Czech Technical University in Prague
Faculty of Nuclear Sciences and Physical Engineering

Department of Solid State Engineering
Specialization: Physical Engineering

Doctoral Thesis

X-ray structural analysis of flavin- and Cu-dependent oxidoreductases for biotechnological and medical purposes

Author: Ing. Leona Švecová
Supervisor: RNDr. Tereza Skálová, Ph.D.
Supervisor Specialist: Mgr. Tomáš Koval', Ph.D.
Year: 2021

Prohlášení

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act. No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Czech Technical University in Prague has the right to conclude a licence agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

V Praze dne .....

.....
Ing. Leona Švecová

Acknowledgements

I would like to thank my supervisor RNDr. Tereza Skálová, Ph.D. for her guidance, support, and patience and Mgr. Tomáš Koval', Ph.D. for his advice and support. I would like to extend my thanks to Ing. Jan Dohnálek, Ph.D. for all his support, advice and opportunity to accomplish the work on this thesis in his research group at the Institute of Biotechnology of the Czech Academy of Sciences. My thanks also belongs to all members of the Laboratory of Structure and Function of Biomolecules for friendly environment, support, and invaluable advice and discussions on results. Also, I would like to thank all staff of the Centre of Molecular Structure at Biocev, especially RNDr. Jiří Pavlíček, Ph.D., Ing. Jan Stránský, Ph.D., RNDr. Petr Pompach, Ph.D., and Tatsiana Charnavets, Ph.D., then Mgr. David Sedlák, Ph.D. from CZ-OPENSCREEN, all staff of beamlines 14.1 and 14.2 at the BESSY II electron storage ring operated by the Helmholtz-Zentrum Berlin, Dr. Isabel Bento from EMBL Hamburg, and Dr. Sofia Ferreira from Refeyn Ltd. for their assistance with data acquisition. Moreover, I would like to thank Novozymes A/S for the possibility to work on an interesting molecular target.

Finally, my gratitude goes to my family and friends for their constant support and patience.

This work was supported by the institutional support of the Institute of Biotechnology of the Czech Academy of Sciences (RVO: 86652036), by the European Regional Development Fund (CZ.02.1.01/0.0/0.0/15_003/0000447, CZ.02.1.01/0.0/0.0/16_013/0001776, and CZ.1.05/1.1.00/02.0109), by the Ministry of Education, Youth and Sports of the Czech Republic (LM2015043 and LM2018127, support of Biocev-CMS), and by the Grant Agency of the Czech Technical University in Prague (SGS19/189/OHK4/3T/14).

Ing. Leona Švecová

Bibliografický záznam

Autor:	Ing. Leona Švecová České vysoké učení technické v Praze, Fakulta jaderná a fyzikálně inženýrská, Katedra inženýrství pevných látek Biotechnologický ústav AV ČR, v. v. i.

Název práce:	**Rentgenostrukturní analýza flavin- a Cu-dependentních oxidoreduktáz pro biotechnologické a medicínské účely**
Studijní program:	Aplikace přírodních věd
Studijní obor:	Fyzikální inženýrství

Školitel:	RNDr. Tereza Skálová, Ph.D. Biotechnologický ústav AV ČR, v. v. i.
Školitel - specialista:	Mgr. Tomáš Koval', Ph.D. Biotechnologický ústav AV ČR, v. v. i.

Akademický rok:	2021
Počet stránek:	146
Klíčová slova:	nadrodina glukóza-methanol-cholin oxidoreduktáz, krystalografický screening fragmentů, bilirubin oxidáza, Trp–His kovalentní vazba

Bibliographic Entry

Author:	Ing. Leona Švecová Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering, Department of Solid State Engineering Institute of Biotechnology CAS, v. v. i.

Title of Dissertation:	**X-ray structural analysis of flavin- and Cu-dependent oxidoreductases for biotechnological and medical purposes**
Degree Programme:	Application of Natural Sciences
Field of Study:	Physical Engineering

Supervisor:	RNDr. Tereza Skálová, Ph.D. Institute of Biotechnology CAS, v. v. i.
Supervisor Specialist:	Mgr. Tomáš Koval', Ph.D. Institute of Biotechnology CAS, v. v. i.

Academic Year:	2021
Keywords:	glucose-methanol-choline superfamily of oxidoreductases, crystallographic fragment screening, bilirubin oxidase, Trp–His covalent bond

Abstrakt

Disertační práce se věnuje studiu dvou bílkovin – nové FAD-dependentní oxidoreduktáze z houby *Chaetomium thermophilum* (*Ct*FDO) degradující lignocelulózu a bilirubin oxidáze z rostlinného patogenu *Myrothecium verrucaria* (*Mv*BOx).

*Ct*FDO patří do nadrodiny glukóz–methanol–cholin oxidoreduktáz. Předpokládaná enzymatická aktivita *Ct*FDO vůči složkám ligninu nebyla potvrzena a její aktivita nebyla prokázána ani vůči dalším 990 látkám. K určení typu substrátu bylo využito postupů krystalografického screeningu fragmentů. Z krystalové struktury *Ct*FDO plyne, že *Ct*FDO má nezvykle velkou široce otevřenou kapsu s aktivním místem obsahujícím unikátní His–Ser pár, který je pravděpodobně zapojený do enzymatické katalýzy. Šest komplexů *Ct*FDO s ligandy umožnilo identifikaci pěti vazebných míst uvnitř této kapsy, která vážou přednostně aromatické uhlovodíky. Konformační flexibilita interagujících aminokyselin ukazuje jejich schopnost přizpůsobit se daným ligandům a umožnuje vazbu substrátu s molekulární hmotností větší než 500 Da. Velikost kapsy a její preference pro vazbu aromatických uhlovodíků naznačují komplexní polyaromatický charakter domnělého substrátu.

*Mv*BOx obsahuje jedinečný typ kovalentní vazby mezi postranními řetězci tryptofanu a histidinu, který nebyl zatím pozorován u žádného studovaného proteinu. Role Trp396–His398 aduktu ve vazbě substrátu, oxidaci a přenosu elektronů nebyla dosud zkoumána a byla motivací pro další studium tohoto enzymu. V této práci je představena první struktura komplexu *Mv*BOx s iontem ferrikyanidu interagujícím s Trp396, která umožnila identifikaci vazebného místa pro substrát v *Mv*BOx. Strukturně-funkční studie v kombinaci s mutagenezí Trp396 odhalily, že nezbytnost Trp–His aduktu závisí na typu substrátu. Zatímco pro oxidaci ferrokyanidu je adukt postradatelný, pro oxidaci substituovaných fenolů je klíčovým prvkem a při oxidaci bilirubinu významně ovlivňuje rychlost reakce. Výsledky naznačují, že různé substráty používají adukt primárně pro vazbu a v menší míře i pro přenos elektronů.

Abstract

The thesis is devoted to study of two proteins – a novel FAD-dependent oxidoreductase from a lignocellulose-degrading fungus *Chaetomium thermophilum* (*Ct*FDO) and bilirubin oxidase from a plant pathogen *Myrothecium verrucaria* (*Mv*BOx).

*Ct*FDO is a member of the superfamily of glucose-methanol-choline oxidoreductases. Its presumed enzymatic activity against lignin components has not been proven and *Ct*FDO did not show activity toward additional 990 compounds either. Therefore, to characterize the nature of the substrate, the technique of crystallographic fragment screening has been utilized. The *Ct*FDO crystal structure revealed an unusually large wide-open active-site pocket containing a unique His–Ser active-site pair likely involved in the enzyme catalysis. Six *Ct*FDO-ligand complexes enabled identification of five subsites inside the pocket binding predominantly aromatic moieties. The conformational flexibility of interacting amino acids suggests their adaptation for a given ligand and allows binding of substrates with a molecular weight

greater than 500 Da. The size of the pocket and its preference for binding aryl moieties indicate the complex polyaromatic nature of the putative substrate.

*Mv*BOx contains a unique type of covalent bond between tryptophan and histidine side chains never observed in any studied protein. The role of the Trp396–His398 adduct in substrate binding, oxidation, and electron transfer has not been examined yet and it was the motivation for further studies of the enzyme. The first structure of *Mv*BOx in complex with ferricyanide ion interacting with Trp396 allowing identification of the substrate binding site in *Mv*BOx is presented here. The structure–function studies combined with mutagenesis of Trp396 revealed that necessity of the Trp–His adduct depends on the substrate type. While the adduct is expendable for oxidation of ferrocyanide, it is a key element for oxidation of substituted phenols and it considerably affects the rate of bilirubin oxidation. The results indicate that various substrates use the adduct primarily for binding and to a smaller extent for electron transfer.

Contents

Introduction

Industrial or "white" biotechnology is dedicated to industrial production using catalytic properties of microorganisms and enzymes. Compared to the conventional routes, the usage of eco-efficient catalysts brings economically and environmentally attractive benefits in form of the need for milder reaction conditions (pH, temperature, less water and energy) and synthesis of easily degradable products while producing less waste. Enzymes have long been used in various fields including medicine, pharmaceutical industry, in food and textile industry, for biofuel production, chemical synthesis, and others.

White biotechnology and other biotechnology branches constantly seek for novel interesting sources of proteins adapted to extreme environments and novel enzymes able to withstand extreme conditions and having desired function towards target compounds. The search for novel interesting enzymes may ultimately lead to samples with hard-to-identify substrate specificity. There are several approaches to identify substrates for the target macromolecule under study. These include computationally intensive *in silico* methods (virtual screening), *in vitro* testing of the enzymatic activity of the macromolecule towards a large number of chemicals (high-throughput screening), or crystallographic screening of weakly binding low-molecular-weight compounds (fragments) to target macromolecule. The last mentioned method, the crystallographic fragment screening, is a powerful tool for mapping the binding sites in desired protein, identification of chemical groups that specifically interact with the macromolecule, and for analysis of fragment-macromolecule interactions. Although the method finds the most common use in design of new drugs, it can be utilized also for identification of chemical structure of putative substrates and for prediction of enzymatic function of the enzyme target with unknown function. In common practice, all three approaches for substrate identification are often combined.

The novel FAD-dependent oxidoreductase from *Chaetomium thermophilum* (*Ct*FDO), the first of reported proteins in this thesis, belongs to the glucose-methanol-choline oxidoreductase superfamily, a large family of functionally diverse proteins. *Chaetomium thermophilum* is a soil-borne fungus of high temperature tolerance (61 °C), which participates in degradation process of the most abundant highly renewable natural resource on the Earth – the lignocellulosic biomass. The organism is a promising source of new heat-tolerant proteins with a wide use in various biotechnological applications. Therefore it has been the subject of considerable scientific and commercial interest in recent years. Significant thermal stability and involvement in lignocellulose degradation are also expected from *Ct*FDO. This work deals with biophysical

and structural analysis of *Ct*FDO with a focus on mapping the binding sites in this protein using crystallographic fragment screening. The structural studies revealed novel features of the active site, which together with the *in vitro* screening of *Ct*FDO activity indicate likely a different type of substrate than is common for the superfamily.

The search for novel enzymes is, however, not the only way with a potential to find samples with better functions. Protein engineering can lead to improved stability, activity or specificity of the sample. Engineered proteins are commonly used in many industrial branches, and the potential for their utilization and interest in them is growing. For targeted protein engineering, the study of the wild type protein, its structure–dynamics–function relationships, and identification of key structural elements are the basis.

Bilirubin oxidase, from a plan pathogen *Myrothecium verrucaria* (*Mv*BOx), the second protein reported in the thesis, belongs to the family of multicopper oxidases, namely to those having four copper ions as cofactors placed in two distant active sites (T1Cu site and trinuclear copper cluster). *Mv*BOx is known especially for its utilization in clinical medicine for recognition of particular kinds of jaundice and for liver function tests. Due to its catalytic activity in a wide pH range, oxidation of a great variety of organic and inorganic substrates, reduction of molecular oxygen to water without release of reactive oxygen species, or high temperature tolerance (60 °C), it is a subject of interest for various biotechnological and industrial applications. *Mv*BOx is a promising candidate for biocathodes in hybrid biofuel cells and biosensors, it also has potential in the textile industry for decolourisation and detoxification of synthetic dyes.

*Mv*BOx has been previously intensively studied using the biochemical, structural, and computational methods. During the last two decades, X-ray crystal structures of native and mutated *Mv*BOx were solved and published. Recent analysis of high-resolution *Mv*BOx structures revealed a unique type of covalent link between tryptophan and histidine side chains in the close vicinity of the T1Cu site. The mechanism of dioxygen binding in the trinuclear copper cluster, electron transfer from T1Cu site to the cluster, and dioxygen reduction to water is identical for *Mv*BOx and other multicopper oxidases and has been extensively studied and is well understood and described. Nevertheless, the binding mechanism of electron-donating substrate and the pathway of electron transfer from the substrate to the T1Cu site, and participation of the covalent bond in the transfer in *Mv*BOx has not been determined yet. The first complex of *Mv*BOx with a ligand together with mutagenesis coupled with structure–function analysis of the covalent bond and surrounding residues is presented in the thesis. The results enabled identification of the position of the substrate binding site in *Mv*BOx and, depending on a substrate type, determination of the role of the covalent link in substrate binding and electron transfer.

Chapter 1

Theoretical background

1.1 Biophysical characterization of proteins

1.1.1 Methods based on protein's mass, size, and charge

Size exclusion chromatography

Size exclusion chromatography (SEC) is liquid chromatography that separates molecules based on their size using a SEC column, which contains a stationary phase composed of small porous particles. As the sample passes through the column, the porous particles retain the analytes depending on their size and shape. Larger molecules pass the column faster as they interact with the pores and the particles themselves much weaker than smaller ones. Peak elution times can be therefore quantitatively correlated to the hydrodynamic diameters of the analytes. For analysis of protein and peptides, UV absorbance at 260 nm and 280 nm is most commonly measured [1].

Mass spectrometry

Mass spectrometry (MS) is a technique that measures the mass-to-charge (m/z) ratio of ions. A sample is ionized by an ion source so that gas-phase analyte ions appear, which are transferred to the mass analyser and separated according to their m/z ratio. The separated ions are then detected by an ion detector. Detected electric signals are processed into mass spectra representing distribution of individual ions presented in a sample as a function of m/z values. The technique is used, for example, for detection of a protein in a sample, for verification of protein sequence and molecular mass, and for quantitative analysis [2].

Mass photometry

Mass photometry (MP) measures the mass of molecules in solution and quantifies them using an interferometric scattering microscope. MP detects individual molecules by their light scattering resulting from their non-specific binding to the surface of the microscope cover glass. The method is based on accurate

measurement of the change in reflecting index of a glass–water interface caused by interference between scattered (by molecule binding to the interface) and reflected (at glass–water interface) light. The reflectivity change is directly proportional to the molecular mass [3].

Dynamic light scattering

Dynamic light scattering (DLS) determines the size distribution of particles in solution. DLS is based on the Rayleigh scattering of a monochromatic light beam (red light laser) on particles in Brownian motion. Once the beam encounters the moving particle, the scattering intensity of the incoming light fluctuates. Small particles move fast and thus the fluctuations of the scattered signal change quickly, large particles cause slower fluctuations. Based on measurement of time-dependent fluctuations, a diffusion coefficient and subsequently a spherical size distribution can be calculated using an autocorrelation function [4].

Sodium dodecyl sulfate polyacrylamide gel electrophoresis

Sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (SDS-PAGE) is a technique for separation of biomolecular particles according to their molecular weight in discontinuous polyacrylamide gel using externally applied electric field. SDS is an anionic detergent with a negative charge among a wide pH range. Once dissolved in protein solution, its negative charges cause protein denaturation and breakage of protein assemblies. Such a sample is loaded on the polyacrylamide gel, which matrix serves as a molecular sieve retarding the movement of protein molecules. Once the electric field is established, molecules binding SDS migrate with different motion rates along the gel to the anode, resulting in their fractionation according to their molecular weight [5].

Isoelectric focusing

Proteins are ampholytes, i.e. they are composed of amino acids, which behave as weak acids and bases. Isoelectric focusing (IEF) is a technique for separation of molecules by their isoelectric points (pI, pH when the net charge of a molecule is zero). Similarly as SDS-PAGE, IEF uses an external electric field to bring the molecules in motion to the attracting electrode. The mobility of individual ampholytes is affected by the presence of immobilized pH gradient gel (polyacrylamide gel matrix with established pH gradient). Once the ampholytes reach the pH region related to their pI, they stop moving [6].

1.1.2 Methods utilizing UV-VIS absorption of proteins

Proteins absorb light at specific wavelengths. The main elements responsible for the absorption are aromatic amino-acid residues (tryptophan, tyrosine, and phenylalanine), disulfide bonds, and peptide bonds, which absorb light at wavelengths of 180-300 nm [7].

Absorption spectroscopy

Absorption spectroscopy measures absorption A of radiation as a function of its wavelength. As the absorption strength of the protein depends linearly on its concentration c according to the Lambert–Beer law

$$A = \varepsilon \cdot c \cdot l, \tag{1.1}$$

where ε is a molar extinction coefficient and l an optical path length, the quantitative analysis can be performed. Not only protein, but also many cofactors absorb the light. For example, the oxidized form of flavin adenine dinucleotide (FAD) absorbs light at wavelengths around 390 nm and 450 nm [8]. Therefore, oxidized and reduced states of FAD-containing proteins can be distinguished [9].

The absorption spectroscopy can be used to monitor a rate of an enzymatic reaction (the transformation of one (or more) substrate to one (or more) product catalysed by an enzyme), provided that the substrate and/or the product absorbs the light and their absorption spectra differ. The increase in concentration of the product or decrease in concentration of the substrate in time can be measured to study the enzyme kinetics [9].

Circular dichroism spectroscopy

Circular dichroism spectroscopy uses the ability of optically active molecules (e.g. proteins) to absorb left- and right-circularly polarized light unequally, resulting in its elliptical polarisation. The effect called circular dichroism (CD) is measured over a range of wavelengths and its spectrum is often reported in degrees of ellipticity given by

$$\tan\theta = \frac{E_L - E_R}{E_L + E_R}, \tag{1.2}$$

where E_L and E_R are magnitudes of electric field vectors of the left- and right-circularly polarized light. In the far-ultraviolet region (180-260 nm), the CD spectrum can be analysed to give the information about protein fold – its secondary and tertiary structure, usually used for comparison of structural changes between two or more samples. The overall CD spectrum of the protein is created by linear combination of CD spectra of individual types of secondary structure elements weighted by their fractional representation [10].

Nano differential scanning fluorimetry

Nano differential scanning fluorimetry (nanoDSF) measures changes in fluorescence of sensitive amino acids (tryptophan and tyrosine) to determine the protein thermal and chemical stability. These residues absorb light in the near-ultraviolet region and their emission is highly sensitive to their close surroundings. As the temperature increases, the non-covalent bonds break, protein unfolds, and the chromophore molecules become exposed. The thermal unfolding transition temperature (T_M) is measured by an increase in the fluorescence. Dependence of a ratio of the integrated fluorescence at 350 nm and 330 nm on temperature (F350/F330) is recorded. Inflection points of recorded curves are equivalent to T_M [11, 12].

1.2 Kinetics of enzyme reaction

The reaction rates (the ratio of the change in concentration of the reactants per unit of time) of the various enzymatic reactions are measured in different ways. The basic model of enzyme kinetics is the Michaelis-Menten equation expressing the hyperbolic behaviour of the initial rate of reaction v_0 as a function of substrate concentration [S] (Fig. 1.1) :

$$v_0 = \frac{V_{max}[S]}{K_M + [S]}, \tag{1.3}$$

where Michaelis constant K_M is a concentration of given substrate at which the reaction rate reaches the half of maximum velocity V_{max}.

For most cases, the enzyme kinetic measurements fit well to the Michaelis-Menten hyperbolic behaviour. Nevertheless, there are some deviations from the hyperbolic dependence of v_0 on [S] [13, 14]. The enzyme kinetics can be driven, for example, by allosteric sigmoidal behaviour (Fig. 1.1). It is caused either by a transfer of ligand-induced conformational changes through the enzyme altering the active site or as a reflection of substrate binding at one binding site affecting the affinity of other sites for their substrates. The effect can be described by Hill equation:

$$v_0 = \frac{V_{max}[S]^h}{{K_{1/2}}^h + [S]^h}, \tag{1.4}$$

where $K_{1/2}$ is the substrate concentration at the half of the maximal velocity V_{max} and h is a Hill slope characterizing the sigmoidal shape (for the $h = 1$, $K_{1/2}$ is equal to K_M value) [13, 14].

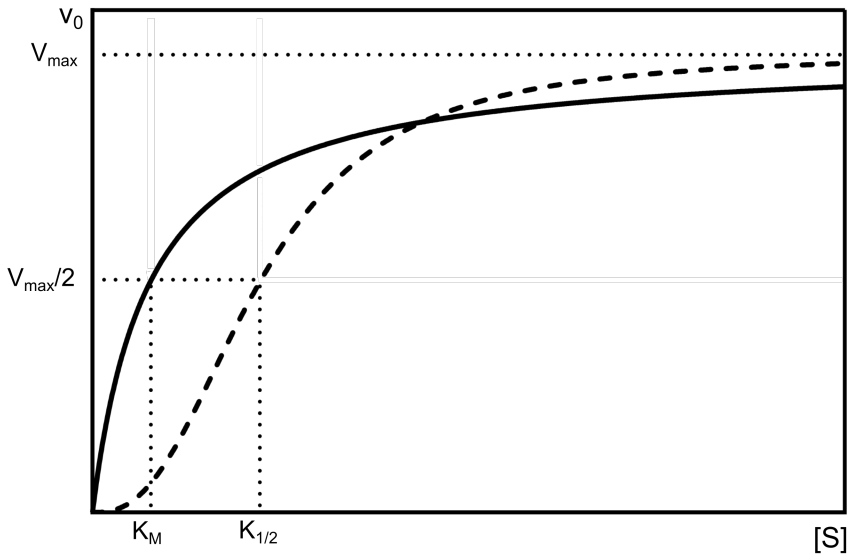

Figure 1.1: Schematic representation of reaction rate v_0 dependence on substrate concentration [S] for enzyme's kinetic behaviour following the Michaelis-Menten equation (solid line) and allosteric sigmoidal (Hill) equation (dashed line).

1.2.1 Activity assays relevant to this study

Colorimetric assays

Spectrophotometric enzyme assays are based on measurement of light absorption by the reaction solution. The increase of product concentration, or the decrease of substrate concentration is calculated by Lambert–Beer law (Eq. 1.1).

In the case that the changes in the reaction are difficult to detect spectrophotometrically, the reaction can be coupled with another observable reaction, i.e. one of the products from the first reaction is utilized in the coupling reaction, which runs immediately after the first reaction and yield coloured product. For example, horseradish peroxidase (HRP) catalyzes the oxidation of chromogenic substrates to coloured products. HRP uses H_2O_2 (product of the first reaction) as the oxidizing agent, therefore, the amount of H_2O_2 can be directly detected.

Luminescence-based assay

Luminescent assays are useful when the reaction product emits photons of light instead of developing a visible colour. An example is the ROS-GLO^{TM} H_2O_2 Assay (Promega, worldwide.promega.com) used for H_2O_2 measurement. It is based on several-steps production of luciferin, initialized with H_2O_2 and terminated by luciferin reaction with luciferase yielding luminescent signal proportional to concentration of H_2O_2 in solution (Fig. 1.2).

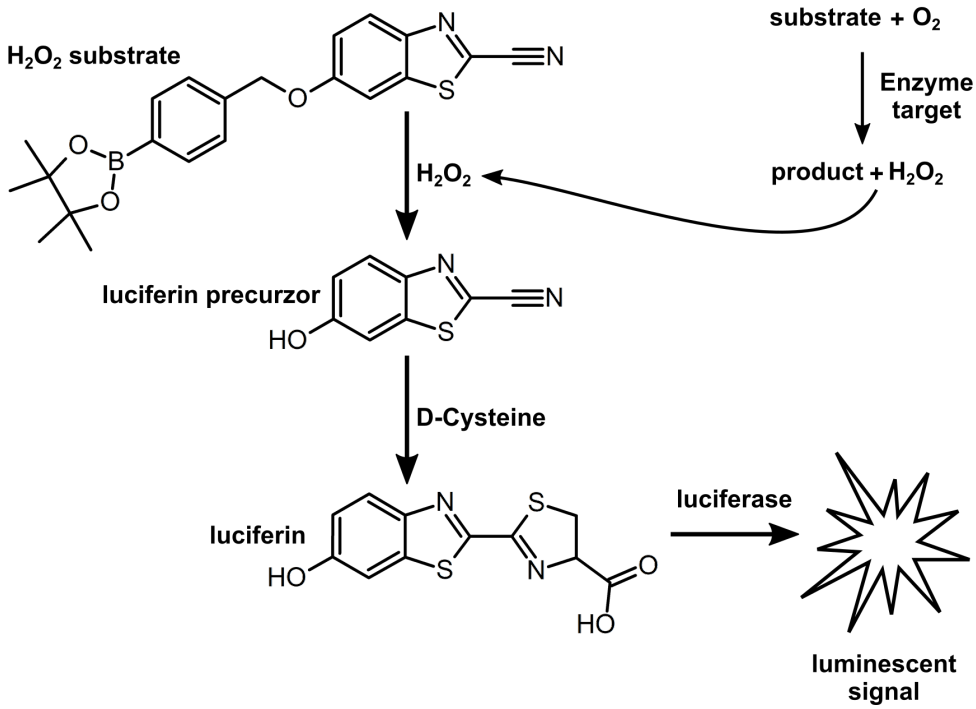

Figure 1.2: ROS-GLO^{TM} Assay mechanism for H_2O_2 measurement. The diagram was prepared according to Promega websites (worldwide.promega.com).

1.3 Protein crystallization

X-ray, electron, and neutron diffraction techniques can be used for structure determination of macromolecules at atomic level. They all require single crystals of the target molecule. Crystallization of proteins usually takes place in an aqueous solution (called mother liquor), composed of solution containing protein and precipitants, and it is based on the creation of a supersaturated state, i.e. the concentration of the protein is in excess over its solubility limit. Once supersaturation state is reached, the equilibrium is then achieved by precipitation of the protein to a solid state (nucleation). In a desired case, the resulting solid state is crystalline. Afterwards, the surrounding protein molecules interact with the nuclei and arrange on their surface (crystal growth).

A crystallization is influenced by many factors, such as the type and the concentration of individual compounds, pH, temperature, pressure, viscosity, etc. The goal of the protein crystallization techniques is to find a suitable settings of these parameters leading to the crystal growth. The crystallization can be generally divided into two phases – initial screening, i.e. identification of chemical, biochemical and physical conditions, which leads to emergence of a crystalline material (3D, 2D crystals, microcrystals), and optimization, i.e. systematic modification of initial parameters to reach crystals of sufficient size and quality.

The supersaturation state can be attained by different approaches. One of them is the vapour diffusion technique. It is based on a transport of aqueous vapour and possibly other volatile substances from the crystallization drop (mixture of protein and reagent solution) to the precipitation solution in the reservoir. The transport is induced by different concentrations of the reagents in the drop and the reservoir. Afterwards, the drop volume decreases and the system in crystallization drop comes in a supersaturated state. The vapour diffusion technique can be realized in two major arrangements – in sitting or hanging drop arrangement (Fig. 1.3). Nowadays, the utilization of crystallization robots capable of setting high amounts of multichambered trays in nanoliter volumes in short time and availability of variable commercially prepared crystallization kits simplify and speed up the protein crystallization screening and subsequent optimization [15, 16].

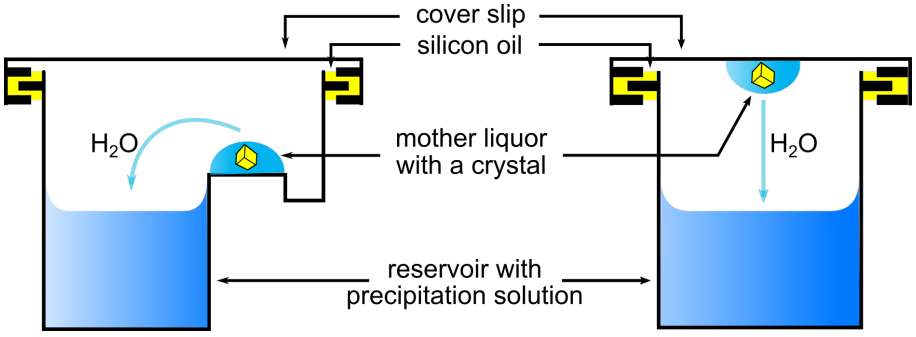

Figure 1.3: Schematic representation of two arrangements of the vapour diffusion method: sitting drop (left) and hanging drop (right).

1.4 Crystals of macromolecules

Compared to crystals of small molecules, not all symmetry operations are allowed in crystals of macromolecules: only left-handed amino acids are present in proteins, therefore the inversion centre and mirror planes are forbidden in protein crystals. Thus, the number of possible space groups is reduced from 230 to 65 (11 point groups) [17]. Up to date, the most frequent space groups of macromolecular crystals are $P2_12_12_1$ (20.1 %), $P2_1$ (14.7 %), and $C2$ (8.8 %) [18]. Protein conformational dynamics and sensitivity of lattice contacts can lead to polymorphism of protein crystals. Different crystal morphologies can be also seen in the same mother liquor sample [15].

Protein crystals are not so compact as crystals of small molecules. Considerable part of protein crystals is created by cavities, channels, and interstitial space containing aqueous solution (26-90 % of the volume) [19]. The high solvent content contributes to fragility of the crystals and the crystals are sensitive to humidity level. Protein crystals are also sensitive to temperature and pressure and can be damaged by radiation during data collection [15].

1.4.1 Preparation of protein–ligand complexes

Crystal structures of protein–ligand complexes contribute to understanding of the protein function and, therefore they are desired. The complexes can be prepared using different strategies. Common methods are co-crystallization, i.e. adding the ligand to the crystallization condition, and soaking ligands into grown protein crystals, which is allowed due to the high solvent content mentioned above [20].

Crystallographic fragment screening

Soaking of various compounds into protein crystals can be used for mapping of protein binding sites and identification of chemical groups interacting with the protein and their combination into a template structure usually for a new inhibitor (drug) development. This method called *crystallographic fragment screening* is based on screening of libraries of weakly binding compounds with low molecular weight (fragments). Most frequently, fragments obey the rule-of-three requiring following physicochemical properties: molecular weight $\leq 300\,\mathrm{Da}$[1], the number of hydrogen-bond donors ≤ 3, the number of hydrogen-bond acceptors ≤ 3, and logarithm of the partition coefficient between octanol and water logP ≤ 3 [21].

1.4.2 Crystal cryoprotection and vitrification

Radiation damage is one of the obstacles for the structural studies of proteins using ionizing radiation. It is primary caused by photoabsorption and inelastic interactions of X-rays with atoms, which lead to releasing of high-energy photoelectrons

[1] $1\,\mathrm{Da} = 1.66053886 \cdot 10^{-27}\,\mathrm{kg}$

and Auger electrons, subsequent cascade of electron-impact ionization, and creation of reactive radicals, which diffuse throughout the crystals. Weak radiation damage can be observed in structural changes such as reduction of metal centres, decarboxylation of acidic residues, or broken disulfide bonds. Stronger radiation damage may fully destroy the crystal. Therefore, protein crystals are usually flash frozen in liquid nitrogen (77 K) to minimize the damage. Cryoprotective agents such as glycerol are added to the protein crystal to prevent crystallization of water in the protein crystal during the crystal vitrification [22–24].

1.5 Principles of X-ray diffraction experiment

Macromolecular X-ray (MX) crystallography has the same principles as X-ray crystallography of small molecules; however, the observation-to-parameter ratio for macromolecules is small especially at low resolution and the structural studies are more computationally demanding due to the high number of atoms in macromolecules. The development of computational technology, crystallographic methods, and instrumentation often primarily motivated by MX crystallography accelerated the rate of structure determination in the last decade, as can be seen in the increasing number of structures recorded in the Protein Data Bank (PDB) [18, 25].

1.5.1 Interaction of X-ray radiation with matter

X-ray crystallography is based on X-ray diffraction on a crystalline material. Considering the wave-particle duality of X-rays, the elastic (Thomson) and inelastic (Compton) scattering, among others, can occur during the electromagnetic interaction of X-ray radiation with matter. While Thomson scattering is important for creation of diffraction maxima, the non-coherent Compton scattering contributes to the background in the diffraction image [26].

For constructive interference of scattered X-rays, following geometrical conditions have to be fulfilled:

$$\begin{aligned} \mathbf{a} \cdot \mathbf{S} &= h \\ \mathbf{b} \cdot \mathbf{S} &= k \\ \mathbf{c} \cdot \mathbf{S} &= l, \end{aligned} \tag{1.5}$$

where $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ are basis vectors of the lattice, h, k, l are the integral numbers, and $\mathbf{S}$ is a scattering vector measuring the change between wave vectors of incident ($\mathbf{k}$) and scattered waves ($\mathbf{k_0}$):

$$\mathbf{S} = \frac{\mathbf{k} - \mathbf{k}_0}{2\pi}. \tag{1.6}$$

The three equations – Laue equations – determine the position of the diffraction maxima, i.e. the peaks in the diffraction pattern [26].

Crystal diffraction can be explained by scattering of radiation on a set of parallel lattice planes in the crystal with Miller indices (h, k, l). In the case of constructive

interference, the relationship between the scattering angle θ and the interplanar distance d_{hkl} for the set of reflecting planes (hkl) can be described by Bragg equation:

$$n\lambda = 2d_{\mathrm{hkl}} \sin\theta, \tag{1.7}$$

where n is an integer expressing the order of the reflection and λ is the wavelength of incident X-rays. The scattering power of an isolated atom is characterized by the atomic scattering factor f, which depends on the scattering amplitude of the atom and on the Bragg angle θ ($f \sim \sin\theta/\lambda$).

Scattering from all atoms in a crystal in direction (hkl) is given by structure factor $F(hkl)$ which characterizes ordering of atoms in the unit cell with volume V:

$$F(hkl) = |F(hkl)|e^{i\varphi(hkl)} = \sum_{j=1}^{N} f_j e^{2\pi i(hx_j+ky_j+lz_j)}, \tag{1.8}$$

where $|F(hkl)|$ and $\varphi(hkl)$ are the amplitude and phase of a wave diffracted from crystal lattice planes (hkl), N is a number of all atoms in the unit cell, and f_j is the scattering factor of the jth atom with positional coordinates x_j, y_j, z_j. The equation (Eq. 1.8) can be alternatively written as Fourier transform of electron density distribution $\rho(x,y,z)$ at position (x,y,z) [27]:

$$F(hkl) = \int_V \rho(x,y,z) e^{2\pi i(hx_j+ky_j+lz_j)} dV. \tag{1.9}$$

The electron density distribution $\rho(x,y,z)$ at position (x,y,z) is then the inverse Fourier transform of the structure factor amplitude $|F(hkl)|$ given by:

$$\rho(xyz) = \frac{1}{V} \sum_{hkl} |F(hkl)|\, e^{-2\pi i(hx+ky+lz)+i\varphi(hkl)}. \tag{1.10}$$

Since the crystal is a periodic system a sum over all (hkl) planes can be used [28].

The measurable quantity in a real diffraction experiment is the intensity $I(hkl)$, which is proportional to the square of the amplitude $|F(hkl)|$ of the complex structure factor $F(hkl)$ according to the relation:

$$I(hkl) \sim FF^* = |F(hkl)|^2 . \tag{1.11}$$

Phase angles $\varphi(hkl)$ are not accessible from the basic diffraction experiment and constitutes the so called *crystallographic phase problem*. Solving the phase problem is one of the crucial steps in structural analysis of macromolecules [28].

Thermal motion of atoms

Atoms in crystal lattice vibrate around their mean positions and the motion decreases their contribution to the structure factor. In MX crystallography, atomic displacements are monitored by atomic displacement parameters (ADPs, thermal factors, B-factors). Small displacements are approximated by harmonic oscillations.

ADPs are related to mean square isotropic displacement $\langle u^2 \rangle$ according to relation [26]:

$$B = 8\pi^2 \langle u^2 \rangle. \tag{1.12}$$

The atomic B-factor can be modelled as isotropic (electron density of the atom is spherical) or anisotropic (ellipsoidal). The B-factors in protein crystal structures typically have values 2-200 Å^2 [25, 26].

1.5.2 X-ray sources

At present, there are three basic types of devices for X-ray generation: laboratory sources, synchrotrons, and X-ray free-electron lasers. In the in-house laboratory sources, the X-ray radiation is generated by bombardment of a metal (e.g. Cu, Mo, Co) anode by accelerated electrons (tens to hundreds of keV) emitted by a heated cathode. During the time, solid-metal anodes were replaced by rotation-anodes or liquid-metal jet anodes, which have even 100 × higher effective brightness than current rotating-anode sources. Although the in-house laboratory sources became more compact and automated, they still have weak intensity compared to synchrotrons and work with a fixed wavelength [29, 30].

Synchrotron sources are indispensable sources of X-ray radiation for MX crystallography, as they generate radiation of high intensity with small divergence, high brightness, polarization properties, and with wavelength-selectable X-rays. The synchrotron radiation is generated by charged relativistic particles (electrons) moving in a circular orbit with bending magnets and insertion devices with periodic magnetic field – undulators and wigglers, where the electrons are accelerated. Once these magnets (Lorentz force) bend the particle trajectory, the particle emits electromagnetic radiation perpendicular to its velocity. Undulators provide partially coherent radiation with the highest brightness and concentrated into a cone with the smallest possible opening angle in both planes. These and other properties of undulators make them very attractive source of radiation for materials and chemical research, including the macromolecular structure studies. Monochromatization of the X-ray beam is performed using monochromator crystals with two-bounce geometry and focusing of the beam using focusing optics such as mirrors, multilayers, or wave plates [31].

X-ray free-electron lasers are the newest type of radiation source used in MX crystallography. Their high intensity and a femtosecond coherent pulse character applied on flux of protein nanocrystals enable time-resolved diffraction studies of protein dynamics. The process adopted to generate the X-ray pulses is based on the self-amplified-spontaneous-emission principle. Bunches of electrons generated in a high-brightness gun are accelerated to high energies in a superconducting linear accelerator. Then the bunches travel through a very long undulator and the energy from the bunch is converted to X-ray pulses. The peak brilliance of the X-rays is about ten orders of magnitude greater than that of any existing synchrotron radiation sources [32].

1.5.3 Detectors

Detectors for MX crystallography have undergone a huge development from X-ray films (until late 1980s), via multiwire proportional counters, imaging plates and detectors based on CCD (charge-coupled device) technology, to an active-pixel complementary metal-oxide semiconductor chips (AP CMOS) and hybrid-pixel photon-counting devices. Detectors of CCD and AP CMOS type and the hybrid pixel detectors dominate at present and are available for both synchrotron and the in-house laboratory systems. CCD and AP CMOS detectors cannot directly record X-rays. The incident X-rays are converted to visible photons via a scintillating screen, which are guided by fibre-optics and then converted to analog signals for readout and digitization. The hybrid-pixel photon-counting devices convert incoming X-rays directly into an electrical signal by a silicon pixel. So they offer advantages such as fast data read-out and low visible noise [29, 33].

1.5.4 Data collection

Diffraction data are collected using a multi-axis goniometer. The protein crystal is mounted on the head of a goniometer and is centred in the X-ray beam. During the diffraction experiment, the protein crystal cooled by a nitrogen gas stream (100 K) is rotated perpendicular to the X-ray beam. Usually, first several initial test images are collected to suggest the data-collection strategy (exposure time, oscillation range, crystal-to-detector distance, etc.). The diffraction data are recorded on a sensitive high-speed X-ray detector. The fast readout time and low noise of photon counting X-ray detectors available on synchrotron beamlines allow the shutterless fine φ-slicing (small rotation angles). It results in more accurate determination of error in the background and improved signal and, as it provides different parts of a partially recorded reflection, it enables assembly of the 3D profile of the reflection [34, 35].

1.6 Diffraction data processing

Processing of X-ray diffraction data results in a list of reflections characterised by Miller indices (h,k,l). Every reflection has its intensity ($I(hkl)$) and standard deviation ($\sigma(hkl)$). Data processing can be divided in four steps: spots finding and indexing, refinement of parameters, integration of reflection intensities, and scaling and merging (data reduction). The spot finding and indexing routine determines the type of crystal lattice and unit-cell parameters and index reflections by Miller indices. Resulting parameters are approximate, therefore, the fit of observed to calculated spot coordinates is further optimized. During the integration step, intensities ($I(hkl)$) from pixels corresponding to each reflection are counted using profile fitting methods [36, 37].

The scaling routine makes the data consistent by minimizing differences between multiple observations of the same (or symmetry equivalent) reflection, i.e. multiple

observations of the same unique reflection are averaged. The result is an intensity value $I(hkl)$ and its error $\sigma(I(hkl))$ for each unique reflection (hkl). The final step of the data reduction is computation of structure factor amplitudes from intensities [36, 38].

There are several statistics introduced to measure the self-consistency of the crystallographic data and to help with a selection of a high resolution cut-off to reach the best model. The indicator of the first rank is Pearson's correlation coefficient ($CC(1/2)$) between the two half-dataset intensities given by equation:

$$CC(1/2) = \frac{\sum_{hkl}(I_A(hkl) - \overline{I_A})(I_B(hkl) - \overline{I_B})}{\sqrt{\sum_{hkl}(I_A(hkl) - \overline{I_A})^2 \sum_{hkl}(I_B(hkl) - \overline{I_B})^2}}, \tag{1.13}$$

where $I_A(hkl)$ and $I_B(hkl)$ are the intensities of single observations of unique reflection (hkl) randomly assigned to half-sets A and B, respectively, and $\overline{I_A}$ and $\overline{I}_B$ are their average intensities in the half-sets [39].

Other useful indicators are merging R-factor R_{merge} (R_{sym}),

$$R_{\text{merge}} = \frac{\sum_{hkl}\sum_{i=1}^{N}|I_i(hkl) - \overline{I(hkl)}|}{\sum_{hkl}\sum_{i=1}^{n} I_i(hkl)}, \tag{1.14}$$

redundancy-independent merging R-factor R_{meas} ($R_{\text{r.i.m}}$)

$$R_{\text{meas}} = \frac{\sum_{hkl}\sqrt{\frac{N}{N-1}}\sum_{i=1}^{N}|I_i(hkl) - \overline{I(hkl)}|}{\sum_{hkl}\sum_{i=1}^{N} I_i(hkl)}, \tag{1.15}$$

precision-indicating R-factor $R_{\text{p.i.m}}$

$$R_{\text{p.i.m}} = \frac{\sum_{hkl}\sqrt{\frac{1}{N-1}}\sum_{i=1}^{N}|I_i(hkl) - \overline{I(hkl)}|}{\sum_{hkl}\sum_{i=1}^{N} I_i(hkl)}, \tag{1.16}$$

and average signal-to-noise ratio I/σ. $I_i(hkl)$ and $\overline{I(hkl)}$ in equations Eqs. 1.14, 1.15, and 1.16 express intensity of individual observation and average intensity of N (redundancy) observations of reflection (hkl), respectively [40]. Mean I/σ is defined as the fraction of measured signal divided by its noise level. It can be expressed as average I/σ of N reflections in resolution shell [26]:

$$I/\sigma(I) = \frac{1}{N}\sum_{hkl}^{N}\frac{|I(hkl)|}{\sigma(I(hkl))} \tag{1.17}$$

1.7 Solving the phase problem

As explained above, only the absolute values (amplitudes) of structure factors are obtained from the standard experiment. The phase of the structure factor is absent. This difficulty is called a phase problem. There are various methods for solving the phase problem: direct mathematical methods, experimental methods (isomorphous replacement methods and single/multiple wavelength anomalous dispersion), and molecular replacement, which was used in this work and is described below in more detail [41].

1.7.1 Molecular replacement

Molecular replacement utilizes the already solved structure of a similar protein. The protein is considered as a rigid body. Selection and preparation of suitable model is probably the most critical step in molecular replacement. The agreement of the model in each spatial position in the unit-cell with experimental structure factor amplitudes is monitored. A good model has a low root-mean-square deviation from unknown structure and high completeness, i.e. it models a high fraction of total scattering from unknown structure with high accuracy. Phases of the correctly placed model are used as initial estimates of phases of the unknown structure. Initial electron density map is then calculated using these structure factor phases and using experimentally observed structure factor amplitudes [26, 41, 42].

Programs for molecular replacement can be based on different search functions. For example, program Phaser uses maximum likelihood-based search functions, MOLREP uses the Patterson search methods [43,44]. MOLREP is also a constituent of MoRDa pipeline, which uses its own internal structure database derived from PDB database and is able of sequential search of different domains or chains, their modification, and creation of oligomeric models [45].

1.8 Model building, Refinement and Validation

1.8.1 Automated building of initial model

When the phase problem is solved, the first model of the protein structure with initial electron density map is obtained. This model should be further refined to reach the best agreement between observed structure factor amplitudes and structure factors computed from the protein structure model. The structure solution is an iterative process. Model building is the process, when parts of the protein not yet included in the model should be built. This can be performed using various software and various procedures. For example, program Resolve identifies fragments of helices and strands by matching templates to the electron density map via procedure based on Fast Fourier transform [46]. Programs for automated model building can be coupled with density modification procedures into pipelines to generate a high-quality model (e.g. PHENIX AutoBuild wizard uses Resolve [47]).

1.8.2 Model building and refinement

The model building process includes interactive local modifications of three-dimensional model ((re)building of residues, correction of waters and ligands positions, etc.) in real space by a graphical program (e.g. by Coot [48]). The aim is the best fit of the model in the crystallographic data given by the electron density map. The model is adjusted using real space geometry regularization tools.

Different types of maps calculated between observed (F_o) and calculated (F_c) struc-

ture factors (observed amplitudes are completed by calculated phases) help with correct interpretation of electron density map. Two basic maps are electron density map $2mF_o$-DF_c, which should show density contours surrounding well-defined atoms in the model, and mF_o-DF_c map, which emphasizes differences between current model and observed data. The maximum-likelihood weights m (figure-of-merit[2] [26]) and D (σ_A-wighting factor) are introduced to reduce the model bias [49, 50]. Other crystallographic maps commonly used are e.g. feature-enhanced map [51] or variable OMIT maps (e.g. Composite omit map [52]). They are used during model building and refinement as a useful tool for displaying weak densities of ligands, alternative conformations, etc.

Interactive manual modifications are improved during global reciprocal space restrained refinement (e.g. in Refmac5 [25], phenix.refine [53]). This part includes refinement of the model atoms parameters (the coordinates x, y, z and individual B-factors) and overall parameters (scale factor, overall B-factor, bulk solvent corrections, anisotropy corrections) against the experimental data. The model is restrained to ideal values of chemical lengths and angles by a set of restraints based on a survey of accurate high-resolution crystal structures of small molecules from the Cambridge Structural Database. The fit between the model and observed data is optimized by minimization of the X-ray target function based either on the least-squares residuals or on the maximum-likelihood formalism [26].

The progress in the structure refinement is monitored using R-factors, which measure the agreement between observed ($|F_{obs}(hkl)|$) and computed model structure factor amplitudes ($|F_{calc}(hkl)|$). To avoid the over-fitting, usually 5 % (testing set) of reflections is excluded from refinement and used for calculation of so-called free R factor (R_{free}). It measures how well the remaining 95 % of reflections (working set) predict the testing set. Both factors, R factor and R_{free} are given by equation:

$$R = \frac{\sum_{hkl} ||F_{obs}(hkl)| - |F_{calc}(hkl)||}{\sum_{hkl} |F_{obs}(hkl)|}. \tag{1.18}$$

Another quality indicators are r.m.s. deviations ($RMSD$) of bond lengths and angles from ideal values given by

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N}(x_j - \overline{x_j})^2}{N-1}}, \tag{1.19}$$

where $\overline{x}$ is the average value of N compared bond lengths (angles) x_j.

1.8.3 Validation

Validation is a rapidly developing area mainly due to low-resolution structures. A good model of protein structure should be consistent from a chemical point

[2]Figure-of-merit (FOM) statistic monitors the correctness of phases and can be expressed by a phase error $\Delta\varphi$. FOM is given by relation $m = \overline{\cos(\Delta\varphi)}$

of view and should be in agreement with all available experimental data. To evaluate the model quality, the indicators can be divided by its global or local character. The global quality indicators inform about the overall quality and geometry of the model (e.g. R and R_{free} factors, average B-factor, r.m.s.d. of bond lengths and angles) and the local indicators inform about individual residues, ligands and their atoms. The local indicators are useful especially during model building and local problems identification, and include, for example, analysis of hydrogen bonds, individual B-factors, rotamers, occupancies, interatomic lengths, torsion angles (via Ramachandran plot [54]), steric clashes (unphysical overlap of any two non-bonding atoms in a protein structure), etc. Several programs, such as MolProbity [55], tools implemented in Coot [48], and others, were developed for validation of models of protein structures. Moreover, some validation servers and programs provide evaluation with respect to previously deposited protein structures.

Chapter 2

Proteins under study

2.1 Organisms relevant to this study

Chaetomium thermophilum

Chaetomium thermophilum (*Ct*) is a thermophilic fungus living in soil, compost heaps, and dung, where it participates in degradation of lignocellulose, the most abundant and renewable biomass resource on the planet[1]. *Ct* has optimal growth temperature between 45 °C and 55 °C (maximum temperature for its growth is 61 °C). During the recent years, *Ct* has become interesting for various scientific studies and also industrial and commercial purposes as a source of new heat-tolerant enzymes, which promise wide use in biotechnological applications [56, 57].

Myrothecium verrucaria

Myrothecium verrucaria (*Mv*) is an ascomycete plant pathogen (cellulolytic fungus) that can be found around the world attacking several plant species such as sunflower, soybean, cucumber, tomato, rice, or corn. It is known for the secretion of high amounts of enzymes, which are able to macerate tissues and degrade cell wall components. *Mv* is a source of chitinases, xylanases, pectinases, extracellular proteases, etc. In industry, it is widely used especially for production of bilirubin oxidase [58, 59].

Aspergillus oryzae

Aspergillus oryzae is a filamentous fungus used especially in East Asian (Japanese and Chinese) fermentation industries (soy sauce, vinegar production). As it is able to secrete large amounts of various degradative enzymes. It has the ability to perform post-translational modifications including glycosylation and so it became one of the most potent expression hosts for proteins of eukaryotic organisms [60, 61].

[1] Lignocellulose is a fundamental constituent and supporting tissue of plant cell wall composed of polysaccharides and aromatic polymer (lignin).

2.2 Oxidases

In living organisms, most of the energy comes from reduction-oxidation (redox) processes, where electrons are transported from an electron donor (first substrate) to an electron acceptor (second substrate). This reaction is catalysed by a class of enzymes called oxidoreductases (EC 1.x.x.x). Oxidases, a subgroup of oxidoreductases, use oxygen as an electron acceptor, and have therefore become valuable enzymes for industrial and biotechnological applications. Molecular oxygen is essential for cellular respiration in all aerobic organisms. The main way they use dioxygen is the four-electron reduction, yielding water molecules. This reaction is their major source of energy [62].

Dioxygen has two unpaired electrons with parallel spin (triplet ground state). Although the overall four-electron reduction of O_2 to H_2O is favourable, the first electron insertion is difficult to perform. This step is thermodynamically allowed but kinetically slow, as it requires spin inversion. Breaking the kinetic spin restriction is possible by strong reducing agents such as activated flavins, hydroquinones, or metals present in enzymes, which are capable to produce free radicals [63–65].

Electron transfer could be realized through a pathway of single-electron redox centres. The efficient electron transfer in proteins falls within the distance of 4-14 Å [66]. Generally, amino acids themselves are miserable mediators of redox reactions, but several amino acids such as tyrosine or tryptophan are redox-active and can form part of the pathway. Nevertheless, in many cases, tightly bound cofactors (flavin adenine dinucleotide (FAD), metal, heme, etc.) play a crucial role in electron transfer. Two main families of oxidoreductases relied on cofactors are the family of flavoproteins and metalloenzymes [67, 68].

2.2.1 Flavoenzymes

Flavoenzymes are structurally diverse proteins with numerous folding architectures. They catalyze a wide range of different types of reactions in many biological processes, such as biodegradation, biosynthesis, bioluminescence, or energy production. In recent years, they have become the subject of interest due to their great potential for use in the pharmaceutical, chemical, and food industries [68]. Flavoenzymes contain covalently or non-covalently bound FAD (Fig. 2.1) or flavin mononucleotide (FMN), which converts between reduced and oxidized state by accepting and donating electrons, respectively (Fig. 2.2). Both states can be distinguished by protein crystallography. The conversion from the oxidized to the reduced state is accompanied by a conformational change of the FAD isoalloxazine ring – a planar conformation in the oxidized state, a bent isoalloxazine ring around an axis passing through the FAD-N^5 and FAD-N^{10} atoms for the reduced state [69] (Fig. 2.2b). The conformational change is associated with a change in the absorption in the UV-VIS region – the oxidized state absorbs wavelengths around 380 and 460 nm yielding the typical yellow colour of the protein (colourless in the reduced state) (Fig. 2.2c) [70].

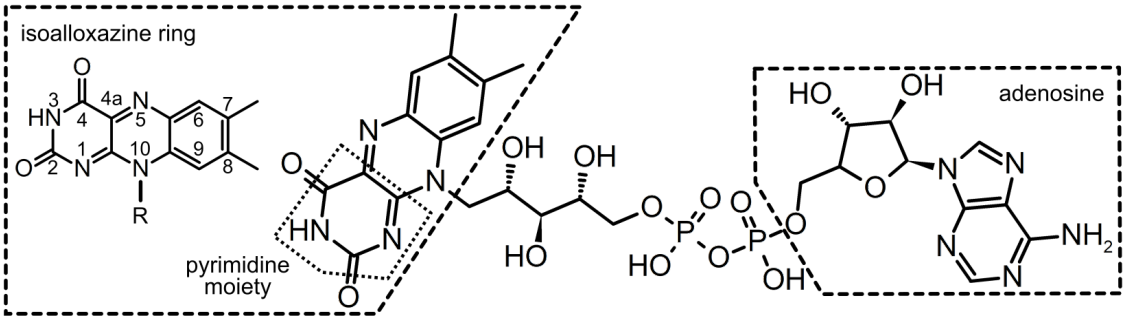

Figure 2.1: Chemical structure of flavin adenine dinucleotide with labelled isoalloxazine ring, pyrimidine moiety, and adenosine moiety.

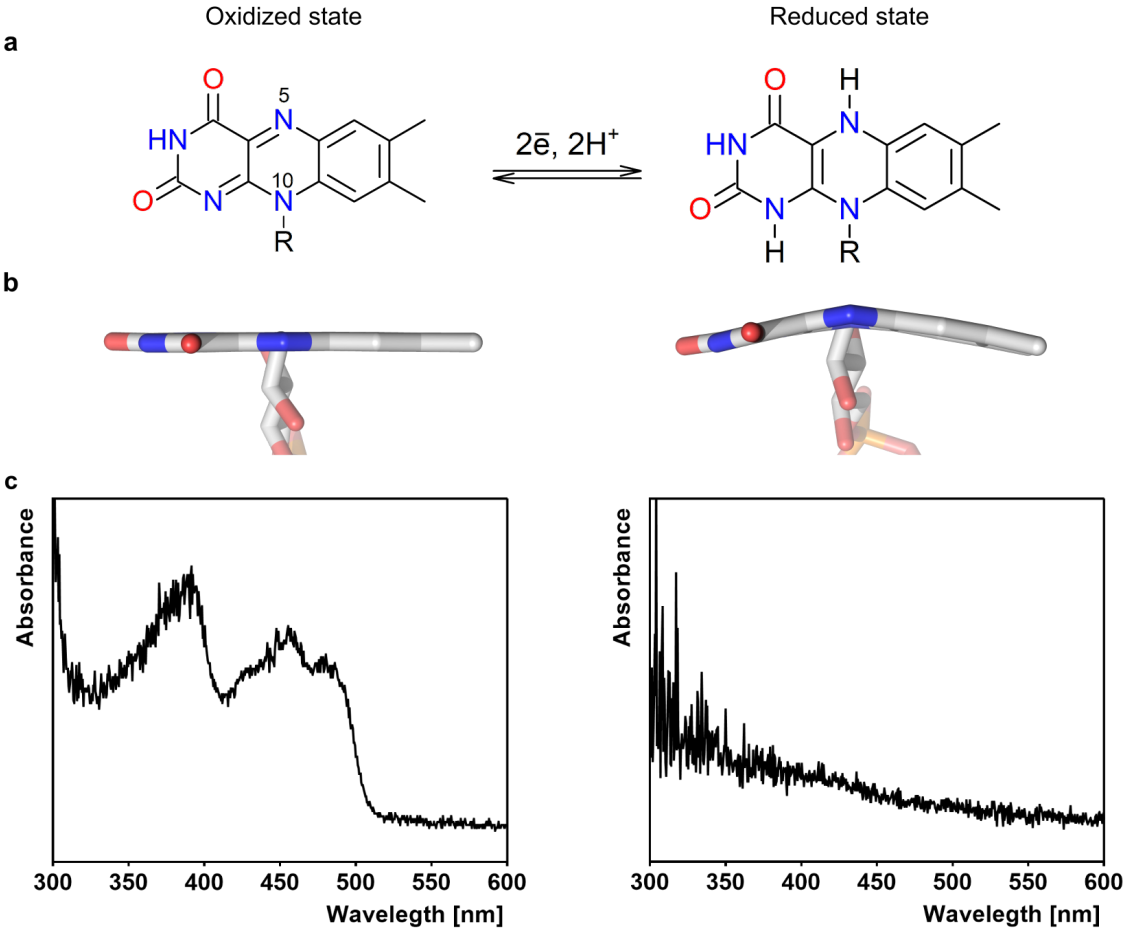

Figure 2.2: Conversion of the isoalloxazine moiety between oxidized (left figures) and reduced state (right figures). (a) Chemical structure depiction for the isoalloxazine ring of the oxidized and reduced form of free FAD (FMN) in solution. (b) A three-dimensional representation of the planar conformation (oxidized state) and the bent conformation around the axis passing through the N^5-N^{10} atoms (reduced form). (c) UV-VIS spectra of FAD-dependent oxidoreductase in a crystal form.

2.2.2 Glucose–methanol–choline superfamily of oxidoreductases

Flavoenzymes belong to several families of flavoproteins. One of them is the glucose–methanol–choline (GMC) oxidoreductase superfamily. Known GMC oxidoreductases act on the hydroxyl group of primary and secondary alcohols, carbohydrates, and sterols. The superfamily includes several well-known oxidases such as glucose oxidase and cholesterol oxidase widely used in medicine to monitor glucose and cholesterol levels in serum [71, 72], choline oxidase used for engineering of biosensors for the detection of choline in serological fluids and food [73], or aryl-alcohol oxidase involved in lignin degradation [74]. In addition to oxidases, the GMC superfamily also includes dehydrogenases, such as wood-degrading cellobiose dehydrogenase [75], or non-redox flavoproteins, such as hydroxynitrile lyase [76]. The majority of oxidoreductases from the GMC superfamily contain dissociable FAD cofactor. The exceptions are choline oxidase, pyranose 2-oxidase, and pyranose dehydrogenase, which have the flavin cofactor covalently bound through a FAD–histidyl covalent linkage [68, 77].

The members of the GMC superfamily are diverse in their function, but they share the core structural elements and the two-domain character. The FAD binding domain contains a conserved N-terminal Gly-X-Gly-X-X-Gly (X represents variable amino-acid residue) sequence motif typical for the initial $\beta\alpha\beta$ segment of the Rossmann fold binding the adenosine diphosphate moiety of FAD. The substrate binding domain contains a conserved C-terminal active-site histidine [78].

Catalytic cycle in GMC oxidoreductases

In flavoenzymes, the site binding the C atom of the substrate undergoing rupture of the C-H bond with transfer of two electrons to FAD is called *the site of oxidative attack*. It is defined as a site in FAD *re*-face at a distance of $\sim$ 3.5-3.8 Å from FAD-N^5 and forming an angle of $\sim$ 96-117 ° with the FAD-N^{10}–FAD-N^5 atoms (Fig. 2.3) [79]. The site in ligand-free structures of GMC oxidoreductases is usually occupied by a water molecule [76, 80–85].

The catalytic cycle in GMC oxidoreductases can be divided into two half-reactions. During the reductive half-reaction, the active-site base deprotonates the hydroxyl group of the substrate and so activates the substrate for two electrons and proton transfer to FAD-N^5, yielding carbonyl group on the substrate, protonated base, and reduced FAD cofactor (anionic hydroquinone).

The oxidative half-reaction involves the re-oxidation of the reduced FAD cofactor by two-electron and proton transfer to a suitable electron acceptor such as oxygen (Fig. 2.4) or quinones. The second proton is provided by the base. In the case of GMC oxidoreductases using dioxygen as the second substrate, after acceptance of the first electron, O_2 is reduced to the superoxide intermediate and then, after acceptance of the second single-electron, to hydrogen peroxide and the oxidized form of FAD. The second step can take place by forming a transient C^{4a}-(hydro)peroxy-flavin or via an outer-sphere second electron transmission (Fig. 2.4) [68, 77, 86, 87].

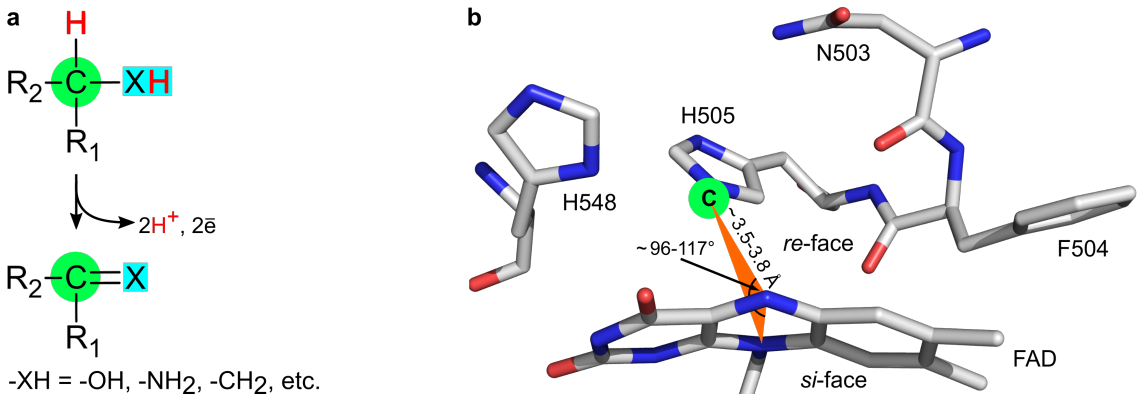

Figure 2.3: The site of oxidative attack in flavoenzymes. (a) General scheme of a substrate dehydrogenation with highlighted C atom (green circle) undergoing the rupture of a C-H bond and activated group (X, cyan rectangle). R_1 and R_2 are unspecific substituents [79]. (b) The site of oxidative attack illustrated on the active site of *Aspergillus flavus* FAD glucose dehydrogenase (PDB id: 4YNT) [88]. The green circle marks the theoretical position of the site. The molecular graphics was created in PyMOL (Schrödinger, LLC).

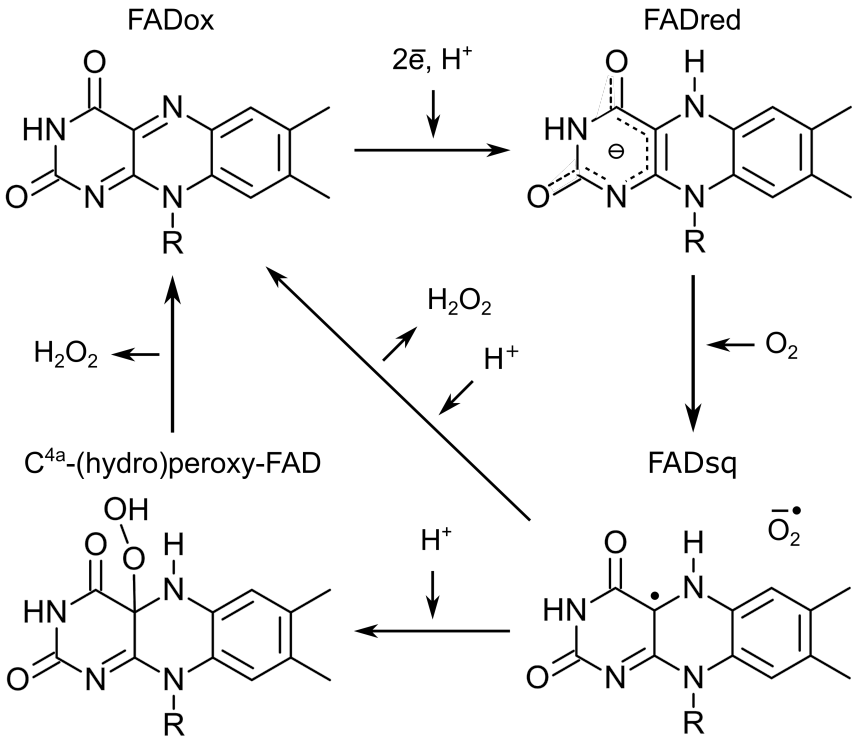

Figure 2.4: Catalytic cycle of FAD cofactor in FAD-dependent oxidases. The oxidized form of FAD (FADox) is reduced by receiving two electrons and proton to the anionic hydroquinone (FADred). After transfer of the first electron to O_2, a superoxide and neutral flavin semiquinone (FADsq) are generated. Then, through a formation of C^{4a}-(hydro)peroxy-FAD or via an outer-sphere second electron transmission, the oxidized state of FAD is restored.

Active-site residues in GMC oxidoreductases

The conserved C-terminal active-site histidine is located ≤ 5 Å from the FAD-N^5 atom in all GMC oxidoreductases. Its conservation indicates important functional roles of this residue [86]. The biochemical, computational, and structural studies of various GMC oxidoreductases have shown that, depending on GMC oxidoreductase, the active-site histidine acts as

1. the catalytic base during the reductive half-reaction in the majority of GMC oxidoreductases [68, 74, 89–94]. The exception is cholesterol oxidase, where the role of catalytic base is arranged by a glutamic acid. The conserved active-site histidine however participates in substrate binding [95]. A different catalytic base (second histidine present in the active site) is suggested also in FAD-dependent hydroxynitrile lyase [76].
2. the catalytic acid during the oxidative half-reaction providing a proton to O_2, and thus promoting the first electron transfer from reduced FAD in *Pleurotus eryngii* aryl-alcohol oxidase [74] and pyranose 2-oxidase [96].
3. a residue stabilizing the C^{4a}-(hydro)peroxy-flavin in pyranose 2-oxidase [89].

The conserved histidine is typically present in active-sites of GMC oxidoreductases in combination with another residue – most often with histidine or asparagine. The residue is located on the *re*-face of the FAD isoalloxazine ring (close to the FAD pyrimidine moiety) and with the conserved histidine it creates a so-called His–His or His–Asn active-site pair [97]. Two different pairs were described so far, namely His–Pro pair in pyridoxine 4-oxidase (PDB id: 3T37) and His–Arg pair in formate oxidase (PDB id: 3Q9T) (Fig. 2.5) [85, 93]. The second residue of the pair is assumed to bind the hydroxyl group of the related substrate in the right position in *Pleurotus eryngii* aryl-alcohol oxidase, pyranose 2-oxidase, pyridoxine 4-oxidase, and others [74, 89, 92, 93, 98].

2.2.3 *Chaetomium thermophilum* FAD-dependent oxidoreductase and the motivation for its study

The biotechnology and white industries are constantly looking for novel enzymes, often new in their functionality and able to withstand extreme conditions. Carbohydrates- and alcohol-oxidizing GMC enzymes from fungal sources attract attention through their application for example in biosensors, in the food industry, or in lignocellulose degradation. The subject of this study, the novel FAD-dependent oxidoreductase from lignocellulose-degrading fungus *Chaetomium thermophilum* (*Ct*FDO) is a great candidate to meet the mentioned expectations. As *Ct*FDO comes from a thermophilic fungus, its high thermal stability was expected. Moreover, its quite unique sequence compared to oxidoreductases from the GMC superfamily with known structure (sequence similarity up to 37 %) brings possibly new interesting structure features, a new residue arrangement in the active site, and possibly new interesting enzymatic characteristics and function.

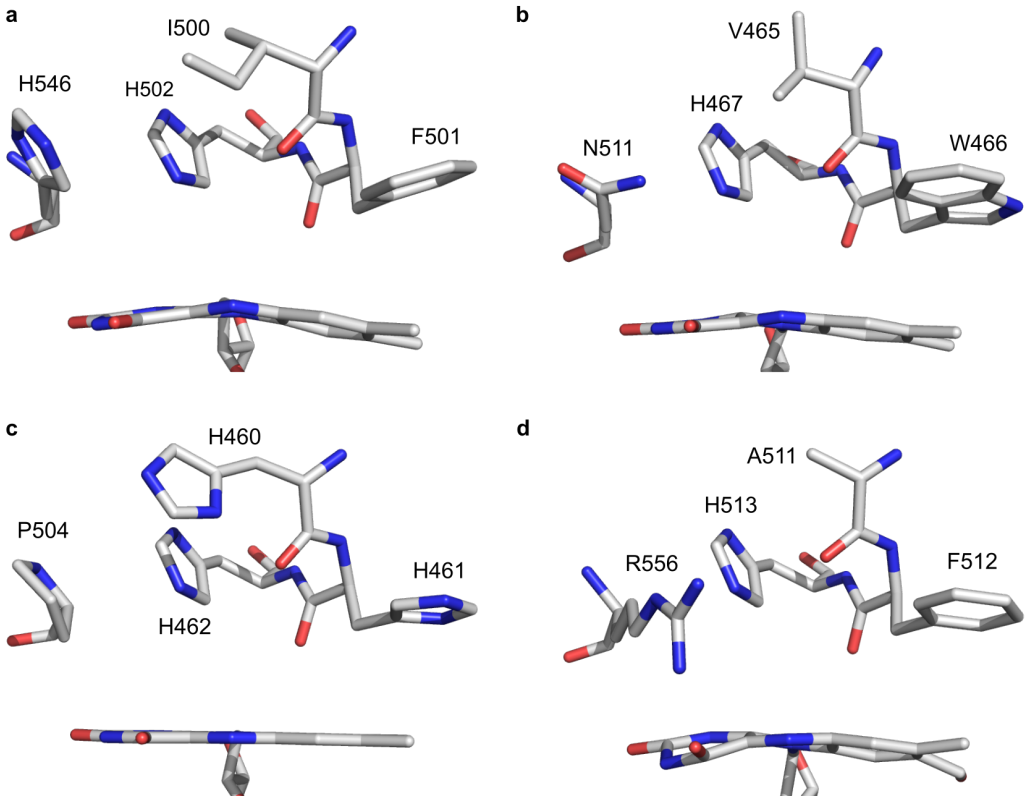

Figure 2.5: Active-site pairs in crystal structures of GMC oxidoreductases. (a) His–His pair in *Pleurotus eryngii* aryl-alcohol oxidase (PDB id: 3FIM), (b) His–Asn pair in 5-hydroxymethylfurfural oxidase from *Methylovorus sp.* (PDB id: 4UDP), (c) His–Arg pair in *Aspergillus oryzae* formate oxidase (PDB id: 3Q9T), and (d) His–Pro pair in *Mesorhizobium loti* pyridoxine 4-oxidase (PDB id: 3T37). The molecular graphics were created using PyMOL (Schrödinger, LLC).

2.2.4 Multicopper oxidase family

Copper is an essential trace element from microorganisms to humans. Due to its ability to accept and surrender electrons, it is a key cofactor in various biological redox reactions [99]. Biological copper centres in proteins have been historically classified into three classes (type I (T1Cu), type II (T2Cu) and type III (T3Cu) copper ions) according to their spectroscopic properties [100].

Multicopper oxidases (MCOs) is a family of enzymes performing oxidation of various substrates with concomitant reduction of O_2 to H_2O without release of reactive oxygen species [101]. They consist of two, three, or six cupredoxin-like domains (Greek key β-barrel) and typically include at least one of each of three types of copper ions, which are organized in a highly conserved motif identical for the majority of enzymes from MCOs. Three- and six-domain MCO enzymes are catalytically active as monomers and the two-domain ones only as homotrimers (e.g. small laccase from *Streptomyces coelicolor* [102]). The family includes ascorbate oxidases, laccases, ceruloplasmin, or bilirubin oxidase [100].

Active sites and catalytic cycle in MCOs

MCOs contain at least four redox-active copper ions distributed in two active sites – the T1Cu active site and the trinuclear copper cluster (TNC) active site (Fig. 2.6a). The substrate is oxidized in the T1Cu site comprising of one T1Cu ion (acceptor of electrons from the electron-donating substrate) coordinated by one cysteine, two histidines, and methionine, leucine, or phenylalanine residue. The orientation of the T1Cu orbital and the cysteine sulphur orbital results in strong covalent bond between them, yielding efficient charge transfer [100]. The charge transfer transition results in an intense absorption band around 600 nm in the UV-VIS spectrum resulting in the typical deep blue colour of the protein [103].

The TNC site is responsible for O_2 reduction to H_2O. It contains one mononuclear T2Cu and binuclear T3Cu coordinated by eight histidines in conserved His–X–His motifs. While T2Cu lacks strong absorption features, T3Cu ions show a charge transfer band around 330 nm in the UV-VIS spectrum coming from bridging OH^- (or dioxygen) between cupric ions. For the dioxygen reduction process, four electrons from the electron-donating substrate and four protons from water molecules are needed. Electrons are transferred from T1Cu through His–Cys–His bridge to the 13 Å distant TNC. The spectroscopic studies on the TNC show that all three copper ions have an open coordination position oriented into the cluster, and therefore, in their reduced state, they provide high affinity for the binding of O_2 [100, 104].

The mechanism of O_2 reduction to H_2O is performed by two two-electron steps (Fig. 2.6b). The process starts with a fully reduced enzyme (Cu^I states), unprotonated histidines coordinating copper ions of TNC, and a water molecule weakly bound to the T2Cu ion. In the first electron-transfer step, the fully reduced enzyme immediately reduces the oxygen molecule by handover of two electrons from copper ions (T2Cu and T3Cu) yielding a peroxide intermediate (PI), which stays bound in the cluster. Due to the second electron-transfer, where electrons from T1Cu and T3Cu are transported, the O-O bond is cleaved and a native intermediate (NI) is generated. All Cu ions are oxidized. In the absence of a substrate, NI slowly decays to the resting oxidized form. On the contrary, in the presence of a substrate, NI is directly reduced to the fully reduced form [105–107].

Bilirubin oxidase from *Myrothecium verrucaria*

Bilirubin oxidase from *Myrothecium verrucaria* (*Mv*BOx, EC 1.3.3.5) is a member of the MCOs. *Mv*BOx was discovered in 1981 and since then, it has been identified in various fungi and bacteria [108, 109]. *Mv*BOx is known especially due to its ability to catalyze the oxidation of bilirubin (the end product of heme metabolism) to biliverdin (green bile pigment). *Mv*BOx is catalytically active in a wide pH range (2.7-8.2 [110]), wherein the pH optimum depends on the nature of particular substrate. Besides bilirubin, *Mv*BOx is able to catalyze the oxidation of both organic (phenolic compounds) and some inorganic substrates (such as potassium ferrocyanide). Although *Mv*BOx was intensively studied, its physiological role in nature remains uncertain.

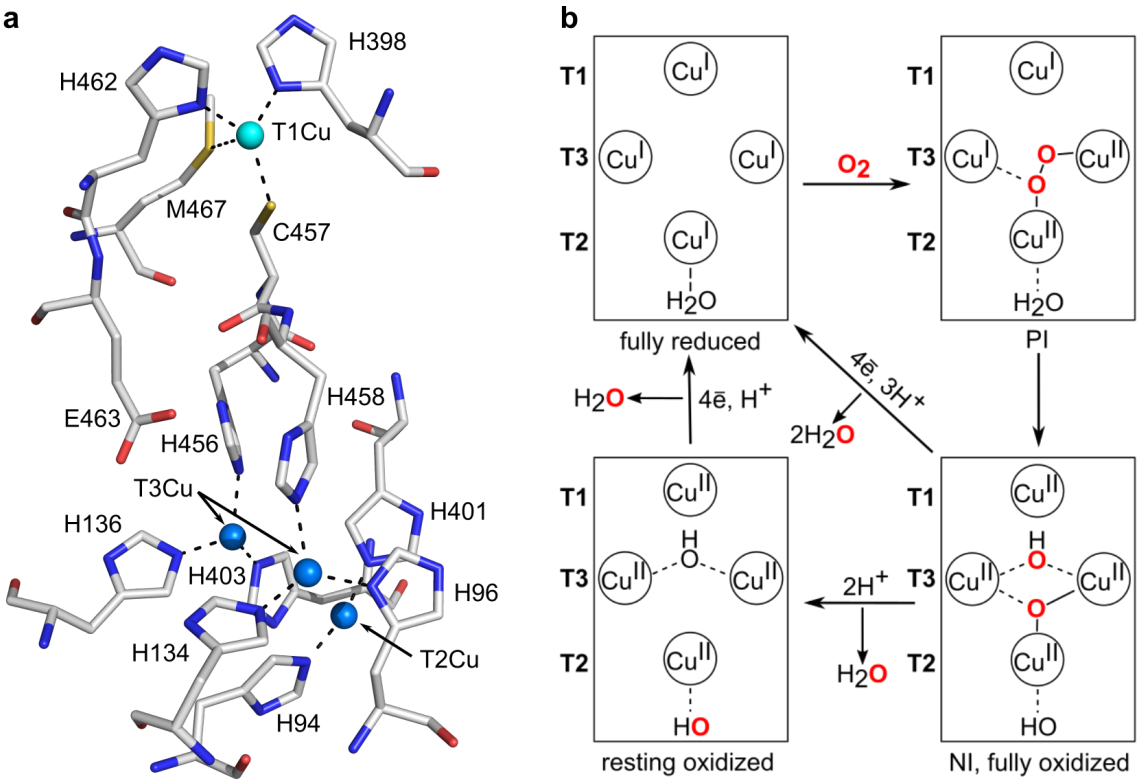

Figure 2.6: T1Cu site, TNC site, and O_2 reduction mechanism in MCOs. (a) The T1Cu and TNC sites illustrated using *Mv*BOx crystal structure (PDB id: 6IQZ). T1Cu ion is shown as a cyan and T2Cu and two T3Cu ions as blue spheres. Molecular graphics was created using PyMOL (Schrödinger, LLC). (b) Mechanism of O_2 reduction created on the basis of known structures of enzymes from the MCO family [105, 106]. T1, T2, and T3 mark copper ions of type I, II, and III, respectively. PI and NI mark peroxide and native intermediate, respectively.

*Mv*BOx is used in medicine for determination of the amount of bilirubin in serum for diagnosis of jaundice and hyperbilirubinemia [103]. *Mv*BOx have become a subject of interest due to its substrate promiscuity, catalytic activity in a wide pH range, thermal stability up to 60 °C, and utilization of molecular oxygen as a second substrate without release of reactive oxygen species. These properties make *Mv*BOx a promising candidate for industrial and biotechnological applications, such as for development of biocathodes in hybrid biofuel cells, biosensors, bioreactors, and biosupercapacitors, for synthetic dye decolourisation, or detoxification of the environment [103, 108, 111, 112].

*Mv*BOx is a glycoprotein (two N-glycosylation sites) composed of 534 amino acids (with molecular mass around 60 kDa) arranged into three cupredoxin-like domains with two active sites containing together four redox-active copper ions. T2Cu and the binuclear T3Cu ions are located at the interface between domains 1 (residues 36-159) and 3 (residues 347-472) [113]. T1Cu ion is placed in domain 3 (Fig. 2.7a) near the protein surface and it is coordinated by cysteine, two histidines, and methionine acting as an axial ligand. Compared to other MCO enzymes, the imidazole moiety of T1-coordinating histidine residues in *Mv*BOx is not accessible for substrate molecules, because an additional layer separates the T1Cu site from the bulk

solvent. The layer contains several amino acid residues including Trp396, which was found to be covalently bound to the imidazole ring of His398 coordinating T1Cu (Fig. 2.7b) [109].

Covalent crosslinks between various side chains of residues (Tyr and Cys, Tyr and Met, Tyr and Trp, Tyr and His, and Cys and His) were observed in several enzymes, including three types of copper-containing oxidases [114]. To the best of our knowledge, the Trp–His covalent crosslink has not been described for other MCOs deposited in PDB and neither for other proteins. The presence of the Trp396–His398 crosslink in *Mv*BOx was first confirmed by high-resolution structure (1.46 Å) of *Mv*BOx wild type (PDB id: 6IQZ, Fig. 2.7b) [109]. The distance between Trp396-$C^{\delta 1}$ and His398-$N^{\varepsilon 2}$ atoms was found to be 1.4 Å. Two other structures of *Mv*BOx wild type were prepared and published so far (PDB id: 2XLL [115] and 3ABG [116]), but the crosslink was not modelled here.

Compared to other MCOs, the T1Cu site in *Mv*BOx has relatively high redox potential. Due to the increase of His398 hydrophobicity by covalent link to the hydrophobic residue (tryptophan), the crosslink raises the redox potential of the T1Cu and thus affects the properties of the T1Cu site. It has been shown that mutation of the axial methionine in the T1Cu site leads to reduction of this potential and consequently the covalent link between Trp396 and His398 is not formed [109].

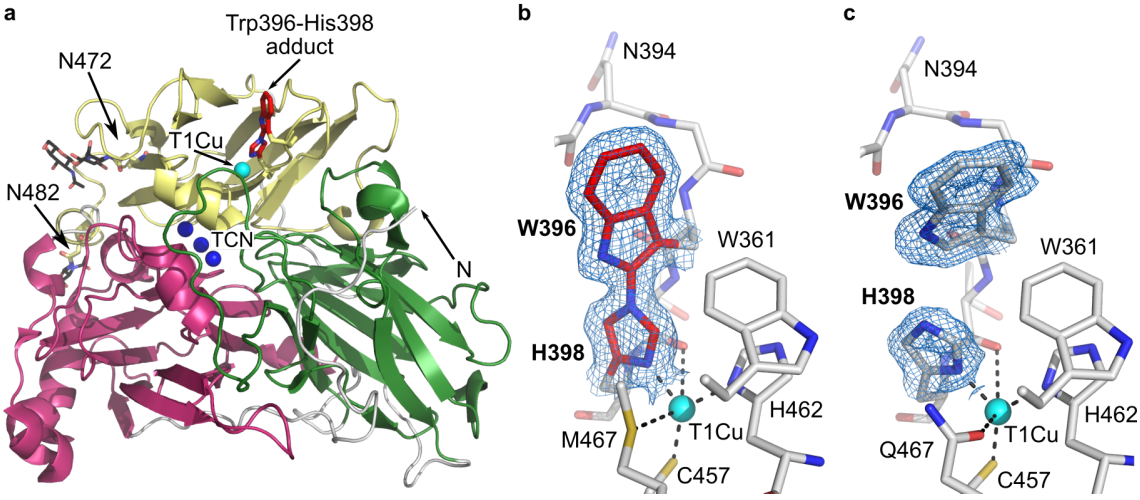

Figure 2.7: High-resolution crystal structure of three-domain bilirubin oxidase. (a) Overall X-ray crystal structure of *Mv*BOx in cartoon representation with colour-distinguished domains (PDB id: 6IQZ) defined with PROMOTIF and HERA program [117,118]. T1Cu (cyan sphere) is placed in domain 3 (yellow), T2Cu and binuclear T3Cu (blue spheres) are placed between domain 1 (warmpink) and domain 3. Domain 2 is coloured in green. The Trp396–His398 adduct is shown as sticks with red C atoms. N marks the N-terminus. N472 and N482 mark glycosylation sites at Asn472 and Asn482. The N-linked carbohydrates are shown as sticks with black C atoms. The view on the T1Cu site in (b) *Mv*BOx wild type (PDB id: 6IQZ [109]) with the Trp396–His398 adduct (red C atoms) and (c) *Mv*BOx-M467Q mutated variant (PDB id: 6IQX [109]), both with the $2mF_o$-DF_c map contoured at 1σ (blue mesh). The T1Cu ion is shown as a cyan sphere. The molecular graphics were created using PyMOL (Schrödinger, LLC).

2.2.5 Motivation for further *Mv*BOx studies

The mechanism of electron transfer from T1Cu to TNC and the dioxygen reduction are very similar for the members of MCO family and were intensively studied using various approaches and described previously [105, 106]. On the other hand, the mechanism of substrate binding and oxidation at the T1Cu site was described only for a few MCOs, as this site varies for each member of the family. The structure of T1Cu site in many MCOs, such as laccases, enables substrate binding in the close vicinity of T1Cu and the oxidized moiety of the substrate has often direct contact to one of the T1Cu-coordinating histidine side chains. The T1Cu site in *Mv*BOx is separated from the bulk solvent by an additional layer of residues including Trp396 linked to His398 by the covalent bond (Fig. 2.8) [109]. Due to the presence of additional residues, the electron pathway from the electron-donating substrate to T1Cu in *Mv*BOx is not straightforward as in other MCOs. This raises the question regarding the electron pathway in *Mv*BOx, the position of the substrate binding site, and the role of Trp396 in substrate binding and electron transfer.

Several X-ray crystal structures of the native and mutated *Mv*BOx were published previously. These are: structure of *Mv*BOx from the basic condition (PDB id: 2XLL [115] and 3ABG [116]) and from the acidic anaerobic condition (PDB id: 6IQZ [109]) and further structure of the Met467Gln mutant (*Mv*BOx-M467Q) prepared anaerobically and aerobically (6IQY and 6IQX [109]) in acidic condition. No ligand binds near the T1Cu site of these structures. We focused on the *Mv*BOx complex preparation, which could answer the questions mentioned above. The motivation for this work was the identification of the substrate binding site in *Mv*BOx by analysis of *Mv*BOx complex with ferricyanide ion, and investigation of the role of the Trp396–His398 crosslink in the reaction mechanism by mutagenesis combined with the structure–function analysis of the additional layer covering the T1Cu site.

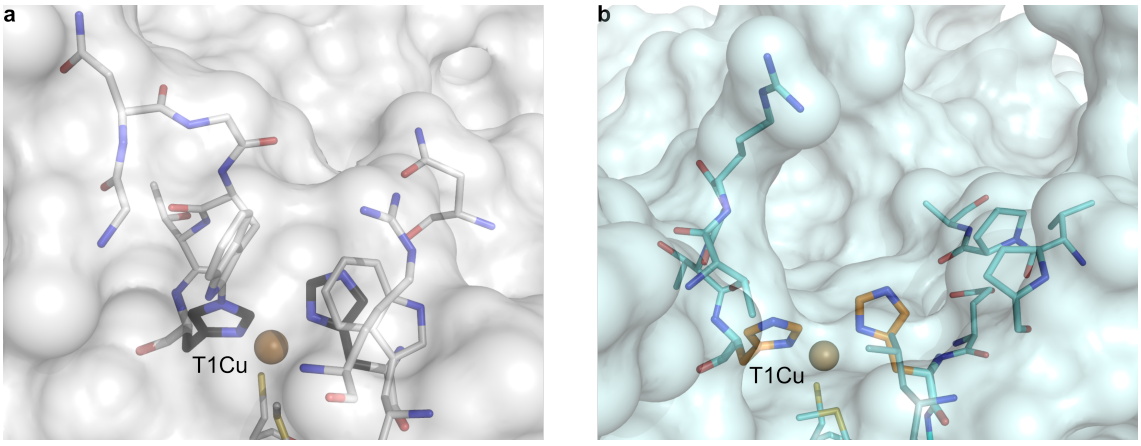

Figure 2.8: T1Cu surroundings in *Mv*BOx and CotA laccase. (a) Structure of T1Cu surroundings in *Mv*BOx (light grey C atoms, PDB id: 6IQZ) with an additional layer of residues preventing direct access of the solvent to the T1Cu coordinating histidine side chains (black C atoms). (b) Structure of T1Cu surroundings in CotA (cyan C atoms, PDB id: 1GSK). T1Cu-coordinating side chains are shown with orange C atoms. T1Cu is shown as a brown sphere. The molecular graphics were created using PyMOL (Schrödinger, LLC).

Chapter 3

Aims of the study

Structure–function study of FAD-dependent oxidoreductase from *Chaetomium thermophilum*

1. Biophysical characterization
2. Catalytic activity testing
3. Determination of crystal structures of *Ct*FDO and its complexes with ligands and their structure–function analysis
4. Description of the nature of putative *Ct*FDO substrate

Structure–function studies of bilirubin oxidase from *Myrothecium verrucaria* and its mutants

1. Structure–function analysis of *Mv*BOx wild type in complex with ferricyanide
2. Mutational study of Trp396
3. Enzymological study of *Mv*BOx mutants with different types of substrates
4. Analysis of the role of Trp396–His398 covalent bond in substrate binding and electron transfer

Chapter 4

Materials and methods

FAD-dependent oxidoreductase from *Chaetomium thermophilum*

4.1 Preparation of *Ct*FDO

4.1.1 Expression and purification of *Ct*FDO

The protein cloning, expression, and purification was performed by Novozymes A/S. The gene encoding the oxidoreductase was cloned using genomic DNA prepared from *Chaetomium thermophilum var. thermophilum* (strain CBS144.50 / DSM 1495). The enzyme was heterologously expressed in *Aspergillus oryzae*. The purification was done using cation-exchange chromatography. The purity of collected fractions was controlled by SDS-PAGE and the absorbance scan. Pure enzyme fractions were pooled and concentrated using Amicon spin filters with 10 kDa cut-off. Production of the samples was not the aim of this thesis. The detailed description of the sample preparation is given in Švecová, *et al.* [119].

4.1.2 Deglycosylation of *Ct*FDO

The deglycosylation of *Ct*FDO was performed according to the assay stated on the Sigma-Aldrich website (www.sigmaaldrich.com) using Endoglycosidase F1[1]. The mixture was incubated at 37 °C for 2 h and at 4 °C overnight. The cleavage was monitored with the SDS-PAGE using NuPAGE 4-12 % Bis-Tris Gel stained with Coomassie Brilliant Blue R-250 (ThermoFisher Scientific). Mark12 Unstained Standard (ThermoFisher Scientific) was used as a marker. The concentration of the deglycosylated *Ct*FDO ($Ct\text{FDO}_{\text{degl}}$) was performed using a Nanosep 10K centrifugal

[1]Endoglycosidase F1 cleaves asparagine-linked oligomannose moieties. The cleavage is done between the two *N*-acetylglucosamine residues in the diacetylchitobiose core of the oligosaccharide (www.sigmaaldrich.com).

device with molecular weight cut-off 10 kDa (Pall Corporation). *Ct*FDO$_{degl}$ was concentrated in the storage buffer (25 mM Tris-HCl pH 7.5 with 100 mM NaCl) to a final concentration of 8 mg·ml^{-1}.

4.1.3 *Ct*FDO sequence analysis

The *Ct*FDO sequence analysis, the secondary structure matching, and prediction of protein function based on the *Ct*FDO$_{degl}$ crystal structure were performed using the following software tools and servers: BLAST [120], PDBeFold [121], ESPript 3.0 [122], and ProFunc [123].

Liquid chromatography tandem mass spectrometry (LC-MS/MS)

300 pmol of *Ct*FDO$_{degl}$ sample diluted in 0.5 M glycine buffer pH 2.3 was injected on a pepsin column and let digested for 3 min. The chromatographic separation of peptides was done by a 1290 series UHPLC system (Agilent Technologies) on reverse phase C18 column. The column was connected to a 15T solariX XR FT-ICR mass spectrometer (Bruker Daltonics) equipped with an electrospray ion source. The data were measured with the spectrometer operating in positive, data dependent mode and processed using the DataAnalysis 4.2 software. Peptides were identified with ProteinScape (Bruker Daltonics) with the Mascot search engine.

Peptide mass fingerprinting using MALDI-TOF

For the sample analysis, the *Ct*FDO$_{degl}$ sample (4μl) loaded onto SDS-PAGE gel lane was used. Gel band was manually cut-out from the gel, chopped into pieces, mixed with acetonitrile, and let dehydrate. The gel pieces were mixed with 50 mM dithiothreitol and let incubate at 60 °C for 30 min. Then the 0.1 M iodoacetamide was added to the pieces and let incubate in dark at room temperature. After 30 min, the pieces were washed by water and repeatedly dehydrated by acetonitrile. The digestion was performed by trypsin at 37 °C overnight. 1μl of tryptic peptide mixture was applied on the stainless steel MALDI target, covered with α-cyano-4-hydroxycinnamic acid as a matrix. The data were measured by 15T solariX XR FT-ICR mass spectrometer (Bruker Daltonics) operating in positive mode. Data were processed with the DataAnalysis 4.2 software (Bruker Daltonics) and mMass software.

4.2 Biophysical characterization of *Ct*FDO

4.2.1 Temperature stability determination

Nano differential scanning fluorimetry (nanoDSF)

The nanoDSF measurements were performed with Prometheus NT.48 (NanoTemper) using the PR.ThermControl software (v2.1.1). The thermal stability of *Ct*FDO

and *Ct*FDO$_{degl}$ was measured in the storage buffer (25 mM Tris-HCl pH 7.5 with 100 mM NaCl) and following buffers: 50 mM citric acid pH 3.0; 50 mM Na acetate pH 4.0; 50 mM Bis-Tris pH 5.5; 50 mM Mes pH 6.0; 50 mM Tris-HCl pH 7.0 and 8.5; 50 mM Hepes pH 7.5; 50 mM Bicine pH 8.0; and 50 mM Glycine pH 9.0; all prepared with 50 mM NaCl. The samples were diluted in buffers to the final concentration of approximately 0.7 mg·ml^{-1}. The measurement ran in the temperature range of 20-95 °C with a temperature slope of 2.5 °C per min and with an excitation power of 45 %. The stability was measured as a dependence of the ratio of integrated fluorescence at 350 and 330 nm and its first derivative on temperature.

Circular dichroism (CD)

The CD signal was measured at 222 nm in temperature range of 20-92 °C with a slope of 1 °C per 12 s using 1 cm quartz cell and a Chirascan^{TM}–plus spectrometer (Applied Photophysics Ltd.). The *Ct*FDO and *Ct*FDO$_{degl}$ samples of concentration approximately 0.2 mg·ml^{-1} were diluted in 50 mM Tris-HCl pH 7.0 with 50 mM NaCl. The resulting spectra were buffer-subtracted.

4.2.2 Determination of size and state of *Ct*FDO in solution

Size-exclusion chromatography (SEC)

SEC was performed using ÄKTA purifier chromatographic system (GE Healthcare). 400 μl of concentrated (8 mg·ml^{-1}) *Ct*FDO was injected onto a Superdex 200 10/300 GL gel filtration column (GE Healthcare) pre-equilibrated with buffer 50 mM Tris-HCl pH 7.5 containing 150 mM NaCl. Elution was monitored by measuring absorbance at 280 nm, 254 nm, and 561 nm. The column was calibrated with bovine thyroglobulin (669 kDa), apoferritin (443 kDa), alcohol dehydrogenase (150 kDa), chicken albumin (44.3 kDa), and Gel Filtration Markers Kit (Sigma Aldrich) containing blue dextran (2000 kDa), bovine albumin (66 kDa), carbonic anhydrase (29 kDa), cytochrome C (12.4 kDa), and aprotinin (6.5 kDa) .

MALDI-TOF mass spectrometry (MALDI-TOF MS)

*Ct*FDO and *Ct*FDO$_{degl}$ at a concentration of 10 nmol·ml^{-1} (1 μl) were deposited onto the MALDI stainless-steel target and after complete evaporation were overlaid with 1 μl of matrix fluid (sinapinic acid (Sigma-Aldrich) and left to air dry at room temperature. MALDI-TOF MS analysis was performed using a MALDI-TOF mass spectrometer Autoflex Speed (Bruker Daltonics) operated in linear positive mode. Data were processed using the FlexAnalysis 3.3 software (Bruker Daltonics).

Mass photometry (MP)

MP experiment was performed by Sofia Ferreira, PhD. from Refeyn company (Refeyn Ltd.). The data were collected on a Refeyn OneMP instrument using the DiscoverMP (v2.2.1) software (Refeyn Ltd.). The *Ct*FDO and *Ct*FDO$_{degl}$ samples were

diluted in 25 mM Tris-HCl pH 7.5 with 75 mM NaCl to a final concentration of 0.0017 mg·ml^{-1} (20 nM) and 0.0016 mg·ml^{-1} (24 nM), respectively. The samples were applied on clean coverslips (High Precision coverslips, No. 1.5, 24·50 mm, Marienfeld) mounted with silicone gaskets (CultureWell™ Reusable Gaskets, Grace biolabs). Mass profiles of the samples were acquired and analysed using DiscoverMP (v2.3.dev12) with default settings.

Multi-angle dynamic light scattering (MADLS)

The MADLS data were collected using Zetasizer Ultra (Malvern Panalytical) with a helium–neon laser (wavelength 632.8 nm). The 45-μl quartz cuvette was filled by 32 μl of *Ct*FDO or *Ct*FDO$_{degl}$ diluted in the storage buffer to a final concentration of 3 mg·ml^{-1}. The measurement was done at 25 °C. The data were analysed and the hydrodynamic diameter was determined with the ZS Xplorer software (Malvern Panalytical).

4.2.3 Other measurements

UV-VIS spectrophotometry of *Ct*FDO in solution

The UV-VIS absorption spectra of *Ct*FDO and *Ct*FDO$_{degl}$ at a concentration of approximately 7.2 mg·ml^{-1} were measured at room temperature using a DeNovix DS-11 microvolume spectrophotometer and 1.2 µl of the samples. The spectrum of the buffer (25 mM Tris-HCl pH 7.5 with 100 mM NaCl) was subtracted.

Isoelectric focusing (IEF)

The isoelectric points of *Ct*FDO and *Ct*FDO$_{degl}$ were determined using IEF. *Ct*FDO (4 μg) was loaded on a 5 % polyacrylamide gel with fixed vertical pH gradient (pH 3-10) (Novex pH 3-10 IEF, ThermoFisher Scientific). The marker was IEF Marker 3-10 (ThermoFisher Scientific). The running voltage was set on 100 V for 1 h, 200 V for 1 h, and then on 500 V for 30 min. The gel was fixed in 12 % trichloracetic acid for 30 min. InstantBlue (Expedeon) was used to stain the gel.

4.3 Structure determination of *Ct*FDO and its complexes with ligands

4.3.1 Crystallization

Screening and optimization of crystallization condition

The initial screening was performed with fully glycosylated *Ct*FDO (8 mg·ml^{-1} in the storage buffer) and with the commercially available crystallization kits Index and Crystal Screen 1 and 2 (Hampton Research). Sitting drop vapour diffusion technique was used with 0.6 μl drops composed of protein and reservoir solution

in a ratio of 1:1. Positive hits appeared with conditions 40 and 91 of Index screen and condition 30 of Crystal Screen 1. However, the optimization of the preliminary hits did not lead to crystals, probably due to a larger number (seven) of glycosylated asparagine residues in *Ct*FDO, as predicted by the NetNGlyc 1.0 Server [124].

The crystallization condition screening was repeated for $Ct\text{FDO}_{\text{degl}}$ (8 mg·ml^{-1}, storage buffer) with the same set-up and using kits Crystal Screen 2 and PEGRx 2 (Hampton Research). After one week, two positive hits were identified with conditions PEGRx 2 number 30 (PEGRx2-30: 0.2 M magnesium formate, 18 % (w/v) polyethylene glycol (PEG) 5000, and 0.1 M sodium acetate, pH 4), and number 37 (PEGRx2-37: 0.2 M sodium chloride, 22 % (w/v) PEG 8000, and 0.1 M sodium acetate, pH 4). These conditions have been further optimized by changing pH and concentration of individual compounds.

The optimization of PEGRx2-37 was unsuccessful. PEGRx2-30 was successfully optimized to the final condition containing 160 mM magnesium formate, 17 % (w/v) PEG 5000, and 0.1 M sodium acetate, pH 5.5. The optimization was performed using hanging drop vapour-diffusion set-up with 1 μl drops with protein to reservoir solution ratio of 1:1. Single crystals usually rectangular in shape grew at 20 °C in 7-14 days (Fig. 4.1). Occasionally, a crystal polymorphism could be seen (Fig. 4.1c). The crystal used for the determination of the ligand-free *Ct*FDO structure (*Ct*FDO:free) was prepared using the reservoir solution enriched by 20 mM cystamine.

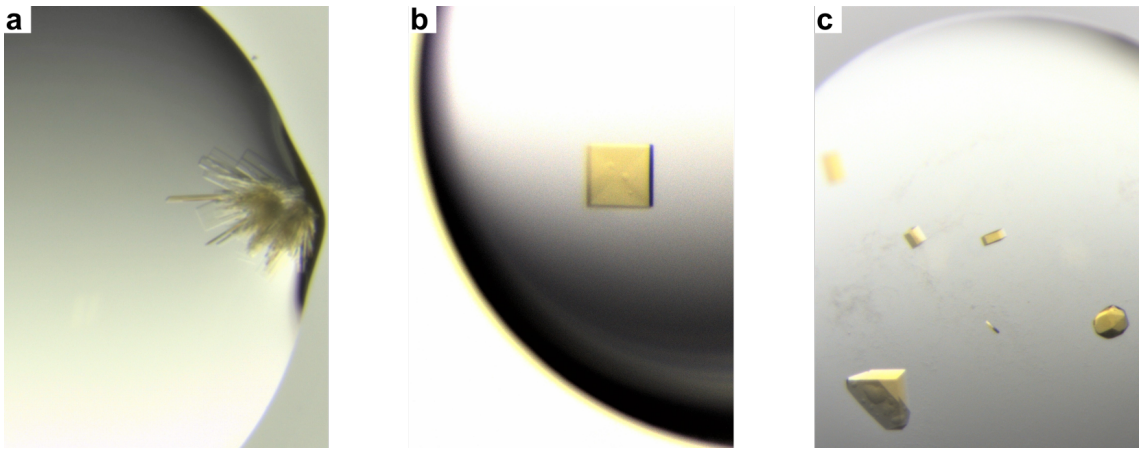

Figure 4.1: Crystallization drops with $Ct\text{FDO}_{\text{degl}}$ crystalline material. (a) $Ct\text{FDO}_{\text{degl}}$ needle cluster growing in original PEGRx2-30 condition, (b) $Ct\text{FDO}_{\text{degl}}$ crystal rectangular in shape growing in final optimized condition, (c) $Ct\text{FDO}_{\text{degl}}$ crystal polymorphism.

Preparation of complexes with ligands

For the preparation of $Ct\text{FDO}_{\text{degl}}$ complexes with ligands, variable organic and inorganic compounds, including those selected to match the expected stoichiometry of the putative lignin degradation substrate/product, were used as ligands. Complexes were prepared using both co-crystallization and soaking methods. The soaking of $Ct\text{FDO}_{\text{degl}}$ crystals was performed using 42 fragments from Frag Xtal Screen (Jena Bioscience GmbH) [125] and other ligands representing variable functional groups

(Table A.1). The fragment screening was performed following the protocol on Jena Bioscience websites (www.jenabioscience.com). Crystals were soaked in the solution created by reservoir solution (0.5μl) and 50 nmol of a fragment for 3 s-70 h (based on crystals stability) at temperature of 20 °C.

Crystals of complexes with fragments methyl4-(aminomethyl)benzoate (MAMB, Fig. A.1a) and 4-oxo-*N*-[1-(3pyridinyl)-ethyl]-2-thiophenebutanamide (PESB, Fig. A.1b) were soaked for 2.5 h and 3 h, respectively. Crystals prepared for soaking in reservoir solution with 2-(1*H*-indol-3-yl)-*N*-[(1-methyl-1*H*-pyrrol-2-yl)-methyl]-ethan-amine (IPEA, Fig. A.1e) and 4-nitrophenol (4NC, Fig. A.1c) grew in crystallization condition containing additional 20 mM $MgCl_2$. Crystals were soaked for 70 h and 20 h in solution containing IPEA and 4NC, respectively.

For co-crystallization, ligands were added to reservoir solution to final concentration of 0.15-60 mM (Table A.2). Complex of *Ct*FDO$_{degl}$ with 4-nitrophenol (4NP, Fig. A.1d) was obtained from the crystallization condition with 20 mM 4NP. Complex with 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS, Fig. A.1f) was obtained from the condition with addition of 8.5 mM ABTS and 20 mM $MgCl_2$.

Data collection, processing, model building, and structure refinement

Crystal of ligand-free *Ct*FDO$_{degl}$ was cryoprotected by soaking in a reservoir solution containing 5 % glycerol and 5 % (v/v) PEG 200 for 22 min. Crystals of complexes were cryoprotected with perfluoropolyether cryo oil (Hampton Research). All crystals were harvested using LithoLoops (Molecular Dimensions) or nylon CryoLoops (Hampton Research) and vitrified and stored in liquid nitrogen.

The diffraction data sets were collected either on Petra III beamline P13 (DESY, Hamburg, Germany) [126], equipped with a Pilatus 6M detector (Dectris), or on Bessy II beamlines 14.1 and 14.2 (Helmholtz Zentrum, Berlin, Germany) [127] equipped with a Pilatus 6M and Pilatus 2M detector (Dectris), respectively (Table 4.1). All diffraction data were recorded at 100 K.

The data were processed and scaled with XDSgui [128] and merged with Aimless [38]. For *Ct*FDO:free, the phase problem was solved by molecular replacement with MoRDa [45] using structure of *Aspergillus flavus* FAD glucose dehydrogenase (*Af*GDH, PDB id: 4YNT [88], seq. identity of 30 %) as a search model. The R_{free} factor of 0.453 (R_{work} = 0.432) was suggested as indicative of a solution. Several parts of the initial model were removed and manually rebuilt in Coot [48] and using the PHENIX Autobuild wizard [47]. The structures of *Ct*FDO$_{degl}$ complexes with MAMB (*Ct*FDO:MAMB) and PESB (*Ct*FDO:PESB) were solved by molecular replacement using *Ct*FDO:free structure and Molrep [44]. Complexes with 4NC (*Ct*FDO:4NC), 4NP (*Ct*FDO:4NP), and ABTS (*Ct*FDO:ABTS) were solved using Phaser [43]. The structure of *Ct*FDO$_{degl}$ complex with IPEA (*Ct*FDO:IPEA) having the same unit-cell and space group as *Ct*FDO:free was directly refined using rigid-body refinement in Refmac5 [25] with *Ct*FDO:free structure as the input model.

Finalization of structures and addition of the cofactor and ligands were manually performed in Coot. All structures, except for *Ct*FDO:IPEA, were refined using re-

strained refinement in Refmac5 (version 5. 8. 0258) [25]. *Ct*FDO:IPEA was refined using phenix.refine (version 1.18.1-3865) [53]. All *Ct*FDO_{degl} structures were refined with the FAD in its reduced form ($FADH_2$). The geometrical restraints for $FADH_2$ and ligands MAMB, PESB, IPEA, 4NC, 4NP, and ABTS were generated using AceDRG [129] and manually checked and edited. External restraints were applied on water molecules in necessary cases. Structures of *Ct*FDO:free, *Ct*FDO:4NC, and *Ct*FDO:ABTS were refined with anisotropic temperature factors[2]. R_{free} calculation was used as the cross validation method. All refinements were done with hydrogen atoms in riding positions. For all structures, the last refinement cycle in Refmac5 and phenix.refine was ran with all reflections. Validation of structures was performed using the validation tools implemented in Coot, Molprobity [55], and wwPDB Validation service [18]. Carbohydrates were checked using the pdb-care program [130]. All structures were deposited in the PDB. The summary of structure determination statistics is presented in Table 4.2.

Preparation of *Ct*FDO crystals with reduced and oxidized state of FAD

For FAD in the oxidized state, crystals grown in crystallization condition mentioned above (160 mM magnesium formate, 17 % (w/v) PEG 5000, and 0.1 M sodium acetate, pH 5.5) were soaked in drops composed of reservoir solution and 20 % glycerol and 20 % (v/v) PEG 200 for 1-3 min. For FAD in the reduced state, crystals grown in the same crystallization condition were soaked in the reservoir solution enriched with 20-150 mM sodium dithionite (DTN), 20 % glycerol, and 20 % (v/v) PEG 200 for a few seconds. The reduction with DTN was confirmed by a colour change of crystals from yellow to colourless.

The crystals were tested for diffraction using D8 Venture diffractometer with a Photon II detector (Bruker) and a MetalJet D2 X-ray source (Excillum). The testing images were collected at 100 K with the wavelength of 1.3418 Å and the exposure time of 40 s per image.

Calculation of the surface electrostatic potential of *Ct*FDO and visualization of its cavities accessible to solvent

Hydrogen atoms were added using PropKa for protonation states at pH 7 and pH 5.5 [131]. The PQR parameter files (containing charge and radius for each atom) were prepared assuming reduced FAD ($FADH^-$) and using AMBER ff99 force field parameters [132]. Surface electrostatic potential distributions were calculated using Adaptive Poisson-Boltzmann Solver APBS [133]. The rendered electrostatic potential was visualized with the APBS plug-in of the PyMOL Molecular Graphics System (Schrödinger, LLC). Cavities and other interior surfaces were identified with HOLLOW [134] and the volume of the active-site pocket was established with 3V server [135].

[2]Complex *Ct*FDO:4NC was refined with anisotropic ADPs except for 452 water molecules. Complex *Ct*FDO:ABTS as refined with anisotropic ADPs except for the water molecules and one *N*-acetyl-D-glucosamine moiety modelled linked to Asn197 of both chains.

Table 4.1: Data-collection parameters and statistics from data processing for *Ct*FDO:free, *Ct*FDO:MAMB, *Ct*FDO:PESB, *Ct*FDO:IPEA, *Ct*FDO:4NC, *Ct*FDO:4NP, and *Ct*FDO:ABTS. Values in parentheses are for the outer resolution shell.

	*Ct*FDO:free	*Ct*FDO:MAMB	*Ct*FDO:PESB	*Ct*FDO:IPEA	*Ct*FDO:4NC	*Ct*FDO:4NP	*Ct*FDO:ABTS
PDB id	6ZE2	6ZE3	6ZE4	6ZE5	6ZE6	6ZE7	7AA2
X-ray source	Petra III, BL P13	Bessy II, BL 14.2	Bessy II, BL 14.1	Petra III, BL P13	Bessy II, BL 14.2	Bessy II, BL 14.1	Petra III, BL P13
Wavelength (Å)	1.0332	0.9184	0.9184	0.9763	0.9184	0.9184	0.9763
Detector type	Pilatus 6M	Pilatus 2M	Pilatus 6M	Pilatus 6M	Pilatus 2M	Pilatus 6M	Pilatus 6M
Crystal-to-detector distance (mm)	141.74	222.69	266.58	343.72	159.95	293.45	243.41
Oscillation width (°)	0.1	0.1	0.1	0.05	0.1	0.1	0.05
No. of used images	1699	1996	1498	3579	2000	1400	2000
Exp. time per frame (s)	0.04	0.1	0.2	0.04	0.2	0.1	0.04
Space group	$P2_1$	$P2_12_12$	$P2_12_12_1$	$P2_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters							
a, b, c (Å)	46.7, 116.8, 109.0	109.7, 115.6, 46.6	93.6, 109.9, 116.1	47.0, 117.0, 109.9	93.6, 109.6, 116.1	92.6, 109.7, 115.6	93.6, 109.8, 116.0
α, β, δ (°)	90.0, 90.8, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.7, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution range (Å)	46.69-1.31 (1.33-1.31)	43.26-2.22 (2.29-2.22)	49.38-1.60 (1.63-1.60)	43.40-1.82 (1.85-1.82)	29.06-1.26 (1.28-1.26)	47.23-1.50 (1.53-1.50)	49.34-1.40 (1.42-1.40)
Measured reflections	856212 (38307)	218724 (20580)	811144 (43706)	358504 (18223)	2316159 (81149)	976283 (48771)	825189 (43026)
Unique reflections	268173 (13444)	30083 (2690)	147085 (7605)	105562 (5247)	314691 (14281)	186872 (9171)	223121 (11468)
Completeness (%)	95.4 (96.8)	99.9 (100.0)	93.5 (98.5)[1]	99.4 (99.7)	98.3 (91.0)	99.5 (99.2)	95.5 (99.8)
Redundancy	3.2 (2.8)	7.3 (7.7)	5.5 (5.7)	3.4 (3.5)	7.4 (5.7)	5.2 (5.3)	3.7 (3.8)
Mean $I/\sigma(I)$	11.2 (2.0)	8.3 (1.8)	8.7 (1.6)	9.6 (1.7)	16.5 (1.7)	10.6 (1.7)	9.2 (1.4)
R_{merge}	0.045 (0.489)	0.167 (1.147)	0.088 (1.009)	0.079 (0.752)	0.060 (0.978)	0.086 (0.908)	0.058 (0.925)
R_{meas}	0.053 (0.599)	0.182 (1.236)	0.097 (1.108)	0.094 (0.888)	0.065 (1.076)	0.096 (1.008)	0.067 (1.079)
R_{pim}	0.029 (0.340)	0.066 (0.441)	0.039 (0.441)	0.051 (0.469)	0.024 (0.439)	0.042 (0.429)	0.033 (0.542)
$CC(1/2)$	0.999 (0.795)	0.996 (0.646)	0.998 (0.659)	0.997 (0.673)	1.000 (0.660)	0.997 (0.668)	0.999 (0.589)
Mosaicity	0.06	0.38	0.19	0.18	0.11	0.08	0.09
Solvent content (%)	48.1	47.5	48.0	48.6	47.9	47.1	47.9
Matthews coefficient ($\mathrm{Å}^3{\cdot}\mathrm{Da}^{-1}$)	2.37	2.34	2.36	2.39	2.36	2.32	2.36
Wilson B factor ($\mathrm{Å}^2$)	12.7	33.5	16.8	22.0	10.0	11.1	16.2

[1]The lower overall completeness is a consequence of omitting regions on the detector due to ice rings.

Table 4.2: Structure-refinement and validation statictics for *Ct*FDO:free, *Ct*FDO:MAMB, *Ct*FDO:PESB, *Ct*FDO:IPEA, *Ct*FDO:4NC, *Ct*FDO:4NP, and *Ct*FDO:ABTS. Asymmetric unit and *N*-acetyl-D-glucosamine are abbreviated to AU and GlcNAc.

	*Ct*FDO:free	*Ct*FDO:MAMB	*Ct*FDO:PESB	*Ct*FDO:IPEA	*Ct*FDO:4NC	*Ct*FDO:4NP	*Ct*FDO:ABTS
PDB id	6ZE2	6ZE3	6ZE4	6ZE5	6ZE6	6ZE7	7AA2
Molecular refinement program	Morda	Molrep	Molrep	-	Phaser	Phaser	Phaser
Refinement program	Refmac5	Refmac5	Refmac5	phenix_refine	Refmac5	Refmac5	Refmac5
R_{work}	0.106	0.178	0.169	0.167	0.133	0.155	0.128
R_{free} (% of reflections)	0.137 (2.0)	0.255 (5.0)	0.203 (5.0)	0.202 (4.8)	0.147 (2.0)	0.182 (5.1)	0.172 (5.0)
Average B factor (Å^2)	17.4	41.3	19.6	26.5	14.3	16.0	19.1
R.m.s.deviations							
Bonds (Å)	0.013	0.010	0.010	0.011	0.014	0.012	0.013
Angles (°)	1.825	1.738	1.658	1.127	1.839	1.776	1.751
Ramachandran favoured[a] (%)	96.1	94.0	95.7	95.6	96.2	96.1	96.2
Ramachandra outliers[a] (%)	0.0	0.3	0.2	0.2	0.0	0.2	0.0
No. of protein chains in AU	2	1	2	2	2	2	2
Target ligands in AU		1x MAMB	2x PESB	3x IPEA	4x 4NC	3x 4NP	2x ABTS
(occupancy)		(1.0)	(0.9)	(1.0)	(0.7-0.8)	(0.8)	(0.6-0.9)
No. of water molecules	1664	310	1611	1272	1884	1489	1653
Other localized moieties							
FAD	2	1	2	2	2	2	2
GlcNAc	11	6	12	10	13	10	13
Mannose	-	-	1	-	-	-	-
Formic acid	2	1	4	4	4	8	2
Mg^{2+}	1	1	1	2	3	5	3
Na^{+}	-	1	2	-	1	1	0
Cl^{-}	-	-	-	-	1	1	1
Acetate ion	-	-	-	-	-	2	-
Tetraethylene glycol	-	-	-	-	-	2	-
Triethylene glycol	-	-	-	-	-	2	-

[a]was calculated by Molprobity.

4.3.2 UV-VIS spectrophotometry on *Ct*FDO crystals

The absorption spectra (200-1100 nm) were acquired using a UV-VIS spectrophotometer HR2000+ES (OceanOptics) (MX-SpectroLab, Bessy II, Helmholtz Zentrum, Berlin, Germany) with the OceanView spectroscopy software. The $Ct\text{FDO}_{\text{degl}}$ crystal used for the measurement had a size approximately 150·150·50 μm and was grown without any ligand and vitrified without cryoprotection. The spectra were recorded before and after crystal irradiation by X-ray beam with a dose of the order of 1 MGy (beamline 14.2, Bessy II, Helmholtz Zentrum, Berlin, Germany) [127]. The crystal was vitrified in a 100 K nitrogen gas stream during the data measurement. The dark background was subtracted at the beginning of the spectra measurement. The recorded spectra were normalized (at 290 and 900 nm) with GraphPad Prism version 7.02 for Windows (GraphPad Software, La Jolla California USA, www.graphpad.com).

4.3.3 *Ct*FDO activity testing

Verification of oxidoreductase activity

The fully glycosylated *Ct*FDO was reduced by 0.25 mM DTN and the following re-oxidation was monitored. The solution (70 μl total volume) contained 280 μg *Ct*FDO, 0.25 mM DTN, 25 mM Tris-HCl pH 7.5, and 100 mM NaCl. The re-oxidation effect was monitored in a wavelength range of 280-600 nm with a UV-VIS spectrophotometer Libra S22 (Biochrom Ltd.) in combination with the Resolution Spectrophotometer PC Software (Biochrom Ltd.) for 16 min. The reaction ran under aerobic conditions at 23 °C and the solution was stirred after 3 and 9 min. The spectra were buffer-subtracted.

Colorimetric activity assay with ABTS

The reaction mixture (100 μl total volume) contained 0.5 μg *Ct*FDO, 5 mM ABTS, and 40 mM Britton-Robinson buffer (pH 4.5-8.9)[3]. The reaction ran at 45 °C for 20 min and then at room temperature for 3 min. Absorbance of the product (at 420 nm) was measured with CLARIOstar Monochromator Microplate Reader (BMG Labtech).

Colorimetric activity assay based on coupled reaction with horseradish peroxidase

The reaction (40 μl total volume) carried out either with 0.25 μg *Ct*FDO in 40 mM Britton-Robinson buffer (pH 3.5-9.8) at 45 °C for 20 min, or with 2 μg *Ct*FDO in 100 mM phosphate buffer with 30 mM NaCl (pH 4.5-9) at 37-45 °C for 30 min. After that and after 3 min cooling down to room temperature, the reaction mixture containing $\sim 6\,\text{U}\cdot\text{ml}^{-1}$ of horseradish peroxidase (HRP, Sigma-Aldrich cat. No

[3]Britton-Robinson buffer is composed of boric acid, acetic acid, and phosphoric acid adjusted with NaOH to required pH.

P8250)[4] [136] was added in a ratio of 1:1. Absorbance of the product (at 540 nm) was measured with CLARIOstar Monochromator Microplate Reader (BMG Labtech). The list of substrates used for *Ct*FDO activity determination and details describing particular reactions are summarized in Table A.3.

Furthermore, a liquid compost extract was prepared and tested as a potential source of *Ct*FDO substrate. About 200 g of one year old compost was blended with 0.5 l of distilled water and incubated at 60 °C for 30 min. The infusion was filtered via series of filters: filter paper, sterile syringe filter unit with 0.22 μm pore size hydrophilic polyethersulfone membrane (Millex), and Nanosep 10K centrifugal device with 10 kDa molecular weight cut-off (Pall Corporation). The reaction composed of 20 μl of the extract, 2 μg of *Ct*FDO and 40 mM Britton-Robinson buffer (pH 4.5, 6, 7.5, 8) ran at 45 °C for 30 min. After that and after following 3 min cooling down to room temperature, the reaction mixture containing HRP was added in a ratio of 1:1. Absorbance of the resulting product was measured at 540 nm using CLARIOstar Monochromator Microplate Reader (BMG Labtech).

Luminescent assay based on coupled reaction with luciferase

Together 990 compounds with a molecular weight in range from 101 to 1550 g·mol^{-1} were selected for high-throughput activity screening (HTS) using ROS-Glo H_2O_2 Assay (Promega, worldwide.promega.com). The criterion for compound selection was to have at least one oxidizable OH^- group. The HTS was performed by Mgr. David Sedlák, Ph.D. from CZ-OPENSCREEN (Institute of Molecular Genetics of the Czech Academy of Sciences). The detailed information about the screening including complete list of used compound is shown in Supplementary Information 2 of article Švecová, *et. al.* [119].

4.3.4 *Ct*FDO$_{degl}$ activity testing

*Ct*FDO$_{degl}$ sample was tested with several representatives of substrates of GMC oxidoreductases using the colorimetric activity assay based on coupled reaction with HRP. The reaction mixture (100 μl total volume) contained 1 μg of *Ct*FDO$_{degl}$, 10 mM substrate (glucose, cellobiose, pyridoxine, choline, benzyl alcohol, 4-methoxybenzyl alcohol, 4-hydroxy-3-methoxybenzyl alcohol, or 5 mM coniferyl alcohol) and buffer with pH between 6 and 8 (100 mM Bis-Tris pH 6 with 50 mM NaCl; 100 mM Tris-HCl pH 7 and 8 with 50 mM NaCl). The reactions ran at 45 °C for 20 min and then were cooled down to the room temperature for 3 min. Then, 100 μl of the reaction mixture containing HRP ($\sim$ 6 U·ml^{-1}, mentioned above) was added. Absorbance of the product (at 540 nm) was measured with CLARIOstar Monochromator Microplate Reader (BMG Labtech).

[4]The reaction mixture was composed of 11 mM *N*-ethyl-*N*-sulfopropyl-m-toluidine, 4 mM 4-aminoantipyrine, and $\sim$ 6 U of HRP

Bilirubin oxidase from *Myrothecium verrucaria*

4.4 Preparation of *Mv*BOx mutated variants

The cloning, expression, and purification of *Mv*BOx wild type (*Mv*BOx-WT) and its mutated variants – Trp396 mutation to alanine (*Mv*BOx-W396A), phenylalanine (*Mv*BOx-W396F), and aspartic acid (*Mv*BOx-W396D) were done by our collaborators from Novozymes A/S. The gene encoding bilirubin oxidase was amplified from *Myrothecium verrucaria* (*Albifimbria verrucaria*) strain ATCC24571. *Mv*BOx and its mutated variants were recombinantly expressed in *Aspergillus oryzae*. The purification was performed by anion-exchange chromatography. The purity of the samples was controlled by SDS-PAGE. Production of the samples was not the aim of this thesis. Details of the samples preparation are given in Koval' and Švecová, *et al.* [137].

4.5 Characterization of *Mv*BOx and its variants with biophysical techniques

4.5.1 Liquid chromatography tandem-mass spectrometry

The *Mv*BOx sample was digested to peptide fragments by serine protease - trypsin. The fragments were analysed by the liquid chromatography tandem-mass spectrometry/mass spectrometry (LC-MS/MS) technique using a 15T solariX FT-ICR mass spectrometer (Bruker Daltonics) operating in positive-ion mode. Fragments of peptides were identified with ProteinScape (Bruker Daltonics) with the Mascot search engine. The correct assignment of the peptide ion was confirmed by collision induced dissociation.

4.5.2 Verification of the fold and stability preservation of *Mv*BOx variants

Circular dichroism

CD spectra of *Mv*BOx wild type and its mutated variants was measured using a Chirascan^{TM}–plus spectrometer (Applied Photophysics Ltd.). The spectra were recorded in the far-ultraviolet region of 190-260 nm with a bandwidth of 1.0 nm using a quartz cell cuvette with a path length of 0.1 cm. The samples were diluted

in water to a final concentration of 0.2 mg·ml^{-1}. The experiment was done at room temperature. The final spectra were buffer-subtracted and reported in degrees of ellipticity.

UV-VIS spectrophotometry

The UV-VIS absorption spectra (300-800 nm) of *Mv*BOx wild type and its mutated variants were measured at room temperature using a DeNovix DS-11 microvolume spectrophotometer. All samples were diluted in 25 mM Tris-HCl pH 7.5 with 250 mM NaCl to the concentration of 3 mg·ml^{-1} and 1.2 μl of each sample was used for the measurement. The resulting spectra were buffer-subtracted.

Nano differential scanning fluorimetry (nanoDSF)

The protein temperature stability of *Mv*BOx wild type and its variants was measured using samples with a concentration of 1.5 mg·ml^{-1} in 50 mM Tris-HCl pH 7.5 with 50 mM NaCl. The fluorescence was measured in the range of 20-95 °C with a slope of 2 °C per minute using Prometheus NT.48 (Nanotemper) and the PR.ThermControl software (v2.1.1). The stability was measured as a dependence of the ratio of integrated fluorescence at 350 and 330 nm and its first derivative on temperature.

4.5.3 *Mv*BOx activity measurements

The activity of *Mv*BOx-WT and *Mv*BOx-W396A, *Mv*BOx-W396F, and *Mv*BOx-W396D against four substrates (potassium ferrocyanide ($K_4Fe(CN)_6$), 2,2'-azinobis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS), bilirubin, and 2,6-dimethoxyphenol (DMP)) was measured. The reactions (100 μl in total volume) were measured in triplicates using black 96-well plates (BRAND) at 27 °C. The reactions were monitored spectrophotometrically with CLARIOstar Monochromator Microplate Reader (BMG Labtech).

Oxidation of $K_4Fe(CN)_6$ (further referred as ferrocyanide, as only $[Fe(CN)_6]^{4-}$ undergoes the oxidation), ABTS, and DMP was monitored as an increase of product formation, proportional to an increase in absorbance at 420 nm, 420 nm, and 468 nm, respectively. The experimental set-up is summarized in Table 4.3. The steady-state kinetic parameters (maximal velocity $V_{\max}$, Michaelis-Menten constant K_M) were evaluated using the Michaelis-Menten equation (Eq. 1.3) implemented in GraphPad Prism version 7.02 for Windows (GraphPad Software, La Jolla California USA, www.graphpad.com). Oxidation of bilirubin to biliverdin was monitored at 440 nm as an absorbance decrease, corresponding to a decrease of bilirubin. The collected data showed "S-shaped" sigmoidal behaviour, therefore, the Hill equation (Eq. 1.4, according to GraphPad Prism 7.02 Software) was applied for calculation of the kinetic parameters. The Hill equation was evaluated by the F test and by the corrected Akaike's Information Criterion (AIC) calculations (GraphPad QuickCalcs Web, https://www.graphpad.com/quickcalcs/aic1/, accessed May 2021) as the best equation fitting the particular data.

Table 4.3: Experimental setup for catalytic activity measurement of *Mv*BOx wild type and its mutated variants with substrates bilirubin, ABTS, ferrocyanide, and DMP. All reactions (100 μl) ran at 27 °C. [S] is a concentration range of the substrate, ε_λ is an extinction coefficient of the measured product/substrate at the given wavelength λ.

	bilirubin	ABTS	ferrocyanide	DMP
[S] (mM)	0.001-0.180	0.2-20	0.1-10	0.5-60
λ (nm)	440	420	420	468
ε_λ ($mM^{-1}cm^{-1}$)*	56.30	36.00	1.04	14.80
buffer	200 mM Tris-HCl pH 8.7	100 mM Na acetate pH 4	50 mM Bis-Tris pH 6 25 mM NaCl	50 mM Bis-Tris pH 6.8 25 mM NaCl
time (min)	3	3	3	4
enzyme (μg)	0.02	0.06	0.15	0.40

The extinction coefficients as given in * [138–141].

4.6 X-ray structure determination

4.6.1 Crystallization

The crystals of *Mv*BOx -WT, *Mv*BOx-W396A , and *Mv*BOx-W396F were prepared in an acidic crystallization condition. The crystals grew in 1 μl drops composed in a 1:1 ratio of reservoir solution (14 % (w/v) PEG 3350 and 0.1 M succinic acid with pH 3.1) and protein solution (sample at a concentration of 25 mg·ml^{-1} in 25 mM Tris-HCl pH 7.5 with 250 mM NaCl) at 18 °C in 1-3 days. Crystals of *Mv*BOx and mutated variants were prepared using a hanging drop vapour-diffusion technique. The crystallization of *Mv*BOx samples was performed by Mgr. Tomáš Koval', Ph.D.

Diffraction data collection and processing

Single crystals of *Mv*BOx-WT and *Mv*BOx-W396A were soaked in reservoir solution containing mixture of cryoprotectants (15 % (v/v) PEG 200, 1 % (v/v) glycerol, 1 % (v/v) ethylene glycol, and 1 % (v/v) propylene glycol for *Mv*BOx-WT and 25 % (v/v) glycerol for *Mv*BOx-W396A) and ligand 10 mM $K_4Fe(CN)_6$ for 90 seconds. Crystals of *Mv*BOx-W396F were soaked in the reservoir solution containing 25 % (v/v) glycerol as the cryoprotectant and 10 mM pyrogallol for 60 seconds. Afterwards, crystals were vitrified and stored in liquid nitrogen.

For *Mv*BOx-WT and *Mv*BOx-W396A crystals, the diffraction data were collected at wavelength of 0.91841 Å on beamline BL 14.1, Bessy II synchrotron radiation source (Helmholtz Zentrum Berlin, Germany) at 100 K [127]. A MAR Mosaic CCD detector and a Dectris Pilatus 6M detector, both in combination with a mini kappa goniometer [142], were used for *Mv*BOx-WT and *Mv*BOx-W396A crystals, respec-

tively. For crystal of *Mv*BOx-W396F a 1.3418 Å wavelength (gallium Kα) was used and the data were collected at 100 K using a D8 VENTURE diffractometer with a Photon II detector (Bruker) and a METALJET D2 X-ray source (Excillum).

The diffraction data were processed and scaled in XDSgui [128] and merged in Aimless [38] (Table 4.4). The phase problem has been solved by molecular replacement using Molrep [44] and previously solved structure of *Myrothecium verrucaria* bilirubin oxidase (PDB id: 2XLL) [115] as a template. All structures were refined with Refmac5 [25] in combination with manual model building and refinement in Coot [48]. R_{free} statistic with 5 % of reflections was used as the cross validation method. The last refinement cycle in Refmac5 was ran with all reflections. Validation of structures was performed by the validation tools implemented in Coot, Molprobity [55], and wwPDB Validation service [18]. Carbohydrates were checked using the pdb-care program [130]. All structures were deposited in the PDB.

The geometrical restraint library for the Trp396–His398 crosslink was created based on the data available in the Cambridge Structural Database (CSD, The Cambridge Crystallographic Data Centre). The CSD database contained two records for the covalent link between the indole-C^{δ_1} and imidazole-N^{ε_2} atoms (CSD ids: CIMGUC and CIMHAJ [143]). The restrains were created using average geometrical parameters calculated from these records. The orientation of the Trp396-C^{δ_1}–His398-N^{ε_2} covalent bond was restrained independently in the planes of the two corresponding planar side chains (Fig. 4.2, Table 4.5).

The structures of complexes prepared by soaking crystals of *Mv*BOx wild type and *Mv*BOx-W396A in solution with $K_4Fe(CN)_6$ (complexes WT:FECN and W396A:FECN, respectively) were refined with product of $K_4Fe(CN)_6$ oxidation - with ferricyanide ($[Fe(CN)_6]^{3-}$). The restraint library for $[Fe(CN)_6]^{3-}$ used by CCP4 package [144] was edited, as it contains incorrect settings for distances between Fe and C atoms. The initial value 1.83 Å was changed to 1.93 Å according to the CSD database. The structure of *Mv*BOx-W396F (W396F) was without presence of desired ligand. The summary of structure determination statistics is presented in Table 4.6.

Calculation of the surface electrostatic potential of *Mv*BOx wild type and model of *Mv*BOx-W396D

The electrostatic potential was calculated for *Mv*BOx-WT and a model of *Mv*BOx-W396D mutant. The model was prepared using coordinates of the wild type structure (PDB id: 6I3J) and by replacement of Trp396 with an aspartic acid in the conformation without any clashes with surrounding residues. Hydrogen atoms were added with PropKa [131] for protonation states at pH 7. The PQR parameter files (charges and atomic radii) were prepared with PDB2PQR pipeline using the AMBER ff99 force field [132]. Calculations of the surface electrostatic potential distribution were performed using the Adaptive Poisson-Boltzmann Solver (APBS) [133]. Molecular graphics were done using the APBS plug-in of the PyMOL Molecular Graphics System (Schrödinger, LLC).

Table 4.4: Data-collection parameters and statistics from data processing for WT:FECN, W396A:FECN, and W396F. The values in the parentheses are for the highest resolution shell.

	WT:FECN	W396A:FECN	W396F
PDB id	6I3J	6I3K	6I3L
X-ray source	Bessy II, BL 14.1	Bessy II, BL 14.1	MetalJet D2
Wavelength (Å)	0.9184	0.9184	1.3418
Detector type	MAR mosaic CCD	Pilatus 6M	Photon II
Detector distance (mm)	313.5	266.7	75.0
No. of oscillation images	217	1000	400
Exposure time per image (s)	2	0.2	90
Oscillation width (°)	0.5	0.1	0.3
Space group	$F222$	$F222$	$F222$
Unit-cell parameters			
a, b, c (Å)	134.4, 203.9, 226.7	136.9, 201.8, 217.9	136.3, 200.7, 217.1
Resolution range (Å)	47.30-2.59 (2.67-2.59)	47.93-1.60 (1.63-1.60)	45.54-2.10 (2.14-2.10)
Measured reflections	313528 (21715)	739043 (37457)	400760 (13579)
Unique reflections	48305 (4411)	195423 (9686)	83782 (3866)
Completeness (%)	100.0 (99.9)	99.6 (99.8)	97.3 (82.4)
Redundancy	6.5 (4.9)	3.8 (3.9)	4.8 (3.5)
Mean $I/\sigma(I)$	13.5 (2.0)	9.2 (1.5)	7.3 (2.1)
R_{merge}	0.144 (0.858)	0.074 (0.635)	0.188 (0.649)
R_{meas}	0.157 (0.963)	0.086 (0.738)	0.212 (0.759)
R_{pim}	0.061 (0.428)	0.044 (0.373)	0.095 (0.386)
$CC(1/2)$	0.994 (0.608)	0.998 (0.700)	0.990 (0.697)
Mosaicity (°)	0.22	0.07	0.13
Solvent content (%)	62.1	61.0	60.4
Mathews coefficient (Å$^3\cdot$ Da^{-1})	3.25	3.20	3.10
Wilson B-factor (Å^2)	30.7	10.9	16.7

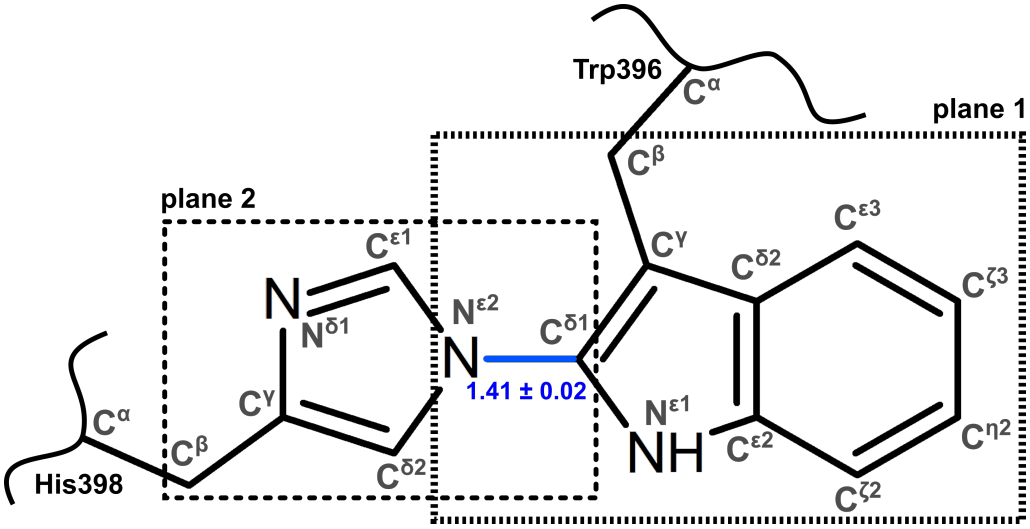

Figure 4.2: Schematic representation of Trp396–His398 crosslink in *Mv*BOx (PDB id: 6I3J). The crosslink is highlighted by a blue line and its length is labelled. The black dashed boxes mark restrained planes.

Table 4.5: Geometric restraints defined for the Trp396–His398 crosslink in *Mv*BOx wild type structure (PDB id: 6I3J).

bond	*bond length* (Å)
Trp396-$C^{\delta1}$–His398-$N^{\varepsilon2}$	1.410 ± 0.020
angles	*bond angle* (°)
Trp396-$C^{\delta1}$–His398-$N^{\varepsilon2}$–His398-$C^{\delta2}$	135.154 ± 3.000
Trp396-$C^{\delta1}$–His398-$N^{\varepsilon2}$–His398-$C^{\varepsilon1}$	114.069 ± 3.000
Trp396-$N^{\varepsilon1}$–Trp396-$C^{\delta1}$–His398-$N^{\varepsilon2}$	109.284 ± 3.000
Trp396-C^{γ}–Trp396-$C^{\delta1}$–His398-$N^{\varepsilon2}$	136.204 ± 3.000
atoms of plane 1	*deviation* (Å)
Trp396-(C^{β}, C^{γ}, $C^{\delta1}$, $N^{\varepsilon1}$, $C^{\delta2}$, $C^{\varepsilon2}$, $C^{\zeta2}$, $C^{\eta2}$, $C^{\varepsilon3}$, $C^{\zeta3}$) and His398-$N^{\varepsilon2}$	± 0.010
atoms of plane 2	
Trp396-$C^{\delta1}$ and His398-(C^{β}, C^{γ}, $N^{\delta1}$, $C^{\varepsilon1}$, $C^{\delta2}$, $N^{\varepsilon2}$)	± 0.010

Table 4.6: Structure refinement and validation statistics for WT:FECN, W396A:FECN, and W396F. Abbreviations used: AU, asymmetric unit and GlcNAc, *N*-acetyl-D-glucosamine.

	WT:FECN	W396A:FECN	W396F
PDB id	6I3J	6I3K	6I3L
R_{work}	0.160	0.131	0.159
R_{free}	0.226	0.154	0.196
Average B factor (Å^2)	36.1	15.5	19.8
R.m.s.d. bonds from ideal (Å)	0.009	0.011	0.009
R.m.s.d. angles from ideal (°)	1.350	1.670	1.539
Ramachandran favoured[a] (%)	94.4	95.77	95.30
Ramachandran outliers[a] (%)	0	0	0
No. of protein chains in AU	2	2	2
Ligands in AU (occupancy)	$9 \times \text{Fe(CN)}_6]^{3-}$	$4 \times \text{Fe(CN)}_6]^{3-}$	-
No. of water molecules	547	1670	1263
Other localized moieties			
GlcNAc	8	8	8
Mannose	6	3	2
Succinic acid	5	7	4
Propylene glycol	1	-	-
Di(hydroxyethyl)ether	-	2	-
Cu^{2+}	8	8	8
Glycerol	-	9	9
Na^{+}	8	2	1
Cl^{-}	1	-	-
K^{+}	-	1	-
Oxygen molecule	-	-	2

[a]was calculated by Molprobity.

Chapter 5

Results

FAD-dependent oxidoreductase from *Chaetomium thermophilum*

5.1 Verification of *Ct*FDO sequence

The *Ct*FDO sequence was verified by LC-MS/MS spectrometry that confirmed 595 amino-acid residues from the expected mature protein sequence (Fig. A.2). The signal peptide (21 residues), twenty following residues at the N-terminus, and eight residues at the C-terminus are missing in *Ct*FDO.

5.2 *Ct*FDO and *Ct*FDO$_{\mathbf{degl}}$ in solution

Both the fully glycosylated (*Ct*FDO) and deglycosylated (*Ct*FDO$_{\mathrm{degl}}$) samples are monomers in solution as was confirmed using several methods (Figs. A.3 and A.4). The molecular weight was determined to be about 85 kDa and 68 kDa for *Ct*FDO and *Ct*FDO$_{\mathrm{degl}}$, respectively (Figs. A.3, A.4, and A.5). A wide peak around 85 kDa in MALDI-TOF spectra for *Ct*FDO shows possible heterogeneity of glycosylation. The theoretical molecular mass of nonglycosylated *Ct*FDO (for the sequence confirmed by LC-MS/MS) is 65 kDa, thus molecular mass of the oligosaccharide moieties is approximately 20 kDa (about 25 % of the whole mass of the *Ct*FDO, i.e. about 2.8 kDa per one oligosaccharide chain). The shift in the MALDI-TOF spectra of about 17 kDa in the molecular mass after *Ct*FDO deglycosylation indicates residual oligosaccharide chain at one glycosylation site. Unsuccessful deglycosylation at Asn197 was confirmed by *Ct*FDO$_{\mathrm{degl}}$ X-ray structure.

Measurements of nanoDSF with *Ct*FDO and *Ct*FDO$_{\mathrm{degl}}$ diluted in various buffers showed that both samples are most stable in solutions with pH between 7.0 and 7.5 (Fig. A.6, Table A.4), however, the melting temperatures for *Ct*FDO$_{\mathrm{degl}}$ are lower compared to *Ct*FDO. For the storage buffer 25 mM Tris-HCl pH 7.5 with 100 mM

NaCl, the cleavage of oligosaccharides caused decrease in melting temperature T_{m1} by about 3.5 °C compared to *Ct*FDO ($T_{m1} = 70.8$ °C) and has an impact on the second peak in the nanoDSF melting curve (Figs. 5.1a and A.6), which is weaker for *Ct*FDO$_{degl}$. A similar decrease in T_{m1} also results from the CD spectroscopy (Fig. 5.1b). The hydrodynamic diameter (D_h) decreased by 10 Å for *Ct*FDO$_{degl}$ compared to *Ct*FDO (about 80 Å) (Fig. 5.1c). The oxidation state (peaks around 390 and 460 nm) of FAD remained preserved upon deglycosylation (Fig. 5.1d). The IEF gel shows a smear between pI 8.0 and 8.3 for *Ct*FDO and four bands between pI 7.4 and 8.3 (approximately at 7.4, 7.5, 7.9, and 8.3 for *Ct*FDO$_{degl}$ (Fig. A.3c).

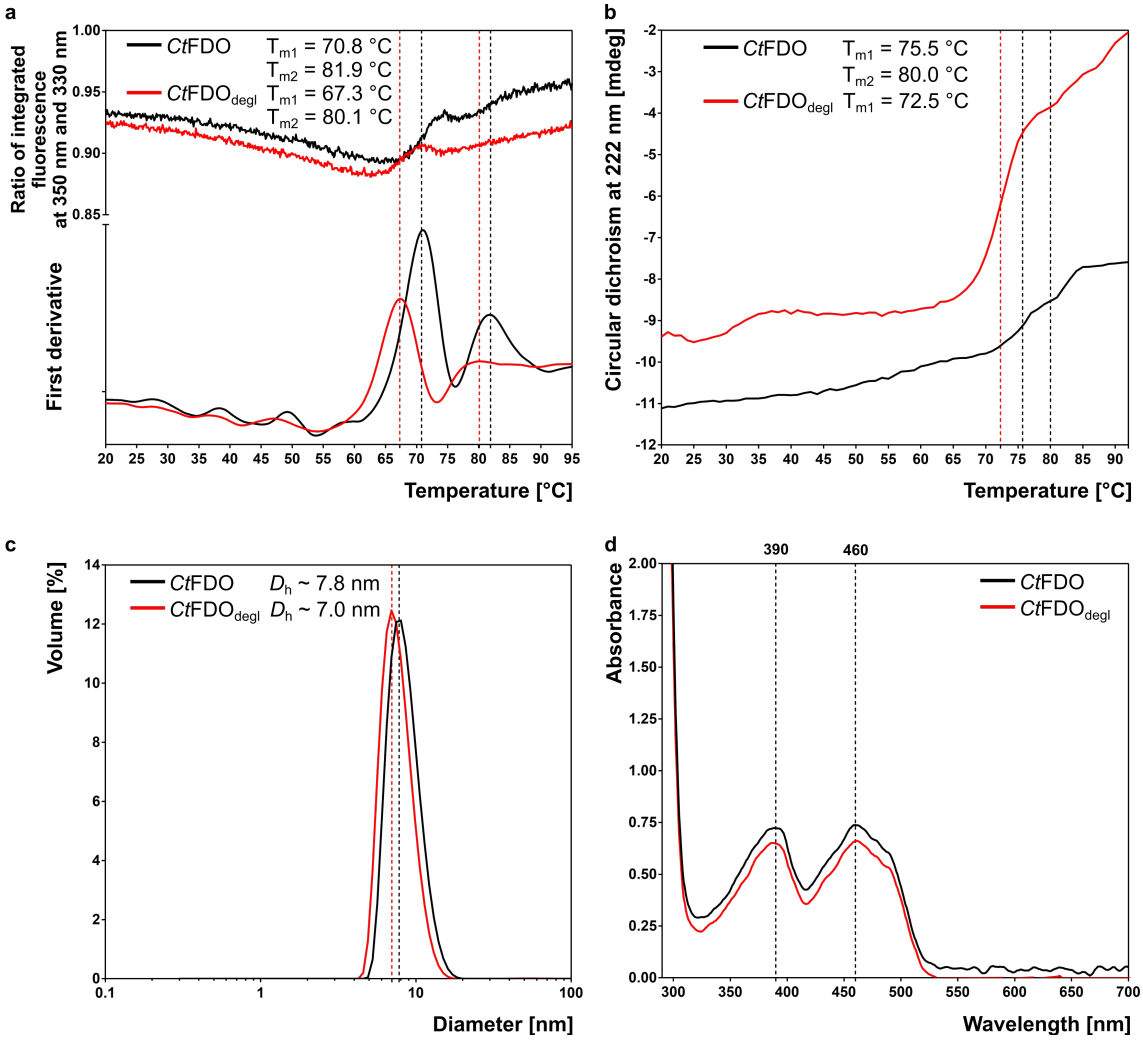

Figure 5.1: Comparison of biophysical properties of *Ct*FDO and *Ct*FDO$_{degl}$. Differences in melting temperatures (T_{m1} and T_{m2}) measured using (a) nanoDSF and (b) CD spectroscopy. While the curves for *Ct*FDO show two peaks, the curves for *Ct*FDO$_{degl}$ show weak or negligible peak for T_{m2} for both methods. (c) Size distribution analysis measured using MADLS indicates hydrodynamic diameters D_h of 7.8 nm and 7.0 nm for *Ct*FDO and *Ct*FDO$_{degl}$, respectively. (d) The UV-VIS absorption spectra measured for *Ct*FDO and *Ct*FDO$_{degl}$ in solution. Both spectra show two characteristic peaks for the oxidized state of FAD at 390 and 460 nm. Differences in peak heights are given by mild differences in concentration of the samples. The spectra were plotted using GraphPad Prism version 7.02 for Windows (GraphPad Software, La Jolla California USA, www.graphpad.com).

5.3 *Ct*FDO activity

The absorption spectra (Fig. 5.2), where the spontaneous re-oxidation of *Ct*FDO after reduction by sodium dithionite (DTN) is apparent, indicates that *Ct*FDO likely uses the molecular oxygen as the acceptor of electrons and protons. The enzymatic activity of *Ct*FDO was tested by two types of activity assays using almost 1000 various compounds from various classes – carbohydrates, aliphatic and aromatic alcohols, aldehydes, vitamins, lignin components, and others. None of the compounds showed significant activity with *Ct*FDO. Some representatives of substrates of GMC oxidoreductases were tested also with $CtFDO_{degl}$, but led to no positive results.

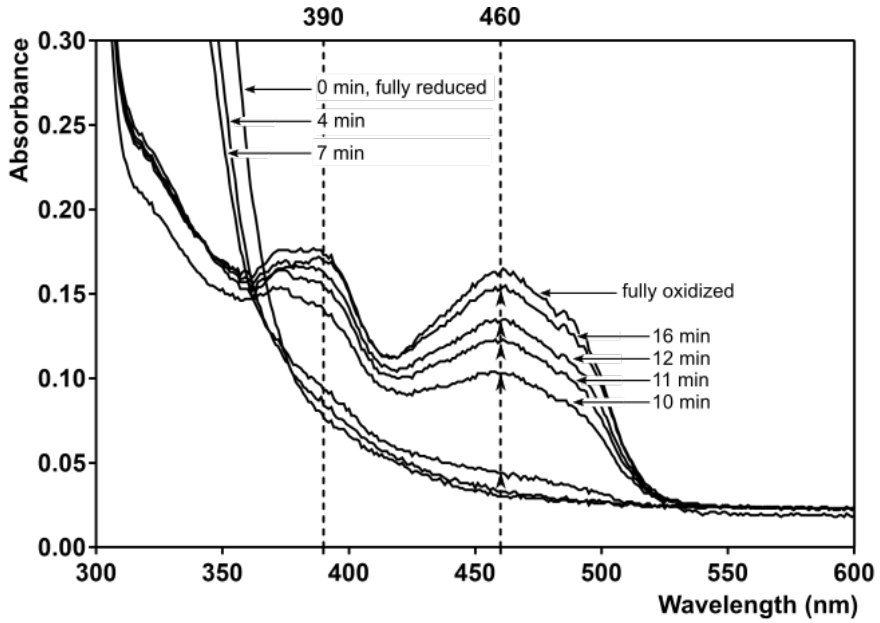

Figure 5.2: UV-VIS absorption spectra of *Ct*FDO monitoring spontaneous re-oxidation of *Ct*FDO by atmospheric molecular oxygen after chemical reduction by DTN. Profiles missing peaks around 390 and 460 nm are indicative for reduced state of FAD. The increased absorbance in the range of wavelengths of 300-360 nm corresponds to absorbance of the reducing agent and its consumption in time. The spectra were plotted with GraphPad Prism version 7.02 [119].

5.4 The overall crystal structure of *Ct*FDO

*Ct*FDO sequence and fold place it into the GMC oxidoreductase superfamily (Fig. A.2). The *Ct*FDO molecule has elongated cylindrical shape with dimensions approximately 74 Å × 53 Å × 43 Å. It consists of two domains: the N-terminal FAD-binding and the C-terminal substrate-binding domain (Fig. 5.3). The core of the former one is composed of a five-stranded parallel β-sheet (β_1, β_2, β_8, β_{11}, and β_{21}) crossed by three-stranded antiparallel β-sheet (β_9, β_{10}, and β_{11}) and surrounded by α-helices. It contains the initial $\beta\alpha\beta$ unit of the Rossmann fold (residues 28-56) with a highly conserved FAD-binding Gly-X-Gly-X-X-Gly sequence motif (residues 32-37). The domain includes the only disulfide bridge in *Ct*FDO interestingly created

by two sequence-adjacent cysteine residues (Cys566 and Cys567) – called a *trans* vicinal disulfide [145]. The substrate-binding domain contains a central six-stranded antiparallel β-sheet lined by helices. The only conserved residue of the domain among GMC oxidoreductases – the active-site histidine – corresponds to His564 in *Ct*FDO [68].

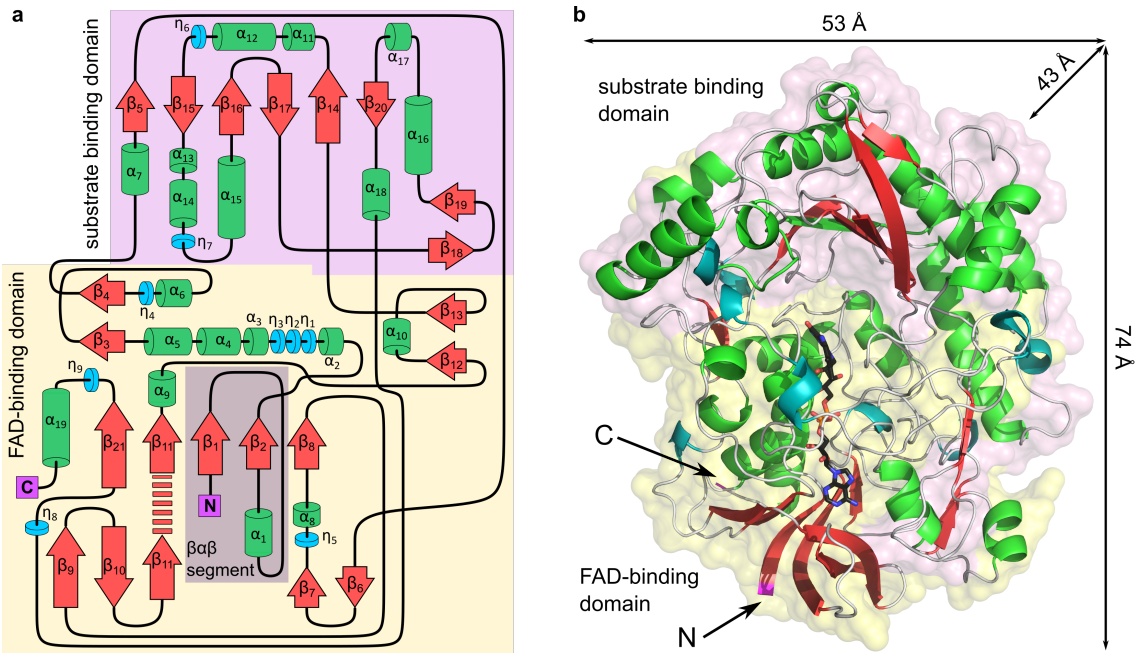

Figure 5.3: Topology and the crystal structure of *Ct*FDO$_{degl}$. (a) Topology diagram of secondary structure elements based on the *Ct*FDO$_{degl}$ crystal structure. The β-strands, α-helices, and 3_{10}-helices are shown as red arrows, green cylinders, and cyan discs (labelled as η), respectively. The pink and yellow background mark the substrate-binding and the FAD-binding domain. The grey box marks the $\beta\alpha\beta$ segment with the Gly-X-Gly-X-X-Gly sequence part conserved in GMC oxidoreductases. (b) The crystal structure of *Ct*FDO$_{degl}$. Colours of the secondary structure elements correspond to colours in the topology diagram. FAD is shown as sticks with black C atoms. The C and N letters mark the C- and N-terminus of the protein. The molecular graphics was created in Pymol (Schödinger, LLC).

The crystal structure of unliganded *Ct*FDO$_{degl}$ (*Ct*FDO:free) contains two molecules of *Ct*FDO$_{degl}$ (referred as chain A and B, r.m.s.d. of their C$^{\alpha}$ atoms is 0.12 Å) in the asymmetric unit. Clear electron density was observed for all amino acids verified by LC-MS/MS spectrometry except for four N-terminal (42-45) and six C-terminal (631-636) residues in both monomers, which were not modelled. The numbering of the amino-acid residues corresponds to the complete expected mature protein sequence including the signal peptide (Fig. A.2).

*Ct*FDO is a heavily glycosylated protein having seven N-glycosylation sites (Fig. 5.4) modified by oligosaccharide moieties with overall weight about 20 kDa. Six N-glycosylation sites (at Asn114, Asn182, Asn197, Asn295, Asn374, and Asn543) were confirmed by *Ct*FDO$_{degl}$ crystal structures and the seventh structurally inconclusive site at Asn46 was confirmed by MALDI-TOF peptide mass fingerprinting (Fig. A.7).

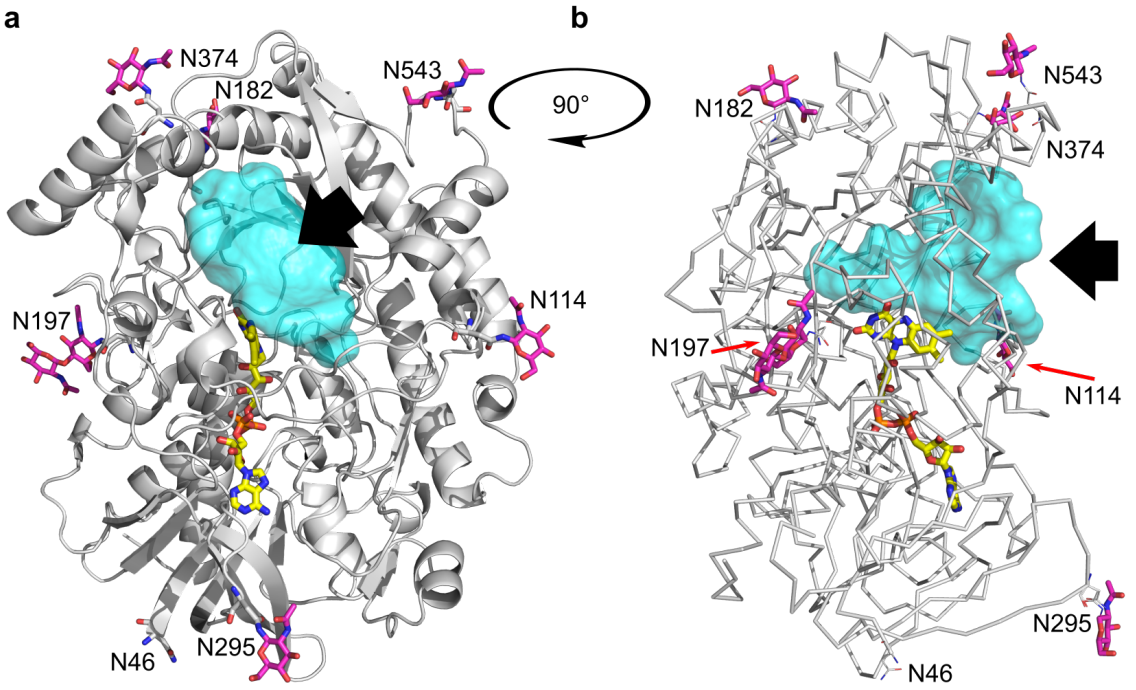

Figure 5.4: The overall crystal structure of *Ct*FDO$_{degl}$. The cofactor and the *N*-acetyl-D-glucosamine units (occupied at least in one of *Ct*FDO$_{degl}$ structures) are shown as sticks with C atoms coloured in yellow and purple, respectively. (a) The view of the entrance to the active-site pocket. (b) The side view of the active-site pocket (cyan surface, calculated using HOLLOW [134]) of *Ct*FDO$_{degl}$ (C^{α} as wires). The black arrow shows the entrance to the pocket. The molecular graphics were created in Pymol (Schödinger, LLC).

5.4.1 Access to the active site

The FAD-binding and substrate-binding domains border the active-site pocket of funnel-like shape and the mutual position and shape of domains cause the creation of a relatively wide-open entrance to the pocket (Fig. 5.4). The entrance is formed by two loops (residues 91-121 and 468-471) and the α_{11}-α_{12}-η_6 sequence part (residues 368-394).The volume of the pocket was established to be approximately 2400 Å^3 (HOLLOW [134], 3V server [135]). The pocket is filled only by water molecules.

5.4.2 FAD in *Ct*FDO, FAD environment and the active site

The FAD cofactor is non-covalently bound to *Ct*FDO directly via hydrogen bonds with thirteen residues (Ile56, Ser57, Glu77, Ala78, Val124, Gly128, Asn132, Ala133, Val135, Val273, Ile597, Ser607, and Met609) and indirectly via twelve water molecules. The pyrimidine moiety of FAD isoalloxazine ring is anchored in *Ct*FDO via hydrogen bonds to the main chain of Ala133, Val135, Met609, and Ser607 and via water-mediated hydrogen bonds to the main chain of Ala610, Leu137, and Ser607 (Fig. 5.5). Composite omit map calculated for FAD shows nonplanar conformation of the isoalloxazine ring (Fig. 5.6a). The angle measured between C^4–N^5–C^6 atoms

of FAD shows a bend of 20.0 ° (21.4 ° in chain B) around the N^5–N^{10} axis indicating reduced state of FAD (Fig. 5.6b) [8]. Absorption spectra of unliganded *Ct*FDO$_{degl}$ crystal show that FAD was oxidized before the exposure to X-rays and reduced after (Fig. 5.6c), therefore, the anionic fully reduced ($FADH^-$) form of the cofactor in all *Ct*FDO$_{degl}$ crystal structures is assumed.

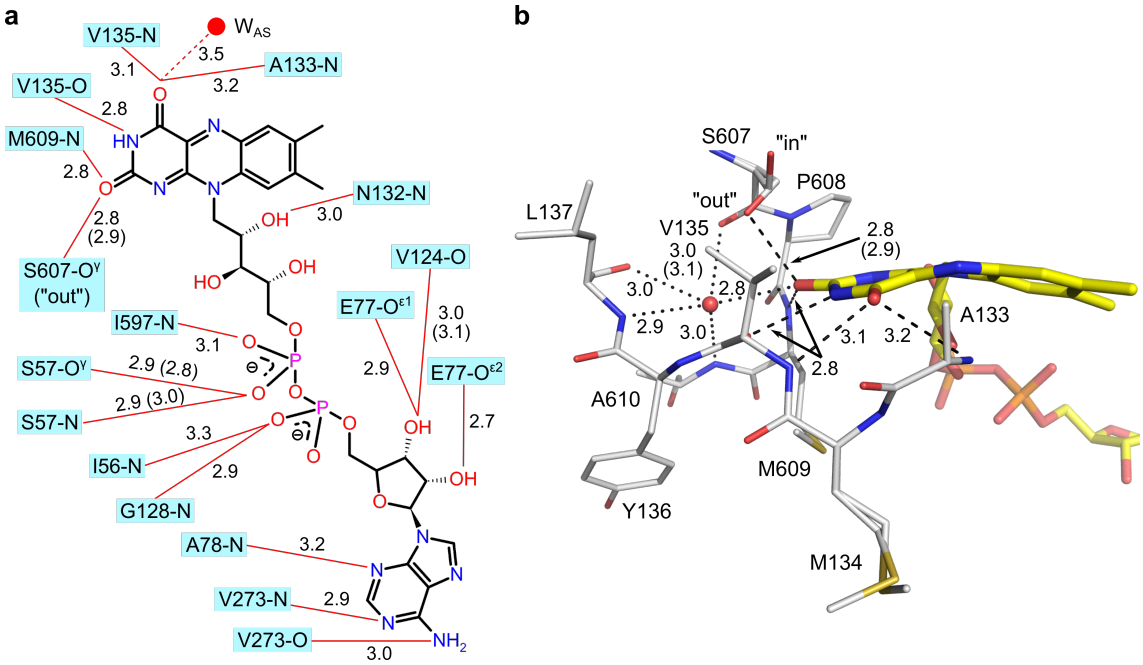

Figure 5.5: *Ct*FDO–FAD interactions. (a) Schematic diagram of *Ct*FDO–FAD interactions. The hydrogen bonds between FAD and *Ct*FDO are indicated by solid red lines and the hydrogen bonds to the active-site water molecule W_{AS} (red sphere) by dashed red line. (b) Anchoring of the FAD (C atoms in yellow) pyrimidine moiety in *Ct*FDO (C atoms in light grey). The water molecule (red sphere) mediating hydrogen bonds between FAD and the peptides is shown with all hydrophilic contacts. The direct hydrogen bonds are indicated as dashed lines and the water-mediated bonds as dotted lines. The graphics was created in Pymol (Schrödinger, LLC). The interatomic distances are given in Å. Different distances for chain B are written in parentheses. The "in" and "out" means two conformers of Ser607.

Five residues are located in the close vicinity of the FAD isoalloxazine ring – Asn562, Ala563, His564, and Ser607 on the *re*-face and Ala133 on the *si*-face of the isoalloxazine ring (Fig 5.6a). The position of His564 corresponds to the highly conserved active-site histidine in GMC oxidoreductase superfamily, which plays the role of the catalytic base during the reductive half-reaction in the majority of GMC enzymes [74, 89–94]. The orientation of the His564 imidazole ring is stabilized by a hydrogen bond to Gln351 (His564-$N^{\delta1}$–Gln351-$O^{\varepsilon1}$). The position of Ser607 (close to the pyrimidine moiety of the isoalloxazine ring) corresponds to the position of His/Asn/Arg/Pro in GMC enzymes, which create with the conserved histidine a His–His/His–Asn/His–Arg/His–Pro pair. The His564, Ser607, and Asn562 residues and the isoalloxazine ring bind a water molecule (W_{AS}) in the *Ct*FDO:free structure (Fig. 5.6).

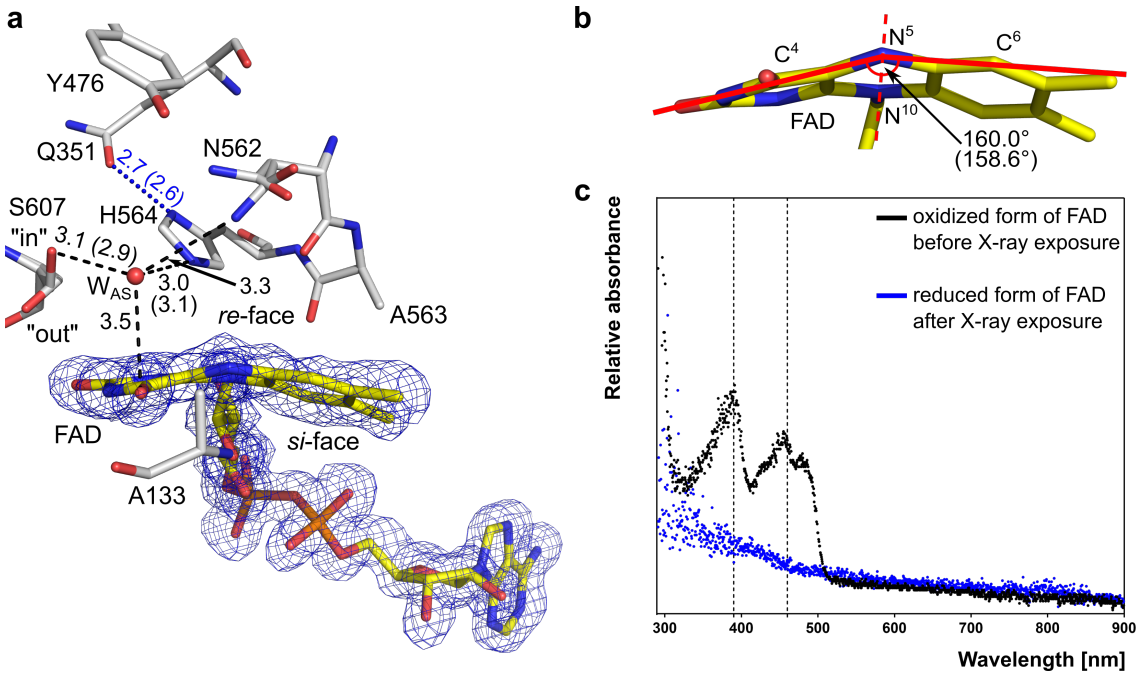

Figure 5.6: The FAD cofactor in *Ct*FDO. (a) The active site of *Ct*FDO:free (chain A, PDB id: 6ZE2) containing FAD (C atoms in yellow) surrounded by Ala133 in its *si*-face and Asn562, Ala563, His564, and Ser607 in its *re*-face, complemented by Tyr476 and Gln351 stabilizing the His564 orientation via His564-$N^{\delta 1}$–Gln351-$O^{\varepsilon 1}$ hydrogen bond (blue dotted line). Residues Ser607 and Asn562 were modelled in two alternative conformations. Conformations of Ser607 are labelled. The cofactor is shown with a simulated-annealing $2mF_o$-DF_c composite omit map contoured at 1σ level (blue mesh; Phenix [52]). The water molecule bridging FAD isoalloxazine ring, Ser607 ("in" conformation), Asn562, and His564 is shown as a red sphere and is labelled W_{AS}. Distances between W_{AS} and *Ct*FDO:free structure are given in Å (dashed lines). Different distances for chain B are written in parentheses. (b) The angle measured between atoms C^4, N^5, and C^6 of FAD indicating reduced state of FAD. The N^5–N^{10} axis of isoalloxazine bend is shown as a red dashed line. Molecular graphics were created in Pymol (Schödinger, LLC). (c) The UV-VIS absorption spectra measured with $Ct\text{FDO}_{\text{degl}}$ crystal before (black) and after (blue) the irradiation by X-rays confirming the oxidized and reduced state of FAD, respectively. The spectra were plotted using GraphPad Prism version 7.02.

5.5 Crystal structures of $Ct\text{FDO}_{\text{degl}}$ complexes

The soaking experiments of various compounds and fragments to $Ct\text{FDO}_{\text{degl}}$ crystals and co-crystallization experiments (Tables A.1 and A.2) resulted in collection of 144 data sets of putative complexes to resolution limit between 1.05 and 3.5 Å. The subsequent structural analysis revealed that four fragments bind to *Ct*FDO, namely methyl4-(aminomethyl)benzoate (MAMB), 4-oxo-*N*-[1-(3pyridinyl)-ethyl]-2-thiophenebutanamide (PESB), 2-(1*H*-indol-3-yl)-*N*-[(1-methyl-1*H*-pyrrol-2-yl)-methyl]-ethan-amine (IPEA), and 4-nitrocatechol (4NC). The co-crystallization experiments yield two additional compounds: 4-nitrophenol (4NP) and 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS) (Fig. A.1). Altogether seven structures of $Ct\text{FDO}_{\text{degl}}$ (including *Ct*FDO:free) were solved and deposited.

5.5.1 Analysis of binding sites in the active site of *Ct*FDO

The investigated ligands in *Ct*FDO$_{degl}$ complexes bind differently in the active site pocket of *Ct*FDO. Based on the structures of complexes, five binding sites inside the active-site pocket were identified: the catalytic site (CS) and subsites S1, S2, S3, and S4 (Fig. 5.7). CS is placed in the narrowing of the active-site pocket between the isoalloxazine ring and residues Ala133, Tyr476, Asn562-His564, and Ser607 (Fig. 5.6). Some residues of CS (Tyr476, Asn562, and Ser607) occasionally occur in alternative conformations reflecting their adaptation to ligand binding (Fig. A.8). Interestingly, Ser607 was observed in two alternative conformations (closest rotamer χ_1[1] = – 65° ("in") and χ_1 = 64° ("out")) in the majority of all 144 *Ct*FDO$_{degl}$ structures. The exceptions are complexes with ABTS and 4NC, where the "in" conformation of Ser607 is incompatible with the ligand binding. Unless occupied by a ligand, CS is binding a water molecule (active-site water molecule, W_{AS}) (Fig. 5.6) having hydrogen bonds to His564-N$^{\varepsilon 2}$, Ser607-O$^{\gamma}$ ("in" conformer), Asn562-N$^{\delta 2}$, and FAD-O^{4}.

The remaining subsites were numbered from the deepest part of the tunnel to its entrance. The most buried part of the pocket formed mainly by hydrophobic residues (Val135, Leu137, Lys231, Ile237, Gln351, Ile403, Val405, Thr464, His466, Leu474, Leu478, and Gly606) is marked as Subsite 1 (S1). Nearby located residues Val135, Ile237, and His466 form Subsite 2 (S2). Binding sites at the entrance to the tunnel were marked as Subsite 3 and 4 (S3 and S4). S3 in a shape of shallow pocket is formed by residues Val117-Ile120 and Asn559-Ser561 and Tyr501. The S4 subsite is formed by Phe94, Pro96, Trp97, and Phe100 (Fig. 5.7).

5.5.2 Analysis of crystal structures of *Ct*FDO$_{degl}$ complexes

The structures of *Ct*FDO$_{degl}$ complexes were solved with the resolution between 1.26 and 2.22 Å (Table 4.1). Although three different space groups ($P2_1$, $P2_12_12$, and $P2_12_12_1$) were observed for the complexes, their alignment with *Ct*FDO:free show only small differences in the overall structure (C^{α} r.m.s.d is 0.10-0.19 Å, PDBeFold [121]). Ligands modelled to the structures were refined with occupancy between 0.6 and 1.0 (Table 4.2).

All crystal structures of *Ct*FDO$_{degl}$ complexes apart from *Ct*FDO:MAMB contain two molecules of *Ct*FDO$_{degl}$ (chain A and B) in the asymmetric unit. An Mg^{2+} ion coordinated by six water molecules binds between proline-rich loops (Pro172-Pro175) of both chains, where it stabilizes their contact. Each chain, except for the structure of *Ct*FDO:MAMB, binds a formic acid molecule in the β-turn created by residues Pro413-Ser415. All structures contain at least two carbohydrate units linked to Asn197, indicating unsuccessful deglycosylation of this site, which is in agreement with the MALDI-TOF data (section 5.2).

[1] χ_1 for serine residue is the dihedral angle between atoms N, C^{α}, C^{β}, and O^{γ}.

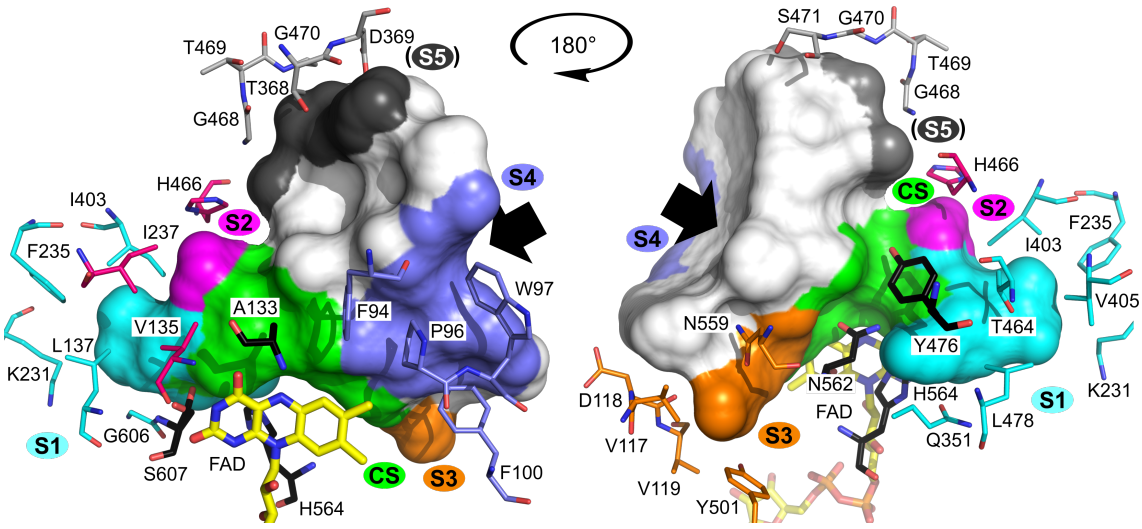

Figure 5.7: The binding subsites inside the active-site pocket of *Ct*FDO. The tunnel (calculated by HOLLOW [134]) is shown from two sides, arrows point to the entrance of the pocket. The CS subsite is highlighted in green, subsite S1 in cyan, S2 in magenta, S3 in orange, S4 in purple, and a putative S5 subsite in dark grey. Chosen residues lining the pocket are shown as sticks with C atoms coloured by colour of the corresponding subsite. Residues forming the CS have C atoms coloured in black. The cofactor is shown by sticks with C atoms in yellow. The molecular graphics were created using Pymol (Schödinger, LLC).

*Ct*FDO:MAMB complex (PDB code: 6ZE3)

The MAMB molecule binds in CS via a hydrogen bond to Tyr476, via water-mediated bonds (via W_{AS} (HOH836)) to the isoalloxazine ring, Asn562, and Ser607 ("in" conformer) (Fig. 5.8a). Its position is further stabilized by van der Waals (vdW) interactions with the isoalloxazine ring, Ala133, Leu399, Tyr476, and Asn562 (Fig. A.9, Table A.5). The S1 site in *Ct*FDO:MAMB complex is occupied by a molecule of formic acid (from crystallization condition, Figs. 5.10 and A.9).

*Ct*FDO:PESB complex (PDB code: 6ZE4)

The aliphatic moiety of PESB molecule binds in CS and the thiophene moiety in subsite S2. PESB has hydrogen bond to Asn562 and via W_{AS} (HOH804 in chain A, HOH817 in B) interacts with isoalloxazine ring, His564, and "in" conformer of Ser607 (Fig. 5.8b). Further stabilization of PESB is established by vdW contacts with the isoalloxazine ring, Ala133, Val135, Leu399, His466, Leu474, Tyr476, and Asn562 (Fig. A.10, Table A.6). The S1 site is occupied by formic acid (Figs. 5.10 and A.10).

*Ct*FDO:IPEA complex (PDB code: 6ZE5)

The pyrrole and the aliphatic moiety of IPEA uses CS, where it replaces W_{AS}, and the indole moiety uses S4. It interacts with Tyr476 via a CH-π bond and a hydrogen bond and with Asn562 via direct and water-mediated hydrogen bonds (Fig. 5.8c). IPEA molecule is further stabilized via vdW interaction with the isoallox-

azine ring, residues Phe94, Pro96, Ile120, Ala133, Val135, Ile403, Tyr476, Asn562, His564, and Ser607 (Fig. A.11, Table A.7). The S1 site is occupied by the formic acid molecule (Figs. 5.10 and A.11).

Except for mentioned IPEA molecule binding in the same way in both chains of *Ct*FDO in *Ct*FDO:IPEA, the complex binds one other IPEA molecule (IPEA-symA) on the contact of chain A (Asp621-Arg628 peptide) and symmetry-related chain A (Trp97 and Leu98). IPEA-symA is bonded via CH-π interaction to the Trp97 indole moiety and via water molecules to Trp97, Asp621, Lys625, Arg628, and Leu629. It has also vdW contact to these residues and, in addition, to Leu98 and Lys624 (Figs. 5.8c and A.11, Table A.7).

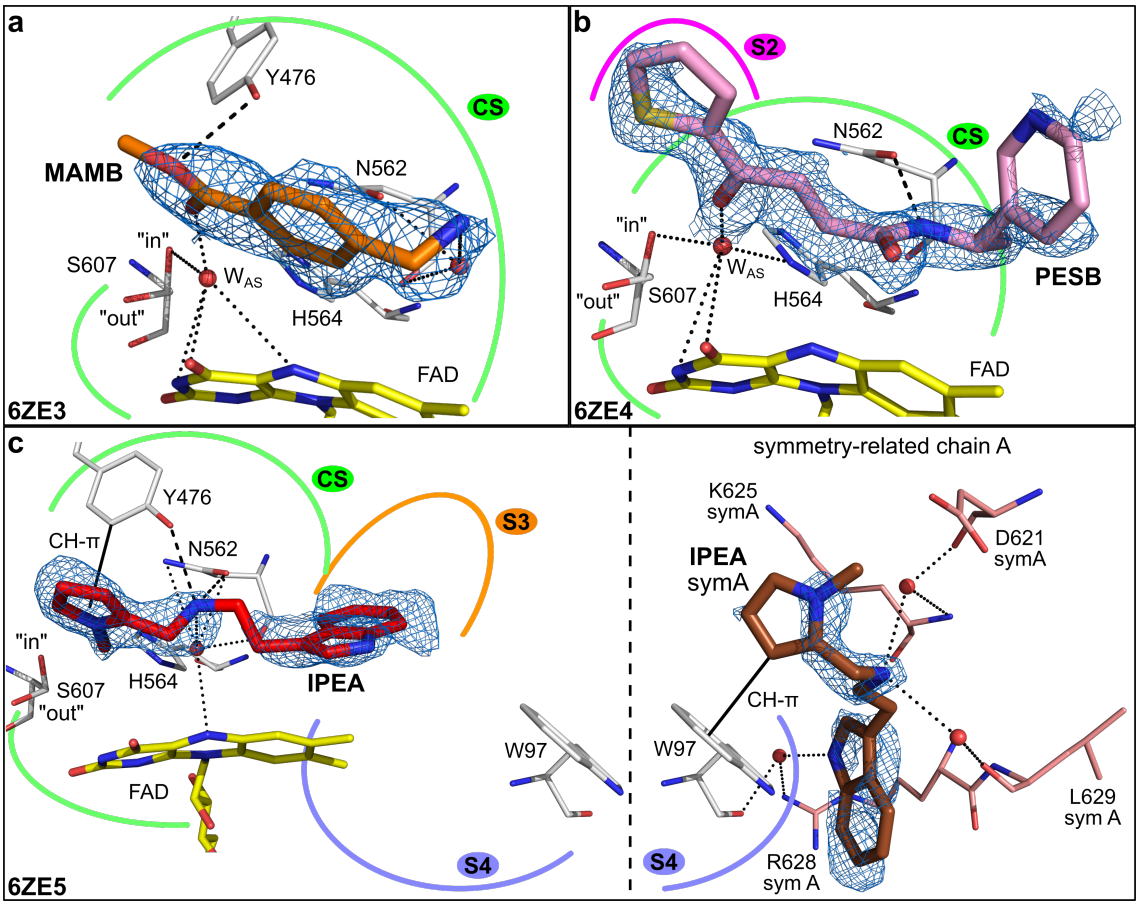

Figure 5.8: Binding of MAMB, PESB, and IPEA molecules to *Ct*FDO. Related PDB codes for complexes are shown in the left bottom corner. Ligands, FAD cofactor (C atoms in yellow), and residues (C atoms in grey) are shown as sticks. Ligands are shown with the simulated-annealing $2mF_o$-DF_c composite omit map (blue mesh; calculated with Phenix [52]). The CS site and subsites S2, S3, and S4 are indicated by green, magenta, orange, and purple curves. Direct hydrogen bonds for ligands are shown as dashed lines and the water-mediated bonds as dotted lines. The active-site water W_{AS} is labelled. The composite omit map calculated (a) for MAMB (C atoms in orange) is contoured at 1σ level, (b) for PESB (C atoms in pink) at 0.8σ level, and (c) for IPEA molecules (C atoms in red and brown) at 1σ level. The molecular graphics were created in Pymol (Schödinger, LLC).

*Ct*FDO:4NC complex (PDB code: 6ZE6)

Two 4NC molecules bind in *Ct*FDO:4NC complex, one in subsite S1 and one other in subsite S3 (Fig. 5.9a). The CS is occupied by the formic acid molecule, where it replaces W_{AS} (Figs. 5.10 and A.12). 4NC in S1 has hydrogen bonds with Gln351, Asn562, and with formic acid molecule, but primarily, it is stabilized via vdW interactions with residues of subsite S1 (Leu137, Lys231, Phe235, Gln351, Ile403, Val405, Thr464, Tyr476, Leu478, Asn562, His564, Gly606, and Ser607). 4NC in subsite S3 is stabilized via direct hydrogen bonds with Ser561, Ala563, and Asn559, water-mediated hydrogen bonds with Ser561, Gln565, and Ile120, and vdW contacts with Phe100, Ile120, Ser561, Asn562, Ala563, Asn559, and the isoalloxazine ring (Table A.8). Position of the aromatic moiety of 4NC corresponds to that of the IPEA indole moiety in *Ct*FDO:IPEA . A strong peak (6σ) in mF_o-DF_c electron density map was found between residues Trp97 of chain A and Lys625 of the symmetry-related chain B. The peak was left unmodelled.

*Ct*FDO:4NP complex (PDB code: 6ZE7)

Each molecule of $Ct\text{FDO}_{\text{degl}}$ (chain A and B) binds one molecule of 4NP in subsite S1 (Fig. 5.9b), where it takes the same position and orientation as 4NC in the S1 subsite of *Ct*FDO:4NC and has the same vdW contacts as 4NC except for that with Gln351. 4NP has also direct hydrogen bonds with the "in" conformer of Ser607 and molecule of formic acid (bound in CS similarly as in *Ct*FDO:4NC (Figs. 5.10 and A.13)) and water-mediated hydrogen bond with Gln351 (Fig. A.13, Table A.8).

One additional 4NP binds in the entrance to the active-site pocket of the chain A (hydroxyl group in subsite S3 and nitro group in subsite S4). It has a hydrogen bond with Ser561 and via a water molecule interacts with Asn559. Further, it has vdW contacts with Phe94, Pro96, Ile120, Ser561-Ala563, and the isoalloxazine ring (Figs. 5.9 and A.13, Table A.8). The aromatic moiety of 4NP corresponds to the same position of 4NC and the indole moiety of IPEA in *Ct*FDO:4NC and *Ct*FDO:IPEA, respectively. A peak of electron density, likely for 4NP, was found also in chain B of *Ct*FDO:4NP, but left uninterpreted. Except for 4NP, *Ct*FDO:4NP binds PEG molecules (tetraethylene glycol and triethylene glycol in each chain) at the crystal contacts in the close vicinity to Trp97.

*Ct*FDO:ABTS complex (PDB code: 7AA2)

One sulfonate moiety of ABTS binds to CS. With respect to assumed $FADH^-$ state of the cofactor, the SO_3 group of ABTS in CS is probably protonated at O^{47}. The other sulfonate of ABTS with the benzothiazoline moiety uses subsite S4 for binding (Fig. 5.9c). This part of ABTS was modelled with 0.6 occupancy, as its electron density indicates its displacement in the direction from the active site and the alternative conformation of Trp97 excludes the full occupancy of this ABTS part. Trp97 and ABTS were modelled with complementary occupancies. ABTS has hydrogen bonds with Asn562, His564, Ser607 ("out" conformer), the isoalloxazine ring, and Arg628 of symmetry-related chain B. It has further water-mediated hydrogen bonds with Gln351, Asn397, Ser398, Asn562, Ser607 and with Asp621

of a symmetry-related chain B. It is also stabilized via numerous vdW interactions (Phe94, Pro96, Trp97, Ala133, Val135, Ser398, Leu399, Tyr476, Asn562, His564, Ser607, and the isoalloxazine ring) (Fig. A.14, Table A.9). An uninterpreted strong peak (4.2σ) in mF_o-DF_c electron density between Trp97 of chain B and Lys625 of symmetry-related chain A likely corresponds to missing C-terminal residues.

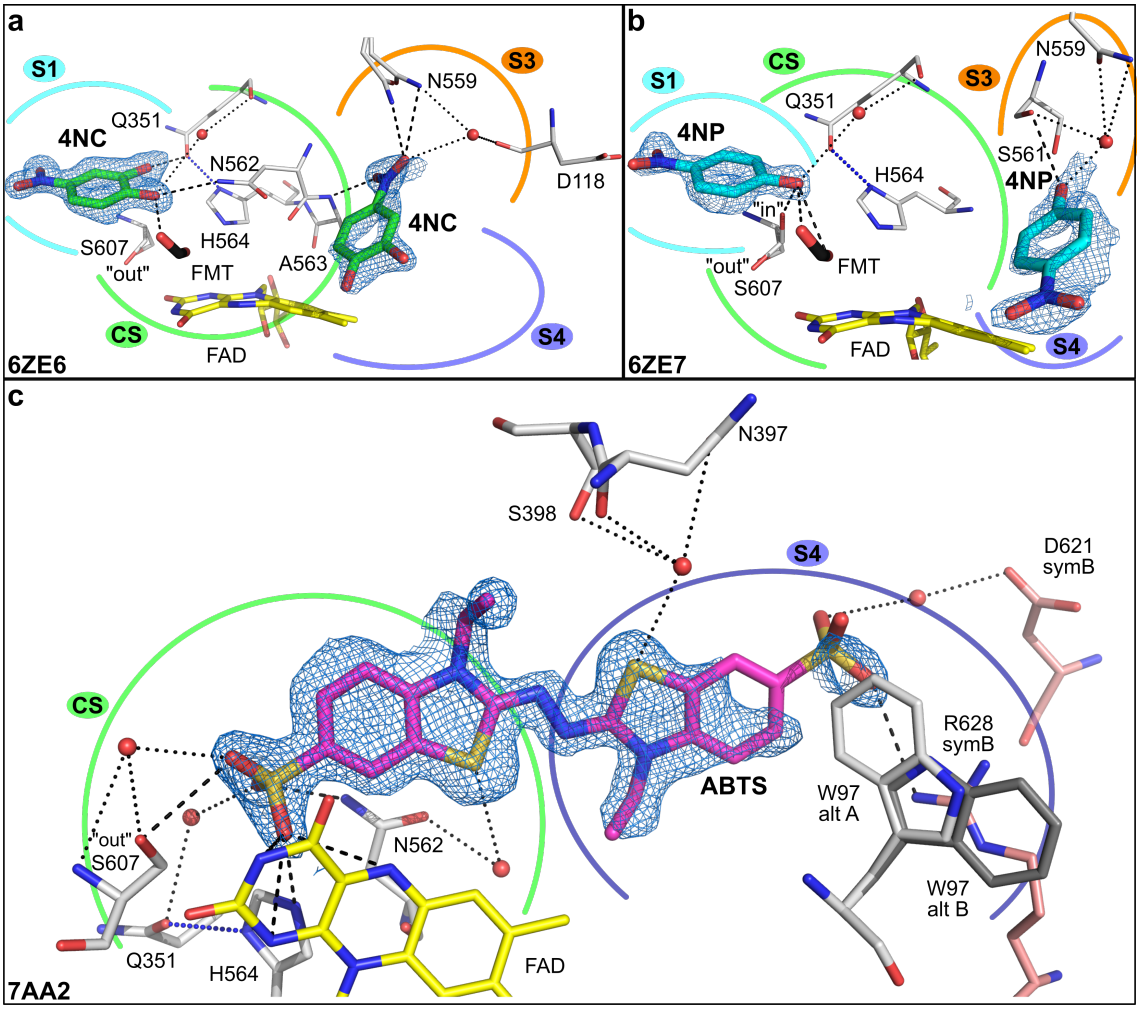

Figure 5.9: Binding of 4NC, 4NP, and ABTS molecules to *Ct*FDO. Related PDB codes are shown in the left bottom corner. Ligands, FAD cofactor (C atoms in yellow) and residues (C atoms in grey) are shown as sticks. Ligands are shown with the simulated-annealing $2mF_o$-DF_c composite omit map (blue mesh; calculated with Phenix [52]). The CS site and subsites S1, S3, and S4 are indicated by green, cyan, orange, and purple curves, respectively. The formic acid molecules are shown with black C atoms. The direct hydrogen bonds for ligands are shown as black dashed lines, the water-mediated bonds as black dotted lines, and the His564-$N^{\delta 1}$–Gln351-$O^{\varepsilon 1}$ bond as blue dotted line. (a) The composite omit map calculated for (a) 4NC molecules (C atoms in green) contoured at 1σ level, (b) 4NP molecules (C atoms in cyan) at 0.8σ level, and (c) ABTS molecule (C atoms in magenta) at 1σ level. The molecular graphics were created in Pymol (Schödinger, LLC).

Formic acid binding in the active-site pocket of *Ct*FDO

The formic acid molecule from crystallization condition was modelled in subsite S1 of *Ct*FDO:MAMB, *Ct*FDO:PESB, and *Ct*FDO:IPEA and in CS of *Ct*FDO:4NC and *Ct*FDO:4NP (Figs. 5.10). In S1, the formic acid molecule has direct hydrogen bonds with Thr464 and Ser607 (both conformations) and water-mediated hydrogen bond with Gln351. In *Ct*FDO:IPEA, it has one additional water-mediated bond to Asn562.

In *Ct*FDO:4NC and *Ct*FDO:4NP, formic acid replaces W_{AS} in CS, and it has direct hydrogen bond to Asn562, His564, 4NC or 4NP in subsite S1, and the isoalloxazine ring, and further water-mediated bond with Tyr476. In the structure of *Ct*FDO:4NP, it has an additional bond to the "in" conformer of Ser607 and in *Ct*FDO:4NC, water-mediated bonds to Asn562, His564, and the isoalloxazine ring.

Formic acid as a negatively charged molecule is probably attracted by the positive electrostatic potential of the active-site pocket at pH of the crystallization condition (pH 5.5, Fig. 5.11). Therefore, formic acid binding was observed in active-site pockets of almost all $Ct\text{FDO}_{\text{degl}}$ crystal structures.

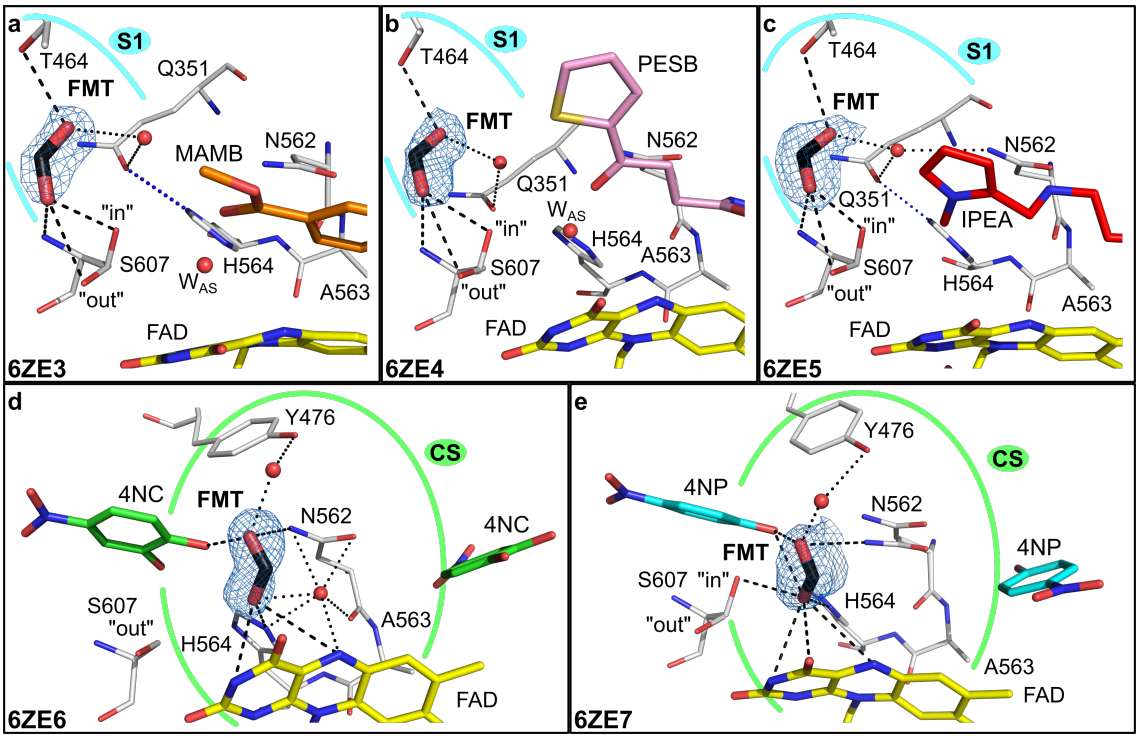

Figure 5.10: Formic acid (FMT) binding in the active-site pocket of $Ct\text{FDO}_{\text{degl}}$ complexes. A simulated-annealing $2mF_o$-DF_c composite omit map (blue mesh; Phenix [52]) calculated for FMT in subsite S1 (cyan curves) of (a) *Ct*FDO:MAMB complex (contoured at 0.8σ level), (b) *Ct*FDO:PESB (1σ), and (c) *Ct*FDO:IPEA (0.8σ), and in CS (green circle) of (d) *Ct*FDO:4NC (1σ), and (e) *Ct*FDO:4NP (1σ). Related PDB codes are shown in the left bottom corner. FMT is shown as sticks with black C atoms. W_{AS} is shown as a labelled red sphere. Ligands, residues, and interactions are displayed as in Figs. 5.8 and 5.9. The molecular graphics were created in Pymol (Schödinger, LLC).

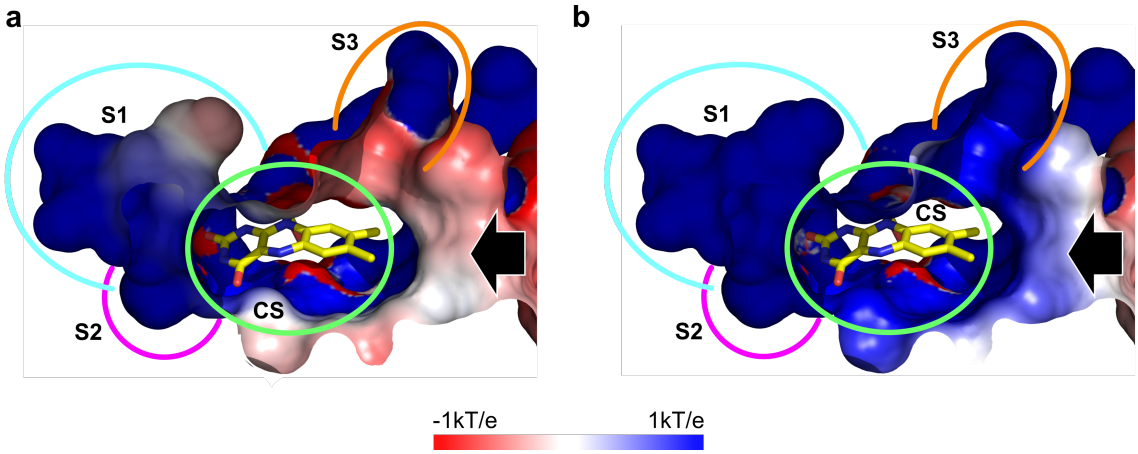

Figure 5.11: Surface electrostatic potential distribution displayed for the active-site pocket of *Ct*FDO (a) at pH 7, (b) pH of crystallization condition (pH 5.5). Both structures are shown in the same orientation. The CS site and subsites S1, S2, and S3 are indicated by green, cyan, magenta, and orange curves, respectively. The cofactor is shown as sticks with yellow C atoms. The protonation state at pH 7 and pH 5.5 were assigned with PropKa [131]. The PQR parameter files were prepared with the PDB2PQR pipeline using the AMBER ff99 force field [132]. Surface electrostatic potential distributions were realized with the Adaptive Poisson-Boltzmann Solver (APBS) [133]. The rendered electrostatic potential was visualized with the APBS plug-in of the PyMOL software (Schrödinger, LLC).

Bilirubin oxidase from *Myrothecium verrucaria*

5.6 Structure of *Mv*BOx and its comparison with previously published structures

New structure of *Mv*BOx wild type (*Mv*BOx-WT) in complex with ferricyanide (WT:FECN, PDB id: 6I3J) was obtained from a crystal growing in a strongly acidic crystallization condition (pH 3.1). *Mv*BOx crystallized in the space group $F222$ with two *Mv*BOx molecules in the asymmetric unit. The structure of WT:FECN is very similar to previously published structures of *Mv*BOx (PDB ids: 2XLL and 6IQZ), although different pH of crystallization conditions were used and each sample crystallized in a different space group (Table 5.1). Conformations of all residues around the T1Cu site, including the side chains, are essentially the same for all three structures (Fig. B.1). Moreover, the glycosylation sites (at Asn472 and Asn482) are preserved in all three structures.

Table 5.1: Comparison of *Mv*BOx-WT structures obtained from different crystallization conditions and published previously with that of WT:FECN (PDB id: 6I3J). pH means the pH of the crystallization condition and AU the asymmetric unit. The alignment (chains A, 533 C^{α} atoms) was performed with PDBeFold [121].

PDB entry	production organism	pH	space group	No. of *Mv*BOx molecules in AU	r.m.s.d. of aligned atoms
6I3J	*Aspergillus oryzae*	3.1	$F222$	2	-
2XLL	*Myrothecium verrucaria*	8.7	$P1$	4	0.30 Å
6IQZ	*Komagataella pastoris*	5.0	$C2$	1	0.25 Å

5.6.1 Trp396–His398 crosslink in *Mv*BOx

The presence of the covalent crosslink between Trp396-$C^{\delta 1}$ and His398-$N^{\varepsilon 2}$ atoms in *Mv*BOx-WT was first confirmed by the X-ray crystal structure (WT:FECN) in the form of a short contact ($\sim$ 1.41 Å) between these atoms supported by the electron density (Fig. 5.12). Later, the crosslink was definitely confirmed by mass spectrometry observing an ion at m/z 649.8327 (4+) corresponding to the modified W^{387}ELINAGNG**W**^{396}T**H**^{398}PIHIHLVDFK^{408} peptide (Fig. B.2). Appropriate geometry of the crosslink was defined using geometrical restraints extracted from X-ray structures of small molecules with this type of bond deposited in the Cambridge Structural Database (CSD, see Materials and Methods for details). Refine-

ment of the WT:FECN structure with the crosslink restraint using only the target value of the distance between Trp396-$C^{\delta 1}$ and His398-$N^{\varepsilon 2}$ atoms (1.41 Å) resulted in little distortions of Trp396 and His398 planar side chains and in a small deviation of the newly defined bond from the planes of both side chains (with opposite signs in chains A and B). For the coplanarity of the bond with each side chain (independently), the bond was further restrained to target values of bond distances and angles derived from the CSD records. With these restraints, the distance between Trp396-$C^{\delta 1}$ and His398-$N^{\varepsilon 2}$ atoms was refined to 1.42 Å (chain A) and 1.40 Å (chain B), i.e. values close to the specified target values for the bond distance (1.41 Å). These settings resulted in successful refinement of the crosslink and to a good agreement with electron density without any difference peaks at this site.

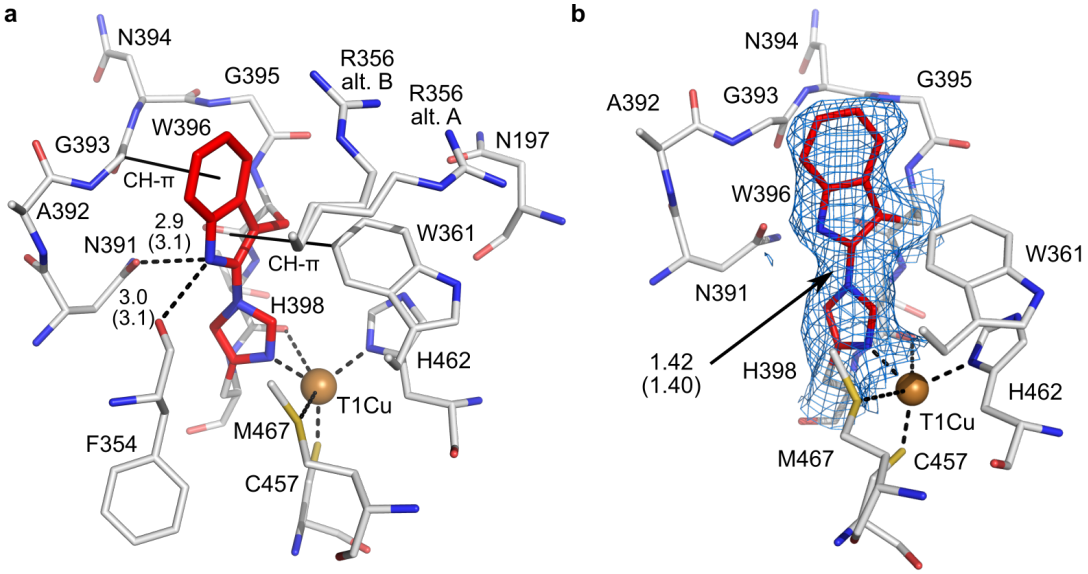

Figure 5.12: Structure of T1Cu site surroundings in *Mv*BOx-WT (PDB id: 6IJ3) from strongly acidic crystallization condition. (a) The Trp–His adduct (C atoms in red) with surrounding residues (C atoms in light grey) in *Mv*BOx-WT. The T1Cu is shown as a brown sphere. The residue Arg356 was modelled in two alternative conformations (alt. A and B). Hydrogen bonds (dashed lines) between the adduct and the surrounding residues are given in Å (different values for chain B are given in parenthesis). The CH-π interactions (solid lines) of the adduct are labelled. (b) The $2mF_o$-DF_c composite omit map (blue mesh, contoured at 1σ; calculated with Phenix [52]) shown for the adduct. Molecular graphics were created using PyMOL (Schrödinger, LLC).

The Trp396 side chain of the Trp–His adduct in *Mv*BOx-WT interacts with the surrounding residues: Trp396-$N^{\varepsilon 1}$ is hydrogen-bonded to the Phe354 main chain oxygen and the side chain oxygen of Asn391, the indole moiety of Trp396 interacts via CH-π hydrogen bonds with the Trp361 side chain and with the Gly393-C^{α} atom. His398 of the adduct is involved in T1Cu coordination. Besides residues mentioned above, the vicinity of the adduct is further formed by the Arg356 side chain and main chain atoms of Gly393-Asn394-Gly395 loop (Fig. 5.12a).

5.7 Mutated variants of *Mv*BOx

5.7.1 Effect of *Mv*BOx mutation on sample stability and fold

Three mutations of Trp396 were designed and produced to evaluate the role of the Trp–His adduct in *Mv*BOx: Trp396 mutation to alanine (*Mv*BOx-W396A) to allow direct solvent access to His398, phenylalanine (*Mv*BOx-W396F) to incorporate an aromatic residue not capable to form a crosslink, and aspartic acid (*Mv*BOx-W396D) to disrupt this site by implementation of negative charge. The point mutation of Trp396 did not affect the composition of the secondary structure of *Mv*BOx and neither the close vicinity of the T1 copper ion. All mutated variants have the same CD profiles as *Mv*BOx-WT (Fig. B.3a), which indicate the same proportion of secondary structure elements (in overall structures). The preservation of the T1Cu site was confirmed by the presence of a peak at 600 nm in all profiles in UV-VIS spectra (Fig. B.3b). On the other hand the nanoDSF measurements show light shift of melting temperatures T_{m} to lower values for *Mv*BOx-W396A and *Mv*BOx-W396F mutated variants (by about 5 °C and 7 °C, respectively) and more significant shift for *Mv*BOx-W396D mutant (by about 20 °C) (Fig. B.4).

5.7.2 Structures of *Mv*BOx-W396A and *Mv*BOx-W396F

Structure of *Mv*BOx-W396A in complex with ferricyanide (W396A:FECN, PDB id: 6I3K) and variant *Mv*BOx-W396F (W396F, PDB id: 6I3L) were obtained from the acidic crystallization condition. All samples crystallized in the same space group $F222$. The structure-based alignment of W396A:FECN and W396F with WT:FECN (r.m.s.d. of 533 C^{α} atoms for chains A is 0.23 Å and 0.21 Å, respectively) confirms that the mutation did not affect the fold and neither the surroundings of the T1Cu site. W396F structure obtained from a crystal soaked in reservoir solution containing pyrogallol (see Materials and Methods) lacks electron density for the ligand. Structure of *Mv*BOx-W396D mutant was not obtained, as the sample did not crystallize.

Coordination of T1Cu

The coordination of T1Cu in WT:FECN is different from those in MCOs without the Trp–His adduct. T1Cu does not bind in the imidazole plane of coordinating His398, as usual for the MCOs (e.g. CotA, PDB id: 1GSK [146]), but His398-imidazole and T1Cu form an angle of about 15-20 ° (Fig. 5.13). The tilt of His398 side chain is accompanied by a longer distance between T1Cu and His462-$N^{\delta 1}$ atoms by about 0.2 Å compared to other MCOs. These features were found only in WT:FECN. Angles and distances between T1Cu and side chains of coordinating histidines in W396A:FECN and W396F correspond to these in CotA and other MCOs that lack the Trp–His adduct.

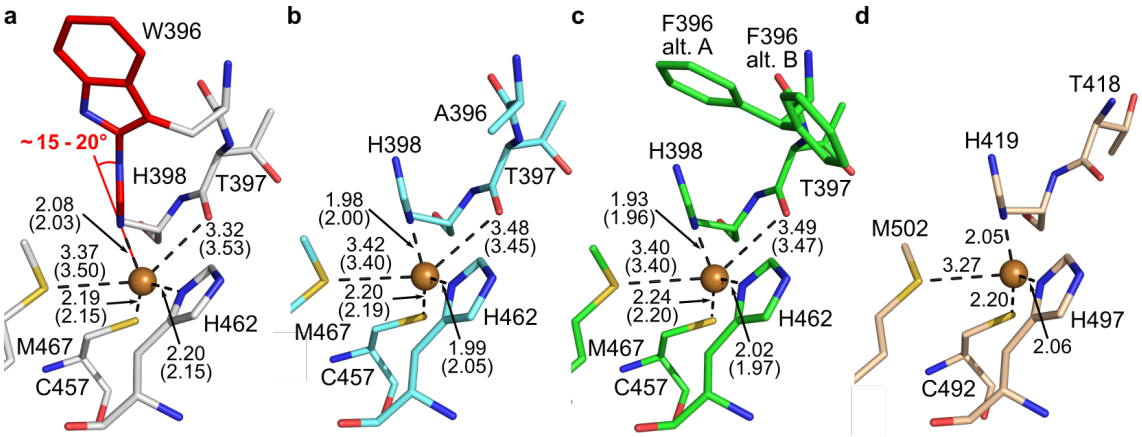

Figure 5.13: T1Cu site in (a) *Mv*BOx-WT (WT:FECN, PDB id: 6I3J, C atoms in light grey), (b) W396A (6I3K, cyan), (c) W396F (6I3L, green), and (d) CotA (1GSK [146], wheat). T1Cu is shown as a brown sphere. Distances are given in Å with 0.01 Å precision to support the discussion about the T1Cu environment changes. Values for chain B are given in parentheses. The residue Phe396 in W396F was modelled in two alternate conformations (alt. A and B). Molecular graphics were created in PyMOL (Schrödinger, LLC).

5.7.3 WT:FECN and W396A:FECN complexes

Structures of WT:FECN (PDB id: 6I3J) and W396A:FECN (6I3K) bind a ferricyanide ion in the active site. Its position was confirmed by the $2mF_o$-DF_c composite omit map and by a peak in an anomalous difference Fourier (Fig. 5.14). Instead of the substrate (ferrocyanide), which was soaked in *Mv*BOx-WT and *Mv*BOx-W396A crystals, the product ferricyanide was modelled as the crystals gradually changed their blue colour to transparent during soaking in the substrate solution (see Materials and Methods). That corresponds to the reduction of the T1Cu site and the concurrent oxidation of ferrocyanide to ferricyanide in the crystals.

In WT:FECN, the ferricyanide ion binds close to the Trp–His adduct, where it interacts with the main chain nitrogens of Asn394 and Gly395 and with the side chain nitrogen of Arg356 (Fig. 5.14a). The shortest observed distance between the ion and T1Cu is 11.4 Å (Fig. 5.14c), i.e. within the 14 Å limit for an effective electron tunnelling in proteins [66]. The ferricyanide ion at this site is not involved in any crystal contacts. That supports the assumption that this is the actual substrate/product binding site. Thus, we refer to this site as oxidation site 1 (OS1, Fig. 5.15).

In W396A:FECN, the ferricyanide ion binds similarly as in WT:FECN, but closer to T1Cu. The shortest observed distance between ferricyanide and T1Cu is 9.2 Å (Fig. 5.14d). Ferricyanide interacts with main chain nitrogens of Asn394 and Gly395 through two of its cyanide moieties in this structure, further with the main chain nitrogen of Ala396, with side chain of Arg356, and via a water molecule with His398 coordinating T1Cu (Fig. 5.14b).

Besides ferricyanide in the OS1 site, WT:FECN as well as W396A:FECN binds other ferricyanide ions. Some of them are involved in formation of crystal contacts, but none are binding in the close vicinity to the T1Cu site.

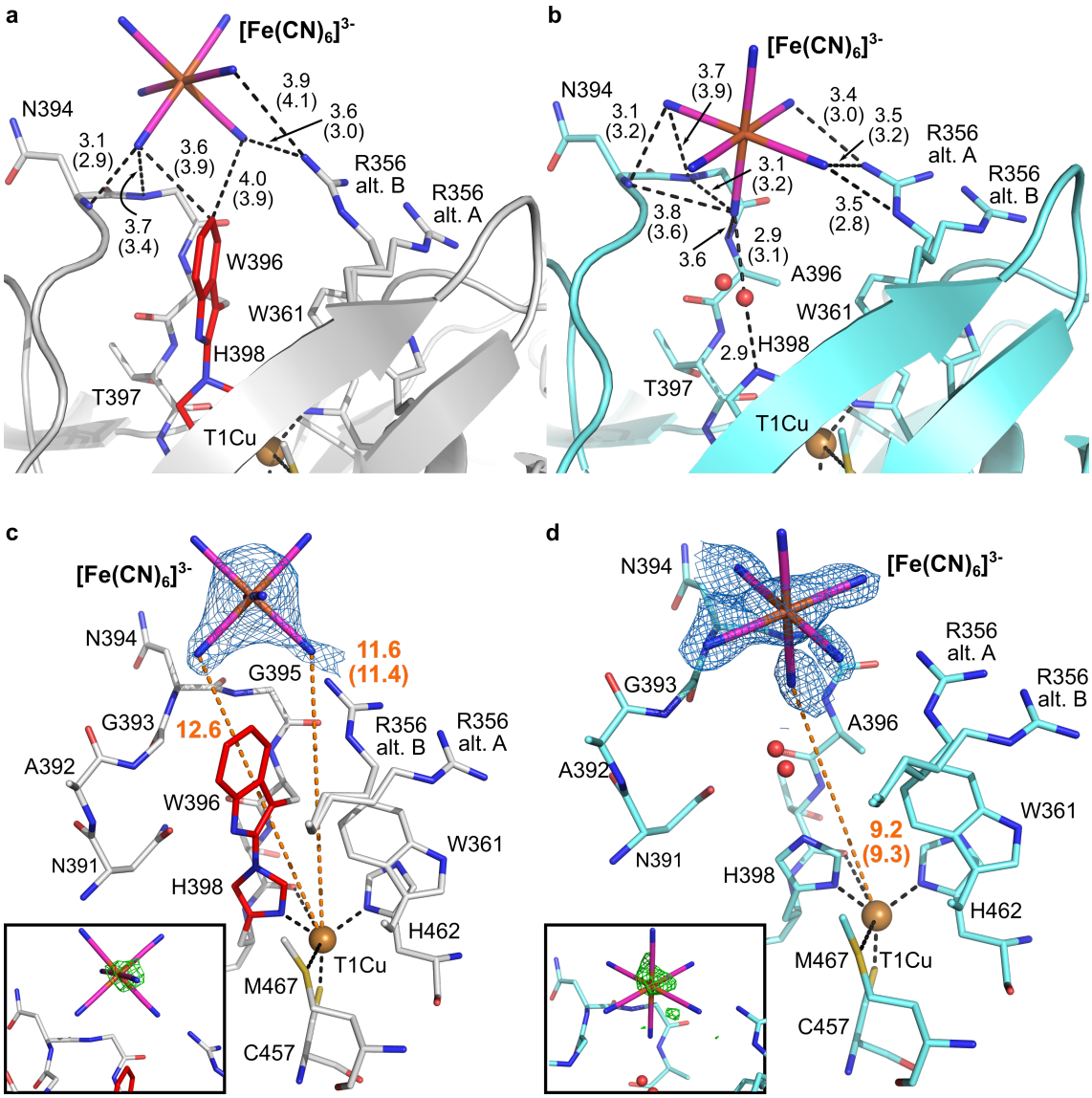

Figure 5.14: The T1Cu site of WT:FECN and W396A:FECN. Binding of ferricyanide ($[Fe(CN)_6]^{3-}$, magenta C atoms) in OS1 of (a) *Mv*BOx-WT (PDB id: 6I3J, C atoms in light grey) and of (b) W396A (PDB id: 6I3K, C atoms in cyan). Interactions between ferricyanide and the enzyme are shown as black dashed lines. A simulated-annealing $2mF_o$-DF_c composite omit map calculated for $[Fe(CN)_6]^{3-}$ in (c) WT:FECN (blue mesh, contoured at 0.8σ level) and (d) W396A:FECN (1.0σ level) (Phenix [52]). The insets in (c) and (d) show an anomalous difference Fourier (green mesh) contoured at 2.5σ level around iron of $[Fe(CN)_6]^{3-}$. Distances between $[Fe(CN)_6]^{3-}$ and T1Cu ions are shown as orange dashed lines. Distances are given in Å. Different values for chain B are given in parentheses. The T1Cu ion is shown as a brown sphere. A water molecule connecting $[Fe(CN)_6]^{3-}$ and His398 in W396A:FECN was modelled in two alternative positions and is shown as a red sphere. The Trp–His adduct is shown with red C atoms. The residue Arg356 was modelled in two alternative conformations (alt. A and B) in both structures. Molecular graphics were created using PyMOL (Schrödinger, LLC).

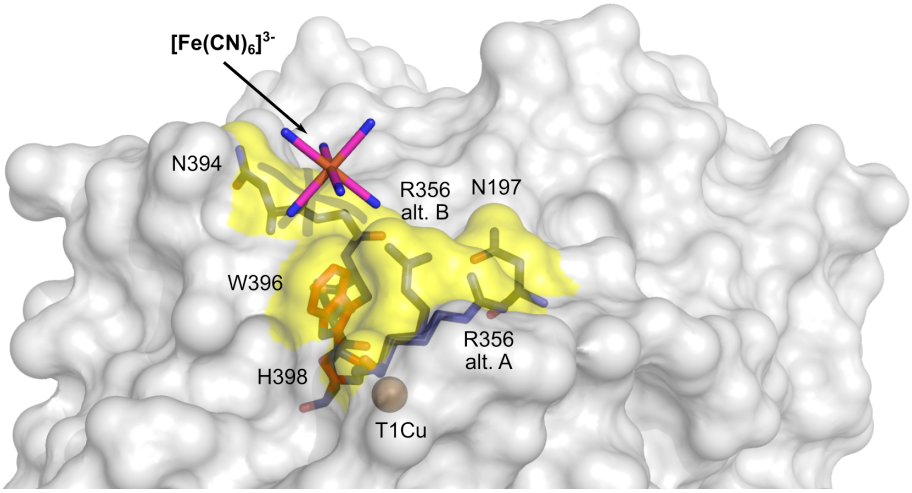

Figure 5.15: Oxidation site 1 (OS1) in *Mv*BOx-WT. The OS1 highlighted (in yellow) on the molecular surface of WT:FECN (PDB id: 6I3J, light grey). Residues forming OS1 and the ferricyanide ion are shown as sticks with black and magenta C atoms, respectively. The adduct is shown with red C atoms. T1Cu is shown as a brown sphere. Molecular graphics was created using PyMOL (Schrödinger, LLC).

5.8 Enzymatic activity of *Mv*BOx wild type and its variants

Enzyme kinetics of *Mv*BOx wild type and mutants was investigated using four different substrates representing inorganic compounds ($K_4Fe(CN)_6$, ferrocyanide), substituted phenols (2,6-dimethoxyphenol, DMP), a standard substrate for MCOs (2,2'-azinobis-(3-ethylbenzothiazoline-6-sulfonic acid), ABTS), and the canonical substrate for bilirubin oxidase (bilirubin) (Fig. B.5). Kinetic profiles are shown in Fig. 5.16 and the kinetic parameters are summarized in Table 5.2.

The activities of *Mv*BOx wild type and its mutated variants exhibit sigmoidal dependency of the reaction velocity v_0 on bilirubin concentration, and thus the data were fitted using the allosteric sigmoidal Hill equation (Eq. 1.4). For substrates ferrocyanide, ABTS, and DMP, the kinetic parameters were calculated using Michaelis-Menten equation (Eq. 1.3), providing the data converge. In other cases, the kinetic parameters were left uncalculated and a simple connecting line was used to link the mean values.

Oxidation of bilirubin is affected by all Trp396 mutations

Reactions of the enzymes with bilirubin were measured by detection of a decrease of bilirubin concentration over time. For determination of kinetic parameters, the values of absorbance decrease with an inverted sign were utilized. The kinetic profiles show sigmoidal dependence on bilirubin concentration, as confirmed using the F test and the AIC calculations test (see Material and Methods). Therefore, the allosteric sigmoidal equation (Eq. 1.4) was used to calculate the kinetic parameters (Fig. 5.16a, Table 5.2). The $K_{1/2}$ values for the wild type and *Mv*BOx-W396F are

comparable ($[0.060 \pm 0.002]$ mM and $[0.079 \pm 0.004]$ mM, respectively). On the other hand, the V_{max} value of *Mv*BOx-W396F ($[8.8 \pm 0.5]$ nmol·min^{-1}·μg^{-1}) is roughly a half of the wild type value ($[15.8 \pm 0.4]$ nmol·min^{-1}·μg^{-1}). Mean values of reaction velocity calculated for bilirubin oxidation with *Mv*BOx-W396A variant were determined with significantly higher errors compared to other *Mv*BOx variants. Activity measurements for *Mv*BOx-W396A at saturating concentrations of bilirubin (i.e $> 160\mu$M) were performed but yielded high errors and the measurements were not interpretable by mean values. The *Mv*BOx-W396D variant exhibited virtually null activity ($[-0.4 \pm 2.8]$ % of *Mv*BOx wild type reaction velocity v_0 with 120 μM bilirubin for 2 min).

Oxidation of ferrocyanide is substantially affected for *Mv*BOx-W396D

The wild type and *Mv*BOx-W396A and *Mv*BOx-W396F variants have similar kinetic parameters for oxidation of ferrocyanide (Fig. 5.16b, Table 5.2). The mutation of Trp396 to aspartic acid resulted in a significant decrease in the enzymatic activity ($[1.27 \pm 0.06]$ % of v_0 of the wild type with 10 mM $K_4Fe(CN)_6$ for 3 min).

Oxidation of ABTS is affected by all tested mutations

The affinity for ABTS considerably decreased for *Mv*BOx-W396A ($[3.1 \pm 0.2]$ mM) and *Mv*BOx-W396F ($[6.8 \pm 0.4]$ mM) compared to the wild type ($[0.30 \pm 0.02]$ mM) (Fig. 5.16c, Table 5.2). Conversely, the V_{max} value ($[30.3 \pm 0.3]$ nmol·min^{-1}·μg^{-1} for *Mv*BOx-WT) increased lightly for both variants ($[37.5 \pm 0.9]$ nmol·min^{-1}·μg^{-1} for *Mv*BOx-W396A and $[33.8 \pm 0.9]$ nmol·min^{-1}·μg^{-1} for *Mv*BOx-W396F). The mutation of Trp396 to aspartic acid resulted in zero activity ($[-0.8 \pm 0.3]$ % of the wild type, reaction velocity for 20 mM ABTS for 3 min).

Oxidation of DMP is diminished by all tested mutations

Unlike the wild type, the mutants show negligible activity to DMP (Fig. 5.16d, Table 5.2), thus a simple connecting line was utilized to link the mean values. The variants *Mv*BOx-W396A, *Mv*BOx-W396F, and *Mv*BOx-W396D reached $[5.0 \pm 2.0]$ %, $[10.0 \pm 3.0]$ %, and $[0.6 \pm 0.3]$ % of the wild type reaction velocity for 30 mM DMP in 4 min, respectively.

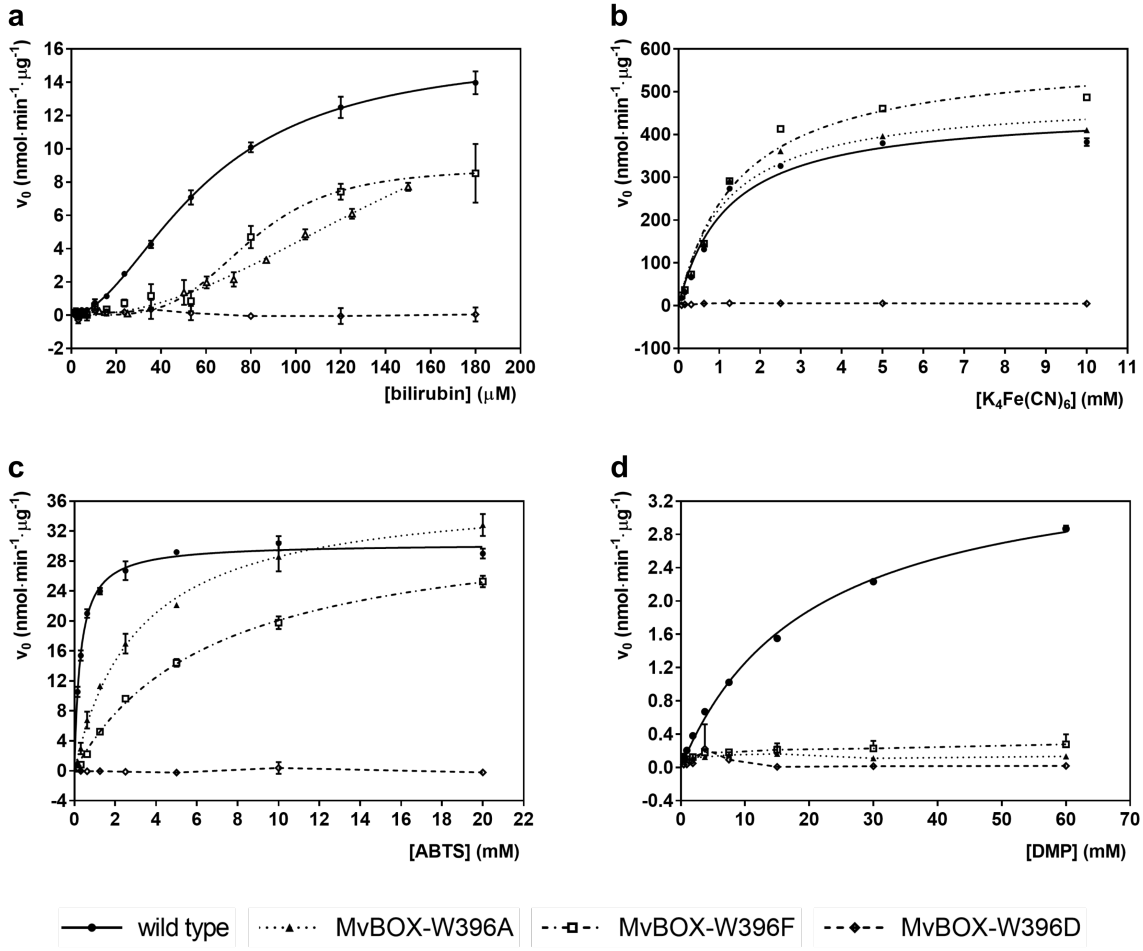

Figure 5.16: Comparison of catalytic activity of *Mv*BOx wild type and its mutated variants. Oxidation of (a) bilirubin, (b) $K_4Fe(CN)_6$, (c) ABTS, and (d) DMP by *Mv*BOx-WT (solid line, •), and *Mv*BOx-W396A (dotted line, ▲), *Mv*BOx-W396F (dot-dash line, □), and *Mv*BOx-W396D (dashed line, ♦). Plots show the reaction velocity v_0 as a function of the substrate concentration. The curves were fitted using the program GraphPad Prism version 7.02 (GraphPad Software). Standard deviations are shown as error bars. The mean values calculated for *Mv*BOx-W396D activity with all substrates and for all mutants with substrate DMP are linked with a simple connecting line [137].

Table 5.2: Kinetic parameters for oxidation of bilirubin, $K_4Fe(CN)_6$, ABTS, and DMP by *Mv*BOx wild type, and mutants *Mv*BOx-W396A and *Mv*BOx-W396F. Calculations were done for the measurements shown in Fig. 5.16 using the Michaelis-Menten equation Eq. 1.3 (K_M, V_{max}) for $K_4Fe(CN)_6$, ABTS, and DMP; and allosteric sigmoidal equation (Eq. 1.4 , $K_{1/2}$, V_{max}) for bilirubin. The parameters for *Mv*BOx-W396A and *Mv*BOx-W396F oxidizing DMP could not be calculated due to almost zero activity. *The value of h lies within a range of 2-4. The exact value could not be evaluated due to the high error present in some of the points obtained for the measurement of bilirubin oxidation by *Mv*BOx-W396F.

Substrate Enzyme	K_M (mM)	$K_{1/2}$ (mM)	V_{max} (nmol·min^{-1}·μg^{-1})	h
Bilirubin				
*Mv*BOx		0.060 ± 0.002	15.8 ± 0.4	1.9 ± 0.1
*Mv*BOx-W396A		0.160 ± 0.050	17.0 ± 6.0	2.2 ± 0.3
*Mv*BOx-W396F		0.079 ± 0.004	8.8 ± 0.5	$(4.1 \pm 0.7)^*$
$K_4Fe(CN)_6$				
*Mv*BOx	1.2 ± 0.2		460 ± 20	
*Mv*BOx-W396A	1.2 ± 0.2		490 ± 20	
*Mv*BOx-W396F	1.5 ± 0.2		590 ± 20	
ABTS				
*Mv*BOx	0.30 ± 0.02		30.3 ± 0.3	
*Mv*BOx-W396A	3.1 ± 0.2		37.5 ± 0.9	
*Mv*BOx-W396F	6.8 ± 0.4		33.8 ± 0.9	
DMP				
*Mv*BOx	20.1 ± 0.9		3.78 ± 0.08	

Chapter 6

Discussion

FAD-dependent oxidoreductase from *Chaetomium thermophilum*

6.1 *Ct*FDO and oxidoreductases from the GMC superfamily

The tertiary structure of *Ct*FDO is similar to numerous enzymes of the GMC oxidoreductase superfamily with C^{α} r.m.s.d. ~ 1.7 Å, and seq. identity $\sim 30\,\%$ (PDBeFold [121]). The following structures were selected as representatives for a detailed comparison (Table 6.1): aryl-alcohol oxidase from *Thermothelomyces thermophilus* (*Mt*AAO), glucose dehydrogenase from *Aspergillus flavus* (*Af*GDH), glucose oxidases (GOXs) from *Aspergillus niger* (*An*GOX) and from *Penicillium amagasakiense* (*Pa*GOX), and aryl-alcohol oxidase from *Pleurotus eryngii* (*Pe*AAO).

Table 6.1: Secondary structure matching of *Ct*FDO with selected GMC enzymes (for chains A) performed with PDBeFold [121].

protein	PDB id	Q-score	C^{α} r.m.s.d.	seq. identity	C^{α} aligned
*Mt*AAO	6O9N	0.64	1.79 Å	37 %	559
*Af*GDH	4YNT	0.64	1.58 Å	29 %	523
*An*GOX	1CF3	0.62	1.61 Å	30 %	521
*Pa*GOX	1GPE	0.60	1.72 Å	28 %	523
*Pe*AAO	3FIM	0.60	1.67 Å	28 %	508

*Ct*FDO, *Mt*AAO, *Af*GDH, and GOXs contain the helical extension (α_{13}-α_{13}, residues 407-457 in *Ct*FDO), missing in *Pe*AAO and in the rest of the GMC oxidoreductases with known structure. They share also the short version of the loop between β_{16}-β_{17} β-strands in *Ct*FDO (residues 468-471), much longer in the majority of the GMC oxidoreductases (Fig. 6.1). *Ct*FDO and aryl-alcohol oxidases (*Mt*AAO

and *Pe*AAO) share the helical insertion α_{11}-α_{12}-η_6 in *Ct*FDO (residues 367-400), which is not present in GOXs and in *Af*GDH either (Fig. 6.1).

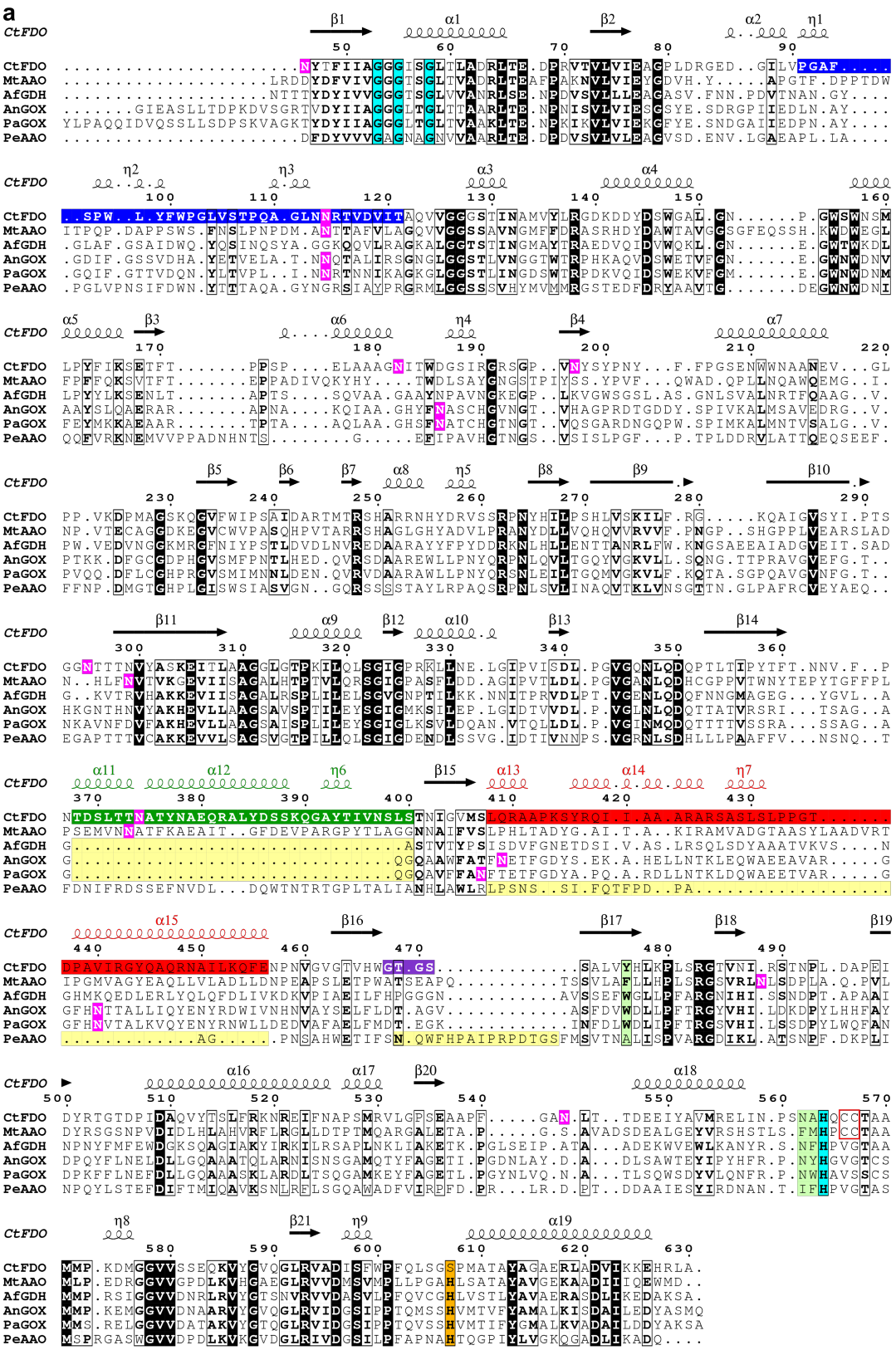

Figure 6.1: (*the legend on next page*)

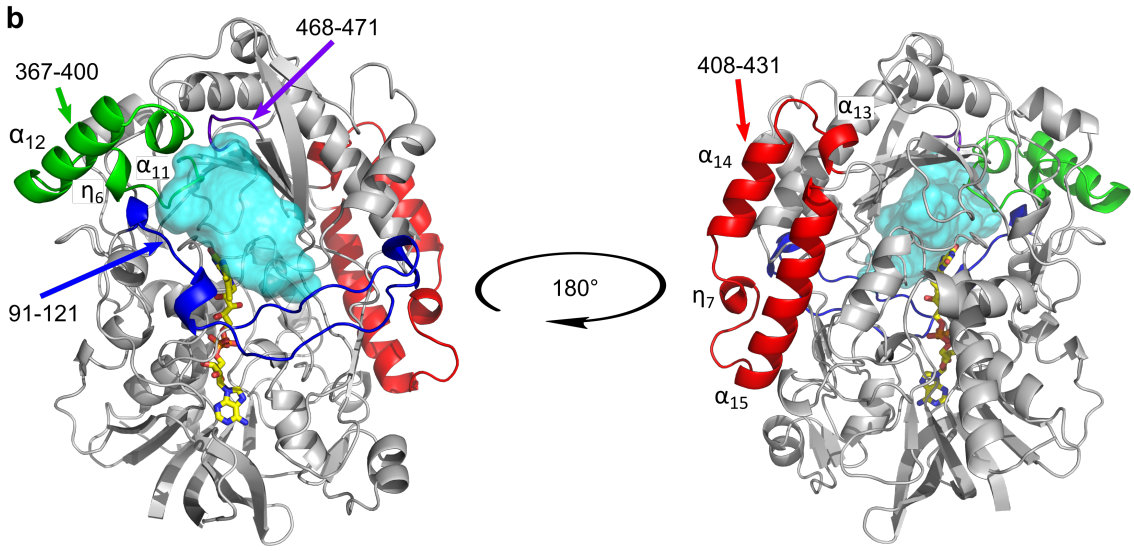

Figure 6.1 *(previous page)*: Structure-based sequence alignment of *Ct*FDO with *Mt*AAO, *Af*GDH, *An*GOX, *Pa*GOX, and *Pe*AAO and comparison of their structural elements. (a) Structure-based sequence alignment of *Ct*FDO:free (PDB id: 6ZE2, chain A) with *Mt*AAO (6O9N), *Af*GDH (4YNT), *An*GOX (1CF3), *Pa*GOX (1GPE, chain A), and *Pe*AAO (3FIM) performed with PDBeFold [121]. Colour-coding in columns: The black and cyan background show invariant residues of compared proteins and the conserved motifs characteristic for the GMC oxidoreductases (Gly-X-Gly-X-X-Gly sequence motif and the conserved histidine), respectively. The second residue in the His-His/His–Ser pair is highlighted with an orange background and other residues present in the active site with lime green background. The sequence parts shown with yellow background correspond to the main secondary-structure differences between the compared proteins. Colour-coding for the *Ct*FDO sequence line and figure (b): The extra secondary-structure elements of *Ct*FDO are coloured in green and red. The purple, blue, and green colours mark the secondary-structure elements forming the entrance to the active site in *Ct*FDO. The N-glycosylation sites in *Ct*FDO and structurally confirmed N-glycosylation sites in compared enzymes are marked by magenta boxes. The *trans* vicinal disulfide in *Ct*FDO and *Mt*AAO is marked by a red rectangle. The graphics was created in ESPript [122]. (b) Crystal structure of *Ct*FDO:free with highlighted structural motifs marked in the structure-based sequence alignment. The active-site pocket is shown in a surface representation (cyan, calculated by HOLLOW [134]. The cofactor is shown as sticks with yellow C atoms. Molecular graphics were created with PyMOL (Schrödinger, LLC).

Surprisingly, *Ct*FDO, *Mt*AAO, and GOXs agree with the position of one glycosylation site (at *Ct*FDO-Asn114, *Mt*AAO-Asn107, *An*GOX-Asn89, and *Pa*GOX-Asn93), which has been proven to contribute to dimer stabilization in GOXs [88]. Dimer formation of *Ct*FDO and *Mt*AAO was not observed. Recombinant *Pe*AAO and *Af*GDH have been prepared nonglycosylated for structural studies [80,88]. Another match is in the presence and position of a *trans* vicinal disulfide formed by two neighbouring cysteine residues in *Ct*FDO (Cys566-Cys567) and *Mt*AAO (Cys581-Cys582) (Fig. 6.2). The disulfide is located near the conserved His564 and likely stabilizes the local protein geometry and interactions [145].

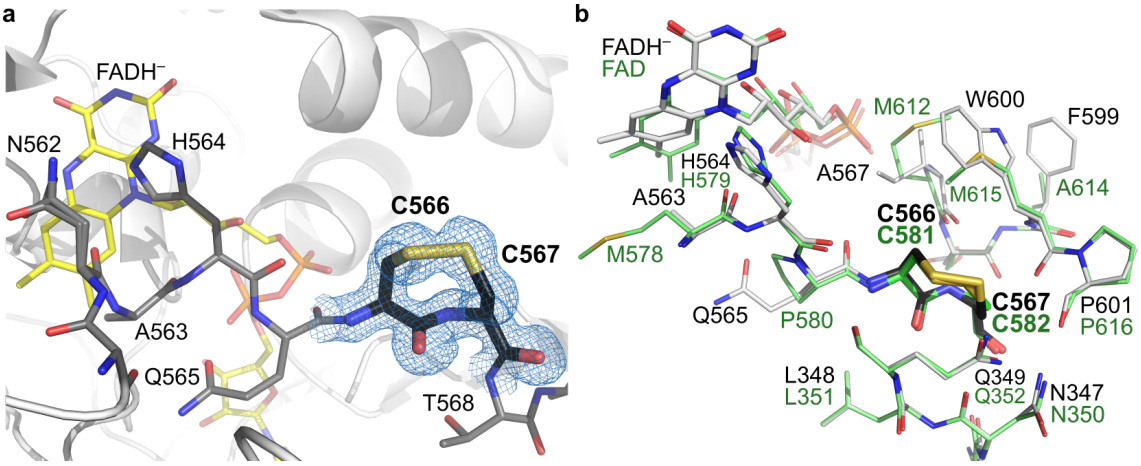

Figure 6.2: A *trans* vicinal disulfide (Cys566-Cys567) in *Ct*FDO. (a) A simulated-annealing $2mF_o$-DF_c composite omit map contoured at 1σ level (blue mesh; Phenix [52]) calculated for the disulfide (C atoms in black). The disulfide surroundings is shown as sticks with light grey C atoms and FAD with yellow C atoms. (b) Superposition of *Ct*FDO:free (C atoms in light grey) and *Mt*AAO (C atoms in palegreen, PDB id: 6O9N). The disulfide in *Ct*FDO:free and *Mt*AAO (Cys581-Cys582) has black and green C atoms, respectively. The molecular graphics were created in Pymol (Schödinger, LLC).

6.2 Features of the active-site pocket

The structural analysis shows that *Ct*FDO has a large active-site pocket wide open to the exterior (Fig. 6.1b), which could possibly accommodate molecules larger than 500 Da as was confirmed by *Ct*FDO:ABTS complex (molecular weight of ABTS is $\sim$ 515 Da). This feature is not common for the members of the GMC oxidoreductase superfamily, where the access to the active site is usually narrowed or completely covered [80, 93, 147–149]. The only exception is the recently published structure of *Mt*AAO, where the wide-open entrance to the catalytic site likely allows the binding and oxidation of both bulky and simple aromatic substrates [82].

The entrance to the active-site pocket of *Ct*FDO is surrounded by three loops: loop of residues 91-121, above mentioned helical insertion α_{11}-α_{12}-η_6 (residues 367-400), and a short loop of residues 468-471 (Fig. 6.1b). For the majority of GMC oxidoreductases, the last loop is much longer and covers the active site. Moreover, it can gate the active site and be partially important for the substrate specificity by taking different conformations during the catalytic cycle as described for *Trametes multicolor* pyranose 2-oxidase (P2O) [150]. This feature for the short loop in *Ct*FDO is excluded. Moreover, the analysis of the *B*-factor distribution in $Ct\text{FDO}_{\text{degl}}$ structures does not indicate considerable motion on the level of secondary structure elements. Therefore, we do not suggest any loop rearrangement related with substrate entry and/or product release.

6.3 Attempts to find substrate for *Ct*FDO using enzymatic assays

Considering that *Ct*FDO originates in cellulose-degrading fungus and its tertiary structure is similar to that of the GMC oxidoreductases, we expected that *Ct*FDO acts on lignocellulosic components. The spectrophotometric observations indicates that *Ct*FDO likely utilizes O_2 as an acceptor of electrons and protons during the oxidative half-reaction. Initial activity tests were performed by Novozymes A/S testing *Ct*FDO with proprietary mixtures of unspecified lignosulfonates. The assay was based on a coupled reaction with horseradish peroxidase and ran at 37 °C overnight. The experiment yielded low response to several mixtures, nevertheless the substrate was not more closely identified. Following tests with similar compounds (lignin components) resulted in no positive hits. The group of tested compounds was extended by substrates of the GMC enzymes and by other various compounds with molecular mass from 70 to 1550 Da (Table A.3, Švecová *et al*. Supplementary Information 2 [119]). Nevertheless, none of these tests led to the identification of any convincing substrate.

The majority of the compounds were tested using ROS-Glo^{TM} H_2O_2 Assay (Promega, worldwide.promega.com) based on luciferin production and its reaction with luciferase. Unfortunately, intermediates in the process of luciferin production contain the benzothiazole unit (Fig. 1.2) and, as we know from the *Ct*FDO:ABTS structure, *Ct*FDO is able to bind benzothiazoline moieties of ABTS. Therefore, it is not excluded that mentioned intermediates inhibit the reaction of *Ct*FDO with tested compounds and, contrariwise, *Ct*FDO inhibits luciferin production and its reaction with luciferase.

6.4 Structure-based identification of binding sites

The crystallographic fragment and compound screening (about 80 molecules, Tables A.2 and A.1) was performed to bring light to the composition of a putative substrate and to analyze binding sites in *Ct*FDO. Altogether, over 130 diffraction data sets with diffraction limit 1.2-3.3 Å were collected, processed, and analyzed. The automated pipelines developed for analysis of multiple datasets for identification of ligand binding could not be used, as their utilization requires series of crystallographic datasets (30-40) of the same crystal system [151] and *Ct*FDO$_{degl}$ crystallizes in three different space groups.

Among 130 solved crystal structures, 13 structures are binding the desired ligand in the active-site pocket of *Ct*FDO. Among them, six unique complexes were identified (some complexes were prepared and measured more times). The remaining 90 % of the structures were without any bound ligand or binding the formic acid molecule in subsite S1. One of them, structure *Ct*FDO:free, was prepared by co-crystallization with cystamine (Table A.2). As no electron density for the ligand was observed and, in addition, cystamine was found to improve the diffraction quality of the *Ct*FDO$_{degl}$ crystal, the structure is presented as the ligand-free structure of *Ct*FDO$_{degl}$.

Based on the *Ct*FDO$_{degl}$ complexes we defined five binding subsites (catalytic site (CS) and subsites S1-S4) inside and in the entrance to the *Ct*FDO active-site pocket. In addition to them, a fifth hypothetical binding subsite (S5) is assumed created by mainly hydrophilic residues Thr368, Asp369, and Gly468-Ser472 peptide at the entrance to the pocket (Fig. 5.7). In the current structures, S5 is filled by water molecules. Different moieties of the bound ligands use the individual subsites differently.

6.5 Catalytic site

6.5.1 Active-site water molecule in GMC oxidoreductases

CS creates a narrowing of the active-site pocket and binds both aromatic (MAMB, pyrrole of IPEA, and benzothiazoline of ABTS) and aliphatic (formic acid, PESB, and IPEA) moieties (Figs. 5.8, 5.9, and 5.10). The ligand-free enzymes from the GMC superfamily usually bind a water molecule in their catalytic site between the FAD isoalloxazine ring and the residues in FAD *re*-face [76, 80–85]. The superposition of ligand-free structures of the GMC enzymes with their complexes with substrates/products/inhibitors shows that the active-site water molecule binds in the close vicinity to the expected position of‘the reactive part of the corresponding substrate in *Pe*AAO (PDB id: 3FIM and 5OC1) [152], pyridoxine 4-oxidase (PNOX, PDB id: 3T37 and 4HA6) [93], *Af*GDH (PDB id: 4YNT and 4YNU) [88], and choline oxidase (PDB id: 2JBV and 4MJW) [149, 153]. This is in agreement with the position in flavoenzymes defined as the site of oxidative attack (see Theoretical background) [79]. Similar features are expected from the water molecule binding in the catalytic site of *Pichia pastoris* alcohol oxidase, FAD-dependent hydroxynitrile lyase from *Prunus amygdalus*, and *Aspergillus oryzae* formate oxidase [76, 83, 85].

The active-site water molecule W$_{AS}$ in *Ct*FDO:free is located in the distance of 3.7 Å (3.8 Å in chain B) to FAD-N^5 atom and makes an angle of $\sim$ 97.5 ° (98.6 ° in chain B) with FAD-N^{10}–FAD-N^5 (Fig. 5.6), i.e. it occupies the position, which meets the definition of the site of oxidative attack in flavoenzymes. W$_{AS}$ in *Ct*FDO$_{degl}$ complexes is often replaced by a ligand – IPEA-C atom in *Ct*FDO:IPEA (Fig. 5.8), ABTS-sulfonate moiety in *Ct*FDO:ABTS (Fig. 5.9), and formic acid in *Ct*FDO:4NC, and *Ct*FDO:4NP (Figs. 5.9 and 5.10). In *Ct*FDO:MAMB and *Ct*FDO:PESB, W$_{AS}$ is shifted by 0.7 Å and 0.2 Å (0.5 Å in chain B) compared to *Ct*FDO:free, respectively, and it mediates hydrogen bonds between the ligand (carbonyl group of MAMB and PESB) and *Ct*FDO. It is likely that W$_{AS}$ mimics the reactive part of a putative substrate of *Ct*FDO.

6.5.2 Residues of the catalytic site

The conserved active-site histidine in the GMC superfamily of oxidoreductases is assumed to play the role of catalytic base initiating the catalytic reaction by abstraction the proton from the hydroxyl group of related substrate in *An*GOX, *Pe*AAO,

5-hydroxymethylfurfural oxidase (HMFO), and others [89–94,98,147,154]. In *Ct*FDO the position of the active-site histidine corresponds to His564 (Fig. 6.1) and it likely plays the same role in *Ct*FDO.

Usually, the conserved histidine is paired with semi-conserved active-site histidine or asparagine residue, creating a His–His or His–Asn pair in the catalytic site of the GMC enzymes. Two different pairs were described so far – a His–Pro and His–Arg pair (see Theoretical background). The semi-conserved residue likely hydrogen bonds the hydroxyl group of related substrates in the right position for catalysis in *Pe*AAO, P2O, PNOX, and others [74, 89, 92, 93, 98].

*Ct*FDO contains a His–Ser (His564–Ser607) pair corresponding to the same position as the His–His/His–Asn/His–Pro/His–Arg pair. Ser607 appears in two conformations in *Ct*FDO: "in" and "out" conformation. The former conformation is involved in binding W_{AS} (*Ct*FDO:free,*Ct*FDO:MAMB, and *Ct*FDO:PESB) and formic acid in the S1 subsite (*Ct*FDO:MAMB, *Ct*FDO:PESB, and *Ct*FDO:IPEA) and in CS (*Ct*FDO:4NP). The "in" conformation of Ser607 excludes binding of small aromatic moieties to subsite S1, as shown in *Ct*FDO:4NC. In all *Ct*FDO structures the "in" conformation was modelled together with its alternative, the "out" conformation. The "in" conformation alone was not observed. On the contrary, the "out" conformation only was modelled in *Ct*FDO:4NC and in *Ct*FDO:ABTS. In *Ct*FDO:4NP, both conformations are observed only due to the reduced occupancy of 4NP in subsite S1.

Ser607 is the first residue at this position, which can switch between two markedly different conformations, as far as we know. We assume that the "in" conformation has a similar role as described for corresponding His and Asn residues. However, as mentioned above, the "in" conformation is not compatible with binding of aromatic moieties to subsite S1. Therefore, it is also possible that the "out" conformer is involved in binding of the first substrate (donor of electrons and protons) to CS with its current binding to the subsite S1 and that the "in" conformer is used to bind the second substrate (electrons acceptor) - oxygen molecule.

*Ct*FDO is not the only protein having the His–Ser pair in the active site. Our search for similar sequences (BLAST [120]) containing the His–Ser active-site pair revealed three additional sequences of GMC superfamily members from fungal species (UniProt ids: A0A175W6J3, B2AMU4, and A0A447BYY0; seq. identity about 50-70 % with *Ct*FDO). Nevertheless, the proteins were not characterized yet.

Another residues of CS worth mentioning are Asn562 and Tyr476 on the FAD *re*-face (Fig. 5.6a). Asn562 corresponds to an asparagine residue in *Af*GDH (Asn503, PDB id: 4YNT), *An*GOX (Asn514, PDB id: 1CF3), and *Pa*GOX (Asn518, PDB id: 1GPE), where it is or is likely to be involved in ligand binding [81, 88]. Tyr476 corresponds to a tyrosine residue in cellobiose dehydrogenase from *Phanerochaete chrysosporium* and *Myricoccum thermophilum* (Tyr609 (PDB id: 1NAA) and Tyr619 (PDB id: 4QI4), respectively), where it probably stabilizes the transition state during the substrate oxidation [155, 156]. In the $Ct\text{FDO}_{\text{degl}}$ complexes, both Asn562 and Tyr476 participate in ligand binding. We suppose that these residues will be involved also in binding and stabilization of the substrate.

6.5.3 Subsite S1

Subsite S1 is a small cavity close to the pyrimidine moiety of FAD and it is formed mainly by hydrophobic residues. The *Ct*FDO complexes show that subsite S1 has capability to bind small inorganic molecules (such as formic acid) and water with Ser607 in "in" conformation, and small aromatic moieties (such as 4NC and 4NP) with the "out" conformation of Ser607. The detailed structural analysis of structural homologues from the GMC superfamily revealed a small cavity at a similar position in HMFO (PDB id: 4UDP). Nevertheless, the cavity binds only water molecules and was not discussed by the authors [147].

6.5.4 Other subsites in *Ct*FDO pocket

Subsite S2 in *Ct*FDO was found to participate in binding of only aromatic moieties (thiophene moiety of PESB). Subsite S3 participates in binding of IPEA (indole part), 4NC (nitro group), and 4NP (hydroxyl group). Subsite S4 binds benzothiazoline of ABTS and it is involved in binding of 4NP (nitro group) and IPEA molecule (indole moiety) from the symmetry-related chain A (IPEA-symA). Analysis of *B*-factors showed higher values for Trp97 belonging to the S4 subsite. In all $Ct\text{FDO}_{\text{degl}}$ structures, the *B*-factor values of Trp97 are higher compared to the residues in the close vicinity and also above the average value of the whole structure (Table 4.2). The higher values indicate certain mobility of Trp97, which was eventually also confirmed by two alternative conformations of Trp97 in *Ct*FDO:ABTS structure (Fig. 5.9c). The first conformation (closest rotamer $\chi_1 = 58\,^\circ$) enables binding of ABTS in the active-site pocket. The second conformer (closest rotamer $\chi_1 = 60\,^\circ$) was observed in the rest of the $Ct\text{FDO}_{\text{degl}}$ structures, where it participates for example in binding of the PEG unit (triethylene glycol) in *Ct*FDO:4NP or IPEA-symA molecule *Ct*FDO:IPEA. Moreover, a peak in electron density (mF_{o}-DF_{c}, not interpreted) was found at the interface of the Trp97 indole moiety (chain A) in *Ct*FDO:4NC, possibly for another 4NC molecule, which could bind to Trp97 via stacking interactions. Thus, we suppose that Trp97 likely serves for binding and stabilization of aromatic moieties of larger substrates, although, the IPEA-symA binding could be a consequence of Trp97 involvement in creation of crystal contacts and the peak in electron density in *Ct*FDO:4NC could be explained as a peak for missing C-terminal residues of the neighbouring chain.

6.6 Putative substrate of *Ct*FDO

The volume of the active-site pocket in *Ct*FDO was determined to be about 2400 Å (HOLLOW [134], 3V server [135]). Assuming that the whole pocket is filled and regarding the density of tested compounds for *Ct*FDO activity (alcohols, carbohydrates, lignin components), the volume roughly corresponds to a molecule with a molecular weight of 1.2-2.5 kDa.

The crystallographic fragment screening resulted in several structurally identified binding subsites in the active-site pocket capable of binding large aromatic compounds, suggesting that the *Ct*FDO substrate is likely a complex polyaromatic compound (Figs. 6.3 and 6.4). This is in line with expectations of a high-molecular weight substrate based on the large volume of the active-site pocket. The substrate could have a lignin-like character, although *Ct*FDO does not show activity against tested lignin components (both small and bulky) in our experiments.

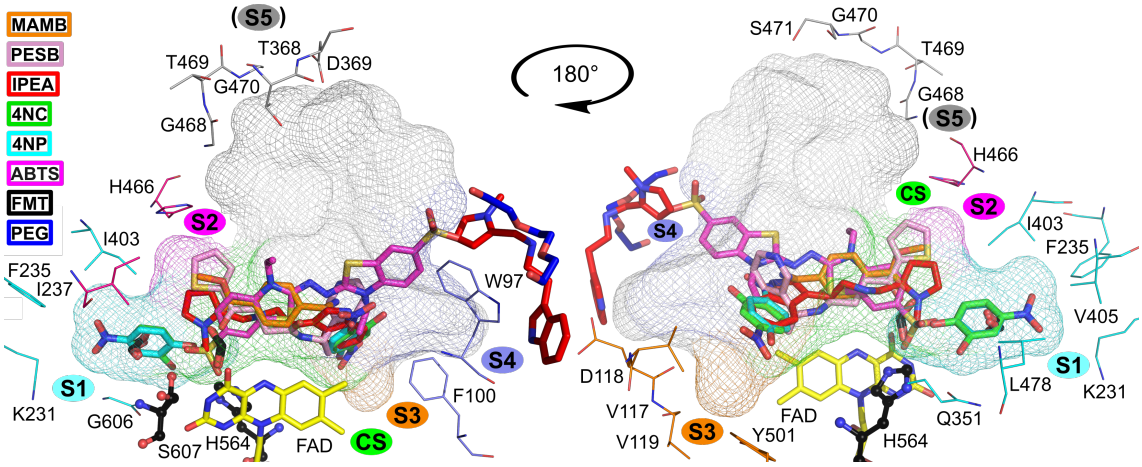

Figure 6.3: Three-dimensional superposition of the active sites of all *Ct*FDO complexes. The superposed ligands are shown in the active-site pocket of *Ct*FDO:free. The pocket (calculated with HOLLOW [134]) is shown in mesh representation. The catalytic site CS is highlighted by green colour and subsites S1 by cyan, S2 by magenta, S3 by orange, S4 by purple, and the hypothetical S5 site by dark grey colour. The selected surrounding residues are shown as lines (the colour of C atoms corresponds to the subsite) and the His–Ser pair (His564 and Ser607) is shown in ball and stick representation (C atoms in black). FAD (C atoms in yellow), MAMB (orange), PESB (pink), IPEA and IPEA-symA (red), 4NC (green), 4NP (cyan), ABTS (magenta), formic acid (FMT, C atoms in black), and triethylene glycol (PEG, C atoms in blue) are shown as sticks. Molecular graphics were created using PyMOL (Schrödinger, LLC).

6.7 FAD state in *Ct*FDO structures

The FAD cofactor in all *Ct*FDO crystal structures was found bent around the FAD-N^5–FAD-N^{10} axis (Fig. 5.6b). This state probably corresponds to the reduced state of FAD, as has been shown by the UV-VIS absorption spectra measured with $CtFDO_{degl}$ crystal before and after irradiation by X-rays (Fig. 5.6c). For all structures of $CtFDO_{degl}$, we expect the $FADH^-$ state of the cofactor. The attempts to preserve the oxidized state of FAD by soaking crystals in solution with higher pH in combination with decreased X-ray dose were unsuccessful. Unfortunatelly even the trials to obtain purposefully reduced state of FAD (chemical reduction of crystals by sodium dithionite), to compare the FAD bending angle with that in discussed $CtFDO_{degl}$ structures were not successful. These trials led to mechanical damage of crystals and to deteriorated diffraction.

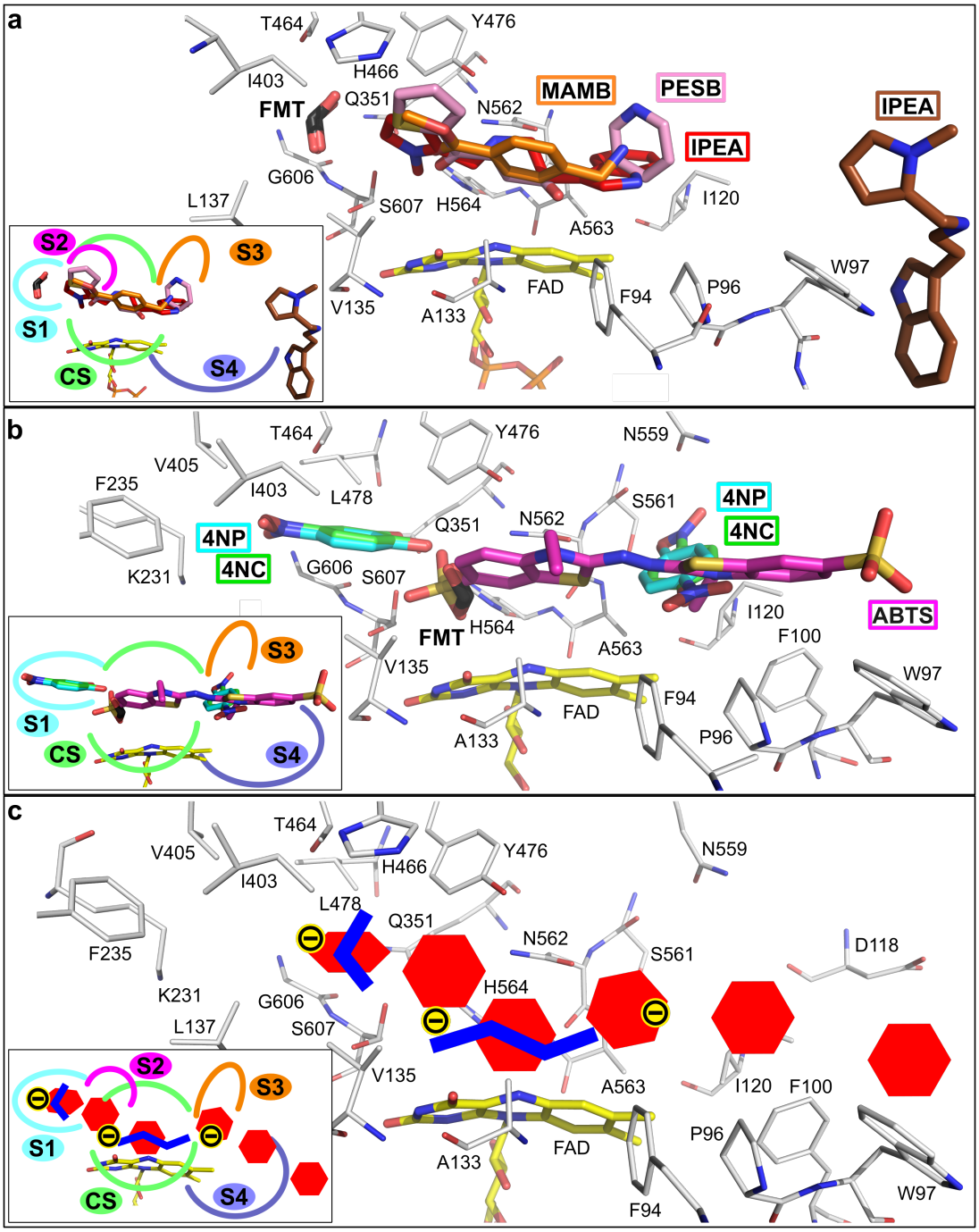

Figure 6.4: Schematic pattern of a putative substrate for *Ct*FDO based on superposition of *Ct*FDO complexes. The superposed ligands are shown in the active-site pocket of *Ct*FDO:free. FAD is shown with C atoms in yellow, formic acid (FMT) in black, and selected residues participating in ligand binding in light grey. (a) Superposition of *Ct*FDO:MAMB (MAMB with orange C atoms), *Ct*FDO:PESB (PESB with pink C atoms), and *Ct*FDO:IPEA (IPEA with red and brown C atoms). (b) Superposition of *Ct*FDO:4NC (4NC with green C atoms), *Ct*FDO:4NP (4NP with cyan C atoms), and *Ct*FDO:ABTS (ABTS with magenta C atoms). (c) Schematic representation of ligands composition. Hexagons, blue lines, and black circles with minus sign represent binding of aromatic rings, aliphatic moieties, and negatively charged moieties, respectively. The pattern indicates polyaromatic composition of a putative substrate. The molecular graphics were created in Pymol (Schödinger, LLC).

6.8 Glycosylation

As mentioned above, *Ct*FDO has the structurally identical position of one glycosylation site (at Asn114) with *Mt*AAO (at Asn107), *An*GOX (at Asn89), and *Pa*GOX (at Asn93) (Fig. 6.1a). The oligosaccharide units modifying Asn89 and Asn93 contribute to stabilization of homodimers of *An*GOX and *Pa*GOX, i.e. the catalytically active forms of *An*GOX and *Pa*GOX (monomeric forms have no catalytic activity) [81, 157]. Although, *Ct*FDO is a monomer (Fig. A.4), considering the missing terminal residues of the protein, it is not excluded that the complete mature protein forms dimers. On the other hand, in terms of primary and tertiary structure similarity, *Ct*FDO structure is more similar to *Mt*AAO, which is catalytically active as monomer, than to *An*GOX and *Pa*GOX [82].

The oligosaccharide moieties of heterologously expressed fully glycosylated *Ct*FDO have overall molecular weight about 20 kDa, i.e. almost 25 % of the molecular weight of the whole *Ct*FDO molecule. It is possible that significant glycosylation of *Ct*FDO may lead to decrease or loss of enzymatic activity as reported previously, for example, for hyperglycosylated glucose oxidase [158].

If any of the above mentioned effects cause any changes to *Ct*FDO activity, the actual catalytic site of *Ct*FDO does not appear to be affected. The ability of FAD to oxidize and reduce was observed and the acceptor of electrons and protons was most likely identified. With respect to the questionable identity of electron and proton donor, it is likely that the true substrate was not in the test group of compounds or the experiment design or reaction conditions were not optimal. There is still some likelihood that the recombinant *Ct*FDO is not capable of binding and/or oxidation of electron donor.

Bilirubin oxidase from *Myrothecium verrucaria*

6.9 Structures of *Mv*BOx in varied pH

The superposition of our *Mv*BOx structure (pH 3.1, PDB id: 6I3J) with those obtained from acidic (pH 5.0, PDB id: 6IQZ) and basic (pH 8.7, PDB id: 2XLL) crystallization conditions shows that the *Mv*BOx structure does not change with pH (Fig. B.1). Negligible structural differences suggest that different pH optima for each type of substrate cannot be associated with structural changes. This is consistent with the findings of Otsuka *et al.*, according to which the specific pH for each substrate is primarily given by the difference in the redox potential between *Mv*BOx and substrates [110].

The presence of Trp396–His398 covalent crosslink in *Mv*BOx was confirmed in the solution (Fig. B.2) and also by the crystal structure obtained from the strongly acidic crystallization condition (reported structure, PDB id: 6I3J). The crosslink was observed also in previously reported structures of *Mv*BOx from acidic (PDB id: 6IQZ) and basic condition (PDB id: 2XLL), as discussed by Atker *et al.* [109]. Based on these results, it can be concluded that the Trp–His crosslink is present in native *Mv*BOx in solution over a wide range of pH values relevant to its function. This is an important fact for the relevance of the mutagenesis–activity studies, which are discussed in this work.

Comparison of *Mv*BOx sequence with sequence homologues identified with BLAST [120] revealed that residues Trp396 and His398 and residues forming their close vicinity are present and conserved in many *Mv*BOx homologues from fungi and bacteria (Fig. 6.5). Preservation of these residues indicates possible presence of the Trp–His crosslink in bilirubin oxidases or closely related enzymes from other organisms.

6.9.1 Coordination of T1Cu in *Mv*BOx structures

The comparison of geometry (lengths and angles) in the close vicinity of T1Cu in *Mv*BOx and other MCOs shows that the Trp396–His398 adduct modifies coordination of T1Cu. A tilt of His398-imidazole moiety coordinating T1Cu is present in our structure of *Mv*BOx-WT and also in the structure of *Mv*BOx-WT published by Akter *et al.* (PDB id: 6IQZ, [109]). It is discernible also in the electron density of *Mv*BOx from the basic condition (PDB id: 2XLL), nevertheless the published model lack this feature (T1Cu is binding in the His398-imidazole plane) probably due to the utilization of standard restrain library. As the tilt is not present in the structure of *Mv*BOx-M467Q mutant, where the adduct is ab-

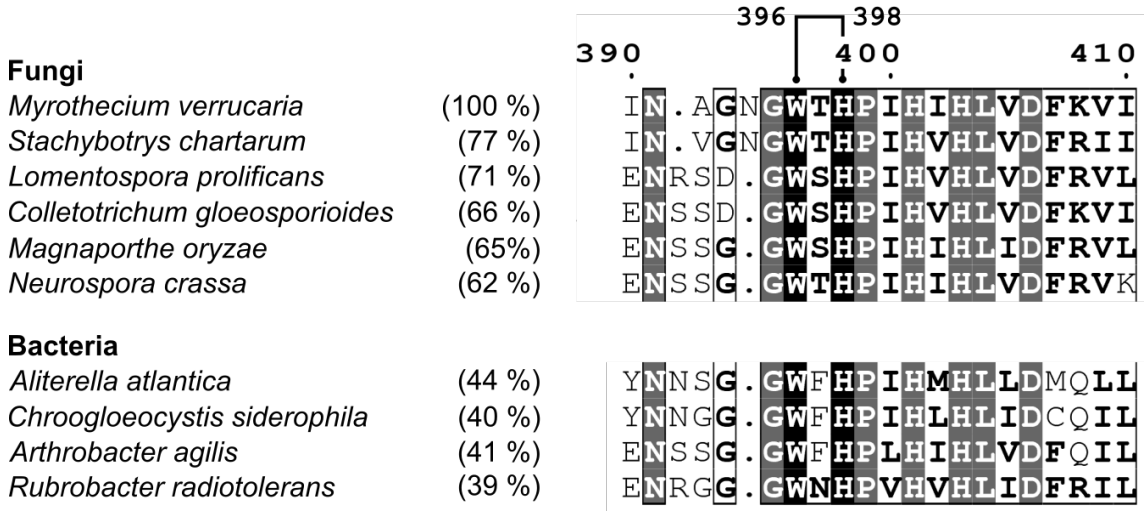

Figure 6.5: Sequence alignment of *Mv*BOx with its homologues from fungi and bacteria with conserved residues Trp396 and His398. Sequences were identified using BLAST [120]. The grey and black background show invariant residues. Sequence identity with *Mv*BOx is shown in parentheses. The graphics was created using ESPript version 3.0 [122].

sent due to the mutation (PDB id: 6IQY, Fig. 2.7c [109]), and it is not observed in the structures W396A:FECN (PDB id: 6I3K) and W396F (PDB id: 6I3L) either, it can be concluded that the T1Cu coordination by the Trp396–His398 adduct cause the His398-imidazole tilt.

The tilt of the His398-imidazole ring is accompanied by a longer distance between T1Cu and the His462-N$^{\delta 1}$ atom in WT:FECN. Longer distance ($\sim$ 2.2 Å) was observed also in the structure from basic condition (PDB id: 2XLL [115]), but not in the structure of *Mv*BOx published by Akter *et al.* [109]. Therefore, the distance between T1Cu and His462-N$^{\delta 1}$ will be a subject of further investigation.

6.9.2 Trp396–His398 crosslink formation

The Trp–His crosslink has been described only in *Mv*BOx wild type as far as we know. The indole moiety of tryptophan is reactive and can undergo irreversible electrochemical oxidation on carbon C$^{\delta 1}$ with an oxidation peak potential 0.64 V (at pH 7) [159]. According to Akter *et al.*, the Trp396–His398 crosslink formation raises the redox potential of the T1Cu, because of the increase of His398 hydrophobicity by covalent link to the hydrophobic Trp398 residue [109].

Our analysis of *Mv*BOx-WT structure excludes other Trp396 standard rotamers in *Mv*BOx-WT, as the observed one is the only acceptable rotamer for this site to avoid any clashes with the surroundings. This applies to both, Trp396 in the structure referred to in this work (PDB id: 6I3J) and also in the *Mv*BOx-M467Q mutant structure without the Trp–His crosslink (PDB id: 6IQY, [109]). It follows that the Trp396 surroundings is in fact in favour of the adduct formation by placing the indole group in close contact with the His398 side chain.

6.9.3 Positively charged site near Trp396–His398 adduct

Previous investigations of the favourable orientation of *Mv*BOx molecules on negatively charged electrode for direct electron transfer revealed that the *Mv*BOx molecules are oriented with the T1Cu site facing to the electrode surface. The surroundings of the T1Cu site is created mainly by basic amino acids which are positively charged under neutral conditions (Fig. 6.6) [160]. Site OS1 is a part of this positively charged surface in *Mv*BOx and it is binding ferricyanide ion in WT:FECN (PDB id: 6I3K) (Fig. 6.6). Therefore, it can be concluded that the crystallographically identified site OS1 is the central site for oxidation of *Mv*BOx substrates. Properties of OS1 must affect substrate binding and catalytic efficiency of *Mv*BOx.

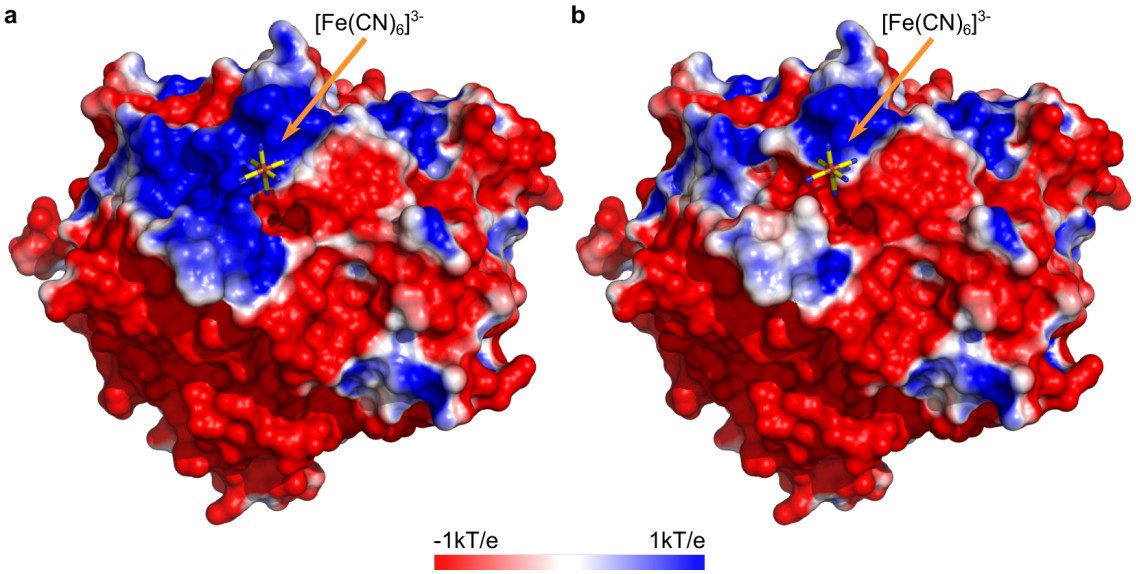

Figure 6.6: Surface electrostatic potential distribution displayed at pH 7 for *Mv*BOx binding ferricyanide ion in the OS1 site. Potential calculated for (a) WT:FECN (PDB id: 6I3J) and (b) model of *Mv*BOx-W396D, which was prepared from the structure of the wild type (PDB id: 6I3J) by mutation of Trp396 to aspartic acid. The Asp396 rotamer was chosen to have no clashes with surrounding residues. Both structures are shown in the same orientation with ferricyanide in the same position in OS1. Ferricyanide ions are labelled and shown as sticks with yellow C atom. The protonation states at pH 7 were assigned with PropKa [131]. The PQR parameter files were prepared with the PDB2PQR pipeline using the AMBER ff99 force field [132]. Surface electrostatic potential distributions were calculated using the Adaptive Poisson-Boltzmann Solver (APBS) [133]. The rendered electrostatic potentials were visualized with the APBS plug-in of the PyMOL software (Schrödinger, LLC).

6.9.4 Structural and functional role of Trp396–His398 adduct in *Mv*BOx

The alignment of *Mv*BOx wild type (WT:FECN), W396A:FECN, and W396F structures showed that the point mutation did not affect the overall structure and neither

the close vicinity of the T1Cu site. The structure of *Mv*BOx-W396D was not determined, as the sample did not crystallize, nevertheless the CD and UV-VIS spectra (Fig. B.3) confirmed that this mutation did not influence the secondary structure composition of *Mv*BOx and the existence of the T1Cu site, either. It follows that the *Mv*BOx structure is not affected by the absence of the Trp396–His398 adduct, nor by the exchange of Trp396 for various residues. Thus the adduct does not appear to have any significant structural role in *Mv*BOx. Trp396 mutation to alanine and phenylalanine does not have significant impact on thermal stability either. In the case of *Mv*BOx-W396D variant a significant decrease of T_{m} (by about 20 °C) was observed. This is likely caused by introducing negative charge to the positively charged site and also by introducing hydrophilic residue (aspartic acid) instead of hydrophobic one (tryptophan) (Fig. 6.6).

6.9.5 Role of the Trp396–His398 adduct in substrate binding and oxidation

Measurement of enzymatic activity of *Mv*BOx and its mutated variants revealed that, the involvement of the Trp396–His398 adduct in substrate binding and oxidation depends on type of substrate (Table 5.2, Fig. 5.16). The measurement also show that different substrates use the adduct in different ways. The adduct (or likely Trp396) plays a crucial role only in the case of DMP as substrate. At the same time, including negatively charged residue (aspartic acid) resulted in a disability to oxidize any tested substrate as the mutation changed the electrostatic potential distribution in OS1 (Fig. 6.6).

Bilirubin

The mutation of Trp396 resulted in considerable decrease of the catalytic efficiency (V_{max}, $K_{1/2}$) for bilirubin. The *Mv*BOx-W396A variant, i.e. mutant with disposed aromatic moiety allowing access to His398 for solvent, has lower binding affinity for bilirubin compared to the wild type. *Mv*BOx-W396F variant, i.e. mutant with introduced aromatic residue incapable to form a crosslink, has affinity comparable with that of the wild type, but a lower maximal reaction velocity (Table 5.2). Interestingly, both mutated variants exhibit significant allosteric effects (Fig. 5.16). According to Kataoka *et al.* [161], the *Mv*BOx-M467Q variant without the crosslink shows substantial decrease in bilirubin oxidation. This could be explained by the use of low substrate concentration (27 μM), which may be in the concentration range, where the activity is influenced by the allosteric effect as observed for *Mv*BOx-W396A and *Mv*BOx-W396F mutants. Considering the differences in structures and in the enzymatic activity, it can be concluded that the Trp396–His398 adduct is likely involved in both, the bilirubin binding and its oxidation, nevertheless, it is not essential for bilirubin binding.

Ferrocyanide

*Mv*BOx-W396A and *Mv*BOx-W396F variants are comparable with the wild type in ferrocyanide oxidation. Based on the W396A:FECN complex, it is clear that the Trp396 is not crucial for ferricyanide (product) binding (Fig. 5.14). Ferricyanide binds in OS1 of both *Mv*BOx wild type and the*Mv*BOx-W396A mutant, however its distance to T1Cu in W396A:FECN structure is about 2 Å shorter than in WT:FECN. In other words, the substitution of Trp396 by alanine allows ferricyanide to bind closer to T1Cu. Both structures have in common that ferricyanide uses the main chain nitrogen atoms of Asn394 and Gly395 and the side chain nitrogens of Arg356 for binding (Fig. 5.14). Ferricyanide bound in OS1 of W396A:FECN has water-mediated bond with His398 coordinating T1Cu. The water molecule could possibly participate in the electron transfer. Previously, it has been shown by Kataoka *et al*. that also the *Mv*BOx-M467Q mutant has retained activity toward ferrocyanide [161].

The *Mv*BOx-W396D mutant shows almost zero activity toward ferrocyanide (about 1 % of the wild type activity). This may be because the presence of a negatively charged aspartic acid side chain either repels the negatively charged ferrocyanide ion and thus interferes with its binding, or it compensates the partial positive charges of the Asn394 and Gly395 main chain nitrogens.

It seems that the adduct itself is not the key component for ferrocyanide binding and for ferrocyanide oxidation either. The geometry of the OS1 and the electrostatic potential of the site enable binding of the ferrocyanide ion close enough to T1Cu (less than 14 Å, [66]) to facilitate an efficient electron transfer (Fig. 5.14). The adduct likely participates in electron transfer in *Mv*BOx, however, electron transfer can be realized in another way. Possible pathways are through the main chain of Asn394-His398 (part of OS1) or through the water molecule in the case of W396A:FECN.

ABTS

The mutation of Trp396 to alanine and phenylalanine led to lower binding affinity for ABTS compared to the wild type, but it has no significant effect on $V_{\max}$ for both mutants. Thus, it is clear that Trp396–His398 adduct is involved in ABTS binding. The adduct can serve for electron transfer, but the electron pathway could be possibly realized also via a main chain loop forming OS1 (residues 394-398).

DMP

All tested mutants exhibit considerable decrease of catalytic activity toward DMP in comparison with the wild type. The kinetic parameters could not be calculated and the attempts to obtain a complex of *Mv*BOx-WT with DMP were unsuccessful. Therefore, it is not clear, if the decreased activity of the mutants toward DMP is a consequence of low substrate affinity, disruption of the electron transfer path by removal of the crosslink, or by the unfavorable difference in the redox potential between DMP and the individual mutants.

6.9.6 Trinuclear copper cluster in *Mv*BOx and its variants

Soaking the *Mv*BOx-WT and *Mv*BOx-W396A crystals by solution containing the substrate potassium ferrocyanide ($K_4Fe(CN)_6$) caused a change in their colour from blue to colourless, which indicates reduction of T1Cu. The colour change of the crystals was accompanied by a colour change of the soaking solution from pale yellow to deep yellow indicating oxidation of the substrate $[Fe(CN)_6]^{4-}$ to the product ferricyanide ($[Fe(CN)_6]^{3-}$). The resolution of complexes WT:FECN (2.59 Å, PDB id: 6I3J) and W396A:FECN (1.60 Å, 6I3K) do not allow to distinguish between ferrocyanide and ferricyanide ions bound in the structures.

The WT:FECN complex was refined with the product ferricyanide ($[Fe(CN)_6]^{3-}$) despite the fact that the trinuclear copper cluster in WT:FECN is fully reduced (Figs. 2.6b and 6.7a), and thus the observed ligand could be ferrocyanide. W396A:FECN was also refined with ferricyanide ions. The TNC of W396A:FECN was modelled binding a water molecule refined with 0.5 occupancy as it has weak $2mF_o$-DF_c map (no peak in composite omit map, Fig. 6.7b). The state of TNC in W396A:FECN corresponds most likely to a mixture of the fully reduced and the resting oxidized state (Fig. 2.6b).

The W396F structure (PDB id: 6I3L, Fig. 6.7c) was modelled binding an dioxygen molecule in the TNC. It likely corresponds to the fully reduced state of TNC immediately after binding dioxygen molecule [105, 107].

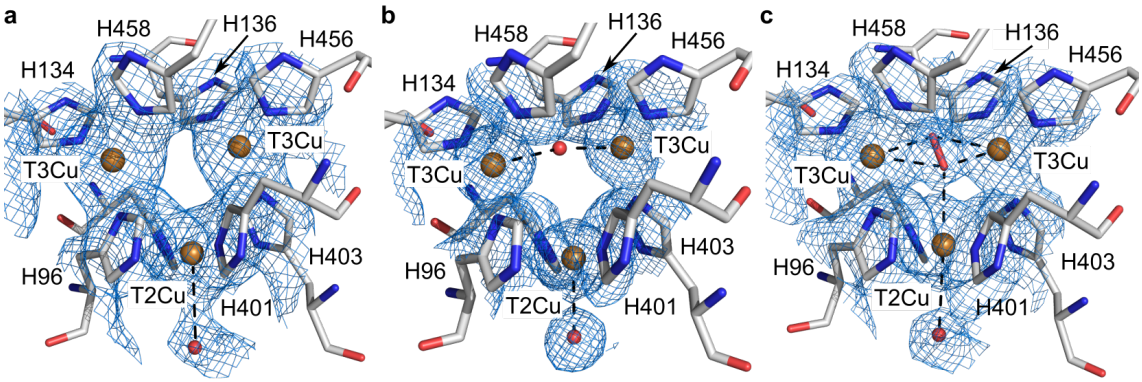

Figure 6.7: Trinuclear copper cluster in structures (a) WT:FECN (PDB id: 6I3J), (b) W396A:FECN (6I3K), and (c) W396F (6I3L). The copper ions are shown as brown spheres with a simulated-annealing $2mF_o$-DF_c composite omit map contoured at 1σ level (blue mesh; calculated using Phenix [52]). The molecular graphics were created in Pymol (Schrödinger, LLC).

Conclusion

FAD-dependent oxidoreductase from *Chaetomium thermophilum*

It was found that *Ct*FDO shares the fold and core structural elements with enzymes from the GMC superfamily of oxidoreductases. Compared to the majority of GMC enzymes, *Ct*FDO has several additional helices, it contains an atypically large wide-open active-site pocket capable to accommodate molecules larger than 500 kDa and which is extended beyond the pyrimidine moiety of FAD to an unusual cavity capable to bind small aromatic molecules. The catalytic site in *Ct*FDO contains a unique His–Ser pair (His564 and Ser607), never observed in related enzymes. Based on the comparison with known structures of GMC oxidoreductases, His564 likely plays the role of general base in *Ct*FDO and residues on the FAD *re*-face (Ser607, Asn562, Tyr476) participate in substrate binding.

The substrate specificity of *Ct*FDO remains unknown. The crystallographic fragments screening performed to identify binding sites in *Ct*FDO yield a series of six *Ct*FDO complexes with ligands. Their structure analysis revealed five binding sites inside the active-site pocket binding primarily aromatic moieties. A complex polyaromatic substrate, likely of a larger lignin-like character, is suggested.

Bilirubin oxidase from *Myrothecium verrucaria*

Myrothecium verrucaria bilirubin oxidase contains a covalent link between the side chains of tryptophan and histidine amino-acid residues, which is unique among structurally characterised enzymes. The formation of this covalent bond between Trp396 and the T1Cu-coordinating His398 is supported by the enzyme fold and local organization of the protein chain and it raises the redox potential of T1Cu.

*Mv*BOx wild type and its variants with Trp396 mutated to alanine, phenylalanine, and aspartic acid were studied and characterized to define the substrate binding site in *Mv*BOx and to evaluate the role of Trp396–His398 adduct in electron pathway from the electron donor to T1Cu ion. First crystal structures of *Mv*BOx wild type in complex with a ligand and structures of new mutated variants *Mv*BOx-W396A and *Mv*BOx-W396F were solved and described. The structure–function analysis of the enzyme samples shows that the mutation of Trp396 affected the enzyme

activity but not the enzyme structure (except the mutated residue). It seems that all tested substrates (bilirubin, ferrocyanide, ABTS, and DMP) must bind in close vicinity to the adduct. The location involved in substrates binding and oxidation – oxidation site 1 – has been defined to be composed of the 393-398 loop including the adduct and of Arg356. The data indicate that the adduct is used by different types of substrates in different ways. *Mv*BOx, even in the absence of the adduct, can oxidize the majority of tested substrates, including bilirubin. Various substrates appear to use the adduct primarily for binding, but the adduct is likely not strictly necessary for electron transfer from the substrate to T1Cu.

Bibliography

[1] Holbrook, R. D., Galyean, A. A., Gorham, J. M., Herzing, A., Pettibone, J. (2015) Overview of Nanomaterial Characterization and Metrology, *Front. Nanosci.* **8**, 47–87.

[2] Ubvan, P. L. (2016) Quantitative mass spectrometry: an overview, *Phil. Trans. R. Soc. A* **374**: 20150382.

[3] Li, Y., Struwe, W. B., Kukura, P. (2020) Single molecule mass photometry of nucleic acids, *Nucleic Acids Res* **48**: 1–47.

[4] Stetefeld, J., McKenna, S. A., Patel, T. R. (2016) Dynamic light scattering: a practical guide and applications in biomedical sciences, *Biophys. Rev.* **8**, 409–427.

[5] Garfin, D. E. (2009) One-Dimensional Gel Electrophoresis, *Methods Enzymol.* **463**, 497–513.

[6] Pergande, M. R., Cologna, S. M. (2017) Isoelectric Point Separations of Peptides and Proteins, *Proteomes* **5**, 4.

[7] Antosiewicz, J. M., Shugar D. (2016) UV–Vis spectroscopy of tyrosine side-groups in studies of protein structure. Part 2: selected applications, *Biophys. Rev.* **8**, 163–177.

[8] Kao, Y.-T., Saxena, C., He, T.-F., Guo, L., Wang, L., Sancar, A., Zhong, D. (2008) Ultrafast dynamics of flavins in five redox states, *J Am Chem Soc.* **130**(39), 13132–13139.

[9] Schmid, F. X. (2001) Biological Macromolecules: UV-visible Spectrophotometry, Encyclopedia of Life Sciences, Introductory Articles *Macmillan Publishers Ltd, Nature Publishing Group*, 1–4.

[10] Greenfield, N. J. (2006) Using circular dichroism spectra to estimate protein secondary structure, *Nat Protoc.* **1**(6), 2876–2890.

[11] Clarke, D. T. (2012) Circular Dichroism in Protein Folding Studies, *Curr Protoc Protein Sci* **70**, 28.3.1–28.3.17.

[12] Alexander, C. G., Wanner, R., Johnson, C. M., Breitsprecher, D., Winter, G., Duhr, S., Baaske, P., Ferguson, N. (2014) Novel microscale approaches for easy, rapid determination of protein stability in academic and commercial settings, *Biochim. Biophys. Acta* **1844**, 2241–2250.

[13] Kodíček, M., Valentová, O., Hynek R. (2018) Biochemie: chemický pohled na biologický svět, 2. vydání, VŠCHT Praha, ISBN 978-80-7592-013-3.

[14] Copeland R. A. (2000) Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis, 2nd edition, Wiley-VCH, Inc., ISBN 0-471-22063-9.

[15] McPherson, A., Gavira J. A. (2014) Introduction to protein crystallization, *Acta Cryst.* **F70**, 2–20.

[16] Nanev, C. N. (2017) Recent Insights into the Crystallization Process; Protein Crystal Nucleation and Growth Peculiarities; Processes in the Presence of Electric Fields, *Crystals* 7, 310, 1–17.

[17] Marek, J., Trávníček, Z. (2002) Monokrystalová rentgenová strukturní analýza, Univerzita Palackého v Olomouci, první vydání, Olomouc, IBSN 80-244-0551-2.

[18] Berman, H. M., Henrick, K., Nakamura, H. (2003) Announcing the worldwide Protein Data Bank, *Nat. Struct. Biol.* **10**: 980.

[19] Weichenberger, C. X., Afonine, P. V., Kantardjieff, K., Rupp, B. (2015) The solvent component of macromolecular crystals, *Acta Cryst.* **D71**, 1023–1038.

[20] Hassell, A. M., An, G., Bledsoe, R. K., Bynum, J. M., Carter III, H. L., Deng, S. J. J., Gampe, R. T., Grisard, T. E., Madauss, K. P., Nolte, R. T., Rocque, W. J., Wang, L., Weaver, K. L., Williams, S. P., Wisely, G. B., Xu, R., Shewchuk, L. M. (2006) Crystallization of protein-ligand complexes, *Acta Cryst.* **D63**, 72–79.

[21] Kirsch, P., Hartman, A. M., Hirsch, A. K. H., Empting, M. (2019) Concepts and Core Principles of Fragment-Based Drug Design, *Molecules* **24**, 4309.

[22] Holton J. M. (2009) A beginner's guide to radiation damage, *J. Synchrotron Rad.* **16**, 133–142.

[23] Nass, K. (2019)Radiation damage in protein crystallography at X-ray free-electron lasers, *Acta Cryst.* **D75**, 211–218.

[24] Pflugrath J. W. (2015) Practical macromolecular cryocrystallography, *Acta Cryst.* **F71**, 622–642.

[25] Murshudov, G. N., Skubak, P. Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., Vagin, A. (2011) REFMAC5 for the refinement of macromolecular crystal structures, *Acta Cryst.* **D67**, 355–367.

[26] Rupp, B. (2009) Biomolecular crystallography: Principles, Practice, and Application to Structural Biology, *New York: Garland Science, Taylor and Francis Group*, 1st edition, ISBN 9780815340812.

[27] International Tables for Crystallography Volume F: Crystallography of Biological Macromolecules, Dordrecht; Boston; London:Published for the International Union of Crystallography, *Kluwer Academic Publishers*.

[28] Taylor, G. (2003) The phase problem, *Acta Cryst.* **D59**, 1881–1890.

[29] Su X. D., Zhang, H., Terwilliger, T. C., Liljas, A., Xiao, J., Dong, Y. (2015) Protein Crystallography from the Perspective of Technology Developments, *Crystallogr. Rev.* **21**(1-2),122–153.

[30] Hemberg, O., Otendal, M., Hertz, H. M. (2003) Liquid-metal-jet anode electron-impact x-ray source, *Appl. Phys. Lett.* **83**(7), 1483–1485.

[31] Shenoy, G. (2003) Basic Characteristics of Synchrotron Radiation, *Structural Chemistry* **14**,3–14.

[32] Neutze, R., Moffat, K. (2012) Time-resolved structural studies at synchrotrons and X-ray free electron lasers: opportunities and challenges, *Curr Opin Struct Biol.* **22**(5), 651–659.

[33] Skarzynski, T. (2013) Collecting data in the home laboratory: evolution of X-ray sources, detectors and working practices, *Acta Cryst.* **D69**, 1283–1288.

[34] Mueller, M., Wang, M., Schulze-Briese, C. (2012) Optimal fine φ-slicing for single-photon-counting pixel detectors, *Acta Cryst.* **D68**, 42–56.

[35] Beis, K., Evans, G. (2018) Protein Crystallography: Challenges and Practical Solutions, Royal Society of Chemistry, ISBN-10: 1782627286.

[36] Powell, H. R. (2017) X-ray data processing, *Biosci Rep*, 1–14.

[37] Powell, H. R. (2017) Integrating macromolecular X-ray diffraction data with the graphical user interface iMOSFLM, *Nat. Protoc.* **12**(7), 1310–1325.

[38] Evans, P. R. (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics, *Acta Cryst.* **D67**, 282–292.

[39] Assmann, G., Brehm, W., Diederichs, K. (2016) Identification of rogue datasets in serial crystallography, *J. Appl. Cryst.* **49**, 1021–1028.

[40] Weiss, M. S. (2001) Global indicators of X-ray data quality, *J. Appl Cryst.* **34**, 130–135.

International Tables for Crystallography. Volume A, Space-Group Symmetry. by Kluwer Academic Publishers, 2002.

[41] Scapin, G. (2013) Molecular replacement then and now, *Acta Cryst.* **D69**, 2266–2275.

[42] Evans, P., McCoy A. (2007) An introduction to molecular replacement, *Acta Cryst.* **D64**, 1–10.

[43] McCoy, A. (2007) Solving structures of protein complexes by molecular replacement with *Phaser*, *Acta Cryst.* **D63**, 32–41.

[44] Vagin, A., Teplyakov, A. (2010) Molecular replacement with *MOLREP*, *Acta Cryst.* **D66**, 22–25.

[45] Vagin, A., Lebedev, A. (2015) MoRDa, an automatic molecular replacement pipeline, *Acta Cryst.* **A71**, s19.

[46] Terwillinger, T. (2004) SOLVE and RESOLVE: automated structure solution, density modification, and model building, *J. Synchrotron Rad.* **11**, 49–52.

[47] Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., Read, R. J., Adams, P. D. (2007) Iterative model building, structure refinement and density modification with the *PHENIX AutoBuild* wizard, *Acta Cryst.* **D64**, 61–69.

[48] Emsley, P., Cowtan, K. (2004) Coot: Model-building tools for molecular graphics, *Acta Cryst.* **D60**, 2126–2132.

[49] Wlodawer, A., Minor, W., Dauter, Z., Jaskolski, M. (2008) Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures, *FEBS J.* **275**(1), 1–21.

[50] Lamb, A. L., Kappock, T. J., Silvaggi, N. R. (2015) You are lost without a map: Navigating the sea of protein structure, *Biochim. Biophys. Acta* **1854**, 258–268.

[51] Afonine, P. V., Moriarty, N. W., Mustyakimov, M., Sobolev, O. V., Terwilliger, T. C., Turk, D., Urzhumtsev, A., Adams, P. D. (2015) FEM: feature-enhanced map, *Acta Cryst.* **D71**, 646–666.

[52] Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Adams, P. D., Read, R. J., Zwart, P. H. Hung, L-W. (2008) Iterative-build OMIT maps: map improvement by iterative model building and refinement without model bias, *Acta Cryst.* **D64**, 515–524.

[53] Kunstleve-Grosse, R.W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine, *Acta Cryst.* **D68**, 352–367.

[54] Hollingsworth, S. A., Karplus, P. A. (2010) A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins, *Biomol Concepts.* **1(3-4)**, 271–283.

[55] Chen, V. B., Arendall 3rd, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., Richardson, D. C. (2010) *MolProbity*: all-atom structure validation for macromolecular crystallography, *Acta Cryst.* **66**, 12–21.

[56] Maheshwari, R., Bharadwaj, G., Bhat, M. K. (2000) Thermophilic fungi: Their physiology end enzymes *Microbiol Mol Biol Rev* **64**(3), 461–488.

[57] Kellner, N., Schwarz, J., Sturm, M., Fernandez-Martinez, J., Griesel, S., Zhang, W., Chait, B. T., Rout, M. P., Kück, U., Hurt, E. (2016) Developing genetic tools to exploit *Chaetomium thermophilum* for biochemical analyses of eukaryotic macromolecular assemblies, *Sci. Rep.* **6**: 20937.

[58] Leath, K. T., Kendall, W. A. (1983) *Myrothecium roridum* and *M. verrucaria* Pathogenic to Roots of Red Clover and Alfalfa, *Plant Dis.* **67**, 1154–1155.

[59] Moreira, F. G., dos Reis, S., Costa, M. A. F., de Souza, C. G. M., Peralta, R. M. (2005) Production of hydrolytic enzymes by the plant pathogenic fungus *Myrothecium verrucaria* in submerged cultures, *Braz J Microbiol* **36**, 7–11.

[60] Machida, M., Yamada, O., Gomi, K. (2008) Genomics of *Aspergillus oryzae*: Learning from the History of Koji Mold and Exploration of Its Future, *DNA Research* **15**, 173–183.

[61] Ntana, F., Mortensen, U. H., Sarazin, C., Figge, R. (2020) Aspergillus: A Powerful Protein Production Platform, *Catalysts* **10**, 1–29.

[62] Torres Pazmiño, D. E., Winkler, M., Glieder, A., Fraaije, M. W. (2010) Monooxygenases as biocatalysts: Classification, mechanism aspects and biotechnological applications, *J. Biotechnol.* **146**, 9–24.

[63] Fridowich, I. (2013) Oxygen: How do we stand it?, *Med Princ Pract* **22**, 131–137.

[64] Augusto, O., Miyamoto, S. (2011) Principles of Free Radical Biomedicine. Volume 1, Chapter II, *Nova Science Publishers, Inc.*, ISBN: 978-1-61209-773-2

[65] Pau, M. Y. M., Lipscomb, J. D., Solomon, E. I. (2007) Substrate activation for O_2 reactions by oxidized metal centers in biology, *PNAS* **104**, 18355–18362.

[66] Moser, C. C., Anderson, J. K., Dutton, P. L. (2010) Guidelines for tunneling in enzymes, *Biochim. Biophys. Acta* **1797**, 1573–1586.

[67] Prabhulkar, S., Tian, H., Wang, X., Zhu, J. J., Li, C. Y. (2012) Engineered Proteins: Redox Properties and Their Applications, *Antioxid Redox Signal* **17**, 1796–1822.

[68] Dijkman, W. P., de Gonzalo, G., Mattevi, A., Fraaije, M. W. (2013) Flavoprotein oxidases: classification and applications, *Appl. Microbiol. Biotechnol.* **97**, 5177–5188.

[69] Zempleni, J., Suttie, J. W., Gregory 3rd, J. F., Stover, P. J.(2013) Handbook of vitamins, *CRC Press, Taylor and Francis Inc., United States*, 5th edition, ISBN 9781466515567.

[70] Kuby, S. A. (1990) A study of Enzymes, Volume II Mechanism of Enzyme Action, *CRC Press*, 1st edition, ISBN 9780849369889.

[71] Bankar, S. B., Bule, M. V., Singhal, R. S., Ananthanarayan, L. (2009) Glucose oxidase — An overview, *Biotechnol. Adv.* **27**(4), 489–501.

[72] Vrielink, A., Ghisla, S. (2009) Cholesterol oxidase: biochemistry and structural features, *FEBS J.* **276**, 6826–6843.

[73] Gadda, G. (2020) Choline oxidases, *The Enzymes* **47**, 137–166.

[74] Hernández-Ortega, A., Lucas, F., Ferreira, P., Medina, M., Guallar, V., Martínez A. T. (2012) Role of Active Site Histidines in the Two Half-Reactions of the Aryl-Alcohol Oxidase Catalytic Cycle, *Biochemistry* **51**, 6595–6608.

[75] Henriksson, G., Johansson, G.,Pettersson, G. (2000) A critical review of cellobiose dehydrogenases, *J. Biotechnol.* **78**, 93–113.

[76] Dreveny, I., Andryushkova, A. S., Glieder, A., Gruber, K., Kratky, C. (2009) Substrate binding in the FAD-dependent hydroxynitrile lyase from almond provides insight into the mechanism of cyanohydrin formation and explains the absence of dehydrogenation activity, *Biochemistry* **48**, 3370–3377.

[77] Pickl, M., Fuchs, M., Glueck S. M. (2015) The substrate tolerance of alcohol oxidases, *Appl. Microbiol. Biotechnol.* **99**, 6617–6642.

[78] Hanukoglu, I. (2015) Proteopedia: Rossmann Fold: A Beta-Alpha-Beta Fold at Dinucleotide Binding Sites, *Biochemistry and Molecular Biology Education* **43**, 206–209.

[79] Fraaije, M. W., Mattevi, A. (2000) Flavoenzymes:Diverse catalysts with recurrent features, *Trends Biochem. Sci.* **25**, 126–132.

[80] Fernández, I. S., Ruíz-Duẽas, F. J., Santillana, E., Ferreira, P., Martínez, M. J., Martínez, Á. T., Romero, A. (2009) Novel structural features in the GMC family of oxidoreductases revealed by the crystal structure of fungal aryl-alcohol oxidase, *Acta Cryst.* **D65**, 1196–1205.

[81] Wohlfahrt, G., Witt, S., Hendle, J., Schomburg, D., Kalisz, H. M., Hecht, H. J. (1999) 1.8 and 1.9 Å resolution structures of the *Penicillium amagasakiense* and *Aspergillus niger* glucose oxidases as a basis for modelling substrate complexes, *Acta Cryst.* **D55**, 969–977.

[82] Kadowaki, M. A. S., Higasi, P. M. E., de Godoy, M. O., de Araújo, E. A., Godoy, A. S., Prade, R. A., Polikarpov, I. (2020) Enzymatic versatility and thermostability of a new aryl-alcohol oxidase from Thermothelomyces thermophilus M77, *Biochim. Biophys. Acta - Gen. Subj.* **1864**, 129681.

[83] Koch, C., Neumann, P., Valerius, O., Feussner, I., Ficner, R. (2016) Crystal structure of alcohol oxidase from Pichia pastoris, *PLoS One* **11**, 1–17.

[84] Bannwarth, M., Bastian, S, Heckmann-Pohl, D., Giffhorn, F., Schulz, G. E. (2004) Crystal structure of pyranose 2-oxidase from the white-rot fungus *Peniophora sp.*, *Biochemistry* **43**, 11683–11690.

[85] Doubayashi, D., Ootake, T., Maeda, Y., Oki, M., Tokunaga, Y., Sakurai, A., Nagaosa, Y., Mikami, B., Uchida, H. (2011) Formate oxidase, an enzyme of the glucose-methanol-choline oxidoreductase family, has a His–Arg pair and 8-formyl-fad at the catalytic site, *Biosci. Biotechnol. Biochem.* **75**, 1662–1667.

[86] Romero, E., Gadda, G. (2014) Alcohol oxidation by flavoenzymes, *BioMol Concepts* **5**(4), 299–318.

[87] Romero, E., Castellanos, J. R. G., Gadda, G., Fraaije, M. W., Mattevi, A. (2018) Same Substrate, Many Reactions: Oxygen Activation in Flavoenzymes, *Chem. Rev.* **118**, 1742–1769.

[88] Yoshida, H., Sakai, G., Mori, K., Kojima, K., Kamitori, S., Sode, K. (2015) Structural analysis of fungus-derived FAD glucose dehydrogenase, *Sci. Rep.* **5**, 1–13.

[89] Wongnate, T. Chaiyen, P. (2013) The substrate oxidation mechanism of pyranose 2-oxidase and other related enzymes in the glucose-methanol-choline superfamily, *FEBS J.* **280**, 3009–3027.

[90] Smitherman, C., Kunchala Rungsrisuriyachai, K., Germann, M. W., Gadda, G. (2015) Identification of the Catalytic Base for Alcohol Activation in Choline Oxidase, *Biochemistry* **54**, 413–421.

[91] Leskovac, V., Trivić, S., Wohlfahrt, G., Kandrač, J., Peričin, D. (2005) Glucose oxidase from *Aspergillus niger*: the mechanism of action with molecular oxygen, quinones, and one-electron acceptors, *Int. J. Biochem. Cell Biol.* **37**, 731–750.

[92] Graf, M. M. H., Sucharitaku, J., Bren, U., Chu, D. B., Koellensperger, G., Hann, S., Furtmuller, P. G., Obinger, C., Peterbauer, C. K., Oostenbrink, C., Chaiyen, P., Dietmar Haltrich, D. (2015) Reaction of pyranose dehydrogenase from *Agaricus meleagris* with its carbohydrate substrates, *FEBS J.* **282**, 4218–4241.

[93] Mugo, A. N., Kobayashi, J., Yamasaki, T., Mikami, B., Ohnishi, K., Yoshikane, Y., Yagi, T. (2013) Crystal structure of pyridoxine 4-oxidase from *Mesorhizobium loti*, *Biochim. Biophys. Acta* **1834**, 953–963.

[94] Sygmund, C., Santner, P., Krondorfer, I., Peterbauer, C. K., Alcalde, M., Nyanhongo, G. S., Guebitz, G. M. Ludwig, R. (2013) Semi-rational engineering of cellobiose dehydrogenase for improved hydrogen peroxide production, *Microb. Cell Fact.* **12**, 38.

[95] Yu, L.-J., Golden, E., Chen, N., Zhao, Y., Vrielink, A., Karto, A. (2017) Computational insights for the hydride transfer and distinctive roles of key residues in cholesterol oxidase, *Sci. Rep.* **7**, 17265.

[96] Wongnate,† T., Surawatanawong, P., Visitsatthawong, S., Sucharitakul, J., Scrutton, N. S., Chaiyen, P. (2014) Proton-Coupled Electron Transfer and Adduct Configuration Are Important for C4a-Hydroperoxyflavin Formation and Stabilization in a Flavoenzyme, *J. Am. Chem. Soc.* **136**, 241–253.

[97] Sützl, L. Foley, G., Gillam, E. M. J., Bodén, M., Haltrich, D. (2019) The GMC superfamily of oxidoreductases revisited: analysis and evolution of fungal GMC oxidoreductases., *Biotechnol Biofuels* **12**: 118.

[98] Rotsaert, F. A. J., Renganathan, V., Gold, M. H. (2003) Role of the Flavin Domain Residues, His689 and Asn732, in the Catalytic Mechanism of Cellobiose Dehydrogenase from *Phanerochaete Chrysosporium*, *Biochemistry* **42**, 4049–4056.

[99] Solomon, E. I., Sundaram, U. M., Machonkin, T. E. (1996) Multicopper Oxidases and Oxygenases, *Chem. Rev.* **96**, 2563–2605.

[100] Kosman, D. J. (2010) Multicopper oxidases: a workshop on copper coordination chemistry, electron transfer, and metallophysiology, *J. Biol. Inorg. Chem* **15**, 15–28.

[101] Arrigoni, R., Arrigoni. O. (2009) Multicopper oxidases: An innovative approach for oxygen management of aerobic organisms, *Rend. Fis. Acc. Lincei* **21**, 71–80.

[102] Skálová, T., Dohnálek, J., Oestergaard, L. H., Oestergaard, P. H., Kolenko, P., Dušková, J., Štěpánková, A., Hašek, J. (2009) The Structure of the Small Laccase from *Streptomyces coelicolor* Reveals a Link between Laccases and Nitrite Reductases, *J. Mol. Biol.* **385**, 1165–1178.

[103] Durand, F., Kjaergaard, C. H., Suraniti, E., Gounel, S, Hadt, R. G., Solomon, E. I. Mano, N. (2012) Bilirubin Oxidase from *Bacillus pumilus*: A promising enzyme for the elaboration of efficient cathodes in Biofuel cells, *Biosens Bioelectron* **35**(1), 140–146.

[104] Solomon, E. I., Augustine, A. J., Yoon, J. (2008) O_2 Reduction to H_2O by the Multicopper oxidases, *Dalton Trans.* **30**, 3921–3932.

[105] Bento, I., Silva, C. S., Chen, Y., Martins, L. O., Lindley, P. F., Soares, C. M. (2010) Mechanism underlying dioxygen reduction in laccases. Structural and modelling studies focusing on proton transfer, *BMC Struct. Biol.* **10**: 28.

[106] Serrano-Posada, H., Centeno-Leija, S., Rojas-Trejo, S. P., Rodríguez-Almazán, C., Stojanoff, V., Rudino-Piera E. (2015) X-ray-induced catalytic active-site reduction of a multicopper oxidase: structural insights into the proton-relay mechanism and O_2-reduction states, *Acta Cryst.* **D71**, 2396–2411.

[107] Siegbahn, P. E. M. (2020) Theoretical Study of O_2 Reduction and Water Oxidation in Multicopper Oxidases, *J. Phys. Chem. A* **124**, 5849–5855.

[108] Mano, N. (2012) Features and applications of bilirubin oxidases, *Appl Microbiol Biotechnol.* **96**, 301–307.

[109] Akter, M., Tokiwa, T., Shoji, M., Nishikawa, K., Shigeta, Y., Sakurai, T., Higuchi, Y., Kataoka, K., Shibata, N. (2018) Redox potential-dependent formation of an unusual His-Trp bond in bilirubin oxidase, *Chemistry* **24**(68), 18052–18058.

[110] Otsuka, K., Sugihara, T., Tsujino, Y., Osakai, T., Tamiya, E. (2007) Electrochemical consideration on the optimum pH of bilirubin oxidase, *Analytical Biochemistry* **370**, 98–106.

[111] Liu, Y., Huang, J., Zhang, J. (2009) Decolorization and biodegradation of remazol brilliant blue R by bilirubin oxidase, *Biosci. Bioeng.* **108**, 496–500.

[112] Miyata, M. Kano, K., Shirai, O., Kitazumi, Y. (2020) Rapid Fabrication of Nanoporous Gold as a Suitable Platform for the Direct Electron Transfer-type Bioelectrocatalysis of Bilirubin Oxidase, *Electrochemistry* **88**(5), 444–446.

[113] Machler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales N. R., Gwadz, M., Hurwitz, D. I., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, C., Zheng, C., Geer, L. Y., Bryant, S. H. (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures, *Nucleic Acids Res.* **45**(D1), D200–D203.

[114] Ravikiran, B., Mahalakshmi, R. (2014) Unusual post-translational protein modifications: the benefits of sophistication, *RSC Adv.* **4**, 33958–33974.

[115] Cracnell, J. A., McNamara, T. P., Lowe, E. D., Blanford C. F (2011) Bilirubin oxidase from *Myrothecium verrucaria*: X-ray determination of the complete crystal structure and a rational surface modification for enhanced electrocatalytic O_2 reduction, *Dalton Trans.* **40**, 6668–6675.

[116] Mizutami, K., Toyoda, M., Sagara, K., Takahashi, N., Sato, A., Kamitaka, Y., Tsujimura, S., Nakanishi, Y., Sugiura, T., Yamaguchi, S., Kano, K., Mikami, B. (2010) X-ray analysis of bilirubin oxidase from *Myrothecium verrucaria* at 2.3 A resolution using a twinned crystal, *Acta Cryst.* **F66**, 765–770.

[117] Hutchinson, E. G., Thornton, J. M. (1990) HERA–a program to draw schematic diagrams of protein secondary structures, *Proteins* **8**, 203–212.

[118] Hutchinson, E. G., Thornton, J. M. (1996) PROMOTIF–a program to identify and analyze structural motifs in proteins, *Protein Sci.* **5**, 212–220.

[119] Švecová, L. Oestergaard, L. H., Skálová, T., Schnorr, K., Koval', T., Kolenko, P., Stránský, J., Sedlák, D., Dušková, J., Trundová, M., Hašek, J., Dohnálek, J. (2021) Crystallographic fragment screening-based study of novel FAD-dependent oxidoreductase from *Chaetomium thermophilum*, *Acta Cryst.* **D77**.

[120] Boratyn, G. M., Schäffer, A. A., Agarwala, R., Altschul, S. F., Lipman, D. J., Madden, T. L.(2012) Domain enhanced lookup time accelerated BLAST, *Biol. Direct* **7**, 1–14.

[121] Krissinel, E., Henrick, K. (2005) Multiple alignment of protein structures in three dimensions, *Lect. Notes Comput. Sci.* **3695** LNBI, 67–78.

[122] Robert, X., Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server, *Nucleic Acids Res.* **42**, 20–324.

[123] Laskowski, R. A., Watson, J. D., Thornton, J. M. (2005) ProFunc: a server for predicting protein function from 3D structure, *Nucleic Acids Res.* **33**, W89–W93.

[124] Gupta, R., Jung, E., Brunak, S. (2004) NetNGlyc 1.0 server - Prediction of N-glycosylation sites in human proteins., Available online: http://www.cbs.dtu.dk/services/NetNGlyc/ (accessed on 1st February 2021).

[125] Huschmann, F. U., Linnik, J., Sparta, K., Ühlein, M., Wang, X., Metz, A., Schiebel, J., Heine, A., Klebe, G., Weiss, M. S., Mueller, U. (2016) Structures of endothiapepsin-fragment complexes from crystallographic fragment screening using a novel, diverse and affordable 96-compound fragment library *Acta Cryst.* **F72**, 346–355.

[126] Cianci, M., Bourenkov, G., Pompidor, G., Karpics, I., Kallio, J., Bento, I.,Roessle, M., Cipriani, F., Fiedlera, S., Schneidera, T. R. (2017) P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing, *J. Synchrotron Rad.* **24**, 323–332.

[127] Mueller, U., Förster, R., Hellmig, M., Huschmann, F.U., Kastner, A., Malecki, P., Pühringer, S., Röwer, M., Sparta, K., Steffien, M., Ühlein, M., Wilk, P., Weiss, M.S. (2015) The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current status and perspectives, *EPJ Plus* **130**: 141.

[128] Kabsch, W. (2010) XDS, *Acta Cryst.* **D66**, 125–132.

[129] Long, F., Nicholls, R. A., Emsley, P., Gražulis, S., Merkys, A., Vaitkus, A., Murshudov, G.N. (2017) AceDRG: A stereochemical description generator for ligands, *Acta Cryst.* **D73**: 112–122.

[130] Lütteke, T., von der Lieth, C. W.(2004) pdb-care (PDB carbohydrate residue check): a program to support annotation of complex carbohydrate structures in PDB files, *BMC Bioinformatics* **5**: 69.

[131] Sondergaard, C.R., Olsson, H.M., Rostkowski, M., Jensen, J.H. (2011) Improvement treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values., *J. Chem. Theory Comput.* 7: 2284–2295.

[132] Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., Baker, N.A. (2004) PDB2PQR: an automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**: W665–W667.

[133] Baker, N.A, Sept, D., Joseph, S., Holst, M.J., McCammon, J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome., *Proc. Natl. Acad. Sci. USA* **98**: 10037–10041.

[134] Ho, B.K., Gruswitz, F. (2008) HOLLOW: Generating accurate representations of channel and interior surfaces in molecular structures., *BMC Struct. Biol.* **8**: 1–6.

[135] Voss, N.R., Gerstein, M. (2010) 3V: Cavity, channel and cleft volume calculator and extractor. *Nucleic Acids Res.* **38**: 555–562.

[136] Nordkvist, M., Nielsen, P. M., Villadsen, J. (2006) Oxidation of Lactose to Lactobionic Acid by a Microdochium nivale Carbohydrate Oxidase: Kinetics and Operational Stability, *Biotechnol. Bioeng.* **97**, 694–707.

[137] Koval', T., Švecová, L., Østergaard, L. H., Skálová, T., Dušková, J., Hašek, J., Kolenko, P., Fejfarová, K., Stránský, J., Trundová, M., Dohnálek, J. (2019) Role of Trp-His Covalent Link in the Unique Substrate Oxidation Environment of Bilirubin Oxidase, Sci. Rep. **9**: 13700.

[138] Wang, J., Yang, C., Chen, X., Bao, B., Zhang, X., Li, D., Du, X., Shi, R., Yang, J., Zhu, R. (2016). A high effective NADH-ferricyanide dehydrogenase coupled with laccase for NAD^+ regeneration, *Biotechnol. Lett.* **38**: 1315–1320.

[139] Reiss, R., Ihssen, J., Thöny-Meyer, L. (2011) *Bacillus pumilus* laccase: a heat stable enzyme with a wide substrate spectrum, *BMC Biotechnology* **11**: 9.

[140] Sigma-Aldrich bilirubin oxidase assay protocol, `<www.sigmaaldrich.com>`

[141] Durand, F., Kjaergaard, C. H., Suraniti, E., Gounel, S., Hadt, R. G., Dolomon, E. I., Mano, N. (2012) Bilirubin oxidase from *Bacillus pumillus*: A promising enzyme for elaboration of efficient cathodes in Biofuel cells, *Biosens Bioelectron.* **35**, 140–146.

[142] Mueller, U., Darowski, N., Fuchs, M. R., Forster, R., Hellmig, M., Paithankar, K. S., Puhringer, S., Steffien, M., Zocher, G., Weiss, M. S. (2012). Facilities for macromolecular crystallography at the Helmholtz–Zentrum Berlin, *J Synchrotron Radiat.* **19**, 442–449.

[143] Pidlypnyi, N., Uhrner, F., Nieger, M., Drafz, M. H. H., Hubner, E. G., Namyslo, J. C. (2013) N-Heterocyclic Carbene–Betaine Interconversions: Tautomeric Equilibria of Imidazolium-Indolates and Indole-Substituted Imidazol-2-ylidenes, *Eur. J. Org. Chem.* **34**, 7739–7748.

[144] Winn, M. D. (2011) Overview of the CCP4 suite and current developments, *Acta Cryst.* **D67**, 235–242

[145] Richardson, J. S., Videau, L. L., Williams, C. J., Richardson, D. C. (2017) Broad Analysis of Vicinal Disulfides: Occurrences, Conformations with *Cis* or with *Trans* Peptides, and Functional Roles Including Sugar Binding, *J. Mol. Biol* **429**, 1321–1335.

[146] Enguita, F. J., Martins, L. O., Henriques, A. O., Carrondo, M. A. (2003) Crystal structure of a Bacterial Endospore Coat Component., *J. Biol. Chem.* **278**, 19416–19425.

[147] Dijkman, W. P,Binda, C., Fraaije, M. W., Mattevi, A. (2015) Structure-based enzyme tailoring of 5-hydroxymethylfurfural oxidase, *ACS Catal.* **5**, 1833–1839.

[148] Hallberg, B. M., Henriksson, G., Pettersson, G., Vasella, A., Divne, C. (2003) Mechanism of the reductive half-reaction in cellobiose dehydrogenase, *J. Biol. Chem.* **278**, 7160–7166.

[149] Salvi, F., Wang, Y. F., Weber, I. T., Gadda. G. (2014) Structure of choline oxidase in complex with the reaction product glycine betaine, *Acta Cryst.* **70**, 405–413.

[150] Spadiut, O., Tan, T. C., Pisanelli, I., Haltrich, D., Divne, C. (2010) Importance of the gating segment in the substrate-recognition loop of pyranose 2-oxidase, *FEBS J.* **277**, 2892–2909.

[151] Pearce, N. M., Krojer, T., Bradley A. R., Collins, P., Nowak, R. P., Talon, R., Marsden, B. D., Kelm, S., Shi, J., Deane, C. M., von Delft, F. (2017) A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density, *Nat. Commun* **8**, 1–8.

[152] Carro, J., Martínez-Júlvez, M., Medina, M., Martínez, A. T., Ferreira, P. (2017) Protein dynamics promote hybride tunneling in substrate oxidation by aryl-alcohol oxidase, *Phys. Chem. Chem. Phys.* **19**(42), 28666–28675.

[153] Quaye, O., Lountos, G. T., Fan, F., Orville, A. M., Gadda, G. (2008) Role of Glu312 in binding and positioning of the substrate for the hydride transfer reaction in choline oxidase, *Biochemistry* **47**, 243–256.

[154] Hernández-Ortega, A., Ferreira, P., Martínez, A. T. (2012) Fungal aryl-alcohol oxidase: a peroxide-producing flavoenzyme involved in lignin degradation, *Appl. Microbiol. Biotechnol.* **93**, 1395–1410.

[155] Hallberg, B. M., Leitner, C., Haltrich, D., Divne, C. (2004) Crystal structure of the 270 kDa homotetrameric lignin-degrading enzyme pyranose 2-oxidase, *J. Mol. Biol.* **341**(3), 781–796.

[156] Tan, T. C., Kracher, D., Gandini, R., Sygmund, C., Kittl, R., Haltrich, D., Hällberg, B. M., Ludwig, R., Divne, C. (2015) Structural basis for cellobiose dehydrogenase action during oxidative cellulose degradation, *Nat. Commun.* **6**: 7542.

[157] Ye, W. N., Combes, D. (1989) The relationship between the glucose oxidase subunit structure and its thermostability, *Biochim. Biophys. Acta* **999**, 86–93.

[158] Romanos, M. A., Scorer, C. A., Clare, J. J. (1992) Foreign gene expression in yeast: a review, *Yeast* **8**, 423–488.

[159] Enache, T. A., Oliveira-Brett, A. M. (2011) Pathways of Electrochemical Oxidation of Indole Compounds., *Electroanalysis* **23**, 1337–1344.

[160] Xia, H., Kitazumi, Y., Shirai, O., Kano, K. (2016) Enhanced direct electron transfer-type bioelectrocatalysis of bilirubin oxidase on negatively charged aromatic compound-modified carbon electrode, *J. Electroanal. Chem.* **763**, 104–109.

[161] Kataoka, K., Kitagawa, R., Inoue, M., Naruse, D., Sakurai, T., Huang, H. (2005) Point Mutations at the Type I Cu Ligands, Cys457 and Met467, and at the Putative Proton Donor, Asp105, in *Myrothecium Verrucaria* Bilirubin Oxidase and Reactions with Dioxygen., *Biochemistry* **44**: 18, 7004–7012.

[162] Laskowski, R. A., Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery, *J. Chem. Inf. Model.* **51**, 2778–2786.

[163] Heppner, D. E., Kjaergaard, C. H., Solomon, E. I. (2014) Mechanism of the Reduction of the Native Intermediate in the Multicopper Oxidases: Insight into Rapid Intramolecular Electron Transfer in Turnover, *J. Am. Chem. Soc.* **136**, 17788–17801.

Appendix

Appendix A
FAD-dependent oxidoreductase from *Chaetomium thermophilum*

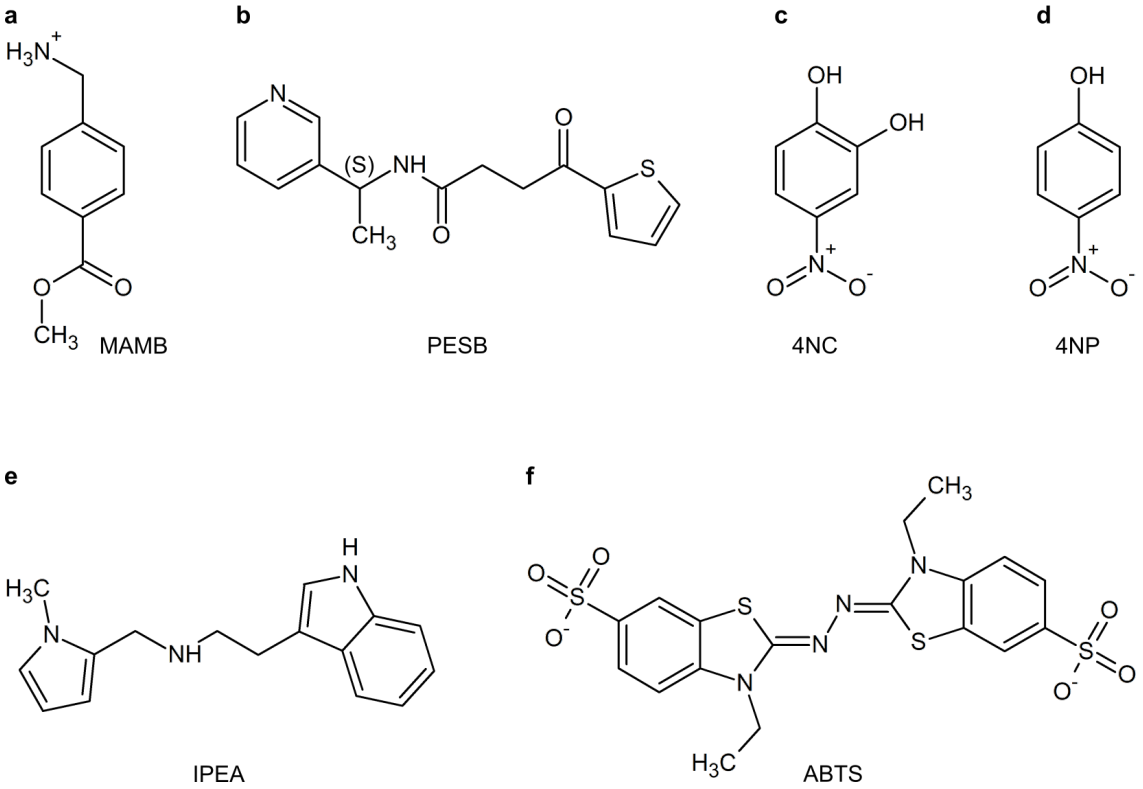

Figure A.1: Structural formulae of ligands binding inside the active-site pocket of *Ct*FDO. (a) methyl 4-(aminomethyl)benzoate (MAMB), (b) 4-oxo-*N*-[1-(3pyridinyl)-ethyl]-2-thiophenebutanamide (PESB), (c) 2-(1*H*-indol-3-yl)-*N*-[(1-methyl-1*H*-pyrrol-2-yl)-methyl]-ethan-amine (IPEA), (d) 4-nitrocatechol (4NC), (e) 4-nitrophenol (4NP), and (f) 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS).

Table A.1: Overview of compounds used for *Ct*FDO$_{degl}$ complexes preparation by soaking. The diffraction data were collected at beamline BL P13 (Petra III, DESY, Hamburg, Germany), BL 14.1, BL 14.2, and BL 14.3 (Bessy II, Helmholz Zentrum, Berlin, Germany), and using Bruker D8 Venture diffractometer with Excillum Metaljet D2 X-ray source (MJ). FXS-no. marks a fragment (50 nmol each) from Frag Xtal Screen (Jena Bioscience). The complexes are indicated by the complex names and by corresponding PDB ids. The DET abbreviation means deteriorated diffraction.

Compound	concentration; soaking time	No. of data sets	beamline	diffraction limit (Å)	ligand localized
Guaiacylglycerol-beta-guaiacyl ether	10 mM; 1 h	1	P13	2.22	no
o-phenylenediamine	10 mM; 4-6 min	2	14.2	2.00-2.25	no
Aminophenol	10 mM; 5 min	1	14.2	2.50	no
3-phenyl-1-propanol	5 mM; 3 s	1	P13	2.10	no
Gallic acid	3.5 mM; 3 s	1	P13	3.80	no
$K_4[(CN)_6]$	20 mM; 2-4 min	2	P13	1.55-2.25	no
Sodium dithionite	10 mM; 10 s	0	14.2	–	DET
5-hydroxymethyl furfural	10 mM; 2-4 min	1	14.2	2.30	no
Tetracycline	1 mM; 2-4 min	0	14.2	–	DET
Fluvastatin	1.3 mM; 1-2 min	3	14.2	2.00-3.20	no
Eltrombopag	0.5 mM; 1 min	2	14.2	2.30	no
AZD-8055	0.5 mM; 1 min	1	14.2	1.90	no
Triamcinolone	1.3 mM; 2 min	1	14.2	2.30	no
FXS-3	70 h	2	P13	1.55-1.95	no
FXS-6	3.5, 24 h	2	14.2	2.00	no
FXS-7	3 h	1	14.1	2.30	no
FXS-9	20 h	1	P13	2.65	no
FXS-10	17 h	4	P13	1.72-2.00	no
FXS-12	4, 70 h	0	P13, 14.1	–	DET
FXS-16	2.5, 22, 70 h	1	P13, 14.2	1.93	no
FXS-18	70 h	2	P13	2.10	no
FXS-24	3, 4, 17 h	5	P13, 14.1	1.60-2.28	*Ct*FDO:PESB (6ZE4)
FXS-28	3, 21 h	2	P13, 14.2	1.67-1.76	no
FXS-29	70 h	2	P13	2.75-3.86	no
FXS-32	2.5, 22 h	1	P13, 14.2	2.22	*Ct*FDO:MAMB (6ZE3)
FXS-35	70 h	2	P13	1.80-1.82	*Ct*FDO:IPEA (6ZE5)
FXS-36	3, 22 h	0	P13, 14.2	–	DET
FXS-38	2.5, 24 h	0	14.2	–	DET
FXS-39	21 h	2	P13	1.72-2.80	no
FXS-41	2.5, 24 h	0	P13, 14.2	–	DET
FXS-42	3, 21 h	2	P13, 14.2	1.85-2.00	no
FXS-43	2.5, 24 h	0	14.2	–	DET
FXS-47	21 h	0	P13	–	DET
FXS-48	17 h	2	14.1	1.75-1.87	no
FXS-49	21 h	2	P13	2.60-3.00	no
FXS-52	21 h	1	P13	1.75	no
FXS-57	21 h	2	P13	1.88-2.30	no
FXS-59	21 h	0	P13	–	DET
FXS-62	21 h	2	P13	1.60-2.00	no
FXS-65	4, 24 h	2	P13, 14.2, MJ	1.65-2.20	no
FXS-67	4, 24 h	3	14.2	1.43	no
FXS-69	21 h	1	P13	2.55	no
FXS-70	4, 24 h	2	P13	1.35-1.52	no
FXS-72	4, 24 h	3	P13, 14.2	1.38-1.97	no
FXS-74	17 h	2	14.1	1.65-2.55	no

Continued on next page

Table A.1 – continued from previous page

Compound	soaking time	No. of data sets	beamline	diffraction limit (Å)	ligand localized
FXS-77	4, 24 h	2	P13, 14.2	1.42-1.70	no
FXS-78	20 h	1	P13	1.50	no
FXS-80	17 h	2	14.1	1.90–2.30	no
FXS-81	21 h	0	P13	-	DET
FXS-82	17, 20 h	4	14.1, 14.2	1.26-1.95	*Ct*FDO:4NC (6ZE6)
FXS-88	21 h	0	P13	–	DET
FXS-90	17 h	4	14.1	1.60-1.85	no
FXS-91	3.5, 24 h	3	P13, 14.2	1.24-1.97	no
FXS-92	17 h	2	14.1	1.50-1.92	no
FXS-94	20 h	1	P13	2.85	no

Table A.2: Overview of compounds used for $Ct\text{FDO}_{\text{degl}}$ complexes preparation by co-crystallization. The diffraction data were collected at beamlines BL P13 (Petra III, DESY, Hamburg, Germany) and BL 14.1, BL 14.2, and BL 14.3 (Bessy II, Helmholz Zentrum, Berlin, Germany). The complexes are indicated by the complex names and by corresponding PDB ids. The DET abbreviation means deteriorated diffraction.

Compound	concentration in reservoir solution	No. of data sets	beamline	diffraction limit (Å)	ligand localized
Cellobiose	6-22 mM	0	14.3	–	DET
D-(+)-galactose	10-15 mM	2	14.1, 14.3	2.10-3.00	no
Glucose	10 mM	1	14.3	1.80	no
Lactose	15 mM	1	14.1	2.60	no
Rhamnose	6-10 mM	0	14.3	–	DET
Raffinose	5 mM	1	14.3	2.40	no
Sucrose	10 mM	1	14.3	2.00	no
Maltose	15 mM	2	P13	1.60-1.70	no
Na-L-ascorbic acid	20 mM	3	14.2	2.8-5	no
D-sorbitol	20 mM	0	14.1	–	DET
Caffeic acid	18 mM	2	P13	2.40	no
Cinnamyl alcohol	22 mM in 0.6 % ethanol	2	P13, 14.1	2.00-2.50	no
p-coumaric acid	20 mM	8	P13, 14.1	1.50-2.60	no
4-methoxybenzyl alcohol	25 mM in 10 % ethanol	3	P13	1.70-1.80	no
4-hydroxy-3-methoxybenzyl alcohol	15 mM	1	P13	1.11	no
crotyl alcohol	20 mM	2	14.1	2.60	no
choline	20 mM	2	14.1, 14.2	1.85-2.30	no
4-nitrophenol	20 mM	2	14.1	1.6	*Ct*FDO:4NP (6ZE7)
2-phenylethanol	20 mM	5	14.1	2.00-3.70	no
ellagic acid	5 mM	0	P13	–	DET
2,6-dichloroindophenol	0.15 mM	1	14.1	1.40	no
$K_4[(CN)_6]$	20 mM	1	14.1	2.00	no
Adenosine monophosphate	20 mM	1	14.2	3.50	no
Nicotinamide adenine dinucleotide	60 mM	2	14.1	1.90	no
Cystamine	20 mM	8	P13, 14.2	1.35-2.20	*Ct*FDO:free (6ZE2)
ABTS	8.5 mM	1	P13	1.40	*Ct*FDO:ABTS (7AA2)

Table A.3: Overview of compounds tested for *Ct*FDO activity. [a]The reaction ran in Britton-Robinson buffer with 0.25 μg of *Ct*FDO at 45 °C for 20 min. [b]The reaction ran in 100 mM phosphate buffer with 30 mM NaCl and 2 μg of *Ct*FDO at 37 °C for 20 min. [c]The reaction as in [a] ran with 0.5 μg of *Ct*FDO. The reaction mixture with HRP was not added. [d]The reaction as in [b] ran at 45 °C.

Compound	concentration in solution	pH
Glucose[b]	5 mM	4.5, 6.0, 8.0
L-Rhamnose[b]	5 mM	4.5, 6.0, 8.0
Cellobiose[b]	5 mM	4.5, 6.0, 8.0
Lactose[b]	5 mM	4.5, 6.0, 8.0
1,4-butanediol[a]	5 mM	3.5, 5.4, 7.1, 8.0
Crotyl alcohol[b]	100 mM	5.0, 6.0, 8.0
Phenol[b]	5 mM	5.0, 6.0, 8.0
Catechol[a]	5 mM	3.5, 5.8, 6.7, 7.7, 8.9
Hydroquinone[b]	5 mM	5.0, 6.0, 8.0
4-nitrophenol[b]	3.8 mM	5.0, 6.0, 8.0
Benzyl alcohol[a]	5 mM	3.5, 5.4, 7.1, 8.0
4-methoxybenzyl alcohol[a]	5 mM	3.5, 5.4, 7.1, 8.0
4-methoxybenzyl alcohol[b]	50 mM	4.5, 5.0, 6.0, 7.0, 8.0, 9.0
4-hydroxy-3-methoxybenzyl alcohol[a]	5 mM	3.5, 5.4, 7.1, 8.0
Cinnamyl alcohol[b]	5 mM	5.0, 6.0, 8.0
p-coumaryl alcohol[b]	9.5 mM	4.5, 6.0, 8.0
Coniferyl alcohol[a]	5 mM	3.5, 5.4, 7.1, 8.0
2-phenyl ethanol[a]	4.4 mM	5.0, 6.0, 8.0
3-phenyl-1-propanol[a]	5 mM	3.5, 5.4, 7.1, 8.0
2,6-dimethoxyphenol[a]	5 mM	3.5, 5.8, 6.7, 7.7, 8.9
4-anisaldehyd[a]	4.3 mM	3.5, 5.4, 7.1, 8.0
Pyriridoxine[a]	5 mM	3.7, 5.4, 7.1, 7.4, 8.0
Choline[a]	5 mM	3.5, 5.4, 7.1, 8.0
5-(hydroxymethyl)furfural[a]	5 mM	3.5, 5.8, 6.7, 7.7, 8.9
Guiacyl-glycerol-guiacyl ether[a]	2.4 mM	3.5, 5.4, 7.1, 8.0
Tannic acid[b]	5 mM	5.0, 7.0
Bilirubin[b]	0.2 mM	8.0
Dithiotreitol[b]	5 mM	5.0, 8.0
ABTS[c]	5 mM	4.5, 5.4, 6.1, 7.1, 8.0, 8.9
Tetracycline[d]	27 mM	4.5, 5.0, 6.0, 7.0, 8.0

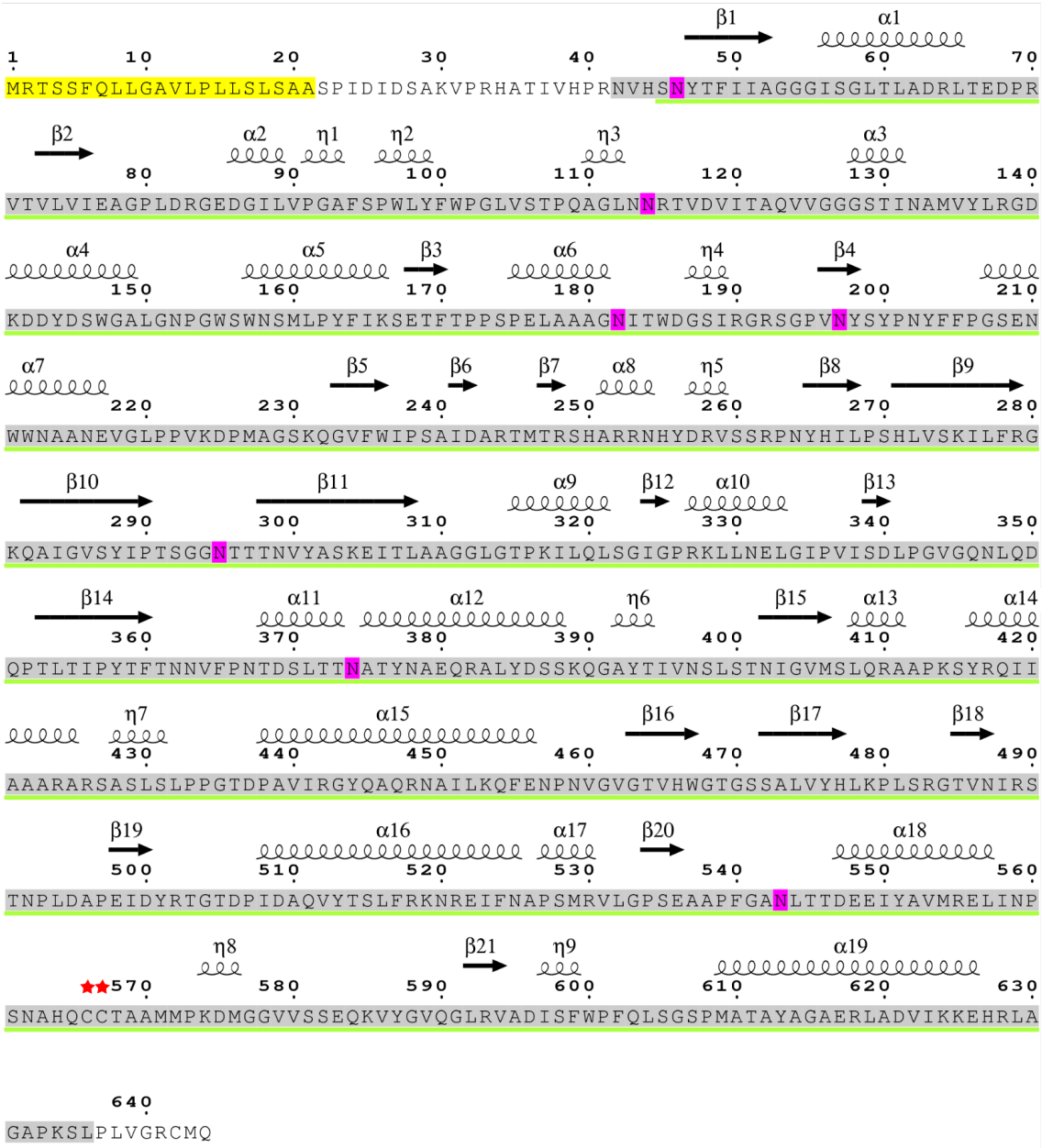

Figure A.2: Complete expected mature sequence of *Ct*FDO including the signal peptide. The sequence part confirmed by LC-MS/MS is shown with grey background (residues 42-636) and confirmed by *Ct*FDO$_{degl}$ structures (residues 45-630) is underlined in green colour. The signal peptide is shown with a yellow background. Residues with a hot pink background correspond to the N-glycosylation sites. Red stars mark the vicinal disulfide bridge in *Ct*FDO. The graphics was created in ESPript version 3.0 [122].

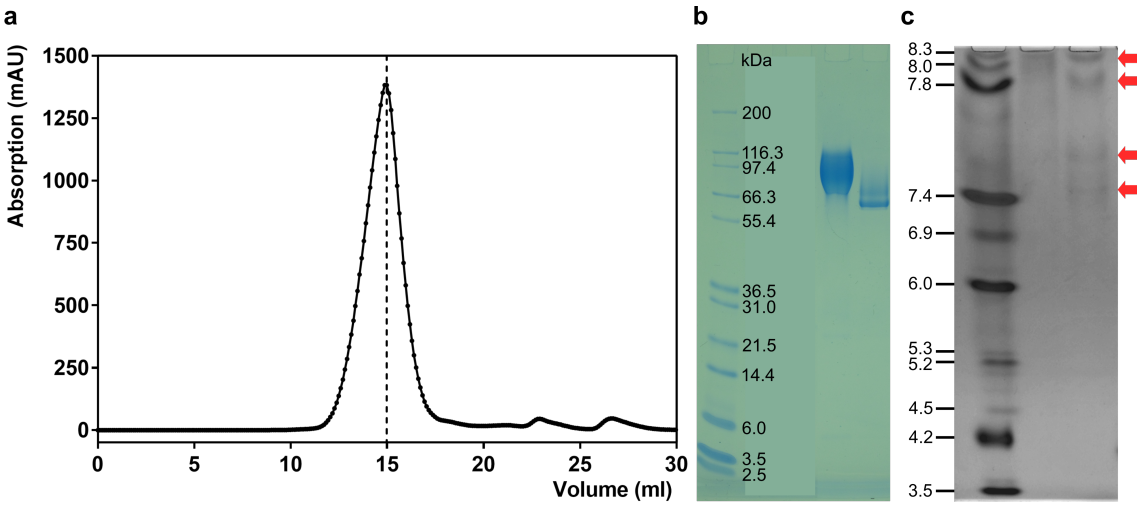

Figure A.3: Control of the purity and deglycosylation of *Ct*FDO and determination of isoelectric point (pI) of *Ct*FDO and *Ct*FDO$_{degl}$. (a) Size-exclusion chromatography of *Ct*FDO performed using ÄKTA purifier chromatographic system and Superdex 200 10/300 GL gel filtration column (GE Healthcare). The absorbance was measured at 280 nm. The peak with an elution volume of 15 ml corresponds to a molecular weight of 85 kDa. (b) Coomassie Brilliant Blue R-250 (ThermoFisher Scientific) stained SDS-PAGE Gel (4-12 % NuPAGE Bis-Tris SDS gradient gel (ThermoFisher Scientific)) of *Ct*FDO (second lane) and *Ct*FDO$_{degl}$ (third lane). The shift in molecular weight indicates successful deglycosylation of *Ct*FDO. The marker (first lane) is Mark12 Unstained Standard (ThermoFisher Scientific). (c) InstantBlue (Expedeon) stained IEF gel (Novex 5 % polyacrylamide gel with fixed vertical pH 3-10 gradient (ThermoFisher Scientific)) of *Ct*FDO (second lane) and *Ct*FDO$_{degl}$ (third lane). The marker (first column) is IEF Marker 3-10 (ThermoFisher Scientific). IEF gel shows a smear between pI 8.0 and 8.3 for *Ct*FDO and four bands between pI 7.0 and 8.3 for *Ct*FDO$_{degl}$.

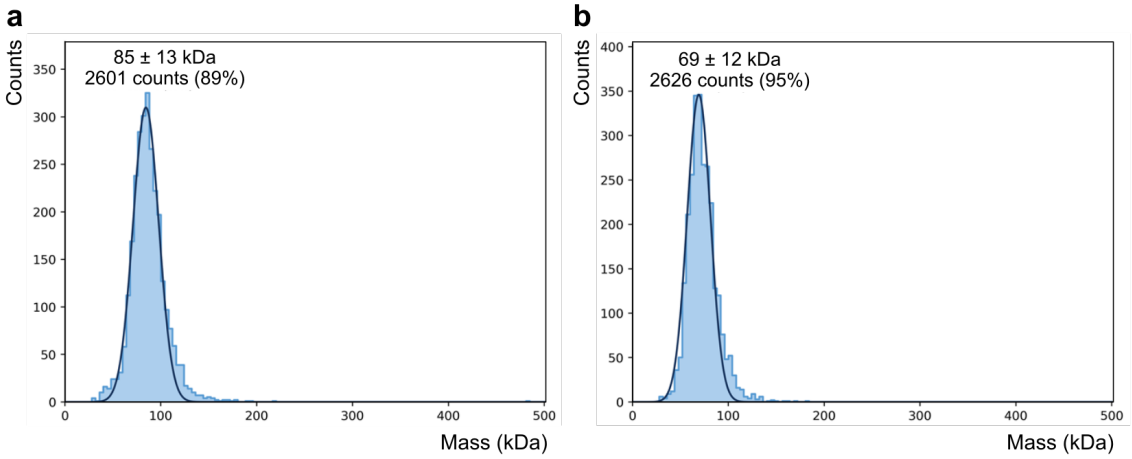

Figure A.4: Mass photometry histograms and kernel density estimates for (a) *Ct*FDO and (b) *Ct*FDO$_{degl}$. Both histograms show a single peak for *Ct*FDO and *Ct*FDO$_{degl}$ corresponding to 85 and 69 kDa containing 89 % and 95 % of population, respectively. The data were collected on a Refeyn OneMP instrument using the DiscoverMP (v2.2.1) software (Refeyn Ltd.).

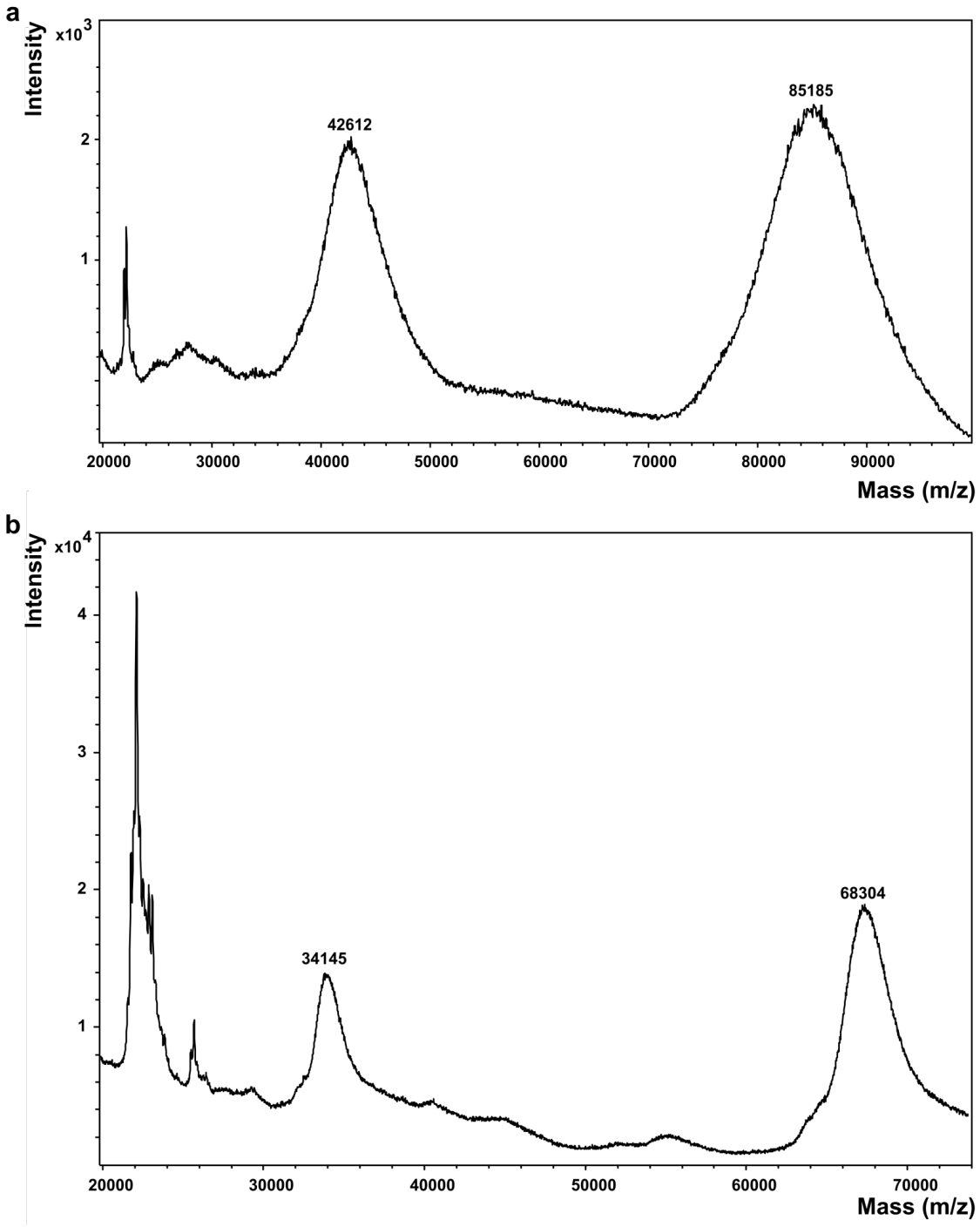

Figure A.5: Mass spectra (MALDI-TOF) of (a) fully glycosylated and (b) deglycosylated form of *Ct*FDO. Molecular mass was determined to be about 85 and 68 kDa for *Ct*FDO and $Ct\text{FDO}_{\text{degl}}$, respectively. A wide peak around 85 kDa in the spectra for *Ct*FDO shows possible heterogeneity of glycosylation. The samples were analysed using a MALDI-TOF mass spectrometer Autoflex Speed (Bruker Daltonics) operated in linear positive mode. Data were processed with the FlexAnalysis (version 3.3) software (Bruker Daltonics).

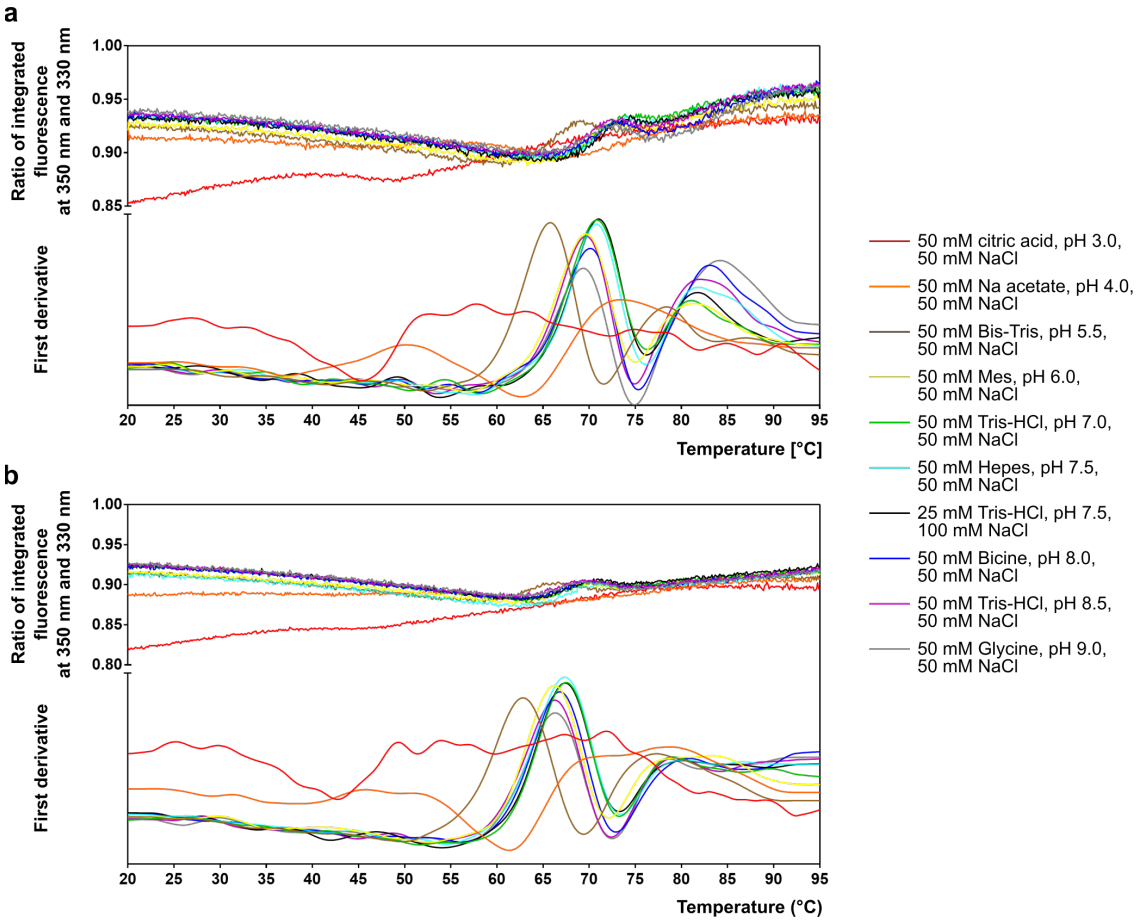

Figure A.6: NanoDSF spectra of *Ct*FDO and *Ct*FDO$_{degl}$ in various buffers (pH 3.0-9.0). Differences in the heights of peaks are given by mild differences in concentration of the samples. The spectra were measured with Prometheus NT.48 (NanoTemper) using the PR.ThermControl software (v 2.1.1) and plotted with the GraphPad Prism (version 7.02) for Windows (GraphPad Software, San Diego, California USA, www.graphpad.com).

Table A.4: Melting temperatures T_{m1} and T_{m2} of *Ct*FDO and *Ct*FDO$_{degl}$ in various buffers (pH 3.0-9.0) measured using nanoDSF. All buffers were in concentration 50 mM and contained 50 mM NaCl, except for 25 mM Tris-HCl pH 7.5 with 100 mM NaCl.

		*Ct*FDO		*Ct*FDO$_{degl}$	
	Buffer	T_{m1}	T_{m2}	T_{m1}	T_{m2}
●	citric acid, pH 3.0	protein precipitation		protein precipitation	
●	Na acetate, pH 4.0	50.2	73.8	ambiguous	
●	Bis-Tris, pH 5.5	65.6	78.4	62.7	77.7
●	Mes, pH 6.0	69.4	81.6	66.1	ambiguous
●	Tris-HCl, pH 7.0	70.7	81.4	67.4	79.1
●	Hepes, pH 7.5	70.7	82.6	67.3	80.4
●	Tris-HCl, pH 7.5	70.8	81.9	67.3	80.1
●	Bicine, pH 8.0	69.8	83.4	66.6	80.6
●	Tris-HCl, pH 8.5	69.4	82.6	66.1	79.4
●	Glycine, pH 9.0	69.1	84.4	66.2	78.5

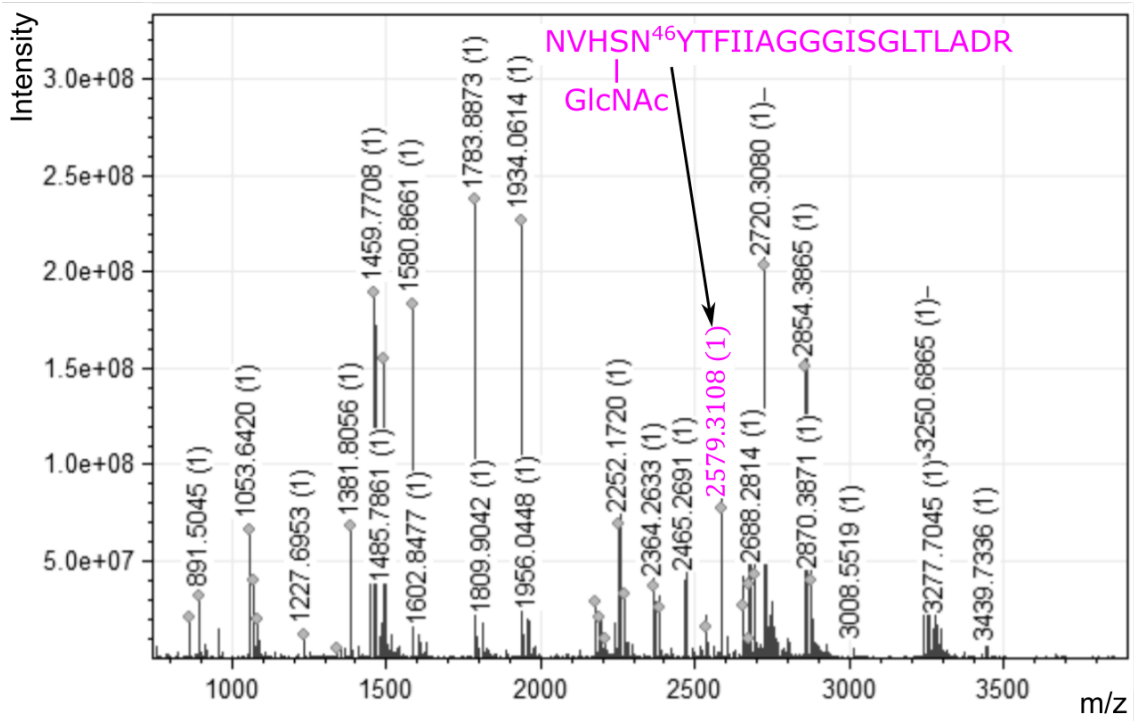

Figure A.7: *Ct*FDO identity verification and confirmation of glycosylation site at Asn46 with MALDI-TOF peptide mass fingerprinting. The ion observed at m/z 2579.3108 corresponds to peptide $N^{42}VHSN^{46}YTFIISGGGISGLTLADR^{64}$ with Asn46 modified by one *N*-acetyl-D-glucosamine unit. PMF was performed using a sample of $CtFDO_{degl}$ loaded onto SDS gel. The data were measured with a 15 T solariX XR FT-ICR mass spectrometer (Bruker Daltonics) operating in positive mode. Data were processed by the DataAnalysis (version 4.2) software and the mMass software (Bruker Daltonics).

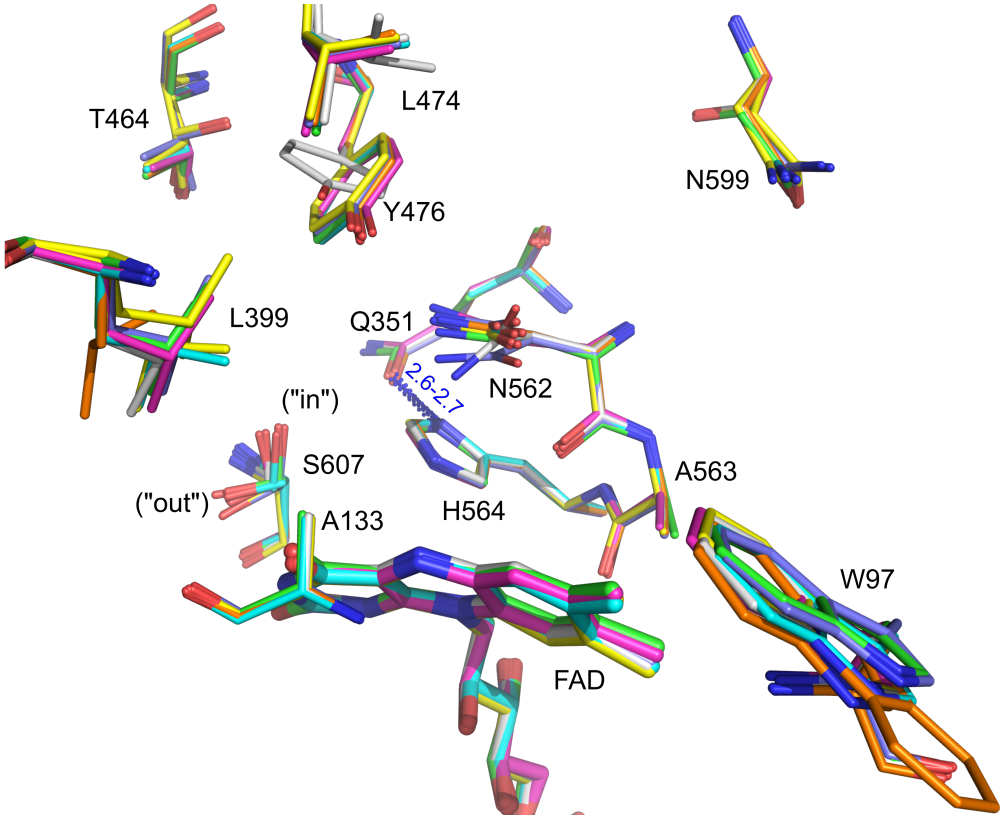

Figure A.8: Superposition of active sites of *Ct*FDO:free and *Ct*FDO complexes with ligands (chains A). The alternative conformations of active-site residues reflect their adaptation to ligand binding. Residues and FAD cofactor are shown as sticks (*Ct*FDO:free - C atoms in light grey, *Ct*FDO:MAMB - green, *Ct*FDO:PESB - cyan, *Ct*FDO:IPEA - magenta, *Ct*FDO:4NC - yellow, *Ct*FDO:4NP - purple, *Ct*FDO:ABTS - orange). The conformations of Ser607 are labelled as "in" and "out". Residues Gln351 and His564 were found having only one conformation stabilized by the His564-$N^{\delta 1}$–Gln351-$O^{\varepsilon 1}$ hydrogen bond (blue dotted line with length given in Å). The molecular graphics was created using Pymol (Schödinger, LLC).

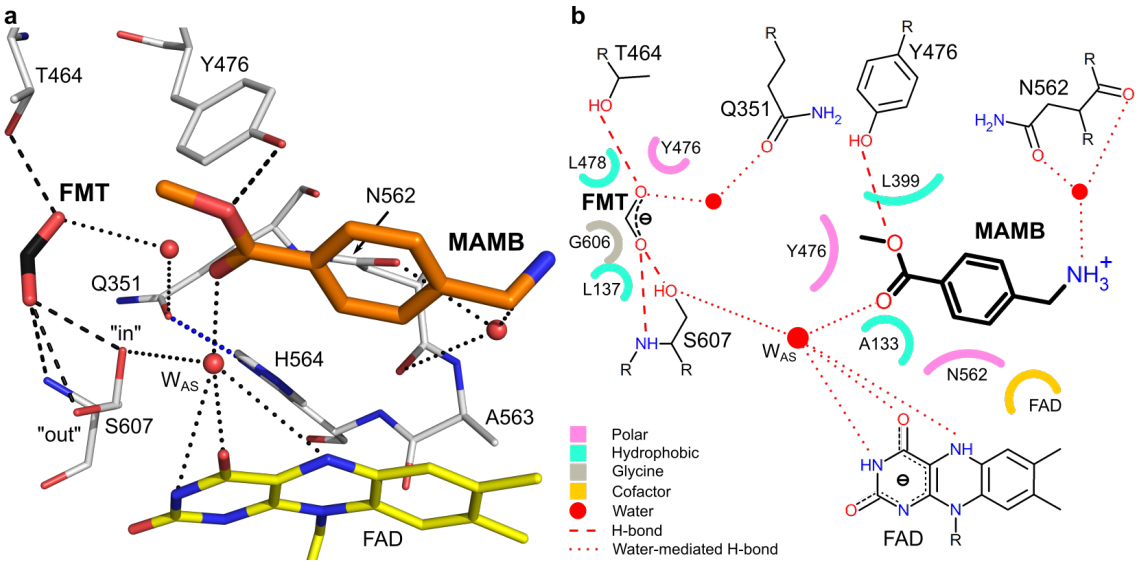

Figure A.9: The active site of *Ct*FDO:MAMB (PDB id: 6ZE3). (a) Three-dimensional structure of the active site binding a formic acid (FMT, C atom in black) and MAMB (orange) molecule. The surrounding residues and FAD have C atoms coloured in light grey and yellow, respectively. The water molecules are shown as red spheres. The active-site water W_{AS} is shown with all hydrophilic contacts and it is labelled. The black dashed and dotted lines mark direct and water-mediated hydrogen bonds, respectively. The blue dotted line marks the His564-$N^{\delta 1}$–Gln351-$O^{\varepsilon 1}$ hydrogen bond. The graphics was created using PyMOL (Schrödinger, LLC). (b) 2D schematic diagram of *Ct*FDO–MAMB and *Ct*FDO–FMT interactions in *Ct*FDO:MAMB based on an interaction diagram generated by LigPlot$^+$ [162]. Direct and water-mediated hydrogen bonds are indicated as red dashed and dotted lines, respectively. Residues with Van der Waals contacts are marked with semicircles.

Table A.5: *Ct*FDO–MAMB interactions in the *Ct*FDO:MAMB complex. W_{AS} is the active-site water molecule.

Hydrogen bond	Distance (Å)	Van der Waals
MAMB-O–Tyr476-OH	3.2	Ala133, Leu399, Tyr476, Asn562, isoalloxazine
MAMB-O^1–W_{AS}–FAD-N^5	2.5–W_{AS}–3.5	
MAMB-O^1–W_{AS}–FAD-O^4	2.5–W_{AS}–2.6	
MAMB-O^1–W_{AS}–FAD-N^3	2.5–W_{AS}–3.2	
MAMB-O^1–W_{AS}–Ser607-O^γ	2.5–W_{AS}–3.5	
MAMB-N–H_2O–Asn562-O	3.6–H_2O–3.4	
MAMB-N–H_2O–Asn562-$O^{\delta 1}$	3.6–H_2O–3.0	

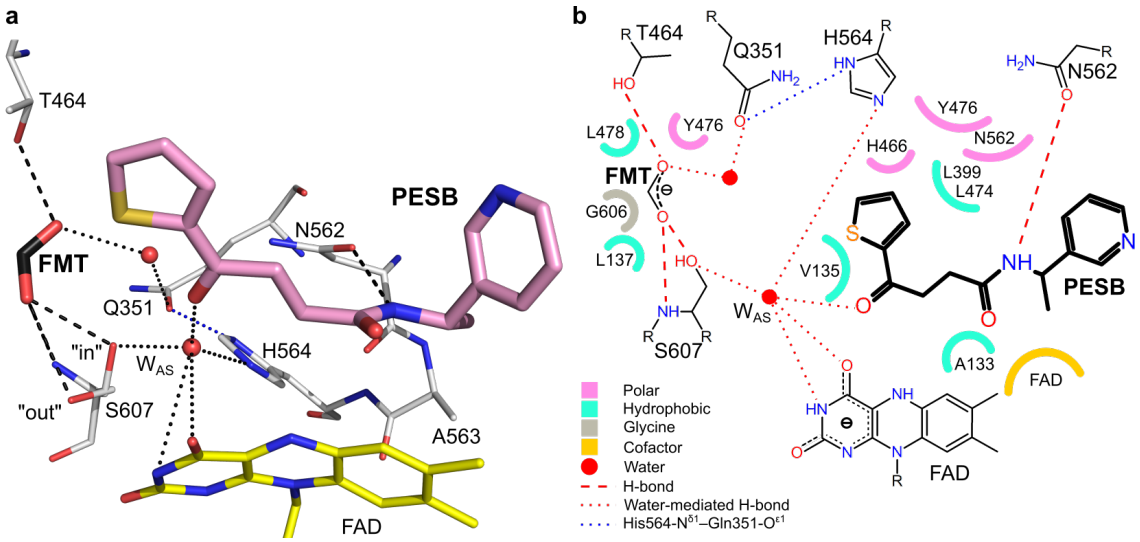

Figure A.10: The active site of *Ct*FDO:PESB complex (PDB id: 6ZE4, chain A) displayed in the same style as Fig. A.9. (a) Three-dimensional structure of the active site binding a formic acid (FMT, C atom in black) and PESB (pink) molecule. The graphics was created using PyMOL (Schrödinger, LLC). (b) 2D schematic diagram of *Ct*FDO–PESB and *Ct*FDO–FMT interactions in *Ct*FDO:PESB based on an interaction diagram generated by LigPlot$^+$ [162]. The active-site water W_{AS} is shown with all hydrophilic contacts.

Table A.6: *Ct*FDO–PESB interactions in the *Ct*FDO:PESB complex. W_{AS} is the active-site water molecule. The values in parentheses are given for chain B.

Hydrogen bond	Distance (Å)	Van der Waals
PESB-N–Asn562-O$^{\delta1}$	2.8	Ala133, Val135, Leu399, His466, Leu474, Tyr476, Asn562, isoalloxazine
PESB-O^1–W_{AS}–FAD-N^3	2.3 (2.5)–W_{AS}–3.5 (3.3)	
PESB-O^1–W_{AS}–FAD-O^4	2.3 (2.5)–W_{AS}–3.3	
PESB-O^1–W_{AS}–His564-N$^{\varepsilon2}$	2.3 (2.5)–W_{AS}–3.0 (2.9)	
PESB-O^1–W_{AS}–Ser607-O$^{\gamma}$	2.3 (2.5)–W_{AS}–2.9	

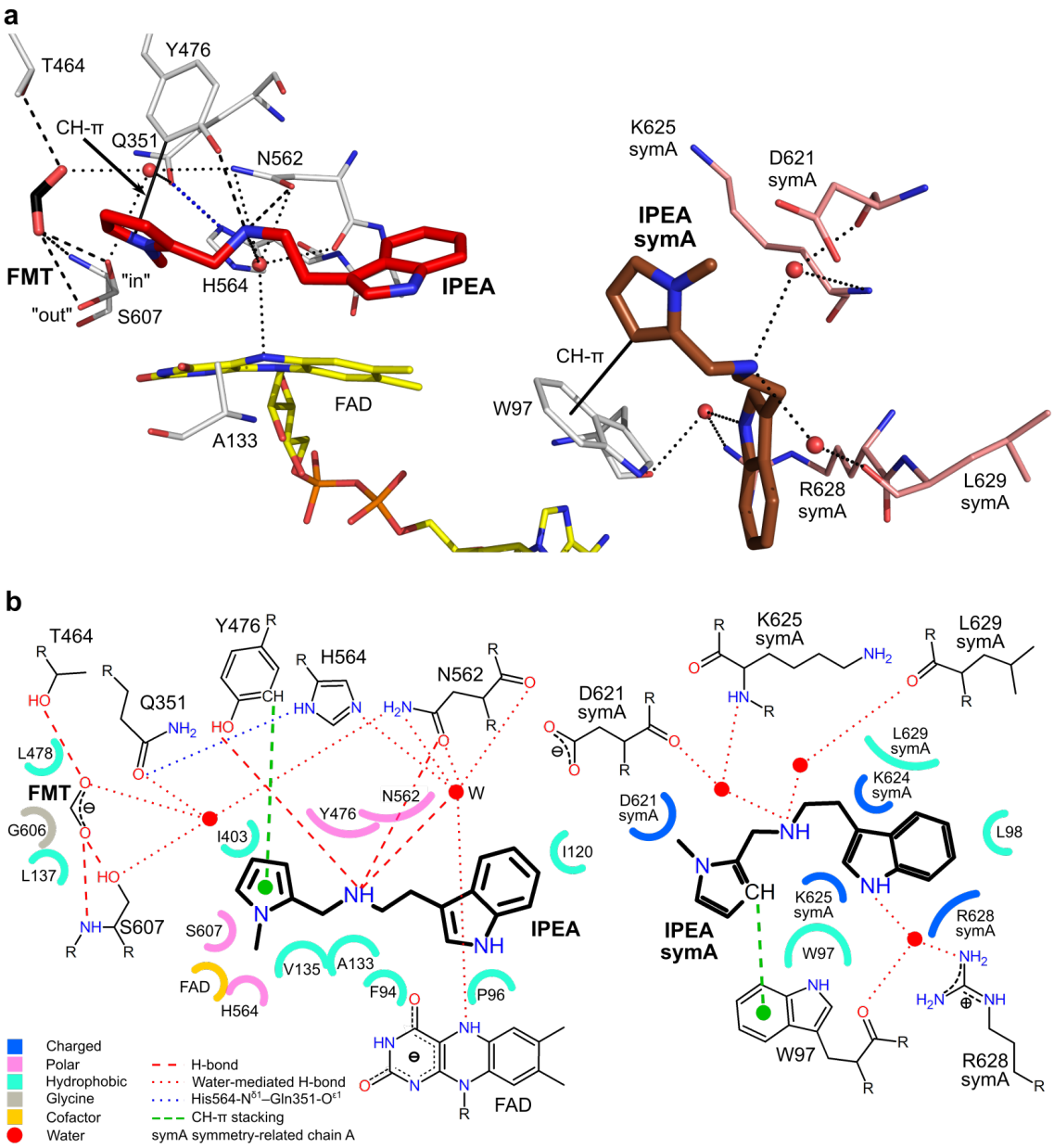

Figure A.11: The active-site pocket of *Ct*FDO:IPEA (PDB id: 6ZE5, chain A) displayed in the same style as Fig. A.9. (a) Three-dimensional structure of *Ct*FDO:IPEA binding a formic acid (FMT, C atom in black) and IPEA (red) molecule in the active-site pocket and second IPEA (brown) molecule on the interface of the symmetry-related chain A (symA, C atoms in salmon) and chain A (light grey). CH-π interactions are shown as solid black lines. The graphics was created using PyMOL (Schrödinger, LLC). (b) 2D schematic diagram of *Ct*FDO–IPEA and *Ct*FDO–FMT interactions in *Ct*FDO:IPEA based on an interaction diagram generated by LigPlot$^+$ [162]. CH-π interactions are shown as green dashed lines. Direct and water-mediated hydrogen bonds are indicated as red dashed and dotted lines, respectively. Residues with Van der Waals contacts are marked with semicircles.

Table A.7: *Ct*FDO–IPEA interactions in the *Ct*FDO:IPEA complex. IPEA and IPEA$_{symA}$ mark IPEA binding in the active-site pocket and on the interface between two molecules of *Ct*FDO, respectively. $^{\bullet}$ and * mark CH–π interaction and interaction with a residue of a symmetry related chain, respectively. The bond lengths in parentheses are given for chain B.

Hydrogen bond	Distance (Å)	Van der Waals
IPEA–Asn562-O$^{\delta 1}$	3.0 (3.2)	Phe94, Pro96, Ile120, Ala133, Val135, Ile403, Tyr476, Asn562, His564, Ser607, isoalloxazine
IPEA-pyrrole–Tyr476$^{\bullet}$	3.3 (3.4)	
IPEA–Tyr476-OH	2.9 (2.8)	
IPEA-N^1–H_2O–Asn562-O	2.9 (2.7)–H_2O–3.0	
IPEA-N^1–H_2O–Asn562-O$^{\delta 1}$	2.9 (2.7)–H_2O–2.8 (3.1)	
IPEA-N^1–H_2O–Asn562-N$^{\varepsilon 2}$	2.9 (2.7)–H_2O–3.4 (3.5)	
IPEA-N^1–H_2O–His564-N$^{\varepsilon 2}$	2.9 (2.7)–H_2O–2.7 (3.0)	
IPEA-N^1–H_2O–FAD-N^5	2.9 (2.7)–H_2O–3.1 (2.9)	
IPEA$_{symA}$-pyrrole–Trp97$^{\bullet}$*	3.3 (3.4)	Leu98*, Trp97*, Asp621, Lys624, Lys625, Arg628, Leu629
IPEA$_{symA}$-N^1–H_2O–Leu629-O	3.0–H_2O–2.6	
IPEA$_{symA}$-N^1–H_2O–Glu621-O	3.1–H_2O–2.3	
IPEA$_{symA}$-N^2–H_2O–Lys625-N	3.1–H_2O–3.5	
IPEA$_{symA}$-N^2–H_2O–Trp97-O**	2.9–H_2O–2.8	

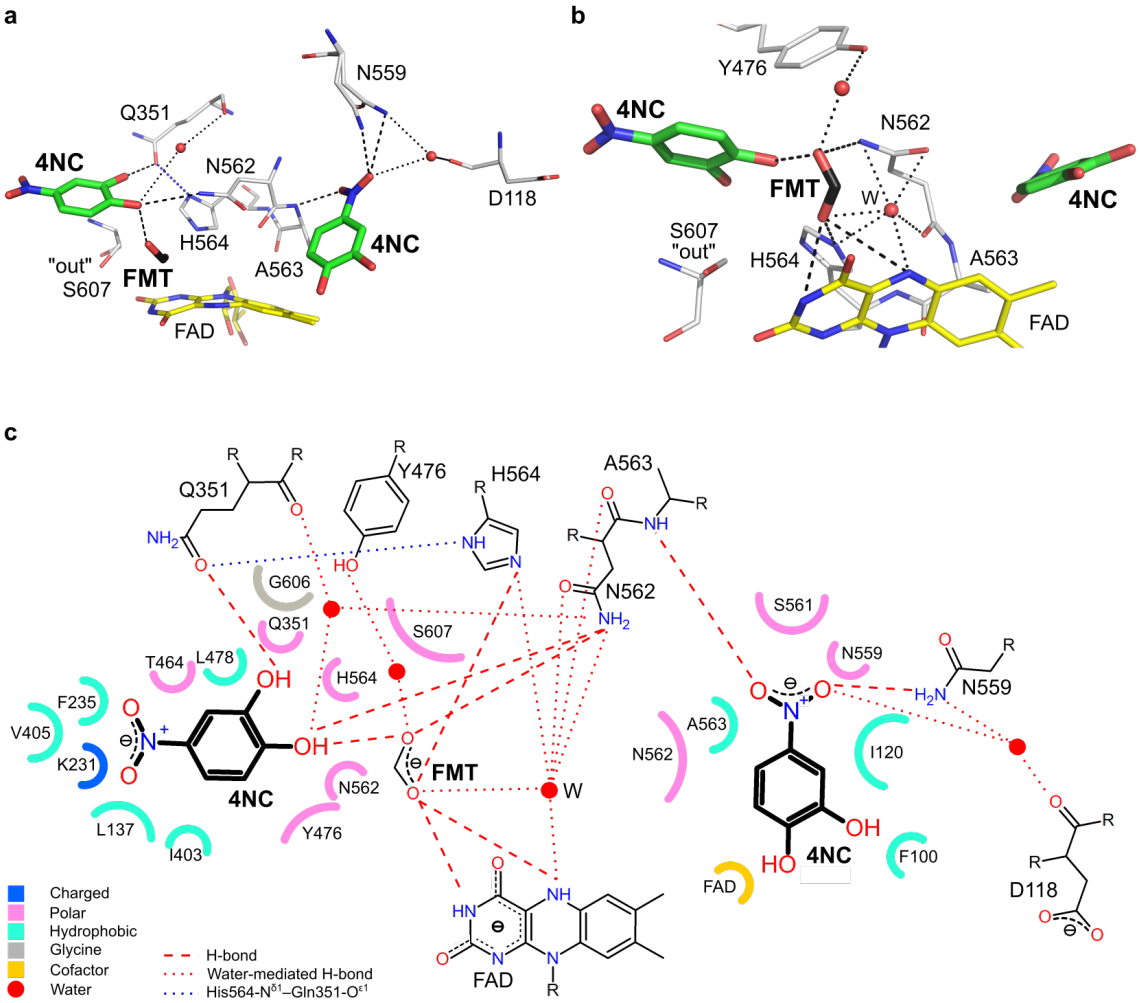

Figure A.12: The active-site pocket of *Ct*FDO:4NC (PDB id: 6ZE6, chain A) displayed in the same style as Fig. A.9. Three-dimensional structure of active-site pocket with hydrogen bonds for (a) 4NC (C atoms in green) molecules and (b) formic acid (FMT, black). The graphics were created with PyMOL (Schrödinger, LLC). (c) 2D schematic diagram of *Ct*FDO–4NC and *Ct*FDO–FMT interactions in *Ct*FDO:4NC based on an interaction diagram generated by LigPlot$^+$ [162]. The water molecule labelled W is shown with all hydrophilic contacts.

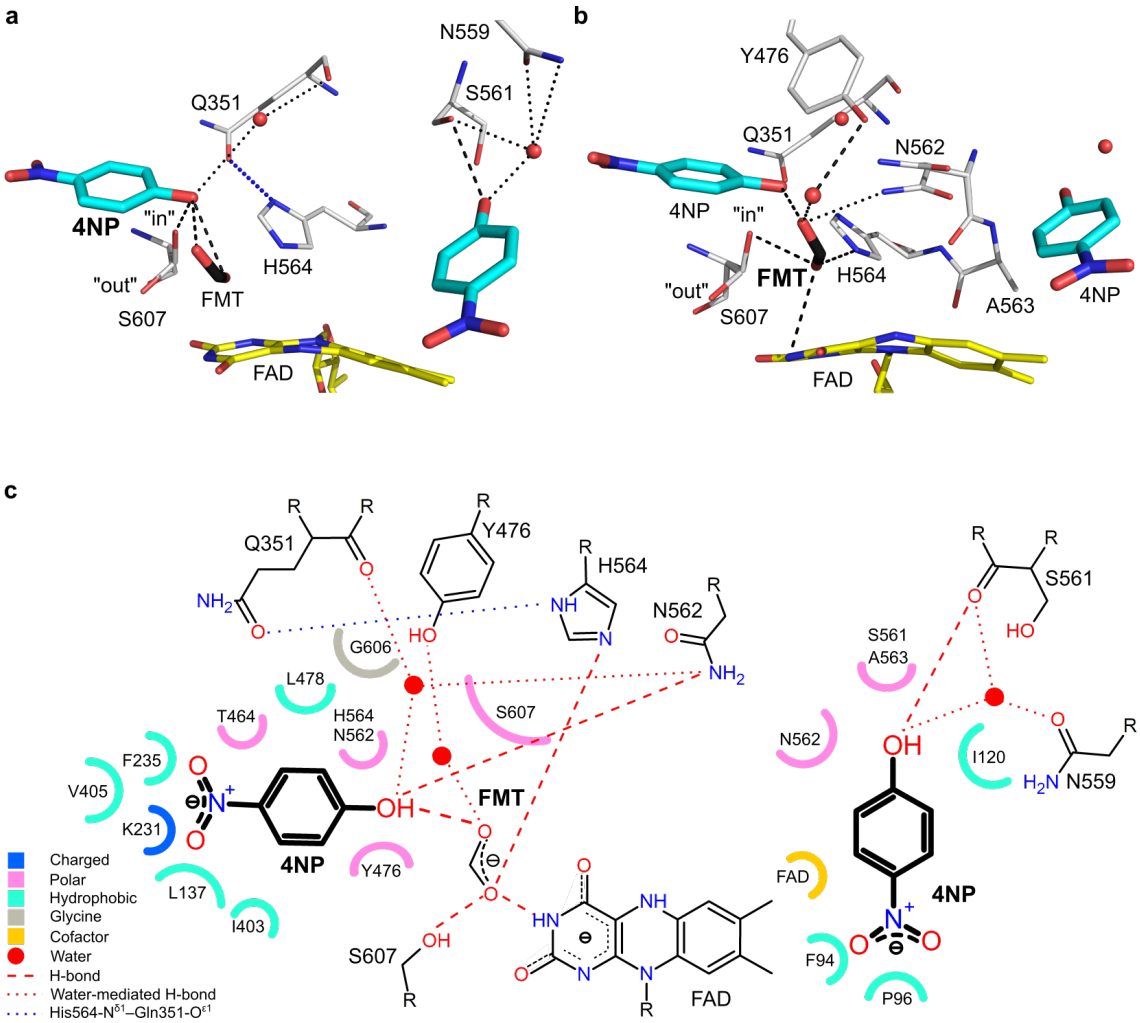

Figure A.13: The active-site pocket of *Ct*FDO:4NP (PDB id: 6ZE7, chain A) displayed in the same style as Fig. A.9. Three-dimensional structure of active-site pocket with hydrogen bonds for (a) 4NP (C atoms in cyan) molecules and (b) formic acid (FMT, black). The graphics were created with PyMOL (Schrödinger, LLC). (c) 2D schematic diagram of *Ct*FDO–4NP and *Ct*FDO–FMT interactions in *Ct*FDO:4NP based on an interaction diagram generated by LigPlot$^+$ [162].

Table A.8: *Ct*FDO–4NC and *Ct*FDO–4NP interactions in the *Ct*FDO:4NC and *Ct*FDO:4NP complex, respectively. $4NC_1$ ($4NP_1$) and $4NC_2$ ($4NP_2$) mark 4NC (4NP) binding in subsite S1 and in the entrance to the pocket, respectively. FMT is the formic acid molecule binding in CS. The bond lengths in parentheses are given for chain B. In the case of two conformations of a residue, a range of bond lengths is given.

Hydrogen bond	Distance (Å)	Van der Waals
4-nitrocatechol		
$4NC_1$-O^7–Asn562-$N^{\delta 2}$	3.3 (3.4)	Leu137, Lys231, Phe235, Gln351, Ile403, Val405, Thr464, Tyr476, Leu478, Asn562, His564, Gly606, Ser607
$4NC_1$-O^7–FMT-O^1	2.6	
$4NC_1$-O^8–Gln351-$O^{\varepsilon 1}$	2.9	
$4NC_1$-O^7–H_2O–Gln351-O	3.5 (3.6)–H_2O–2.9 (3.0)	
$4NC_1$-O^7–H_2O–Asn562-$N^{\delta 2}$	3.5 (3.6)–H_2O–2.9 (2.8)	
$4NC_2$-O^{11}–Ala563-N	3.1 (3.5)	Phe100, Ile120, Ser561, Asn562, Ala563, Asn559, isoalloxazine
$4NC_2$-O^{10}–Asn559-$N^{\delta 2}$	3.0-3.4 (3.3)	
$4NC_2$-O^{10}–H_2O–Asn559-$N^{\varepsilon 2}$	3.1 (2.9)–H_2O–3.1-3.8 (3.0)	
$4NC_2$-O^{10}–H_2O–Asp118-O	3.1 (2.9)–H_2O–2.7	
4-nitrophenol		
$4NP_1$-OH–FMT-O^1	2.6	Leu137, Lys231, Phe235, Gln351, Ile403, Val405, Thr464, Tyr476, Leu478, Asn562, His564, Gly606, Ser607
$4NP_1$-OH–Asn562-$N^{\delta 2}$	3.3 (3.4)	
$4NP_1$-OH–H_2O–Gln351-O	3.6 (3.5)–H_2O–3.0	
$4NP_1$-OH–H_2O–Asn562-$N^{\delta 2}$	3.6 (3.5)–H_2O–2.8-3.5 (2.7-3.6)	
$4NP_2$-OH–Ser561-O	3.1	Phe94, Pro96, Ile120, Ser561, Asn562, Ala563, Asn559, isoalloxazine
$4NP_2$-OH–H_2O–Asn559-$N^{\delta 1}$	2.3–H_2O–2.9	
$4NP_2$-OH–H_2O–Ser561-O	2.3–H_2O–2.9	

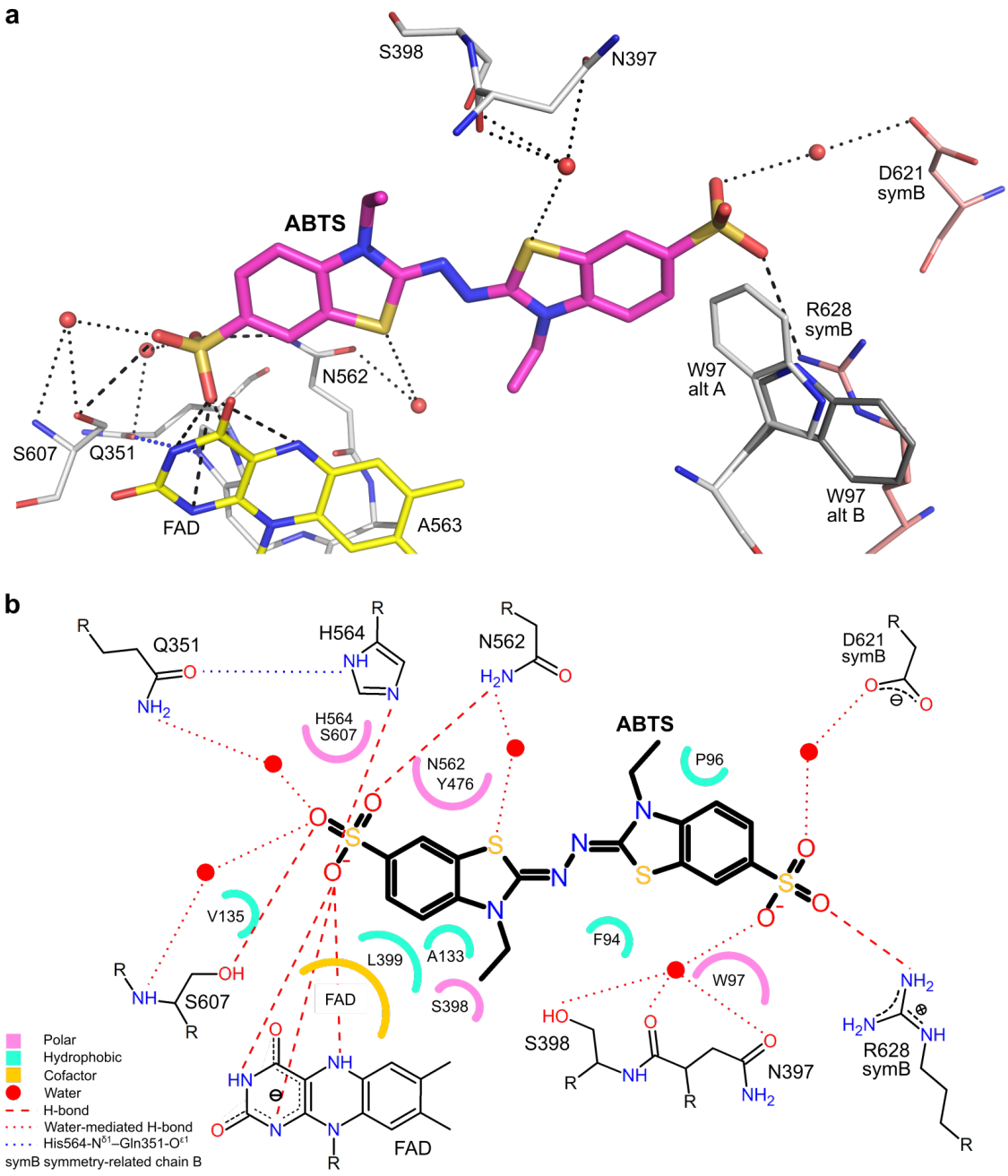

Figure A.14: The active-site pocket of *Ct*FDO:ABTS (PDB id: 7AA2, chain A) displayed in the same style as Fig. A.9. (a) Three-dimensional structure of the pocket (C atoms in light grey) binding ABTS (magenta C atoms) molecule. Symmetry-related chain B (symB) is shown with C atoms coloured in salmon. Residue Trp97 was modelled with two alternative conformations: alt A (C atoms in light grey) and alt B (grey). The graphics was created using PyMOL (Schrödinger, LLC). (b) 2D schematic diagram of *Ct*FDO–ABTS interactions in *Ct*FDO:ABTS based on an interaction diagram generated by LigPlot$^+$ [162].

Table A.9: *Ct*FDO–ABTS interactions in *Ct*FDO:ABTS complex. The bond lengths in parentheses are given for chain B. * marks interaction with a residue of a symmetry related chain.

Hydrogen bond	Distance (Å)	Van der Waals
ABTS-O^{47}–FAD-N^{1}	3.5 (3.6)	
ABTS-O^{47}–FAD-N^{3}	3.1	
ABTS-O^{47}–FAD-N^{5}	3.4	
ABTS-O^{47}–His564-$N^{\varepsilon 2}$	2.6 (2.7)	
ABTS-O^{45}–Ser607-O^{γ}	3.3 (3.4)	
ABTS-O^{46}–Asn562-$N^{\delta 2}$	3.0	Phe94, Pro96,
ABTS-O^{39}–Arg628-$N^{\eta 2}$*	3.4 (-)	Trp97, Ala133,
ABTS-O^{46}–H_2O–Gln351-$O^{\varepsilon 1}$	2.9–H_2O–3.2 (3.3)	Val135, Ser398,
ABTS-S^{18}–H_2O–Asn397-$N^{\delta 1}$	3.3 (3.6)–H_2O–3.5 (3.4)	Leu399, Tyr476,
ABTS-S^{18}–H_2O–Asn397-O	3.3 (3.6)–H_2O–2.9 (3.0)	Asn562, His564,
ABTS-S^{18}–H_2O–Ser398-O^{γ}	3.3 (3.6)–H_2O–3.4 (3.7)	Ser607,
ABTS-S^{9}–H_2O–Asn562-$N^{\delta 1}$	3.5 (3.6)–H_2O–3.0 (3.2)	isoalloxazine
ABTS-S^{9}–H_2O–Asn562-O	3.5 (3.6)–H_2O–3.2 (3.3)	
ABTS-O^{45}–H_2O–Ser607-O^{γ}	3.0–H_2O–3.0 (2.8)	
ABTS-O^{45}–H_2O–Ser607-N	3.0–H_2O–3.6 (3.4)	
ABTS-O^{48}–H_2O–Asp621-$O^{\delta 2}$*	3.3–H_2O–3.1 (-)	

Appendix B

Bilirubin oxidase from *Myrothecium verrucaria*

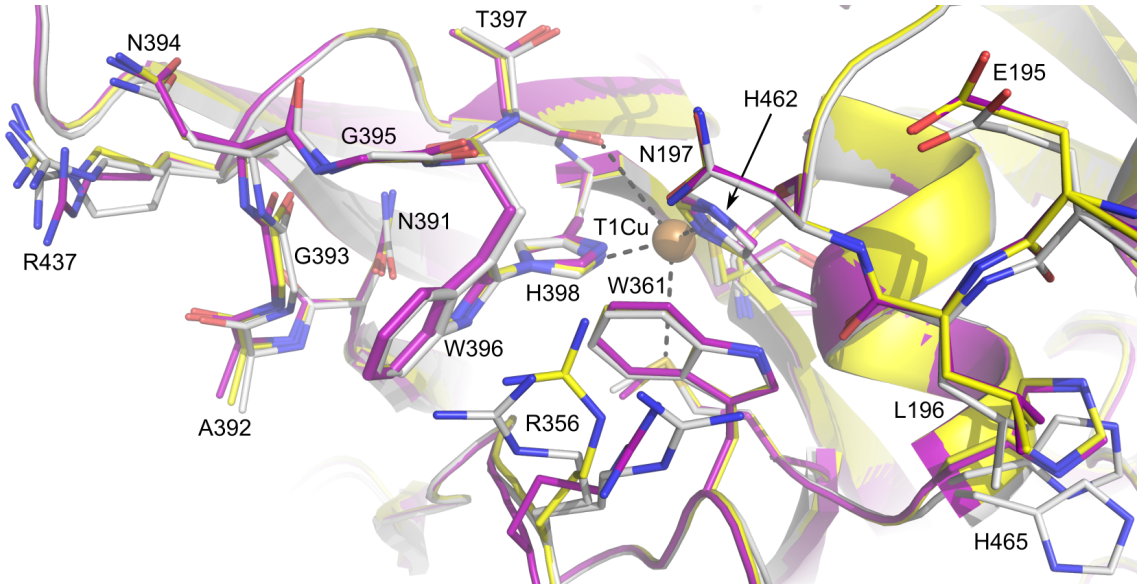

Figure B.1: Structure alignment of surroundings of oxidation site 1 in published structures of *Mv*BOx wild type with that of WT:FECN. Structure of WT:FECN (pH 3.1, PDB id: 6I3J), *Mv*BOx from basic condition (pH 8.7, PDB id: 2XLL), and acidic condition (pH 5.0, PDB id: 6IQZ) has C atoms coloured in light grey, purple, and yellow, respectively. T1Cu is shown as a brown sphere with coordination bonds (black dashed lines) to surrounding residues. Molecular graphics was created using PyMOL (Schrödinger, LLC).

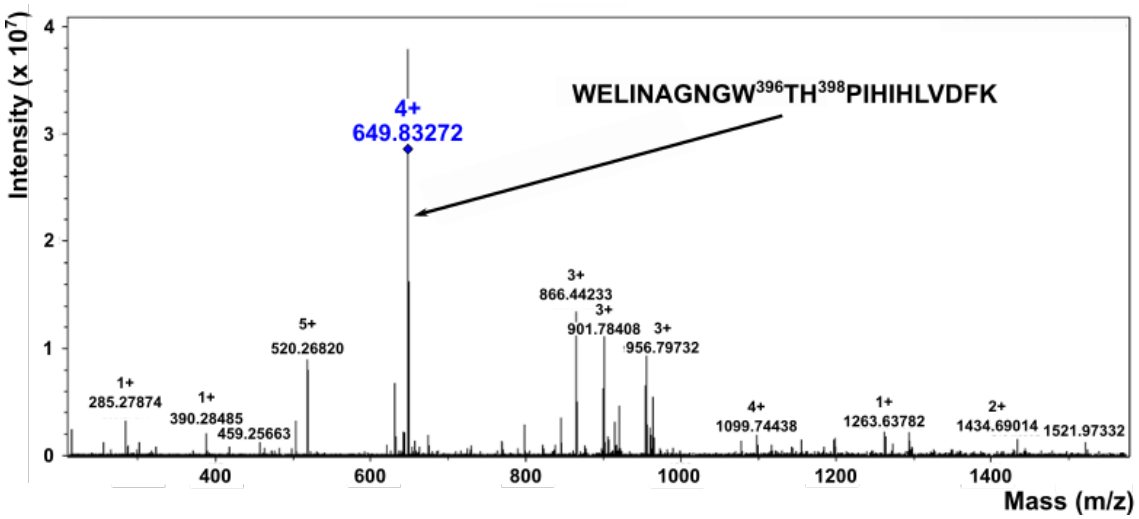

Figure B.2: LC-MS/MS spectrum confirming the existence of the Trp396–His398 crosslink in *Mv*BOx. The fragments from *Mv*BOx digestion by trypsin were analysed using a 15T solariX FT-ICR mass spectrometer (Bruker Daltonics) operating in positive-ion mode. Fragments of peptides were identified with ProteinScape (Bruker Daltonics) with the Mascot search engine. The observation of ion at m/z 649.8327 (4+) corresponds to the W^{387}ELINAGNGW^{396}TH^{398}PIHIHLVDFK^{408} peptide modified by the crosslink.

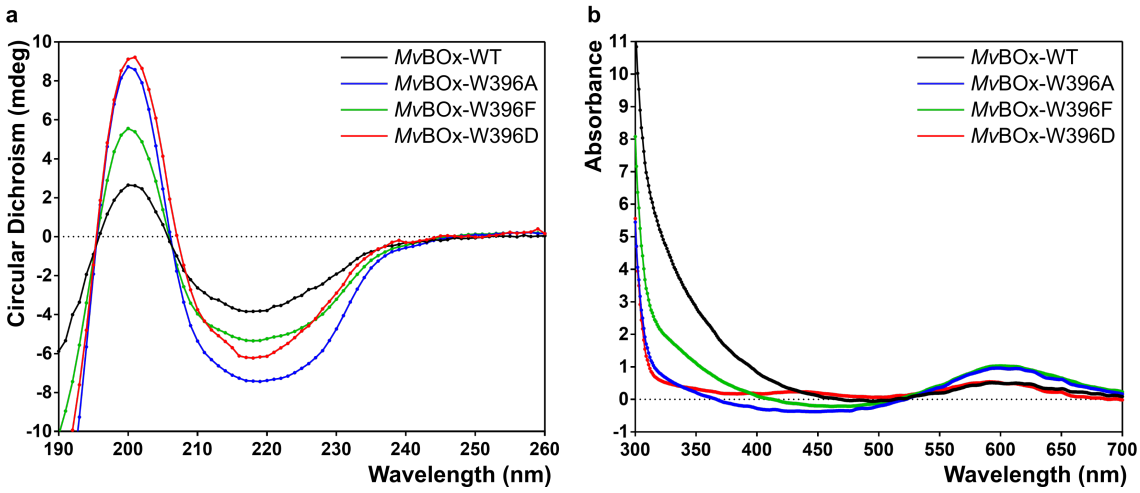

Figure B.3: Confirmation of secondary structure and T1Cu site preservation in *Mv*BOx wild type (black curve), *Mv*BOx-W396A (blue), *Mv*BOx-W396F (green), and *Mv*BOx-W396D (red). Different height of peaks reflects only varied concentration of the samples. (a) CD spectra measured with 0.1 cm quartz cell and a Chirascan^{TM}–plus spectrometer (Applied Photophysics). The final spectra are expressed in degrees of ellipticity. Peaks at the same position in each profile indicate the same proportion of secondary structure elements in all samples. Thus, the mutation of Trp396 to alanine, phenylalanine, and glutamic acid did not change the secondary structure. (b) UV-VIS absorption spectra measured for *Mv*BOx wild type and its variants. All samples has absorbance peak around 600 nm typical for oxidized T1Cu site in multicopper oxidases. The spectra were measured with DeNovix DS-11 microvolume spectrophotometer. The spectra were plotted using the GraphPad Prism version 7.02.

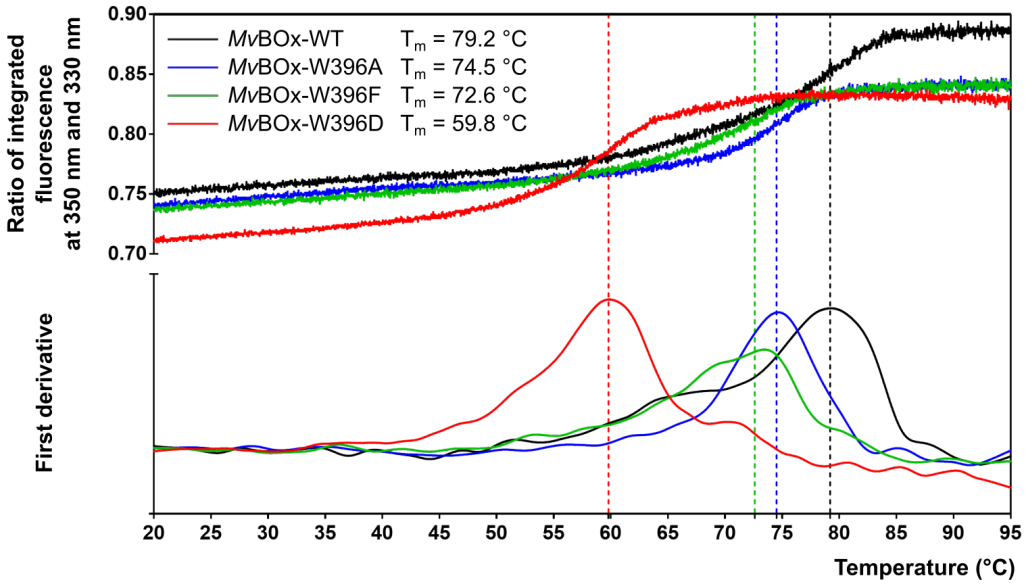

Figure B.4: Thermal stability of *Mv*BOx wild type and its mutated variants in 50 mM Tris-HCl pH 7.5 with 50 mM NaCl. The nanoDSF spectra were measured with Prometheus NT.48 (Nanotemper) using the PR.ThermControl software (v2.1.1) and plotted with the GraphPad Prism version 7.02. The stability is expressed as a dependence of the ratio of integrated fluorescence at 350 and 330 nm (F350/F330) and its first derivative on temperature. Differences in heights of peaks are given by mild differences in concentration of the samples.

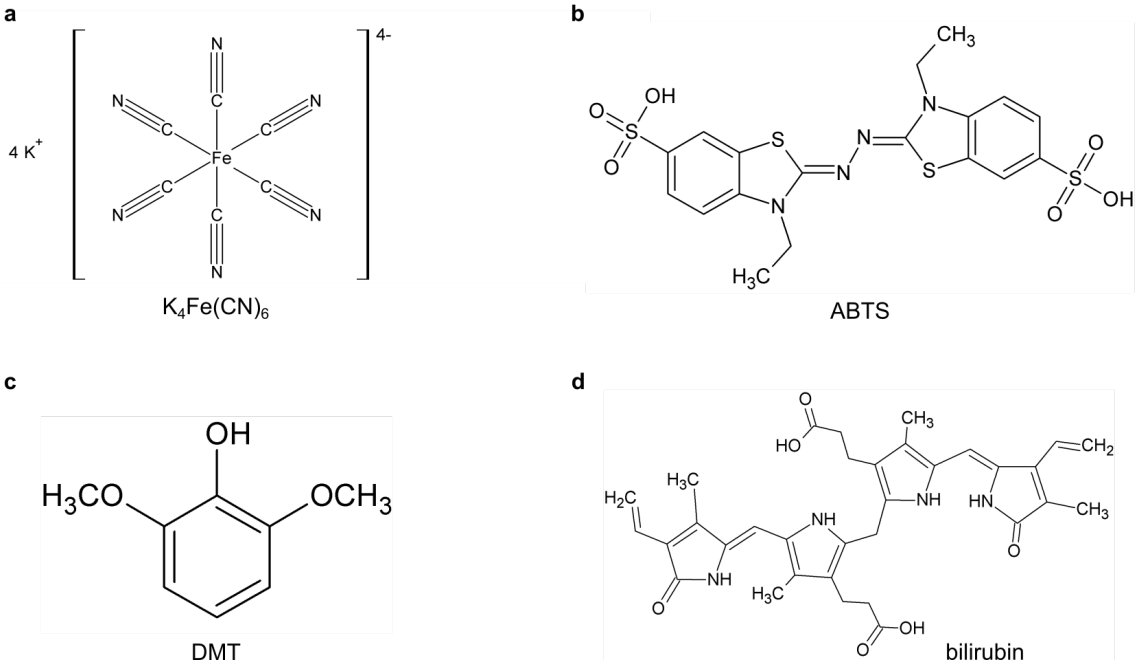

Figure B.5: Structural formulae of substrates of *Mv*BOx used for activity measurements. (a) Potassium ferrocyanide ($K_4Fe(CN)_6$), (b) 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS), (c) 2,6-dimethoxyphenol (DMP), and (d) bilirubin.

List of Abbreviations

Table A.10: List of abbreviations used in the thesis.

abbreviation	full name
4NC	4-nitrocatechol
4NP	4-nitrophenol
ABTS	2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)
*Af*GDH	*Aspergillus flavus* FAD glucose dehydrogenase
AIC	Akaike's Information Criterion
*An*GOX	*Aspergillus niger* glucose oxidase
APBS	Adaptive Poisson-Boltzmann Solver
CD	circular dichroism
CS	catalytic site
CSD	Cambridge Structural Database
Ct	*Chaetomium thermophilum*
*Ct*FDO	*Chaetomium thermophilum* FAD-dependent oxidoreductase
$Ct\text{FDO}_{\text{degl}}$	deglycosylated form of *Ct*FDO
*Ct*FDO:free	structure of ligand-free *Ct*FDO
*Ct*FDO:ligand	structure of *Ct*FDO in complex with a given ligand
DLS	dynamic light scattering
DMP	2,6-dimethoxyphenol
DTN	sodium dithionite
FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide
GlcNAc	*N*-acetyl-D-glucosamine
GMC	glucose-methanol-choline
GOXs	*An*GOX and *Pa*GOX
HMFO	5-hydroxymethylfurfural oxidase from *Methylovorus sp.*
HRP	horseradish peroxidase
HTS	high-throughput screening
IEF	isoelectric focusing
IPEA	2-(1*H*-indol-3-yl)-*N*[(1-methyl-1*H*-pyrrol-2-yl)-methyl]ethan-amine

Continued on next page

Table A.10 – continued from previous page

abbreviation	full name
IPEA-symA	IPEA binding on the contact of chain A and symmetry-related chain A in *Ct*FDO:IPEA
LC-MS	liquid chromatography tandem-mass spectrometry
MADLS	multi-angle dynamic light scattering
MAMB	methyl4-(aminomethyl)benzoate
MCO	multicopper oxidase family
MP	mass photometry
MS	mass spectrometry
*Mt*AAO	*Thermothelomyces thermophilus* aryl-alcohol oxidase
Mv	*Myrothecium verrucaria*
*Mv*BOx	bilirubin oxidase from *Myrothecium verrucaria*
*Mv*BOx-WT	*Mv*BOx wild type
*Mv*BOx-M467Q	*Mv*BOx with Met467 mutated to glutamine
*Mv*BOx-W396A	*Mv*BOx with Trp396 mutated to alanine
*Mv*BOx-W396D	*Mv*BOx with Trp396 mutated to aspartic acid
*Mv*BOx-W396F	*Mv*BOx with Trp396 mutated to phenylalanine
MX crystallography	macromolecular X-ray crystallography
nanoDSF	nano differential scanning fluorimetry
NI	native intermediate
P2O	*Trametes multicolor* pyranose 2-oxidase
*Pa*GOX	*Penicillium amagasakiense* glucose oxidase
PDB	Protein Data Bank
*Pe*AAO	*Pleurotus eryngii* aryl-alcohol oxidase
PEG	polyethylene glycol
PESB	4-oxo-*N*-[1-(3pyridinyl)ethyl]-2-thiophenebutanamide
PI	peroxide intermediate
PNOX	pyridoxine 4-oxidase
S1–S5	Subsite 1–5
SEC	size exclusion chromatography
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
T1Cu, T2Cu, T3Cu	copper ion of type I, II, III
TNC	trinuclear copper cluster
W_{AS}	active-site water molecule
W396A:FECN	crystal structure of *Mv*BOx-W396A in complex with ferricyanide
W396F	crystal structure of *Mv*BOx-W396F
WT:FECN	crystal structure of *Mv*BOx in complex with ferricyanide