

**Structural Characterization of
Flavin-dependent Monooxygenases
from *Zonocerus variegatus*
and the Human
Mitochondrial Amidoxime Reducing Component
(mARC)
– Enzymes involved in Biotransformation**

Dissertation

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften

(Dr. rer. nat.)

der Mathematisch-Naturwissenschaftlichen Fakultät

der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Christian Kubitza

Kiel 2018

Dekanin: Prof. Dr. Natascha Oppelt

Erster Gutachter: Prof. Dr. Axel Scheidig

Zweiter Gutachter: Prof. Dr. Bernd Clement

Tag der mündlichen Prüfung: 20.07.2018

Zum Druck genehmigt: 20.07.2018

*“We cannot change the cards we are dealt,
just how we play the hand.”*

– Randy Pausch

Abstract

Organisms from all kingdoms of life are permanently exposed to substances that are foreign and oftentimes even harmful for their bodies. These compounds are called xenobiotics and are inhaled through the air, ingested with food or water or deliberately administered in the form of therapeutic drugs. But also the accumulation of toxic byproducts of their own metabolism, so-called endobiotics, may be dangerous. Both, exo- and endobiotics have to undergo metabolic processes in order to be eliminated from the organism, which is achieved by enzymatically converting them into more hydrophilic substances. This process is called biotransformation, which is dependent on a viable homeostasis sustained by a variety of enzymes. Therefore, it is of significant importance to increase our knowledge about enzymes involved in biotransformation processes and related detoxification pathways. This especially includes knowledge about these enzymes' three-dimensional structures.

This cumulative thesis focuses on the crystallization and structure determination of two opposing representatives of biotransformation enzymes, the flavin-dependent monooxygenase from the grasshopper *Zonocerus variegatus* (ZvFMO) and the human mitochondrial amidoxime reducing component (mARC). FMOs are able to oxygenate heteroatom-containing substances and thereby produce hydroxylated compounds in most cases, whereas mARC is part of a three-component enzymatic system that is very effectively reducing a broad range of heteroatom-hydroxylated endo- and xenobiotics. Both enzymes belong to the class of oxidoreductases but represent counteracting enzyme families.

Different high-resolution ZvFMO crystal structures provided detailed insight into the coordination of the bound cofactors and related conformational rearrangements of the enzyme. Furthermore, a hitherto unknown dimeric arrangement of subunits was observed for the FMO isoforms found in *Zonocerus variegatus*. To further investigate significant differences in the catalytic activity of ZvFMOs, a variety of ZvFMOa variants with amino acid exchanges near the active site were generated. Thereby, significant impacts on enzyme activity were observed even for single exchanges within the substrate entry site.

In order to crystallize human mARC, a special fusion protein strategy was successfully applied. The elucidated mARC crystal structure was the first MOSC protein that clearly showed the coordination of its molybdenum cofactor and exhibited a topology, which contradicts *in silico* predictions of structural domains currently used for online databases. The large substrate spectrum could be correlated to a surface-exposed active site, which is restricted by only a few amino acid residues. Parologue-specific amino acid residues were identified, which allow for the secure discrimination between mARC1 and mARC2 enzymes. Moreover, structural features indicate that mARC enzymes represent an evolutionary link between the two molybdoenzyme superfamilies of sulfite oxidases and xanthine oxidases.

Kurzzusammenfassung

Organismen aller Lebensreiche sind ständig verschiedensten körperfremden und oft auch schädlichen Substanzen ausgesetzt. Diese Verbindungen werden Xenobiotika genannt und werden über die Atemluft inhaliert, über Nahrung oder Wasser aufgenommen oder willentlich in der Form von Medikamenten eingenommen. Doch auch die Anreicherung von toxischen Nebenprodukten des eigenen Metabolismus, sogenannten Endobiotika, kann gefährlich sein. Sowohl Xeno- als auch Endobiotika müssen metabolische Prozesse durchlaufen, um vom Organismus ausgeschieden werden zu können. Erreicht wird dies durch enzymatische Umwandlung in hydrophilere Substanzen. Dieser Prozess wird als Biotransformation bezeichnet, welcher auf einem metabolischen Gleichgewicht beruht, das von verschiedenen Enzymen aufrechterhalten wird. Es ist daher wichtig, unser Wissen über Biotransformationsenzyme und damit verbundene Detoxifizierungsmöglichkeiten zu vertiefen. Dies gilt insbesondere auch für Kenntnisse derer dreidimensionalen Strukturen.

Diese kumulative Dissertation fokussiert sich auf die Kristallisation und Strukturaufklärung zweier gegensätzlicher Vertreter von Biotransformationsenzymen, der flavinabhängigen Monooxygenase aus der Heuschrecke *Zonocerus variegatus* (ZvFMO) und der menschlichen mitochondrialen Amidoxim-reduzierenden Komponente (mARC). FMOs katalysieren meist die Hydroxylierung von Heteroatom-haltigen Verbindungen, während mARC-Enzyme eine große Effizienz bei der Reduktion eines breiten Spektrums an Heteroatom-hydroxylierten Xeno- und Endobiotika aufweist. Beide Enzyme gehören zu den Oxidoreduktasen, repräsentieren aber kontrahierende Enzymfamilien.

Verschiedene hochaufgelöste ZvFMO-Strukturen lieferten Erkenntnisse in die Koordination gebundener Kofaktoren und damit einhergehende Konformationsänderungen. Weiterhin konnte eine bisher unbekannte Dimerformation dieses Enzyms beobachtet werden. Um tiefere Einblicke in die katalytische Aktivität von ZvFMOs zu erhalten, wurden verschiedene Enzymvarianten untersucht. Dabei wurde festgestellt, dass vor allem im Substrat-eingangsbereich schon Einzelaustausche von Aminosäuren einen signifikanten Einfluss auf die Enzymaktivität haben.

Die mittels einer speziellen Fusionsprotein-Strategie gelöste mARC-Kristallstruktur war die erste MOSC-Proteinstruktur, die Einsicht in die Koordination des Molybdän-Kofaktors gab und eine Topologie aufwies, die bisherigen *in silico* -Vorhersagen widersprach. Das breite Substratspektrum dieses Enzyms konnte dem oberflächenexponierten aktiven Zentrum zugeschrieben werden, welches nur durch wenige Aminosäuren begrenzt wird. Weiterhin wurden Paralog-spezifische Seitenketten identifiziert, die eine sichere Unterscheidung zwischen mARC1 und mARC2-Proteinen erlauben. Zudem weisen einige strukturelle Merkmale darauf hin, dass mARC-Enzyme ein evolutionäres Bindeglied zwischen zwei Molybdo-Enzymfamilien darstellen, den Sulfitoxidasen und den Xanthinoxidasen.

Table of contents

1	Introduction.....	1
1.1	Structure determination by X-ray crystallography.....	1
1.1.1	Protein crystallization.....	2
1.1.2	Data collection, phase calculation and model building.....	3
1.2	Biotransformation.....	5
1.2.1	Implications for drug metabolism and development.....	6
1.3	Cytochrome P450 monooxygenases.....	7
1.4	Flavin-dependent monooxygenases.....	8
1.4.1	Pyrrolizidine alkaloid <i>N</i> -oxygenases.....	10
1.5	Molybdenum-dependent enzymes.....	11
1.5.1	Nitrate reductase (NR).....	12
1.5.2	Sulfite oxidase (SO).....	12
1.5.3	Xanthine oxidoreductase (XOR).....	13
1.5.4	Aldehyde oxidase (AO).....	13
1.5.5	Mitochondrial amidoxime-reducing component (mARC).....	14
1.6	Molybdoenzyme superfamilies.....	16
2	Scope.....	17
3	Results and Discussion.....	19
3.1	Crystal structure of pyrrolizidine alkaloid <i>N</i> -oxygenase from the grasshopper <i>Zonocerus variegatus</i>	20
3.2	Human mitochondrial amidoxime reducing component (mARC): An electrochemical method for identifying new substrates and inhibitors.....	32
3.3	The Involvement of the Mitochondrial Amidoxime Reducing Component (mARC) in the Reductive Metabolism of Hydroxamic Acids.....	46
3.4	T4 Lysozyme-facilitated crystallization of the human molybdenum cofactor-dependent enzyme mARC.....	73
3.5	First crystal structure of human mARC1 reveals its exceptional position among eukaryotic molybdenum enzymes.....	83
4	Summary.....	113
4.1	Structural characterization of ZvFMOs.....	113
4.2	Crystallization and structural characterization of mARC.....	115
5	Concluding Remarks and Prospects.....	118
6	References.....	120
7	Appendix.....	128
7.1	List of Abbreviations.....	128
7.2	List of Figures.....	129
	Curriculum Vitae.....	130
	Erklärung.....	131
	Danksagung.....	133

1 Introduction

1.1 Structure determination by X-ray crystallography

The three-dimensional structure of proteins dictates their properties. Therefore, structure determination is a crucial step towards fully understanding the physiological functions, dynamics and interactions of any protein of interest. For example, crystal structures of proteins in complex with their specific ligands or substrates can provide knowledge about an enzyme's active site, substrate-binding site, cofactor coordination and, ultimately, its reaction cycle. In addition to a better understanding of a protein's general properties, protein structures give insight into certain diseases, which are correlated to single nucleotide polymorphisms (SNPs). Furthermore, rational drug design, which aims for generating highly-specific inhibitors only targeting one enzyme, is strictly dependent on the exact knowledge of the active site and substrate-binding site geometry.

Structure determination of proteins can be achieved by different methods, which have been established and significantly improved over the past decades. In the case of small proteins of interest (with an average molecular weight of 20 to 35 kDa), nuclear magnetic resonance spectroscopy (NMR) is the most promising method¹. Structure determination of large biomacromolecules and protein complexes may be achieved by electron microscopy (EM)². Most three-dimensional protein structures deposited in the protein data bank (PDB) were determined by the so far most powerful method, X-ray crystallography, which can provide details at atomic resolution. However, this approach includes a non-trivial and often longsome process composed of various crucial steps from protein expression to final model refinement. The most important and likewise most demanding step is crystallizing the target protein. Crystals increase the amount of X-rays being diffracted upon interaction with electrons surrounding the nuclei of each atom within the protein. Thereby, diffraction patterns can be obtained, which are reciprocal to the periodic electron density distribution within the crystal and can be used to determine the three-dimensional structure of the crystallized protein molecules or complexes at high resolution.

1.1.1 Protein crystallization

With the exception of the emerging technology of X-ray free electron lasers (XFEL),³ the proportion of X-ray beams diffracted by a single protein molecule is too insignificant in order to detect them. Within a protein crystal, however, molecules are periodically arranged in all three dimensions of space, which amplifies the diffraction power and highly increases the chances for X-ray scattering so that specific diffraction patterns can be obtained. Therefore, protein crystallization is a keystone in X-ray crystallography.⁴

However, macromolecular crystallization is non-predictable and depends on various trial-and-error approaches as well as different crystallization methods in order to find suitable conditions for individual proteins to crystallize. This process may be troublesome because of the high flexibility and large size of protein molecules as well as the complex nature of intermolecular interactions. Furthermore, the target protein solution has to be exceptionally pure and must endure high concentrations. The supersaturation of a protein in solution will lead to the formation of amorphous precipitate in most cases. But if suitable crystallization conditions are found, nucleation may occur within a thermodynamically metastable state by periodic self-assembly of protein molecules. Once these small crystallites are formed, they will grow into larger crystals until an equilibrium between soluble protein within the crystallization solution and solid protein in the form of crystals is reached.

Protein crystallization is dependent on temperature, pH, chaotropic or kosmotropic salts, precipitating agents that compete for water molecules and, among others, additives in the form of cofactors, substrates, inhibitors or just small organic compounds that reduce the protein's flexibility or facilitate crystal contact formation. Usually, initial crystallization conditions have to be optimized, altering the individual variables and concentrations of precipitants in order to obtain crystals of better quality.⁵ This is achieved by successive crystallization trials using grid screens, which means that one or two parameters are slightly altered simultaneously. The condition that yields crystals of higher quality is then refined in further iterative steps, which makes the whole process quite time-consuming, depending on the crystals' growth rate.^{4,6} However, once diffraction-quality single crystals are obtained, they can be used for structure determination.

1.1.2 Data collection, phase calculation and model building

Nowadays, diffraction data are usually collected at synchrotron facilities. These are cyclic electron accelerators able to produce highly-brilliant X-rays of tunable wavelength (0.6 - 2.7 Å). This enables efficient and rapid data acquisition, because the X-ray beam only targets a very small area of the protein crystal while exposing it to a very high number of photons per second. Furthermore, the availability of highly sensitive detectors at synchrotron facilities allows for an immediate read-out of collected data so that complete datasets can be obtained within seconds. This is achieved by exposing the crystals from different angles and collecting the resulting diffraction patterns.

These diffraction patterns of a protein are a reciprocal representation of its electron density distribution. With every atom of the protein molecule contributing to the intensity of each reflection within the diffraction pattern, the resulting electron density distribution is being accounted for by every detected reflection. The position of individual atoms within the target protein can be modelled into these three-dimensional electron density maps, which are calculated by applying an inverse Fourier transformation to the collected diffraction data. The following equation describes the reciprocal correlation between electron density and diffraction data:⁷

$$\rho_{xyz} = \frac{1}{V} \sum_{hkl} |F_{hkl}| \cdot \exp^{-2\pi i(hx+ky+lz)} \cdot \exp^{i\alpha(hkl)}$$

This equation represents the electron density ρ_{xyz} as a Fourier sum containing the structure factor amplitudes $|F_{hkl}|$, the Miller indices h , k and l as well as the unknown phases α_{hkl} of a series of three-dimensional waves. V is the volume of the unit cell and $|F_{hkl}|$ can be calculated from the measured intensities belonging to the respective reflection hkl through the following equation:

$$|F_{hkl}| \sim \sqrt{I}$$

Since the phase angles α_{hkl} of each reflection, which are crucial for structure determination, cannot directly be obtained from the collected data, these have to be determined using additional approaches. This 'phase problem' can be solved by experimental phasing methods using *isomorphous replacement* or *anomalous dispersion*. Furthermore, phases can be calculated from known crystal structures that are homologous to the target protein and share the same overall fold. This technique is called *molecular replacement* and is the simplest among them, but it is prone to model bias and has to be applied with caution. However, all of these methods only provide initial estimates of the real phases, which have to

be improved iteratively in order to obtain well-interpretable electron density maps.⁸ This can also be achieved by combining different phasing methods.

Once initial phases are obtained, electron density maps can be calculated from diffraction data by an inverse Fourier transformation of the experimental structure factor amplitudes $|F_{obs}|$ and their phases, which results in a basic electron density map (F_{obs}, α_{calc}) . This electron density distribution, which gives a first approximation of the true structure, may then be interpreted with atomic coordinates of the target protein. An additional electron density map, the difference map $(F_{obs} - F_{calc}, \alpha_{calc})$ that is obtained using differences between observed and calculated structure factor amplitudes, displays divergences between true and currently modelled structures. It exhibits positive map contours for parts existing in the structure, but not included in the model, and negative map contours for parts wrongly introduced into the model. Usually, $(2F_{obs} - F_{calc}, \alpha_{calc})$ electron density maps are used as a superposition of both maps described above in order to combine their features. In an iterative approach, the model is refined based on the calculated electron-density maps. Improved phases are calculated from the better-quality model and used to calculate more precise electron density maps, which can be interpreted more easily and help to refine the model again.⁹ After consecutive steps of optimization, the final protein structure is evaluated by various quality criteria and can finally be interpreted based on the features it displays.

1.2 Biotransformation

Every living organism is permanently exposed to a multitude of xenobiotics, compounds that are foreign and in many cases harmful to their body. These may be ingested with food or water, inhaled through the air or purposely administered in the form of therapeutic drugs. Likewise, every species has to deal with toxic byproducts of their own metabolism, so-called endobiotics, which are often lipophilic and tend to accumulate within the tissue. Both, exo- and endobiotics have to undergo metabolic processes in order to be eliminated from the organism, which is achieved by enzymatically converting them into more hydrophilic substances that can easily be excreted. These processes are termed biotransformation and are mainly located within the liver, but also occur in kidneys, intestine, spleen, lungs, brain, muscles, skin or blood.¹⁰

There are different phases of biotransformation. Phase I reactions include oxidative, reductive and hydrolytic reactions and are termed functionalization reactions. Thus, they are used to introduce or unmask a polar functional group (-OH, -COOH or -NH₂), so that the target molecule's hydrophilicity is increased and excretion is favored. Phase I biotransformation enzymes catalyzing oxidation reactions include the large family of microsomal cytochrome P450 monooxygenases (CYP), flavin-dependent monooxygenases (FMO), alcohol dehydrogenases and aldehyde dehydrogenases. Reduction reactions are mediated by reductive dehalogenases, nitro-reductases, azo-reductases and the recently discovered mitochondrial amidoxime-reducing component (mARC) enzyme system. Epoxide hydrolases, esterases and amidases catalyze hydrolysis of xeno- and endobiotics. Products of phase I reactions are suitable substrates for conjugation reactions with hydrophilic endogenous compounds that further increase water-solubility. These reactions are catalyzed by phase II biotransformation enzymes, which comprise glucuronyl-, glutathione-, acetyl- and sulfotransferases.¹¹ In some cases, a third phase of biotransformation is reported, describing the active transport of readily-transformed conjugates out of the single cell by multidrug resistance-related proteins or ATP-binding cassette transporters.

Biotransformation enzymes catalyzing phase I reactions usually display a low substrate specificity, which greatly enhances the substrate spectrum of these enzymes. Thus, it is easier for the organism to quickly adapt to new encountered xenobiotics and to excrete them as early as possible. However, although biotransformation reactions are generally considered a detoxification process, they may also lead to bioactivation or biotransformation, thereby generating metabolites that are reactive and more dangerous than the parent compound.

1.2.1 Implications for drug metabolism and development

Biotransformation in the context of drug metabolism is often found as 'drug-metabolizing enzymatic system' in the literature and plays a crucial role in drug discovery and development. Pharmaceutical drugs are required to cure diseases but at the same time are foreign compounds to humans or animals. Therefore, the organism treats them like any other xenobiotic and tries to get them eliminated from the body as fast as possible. While this negatively affects the bioavailability of drugs and thereby decreases the therapeutic effect, it is also desirable to excrete these substances from the body before they can cause any unintended side effects.¹² For many compounds, it is hardly predictable, if biotransformation reactions result in detoxification or bioactivation and ultimately toxification for the organism. For example, the CYP-mediated *N*-oxygenation of amines can produce hydrophilic compounds that can easily be secreted and thus promote detoxification. But equally, *N*-oxygenation may yield cytotoxic and mutagenic products.¹³

The prodrug concept takes advantage of bioactivation processes. Prodrugs are inactive pharmaceuticals that are converted into the active drug by biotransformation enzymes. The development and administration of prodrugs is often used to improve pharmacokinetic or pharmacodynamic properties of the active metabolite, such as resorption, effect duration, water-solubility, bioavailability or targeting.¹⁴

The elucidation of metabolic processes involved in the turnover of new drug candidates with regard to bioavailability, systemic clearance and toxicology is nowadays mandatory in pre-clinical research and drug development.¹⁵ Biotransformation studies comprise *in vivo* as well as *in vitro* research. *In vivo* studies performed in animals and humans are inevitable in the areas of absorption, distribution, metabolism and excretion (ADME) and give insight into complex processes in the organism, such as systemic influences, chronicle effects and tissue-specific aspects. However, in order to keep the number of *in vivo* studies at a minimum for ethical reasons and because of better analytical possibilities, *in vitro* research is equally important. Various *in vitro* systems of different complexity have been established over the past decades. They include isolated enzymes and subcellular fractions¹⁶, cell culture¹⁷ with different permanent cell lines or primary cells, tissue sections¹⁸ and perfused organs.¹⁹ The possibility to selectively alter individual parameters, easier assessment and validation of experiments, lower variance of results and easier handling are great advantages of *in vitro* studies. Thus, the metabolism of potential drug candidates by a number of different biotransformation enzymes can be thoroughly investigated.

1.3 Cytochrome P450 monooxygenases

The major enzyme family associated with biotransformation and drug metabolism is the superfamily of cytochrome P450 enzymes. To date, more than 300 different CYP enzymes have been identified throughout all phyla²⁰ and have been grouped into several families and subfamilies based on the amino acid sequence. They own their name to a characteristic, large UV absorbance at 450 nm upon treatment with carbon monoxide, which binds to CYPs with high affinity. CYPs are also found in the literature as microsomal mixed function oxidases (MFOs). These membrane-bound monooxygenases are mainly localized in the smooth endoplasmic reticulum of liver and other tissues and contain heme as a prosthetic cofactor, which comprises an iron-porphyrin unit as the oxidizing site.^{12,21} In contrast to other cellular hemoproteins, CYPs exhibit a thiol-group of a highly conserved cysteine residue serving as a ligand for the heme-iron. This results in an alteration of the porphyrin ring's electron density, providing an electronic center for the activation of molecular oxygen.²²

The CYP protein superfamily is responsible for the oxidation of a broad range of hydrophobic organic compounds, turning them into more polar metabolites for subsequent excretion. A single CYP can metabolize many chemicals with different structure, showing different degrees of affinity towards them. In principle, these enzymes catalyze the transfer of one atom of molecular oxygen to a substrate, thereby producing the oxidized compound along with a water molecule.²³ However, at least 40 different types of reactions have been described.²⁴ The detailed catalytic mechanisms of different metabolic reactions vary due to the presence of different functional groups within the substrates. Other CYP-mediated reactions include epoxidations, deaminations, aliphatic and aromatic hydroxylations as well as *O*-, *N*- and *S*-dealkylations. Furthermore, there are CYPs involved in the biosynthesis of physiologically active compounds, such as steroid hormones, prostaglandins, bile acids or vitamins. Hence, they play a crucial role in development and maintenance of homeostasis.²⁵ Moreover, CYPs can be induced by xenobiotics and other chemicals, including CYP substrates. This means, CYP enzymatic activities can be increased by *de novo* transcription, mRNA stabilization or protein stabilization.¹¹

The reaction cycle of CYPs requires the presence of a second enzyme delivering electrons needed for reduction of the P450-substrate complex prior to oxygen activation. Those accessory proteins are flavoproteins (such as NADPH-P450 reductase) which are using NADPH + H⁺ to reduce their FAD and/or FMN cofactor. Subsequently, they further transfer the electrons through other mediators (e.g. cytochrome b5) to the P450 hemoprotein where the substrate may then be oxygenized.²²

1.4 Flavin-dependent monooxygenases

Flavin-dependent monooxygenases (FMOs) comprise a second, prominent family of eukaryotic monooxygenases,²⁶ which are expressed by organisms throughout all phyla.²⁷ However, the number of FMO-coding genes and expressed isoforms varies from only one (as found in yeast) over five active enzymes in vertebrates up to a large number of different FMOs that are expressed in plants (e.g. there are 29 genes in *Arabidopsis thaliana*). Unlike CYPs, FMOs are usually not induced or down-regulated by the presence of xenobiotics.²⁸ Hence, their expression levels and relative activity are more dependent on genetic factors than on environmental influence.²⁹ The general function of FMOs is considered to be detoxification catalysts for xenobiotics. They catalyze the oxygenation of a wide range of heteroatom-containing compounds, including amine- sulfide-, phosphorous- and other nucleophilic substances, thereby converting them into polar, readily excretable metabolites.^{30,31} Although in most cases FMOs seem to sacrifice considerable enzyme velocity in exchange for a broader substrate spectrum and a less specific substrate binding domain,³² the function of individual FMOs can be quite selective. For example, human FMO3 is catalyzing the detoxification and deoderation of trimethylamine with high specificity.³³

In contrast to CYPs, FMOs are stand-alone enzymes and therefore independent from electron-delivering accessory proteins. They are provided with reducing equivalents directly from their cosubstrate NADPH.³⁴ This enzyme class is subdivided into external flavoprotein monooxygenases (EC 1.14.13.#) and internal monooxygenases (EC 1.13.12.#). In contrast to their external counterparts, the flavin cofactor of internal FMOs is reduced by the substrate itself instead of using reducing equivalents from NADH or NADPH prior to substrate oxygenation.

A further classification of external FMOs into six subclasses (A-F) is discriminated by sequence similarity and the presence of specific protein sequence motifs. FMOs of classes A and B are encoded by a single gene, whereas enzymes of classes C-F comprise at least two components – a monooxygenase and a reductase. While there is only one dinucleotide binding domain (Rossmann fold) present in FMOs of class A and these enzymes release NADP⁺ immediately after reduction of their FAD cofactor, members of class B FMOs are composed of two dinucleotide binding domains (used for binding FAD as a prosthetic group and recruiting NADPH)³⁵ and keep the oxidized cosubstrate bound during catalysis. Subclasses C-F contain a different set of flavin cofactors (FAD or FMN) and further differ in their sequence identity and motives.³⁶

In this work, only the “two dinucleotide binding domain” containing FMOs of class B are further discussed. Their reaction cycle comprises at least five consecutive steps that are depicted in Figure 1.

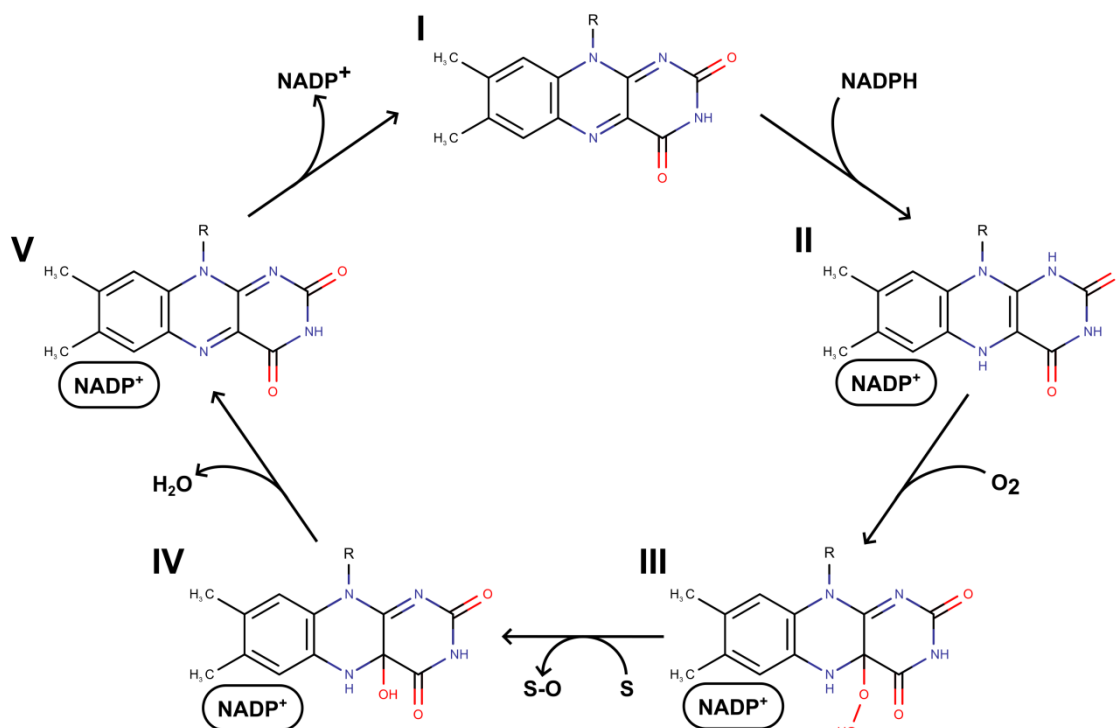


Figure 1: FMO catalytic cycle. Adapted from Alfieri *et al.*, 2008.³⁷

In principle, these enzymes bind NADH or NADPH in order to reduce their prosthetic FAD cofactor and subsequently oxygenate a substrate. The sequence of cofactor and substrate recruitment remained elusive for a long time and different mechanisms have been discussed in the past decades. However, with an increasing number of FMO structures crystallized in different stages during catalysis, the following reaction cycle, which was already postulated on the basis of kinetic studies,³¹ is now widely accepted and distributed in the literature.

In a first step, NAD(P)H is bound and used to reduce the FAD cofactor (I). Upon reaction of reduced FAD (II) with molecular oxygen, a C4a-hydroperoxy-FAD intermediate is formed (III). This intermediate is capable of inserting one oxygen atom into a substrate compound and ends up as C4a-hydroxy-FAD (IV), which may then release the second oxygen atom as part of a water molecule, restoring the oxidized FAD cofactor (V).³¹ During the whole catalytic cycle, the oxidized NAD(P)⁺ remains bound to the enzyme and is most likely involved in the binding of substrates and the stabilization of reaction intermediates, e.g. the activated C4a-hydroperoxy-FAD.³⁸⁻⁴⁰ Only after product release, the NAD(P)⁺ cosubstrate is exchanged for its reduced equivalent.⁴¹

1.4.1 Pyrrolizidine alkaloid *N*-oxygenases

Plants and herbivores are engaged in a permanent co-evolutionary arms race against each other. While plants develop and improve defense mechanisms to protect themselves from harmful threats such as pathogenic microbes or herbivorous insects, the latter counteract with ever-developing strategies to circumvent their food plants' defense.⁴² As part of their chemical defense, certain angiosperm plant species have developed pyrrolizidine alkaloids (PAs).⁴³ PAs are products of the plants' secondary metabolism and encompass several hundred different structures. Usually, they are produced in their non-toxic *N*-oxide form. However, upon ingestion by a vertebrate or insect herbivore, they are converted into the protoxic free base in the reducing gut milieu. PAs in their free base form are lipophilic compounds that can easily permeate membranes and serve as xenobiotic substrates that are recognized by biotransformation enzymes like cytochrome P450 monooxygenases. CYP-mediated bioactivation leads to the formation of pyrrolic compounds, which are capable of reacting with proteins or nucleic acids and are therefore cell-toxic.^{44,45}

Hence, herbivores utilizing PA-producing plants as a food source need to have detoxification strategies, which are in most cases based on specialized enzymatic systems. PAs may be hydrolyzed into the non-toxic necine base by carboxylesterases,⁴⁶ conjugated with glutathione in order to form a non-toxic excretable compound⁴⁷ or *N*-oxygenated by FMOs (Figure 2).^{48,49} The latter strategy is utilized by insects like the African locust *Zonocerus variegatus* or the cinnabar moth *Tyria jacobaeae*. They developed specialized FMOs, so-called pyrrolizidine alkaloid *N*-oxygenases (PNOs) that can convert the pro-toxic free base form of PAs (a tertiary amine) into the non-toxic *N*-oxide, which can then be safely retained within these insects.^{26,43,50} Moreover, these *N*-oxygenated metabolites serve these insects' own chemical defense against their predators.^{51,52}

To date, only a few PNOs are biochemically characterized, whereas there is no structural information available yet.

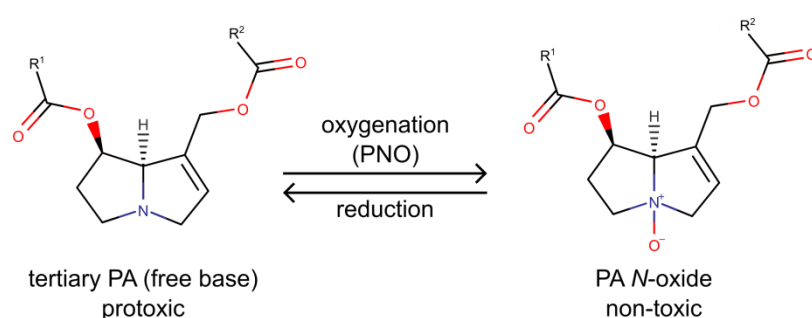


Figure 2: PNO-mediated *N*-oxygenation of PAs. Adapted from Langel & Ober, 2011.⁵³

1.5 Molybdenum-dependent enzymes

The trace element molybdenum is assimilated in the form of molybdate ions (MoO_4^{2-}) via specific transporters and is incorporated into enzymes in the form of the molybdenum cofactor (Moco). This pyranopterin-molybdenum complex and the iron-molybdenum cofactor, which is exclusively found in the bacterial nitrogenase, are the only biologically active molybdenum-dependent cofactors.⁵⁴ Due to these cofactors' redox-capabilities, molybdoenzymes catalyze various redox-reactions and are essential for carbon, nitrogen and sulfur metabolism.⁵⁵ Thus, more highly developed organisms cannot cope with molybdenum depletion.⁵⁶ During Moco-catalyzed reactions, two electrons are being transferred together with an oxygen atom derived from either a water molecule or a substrate, depending on whether an oxidation or reduction is catalyzed. Throughout the catalytic cycle the molybdenum's oxidation state varies from +IV to +VI.⁵⁷

Most of the >50 discovered molybdoenzymes are found in bacteria, whereas only seven were identified in eukaryotes (and only four in mammals).⁵⁷ The eukaryotic enzymes comprise pyridoxal oxidase (only found in *Drosophila melanogaster*⁵⁸), nicotinate hydroxylase (only found in *Aspergillus nidulans*⁵⁹), nitrate reductase, sulfite oxidase, xanthine oxidoreductase, aldehyde oxidase and the recently discovered mARC.⁵⁷ With the exception of pyridoxal oxidase and nicotinate hydroxylase, which are no further discussed in this thesis, the eukaryotic Moco-dependent enzymes are shortly introduced in the following chapters.

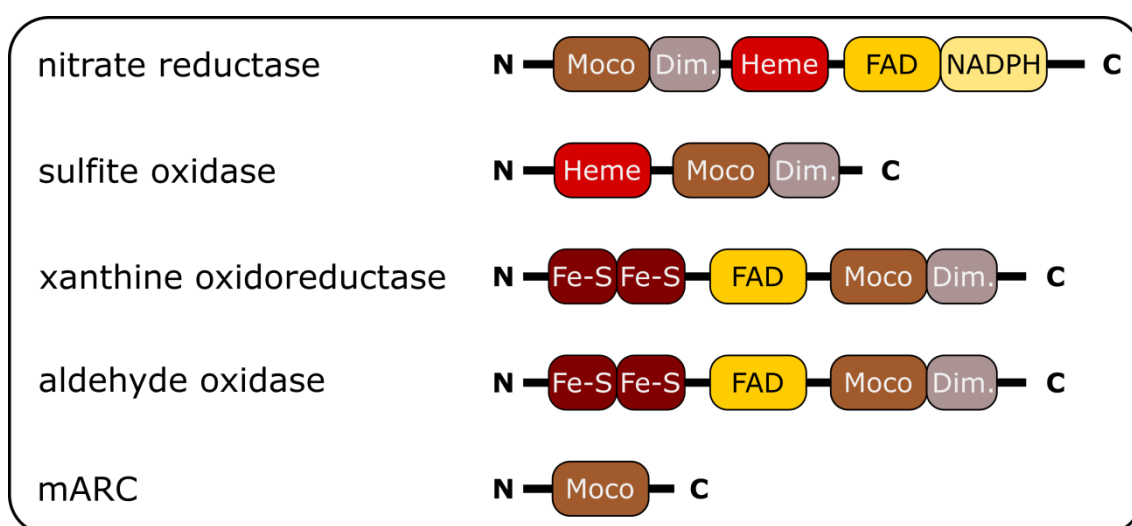


Figure 3: Domain arrangement of eukaryotic Moco-dependent enzymes. Moco: Moco-binding domain; Dim.: dimerization domain; Heme: heme-binding domain; FAD: FAD-binding domain; NADPH: NADPH-binding domain; Fe-S: Fe-S cluster binding domain. Sulfite oxidase as depicted represents the animal form; the plant enzyme lacks the heme-binding site. Adapted from Mendel and Kruse, 2012.⁵⁶

1.5.1 Nitrate reductase (NR)

The eukaryotic NRs display the same domain architecture and comprise three individual domains, connected by two solvent-exposed linker regions called hinge I and hinge II. The *N*-terminal domain binds the Moco as a prosthetic group and additionally contains a dimerization motif. It is followed by a heme-binding cytochrome b_5 domain and a *C*-terminal domain comprising binding sites for the FAD prosthetic group as well as the NAD(P)H cofactor (Figure 3).⁵⁶ The reaction cycle starts with a reductive half-reaction (reduction of FAD by NADPH), followed by an oxidative half-reaction, where electrons are transferred from FAD to the Moco active site via cytochrome b_5 . There, nitrate is reduced to nitrite in a final reaction step.⁶⁰ This enzyme is exclusively found in autotrophic species, such as fungi, algae and plants. In addition to nitrate reduction, NR also shows a low intrinsic capability of reducing nitrite to nitric oxide (NO) and therefore might play a subordinate role in NO signaling and related physiological functions.^{61,62}

1.5.2 Sulfite oxidase (SO)

There are different forms of SO among eukaryotes: the plant SO solely comprises a Moco-binding domain⁶³, whereas the counterpart found in animals displays an additional *N*-terminal heme-binding cytochrome b_5 domain (Figure 3).⁵⁷ This dimeric enzyme catalyzes the oxidation of sulfite to sulfate ions and therefore contributes to sulfite detoxification. Although the reaction mechanism is the same for animal and plant SO and comprises a two electron transfer from the sulfite to the Moco active site, the utilized electron acceptors as well as the localization of these enzymes differ. Plant SO is a peroxisomal protein⁶⁴ that uses molecular oxygen as electron acceptor and produces hydrogen peroxide during sulfite oxidation.⁶⁵ In contrast, human SO is localized in the intermembrane space of mitochondria and utilizes cytochrome *c* as electron acceptor. In addition to sulfite detoxification, human SO is the key enzyme that catalyzes the final step in the degradation of sulfur-containing amino acids.⁵⁶ The most essential physiological role of plant SO, however, is the maintenance of the sulfate-sulfite cycle, which is crucial for a viable sulfur distribution in the cell.⁶⁶

1.5.3 Xanthine oxidoreductase (XOR)

In a two-step reaction, dimeric XOR catalyzes the oxidation of hypoxanthine to xanthine and, subsequently, to uric acid. Therefore, it plays an essential role in purine metabolism. This cytosolic enzyme consists of two [2Fe-2S] clusters, a FAD binding domain and a C-terminal domain that binds the Moco and contains the dimerization motif (Figure 3). As part of the reaction cycle electrons are abstracted from the substrate upon conversion at the Moco active site, relayed to the FAD cofactor via the intramolecular chain of [2Fe-2S] clusters and finally transferred to either NAD^+ (xanthine dehydrogenase; XDH) or molecular oxygen (xanthine oxidase; XO).^{56,67} In addition to hypoxanthine and xanthine, XOR accepts a broad substrate spectrum, including a large number of aldehydes and aromatic heterocycles, which are hydroxylated upon conversion. While the plant XDH solely accepts NAD^+ as a cosubstrate, there are two animal XOR forms (XDH and XO), which can be converted into each other. There are several physiological functions proposed for plant XDH, such as involvement in ROS metabolism⁶⁸, drought stress⁶⁷ and natural senescence.^{69,70} Furthermore, human and animal XOR have been shown to play a role in inflammatory response to ischemia-reperfusion⁷¹ as well as in the formation of milk fat droplets.⁷²

1.5.4 Aldehyde oxidase (AO)

The domain architecture of AO resembles that of XOR, with whom it also shares a high degree of sequence similarity (Figure 3). This is because during evolution AOs have derived from XORs upon a gene duplication event.⁷³ Nonetheless, these are distinct enzymes that display differences in the substrate binding site and their physiological cofactors.⁵⁵ Furthermore, the substrate spectrum of AOs is even broader than that of XORs. It covers purines, pteridines, aldehydes and aromatic as well as aliphatic heterocycles. AOs are dimeric, cytosolic enzymes that exclusively accept molecular oxygen as cosubstrate during catalysis. While mammalian AOs may produce superoxide anions as well as hydrogen peroxide as byproducts, plant AOs solely yield hydrogen peroxide. To date, despite their potential role in detoxification processes, the physiological role of mammal AOs remains uncertain, whereas different functions of plant AOs are described.⁵⁶ In *Arabidopsis thaliana*, the isoform AO δ catalyzes the oxidation of abscisic aldehyde to abscisic acid,⁷⁴ a phytohormone that is essential for various stress responses and in plant development.⁷⁵ Furthermore, AOs are implicated in the biosynthesis of another fundamental phytohormone, auxin.

1.5.5 Mitochondrial amidoxime-reducing component (mARC)

As mentioned before, drugs are a subset of xenobiotics and numerous of them contain nitrogen in different functional groups, especially amidine moieties. They include trypsin-like serine protease inhibitors like factor Xa inhibitors,⁷⁶ thrombin inhibitors, factor VIIa inhibitors⁷⁷ and urokinase-type plasminogen activators.⁷⁸ Furthermore, they are used as anti-parasitic,⁷⁹ antibacterial⁸⁰ and anti-malarial agents⁸¹ as well as antiplatelet GPIIb/IIIa-receptor antagonists.⁸² While the strongly basic amidine group is essential for the interaction with the target proteins, it suffers from poor oral bioavailability. Thus, lowering its high basicity by *N*-hydroxylation (thereby generating an amidoxime group) became a prodrug strategy to temporarily mask the charged moiety and improve intestinal absorption.

In the course of prodrug development an *N*-reductive enzymatic system comprising cytochrome *b*₅, cytochrome *b*₅ reductase and a third, so far unknown, enzyme was held responsible for the conversion and activation of amidoxime-containing compounds. In 2006, this enzyme was isolated from pig liver mitochondria and named after its localization and first obvious function, mitochondrial amidoxime-reducing component (mARC).⁸³ It was identified as a molybdenum-dependent enzyme (the fourth found in humans) and was found to be expressed as two different paralogues, which share strong similarities on nucleotide and amino acid level, in all annotated mammals.⁸⁴ Predictions on structural features suggest that mammalian mARC proteins not only harbor a molybdenum-binding Moco sulfurase C-terminal (MOSC) domain (Figure 3), but also an *N*-terminal β -barrel subdomain that has been discussed to be involved in substrate interaction.⁸⁵ Moreover, mammalian mARC proteins carry another extension on their extreme *N*-terminal end that is likely to be required for mitochondrial targeting.⁸⁶ mARC proteins represent the simplest eukaryotic molybdenum enzymes in that they have the lowest molecular weight (around 35 kDa) and bind Moco as the only prosthetic group. The present knowledge on these enzymes suggests that they are not active as stand-alone proteins, but rather act in concert with other redox-active proteins such as cytochrome *b*₅ and NADH cytochrome *b*₅ reductase.⁸⁴ Yet, the existence of other physiological redox partners and the performance of alternative physiological reactions should be considered, in particular since the localization of mARC on the outer mitochondrial membrane principally allows contact with putative redox proteins of the cytosol.

While its physiological function is as yet largely unknown, mARC has been demonstrated to represent the central part of the aforementioned three-component system that catalyzes the reduction of various *N*-hydroxylated substrates. Electrons (derived from NADH) are delivered from the FAD of cytochrome *b*₅ reductase through the heme of cytochrome *b*₅ to the Moco of mARC, which is likely to provide the substrate binding site. Interestingly, this electron

transport chain resembles that of eukaryotic nitrate reductase, despite the fact that it consists of separate proteins while nitrate reductase combines all cofactors in a single polypeptide chain. The proposed mARC reaction cycle is depicted in Figure 4.

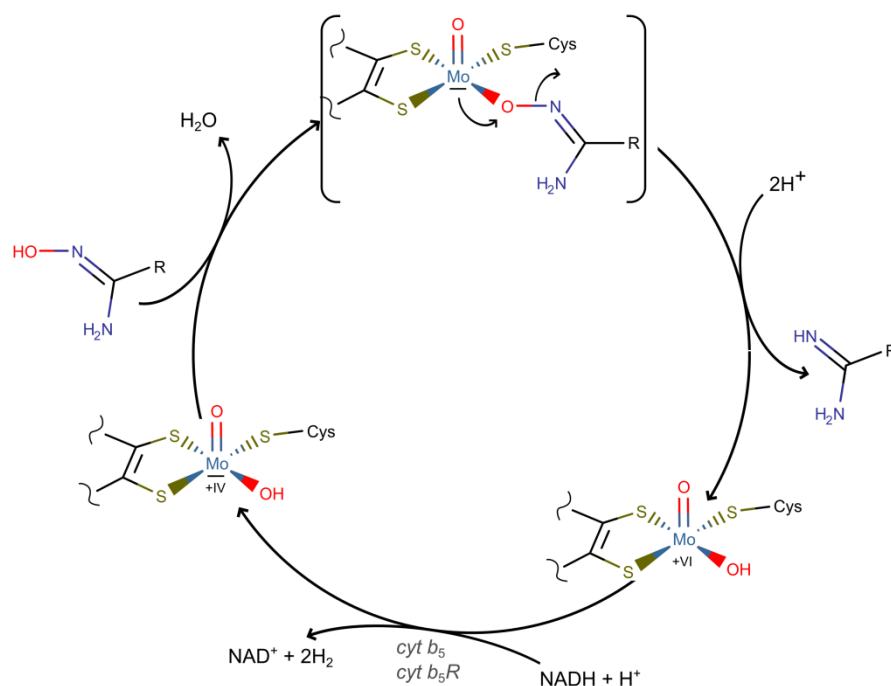


Figure 4: Hypothetical reaction cycle of amidoxime reduction by mARC. The scheme was adapted from Havemeyer *et al.*, 2014.⁸⁷

Over the course of the past decade, mARC was found to be an extremely effective *N*-reductive generalist. In addition to amidoximes, it accepts a broad range of *N*-hydroxylated and *N*-oxygenated compounds. These include hydroxylamines, hydroxyguanidines, oximes, sulfohydroxamic acids and even *N*-oxides.^{84,88-90} Moreover, mARC could be directly linked to detoxification processes. It was shown to effectively reduce toxic and mutagenic *N*-hydroxylated nucleobase analogues and thereby protects cells from apoptotic effects.^{91,92} Furthermore, mARC was identified as a direct counterpart to the CYP-mediated hydroxylation and toxification of the antibiotic sulfamethoxazole, which is converted into the hypersensitivity-inducing hydroxylamine upon phase I biotransformation reactions and needs to be reduced back to the active antibiotic.^{93,94} Other physiological functions that have been addressed to mARC enzymes are the reduction of *N*^ω-hydroxy-L-arginine that is an intermediate in NO synthesis and a regulator in the NO signal pathway,^{95,96} and an involvement in energy metabolism. Although the direct link is yet to be discovered, a correlation between mARC expression levels and the glucose metabolism was identified. Especially mARC2 was found to be connected to type 1 *diabetes mellitus*.⁹⁷ Moreover, there are striking indications for an involvement of this enzyme in lipogenesis.^{89,98,99}

1.6 Molybdoenzyme superfamilies

The eukaryotic Moco-dependent enzymes are subdivided into two separate molybdoenzyme superfamilies, depending on the coordination sphere of the central molybdenum ion. Enzymes of the sulfite oxidase (SO) family contain a form of Moco, where the molybdenum is coordinated by the dithiolene sulfurs of the molybdopterin (MPT) ring system, two oxygen ligands and a proteinogenic cysteine-sulfur (Figure 5, left). Among eukaryotic molybdoenzymes, this superfamily comprises SO itself and plant NR. However, the SO-Moco may be converted by the enzyme Moco sulfuryase, which exchanges the proteinogenic cysteine-sulfur for a terminal sulfur ligand (Figure 5, right). The resulting Moco form is exclusively incorporated into members of the xanthine oxidase (XO) superfamily.^{56,57}

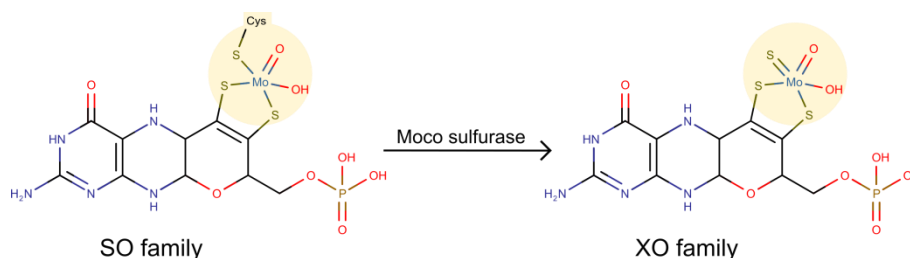


Figure 5: Molybdenum cofactors of eukaryotic molybdenum-dependent enzymes. *Left:* Moco as it is present in members of the SO family. *Right:* Moco bound by members of the XO family.

In contrast to the molybdenum center of XO family enzymes, the central molybdenum ion of mARC proteins is not coordinated by a terminal sulfur. This has been demonstrated on the one hand by cyanide treatment, which neither released sulfur in the form of thiocyanate nor significantly affected the activities of recombinant human mARC proteins.¹⁰⁰ On the other hand, partially reduced human mARC proteins developed EPR signals characteristic of the Mo(V) (d^1) state,¹⁰¹ closely resembling the so-called “low pH” EPR signal seen with SO family enzymes, which do not have a terminal sulfur ligand but a cysteine-sulfur derived from the protein. It was thus concluded that mARC proteins indeed provide a cysteine-sulfur to the molybdenum center, thereby leading to classify mARC enzymes as new members of the SO family.

In addition to the molybdenum coordination sphere, Rothery *et al.* investigated correlations between the MPT conformations found in crystal structures of various molybdoenzymes and their belonging to either the SO or XO superfamily. Interestingly, they provided evidence that certain dihedral angles within the MPT are specific markers, which allow discrimination between both enzyme superfamilies based on the cofactors' conformation. Their results clearly correlated with previous annotations based on the ligand sphere of the molybdenum center.¹⁰²

2 Scope

The current thesis is devoted to the structure determination of different biotransformation enzymes by X-ray crystallography. It is subdivided into two topics: (i) the discussion of crystal structures of flavin-dependent monooxygenases from the African locust *Zonocerus variegatus* (ZvFMO) and (ii) biochemical and structural characterization of the human mitochondrial amidoxime reducing component (mARC).

Our current structural knowledge about enzymes belonging to the large family of “two dinucleotide binding domain” containing flavin-dependent monooxygenases is so far limited to crystal structures of bacterial, fungal or yeast enzymes. The main objective of the structure determination of different ZvFMOs was to gain insight into a representative protein from a more highly developed organism. It was supposed that these enzymes, which are closer related to the pharmaceutically relevant human FMOs, might exhibit different structural features or oligomeric arrangements compared to currently known FMOs from lower developed organisms. Therefore, hitherto unknown aspects of this enzyme family could be uncovered that shed light on the complex balance between their broad substrate spectrum on the one hand and sufficient specificity on the other hand.

The FMOs from *Zonocerus variegatus* were chosen as promising candidates for this fundamental research because they provide a number of advantages. Most importantly, in contrast to many other eukaryotic FMOs, they are expressed as soluble proteins and as such were supposed to be crystallized with less effort compared to membrane or membrane-associated proteins. Secondly, they derive from an organism that is more highly developed than prokaryotes or basic eukaryotes like yeast or fungi, but is still less complex in comparison to humans and therefore allows for an investigation of the evolutionary development of this enzyme family. Furthermore, there are three isoforms of ZvFMOs with different specific enzyme activities towards their designated substrates, pyrrolizidine alkaloids. Based on the comparison of the three-dimensional structures of these wildtype enzymes, rational design and characterization of ZvFMO variants were conducted in order to identify amino acids of the active or substrate binding site that specifically modify enzyme activity and substrate affinity. Moreover, crystal structures of ZvFMOs in complex with different substrates and/or trapped in different states of the catalytic cycle might ultimately complete our understanding of the underlying reaction mechanism and associated conformational changes within FMOs.

Therefore, it was the objective to identify individual crystallization conditions for the ZvFMO isoforms, perform X-ray diffraction experiments and data collection at synchrotron facilities,

determine the three-dimensional structures and compare as well as interpret them with regard to the aforementioned aspects.

The second focus of this work is the crystallization strategy, biochemical/biophysical characterization as well as structure determination and discussion of human mARC. This molybdenum-dependent biotransformation enzyme plays a major role in *N*-reductive metabolism and is of high pharmaceutical relevance since it is involved in detoxification processes, prodrug activation but also drug inactivation and needs to be monitored during pre-clinical research when evaluating the metabolic stability of novel drug candidates.

In order to have sufficient amounts of active protein for all subsequent analyses, including activity assays, substrate screenings and crystallization, the expression and purification protocols for new soluble mARC constructs had to be established and optimized. Especially, the hitherto poor saturation of human mARC with its molybdenum cofactor was to be improved. Higher yields of both mARC paralogues enabled high-throughput screenings for crystallization conditions as well as the utilization of diverse crystallization strategies, e.g. a fusion-protein approach. Furthermore, the recombinant enzymes were to be used to establish cyclic voltammetry-based electrochemical screening methods for potential substrates or inhibitors.

The ultimate goal was the elucidation of the three-dimensional crystal structure of human mARC, which is the only human molybdoenzyme that is not structurally characterized yet. Moreover it belongs to a subfamily of MOSC proteins that are proposed to comprise an *N*-terminal β -barrel domain and a *C*-terminal domain resembling that of Moco sulfurases. However, these potential structural motifs are solely based on *in silico* predictions. Evidence based on crystal structures is still missing, since no protein containing both of these domains has been crystallized so far. Furthermore, to be elucidated three-dimensional structures of mARC enzymes should give insight into differences between the two paralogues and might explain their slightly different substrate preferences. Detailed knowledge about the substrate binding site as well as the active site of these enzymes might even allow predictions on the turnover efficiency of different substrates. Finally, structural observations were to be correlated to previous biochemical and biophysical characterizations of mARC enzymes and used to explain these findings in more detail.

In order to fulfil this main objective, soluble mARC paralogues were to be heterologously expressed, purified and crystallized, X-ray diffraction data were to be collected, suitable phasing strategies were to be applied, model building was to be performed and the crystal structure was to be analyzed, interpreted and correlated to our current knowledge of these molybdenum-dependent biotransformation enzymes.

3 Results and Discussion

This work is a cumulative dissertation. The thesis comprises three peer-reviewed scientific papers and two currently submitted manuscripts. All of these were prepared during my PhD work. Each publication is shortly introduced and then reproduced as published in scientific journals or as they were submitted for peer-review, respectively. Supplementary material is included for papers and manuscripts.

The first publication deals with the structure determination of insect flavin-dependent monooxygenases and their comparison to formerly characterized members of this enzyme family as well as implications for the 'arms race' between the host organism and its food plants. The last four papers address the human mitochondrial amidoxime-reducing component (mARC). They include a newly-developed voltammetric assay which was designed to enable fast identification of so far unknown substrates and inhibitors of this enzyme, the identification of hydroxamic acid compounds as a new class of mARC substrates, the crystallization strategy of a mARC-T4 lysozyme fusion protein and the discussion of the three-dimensional structure of human mARC, which was determined by X-ray crystallography during my PhD work.

3.1 Crystal structure of pyrrolizidine alkaloid *N*-oxygenase from the grasshopper *Zonocerus variegatus*

Christian Kubitz, Annette Faust, Miriam Gutt, Luzia Gäth, Dietrich Ober and Axel J. Scheidig

Acta Crystallographica Section D **2018**, *74*, 422-433.

DOI: 10.1107/S2059798318003510

Flavin-dependent monooxygenases contribute to an efficient biotransformation and detoxification system. Multiple isoforms of these enzymes are expressed throughout all kingdoms of life. However, our structural knowledge of this enzyme family was so far limited to crystal structures of FMOs found in bacteria, yeast or fungi.

In the following paper,¹⁰³ the first crystal structure of a more highly-developed organism, the African locust *Zonocerus variegatus*, is described. Three different soluble FMO isoforms could be identified in this organism and were heterologously expressed in *E. coli*. Two isoforms, ZvPNO and ZvFMOa, could be successfully crystallized and their three-dimensional structure was determined by X-ray diffraction. Different high-resolution datasets of ZvPNO provided detailed insight into the coordination of the two dinucleotide cofactors and related conformational rearrangements of the enzyme. Despite a high conservation of the overall fold among all structurally characterized FMOs, a hitherto unknown dimeric arrangement of subunits was observed for the FMO isoforms found in *Zonocerus variegatus*. This feature was not only observed for the ZvPNO, but also for the ZvFMOa structure, which was determined at a lower resolution. To further investigate significant differences in the catalytic activity of ZvFMOs, a variety of ZvFMOa variants with amino acid exchanges near the active site were generated. Thereby, significant impacts on the catalytic activity were observed even for single exchanges within the substrate entry site. The atomic coordinates and structure factors were deposited in the Protein Data Bank (<http://www.pdb.org/>) under accession numbers 5NMW and 5NMX.

Together with Luzia Gäth, whom I supervised for her master thesis, I purified heterologously expressed ZvFMO variants. I performed crystallization, data collection, data processing, phase calculation, structure refinement and analysis of the isoforms ZvPNO as well as ZvFMOa. Furthermore, I wrote the manuscript together with Prof. Dr. Ober and Prof. Dr. Scheidig.



ISSN 2059-7983

Crystal structure of pyrrolizidine alkaloid *N*-oxygenase from the grasshopper *Zonocerus variegatus*

Christian Kubitz,^a Annette Faust,^a Miriam Gutt,^b Luzia Gäth,^a Dietrich Ober^b and Axel J. Scheidig^{a*}

Received 15 November 2017

Accepted 28 February 2018

^aStructural Biology, Zoological Institute, Kiel University, Am Botanischen Garten 1–9, 24118 Kiel, Germany, and^bBiochemical Ecology and Molecular Evolution, Botanical Institute, Kiel University, Am Botanischen Garten 1–9, 24118 Kiel, Germany. *Correspondence e-mail: axel.scheidig@strubio.uni-kiel.de

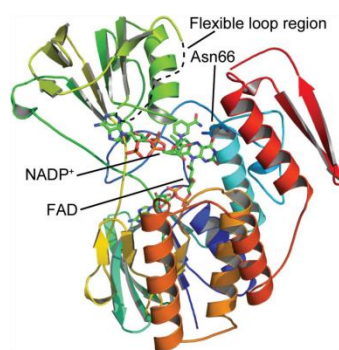
Edited by M. Rudolph, F. Hoffmann-La Roche Ltd, Switzerland

Keywords: flavin-dependent monooxygenase; FMO; pyrrolizidine alkaloids; PNO; *Zonocerus variegatus*.

PDB references: ZvPNO complex with FAD and NADP⁺, 5nmx; ZvPNO complex with FAD, 5nmw

Supporting information: this article has supporting information at journals.iucr.org/d

The high-resolution crystal structure of the flavin-dependent monooxygenase (FMO) from the African locust *Zonocerus variegatus* is presented and the kinetics of structure-based protein variants are discussed. *Z. variegatus* expresses three flavin-dependent monooxygenase (ZvFMO) isoforms which contribute to a counterstrategy against pyrrolizidine alkaloids (PAs). PAs are protoxic compounds produced by some angiosperm lineages as a chemical defence against herbivores. *N*-Oxygenation of PAs and the accumulation of PA *N*-oxides within their haemolymph result in two evolutionary advantages for these insects: (i) they circumvent the defence mechanism of their food plants and (ii) they can use PA *N*-oxides to protect themselves against predators, which cannot cope with the toxic PAs. Despite a high degree of sequence identity and a similar substrate spectrum, the three ZvFMO isoforms differ greatly in enzyme activity. Here, the crystal structure of the *Z. variegatus* PA *N*-oxygenase (ZvPNO), the most active ZvFMO isoform, is reported at 1.6 Å resolution together with kinetic studies of a second isoform, ZvFMOa. This is the first available crystal structure of an FMO from class B (of six different FMO subclasses, A–F) within the family of flavin-dependent monooxygenases that originates from a more highly developed organism than yeast. Despite the differences in sequence between family members, their overall structure is very similar. This indicates the need for high conservation of the three-dimensional structure for this type of reaction throughout all kingdoms of life. Nevertheless, this structure provides the closest relative to the human enzyme that is currently available for modelling studies. Of note, the crystal structure of ZvPNO reveals a unique dimeric arrangement as well as small conformational changes within the active site that have not been observed before. A newly observed kink within helix α 8 close to the substrate-binding path might indicate a potential mechanism for product release. The data show that even single amino-acid exchanges in the substrate-entry path, rather than the binding site, have a significant impact on the specific enzyme activity of the isoforms.



© 2018 International Union of Crystallography

1. Introduction

Flavin-dependent monooxygenases (FMOs), alongside cytochrome P450 enzymes, are one of two prominent families of monooxygenases in eukaryotes (Cashman, 2001). They are found throughout all phyla (Hao *et al.*, 2009), although they are expressed in different numbers of isoforms, ranging from only one FMO in yeast to five isoforms in vertebrates and to a large gene family in plants. FMOs directly recruit reducing equivalents from NAD(P)H without requiring any accessory proteins for enzyme activity (Phillips & Shephard, 2008).

FMOs, at least in vertebrates, mainly contribute to an efficient biotransformation and detoxification system for xenobiotics, converting nucleophilic heteroatom-containing

research papers

compounds (such as amines, amides, thiols and sulfides) into polar, readily excretable metabolites (Cashman, 2001; Ziegler, 2002).

During evolution, different insect lineages have developed specialized FMOs, pyrrolizidine alkaloid *N*-oxygenases (PNOs), as a counterstrategy to cope with pyrrolizidine alkaloids (PAs), which are toxic compounds that are produced by certain angiosperm species as part of their chemical defence against herbivores (Hartmann & Ober, 2008). PAs are produced by plants in their nontoxic polar *N*-oxide form. After ingestion by a vertebrate or insect herbivore, PAs are converted into the protoxic free base in the reducing gut milieu (Fig. 1). Owing to their lipophilic properties, free bases easily permeate membranes and are bioactivated to reactive pyrrolic compounds by cytochrome P450 monooxygenases that are part of the xenobiotic metabolism of the herbivore (Fu *et al.*, 2004).

PA *N*-oxygenases are the only functionally characterized FMOs found to date in insects. They are one of various strategies of adapted insects to avoid high concentrations of toxic PAs in their haemolymph by enzymatically stabilizing the nontoxic *N*-oxide (Fig. 1). This mechanism is realized in the larvae of arctiid moths (Hartmann *et al.*, 1990; Sehlmeier *et al.*, 2010; Naumann *et al.*, 2002), *Longitarsus* flea beetle species (Narberhaus *et al.*, 2003) and the grasshopper genus *Zonocerus* (Bernays *et al.*, 1977; Wang *et al.*, 2012). Identical selection pressure upon different insect lineages has led to the independent yet convergent evolution of FMOs with almost

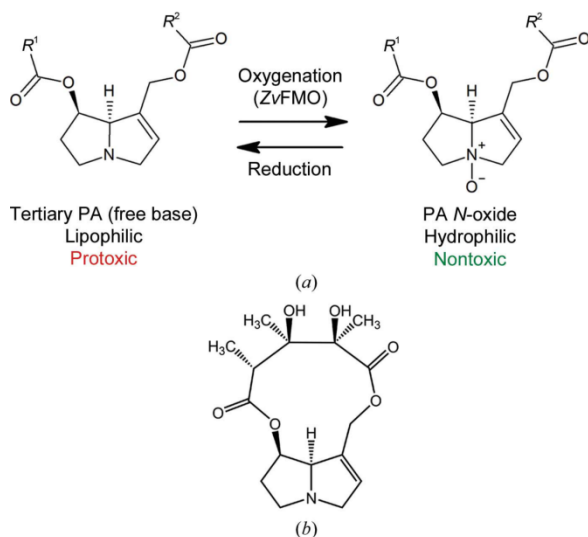


Figure 1

N-Oxygenation of pyrrolizidine alkaloids by ZvFMOs. Upon ingestion, plant-derived PAs are reduced to their protoxic tertiary amine form within the gut of the herbivore (Langel & Ober, 2011). (a) In adapted insects such as *Zonocerus*, protoxic PAs are efficiently converted to the respective PA *N*-oxides by ZvFMOs. PAs are esters of the bicyclic necine base moiety with one or more necic acids that are represented by R^1 and R^2 (Hartmann & Witte, 1995). (b) During this study, the PA monocrotaline was used as a substrate to characterize the different FMOs from *Zonocerus*.

identical substrate specificity in these organisms (Wang *et al.*, 2012).

Three isoforms of flavin-dependent monooxygenase from the locust *Z. variegatus* (ZvFMOs) have been identified, named ZvFMOa, ZvFMOc and ZvPNO. They all share a similar substrate spectrum and are able to oxygenate a variety of different PAs. However, they exhibit quite different enzyme activities. While ZvPNO is the most potent isoform, ZvFMOa and ZvFMOc show a specific activity that is eightfold to 300-fold lower, depending on the PA substrate. The high degree of amino-acid sequence identity (77–77.8%) between the various ZvFMOs combined with their similar substrate spectra indicates a duplication event of a common ancestor FMO-coding gene that already possessed PA *N*-oxygenating capabilities and was subsequently recruited and optimized for plant-derived PAs (Wang *et al.*, 2012).

The FMO reaction cycle had been investigated before the first structures became available, and is widely distributed in the literature. Class B FMOs such as those from *Z. variegatus* consist of two domains, each of which contains a dinucleotide-binding domain in the form of a Rossmann fold, which is responsible for cofactor binding. In its native state, oxidized FAD is tightly bound to the enzyme as a prosthetic group. NADPH is recruited as a co-substrate and transfers reducing equivalents to FAD. Upon the reaction of reduced FAD by molecular oxygen, a C4a-hydroperoxy-FAD intermediate is formed. This intermediate is capable of inserting one O atom into a substrate compound and ends up as C4a-hydroxy-FAD, which may then release the second O atom as part of a water molecule, restoring the oxidized FAD cofactor (Ziegler, 2002; Robinson *et al.*, 2013). A cocked-gun mechanism has been proposed for the reaction cycle of FMOs, meaning that the activated C4a-hydroperoxy-FAD intermediate can be stabilized within the enzyme until a substrate accesses the active site and is then immediately oxygenated (Cashman, 1995; Ziegler, 1993). The presence of NADP⁺ seems to be crucial for the stabilization of this intermediate and therefore it has to remain bound to the enzyme throughout the whole catalytic cycle, making it the last compound to be released from the enzyme (Beatty & Ballou, 1981).

Here, we present two crystal structures: (i) the pyrrolizidine alkaloid *N*-oxygenase (ZvPNO) from *Z. variegatus* with bound FAD and (ii) ZvPNO in complex with FAD and NADP⁺. In addition, we present data on ZvFMOa variants, which are discussed in terms of a better understanding of substrate conversion as well as the evolutionary background and specialization of FMOs in insects. The crystal structures presented in this study are the first available FMO structures to originate from more highly developed eukaryotes. Our data show a flexible helix close to the active site which is likely to be involved in substrate binding and/or product release and has not been described before in other FMO structures. Although ZvFMO isoforms are dimeric in solution, as is frequently observed for other members of the FMO family, their specific arrangement of the two subunits is unique. Of note, the contact area is rather small; nevertheless, the dimer is very stable in solution.

research papers

2. Experimental procedures

2.1. Generation of expression plasmids for ZvFMO isoforms and variants

The identification of three different ZvFMO isoforms (ZvFMOa, ZvFMOc and ZvPNO) and the generation of individual expression plasmids have been described previously (Wang *et al.*, 2012). Expression vectors for the ZvFMOa variants were generated according to the QuikChange site-directed mutagenesis protocol (Stratagene, La Jolla, California, USA) using Phusion high-fidelity DNA polymerase (Fermentas) with the pET-22b-ZvFMOa plasmid as a template and the following primers: ZvFMOa variant F307Y, forward, 5'-CCT CAC GAC GCC TGT TAT TCG ATC TTG TTT GAT C-3'; reverse, 5'-G ATC AAA CAA GAT CGA ATA ACA GGC GTC GTG AGG-3'; ZvFMOa variant Y356A, forward, 5'-CGC CCG CAC TTC ATG GCT AAC CGT CAG TGG AAG-3'; reverse, 5'-CTT CCA CTG ACG GTT AGC CAT GAA GTG CGG GCG-3'; ZvFMOa variant P388S, forward, 5'-G TTC GAT GAT CTG GCT TCT GGT TTG ACG AAG GAC-3'; reverse, 5'-GTC CTT CGT CAA ACC AGA AGC CAG ATC ATC GAA C-3'; ZvFMOa variant F354V, forward, 5'-GGC TTC CGC CCG CAC GTC ATG TAT AAC CGT CAG-3'; reverse, 5'-CTG ACG GTT ATA CAT GAC GTG CGG GCG GAA GCC-3'; ZvFMOa double variant F354V/Y356A, forward, 5'-CGC CCG CAC GTC ATG GCT AAC CGT CAG TGG AAG-3'; reverse, 5'-CTT CCA CTG ACG GTT AGC CAT GAC GTG CGG GCG-3' (the underlined codons indicate the sites of mutagenesis). The template plasmid was digested with the restriction endonuclease DpnI prior to transformation of the reaction mixture into *Escherichia coli* XL1 Blue cells for subsequent screening and control sequencing.

2.2. Expression and purification of ZvFMOs

ZvFMO isoforms as well as their variants were heterologously expressed in *E. coli* BL21(DE3) cells harbouring the respective recombinant plasmids. Cells were grown at 210 K in LB medium containing 100 µg ml⁻¹ ampicillin (for pET-22b-derived vectors) or 30 µg ml⁻¹ kanamycin (for pET-28a-derived vectors). Protein expression was induced by adding 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) to the bacterial cultures at an OD₆₀₀ of 0.6 followed by incubation at 297 K for 16 h. The cells were lysed using an EmulsiFlex-C3 (Avestin, Mannheim, Germany) and subsequently centrifuged at 75 600g for 1 h. His₆-tagged recombinant proteins were purified from the resulting crude extract *via* immobilized metal ion-affinity chromatography (IMAC) using a HisTrap HP 5 ml column and an ÄKTApurifier FPLC system (both from GE Healthcare, Freiburg, Germany). The lysis and equilibration buffer was composed of 50 mM NaH₂PO₄ pH 8.0, 300 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoethanol, 200 µM phenylmethanesulfonyl fluoride (PMSF). Proteins were eluted *via* a linear gradient over five column volumes to a final concentration of 500 mM imidazole. Eluted target protein fractions were pooled and subsequently subjected to size-exclusion chromatography (HiLoad 16/60

Superdex 200 pg, GE Healthcare) with a running buffer composed of 20 mM glycine pH 9.0, 200 mM NaCl, 5 mM DTT. Purified protein was concentrated to approximately 6 mg ml⁻¹ and stored in 20 mM glycine pH 9.0, 1 mM DTT at 193 K.

2.3. SEC-MALS analysis

In order to determine the oligomeric states of the purified ZvFMO proteins, HPLC-based size-exclusion chromatography (SEC) with refractive-index (RI) and multi-angle laser light-scattering (MALS) detectors was performed. Protein solutions were diluted to 4 mg ml⁻¹ in phosphate-buffered saline (PBS) buffer pH 7.0 prior to SEC-MALS analysis. The HPLC-SEC system consisted of an online vacuum degasser (S 8515; SRI Instruments Europe), a high-pressure quaternary pump (G1311A; Agilent Technologies), a manual injector (7725i; Rheodyne), a pre-column (WTC-030N5G, 4.6 × 50 mm; Wyatt Technology), a SEC column (WTC-030N5, 4.6 × 300 mm; Wyatt Technology), a MALS detector (mini-DAWN TREOS, λ = 658 nm; Wyatt Technology) and a RI detector (G1362; Agilent Technologies). Data acquisition and processing were carried out using the ASTRA software (Wyatt Technology). The mobile phase for the SEC-MALS analysis was PBS (adjusted to pH 7.0 with phosphoric acid) and the flow rate was 0.4 ml min⁻¹. Additionally, the SEC column was calibrated with a protein-standard mixture containing bovine thyroglobulin, bovine gamma globulin, chicken ovalbumin and bovine ribonuclease A. Retention times were correlated with molecular weight and used to determine the oligomeric states of the ZvFMOs.

2.4. Crystallization

Initial crystallization hits were identified by high-throughput screening performed at the SPC Facility at EMBL Hamburg. Crystallization experiments for the refinement of the initial conditions were carried out using the hanging-drop vapour-diffusion method at 291 K. ZvPNO crystals were obtained by mixing protein solution (5.6 mg ml⁻¹ in 20 mM glycine-NaOH pH 9.0, 1 mM DTT) with an equal amount of precipitant solution [20 mM Tris-HCl pH 7.0, 200 mM MgCl₂, 15% (w/v) PEG 3350]. In order to generate crystals of ZvPNO in complex with the oxidized NADP cofactor, 1 mM NADP⁺ was added to the precipitant solution for co-crystallization. Crystals of wild-type ZvFMOa were obtained by mixing the protein solution (6.5 mg ml⁻¹ in 20 mM glycine-NaOH pH 9.0, 1 mM DTT) with an equal amount of precipitant solution [20 mM bis-tris propane-NaOH pH 7.5, 200 mM NaNO₃, 22% (w/v) PEG 3000, 1 mM TMA-HCl, 1 mM NADP⁺]. Drops were equilibrated against reservoir solution. Crystals appeared after 2–4 d in the form of mostly rod-shaped clusters or displaying multiple lattices. A few single crystals could be isolated and were briefly equilibrated in cryoprotectant solution prior to flash-cooling in liquid nitrogen. ZvPNO-FAD crystals were equilibrated in 3.2 M trimethylamine N-oxide (TMAO) for cryoprotection. Crystals of ZvPNO in complex with oxidized NADP cofactor as well as ZvFMOa crystals

research papers

Table 1
Data collection and processing.

Values in parentheses are for the highest resolution shell.

Data set	ZvPNO-FAD	ZvPNO-FAD-NADP ⁺	ZvFMOa
PDB code	5nmw	5nmx	
Data collection			
Diffraction source	P14, PETRA III, EMBL Hamburg	P14, PETRA III, EMBL Hamburg	P14, PETRA III, EMBL Hamburg
Wavelength (Å)	0.976261	0.976200	0.976300
Temperature (K)	100	100	100
Detector	PILATUS 6M	PILATUS 6M	PILATUS 6M
Crystal-to-detector distance (mm)	395.2	320.1	463.0
Rotation range per image (°)	0.1	0.1	0.1
Exposure time per image (s)	0.1	0.1	0.1
Space group	<i>P1</i>	<i>P1</i>	<i>C222₁</i>
<i>a</i> , <i>b</i> , <i>c</i> (Å)	73.9, 76.3, 80.9	74.1, 76.1, 81.7	89.6, 212.2, 166.8
α , β , γ (°)	72.0, 81.5, 81.2	71.8, 81.6, 82.0	90, 90, 90
Resolution range [†] (Å)	76.46–1.89 (1.92–1.89)	77.13–1.60 (1.86–1.60)	89.53–3.00 (3.16–3.00)
Total No. of reflections [†]	268372	1034819	181082
No. of unique reflections [†]	119135	213835	32199
Completeness [†] (%)	90.2 (62.8)	96.1 (47.5)	99.8 (99.9)
Multiplicity [†]	2.3	4.9	5.6
$\langle I/\sigma(I) \rangle$ [†]	8.8 (2.0)	39.3 (2.3)	5.9 (0.7)
Mean <i>I</i> half-set correlation $CC_{1/2}$ [†]	99.5 (62.2)	99.7 (43.3)	98.9 (30.4)
R_{merge} [†]	0.07 (0.64)	0.08 (0.83)	0.28 (2.38)
R_{meas} [†]	0.09 (0.83)	0.09 (1.02)	0.35 (2.93)
$R_{\text{p.i.m.}}$ [†]	0.06 (0.52)	0.04 (0.58)	0.20 (1.69)
Matthews coefficient (Å ³ Da ⁻¹)	2.16	2.19	2.71
Solvent content (%)	43.18	43.77	54.61
No. of molecules in asymmetric unit	4	4	3
Refinement			
Resolution (Å)	76.46–1.89 (1.94–1.89)	77.13–1.60 (1.64–1.60)	
R_{work} [‡] (%)	19.1 (27.2)	17.4 (31.6)	
R_{free} [‡] (%)	21.9 (33.1)	20.2 (35.4)	
<i>MolProbity</i> score [§]	1.70	2.05	
No. of non-H atoms			
Protein	13239	13310	
FAD cofactor	212	212	
NADP ⁺ cofactor	—	192	
Mg ²⁺	2	3	
Water	1207	1572	
Total	14660	15289	
Overall <i>B</i> factor from Wilson plot [¶] (Å ²)	35.8	27.0	
Average <i>B</i> factor ^{††} (Å ²)			
Protein (main chain)	29.7	18.3	
Protein (side chain)	32.3	22.9	
Protein (whole chain)	31.0	20.7	
FAD cofactor	25.9	12.8	
NADP ⁺ cofactor	—	29.4	
Mg ²⁺	34.9	23.4	
Water	35.0	28.6	
Ramachandran plot ^{‡‡}			
Favoured (%)	96.4	96.2	
Allowed (%)	3.6	3.7	
Outliers (%)	0.0	0.1	
R.m.s. deviations ^{‡‡}			
Bonds (Å)	0.006	0.023	
Angles (°)	1.000	2.305	

[†] Values as provided by *AIMLESS* after data processing, merging and scaling. [‡] Calculated by *REFMAC5* (Murshudov *et al.*, 2011). [§] Calculated using the *MolProbity* server (Chen *et al.*, 2010). [¶] Calculated using *SFHECK* (Vaguine *et al.*, 1999). ^{††} Calculated using *BAVERAGE*. ^{‡‡} Calculated using *RAMPAGE*. All programs used here are implemented within the *CCP4* program package (Winn *et al.*, 2011).

were equilibrated in 67% (*v/v*) reservoir solution in deionized water supplemented with a final concentration of 27.4% (*w/v*) PEG 3000.

2.5. Co-crystallization and soaking experiments

Co-crystallization experiments were carried out using the abovementioned crystallization conditions additionally supplied with 0.5–1 *mM* senecionine, 0.5–7.6 *mM* atropine, 3–33 *mM* homatropine hydrobromide or 14–154 *mM* atropine sulfate. The same PA substrate solutions and concentrations were used for soaking experiments. Soaking was performed by transferring crystals into the respective cryoprotectant solution (supplied with different amounts of substrates) for 5 min at 291 K before flash-cooling in liquid nitrogen.

2.6. ZvFMO activity assay

The activity of ZvFMO was determined photometrically using an Ultrospec 2100 pro (GE Healthcare, Freiburg, Germany). All substances were dissolved or diluted in glycine buffer (100 *mM* glycine–NaOH, pH 9.0). 340 μl of diluted enzyme (0.2 mg ml^{-1}) was incubated in a UV cuvette for 1 min at 303 K. 40 μl of 2 *mM* NADPH solution was added and incubated at 30°C for another minute. The enzymatic reaction was started by adding 20 μl of monocrotaline solution (2 mg ml^{-1}). NADPH turnover was monitored at 340 nm and directly correlated to *N*-oxygenation of monocrotaline.

2.7. Data collection, structure determination and representation

X-ray diffraction data were collected at 100 K on MX beamline P14, EMBL/DESY PETRA III (Hamburg, Germany) equipped with a PILATUS 6M detector. Diffraction data were indexed and integrated using the *XDS* software (Kabsch, 2010). Space-group determination, data scaling and merging were performed by the *AIMLESS* software as part of the *CCP4* program suite (Winn *et al.*, 2011), while applying the free-*R* flag to 5% of reflections. The structure of ZvPNO without cofactors was solved by molecular replacement (MR) using the *MOLREP* software as implemented in the *CCP4* program suite. The starting model for MR was built based on the bacterial FMO from *Methylophaga aminisulfidivorans* (PDB entry 2xve, subunit A,

research papers

33.3% sequence identity; Cho *et al.*, 2011) using *MODELLER* on the *HHpred* server (Söding *et al.*, 2005; Söding, 2005). For the rotation and translation search the model was manually trimmed to omit less well defined loop regions. The *MOLREP* scores for stepwise positioning of four molecules within the asymmetric unit were 0.136, 0.164, 0.200 and 0.224. The resulting ZvPNO model was iteratively completed by alternating refinement steps using *REFMAC5* (Murshudov *et al.*, 2011) as well as manual inspection, modification and insertion of cofactors and water molecules using *Coot* (Emsley & Cowtan, 2004; Emsley *et al.*, 2010). The refined model was used as the starting model for subsequent data sets from ZvFMO isoforms and variants. Refinement statistics are summarized in Table 1. The final model analysis, imaging and ray tracing were performed using *PyMOL* (v.1.8; Schrödinger).

3. Results

3.1. Expression, purification, crystallization and data collection

All ZvFMO isoforms were overproduced in *E. coli* and purified *via* FPLC. The enzymes eluted as stable, dimeric, 421–425-residue proteins (depending on the isoform) with a characteristic flavoprotein absorption maximum at 450 nm. The oligomeric state was separately confirmed by SEC-MALS analysis. All isoforms were subjected to crystallization trials, yet only ZvPNO yielded high-quality crystals that were suitable for X-ray diffraction experiments and subsequent structure determination. Crystal structures of the ZvPNO–FAD and ZvPNO–FAD–NADP⁺ complexes were determined to resolutions of 1.9 and 1.6 Å, respectively. 12 C-terminal residues (including the hexahistidine affinity tag) as well as residues 216–223 could not be modelled owing to the absence of appropriate electron density. The electron-density maps for the rest of the model were of high quality. Co-crystallization and soaking experiments with low substrate concentrations

yielded crystals that were suitable for structure determination. However, no additional electron density could be observed within the active site which could be interpreted as substrate. Using higher amounts of different substrates for co-crystallization resulted in the formation of amorphous precipitate rather than protein crystals. Crystals which were soaked with higher substrate concentrations immediately showed cracks when transferred to the soaking solution and began to dissolve after a few seconds. Protein crystals as well as data sets were also obtained for the isoform ZvFMOa, but owing to poor crystal quality and the resulting poor data statistics no reliable model could be determined for this isoform with regard to specific side-chain arrangements. However, the obtained diffraction data were sufficient for MR phasing, C^α tracing and analysis of the crystal packing as well as the dimeric arrangement. No crystallization condition was identified for the ZvFMOc isoform, even after screening using ZvPNO crystals for cross-seeding.

3.2. Overall structure of ZvPNO

ZvPNO is composed of two structural domains, each of which exhibits a dinucleotide-binding Rossmann fold crucial for binding either the FAD or the NADP cofactor. While FAD is tightly bound to the large structural domain composed of amino-acid residues 1–155 and 260–425, NADP⁺ is more loosely bound to the small domain (residues 156–259), where it can easily be exchanged for its reduced equivalent. The substrate-binding site is located within the cleft formed by the two domains, in direct proximity to the flavin moiety of the FAD, the ribose of NADP⁺ and the highly conserved amino-acid residue Asn66 (Fig. 2).

3.3. FAD binding

The FAD cofactor is deeply anchored within the large structural domain *via* the formation of hydrogen bonds to

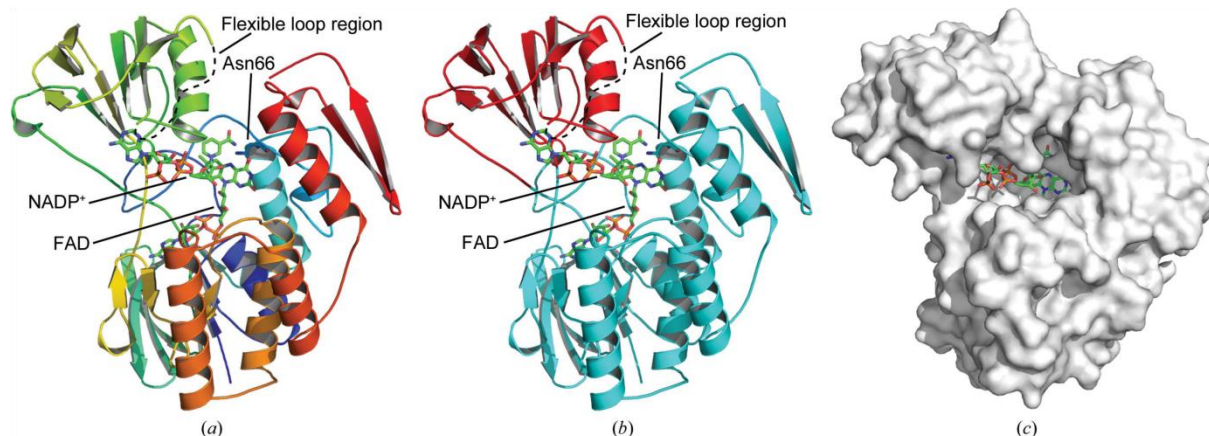


Figure 2

Representation of the overall structure of ZvPNO (PDB entry 5nmx) with the FAD and NADP⁺ cofactors as well as Asn66 depicted as stick models. (a) Cartoon representation of one ZvPNO subunit. The protein backbone is rainbow-coloured from the N-terminus (blue) to the C-terminus (red). The flexible loop region between residues 216 and 223 could not be traced. (b) Colour-coded structural domains. Cyan, large FAD-binding domain (residues 1–155 and 260–425); red, small NADP⁺-binding domain (residues 156–259). (c) Surface representation with visible substrate-entrance path leading towards the active site in front of the isoalloxazine-ring system of FAD.

research papers

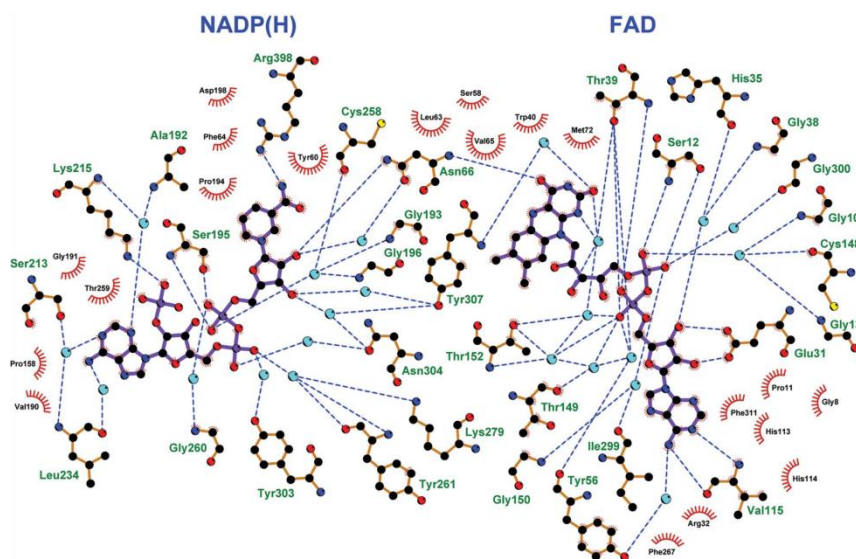


Figure 3

Cofactor-binding site of ZvPNO with bound FAD and NADP⁺. The cofactors are shown in ball-and-stick representation; the bonds are indicated in purple. The protein residues are represented with side chains in ball-and-stick representation; the bonds are indicated in yellow. Hydrogen bonds are shown as blue dashed lines, and the spoked arcs represent protein residues that form hydrophobic interactions with the cofactors. The cyan spheres indicate water molecules which provide bridged hydrogen bonds between amino-acid residues and the cofactors. For clarity, the lengths of the hydrogen bonds are not given. The representation was derived from an analysis with *LigPlot*⁺ (Laskowski & Swindells, 2011).

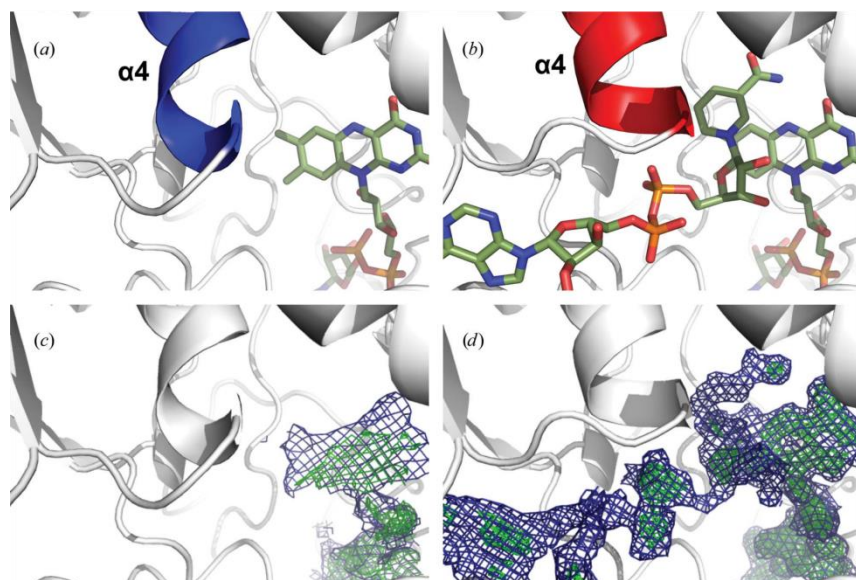


Figure 4

Conformational adjustment of helix $\alpha 4$ upon binding of the NADP cofactor. (a, b) Cofactors are represented as stick models. Helix $\alpha 4$ is colour-coded according to its conformational states: (a) blue, without NADP cofactor (PDB entry 5nmw); (b) red, with bound NADP cofactor (PDB entry 5nmx). (c, d) Composite OMIT maps of cofactors: (c) electron density indicating the FAD cofactor in PDB entry 5nmw; (d) electron density indicating the FAD and NADP⁺ cofactors in PDB entry 5nmx. Blue, $2F_o - F_c$ map contoured at 1.0σ ; green, $F_o - F_c$ map contoured at 2.5σ . Composite OMIT maps were generated using *PHENIX* (Adams *et al.*, 2010). Simulated-annealing cycles were performed to remove model bias.

polar or charged amino-acid side chains (residues Ser12, Glu31 and Thr39) as well as N and O atoms of the protein backbone (residues Ser12, Thr39, Asn66 and Val115). It is further held in position by numerous coordinated water molecules and hydrophobic residues, which define a major part of the binding site for this cofactor (Fig. 3). The typical dinucleotide-binding motif GXXGXXG is represented by residues 8–13. Only the reactive flavin moiety extends towards the solvent-exposed cleft encircled by the two structural domains and is accessible to NADPH, molecular oxygen and substrates.

3.4. NADP⁺ binding

The NADP⁺ cofactor is bound to the GXXGXXG motif (residues 191–196) of the smaller structural domain in an extended conformation *via* hydrogen bonds to its diphosphate moiety. The 2' phosphate of NADP⁺ is coordinated by Lys223 and His351. Interactions between the nicotinamide moiety and Phe64 and Arg398 as well as between the ribose moiety and Asn66 further support the binding and positioning of the cofactor (Fig. 3). In comparison to the ZvPNO–FAD complex, the presence of NADP⁺ leads to a small conformational change of helix $\alpha 4$ such that it provides additional cofactor stabilization by interactions between a positive partial charge of the helix dipole and the diphosphate of NADP⁺ (Fig. 4). The conformational change in the preceding loop region is caused by a flipping alanine, which reduces steric hindrance when binding the NADP⁺ cofactor.

3.5. Active site and substrate-binding cleft

The active site of ZvPNO is located in direct proximity to the functional moieties of both dinucleotide cofactors in combination

research papers

with Asn66, which is highly conserved among FMOs. The ribose moiety of NADP⁺ and Asn66 are supposed to coordinate molecular oxygen and stabilize the C4a-hydroperoxy-FAD intermediate (Eswaramoorthy *et al.*, 2006; Olucha *et al.*, 2011; Hille *et al.*, 2013). This is supported by an elongated

electron-density feature found in this position in at least one chain of the ZvPNO-FAD-NADP⁺ crystal structure, which was interpreted as two neighbouring water molecules but could also account for molecular oxygen. The rest of the substrate-binding pocket is formed by Asn66, Leu67, Pro194, Phe383, Val386, Tyr307 and Tyr303–Ala305, while the entrance is further restricted by a loop composed of Val354, Met356, Asn357 and Gln359 (Fig. 5).

As in most enzymes, the binding pocket of ZvPNO is not static but exhibits some flexibility. Small conformational changes were observable on comparing different protein subunits within the unit cell, accounting for some degree of structural plasticity. The most prominent among them is helix $\alpha 8$ composed of Pro377–Lys392, which can be kinked to a certain degree. Thereby, the space available for substrate binding is partly occupied (Fig. 6). The flexibility of this part of helix $\alpha 8$ is further indicated by higher *B* factors. In addition, the electron density for the side-chain atoms of Leu67 reveals different rotamer conformations when comparing the four protein molecules comprising the asymmetric unit. Depending on the conformation of Leu67, the volume of the binding pocket at the reaction centre is slightly reduced or enhanced, respectively.

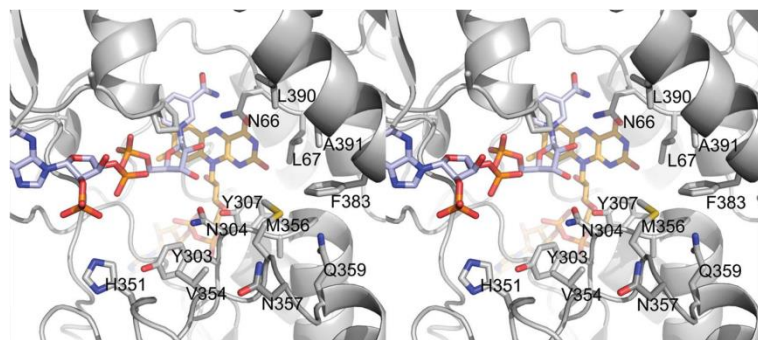


Figure 5

Stereo representation of the ZvPNO substrate-binding site. The protein backbone is shown in cartoon representation; cofactors and residues forming the active site and substrate-binding pocket are shown in stick representation. C atoms and backbone cartoon are in grey, N atoms are in blue, O atoms are in red and S atoms are in yellow; NAD⁺ C atoms are in light blue and FAD C atoms are in light orange.

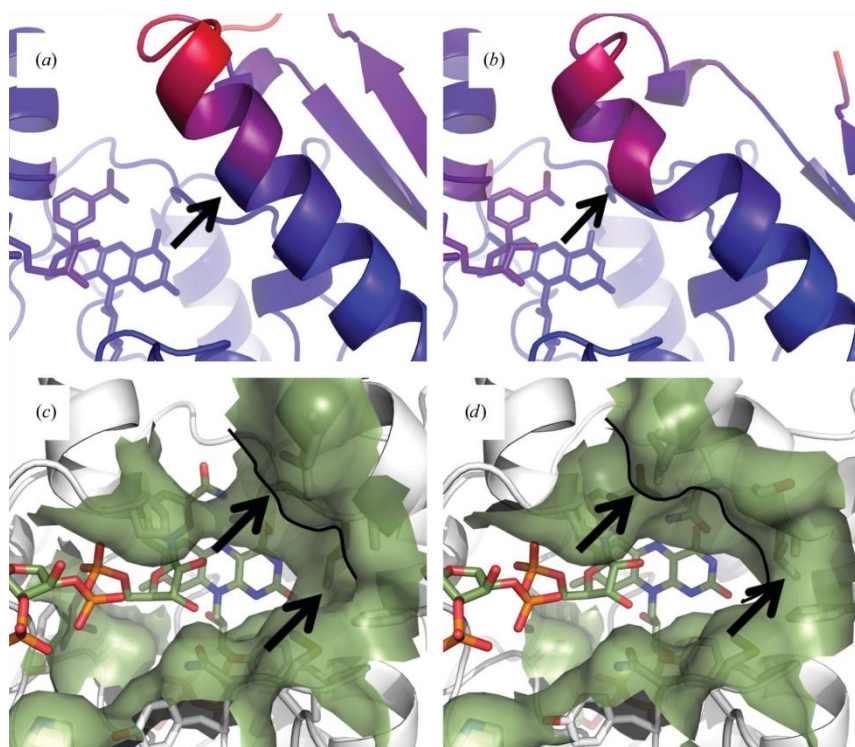


Figure 6

Binding-pocket limitations caused by conformational changes. (a, b) Cartoon representation of helix $\alpha 8$, colour-coded by *B*-factor distribution (spectrum from blue, 10 Å², to red, 65 Å²). Left, chain *D*; right, chain *C* as found in the asymmetric unit of PDB entry 5nmx. Arrows highlight conformational changes. (c, d) Surface representation of residues forming the binding pocket; arrows highlight conformational changes that regulate the size of the substrate-binding pocket.

3.6. Dimer formation

Despite the differences in sequence, the overall structures of the flavin-dependent mono-oxygenases are very similar. This indicates the need for high conservation of the three-dimensional structure for this type of reaction. For the biological assembly, a broader variation can be observed. Many functional assemblies of FMOs are known to be homo-oligomers (dimers, tetramers or hexamers), but the interfaces that contribute to oligomer formation seem to be

research papers

unpredictable. The published crystal structures of dimeric FMOs reveal quite different dimeric arrangements (Malito *et al.*, 2004; Leisch *et al.*, 2012; Eswaramoorthy *et al.*, 2006; Cho *et al.*, 2011), none of which are shared by the structures presented in this study. For all ZvFMO isoforms the major fraction was identified as a dimer by SEC-MALS analysis and calibrated SEC (Fig. 7; results are only shown for ZvPNO). Symmetrical homodimers with twofold symmetry can be observed within the ZvPNO crystal structure, however, with a new type of orientation (Fig. 8). Contacts are provided by hydrogen bonds and salt bridges between the large domain of one subunit (residues Trp47, Val55, Glu116, Trp153 and Arg264) and the small domain of the symmetry mate (residues Arg167, Glu239 and Trp240). With the exception of Val55, which is only involved *via* its backbone atoms, these residues are shared by all three ZvFMO isoforms. Of note, Arg167 is also part of the signature sequence (residues 166–176), which is highly conserved among FMOs (Fraaije *et al.*, 2002). Despite the few contact sites and the rather small interface area of approximately 900 Å² (calculated by *PDBePISA*; Krissinel & Henrick, 2007) a crystallographic artefact can be ruled out by the nonrefined crystal structure of ZvFMOa. It exhibits the

same dimeric arrangement of subunits, but a different crystal packing of these dimers. The substrate-binding sites of the two enzyme molecules forming the biologically active ZvFMO dimer are not occupied by the interface area and are fully solvent-accessible. The two active sites point in opposite directions, which enhances the possibility of cofactor and substrate recruitment in a rather orientation-independent manner. Based on this new dimeric arrangement, it seems that the relative orientation of the protein subunits is not crucial for the enzymatic reaction.

3.7. Sequence alignment of ZvFMO isoforms

All ZvFMO isoforms share high sequence identity (77.7% between ZvPNO and ZvFMOa; 77.1% between ZvPNO and ZvFMOc; 82.8% between ZvFMOa and ZvFMOc). Still, their specific enzyme activities significantly diverge from each other, probably owing to differences in the substrate-binding pocket or substrate-entry path (Fig. 9). In contrast, the sequence identities of ZvPNO compared with its modelling templates from *M. aminisulfidivorans* (34% sequence identity; r.m.s.d. of 2.5 Å) and *Schizosaccharomyces pombe* (27% sequence identity; r.m.s.d. of 2.3 Å) are rather low.

3.8. ZvFMOa variants

Variants of the less active isoform ZvFMOa were generated to identify amino-acid residues which have an impact on the specific activity of this enzyme and are likely to be involved in substrate binding or turnover. We investigated residues which are part of the substrate-binding pocket and differed between the ZvPNO and ZvFMOa isoforms. Based on initial docking experiments with the crystal structure of ZvPNO and different PA substrates, Tyr307 was identified as the most promising part of the binding pocket, potentially forming hydrogen bonds to a variety of PA substrates. Additionally, when comparing the crystal structures of ZvPNO with a homology model of ZvFMOa, a tyrosine residue (Tyr356) blocking the substrate entrance in the latter model was observed. Thus, the ZvFMOa variant F307Y and a double variant F307Y/Y356A

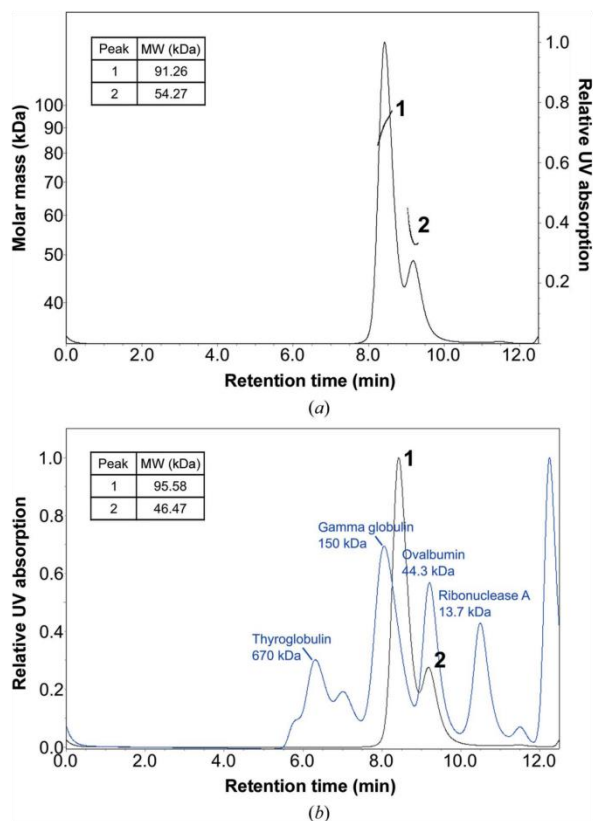


Figure 7
SEC-MALS analysis of ZvPNO. (a) SEC-MALS analysis of ZvPNO with mean molecular weights calculated by the *ASTRA* software for observable peaks. (b) Calibrated SEC analysis with molecular weights calculated from the retention times of observable ZvPNO peaks.

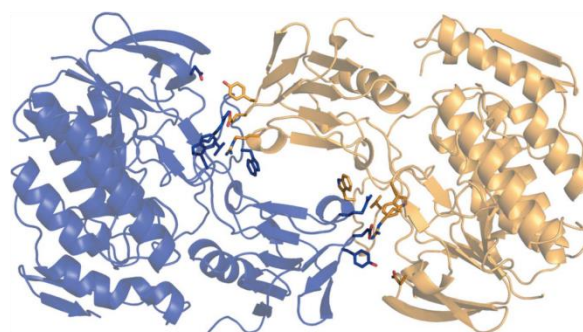


Figure 8
Dimeric arrangement of ZvPNO. Cartoon representation of two ZvPNO subunits (blue and yellow). Interface residues are depicted as stick models. The view is along the noncrystallographic twofold dyad. The surface-complementarity score was calculated to be 0.8 (Lawrence & Colman, 1993).

research papers

were investigated. In contrast to preliminary substrate-docking predictions, the ZvFMOa F307Y variant resulted in a decline in specific activity by a factor of 0.81 in comparison with wild-type ZvFMOa (Fig. 10), while the double variant exhibited a slight increase by a factor of 1.51. Consequently, a variant containing a single Y356A exchange was generated

and analysed. The specific activity of this variant was increased by a factor of 2.45 compared with the wild-type enzyme. Therefore, another variant (F354V) was created which likewise contained an amino-acid exchange within the loop limiting the entrance to the binding site. This exchange also resulted in an increased specific activity (by a factor of 2.75).

However, combining both favourable amino-acid exchanges does not result in a higher level of specific activity, but rather diminishes their individual effects. Still, the observed specific activity for this double variant is increased by a factor of 1.5 compared with the control experiment. The beneficial effect of a more accessible substrate-entry path might possibly be partly antagonized by a loss of substrate affinity when both bulky aromatic amino acids are exchanged for smaller substituents. Based on our identification of the flexible helix $\alpha 8$, which might have an influence on substrate binding and turnover, we also generated a P388S variant. However, the specific activity of this variant did not differ significantly from that of wild-type ZvFMOa, indicating that the presence of proline within the ZvFMOa helix does not interfere with its potential role in substrate turnover.

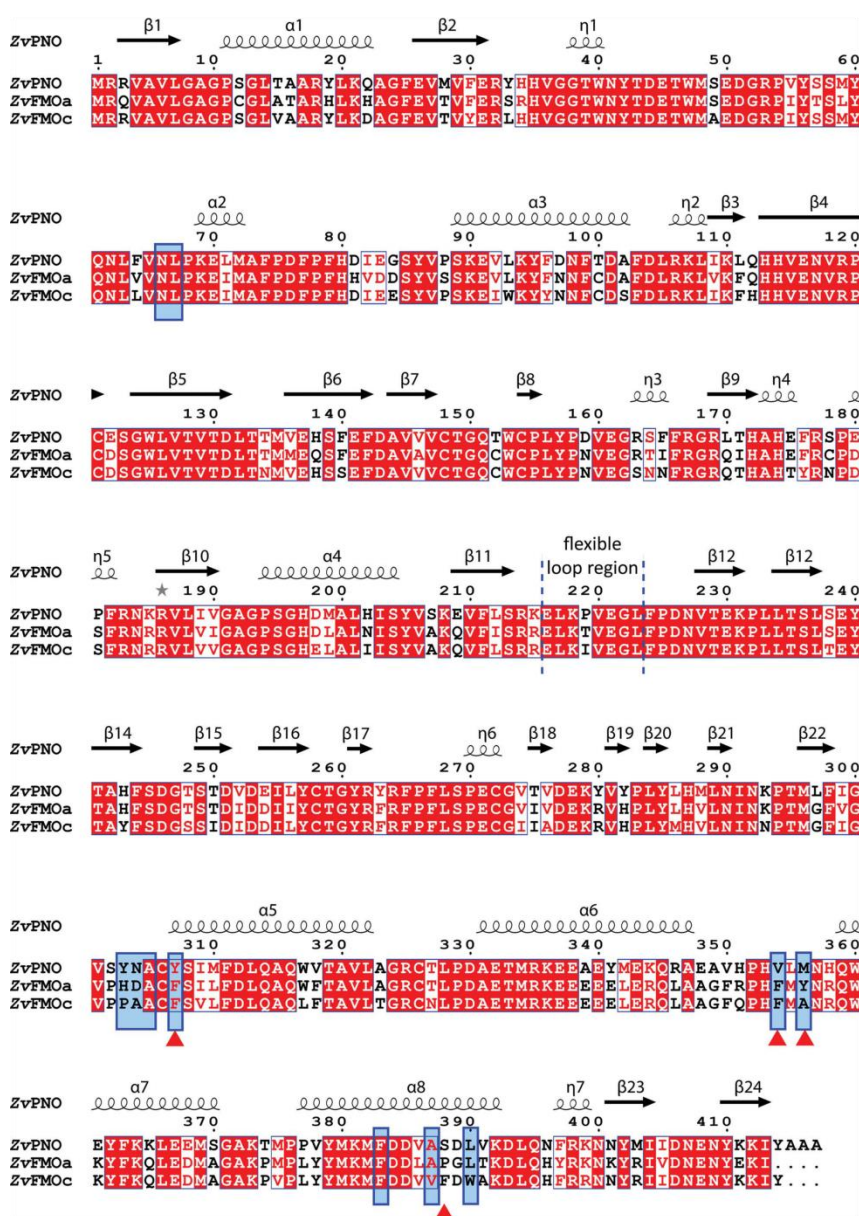


Figure 9

Structure-based sequence alignment of ZvFMO isoforms. Red background, identical amino acids; bold letters, similar amino acids; blue boxes, residues forming the predicted substrate-binding pocket. The multiple sequence alignment was performed with *Clustal Omega* (Sievers *et al.*, 2011); the final figure was prepared using *ESPrpt* (Robert & Gouet, 2014). The secondary structure was derived by *DSSP* (Kabsch & Sander, 1983) and is depicted above the sequences. Residues which were exchanged by site-directed mutagenesis are highlighted by a triangle below the sequence.

research papers

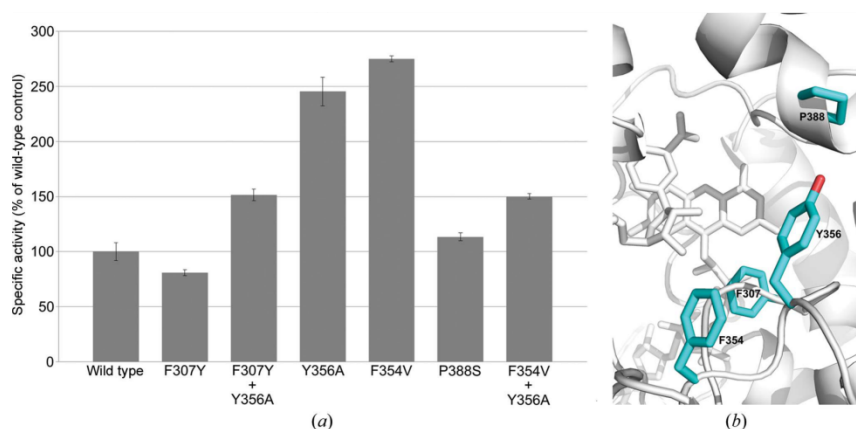


Figure 10
Specific activity of ZvFMOa variants towards monocrotaline. (a) The enzyme activities of the variants were determined photometrically in triplicate. (b) Cyan stick models show the residues in the binding pocket of wild-type ZvFMOa which were exchanged to create variants.

plasticity within this pocket can be envisaged. Substrates are presumably not firmly bound to the enzyme but are instantly oxygenated upon reaching the active site. These findings coincide with the broad PA substrate spectrum of ZvFMOs as well as the proposed cocked-gun mechanism for these enzymes. It might also be possible that an active C4a-hydroperoxy-FAD intermediate structure may preferably bind the PA ligand compared with the oxidized FAD structure as presented. These circumstances might be the reason for our unsuccessful co-crystallization and soaking experiments and the lack of structural evidence for an enzyme–substrate complex.

Still, the active site and substrate-binding cleft of ZvPNO show some flexibility, especially in helix $\alpha 8$, which either reduces or enhances the volume of the binding pocket depending on its bending angle. This flexibility might account for either one of two features: (i) it might contribute to an induced fit to best adapt to different PA substrates or (ii) it might serve as some kind of product-removal mechanism by mechanically pushing the oxygenated PA out of the active site.

Of note, only the space in direct proximity to the reactive flavin moiety is quite limited, while the rest of the substrate-binding pocket is largely solvent-exposed and exhibits few steric boundaries. Therefore, the small active site only accepts small compounds such as molecular oxygen, which is needed for the oxygenation reaction, and is probably selective for structural features that resemble a necine base, which is shared by all PA substrates. Otherwise, there seems to be very limited selectivity regarding accepted substrates apart from the steric hindrance present in the entrance site of the binding cleft. This is backed up by our findings that single amino-acid exchanges within the ZvFMOa sequence could significantly increase its specific activity towards one of the sterically most demanding PA substrates, monocrotaline.

Therefore, ZvFMOs are not specialists with high efficiency towards a single substrate, but rather generalists which accept a variety of compounds with similar features. This harbours a

great advantage for the host organism *Zonocerus*, which can easily adapt to a variety of PAs produced by food plants and thereby circumvent their chemical defence against herbivory as part of the so-called arms race between plants and their attackers (Pieterse & Dicke, 2007). Evolutionary pressure caused by ever-developing PAs in the host plants might be a reason for this organism to retain three isoforms of ZvFMOs with differences in their substrate-binding cleft in its genome, although two of them are far less active towards currently known PAs. A few mutations during the course of evolution might turn them into

highly active enzymes capable of oxygenating PAs, as our studies of ZvFMOa variants have already indicated. However, it cannot be ruled out that the real substrates for isoform A and isoform C have not yet been identified.

Owing to their structural and catalytic properties, the ZvFMO isoforms belong to the class B flavoprotein monooxygenases (van Berkel *et al.*, 2006). To date, our structural knowledge about this specific FMO subclass is solely based on FMO crystal structures from bacteria, yeast and fungi (Malito *et al.*, 2004; Eswaramoorthy *et al.*, 2006; Cho *et al.*, 2011; Mirza *et al.*, 2009; Olucha *et al.*, 2011; Yachnin *et al.*, 2012; Leisch *et al.*, 2012; Jensen *et al.*, 2012, 2014; Franceschini, Fedkenheuer *et al.*, 2012; Franceschini, van Beek *et al.*, 2012; Binda *et al.*, 2015; Setser *et al.*, 2014; Ferroni *et al.*, 2016; Romero *et al.*, 2016; Fürst *et al.*, 2017). Therefore, the structures presented here are the first to derive from a more highly developed organism and give insight into the evolutionary conservation of FMOs. Although displaying a unique dimer interface, the overall structure of the FMO subunits has remained largely unchanged throughout evolution. However, the detailed mechanisms of cofactor and substrate binding as well as product release from the active site, which are conducted by minor and major conformational changes within the active site, are still not fully understood. The identification of a kink within the flexible helix $\alpha 8$ presented in this work might contribute to a better future understanding of the complete machinery.

Because of its closer evolutionary relationship, the crystal structure of ZvPNO will serve as a more suitable template for the related human enzyme involved in the biotransformation of currently available xenobiotics.

Acknowledgements

We gratefully acknowledge access to the core facilities of the BiMo/LMB of Kiel University. We thank Felix Helfrich, Sebastian Krossa and Christina Hopf for helpful discussions

research papers

and support during data collection. Diffraction data were collected on beamlines P13 and P14 operated by EMBL at the PETRA III storage ring. We are grateful to the beamline staff for providing assistance in using the beamlines.

Funding information

The research leading to these results received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under BioStructX (grant agreement No. 283570). Additionally, we are grateful for access to the HTX crystallization facility by means of a grant from P-Cube and BioStructX at the EMBL Outstation Hamburg. Beamtime at P14 at the EMBL Outstation Hamburg was also funded by a BioStructX grant. This work was funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) grant OBI62/9-1.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Beatty, N. B. & Ballou, D. P. (1981). *J. Biol. Chem.* **256**, 4619–4625.
- Berkel, W. J. H. van, Kamerbeek, N. M. & Fraaije, M. W. (2006). *J. Biotechnol.* **124**, 670–689.
- Bernays, E., Edgar, J. & Rothschild, M. (1977). *J. Zool.* **182**, 85–87.
- Binda, C., Robinson, R. M., Martin Del Campo, J. S., Keul, N. D., Rodriguez, P. J., Robinson, H. H., Mattevi, A. & Sobrado, P. (2015). *J. Biol. Chem.* **290**, 12676–12688.
- Cashman, J. R. (1995). *Chem. Res. Toxicol.* **8**, 166–181.
- Cashman, J. R. (2001). *Enzyme Systems that Metabolise Drugs and Other Xenobiotics*, edited by C. Ioannides, pp. 67–93. Chichester: John Wiley & Sons.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- Cho, H. J., Cho, H. Y., Kim, K. J., Kim, M. H., Kim, S. W. & Kang, B. S. (2011). *J. Struct. Biol.* **175**, 39–48.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Eswaramoorthy, S., Bonanno, J. B., Burley, S. K. & Swaminathan, S. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 9832–9837.
- Ferromi, F. M., Tolmie, C., Smit, M. S. & Opperman, D. J. (2016). *PLoS One*, **11**, e0160186.
- Fraaije, M. W., Kamerbeek, N. M., van Berkel, W. J. & Janssen, D. B. (2002). *FEBS Lett.* **518**, 43–47.
- Franceschini, S., Fedkenheuer, M., Vogelaar, N. J., Robinson, H. H., Sobrado, P. & Mattevi, A. (2012). *Biochemistry*, **51**, 7043–7045.
- Franceschini, S., van Beek, H. L., Pennetta, A., Martinoli, C., Fraaije, M. W. & Mattevi, A. (2012). *J. Biol. Chem.* **287**, 22626–22634.
- Fu, P. P., Xia, Q., Lin, G. & Chou, M. W. (2004). *Drug Metab. Rev.* **36**, 1–55.
- Fürst, M. J., Savino, S., Dudek, H. M., Gómez Castellanos, J. R., Gutiérrez de Souza, C., Rovida, S., Fraaije, M. W. & Mattevi, A. (2017). *J. Am. Chem. Soc.* **139**, 627–630.
- Hao, D. C., Chen, S. L., Mu, J. & Xiao, P. G. (2009). *Genetica*, **137**, 173–187.
- Hartmann, T., Biller, A., Witte, L., Ernst, L. & Boppré, M. (1990). *Biochem. Syst. Ecol.* **18**, 549–554.
- Hartmann, T. & Ober, D. (2008). *Induced Plant Resistance to Herbivory*, edited by A. Schaller, pp. 213–231. Dordrecht: Springer.
- Hartmann, T. & Witte, L. (1995). *Alkaloids: Chemical and Biological Perspectives*, edited by S. W. Pelletier, Vol. 9, pp. 155–233. New York: Elsevier.
- Hille, R., Miller, S. & Palfey, B. (2013). Editors. *Handbook of Flavoproteins*. Berlin: De Gruyter.
- Jensen, C. N., Ali, S. T., Allen, M. J. & Grogan, G. (2014). *J. Mol. Catal. B Enzym.* **109**, 191–198.
- Jensen, C. N., Cartwright, J., Ward, J., Hart, S., Turkenburg, J. P., Ali, S. T., Allen, M. J. & Grogan, G. (2012). *Chembiochem*, **13**, 872–878.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Langel, D. & Ober, D. (2011). *Phytochemistry*, **72**, 1576–1584.
- Laskowski, R. A. & Swindells, M. B. (2011). *J. Chem. Inf. Model.* **51**, 2778–2786.
- Lawrence, M. C. & Colman, P. M. (1993). *J. Mol. Biol.* **234**, 946–950.
- Leisch, H., Shi, R., Grosse, S., Morley, K., Bergeron, H., Cygler, M., Iwaki, H., Hasegawa, Y. & Lau, P. C. (2012). *Appl. Environ. Microbiol.* **78**, 2200–2212.
- Malito, E., Alfieri, A., Fraaije, M. W. & Mattevi, A. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 13157–13162.
- Mirza, I. A., Yachnin, B. J., Wang, S., Grosse, S., Bergeron, H., Imura, A., Iwaki, H., Hasegawa, Y., Lau, P. C. K. & Berghuis, A. M. (2009). *J. Am. Chem. Soc.* **131**, 8848–8854.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Narberhaus, I., Theuring, C., Hartmann, T. & Dobler, S. (2003). *J. Comp. Physiol. B*, **173**, 483–491.
- Naumann, C., Hartmann, T. & Ober, D. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 6085–6090.
- Olucha, J., Meneely, K. M., Chilton, A. S. & Lamb, A. L. (2011). *J. Biol. Chem.* **286**, 31789–31798.
- Phillips, I. R. & Shephard, E. A. (2008). *Trends Pharmacol. Sci.* **29**, 294–301.
- Pieterse, C. M. & Dicke, M. (2007). *Trends Plant Sci.* **12**, 564–569.
- Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.
- Robinson, R., Badieyan, S. & Sobrado, P. (2013). *Biochemistry*, **52**, 9089–9091.
- Romero, E., Castellanos, J. R., Mattevi, A. & Fraaije, M. W. (2016). *Angew. Chem. Int. Ed.* **55**, 15852–15855.
- Sehlmeyer, S., Wang, L., Langel, D., Heckel, D. G., Mohagheghi, H., Petschenka, G. & Ober, D. (2010). *PLoS One*, **5**, e10435.
- Setser, J. W., Heemstra, J. R. Jr, Walsh, C. T. & Drennan, C. L. (2014). *Biochemistry*, **53**, 6063–6077.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D. & Higgins, D. G. (2011). *Mol. Syst. Biol.* **7**, 539.
- Söding, J. (2005). *Bioinformatics*, **21**, 951–960.
- Söding, J., Biegert, A. & Lupas, A. N. (2005). *Nucleic Acids Res.* **33**, W244–W248.
- Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* **D55**, 191–205.
- Wang, L., Beuerle, T., Timbilla, J. & Ober, D. (2012). *PLoS One*, **7**, e31796.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yachnin, B. J., Sprules, T., McEvoy, M. B., Lau, P. C. K. & Berghuis, A. M. (2012). *J. Am. Chem. Soc.* **134**, 7788–7795.
- Ziegler, D. M. (1993). *Annu. Rev. Pharmacol. Toxicol.* **33**, 179–199.
- Ziegler, D. M. (2002). *Drug Metab. Rev.* **34**, 503–511.

3.2 Human mitochondrial amidoxime reducing component (mARC): An electrochemical method for identifying new substrates and inhibitors

Palraj Kalimuthu, Antje Havemeyer, Bernd Clement, Christian Kubitza, Axel J. Scheidig and Paul V. Bernhardt

Electrochemistry Communications **2017**, *84*, 90-93.

DOI: 10.1016/j.elecom.2017.10.003

Despite their function as drug metabolizing enzymes, modulation of the NO pathway by reduction of *N*(ω)-hydroxy-L-arginine, detoxification of *N*-hydroxylated compounds and involvement in lipid metabolism, the physiological role of mARC enzymes remains largely elusive. High-throughput screenings are limited by the availability of low-yield recombinant soluble mARC proteins and their electron-delivering redox partner proteins cytochrome b_5 (Cyb5) and cytochrome b_5 reductase (Cyb5R). This three-component enzyme system needs to be heterologously expressed and purified separately prior to *in vitro* reconstitution and subsequent screening assays aiming for the identification of new mARC substrates or inhibitors. Furthermore, putative hits obtained by such assays need to be cross-validated in order to ensure they are indeed related to mARC activity rather than interaction with Cyb5 or Cyb5R.

In the following paper,¹⁰⁴ a novel electrochemical approach is described, which enables screening for new mARC substrates and inhibitors with less effort and lower amounts of recombinant protein. In this assay electrons are delivered directly to Cyb5, which is coupled to a gold electrode, and subsequently transferred to mARC. Since there is no need for recombinant Cyb5R using this approach, the effort spent in protein production and purification as well as cross-validation can be minimized. Substrates or inhibitors can be identified electrochemically by a voltammetric shift of a symmetric reversible Cyb5 response to a sigmoidal catalytic curve, which is due to an electrocatalytic cycle being established. This method was established for the model substrate benzamidoxime but can in principal be applied to any given potential mARC substrate or inhibitor.

I performed the expression and purification of soluble human mARC and Cyb5 proteins, which were subsequently used for the voltammetric assays performed by Dr. Palraj Kalimuthu.



Contents lists available at ScienceDirect

Electrochemistry Communications

journal homepage: www.elsevier.com/locate/elecom

Human mitochondrial amidoxime reducing component (mARC): An electrochemical method for identifying new substrates and inhibitors

Palraj Kalimuthu^a, Antje Havemeyer^b, Bernd Clement^b, Christian Kubitzka^c, Axel J. Scheidig^c, Paul V. Bernhardt^{a,*}^a School of Chemistry and Molecular Biosciences, University of Queensland, Brisbane 4072, Australia^b Pharmazeutisches Institut, Universität Kiel, Gutenbergstrasse 76, 24118 Kiel, Germany^c Zoologisches Institut – Strukturbiologie, Universität Kiel, Am botanischen Garten 11, 24118 Kiel, Germany

ARTICLE INFO

Keywords:
mARC
Molybdenum
Enzyme
Voltammetry
Cytochrome

ABSTRACT

As recently as 2006 the mitochondrial amidoxime reducing component (mARC) was identified as the fourth and last Mo enzyme present in humans. Its physiological role remains unknown. mARC is capable of reducing a variety of *N*-hydroxylated compounds such as amidoximes to their corresponding amidine and there is considerable interest in this enzyme from a pharmaceutical perspective. mARC is a target for *N*-hydroxylated prodrugs that may be reductively activated intracellularly to release potent drugs such as cationic amidinium ions, which exhibit a broad spectrum of activity as antithrombotics and against various bacteria and parasites. In the quest for a rapid screen of new mARC substrates and inhibitors we present an electrochemical method which utilizes the natural electron partner of mARC, cytochrome *b*₅, coupled to an electrochemical electrode. Mediated electron transfer from the electrode via cytochrome *b*₅ to mARC results in a catalytic current in the presence of substrate.

1. Introduction

In 2006, the mitochondrial amidoxime reducing component (mARC) was the fourth, and last, Mo-dependent enzyme identified in humans [1]. Like the well-studied human sulfite oxidase, mARC is also located in mitochondria, while the remaining two Mo enzymes (xanthine oxidase and aldehyde oxidase) are found in the cytosol [2]. In addition to its mitochondrial localization, mammalian mARC has been detected in peroxisomes as well [3]. Although no crystal structure of a mARC enzyme has been reported, sequence analysis and spectroscopy have shown that it belongs in the sulfite oxidase family [4] according to the enduring Mo enzyme classification originally proposed by Hille [5] (Scheme 1).

The 35 kDa mARC enzyme bears no cofactors other than the Mo active site (Scheme 1) and, in its reduced Mo^{IV} form, catalyses the reduction of *N*-hydroxylated compounds including, but not limited to, amidoximes and hydroxylamines to their corresponding amidines or amines [4]. mARC is found in two isoforms (mARC1 and mARC2), which share 80% sequence similarity [6], but, depending on the species and tissue, only one mARC protein is expressed predominately [7,8]. The physiological function of both mARC1 and mARC2 in humans and other eukaryotes is still unknown. The enzymes appear to be involved

in metabolic detoxification reactions [7,8], the NO pathway by aerobic reduction of the NO-precursor *N* ω -hydroxy-L-arginine or anaerobic reduction of nitrite to NO. [9] Moreover, mARC is implicated in energy and lipid metabolism as well as metabolic disorders as diabetes mellitus [4,10].

There is much interest in human mARC from the perspective of drug metabolism. By analogy with the cytochromes P450, which frequently hydroxylate and deactivate xenobiotic drug-like compounds, mARC is able to reduce *N*-hydroxylated compounds in a complementary way. This may be turned to advantage in drug design [11] through the administration of *N*-hydroxylated prodrugs which, after reductive dehydroxylation in vivo by mARC, can release active drugs already delivered to the target cell. This is particularly useful for very active amidines (Scheme 2), which, due to their high basicity, are protonated at physiological pH (to give the amidinium conjugate acid) and unable to cross the cell membrane.

However, in their charge-neutral amidoxime form they are sufficiently hydrophobic to penetrate the cell by passive diffusion and act as a ‘Trojan horse’ that is activated intracellularly by mARC catalyzed reduction. In other cases toxic *N*-hydroxylated metabolites formed by P450 enzymes are reduced by mARC to their corresponding dehydroxylated parent compound. Due to its involvement in many

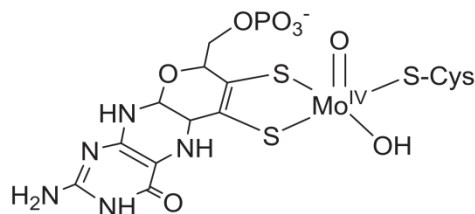
* Corresponding author.

E-mail address: p.bernhardt@uq.edu.au (P.V. Bernhardt).<http://dx.doi.org/10.1016/j.elecom.2017.10.003>

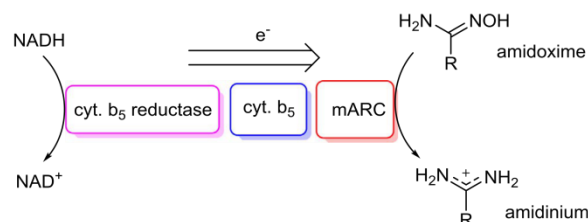
Received 26 August 2017; Received in revised form 27 September 2017; Accepted 3 October 2017

Available online 05 October 2017

1388-2481/© 2017 Elsevier B.V. All rights reserved.



Scheme 1. Active site of mARC in its Mo^{IV} form (overall charge of complex ion not shown).



Scheme 2. Electron flow during mARC catalysis.

activation, deactivation and detoxification reactions, mARC is a new and essential addition to the list of known drug metabolizing enzymes [11].

The overall electron transfer sequence for mARC catalysis is illustrated in Scheme 2. Reducing equivalents supplied by NADH are passed to the flavoprotein cytochrome *b*₅ reductase, then, one at a time, to the heme protein cytochrome *b*₅ before being relayed to mARC to prime the Mo^{IV} active site for substrate reduction [8]. The natural substrate (or substrates) for mARC1 and mARC2 remain unknown. This is complicated by the fact that mARC is able to reduce a variety of organic compounds with N-O functionality [4].

Being an Mo-dependent oxidoreductase, mARC may be examined by electrochemistry [12] and that is the focus of this investigation. In the quest for a rapid and high throughput assay of potential drug-like mARC substrates and inhibitors, we present a new cyclic voltammetry (CV) methodology that utilizes miniscule (picomole) quantities of enzyme and potentially may screen many new drug candidates rapidly and efficiently.

2. Materials and methods

A gold working electrode (BAS Inc.) was chemically modified with the bifunctional thiol 3-mercaptopropionic acid (MSA) following a literature protocol [13]. A Pt wire counter electrode and Ag/AgCl reference electrode were incorporated into a BAS C3 cell stand and attached to a BAS100 potentiostat. All electrochemical experiments were carried out under an atmosphere of dinitrogen. Human mARC1 and cytochrome *b*₅ were recombinantly expressed in *E. coli* and purified as described [6]. A mixture of 4 μ L of recombinant human cytochrome *b*₅ (17 μ M) and 2 μ L of 0.25% chitosan solution (in 1% acetic acid) was pipetted onto the conducting surface of an inverted, freshly prepared Au/MSA working electrode and allowed to evaporate to a film at 4 $^{\circ}$ C. After that 2 μ L of mARC1 (250 μ M) was dispensed onto the same Au/MSA/chitosan-cyt.*b*₅ electrode and again allowed to evaporate over 1 h at 4 $^{\circ}$ C to a thin film. To prevent protein loss the enzyme modified electrode surface was carefully covered with a perm-selective dialysis membrane (molecular weight cut off 3500 Da), pre-soaked in water. The dialysis membrane was pressed onto the electrode with a Teflon cap and fastened to the electrode with a rubber O-ring to prevent leakage of the internal membrane solution. The resulting enzyme modified electrode was stored at 4 $^{\circ}$ C in 100 mM phosphate buffer (pH 6.0) when not in use. The proteins cytochrome *b*₅ and mARC were confined to a thin layer

beneath the membrane with chitosan while the substrates were able to diffuse across the membrane. Phosphate buffer (100 mM) was used for experiments at pH 6. For pH-dependent experiment, the mixture of buffers (25 mM citric acid buffer pH 3.0–6.2, 25 mM Bis-Tris buffer pH 5.8–7.2, 25 mM Tris buffer pH 7.0–9.0 and 25 mM CHES buffer pH 8.6–10.0) were used and the desired pH was obtained by titration with dilute acetic acid or NaOH. All solutions were prepared with ultrapure water (resistivity 18.2 M Ω ·cm).

3. Results and discussion

We employed the native mARC1 partner, outer mitochondrial membrane cytochrome *b*₅ (cyt. *b*₅), as a mediator of electron transfer between a chemically modified Au working electrode and the enzyme. This enables catalytic electrochemistry to take place at the relatively high potential of the cyt. *b*₅ redox couple thus avoiding non-specific reduction of interfering species at lower potentials. The working electrode replaces NADH and cyt. *b*₅ reductase in Scheme 2.

UV-vis monitored spectroelectrochemistry of cyt. *b*₅ was conducted at pH 8 using a published set of electron transfer mediators [14,15] (see Supporting Information Fig. S1 for details) yielding a Fe^{III/II} redox potential of $-50 (\pm 5 \text{ mV})$ vs SHE which is comparable with previous investigations of human and other vertebrate cyt. *b*₅ proteins [16–18].

Cyclic voltammetry of cyt. *b*₅ (Fig. 1) was achieved using a gold working electrode chemically modified with a self-assembled monolayer (SAM) of mercaptosuccinic acid (MSA). As the overall goal was electrochemically driven catalysis of mARC1, we trapped both cyt. *b*₅ and mARC1 under a semi-permeable membrane (MW cutoff 3.5 kDa) which enabled small volumes (microliters) of protein solution (micromolar concentration) to be employed. An optimal cyt. *b*₅ response was obtained with the promoter chitosan (poly-D-glucosamine) also present. Chitosan is electro-inactive but greatly enhances the cyt. *b*₅ current response without altering the redox potential relative to experiments carried out in its absence (Supporting Information Fig. S2). The protonation constants of the glucosamine monomers within chitosan lie in the range pK_a 6.5–7 [19,20] so that at all pH values investigated here chitosan bears a net positive charge.

A stable and reversible ferric/ferrous cyt. *b*₅ voltammetric response is apparent in Fig. 1 within the range $4.56 < \text{pH} < 6.55$ and the redox potential is almost pH independent (-1 mV at pH 4.56 and -19 mV vs SHE at pH 6.55). It has been noted previously that the

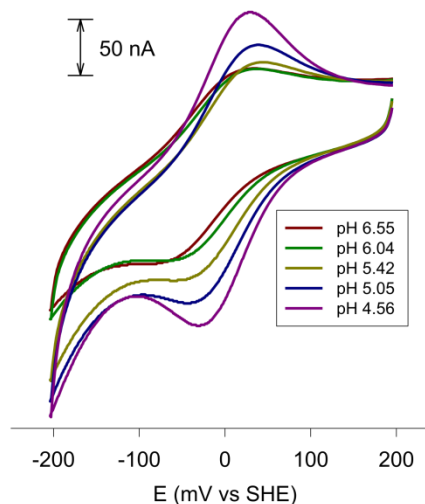


Fig. 1. CVs obtained at a Au/MSA/chitosan-cyt.*b*₅/mARC1 electrode at various pH values (0.1 M mixed buffer solution titrated with AcOH/NaOH) and a scan rate of 5 mV s⁻¹.

apparent redox potential of cyt. b_5 is sensitive to additives such as polylysine and Mg^{2+} ions [16–18] so the modest (~ 30 – 50 mV) anodic shift of the CV redox potential relative to the spectroelectrochemical titration (Supporting information Fig. S1) is consistent with these reports. The scan rate dependence of the current is indicative of a diffusion controlled voltammetric response (Supporting Information Fig. S3) although cyt. b_5 is confined to a small volume beneath the membrane covering the Au working electrode. No CV response from mARC1 was obtained under these conditions.

The peak currents (i_p) increase by a factor of 1.7 going from pH 6.6 to pH 4.6 (Fig. 1) and also the peak-to-peak separation decreases, which are both indicative of an enhanced heterogeneous electron transfer rate. The current variation in Fig. 1 is most likely related to an increased positive charge on the chitosan promoter as the pH is lowered, which enhances electrostatic attraction to the negatively charged cyt. b_5 protein (pI 4.3) and to the SAM surface carboxylates. Below pH 4.5, the cyt. b_5 signal vanishes. This is most likely a combination of the carboxylate groups within the SAM being protonated and the protein itself assuming a net positive charge which neutralises the cation-anion attractive forces between the SAM, chitosan and cyt. b_5 . Similarly, above pH 7 the cyt. b_5 CV response is lost apparently due to deprotonation of chitosan and loss of the same favourable electrostatic forces at the electrode surface.

Upon addition of benzamidoxime, a known mARC1 substrate, the symmetric reversible cyt. b_5 response at pH 6.0 (Fig. 2a) transforms into a sigmoidal waveform (Fig. 2b). This is due to an electrocatalytic cycle being established. At potentials below the ferric/ferrous cyt. b_5 redox couple, a feedback loop is established whereby (ferrous) cyt. b_5 is re-oxidised by (benzamidoxime-oxidised) mARC1 and the ensuing ferric cyt. b_5 is again reduced electrochemically. As long as benzamidoxime is present the cathodic current is sustained and the sigmoidal waveform is indicative of an electrochemical steady state. Smaller additions of benzamidoxime reveal a more gradual transition from the reversible transient voltammetry of cyt. b_5 to the ultimate sigmoidal catalytic wave (Supporting Information Fig. S4). In the absence of mARC1 the CV of cyt. b_5 is unchanged on addition of benzamidoxime (Supporting Information Fig. S5), which shows that the mediator is incapable of reducing benzamidoxime.

The pH dependence of the catalytic current was examined at a constant and saturating concentration of benzamidoxime (2.4 mM) (Fig. 3). The catalytic current reflects the intrinsic pH dependence of

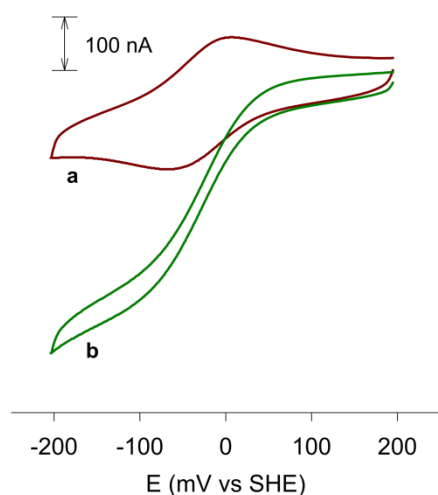


Fig. 2. CVs obtained at a Au/MSA/chitosan-cyt. b_5 /mARC1 electrode (a) in the absence and (b) in the presence of 4 mM benzamidoxime in 0.1 M phosphate buffer solution (pH 6.0) and a scan rate of 5 mV s^{-1} .

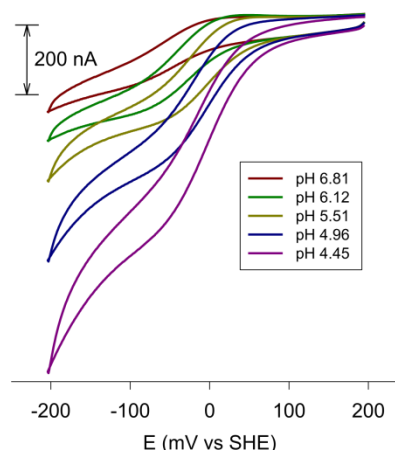


Fig. 3. CVs obtained at a Au/MSA/chitosan-cyt. b_5 /mARC1 electrode at various pH values in the presence of 2.5 mM benzamidoxime (0.1 M mixed buffer solution titrated with AcOH/NaOH) and a scan rate of 5 mV s^{-1} .

mARC catalysis. However the catalytic current is also limited by the amount of cyt. b_5 mediator at the electrode which is moderately pH dependent (Fig. 1). For this reason, the catalytic currents in Fig. 3 were normalized by taking the ratio of the observed catalytic current (i_{cat}) and the corresponding peak current in the absence of benzamidoxime (i_p) in order to factor out any variations in the amount of electroactive cyt. b_5 as a function of pH. The result of this analysis (Supporting Information Fig. S6F) shows that mARC1 activity increases threefold going from pH 7 to pH 4.5. No comparable pH-dependent biochemical assay data have been published for mARC to date but this work demonstrates that mARC can function efficiently in weakly acidic solutions as low as pH 4.5.

To further illustrate that mARC1 is functioning natively, the temperature dependence of the catalytic current was investigated in the presence of a saturating concentration of benzamidoxime. As expected the current increases markedly with temperature and an Arrhenius plot (Supporting Material Fig. S7) yielded an activation energy for the catalytic reaction of 32 kJ mol^{-1} which is reasonable on the basis of comparable enzyme catalyzed processes [21]. The same electrochemical methodology was applied to the analogous substrate *p*-tri-fluoromethyl benzamidoxime. The results (Supporting Information Fig. S8) are qualitatively similar to those obtained for benzamidoxime albeit with higher apparent K_M values.

4. Conclusions

In principle this novel electrochemical procedure can be applied to any potential mARC substrate or inhibitor. No physiologically relevant inhibitors of mARC1 have been identified to date and this approach provides a new way forward for rapid and high throughput identification of these important compounds from a drug metabolism perspective. Reversible or irreversible inhibitors may be identified quickly by exchanging the analyte solution of Au/MSA/chitosan-cyt. b_5 /mARC electrode with fresh buffer and then a mARC substrate to re-establish activity.

Currently, novel drug candidates are optimized in a way that avoids drug metabolism by P450 enzymes. However, the risk of undesirable metabolism by other enzymes such as mARC also needs to be addressed in the assessment of any new drug candidate [22,23]. This is the first example using such a simple and rapid analytical test for evaluating the involvement of this key drug metabolizing enzyme.

Acknowledgement

The Australian Research Council is acknowledged for financial support through Discovery Project DP150103345.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.elecom.2017.10.003>.

References

- [1] A. Havemeyer, F. Bittner, S. Wollers, R. Mendel, T. Kunze, B. Clement, J. Biol. Chem. 281 (2006) 34796–34802.
- [2] J.M. Klein, J.D. Busch, C. Potting, M.J. Baker, T. Langer, G. Schwarz, J. Biol. Chem. 287 (2012) 42795–42803.
- [3] M. Islinger, G.H. Lüers, K.W. Li, M. Loos, A. Völkl, J. Biol. Chem. 282 (2007) 23055–23069.
- [4] G. Ott, A. Havemeyer, B. Clement, J. Biol. Inorg. Chem. 20 (2015) 265–275.
- [5] R. Hille, Chem. Rev. 96 (1996) 2757–2816.
- [6] B. Wahl, D. Reichmann, D. Niks, N. Krompholz, A. Havemeyer, B. Clement, T. Messerschmidt, M. Rothkegel, H. Biester, R. Hille, R.R. Mendel, F. Bittner, J. Biol. Chem. 285 (2010) 37847–37859.
- [7] N. Krompholz, C. Krischkowski, D. Reichmann, D. Garbe-Schönberg, R.R. Mendel, F. Bittner, B. Clement, A. Havemeyer, Chem. Res. Toxicol. 25 (2012) 2443–2450.
- [8] B. Plitzko, G. Ott, D. Reichmann, C.J. Henderson, C.R. Wolf, R. Mendel, F. Bittner, B. Clement, A. Havemeyer, J. Biol. Chem. 288 (2013) 20228–20237.
- [9] C.E. Sparacino-Watkins, J. Tejero, B. Sun, M.C. Gauthier, J. Thomas, V. Ragireddy, B.A. Merchant, J. Wang, I. Azarov, P. Basu, M.T. Gladwin, J. Biol. Chem. 289 (2014) 10345–10358.
- [10] A. Llamas, A. Chamizo-Ampudia, M. Tejada-Jimenez, A. Galvan, E. Fernandez, Biofactors 43 (2017) 486–494.
- [11] S. Gruenewald, B. Wahl, F. Bittner, H. Hungeling, S. Kanzow, J. Kotthaus, U. Schwering, R.R. Mendel, B. Clement, J. Med. Chem. 51 (2008) 8173–8177.
- [12] P.V. Bernhardt, Chem. Commun. 47 (2011) 1663–1673.
- [13] J. Tkac, J.J. Davis, J. Electroanal. Chem. 621 (2008) 117–120.
- [14] P.V. Bernhardt, K.-I. Chen, P.C. Sharpe, J. Biol. Inorg. Chem. 11 (2006) 930–936.
- [15] F.M.C. He, P.V. Bernhardt, J. Biol. Inorg. Chem. 22 (2017) 775–788.
- [16] M. Rivera, M.A. Wells, F.A. Walker, Biochemistry 33 (1994) 2161–2170.
- [17] M. Rivera, R. Seetharaman, D. Girdhar, M. Wirtz, X. Zhang, X. Wang, S. White, Biochemistry 37 (1998) 1485–1494.
- [18] T. Aono, Y. Sakamoto, M. Miura, F. Takeuchi, H. Hori, M. Tsubaki, J. Biomed. Sci. 17 (2010) 90.
- [19] S. Cataldo, F. Crea, A. Gianguzza, A. Pettignano, D. Piazzese, J. Mol. Liq. 148 (2009) 120–126.
- [20] D.W. Lee, C. Lim, J.N. Israelachvili, D.S. Hwang, Langmuir 29 (2013) 14222–14229.
- [21] A. Fersht, Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding, W. H. Freeman and Company, New York, 1999.
- [22] J.M. Hutzler, R.S. Obach, D. Dalvie, M.A. Zientek, Expert Opin. Drug Metab. Toxicol. 9 (2013) 153–168.
- [23] M. Strolin Benedetti, R. Whomsley, E. Baltes, Expert Opin. Drug Metab. Toxicol. 8 (2006) 895–921.

Electronic Supporting Information

Human Mitochondrial Amidoxime Reducing Component (mARC): electrochemical identification of new substrates and inhibitors

Palraj Kalimuthu, Antje Havemeyer, Bernd Clement, Christian Kubitza, Axel J. Scheidig and Paul V. Bernhardt*

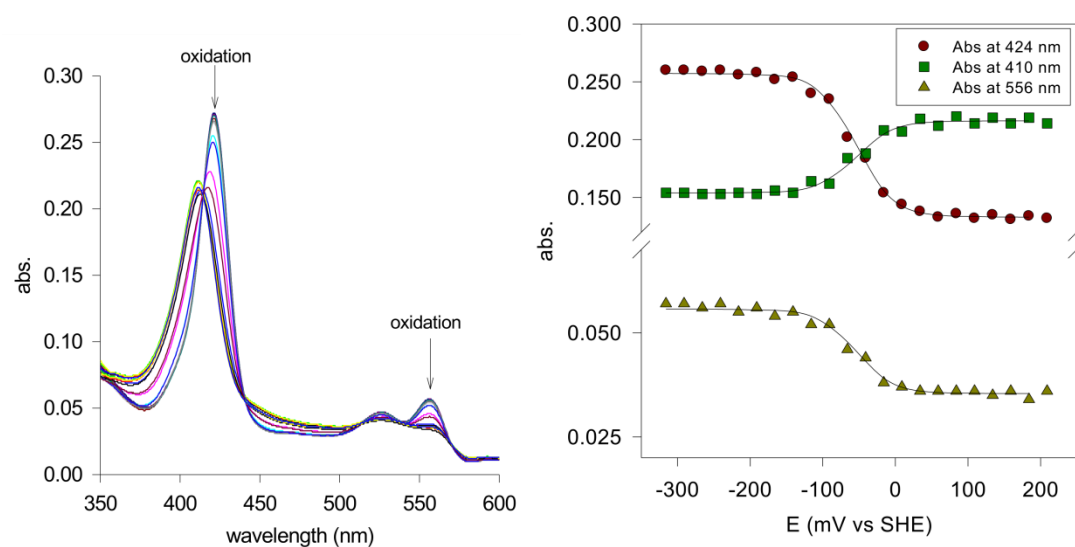
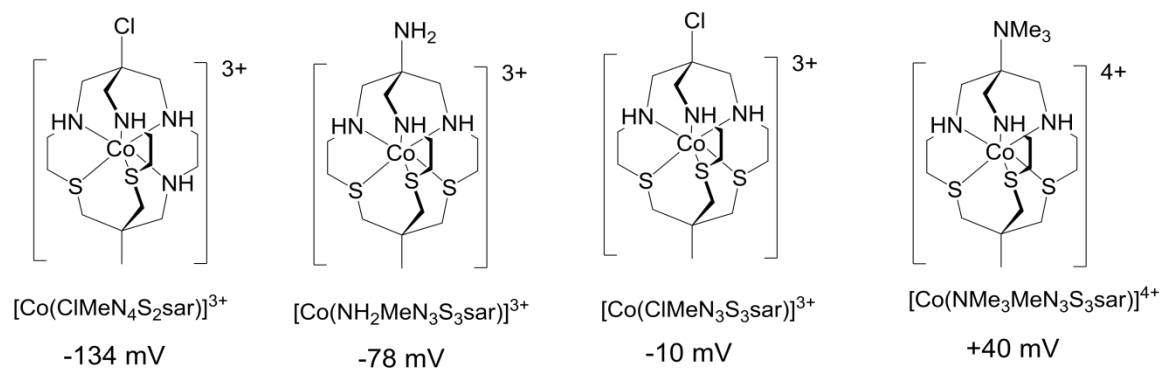


Figure S1. Spectroelectrochemical data for (a) cyt. b_5 ($7.7 \mu\text{M}$) at pH 8 (Tris buffer). Mediators present are shown below and present at concentrations of $10 \mu\text{M}$.



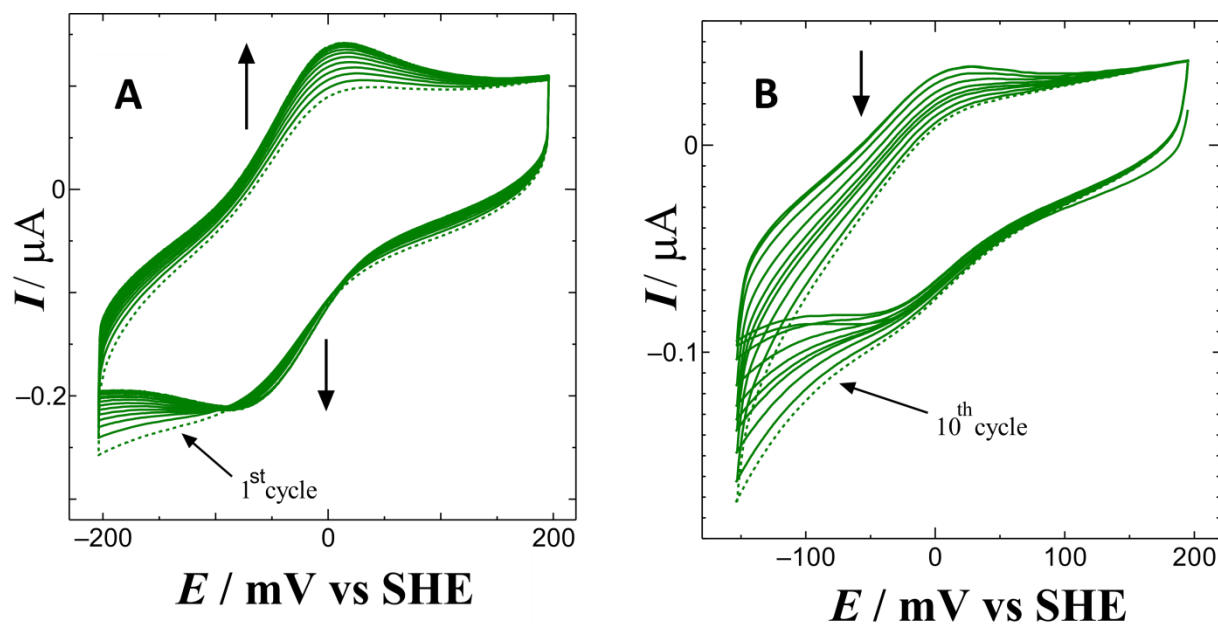


Figure S2. Continuous CVs obtained for cyt. *b*₅ (A) in the presence and (B) in the absence of a dialysis membrane at Au/MSA/chitosan electrode: 100 mM phosphate buffer solution (pH 6.0) at a scan rate of 10 mV s^{-1} . The vertical arrows indicate the direction of change with successive cycles. Note the different current vertical axes.

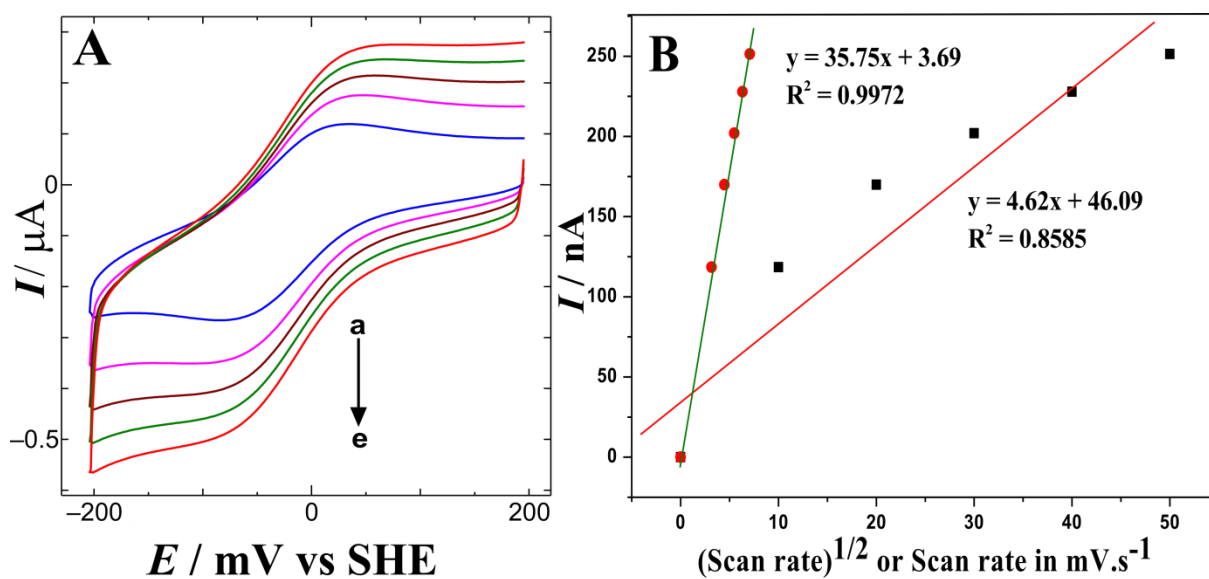


Figure S3. (A) CVs obtained for Au/MSA/chitosan-cyt b_5 electrode in 0.1 M phosphate buffer solution (pH 6.0) at different scan rates (a) 10, (b) 20, (c) 30, (d) 40 and (e) 50 mV s^{-1} and (B) Plot obtained for the base line subtracted oxidation current vs scan rate (black squares) or $(\text{scan rate})^{1/2}$ (red circles).

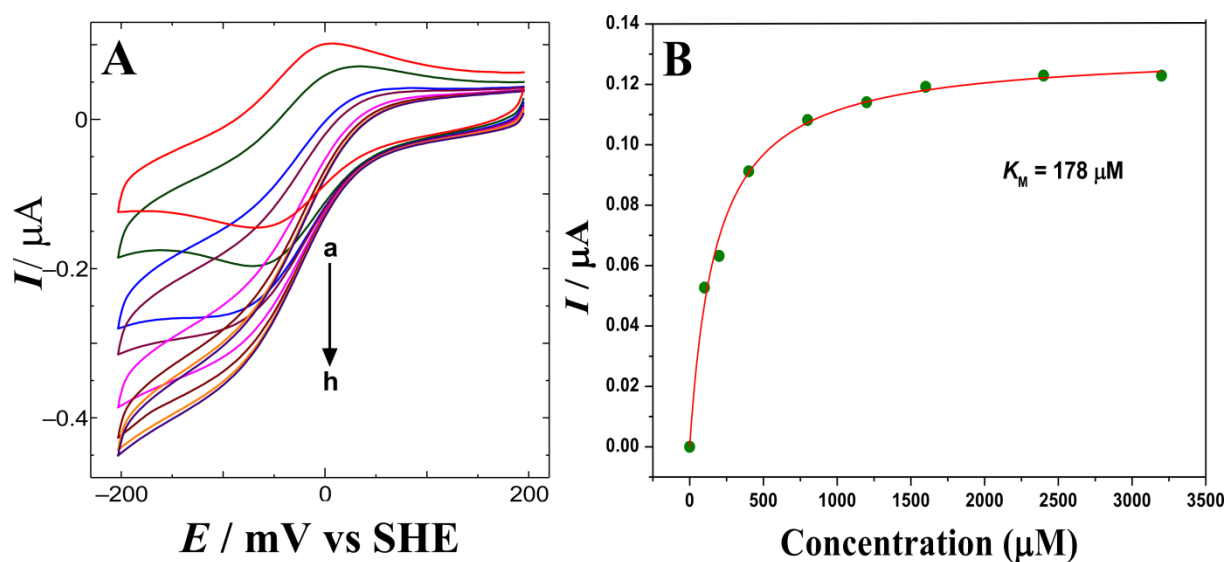


Figure S4. (A) CVs obtained for the increasing concentration of benzamidoxime (a) 0, (b) 100, (c) 200, (d) 400, (e) 800, (f) 1600, (g) 2400 and (h) 3200 μM at Au/MSA/chitosan-cyt b_5 /mARC1 electrode in 0.1 M phosphate buffer solution pH 6.0 at a scan rate of 5 mV s^{-1} . (B) Michaelis-Menten plot for the baseline subtracted electrocatalytic reduction current at -100 mV (i_{cat}) as a function of benzamidoxime concentration ($[S]$) and modelled with Equation (1)

$$i_{\text{cat}} = \frac{i_{\text{max}}[S]}{K_{\text{M,app}} + [S]} \quad (1)$$

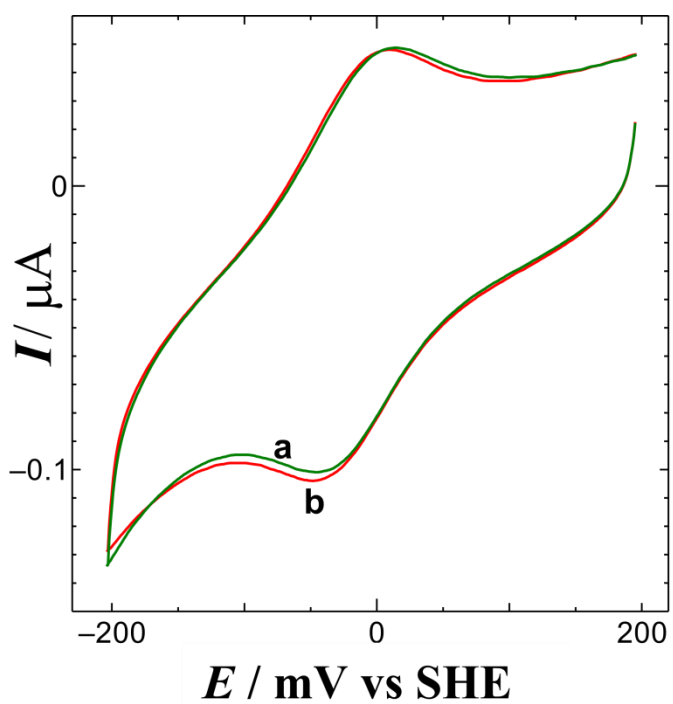


Figure S5. CVs obtained in the (a) absence and (b) presence of 4 mM benzamidoxime at a Au/MSA/chitosan-cyt b_5 electrode (no mARC1) in 0.1 M phosphate buffer solution (pH 6.0) at a scan rate of 5 mV s^{-1} .

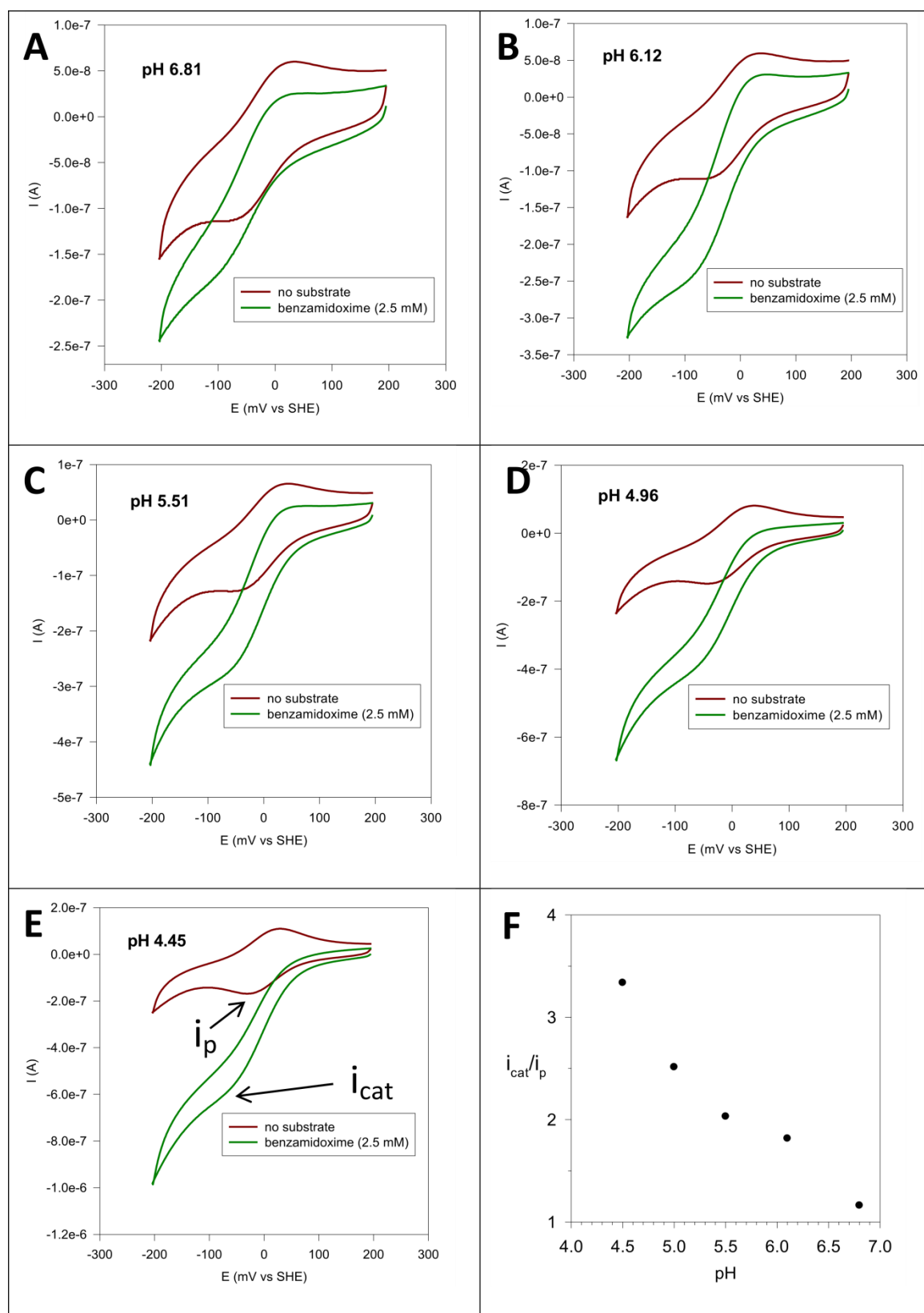


Figure S6. (A-E) Superposition of non-catalytic and catalytic CVs measured with the Au/MSA/chitosan-cyt b_5 /mARC1 electrode at different pH values and (F) normalised catalytic current (catalytic current i_{cat} divided by non-turnover current i_p) as a function of pH. Note that $i_{cat}/i_p = 1$ represents no catalysis.

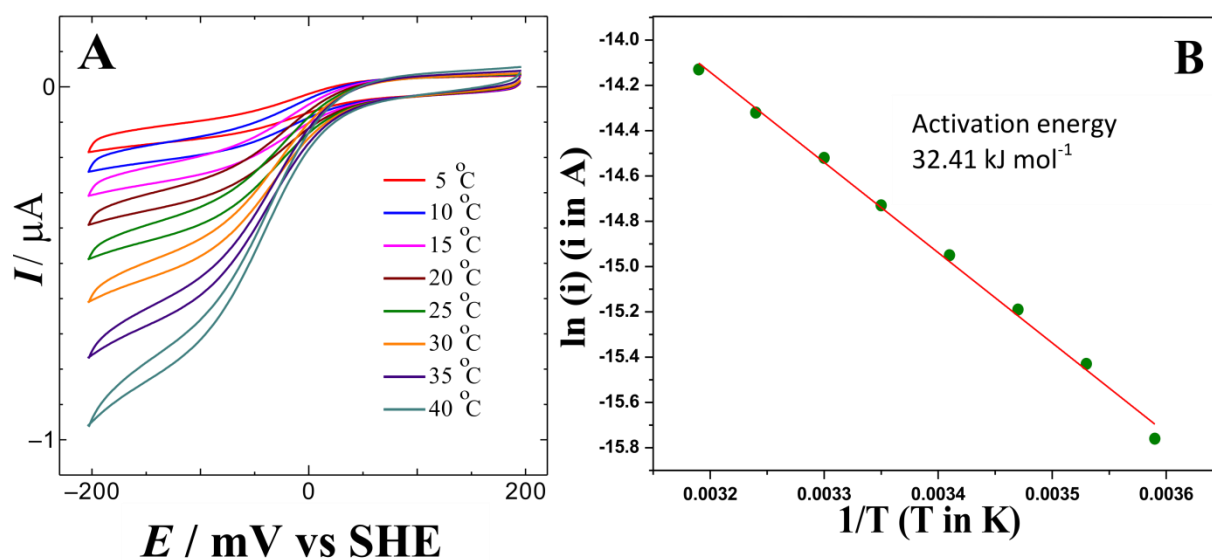


Figure S7. (A) CVs obtained with the Au/MSA/chitosan-cyt b_5 /mARC1 electrode in the presence of 3.2 mM benzamidoxime at different temperatures: 100 mM phosphate buffer (pH 6.0) and scan rate of 5 mV s^{-1} . (B) Arrhenius plot of the natural logarithm of maximum catalytic current versus the reciprocal of absolute temperature (K).

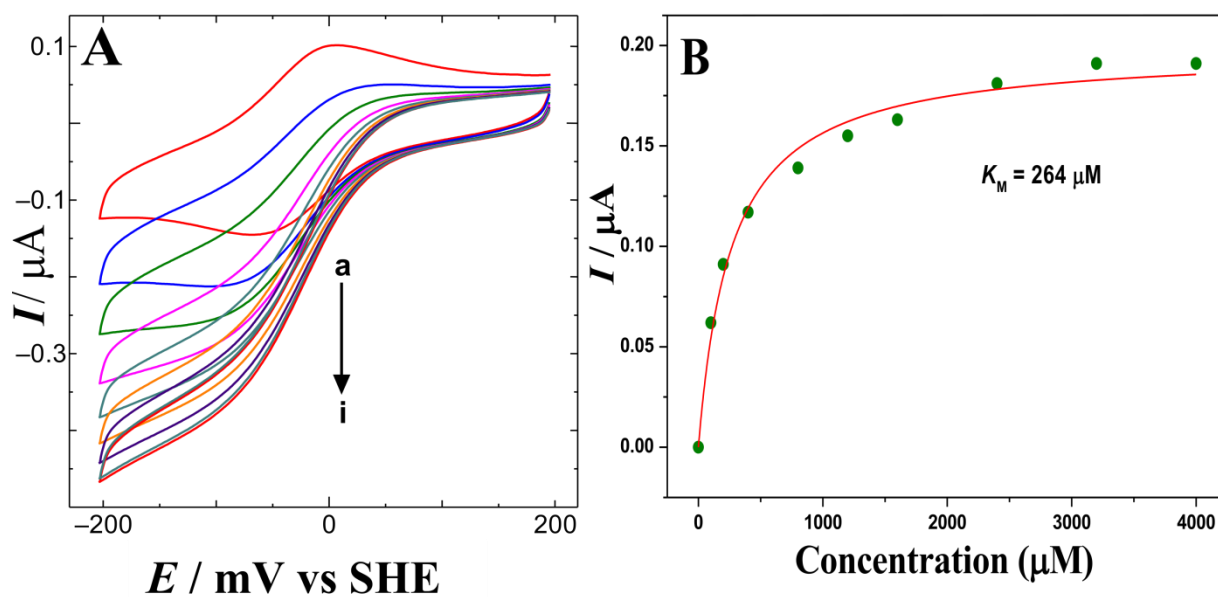


Figure S8. (A) CVs obtained for the increasing concentration of substrate 4-(trifluoromethyl)benzamidoxime (a) 0, (b) 100, (c) 200, (d) 400, (e) 800, (f) 1600, (g) 2400, (h) 3200 and (i) 4000 μM at a Au/MSA/chitosan-cyt b₅/mARC1 electrode in 0.1 M phosphate buffer solution (pH 6.0) and scan rate of 5 mV s^{-1} . (B) Michaelis-Menten plot for the baseline subtracted electrocatalytic reduction current at -100 mV as a function of substrate concentration with data fit to Equation (1).

3.3 The Involvement of the Mitochondrial Amidoxime Reducing Component (mARC) in the Reductive Metabolism of Hydroxamic Acids

Carsten Ginsel, Birte Plitzko, Danilo Friepp, Diana A. Stofa, Manfred Jung, Christian Kubitz, Axel Scheidig, Antje Havemeyer, Bernd Clement

Manuscript submitted to *Drug Metabolism and Disposition*: May 18^h, 2018

Hydroxamic acids are compounds with a broad range of pharmacological applications. They were shown to exhibit anti-inflammatory¹⁰⁵ as well as antibiotic activities.^{106,107} Owing to their strong cation-chelating capabilities, they also have inhibitory effects on metalloproteinases and histone deacetylases.^{108,109} Therefore, drug candidates, such as antineoplastics, often contain hydroxamic acid functional groups.^{110,111} However, in order to effectively exert their pharmacological activities, these compounds require some degree of metabolic stability, so that they are not immediately inactivated by biotransformation enzymes.

The metabolism of orally administered hydroxamic acid compounds leads to the formation of glucuronides, the corresponding carboxylic acids or the amides as the main metabolites.¹¹² Since hydroxamic acids belong to the diverse group of *N*-hydroxylated compounds, they were considered potential substrates of the mARC enzymatic system, which was supposed to be responsible for the reduction to the corresponding amides. In the following manuscript, hydroxamic acids are confirmed to be mARC substrates by HPLC-based activity assays using benzhydroxamic acid as a model substrate. Furthermore, the mARC-dependent reduction of three different drugs containing hydroxamic acid moieties is investigated as well as the in-/ability of this enzyme system to convert the toxic metabolite *N*-hydroxyphenacetin to the pharmacologically active compound.

I performed the expression and purification of soluble human mARC, cytochrome b₅ and NADH cytochrome b₅ reductase. This three-component recombinant enzymatic system was subsequently used for the activity assays performed by Carsten Ginsel (group of Prof. Dr. Clement) to investigate the mARC-dependent conversion of hydroxamic acid compounds.

**The Involvement of the Mitochondrial Amidoxime Reducing Component (mARC) in the
Reductive Metabolism of Hydroxamic Acids**

**Carsten Ginsel, Birte Piltzko, Danilo Froriep, Diana A. Stolfa, Manfred Jung, Christian Kubitza,
Axel J. Scheidig, Antje Havemeyer, Bernd Clement***

CG, BP, DN, AH, BC:

Pharmaceutical Institute

Department of Pharmaceutical & Medicinal Chemistry

Christian-Albrechts-Universität zu Kiel

Gutenbergstraße 76, 24118 Kiel (Germany)

E-Mail: bclement@pharmazie.uni-kiel.de

MJ, DAS:

Institute of Pharmaceutical Sciences

Albert-Ludwig-Universität Freiburg

Albertstraße 25, 79104 Freiburg (Germany)

CK, AJS:

Zoological Institute – Structural biology

Christian-Albrechts-Universität zu Kiel

Am Botanischen Garten 1-9, 24118 Kiel (Germany)

Running Title: Reduction of hydroxamic acids by mARC

To whom correspondence should be addressed: Bernd Clement, Department of Pharmaceutical and Medicinal Chemistry, Pharmaceutical Institute, Christian-Albrechts-University of Kiel, Gutenbergstraße 76, 24118 Kiel, Germany, Tel.: +494318801129; Fax: +494318801352; Email: bclement@pharmazie.uni-kiel.de

Text pages:

Tables: 1

Figures: 5

References: 45

Words Abstract: 243

Words Introduction: 664

Words Discussion: 665

Nonstandard abbreviations:

ACN = acetonitrile, BCA = bicinchoninic acid, bs = broad singlet, CI= chemical ionization, CYB5B = cytochrome b5B, CYB5R = NADH-cytochrome b5-reductase, d = doublet, dd = double doublet, DMSO = dimethyl sulfoxide, DS116 = N¹-phenyloctanediamide, DS92 = 8-oxo-8-(phenylamino)octanoic acid, dt = double triplet, EI = electron ionization, ESI = electrospray ionization, FBS = fetal bovine serum, HDACs = histone deacetylases, HPLC = high performance liquid chromatography, J = coupling constant in Hz, m = multiplet, LOQ = limit of quantification, mARC = mitochondrial amidoxime reducing component, MMP = matrix metalloproteinase, mp = melting point, NMR = nuclear magnetic resonance, PCI = positive chemical ionization, rt = retention time, s = singlet, t = triplet, TLC = thin layer chromatography, UV = ultra violet, δ = chemical shift in ppm

Abstract

The mitochondrial amidoxime reducing component is a recently discovered molybdenum enzyme in mammals which, in concert with the electron transport proteins cytochrome b5 and NADH cytochrome b5 reductase, catalyzes the reduction of *N*-oxygenated structures. This three component enzyme system plays a major role in *N*-reductive drug metabolism. Belonging to the group of *N*-hydroxylated structures, hydroxamic acids are also potential substrates of the mARC-system. Hydroxamic acids show a variety of pharmacological activities and are therefore often found in drug candidates. They can also exhibit toxic properties as is the case for many aryl hydroxamic acids formed during the metabolism of arylamides. Biotransformation assays using recombinant human proteins, subcellular porcine tissue fractions as well as human cell culture were performed. Here the mARC-dependent reduction of the model compound benzhydroxamic acid is reported in addition to the reduction of three drugs. In comparison to other known substrates of the molybdenum depending enzyme system (e.g. amidoxime prodrugs) the conversion rates measured here are slower, thereby reflecting the mediocre metabolic stability and oral bioavailability of distinct hydroxamic acids. Moreover, the toxic *N*-hydroxylated metabolite of the analgesic phenacetin, *N*-hydroxyphenacetin, is not reduced by the mARC-system under the chosen conditions. This confirms the high toxicity of this component, as it needs to be detoxified by other pathways.

This work highlights the need to monitor the *N*-reductive metabolism of new drug candidates by the mARC-system when evaluating the metabolic stability of hydroxamic acid-containing structures or the potential risks of toxic metabolites.

Introduction

Hydroxamic acids are a class of substances with a variety of biological activities, including antibiotic (Barb et al., 2007; Halouska et al., 2014) and anti-inflammatory activities (Brogden et al., 1975), along with inhibitory properties towards metalloproteinases (Dalvie et al., 2008; Verma, 2012) and histone-deacetylases (HDACs) (Dokmanovic et al., 2007; Zhang and Zhong, 2014). The latter is of particular interest, as inhibitors of HDACs show antineoplastic activities and are currently applied in cancer treatment with new candidates under development (Wagner et al., 2010; Zhang and Zhong, 2014). Examples are vorinostat (suberoylanilide hydroxamic acid = SAHA) which is marketed as Zolinza[®] for the treatment of T-cell lymphoma and Panabinstat (Farydak[®]) which has just been approved (Raedler, 2016).

Applying compounds with hydroxamic acids in therapy requires a reasonable level of metabolic stability. One expectable metabolic pathway, besides conjugation reactions, is hydrolysis to the corresponding carboxylic acid (Liu et al., 2014). As shown in Figure 1, another likely metabolic conversion is the reduction of hydroxamic acids to the corresponding amide (Lowenthal, 1954; Hirsch and Kaplan, 1961; Kitamura and Tatsumi, 1985; Kiesel et al., 2013). For example, metabolism studies of CP544439, a hydroxamic acid-containing MMP inhibitor, revealed the formation of the amide, the glucuronide and the carboxylic acid as the main metabolites of the orally administered drug (Dalvie et al., 2008). Besides their pharmacological advantages, aryl hydroxamic acids have been shown to possess toxic and mutagenic properties (Miller et al., 1961; Vaught et al., 1981). For example, the analgesic drug 'Phenacetin' is *N*-hydroxylated during metabolism to yield *N*-hydroxyphenacetin (Hinson and Mitchell, 1976; Wirth et al., 1980). Phenacetin was withdrawn from the market because it was found to induce severe renal papillary necrosis and tumors of the renal pelvis and bladder in humans (Liu et al., 1972; Bengtsson et al., 1978). *N*-hydroxyphenacetin has been held responsible for these severe effects as it conducts similar pathways to many aromatic amines leading to the formation of DNA adducts (Vaught et al., 1981). *In vivo* and *in vitro*, *N*-hydroxyphenacetin is metabolized to Phenacetin (Fischbach and Lenk, 1985). The enzymatic basics of the reductive metabolism remain undetermined. The involvement of the cytochrome P450 isoform 2S1 in this reductive metabolism could be excluded (Wang and Guengerich, 2013). It is well accepted that the mitochondrial amidoxime-reducing component 'mARC', together with the electron transport proteins cytochrome b5 and NADH-cytochrome b5 reductase, forms an *N*-reductive

three component enzyme system located in the outer mitochondrial membrane which plays a major role in *N*-reductive drug metabolism. It has been shown after *in vitro* reconstitution and in cell culture that the mARC-system is responsible for the reduction of various *N*-hydroxylated compounds like amidoximes, *N*-hydroxyguanidines or sulfhydroxamic acids, hydroxylamines and *N*-oxides (Plitzko et al., 2013; Ott et al., 2015). Our laboratory has recently shown that mARC is a mitochondrial, molybdenum-containing enzyme (Havemeyer et al., 2006). All of the currently analyzed and completely annotated mammalian genomes code for two mARC proteins (mARC1 and mARC2) which share a high degree of sequence identity to each other (Wahl et al., 2010). Though the endogenous function of mARC is still not fully understood, mARC proteins are assumed to be involved in detoxification of mutagenic and toxic aromatic hydroxylamines like *N*-hydroxylated DNA-base analogs (Krompholz et al., 2012; Plitzko et al., 2015). The involvement in energy and NO metabolism has been discussed (Kotthaus et al., 2011; Neve et al., 2012; Jakobs et al., 2014; Sparacino-Watkins et al., 2014). In this study we examined whether the mARC-system is also capable of reducing hydroxamic acids and is thus involved in the metabolic conversion of this substance class. We assayed the *N*-reduction of the model substrate benzhydroxamic acid and of four relevant pharmaceutical compounds (vorinostat, bufexamac and CP544439 as drugs and *N*-hydroxyphenacetin as a toxic drug metabolite) by performing biotransformation assays in the reconstituted recombinant human *N*-reductive system and in subcellular porcine liver fractions. Additionally, by performing metabolism studies and RNAi-mediated down-regulations of mARC in HEK-293 cells, the physiological relevance of the *N*-reduction of hydroxamic acids in human cell metabolism was evaluated.

Material and Methods

Reagents and cell lines. Unless otherwise stated all chemicals were purchased from Carl Roth GmbH & co. KG (Karlsruhe, Germany), Sigma Aldrich or Fluka and used without further purification. Methanol (HPLC grade) was from JT Baker (Deventer, Netherlands). Acetonitrile (HPLC grade) was from Honeywell (Seelze, Germany). Benzoic acid and chlorosulfonic acid were from Merck KGaA (Darmstadt, Braunschweig). Ethyl 4-aminotetrahydro-2H-pyran-4-carboxylate hydrochloride and 4-fluorophenoxybenzene were from abcr (Karlsruhe, Germany). 4-nitrophenetole, 4-butoxyphenylacetic acid and sodium 1-octanesulfonate were from TCI (Zwijndrecht, Belgium). HEK-293 human embryonic kidney cells were purchased from Cell Lines Service (Eppelheim, Germany). Opti-MEM, minimum essential medium, sodium pyruvate solution, sodium bicarbonate, minimum Eagles's medium nonessential amino acids, FBS, trypsin, L-glutamine, PBS, Lipofectamine RNAiMAX, Stealth Select RNAi siRNA targeting human MOSC1 (MOSC1HSS127704), and Stealth Select RNAi siRNA negative control were obtained from Invitrogen (Germany). ONTARGETplus SMARTpool siRNA targeting human MOSC2 was purchased from Thermo Scientific. Complete protease inhibitor cocktail was acquired from Roche Applied Science (Mannheim, Germany). Benzamidoxime was synthesized from benzonitrile and hydroxylamine (Krüger, 1885). For the synthesis of CP544439 and its metabolites see supplemental data.

Synthesis of vorinostat and metabolites. Chromatographic separations were performed on silica gel (15–40 mesh, Merck) using flash methodology. Reaction progress was monitored by analytical TLC on pre-coated silica gel (Kieselgel 60 F254) plates, and spots were detected by UV light (λ 254 nm). mp of the final target HDAC inhibitors were determined by the open capillary method on a Stuart-Scientific SMP3 electrothermal apparatus and are uncorrected. ^1H NMR spectra were recorded in the indicated deuterated solvents on a Bruker Avance DRX 400 MHz spectrometer and ^{13}C NMR on a Varian 100 MHz. Signals due to OH and NH protons were located by deuterium exchange with D_2O . EI- and CI-mass spectra were measured with a TSQ700 mass spectrometer (Thermoelectron). ESI- and PCI-mass spectra were recorded with a LCQ-Advantage mass spectrometer. In all cases, spectroscopic data are in agreement with known compounds and assigned structures. HPLC purity determinations were performed on a JASCO HPLC system under isocratic conditions, using a Phenomenex Synergi Hydro RP-C18 column (250 mm x 4.6 mm, 4 μm particle size). Elution was performed using 0.05 % of TFA in water/ACN 60/40 (v/v), at room temperature. The purity of all tested compounds was $\geq 98\%$, as measured by HPLC.

Injection volumes were 2 μ L, flow rate was 0.5 ml/min, detection was performed with UV ($\lambda = 254$ nm). All chromatographic and spectroscopic data were in accordance with literature data.

Synthesis of vorinostat: vorinostat was synthesized according to (Mai et al., 2001). Retention time for HPLC was 7.8 min.

Synthesis of DS92: DS92 was synthesized by a modification of the procedure described in the literature (Suzuki et al., 2005). Suberic acid (5.00 g, 28.7 mmol) was slowly added to neat freshly distilled aniline (2.70 ml, 30.1 mmol) and the mixture was stirred at 185 °C and left at the same temperature for 2 h. The mixture was cooled to room temperature and a NaOH 2 N aqueous solution was added to set the pH of the mixture at ~8. The suspension was filtered and the filtrate was acidified to pH~2, obtaining a white precipitate that was collected and suspended in warm water (50 °C). The insoluble part was filtered and washed with hot water, yielding DS92 as white pure precipitate (3.93 g, 55 %). $R_f = 0.35$ (EtOAc/Cy, 8:2); Retention time for HPLC was 14.5 min; mp = 124-126 °C; $^1\text{H NMR}$ (400 MHz, $d_6\text{DMSO}$): δ 1.23-1.35 (m, 4H), 1.42-1.52 (m, 2H), 1.54-1.63 (m, 2H), 2.20 (t, J = 7.4 Hz, 2H), 2.29 (t, J = 7.4 Hz, 2H), 7.01 (tt, J = 7.6; J = 1.1 Hz, 1H), 7.28 (t, J = 7.6 Hz, 2H), 7.58 (dd, J = 8.7; J = 1.1 Hz, 2H), 9.85 (s, 1H), 11.97 (br s, 1H); LRMS (ESI) m/z 248.1 [M-H] $^-$

Synthesis of DS116: DS116 was obtained by a fast two-step synthesis starting from the carboxylic acid derivative DS92, which is transformed using thionyl chloride into the corresponding acyl chloride, further reacted into the target according to an already described procedure (Wright and Corbett, 1993). To a solution of 8-oxo-8-(phenylamino)octanoic acid (0.30 g, 1.2 mmol) in dry DCM (3 ml) thionyl chloride (330 μ l, 4.6 mmol) was slowly added at room temperature. The solution was heated to reflux for 3 h and later a stream of nitrogen was used to remove the solvent and the excess of thionyl chloride. To the crude product was added a solution of concentrated aqueous ammonia solution (1.03 ml, 26.4 mmol) and NH_4Cl (0.22 g, 4.1 mmol) in 1 ml of water and the mixture was stirred at 20 $^\circ\text{C}$ for 8 h. The reaction was quenched with NaBH_4 (23 mg, 0.6 mmol) and after 1 h at room temperature HCl 2 N aqueous solution was added dropwise to pH~2. The precipitate was collected to give DS116 as white powder (0.24 g, 80 %). $R_f = 0.30$ (DCM/ Et_2O /MeOH, 9.5:1:0.5); Retention time for HPLC was 9.18 min; mp = 161-163 $^\circ\text{C}$; $^1\text{H-NMR}$ (400 MHz, DMSO- d_6): δ 1.22-1.34 (m, 4H), 1.44-1.52 (m, 2H), 1.63-1.53 (m, 2H), 2.03 (t, J = 7.4 Hz, 2H), 2.29 (t, J = 7.4 Hz, 2H), 6.68 (br s, 1H), 7.01 (t, J = 7.4, 1H), 7.22 (br s, 1H), 7.28 (t, J = 7.8 Hz, 2H), 7.58 (d, J = 7.8 Hz, 2H), 9.85 (s, 1H); LRMS (ESI) m/z 247.1 [M-H] $^-$.

Synthesis of *N*-hydroxyphenacetin. *N*-hydroxyphenacetin was prepared using a previously reported procedure with minor modifications (Hinson and Mitchell, 1976). 4-Nitrophenetole (2.0 g, 11.96 mmol) and NH_4Cl (0.64 g, 11.96 mmol) were dissolved in 40 ml of a $\text{C}_2\text{H}_5\text{OH}/\text{H}_2\text{O}$ mixture (4:1, v/v) at room temperature. After the addition of Zn dust (3.2 g, 48.94 mmol) the reaction mixture was stirred for 10 minutes. The Zn dust was filtered and washed with 40 ml $(\text{C}_2\text{H}_5)_2\text{O}$. The ethereal phase was washed with 50 ml brine and separated from the aqueous phase. NaHCO_3 (1.6 g, 19.04 mmol) was suspended in 5 ml H_2O at 0 $^\circ\text{C}$ and the *N*-hydroxyphenetidine containing ethereal phase was added. 500 μ l of 2.5 % acetyl chloride in $(\text{C}_2\text{H}_5)_2\text{O}$ were slowly added over 1 h. The formation of the product was TLC controlled. *N*-hydroxyphenacetin gave a red, *N*-hydroxyphenetidine a blue spot with FeCl_3 . After disappearance of the blue spot, 20 ml of H_2O was added and twice extracted with $(\text{C}_2\text{H}_5)_2\text{O}$. The combined layers were washed with H_2O and twice with brine. The product was extracted with 100 ml of cold 0.2 M NH_3 -solution. After neutralization to pH 7.0 with NaHCO_3 the product was extracted in 300 ml $(\text{C}_2\text{H}_5)_2\text{O}$, dried over anhydrous Na_2SO_4 and the solvent was evaporated to dryness. *N*-hydroxyphenacetin was recrystallized with $(\text{C}_2\text{H}_5)_2\text{O}$ /hexane, 15 % yield from 4-nitrophenetole. $^1\text{H-NMR}$ (300 MHz, DMSO- d_6): δ 10.48 (s, 1H,

hydroxyl), 7.45 (d, J = 8.7 Hz, aromatic, 2H), 6.91 (d, J = 9.1 Hz, aromatic, 2H), 4.01 (q, J = 7.0 Hz, 2H, H₃CH₂-), 2.14 (s, 3H, CH₃CO-), 1.31 (t, J = 7.0 Hz, 3H, CH₃CH₂-); LC-MS (ESI), m/z 196 [M+H]⁺, 178, 150.

Protein sources. Subcellular porcine tissue fractions were purified as described earlier (Ott et al., 2014). Expression and purification of human mARC1 (reference sequence NP_073583) and mARC2 (reference sequence NP_060368), CYB5B (reference sequence NP_085056) and CYB5R isoform3 (reference sequence NP_000389) was carried out in *Escherichia coli* as described by Wahl and coworkers (Wahl et al., 2010). Protein content was determined using (BCA) protein assay kit (Pierce, Rockford, IL, USA) according to the manufacturer's protocol. Heme content in CYB5B was determined according to the method of Estabrook by recording the difference spectrum of oxidized and NADH-reduced protein (Estabrook and Werringloer, 1978). FAD-content in CYB5R was measured at 450 nm according to Whitby after sample preparation by heating at 100 °C for 10 min and centrifugation at 22000 g for 5 min at room temperature (Whitby, 1953).

In vitro N-reductive activity assay. *In vitro* biotransformation assay was carried out at 37 °C in a shaking water bath. Incubation mixture consisted of 100 µg of porcine subcellular fractions in 100 mM potassium phosphate buffer, pH 6.0 or in case of recombinantly expressed human proteins of 7.5 µg mARC1 or mARC2, CYB5B resulting in 75 pmol heme and CYB5R resulting in 7.5 pmol FAD in 20 mM MES buffer, pH 6.0. Different substrate concentrations were used and for CP544439, vorinostat and bufexamac 4.0 %, 4.8 % and 8.0 % DMSO were added respectively. After 3 min of pre-incubation, the reaction was started by adding 1 mM NADH, resulting in a total volume of 150 µl. Incubation was stopped after 15 min by adding 150 µl of cold methanol. Afterwards, samples were shaken for 5 min at room temperature and centrifuged for 5 min with 9500 g at room temperature. Supernatants were analyzed by HPLC. For the determination of kinetic parameters with recombinantly expressed human protein the DMSO concentration was kept at the same level and only the substrate concentration was modified.

HPLC analysis. The flow rate was kept at 1.0 ml/min and the injection volume was 10 µl for all performed HPLC analysis. All benzhydroxamic acid and CP544439 related samples were measured on a Waters e2695 Separation Module with a Waters 2998 Photodiode Array Detector and Waters Empower 2 Build 2154 as integration software.

For the separation of benzhydroxamic acid ($rt = 5.0 \pm 0.2$ min) benzamide ($rt = 8.1 \pm 0.2$ min) and benzoic acid ($rt = 12.7 \pm 0.1$ min) a Phenomenex Gemini NX-C18 ($5 \mu\text{m}$), 150×4.6 mm with a Phenomenex C18 4×3.0 mm pre-column was used. The mobile phase consisted of 50 mM KH_2PO_4 , pH 4.6, 10 mM tetramethylammonium chloride and 10 % acetonitrile (v/v). Detection wavelength was 210 nm.

For the separation of CP544439 ($rt = 4.8 \pm 0.3$ min), deoxy CP544439 ($rt = 5.9 \pm 0.2$ min) and the carboxylic derivative ($rt = 7.0 \pm 0.2$ min) a Waters Sunfire C18, $3.5 \mu\text{m}$, 150×4.6 mm with a Phenomenex C18 4×3.0 mm pre-column was used. Solvent A (0.2 % formic acid in H_2O (v/v)) and Solvent B (0.2 % formic acid in acetonitrile (v/v)) were used. Starting with 60 % A, the gradient changed linearly from 3 min to 7 min to 10 % A. At 11 min A was set to 60 % over 0.5 min. Total runtime was 16 min. Detection wavelength was 247 nm. The column temperature was maintained at 25°C and sample storage temperature at 18°C .

All vorinostat, bufexamac, *N*-hydroxyphenacetin and benzamidoxime related samples were measured on a Waters HPLC system consisting of a Waters 717 autosampler, a Waters 1525 pump and a Waters 2487 dual absorbance detector at room temperature. A Phenomenex Gemini NX-C18 ($5 \mu\text{m}$), 150×4.6 mm with a Phenomenex C18 4×3.0 mm pre-column was used with exception for benzamidoxime where a LiChrospher® 60 RP-select B ($5 \mu\text{m}$), 250×4 mm column combined with a LiChrospher 60 RP-select pre-column was used.

For separation of vorinostat ($rt = 6.7 \pm 0.0$ min), DS116 ($rt = 8.9 \pm 0.1$ min) and DS92 ($rt = 18.9 \pm 0.2$ min) the mobile phase consisted of 100 mM KH_2PO_4 and 23 % acetonitrile (v/v). The detection wavelength was set to 254 nm.

For separation of bufexamac ($rt = 16.9 \pm 0.1$ min), deoxy bufexamac ($rt = 20.9 \pm 0.1$ min) and the carboxylic derivative ($rt = 28.5 \pm 0.2$ min) the mobile phase consisted of 50 mM KH_2PO_4 , pH 4.4, 10 mM tetramethylammonium chloride and 45% methanol (v/v). Detection wavelength was set to 228 nm.

For separation of *N*-hydroxyphenacetin ($rt = 7.8 \pm 0.2$ min) and phenacetin ($rt = 9.1 \pm 0.0$ min) the mobile phase consisted of 1 % formic acid in H_2O (v/v) and 22.5 % acetonitrile (v/v). Detection wavelength was 245 nm.

For separation of benzamidoxime ($rt = 9.1 \pm 0.3$ min) and benzamidine ($rt = 15.9 \pm 0.1$ min) the mobile phase consisted of 10 mM sodium 1-octanesulfonate and 20 % acetonitrile. Detection wavelength was 229 nm.

Cell Culture. HEK-293 cells were maintained in minimum essential medium supplemented with 10% FBS, 2 mM L-glutamine, 0.1 mM nonessential amino acids, 1 mM sodium pyruvate, and 1.5 g/L sodium bicarbonate. The cell line was incubated at 37 °C in 5 % CO₂.

siRNA Transfection and Design of Knockdown Experiments. HEK-293 cells were reverse transfected and mARC-protein down-regulated according to the previous described procedure (Plitzko et al., 2015).

N-reductive metabolism of benzhydroxamic acid in HEK 293 cells. For N-reduction studies in HEK-293 the culture medium was removed, and cells were carefully washed and pre-incubated with substrate-free incubation buffer (Hanks' balanced salt solution containing 10 mM HEPES, pH 7.4) at 37 °C for 10 min. After removing the substrate-free incubation buffer, the vital cells were then incubated with benzhydroxamic acid-containing incubation buffer (3 mM, 0.5% (v/v) DMSO) at 37 °C for 180 min. After the designated time, the culture supernatant was carefully removed, centrifuged to eliminate cellular debris and analyzed by HPLC as described above.

Total Cellular Protein Extraction. Cellular protein was harvested and protein contents determined as previously described (Plitzko et al., 2015).

Western Blot analysis. SDS-PAGE and Western Blot analysis to verify down-regulation of mARC-protein in HEK-293 cells was carried out as described previously (Plitzko et al., 2015).

Statistical Analysis. Statistical analyses were carried out using the SigmaPlot 11 software (Systat Software Inc.). The significance of observed differences was evaluated by Bonferroni test. A probability less than 5% was considered to be significant. All experimental values are given as means \pm S.D.

Results

From all analyzed subcellular fractions, *N*-reductive activity in mitochondria was found to be enhanced compared to other tissue fractions for all *N*-unsubstituted hydroxamic acids (Figure 3). These findings reflect the enrichment of mARC depending enzyme activities in this fraction as published earlier (Krompholz et al., 2012). In the microsomal and cytosolic fractions only minor/no *N*-reductive activity was detectable. By contrast, no reduction of *N*-hydroxyphenacetin could be detected after incubation with the mitochondrial fraction and only minor reduction rates were detected within the cytosolic fraction (data not shown). Besides reduction to the amide, hydrolysis to the corresponding carboxylic acids was monitored and was found to be more pronounced in microsomes compared to other fractions; this indicates the involvement of a microsomal enzymatic system or a non-enzymatic reaction (Figure 3A, D). In the case of bufexamac and CP544439 the microsomal fraction was the only fraction where hydrolysis could be detected to a small extent (data not shown).

To prove the involvement of mARC in the reduction of hydroxamic acids, a cell based siRNA experiment was performed. In HEK-293 cells reductive conversion of benzhydroxamic acid to benzamide occurs in a time-dependent (data not shown) and substrate-dependent manner (Fig. 4A) and followed Michaelis-Menten kinetics ($v_{\max} = 0.06 \pm 0.01 \text{ nmol benzamide} \cdot \text{min}^{-1} \cdot \text{mg protein}^{-1}$). By siRNA-mediated down-regulation of mARC1, benzamide formation decreased dramatically to approximately 30% compared to the negative control (Fig. 4B). Knockdown of mARC2 in HEK-293 cells did not affect the reduction of benzhydroxamic acid in HEK-293 cells. This same behavior is also observed with the model substrate benzamidoxime and is attributed to the low level of mARC2-protein expression in HEK-293 cells (Plitzko et al., 2013). Simultaneous knockdown of both mARC-proteins led to a small, further decrease in *N*-reductive activity than was observed in the mARC1-only knockout experiments. To further elucidate the differences between mARC1 and mARC2 the kinetic parameters v_{\max} and K_M were determined with the *in vitro* reconstituted recombinant *N*-reductive system (Figure 5A-D). All hydroxamic acids were clearly reduced to the corresponding amides with the exception of *N*-hydroxyphenacetin. These results are consistent with previous observations in subcellular fractions. Benzhydroxamic acid, bufexamac, CP544439 and vorinostat were exclusively reduced to their corresponding amides; none of the corresponding carboxylic acid products were detected. The reductions obey Michaelis-Menten kinetics for both mARC-proteins. The calculated K_M and v_{\max} values are presented in Table 1. The conversion rates of hydroxamic acids are

lower in comparison to the model compound benzamidoxime (e.g. for vorinostat, v_{\max} is 15 times slower). For both incubation types with subcellular fractions and recombinant expressed proteins, highest v_{\max} values were obtained for the model compound benzhydroxamic acid in comparison to all other studied hydroxamic acids.

According to the v_{\max} values, CP544439 and vorinostat are reduced to the same amount and smallest conversion rates were detected for bufexamac. Vorinostat, bufexamac and CP544439 required the use of DMSO as a solubilizer. Due to the negative influence of high DMSO concentrations in the incubation mixture (see supplemental data), only the minimum amount of DMSO required for solubility was added. In the case of benzhydroxamic acid v_{\max} was about four times higher with mARC2 than with mARC1 but the K_M increased seven-fold. Only slight differences of the kinetic parameters for vorinostat and bufexamac were detected for both mARC forms, whereas CP544439 was exclusively reduced by mARC1. In accordance with the results obtained with subcellular liver fractions and RNAi studies all *N*-unsubstituted hydroxamic acids are reduced by mARC.

Discussion

The mARC-containing three component enzyme system is responsible for the reduction of various *N*-hydroxylated compounds (Ott et al., 2015) and the results of our recent investigations demonstrate clearly that hydroxamic acids belong to this class of compounds. The hydroxamate moiety exhibits strong cation chelating properties and thereby possesses the ability to affect a variety of enzymes. Hydroxamic acid moieties are found in a multitude of drugs and drug candidates (Halouska et al., 2014; Zhang and Zhong, 2014). In addition, aryl hydroxamic acids have been shown to possess toxic and mutagenic properties (Miller et al., 1961; Vaught et al., 1981). Therefore, the investigation of the metabolic fate of hydroxamic acids is of particular relevance for further drug developments and for the understanding of detoxification pathways. We could demonstrate that hydroxamic acids can serve as substrates for the mARC-system. The model compound benzhydroxamic acid, as well as three other drugs (vorinostat, bufexamac and CP544439) are reduced to the corresponding amides by porcine mitochondria and the reconstituted recombinant human mARC-system (Table 1). The *N*-reduction observed *in vitro* is also evident in intact human cell metabolism and is mARC-dependent as the siRNA-mediated down-regulation leads to a dramatic decrease in *N*-reductive activity (Fig. 4B). However, the *N*-reductive conversion of hydroxamic acids is lower compared to the model compound (benzamidoxime). Amidoximes, used as pro-drugs for amidines, are rapidly reduced *in vivo* (Clement et al., 1992). In the case of hydroxamic acids, reduction leads to inactivation because the amide is not able to form strongly chelating complexes. As metabolism studies for CP544439 reveal, the main metabolism pathways of hydroxamic acids *in vivo* are glucuronidation, reduction and hydrolysis (Dalvie et al., 2008). The reduction to the amide is of great physiological relevance, especially in rats, where the amide is the most prevalent metabolite. It has been demonstrated that the aldehyde oxidase is capable of reducing hydroxamic acids to amides (Sugihara et al., 1983a, 1983b; Sugihara and Tatsumi, 1986). Reduction of CP544439 was proposed to be catalyzed by this enzyme as well; studies with human cytosolic liver fractions using the artificial electron donor *N*-methylnicotinamide resulted in a very low conversion rate ($1.6 \text{ pmol} \cdot \text{min}^{-1} \cdot \text{mg protein}^{-1}$) (Obach, 2004). The aldehyde oxidase is located in the cytosol, but we could not detect any *N*-reductive activity in the porcine liver cytosol under our tested conditions (Fig. 5C). Our determined conversion rate with the mitochondrial fraction for CP544439 is about $2.5 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg protein}^{-1}$, which is more than 1500 times higher than the previous described cytosolic conversion rate. This finding indicates the important role of mARC in the

reduction of hydroxamic acids to the amides. Interestingly, vorinostat has sufficient metabolic stability for therapy and is applied orally, but high doses of 400 mg per day are necessary (Mann et al., 2007). The enhanced lipophilicity of vorinostat compared to benzhydroxamic acid and benzamidoxime could be a pivotal characteristic. The low conversion rates determined for bufexamac, which is a very lipophilic ether compound, support this theory (Fig. 5B). Studies to further elucidate structure activity relationships should be done to evaluate whether lipophilicity is an important feature for enzymatic conversion of hydroxamic acids by the mARC-system and to detect candidates with better stability towards *N*-reduction.

In the case of *N*-hydroxyphenacetin no reduction to the amide by mARC was observed. This is the first investigated hydroxamic acid studied by us so far which is not reduced. However, this could be an explanation for the high toxicity of this compound as it needs to be detoxified by other pathways. The most apparent difference between *N*-hydroxyphenacetin and the other studied hydroxamic acids is the substitution of the nitrogen's hydrogen with the sterically demanding phenyl group. Due to this steric hindrance mARC might not be able to bind and reduce such kind of substrates. This hypothesis needs to be proved by further structure activity studies.

In conclusion, to properly evaluate the metabolic stability of new hydroxamic acid containing drug candidates, metabolism by the mitochondrial mARC-system must be considered.

Acknowledgement

DAS and MJ received funding from the European Union's Seventh Framework Program for research, technological development and demonstration under grant agreements nr. 241865 (SEtTReND).

Thanks to Timothy Zerk for proofreading the article.

Authorship contribution

Participated in research design: Clement, Scheidig, Havemeyer

Conducted experiments: Ginsel, Plitzko

Contributed new reagents or analytical tools: Ginsel, Froriep, Stolfa, Jung

Expression and purification of recombinant proteins: Kubitza

Performed data analysis: Ginsel, Plitzko

Wrote or contributed to the writing of the manuscript: Ginsel, Plitzko, Clement

Figure legends

Figure 1. Metabolism of hydroxamic acids. Hydroxamic acids can either undergo reduction to the corresponding amide or hydrolysis to the corresponding carboxylic acid. A hydrolysis of the amide to the carboxylic acid might be possible.

Figure 2. Structures of all studied compounds.

Figure 3. *N*-reduction of benzhydroxamic acid, bufexamac, CP544439 and vorinostat in hepatic subcellular fractions. Biotransformation assay consisted of 100 µg protein, 1 mM NADH and either 1.0 mM vorinostat and 4.8 % DMSO, 3.0 mM benzhydroxamic acid, 0.5 mM CP544439 and 4.0 % DMSO or 1.0 mM bufexamac and 8.0 % DMSO. Incubation was carried out for 15 min and stopped by addition of methanol. Activities are means ± SD of two biological determinations. (A) benzhydroxamic acid (B) bufexamac, (C) CP544439, (D) vorinostat, * = under limit of quantification, n.d. not detectable

Figure 4. *N*-reductive metabolism of benzhydroxamic acid in HEK-293 cells. *N*-reductive activities of cells were determined as described in materials and methods. Results are presented as means ± S.D. (n=3). **(A)** Substrate dependent metabolism. Incubation time was 180 min. **(B)** Effect of mARC knockdown on *N*-reduction. HEK-293 cells were transfected with 20 nM mARC siRNA or non-targeting (NC) siRNA. The siRNA-mediated down-regulations of the proteins of interest were verified by western blot using anti-mARC1, anti-mARC2 or anti-calnexin antibody. Calnexin levels were used as loading control. mARC2-protein could not be detected. *N*-reductive activities were determined on day 4 after transfection. *** p < 0.001.

Figure 5. Substrate saturation curves. Biotransformation assay consisted of 7.5 µg hmARC, 75 pmol CYB5B and 7.5 pmol CYB5R. Incubation time was 15 minutes. Quantification was done by HPLC analysis. Activities are means ± S.D. of two biological determinations. (A) benzhydroxamic acid, (B) bufexamac, (C) CP544439, (D) vorinostat

Table 1: N-reduction of vorinostat, benzhydroxamic acid, bufexamac, CP544439, N-hydroxyphenacetin and benzamidoxime by the reconstituted recombinant mARC-system.

Biotransformation assays were carried out as described in material and methods. Activities are means \pm S.D. of two biological determinations.

^a LOQ = 0.9 nmol/(min \times mg protein), ^b LOQ = 3.5 nmol/(min \times mg protein).

Substrate		K_m (mM)	V_{max} (nmol amide/ (min \times mg protein))
vorinostat	mARC1	1.45 \pm 0.37	46.3 \pm 5.5
	mARC2	2.73 \pm 0.18	19.3 \pm 0.8
benzhydroxamic acid	mARC1	0.31 \pm 0.11	73.7 \pm 7.0
	mARC2	2.12 \pm 0.28	313.5 \pm 22.5
bufexamac	mARC1	1.09 \pm 0.15	13.4 \pm 1.1
	mARC2	1.07 \pm 0.15	8.2 \pm 0.7
CP544439	mARC1	0.25 \pm 0.06	50.2 \pm 4.4
	mARC2	-/-	-/- ^a
N-hydroxyphenacetin	mARC1	-/-	-/- ^b
	mARC2	-/-	-/- ^b
benzamidoxime	mARC1	0.63 \pm 0.06	674.8 \pm 25.1
	mARC2	0.54 \pm 0.03	549.6 \pm 12.1

Figure 1

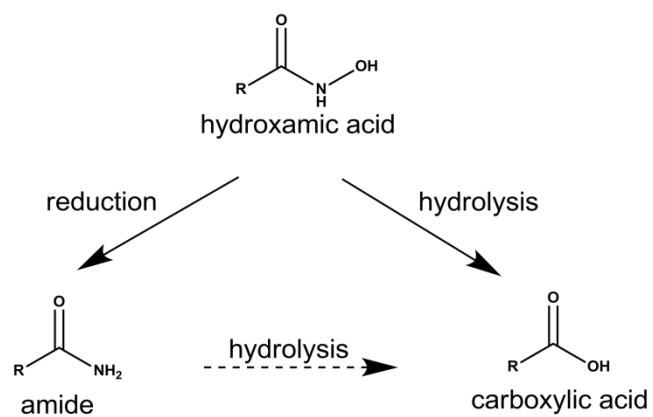


Figure 2

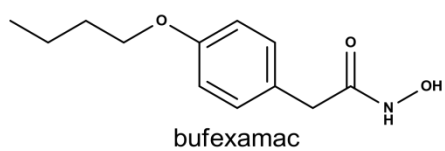
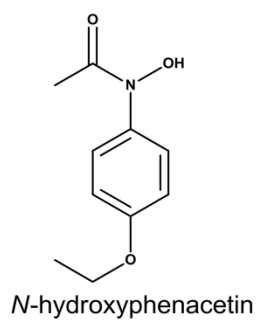
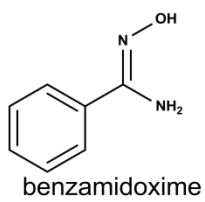
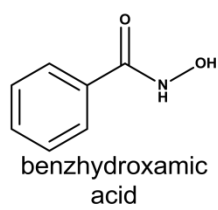
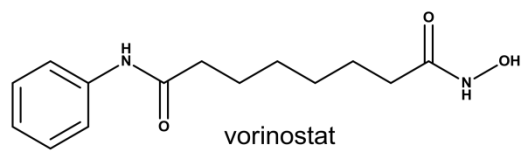
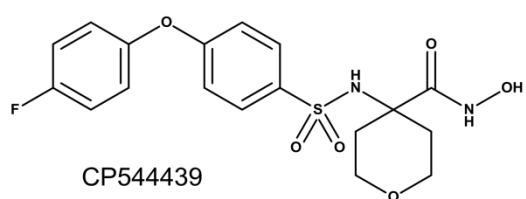


Figure 3

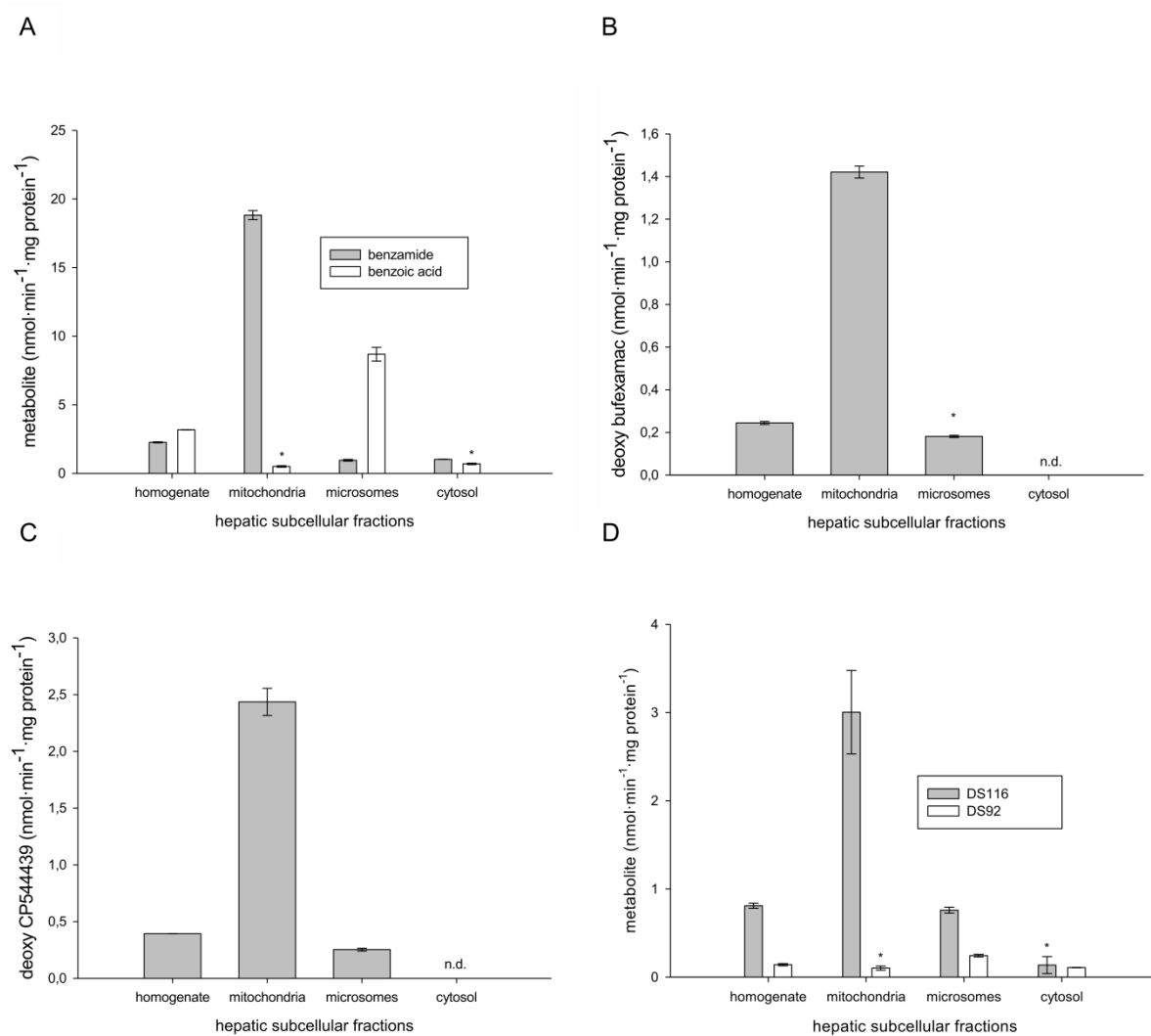


Figure 4

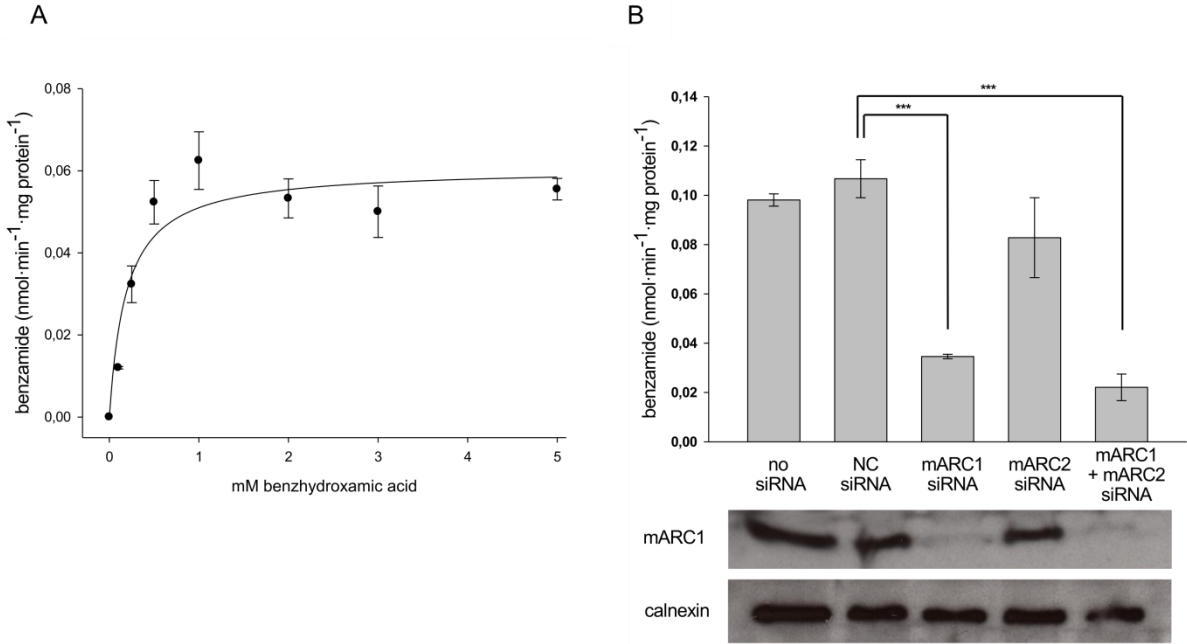
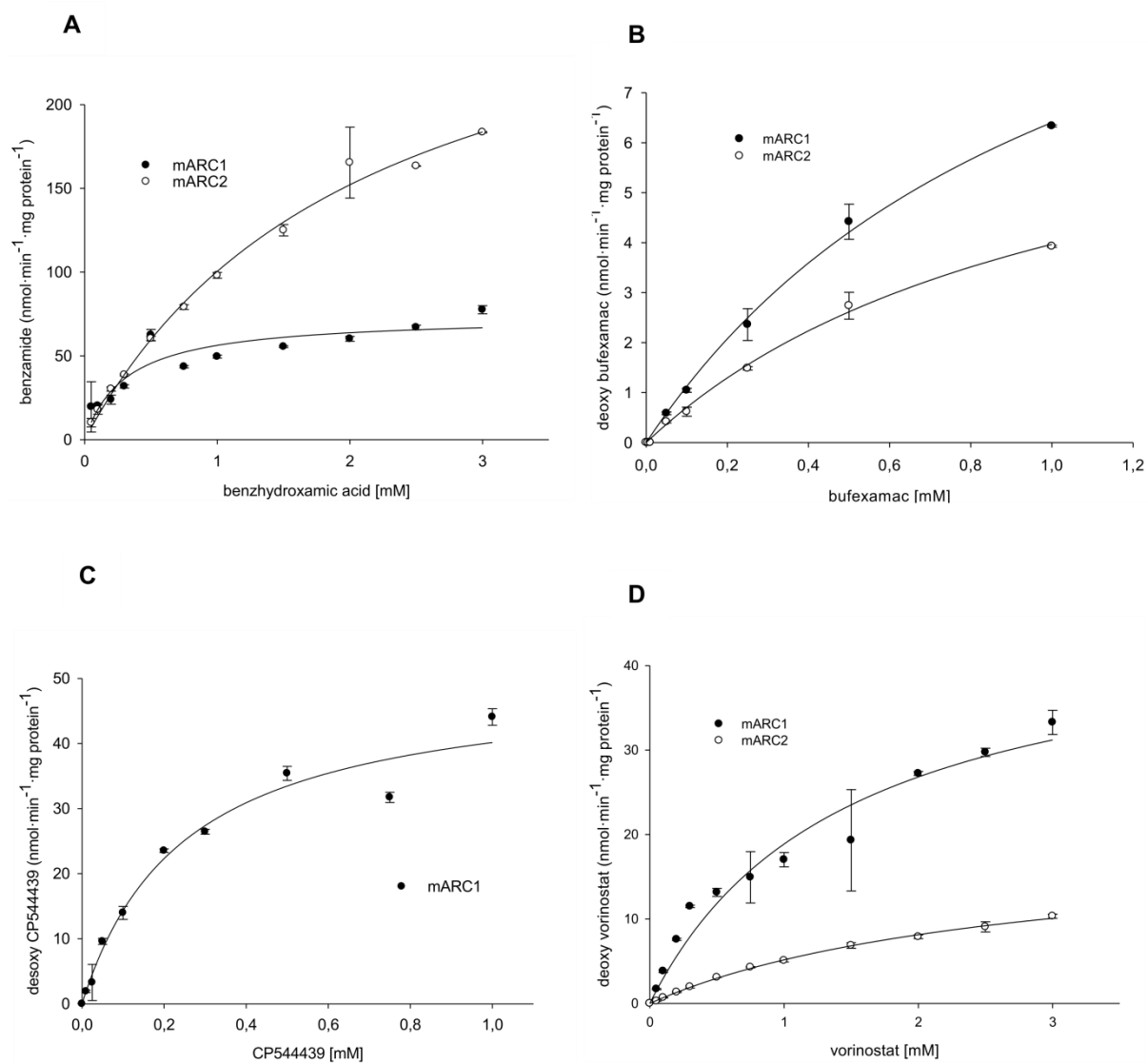


Figure 5



References

- Barb AW, McClerren AL, Snehelatha K, Reynolds CM, Zhou P, and Raetz CRH (2007) Inhibition of lipid A biosynthesis as the primary mechanism of CHIR-090 antibiotic activity in *Escherichia coli*. *Biochemistry* **46**:3793–3802 doi: 10.1021/bi6025165.
- Bengtsson U, Johansson S, and Angervall L (1978) Malignancies of the urinary tract and their relation to analgesic abuse. *Kidney Int.* **13**:107–113 doi: 10.1038/ki.1978.13.
- Brogden RN, Pinder RM, Sawyer PR, Speight TM, and Avery GS (1975) Bufexamac: a review of its pharmacological properties and therapeutic efficacy in inflammatory dermatoses. *Drugs* **10**:351–356.
- Clement B, Immel M, Terlinden R, and Wingen F-J (1992) Reduction of Amidoxime Derivatives to Pentamidine in vivo. *Arch. Pharm. (Weinheim)* **325**:61–62 doi: 10.1002/ardp.19923250114.
- Dalvie D, Cosker T, Boyden T, Zhou S, Schroeder C, and Potchoiba MJ (2008) Metabolism distribution and excretion of a matrix metalloproteinase-13 inhibitor, 4-(4-(4-fluorophenoxy)-benzenesulfonylamino)tetrahydropyran-4-carboxylic acid hydroxyamide (CP-544439), in rats and dogs: Assessment of the metabolic profile of CP-544439 in plasma and urine of humans. *Drug Metab. Dispos.* **36**:1869–1883 doi: 10.1124/dmd.108.022566.
- Dokmanovic M, Clarke C, and Marks PA (2007) Histone deacetylase inhibitors: Overview and perspectives. *Mol. Cancer Res.* **5**:981–989 doi: 10.1158/1541-7786.MCR-07-0324.
- Estabrook RW and Werringloer J (1978) The measurement of difference spectra: Application to the cytochromes of microsomes. *Methods Enzymol.* **52**:212–220.
- Fischbach T and Lenk W (1985) The metabolism of N-hydroxyphenacetin in vitro and in vivo. *Xenobiotica* **15**:915–927 doi: 10.3109/00498258509045045.
- Halouska S, Fenton RJ, Zinniel DK, Marshall DD, Barletta RG, and Powers R (2014) Metabolomics analysis identifies d-Alanine-d-Alanine ligase as the primary lethal target of d-Cycloserine in mycobacteria. *J. Proteome Res.* **13**:1065–1076 doi: 10.1021/pr4010579.
- Havemeyer A, Bittner F, Wollers S, Mendel R, Kunze T, and Clement B (2006) Identification of the missing component in the mitochondrial benzamidoxime prodrug-converting system as a novel molybdenum enzyme. *J. Biol. Chem.* **281**:34796–34802 doi: 10.1074/jbc.M607697200.
- Hinson JA and Mitchell JR (1976) N-Hydroxylation of phenacetin by hamster liver microsomes. *Drug Metab. Dispos.* **4**:430–435.
- Hirsch PF and Kaplan NO (1961) The conversion of pyridine hydroxamic acids to amides by mouse liver mitochondria. *J. Biol. Chem.* **236**:926–930.
- Jakobs HH, Mikula M, Havemeyer A, Strzalkowska A, Borowa-Chmielak M, Dzwonek A, Gajewska M, Hennig EE, Ostrowski J, and Clement B (2014) The N-reductive system composed of mitochondrial amidoxime reducing component (mARC), cytochrome b5 (CYB5B) and cytochrome b5 reductase (CYB5R) is regulated by fasting and high fat diet in mice. *PLoS ONE* **9**:e105371 doi: 10.1371/journal.pone.0105371.
- Kiesel BF, Parise RA, Tjørnelund J, Christensen MK, Loza E, Tawbi H, Chu E, Kummur S, and Beumer JH (2013) LC-MS/MS assay for the quantitation of the HDAC inhibitor belinostat and five major metabolites in human plasma. *J. Pharm. Biomed. Anal.* **81-82**:89–98 doi: 10.1016/j.jpba.2013.03.025.
- Kitamura S and Tatsumi K (1985) Purification of N-hydroxy-2-acetylaminofluorene reductase from rabbit liver cytosol. *Biochem. Biophys. Res. Commun.* **133**:67–74 doi: 10.1016/0006-291X(85)91842-X.
- Kotthaus J, Wahl B, Havemeyer A, Kotthaus J, Schade D, Garbe-Schönberg D, Mendel R, Bittner F, and Clement B (2011) Reduction of N(ω)-hydroxy-L-arginine by the mitochondrial amidoxime reducing component (mARC). *Biochem. J.* **433**:383–391 doi: 10.1042/BJ20100960.
- Krompholz N, Krischkowski C, Reichmann D, Garbe-Schönberg D, Mendel R-R, Bittner F, Clement B, and Havemeyer A (2012) The mitochondrial Amidoxime Reducing Component (mARC) is involved in detoxification of N-hydroxylated base analogues. *Chem. Res. Toxicol.* **25**:2443–2450 doi: 10.1021/tx300298m.
- Krüger P (1885) Ueber Abkömmlinge des Benzenylamidoxims. *Ber. Dtsch. Chem. Ges.* **18**:1053–1060 doi: 10.1002/cber.188501801219.

- Liu L, Detering J-C, Milde T, Haefeli WE, Witt O, and Burhenne J (2014) Quantification of vorinostat and its main metabolites in plasma and intracellular vorinostat in PBMCs by liquid chromatography coupled to tandem mass spectrometry and its relation to histone deacetylase activity in human blood. *J. Chromatogr. B. Analyt. Technol. Biomed. Life. Sci.* **964**:212–221 doi: 10.1016/j.jchromb.2014.02.014.
- Liu T, Smith GW, and Rankin JT (1972) Renal pelvic tumour associated with analgesic abuse. *Can. Med. Assoc. J.* **107**:768 passim.
- Lowenthal J (1954) Enzymatic Conversion of Salicylhydroxamic Acid to Salicylamide. *Nature* **174**:36–37 doi: 10.1038/174036a0.
- Mai A, Esposito M, Sbardella G, and Massa S (2001) A new facile and expeditious synthesis of N-hydroxy-N'-phenyloctanediamide, a potent inducer of terminal cytodifferentiation. *Org. Prep. Proced. Int.* **33**:391–394 doi: 10.1080/00304940109356608.
- Mann BS, Johnson JR, Cohen MH, Justice R, and Pazdur R (2007) FDA approval summary: Vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. *Oncologist* **12**:1247–1252 doi: 10.1634/theoncologist.12-10-1247.
- Miller EC, Miller JA, and Hartmann HA (1961) N-Hydroxy-2-acetylaminofluorene: a metabolite of 2-acetylaminofluorene with increased carcinogenic activity in the rat. *Cancer Res.* **21**:815–824.
- Neve EPA, Nordling A, Andersson TB, Hellman U, Diczfalusy U, Johansson I, and Ingelman-Sundberg M (2012) Amidoxime reductase system containing cytochrome b5 type B (CYB5B) and MOSC2 is of importance for lipid synthesis in adipocyte mitochondria. *J. Biol. Chem.* **287**:6307–6317 doi: 10.1074/jbc.M111.328237.
- Obach RS (2004) Potent inhibition of human liver aldehyde oxidase by raloxifene. *Drug Metab. Dispos.* **32**:89–97 doi: 10.1124/dmd.32.1.89.
- Ott G, Havemeyer A, and Clement B (2015) The mammalian molybdenum enzymes of mARC. *J. Biol. Inorg. Chem.* **20**:265–275 doi: 10.1007/s00775-014-1216-4.
- Ott G, Plitzko B, Krischkowski C, Reichmann D, Bittner F, Mendel RR, Kunze T, Clement B, and Havemeyer A (2014) Reduction of sulfamethoxazole hydroxylamine (SMX-HA) by the mitochondrial amidoxime reducing component (mARC). *Chem. Res. Toxicol.* **27**:1687–1695 doi: 10.1021/tx500174u.
- Plitzko B, Havemeyer A, Kunze T, and Clement B (2015) The pivotal role of the mitochondrial amidoxime reducing component 2 in protecting human cells against apoptotic effects of the base analog N6-hydroxylaminopurine. *J. Biol. Chem.* **290**:10126–10135 doi: 10.1074/jbc.M115.640052.
- Plitzko B, Ott G, Reichmann D, Henderson CJ, Wolf CR, Mendel R, Bittner F, Clement B, and Havemeyer A (2013) The involvement of mitochondrial amidoxime reducing components 1 and 2 and mitochondrial cytochrome b5 in N-reductive metabolism in human cells. *J. Biol. Chem.* **288**:20228–20237 doi: 10.1074/jbc.M113.474916.
- Raedler LA (2016) Farydak (Panobinostat): First HDAC Inhibitor Approved for Patients with Relapsed Multiple Myeloma. *Am Health Drug Benefits* **9**:84–87.
- Sparacino-Watkins CE, Tejero J, Sun B, Gauthier MC, Thomas J, Ragireddy V, Merchant BA, Wang J, Azarov I, Basu P, and Gladwin MT (2014) Nitrite reductase and nitric-oxide synthase activity of the mitochondrial molybdopterin enzymes mARC1 and mARC2. *J. Biol. Chem.* **289**:10345–10358 doi: 10.1074/jbc.M114.555177.
- Sugihara K, Kitamura S, and Tatsumi K (1983a) Evidence for reduction of hydroxamic acids to the corresponding amides by liver aldehyde oxidase. *Chem. Pharm. Bull.* **31**:3366–3369 doi: 10.1248/cpb.31.3366.
- Sugihara K, Kitamura S, and Tatsumi K (1983b) Involvement of liver aldehyde oxidase in conversion of N-hydroxyurethane to urethane. *J. Pharmacobio-Dyn.* **6**:677–683 doi: 10.1248/bpb1978.6.677.
- Sugihara K and Tatsumi K (1986) Participation of liver aldehyde oxidase in reductive metabolism of hydroxamic acids to amides. *Arch. Biochem. Biophys.* **247**:289–293.

- Suzuki T, Matsuura A, Kouketsu A, Hisakawa S, Nakagawa H, and Miyata N (2005) Design and synthesis of non-hydroxamate histone deacetylase inhibitors: identification of a selective histone acetylating agent. *Bioorg. Med. Chem.* **13**:4332–4342 doi: 10.1016/j.bmc.2005.04.002.
- Vaught JB, McGarvey PB, Lee MS, Garner CD, Wang CY, Linsmaier-Bednar EM, and King CM (1981) Activation of N-hydroxyphenacetin to mutagenic and nucleic acid-binding metabolites by acyltransfer, deacylation, and sulfate conjugation. *Cancer Res.* **41**:3424–3429.
- Verma RP (2012) Hydroxamic acids as matrix metalloproteinase inhibitors. *EXS* **103**:137–176 doi: 10.1007/978-3-0348-0364-9_5.
- Wagner JM, Hackanson B, Lübbert M, and Jung M (2010) Histone deacetylase (HDAC) inhibitors in recent clinical trials for cancer therapy. *Clin. Epigenet.* **1**:117–136 doi: 10.1007/s13148-010-0012-4.
- Wahl B, Reichmann D, Nix D, Krompholz N, Havemeyer A, Clement B, Messerschmidt T, Rothkegel M, Biester H, Hille R, Mendel RR, and Bittner F (2010) Biochemical and spectroscopic characterization of the human mitochondrial amidoxime reducing components hmARC-1 and hmARC-2 suggests the existence of a new molybdenum enzyme family in eukaryotes. *J. Biol. Chem.* **285**:37847–37859 doi: 10.1074/jbc.M110.169532.
- Wang K and Guengerich FP (2013) Reduction of aromatic and heterocyclic aromatic N-hydroxylamines by human cytochrome P450 2S1. *Chem. Res. Toxicol.* **26**:993–1004 doi: 10.1021/tx400139p.
- Whitby LG (1953) A new method for preparing flavin-adenine dinucleotide. *Biochem. J.* **54**:437–442.
- Wirth PJ, Dybing E, Bahr C von, and Thorgeirsson SS (1980) Mechanism of N-hydroxyacetylarylamine mutagenicity in the Salmonella test system: Metabolic activation of N-hydroxyphenacetin by liver and kidney fractions from rat, mouse, hamster, and man. *Mol. Pharmacol.* **18**:117–127.
- Wright SW and Corbett RL (1993) An Efficient Preparation of 2H-(5,4-b)Pyridoisothiazolone. *Org. Prep. Proced. Int.* **25**:247–249 doi: 10.1080/00304949309457957.
- Zhang J and Zhong Q (2014) Histone deacetylase inhibitors and cell death. *Cell. Mol. Life Sci.* **71**:3885–3901 doi: 10.1007/s00018-014-1656-6.

3.4 T4 Lysozyme-facilitated crystallization of the human molybdenum cofactor-dependent enzyme mARC

Christian Kubitza, Carsten Ginsel, Florian Bittner, Antje Havemeyer, Bernd Clement and Axel J. Scheidig

Acta Crystallographica Section F **2018**, 74 (article in press)

DOI: 10.1107/S2053230X18006921

The crystallization of a target protein marks the bottleneck of structure determination by X-ray crystallography. Trying to find the ideal conditions for a protein to crystallize is in most cases only achievable by a time-consuming, unpredictable trial-and-error approach with uncertain outcome. However, within the past decades some promising methods have been established in order to facilitate the crystallization of formerly non-crystallizable proteins. One of them is fusing the target protein to another protein that is easily crystallizable on its own and might drive the crystallization of the whole fusion protein by providing additional crystal contacts that cannot be formed by the target protein alone.

The following paper¹¹³ describes the crystallization strategy of the human mARC1 enzyme, which was not crystallizable by using conservative approaches. Therefore, the fusion protein strategy was applied by tethering hmARC1 to the phage enzyme T4 lysozyme (T4L), which is often used to crystallize proteins belonging to the family of G protein-coupled receptors. Based on *in silico* predictions of mARC secondary and tertiary structure elements, four fusion protein constructs were designed. Two of them had the T4L moiety fused to hmARC1 either *N*- or *C*-terminally. For the third and fourth construct, T4L was integrated into either a potential three-stranded antiparallel β -sheet or between two α -helices predicted for the hmARC1 moiety. The fusion proteins were expressed in *E. coli*, purified by affinity chromatography, assayed for catalytic activity and subjected to crystallization trials. Protein crystals were obtained for one of the fusion constructs and could be used for X-ray diffraction experiments with synchrotron radiation. Thereby, high-quality datasets were obtained, which formed the foundation of later structure determination.

I performed *in silico* predictions of hmARC secondary and tertiary structure elements, designed the fusion constructs, generated them using molecular biology techniques, performed the expression, purification, crystallization, diffraction experiments and data analysis. Furthermore, I wrote the manuscript together with Prof. Dr. Scheidig.



ISSN 2053-230X

Received 15 February 2018

Accepted 4 May 2018

Edited by N. Sträter, University of Leipzig, Germany

Keywords: T4 lysozyme; fusion protein; crystallization strategy; carrier-driven crystallization.



© 2018 International Union of Crystallography

T4 lysozyme-facilitated crystallization of the human molybdenum cofactor-dependent enzyme hmARC

Christian Kubitza,^a Carsten Ginsel,^b Florian Bittner,^c Antje Havemeyer,^b Bernd Clement^b and Axel J. Scheidig^{a*}

^aStructural Biology, Zoological Institute, Kiel University, Am Botanischen Garten 1–9, 24118 Kiel, Germany,

^bPharmaceutical Institute, Kiel University, Gutenbergstrasse 76, 24118 Kiel, Germany, and ^cJulius Kuehn Institute, Federal Research Centre for Cultivated Plants, Erwin-Baur-Strasse 27, 06484 Quedlinburg, Germany. *Correspondence e-mail: axel.scheidig@strubio.uni-kiel.de

The human mitochondrial amidoxime reducing component (hmARC) is a molybdenum cofactor-dependent enzyme that is involved in the reduction of a diverse range of N-hydroxylated compounds of either physiological or xenobiotic origin. In this study, the use of a fusion-protein approach with T4 lysozyme (T4L) to determine the structure of this hitherto noncrystallizable enzyme by X-ray crystallography is described. A set of four different hmARC-T4L fusion proteins were designed. Two of them contained either an N-terminal or a C-terminal T4L moiety fused to hmARC, while the other two contained T4L as an internal fusion partner tethered to the hmARC enzyme between two predicted secondary-structure elements. One of these internal fusion constructs could be expressed and crystallized successfully. The hmARC-T4L crystals diffracted to 1.7 Å resolution using synchrotron radiation and belonged to space group $P2_12_12_1$ with one molecule in the asymmetric unit. Initial attempts to solve the structure by molecular replacement using T4L did not result in electron-density distributions that were sufficient for model building and interpretation of the hmARC moiety. However, this study emphasizes the utility of the T4L fusion-protein approach, which can be used for the crystallization and structure determination of membrane-bound proteins as well as soluble proteins.

1. Introduction

Difficulties in obtaining diffraction-quality crystals of a target protein have always been a bottleneck in structure determination *via* X-ray crystallography and will probably remain a limiting factor in structural biology. In recent years, different tools and strategies have been developed to mediate the crystallization of challenging proteins which could not be crystallized by any conservative approach. These tools include surface-entropy reduction (Cooper *et al.*, 2007; Goldschmidt *et al.*, 2007), lysine methylation (Walter *et al.*, 2006), *in situ* proteolysis (Dong *et al.*, 2007) and metal-mediated crystallization (Laganowsky *et al.*, 2011). All of these manipulations slightly alter the surface characteristics of the target protein and aim to promote the formation of crystal contacts. Another, less subtle, strategy is chaperone-assisted crystallization. This approach relies on specific antibodies (or derivatives thereof) or other high-affinity binding-partner molecules to form stable, more rigid and, ultimately, crystallizable complexes with the target protein (Bukowska & Grütter, 2013).

A related strategy to the use of antibodies is based on the use of fusion proteins to assist crystallization. In principle, this approach has many advantages. The proteins are easily crystallizable on their own and provide additional surface area favourable for crystal lattice formation (Bell *et al.*, 2013), and

research communications

some fusion partners can also be used as affinity tags for purification. The use of fusion tags has also been shown to enhance the solubility of the target protein and might therefore favour higher expression yields (Stevens, 2000). Ultimately, the three-dimensional structures of the fusion partners can be used as starting models to solve the phase problem by molecular replacement (Niemann *et al.*, 2001). This is of especially high value if no structural homologues of the target protein have been determined. However, there are also some drawbacks to be considered in using this strategy. First of all, the fusion partner might interfere with the three-dimensional arrangement of the target protein. This might result in non-natural altered conformations or even loss of function. Furthermore, the linker sequence between the fusion partners has to be chosen carefully. It needs to provide a certain degree of rigidity to avoid conformational heterogeneity, which would otherwise have negative effects on crystallization (Kobe *et al.*, 2015).

Several protein structures have successfully been determined using a fusion-protein-assisted crystallization approach. The fusion proteins used for this strategy include glutathione *S*-transferase (GST; Kuge *et al.*, 1997), thioredoxin (TRX; Corsini *et al.*, 2008), green fluorescent protein (GFP; Suzuki *et al.*, 2010), barnase (Niemann *et al.*, 2006), an engineered sterile- α motif (SAM; Nauli *et al.*, 2007), maltose-binding protein (MBP; Kobe *et al.*, 1999) and T4 lysozyme (T4L; Rosenbaum *et al.*, 2007). Although the fusion-protein strategy does not seem to be applicable to all target proteins of choice, the number of crystal structures solved by this approach is increasing owing to the fact that intensive research is being conducted in order to improve the method. One of the most thoroughly investigated systems is the MBP-mediated crystallization approach, which is recommended if the target protein starts with an N-terminal α -helix (Jin *et al.*, 2017).

The use of T4L as a crystallization-facilitating fusion partner became popular with the structure determination of the human β_2 -adrenergic G-protein-coupled receptor (GPCR; Cherezov *et al.*, 2007), and has since been successfully applied to a number of different GPCRs (see, for example, Doré *et al.*, 2014; Chien *et al.*, 2010; Haga *et al.*, 2012; Miller-Gallacher *et al.*, 2014; Srivastava *et al.*, 2014; Wu *et al.*, 2010). To date, more than 90 crystal structures of T4L fusion proteins have been deposited in the PDB. Although the majority of the deposited protein structures belong to the membrane-protein family of GPCRs, there are a few individual cases of soluble proteins which could also be structurally characterized using this approach (see, for example, Bhabha *et al.*, 2014; Scott *et al.*, 2017; Baumlova *et al.*, 2014). T4L was found to be an optimal fusion partner, since it is a well folded soluble protein which can be crystallized under many conditions. Even more importantly, both termini of this protein are in close proximity to each other and therefore allow internal fusion into a target protein, tethering the T4L at two ends within a loop. This greatly reduces the risk of conformational heterogeneity, since the fusion partners can be quite rigidly linked to each other. However, Zou and coworkers were able to show that even N-terminal T4L fusions can facilitate crystallization, at least in the case of the aforementioned β_2 -adrenergic GPCR (Zou *et al.*, 2012). Thorsen and coworkers recently described two ways to maximize T4L rigidity and to improve the utility of the T4L fusion approach (Thorsen *et al.*, 2014). They introduced two disulfide bridges in the interface between the lobes or, in another approach, created a 'minimal T4L' which no longer contains the N-terminal lobe. Several other reported T4L crystal structures contain molecules between two lobes of this protein which are derived from the crystallization buffer and decrease the overall flexibility.

Here, we report the crystallization of a soluble protein which contains unmodified T4L as an internal fusion partner tethered between two strands of a predicted antiparallel β -sheet. The aim was the structure determination of the human mitochondrial amidoxime-reducing component (hmARC). This enzyme was discovered in 2006 as a fourth human molybdenum cofactor (Moco)-containing enzyme in addition to sulfite oxidase, aldehyde oxidase and xanthine oxidoreductase (Havemeyer *et al.*, 2006). In the presence of NADH, mARC proteins exert N-reductive activity towards N-hydroxylated substrates together with the two electron-transport proteins cytochrome b_5 and NADH cytochrome b_5 reductase. This enzyme system is located at the outer

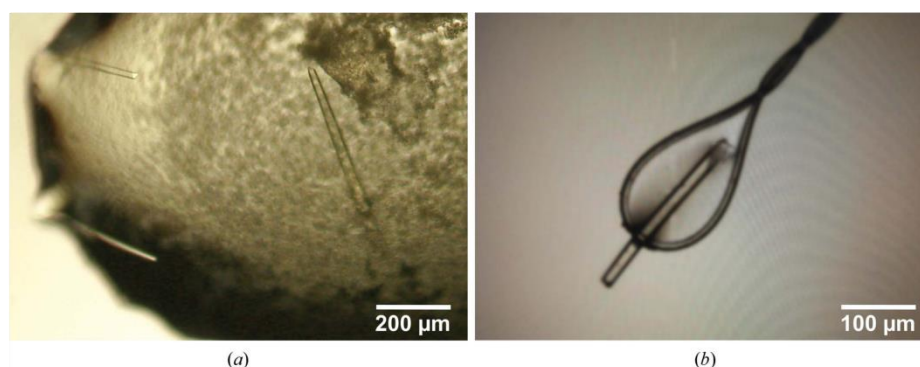


Figure 1 Spherulite formation in hmARC1 crystallization setups. (a) Spherulites obtained in an initial crystallization trial. (b) Spherulites obtained by streak-seeding.

research communications

mitochondrial membrane and is expressed in every tissue studied to date (Ott *et al.*, 2015). mARC enzymes belong to a subfamily of Moco sulfuryase C-terminal domain (MOSC)-like proteins which has not yet been structurally characterized. Here, we present our strategy to generate T4L-hmARC fusion proteins for crystallization purposes.

2. Materials and methods

2.1. Preliminary design of soluble hmARC constructs and crystallization trials

Human mARC proteins are physiologically expressed with a predicted N-terminal transmembrane helix, potentially anchoring these proteins to the outer mitochondrial membrane. There are two isoforms, hmARC1 and hmARC2, which have a sequence identity of 65.7% (Wahl *et al.*, 2010). A soluble, N-terminally truncated variant of hmARC1 was designed and expressed for *in vitro* studies and crystallization purposes. In addition to the N-terminal truncation of the first 52 residues, this construct was equipped with a C-terminal His₆ tag, which was used for affinity chromatography. Several

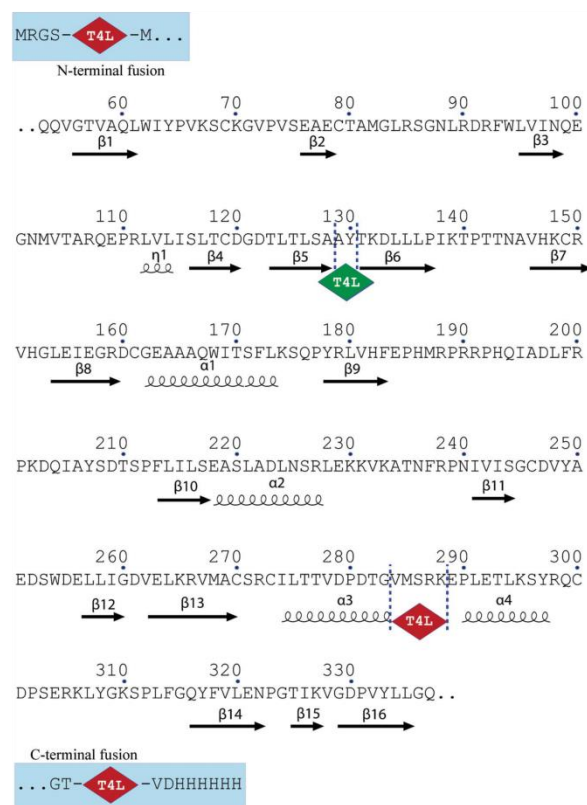


Figure 2
Secondary-structure prediction for hmARC1. T4L insertion sites for the generation of the fusion proteins are indicated by red or green rhombi. Red, constructs resulting in non-expressible or inactive fusion proteins; green, construct resulting in a functional and crystallizable fusion protein.

crystallization approaches and the setting up of various commercial crystallization screens with different protein concentrations did not result in the formation of protein crystals. At best, spherulites could be obtained (Fig. 1a), but neither additive screening nor seeding techniques (Fig. 1b) were sufficient to improve these initial conditions to yield diffraction-quality protein crystals.

2.2. Modelling of hmARC1 and *in silico* design of T4L fusion constructs

Based on recent structural publications for challenging proteins, such as GPCRs, we aimed at the rational design of fusion proteins comprised of hmARC1 (33.5 kDa) and phage T4 lysozyme (T4L; 18.3 kDa) in order to facilitate crystallization. Since the N- and C-terminal residues of T4L are in close proximity, this protein is well suited for integration between two secondary-structure elements of the target protein. This enhances the chance of a more rigid linkage between the fusion partners, which might be essential for crystallization. Therefore, secondary-structure prediction and homology modelling (Fig. 2) were carried out for hmARC1 using *MODELLER* within the *HHpred* server (Söding, 2005; Söding *et al.*, 2005) as well as the *Phyre²* protein-fold recognition server (Kelley *et al.*, 2015). There are no close structural homologues to the hmARC enzymes, but other family members of the MOSC family, such as YuaD from *Bacillus subtilis* (PDB entry 1oru; Midwest Center for Structural Genomics, unpublished work) and a MOSC_N domain-containing protein of unknown function (PDB entry 2exn; Rossi *et al.*, 2005), share sufficient sequence identity to perform initial secondary- and tertiary-structure predictions. All generated models were analyzed for common features as well as structural differences. The most promising elements were merged into a final model, which was used to determine suitable insertion sites for the T4L fusion partner. It should be noted that this model did not represent the whole hmARC enzyme, but contained a significant portion of unmodelled gaps owing to the lack of homologues. However, some secondary-structure elements were predicted with high confidence. We designed four different hmARC1-T4L fusion proteins, with two of them having T4L fused to the hmARC1 sequence either N- or C-terminally (named N-T4L-hmARC1 and hmARC1-T4L-C, respectively). Additionally, we decided to integrate T4L into a potential three-stranded antiparallel β-sheet, replacing a predicted two-residue β-turn (residues Ala129 and Tyr130) with lysozyme (hmARC1-T4L-β). The fourth construct (hmARC1-T4L-α) had T4L internally fused between two predicted α-helices, replacing residues Val284–Lys288.

2.3. Cloning

The N-terminally truncated hmARC1 construct with a C-terminal His₆ tag (hmARC1_{N-del}) was present in a modified pQE80L plasmid (without its original 5' His₆ tag-coding sequence), where it was cloned between BamHI and KpnI restriction sites. To generate the terminal T4L fusion

research communications

Table 1
Oligonucleotide primers used for cloning.

Bold, restriction sites; underlined, complementary to T4L-coding sequence; italic, complementary to the hmARC1-coding sequence at the target site.

N-T4L-hmARC1_fwd	TAGCT GGATCC ATGAATATATTTGAAATG
N-T4L-hmARC1_rev	AGCTA GGATCC ATACGCGTCCCAAGTGC
hmARC1-T4L-C_fwd	TAGCT GGTACC ATGAATATATTTGAAATG
hmARC1-T4L-C_rev	AGCTA GTCGAC ATACGCGTCCCAAGTGC
hmARC1-T4L- α _fwd	GCATTTTAACCACAGTGGACCCAGACACCGGTAT GAATATATTTGAAATG
hmARC1-T4L- α _rev	GGCGATAACTCTTCAGTGTTCACGCGGTTTCATA CGCGTCCCAAGTGC
hmARC1-T4L- β _fwd	CGGATGGTGACACCTTGACTCTCAGTCAATGAA TATATTTGAAATG
hmARC1-T4L- β _rev	CGGTTTTGATAGGCAGTAGTAGGTCCTTTGTATA CGCGTCCCAAGTGC

constructs, extension polymerase chain reactions (PCRs) were performed with the T4L-coding sequence (as provided by Addgene plasmid 18110) as a template to introduce BamHI restriction sites (for N-terminal fusion) or KpnI and SalI restriction sites (for C-terminal fusion). Subsequently, the PCR products were digested with the respective restriction enzymes and subcloned into the hmARC1-containing pQE80L vector to generate the N-T4L-hmARC1 and hmARC1-T4L-C constructs. The hmARC1-T4L- α and hmARC1-T4L- β fusion proteins were generated by overlap extension PCR within a restriction-free cloning approach (Bond & Naus, 2012). This allows the insertion of any sequence into any position within a template plasmid, yet this technique is independent of restriction-enzyme recognition sites. Therefore, a pair of hybrid primers were designed which contained complementary sequences to both the T4L insert and the hmARC1-coding sequence at the insertion site on the target plasmid. These primers were used to amplify the T4L-coding sequence from a source vector under high-fidelity conditions in a primary PCR (see Table 1). The resulting product was purified and used as a megaprimer in a secondary PCR with the hmARC1-coding target plasmid as a template. Following the secondary PCR, any template plasmid was degraded using the DpnI restriction enzyme and the purified product was ligated overnight. All fusion constructs were subsequently transformed into competent *Escherichia coli* XL1 Blue and TP1000 cells (Palmer *et al.*, 1996).

2.4. Protein expression and purification

The soluble, C-terminally His₆-tagged hmARC1-T4L fusion proteins were heterologously expressed in *E. coli* TP1000 cells (see Table 2). This protein-expression strain is unable to synthesize the dinucleotide molybdenum cofactor, but rather enriches the mononucleotide Moco as found within mARC proteins (Palmer *et al.*, 1996). A volume of 13 ml from an overnight culture was used to inoculate 2 l LB medium supplemented with 5 mM ammonium chloride, 1 mM sodium molybdate and 130 $\mu\text{g ml}^{-1}$ ampicillin, resulting in an OD₆₀₀ of approximately 0.02. The bacterial culture was incubated at 310 K and shaken at 90 rev min⁻¹ until the OD₆₀₀ reached 0.1. Protein expression was induced with freshly prepared 15 μM isopropyl β -D-1-thiogalactopyranoside (IPTG) solution and

Table 2
Protein-production information for hmARC1-T4L- β .

UniProt accession No.	Q5VT66 (hmARC1), D9IEF7 (T4 lysozyme)
Source organism	Human (hmARC1), <i>Enterobacteria phage T4</i> (T4 lysozyme)
Expression vector	pQE80
Expression host	<i>E. coli</i> TP1000
Amino-acid sequence of the construct for crystallization†	MRGSMQQVGTVAQLWLYPVKSKGVVSEAE ECTAMGLRSGNLRDRFNLVINQEGNMVT ARQEPRLVLI SLTCDGDTLLSANNIFE MLRIDEGRLRLKIYKDTGEGYITIGIHL TKSPSLNAAKSELDKAI GRNCGNIVITKD EAEKLFNQDVAAVRGILRNALKPVYD SLDAVRRCALINMVFQMGETGVAGFTNS LRMLQQKRWDEAAVNLA KSRWYNQTPNR AKRVIITTFRTGTWDAYTKDLLLPKTP TNAVHKCRVHGLEIEGRDCGEEAAQWIT SFLKSQPYRLVHFEFPMRPRRPHQIADL FRPKDQIAYS DTS PFLILSEASLADLNS RLEKKVKATNFRPNIVISGCDVVAEDSW DELLIGDVELKRVMACSRCLITVDPDT GVMSRKEPLETLKSYRQCDPSEKLYGK SPLFGQYFVLENPGT IKVGD PVYLLGQG TVDDHHHHHH
Molecular weight (kDa)	51.8 (hmARC1-T4L- β), 33.5 (hmARC1 portion), 18.3 (T4L portion)

the temperature was decreased to 295 K. After 20 h, the cells were harvested at 10 000g. Bacterial pellets were resuspended in running buffer (50 mM sodium dihydrogen phosphate, 300 mM sodium chloride, 25 mM sodium molybdate, 10 mM imidazole pH 8.0), lysed using an EmulsiFlex-C3 (Avestin, Mannheim, Germany) and subsequently centrifuged at 75 600g for 1 h. His₆-tagged recombinant proteins were purified from the resulting crude extract *via* immobilized metal ion affinity chromatography (IMAC) using a HisTrap HP 5 ml column and an ÄKTApurifier FPLC system (both from GE Healthcare, Freiburg, Germany). After loading the crude extract, nonspecifically bound proteins were washed from the nickel column with 20 mM imidazole in running buffer. The target protein was eluted from the column by applying a one-step elution gradient to 125 mM imidazole in running buffer. Elution fractions were pooled, concentrated and subjected to a 5 ml HiTrap Desalting column (GE Healthcare, Freiburg, Germany) in order to exchange the buffer for a low-salt buffer (50 mM Tris-HCl, 5 mM sodium chloride pH 7.0) for subsequent ion-exchange chromatography (IEC). The protein was loaded onto a 1 ml HiTrap SP XL column (GE Healthcare, Freiburg, Germany). Impurities that were still present after IMAC purification were usually washed through the column, while the target protein was bound with high affinity. Fractions of pure fusion protein were eluted *via* a linear gradient over 20 column volumes to a final concentration of 500 mM sodium chloride. Elution fractions containing the fusion protein were pooled, concentrated to 14.8 mg ml⁻¹ and supplemented with 7% (v/v) glycerol prior to storage at 193 K.

2.5. HPLC-based activity assay

In order to verify the enzymatic activity of the hmARC1-T4L fusion proteins, an HPLC assay was performed. In concert with the flavin-containing NADH cytochrome *b*₅

Table 3
Crystallization.

Method	Sitting drop
Plate type	96-well
Temperature (K)	291
Protein concentration (mg ml ⁻¹)	14.8
Buffer composition of protein solution	50 mM Tris-HCl pH 7.0, 250 mM NaCl, 7% glycerol
Composition of reservoir solution	100 mM bis-tris propane-HCl pH 6.5, 200 mM Na ₂ MoO ₄ , 27.5% PEG 3350
Volume and ratio of drop	200 nl protein solution, 1:1 ratio
Volume of reservoir (μl)	50

reductase and the haem-containing cytochrome *b₅*, hmARC is able to reduce N-hydroxylated structures such as the model substrate benzamidoxime, which is converted to benzamidine (Havemeyer *et al.*, 2006). The incubation mixture consisted of 3.75 μg hmARC1 (or 5.77 μg fusion protein), 37.5 pmol cytochrome *b₅* and 3.75 pmol cytochrome *b₅* reductase in 20 mM MES buffer pH 6.0. The benzamidoxime concentration was set to 3 mM. After pre-incubation for 3 min at 310 K in a shaking water bath, the reaction was started under aerobic conditions by the addition of 1 mM NADH, yielding a total volume of 150 μl. Incubation was carried out for 15 min at 310 K and was stopped with 150 μl cold methanol, followed by 5 min shaking and 5 min centrifugation at 9500g. An isocratic HPLC method was used for the quantification of benzamidine. The samples were separated on a LiChrospher 60 RP-Select B (5 μm) 250-4 column with a RP-Select B 4 × 4 mm guard column at 295 K. The mobile phase consisted of 10 mM sodium 1-octanesulfonate, 20% acetonitrile. The flow rate was set to 1 ml min⁻¹ and the detection wavelength to 229 nm. The retention times were 7.6 ± 0.2 min for benzamidoxime and 16.4 ± 0.3 min for benzamidine.

2.6. Crystallization of hmARC1-T4L fusion proteins

Initial crystallization experiments with the commercial screens JCSG-*plus* and PACT-*premier* (Molecular Dimensions, Suffolk, England) were carried out using the sitting-drop vapour-diffusion method at 291 K. 200 nl protein solution (14.8 mg ml⁻¹ in 50 mM Tris-HCl, 250 mM NaCl, 7% glycerol pH 7.0) was mixed with an equal amount of precipitant solution using a Mosquito HTS pipetting robot and 96-well sitting-drop plates (both from TTP Labtech, Melbourn, England). Crystals of the hmARC1-T4L-β construct were obtained with a precipitant condition consisting of 100 mM bis-tris propane-HCl pH 6.5, 20% (w/v) PEG 3350 and 200 mM of various organic or inorganic salts, such as sodium malonate, lithium sulfate, sodium sulfate or ammonium sulfate. To improve the crystal size and quality, 96-well grid screens were set up in order to fine-tune the initial identified crystallization conditions. Different salts (at a concentration of 200 mM), including sodium molybdate, were screened using MES-NaOH or bis-tris propane-HCl buffers at different pH values (5.0–8.0) as well as different concentrations of PEG 3350 or PEG 4000 [15–35% (w/v)]. Crystals were taken out of the crystallization drop using nylon loops and were flash-cooled by transfer into liquid nitrogen without adding

Table 4
Data collection and processing.

Values in parentheses are for the highest resolution shell.	
Diffraction source	P14, PETRA III, Hamburg
Wavelength (Å)	0.9789
Temperature (K)	100
Detector	PILATUS 6M
Crystal-to-detector distance (mm)	188.54
Rotation range per image (°)	0.1
Exposure time per image (s)	0.00146
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁
<i>a</i> , <i>b</i> , <i>c</i> (Å)	61.1, 74.8, 110.7
α , β , γ (°)	90.0, 90.0, 90.0
Resolution range [†] (Å)	42.52–1.65 (1.68–1.65)
Total No. of reflections [†]	723415 (18654)
No. of unique reflections [†]	60665 (2860)
Completeness [†] (%)	98.4 (95.9)
Multiplicity [†]	11.9 (6.5)
$\langle I/\sigma(I) \rangle$ [†]	9.0 (0.4)
Mean $\langle I \rangle$ half-set correlation <i>CC</i> _{1/2} [†]	0.99 (0.15)
<i>R</i> _{merge} [†]	0.21 (3.85)
<i>R</i> _{meas} [†]	0.22 (4.17)
<i>R</i> _{p.i.m.} [†]	0.06 (1.59)

[†] Values as provided by *AIMLESS* after data processing, merging and scaling.

additional cryoprotectant. The most important information on crystallization is summarized in Table 3.

2.7. Data collection, processing and phasing

X-ray diffraction data were collected at 100 K on the EMBL MX beamline P14 at PETRA III, DESY, Hamburg, Germany equipped with a PILATUS 6M detector. High-resolution data sets with sufficient quality could only be obtained by combining a double-focused beam (5 × 10 μm) at 100% transmission with helical data collection along the longitudinal axis of the crystals in order to minimize radiation damage. Diffraction data were indexed and integrated using *XDS* (Kabsch, 2010). Space-group determination, data scaling and merging were performed using *AIMLESS* (Evans & Murshudov, 2013) as part of the *CCP4* suite (Winn *et al.*, 2011), while applying the free *R* flag to 5% of the reflections. Data-collection and processing statistics are summarized in Table 4. Since there were neither sufficient anomalous scattering data derived from potentially bound molybdenum nor any structural homologue of hmARC, phasing approaches were limited to molecular replacement (MR) using T4 lysozyme as a search template within *MOLREP* (Vagin & Teplyakov, 2010).

3. Results and discussion

N-terminally truncated hmARC1 was cloned into a modified pQE80 vector providing a C-terminal His₆ tag. This hmARC1_N-del construct was used as a template to generate fusion proteins comprised of hmARC and T4 lysozyme. In order to reduce conformational heterogeneity of the fusion proteins, T4L was inserted into short loop regions between two predicted α -helices (hmARC1-T4L- α) or between two β -strands of a predicted three-stranded β -sheet (hmARC1-T4L- β). The internal insertion sites were chosen based on secondary- and tertiary-structure predictions performed by

research communications

homology-modelling servers. While proteins with the T4L moiety as an N- or C-terminal fusion partner could not be overexpressed in *E. coli*, fusion proteins containing T4L as an internal fusion partner within the hmARC1 sequence were expressed as soluble proteins with sufficient yields (1–2 mg l⁻¹). Purification *via* affinity and ion-exchange chromatography resulted in protein fractions of high purity.

The enzyme activity of the fusion proteins was determined by benzamidoxime turnover within an HPLC-based activity assay. The hmARC1-T4L- α construct did not exert any detectable activity at all. It is most likely that the T4L moiety induced conformational changes within the tertiary structure of hmARC, rendering the enzyme inactive. Likewise, the internal fusion partner might either interfere with the active site of the enzyme or inhibit electron transfer between hmARC and its redox partners. Either way, this construct was discarded from further studies and crystallization trials because it did not reflect the hmARC enzyme under near-physiological conditions. However, the enzyme activity of the hmARC1-T4L- β construct (536 \pm 20 nmol benzamidoxime per minute per milligram of protein) was comparable to that of soluble hmARC1 (777 \pm 8 nmol benzamidoxime per minute per milligram of protein). Despite the bulky fusion partner, this enzyme still possesses \sim 70% of the specific activity of the N-terminally truncated wild-type enzyme and was therefore regarded as a promising candidate for crystallization and structure determination.

Initial screens identified conditions containing 100 mM bis-tris propane adjusted to various pH values, 20% (w/v) PEG 3350 and various salts at 200 mM to be suitable for the crystallization of hmARC1-T4L- β . These initial conditions were refined in grid screens in order to improve crystal size and quality. The optimized crystallization conditions (100 mM bis-tris propane pH 6.5, 200 mM sodium molybdate, 27.5% PEG 3350) yielded rod-shaped protein crystals with a length of 200–400 μ m but with a diameter of only 10–15 μ m (Fig. 3). Of note, only precipitating solutions containing bis-tris propane buffer produced protein crystals. Apart from its buffering capabilities, this molecule might serve as a bidentate additive that is necessary to stabilize the two flexible lobes of the T4L moiety.

Similar conditions containing MES buffer at the same pH did not result in crystal formation.

Crystals of the hmARC-T4L- β construct diffracted to 2.5–1.7 Å resolution using double-focused synchrotron radiation. These crystals were very susceptible to radiation damage, which hindered the collection of complete data sets using traditional data-collection strategies. However, given the length of the crystals and their homogenous diffraction quality, they could be subjected to a continuous helical data-collection strategy ('4D-scan'), which was achieved on beamline P14 at EMBL Hamburg. Complete data sets of very good quality were obtained.

Owing to the lack of structural homologues of the hmARC1 enzyme, phasing was limited to MR using the T4L moiety as a search template. Indeed, the position of the T4L could be found using *MOLREP* as implemented within the *CCP4* suite, and convincing electron-density maps could be obtained for this part of the fusion protein using rigid-body as well as restraint refinement steps in *REFMAC5* (Murshudov *et al.*, 2011; Fig. 4). However, the derived phases and calculated electron-density distributions were not sufficient to build and refine the hmARC1 enzyme. Using the hmARC1 model generated by *MODELLER* or *Phyre*² as an MR search template did not result in convincing solutions. Therefore, additional data will be needed in order to overcome the phase problem and solve the crystal structure of hmARC1. Experimental phasing methods, such as single-wavelength or multi-wavelength anomalous dispersion, will have to be applied to selenomethionine- or heavy-atom-derivatized protein crystals and may be combined with the phases already obtained by MR using T4L as a search template. Alternatively, a better hmARC1 homology model (or at least of parts of the enzyme) will have to be generated and used as an MR template in addition to the already fixed T4L position. This way, phases could be obtained and improved in an iterative approach in order to generate electron-density distributions that are sufficient for model building and interpretation of the crystal structure of hmARC1.

This study emphasizes the utility of the T4L fusion-protein approach, which is not only useful for the crystallization and

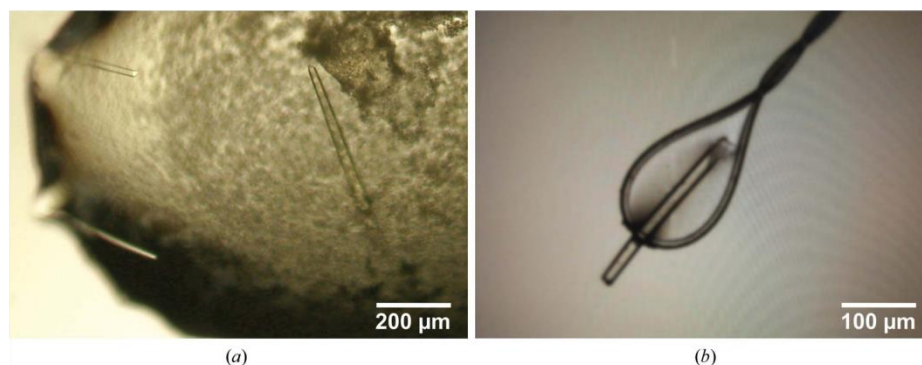


Figure 3 Crystals of hmARC1-T4L- β . (a) Crystals grown in 100 mM bis-tris propane-HCl pH 6.5, 20% (w/v) PEG 3350, 200 mM sodium molybdate. (b) A harvested crystal in a nylon loop.

research communications



Figure 4
Electron-density map showing T4L molecules, contoured at 2.0σ . Phase information was obtained by molecular replacement using T4L as the template (PDB entry 2061; Blaber *et al.*, 1993). Owing to insufficient phase information, hmARCl molecules are not yet distinguishable within the unit cell.

structure determination of GPCR membrane proteins but can also be applied to soluble proteins. The ability to use T4L as an internal fusion partner by replacing a small loop of the target protein makes this strategy applicable to a wide range of different proteins. Here, we show that it might be highly advantageous to choose an insertion site within a predicted β -sheet, since this greatly reduces the risk of conformational heterogeneity of the fusion protein. Therefore, we propose that the T4L fusion-protein approach should be investigated with a number of different proteins for which traditional strategies have failed to yield good crystals. We were able to show that even unmodified full-length T4L is suitable as a fusion partner in order to obtain high-quality crystals with good diffraction properties. Furthermore, we propose that the utility of T4L fusion proteins in structural biology has not yet been fully exploited and might contribute to significant advancements in fusion-driven protein crystallization.

Acknowledgements

The plasmid containing the coding sequence for wild-type T4 lysozyme was obtained from Addgene (plasmid 18110) as deposited by Brian Matthews (Eugene, USA). We gratefully acknowledge access to the core facilities of the BiMo/LMB of Kiel University. We thank Felix Helfrich for helpful discussions and support during data collection. Diffraction data were collected on beamline P14 operated by EMBL at the PETRA III storage ring. We are grateful to the beamline staff for providing assistance in using the beamlines.

Funding information

The research leading to these results received funding from the European Community's Seventh Framework Programme

(FP7/2007–2013) under BioStructX (grant agreement No. 283570). Additionally, we are grateful for access to the HTX crystallization facility by means of a grant from P-Cube and BioStructX at the EMBL Outstation Hamburg. Beamtime at P14 at the EMBL Outstation Hamburg was also funded by a BioStructX grant.

References

- Baumlova, A., Chalupska, D., Różycki, B., Jovic, M., Wisniewski, E., Klima, M., Dubankova, A., Kloer, D. P., Nencka, R., Balla, T. & Boura, E. (2014). *EMBO Rep.* **15**, 1085–1092.
- Bell, M. R., Engleka, M. J., Malik, A. & Strickler, J. E. (2013). *Protein Sci.* **22**, 1466–1477.
- Bhabha, G., Cheng, H.-C., Zhang, N., Moeller, A., Liao, M., Speir, J. A., Cheng, Y. & Vale, R. D. (2014). *Cell*, **159**, 857–868.
- Blaber, M., Lindstrom, J. D., Gassner, N., Xu, J., Heinz, D. W. & Matthews, B. W. (1993). *Biochemistry*, **32**, 11363–11373.
- Bond, S. R. & Naus, C. C. (2012). *Nucleic Acids Res.* **40**, W209–W213.
- Bukowska, M. A. & Grütter, M. G. (2013). *Curr. Opin. Struct. Biol.* **23**, 409–416.
- Cherezov, V., Rosenbaum, D. M., Hanson, M. A., Rasmussen, S. G. F., Thian, F. S., Kobilka, T. S., Choi, H.-J., Kuhn, P., Weis, W. I., Kobilka, B. K. & Stevens, R. C. (2007). *Science*, **318**, 1258–1265.
- Chien, E. Y. T., Liu, W., Zhao, Q., Katritch, V., Han, G. W., Hanson, M. A., Shi, L., Newman, A. H., Javitch, J. A., Cherezov, V. & Stevens, R. C. (2010). *Science*, **330**, 1091–1095.
- Cooper, D. R., Boczek, T., Grelewski, K., Pinkowska, M., Sikorska, M., Zawadzki, M. & Derewenda, Z. (2007). *Acta Cryst. D* **63**, 636–645.
- Corsini, L., Hothorn, M., Scheffzek, K., Sattler, M. & Stier, G. (2008). *Protein Sci.* **17**, 2070–2079.
- Dong, A. *et al.* (2007). *Nature Methods*, **4**, 1019–1021.
- Doré, A. S., Okrasa, K., Patel, J. C., Serrano-Vega, M., Bennett, K., Cooke, R. M., Errey, J. C., Jazayeri, A., Khan, S., Tehan, B., Weir, M., Wiggan, G. R. & Marshall, F. H. (2014). *Nature (London)*, **511**, 557–562.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst. D* **69**, 1204–1214.
- Goldschmidt, L., Cooper, D. R., Derewenda, Z. S. & Eisenberg, D. (2007). *Protein Sci.* **16**, 1569–1576.
- Haga, K., Kruse, A. C., Asada, H., Yurugi-Kobayashi, T., Shiroishi, M., Zhang, C., Weis, W. I., Okada, T., Kobilka, B. K., Haga, T. & Kobayashi, T. (2012). *Nature (London)*, **482**, 547–551.
- Havemeyer, A., Bittner, F., Wollers, S., Mendel, R., Kunze, T. & Clement, B. (2006). *J. Biol. Chem.* **281**, 34796–34802.
- Jin, T., Chuenchor, W., Jiang, J., Cheng, J., Li, Y., Fang, K., Huang, M., Smith, P. & Xiao, T. S. (2017). *Sci. Rep.* **7**, 40991.
- Kabsch, W. (2010). *Acta Cryst. D* **66**, 125–132.
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. (2015). *Nature Protoc.* **10**, 845–858.
- Kobe, B., Center, R. J., Kemp, B. E. & Poulos, P. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 4319–4324.
- Kobe, B., Ve, T. & Williams, S. J. (2015). *Acta Cryst. F* **71**, 861–869.
- Kuge, M., Fujii, Y., Shimizu, T., Hirose, F., Matsukage, A. & Hakoshima, T. (1997). *Protein Sci.* **6**, 1783–1786.
- Laganowsky, A., Zhao, M., Soriaga, A. B., Sawaya, M. R., Cascio, D. & Yeates, T. O. (2011). *Protein Sci.* **20**, 1876–1890.
- Miller-Gallacher, J. L., Nehmé, R., Warne, T., Edwards, P. C., Schertler, G. F. X., Leslie, A. G. W. & Tate, C. G. (2014). *PLoS One*, **9**, e92727.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst. D* **67**, 355–367.
- Nauli, S., Farr, S., Lee, Y.-J., Kim, H.-Y., Faham, S. & Bowie, J. U. (2007). *Protein Sci.* **16**, 2542–2551.
- Niemann, H. H., Knetsch, M. L., Scherer, A., Manstein, D. J. & Kull, F. J. (2001). *EMBO J.* **20**, 5813–5821.

research communications

- Niemann, H. H., Schmoltdt, H. U., Wentzel, A., Kolmar, H. & Heinz, D. W. (2006). *J. Mol. Biol.* **356**, 1–8.
- Ott, G., Havemeyer, A. & Clement, B. (2015). *J. Biol. Inorg. Chem.* **20**, 265–275.
- Palmer, T., Santini, C. L., Iobbi-Nivol, C., Eaves, D. J., Boxer, D. H. & Giordano, G. (1996). *Mol. Microbiol.* **20**, 875–884.
- Rosenbaum, D. M., Cherezov, V., Hanson, M. A., Rasmussen, S. G. F., Thian, F. S., Kobilka, T. S., Choi, H.-J., Yao, X.-J., Weis, W. I., Stevens, R. C. & Kobilka, B. K. (2007). *Science*, **318**, 1266–1273.
- Rossi, P., Ramelot, S. S. T., Xiao, R., Ho, C. K., Ma, L.-C., Acton, T. B., Kennedy, M. A. & Montelione, G. T. (2005). *J. Biomol. NMR*, **33**, 197.
- Scott, D. C. *et al.* (2017). *Nature Chem. Biol.* **13**, 850–857.
- Söding, J. (2005). *Bioinformatics*, **21**, 951–960.
- Söding, J., Biegert, A. & Lupas, A. N. (2005). *Nucleic Acids Res.* **33**, W244–W248.
- Srivastava, A., Yano, J., Hirozane, Y., Kefala, G., Gruswitz, F., Snell, G., Lane, W., Ivetac, A., Aertgeerts, K., Nguyen, J., Jennings, A. & Okada, K. (2014). *Nature (London)*, **513**, 124–127.
- Stevens, R. C. (2000). *Structure*, **8**, R177–R185.
- Suzuki, N., Hiraki, M., Yamada, Y., Matsugaki, N., Igarashi, N., Kato, R., Dikic, I., Drew, D., Iwata, S., Wakatsuki, S. & Kawasaki, M. (2010). *Acta Cryst. D* **66**, 1059–1066.
- Thorsen, T. S., Matt, R., Weis, W. I. & Kobilka, B. K. (2014). *Structure*, **22**, 1657–1664.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst. D* **66**, 22–25.
- Wahl, B., Reichmann, D., Niks, D., Krompholz, N., Havemeyer, A., Clement, B., Messerschmidt, T., Rothkegel, M., Biester, H., Hille, R., Mendel, R. R. & Bittner, F. (2010). *J. Biol. Chem.* **285**, 37847–37859.
- Walter, T. S., Meier, C., Assenberg, R., Au, K. F., Ren, J., Verma, A., Nettleship, J. E., Owens, R. J., Stuart, D. I. & Grimes, J. M. (2006). *Structure*, **14**, 1617–1622.
- Winn, M. D. (2010). *Acta Cryst. D* **67**, 235–242.
- Wu, B., Chien, E. Y. T., Mol, C. D., Fenalti, G., Liu, W., Katritch, V., Abagyan, R., Brooun, A., Wells, P., Bi, F. C., Hamel, D. J., Kuhn, P., Handel, T. M., Cherezov, V. & Stevens, R. C. (2010). *Science*, **330**, 1066–1071.
- Zou, Y., Weis, W. I. & Kobilka, B. K. (2012). *PLoS One*, **7**, e46039.

Correction of the Proofs

The previous paper represents an 'early view article' based on the manuscript proofs. Figure 1 is falsely depicted in this article and needs to be replaced for the following figure:

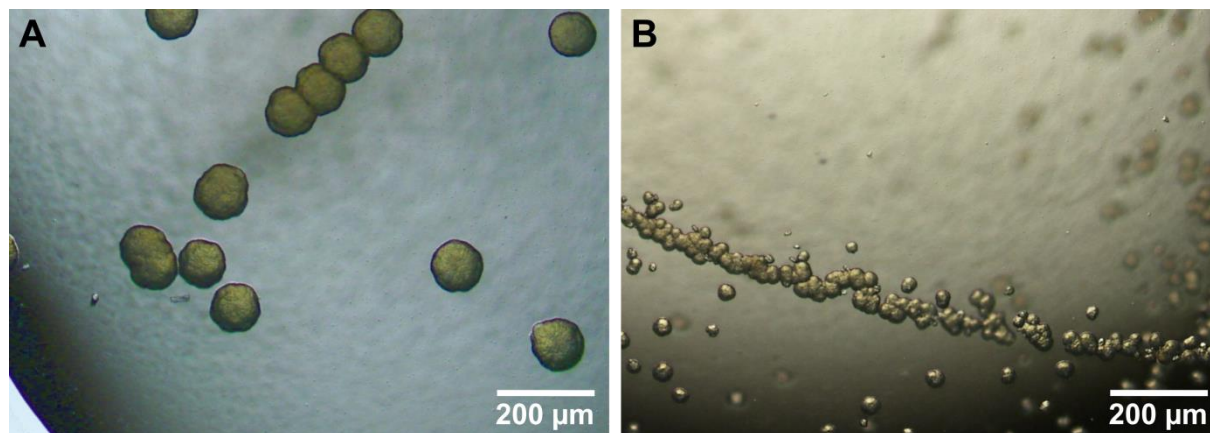


Figure 1: Spherulite formation in hmARC1 crystallization setups. **A**, Spherulites obtained in an initial crystallization trial. **B**, Spherulites obtained by streak seeding.

A request for the exchange of this figure has already been sent to the journal's editor, but no updated version of the paper was available at the time this dissertation was submitted.

3.5 First crystal structure of human mARC1 reveals its exceptional position among eukaryotic molybdenum enzymes

Christian Kubitz, Florian Bittner, Antje Havemeyer, Bernd Clement and Axel J. Scheidig

Manuscript submitted to *PNAS*: May 18th, 2018

The molybdoenzyme mARC was identified in 2006 as a formerly unknown biotransformation enzyme in the context of xenobiotic conversion. It was discovered to be an extremely effective reductase with a broad substrate spectrum comprising a wide range of *N*-hydroxylated compounds. This makes it a potent counterpart of CYP- or FMO-mediated oxygenation reactions and ensures a viable homeostasis within the metabolic cycle. mARC belongs to a subfamily of MOSC proteins, which were not structurally characterized before.

In the following manuscript, which is currently under consideration in *PNAS*, the high-resolution crystal structure of human mARC1 is discussed in detail. It is the first structure of a MOSC protein that clearly shows the coordination of its molybdenum cofactor and exhibits a topology, which contradicts *in silico* predictions of structural domains currently made by online databases. Amino acid residues, which are highly conserved among mARC proteins from various species, were shown to play a crucial role for the integrity of the three-dimensional fold of this enzyme. The large substrate spectrum could be correlated to a surface-exposed active site, which is restricted by only a few residues. While some of them, such as D209, are highly conserved among all mARC homologues and are crucial for the catalytic cycle, others were found to be paralogue-specific and might serve to discriminate between mARC1 and mARC2 in future annotations of these enzymes. Furthermore, the pyranopterin backbone of the molybdenum cofactor within the mARC1 crystal structure was observed in an unexpected conformation. Based on these findings, the hypothesis was made that mARC enzymes represent an evolutionary link between molybdoenzymes belonging either to the SO- or the XO-superfamily.

The atomic coordinates and structure factors were deposited in the Protein Data Bank (<http://www.pdb.org/>) under accession number 6FW2.

I performed the phase calculation, structure determination as well as refinement of previously obtained datasets and I analyzed the final three-dimensional model in the context of previous biochemical and biophysical characterizations of mARC proteins. Furthermore, I wrote the manuscript with the help of Dr. Bittner, Dr. Havemeyer, Prof. Dr. Clement and Prof. Dr. Scheidig.

First crystal structure of human mARC1 reveals its exceptional position among eukaryotic molybdenum enzymes

Christian Kubitza^a, Florian Bittner^b, Antje Havemeyer^c, Bernd Clement^c, Axel J. Scheidig^a

Affiliations:

^aStructural Biology, Zoological Institute, Kiel University, Am Botanischen Garten 1-9, Kiel, 24188, Germany.

^bJulius Kuehn Institute, Federal Research Centre for Cultivated Plants, Erwin-Baur-Str. 27, Quedlinburg, 06484, Germany.

^cPharmaceutical Institute, Kiel University, Gutenbergstr. 76, Kiel 24118, Germany.

Corresponding Author:

Axel J. Scheidig,

Structural Biology, Zoological Institute, Kiel University, Am Botanischen Garten 1-9, Kiel, 24188, Germany

Tel: #49 431 880 4286, E-Mail: axel.scheidig@strubio.uni-kiel.de

Classification:

BIOLOGICAL SCIENCES

Biochemistry and Pharmacology

Keywords:

biotransformation, detoxification, drug metabolism, molybdenum cofactor, MOSC

Author contributions:

A.J.S. and B.C. designed research. C.K. performed research and analyzed data. F.B. and A.H. conducted pilot studies and contributed materials. C.K., F.B., A.H., B.C. and A.J.S. wrote the paper

ABSTRACT

Biotransformation enzymes ensure a viable homeostasis by regulating reversible cycles of oxidative and reductive reactions. The metabolism of nitrogen-containing compounds is of high pharmaceutical and toxicological relevance because *N*-oxygenated metabolites derived from reactions mediated by cytochrome P450 (CYP) enzymes or flavin-dependent monooxygenases (FMO) are in some cases highly toxic or mutagenic. The molybdenum-dependent mitochondrial amidoxime-reducing component (mARC) was found to be an extremely efficient counterpart, which is able to reduce the full range of *N*-oxygenated compounds and thereby mediates detoxification reactions. However, the three-dimensional structure of this enzyme is unknown. Here we present the high-resolution crystal structure of human mARC. We give detailed insight into the coordination of its molybdenum cofactor (Moco), the catalytic mechanism and its ability to reduce a wide range of *N*-oxygenated compounds. The identification of two key residues will allow future discrimination between mARC paralogues and ensure correct annotation. Since our structural findings contradict *in silico* predictions that are currently made by online databases, we propose new domain definitions for members of the superfamily of Moco sulfuryase C-terminal (MOSC) domain-containing proteins. Furthermore, we present evidence for an evolutionary role of mARC for the emergence of the xanthine oxidase protein superfamily. We anticipate the hereby presented crystal structure to be a starting point for future descriptions of MOSC proteins, which are currently poorly structurally characterized.

SIGNIFICANCE STATEMENT

The involvement of biotransformation enzymes in drug metabolism is one of the most crucial objectives during preclinical research, since they ultimately determine the bioavailability of medicinal drugs. The mARC *N*-reductive enzyme system was found to be a highly effective counterpart to the most prominent biotransformation enzymes, CYP450, and is involved in activation of amidoxime prodrugs as well as inactivation of other drugs containing *N*-hydroxylated functional groups. Owing to its potent *N*-reductive capacity towards a broad range of compounds, including mutagenic *N*-oxygenated nucleobase analogues, mARC plays a crucial role in pharmacology. Our crystal structure of human mARC forms the basis for predictions on the metabolism of drug candidates and structure-activity-relationships. Moreover, it indicates the evolutionary development of different molybdoenzyme families.

Text

To enable the body creating a viable environment homeostasis, complex biochemical transformations such as reversible metabolic cycles of oxidative and reductive reactions are required. In this respect, metabolisms involving nitrogen are of high pharmaceutical and toxicological relevance since a number of nitrogen-containing functionalities can undergo *N*-oxygenations to *N*-oxides or *N*-hydroxylated compounds (NHC) by CYP- or FMO-catalysed xenobiotic metabolism. The resulting metabolites have different pharmacological properties, and in some cases even highly toxic, mutagenic or carcinogenic *N*-hydroxylated metabolites are produced (1, 2). Thus, retro-reduction of such first generation metabolites to their parent compounds can be regarded as a detoxification reaction. In this context, the mitochondrial amidoxime-reducing component (mARC) was discovered in our laboratory in 2006 as a so far unknown molybdenum-containing protein (3). It was identified to be an extremely effective reductase for a multitude of *N*-oxygenated molecules such as hydroxylamines, hydroxyamidines (“amidoximes”), hydroxyguanidines, oximes, *N*-oxides, hydroxamic acids, and sulfohydroxamic acids (4). mARC therefore plays a pivotal role as a counterpart to CYP- and FMO-mediated oxygenation reactions in metabolic cycles. Furthermore, recent studies suggest that mARC is important for organisms to ensure reductive detoxification strategies, e.g. of toxic hydroxylamines (5) or mutagenic *N*-hydroxylated nucleobases (6, 7). After its discovery, subsequent studies have depicted that the enzyme is able to reduce the full range of *N*-oxygenated compounds, including the capacity to reduce inorganic nitrite to nitric oxide (NO) (8) and *N*-hydroxy-L-arginine to arginine (9). All annotated genomes of mammals appear to possess two copies of mARC genes, with both copies showing strong similarities on nucleotide and amino acid levels, thus making a discrimination difficult, but defining them as paralogous proteins. Nevertheless, different substrate preferences (10) as well as different tissue-specific expression levels have been shown for the two paralogues (6). Beside sulfite oxidase, aldehyde oxidase and xanthine oxidoreductase, mARC is only the fourth molybdenum-containing enzyme found in humans and other mammals (11). With a molecular weight of around 35 kDa, it represents the simplest form among them, only binding the molybdenum cofactor (pyranopterin as prosthetic group coordinated to molybdenum; further referred to as Moco) and its substrates. In the presence of NADH, mARC proteins exert *N*-reductive activity towards NHC in concert with the two electron transport proteins cytochrome *b*₅ and NADH cytochrome *b*₅ reductase. Even though mARC proteins alone have been found to be associated not only to mitochondria but also to peroxisomes, the entire three-component system is located on the outer mitochondrial membrane and is expressed in every tissue studied so far (4). The catalytic cycle was proposed to be analogue to the described mechanism of nitrate reduction by nitrate reductase (12, 13) (Fig. 1).

Based on *in silico* analysis, Anantharaman and Aravind (2002) proposed that the *C*-terminal domain of molybdenum cofactor sulfurases represents a common feature among various enzymes in eukaryotes and prokaryotes. According to this, proteins carrying this domain were referred to as molybdenum cofactor sulfurase *C*-terminal (MOSC) domain-containing proteins (14). In humans, only the Moco sulfurase itself and mARC belong to this quite diverse protein family (11). Structural knowledge of MOSC proteins is so far limited to the bacterial Moco-dependent proteins YuaD and the recently published YiiM (15). Yet, these crystal

structures are lacking the molybdenum cofactor and therefore only allow speculations about the nature of the active site and substrate binding mode.

Here we present the high-resolution crystal structure of human mARC, which for the first time reveals detailed insights into MOSC proteins and their coordination of the Mo-molybdopterin cofactor. In addition, the structure provides evidences for the evolutionary link between sulfite oxidase and xanthine oxidase family of molybdoenzymes. This structure is crucial for a deeper understanding of the observed substrate spectrum in drug metabolism.

RESULTS

Crystal structure of human mARC1. The crystal structure of the fusion protein comprising T4 lysozyme (T4L) and *N*-terminally truncated human mARC1 (hmARC1) was determined by molecular replacement (Fig. S1). The 1.78 Å data set was refined to a final R-factor of 16.9 % ($R_{\text{free}} = 20.8$ %). The final model consists of 444 residues (with 283 belonging to hmARC1 and 161 to T4L), one Moco, four molybdate ions, one phosphate ion, one bis-TRIS propane molecule and 439 water molecules in the asymmetric unit. The ions and small molecules are derived from the purification and crystallization buffers, respectively. The *C*-terminal His₆ tag is disordered and could not be modelled due to the absence of appropriate electron density. Residues D301 - K310 are poorly defined and display high *B*-factors, yet their positions could be traced at low σ -contouring. The electron density maps for the rest of the model were of high quality and could be modelled with high confidence.

hmARC1 comprises two structural domains which consist of non-successive secondary structure elements. Most parts of the protein are dominated by β -strands which form two three-stranded anti-parallel β -sheets (Fig. 2A/C, light and dark blue, respectively), a small four-stranded antiparallel β -barrel (yellow) and a large, seven-stranded mostly anti-parallel β -barrel (red), which is slightly deformed. Three α -helices of different lengths and four 3_{10} -helices complete the crystal structure. Two structural domains can be distinguished: one is comprised of β -strands 4-10 as well as helix α_1 , the other encloses the large β -barrel, a three-stranded β -sheet forming a “lid” of the barrel, helices α_2 and α_3 as well as the 3_{10} -helices $\eta_2 - 4$. Buried within the cleft between the two domains lies the Moco, accompanied by 3_{10} -helix η_1 .

Structural insight into MOSC proteins. MOSC domains were revealed by computational analysis as a novel, yet ancient, superfamily of β -strand-rich domains, which occur either as stand-alone forms or fused to other domains. They were predicted to be sulphur-carrier domains, which receive formerly enzymatically abstracted sulphur on a highly-conserved cysteine residue and further deliver it for the formation of diverse metal-sulphur clusters (14). In eukaryotes, this domain superfamily solely comprises the two mARC proteins and the Moco sulfurase (11), while in prokaryotes the Moco-dependent enzymes YcbX and YiiM, among others, can also be found (14). In addition to the MOSC domain, larger family members also share a distinct N-terminal domain, which is referred to as MOSC_N domain. This domain was nowhere detectable as a stand-alone form, but was predicted to adopt a β -barrel-like structure and to be involved in substrate recognition and binding, while the MOSC domain

contains two conserved cysteines needed for sulphur transition. In hmARC1, the MOSC_N and MOSC domains are formed by residues 56 – 175 and 187 – 335, respectively. Residues of the MOSC domain, which enclose the β -strands 11 – 17 and all helices but α 1, bind the Moco and contain the two conserved cysteines, with one of them coordinating the central molybdenum ion. However, in contrast to predictions and current annotations in databases, the MOSC_N domain consists of β -strands 1 – 9 and helix α 1, which do not form a β -barrel-like structure. In fact, only β 1 is part of the large β -barrel, which is present within the hmARC1 crystal structure, but mainly comprises residues of the MOSC domain. Still, the four β -strands 4 and 8 – 10 from the MOSC_N domain form a small β -barrel. While there are two domains distinguishable within the crystal structure of hmARC1, these do not correlate with the two computationally predicted MOSC and MOSC_N-domains. Tertiary structure elements like the large β -barrel or a three-stranded β -sheet (β 2 – 3 and β 16) rather comprise residues from both predicted domains. We propose a clarification for the definition of the MOSC_N as well as MOSC domain and therefore for the MOSC-protein family, which not only takes into account predictions based on sequence analysis but also the newly identified structural arrangement and composition of conserved domains.

Molybdenum cofactor coordination. The Moco is tightly bound within the core region of hmARC1 by mostly positively charged amino acids and residues carrying a hydroxyl group in their side chain (Fig. S2). Besides the two dithiolene sulphurs from the molybdopterin backbone and two oxygen ligands, the central molybdenum is coordinated by C273, forming a slightly distorted trigonal bipyramidal geometry. This cysteine residue is highly conserved among all mARC proteins throughout all species and is part of a common CxxC motif. The most prominent Moco-coordinating side chain is R92 which interacts with several atoms of the cofactor via polar and ionic interactions, keeping it strictly in place. The pterin ring system is further bound by residues T210, S211, P212, R238, N240 and Y317, while the phosphate moiety is coordinated by K67, S68, R92 and R238 (Fig. 3). This multitude of specific interactions between the cofactor and surrounding protein residues allows for a tight coordination of the molybdopterin backbone within the core of the enzyme. However, the reactive site comprising the molybdenum and its ligands remains exposed to the surface of the protein, where NHC are being recruited and subsequently reduced.

Assignment to xanthine oxidase or sulfite oxidase family. Eukaryotic molybdenum enzymes are currently classified into two different families, which are distinguishable from each other by the composition of the fivefold coordination sphere of the molybdenum centre. In both families, the molybdenum is coordinated by the dithiolene sulphurs of the molybdopterin as well as two oxygen (either oxo or hydroxyl) ligands. Members of the sulfite oxidase (SO) family contain a proteinogenic cysteine-sulphur as the fifth ligand, while molybdenum enzymes of the xanthine oxidase (XO) family carry an inorganic “terminal sulphur” ligand, which is essential for catalytical activity (11). Abstracting this terminal sulphur ligand from members of the XO family by cyanide treatment leads to inactivation of the respective enzyme (16, 17). Human mARC proteins, however, did not release any sulphur in form of thiocyanate after cyanide treatment, excluding their belonging to the XO family (18). In contrast, mARC homologues from different source organisms were assigned to the SO family of molybdenum enzymes by pulsed electron paramagnetic resonance (EPR) (19), X-ray absorption near-edge structure

(XANES) and extended X-ray absorption fine structure (EXAFS) at the molybdenum K-edge (20) as well as biochemical studies and activity assays performed with mARC variants (21). Taken together, these studies clearly identified a protein-derived cysteine as the fifth ligand, which is indeed confirmed by our hmARC1 crystal structure.

Apart from physical and biochemical investigations, the pyranopterin conformations of currently available protein structures of mononuclear molybdenum and tungsten enzymes were analysed (22). By deriving a distortion coordinate based on dihedral angles within the prosthetic group, the authors could show that the pyranopterin conformation can be correlated with their former biochemical assignment to either the XO or SO enzyme family. Interestingly, despite being well-characterized as a member of the SO family, the hmARC1 crystal structure reveals a pyranopterin conformation with dihedral angles of $\alpha = -43.3^\circ$ and $\beta = 73.4^\circ$ (Fig. S3). Remarkably, this angle combination would assign hmARC to the XO family based on the analysis of Rothery *et al.* (2012) (22).

Therefore, we suggest that mARC proteins should not be directly assigned to either of the two currently proposed and distinct mononuclear molybdenum enzyme families, since it exhibits combined characteristics of both, SO and XO family.

Composition of the active site and substrate binding area. While the molybdopterin moiety itself is tightly anchored within the enzyme, the reactive centre is well accessible from the solvent area. There are few spacial limitations to the active site and substrate binding area, which are mainly composed of residues C273, D209, R272, S271, R107, Y317, T210, H152 and S311 (Fig. S4). With the exception of T210, H152 and S271, these residues are highly-conserved throughout mARC proteins from different organisms, suggesting quite similar substrate spectra among them. However, we were able to identify two potential key residues close to the active site, which are indicative for either mARC1 or mARC2. While H152 did not appear to be conserved among all analysed mARC enzymes, it is indeed highly-conserved among the mARC1 orthologues, while in mARC2 there is a likewise conserved phenylalanine in this position. The second discriminator between the two paralogues is S271. Among mARC1 sequences, mostly serine or threonine residues can be found in this position, whereas mARC2 proteins predominantly display a proline (Fig. S5).

The solvent-exposed reactive centre as well as the absence of any highly-specific substrate binding site clearly account for the broad substrate spectrum of mARC enzymes. Also, given the presence of some residues with alternative conformations (C273 and D209), the active site allows for some plasticity, making it even more adjustable towards different kinds of substrates. hmARC1 displays positively charged residues surrounding the active site (Fig. 4). Since a variety of substrates contain nitrogen atoms with at least a partial positive charge, this was unexpected. However, there is one negatively charged, highly-conserved, residue (D209) in direct proximity to the molybdenum center, which is crucial for catalytic activity (discussed below) and is most likely involved in the binding of the hydroxylated nitrogen of any given substrate. The surrounding positively charged patch might instead be crucial for the interaction between mARC1 and its electron-delivering redox-partner protein cytochrome b5, which displays negatively charged residues on its surface surrounding the heme cofactor (information derived from PDB entry 3NER (23)).

Comparison with studies of crARC variants. Recently, the mARC homologue from the green alga *Chlamydomonas reinhardtii* (crARC) was intensively investigated with regards to highly conserved amino acid residues among all mARC enzymes. Different variants, where the residue of choice was replaced by alanine, were studied for their impact on enzyme activity, Moco chelation, protein oligomerization and potential electron transfer between crARC and its redox partner proteins (24). The authors identified three residues which are essential for reduction activity towards NHC. These residues correspond to hmARC1 residues D209, F237 and R298 (identified by sequence alignment between crARC and hmARC1; Fig. S6). Another two residues (corresponding to hmARC1 R298 and F237) were shown to have an impact on Moco coordination, while some others (corresponding to hmARC1 L180, R238, E251 and E289) are supposed to be involved in the electron transport between cytochrome b5 and mARC. Our crystal structure of hmARC1 strongly supports these findings.

While the decreased *N*-reductive activity in two of three variants correlates with decreased cofactor-binding ability, the identified aspartic acid residue seems to have a Moco-independent impact on mARC enzymatic activity. This might be due to a direct involvement within the catalytic cycle. D209 is situated in direct proximity to the molybdenum ion and its hydroxo ligand (Fig. S7A). We propose, this residue is responsible for binding the hydroxylated nitrogen atom of any NHC and coordinating it in such a way that allows for the hydroxo ligand of the Moco to be exchanged for the NHC, resulting in a reaction intermediate (see Fig. 1). Furthermore, D209 is the only negatively charged residue within a mostly alkaline substrate binding area, which further supports its role in recruiting hydroxylated nitrogen compounds.

F237 is part of the Moco binding site but not directly involved in cofactor coordination, while R298 is not found in direct proximity of the cofactor. However, both residues are essential for the structural integrity of the same mARC domain. F237 is the central amino acid of a hydrophobic core between the large β -barrel and helices α_3 , α_4 and η_4 , securing the three-dimensional arrangement of the Moco binding site (Fig. S7B). R298 has a similar function: it is the central residue which connects the large β -barrel and helices α_3 and η_3 via polar and ionic interactions as well as cation- π -stacking (Fig. S7C). Replacement of either of these residues probably results in a partial collapse of these domains and leads to the observed decrease or even loss of Moco-binding capacity and ultimately the loss of NHC reductive activity.

Chamizo-Ampudia *et al.* simultaneously tested crARC variants for their NADH-dependent as well as dithionite-dependent reductive activity. The former assay is dependent on the partner redox proteins cytochrome b5 reductase and cytochrome b5. In contrast, dithionite can directly deliver electrons to the Moco, enabling crARC to perform *N*-reduction without the assistance of other proteins. Variants, where residues corresponding to hmARC1 L180, R238, E251 and E289 were exchanged for alanine, showed decreased NADH-dependent reductive activity, while dithionite-dependent activity was still comparable to the wildtype enzymes. Therefore, the authors concluded that these residues might be involved in the interaction between the three redox partner proteins and/or the electron transfer between them. While our structure does not allow a reliable prediction of a protein-protein interaction site, the investigated residues are indeed essential for structural integrity. Like F237 and R298 they connect different secondary

structure elements and might therefore stabilize conformations of the mARC enzyme that allow for redox partners to bind or electrons to be transported to the Moco (Fig. S7D, E).

Of note, variants corresponding to hmARC1 N240, D252 and L294 were discarded from the authors' investigations, because the exchange for an alanine drastically changed the conformation of the enzyme (as determined by fluorescence emission by excitation of aromatic amino acids). Within the hmARC1 crystal structure, these can also be identified as key residues, which play a significant role in the tertiary structure arrangement (Fig. S7F - H).

DISCUSSION

The high-resolution crystal structure of hmARC1 allows detailed insight into the fourth mammalian molybdenum-dependent enzyme. For the first time, previous biochemical characterizations of mARC proteins, such as the homologue found in *Chlamydomonas reinhardtii*, can be directly correlated to the three-dimensional architecture of the enzyme. The role of different highly-conserved residues has been investigated in variant studies before (24). We can now emphasize these results by identifying most of these residues to be crucial for the structural integrity of mARC proteins, while others are directly or indirectly involved in cofactor binding or catalytic activity.

Furthermore, we provide the first structural interpretation of a MOSC protein and its computationally predicted MOSC and MOSC_N domains. While there are two domains distinguishable within the hmARC1 crystal structure, these do not fully correlate to the *in silico* domain definitions. Opposed to former hypotheses, the MOSC_N domain does not represent a β -barrel-like fold. Although there is a large β -barrel, which contributes to a major part of one structural domain, this one mainly comprises β -strands from the predicted MOSC domain in concert with the very N-terminal β -strand of the MOSC_N domain. Since both predicted domains of MOSC proteins are intertwined on a structural level, we propose a new definition of MOSC domains to be deposited in databases, which take into account sequence motifs as well as the now available structural information. Of note, the recently published crystal structures of the bacterial enzyme YiiM (15) (PDB entries 5YHH and 5YHI) also contribute to the MOSC superfamily description. However, they belong to another subfamily, are lacking their essential molybdopterin prosthetic group and the MOSC_N/MOSC-specific topology is not discussed.

The hmARC1 crystal structure revealed an unexpected conformation of the bound Moco, which is usually exclusively observed in molybdenum-dependent enzymes belonging to the XO family. However, mARC proteins from different organisms have been clearly identified as members of the SO family based on biochemical and biophysical features. Therefore, in contrast to previous assumptions, XO and SO families of Moco-containing enzymes might not be strictly separated from each other – at least not in all cases. On the other hand, mARC proteins might represent an evolutionary link between the two enzyme families. This hypothesis is supported by the high sequence similarity between mARC and the C-terminus of Moco sulfurase. The latter protein is composed of two functional domains and catalyses the final maturation step, the sulphuration, of the Moco as it is found within enzymes of the XO family

(17, 25). In a multi-step reaction, this enzyme abstracts sulphur from a free cysteine with its cysteine desulphurase domain, transfers it to a cysteine residue within the MOSC domain and subsequently sulphurates recruited SO-Moco. Finally, the sulphurated Moco is released and inserted into enzymes of the XO family (26). Two things can be concluded from this reaction: i) members of the XO family emerged later than the SO-enzymes and ii) Moco sulfurase needs to be able to bind both, SO-type as well as XO-type Moco, via its MOSC domain. It directly links both families of molybdenum enzymes to each other. There are indications that Moco sulfurases evolved from MOSC proteins like mARC by domain fusion with the aforementioned cysteine desulphurase-like domains (14). Conclusively, mARC enzymes could be the ancestors of Moco sulfurase proteins and represent an evolutionary link between different families of Moco-dependent enzymes, while still retaining their ability to bind the SO-type of Moco, however, in a conformation usually observed within members of the XO family.

The active site of hmARC1 is revealed to be almost completely solvent-exposed. Like many other enzymes involved in biotransformation, mARC proteins are generalists rather than specialists, which provide a fast and efficient detoxification system for a variety of NHC. This necessitates a free access of any substrate to the active site, which is not buried inside a restricting binding pocket. The limited substrate specificity of mARC enzymes is realized by only a few residues surrounding the reactive molybdenum centre, which are strictly conserved throughout different organisms. This explains why it is very difficult to derive structure-activity relationships (27). Thus, functional groups with *N*-hydroxylated components are reduced irrespective of the rest of the molecule. For example, ximelagatran, a thrombin inhibitor and a big molecule with many other functional groups is reduced to a similar extent as a simple *N*-hydroxylated benzamidine (benzamidoxime) (28).

By comparison of mARC sequences from different eukaryotes, we could identify two residues close to the active site, which appear to be paralogue-specific and can therefore be used to discriminate between mARC1 and mARC2. These findings might explain the few differences in their substrate preferences and specific enzymatic activity, despite their mostly overlapping substrate spectrum. Furthermore, these paralogue-specific residues might be used for future annotations of mARC enzymes, which are not yet deposited in databases.

Whereas mARC is undoubtedly a drug metabolizing enzyme and is also involved in the detoxification of NHC, its physiological role is not quite clear. Besides the modulation of the NO pathway (9, 29, 30), evidence has been accumulated that mARC plays a role in lipid metabolism (31).

METHODS

Protein expression / purification / crystallization. The design, protein expression, purification, functional characterization and crystallization of the hmARC1-T4L fusion construct have been described in detail before ([article in press; Acta Crystallographica Section F](#)). Briefly, the *N*-terminally truncated fusion protein was expressed in *E. coli* TP1000 cells and purified by sequential affinity and cation exchange chromatography. Purified protein was analysed for *N*-reductive activity and subjected to hanging-drop vapour-diffusion crystallization setups, yielding hmARC1-T4L crystals of the orthorhombic space group P2₁2₁2₁ which diffracted to a resolution of 1.65 - 3.5 Å.

Data collection / Phasing / Model building / Refinement. Diffraction data were collected at beamline P14 (EMBL, DESY PETRA III, Hamburg, Germany) by using a Pilatus 2M detector. Data were collected at 100 K, a detector distance of 136.2 mm, a wavelength of 0.9789 Å, an oscillation range of 0.1° and an exposure time of 0.00146 s per frame. Due to needle-shaped crystal morphology and in order to limit radiation damage, a helical data collection strategy along the longitudinal axis of the crystals was chosen. The best dataset derived from a single crystal which diffracted to a resolution of 1.78 Å.

Phasing was performed by molecular replacement (MR) using Molrep (32). Due to the lack of structurally characterized mARC-homologues, MR had to be performed in sequential steps using partial models of the fusion protein. First, T4L (PDB entry 206L) was used as a search template. A convincing solution with just one molecule within the asymmetric unit was found. After rigid body refinement, performed by Refmac5 (33), reliable electron density distribution was found for the lysozyme molecule. Some additional electron density was visible, which had to account for hmARC1, yet it was poorly defined and did not allow modelling the molecule. To improve phases, additional molecular replacement runs were performed with the T4L as a fixed input model and additional structure elements of truncated hmARC1 models. These incomplete homology models were generated using MODELLER within the HHpred server (34, 35) and the crystal structure of the bacterial Moco-dependent enzyme YuaD (PDB entry 1ORU; sequence identity to hmARC1: 25.38 %). MR solutions were found using a model containing several β -strands as well as an α -helix (residues W94 – N98 and S116 – F183) and, subsequently, a model containing the predicted β -barrel. The position of the β -barrel was found, yet the orientation was inverted and residues were incorrectly assigned. However, due to improved phases, the electron density distribution was sufficiently well defined to allow for manual inspection in Coot (36) and correctly building the hmARC1 C $_{\alpha}$ -chain step by step. Iterative refinement cycles were performed using Coot and Refmac to complete the model. 97.9 % of all modelled residues lie within the favoured region and 2.1 % lie within the allowed region of the Ramachandran plot. There are no outliers. Refinement statistics are summarized in Table S1.

Visualization. The pH-dependent electrostatic potential maps at pH 7.4 were calculated by using the software APBS with the PARSE force field (37). Input files in PQR-format for APBS were generated from files containing hmARC1 atom coordinates in PDB-format with the program PDB2PQR in order to calculate pK_a values and set protonation states for titratable groups accordingly (38). Composite omit maps were generated using the program Phenix (39).

Simulated annealing cycles were performed to erase model bias. The representation of residues involved in binding the Moco was derived from an analysis with LigPlot+ (40). The WebLogo representation of conserved mARC paralogue discriminators was created using the Online-Tool at <http://weblogo.berkeley.edu/logo.cgi> (41). The multiple sequence alignment of 58 mARC1 and 65 mARC2 proteins from different mammalian organisms as well as the alignment of hmARC1 with crARC was performed with Clustal Omega (42). The mARC sequences were derived from the Ensemble website (www.ensembl.org), GeneTree StableID ENSGT00530000063150, node_id 20095130 (mARC1 orthologues) and 20094778 (mARC2 orthologues) as of march 17th, 2018. Incomplete sequences were sorted out before analysis. The final figure for the hmARC1/crARC alignment was prepared using the program ESPript (43). The secondary structure of hmARC1 was derived by DSSP (44). All visualization and preparation of 3D structural images was performed using the program PyMOL (45).

Acknowledgments: We thank Tracy Palmer and Grant Buchanan (University of Dundee) for sharing *E. coli* strain TP1000. We gratefully acknowledge access to the core facilities of the BiMo/LMB of the University Kiel. Diffraction data were collected on beamline P14 operated by EMBL at the PETRAIII storage ring (Hamburg, Germany).

Funding: Our research received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under BioStructX (grant agreement N°283570). Additionally, we are grateful for access to the HTX crystallization facility and beamtime at P14 by means of a grant from BioStructX at the EMBL outstation Hamburg.

Competing interests: Authors declare no competing interests.

Data and materials availability: The atomic coordinates and structure factors have been deposited in the Protein Data Bank (<http://www.pdb.org/>) under accession number 6FW2.

References:

1. Foti RS & Dalvie DK (2016) Cytochrome P450 and Non-Cytochrome P450 Oxidative Metabolism: Contributions to the Pharmacokinetics, Safety, and Efficacy of Xenobiotics. *Drug Metab Dispos* 44(8):1229-1245.
2. Hlavica P (2002) N-oxidative transformation of free and N-substituted amine functions by cytochrome P450 as means of bioactivation and detoxication. *Drug Metab Rev* 34(3):451-477.
3. Havemeyer A, *et al.* (2006) Identification of the missing component in the mitochondrial benzamidoxime prodrug-converting system as a novel molybdenum enzyme. *J Biol Chem* 281(46):34796-34802.
4. Ott G, Havemeyer A, & Clement B (2015) The mammalian molybdenum enzymes of mARC. *J Biol Inorg Chem* 20(2):265-275.
5. Ott G, *et al.* (2014) Reduction of sulfamethoxazole hydroxylamine (SMX-HA) by the mitochondrial amidoxime reducing component (mARC). *Chem Res Toxicol* 27(10):1687-1695.
6. Krompholz N, *et al.* (2012) The mitochondrial Amidoxime Reducing Component (mARC) is involved in detoxification of N-hydroxylated base analogues. *Chem Res Toxicol* 25(11):2443-2450.
7. Plitzko B, Havemeyer A, Kunze T, & Clement B (2015) The pivotal role of the mitochondrial amidoxime reducing component 2 in protecting human cells against apoptotic effects of the base analog N6-hydroxylaminopurine. *J Biol Chem* 290(16):10126-10135.
8. Llamas A, Chamizo-Ampudia A, Tejada-Jimenez M, Galvan A, & Fernandez E (2017) The molybdenum cofactor enzyme mARC: Moonlighting or promiscuous enzyme? *Biofactors* 43(4):486-494.
9. Kotthaus J, *et al.* (2011) Reduction of N(omega)-hydroxy-L-arginine by the mitochondrial amidoxime reducing component (mARC). *Biochem J* 433(2):383-391.
10. Jakobs HH, *et al.* (2014) The mitochondrial amidoxime reducing component (mARC): involvement in metabolic reduction of N-oxides, oximes and N-hydroxyamidinohydrazones. *ChemMedChem* 9(10):2381-2387.
11. Hille R, Nishino T, & Bittner F (2011) Molybdenum enzymes in higher organisms. *Coord Chem Rev* 255(9-10):1179-1205.
12. Fischer K, *et al.* (2005) Structural basis of eukaryotic nitrate reduction: crystal structures of the nitrate reductase active site. *Plant Cell* 17(4):1167-1179.
13. Havemeyer A, Lang J, & Clement B (2011) The fourth mammalian molybdenum enzyme mARC: current state of research. *Drug Metab Rev* 43(4):524-539.
14. Anantharaman V & Aravind L (2002) MOSC domains: ancient, predicted sulfur-carrier domains, present in diverse metal-sulfur cluster biosynthesis proteins including Molybdenum cofactor sulfurases. *FEMS Microbiol Lett* 207(1):55-61.
15. Namgung B, Kim JH, Song WS, & Yoon SI (2018) Crystal structure of the hydroxylaminopurine resistance protein, YiiM, and its putative molybdenum cofactor-binding catalytic site. *Sci Rep* 8(1):3304.
16. Massey V & Edmondson D (1970) On the mechanism of inactivation of xanthine oxidase by cyanide. *J Biol Chem* 245(24):6595-6598.
17. Wahl RC, Warner CK, Finnerty V, & Rajagopalan KV (1982) Drosophila melanogaster ma-1 mutants are defective in the sulfuration of desulfo Mo hydroxylases. *J Biol Chem* 257(7):3958-3962.
18. Wahl B, *et al.* (2010) Biochemical and spectroscopic characterization of the human mitochondrial amidoxime reducing components hmARC-1 and hmARC-2 suggests the

- existence of a new molybdenum enzyme family in eukaryotes. *J Biol Chem* 285(48):37847-37859.
19. Rajapakshe A, *et al.* (2011) Structural studies of the molybdenum center of mitochondrial amidoxime reducing component (mARC) by pulsed EPR spectroscopy and ¹⁷O-labeling. *Biochemistry* 50(41):8813-8822.
 20. Giles LJ, *et al.* (2014) Molybdenum site structure of MOSC family proteins. *Inorg Chem* 53(18):9460-9462.
 21. Chamizo-Ampudia A, Galvan A, Fernandez E, & Llamas A (2011) The *Chlamydomonas reinhardtii* molybdenum cofactor enzyme crARC has a Zn-dependent activity and protein partners similar to those of its human homologue. *Eukaryot Cell* 10(10):1270-1282.
 22. Rothery RA, Stein B, Solomonson M, Kirk ML, & Weiner JH (2012) Pyranopterin conformation defines the function of molybdenum and tungsten enzymes. *Proc Natl Acad Sci U S A* 109(37):14773-14778.
 23. Parthasarathy S, *et al.* (2011) Accommodating a nonconservative internal mutation by water-mediated hydrogen bonding between beta-sheet strands: a comparison of human and rat type B (mitochondrial) cytochrome b5. *Biochemistry* 50(24):5544-5554.
 24. Chamizo-Ampudia A, Galvan A, Fernandez E, & Llamas A (2017) Study of Different Variants of Mo Enzyme crARC and the Interaction with Its Partners crCytb5-R and crCytb5-1. *Int J Mol Sci* 18(3).
 25. Amrani L, *et al.* (2000) Comparison of the sequences of the *Aspergillus nidulans* hxB and *Drosophila melanogaster* ma-1 genes with nifS from *Azotobacter vinelandii* suggests a mechanism for the insertion of the terminal sulphur atom in the molybdopterin cofactor. *Mol Microbiol* 38(1):114-125.
 26. Schwarz G & Mendel RR (2006) Molybdenum cofactor biosynthesis and molybdenum enzymes. *Annu Rev Plant Biol* 57:623-647.
 27. Bauch E, *et al.* (2015) Electrochemical and mARC-catalyzed enzymatic reduction of para-substituted benzamidoximes: consequences for the prodrug concept "amidoximes instead of amidines". *ChemMedChem* 10(2):360-367.
 28. Gruenewald S, *et al.* (2008) The fourth molybdenum containing enzyme mARC: cloning and involvement in the activation of N-hydroxylated prodrugs. *J Med Chem* 51(24):8173-8177.
 29. Sparacino-Watkins CE, *et al.* (2014) Nitrite reductase and nitric-oxide synthase activity of the mitochondrial molybdopterin enzymes mARC1 and mARC2. *J Biol Chem* 289(15):10345-10358.
 30. Yang J, *et al.* (2015) Oxyl and hydroxyl radical transfer in mitochondrial amidoxime reducing component-catalyzed nitrite reduction. *J Am Chem Soc* 137(16):5276-5279.
 31. Neve EP, *et al.* (2012) Amidoxime reductase system containing cytochrome b5 type B (CYB5B) and MOSC2 is of importance for lipid synthesis in adipocyte mitochondria. *J Biol Chem* 287(9):6307-6317.
 32. Vagin A & Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):22-25.
 33. Murshudov GN, *et al.* (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):355-367.
 34. Soding J, Biegert A, & Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33(Web Server issue):W244-248.
 35. Soding J (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics* 21(7):951-960.
 36. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.

37. Jurrus E, *et al.* (2018) Improvements to the APBS biomolecular solvation software suite. *Protein Sci* 27(1):112-128.
38. Dolinsky TJ, Nielsen JE, McCammon JA, & Baker NA (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* 32(Web Server issue):W665-667.
39. Adams PD, *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213-221.
40. Laskowski RA & Swindells MB (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* 51(10):2778-2786.
41. Crooks GE, Hon G, Chandonia JM, & Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14(6):1188-1190.
42. Sievers F, *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539.
43. Robert X & Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42(Web Server issue):W320-324.
44. Kabsch W & Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22(12):2577-2637.
45. Schrödinger, LLC (2015) The PyMOL Molecular Graphics System, Version 1.8.
46. Collaborative Computational Project N (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50(Pt 5):760-763.

FIGURE LEGENDS

Fig. 1. Catalytic cycle of mARC enzymes. Reducing equivalents supplied by NADH are passed to cytochrome *b*₅ reductase, then to cytochrome *b*₅ before being relayed to mARC to prime the Mo^{IV} active site for substrate reduction. It is assumed that *N*-hydroxylated substrates are reduced by cleavage of the *N*-*O* bond, in analogy to the described mechanism of nitrate reduction by nitrate reductase (12). This is accomplished by protonation of the hydroxyl group and subsequent leaving of one water molecule (13).

Fig. 2. Crystal structure and topology of hmARC1. A, Cartoon representation of hmARC1 coloured by secondary structure elements. B, Cartoon representation coloured by *in silico* predictions of MOSC_N (orange) and MOSC domain (blue). C, Topology model of hmARC1. Secondary structure elements are coloured as in Fig. 2A. Triangles represent β -strands, large circles represent α -helices and small circles represent 3_{10} -helices. Domains are highlighted in orange (MOSC_N) and blue (MOSC). The insertion site of the crystallization-facilitating fusion partner T4 lysozyme is indicated by a brown rhombus.

Fig. 3. Coordination of the molybdenum cofactor. Residues interacting with the Moco are depicted as grey boxes. Dashed lines represent hydrogen bonds between polar atoms. Blue circles indicate water molecules.

Fig. 4. Electrostatic potential of the surface surrounding the active site. The electrostatic potential is represented as a colour gradient from red ($-30 \text{ k}_B\text{T}/e_c$) over white ($0 \text{ k}_B\text{T}/e_c$) to blue ($+30 \text{ k}_B\text{T}/e_c$).

FIGURES

Figure 1

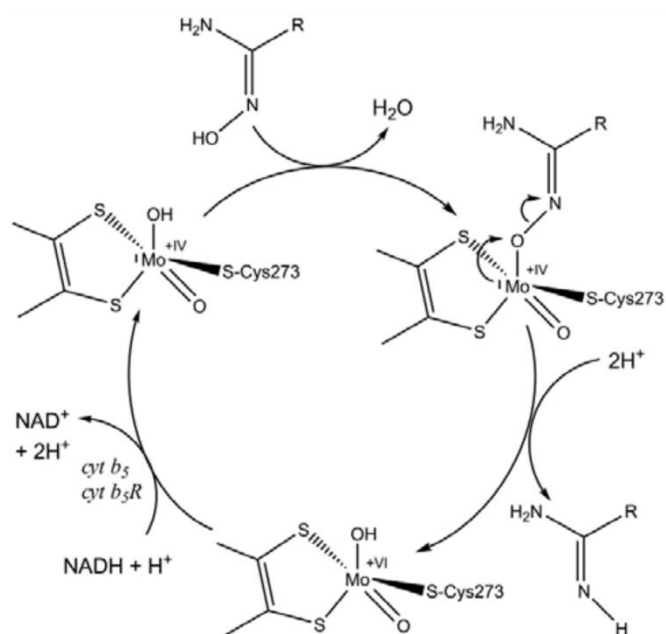


Figure 2

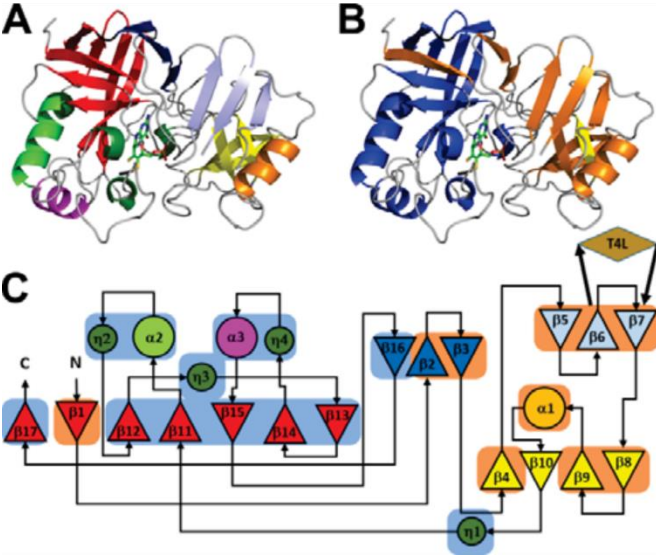


Figure 3

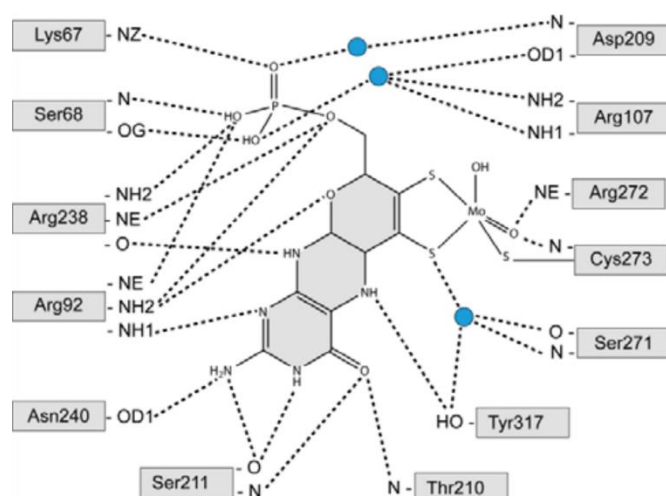
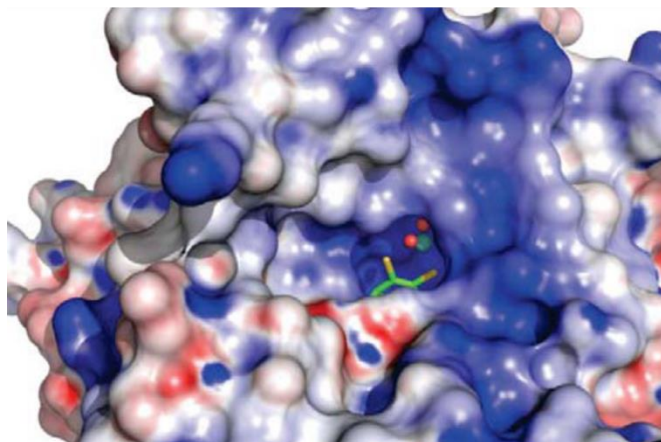


Figure 4



SUPPLEMENTARY INFORMATION (SI)**SUPPLEMENTARY FIGURE LEGENDS**

Fig. S1. Interactions between the T4L and hmARC1 fusion partners. A, Cartoon representation of the hmARC1-T4L fusion protein. The molybdenum cofactor and small molecules are depicted as stick models. Despite the internal integration of the T4L moiety (red) into the hmARC1 sequence (blue), both proteins are clearly separated from each other. The two lobes of T4L are rigidly connected with each other by a tightly bound bis-TRIS propane molecule. B, Interaction sites between hmARC1 and T4L. In addition to the peptide bonds that covalently tether both moieties to each other, the proteins interact via a few polar direct and indirect contacts. Interacting residues are shown as grey boxes, blue circles represent water molecules and dashed lines indicate hydrogen bonds. C, Crystal packing of the hmARC1-T4L fusion protein. Blue, hmARC1 molecules; red, T4L molecules. The view is along the b-axis of the unit cell. hmARC1 and T4L molecules are packed alternating within the crystal, indicating that the T4L fusion partner is facilitating the crystal contacts.

Fig. S2. Molybdenum cofactor binding site. The hmARC1 protein backbone is depicted in cartoon representation. Residues interacting with the Moco are shown as sticks and coloured according to different atom types (grey, carbon; red, oxygen; blue, nitrogen; yellow, sulphur; orange, phosphorus). A, Composite omit map of the Moco. Blue, 2Fo-Fc map, contoured at 1.0 σ ; green, Fo-Fc map, contoured at 2.5 σ . B, Representation of the cofactor as it was modelled into the electron density map. The molybdopterin backbone is depicted in stick representation, the molybdenum ion and its oxygen ligands as spheres.

Fig. S3. Molybdopterin conformation observed in hmARC1. A, 2D representation of the Moco. Green highlights, atoms used for determination of specific dihedral angles. B, 3D depiction of the Moco conformation as observed in the hmARC1 crystal structure. The atoms used for dihedral angle determination are labeled with different letters. C, Representation of the α and β dihedral angles. α is defined by atoms a-b-c-d; β is defined by a-b-c-e.

Fig. S4. Representation of hmARC1 active site. The enzyme is shown in cartoon as well as semi-transparent surface representation. The Moco is depicted in stick (molybdopterin backbone) and sphere representation (molybdenum ion and oxygen ligands). Residues in close proximity to the reactive molybdenum centre are shown in stick representation and coloured according to different atom types (green, carbon; red, oxygen; blue, nitrogen; yellow, sulphur; orange, phosphorus).

Fig. S5. WebLogo representation of identified mARC paralogue discriminators. Arrows indicate the position of the amino acid residue, which is highly conserved among either mARC1 or mARC2 enzymes, respectively. The height of the letters indicates the percentage with which a residue occurs in this position among all analysed sequences.

Fig. S6. Sequence comparison of hmARC1 and crARC. Identical residues are shown as white letters with red background, and similar residues are shown as red letters with white background. The residues 2-52 of hARC1 (light blue background) were not part of the crystallized construct. The secondary structure of hmARC1 is depicted above the sequences. α : α -helix; β : β -sheet; η : 3_{10} -helix; T: turn.

Fig. S7. Position and function of residues that are highly conserved among mARC enzymes from different organisms. The hmARC1 backbone is shown in cartoon representation and coloured in white. Highly conserved mARC-residues (red) and neighbouring interacting residues (green) are depicted in stick representation and coloured according to different atom types (red, oxygen; blue, nitrogen; yellow, sulphur; orange, phosphorus). Secondary structure elements of interest are highlighted by blue labels. A, The protein is additionally shown in a semi-transparent surface representation and the Moco is depicted in stick (molbdopterin backbone) and sphere representation (central molybdenum ion and oxygen ligands). The catalytically crucial residue D209 is located in direct proximity to the reactive centre and supposedly recruits the nitrogen atom of *N*-hydroxylated substrates to the active site. B, F237 is the central residue within a large hydrophobic core, which connects the large β -barrel and helices α_3 , α_4 and η_4 . C, R298 interacts with residues Y249, D252 and L313, thereby connecting helix α_3 with the large β -barrel and 3_{10} -helix η_3 . D, L180 is the central residue within a hydrophobic core, which connects helix α_1 with the small β -barrel and the β -sheet comprised of strands β_5 to β_7 . E, Residue E251 interacts with S217 and S220, thereby connecting the large β -barrel with helix α_2 . F, N240 interacts with the neighbouring residues D91 and R92 as well as the Moco by polar contacts and stabilizes the connection between the large β -barrel and the Moco binding site. G, D252 is the central residue between R298 and N316 and thereby fixates the relative position of helix α_3 towards the large β -barrel. H, L294 is the central residue within a hydrophobic core, which connects the helices α_2 , α_3 and η_2 .

SUPPLEMENTARY FIGURES

Figure S1

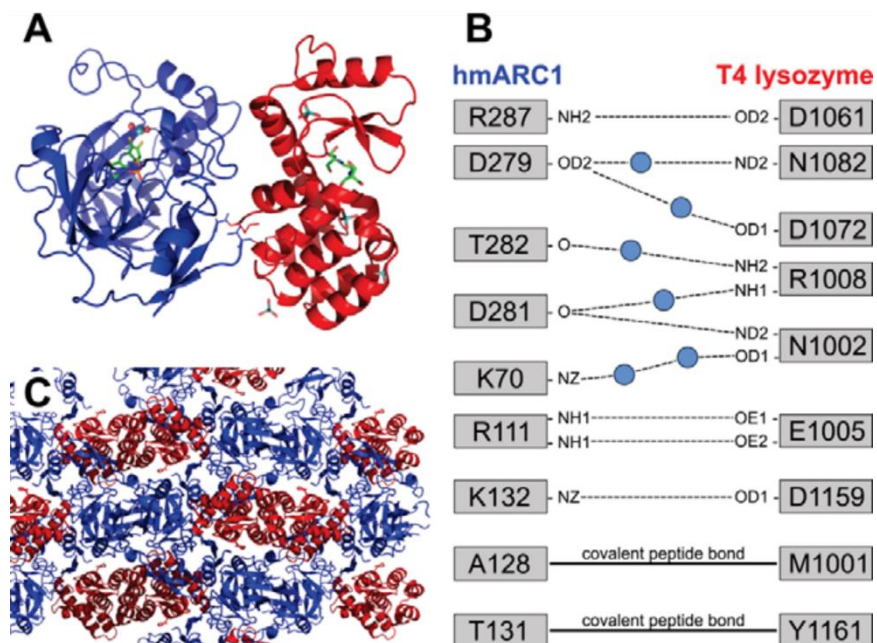


Figure S2

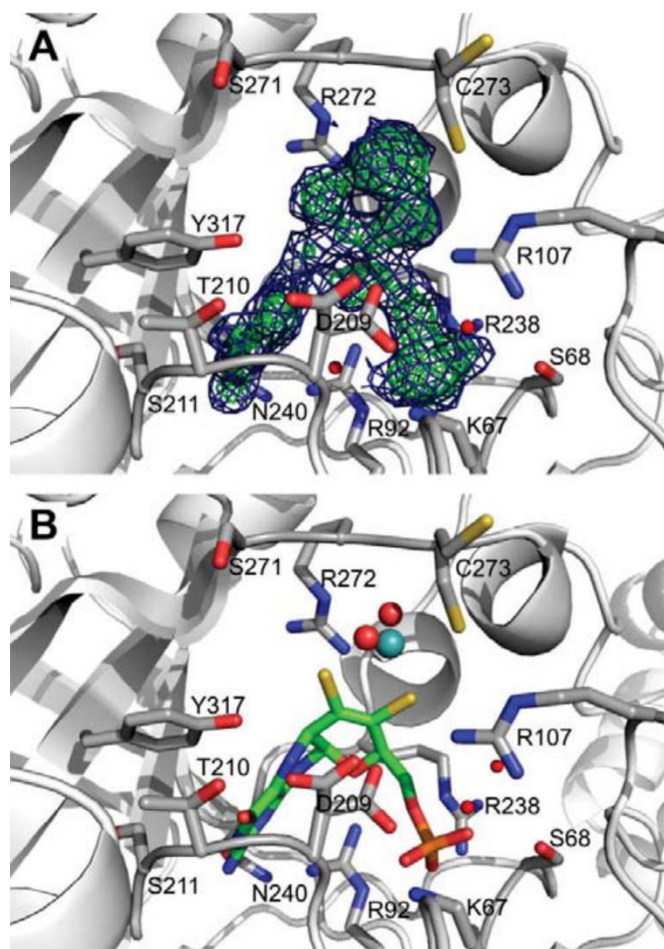


Figure S3

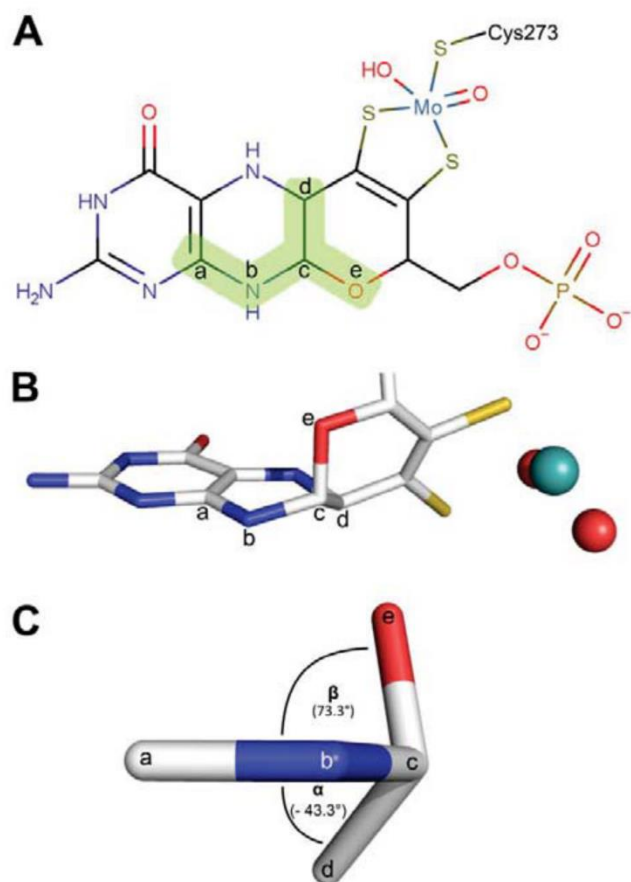


Figure S4

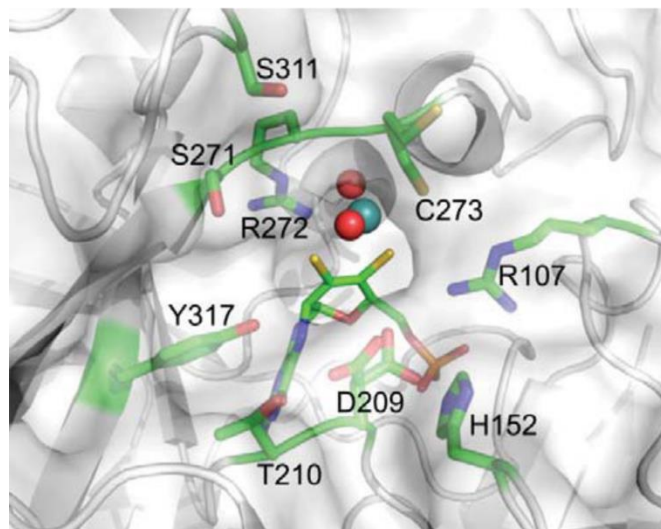


Figure S5

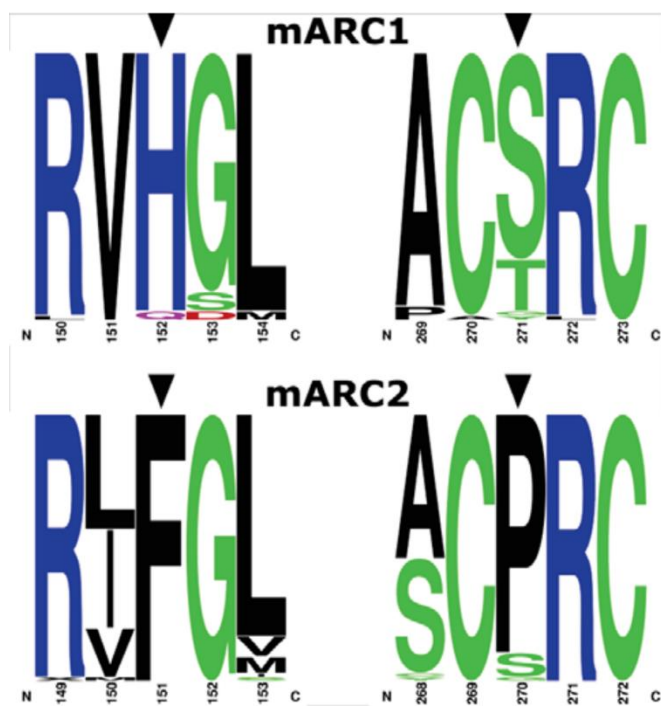


Figure S6

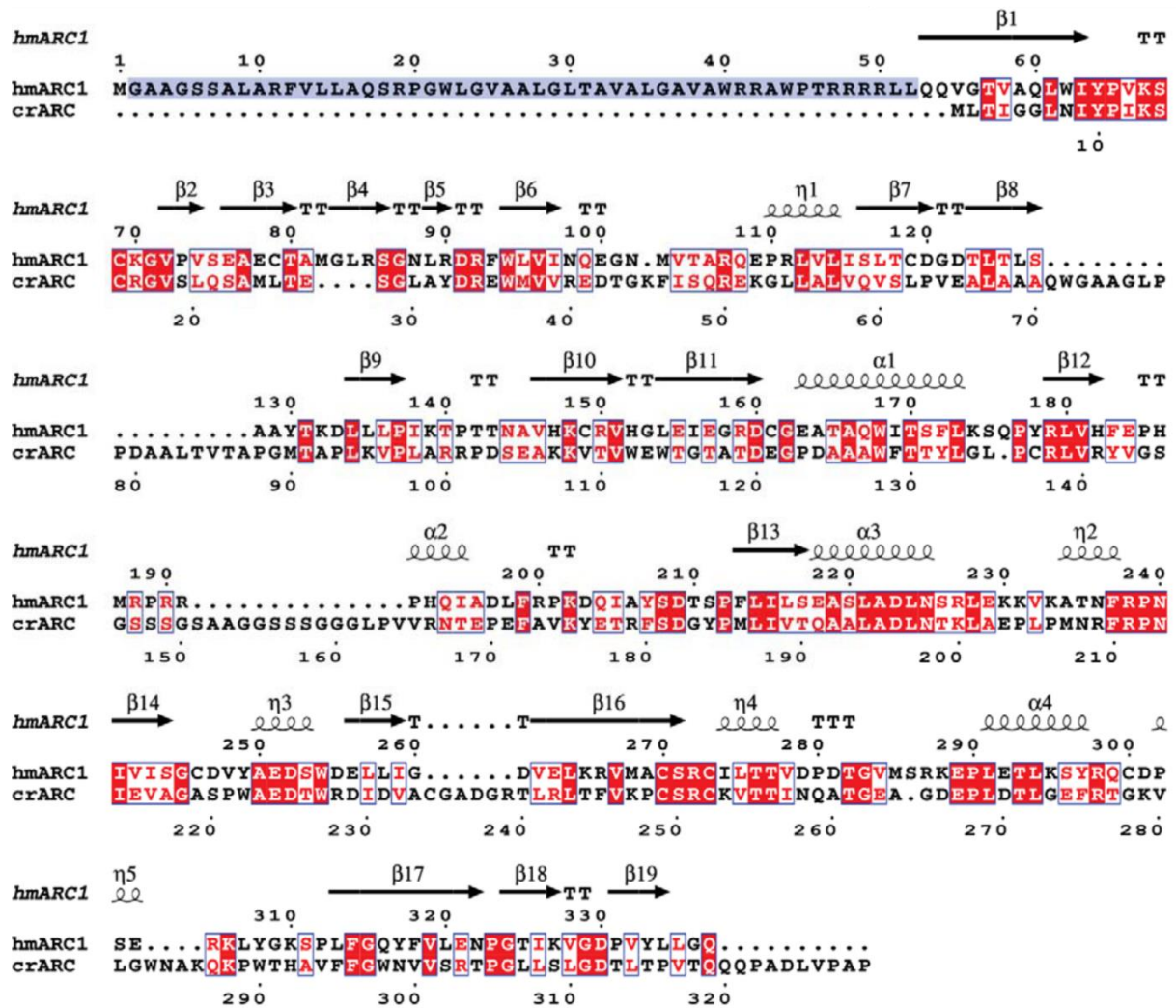


Figure S7

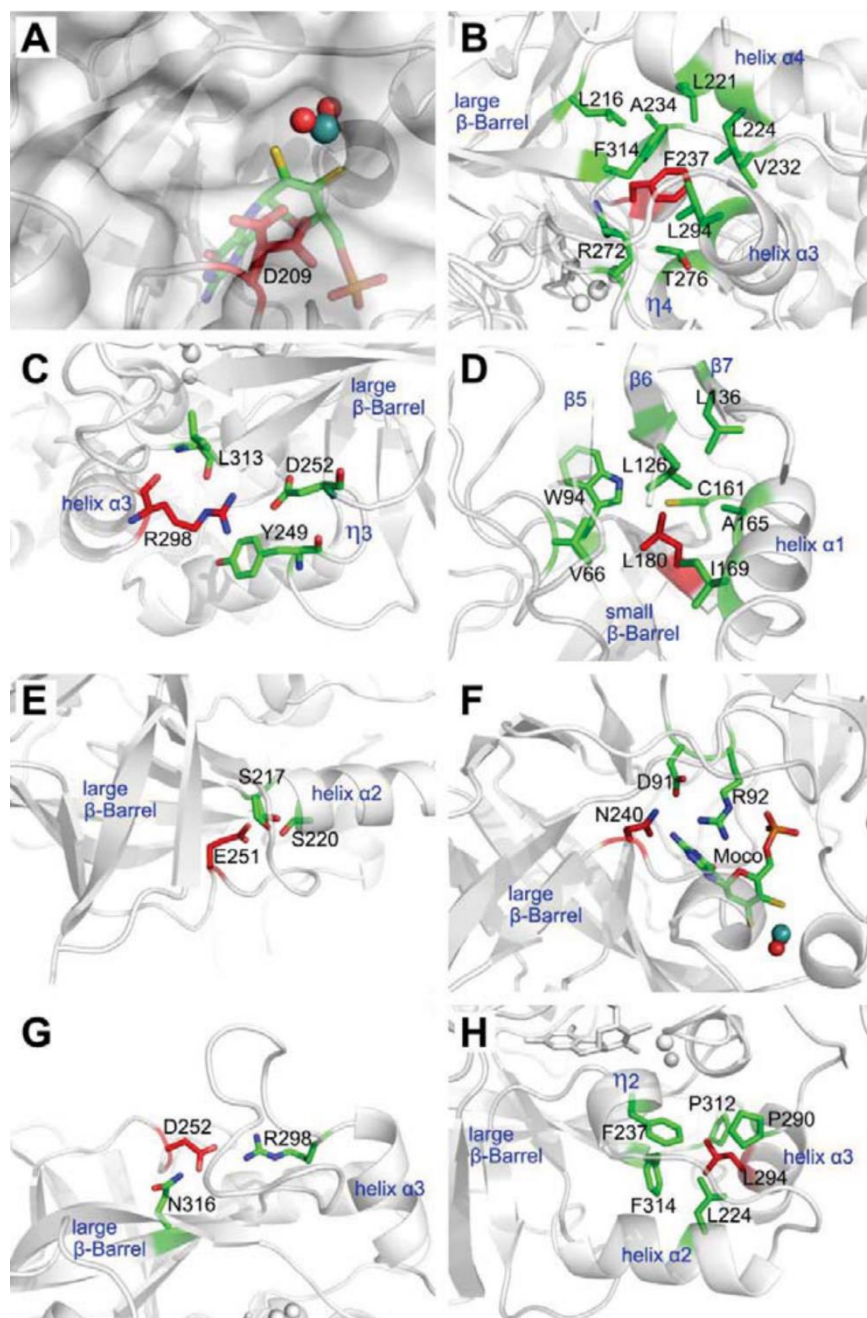


Table S1

Table S1. Data collection and refinement statistics

Data collection	
Space group*	P2 ₁ 2 ₁ 2 ₁
Cell dimensions*	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	61.13, 74.82, 110.73
α , β , γ (°)	90.00, 90.00, 90.00
Resolution (Å)*	43.57 – 1.78 (1.81 – 1.78)
<i>R</i> _{merge} *	0.08 (0.40)
<i>I</i> / σI^2	10.9 (2.6)
Mean $\langle I \rangle$ half-set correlation CC(1/2)*	99.5 (82.9)
Completeness (%)*	91.9 (23.9)
Redundancy*	6.8 (4.2)
Refinement	
Resolution (Å) [†]	43.57 – 1.78 (1.82 – 1.78)
No. reflections [†]	311122
<i>R</i> _{work} / <i>R</i> _{free} [†]	16.9 / 20.8 (27.0 / 34.6)
No. atoms	
Protein	3562
Molybdenum cofactor	27
Bis-TRIS propane	19
Molybdate	20
Phosphate	5
Water	439
<i>B</i> -factors [‡]	
Protein	29.1
Molybdenum cofactor	23.5
Bis-TRIS propane	26.1
Molybdate	41.1
Phosphate	57.8
Water	34.6
R.m.s. deviations [†]	
Bond lengths (Å)	0.0241
Bond angles (°)	2.7570

Values in parentheses are for highest-resolution shell. *Values as provided by AIMLESS after data processing, merging and scaling. [†]Calculated by REFMAC5 (33). [‡]Calculated using the program BAVEAGE. All programs used here are implemented within the CCP4 program package (46).

4 Summary

4.1 Structural characterization of ZvFMOs

One focus of this thesis was the structure determination of flavin-dependent monooxygenases from the African locust *Zonocerus variegatus* and subsequent structure-based analysis of the catalytic mechanism. Furthermore, structural differences in the composition of the active site and substrate entry path should be investigated in relation to quite diverse specific enzyme activities observed for the three ZvFMO paralogues.

All paralogues could be heterologously expressed in *E. coli*, were purified by affinity and size exclusion chromatography and exhibited enzymatic activity towards PA substrates. However, the isoform ZvFMOc appeared to be less stable than the others, which coincides with the crystallization success of these enzymes: suitable individual crystallization conditions could be established for ZvFMOa and ZvFMOd (named ZvPNO), whereas various crystallization attempts remained unsuccessful for isoform ZvFMOc. The obtained protein crystals could be used for X-ray diffraction experiments conducted at beamlines of the PETRA III synchrotron facility (DESY, Hamburg) and yielded datasets suitable for structure determination.

High-resolution structures of the ZvPNO-FAD and ZvPNO-FAD-NADP⁺ complexes were determined using molecular replacement and were refined to 1.9 and 1.6 Å, respectively. With the exception of a flexible loop region and the first twelve C-terminal amino acid residues of the expression construct, which could not be modelled due to the absence of appropriate electron density, the determined ZvPNO structures gave detailed insight into the three-dimensional arrangements of secondary structure elements as well as the position of individual side chains. Comparison of ZvPNO crystal structures with and without bound NADP⁺ cosubstrate revealed conformational changes within the enzyme that occur during the reaction cycle, such as the elongation of helix α 4, which facilitates NADPH/NADP⁺ coordination, and a bending of helix α 8, which might either contribute to an induced fit upon substrate binding or contribute to a mechanical product release mechanism.

Datasets obtained for ZvFMOa crystals were of lower quality and thus, no high-resolution model of this isoform could be generated. However, the data were sufficient for molecular replacement and determining a low-resolution structure of this enzyme. At least, the peptide backbone could be traced and modelled with confidence and even the position of some side chains within the substrate binding area and active site were observable. Therefore, the crystal structures of two different ZvFMO isoforms with differences in their specific enzymatic

activities could be directly compared to each other and to formerly solved crystal structures of FMOs from other species. While the overall fold of FMOs appears to be highly conserved throughout evolution from prokaryotes over simple eukaryotes up to more highly developed eukaryotes like *Zonocerus variegatus*, this solely applies to FMO subunits, whereas FMOs from different species exist in a multitude of different oligomeric arrangements. Moreover, the interaction sites of subunits within these oligomers can be quite diverse. ZvFMOs were identified to be homodimers by calibrated size exclusion chromatography as well as multi-angle laser light scattering (MALS). Their dimeric arrangement, which was observed in the crystal structures of both, ZvPNO and ZvFMOa, is so far unique and has not been described before.

Differences in the amino acid composition of the substrate binding area of ZvFMO isoforms were supposed to be responsible for their diverging specific enzyme activities. By comparing the crystal structures it was observed that the substrate entry path of isoform ZvFMOa is narrower than in ZvPNO, which is owed to the presence of sterically more demanding residues like phenylalanine and tyrosine. Therefore, different ZvFMOa variants were generated and analyzed in relation to their specific enzyme activity towards the substrate monocrotaline. Among these, loss-of-function as well as gain-of-function variants were identified, which led to the observation that modifications of the substrate entry path have a large impact on specific enzyme activity. Even single amino acid exchanges were shown to have significant effects.

In summary, two out of three ZvFMO isoforms could be successfully crystallized and structurally characterized. The determined crystal structures gave detailed insight into the coordination of the FAD prosthetic group as well as the NADPH/NADP⁺ cosubstrate. Moreover, conformational changes of the enzyme were observed that contribute a deeper understanding of the FMO catalytic cycle and underlying molecular mechanisms. The novel dimeric arrangement observed for ZvFMOs helps to complete our knowledge about possible interaction sites and oligomeric quaternary structures of this enzyme class. Furthermore, rationally designed ZvFMOa variants shed light on substrate specificity and indicate that even single mutation events during evolution may affect enzymatic activity and therefore help the host organism to quickly adapt to new xenobiotic substrates. The crystal structures of ZvFMOs presented in this thesis are the first available FMO structures to originate from more highly developed eukaryotes. Owing to their closer evolutionary relationship, they may serve as suitable templates for homology models of human FMOs.

4.2 Crystallization and structural characterization of mARC

The second focus of this dissertation was devoted to further investigations on human mARC enzymes. In order to identify new substrates, especially by applying novel methods that aim for high-throughput screenings, expression constructs and conditions were to be optimized for better protein yields. The main objective, however, was the elucidation of the three-dimensional mARC structure by X-ray crystallography. Therefore, suitable crystallization conditions were to be found and diffraction experiments had to be optimized in order to obtain complete datasets of sufficient quality for subsequent structure determination and interpretation.

Both human mARC paralogues (hmARC1 and hmARC2) could be heterologously expressed in the *E. coli* TP1000 strain. The overall yield of soluble protein could be improved by creating a new expression construct that was *N*-terminally truncated by the first 52 amino acid residues of wildtype mARC and carried a *C*-terminal hexahistidine tag. Purification by affinity chromatography and subsequent anion exchange chromatography yielded enzyme fractions of high purity and with slightly improved cofactor-saturation. These constructs were used for all biochemical and biophysical characterizations that are part of this thesis, except for structure determination.

In cooperation with the group of Prof. Dr. Paul V. Bernhardt (University of Queensland, Brisbane, Australia) an electrochemical assay for substrate and inhibitor screening, which is based on cyclic voltammetry, was successfully established. The principal utility and effectiveness of this method, which detects current changes at a working electrode (coupled to cytochrome b_5 and mARC) upon contact with a substrate dissolved in the surrounding analyte solution, was demonstrated using benzamidoxime as a model substrate.

In the group of Prof. Dr. Clement, hydroxamic acids could be identified as an additional class of *N*-hydroxylated mARC substrates by HPLC-based activity assays performed with recombinant enzymes and benzhydroxamic acid as a model substrate. It was shown that mARC is capable of reducing and thereby inactivating drugs containing hydroxamic acid functional groups. These findings highlight the need to monitor the *N*-reductive metabolism of new hydroxamic acid based drug candidates. However, it was found that the toxic metabolite *N*-hydroxyphenacetin was not reduced by the mARC enzyme systems, probably due to its additional substituent at the nitrogen atom, which might sterically interfere with the mARC active site and prevent its conversion. Future research will show, if *N*-substitution of hydroxamic acid drug candidates might enhance their metabolic stability while still retaining pharmacological activity.

First crystallization attempts with soluble mARC protein, as it was used for biochemical assays, were not successful. At best, spherulites could be obtained using hanging- or sitting-drop vapour diffusion crystallization methods. These spherulite-producing conditions could not be improved in order to yield crystals that were appropriate for diffraction experiments. Therefore, a fusion-protein assisted crystallization strategy was utilized that became famous with the first crystal structures of G protein-coupled receptors: introducing an internal T4 lysozyme-fusion into the native target protein sequence. Based on secondary and tertiary structure predictions for human mARC1, suitable insertion sites for T4L were chosen and fusion constructs were generated. One of these retained mARC-specific activity after expression and purification and was successfully crystallized. In addition to obtaining diffraction-quality crystals of mARC1, it was shown that this approach is not only useful for membrane proteins but can also be utilized for the crystallization of soluble proteins.

Phasing of datasets obtained for the mARC1-T4L fusion protein was achieved by iterative approaches using molecular replacement. Using the T4L crystal structure as a search model generated initial electron density maps that were easily interpretable with this fusion partner and already indicated the position of mARC. Keeping the T4L moiety fixed and taking into account additional search models for molecular replacement, which represented tertiary structure elements predicted for mARC by homology modelling, resulted in significantly improved calculated phases. These were sufficient to generate electron density maps suitable for model building and allowed determination and refinement of the high-resolution crystal structure of human mARC1.

The mARC1 crystal structure presented in this thesis was the first structure of a MOSC-family protein to be comprised of the MOSC domain (that all members have in common) as well as the MOSC_N domain, which had not been structurally characterized so far. Contrary to *in silico* predictions made by Anantharaman and Aravind⁸⁵ as well as structure motifs currently predicted by online databases, the *N*-terminal mARC domain does not exhibit a β -barrel fold. Instead, it comprises several β -sheets as well as helix α 1 and only contributes a single β -strand to a seven-stranded barrel that is mainly formed by strands of the MOSC domain. Furthermore, the mARC1 crystal structure is the first of the MOSC family to contain the molybdenum cofactor, which is tightly bound as a prosthetic group in between MOSC_N and MOSC domain and only exposes the active molybdenum site to the surface of the enzyme. This is opposed to evolutionary related molybdoenzymes only containing the MOSC domain, such as the recently structurally characterized YiiM, which supposedly has the Moco bound to the surface of the protein, only coordinated by a few residues.¹¹⁴

The mARC1 active site and also the substrate binding area were found to be completely surface-exposed and to be easily accessible. There are only a few spacial limitations in direct proximity to the reactive molybdenum center, which is most likely the reason for the broad substrate spectrum of this enzyme and its low specificity towards a quite diverse range of *N*-hydroxylated compounds. Two amino acid residues that are located in the potential substrate binding area were identified to be key residues that can be used to discriminate between mARC paralogues and might also be responsible for their in-/ability to reduce *N*-oxides. The binding site of mARC1 is slightly more polar and might therefore be able to bind and reduce *N*-oxides, whereas the more hydrophobic pocket of the mARC2 paralogue does not exhibit enzymatic activity towards these substances.

Previously performed extensive variant studies with the mARC homologue from *Chlamydomonas reinhardtii* aimed to elucidate the role of amino acid residues that are highly conserved among all mARC enzymes.¹¹⁵ While some of them were supposed to be involved in the catalytic cycle or the coordination of the cofactor, others had a clear impact on enzyme activity but could not clearly be addressed to underlying functions or mechanisms. The crystal structure of human mARC1 ultimately revealed the importance of these highly-conserved residues. Some are indeed responsible for catalytic activity or cofactor binding, but in most cases these residues were found to be located in the middle of hydrophobic cores or are part of hydrogen bond networks, depending on their hydrophobic or hydrophilic nature. They were identified to be crucial for the structural integrity of individual domains and the overall three-dimensional fold.

The molybdopterin backbone of the prosthetic Moco was found to exhibit an unexpected conformation. According to the findings of Rothery *et al.*,¹⁰² the observed dihedral angles would suggest an assignment of human mARC to the XO-family of molybdenum-dependent enzymes. However, based on previous biochemical investigations and biophysical characterizations, mARC was clearly assigned to the SO-family owing to the identified cysteine ligand and the absence of a terminal sulfur ligand. These opposing observations led to the hypothesis that mARC enzymes could be an evolutionary link between both molybdoenzyme families and might therefore represent the ancestors of the XO-family, which will need to be addressed in future research with respect to the evolution of molybdoenzymes.

5 Concluding Remarks and Prospects

This dissertation reports the crystallization strategies and three-dimensional structures of human mARC1 as well as FMOs from *Zonocerus variegatus* together with further biochemical characterizations of these enzymes. Thereby, it provides novel fundamental knowledge about different representatives of biotransformation enzymes. Still, the results discussed in this thesis can and should be understood as a starting point for further investigations and experiments that can now be rationally designed based on the elucidated crystal structures.

Additional ZvFMO variants could be generated in order to identify even more amino acid residues that have a significant impact on enzyme activity. Furthermore, activity assays should be performed with a number of different PAs in order to investigate PA specificity. Comparison of ZvFMOs to PNOs from other insects such as *Tyria jacobaeae* might explain why these enzymes are still quite specific towards PA substrates, despite their largely exposed active site that is theoretically accessible by many other compounds. To further elucidate the detailed underlying enzyme mechanism, the crystal structure of a ZvFMO in complex with one of its substrates or a specific inhibitor would be of great advantage. However, all previous co-crystallization and soaking experiments conducted so far were not successful and will need some careful improvement or require a different strategy.

To complete our knowledge about human mARC enzymes, the crystallization of the mARC2 paralogue should be attempted using the same fusion-protein strategy that has proven successful for mARC1. Comparison of both enzymes, especially of their substrate binding area and active sites, could help to identify features that are responsible for their in-/ability to catalyze the reduction of *N*-oxides. These different substrate preferences might be attributed to the two key residues identified in this work that can be used to discriminate between both paralogues. In order to check this hypothesis, potential gain-of-function variants of the mARC2 paralogue should be generated that have exchanged either F151 for a histidine or P270 for a serine or contain both exchanges.

Furthermore, the mARC catalytic cycle should be thoroughly investigated. Thus, crystal structures of this enzyme trapped in different states during catalysis should be a main objective of future research. Owing to the very few limitations of substrate binding on the surface of the enzyme, crystals of mARC in complex with a substrate will be hard to obtain. However, Patrick Indorf (AG Prof. Dr. Clement) recently identified *N*-hydroxyurethane as a potential mARC1-specific inhibitor, which would be ideal for co-crystallization or soaking experiments. This compound might bind to S271 via a carbamylation reaction, orient its

N-hydroxylated moiety towards the active site and therefore give insight into the binding mode of substrates and inhibitors.

Another aspect to be considered in the future is the evolutionary development of different molybdoenzyme families. The mARC crystal structure provides first indications for an evolutionary link between the SO and the later-emerging XO family. To further investigate this hypothesis the enzyme Moco sulfurase should be structurally characterized. This enzyme is the direct biochemical link between the both aforementioned families of molybdenum-dependent enzymes, because it accepts the SO Moco as a substrate and catalyzes its sulfuration to give the form of Moco that is coordinated within enzymes belonging to the XO family.¹¹⁶ Moco sulfurases comprise two functional domains, an *N*-terminal cysteine desulfurase domain and a *C*-terminal (MOSC) domain, where the Moco is bound and converted. Since the *C*-terminus of these enzymes shares a high sequence identity with mARC proteins, it was proposed that Moco sulfurases evolved from mARC-like proteins through domain fusion with a cysteine desulfurase-like domain.⁸⁵ Thus, Moco sulfurases might not only be a direct biochemical but also evolutionary link between the SO and XO family. Elucidating the three-dimensional structure of these enzymes might give further insight into the development of Moco-dependent proteins.

The crystal structures reported in this PhD thesis belong to two opposing families of biotransformation enzymes: monooxygenases and reductases. FMOs and mARC, among other crucial biotransformation enzymes, contribute to xenobiotic metabolism by maintaining a viable redox equilibrium. Recently, Patrick Indorf and Jennifer Schneider from the group of Prof. Dr. Clement identified pyrrolizidine alkaloid *N*-oxides and trimethylamine *N*-oxide as substrates of human mARC1. These substances have so far been known as typical products generated by FMOs: the human FMO3 is well known for its specific conversion of trimethylamine into the non-odorous *N*-oxide, and ZvFMOs oxygenize PAs as part of a detoxification process. Therefore, both enzymes discussed in this dissertation are representatives of direct physiological counterparts, even though they don't originate from the same organism. However, this emphasizes the need for a better understanding of the many facets of xenobiotic conversion. Toxication or detoxification of xeno- and endobiotics is oftentimes just a matter of tissue localization and expression levels of different biotransformation enzymes. Thus, it is not only important to study these enzymes on a molecular level and determine their specific activities *in vitro* but to take the whole organism into perspective, when investigating the metabolism of xenobiotics, especially novel drug candidates.

6 References

- 1 Ishima, R. & Torchia, D. A. Protein dynamics from NMR. *Nature structural biology* **7**, 740-743, doi:10.1038/78963 (2000).
- 2 Adrian, M., Dubochet, J., Lepault, J. & McDowell, A. W. Cryo-electron microscopy of viruses. *Nature* **308**, 32-36 (1984).
- 3 Chapman, H. N. *et al.* Femtosecond diffractive imaging with a soft-X-ray free-electron laser. *Nature Physics* **2**, 839-843, doi:10.1038/nphys461 (2006).
- 4 Rupp, B. *Biomolecular crystallography : principles, practice, and application to structural biology.* (Garland Science, 2010).
- 5 Messerschmidt, A. *X-ray crystallography of biomacromolecules : a practical guide.* (Wiley-VCH, 2007).
- 6 McPherson, A. Introduction to protein crystallization. *Methods* **34**, 254-265, doi:10.1016/j.ymeth.2004.03.019 (2004).
- 7 McPherson, A. *Introduction to macromolecular crystallography.* 2nd ed edn, (Wiley-Blackwell, 2009).
- 8 Rhodes, G. *Crystallography made crystal clear : a guide for users of macromolecular models.* 3rd edn, (Elsevier/Academic Press, 2006).
- 9 Wlodawer, A., Minor, W., Dauter, Z. & Jaskolski, M. Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *The FEBS journal* **275**, 1-21, doi:10.1111/j.1742-4658.2007.06178.x (2008).
- 10 Mutschler, E., Geisslinger, G., Kroemer, H. K. & Schäfer-Korting, M. *Mutschler Arzneimittelwirkungen: Lehrbuch der Pharmakologie und Toxikologie.* (Wiss. Verlag-Ges., 2008).
- 11 Testai, E. The drug-metabolizing enzymatic system and the experimental tools used for in vitro toxicology for metabolic studies. *Cell Biol Toxicol* **17**, 271-285 (2001).
- 12 Taxak, N. & Bharatam, P. V. Drug metabolism. *Resonance* **19**, 259-282 (2014).
- 13 Hlavica, P. N-oxidative transformation of free and N-substituted amine functions by cytochrome P450 as means of bioactivation and detoxication. *Drug Metab Rev* **34**, 451-477, doi:10.1081/DMR-120005646 (2002).
- 14 Etmayer, P., Amidon, G. L., Clement, B. & Testa, B. Lessons learned from marketed and investigational prodrugs. *J Med Chem* **47**, 2393-2404, doi:10.1021/jm0303812 (2004).
- 15 Shu, Y. Z., Johnson, B. M. & Yang, T. J. Role of biotransformation studies in minimizing metabolism-related liabilities in drug discovery. *AAPS J* **10**, 178-192, doi:10.1208/s12248-008-9016-9 (2008).
- 16 Constantino, L. *et al.* Metabolism of primaquine by liver homogenate fractions. Evidence for monoamine oxidase and cytochrome P450 involvement in the oxidative deamination of primaquine to carboxyprimaquine. *Exp Toxicol Pathol* **51**, 299-303, doi:10.1016/S0940-2993(99)80010-4 (1999).

- 17 Nussler, A. K. *et al.* The suitability of hepatocyte culture models to study various aspects of drug metabolism. *ALTEX* **18**, 91-101 (2001).
- 18 Salonen, J. S. *et al.* Comparative studies on the cytochrome p450-associated metabolism and interaction potential of selegiline between human liver-derived in vitro systems. *Drug Metab Dispos* **31**, 1093-1102, doi:10.1124/dmd.31.9.1093 (2003).
- 19 Vuppugalla, R. & Mehvar, R. Short-term inhibitory effects of nitric oxide on cytochrome P450-mediated drug metabolism: time dependency and reversibility profiles in isolated perfused rat livers. *Drug Metab Dispos* **32**, 1446-1454, doi:10.1124/dmd.104.001487 (2004).
- 20 Nelson, D. R. *et al.* P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* **6**, 1-42 (1996).
- 21 Hayaishi, O. *Oxygenases*. (Academic Press, 1962).
- 22 Hasler, J. A. *et al.* Human cytochromes P450. *Molecular Aspects of Medicine* **20**, 1–137, doi:10.1016/s0098-2997(99)00005-9 (1999).
- 23 Bryant, B. & Knights, K. *Pharmacology for Health Professionals ebook*. (Elsevier Health Sciences, 2014).
- 24 Coon, M. J., Vaz, A. D. & Bestervelt, L. L. Cytochrome P450 2: peroxidative reactions of diversozymes. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **10**, 428-434 (1996).
- 25 Porter, T. D. & Coon, M. J. Cytochrome P-450. Multiplicity of isoforms, substrates, and catalytic and regulatory mechanisms. *J Biol Chem* **266**, 13469-13472 (1991).
- 26 Naumann, C., Hartmann, T. & Ober, D. Evolutionary recruitment of a flavin-dependent monooxygenase for the detoxification of host plant-acquired pyrrolizidine alkaloids in the alkaloid-defended arctiid moth *Tyria jacobaeae*. *Proc Natl Acad Sci U S A* **99**, 6085-6090, doi:10.1073/pnas.082674499 (2002).
- 27 Hao da, C., Chen, S. L., Mu, J. & Xiao, P. G. Molecular phylogeny, long-term evolution, and functional divergence of flavin-containing monooxygenases. *Genetica* **137**, 173-187, doi:10.1007/s10709-009-9382-y (2009).
- 28 Cashman, J. R. & Zhang, J. Human flavin-containing monooxygenases. *Annual review of pharmacology and toxicology* **46**, 65-100, doi:10.1146/annurev.pharmtox.46.120604.141043 (2006).
- 29 Phillips, I., Francois, A. & Shephard, E. The Flavin-Containing Monooxygenases (FMOs): Genetic Variation and its Consequences for the Metabolism of Therapeutic Drugs. *Current Pharmacogenomics* **5**, 292–313, doi:10.2174/157016007782793683 (2007).
- 30 Cashman, J. R. in *Enzyme Systems that Metabolise Drugs and Other Xenobiotics* (ed Costas Ioannides) 67–93 (John Wiley & Sons, Ltd, 2001).
- 31 Ziegler, D. M. An overview of the mechanism, substrate specificities, and structure of FMOs. *Drug Metab Rev* **34**, 503-511, doi:10.1081/DMR-120005650 (2002).
- 32 Jakoby, W. B. & Ziegler, D. M. The enzymes of detoxication. *J Biol Chem* **265**, 20715-20718 (1990).

- 33 Lang, D. H. *et al.* Isoform specificity of trimethylamine N-oxygenation by human flavin-containing monooxygenase (FMO) and P450 enzymes: selective catalysis by FMO3. *Biochem Pharmacol* **56**, 1005-1012 (1998).
- 34 Phillips, I. R. & Shephard, E. A. Flavin-containing monooxygenases: mutations, disease and drug response. *Trends Pharmacol Sci* **29**, 294-301, doi:10.1016/j.tips.2008.03.004 (2008).
- 35 Ojha, S., Meng, E. C. & Babbitt, P. C. Evolution of function in the "two dinucleotide binding domains" flavoproteins. *PLoS computational biology* **3**, e121, doi:10.1371/journal.pcbi.0030121 (2007).
- 36 van Berkel, W. J., Kamerbeek, N. M. & Fraaije, M. W. Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J Biotechnol* **124**, 670-689, doi:10.1016/j.jbiotec.2006.03.044 (2006).
- 37 Alfieri, A., Malito, E., Orru, R., Fraaije, M. W. & Mattevi, A. Revealing the moonlighting role of NADP in the structure of a flavin-containing monooxygenase. *Proc Natl Acad Sci U S A* **105**, 6572-6577, doi:10.1073/pnas.0800859105 (2008).
- 38 Ziegler, D. M. Recent studies on the structure and function of multisubstrate flavin-containing monooxygenases. *Annu Rev Pharmacol Toxicol* **33**, 179-199, doi:10.1146/annurev.pa.33.040193.001143 (1993).
- 39 Cashman, J. R. Structural and catalytic properties of the mammalian flavin-containing monooxygenase. *Chem Res Toxicol* **8**, 166-181 (1995).
- 40 Robinson, R., Badieyan, S. & Sobrado, P. C4a-hydroperoxyflavin formation in N-hydroxylating flavin monooxygenases is mediated by the 2'-OH of the nicotinamide ribose of NADP(+). *Biochemistry* **52**, 9089-9091, doi:10.1021/bi4014903 (2013).
- 41 Beaty, N. B. & Ballou, D. P. The oxidative half-reaction of liver microsomal FAD-containing monooxygenase. *J Biol Chem* **256**, 4619-4625 (1981).
- 42 Pieterse, C. M. & Dicke, M. Plant interactions with microbes and insects: from molecular mechanisms to ecology. *Trends Plant Sci* **12**, 564-569, doi:10.1016/j.tplants.2007.09.004 (2007).
- 43 Hartmann, T. & Ober, D. in *Induced Plant Resistance to Herbivory* 213-231 (Springer, 2008).
- 44 Fu, P. P., Xia, Q., Lin, G. & Chou, M. W. Pyrrolizidine alkaloids--genotoxicity, metabolism enzymes, metabolic activation, and mechanisms. *Drug Metab Rev* **36**, 1-55, doi:10.1081/DMR-120028426 (2004).
- 45 Ober, D. & Kaltenecker, E. Pyrrolizidine alkaloid biosynthesis, evolution of a pathway in plant secondary metabolism. *Phytochemistry* **70**, 1687-1695, doi:10.1016/j.phytochem.2009.05.017 (2009).
- 46 Dueker, S. R., Lame, M. W. & Segall, H. J. Hydrolysis of pyrrolizidine alkaloids by guinea pig hepatic carboxylesterases. *Toxicology and applied pharmacology* **117**, 116-121 (1992).
- 47 Lin, G., Cui, Y. Y. & Hawes, E. M. Microsomal formation of a pyrrolic alcohol glutathione conjugate of clivorine. Firm evidence for the formation of a pyrrolic metabolite of an otonecine-type pyrrolizidine alkaloid. *Drug metabolism and disposition: the biological fate of chemicals* **26**, 181-184 (1998).

- 48 Huan, J. Y., Miranda, C. L., Buhler, D. R. & Cheeke, P. R. Species differences in the hepatic microsomal enzyme metabolism of the pyrrolizidine alkaloids. *Toxicology letters* **99**, 127-137 (1998).
- 49 Williams, D. E., Reed, R. L., Kedzierski, B., Ziegler, D. M. & Buhler, D. R. The role of flavin-containing monooxygenase in the N-oxidation of the pyrrolizidine alkaloid senecionine. *Drug metabolism and disposition: the biological fate of chemicals* **17**, 380-386 (1989).
- 50 Lindigkeit, R. *et al.* The two faces of pyrrolizidine alkaloids: the role of the tertiary amine and its N-oxide in chemical defense of insects with acquired plant alkaloids. *Eur J Biochem* **245**, 626-636 (1997).
- 51 Sehmeyer, S. *et al.* Flavin-dependent monooxygenases as a detoxification mechanism in insects: new insights from the arctiids (Lepidoptera). *PLoS One* **5**, e10435, doi:10.1371/journal.pone.0010435 (2010).
- 52 Hartmann, T. & Ober, D. in *Biosynthesis Vol. 209 Topics in Current Chemistry* (eds Finian J. Leeper & John C. Vederas) 207–243 (Springer Berlin Heidelberg, 2000).
- 53 Langel, D. & Ober, D. Evolutionary recruitment of a flavin-dependent monooxygenase for stabilization of sequestered pyrrolizidine alkaloids in arctiids. *Phytochemistry* **72**, 1576-1584, doi:10.1016/j.phytochem.2010.12.014 (2011).
- 54 Hu, Y. & Ribbe, M. W. Biosynthesis of Nitrogenase FeMoco. *Coord Chem Rev* **255**, 1218-1224, doi:10.1016/j.ccr.2010.11.018 (2011).
- 55 Hille, R., Hall, J. & Basu, P. The mononuclear molybdenum enzymes. *Chem Rev* **114**, 3963-4038, doi:10.1021/cr400443z (2014).
- 56 Mendel, R. R. & Kruse, T. Cell biology of molybdenum in plants and humans. *Biochim Biophys Acta* **1823**, 1568-1579, doi:10.1016/j.bbamcr.2012.02.007 (2012).
- 57 Hille, R., Nishino, T. & Bittner, F. Molybdenum enzymes in higher organisms. *Coord Chem Rev* **255**, 1179-1205, doi:10.1016/j.ccr.2010.11.034 (2011).
- 58 Warner, C. K. & Finnerty, V. Molybdenum hydroxylases in Drosophila. II. Molybdenum cofactor in xanthine dehydrogenase, aldehyde oxidase and pyridoxal oxidase. *Mol Gen Genet* **184**, 92-96 (1981).
- 59 Lewis, N. J., Hurt, P., Sealy-Lewis, H. M. & Scazzocchio, C. The genetic control of the molybdoflavoproteins in *Aspergillus nidulans*. IV. A comparison between purine hydroxylase I and II. *Eur J Biochem* **91**, 311-316 (1978).
- 60 Skipper, L., Campbell, W. H., Mertens, J. A. & Lowe, D. J. Pre-steady-state kinetic analysis of recombinant Arabidopsis NADH:nitrate reductase: rate-limiting processes in catalysis. *J Biol Chem* **276**, 26995-27002, doi:10.1074/jbc.M100356200 (2001).
- 61 Beligni, M. V. & Lamattina, L. Nitric oxide stimulates seed germination and de-etiolation, and inhibits hypocotyl elongation, three light-inducible responses in plants. *Planta* **210**, 215-221, doi:10.1007/PL00008128 (2000).
- 62 Durner, J. & Klessig, D. F. Nitric oxide as a signal in plants. *Curr Opin Plant Biol* **2**, 369-374 (1999).
- 63 Eilers, T. *et al.* Identification and biochemical characterization of *Arabidopsis thaliana* sulfite oxidase. A new player in plant sulfur metabolism. *J Biol Chem* **276**, 46989-46994, doi:10.1074/jbc.M108078200 (2001).

- 64 Nowak, K. *et al.* Peroxisomal localization of sulfite oxidase separates it from chloroplast-based sulfur assimilation. *Plant Cell Physiol* **45**, 1889-1894, doi:10.1093/pcp/pch212 (2004).
- 65 Hansch, R. *et al.* Plant sulfite oxidase as novel producer of H₂O₂: combination of enzyme catalysis with a subsequent non-enzymatic reaction step. *J Biol Chem* **281**, 6884-6888, doi:10.1074/jbc.M513054200 (2006).
- 66 Randewig, D. *et al.* Sulfite oxidase controls sulfur metabolism under SO₂ exposure in *Arabidopsis thaliana*. *Plant Cell Environ* **35**, 100-115, doi:10.1111/j.1365-3040.2011.02420.x (2012).
- 67 Yesbergenova, Z. *et al.* The plant Mo-hydroxylases aldehyde oxidase and xanthine dehydrogenase have distinct reactive oxygen species signatures and are induced by drought and abscisic acid. *Plant J* **42**, 862-876, doi:10.1111/j.1365-313X.2005.02422.x (2005).
- 68 Montalbini, P. Inhibition of Hypersensitive Response by Allopurinol Applied to the Host in the Incompatible Relationship between *Phaseolus vulgaris* and *Uromyces phaseoli*. *Journal of Phytopathology* **134**, 218-228, doi:doi:10.1111/j.1439-0434.1992.tb01230.x (1992).
- 69 Hesberg, C., Hansch, R., Mendel, R. R. & Bittner, F. Tandem orientation of duplicated xanthine dehydrogenase genes from *Arabidopsis thaliana*: differential gene expression and enzyme activities. *J Biol Chem* **279**, 13547-13554, doi:10.1074/jbc.M312929200 (2004).
- 70 Pastori, G. M. & Del Rio, L. A. Natural Senescence of Pea Leaves (An Activated Oxygen-Mediated Function for Peroxisomes). *Plant Physiol* **113**, 411-418 (1997).
- 71 Agarwal, A., Banerjee, A. & Banerjee, U. C. Xanthine oxidoreductase: a journey from purine metabolism to cardiovascular excitation-contraction coupling. *Crit Rev Biotechnol* **31**, 264-280, doi:10.3109/07388551.2010.527823 (2011).
- 72 Vorbach, C., Scriven, A. & Capecchi, M. R. The housekeeping gene xanthine oxidoreductase is necessary for milk fat droplet enveloping and secretion: gene sharing in the lactating mammary gland. *Genes Dev* **16**, 3223-3235, doi:10.1101/gad.1032702 (2002).
- 73 Rodriguez-Trelles, F., Tarrío, R. & Ayala, F. J. Convergent neofunctionalization by positive Darwinian selection after ancient recurrent duplications of the xanthine dehydrogenase gene. *Proc Natl Acad Sci U S A* **100**, 13413-13417, doi:10.1073/pnas.1835646100 (2003).
- 74 Seo, M. *et al.* The *Arabidopsis* aldehyde oxidase 3 (AAO3) gene product catalyzes the final step in abscisic acid biosynthesis in leaves. *Proc Natl Acad Sci U S A* **97**, 12908-12913, doi:10.1073/pnas.220426197 (2000).
- 75 Verslues, P. E. & Zhu, J. K. Before and beyond ABA: upstream sensing and internal signals that determine ABA accumulation and response under abiotic stress. *Biochem Soc Trans* **33**, 375-379, doi:10.1042/BST0330375 (2005).
- 76 Abbenante, G. & Fairlie, D. P. Protease inhibitors in the clinic. *Med Chem* **1**, 71-104 (2005).
- 77 Shirk, R. A. & Vlasuk, G. P. Inhibitors of Factor VIIa/tissue factor. *Arterioscler Thromb Vasc Biol* **27**, 1895-1900, doi:10.1161/ATVBAHA.107.148304 (2007).

- 78 Meyer, J. E. *et al.* The Oral Serine Protease Inhibitor WX-671 - First Experience in Patients with Advanced Head and Neck Carcinoma. *Breast Care (Basel)* **3**, 20-24, doi:10.1159/000151736 (2008).
- 79 Werbovets, K. Diamidines as antitrypanosomal, antileishmanial and antimalarial agents. *Curr Opin Investig Drugs* **7**, 147-157 (2006).
- 80 Fuller, A. T. Antibacterial action of some aromatic amines, amidines, amidoximes, guanidines and diguanidines. *Biochem J* **41**, 403-408 (1947).
- 81 Salom-Roig, X. J., Hamze, A., Calas, M. & Vial, H. J. Dual molecules as new antimalarials. *Comb Chem High Throughput Screen* **8**, 49-62 (2005).
- 82 Weller, T. *et al.* Orally active fibrinogen receptor antagonists. 2. Amidoximes as prodrugs of amidines. *J Med Chem* **39**, 3139-3147, doi:10.1021/jm9509298 (1996).
- 83 Havemeyer, A. *et al.* Identification of the missing component in the mitochondrial benzamidoxime prodrug-converting system as a novel molybdenum enzyme. *J Biol Chem* **281**, 34796-34802, doi:10.1074/jbc.M607697200 (2006).
- 84 Ott, G., Havemeyer, A. & Clement, B. The mammalian molybdenum enzymes of mARC. *J Biol Inorg Chem* **20**, 265-275, doi:10.1007/s00775-014-1216-4 (2015).
- 85 Anantharaman, V. & Aravind, L. MOSC domains: ancient, predicted sulfur-carrier domains, present in diverse metal-sulfur cluster biosynthesis proteins including Molybdenum cofactor sulfurases. *FEMS Microbiol Lett* **207**, 55-61 (2002).
- 86 Klein, J. M. *et al.* The mitochondrial amidoxime-reducing component (mARC1) is a novel signal-anchored protein of the outer mitochondrial membrane. *J Biol Chem* **287**, 42795-42803, doi:10.1074/jbc.M112.419424 (2012).
- 87 Havemeyer, A., Lang, J. & Clement, B. The fourth mammalian molybdenum enzyme mARC: current state of research. *Drug Metab Rev* **43**, 524-539, doi:10.3109/03602532.2011.608682 (2011).
- 88 Gruenewald, S. *et al.* The fourth molybdenum containing enzyme mARC: cloning and involvement in the activation of N-hydroxylated prodrugs. *J Med Chem* **51**, 8173-8177, doi:10.1021/jm8010417 (2008).
- 89 Jakobs, H. H. *et al.* The N-reductive system composed of mitochondrial amidoxime reducing component (mARC), cytochrome b5 (CYB5B) and cytochrome b5 reductase (CYB5R) is regulated by fasting and high fat diet in mice. *PLoS One* **9**, e105371, doi:10.1371/journal.pone.0105371 (2014).
- 90 Havemeyer, A. *et al.* Reduction of N-hydroxy-sulfonamides, including N-hydroxy-valdecoxib, by the molybdenum-containing enzyme mARC. *Drug Metab Dispos* **38**, 1917-1921, doi:10.1124/dmd.110.032813 (2010).
- 91 Krompholz, N. *et al.* The mitochondrial Amidoxime Reducing Component (mARC) is involved in detoxification of N-hydroxylated base analogues. *Chem Res Toxicol* **25**, 2443-2450, doi:10.1021/tx300298m (2012).
- 92 Plitzko, B., Havemeyer, A., Kunze, T. & Clement, B. The pivotal role of the mitochondrial amidoxime reducing component 2 in protecting human cells against apoptotic effects of the base analog N6-hydroxylaminopurine. *J Biol Chem* **290**, 10126-10135, doi:10.1074/jbc.M115.640052 (2015).
- 93 Cribb, A. E., Spielberg, S. P. & Griffin, G. P. N4-hydroxylation of sulfamethoxazole by cytochrome P450 of the cytochrome P4502C subfamily and reduction of

- sulfamethoxazole hydroxylamine in human and rat hepatic microsomes. *Drug Metab Dispos* **23**, 406-414 (1995).
- 94 Ott, G. *et al.* Reduction of sulfamethoxazole hydroxylamine (SMX-HA) by the mitochondrial amidoxime reducing component (mARC). *Chem Res Toxicol* **27**, 1687-1695, doi:10.1021/tx500174u (2014).
- 95 Boucher, J. L., Moali, C. & Tenu, J. P. Nitric oxide biosynthesis, nitric oxide synthase inhibitors and arginase competition for L-arginine utilization. *Cell Mol Life Sci* **55**, 1015-1028 (1999).
- 96 Kotthaus, J. *et al.* Reduction of N(omega)-hydroxy-L-arginine by the mitochondrial amidoxime reducing component (mARC). *Biochem J* **433**, 383-391, doi:10.1042/BJ20100960 (2011).
- 97 Malik, A. N., Rossios, C., Al-Kafaji, G., Shah, A. & Page, R. A. Glucose regulation of CDK7, a putative thiol related gene, in experimental diabetic nephropathy. *Biochem Biophys Res Commun* **357**, 237-244, doi:10.1016/j.bbrc.2007.03.132 (2007).
- 98 Neve, E. P. *et al.* Expression and Function of mARC: Roles in Lipogenesis and Metabolic Activation of Ximelagatran. *PLoS One* **10**, e0138487, doi:10.1371/journal.pone.0138487 (2015).
- 99 Neve, E. P. *et al.* Amidoxime reductase system containing cytochrome b5 type B (CYB5B) and MOSC2 is of importance for lipid synthesis in adipocyte mitochondria. *J Biol Chem* **287**, 6307-6317, doi:10.1074/jbc.M111.328237 (2012).
- 100 Wahl, B. *et al.* Biochemical and spectroscopic characterization of the human mitochondrial amidoxime reducing components hmARC-1 and hmARC-2 suggests the existence of a new molybdenum enzyme family in eukaryotes. *J Biol Chem* **285**, 37847-37859, doi:10.1074/jbc.M110.169532 (2010).
- 101 Yang, J. *et al.* Oxy and hydroxyl radical transfer in mitochondrial amidoxime reducing component-catalyzed nitrite reduction. *J Am Chem Soc* **137**, 5276-5279, doi:10.1021/jacs.5b01112 (2015).
- 102 Rothery, R. A., Stein, B., Solomonson, M., Kirk, M. L. & Weiner, J. H. Pyranopterin conformation defines the function of molybdenum and tungsten enzymes. *Proc Natl Acad Sci U S A* **109**, 14773-14778, doi:10.1073/pnas.1200671109 (2012).
- 103 Kubitza, C. *et al.* Crystal structure of pyrrolizidine alkaloid N-oxygenase from the grasshopper *Zonocerus variegatus*. *Acta Crystallographica Section D* **74**, 422-432, doi:10.1107/S2059798318003510 (2018).
- 104 Kalimuthu, P. *et al.* Human mitochondrial amidoxime reducing component (mARC): An electrochemical method for identifying new substrates and inhibitors. *Electrochemistry Communications* **84**, 90-93, doi:10.1016/j.elecom.2017.10.003 (2017).
- 105 Brogden, R. N., Pinder, R. M., Sawyer, P. R., Speight, T. M. & Avery, G. S. Bufenamac: a review of its pharmacological properties and therapeutic efficacy in inflammatory dermatoses. *Drugs* **10**, 351-356 (1975).
- 106 Barb, A. W. *et al.* Inhibition of lipid A biosynthesis as the primary mechanism of CHIR-090 antibiotic activity in *Escherichia coli*. *Biochemistry* **46**, 3793-3802, doi:10.1021/bi6025165 (2007).

- 107 Halouska, S. *et al.* Metabolomics analysis identifies d-Alanine-d-Alanine ligase as the primary lethal target of d-Cycloserine in mycobacteria. *J Proteome Res* **13**, 1065-1076, doi:10.1021/pr4010579 (2014).
- 108 Dokmanovic, M., Clarke, C. & Marks, P. A. Histone deacetylase inhibitors: overview and perspectives. *Mol Cancer Res* **5**, 981-989, doi:10.1158/1541-7786.MCR-07-0324 (2007).
- 109 Verma, R. P. Hydroxamic acids as matrix metalloproteinase inhibitors. *EXS* **103**, 137-176, doi:10.1007/978-3-0348-0364-9_5 (2012).
- 110 Wagner, J. M., Hackanson, B., Lubbert, M. & Jung, M. Histone deacetylase (HDAC) inhibitors in recent clinical trials for cancer therapy. *Clin Epigenetics* **1**, 117-136, doi:10.1007/s13148-010-0012-4 (2010).
- 111 Zhang, J. & Zhong, Q. Histone deacetylase inhibitors and cell death. *Cell Mol Life Sci* **71**, 3885-3901, doi:10.1007/s00018-014-1656-6 (2014).
- 112 Dalvie, D. *et al.* Metabolism distribution and excretion of a matrix metalloproteinase-13 inhibitor, 4-[4-(4-fluorophenoxy)-benzenesulfonylamino]tetrahydropyran-4-carboxylic acid hydroxyamide (CP-544439), in rats and dogs: assessment of the metabolic profile of CP-544439 in plasma and urine of humans. *Drug Metab Dispos* **36**, 1869-1883, doi:10.1124/dmd.108.022566 (2008).
- 113 Kubitza, C. *et al.* T4 lysozyme-facilitated crystallization of the human molybdenum cofactor-dependent enzyme mARC. *Acta Crystallographica Section F* **74**, doi:doi:10.1107/S2053230X18006921 (2018).
- 114 Namgung, B., Kim, J. H., Song, W. S. & Yoon, S. I. Crystal structure of the hydroxylaminopurine resistance protein, YiiM, and its putative molybdenum cofactor-binding catalytic site. *Sci Rep* **8**, 3304, doi:10.1038/s41598-018-21660-y (2018).
- 115 Chamizo-Ampudia, A., Galvan, A., Fernandez, E. & Llamas, A. Study of Different Variants of Mo Enzyme crARC and the Interaction with Its Partners crCytb5-R and crCytb5-1. *Int J Mol Sci* **18**, doi:10.3390/ijms18030670 (2017).
- 116 Schwarz, G. & Mendel, R. R. Molybdenum cofactor biosynthesis and molybdenum enzymes. *Annu Rev Plant Biol* **57**, 623-647, doi:10.1146/annurev.arplant.57.032905.105437 (2006).

7 Appendix

7.1 List of Abbreviations

ADME	absorption, distribution, metabolism, excretion
AO	aldehyde oxidase
ATP	adenosine triphosphate
Cyb5	cytochrome <i>b</i> ₅
Cyb5R	NADH cytochrome <i>b</i> ₅ reductase
CYP	cytochrome P450 monooxygenase
EM	electron microscopy
EPR	electron paramagnetic resonance spectroscopy
FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide
FMO	flavin-dependent monooxygenase
MALS	multi-angle laser light scattering
mARC	mitochondrial amidoxime reducing component
MFO	mixed function oxidase
Moco	molybdenum cofactor
MOSC	Moco sulfurase C-terminal domain
MPT	molybdopterin
NAD ⁺ / NADH	nicotinamide adenine dinucleotide (oxidized / reduced)
NADP ⁺ / NADPH	nicotinamide adenine dinucleotide phosphate (oxidized / reduced)
NMR	nuclear magnetic resonance spectroscopy
NO	nitric oxide
NR	nitrate reductase
PA	pyrrolizidine alkaloid
PDB	protein data bank
PNO	pyrrolizidine alkaloid <i>N</i> -oxygenase
SNP	single nucleotide polymorphism
SO	sulfite oxidase
T4L	T4 lysozyme (endolysin from enterobacteria phage T4)
XDH	xanthine dehydrogenase
XFEL	X-ray free electron laser
XO	xanthine oxidase
XOR	xanthine oxidoreductase

7.2 List of Figures

Figure 1: FMO catalytic cycle.	9
Figure 2: PNO-mediated <i>N</i> -oxygenation of PAs.....	10
Figure 3: Domain arrangement of eukaryotic Moco-dependent enzymes.	11
Figure 4: Hypothetical reaction cycle of amidoxime reduction by mARC.....	15
Figure 5: Molybdenum cofactors of eukaryotic molybdenum-dependent enzymes.	16

Curriculum Vitae

Persönliche Daten

Name: Christian Kubitza
Geburtsdatum: 04.04.1988
Geburtsort: Stendal
Wohnort: 24118 Kiel
Staatsangehörigkeit: deutsch

Schulbildung und wissenschaftlicher Werdegang

08/2000 – 07/2007 Markgraf-Albrecht-Gymnasium, Osterburg
07/2007 **Allgemeine Hochschulreife (Abitur)**
08/2007 – 04/2008 Zivildienst
10/2008 – 09/2011 **Bachelor of Science Biochemie und Molekularbiologie,**
Christian-Albrechts-Universität zu Kiel
10/2011 – 09/2013 **Master of Science Biochemie und Molekularbiologie,**
Christian-Albrechts-Universität zu Kiel
Seit 10/2013 **Wissenschaftlicher Mitarbeiter und Promotion,**
Arbeitsgruppe von Prof. Dr. Axel Scheidig,
Zoologisches Institut, Strukturbiologie,
Christian-Albrechts-Universität zu Kiel

Erklärung

Die vorliegende Arbeit wurde unter Anleitung von Herrn Prof. Dr. Axel Scheidig im Fachbereich Strukturbiologie des Zoologischen Instituts der Christian-Albrechts-Universität zu Kiel im Zeitraum von Oktober 2013 bis Mai 2018 angefertigt.

Hiermit erkläre ich, Christian Kubitza, dass der Inhalt und die Form dieser Abhandlung, abgesehen von der Beratung durch meinen Betreuer, selbständig von mir erarbeitet wurden. Es wurden keine weiteren Quellen als angegeben herangezogen. Die Arbeit hat an keiner Stelle im Rahmen eines Prüfungsverfahrens vorgelegen und ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden. Es handelt sich um meinen ersten Promotionsversuch.

Kiel, Mai 2018

Christian Kubitza

Die Ergebnisse dieser Arbeit wurden bereits in folgenden Beiträgen veröffentlicht:

Kalimuthu, P.; Havemeyer, A.; Clement, B.; Kubitza, C.; Scheidig, A.J.; Bernhardt, P.V. Human mitochondrial amidoxime reducing component (mARC): An electrochemical method for identifying new substrates and inhibitors. *Electrochemistry Communications* **2017**, *84*, 90-93, DOI: 10.1016/j.elecom.2017.10.003

Kubitza, C.; Faust, A.; Gutt, M.; Gäth, L.; Ober, D.; Scheidig, A.J. Crystal structure of pyrrolizidine alkaloid N-oxygenase from the grasshopper *Zonocerus variegatus*. *Acta Crystallographica Section D* **2018**, *74*, 422-432, DOI:10.1107/S2059798318003510

Kubitza, C.; Ginsel, C.; Bittner, F.; Havemeyer, A.; Clement, B.; Scheidig, A.J. T4 Lysozyme-facilitated crystallization of the human molybdenum cofactor-dependent enzyme mARC. *Acta Crystallographica Section F* **2018**, *74*, DOI:10.1107/S2053230X18006921

Weitere Ergebnisse dieser Arbeit wurden bei Fachzeitschriften eingereicht und werden voraussichtlich in diesen oder anderen veröffentlicht werden:

Christian Kubitza, Florian Bittner, Antje Havemeyer, Bernd Clement and Axel J. Scheidig
First crystal structure of human mARC1 reveals its exceptional position among eukaryotic molybdenum enzymes

Manuscript submitted to *PNAS*: May 18^h, 2018

Carsten Ginsel, Birte Plitzko, Danilo Froiep, Diana A. Stofa, Manfred Jung, Christian Kubitza, Axel Scheidig, Antje Havemeyer, Bernd Clement

The Involvement of the Mitochondrial Amidoxime Reducing Component (mARC) in the Reductive Metabolism of Hydroxamic Acids

Manuscript submitted to *Drug Metabolism and Disposition*: May 18^h, 2018

Danksagung

Die vorliegende Arbeit entstand am Zoologischen Institut der Christian-Albrechts-Universität zu Kiel in der Abteilung Strukturbioogie auf Anregung und unter der Leitung von

Herrn Prof. Dr. Axel Scheidig.

Bei ihm als meinen Doktorvater möchte ich mich an erster Stelle für die freundliche Aufnahme in seinen Arbeitskreis, das interessante Promotionsthema und seine ständige Diskussions- und Hilfsbereitschaft danken – nicht zuletzt aber auch für gelegentliche Motivationsschübe, das entgegengebrachte Vertrauen und die Freiheiten, die ich während meiner Promotionszeit genießen konnte.

Ein großer Dank gilt auch Prof. Dr. Bernd Clement für jahrelange Unterstützung, Vertrauen und Zuversicht bei der Bearbeitung des mARC-Projektes sowie für die Integration in seinen Arbeitskreis in Form der monatlichen „mARC-Runde“.

Herrn Prof. Dr. Dietrich Ober danke ich in gleichem Maße für die engagierte Unterstützung, hilfreiche Anregungen und stete Freundlichkeit für die wertvolle, jahrelange Kooperation beim FMO-Projekt.

Ich danke allen, die mich während meiner Promotion bei der Bearbeitung meiner Themen durch eigenständige, verlässliche Laborarbeit unterstützt und zum letztendlichen Gelingen der verschiedenen Projekte beigetragen haben. Dies gilt insbesondere für Luzia Gäth, Kathrin Funk, Miriam Gutt und Carsten Ginsel, aber auch für die tatkräftige Unterstützung durch Brigitte Bittner, Annika Brauer und Brigitte Schemmerling.

Herzlicher Dank gilt natürlich auch jedem einzelnen meiner Kollegen, die mich bereits viele Jahre lang im Arbeitskreis Scheidig auf meinem wissenschaftlichen Werdegang vom Bachelorstudenten bis hin zur Promotion durch Höhen und Tiefen begleitet haben und stets zu einem angenehmen Arbeitsklima beigetragen haben: Christoph, Monika, Susi, Brigitte, den Sonjas, Annette, Uli, Renato, Leslye, Sebastian, Chrissy, Carsten und Felix.

Auch Thomas, Jenny, Patrick, Sophia, Louisa, Inken, Birte, Eva, Helene, Carmen, Merle und Heyka danke ich für die freundliche Aufnahme in den Arbeitskreis Clement bei etlichen gemeinsam verbrachten Gruppensitzungen. Mein besonderer Dank gilt hier Antje Havemeyer, die von Beginn an wertvolle Ansprechpartnerin und Kontaktperson im Pharmazeutischen Institut für mich war und auf deren permanente, freundliche und umgehende Unterstützung ich immer zählen konnte.

Vielen Dank auch an Florian Bittner als äußerst wertvollen Kooperationspartner für die kompetente, unermüdliche und stets gut-gelaunte Hilfestellung bei Fragen zum mARC-Projekt und verwandten Themen sowie seine ständige Diskussionsbereitschaft und wissenschaftlichen Anregungen.

Zu guter Letzt bin ich denjenigen zu größtem Dank verpflichtet, die mich in den letzten Jahren nicht nur durch Studium und Promotion begleitet haben, sondern mein Leben darüber hinaus bereichert haben: meinen Freunden. Leider kann ich euch nicht alle namentlich erwähnen, obwohl ihr es definitiv verdient hättet. Ihr seid es, die mein Leben lebenswert machen, für deren Unterstützung und Wertschätzung ich mich gar nicht dankbar genug zeigen kann und die gemeinsam mit mir diese Welt ergründen und unseren Platz in ihr suchen. Meine höchste Dankbarkeit gilt dabei denen von euch, auf die ich in jeder Lebenslage bauen kann und die für mich über die letzten Jahre zur Familie geworden sind: Felix, Ruven, Lennart und Toffi.