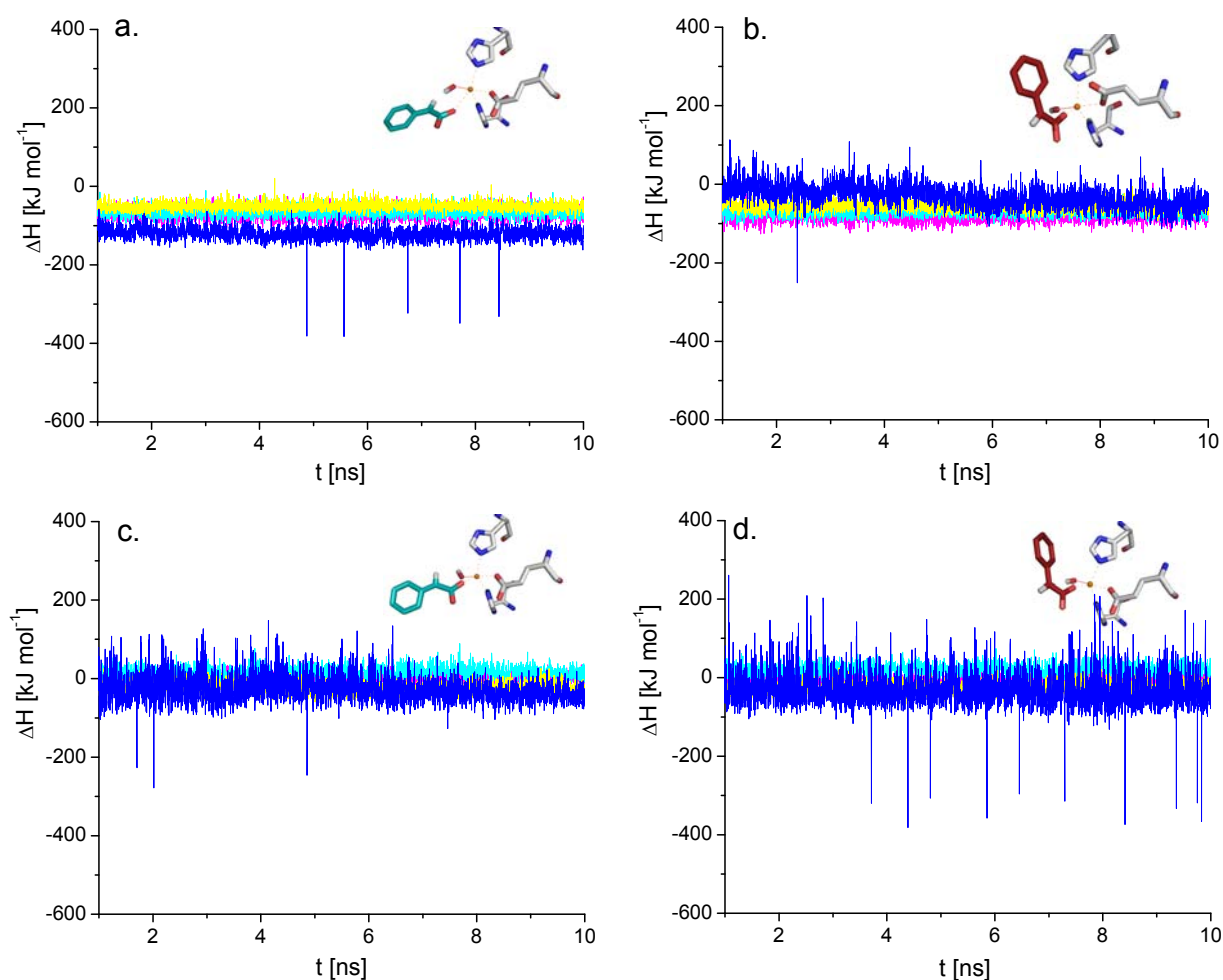
Figure S10. Potential energies of the intermediate PA radical ligand (blue) and the amino acid residues in positions 221, 223 and 359 (yellow, cyan, magenta), derived from MD-simulations of the reaction intermediates in complex with the respective protein structures. (a., b.) *Sc*HMS wild type and (c., d.) *Sc*HMS triple variant Ser221Met_Val223Phe_Tyr359Ala structural models were equilibrated at 298 K, whereby the initial structures had the intermediate ligand positioned in optimal *pro*-(*S*)- (a., c.) and *pro*-(*R*)- (b., d.) hydroxylation geometry.
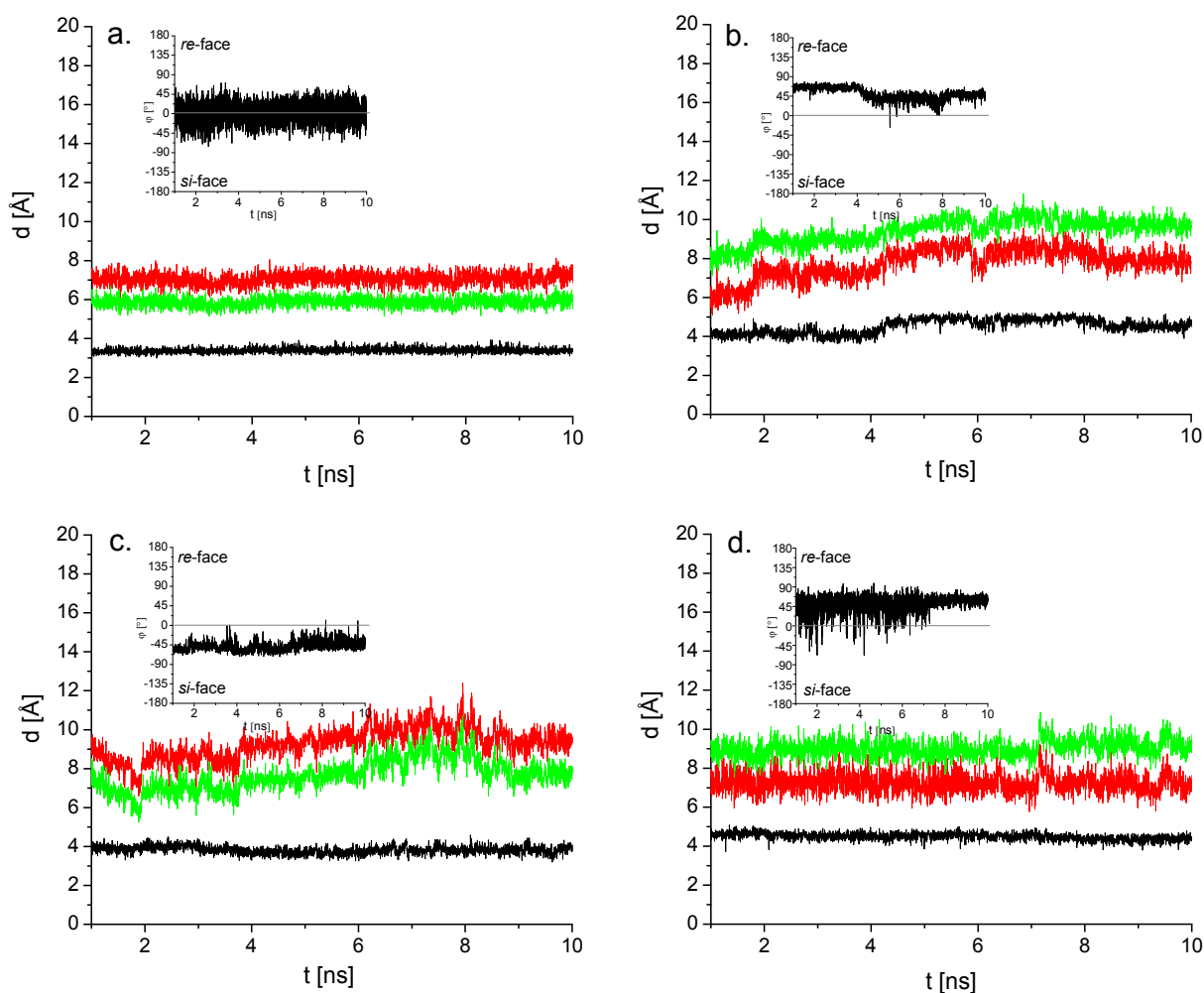
Figure S11. Distance of Fe^{III}-OH oxygen to the reactive sp^2 hybridized carbon center of the PA radical ligand (black) and distances of the sp^2 carbon center's substituents, namely the C1 carbon atom of the aromatic substituent C1$_{aromatic}$ (red) and the hydrogen atom H-α (green) to the C-β of residue Phe350 of the protein structure during MD simulations from Figure S10. C-β of Phe350 roughly coincides with the vector that is spanned by C1$_{aromatic}$ and H-α and the constant relative distances of these two ligand atoms to this Phe350 'anchor' indicate the absence of rotational movement around the axis defined by the ligand's carboxylate carbon atom and the sp^2 hybridized carbon center. Panels a. and b. show ScHMS wild type and panels c. and d. show the ScHMS triple variant Ser221Met_Val223Phe_Tyr359Ala, whereby the initial structures had the intermediate PA radical ligand positioned in 'ideal' *pro-(S)-* (a., c.) and *pro-(R)-* (b., d.) hydroxylation geometries. Insets show the dihedral angle that is spanned by the sp^2 hybridized carbon atom of the ligand, its carboxylate and hydrogen substituents and the oxo atom of the Fe^{IV}=O center, and it is a measure for the pro-chiral orientation of the ligand's face for oxygenation.
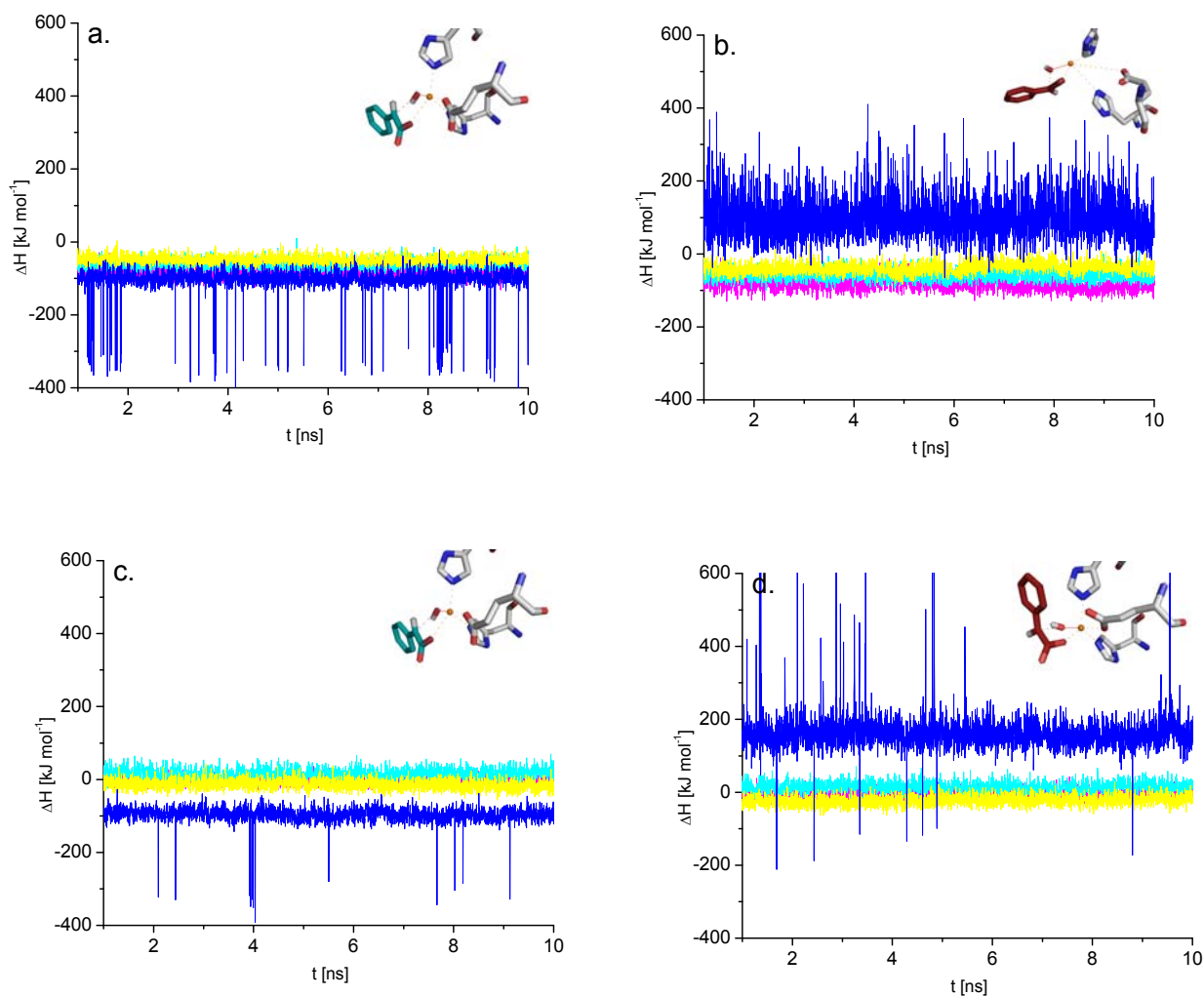
Figure S12. Potential energies of the reaction intermediate (blue) and of the amino acid residues in positions 221, 223 and 359 (yellow, cyan, magenta) during the transition state of rebound hydroxylation, derived from MD-simulations of the reaction intermediates in complex with the respective protein structures. (a., b.) *Sc*HMS wild type and (c., d.) *Sc*HMS triple variant Ser221Met_Val223Phe_Tyr359Ala structural models were equilibrated at 298K, whereby the initial structures had the intermediate ligand in optimal *pro*-(*S*)- (a., c.) and *pro*-(*R*)- (b., d.) rebound hydroxylation geometry.
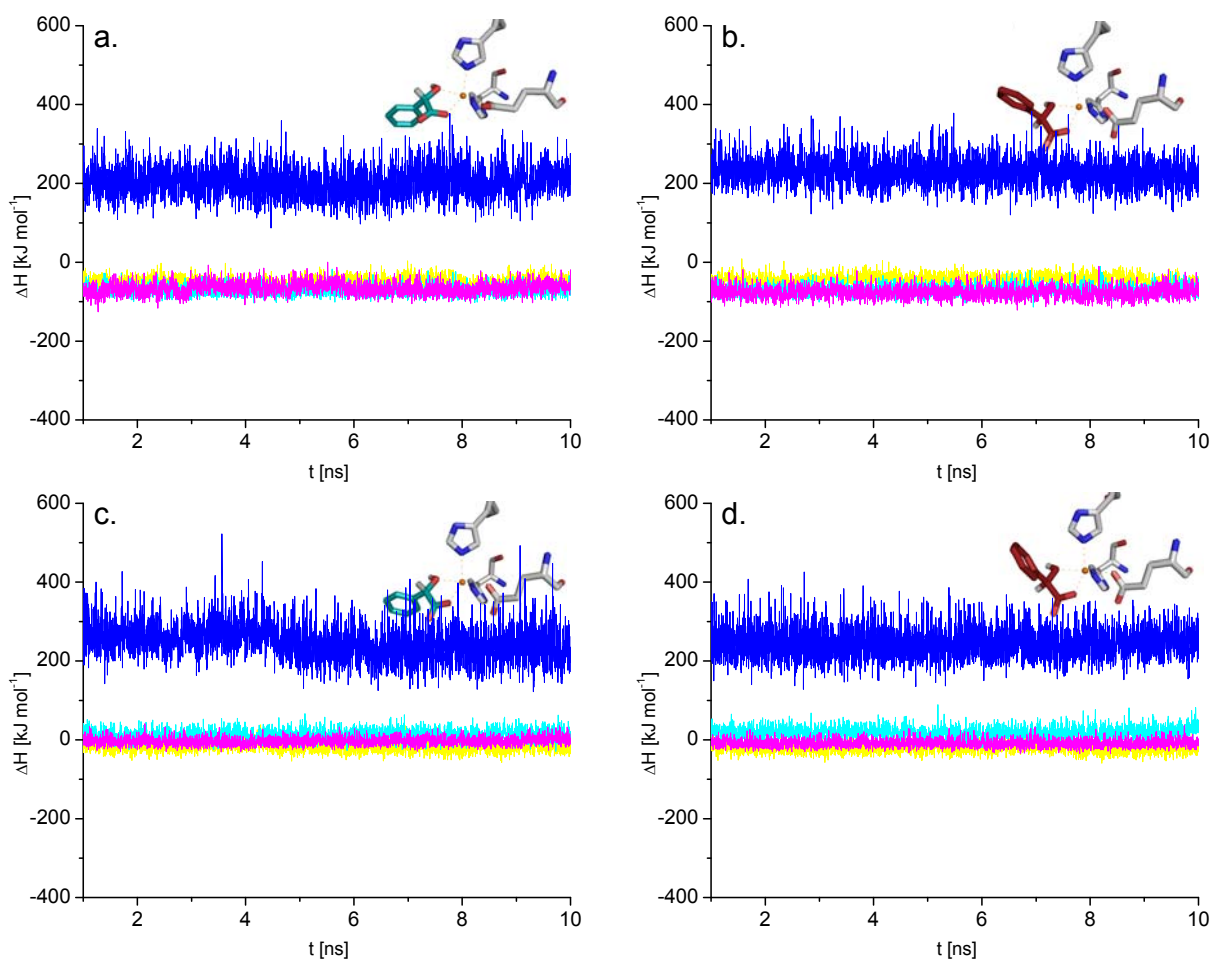
Figure S13. Potential energies of the product ligand (blue) and the amino acid residues in positions 221, 223 and 359 (yellow, cyan, magenta), derived from MD-simulations of the enzyme product complexes. (a., b.) *Sc*HMS wild type and (c., d.) *Sc*HMS triple variant Ser221Met_Val223Phe_Tyr359Ala structural models were equilibrated at 298 K, whereby the initial structures had the ligand coordination geometry that is also found in the product ligated *Ao*HMS wild-type and *Sc*HMS triple variant crystal structures, which have the respective (*S*)- (a., c.) and (*R*)-enantiomer (b., d.) bound.

F. References

[1]   K. Arnold, L. Bordoli, J. Kopp, T. Schwede, *Bioinformatics (Oxford, England)* **2006,** 22, 195-201.

[2]   F. Kiefer, K. Arnold, M. Kunzli, L. Bordoli, T. Schwede *Nucleic acids Res.* **2009,** 37, D387-392.

[3]   R. A. Laskowski, M. W Macarthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **1993,** 26, 283–291.

[4]   J. Brownlee, P. He, G. R. Moran, D. H. Harrison, *Biochemistry* **2008,** 47, 2002-2013.

[5]   W. F. van Gunsteren *et al.* Vdf Hochschulverlag AG an der ETH Zürich, Zürich, **1996**.

[6]   S. C. Lovell, J. M. Word, J. S. Richardson, D. C. Richardson, *Proteins* **2000,** 40, 389-408.

[7]   W. Wang, B. A. Malcolm, *Biotechniques* **1999,** 26, 680-682.

[8]   G. D. Straganz, A. Slavica, H. Hofer, U. Mandl, W. Steiner, B. Nidetzky, *Biocat. Biotrans.* **2005,** 23, 261-269.

[9]   A. M. Brzozowski, J. Walton, *J. Appl. Cryst.* **2001,** 34, 97-101.

[10]   A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L.C. Storoni, R.J. Read. *J. Appl. Cryst.* **2007,** 40, 658-674.

[11]   P. Emsley, K. Cowtan, *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2004,** 60, 2126-2132.

[12]   G. N. Murshudov, A. A. Vagin, E. J.Dodson, *Acta Crystallogr., Sect D: Biol. Crystallogr.* **1997,** 53, 240-255.

[13]   S. P. de Visser, *J. Am. Chem. Soc.* **2010,** 132, 1087–1097.

[14]   D. Kumar, W. Thiel, S. P.de Visser, J. *Am. Chem. Soc.* **2011,** 133, 3869–3882.

[15]   M. J. Frisch, *et al.* Gaussian, Inc., Wallingford CT, **2010**.

[16]   R. Latifi, M. Bagherzadeh, S. P. de Visser, *Chem. Eur. J.* **2009,** 15, 6651–6662.

[17] D. Kumar, S. P. de Visser, S. Shaik, *Chem. Eur. J.* **2005,** 11, 2825–2835.

[18] S. P. de Visser, K. Oh, A.-R. Han, W. Nam, *Inorg. Chem.* **2007,** 46, 4632–4641.

[19] A. K. Vardhaman, C. V. Sastri, D. Kumar, S. P. de Visser, *Chem. Commun.* **2011,** 47, 11044–11046.

[20] A. D. Becke, *J. Chem. Phys.* **1993,** 98, 5648–5652.

[21] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988,** 37, 785–789.

[22] P. J. Hay, W. R. Wadt, *J. Chem. Phys.* **1985,** 82, 270–283.

[23] W. J. Hehre, R. Ditchfield, J. Pople, *A. J. Chem. Phys.* **1972,** 56, 2257–2261.

[24] Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **2007,** 120, 215–241.

[25] YASARA (9.11.9 9.11.9) *Vienna, Austria: YASARA Biosciences GmbH;* **2012**.

[26] J. Wang, P. Cieplak, P. A. Kollman, *J. Comput. Chem.* **2000,** 21, 1049-1074.

[27] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995,** 117, 5179-5197.

[28] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys. B* **1995,** 103, 8577-8593.

[29] A. Jakalian, D. B. Jack, C. Bayly, *J. Comput. Chem.* **2002,** 23, 1623-1641.

[30] R. W. Hooft, G. Vriend, C. Sander, E. E. Abola, *Nature* **1996,** 381, 272.

[31] E. Kriegler, J. E. Nielsen, C. A. Spronk, G. Vriend, *J. Mol. Graph. Model.* **2006,** 25, 481-486.

[32] C. M. L. Di Giuro, C. Konstantinovics, U. Rinner, C. Nowikow, E. Leitner, Grit. D. Straganz, *PLOS ONE* **2013,** accepted

Curriculum vitae

Personal data:

Name: Sarah Pratter

Address: Muchargasse 12, 8010 Graz

Date of birth: 10. 02. 1985

Place of birth: Wagna, Austria

E-Mail: sarah.pratter@gmx.at

Education:

09/1995 – 06/2003   BG/BRG Leibnitz

10/2003 – 11/2005   Diploma program (first part) in chemistry,
                     University of Graz, Austria

11/2005 – 06/2009   Diploma program (second part) in biochemistry and
                     molecular biology, University of Graz, Austria
                     Diploma thesis at the Institute of Biochemistry

08/2009 – 07/2013   PhD program DK Molecular Enzymology at the Institute of
                     Biotechnology and Biochemical Engineering,
                     Graz University of Technology, Austria

Professional experience:

07/2006 – 08/2006:   Austrian Agency for Health and Food Safety (AGES) Graz,
                     Austria

06/2007 – 03/2008:   Ingenieurgemeinschaft KZR (water analytics), Seiersberg,
                     Austria

Research stays:

03/2012 – 05/2012   Laboratory of Prof. E. Solomon, Stanford University, USA

09/2012 – 12/2012   Laboratory of Dr. G. Grogan, University of York, England

Teaching activities:

Winter term 2008/09   Tutor in the lab course "Basic Chemistry", University of Graz

Summer term 2009   Tutor in the lab course "Biochemical Practice", University of Graz

Summer term 2011   Tutor in the lab course "Biotechnology", Graz University of Technology

Poster and oral presentations:

06/2010   Research 2010, Graz

04/2010   Graduate Seminar DK Molecular Enzymology Loipersdorf, Austria

12/2010   Pacifichem 2010 Honolulu, Hawaii, USA

02/2011   5th Doc Day NAWI Graz Doctoral Schools Molecular Biology and Biochemistry & Molecular Biosciences and Biotechnology, Graz

03/2011   Graduate Seminar DK Molecular Enzymology, Graz

07/2012   8th Doc Day NAWI Graz Doctoral Schools Molecular Biology and Biochemistry & Molecular Biosciences and Biotechnology, Graz

12/2012   "Enzymes and Signals" joint research summit with the DK Molecular Mechanisms in Cell Signaling (Vienna), Graz

Patents:

Straganz, G.D., Konstantinovic C.C., Pratter, S.M., DiGiuro C.M.L. *R-Mandelate Synthase*. Vol. EP 11183652.4 2.10.2011.