Molecular Simulations on Electron Transfer Proteins

Dissertation
zur Erlangung des Grades
des Doktors der Naturwissenschaften
der Naturwissenschaftlich-Technischen Fakultät
der Universität des Saarlandes

von

Jan Leo Riehm

Saarbrücken

September 2018

Tag des Kolloquiums:	13.06.2018
Dekan:	Prof. Dr. Guido Kickelbick
Berichterstatter:	PD Dr. Michael Hutter
	Prof. Dr. Rita Bernhardt
Vorsitz:	Prof. Dr. Gregor Jung
Akad. Mitarbeiter:	Dr. Britta Diesel

Abstract

Cytochrome P450s are versatile biocatalysts with the ability to accept a vast range of substrates. However, to enable their high potential they are in need of electrons. Those are delivered via complex electron transfer chains including an electron donor (NAD(P)H) and one or two transfer proteins. Transfer proteins are exchangeable from different organisms and work in various combinations. To gain knowledge and understanding about possible interactions of redox partners and subsequent transport of electrons to cytochromes P450 (P450s) molecular modelling methods are vital. Increasing computational power makes *in silico* methods like molecular docking and molecular dynamics simulations more and more suitable to analyze biological systems like P450s. With these methods we were able to establish changes in progesterone hydroxylation and subsequently shift and increase the product spectrum. The introduction of modelled glycine and proline linker structures could help to explain the interactions of proteins and the accompanied electron transfer. Also we investigated combinations of redox partners to optimize the redox chain of CYP106A2. To lower the redox potential of adrenodoxin, a suitable redox partner of CYP106A2, methods of rational protein design were applied.

Zusammenfassung

Cytochrom P450 Enzyme sind breit einsetzbare Biokatalysatoren, die eine Vielzahl an Substraten akzeptieren können, eine Fähigkeit die sie sehr attraktiv für eine biotechnologische Anwendung macht. Trotz ihrer vielseitigen Einsetzbarkeit gibt es jedoch einige Limitierungen, die es zu überwinden gilt. Hierzu zählen beispielsweise die komplexen Elektronentransferprozesse oder die niedrige Elektronentransferrate. Die zur Elektronenübertragung benötigten Transferpartner sind über verschiedene Organismen hinweg austauschbar und kombinierbar. Mit stetiger Erhöhung der Rechenleistungen werden *in silico* Methoden wie molekulares Docking oder Moleküldynamiksimulationen zur Analyse biologischer Systeme immer leichter einsetzbar. Im Rahmen unserer Forschungsarbeiten konnten wir damit Änderungen der Progesteronhydroxylierung bewirken, die das Produktspektrum verschob und den Ertrag wesentlich erhöhte. Die Modellierung von Glycin- und Prolin-Linker-Strukturen trug dazu bei, Interaktionen der Redoxpartner und den einhergehenden Elektronentransfer zu erklären. Des Weiteren wurden verschiedene Kombinationsmöglichkeiten von Redoxpartnern zur Optimierung der Redoxkette von CYP106A2 untersucht. Um das Redoxpotential von Adrenodoxin, einem passenden Redoxpartner des CYP106A2, herabzusetzen, kamen Methoden des rationalen Proteindesigns zum Einsatz.

Acknowledgement

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Lehrstuhl für Computational Biology der Universität des Saarlandes im Rahmen des vom Bundesministerium für Bildung und Forschung geförderten Projekts **SupraRedoxModul** - *Supramolecular assembly of redox systems and cytochromes P450*. Sie wäre ohne die Mithilfe von verschiedenen Menschen nicht möglich gewesen. Sei es entweder direkt durch Ratschläge oder indirekt durch ihre Untersützung gewesen.

An erster Stelle möchte ich mich bei meinem Doktorvater PD Dr. Michael Hutter bedanken. Er gab mir die Möglichkeit, mich in diesem spannenden Themenfeld zu bewegen und zu forschen. Ich danke ihm für die jederzeitige Unterstützung und erfrischenden Diskussionen, die mir im dunklen Wald der Chemoinformatik stets den richtigen Weg aufzeigten. Unvergessen bleiben auch die Fahrten nach Erlangen zum Molecular Modelling Workshop mit anschließendem Bier im Bierkeller.

Des Weiteren bedanke ich mich bei unseren Projektpartnern für die fruchtbaren Diskussionen, den Ratschlägen und netten Abenden bei unseren Projekttreffen. Explizit sind hier Prof. Dr. Rita Bernhardt und Tanja Sagadin von der Universität des Saarlandes, Prof. Dr. Vlada Urlacher, Dr. Patrick Bakkes und Stefan Biemann der Universität Düsseldorf und Prof. Dr. Roger Gläser, Prof. Dr. Einicke und Kerstin Thiele der Universität Leipzig zu nennen.

Ich bedanke mich ebenfalls bei Prof. Dr. Helms, dem Lehstuhlleiter, für die hilfreichen Ratschläge und angenehmen Unterhaltungen jenseits der Wissenschaft.

Nicht zuletzt bedanke ich mich bei allen Mitarbeitern des Lehrstuhls, insbesondere der Mensatruppe um Thorsten Will, Kerstin Reuter und Rahmad Akbar.

Ein besonderer Dank gilt auch unserer Sekretärin Kerstin Gronow-Pudelek, die mir bei allen Problemen administrativer Natur beigestanden hat.

Zum Schluss möchte ich mich nun bei meiner Familie und Freunden bedanken. Meinen Eltern Hans-Joachim und Elfi, die nun wahrscheinlich sehr erleichtert aufatmen, da die Arbeit *endlich* fertig ist. Natürlich auch bei Jonas, Lena, Nicholas, Peter und Monika.

Mein allergrößter Dank gilt selbstverständlich meiner Frau Franka. Sie hat mir während meiner Zeit als Doktorand das Leben erheblich erleichtert und verschönert und mich wahnsinnig unterstützt. Dafür bin ich ihr sehr dankbar. Danke Franka!
Darüberhinaus unserer Tochter Karla, die nun nachdem die Arbeit fertig ist, scheinbar endlich durchschläft.

Contents

Contents

Chapter 1

Introduction

1.1 CYP P450 Enzymes

The discovery of cytochrome P450 enzymes (P450s or CYPs) as such lays back around 60 years. In 1958, shortly after the discovery of Mason [1] and Hayashi [2], Klingenberg [3] and Garfinkel [4] described a carbon monoxide (CO-) binding pigment in rats and pigs liver microsomes. Only six years later its hemoproteioic nature was proven by Omura and Sato [5] and the term *Cytochrome P450* first showed up in the scientific literature. The term P450 originates from the word pigment and the characteristic absorption maximum of 450 nm shown by the reduced CO-complex. During the 1960s a lot of fundamental discoveries were made, e.g. 1963 Estabrook and coworkers discovered the oxygenase function in steroid metabolism [6], in 1965 Cooper learnt about the oxidation of diverse drugs in liver microsomes and Appleby came up with the first P450 spectrum in bacterial cells in 1967. In the subsequent year a major breakthrough was achieved by Gunsalus by the discovery of the bacterial P450 monooxygenase P450-CAM from *Pseudomonas putida* [7]. In contrast to their mammalian counterparts, bacterial P450 enzymes are soluble, not attached to a membrane and easy to purify. This allows for a subsequent production in large quantities. From these findings many mechanistic and spectroscopic measurements were seized up to the creation of the first crystallographic structure of *P450-CAM* in 1985 (*CYP101A1*).
Due to technological improvement of genetic techniques in the 1980s more and more P450 genes were described. To keep track a unique naming convention had to be established. In 1987 Nebert and Gonzales [8] launched a standardised P450 nomenclature system based on sequence comparison. Here the term CYP is followed by an arabic number representative for the CYP family with a minimum sequence identity of 40%. The subfamily with a sequence identity of at least 55% is marked by a letter and the individual gene again by a number. In 1995 David Nelson published a webpage containing information of about 35,166 [9](as of April 2016) individual genes

based on that nomenclature. The database is hand curated and contains information from all kingdoms of life, such as animals, plants, fungi, protozoa, bacteria, archaea and also viruses. Figure 1.1 shows the diverse gene distribution with the highest number seen in plants with nearly 14000. Despite having sequence identities as low

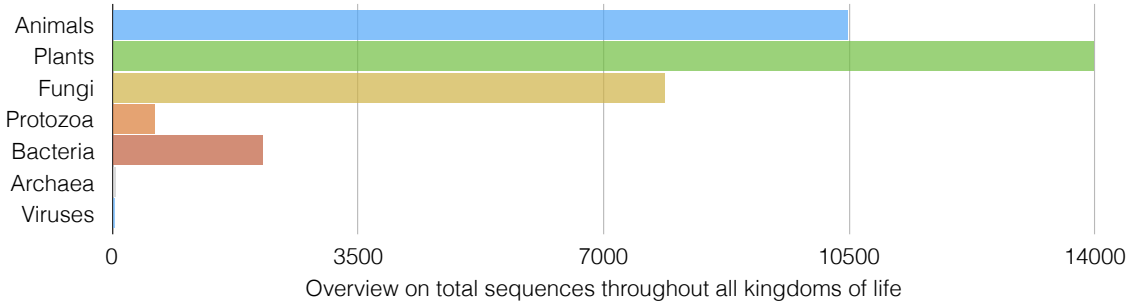

Figure 1.1: Distribution of P450 genes throughout all kingdoms of life (taken from [9]).

as 16% they still share the same structural fold since the very beginning. The only conserved amino acid in all P450s is a cysteine that serves as a fifth ligand to the heme iron [10].

1.1.1 Function

P450 enzymes are members of the heme *b* containing superfamily of monooxygenases. Their ability to transform a broad variety of substrates such as fatty acids, steroids, prostaglandines, pharmaceutical drugs, anaesthetics, ethanol, pesticides or carcinogens as well as their involvement in hydroxylation, N-, O- and S-dealkylation, sulphoxidation, epoxidation, deamination, desulphuration or peroxidation makes them a pretty promising target for biotechnological applications.
Despite its manifold area of application, in reality, biotechnological usability is limited due to e.g. low activity of the enzyme, lacking amount of substrate bound crystal structures for precise protein design, the need for electron transfer partners and the limitations in electron transport. An excellent snapshot in time on chances and limitations concerning biotechnological application of CYPs is provided in the review of Bernhardt and Urlacher [11]. When using whole cell systems one is confronted with problems such as NAD(P)H depletion, substrate solubility, as well as substrate or product toxicity. During the actual catalytic cycle P450s transfer a single atom of molecular oxygen to their substrate while reducing the other atom to water (eq. 1.1).

$$RH + O_2 + NAD(P)H + H^+ \longrightarrow ROH + H_2O + NAD(P)^+ \tag{1.1}$$

For this reaction they are in need of external electron donors. Depending on the kind of electron transfer from NAD(P)H, P450s are divided into four classes [10]. Whereas class I and II enzymes receive electrons from a ferredoxin and a FAD containing

reductant or just a FMN/FAD containing NADPH-P450 reductase respectively, class III enzymes are self sufficient and do not need a donor. Class IV P450s get their required electrons directly from NAD(P)H. Figure 1.2 shows a simplified scheme of the hydroxylation process in P450s. In the first step ferric state iron gets reduced upon electron acceptance into the ferrous state. After the subsequent binding of dioxygen the second electron is introduced which leads to the cleavage of the O-O bond. One of the oxygen atoms gets reduced to water whereas the other forms a highly reactive ferryl intermediate (Fe=O). This intermediate adopts the hydrogen from the RH molecule. The resulting free unstable $R\cdot$ radical accepts the OH group while the ferryl iron transforms back to the ferric state.

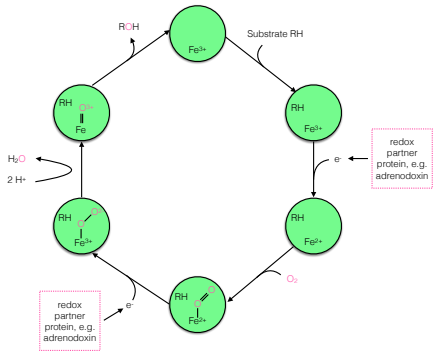

Figure 1.2: Simplified scheme of the catalytic cycle, common in almost every P450 enzyme

1.1.2 Electron Transfer

Electron transport to P450s is mediated through a complex multiprotein system. An excellent overview of the various classes of electron transport chains was published by Hannemann et al. in 2007 [12]. In general, electron transport systems can be divided in two main classes (figure 1.3). Class I contains membrane-bound adrenal mitochondrial P450s. Electron transfer from NAD(P)H includes a flavin-containing reductase (ferredoxin/adrenodoxin reductase) and an iron-sulfur protein (ferredoxin/adrenodoxin) [13]. On the other hand, in class II systems electron transfer is realized with just one redox partner, a Flavin-Adenin-Dinucleotide (FAD) and Flavin mononucleotide (FMN) containing reductase. In humans, 50 out of 57 P450s are microsomal and are located in the endoplasmatic reticulum (ER). The remaining ones are mitochondrial [13]. Most bacterial P450s belong to class I. They also use two redox partners but are unlike their eukaryotic counterparts soluble and not attached to the membrane. In the 1980s a first fused system was found [14, 15], which was followed by other transport chains leading to a total of 10 different classes.

The part as mediator of electron transport to P450s is taken over in many cases by ferredoxins. Ferredoxins are small (6-25 kDa) acidic proteins that are negatively charged at neutral pH. Their key feature is a iron-sulfur cluster as prosthetic group. Depending on the number of iron-sulfur clusters they can be divided into different classes ([2Fe-2S],[3Fe-3S] or [4Fe-4S] and combinations of them). [2Fe-2S] ferrdoxins are subdivided into three classes namely plant-type, thioredoxin-type and vertebrate-type (figure 1.4) [16].

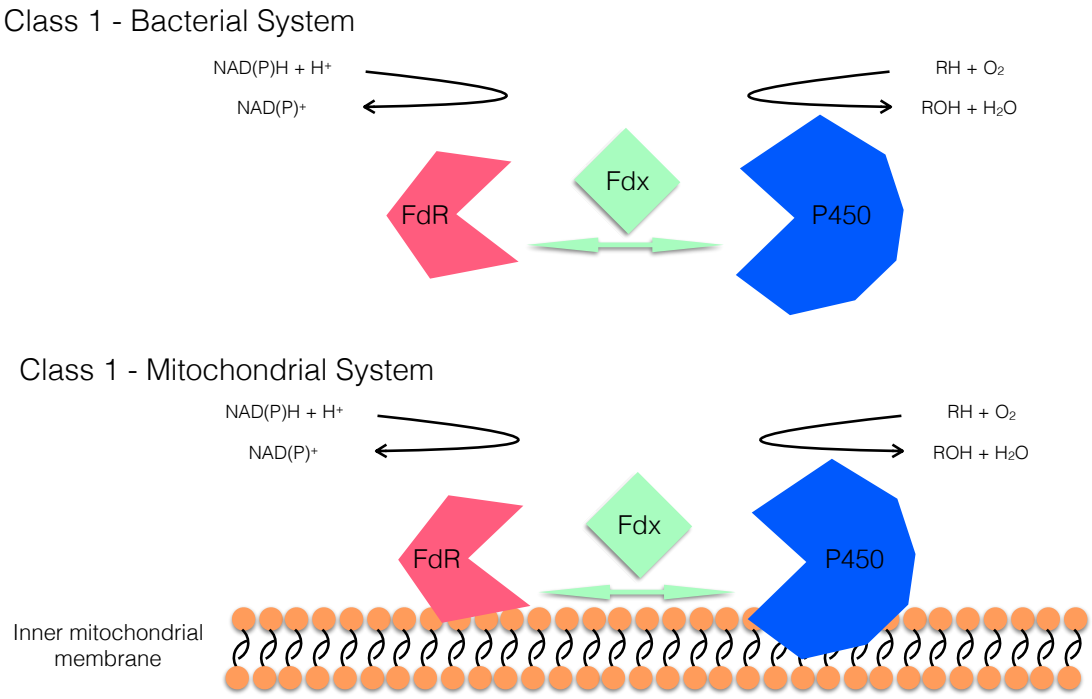

Figure 1.3: Overview of the different class 1 electron transport mechanisms

1.1.3 CYP106A2

CYP106A2 is a bacterial P450 from *Bacillus megaterium* ATCC 13368 with a molecular mass of 47.5 kDa and a length of 410 amino acids. Despite the lack of knowledge of its native function, it has the ability to hydroxylate not only steroids, but also terpenoids or terpenes. In 1958 McAleer et. al [17] performed a 15β-hydroxylation of progesterone through *Bacillus megaterium*. However, this was only proved more than 20 years later [18]. Further investigation of Berg and coworkers identified the hydroxylation-complex consisting of a P450, a NADPH-dependent megaredoxin reductase and megaredoxin, an iron-sulfur cluster protein [18] - [23]. The recently identified CYP106A2 from *Bacillus megaterium* has a sequence identity of 63% to CYP106A1.

Neither its redox partners are cloned yet, nor is its native substrate known. However, it can interact with the adrenal redox system such as adrenodoxin (Adx) and adrenodoxin reductase (AdR) [20, 24] or with other electron transfer systems of *Bacillus subtilis* [25] or *Pseudomonas putida* (Putidaredoxin and Putidaredoxin reductase) [26]. Previous studies concentrated on the conversion and testing of $3-oxo-\Delta^4$-steroids mostly in the 15β position. Recently the first report of a diterpene hydroxylation was published [27] as well as findings about conversions of pentacyclic triterpene $11-keto-\beta$ boswellic acid (KBA) [28] or the triterpenoid dipterocarpol [29]. In 2016 Janocha et al. [30] published the first crystal structure of CYP106A2 with and without a bound substrate (abietic acid) which was used in part of my studies presented in the results chapters of this thesis.
Since CYP106A2 is a well established hydroxylase it could be an excellent candidate for biotechnological, pharmaceutical or industrial application.

1.2 Adrenodoxin, Etp1^{fd} and other ferredoxins

The name adrenodoxin (Adx) was derived from the tissue it was first isolated from, the adrenal glands. Those are the manufacturing facilities of glucocorticoid-, mineralcorticoid- and androgen-type steroid hormones [31][32][33]. As shown in figure 1.4 it is a vertebrate-P450-type ferredoxin.

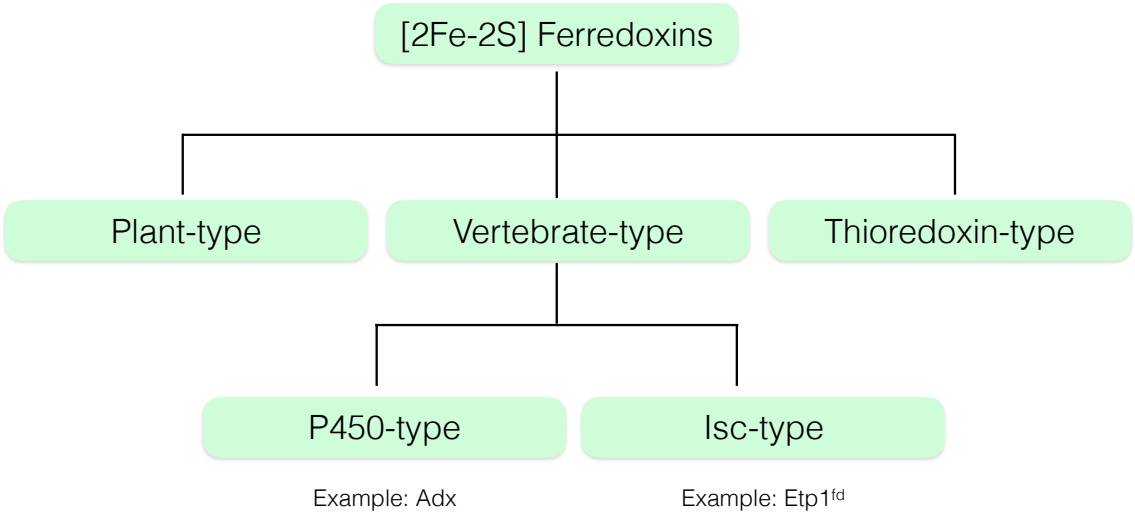

Figure 1.4: Overview of the different classes of [2Fe-2S] ferredoxins and the typical allocation of Adx and Etp1^{fd}.

Adx transfers electrons from NAD(P)H-dependent adrenodoxin reductases (AdR) to CYP11 family CYPs in adrenal mitochondria as well as to microsomal CYPs of families CYP17, CYP19 and CYP21 upon reconstruction. CYP11A1 is capable of cleaving the side chain from cholesterol in three consecutive hydroxylation steps

resulting in the precursor steroid hormone of all mammals: Pregnenolone. Additionally Adx mediates electrons to CYP11B1 and CYP11B2 for subsequent production of cortisol (glucocorticoid) and aldosterone (mineralglucocorticoid) [34][35][36]. Functionally, Adx is a [2Fe2S] type ferredoxin, whereby the side chains of four cysteines as well as the two bridging sulfur atoms form an approximately tetrahedral coordination around each of the two iron atoms of the iron-sulfur cluster. This is known from several X-ray crystallographic structures of oxidized bovine Adx (pdb entry codes 1AYF, 1CJE). Of interest are also the one of the isolated human Adx (3P1M) and that of the complex with CYP11A1 (3N9Y), showing the binding interface [37]. Comparison of the structural changes upon reduction of bovine Adx determined by NMR spectroscopy suggests that the largest deviations occur in the region 68-85, which comprises the primary interaction region to the cytochrome [38]. In experimental and theoretical studies mostly human and bovine Adx are used. Both their mRNAs encode for a N-terminal extension of the mature protein. In bovine Adx the extension is about 58 amino acids long and serves as a mitochondrial targeting sequence, which is cleaved upon mitochondrial entry. The resulting mature ferredoxin weighs about 14.4 kDa and has a length of 128 amino acids. An Adx-like ferredoxin from *Schizosaccharomyces pombe* can be found as being part of the electron transport protein 1 (Etp1) [39]. Etp1 is composed of an approximately 400 amino acid long N-terminal (homologous to COX15 in humans) and a C-terminal ferredoxin domain (homologous to Adx) part named $Etp1^{fd}$. Its native task includes the support of assembling iron-sulfur clusters as well as the synthesis of heme A [40]. In 2002 the ability of electron transport in steroid hydroxylation processes in bacterial and mammal P450s was discovered [39]. In certain reactions Adx can be replaced by $Etp1^{fd}$. In addition to that there are other P450s and lsc-associated ferredoxins that share a surprisingly high sequence identity (even having different physiological functions, e.g. Adx is quite similar to $Etp1^{fd}$ (sequence similarity of 39% [40]) or to Pdx (sequence similarity of 40% [41])). Also they share some highly conserved regions which is shown in figure 1.5.Three of them are placed around the iron-sulfur clusters of the proteins.

To quantify for the efficiency of electron transfer from a ferredoxin to a P450 the midpoint redox potential indicated in mV can be used. From the thermodynamic point of view it is obvious that the lower the value, the faster the electron transfer and the faster the reaction should be. An overview over the previous mentioned ferredoxins is given in table 1.1.

1.3 Steroid Hormones

Steroids are small molecules that have regulatory functions and belong to the group of lipids. They all have a characteristic base structure consisting of 17 carbon atoms arranged in three cyclohexane rings and one cyclopentane ring (fig. 1.6). Molecules

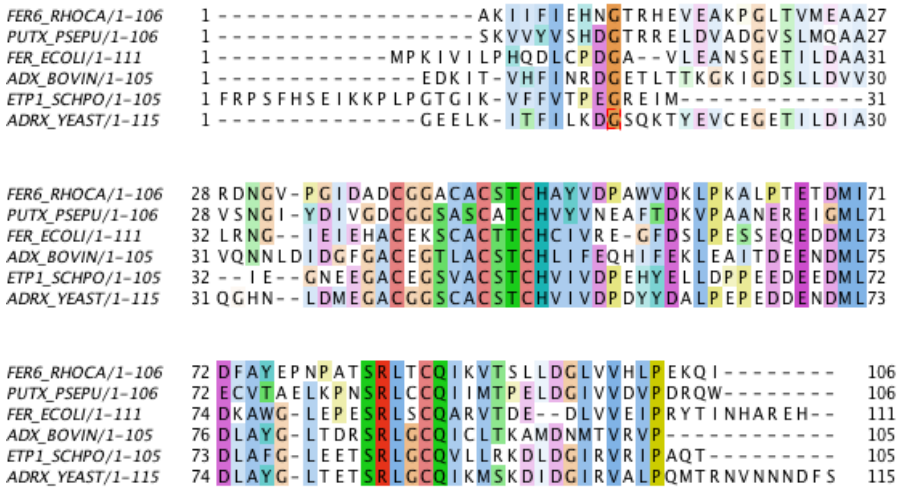

Figure 1.5: Sequence alignment of different ferredoxins. There are three conserved regions marked in the picture (bold and pale colors indicate highly and lower conserved regions, respectively, within the sequence; colors are just for clarification).

Ferredoxine (Organism)	Redox potential[mV]	PDB id
Adx WT (Bos taurus) [42]	-274	1AYF chain B
4-128 (Bos taurus) [43]	-267	
4-114 (Bos taurus) [44]	-342	
4-108 (Bos taurus) [45]	-344	
Etp1 (S. pombe) [46]	-353	2WLB chain B
FdX (E.coli) [47]	-380	1I7H chain A
FdxE / FdVI (R. capsulatus) [48]	-306	1UWM chain A
PdX (P. putida) [49]	-240	1XLQ chain A

Table 1.1: Overview of different P450 type and lsc-type Ferredoxins with corresponding redox potentials

like sterols, bile acids, steroid hormones, vitamin D and also many other secondary metabolites again belong to the group of steroids. They were found and identified in animals, plants and as well as in lower eukaryotes such as yeast or fungi [50, 51]. In humans the steroid hormones are synthesized in the adrenal glands or gonads. They steer the water/electrolyte homeostasis (aldosterone), the energy mobilization and immune response (cortisol) or account for sexual characteristics (testosterone, estrone). Even in gestational reactions (progesterone), cell communication processes and cell proliferation they play a major role [52]. The central steroid hormone is cholesterol. It is the common precursor to all mammalian steroid hormones and all others evolve from it. A rough scheme of the complexity involved in steroid hormone synthesis originating from cholesterol is shown in figure 1.7. As mentioned they play a central role as regulatory substances. Hence they are very interesting for pharma-

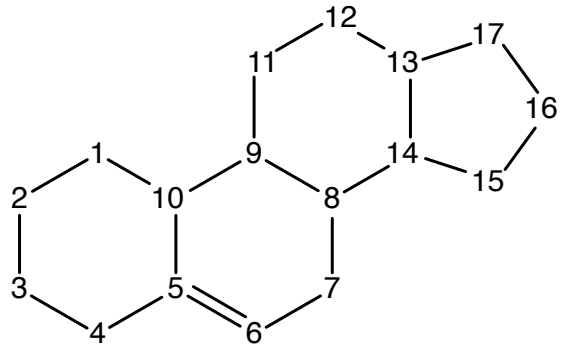

Figure 1.6: Basic structure of steroids. Consisting of 17 carbon atoms arranged in three cyclohexane rings and one cyclopentane ring

ceutical application. Representing one of the largest sectors of the pharmaceutical industry (besides antibiotics), they list more than 300 approved drugs [53, 54]. In addition to it the alternatively hydroxylated compounds have advantages towards their natural counterparts such as lowering occurring side effects, higher potency or other activating or inhibiting properties [55]. A big challenge in future research is to create new efficient ways to synthesize steroid hormones by biotechnological means.

1.4 Aim and Scope of This Work

The global aim of this work was the increase of selectivity and activity of P450 enzymes and the optimal adaption of redox partner interactions. The first part includes engineering of well known CYPs like CYP106A2 towards an efficient and selective mode to increase the selectivity of certain substrate conversions. Here we focused mostly on progesterone and its diverse hydroxylation possibilities. To guarantee an efficient way of electron transfer its redox partners must fit together nicely. Therefore engineering on the surface and the inner structure was performed in order to increase the fitting of the surfaces. Our suggestions were tested and partially verified in the wet lab.
The thesis first gives an introduction on the biological background. Here an up to date overview on CYP P450s, electron transfer proteins as well as common ligands is provided.
The following chapter deals with the methodology that was used to theoretically investigate and tackle occurring problems for optimizing protein-protein interactions or protein-ligand interactions in P450 specific environments at the computational level. It provides the theoretical background that needs to be understood to apply the right methods for the given cases.
In chapters three to six I show the results of published and soon to be published work in joint cooperations with our coworkers. Chapter three summarizes the investiga-

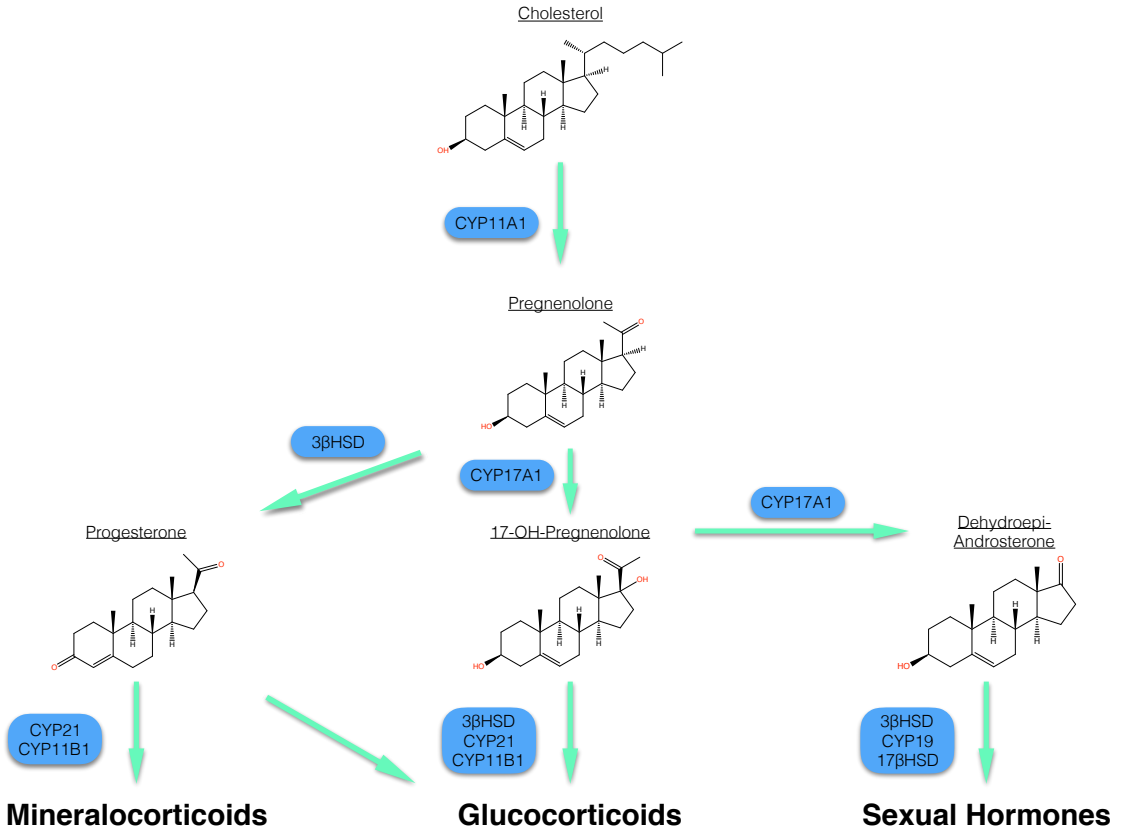

Figure 1.7: Simplified scheme of steroid hormones synthesis pathways originating from cholesterol.

tion of different combinations of ferredoxins and ferredoxin reductases towards an optimal redox system for progesterone synthesis through CYP106A2. Here we were also able to show why Etp1 binds more optimal to CYP106A2 than Adx and suggested a mutation to improve Adx' binding to CYP106A2. The results summarized in chapter four and published in *Steroids* [55] suggested theoretically derived mutations in the amino acid sequence of CYP106A2 for shifting 15β-hydroxy-progesterone production towards 6β-, 9α- and 11α-hydroxy-progesterone. Subsequent Molecular Dynamics simulations were carried out to show the behaviour of progesterone in the binding pocket over time. Closing the section on improvement of Adx' redox potential through mutations is done in chapter five. Here an approach from 1997 of Zhou [56] was revisited. They state that the change of the direction of the dipole moment can increase or decrease the redox potential on iron-sulfur cluster proteins, respectively. A theoretical investigation was performed with suggestions for mutations on Adx. An experimental verification still needs to be carried out. Chapter six deals with redox complexes of electron transfer proteins that were fused with different linker constructs. Here we were able to explain why certain linkers improved the performance of the fused proteins whereas others did not.
A conclusion at the end summarizes my efforts and provides also insights into failed

attempts to generate an universal system for predicting mutations to improve the redox potential for Adx. Subsequently I give an outlook about possible future work in this field.

Chapter 2

Materials and Methods

In this chapter I will describe the different computational methods that were used during my thesis. It introduces the concept of molecular dynamics (MD) simulations and guides through the basic physics lying underneath ending up in the full algorithm that performs such a simulation. This is followed by the introduction to Docking and density functional theory. The chapter is rounded up by an overview over the Adaptive Poisson-Boltzmann Solver developed by Nathan Baker and co workers [57].

2.1 Molecular Dynamics Simulations

MD simulation is a molecular modeling technique to describe the dynamics, conformational changes or thermodynamics in a molecular system. They end up in a trajectory that documents how the positions and velocities of particles change inside a defined time frame. The systems configurations are generated by the numerical solution to *Newton's second law*:

$$F = ma \tag{2.1}$$

where the force F is a result of mass of the particle times its acceleration. Since the acceleration is the first derivative of the velocity v and the second derivative of position r with respect to time t one can rewrite equation 2.1 explicitly for each particle as:

$$F_{r_i} = m_i \frac{d^2 r_i}{dt^2} \tag{2.2}$$

Due to the size of the systems and the coupling between the particles equation 2.2 cannot be solved analytically. The integration of the equation can be solved with a numerical solution such as *finite differences method*. Here the dynamic properties like acceleration, velocities etc. and positions are approximated with a Taylor series expansion.
In molecular dynamics simulations the most common method for solving Newton's

equation of motion is the *Verlet* algorithm [58]. The *Verlet* algorithm computes the positions of the particles for the next time step from the position of the particle from the previous time step $t - \Delta t$:

$$r(t - \Delta t) = r(t) - \Delta t v(t) + \frac{1}{2}\Delta t^2 a(t) - \dots \tag{2.3}$$

and the acceleration and position of the particle from the actual time step t:

$$r(t + \Delta t) = r(t) + \Delta t v(t) + \frac{1}{2}\Delta t^2 a(t) + \dots \tag{2.4}$$

combined the two equations result in:

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + \Delta t^2 a(t) \tag{2.5}$$

A drawback as it is in this basic form of the algorithm is the lack of explicit naming of the velocity of the particles. The velocities are important when it comes to calculation of the kinetic energy of the system and thus the temperature. A variation of the *Verlet* algorithm that overcomes this difficulty is the *leap-frog* algorithm [59]. It calculates the position of the next time step from the particles actual position and the velocity from a time step $t + \frac{1}{2}\delta t$:

$$r(t + \Delta t) = r(t) + \Delta t v(t + \frac{1}{2}\Delta t) \tag{2.6}$$

Before that the velocity at that time step are first updated:

$$v(t + \frac{1}{2}\Delta t) = v(t - \frac{1}{2}\Delta t) + \Delta t a(t) \tag{2.7}$$

Setup of a MD simulation requires an initial configuration of the system. A starting geometry can be derived from a theoretical model, an experiment (crystal structure) or combination of both. The assignments for the velocities is usually done by randomly picking values from a *Maxwell-Boltzmann* distribution, which is essentially a Gaussian distribution:

$$p(v_{ix}) = (\frac{m_i}{2\pi k_b T})^{\frac{1}{2}} \times exp[-\frac{1}{2}\frac{m_i v_{ix}^2}{k_b T}] \tag{2.8}$$

Equation 2.8 gives the probability that atom i with mass m_i has a velocity of v_{ix} in x direction at temperature T.
Generally before a MD simulation itself, some preparatory steps need to be carried out. A typical MD simulation comprises of the following parts:

1. Define and prepare starting geometry
2. Energy Minimization
3. Equilibration

4. Production MD
5. Analysis

Depending on the kind of experiment first one needs to define and curate (check for missing atoms etc.) the starting geometry (e.g. crystal structure from Protein Data Bank (www.rcsb.org [60])). In subsequent steps the experiment gets discretized and specified in detail. First the appropriate force field has to be chosen which has to fit the users needs. GROMACS [61] for example lets the user choose from a range of different force fields (e.g. AMBER[62], CHARMM[63], GROMOS[64] or OPLSAA[65]). After that a spatial specification is needed where the user decides for a shape where and in what medium the molecule is simulated in (solvents, vacuum). To obtain a net neutral charge of the system, positively or negatively charged ions can be added. From that on the energy minimization step can be started where steric clashes or inappropriate geometries are resolved (they emerge for example from the principal limitations of crystallographic coordinates of atoms). The molecule is now mostly optimized within itself but not with respect to its surrounding. Necessary Equilibration is performed in two phases. The first phase comprises the temperature setting whereas the second comprises the pressure settings of the whole system. Under a NVT (**N**umber of particles; **V**olume; **T**emperature) ensemble the system is stepwise heated up to the temperature the simulation should be performed at. This is usually realized via a certain thermostat (e.g. [66]). To reflect that also the (atmospheric) pressure does not vary, an analog equilibration proceeds. Here the preferred pressure is conducted under a NPT (**N**umber of particles; **P**ressure; **T**emperature) ensemble via a certain barostat (e.g. [67]).
The now equilibrated system can be used to run a production MD simulation for a defined time frame.
The time needed depends on the size of the molecule(s). E.g. a equilibrating time of 50-100 ps for NVT and NPT suffices a 130 amino acid long protein. The subsequent MD simulation should last at least 20 ns. A standard time step size is 2 fs (for this case the simulation comprises of 10,000,000 computation steps). Since such a computation would last weeks to months, MD simulations are usually performed on computer clusters or supercomputers that are highly parallelizable. Subsequent

Molecular motion	Time [s]
Atomic fluctuations/Side-chain motion	10^{-15} - 10^{-12}
Loop & Rigid-body Motion	10^{-9} - 10^{-6}
Domain and Subunit Motion	10^{-6} - 10^{-3}
Protein (Un-) Folding	10^{-3} - 10^{4}

Table 2.1: Scheme of captured motion of molecules for different time spans

to simulation the computed trajectories (snapshots of the simulation at fixed time frames) are analyzed. Here for example the RMSD value over time or the spatial behaviour of the molecule(s) is(are) of interest.

2.2 Docking

Computer aided molecule design has not only become a crucial part in development of new drugs but also is a helpful method for biotechnological purposes. An important tool in this area is molecular docking [68]. Docking allows for investigation of molecular interactions in terms of geometry as well as electrostatics and estimates the binding affinity and the mutual orientation between a receptor and ligand by means of the underlying scoring function. The conformational space that has to be sampled is huge, even if both protein and ligand are treated as rigid bodies. Not only the three degrees of freedom from ligand relative to the protein have to be considered but also the additional internal degrees of the ligand itself that arise from rotatable bonds. Additionally, induced-fit phenomena are neglected, when considering the protein as rigid body. The subsequent evaluation by a scoring function [69] of each protein-ligand conformation should eventually come up with a pose that could occur in nature. But there are different ways how one can evaluate a score for a pose. Some scoring functions are based on the underlying force field, on knowledge or are even empirical and usually include a weighting for individual terms that were fitted and trained on known sets of protein-ligand complexes. Therefore not every evaluation will work as desired on each kind of receptor ligand complex [68] [70].
Since the early days of docking with rigid protein and ligand (e.g. DOCK [71]), algorithms more and more incorporated full flexibility for ligands and partially receptors. In the following paragraph concepts behind automated docking will be explained by means of the widely used AutoGrid4 and AutoDock4 [72] software. The visual molecule inspecting program AutoDockTools can produce input files for both command line tools AutoGrid4 and AutoDock4. Hereby the the actual search pace will be exactly defined. In the beginning a grid placed into the receptor where on each grid point the electrostatic potential between every atom type occurring in the ligand and the receptor is calculated (grid density can be varied). After the precomputation step the actual docking step can be performed. The optimization of the receptor-ligand pose and their corresponding interaction energy is performed by a Lamarckian Genetic Algorithm (LGA) [73]. Generally a GA implements Darwins law of evolution by natural selection. A conformation is assessed with a fitness score via the scoring function and the best ones pass on their *genes* (conformation) to the next generation. The conformation then will be recombined, randomly mutated or left unchanged. The score estimates the binding free energy ΔG of the complex in a solvated state with a semi-empirical free energy force field. The force field in particular is already trained and parameterized using a large number of protein-inhibitor complexes. It uses six pair-wise evaluations (V) and an estimation of the conformational entropy lost due to binding (ΔS_{conf}):

$$\Delta G = (V^{L-L}_{bound} - V^{L-L}_{unbound}) + (V^{P-P}_{bound} - V^{P-P}_{unbound}) + (V^{P-L}_{bound} - V^{P-L}_{unbound} + \Delta S_{conf}) \quad (2.9)$$

For each term in 2.9 energetic terms have to be computed including evaluations for repulsion/dispersion, hydrogen bonding, electrostatics and desolvation, respectively.

$$\begin{aligned}\Delta G = &\Delta G_{vdW} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) \\ &+ \Delta G_{hbond} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}\right) \\ &+ \Delta G_{elec} \sum_{i,j} \frac{q_i q_j}{e(r_{ij}) r_{ij}} \\ &+ \Delta G_{sol} \sum_{i,j} (S_i V_j + S_j V_i) e^{-r_{ij}^2/2\sigma^2} \\ &+ \Delta G_{tor} N_{tor}\end{aligned} \tag{2.10}$$

The first term describes the repulsion/dispersion value (van der Waals interaction) between all atom pairs i and j using a Lennard-Jones potential. To account for hydrogen bonding a 12-10 potential is used with the integration of *E(t)*, defining the direction of the bond with angle *t*. The third term denotes the Coulomb interactions and the fourth term the desolvation potential. Hereby S_i gives the required energy that is needed to transfer atom i from a solvated state to a buried one. V_i estimates the amount of desolvation due to ligand complexation. Upon bonding, there is a loss of entropy that is proportional to the number of sp^3 bonds, N_{tor}. The weighting constants ΔG were derived from a set of experimentally known binding constants.

2.3 Density Functional Theory

To almost all physical properties of a system, the electronic ground state is fundamentally important. The Density Functional Theory (DFT) approaches a solution to the partial differential equation that describes the behaviour of atoms and molecules over time, namely the *Schrödinger equation*. In its most general form, the time-dependent form reads as:

$$i\hbar \frac{\delta}{\delta t} \Psi(r,t) = \hat{H} \Psi(r,t) \tag{2.11}$$

where i denotes the imaginary unit, $\hbar$ the *reduced Planck constant*, $\frac{\delta}{\delta t}$ the partial derivative with respect to time of the wave function Ψ of the system with r and t being the spatial vector and time. $\hat{H}$ is the *Hamiltonian operator*. Through minimization of the total energy of the system the ground state wave function can be found. The wave function Ψ is a many-body function that contains coordinates for all degrees of freedom of a system. Analytical solutions to the *Schrödinger equation* cannot exceed the two-body form such as one hydrogen atom with one nucleus and one electron. The complexity increases by three times the degrees of

freedom. For example for an oxygen atom with its electrons this would mean to solve a 24-dimensional partial differential equation [74].
Classical mechanics are not capable of describing or solving this problem. To be able to do that one has to specify the location of an atom, hence the location of the nucleus (protons and neutrons) and the electrons. Since protons and neutrons are much heavier then the electrons (approx. 1800 times [75]), the electrons respond much more rapidly to changes in their surroundings. Applying the *Born-Oppenheimer* approximation [76] the computation of the wave function can be done by separating the wave function into its electronic and nucleic components.

$$\Psi_{tot} = \Psi_{elec} \times \Psi_{nuc} \tag{2.12}$$

Following the *Born-Oppenheimer* approximation and neglecting the nuclei one could write the Hamiltonian for the electrons from equation 2.11 as:

$$i\hbar\frac{\delta}{\delta t}\Psi(r,t) = [\frac{\hbar^2}{2m_e}\sum_i \nabla_i^2 + \sum_i V_{ext}(r_i) + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{|r_i - r_j|}]\Psi(r,t) \tag{2.13}$$

where the $\hat{H}$ is replaced by terms for the kinetic energy of each electron, the interaction energy between the electrons and the collection of nuclei and the interaction energy between different electrons. In 1964 Hohenberg and Kohn [77] stated that the external potential E is a unique functional of the electron density $\rho(r)$. A functional allows a function to be mapped to a number. This leads to a simplification that reduces the number of variables to handle from 3N (N as number of electrons) to 3 (electron density). In DFT the energy functional is defined as follows:

$$E[\rho(r)] = \int V_{ext}(r)\rho(r)dr + F[\rho(r)] \tag{2.14}$$

$V_{ext}(r)$ accounts for the interaction of electrons with an external potential and $F[\rho(r)]$ denotes the sum of the kinetic energy. Essentially the minimum energy complies with the ground state electron density. Since the wave function $F[\rho(r)]$ is not known, Kohn and Sham 1965 [78] came up with the approximation:

$$F[\rho(r)] = E_{KE}[\rho(r)] + E_H[\rho(r)] + E_{\chi C}[\rho(r)] \tag{2.15}$$

Here the wave function is equal to the sum of the kinetic energy (from non-interacting electrons with the same density as the real system) which writes as:

$$E_{KE}[\rho(r)] = \sum_{i=1}^{N}\int \Psi_i(r)(-\frac{\nabla^2}{2})\Psi(r)dr \tag{2.16}$$

Electron-electron Coulombic energy which is also known as Hartree electrostatic energy is described by the second term of equation 2.15. The Hartree approach fails to incorporate the correlation of electron movement and therefore is only a part

of equation 2.15. The Hartree energy term integrates the interaction between two charge densities and can be rewritten as:

$$E_H[\rho(r)] = \frac{1}{2}\int\int \frac{\rho(r_1)\rho(r_2)}{|r_1 r_2|} dr_1 dr_2 \tag{2.17}$$

The last term in equation 2.15 denotes contributions from exchange and correlation that are typically computed from a set of functions, namely basis functions that are linearly combined to create molecular orbitals. In practise, the amount of correlation and exchange can be optimized to the problem at hand and therefore a multitude of DFT functional has been developed.
If we now let equations 2.16 and 2.17 flow together we end up with the full Kohn Sham equation to compute the energy of a N-electron system:

$$\begin{aligned} E[\rho(r)] = & \sum_{i=1}^{N} \int \Psi_i(r)(-\frac{\nabla^2}{2})\Psi(r)dr \\ & + \frac{1}{2}\int\int \frac{\rho(r_1)\rho(r_2)}{|r_1 r_2|} dr_1 dr_2 \\ & + E_{\chi C}[\rho(r)] \\ & - \sum_{A=1}^{M} \int \frac{Z_A}{|r - R_A|}\rho(r)dr \end{aligned} \tag{2.18}$$

2.4 Adaptive Poisson-Boltzmann Solver

To calculate electrostatic interactions between proteins or proteins and ligands I used the Adaptive Poisson Boltzmann Solver software developed by Nathan Baker [57]. The basics for electrostatics consist on the Poisson equation and the Boltzmann distribution of ions in solution and continuum models.
To investigate the electrostatic properties of biological macromolecules (e.g. proteins) the Poisson equation is widely used. Hereby the solute is treated as body of constant dielectric (between 2 and 20) whereas the solvent is represented as continuum of high dielectric. The Poisson equation relates the variation in the potential ϕ within dielectric constant ϵ to the charge density ρ:

$$\nabla\epsilon(r)\nabla\phi(r) + 4\pi\rho(r) = 0 \tag{2.19}$$

Now the Boltzmann distribution describes a probability that a system is in a certain state that is defined by a function of that states temperature, energy and number of atoms. The distribution is defined as:

$$n(r) = Ne^{-\frac{q\phi(r)}{kT}} \tag{2.20}$$

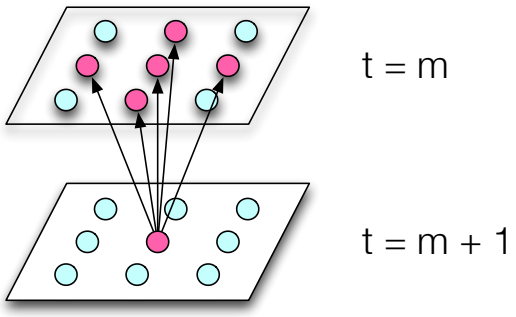

Figure 2.1: Stencil for discretization in two-dimensional case. Value of timestep t is dependent on time step $t - \Delta t$

Merging equations 2.19 and 2.20 result in the well-known Poisson-Boltzmann equation (PBE), the most frequently used continuum modeling equation for protein-solvent systems:

$$\nabla[\epsilon(r)\nabla\phi(r)] - \kappa'^2\phi(r) + \frac{4\pi\rho(r)}{kT} = 0 \tag{2.21}$$

As any other linear or nonlinear partial differential equation the PBE is hard to solve analytically. This is only feasible to simple geometries such as spheres, cylinders or plains. Numerical solutions such as finite differences, elements or volumes methods for approximation are widely used. The steps for solving the PBE involve:

1. Discretization of the spatial domain
2. Assignment of values for electrostatic potential, charge density, dielectric constant and ionic strength to each discrete grid point
3. Determination of derivatives for PBE by numerical methods

An example for a discretization method would be the finite differences method. Here the actual potential of a grid point is dependent on the surrounding grid points as well as the previous time step. In a two-dimensional space (for simplification) that would be four neighbouring grid points as shown in figure 2.1. The potential at each grid point would be:

$$\phi_0 = \frac{\sum \epsilon_i\phi_i + 4\pi\frac{q_0}{h}}{\sum \epsilon_i + \kappa_0'^2 f(\phi_0)} \tag{2.22}$$

The total energy of a system is computed from two necessary simulations, namely the system in vacuum and in solution. The solvation energy is computed from point charges located on the particles and different potentials:

$$\Delta G_{elec} = \frac{1}{2}\sum_1 q_i(\phi_i^{80} - \phi_i^1) \tag{2.23}$$

A very simple to use tool to compute electrostatic properties for biomolecules is the Adaptive Poisson Boltzmann Solver (APBS). It can be used via web server or

locally. Configuration of input files, like assignment of charges, can be done by a co-program named pdb2pqr [79] [80]. This generates charge files ready to use for APBS. APBS solves the PBE with a finite elements approach.

Chapter 3

Towards an optimal redox chain for CYP106A2 from Bacillus megaterium ATCC 13368

The results and discussion sections (3.1 and 3.2) as well as figures 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 3.10 and table 3.1 were adapted unaltered from our submitted manuscript [81], whereas the experimental section of our coworkers is omitted here. The work presented in this chapter emerged from a collaboration together with the group of Prof. Dr. Rita Bernhardt and is in revision process. It aims towards an optimal redox chain for CYP106A2 from *Bacillus megaterium* ATCC13368 for progesterone conversion.
Progesterone was chosen as substrate since it is well studied as a substrate for CYP106A2. Besides the main product 15β-hydroxy-progesterone other monohydroxylated products like 6β-, 9α- or 11α-hydroxyprogesterone can emerge in low quantities [82, 83]. Additionally polyhydroxylated progesterone products were also found.
The most preferred redox system used for substrate conversion with CYP106A2 is bovine Adrenodoxin (Adx) and its reductase Adrenodoxin reductase (AdR). However, in order to achieve an optimal redox system with a higher yield of substrate conversion, 11 different combinations of the following ferredoxins or flavodoxins were tested:

1. Ferredoxins/Flavodoxins:
   (a) Adx (*Bos taurus*)
   (b) Adx(4-108) (*Bos taurus*)
   (c) Etp1(516-618) (*Schizosaccharomyces pombe*)
   (d) Fdx2 (*Bacillus megaterium*)

   (e) FldA (*Escherichia coli*)

   (f) YkuN (*Bacillus subtilis*)

2. Reductases

   (a) AdR (*Bos taurus*)

   (b) Arh1 (*Schizosaccharomyces pombe*)

   (c) BmCPR (*Bacillus megaterium*)

   (d) Fpr (*Escherichia coli*)

It was shown that the choice of redox partner not only affects the rate of substrate conversion but also strongly influences the product pattern in terms of multi-product conversion. The best fitting system yielded a high rate of 15-β-hydroxy-progesteron and kept the conversion of unwanted multihydroxylated products low.
Manual investigation of the crystal structures of Etp1 and Adx and subsequent docking experiments lead to the identification of significant alterations in their similar three dimensional conformation that probably account for faster electron transfer rates of Etp1.

3.1 Results

To test the ability of the various redox chains for their efficiency to transfer electrons from NAD(P)H to CYP106A2, in vitro experiments using various heterologous redox systems were conducted. The typically used redox system for CYP106A2 is the truncated form of bovine adrenodoxin, Adx(4-108), and the corresponding reductase, AdR, which was therefore used as "parent" system. We investigated different naturally occuring redox protein systems as well as combinations between them. The redox system of *Schizosaccaromyces pombe*, consisting of the ferredoxin Etp1 (electron transfer protein 1) and ferredoxin reductase Arh1 (adrenodoxin reductase homolog 1), is completely soluble and easily expressable [84, 85]. With the bacterial CYP105A1 from *Streptomyces griseolus*, the reduction efficiency was shown to be 100% for the combination of Etp1/Arh1 compared with only ∼25% when using Adx(4-108)/AdR [86]. The lower redox potential of Etp1 of -353 mV instead of -344 mV is likely to be a major driving force that would explain this. Other possiblities to investigate are the full-length bovine adrenodoxin, which works better with CYP11A1 [87], than the truncated form (Adx(4-108)) that is mostly used for CYP106A2, as well as redox partners from different Bacillus strains, such as the ferredoxin isolated from *Bacillus megaterium DSM319*, Fdx2 [88], the typical short-chain flavodoxin YkuN from *Bacillus subtilis* [89], the *Bacillus megaterium* cytochrome P450 reductase BmCPR [90], or redox partners from *Escherichia coli* like the flavodoxin FldA and NADPH-dependent flavodoxin/ferredoxin reduc-

tase Fpr [91, 92, 93] (fig. 3.1). First tests were conducted with a fixed ratio

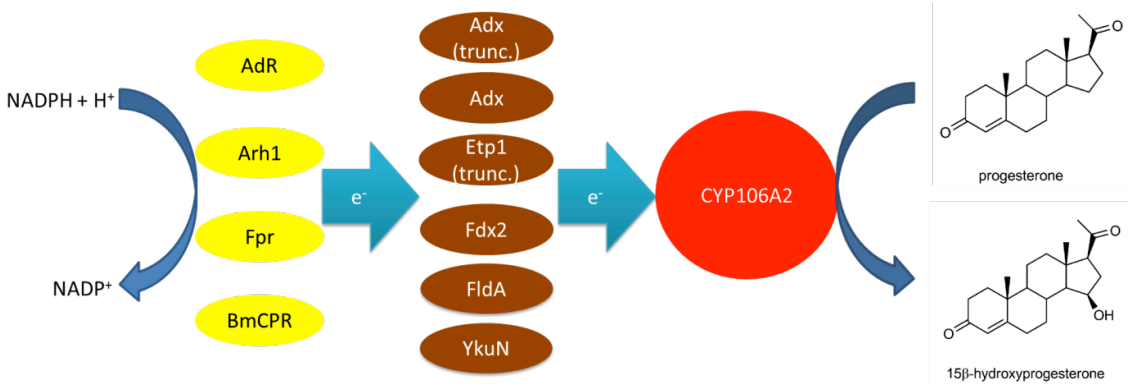

Figure 3.1: Schematic drawing of the investigated redox partner proteins. Redox equivalents are transferred from NADPH via a reductase (AdR, Arh1, Fpr, BmCPR) and a ferredoxin (Adx, Adx(4-108), Etp1(516-618), Fdx2), or flavodoxin (FldA, YkuN)) to the progesterone converting enzyme CYP106A2 (taken from [81])

of CYP:ferredoxin:reductase of 1:20:2 (exceptions for CYP:FldA:Fpr=1:50:50 and CYP:YkuN:Fpr=1:10:10) for 30 min to check the compatibility of redox proteins with CYP106A2 (fig. 3.2). The product patterns of the progesterone conversion differentiate clearly between the tested combinations of redox partners. Some combinations have similar overall conversion rates but produce more than 60 % polyhydroxylated progesterones (Adx(4-108)/Arh1), whereas others tend to give more than 70 % 15β-OH-P (Etp1(516-618)/AdR). Also redox partner combinations which have slower conversion rates are distinguisable (Adx/AdR). Using the above mentioned conditions (ratio of enzymes, temperature, time, in vitro components), it is clearly visible that the ferredoxin Adx(4-108) (combined with AdR) produces an abundance of polyhydroxylated products, whereas the wildtype Adx as well as the FldA have an overall lower conversion of progesterone (figs. 3.2 and 3.3). Etp1(516-618) is showing a very high conversion (similar to Adx(4-108)/AdR) but a low amount of polyhydroxylated progesterone products. The combination of Adx(4-108) and BmCPR displays a high conversion also exhibiting low amounts of polyhydroxylated products. Another way to get a deeper insight into the redox partner exchange is to look at the absolute amounts of converted progesterone and its products, which are displayed in figure 3.3. The combination of Etp1(516-618) and AdR was able to convert the total of 199 μM progesterone in the 250 μl in vitro assay within 30 min. Almost all truncated ferredoxins were able to convert between 180 and 190 μM progesterone. When Adx(4-108) or Etp1(516-618) were reconstituted with BmCPR, only 10 to 40 μM progesterone were not converted. Full-length Adx(1-128) or Fdx2 were able to convert between 118 and 143 μM progesterone to its products, whereas FldA (in combination with the corresponding reductase, Fpr) only converted 53 μM of the substrate. In contrast, when Fpr was combined with the flavodoxin from *Bacillus subtilis*, YkuN, 183 μM of progesterone was converted. All combinations (except FldA/Fpr) tested did not remarkably change the amount of

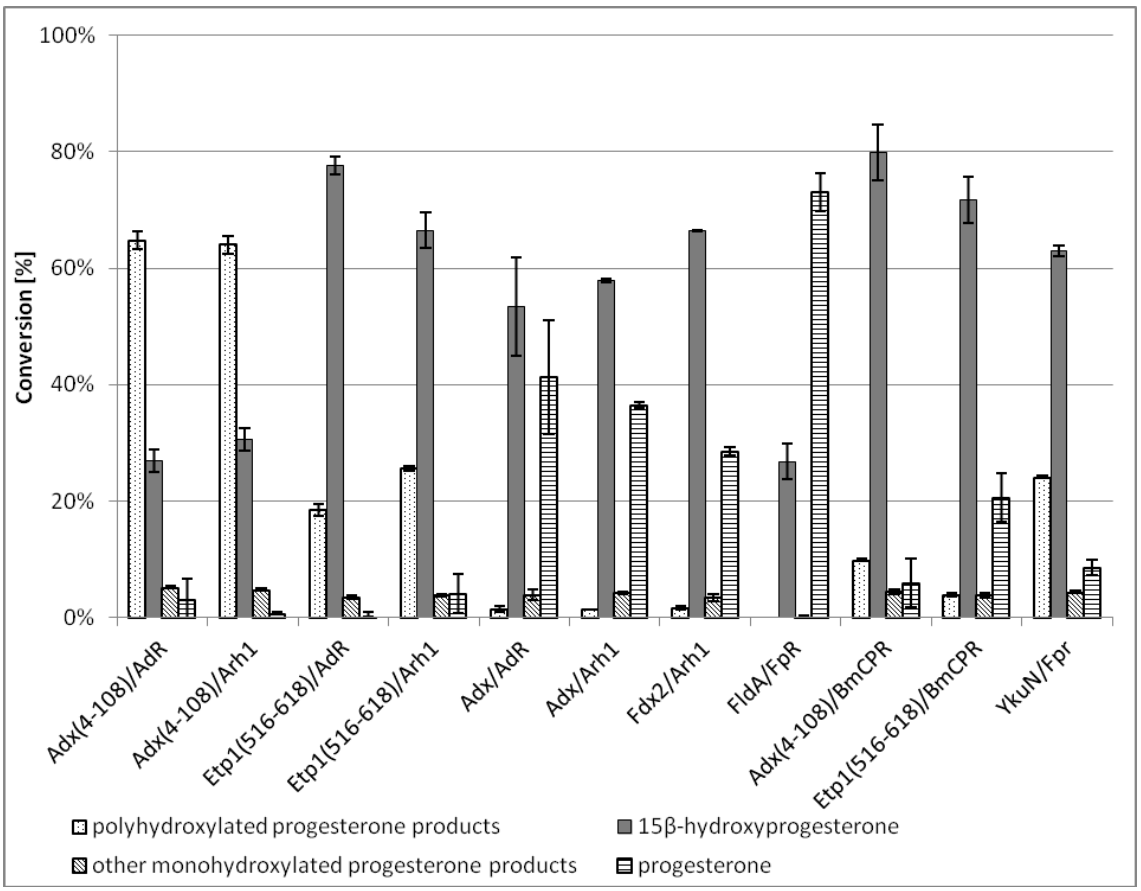

Figure 3.2: Comparison of in vitro conversions of progesterone with CYP106A2 and different redox partner combinations. Reactions were run for 30 minutes while keeping the ration of CYP:ferredoxin:reductase at 1:20:2 (exept for CYP:FldA/Fpr=1:50:50 and CYP:YkuN:Fpr=1:10:10); dotted bars indicate polyhydroxylated progesterone products, grey bars indicate the main product 15β-OH-P, diagonally striped bars indicate other monohydroxylated progesterone products (e.g. 11α-, 9α-, and 6β-hydroxyprogesterone), horizontally striped bars indicate the remaining substrate progesterone. The relative particular product level was calculated by using the relative peak area of the specific product compared to the total peak areas of educt and products (taken from [81]).

the other monohydroxlated progesterone products (6β-, 9α-, and 11α), which always was between 3 and 5 % under the conditions being tested. Switching the established bovine reductase, AdR, to the yeast Arh1 resulted in only little changes: 27 % of the total product consists of the main product, 15β-OH-P, when using AdR as reductase, and 31 % when using Arh1. Also, a difference in polyhydroxylated progesterone is not notable, which is 65 % and 64 % when using AdR and Arh1, respectively. In both cases the other monohydroxylated progesterone products (6β, 9α, 11α) were formed at a level of 5 %. When replacing the ferredoxin from Adx(4-108) to Etp1(516-618), the change in the product distribution is remarkable. The combination of Etp1(516-618) and AdR resulted in 78 % main product and only 18

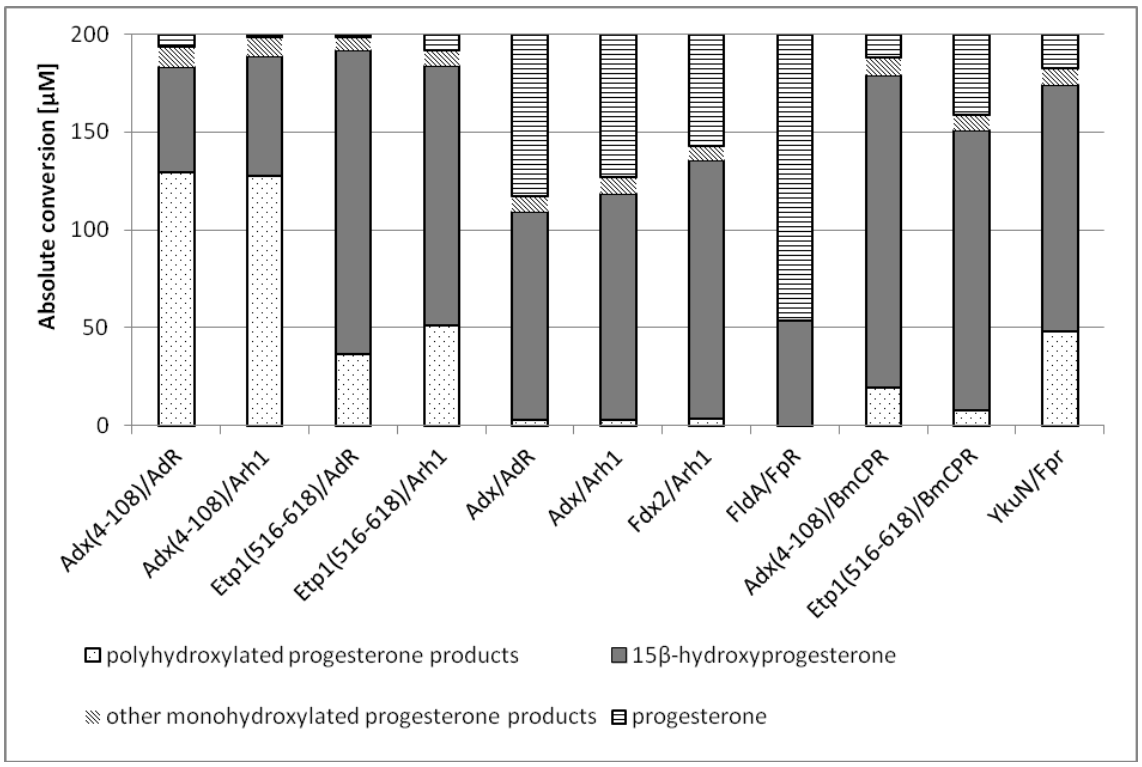

Figure 3.3: Comparison of the absolute amounts of detectable substrate and products after in vitro conversion of progesterone with CYP106A2 and different redox partner combinations. Reactions were run for 30 minutes while keeping the ration of CYP:ferredoxin:reductase at 1:20:2 (exept for CYP:FldA/Fpr=1:50:50 and CYP:YkuN:Fpr=1:10:10); dotted bars indicate polyhydroxylated progesterone products, grey bars indicate the main product 15β-OH-P, diagonally striped bars indicate other monohydroxylated progesterone products (e.g. 11α-, 9α-, and 6β-hydroxyprogesterone), horizontally striped bars indicate the remaining substrate progesterone (taken from [81]).

% polyhydroxylated progesterone, which is a 2,6-fold increase in the main product when compared with Adx(4-108)/AdR. Also the source of reductase seems to affect the product pattern. The use of Arh1 resulted in 66 % main product and 26 % poly-OH-progesterone compared with 78 % and 18 % when AdR was used (fig. 3.2). Applying the full-length Adx(1-128), also completely changes the product formation: there is still 41 % and 36 % substrate left (in combination with AdR and Arh1, respectively), but there is also 53 % and 58 % of 15β-OH-P formed. The amount of polyhydroxylated products is negligable. The use of Fdx2 from Bacillus megaterium results in 66 % main product and 29 % of unconverted substrate. No polyhydroxylated products were formed. In contrast, the redox partners originating from E. coli, FldA and Fpr, were not able to provide electrons to CYP106A2 when a ratio of CYP:ferredoxin:reductase of 1:20:2 was applied, which resulted in no product formation (data not shown). Therefore, the ratio was increased to 1:50:50, which lead to a conversion of 27% (all main product). Exchanging the FldA with the YkuN from

Bacillus subtilis (at a ratio of 1:10:10), a production of 24 % poly-OH-progesterone and 63 % main product was observed. Surprisingly, the use of the very recently cloned and characterized *Bacillus megaterium* reductase BmCPR [90] resulted in 80 % and 72 % 15β-OH-P and only 10 % and 4 % poly-OH-P formation, respectively. These data demonstrate that the application of different redox systems can alter not only the conversion rate of the substrate, but also the product distribution. To gain deeper insight into the time dependence of the progesterone conversion and product distribution by CYP106A2, the combinations of the ferredoxins Adx(4-108) and Etp1(516-618) with AdR and Arh1 were investigated in more detail.

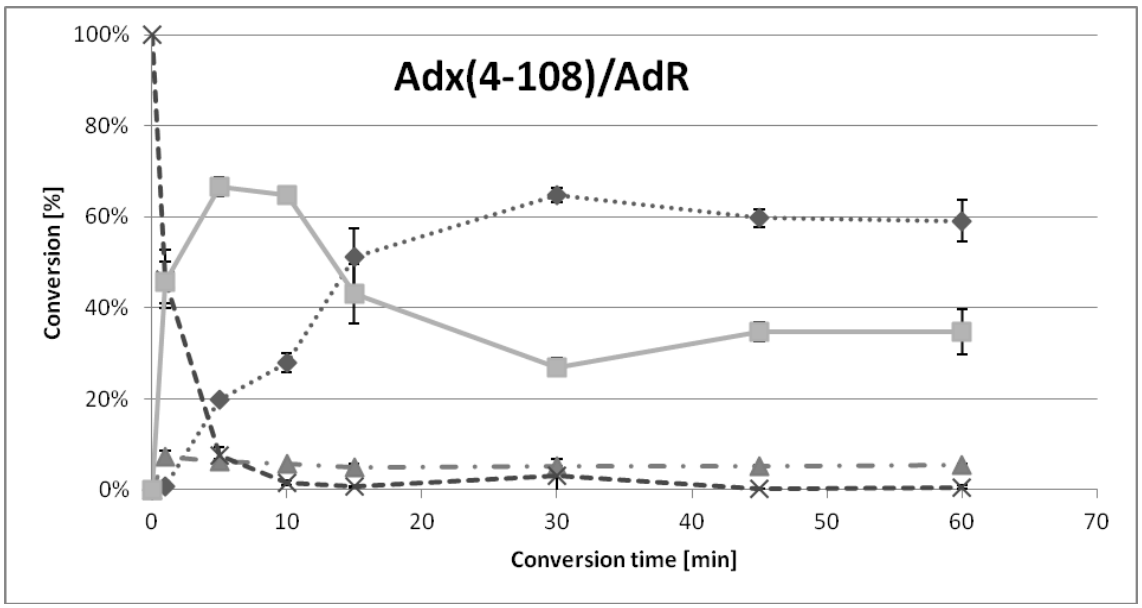

Figure 3.4: Time-dependent in vitro conversion of progesterone by CYP106A2 and Adx(4-108)/AdR. Conversions were stopped after 1, 5, 10, 15, 30, 45, or 60 minutes; dotted lines indicate polyhydroxylated progesterone products, grey lines indicate the main product 15β-OH-P, dashed with single dot lines indicate other monohydroxylated progesterone products, dashed lines indicate the remaining substrate progesterone. Within 10 minutes, 93 % of progesterone has been converted and 65 % of 15β-OH-P was produced. Further conversion leads to an increase in polyhydroxylated progesterone products and a decrease of the main product, 15β-OH-P. The percentage of other monohydroxylated did not change within the considered time interval (taken from [81]).

When using Adx(4-108)/AdR (fig. 3.4), the steep increase of 15β-OH-P formation up to 67% (5 min) is followed by a decrease to 34% after 60 min. At the same time, the amount of polyhydroxylated progesterone is slowly increasing from 7% after 5 min to 59% after 60 min. This suggests that the main product is partially further hydroxylated to result in polyhydroxylated end products. After changing AdR for Arh1 (see fig. 3.5), the time dependence of progesterone conversion was almost identical. First, the main product increased to 76% (5 min), dropping to 27% after 60 min, while poly-OH-progesterone is increasing up to 66% (60 min). Using Etp1(516-

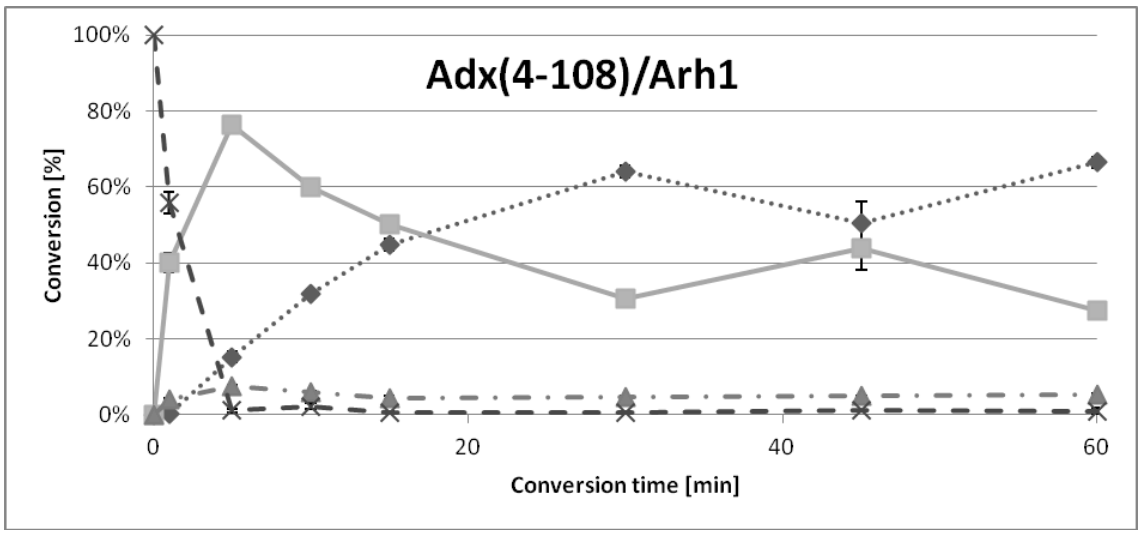

Figure 3.5: Time-dependent in vitro conversion of progesterone by CYP106A2 and Adx(4-108)/Arh1. Conversions were stopped after 1, 5, 10, 15, 30, 45, or 60 minutes; dotted lines indicate polyhydroxylated progesterone products, grey lines indicate the main product 15β-OH-P, dashed with single dot lines indicate other monohydroxylated progesterone products, dashed lines indicate the remaining substrate progesterone. The conversion of progesteron is almost complete after 5 min (1 % left), whereas the rate of 15β-OH-P (76%) decreases fast after 5 min while increasing amounts of polyhydroxylated progesterones are formed (up to 66 %). Also, further monohydroxylated products range at about 4 % over the time course of 60 minutes (taken from [81]).

618) as a ferredoxin (fig. 3.6) with AdR as corresponding reductase, a steep increase of 15β-OH-P formation to 75% within in the first 5 min is followed by a slow increase up to 81% after 60 min. In this case, a slow rise of poly-OH-progesterone (not reducing the amount of the main product) to 14% after 60 min occurs. Changing the reductase for Arh1 (fig. 3.7), the highest main product formation is observed after 10 min (78%), dropping slightly to 69% after 60 min. This is accompanied by an increase of poly-OH-prog from 6% after 10 min to 26% after 60 min.

Etp1(516-618) from Schizosaccharomyces pombe, which exhibits 81% sequence identity to the shortened, yet fully functional Adx(4-108), is also highly similar in its three-dimensional structure as the comparison of the high resolution crystal structures shows (see figure 3.8). Particularly in region around the iron-sulfur cluster any substantial deviations of the protein backbone are absent. Our docking studies, however, showed that subtle structural differences in the Adx(4-108) prevent it from adopting the same binding orientation to the cytochrome as it was found for Etp1 (see figure 3.9). It turned out that the residues around position 14 and 41 of Adx would sterically clash with the cytochrome. Furthermore, Tyr82 cannot form hydrogen-bonds to the cytochrome, whereas the corresponding Phe82 of Etp1 forms hydrophobic contacts, whereby penetrating deeply through the cytochrome's surface. There, residues Phe107, Pro359, and Leu356 encase the side chain of Phe82 (see

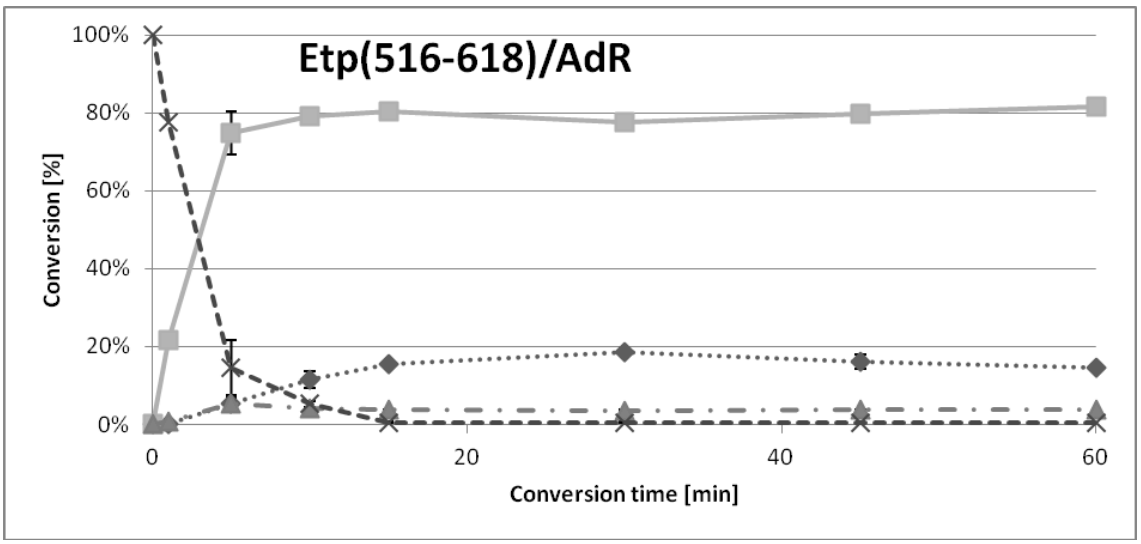

Figure 3.6: Time-dependent in vitro conversion of progesterone by CYP106A2 and Etp1(516-618)/AdR. Conversions were stopped after 1, 5, 10, 15, 30, 45, or 60 minutes; dotted lines indicate polyhydroxylated progesterone products, grey lines indicate the main product 15β-OH-P, dashed with single dot lines indicate other monohydroxylated progesterone products, dashed lines indicate the remaining substrate progesterone. Within 10 minutes of conversion, 95 % of the progesterone is converted and 79% of 15β-OH-P is produced. The amount of 15β-OH-P increases slightly up to 81 %, as polyhydroxylated products are stay between 11 % and 18 % total. Here, further monohydroxylated products range at about 4 % over the time course of 60 minutes (taken from [81]).

figure 3.10). As a consequence the distance between the heme iron of the cytochrome and the iron-sulfur cluster is substantially shorter in the complex with Etp1 (19 Å) than in that with Adx(4-108) (24 Å). The experimentally observed hydroxylation pattern of progesterone, in particular the formation of polyhydroxylated products, requires the reorientation of mono-hydroxylated progesterone. To be able to dissect this process of reorientation from that of the electron transfer by the ferrodoxins, we performed molecular dynamics simulation of 15β-OH-progesterone in CYP106A2. The correspoding results show that reorientation of the mono-hydroxylated product occurs within less than 100 ns. These molecular movements lead to conformations in which hydroxylation of other positions (11α-, 9α-, and 6β) are possible, assuming that a distance of less than 5 Å between the respective carbon atom and the iron of the heme is sufficiently close for reaction. To proof experimentally the influence of the mode of ferredoxin binding regarding the electron transfer velocity, stopped-flow measurements were performed. Differences in the reduction traces reflect the different binding behavior that were indicated by the conversion analysis and molecular docking calculations. The resulting apparent rate constants (k_{app}) were calculated and are presented in table 3.1.

Applying the redox system of Etp1(516-618)/Arh1, the apparent rate constant is 0.34 s^{-1}, whereas the combination of Adx(4-108) and AdR leads to a lower k_{app} of

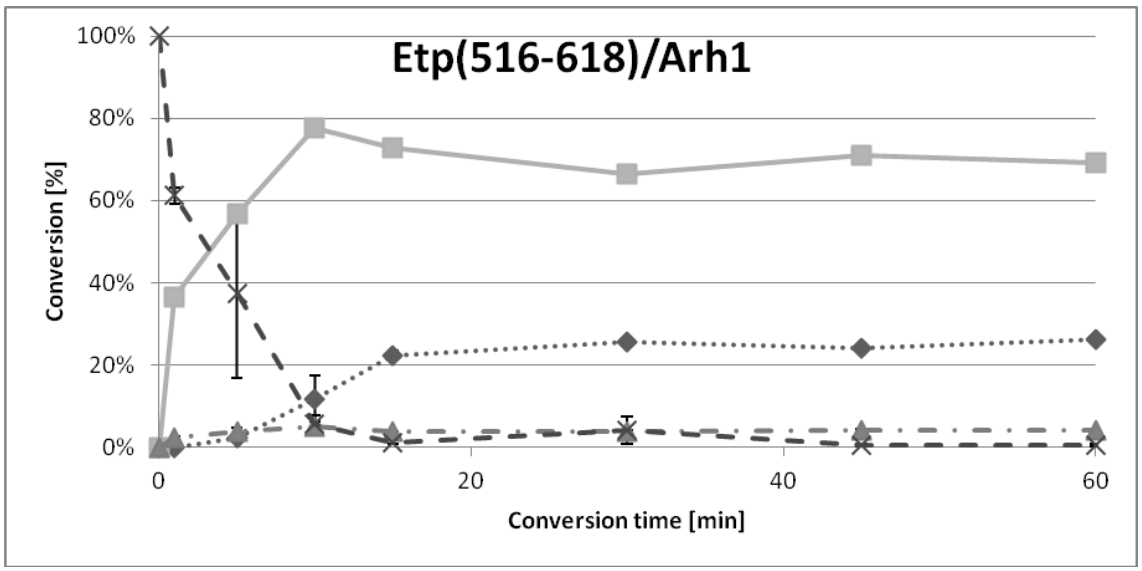

Figure 3.7: Time-dependent in vitro conversion of progesterone by CYP106A2 and Etp1(516-618)/Arh1. Conversions were stopped after 1, 5, 10, 15, 30, 45, or 60 minutes; dotted lines indicate polyhydroxylated progesterone products, grey lines indicate the main product 15β-OH-P, dashed with single dot lines indicate other monohydroxylated progesterone products, dashed lines indicate the remaining substrate progesterone. Progesterone conversion is almost complete after 15 minutes (1 % left). Most 15β-OH-P (78%) is formed after 10 minutes, and slowly decreasing to 69 % within 60 minutes of conversion (less than the combination of Etp1(516-618)/AdR, see fig. 3.6). The formation of polyhydroxylated progesterone products is slowly increasing and ranging in between 24 % and 26 % after 60 minutes of conversion (higher than with the combination Etp1(516-618)/AdR, see fig. 3.6). Formation of other monohydroxylated progesterone products (than 15β-OH-P) remains at 4 % after 5 minutes of conversion time (taken from [81]).

Redox combinations	$k_{app}[s^{-1}]$
Adx(4-108)/AdR	0.29176 $\pm$ 0.00107
Etp1(516-618)/Arh	0.34322 $\pm$ 0.00072

Table 3.1: Table 1: Apparent rate constants (k_{app}) for the reduction of CYP106A2 by the two redox partner combinations Adx(4-108)/AdR or Etp1(516-618)/Arh using 400 μM NADPH, 2 μM ferredoxin reductase (AdR or Arh1), and 20 μM ferredoxin (Adx(4-108) or Etp1(516-618)) in syringe A and 2 μM CYP106A2 and 200 μM progesterone in syringe B. The equations used to fit with the data points were: $f(x) = a * (1 - e - b * x)$. The given constants correspond to the respective b in the exponent of the equations (taken from [81]).

0.29 s^{-1}. This is in full agreement with the calculated shorter distance of the active site of the ferredoxin Etp1(516-618) and CYP106A2 compared with Adx(4-108), as well as the faster overall conversion of progesterone when using Etp1(516-618)/Arh1.

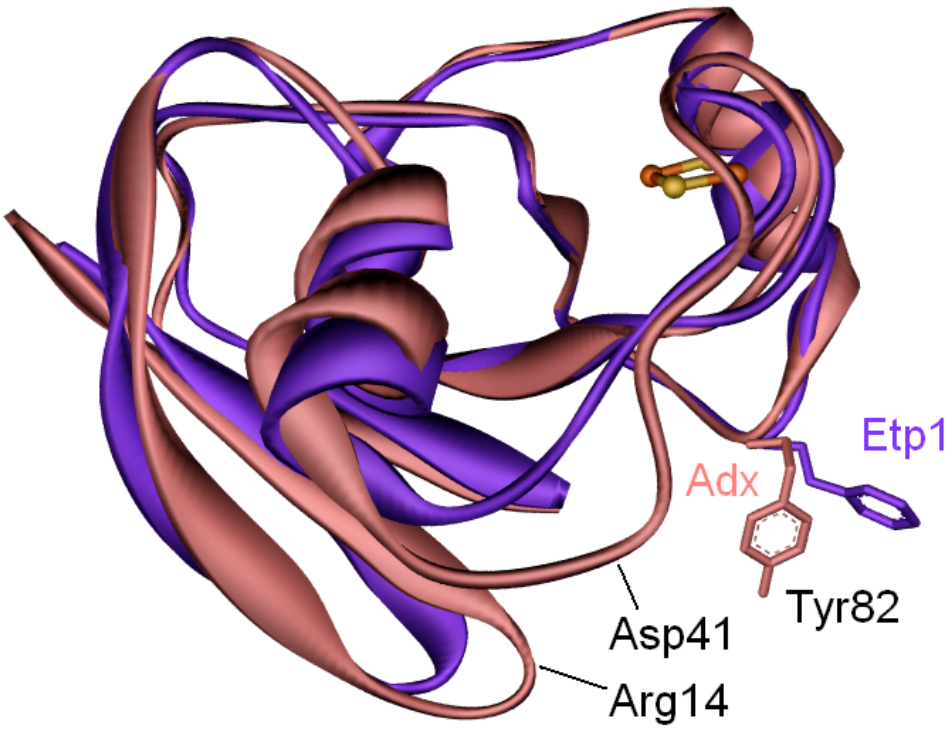

Figure 3.8: Superimposition of AdX(4-108) to the obtained docking conformation of Etp1 (violet), which has its FeS-cluster in closest distance to CYP106A2. The loops of Adx (coral) around residues 14 and 41 would clash with the cytochrome (omitted for clarity) (taken from [81]).

3.2 Discussion

As every cytochrome P450 is in need of electrons, the choice of the corresponding redox partners is crucial for the proper function of CYPs [86]. However, not every electron transfer protein chain is capable of delivering electrons to different CYPs. Therefore, a universal redox system, being able to reconstitute the activity of a majority of CYPs, is the ultimate goal of current investigations [11]. There is a variety of parameters one should consider when looking for such a universal system, for example expressability, (thermo-) stability or electron transfer rate. Also, mutants of CYPs that show altered interaction sites to the ferredoxins could be a way to change the substrate conversions. Pikuleva et al. [94] demonstrated that a replacement exchanging conserved lysines with alanines at positions 354 and 358 of CYP27A1 within the presumed adrenodoxin-binding site resulted in an approximately 20-fold increase in the apparent K_s values for adrenodoxin. In this study, we had a closer look at redox systems from different organisms that have already

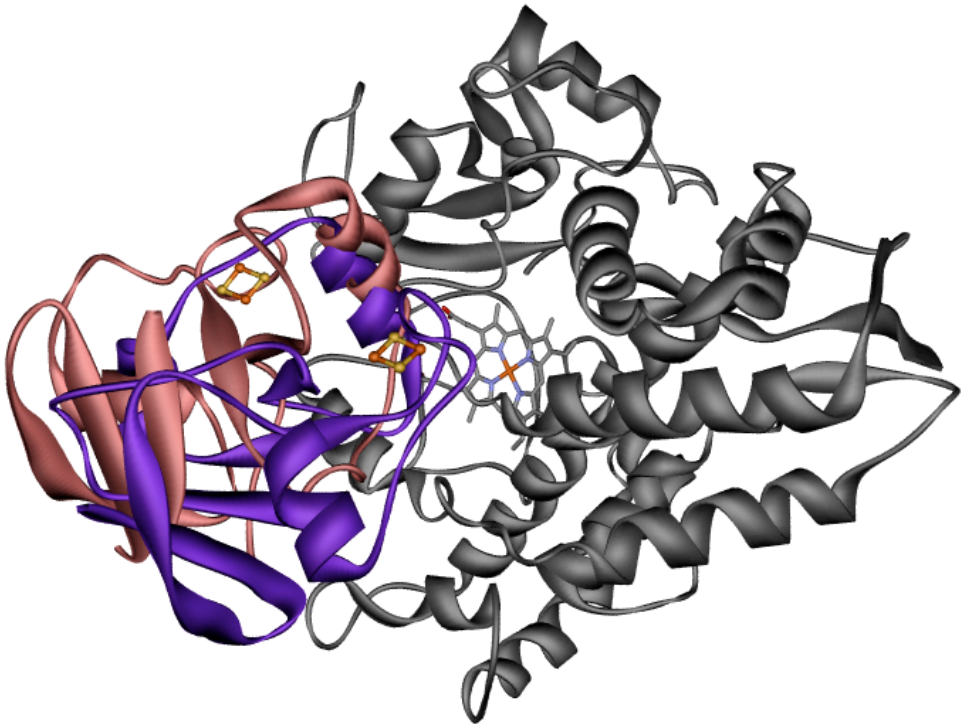

Figure 3.9: Overlay of the energetically most favorable docking conformations of AdX(4-108) and of Etp1 to CYP106A2. Adx (coral) is not able to bind in the same orientation as Etp1 (violet) due to subtle but distinct structural differences (taken from [81]).

been used with diverse CYPs: Adx(wildtype), Adx (4-108) and AdR from bovine origin, Etp1(516-618) and Arh1 from the yeast *Schizosaccharomyces pombe*, FldA and Fpr from *Escherichia coli*, Fdx2 and BmCPR from *Bacillus megaterium DSM 319*, and YkuN from *Bacillus subtilis 168*. When looking for ideal redox partners for CYPs, identification of an efficient redox system for a specific CYP should be performed. This can be further optimized by methods of protein engineering and finally tested with other CYP isoforms. As a first step to create a universal redox system for CYPs, we tested the applicability of these diverse redox partners to the well-studied CYP106A2 [27, 82, 25, 95, 30]. The experiments showed clearly that all used combinations of redox partners were able to provide electrons to CYP106A2. However, FldA and YkuN in combination with Fpr needed higher ratios than the usual CYP:ferredoxin/flavodoxin:reductase ratio of 1:20:2 (figure 3.2). Only the truncated forms of Adx and Etp1, combined with AdR or Arh1, were able to support CYP in a way that almost all provided progesterone (200 μM) was converted within 30 minutes. Full-length Adx(1-128) as well as Fdx2 appeared to have slower electron transfer rates towards CYP106A2 as their use results in 57 to 82 μM leftover substrate after 30 minutes of conversion. The selection of the

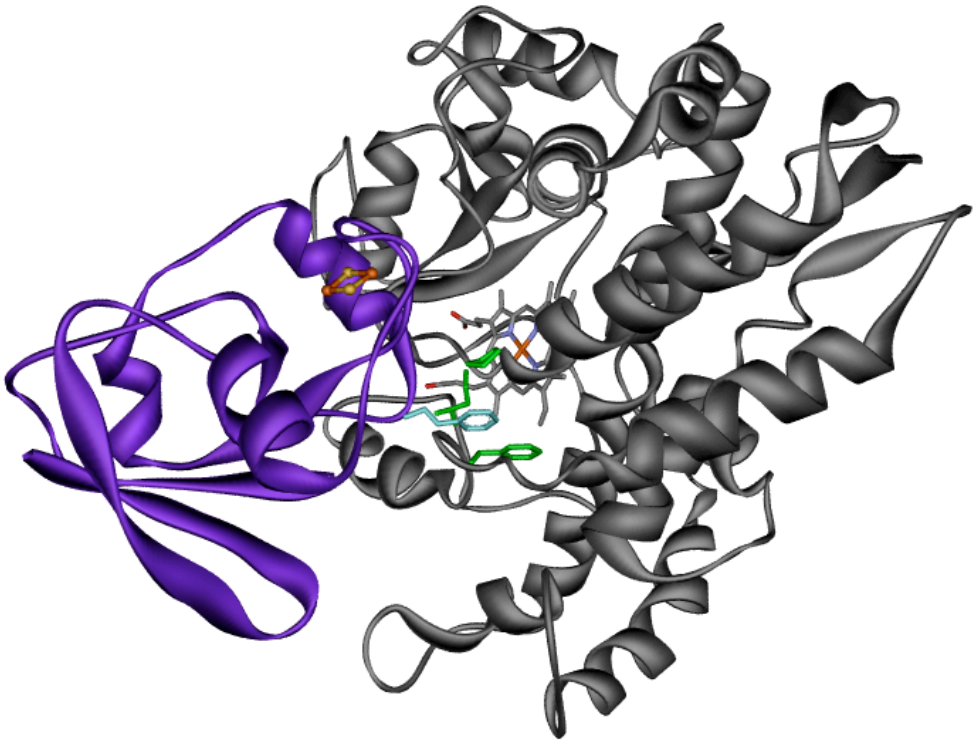

Figure 3.10: Obtained docking conformation of Etp1 (violet) in which its FeS-cluster is closest to the heme group of CYP106A2 (grey). Phe82 (cyan) penetrates into the cytochrome forming hyrophobic contacts with Phe107, Pro359, and Leu356 (green) (taken from [81]).

reductase in the redox system has also some influence on conversion as it affects the overall turnover yield as seen with Etp1(516-618)/AdR (100%) compared with Etp1(516-618)/BmCPR (79 %). The major, but most unexpected result testing the electron transport ability of the different redox systems, was the observation that the exchange of the redox partners resulted in variable product distributions. The hydroxylation of progesterone by CYP106A2 reconstituted with the commonly used redox system Adx(4-108) and AdR produces monohydroxylated progesterones (15β-, 6β-, 9α-, and 11α-progesterone, 15β-OH-P being the main product of wild-type CYP106A2) and polyhydroxylated progesterone products. When using this "parent" system of Adx(4-108) and AdR for 30 minutes of conversion, 65 % of the products are polyhydroxylated progesterones, 27 % are 15β-OH-P, 5 % are other monohydroxylated progesterones, and 3 % is remaining substrate. This ratio dramatically changed when using Etp1(516-618) as ferredoxin instead of Adx(4-108) which results in only 18 % poly-OH-P, 78 % 15β-OH-P, 3 % other monohydroxylated progesterones, and 0 % remaining substrate. Similar results were obtained when using full-length Adx(1-128) and AdR (where the formation of poly-OH-P al-

most completely drops to 1% and 15β-OH-P is up to 53 %; but in this case, still 41 % of the substrate is not converted after 30 minutes) or Fdx2 and Arh1 (where 2 % poly-OH-P and 66 % 15β-OH-P is produced, but 29 % substrate is not converted). Obviously, in the later two cases the reaction rate is too slow to support further hydroxylation of 15β-OH-P to polyhydroxylated products. To inspect the time dependency of the product formation, the conversions by CYP106A2 with the combinations of Adx(4-108) or Etp1(516-618) with AdR or Arh1 were stopped after 1, 5, 10, 15, 30, 45, and 60 minutes (figures 3.4-3.7) to be investigated in detail. These experiments show that the conversion of progesterone by the ferredoxins is typically complete within 10 minutes (Etp1(516-618)/AdR) to 30 min (Adx(4-108)/AdR and Adx(4-108)/Arh1). This proves a very efficient substrate conversion. The product distribution, however, is varying heavily within the investigated time frame. While the amount of 15β-OH-P is raising in every combination to at least 67 % within 5 to 10 minutes, it is afterwards dropping substantialy down to around 30 % when using Adx(4-108). This harsh decrease is accompanied by a steep increase of poly-OH-P up to 66 % (Adx(4-108)/Arh1). Using Etp1(516-618), the ratio of 15β-OH-P is either decreasing only slightly after 10 minutes (with Arh1) or remains constant (with AdR). Based on these results, it becomes clear that the different redox partner combinations tested here can be used for different purposes, depending on the overall product yield and pattern one is aiming for. For approaches going for fast substrate conversions, truncated versions of ferredoxin should be used as these are able to process over 90 % of 200 μM substrate within 10 to 15 min to the corresponding products. Adx(4-108) would be preferable to those aiming for many polyhydroxylated progesterone products as with increasing conversion time one is able to obtain almost 130 μM polyhydroxylated progesterone products out of 200 μM substrate (59 % conversion). Polyhydroxylated products can be of big interest, because they represent a high diversity and, therefore, different possibilities for further alteration of the steroid molecule. Since many attempts are focussed on high yields of certain monohydroxylated products, application of truncated Etp1(516-618) is useful to obtain high absolute and relative yields (shown here for the main product 15β-OH-P) within only 5 minutes. The least percentage of polyhydroxylated progesterones after 30 minutes was found with the redox partner system FldA/Fpr. However, the facts that only a ratio of CYP106A2/FldA/Fpr of 1:50:50 was able to hydroxlate progesterone and that only 27 % of product were found, characterize this system as low efficient. Systems using Adx(1-128) as redox partner are further candidates for the production of monohydroxylated progesterones because they produce only 1% of polyhydroxylated products. However, also in this case, the activity is rather low (max. 62 % conversion after 30 minutes). A longer conversion time (60 minutes) is increasing the overall product yield (data not shown). As almost every redox partner combination alters the product pattern of the conversion of progesterone by CYP106A2, the comparison especially of the ferredoxins as the direct electron transfer proteins to CYP106A2 is of interest. The crystal structures of the bovine Adx(4-108) and the yeast Etp1(516-618) have been solved by Müller [96, 40] and an evaluation of differences and similarities is, therefore, possible. Both enzymes

show an amino acid identity of 46 %, with overall highly conserved regions. Despite the similarity, there are also some differences: The thermal transition temperature T_m of Etp(516-618) is 65 °C, about 14 °C higher than the T_m of Adx(4-108) of 51 °C [40]. When comparing protein-protein dockings of Adx(4-108) or Etp1(516-618) and CYP106A2 (fig. 3.8), different binding modes of the chosen ferredoxins become obvious: Adx(4-108) prefers binding close to the heme using a different interaction region compared with Etp(516-618) (see figs. 3.8 and 3.9). Phe82 of the Etp1(516-618) is in a sandwich position with Phe107, Pro359, and Leu356 of the CYP106A2 and enables Phe82 to come close to the heme (see fig. 3.10). In contrast, Tyr82 of Adx(4-108) does not bind to the CYP, because the corresponding orientation is sterically hindered by the residues Arg14 and Asp41, as they would clash with the CYP. On the basis of the possible orientations, the calculated distance between the FeS-Cluster and the heme-iron atom is longer (24 Å) for the binding of the Adx(4-108) than for Etp1(516-618) (19 Å) in the closest conformation found, respectively. This condition allows Etp1(516-618) to approach closer and, therefore, enables faster electron transfer. On the other hand, the binding of the Adx(4-108) could be stronger than that of Etp1(516-618), for example due to larger electrostatic contributions, resulting in a slower release of Adx(4-108). This may lead to a longer disposition of the substrate in the active site of the CYP106A2 and, therefore, also more than one single hydroxylation product. This can be caused by the kind of contact between the ferredoxin and the CYP, meaning an intense binding and slow transfer of electrons could not allow the then-build product to exit the active site of the CYP, reverse that a looser binding and fast e-transfer is enabling the monohydroxylated progesterone to leave the binding pocket before further hydroxylation is performed. In the absence of experimental data on the actual life time of these transient complexes between the respective ferredoxins and the CYP, corresponding assumptions must be considered with some care. Nevertheless, our molecular dynamic simulations of 15β-OH-progesterone in the binding pocket of CYP106A2 show that reorientation of the mono-hydroxylated product occurs within 100ns, leading to conformations in which hydroxylation of other positions (11α-, 9α-, and 6β) is likely. This time scale is rather fast compared to that of the product release, which was not observed within the simulation time of 200 ns. Thus, the observed product spectra are governed by several factors, which are not independent of each other. Taken together, our results demonstrate that different aspects play a role when a decision has to be undertaken which redox system is the most preferable one for a certain cytochrome P450. We were able to show that the use of different redox systems manipulates the rate and the yield of conversion of progesterone by CYP106A2. Also product distribution by changing the redox partners can be altered between high levels of polyhydroxylated or monohydroxylated progesterone products. The application of the *Schizosaccharomyces pombe* Etp1(516-618) results in high quantity of the main monohydroxylated product, achieving 155 μM 15β-OH-P out of 200 μM substrate (78 %) within 30 minutes of conversion, whereas the bovine Adx(4-108) produces many polyhydroxylated progesterone products (129 μM or 65 %). Working with the full-length Adx(1-128), one gets high amounts of monohydroxylated progesterone

products (113 μM) with the downside of a slow conversion (only 59 %). These experimental data were supported by molecular docking revealing a shorter distance between the FeS-cluster of the ferredoxin and the CYP heme for Etp1(516-618) than for Adx(4-108). Backing the molecular docking data, it was shown by stopped flow measurements that the apparent rate constant (k_{app}) of Etp1(516-618)/Arh1 was higher than that of Adx(4-108)/AdR, therefore resulting in a faster electron transfer between Etp1(516-618) and CYP106A2 than between Adx(4-108) and CYP106A2, explaining the above mentioned product pattern in conversions with progesterone. Assuming a longer life time of the complex (Adx(4-108) with CYP106A2, the monohydroxylated products are able to reorient more often applying Adx(4-108) than Etp1(516-618) resulting in more hydroxylations and, therefore, higher amounts of polyhydroxylated products. This knowledge can be employed to modify or tune the amount of the given products. Thus, applying the studied redox systems from bacterial, fungal, and mammalian origin, various effects on the product patterns have been observed and are promising for tuning biotechnological reactions for different purposes.

Chapter 4

Engineering of CYP106A2 for steroid 9α - and 6β-hydroxylation

The results and discussion section (4.1), the conclusion section (4.2) as well as figures 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8 and tables 4.1, 4.2, 4.3 were adapted unaltered from our publication [55], whereas the experimental section is omitted here.
Bacillus megaterium cytochrome CYP106A2 also known as 15β-hydroxylase is a very interesting target for biotechnological applications due to its broad range of steroid acceptance. In past experiments directed evolution or site-directed mutagenesis were successfully applied. In publications from Lisurek et al [82, 97] CYP106A2 was engineered towards the conversion of progesterone. The main product was 15-OH-progesterone but side products were also found namely 6β-, 9α- and 11α-OH-progesterone (which were not only monohydroxylated but also di- and polyhydroxylated; fig. 4.1). Exchange of the amino acids T89N/A395I and A160T/A395I/R409L lead to an increase of 11α-OH-progesterone from 27.7% to 80.9% while production of 15β-OH-progesterone decreased from 50.4% to just 4.8% with a mutation of T89N/A395I [98]. 11α-OH-progesterone is a very interesting candidate for the pharmaceutical application. Not only does it exhibit an antiadrenogenic activity with minimal estrogenic or progestional side effects [99, 100] but also influences blood pressure through selective inhibition of 11β-hydroxysteroid dehydrogenase type 2 [101, 102]. Additionally 11α-OH-progesterone serves as a substrate for cortisone and hydrocortisone production [103, 104]. 6β-OH-progesterone was proven to be an intermediate for inhibitors of growth of breast cancer cells [105] and inhibits 5α reductase activity in male rat liver microsomes [106]. 9α-OH-progesterone serves also as an important intermediate for a substrate that exhibits glucocorticoid and progestional activity [82, 107].
Due to complicated reaction processes in chemical synthesis of this compounds a biotechnological synthesis is favored. Changing the selectivity of natural CYP106A2 with site-directed mutagenesis or directed evolution towards a 9α- or 6β-OH-progesterone smooth the way for biotechnological production of pharmaceutical

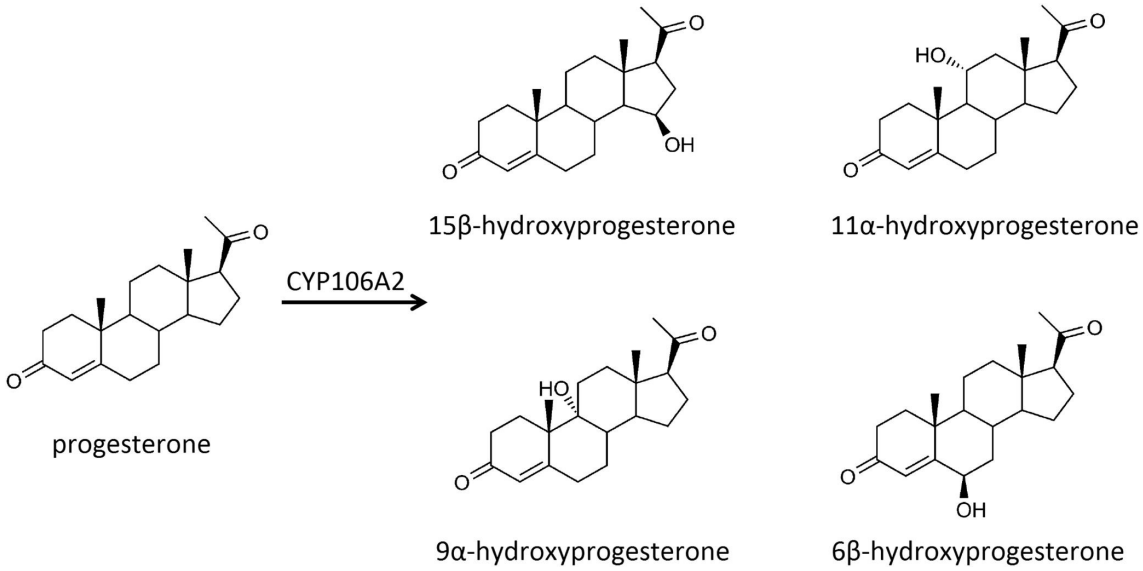

Figure 4.1: CYP106A2-catalyzed conversion of progesterone to 15β-, 11α-, 9α- and 6β-hydroxyprogesterone (taken from [55]).

compounds and industrial application of the enzyme.
This study describes the engineering of CYP106A2 towards the production of 6β- as well as 9α-hydroxyprogesterone. The mutant proteins were selected from a saturation mutagenesis library with performing a site-directed mutagenesis of them. For 9α-hydroxylase activity a 11-fold increase compared to the wild type (WT) could be measured, using a whole cell system. During the studies a first crystal structure for CYP106A2 was published [30] which allowed for docking experiments of progesterone to variants of CYP106A2 that would change the selectivity from 15β- to 6β-OH-progesterone. Our former docking experiments show two different conformations. One that allows for a 15β- or 6β hydroxylation and an other one that makes the α-positions accessible. For the latter the $[C_3 = O]$ bond builds a hydrogen bond with threonine at position 247. Therefore we proposed three different mutations:

Threonine at position 247 to alanine
Breaks hydrogen bond at position 247 and makes alternative conformation more likely

Alanine at position 243 to serine
Stabilizes alternative conformation by creating hydrogen bond with $[C_{20} = O]$

Phenylalanine at position 173 to alanine
Creates more space above heme group to place progesterone more "optimal"

4.1 Results and Discussion

The bacterial enzyme CYP106A2 was regarded as suitable platform to change the steroid hydroxylation regioselectivities and to improve the catalytic efficiencies of the obtained mutant proteins. Therefore, a library of CYP106A2 with different mutations, containing more than 13000 clones was created based on the replacements at positions A395 and G397 by a combination of site-directed and saturation mutagenesis of active site residues, followed by several rounds of site-directed mutagenesis based on a low-homology model of CYP106A2 [98, 108]. The library was screened using 200 μM progesterone as a substrate followed by HPLC to identify mutant proteins with increased 9α-hydroxylase activity. Sixteen mutant proteins with improved activity were selected from this library. HPLC analysis identified the yields of 15β-, 11α-, 9α- and 6β-hydroxyprogesterone for each mutant protein (table 4.1). In addition to the single-hydroxylated products in dependence on the reaction time also small amounts of di- and polyhydroxylated products were formed (data not shown). Ten mutant proteins formed 9α-hydroxyprogesterone as a main product. It can be seen from table 4.1 that CYP106A2 A395W/G397K showed the highest (55.9%) and A395H/G397Y, respectively, the lowest (38.6%) 9α-hydroxyprogesterone formation. This comprises an at least 8-fold improvement compared to the wild type. In contrast, CYP106A2 A395R/G397R possessed the greatest decrease of 15β-hydroxylase activity, from 67.9% to 8.8%. The other six mutant proteins

Nr	Mutation	15β-OH-P [%]	11α-OH-P [%]	9α-OH-P[%]	6β-OH-P [%]
	WT	67.9	18.0	5.4	8.7
1	A395T/G397K	11.2	20.0	45.1	23.7
2	A395H/G397Y	22.9	15.2	38.6	23.3
3	A395I	28.7	50.9	10.2	10.2
4	A395E/G397V	9.4	15.9	48.6	26.1
5	A395R/G397K	9.1	12.9	44.8	33.2
6	A395M/G397S	13.1	65.6	14.4	6.9
7	A395R/G397I	13.2	12.7	47.0	27.2
8	A395R/G397P	38.4	57.0	2.4	2.2
9	A395K	16.5	37.2	42.4	4.0
10	G397W	27.3	36.4	27.3	9.1
11	A395F	32.5	38.3	11.5	17.7
12	A243V	78.0	8.2	3.6	10.2
13	A395W/G397K	10.5	19.2	55.9	14.3
14	A395E/G397Stop	13.4	15.5	45.4	25.8
15	A395S/G397Y	18.5	13.7	41.1	26.8
16	A395R/G397R	8.8	14.9	45.8	30.6

Table 4.1: Progesterone hydroxylation catalyzed by CYP106A2. Relative values of the product pattern (15β-, 11α-, 9α- and 6β-hydroxyprogesterone) after 24 h in vivo conversion of 200 μM progesterone catalyzed by wild type CYP106A2 and its 16 mutant proteins, which were selected from a library. Values show percentages with the sum of all monohydroxylated products set as 100% (taken from [55]).

(A395I, A395M/G397S, A395R/G397P, G397W, A395F and A395V) exhibited a smaller 9α-hydroxylase activity compared to A395W/G397K and A395H/G397Y, whereby two mutant proteins (A395R/G397P and A243V) showed almost no 9α-hydroxyprogesterone formation at all. In addition, A243V showed a higher selectivity towards 15β-hydroxyprogesterone formation compared to the wild type enzyme. The four best mutant proteins A395E/G397V, A395R/G397K, A395W/G397K, and A395R/G397R with the highest 9α-hydroxyprogesterone formation were selected for further mutation via site-directed mutagenesis to improve their regioselectivity. Therefore, seven single replacements (D217V, A243V, A106T, F165L, T89N, T247V and T247W) were inserted into the four mutant proteins. These mutations were shown before to influence the regioselectivity and catalytic efficiency in a positive way. Rauschenbach and Virus could show that they exhibit higher turnover rates towards progesterone conversion compared to the wild type [109, 95]. In addition, it was reported that the highly conserved threonine residue in the I-helix of the active side near the heme of P450s directly affects the catalytic efficiency due to its participation in the proton shuttle which is important for proton delivery and molecular oxygen activation [97]. Therefore, the mutations T247V and T247W were also carried out. The insertion of these seven amino acid replacements into the four "parent mutant proteins" led to 28 new mutant proteins that were also tested for in vivo conversion of 200 μM progesterone, respectively. Five mutant proteins showed almost no conversion. Due to their low activity, these clones were not followed up in further experiments. Sixteen mutant proteins possessed a 9α-hydroxylase activity similar to that of the parent mutant proteins. There were also two mutant proteins (A243V/A395E/G397V and T89N/A395W/G397K) with a strong shift in regioselectivity, but they also showed an increase in di- and polyhydroxylated progesterone. These products were observed in the HPLC chromatograms (between 3.5 and 7.5 min) with the peak size depending on the conversion time as well as the corresponding mutant protein. By shortening the reaction time, the amounts of side product can be reduced [98]. Interestingly, five mutant proteins displayed a low amount of di- and polyhydroxylated progesterone and conversely a high amount of 9α-hydroxyprogesterone (D217V/A395E/G397V, F165L/A395E/G397V, A106T/A395R/G397K, A243V/A395R/G397R and A106T/A395R/G397R) (fig 4.2). The relative amounts of hydroxylated metabolites after conversion of progesterone in the presence of the best five of the 28 mutant proteins of CYP106A2 are shown in table 4.2. CYP106A2 F165L/A395E/G397V had the strongest shift in the regioselectivity showing a formation of 59.7% 9α-hydroxyprogesterone. This represents an improvement of the yield of 9α-hydroxyprogesterone by 15% compared to the initial CYP106A2 A395E/G397V, but a 12-fold improvement compared to the wild type CYP106A2. In turn, the amount of 15β-hydroxyprogesterone decreased from 67.9% to 11.3%. Mutant proteins D217V/A395E/G397V, A106T/A395R/G397K, and A106T/A395R/G397R showed a shift in regioselectivity of about 10% compared to the parent mutant proteins A395E/G397V, A395R/G397K, and A395R/G397R, respectively, and of about 51% compared to the wild type. CYP106A2 A243V/A395R/G397R pos-

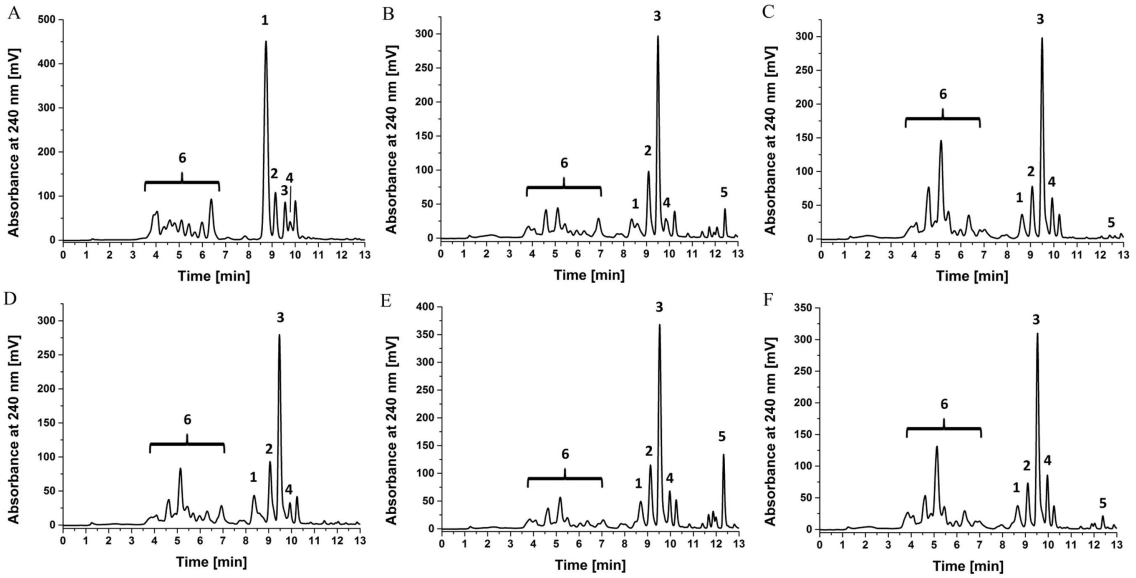

Figure 4.2: HPLC analysis of the in vivo conversion of 200 μM progesterone by wild type CYP106A2 (A) and the mutant proteins D217V/A395E/G397V(B), F165L/A395E/G397V (C), A106T/A395R/G397K (D), A243V/A395R/G397R (E), A106T/A395R/G397R (F). The figure shows di- and polyhydroxylated(6) and monohydroxylated products (15β- (1), 11α- (2), 9α- (3) and 6β-hydroxyprogesterone (4)), as well as the substrate progesterone (5). HPLC analysis was performed with H2O:ACN (90:10) as a mobile phase and a CC 125/4 nucleodur 100-5 C18 ec column at 40 °C (taken from [55]).

Mutation	15β-OH-P [%]	11α-OH-P [%]	9α-OH-P[%]	6β-OH-P [%]	P[%]
WT	67.9	18.0	5.4	8.7	0.0
D217V/A395E/G397V	7.5	23.2	55.9	7.9	5.4
F165L/A395E/G397V	11.3	17.2	59.7	11.0	0.9
A106T/A395R/G397K	13.0	22.4	57.4	6.8	0.4
A243V/A395R/G397R	11.8	17.6	49.7	8.3	12.7
A106T/A395R/G397R	12.9	14.9	56.2	13.3	2.7

Table 4.2: Selection of progesterone hydroxylation catalyzed by improved mutated CYP106A2. Relative values in % of progesterone and the monohydroxylated products 15β-, 11α-, 9α- and 6β-hydroxyprogesterone after 24 h in vivo conversion of 200 μM progesterone catalyzed by wild type CYP106A2 and the mutant proteins D217V/A395E/G397V, F165L/A395E/G397V, A106T/A395R/G397K, A243V/A395R/G397R and A106T/A395R/G397R, respectively. Values show percentages with the sum of the substrate progesterone and all monohydroxylated products set as 100% (taken from [55]).

sessed the smallest enhancement of 9α-hydroxyprogesterone formation (49.7%), but a large decrease of 15β-hydroxyprogesterone from 67.9% to 11.8% compared to the CYP106A2 wild type. Taken together, this data clearly demonstrate that using a combination of site-directed mutagenesis/saturation mutagenesis not only changes

the selectivity from the 15β- to the 11α-position of progesterone as described before, [98] but this approach can also be applied to change the regioselectivity of progesterone hydroxylation from position 15 in the D ring to position 9 in the B ring. For creating mutant proteins leading to 6β-hydroxyprogesterone production, the recently solved crystal structure of CYP106A2 was used [30]. After performing docking of progesterone into this structure, several amino acids in the enzyme were identified, whose mutation should lead to a change in the selectivity from 15β to the 6β position leading to the production of 6β-hydroxyprogesterone. It was suggested that the change from threonine to alanine at position 247 should interrupt a hydrogen bond, which should then prefer an alternative conformation of the substrate within the active site. Moreover, serine instead of alanine in position 243 was expected to stabilize a corresponding conformation in which 6β hydroxylation is favored by introducing a new hydrogen bond with the carbonyl oxygen in position 21 (fig. 4.3). To support this assumption, we carried out molecular dynamics simulations of progesterone in a homology model of this CYP106A2 A243S. Fig. 4.4 and fig. 4.5 show the time course of the distances of the hydrogen atoms in 15β and 6β position from the iron atom of the heme group, respectively. Whereas both curves show a decrease of the distance from the beginning of the simulation to about 5 ns, the 6β position remains closer to the heme iron within the subsequent 5 ns. The shortest distance for the 15β position was observed after 19.2 ns (2.6 Å), whereas the minimum distance of the 6β position was found already after 6.8 ns (3.0 Å). Likewise, the average distances are 5.7 and 5.5 Å, respectively, over the total simulation time of 20 ns. The maximum distances were seen rather early between 0.5 and 3.0 ns, both being beyond 9.2 Å. Obviously this short time span of 5 to 10 ns in which the 6β position is about 1 Å closer to the heme iron than the 15β position seems to determine the product spectrum. Finally, changing phenylalanine against the smaller alanine at position 173 should create more space over the heme for an optimal conformation of progesterone for 6β-hydroxylation. Single mutations as well as combinations of these three changes were inserted into CYP106A2. The obtained mutant proteins were used for the conversion of progesterone. CYP106A2 F173A/T247A did not show any significant conversion, whereas mutant proteins F173A, T247A, A243S/T247A and F173A/A243S converted progesterone to the main product 9α-hydroxyprogesterone. However, CYP106A2 with the replacements A243S and F173A/A243S/T247A, respectively, showed 6β-hydroxyprogesterone as main product with a 9- and 6-fold improvement, respectively, compared to the wild type. The amount of other products, 15β-, 11α- and 9α-hydroxyprogesterone, of CYP106A2 A243S is much lower than that of the triple mutant protein (table 4.3). Fig. 4.6 shows the chromatogram of CYP106A2 A243S compared to the wild type. It can be seen that site-directed mutagenesis of CYP106A2 based on the results of docking of progesterone into the crystal structure of this protein, led to a nearly complete change of the regioselectivity of hydroxylation from the 15β- to the 6β-position, i.e. from the D-ring to the B-ring. More than 80% 6β-hydroxylation was obtained. To find out whether mutant proteins, which are changing the regioselectivity of progesterone hydroxylation from the 15 to the 11, 9, or 6 position, show

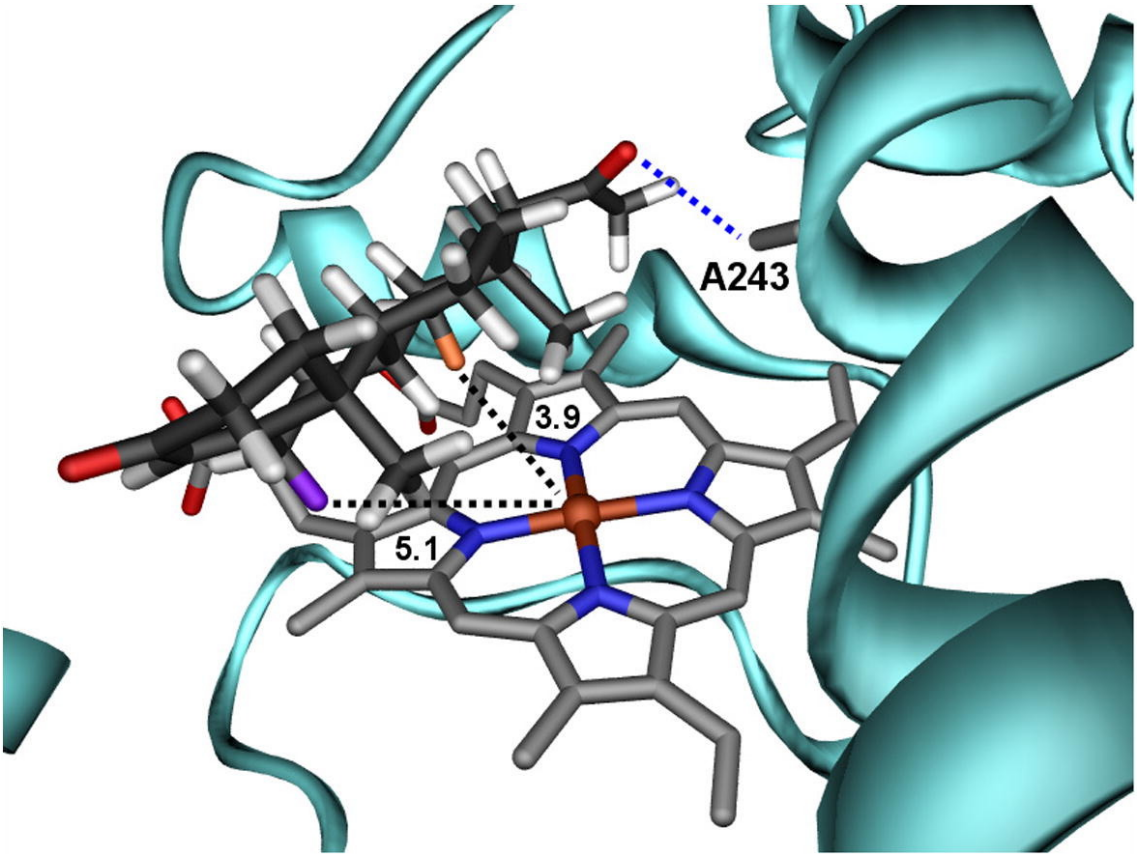

Figure 4.3: Docking of progesterone into the X-ray structure of wild-type CYP106A2 showed position 15β (orange) being closest to the heme iron (3.9 Å). However, hydroxylation was experimentally also observed in 6β position (purple) (5.1 Å). The replacement of alanine by serine in position 243 would enable a new hydrogen-bond (blue dotted line) to the carbonyl oxygen in position 21. As a further consequence the progesterone is also shifted towards this residue thereby bringing position 6β in closer contact to the heme iron. The corresponding mutant protein shifted the main product from 15β to 6β-hydroxy-progesterone (taken from [55]).

another substrate space than the wild type, a natural product library was screened with CYP106A2 T89N/A395I [98], F165L/A395E/G397V, and A243S, respectively. Screening of the library with these three mutant proteins showed the same twelve compounds as already published by Schmitz et al. [29] for wild type CYP106A2 so that obviously no change in the substrate space occurred. To investigate a possible change in the regioselectivity of other known CYP106A2 substrates, dipterocarpol was converted with the three mutant proteins and the results were compared to those of the wild type, which formed two products, 7β-hydroxydipterocarpol and 7β, 11α-dihydroxydipterocarpol. Interestingly, the three mutant proteins converted dipterocarpol only into one product, 7β-hydroxydipterocarpol, indicating an increased specificity of hydroxylation towards the 7β-position. Fig. 4.7 shows exemplarily the thin layer analysis of CYP106A2 T89N/A395I. Finally, 11-deoxycorticosterone,

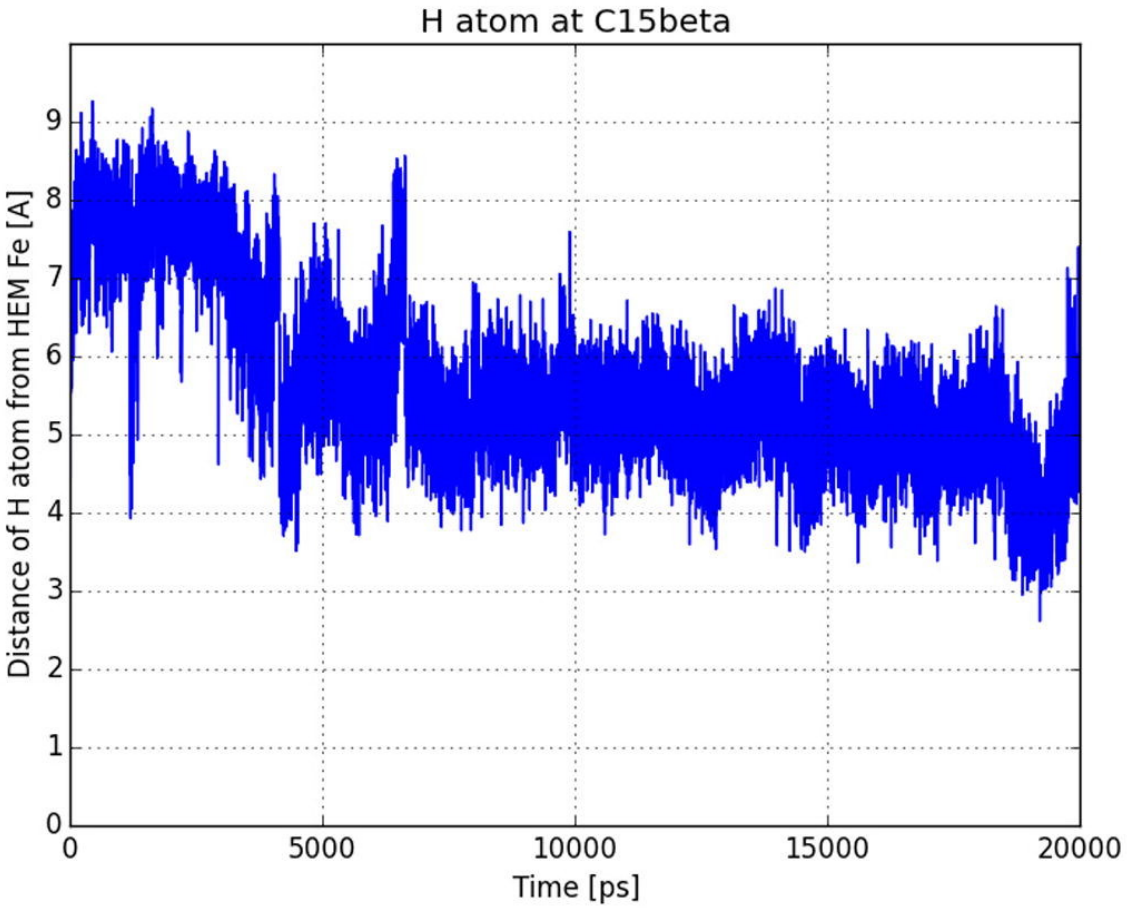

Figure 4.4: Time course of the distance (given in Å) between the hydrogen atom in position 15β of progesterone and the iron atom of heme of the heme group during a 20 ns molecular dynamic simulation of CYP106A2 A243S (taken from [55]).

was converted in vivo in the presence of wild type CYP106A2 and the six best mutant proteins created in this work (D217V/A395E/G397V, F165L/A395E/G397V, A106T/A395R/G397K, A243V/A395R/G397R or A106T/A395R/G397R, A243S). Fig. 4.8 shows the HPLC chromatogram of the conversion of DOC with wild type CYP106A2 and with CYP106A2 A106T/A395R/G397R. The wild type produced 15β-hydroxy-DOC as the main product as well as 6β-hydroxy-DOC and 7β, 15β-dihydroxy-DOC as minor products [110]. In case of the mutant proteins the peaks 2 (7β, 15β-dihydroxy-DOC) and 3 (15β-OH-DOC) drastically decreased whereas peak 4 was enhanced, demonstrating an increased production of 6β-OH-DOC. To date product 1 has not been identified because the amount is not sufficient for NMR analysis. We assume that this peak comprises either a mixture of several monohydroxylated DOC products or di- and/or polyhydroxylated DOC. Taken together, it can be concluded that the mutations introduced by directed evolution and site-directed mutagenesis led also to a change in the regioselectivity of other steroidal substrates besides progesterone resulting in an increase in the selectivity of

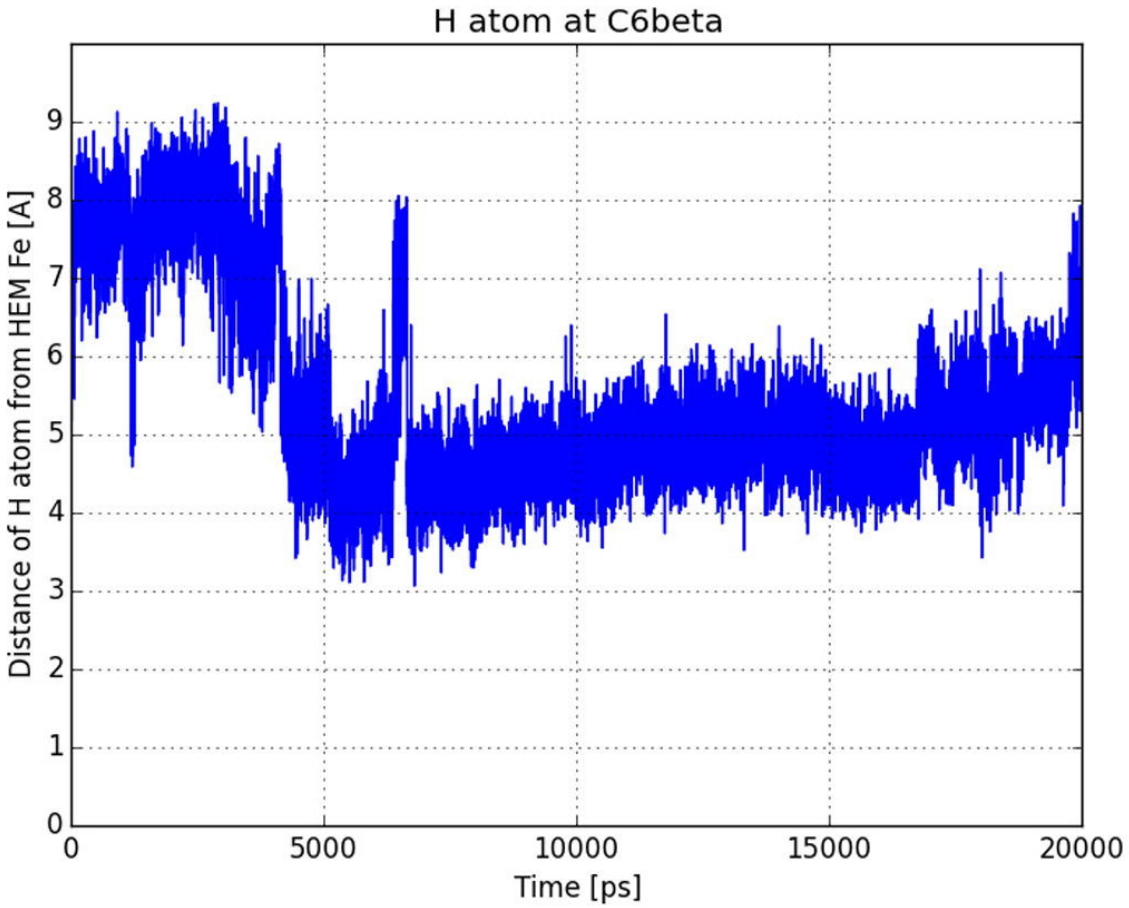

Figure 4.5: Corresponding time course of the distance for the hydrogen atom in position 6β of progesterone. Between 5 and 10 ns this position is about 1 Å closer to the heme iron atom and was experimentally found to be the main hydroxylation product of CYP106A2 A243S (taken from [55]).

dipterocarpol hydroxylation towards the 7β position and in the new main product 6β-hydroxy-DOC instead of 15β-hydroxy-DOC. The strongest shift in selectivity towards DOC hydroxylation was shown by CYP106A2 D217V/A395E/G397V displaying an improvement of 6β-OH-DOC by 44% and, on the other hand, showing a decrease of 15β-OH-DOC by 36% compared to CYP106A2 wild type.

4.2 Conclusion

In this work we were able to demonstrate that a combination of site-directed mutagenesis and saturation mutagenesis of CYP106A2 is able to influence the regioselectivity of progesterone hydroxylation, shifting it from C15 in the D ring to C9 and to C6 in the B ring, respectively. Besides changes in the regioselectivity, also

Mutation	15β-OH-P [%]	11α-OH-P [%]	9α-OH-P[%]	6β-OH-P [%]	P[%]
WT	67.9	18.0	5.4	8.7	0.0
F173A	9.7	32.5	44.1	5.0	8.6
A243S	3.4	4.6	2.4	82.7	6.9
T247A	11.7	13.1	45.7	24.3	5.3
F173A/A243S	12.5	15.5	51.8	15.9	4.3
A243S/T247A	14.6	13.2	48.0	20.7	3.4
F173A/A243S/T247A	13.6	6.5	8.7	57.2	19.1

Table 4.3: Progesterone hydroxylation catalyzed by improved mutated CYP106A2. Relative values in % of progesterone as well as the monohydroxylated products 15β-, 11α-, 9α- and 6β-hydroxyprogesterone after 24 h in vivo conversion of 200 μM progesterone by wild type CYP106A2 and the mutant proteins F173A, A243 S, T247A, F173A/A243S, A234S/T247A and F173A/A243S/T247A, F165L/A395E/G397V, A106T/A395R/G397K, A243V/A395R/G397R and A106T/A395R/G397R, respectively. Values show percentages with the sum of the substrate progesterone and all monohydroxylated products set as 100 % (taken from [55]).

a change in the stereo-selectivity from the 15β to the 9α position was obtained. It was shown that the hydroxylation position in the steroid depends on the mutation introduced into the protein. Whereas mutant proteins with a change in the selectivity from 15β- to 11α-hydroxyprogesterone as described previously [98] showed the largest shift in mutant proteins A106T/A395I, A106T/A395I/R409L, and T89N/A395I, CYP106A2 F165L/A395E/G397V demonstrated the strongest shift from the 15β towards 9α position. Interestingly, a nearly complete shift in the regioselectivity of progesterone hydroxylation from the 15β to the 6β position was obtained with CYP106A2 A243S, which was constructed by rational protein design/site-directed mutagenesis based on the recently available X-ray structure of the wild type [98]. In addition, we were able to demonstrate that also other steroids similar to progesterone, such as DOC and dipterocarpol, are hydroxylated with changed selectivity when using the above mentioned mutant proteins. Finally, the approach described here, especially when using the 3D structure of CYP106A2 as a basis for the rational design of mutant proteins, leads to desired changes in the regio- and stereo-selectivity of progesterone hydroxylation. The obtained products, 9α- and 6β-hydroxyprogesterone, are of biotechnological importance for a sustainable production of pharmaceutically valuable compounds.

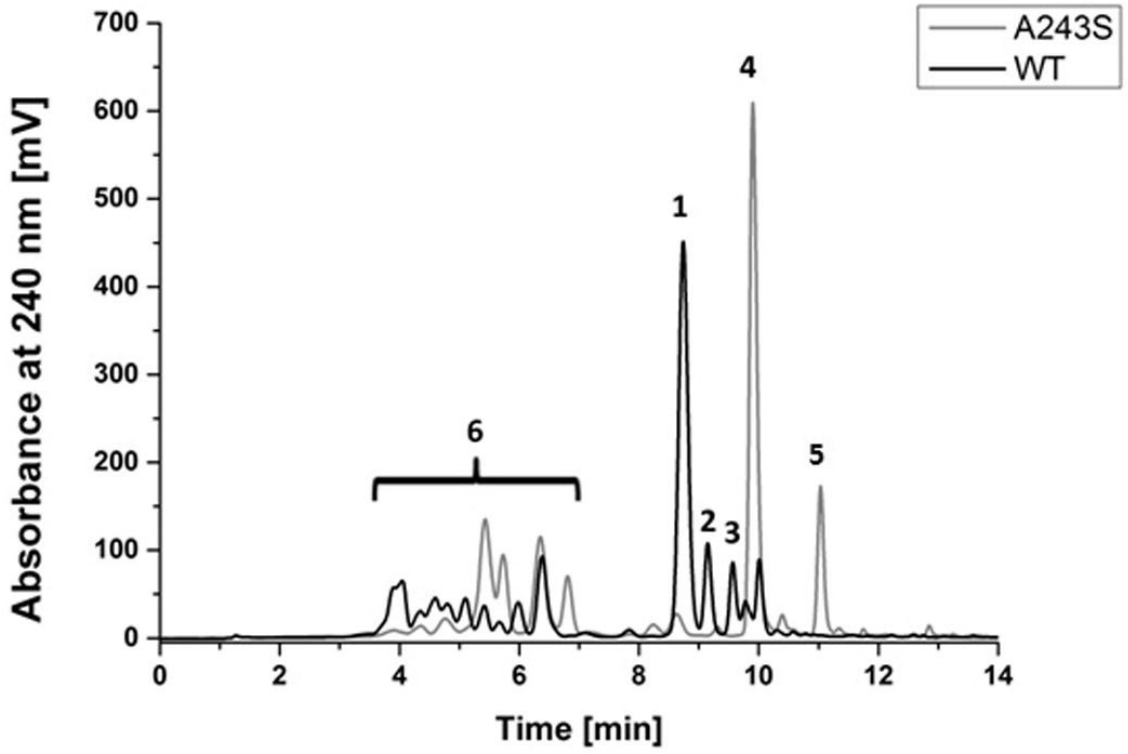

Figure 4.6: HPLC analysis of the in vivo conversion of 200 μM progesterone by wild type CYP106A2 and with the mutation A243S. The figure shows di- and polyhydroxylated (6) and monohydroxylated products (15β- (1), 11α- (2), 9α- (3) and 6β-hydroxyprogesterone (4)). Product 5 only appears in the mutant protein and was not identified yet. HPLC analysis was performed with H2O:ACN (90:10) as a mobile phase and a CC 125/4 nucleodur 100-5 C18 ec column at 40 °C (taken from [55]).

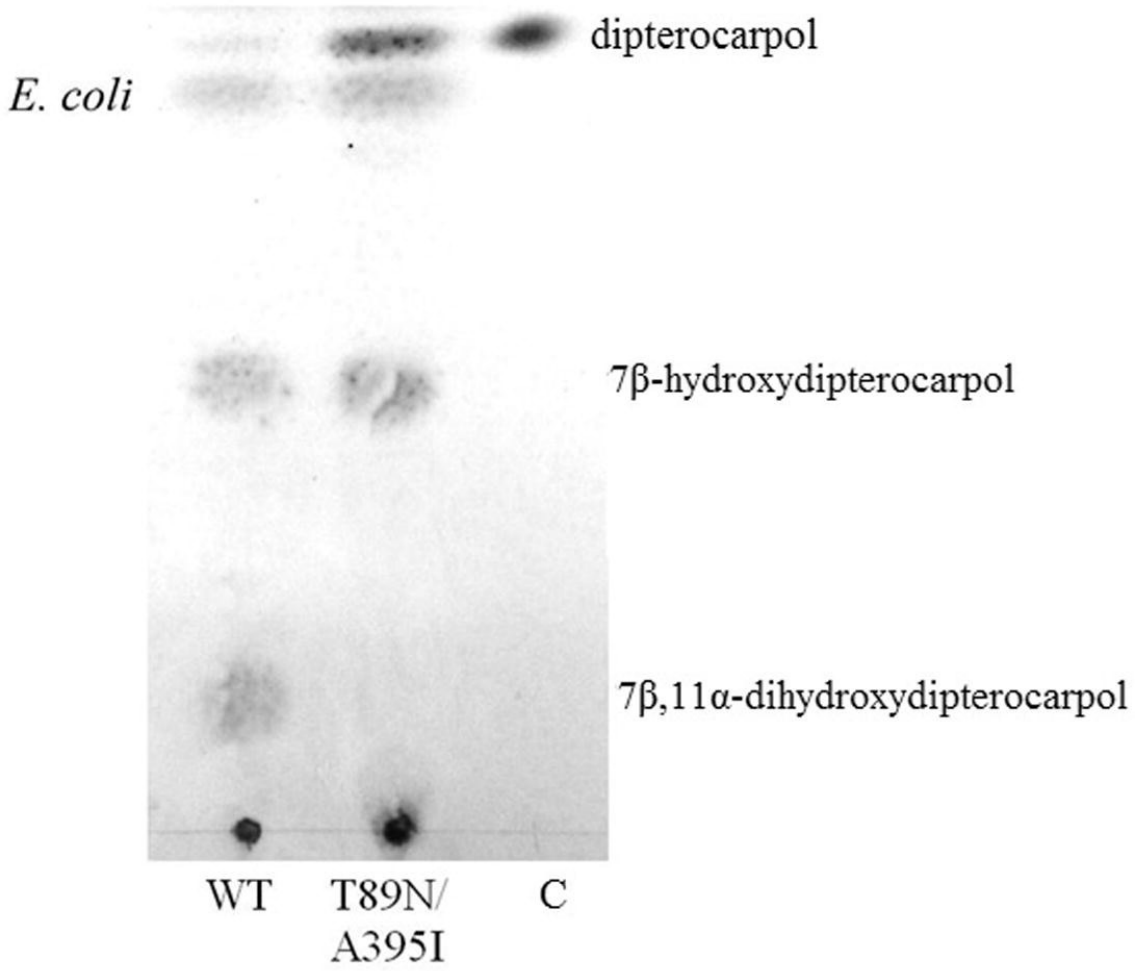

Figure 4.7: Thin layer analysis of the in vivo conversion of 200 μM dipterocarpol by wild type CYP106A2, the CYP106A2 T89N/A395I and a control without any enzyme. The figure shows the two products, 7β-OH-dipterocarpol and 7β,11α-OH-dipterocarpol. Thin layer analysis was performed with hexane/EE 1:1 as a mobile phase (taken from [55]).

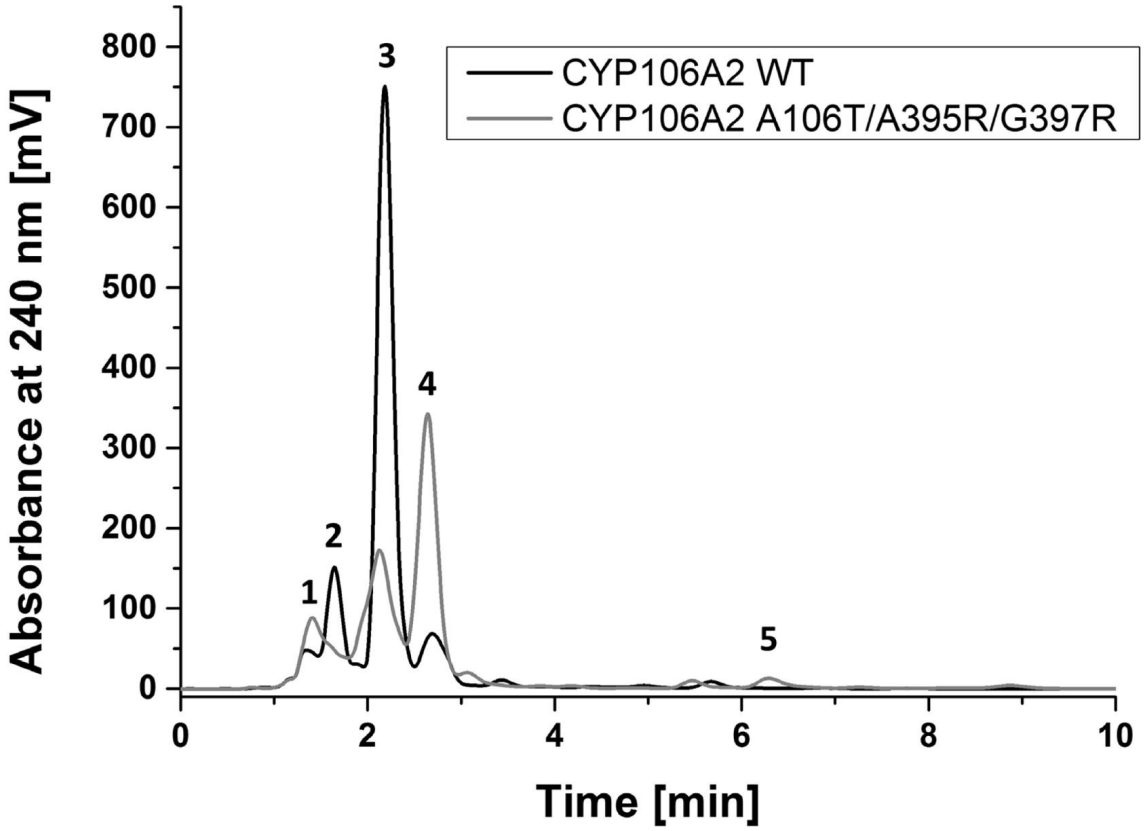

Figure 4.8: HPLC analysis of the in vivo conversion of 200 μM DOC by wild type CYP106A2 and with the mutation A106T/A395R/G397R. The figure shows the products 15β-OH-DOC (3), 6β-OH-DOC (4) and 7β,15β-dihydroxy-DOC (2), as well as the substrate DOC (5). Due to the small amount product 1 is not identified yet. HPLC analysis was performed with H2O:ACN (60:40) as a mobile phase and a CC 125/4 nucleodur 100-5 C18 ec column at 40 °C (taken from [55]).

Chapter 5

Rational Design of Mutants that Lower the Redox Potential of Bovine Adx

In this study we dealt with new ways towards implementing a system to check the influence of an amino acid exchange on the redox potential of Adx. We were also able to give proof how the redox potential is affected in the Adx deletion mutant TΔ49. More precisely what the distortion of the iron-sulfur cluster accounts for. Experimental validation for our suggested amino acid exchange still needs to be performed. The difficulties in computing redox potentials for ferredoxins like Adx is to account for the structural change upon reduction of the iron-sulfur cluster. Unfortunately, no conclusive information regarding the iron-sulfur cluster and its proximity is available from NMR measurements, because this region is blurred due to the paramagnetic properties of the reduced state. Moreover, the geometry of the iron-sulfur cluster was assumed to be fixed. Existing quantum chemical calculation on model systems of the iron-sulfur cluster, however, strongly suggest that the iron-iron distance increases in the reduced state, independently of the density functional being used [111, 112]. This can be attributed to the fact that the iron-sulfur cluster compensates the increased electron density by a distinct expansion. Bernhardt and coworkers studied a series of Adx mutants and collected experimental data regarding their catalytic activity [113]. They observed that mutants with lower redox potential than the wild-type showed increased values of k_{cat}. In particular, the T49Δ mutant where threonine at position 49 was deleted, exhibited the most pronounced increase of K_M that goes along with a decrease of -87 mV (0.087 eV) of the redox potential, whereas all other exchanges of single amino acids did not lead to similar strong effects [114]. Substantial decrease of the redox potential was also observed for deletions of multiple residues at the C-terminal end, e.g. those beyond position 108, which do not seem to be necessary for the function. Moreover, residues above position 112 were not resolved in the above mentioned X-ray structures, suggesting increased flexibility

and the absence of secondary structural elements. So far, no systematic study to elucidate the underlying mechanisms that influence the redox potential by the above mentioned mutations has been carried out, although theoretical approaches have been presented, which account for mutations of single amino acids [56].

5.1 Theoretical Background

Experimentally, the redox potential E_0 is measured in comparison to a reference potential applying a dye photoreaction, whereby Safranin T functions serves both as indicator and mediator [113]. Using Nernst's equation the free energy ΔG of the reduction process for Adx is thus the difference in solvation energy ΔG_{solv} between the reduced and the oxidized protein plus the absolute potential of the standard hydrogen electrode E_{SHE}, whereby $E_{SHE} = 4.44 \pm 0.02$ V (eq. 5.1):

$$\Delta G = -zGE^0 = \Delta G_{solv} + FE_{SHE} \tag{5.1}$$

Here, z = 1 since only one electron is transferred, and F is the *Faraday* constant. In more detail the change in solvation energy comprises the change of the inner-sphere energy ΔG_{in} due to the different electron configuration, which is mainly localized on the iron-sulfur cluster, and the change of the outer sphere energy ΔG_{out} that arises from conformational changes of the surrounding protein (eq. 5.2):

$$\Delta G_{solv} = \Delta G_{in} + \Delta G_{out} \tag{5.2}$$

The change in inner-sphere energy ΔG_{in} corresponds to the adiabatic gas phase ionization potential of the reduced iron-sulfur cluster IP_{red}, which can be computed from density functional calculations [111]. Most of the known mutations of bovine Adx comprise exchanges of a single amino acid on its surface or deletions at the N-terminal part, respectively the C-terminal part that are unlikely to change the local geometry of the iron-sulfur cluster [113]. Therefore ΔG_{in}, respectively IP_{red} can be assumed to be the same for the wild-type and these mutants, whereas the major difference that causes the altered redox potential arises from ΔG_{out}. However for mutants that induce a drastic conformational change to the geometry of the iron-sulfur cluster, i.e. the T49Δ mutant, which leads to the rearrangement of the side chain of Cys46, a substantial fraction of the observed change in redox potential is likely to arise from the change of IP_{red} . So far, no corresponding investigation of this effect has been carried out. To explain the underlying electrostatic interactions that eventually give rise to the alterations in energy of iron-sulfur proteins, respectively the redox potential of mutants, several theoretical approaches have been applied: Poisson-Boltzmann theory requires the accurate three-dimensional structure representation of both the oxidized as well as the reduced Adx solution, along with appropriate assignment of atomic charges and protonation states for each amino

acid [111, 115]. Furthermore, a lower dielectric constant can be used for the region of the iron-sulfur cluster than for the surrounding protein [116]. Much earlier, Churg and Warshel applied the protein dipoles-Langevin dipoles model to compute differences in solvation energy, whereby a similar finite grid was used [117]. Related to both methods is the work of Ichiye and coworkers that makes use of the electrostatic potential [118]. An intriguingly simple approach to explain the effect of single point mutations was presented by Zhou [56]. Here, the protein (Adx) is treated as sphere embedded in a medium of high dielectric constant (water), similar to the Tandford and Kirkwood theory of solvation. The iron-sulfur cluster is represented as point charge and the respective amino acid by its resulting dipole vector (fig. 5.1). According to this approach the change in redox potential depends on the orientation of the dipole vector of this particular residue with respect to the vector that connects the amino acid with the iron-sulfur cluster, as well as the corresponding distance and the dielectric constant where the residue is located. In cases where mutations of amino acids on the protein surface are carried out, the expected changes in redox potential should therefore be minor, because polar or charged side chains are fully solvated and thus effectively shielded by water due to its high dielectric constant. Conversely, the change in redox potential is strongest, if buried residues in close vicinity of the iron-sulfur cluster are exchanged, since they are now in the same medium of low dielectric constant. If the resulting dipole vector points towards the cluster an increase of the redox potential is expected and conversely a decrease if the dipole vectors points away. In cases where the dipole vector is almost perpendicular to the connecting vector (α around 90°), only a marginal effect is predicted. According to Zhou, the contribution E_i of the considered residue i to the redox potential is given as (eq. 5.3):

$$E_i \approx \frac{e \cdot \mu \cdot r}{\epsilon_p \cdot r^3} \tag{5.3}$$

containing the scalar product of μ (dipole vector of residue i) and r (vector connecting residue i and the iron-sulfur cluster). This term is multiplied by the electron charge e and divided by the third power of the distance of vector r times the dielectric constant of the protein ϵ_p, which is around four and thus much smaller than that of the aqueous surrounding ϵ_s. From these physical considerations it transpires that for the rational design of mutants with altered redox potential two factors are important: First, the location of the particular amino acid should be close to the iron-sulfur cluster and not on the protein surface, and second, the orientation of its dipole vector is crucial. The required three-dimensional atomic coordinates of the mutated amino acid can be obtained from corresponding homology models of the wild-type Adx.

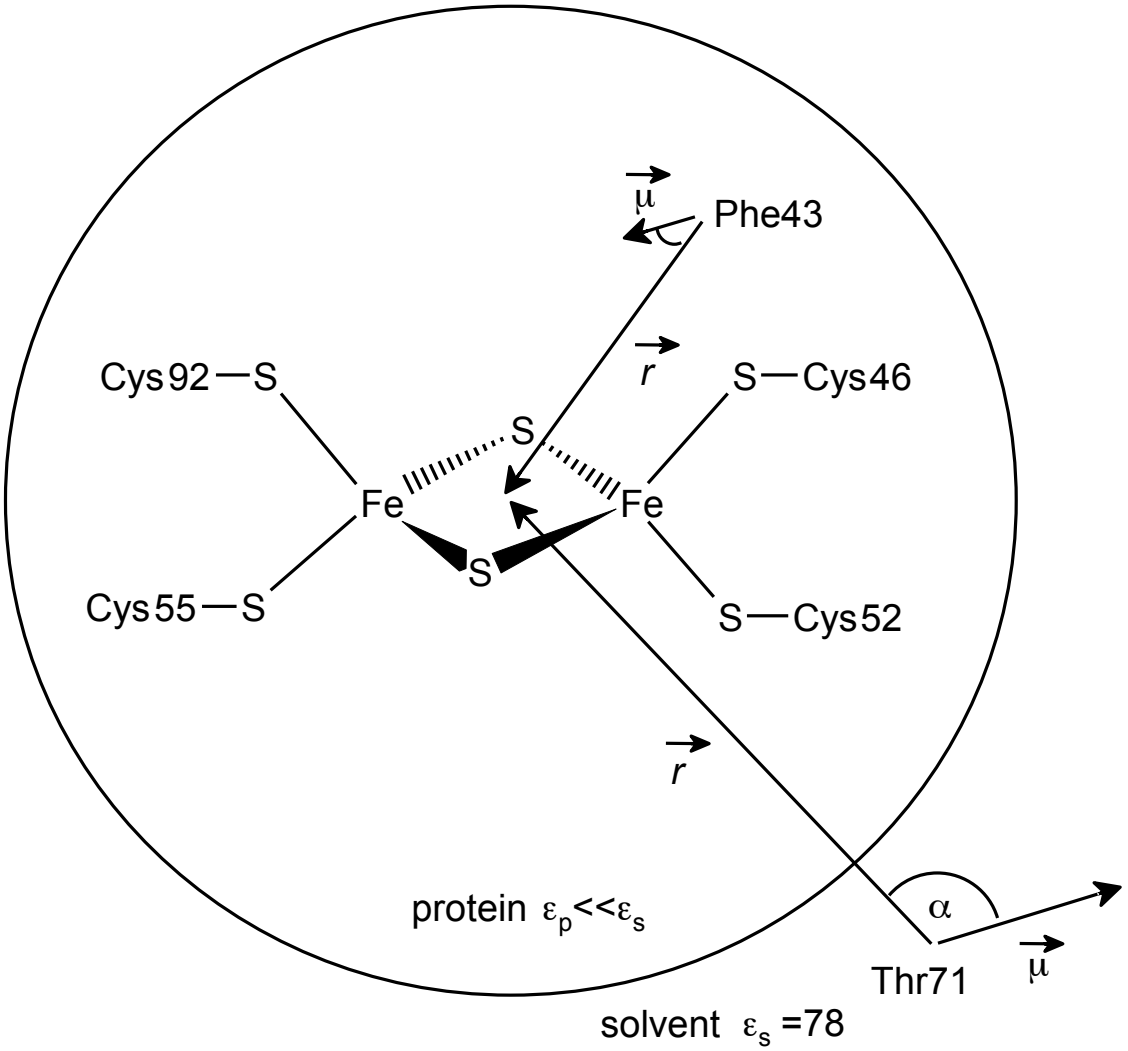

Figure 5.1: The redox protein is approximated as sphere of low dielectric constant ϵ_P surrounded by water. Mutation of amino acids on the protein surface do not change the redox potential notably, because their charges are strongly shielded from those of the iron-sulfur cluster due to the much larger ϵ_S of the solvent. The changes are expected to be most pronounced if the mutated residue is buried inside the protein (same low dielectric constant) and its dipole vector μ either points towards the iron-sulfur cluster ($\alpha \ll 90°$, leading to an increase of the redox potential), or pointing away ($\alpha \gg 90°$, decrease of the redox potential), whereas a perpendicular orientation ($\alpha \approx 90°$) would cause the scalar product between μ and r to vanish (see theory section for details)

5.2 Computational Methods

5.2.1 Density Functional Calculations

In order to obtain an estimate of the electronic reorganization energy upon reduction of the iron-sulfur cluster, we generated model systems of the Adx wild-type and

of the corresponding T49Δ mutant that were used in subsequent density functional calculations. For the model of the Adx wild-type the initial atomic coordinates of the $Fe_2S_2(SCH_2CH_3)_4$ fragment were taken from the crystal structure of bovine Adx (pdb entry 1AYF, chain B). The positions of the carbon atoms correspond to those of the C_β and C_α atoms of the four ligating cysteines, respectively. The starting geometry for the same fragment of the T49Δ mutant was taken from a homology model of bovine Adx as obtained from the SWISS-MODEL server [119, 120]. The deletion of threonine in position 49 causes a structural rearrangement that affects residues 46-50, whereby the side chain of Cys46 is one of the four cysteines that ligate the iron-sulfur cluster. Both model systems were energetically optimized in the oxidized state, respectively in the reduced state, applying the PBE0 density functional [121] and the def2-TZVP basis set [122] to all atoms as implemented in NWChem (Version 6.1) [123]. During geometry optimization the positions of the carbon atoms were kept frozen to their positions in the crystal structure, respectively to those in the homology model. To account for the antiferromagnetic coupling, the broken symmetry approach [124, 111] was applied, using the cdft option of NWChem to enforce the corresponding spin state for the two $FeS(SCH_2CH_3)_2$ halves of the iron-sulfur cluster. Subsequently, single point calculations on the obtained geometries were carried out using the larger def2-TZVPP basis set [122] and also applying following density functionals: BP86 [125, 126], PW91 [126], B3LYP [127, 128], M06 [129], as well as B(5%HF)P86 [130, 131]. Whereas BP86 and PW91 are local functionals, PBE0, B3LYP, M06, and B(5%HF)P86 are non-local hybrid functionals that involve various amounts of Hartree-Fock exchange. The latter was introduced by Szilagyi et al. [130, 131] based on superior performance for iron-sulfur clusters. In contrast to PB86 the exchange part consists of 5% Hartree-Fock and 95% of the Becke88 functional. Within NWChem the B(5%HF)P86 formalism was specified by the keyword line "xc becke88 0.95 hfexch 0.05 perdew86 1.0". Furthermore the keyword "grid fine" was used to request a higher grid resolution than the default. The def2-TZVPP basis set likewise showed to be a reasonable trade-off regarding computational demand, particularly in comparison to the 6-311+G(d) basis set [131]. The electronic reorganization energy upon one-electron reduction of the Adx was obtained as the difference in energy between the oxidized and reduced states of the two models systems. This energy corresponds to the adiabatic ionization potential of the iron-sulfur cluster in the reduced state (IP_{red}). The contribution to the shift of the redox potential in the T49Δ mutant due to the involved structural rearrangement of the iron-sulfur cluster can likewise be obtained from the difference of the adiabatic ionization potentials between wild-type and mutant. Results for the various density functionals are summarized in table 5.2.

distance/angle/dihedral angle	wild-type	T49Δ mutant
Fe1...Fe2	2.8856 (2.9503)	2.8075 (2.8494)
Fe1-S1	2.2369 (2.4903)	2.3388 (2.2087)
Fe1-S2	2.2369 (2.3830)	2.2180 (2.3078)
Fe2-S1	2.2251 (2.2078)	2.2291 (2.4709)
Fe2-S2	2.3525 (2.2829)	2.3585 (2.3377)
S1-Fe1-S2	101.8° (96.2°)	104.4° (108.7°)
Fe1-S2-Fe2	77.9° (78.4°)	75.6° (75.7°)
S2-Fe2-S1	102.1° (107.7°)	103.4° (99.5°)
Fe2-S1-Fe1	78.2° (77.5°)	75.8° (74.8°)
Fe1-S1-Fe2-S2	2.1° (2.8°)	6.8° (8.9°)

Table 5.1: Geometric data of the optimized iron-sulfur clusters[a]
[a] distances given in Angstrom for the oxidized states. Values for the reduced states are given in parenthesis.

5.3 Results

5.3.1 Electronic effect of the distorted iron-sulfur cluster

Structural optimization of the $Fe_2S_2(SCH_2CH_3)_4$ model compounds showed the expected increase of the iron-iron distance in the reduced state of 0.065 Å for the wild-type model and of 0.042 Å for the T49Δ mutant model. This is similar to the value used by Mouesca and coworkers (0.04 Å) for their $Fe_2S_2(SCH_3)_4$ model compounds based on structural data from synthetic analogs [111]. For the same molecular system Shoji and coworkers obtained a larger increase of 0.1 Å using a double zeta basis set and the B3LYP density functional [112]. The superposition of the wild-type and T49Δ mutant iron-sulfur cluster shows that the deletion of Thr49 mostly affects the side chain of Cys46 and in particular of S_γ, which is now in a different position with respect to the distal iron atom (see figure 5.2). Structural details of our optimized $Fe_2S_2(SCH_2CH_3)_4$ model systems are given in table 5.1. The calculated adiabatic ionization potentials IP_{red} and therefore also the contributions to the redox potential of the T49Δ mutant show a strong dependence on the density functional being used (see table 5.2). The PBE0 formalism seems to overestimate the electronic reorganization energy upon one electron reduction. The effect is, however, less pronounced for single point calculations using the larger TZVPP basis set. The overestimation is also present in the other non-local density functionals, which involve large amounts of Hartree-Fock exchange: M06 (27%), PBE0 (25%), and B3LYP (20%). These predict contributions of more than -90mV for the T49Δ mutant. The other used density functionals that either do not make use of Hartree-Fock exchange (BP86 and PW91), or only to a much smaller extent (5% in B(5%HF)P86), yield much smaller contributions (around -50mV). Therefore we attribute the overestimation to the large extent of Hartee-Fock exchange in the non-local density functionals. This is in line with the finding that the popular B3LYP

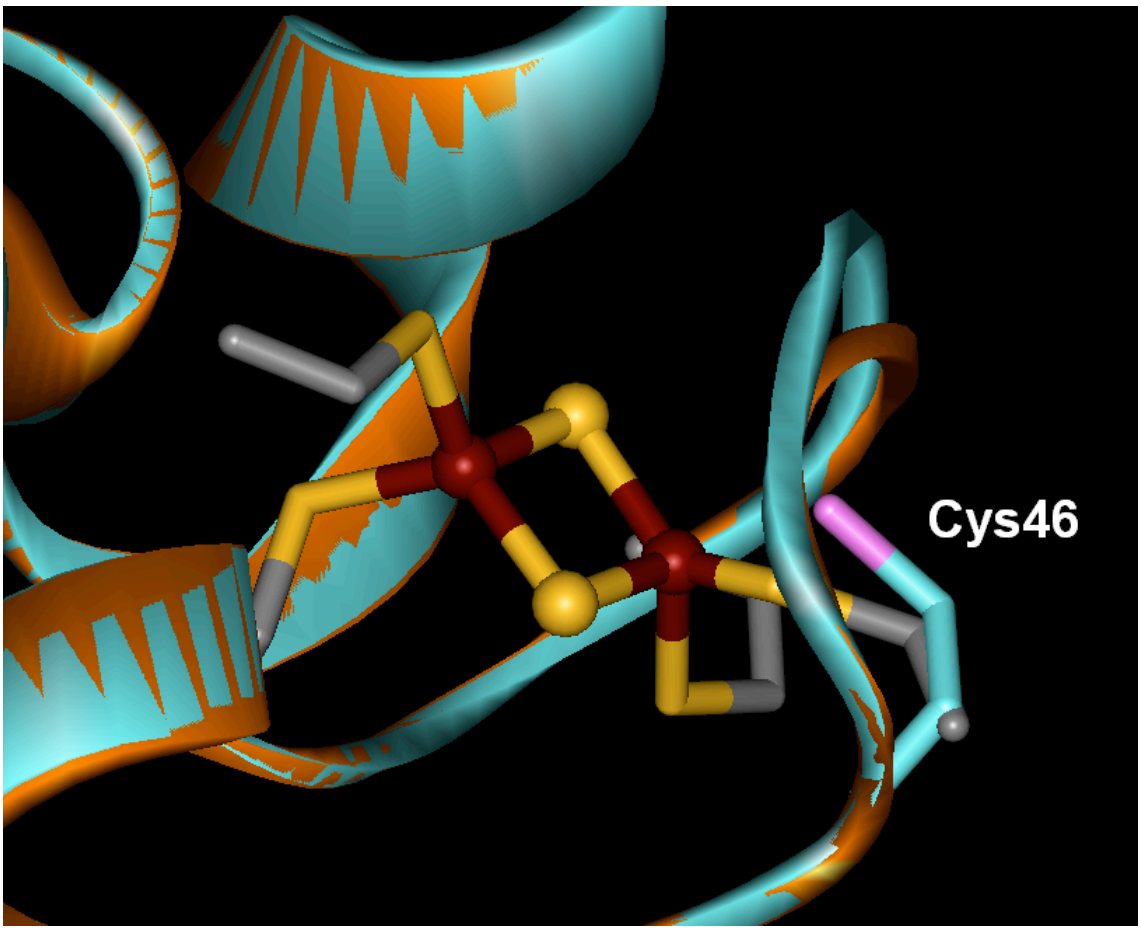

Figure 5.2: Superposition of the wild-type (cyan) and T49Δ mutant (orange) of Adx focusing on the region around the [2Fe2S] cluster. The deletion of Thr49 mostly affects the side chain of Cys46, whereby its S_γ atom (pink) coordinates the distal iron atom in a different angle. The structural changes lead to a substantially lower redox potential of this mutant

functional is less well suited for the calculation of electronic properties of iron-sulfur clusters [131]. Using the non-local density functionals, the obtained difference of the adiabatic ionization potentials between wild-type Adx and the T49Δ mutant is within 10mV, giving rise to a contribution between -44mV and -54mV to the experimentally observed -87mV change in redox potential. Thus the conformational changes in the T49Δ mutant, which predominately affect the side chain of Cys46, which is part of the coordination sphere of one iron atom, substantially contribute to the lower redox potential.

Level of Theory	$Fe_2S_2(SCH_2CH_3)_4[WT]^c$	$Fe_2S_2(SCH_2CH_3)_4[T49\Delta]^d$
PBE0/def2-TZVP	-4.335 eV	-4.505 eV (-171 mV)
PBE0/def2-TZVPP	-4.387 eV	-4.482 eV (-95 mV)
M06/def2-TZVPP	-4.520 eV	-4.632 eV (-112 mV)
B3LYP/def2-TZVPP	-4.333 eV	-4.427 eV (-94 mV)
PB86/def2-TZVPP	-4.714 eV	-4.758 eV (-44 mV)
PW91/def2-TZVPP	-4.822 eV	-4.872 eV (-50 mV)
B(5%HF)P86/def2-TZVPP	-4.619 eV	-4.673 eV (-54 mV)

Table 5.2: Calculated adiabatic ionization potentials (IP_{red}) of the iron-sulfur clusters and corresponding contribution to the shift of the Adx redox potential[a]
[a] The contribution to the shift of the redox potential is given in parenthesis. The experimentally observed shift for the T49Δ mutant is -87 mV.
[b] Single point calculations on the energetically optimized geometries of the oxidized and reduced states of the respective model systems at PBE0/def2-TZVP level of theory (please see computational methods section for details).
[c] Adx wild-type model.
[d] Adx T49Δ model.

5.3.2 Effect of mutations in position 108

So far shorter mutants down to position 108 have been made [113]. It was shown that the proline in this position is important for folding and stabilization of the protein, because mutants where the residues above position 107 are deleted exhibit higher sensitive to denaturation by urea [45]. Corresponding mutations create a cavity in the protein, because van der Waals contacts to neighboring amino acids (His56, Ile58, and Val107) are lost. Therefore even larger molecules present in the solvent, such as urea, now have access to the interior. Likewise we observed that other amino acids in position 108 such as arginine or lysine show only about 10% of the expression level of other Adx 4-108 mutants that contain the terminal proline.

5.3.3 Rational design of mutants

According to Zhou [56], a strong change of the redox potential would be expected by replacing a hydrophobic amino acid in the interior of Adx by a polar residue close to the FeS cluster. Furthermore it is necessary that the resulting dipole vector of the polar side chain either points towards the FeS cluster (resulting in an increase of the redox potential) or conversely away (leading to a decrease of the redox potential). From our structural alignment of Etp1 and Adx we identified Phe43 as promising position, because it is close to the FeS cluster and also buried within the protein. Electrostatic shielding is also stronger which should give a big impact on the redox potential. Furthermore Phe43 is replaced by leucine in Etp1, which indicates some variability in this position. To experimentally verify this simple model based on the theories of Born and Kirkwood regarding point charges and dipoles, we proposed

two amino acids, namely Asparagine and Glutamine, that should have opposite influences on the redox potentials, since their dipole moments point to different directions. This postulation was not experimentally evaluated so far.

5.4 Conclusion

In our efforts for designing new mutants of bovine Adx to lower its redox potential we first exploited why certain mutations have severe effects on the redox potential and how we could propose new mutants in the future.
Our computed adiabatic ionization potentials of the [2Fe2S] cluster in the geometries adopted in the wild-type and the T49Δ deletion mutant of bovine adrenodoxin strongly suggest that the experimentally observed difference between the corresponding redox potentials is essentially due to the changes in the geometry around the iron-sulfur cluster. The arrangement of the ligating cysteines in the protein environment causes the reduced state to become energetically less destabilized, thus facilitating electron uptake, which results in lower redox potentials, particularly for the T49Δ deletion mutant.
In an attempt to introduce new mutations we investigated the approach that was formulated by Zhou[56]. They used a simplified sphere model for protein and introduced the idea of setting dipole directions of amino acids in relation to the redox potential. The concept is extremely interesting and wasn't carried out on Adx so far, despite many mutations were tested over the years. Especially since an exchange of amino acid phenylalanine at position 43 wasn't considered yet. Hence for evaluation of our postulation an experimentally verification is still due. Building up on this one could be able to predict results of exchanges of amino acids. Furthermore this could be extended to other ferre- or flavodoxins affecting other CYP P450 redox systems.

Chapter 6

Engineering of versatile redox partner fusion enzymes that support monooxygenase activity of functionally diverse cytochrome P450s

The results and discussion sections (6.1 and 6.2) as well as figures 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7 and tables 6.1, 6.2, 6.3 were adapted unaltered from our publication [132], whereas the experimental section is omitted here.
This chapter was published as "Engineering of versatile redox partner fusion enzymes that support monooxygenase activity of functionally diverse cytochrome P450s" in collaboration with the labs of Prof. Dr. Vlada Urlacher from Düsseldorf and Prof. Dr. Bernhardt from Saarbrücken.
Not only does the dependence on NAD(P)H as an electron donor and redox partners [11] limit the application of CYPs as biocatalysts but also the identification of physiological redox partners [11, 133]. Therefore surrogate electron transfer systems are used to characterize the functional and biocatalytic application of CYPs. In different approaches the fusion of CYPs with redox partners have been established to reduce complexity as well as enhance the performance of the CYPs through faster electron transfer [134, 135, 91, 136, 137, 138, 139].
In this work the flavodoxin YkuN from *B. subtilis* was fused to the N-terminus of the flavodoxin reductase Fpr from *E. coli*. To gain insights to the influence of the linker region we varied the type as well as the length of it. Tested were flexible Glycine $(GGGGS)_n$ and stiffer Proline linkers $([E/L]PPPP)_n$ of various lengths (with n = 1 - 5). Subsequently their activity was measured by determining their ability to support catalysis to a variety of CYPs in vitro. A proof why certain linkers work better than others was supported by MD simulations.

6.1 Results

6.1.1 Construction of redox partner fusion enzymes consisting of *B. subtilis* YkuN and E. coli Fpr

The flavodoxin YkuN from *Bacillus subtilis* is a promiscuous electron carrier capable of transferring electrons to a variety of P450s, including endogenous CYP107H1 (P450 BioI) and CYP109B1, but also to heterologous P450s such as CYP154A8 from *Nocardia farcinica* or CYP154E1 from *Thermobifida fusca YX* [92, 89, 140]. Herein, YkuN is typically paired with the flavodoxin reductase Fpr from *E. coli*, which provides the NADPH-derived electrons. The "mixed" redox pair Fpr/YkuN often outperforms physiological redox pairs such as *E. coli* Fpr/FldA or *P. putida* PdR/Pdx [92, 140]. Consistent with this, Fpr/YkuN supported substantially higher conversion of myristic acid by CYP109B1 than the Fpr/FldA system (Fig. 6.1(a)).

Importantly, even in the case of natural CYP redox chains, such as PdR:Pdx:P450cam, excess of the ferredoxin Pdx is required for productive CYP catalysis in vitro [141]. Similarly, for Fpr/YkuN/CYP systems, a respective stoichiometry of 1:10:1 is usually sufficient to efficiently drive CYP catalysis in vitro [92, 140]. Thus surplus of flavodoxin or ferredoxin overcomes apparent rate-limiting steps in electron transfer [92, 141, 142]. In support of this, a 1:10:1 Fpr/YkuN/-CYP109B1 system achieved 36% higher conversion of myristic acid than a 1:1:1 system (fig. 6.1(a)), which is consistent with the function of YkuN as an electron shuttle that transfers the NADPH-derived electrons from the reductase to the CYP [89, 142]. A stimulatory effect was also noted for the alternative system using the *E. coli* flavodoxin FldA (fig. 6.1(a)). Nevertheless, independent of the stoichiometry applied the Fpr/YkuN redox pair proved to be superior to Fpr/FldA in supporting CYP109B1 catalysis, which makes YkuN a promising candidate for covalent fusion to Fpr. YkuN was fused to the N-terminus of Fpr, leaving the C-terminus of Fpr free, as it was demonstrated previously that attachment of FldA to the C-terminus of Fpr leads to reduced enzyme function, whereas constructs with the fusion partners in reversed order were substantially more active [91]. Attachment of the flavodoxin to the C-terminus of Fpr likely interferes with the function of the aromatic residues at the extreme C-terminus of Fpr, which are important for nucleotide binding and electron exchange [143, 144, 145]. It is worthy of note that homologous catalytic residues within the reductase domain of natural fusion proteins such as CPR and self-sufficient CYPs such as CYP102A1 are also found at the C-terminus [12, 146]. To optimize activity of the YkuN-Fpr (YR) fusion construct both flexible $(GGGGS)_n$ and rigid $([E/L]PPPP)_n$ linkers of different lengths (n = 1 - 5) were placed between the fusion partners. Thus, 11 different fusion constructs were created that were

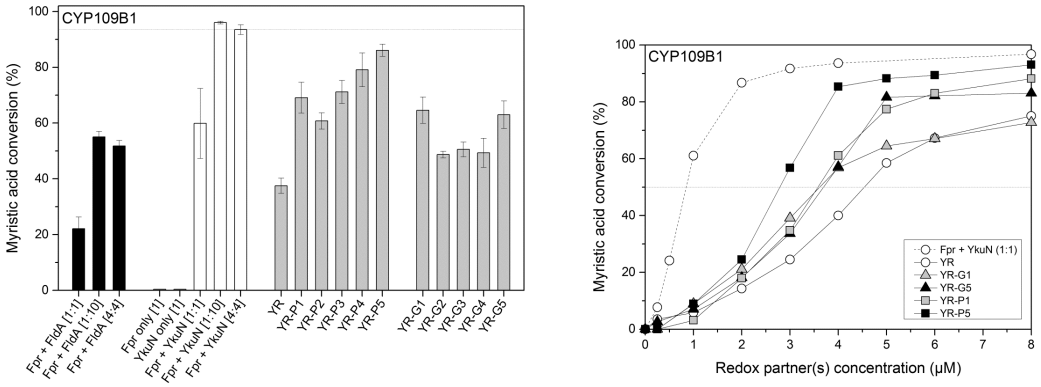

Figure 6.1: Myristic acid conversion by B. subtilis CYP109B1 supported by different redox partners. Reactions were started by the addition of a mixture of NADPH and myristic acid and allowed to proceed for 120 min under the support of an NADPH regenerating system. (a) CYP109B1 (1 μM) conversion reactions were carried out in the presence of non-fused redox partners Fpr and FldA (black bars), Fpr and YkuN (white bars) or with different YkuN-Fpr fusion constructs (grey bars). Reductase (Fpr) and flavodoxin (FldA or YkuN) together with CYP109B1 were employed at respective ratios of 1:1:1, 1:10:1 and 4:4:1. Reactions conducted with YkuN-Fpr (YR) fusion constructs and CYP109B1 were carried out at a respective ratio of 4:1. YR indicates the linker-less YkuN-Fpr fusion construct, whereas linker designations P1 - P5 and G1 - G5 correspond to linker sequences $(GGGGS)_n$ and $([E/L]PPPP)_n$ of different lengths (n = 1 - 5). The data presented are average values of 3 - 6 independent conversion reactions with indicated standard deviation. (b) Myristic acid conversion by CYP109B1 in the presence of different concentrations of selected fusion constructs or non-fused Fpr/YkuN. The ratio of non-fused redox partners was maintained at 1:1 (taken from [132]).

expressed in *E. coli*. The corresponding fusion proteins were subsequently isolated from the soluble protein fraction and purified to near homogeneity by IMAC, typically yielding 1 - 2 mg of purified fusion protein from 40 ml culture (OD_{600} = 4 - 5).

6.1.2 Ykun-Fpr fusion constructs support monooxygenase activity of CYP109B1

The ability of the fusion constructs (4 μM) to support CYP monooxygenase activity was tested in vitro with CYP109B1 (1 μM) using myristic acid as substrate (fig. 6.1(a)). Because the stoichiometry of the fused redox partners is fixed at 1:1, control reactions were conducted with 4 μM each of non-fused Fpr and YkuN. Independent of the flavodoxin used (FldA or YkuN), myristic acid conversion achieved with the 4:4:1 reconstituted system was nearly equal to that obtained with the commonly used 1:10:1 system (fig. 6.1(a)). Thus, in total less redox protein was required to

obtain a similar level of conversion. Importantly, the different fusion constructs were all able to support CYP109B1 catalysis (fig. 6.1(a)), which indicates that the covalently attached Fpr and YkuN have retained their ability to deliver the necessary electrons to CYP109B1. The linker-less fusion YR exhibited appreciable activity, supporting ∼38% conversion of myristic acid (fig. 6.1). Nevertheless, conversion was substantially lower than with equivalent amounts of non-fused Fpr/YkuN, suggesting reduced activity upon direct attachment of the redox partners. Introduction of a linker between YkuN and Fpr however substantially improved activity. In particular the fusion constructs YR-P4 and YR-P5 supported high myristic acid conversion (79% and 86%, respectively), which is close to the 94% conversion achieved with non-fused Fpr/YkuN (fig. 6.1(a)). Quantitative product analysis revealed that both the regioselectivity of CYP109B1 for myristic acid hydroxylation as well as product distribution was not affected by the fusion constructs.In all cases, the carbon atoms C-11 (ω_3) and C-12 (ω_2) of myristic acid were preferentially hydroxylated. Interestingly, the fusion constructs carrying proline-rich linkers (P1 - P5) in all cases outperformed their glycine-rich (G1 - G5) counterparts (fig. 6.1(a)). A similar phenomenon has previously been observed upon fusion of E. coli FldA to Fpr [91]. Moreover, the performance of the P-linker constructs seems to be dependent on the linker length, supporting higher conversion with increased linker length, whereas for the G-linker constructs a dependence on linker length was not evident (fig. 6.1(a)). Spectral analysis of the different YkuN-Fpr fusion constructs revealed nearly identical absorbance characteristics (fig. 6.2), which indicates that the fusion proteins have a similar co-factor content (Fpr-FAD and YkuN-FMN) as well as a similar cofactor-protein environment. It seems therefore unlikely that the differences in activity between the fusion constructs relate to linker induced effects on cofactor binding or protein folding. The performance of the fusion constructs was further investigated using different redox protein concentrations (0 - 8 μM). In all cases, sigmoid curves were obtained over the range of concentrations tested (fig. 6.1(b)). The dependence of CYP109B1 activity on the fusion protein concentration was clearly different from that of an equivalent system of non-fused Fpr/YkuN maintained at a 1:1 ratio, and also differed among the fusion constructs, which indicates rather complex behaviour (fig. 6.1(b)). Overall, YR-P5 proved the most effective fusion construct, as it supported nearly complete conversion of myristic acid at protein concentrations of 4 - 8 μM, much like the non-fused redox partners. However, at concentrations <4 μM, the performance of YR-P5 dropped well below that of the non-fused Fpr/YkuN (fig. 6.1(b)). A reduced performance at low fusion protein concentration was also observed for the other tested fusion constructs (fig. 6.1(b)). As a measure for the overall performance of the redox partners, the concentration of the redox enzymes at which 50% of myristic acid (EC_{50}) was converted by CYP109B1 was estimated from fig. 6.1(b). The approximate EC_{50} are 0.8 μM for non-fused Fpr/YkuN, 2.7 μM for YR-P5, 3.5 μM for both YR-P1 and YR-G1, 3.6 μM for YR-G5 and 4.5 μM for the linker-less construct YR, which demonstrates that YR-P5 is the superior fusion construct.

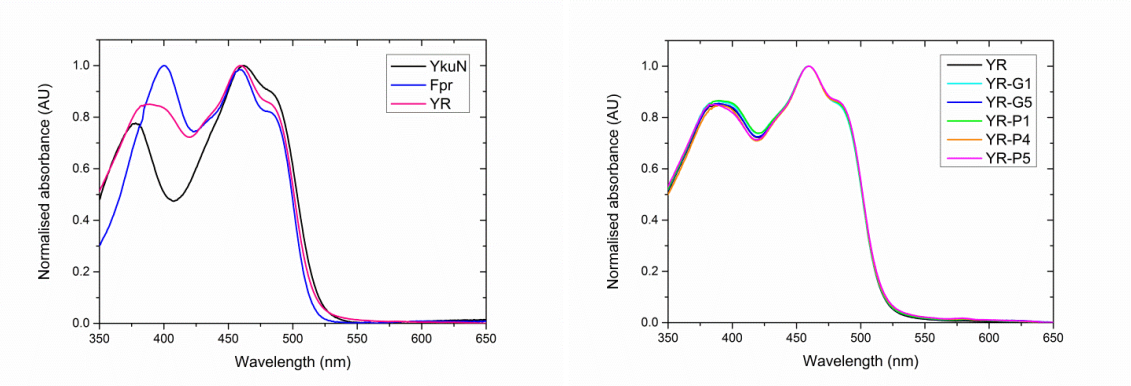

Figure 6.2: Spectral properties of the individual and fused redox partners, YkuN and Fpr. Arbitrary amounts of indicated redox protein were diluted in appropriate buffer (A, 50mM KP_i pH 7.5, 500 mM NaCl, 20% (v/v) glycerol, and B, 50 mM Tris-HCl pH 7.5), typically yielding final concentrations in the range of 4 - 8 μM. Spectra were recorded at 25°C using a thermostated Perkin Elmer Lambda 35 spectrometer. Presented normalised spectra are the average of 3 recorded spectra that were corrected for buffer absorbance (taken from [132]).

6.1.3 Intrinsic properties of the $(GGGGS)_n$ and $([E/L]PPPP)_n$ linkers

To further evaluate the properties of the $(GGGGS)_n$ and $([E/L]PPPP)_n$ linkers molecular dynamics (MD) simulations were carried out. MD simulations of the G-linkers revealed that starting from a linear extended conformation these linkers readily fold into more compact random structures (fig. 6.3(a)). Collapse of the extended conformation is particularly evident for linkers G3, G4 and G5, which consist of 3, 4 and 5 (GGGGS)-repeats, respectively. These linkers exhibited a substantial (>1 Å) decrease in the radius of gyration (Rg) over time (fig. 6.3(b)). Moreover, the Rg of the G4 and G5 linker exhibited large oscillations over time, which indicates substantial structural rearrangements and high structural flexibility. Structural behaviour of the $([E/L]PPPP)_n$ linkers on the other hand was clearly different. The Rg of the different P-linkers remained largely constant over time, with only minor oscillations observed for linkers P3 - P5 (fig. 6.3(c)). Thus, the P-linkers effectively maintained their linear conformation. Consequently, the Rg of the P-linkers exhibited a dependence on the linker length. For each additional ([E/L]PPPP)-segment the Rg increased by ∼0.4 - 0.5 Å (fig. 6.3(c)). Linker P5, which consists of 25 residues forms an exception to this as it exhibited an Rg similar to that of the P4 linker, i.e. ∼1.8 Å (fig. 6.3(c)). Overall bending of the long P-linkers may have contributed to this occurrence (fig. 6.3(a)).

6.1.4 Influence of auxiliary YkuN on CYP109B1 catalysis driven by the YkuN-Fpr fusion constructs

In comparison to the separate redox enzymes the fusion constructs exhibit a reduced performance at low protein concentration (fig. 6.1(b)), which may indicate that one (or both) of the fused redox partners is limiting the P450-catalysed reaction. With reconstituted P450 systems the electron shuttle ferredoxin/flavodoxin is usually a major limiting factor that is typically overcome by applying an excess of this electron transfer protein [92, 141, 142]. To investigate this, the different fusion constructs (1μM) were incubated together with CYP109B1 (1μM) in the absence or presence of an excess of YkuN (4 or 9μM). In all cases, mixing the fusion proteins with additional YkuN led to an increase in myristic acid conversion (fig. 6.4), which confirms that YkuN is limiting the reaction. By increasing the concentration of the fusion constructs itself in the P450-catalysed reaction, this apparent limitation was also (partly) overcome (fig. 6.1(b)).

Since the reductase (Fpr) domain of the fusion constructs is capable of donating electrons to the auxiliary electron shuttle YkuN, this also indicates that no strict functional coupling between the Fpr and YkuN domain exists within the fusion constructs. Generally, electron transfer rates are thought to be controlled in part by the distance between the redox centres [147]. Considering the YkuN-Fpr fusion constructs it is likely that both the length and the structural properties of the linker contribute to the relative distance between the respective redox centres. Regarding the rigid nature of the $([E/L]PPPP)_n$ linkers (fig. 6.3) it seems plausible that in particular long P-linkers might facilitate a higher degree of separation of the fusion partners. In line with this, the increase in conversion upon YkuN addition was much more pronounced for fusion constructs carrying the rigid P-linkers (fig. 6.4), which more effectively maintain an extended conformation than their glycine-rich counterparts (fig. 6.3). In fact, the highest increase in myristic acid conversion governed by the addition of YkuN was observed for the YR-P5 construct which carries the longest P-linker tested (25 residues), achieving a similar conversion as the comparable 1:10 system of non-fused Fpr/YkuN (fig. 6.4).

6.1.5 NADPH oxidation rates and coupling efficiencies measured with the YkuN-Fpr fusion enzymes

The overall performance of a reconstituted P450-redox partner system depends amongst others on its NAD(P)H oxidation rate and the coupling efficiency between NAD(P)H oxidation and P450 substrate conversion. Since delivery of electrons to CYP109B1 depends on functional reductase (Fpr) as well as flavodoxin (YkuN), a loss of electrons may occur either upon electron transfer from Fpr to YkuN or from YkuN to CYP109B1, which in both cases will lead to uncoupling between NADPH oxidation and product formation.

The NADPH oxidation rate and the coupling efficiency were determined for various reconstituted systems (Table 6.1). CYP109B1 reconstituted with the non-fused redox partners exhibited the highest NADPH oxidation rate, i.e. 26 $nmol(nmol \times P450)^{-1}min^{-1}$, along with a coupling efficiency of 28.3%. In contrast, for the fusion constructs substantially lower NADPH consumption rates were observed, whereas the coupling efficiency was markedly improved (table 6.1). Direct attachment of YkuN to Fpr (YR) led to a ∼1.8-fold improvement of the coupling efficiency (49.8%) and a more than 8-fold decrease in NADPH oxidation rate (table 6.1). In comparison to YR, the superior construct YR-P5 (fig. 6.1) showed a similar coupling (48.9%), whereas the NADPH oxidation rate was 3.6-fold higher, i.e. 11.3 $nmol(nmol \times P450)^{-1}min^{-1}$ (table 6.1). The highest coupling efficiency of 81.2% was however observed for YR-G5, which, notably, exhibited the lowest NADPH oxidation rate of 1.4 $nmol(nmol \times P450)^{-1}min^{-1}$.

Reconstituted system [a]	NADPH oxidation rate [b]	Coupling efficiency (%)
Fpr / YkuN / CYP109B1 - [4:4:1]	26.0 ± 1.9	28.3 ± 3.6
YR / CYP109B1 - [4:1]	3.1 ± 0.3	49.8 ± 8.3
YR-P1 / CYP109B1 - [4:1]	12.7 ± 1.3	62.1 ± 5.1
YR-P5 / CYP109B1 - [4:1]	11.3 ± 1.9	48.9 ± 3.4
YR-G1 / CYP109B1 - [4:1]	6.6 ± 0.3	72.4 ± 3.6
YR-G5 / CYP109B1 - [4:1]	1.4 ± 0.2	81.2 ± 4.1

Table 6.1: NADPH oxidation rate and coupling efficiency of the CYP109B1-catalysed conversion of myristic acid, supported by different redox partners. Presented data represent average values of at least three independent myristic acid conversion reactions.
[a] Values in brackets indicate applied ratio as well as final concentration (μM) of indicated proteins.
[b] Rates are given in nmol NADPH per nmol CYP109B1 per minute. The background NADPH consumption rate in the absence of redox partner(s) was 0.1 ± 0.0 (taken from [132]).

6.1.6 Reduction of CYP109B1 Fe^{3+}-heme by the YkuN-Fpr fusion constructs

Since the YkuN-Fpr fusion constructs exhibited attenuated NADPH oxidation rates along with improved coupling, the question arises whether the crucial step of electron transfer to the P450 is also influenced upon attachment of YkuN to Fpr. The reduction of the Fe^{3+} to Fe^{2+}-heme can be conveniently monitored by the addition of carbon monoxide (CO), yielding a stable complex that exhibits the characteristic absorbance maximum at 450 nm [148]. Indeed, absorbance spectroscopy of the anaerobic reduction of CYP109B1 in the presence of CO by the different redox partners revealed the formation of the typical absorbance peak at ∼450 nm, which is governed by the reduction of Fe^{3+} to Fe^{2+}-heme and the subsequent formation of the stable heme-Fe^{2+}-CO complex. Typical traces for the time dependent reduction are

shown in fig. 6.5. CYP reduction occurred rather slowly under the tested conditions (fig. 6.5). Data were fit to a bi-exponential function, revealing a slow and a fast phase; the corresponding reduction rates are shown in table 6.2. Reduction occurred

Reconstituted system [a]	Heme-iron reduction rate [b]	
	$k_1(s^{-1})$	$k_2(s^{-1})$
Fpr / FldA / CYP109B1 - [4:4:1]	0.004 ± 0.001	0.027 ± 0.008
Fpr / YkuN / CYP109B1 - [4:4:1]	0.009 ± 0.001	0.078 ± 0.023
YR / CYP109B1 - [4:1]	0.002 ± 0.001	0.030 ± 0.002
YR-P5 / CYP109B1 - [4:1]	0.008 ± 0.002	0.045 ± 0.005
YR-G5 / CYP109B1 - [4:1]	0.007 ± 0.001	0.035 ± 0.002

Table 6.2: Reduction of B. subtilis CYP109B1 $Fe3^{+}$-heme by different redox partners.
[a] Values in brackets indicate applied ratio as well as final concentration (μM) of the proteins in the employed reconstituted systems.
[b] The heme-iron reduction rates were measured as described in the Methods section. Kinetic traces were fit to a bi-exponential function, revealing a slow and a fast phase. Typical kinetic traces and corresponding fits are shown in the online Supplemental Information. Presented reduction rates represent average values of at least three independent reactions carried out under anaerobic conditions at 20°C (taken from [132]).

fastest by the non-fused Fpr/YkuN, exhibiting a k^1 of 0.009 s^{-1} and k^2 of 0.078 s^{-1}. For the linker-less fusion YR corresponding rates were 4.5 and 2.6-fold decreased, respectively. In contrast, for the fusion constructs carrying the long P5 or G5 linker, $k_1$was nearly identical to that of the non-fused Fpr/YkuN and k_2 was decreased to a lesser extent (table 6.2). Of the tested fusion constructs, the superior YR-P5 also exhibited the highest heme-iron reduction rate ($k^1 = 0.008$ s^{-1} and $k^2 = 0.045$ s^{-1}). Thus, introduction of the P5 linker (25 residues) between YkuN and Fpr led to a substantial increase of the heme-iron reduction rate.

6.1.7 Versatility of the YkuN-Fpr redox fusion enzymes

From a biotechnological point of view, a redox partner system desirably should have the ability to transfer electrons to different terminal acceptors. To demonstrate cross-reactivity, selected fusion constructs were additionally tested with CYP154E1 from *Thermobifida fusca YX*, CYP106A2 from *Bacillus megaterium ATCC 13368* and bovine CYP21A2. CYP154E1 is a versatile monooxygenase that converts a large variety of substrates, including fatty acids and alcohols, as well as acyclic and bulky cyclic terpenoids [140, 149, 150]. Since Fpr/YkuN are known to effectively support CYP154E1 catalysis [140, 149, 150] it is expected that the YkuN-Fpr fusion constructs are also able to act as surrogate redox partners. To investigate this, selected fusion constructs were tested for their ability to support CYP154E1 catalysis using β-ionone as substrate (fig. 6.6). Ionones are cyclic terpenoids that are key fragrance components used for the production of perfumes, cosmetics and other fine chemicals [151]. Moreover, β-ionone is an important intermediate in the

manufacturing of vitamin A, while its oxygenated derivative 4-hydroxy-β-ionone is a key intermediate in the synthesis of carotenoids and the plant hormone abscisic acid [152, 153, 154].

Non-fused Fpr/YkuN (1:10 or 4:4 system) effectively supported CYP154E1 catalysis, achieving >70% conversion of β-ionone in 1 h (fig. 6.6). Thus, in addition to the previously reported terpenoid substrates [150], β-ionone is also converted by CYP154E1. As expected, the fusion constructs were able to support CYP154E1 catalysis, of which YR-P5 again proved most effective at driving CYP catalysis (fig. 6.6). In the presence of YR-P5 ∼1 mM of β-ionone was converted in 1 h, yielding a turnover of ∼16 min^{-1}, while non-fused Fpr/YkuN (4:4) supported a turnover of ∼24 min^{-1}. The YR-P5/CYP154E1 (4:1) system achieved a coupling efficiency of ∼50%, whereas with the non-fused enzymes ∼44% was measured. Furthermore, under cofactor regeneration conditions, total turnover numbers (TTN) were 1582 with YR-P5 and 1948 with the non-fused Fpr/YkuN. Regardless whether Fpr and YkuN were fused or not, in all cases a major oxidation product was formed, which was identified as 4-hydroxy-β-ionone by comparison to MS analysis of an authentic reference substance [155].

To further substantiate functional promiscuity, the best construct YR-P5 was tested with CYP106A2 and CYP21A2 using progesterone as substrate in both cases. It is of note that these CYPs have not been tested previously with Fpr/YkuN as surrogate redox partners. CYP106A2 is a regio- and stereoselective 15β-hydroxylase of 3-oxo-Δ^4-steroids [19, 97] that recently has been demonstrated to also accept 3-hydroxy-Δ^5-steroids as well as di- and triterpenes as substrates [156, 28]. Herein, the bovine adrenodoxin reductase (AdR) and adrenodoxin (Adx) typically serve as surrogate redox partners. Interestingly Fpr/YkuN and its fused derivative YR-P5 were able to functionally substitute for AdR/Adx in the CYP106A2-mediated conversion of progesterone (table 6.3). Moreover, performance of the 4:4:1 Fpr/YkuN/CYP106A2 system was virtually identical to that of the 4:4:1 AdR/Adx/CYP106A2 system, both with respect to progesterone conversion and product distribution (table 6.3).

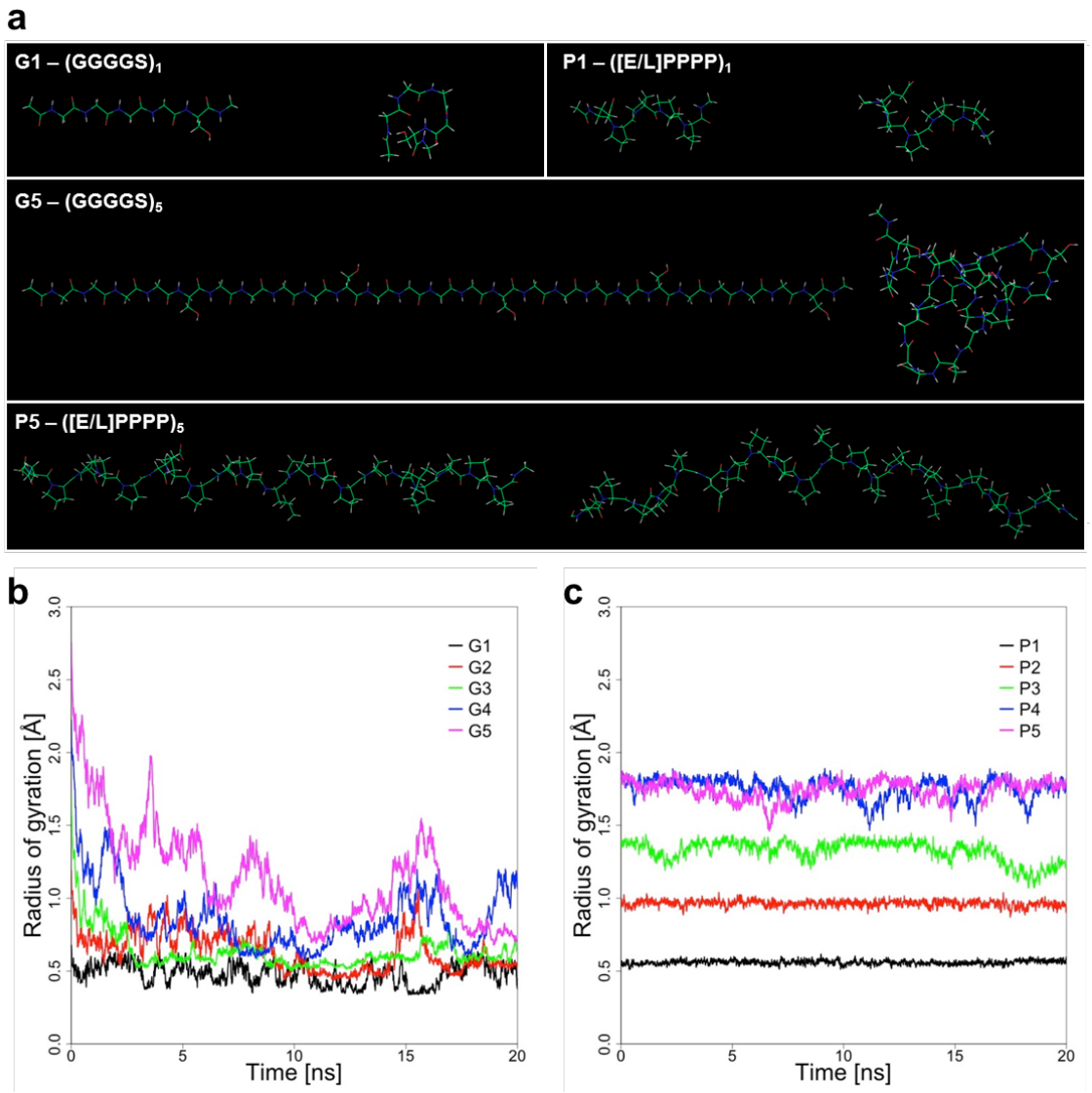

Figure 6.3: MD simulations of the $(GGGGS)_n$ and $([E/L]PPPP)_n$ linkers used to functionally connect YkuN and Fpr. (a) Structural snapshots of $(GGGGS)_n$ and $([E/L]PPPP)_n$ linkers of length 1 and 5. Depicted in each panel are the starting conformations (left) and the lowest energy conformation found during the 20 ns MD simulations (right). Development of the radius of gyration (Rg) during MD simulations for $(GGGGS)_n$ and $([E/L]PPPP)_n$ linkers of lengths (n = 1 - 5) are shown in (b) and (c), respectively. The strong decrease in Rg for $(GGGGS)_n$ linkers reflects the hydrophobic collapse, whereas the $([E/L]PPPP)_n$ linkers remain in an extended conformation (taken from [132]).

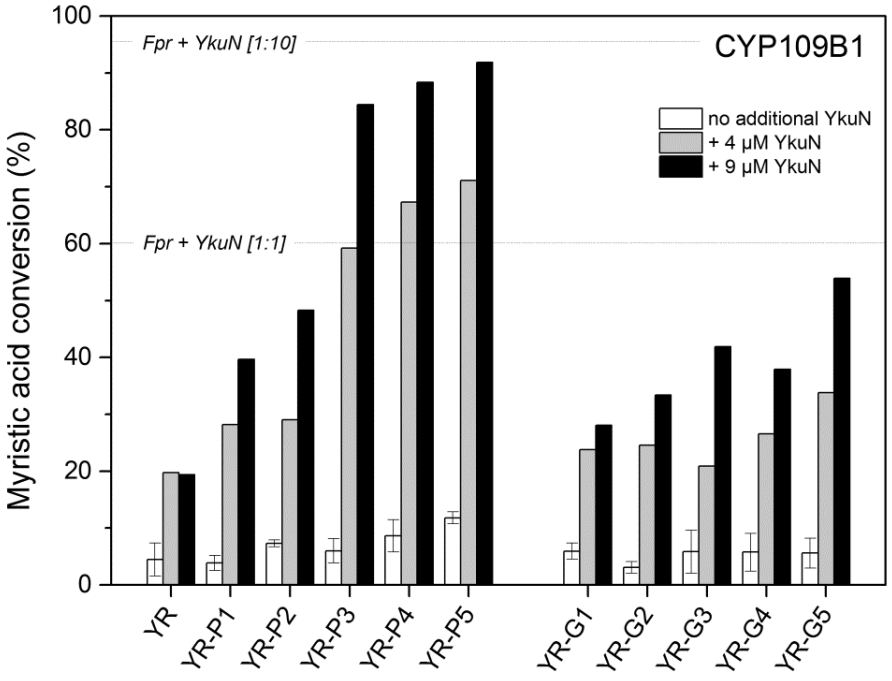

Figure 6.4: Influence of auxiliary YkuN on CYP109B1 catalysis driven by the different YkuN-Fpr fusion constructs. Myristic acid conversion with CYP109B1 (1 μM) was carried out in the presence of the different Ykun-Fpr fusion constructs either in the absence (-YkuN) or presence of additional YkuN (4 or 9 μM). In all cases reactions were started by the addition of a mixture of NADPH and myristic acid and allowed to proceed for 120 min under the support of an NADPH regenerating system. The Fpr/YkuN/CYP109B1 system reconstituted at 1:10:1 and 1:1:1 ratio, respectively achieved 96% and 60% conversion of myristic acid in 120 min (taken from [132]).

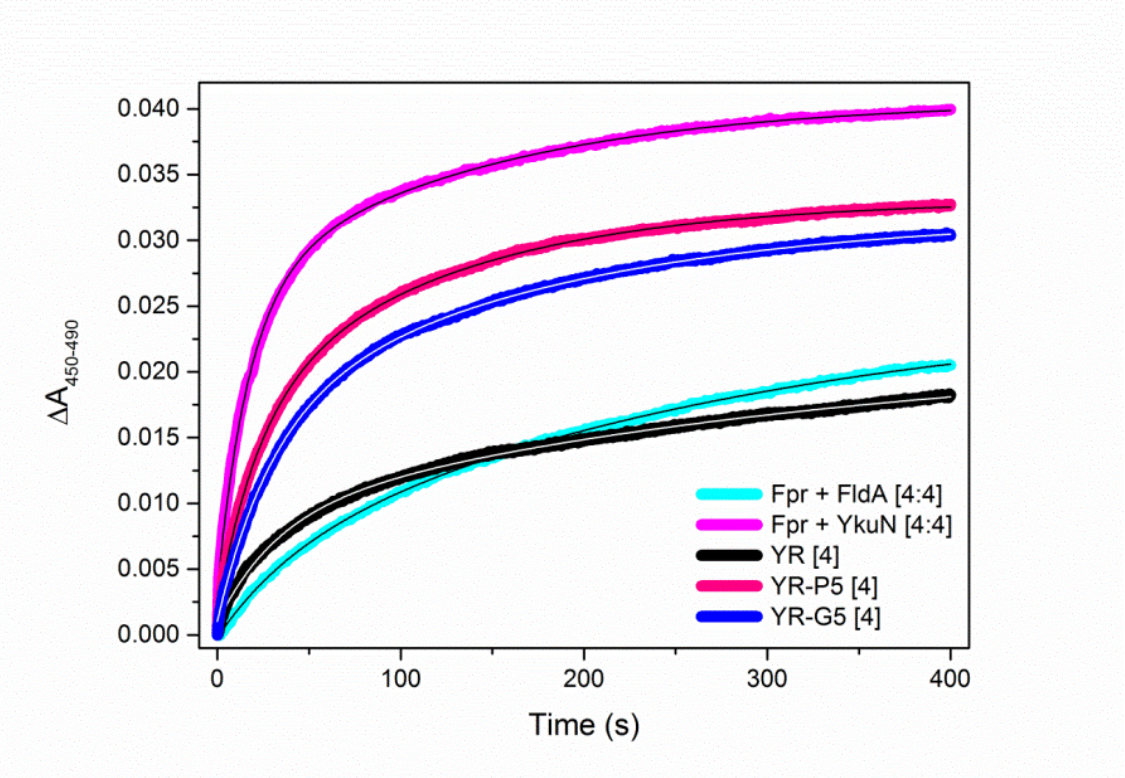

Figure 6.5: CYP109B1 heme iron reduction by different redox partners (taken from [132]).

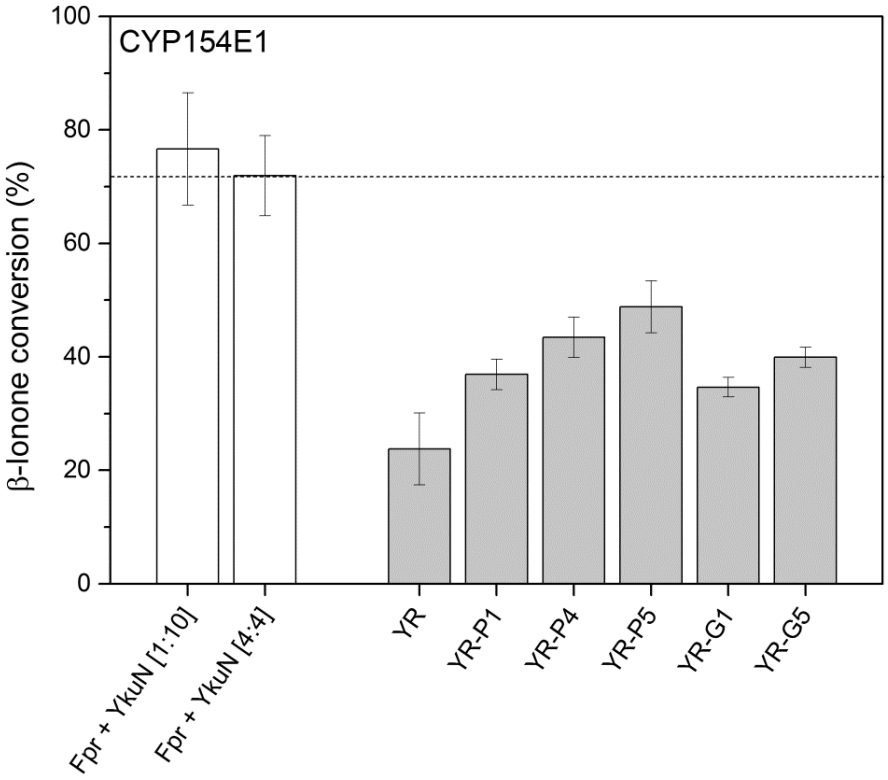

Figure 6.6: Conversion of β-ionone by *T. fusca YX* CYP154E1 supported by selected fusion constructs. Conversion reactions were carried out in the presence of 1 μM CYP154E1 and 4 μM of indicated fusion constructs. Control reactions with non-fused Fpr/YkuN were conducted with 1 μM CYP154E1 together with either 1 μM Fpr + 10 μM YkuN [1:10], or 4 μM Fpr + 4 μM YkuN [4:4]. In all cases the initial concentration of the substrate β-ionone was 2 mM. Reactions were started by the addition of NADP+ under support of an NADPH regenerating system and stopped after 1 h (taken from [132]).

Redox partners	Ratio [a]	Conversion time (min)	Progesterone conversion(%)	Compounds(%)			
				Prog [b]	15 βOH [b]	mono-OH [b]	poly-OH [b]
AdR/Adx	(4:4:1)	30	96.0 ± 0.5	4.0 ± 0.5	81.7 ± 1.0	8.7 ± 1.3	5.6 ± 0.9
		120	99.0 ± 0.1	1.0 ± 0.1	81.1 ± 0.5	6.9 ± 0.5	11.0 ± 0.7
	(10:10:1)	30	96.0 ± 0.6	4.0 ± 0.6	81.5 ± 0.5	7.9 ± 0.1	6.6 ± 0.2
		120	96.7 ± 1.0	3.3 ± 1.0	82.6 ± 1.6	7.9 ± 0.3	6.1 ± 1.4
Fpr/YkuN	(4:4:1)	30	92.7 ± 0.7	7.3 ± 0.7	80.4 ± 0.5	4.0 ± 0.4	8.3 ± 0.1
		120	98.9 ± 0.0	1.1 ± 0.0	81.9 ± 0.0	4.3 ± 0.0	12.8 ± 0.0
	(10:10:1)	5(500μM)[c]	84.7 ± 5.4	15.3 ± 5.4	80.4 ± 3.1	2.0 ± 0.2	2.2 ± 2.2
		30	91.5 ± 1.3	8.5 ± 1.3	63.0 ± 1.0	4.4 ± 0.1	24.1 ± 0.2
		60	94.0 ± 0.1	6.0 ± 0.1	58.4 ± 0.7	5.2 ± 0.7	30.5 ± 0.1
		90	98.5 ± 0.4	1.5 ± 0.4	68.2 ± 4.2	6.5 ± 0.4	23.8 ± 4.7
		120	97.8 ± 0.2	2.1 ± 0.2	66.3 ± 2.6	6.4 ± 0.4	25.2 ± 3.0
YR-P5	(4:1)	30	19.5 ± 2.6	80.5 ± 2.6	19.1 ± 2.1	0.4 ± 0.5	-
		60	29.6 ± 5.4	70.4 ± 5.4	29.2 ± 4.9	0.4 ± 0.6	-
		90	39.7 ± 16.2	60.3 ± 16.2	39.7 ± 16.2	-	-
		120	45.9 ± 15.2	54.1 ± 15.2	44.0 ± 13.4	1.9 ± 1.9	-
	(10:1)	5(500μM)[c]	29.7 ± 0.6	70.3 ± 0.6	29.7 ± 0.6	-	-
		30	75.9 ± 3.8	24.1 ± 3.8	71.3 ± 4.2	3.4 ± 0.5	1.2 ± 0.9
		60	80.0 ± 0.9	20.0 ± 0.9	74.9 ± 1.4	3.2 ± 0.8	1.8 ± 0.3
		90	90.7 ± 5.3	9.3 ± 5.3	81.5 ± 3.7	3.6 ± 0.8	5.6 ± 1.0
		120	94.9 ± 0.7	5.1 ± 0.7	84.9 ± 0.7	4.0 ± 1.3	6.1 ± 0.2

Table 6.3: Product distribution for the CYP106A2-catalysed conversion of progesterone supported by different redox partners. Data represent average values of three independent conversion reactions with indicated standard deviation.

[a] Redox partner-CYP106A2 ratio of reconstituted system.

[b] *B. megaterium* CYP106A2 hydroxylates progesterone at positions 15β, 6β, 11α and 9α [82]; mono-OH, monohydroxylated progesterone at positions other than 15β; poly-OH, di- or polyhydroxylated progesterone.

[c] To assess turnover, conversion reactions were carried out for only 5 min in the presence of 500 μM progesterone (taken from [132]).

With both systems nearly all progesterone was converted after 120 min and 15β-hydroxyprogesterone was the main product formed (table 6.3). The fusion construct YR-P5 also supported CYP106A2 catalysis, albeit with lower efficacy (table 6.3). The YR-P5/CYP106A2 (4:1) system achieved ∼46% conversion after 120 min, producing almost exclusively 15β-OH-progesterone (table 6.3). By increasing the YR-P5 concentration (10:1 system) conversion could be enhanced to near completion (95%) after 120 min, while only minor amounts of undesired polyhydroxylated progesterone were formed (up to 6%). To assess turnover rates, reactions were carried out under conditions in which essentially no overoxidation products were formed (5 min reactions using 500 μM progesterone). With the Fpr/YkuN/CYP106A2 system (10:10:1) nearly 85% of the progesterone was converted, thus achieving a turnover of 169 min^{-1}, whereas with the YR-P5/CYP106A2 system (10:1) nearly 30% progesterone conversion was achieved, thus yielding a turnover of 59.4 min^{-1}. To further illustrate versatility of YR-P5, bovine CYP21A2, which is a membrane-bound microsomal P450, was used as terminal electron acceptor. CYP21A2 is involved in the biosynthesis of steroid hormones, and typically catalyses the hydroxylation of the carbon atom 21 in steroids [157]. In general, CYP21A2 obtains the necessary electrons from NADPH via an FAD and FMN-containing microsomal CPR [12]. However, it is also able to use AdR and Adx as redox partners [158]. Notably, YR-P5 supported CYP21A2 catalysis; with YR-P5 40% of the (200 μM) progesterone was converted to 21-hydroxyprogesterone in 30 min, while ∼80% conversion was achieved with non-fused Fpr/YkuN (Fig. 6.7). Taken together, the Fpr/YkuN redox pair and its fused derivative YR-P5 are able to support the activity of the steroidogenic CYP106A2 and CYP21A2 in vitro, which indicates that these bacterial electron transfer systems may serve as alternatives to the mammalian electron transfer systems AdR/Adx and CPR.

6.2 Discussion

Flavodoxins are promiscuous electron carriers that donate electrons to structurally and functionally diverse enzymes, including pyruvate-formate lyase [159], ribonucleotide reductase [160], key enzymes in photosynthesis [161], nitrogen fixation [162], methionine [163] and biotin [164] synthesis, but also P450s [92, 89, 142, 165, 166]. The mixed redox pair Fpr/YkuN outperformed the physiological redox pair Fpr/FldA from *E. coli* in supporting CYP109B1 catalysis (fig. 6.1(a)), which suggests that electron transfer by the flavodoxin YkuN is more effective. Since YkuN and CYP109B1 both originate from *B. subtilis* it is possible that YkuN is the physiological redox partner of CYP109B1 and therefore a more favourable electron carrier than the heterologous FldA.
Typically, the midpoint potentials of the oxidised/semiquinone (E1') and semiquinone/hydroquinone (E2') couples reported for the short-chain flavodoxin YkuN are higher than those for the long-chain flavodoxin FldA, i.e. -105 mV and

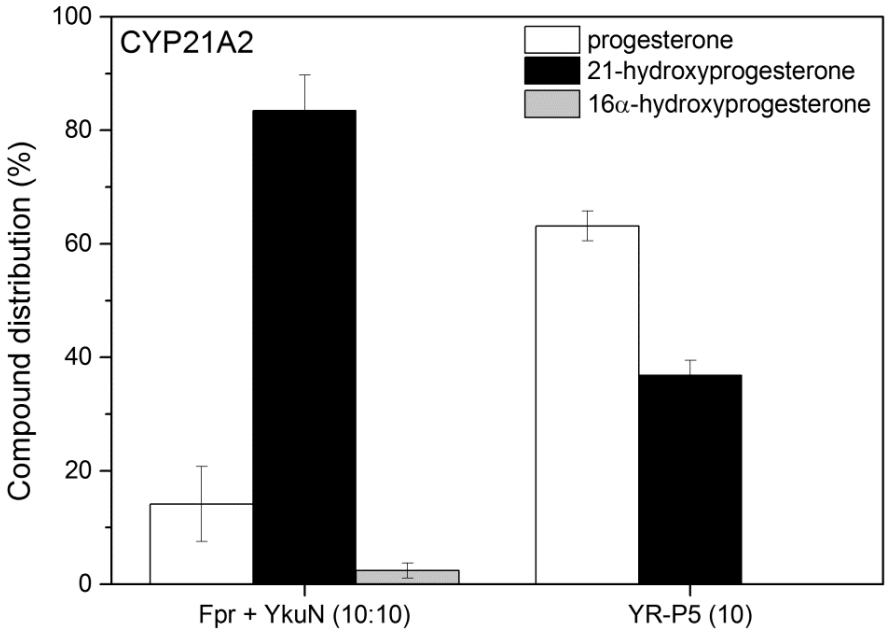

Figure 6.7: Conversion of progesterone by bovine CYP21A2 supported by Fpr/YkuN and their fused derivative, YR-P5. In all cases conversion reactions were carried out in the presence of 0.5 μM CYP21A2, 200 μM progesterone and 5 μM of indicated redox partner(s). Corresponding ratios of redox partner(s) relative to CYP21A2 are indicated in brackets. Reactions were started by the addition of NADPH and supported by an NADPH regenerating system. Reactions were stopped after 30 min (taken from [132]).

-382 mV [89] vs. -254 mV and -433 mV, respectively [166]. E1' and E2' of the employed *E. coli* reductase Fpr are -308 mV and -268 mV, respectively [166]. Indeed, the anaerobic reduction of CYP109B1 heme-iron by YkuN occured 2-3 times faster than by FldA (table 6.2). Similarly, stopped-flow experiments monitoring the reduction of *B. subtilis* CYP107H1 have indicated that YkuN is more effective than FldA in first electron transfer to palmitoleate-bound CYP107H1, as evidenced by a ∼12-fold higher k_{red} (2.38 s_{-1} vs. 0.19 s_{-1}) [89]. Thus, accelerated electron transfer likely contributed to the overall better performance of YkuN compared to FldA in the CYP109B1-catalysed conversion of myristic acid (table 6.2, fig. 6.1(a)).

The inherent dependence of P450s on electron transfer proteins presents a challenging limitation in their biotechnological exploitation. To simplify redox chains and to improve the catalytic properties of P450 systems, a variety of man-made P450 fusion enzymes have been successfully created using a variety of molecular approaches, including "Molecular Lego"[136], "LICRED"[167], and "PUPPET"[168]. Here, the versatility and superior properties of YkuN were exploited to generate redox fusion

enzymes capable of driving catalysis of different CYPs, using the previously established DuaLinX procedure [91] for linker engineering. The catalytic performance of a reconstituted P450 system is often limited by a low coupling efficiency. Uncoupling events waste expensive reduced cofactors (NAD(P)H) and lead to the generation of reactive oxygen species that can cause enzyme inactivation [169]. With non-physiological redox chains the coupling efficiency is often particularly poor (less than 20%) [92] and also when P450s catalyse reactions with non-physiological substrates, coupling efficiencies are frequently severely diminished (<10%)[169]. Achieving a high coupling efficiency is thus a particularly challenging task.

Direct attachment of YkuN to Fpr via genetic fusion led to a 1.8-fold improvement of the coupling efficiency in CYP109B1-catalysed reactions (table 6.1). Coupling efficiency could be even further improved by insertion of an appropriate linker between the fusion partners. In case of the YR-G5 construct a coupling efficiency as high as 81.2% was achieved. Notably, in all cases the CYP109B1 systems reconstituted with the different YkuN-Fpr fusion constructs exhibited a higher coupling efficiency than those reported previously using a variety of different redox partners (1.8-45.8%) [92], which is an advantage under cofactor regeneration condition. However, the NADPH oxidation rate was markedly reduced.

Despite improved coupling, overall performance of the parental linker-less YkuN-Fpr fusion construct in P450-catalysed reactions was rather modest in comparison to the non-fused enzymes (figs. 6.1 and 6.6). A >8-fold reduction in NADPH oxidation rate (table 6.1) combined with slower electron transfer to the CYP109B1 (table 6.2) likely contributed to the reduced performance of the YR fusion construct. Insertion of a suitable linker between the fusion partners however, substantially improved the overall performance. Herein, long $([E/L]PPPP)_n$ linkers (n = 4 - 5) were particularly effective (figs. 6.1 and 6.6), which is consistent with previous findings for fusions between *E. coli* FldA and Fpr [91]. At protein concentrations $\leq$4 μM, the superior construct YR-P5 acted as a nearly equivalent substitute for non-fused Fpr/YkuN in driving CYP109B1 catalysis (fig. 6.1). Insertion of the P5 linker between the fused YkuN and Fpr resulted in a 1.5-fold increase in heme iron reduction rate (table 6.2, k_2) along with a 3.6-fold higher NADPH oxidation rate, while maintaining a high coupling efficiency of nearly 50% (table 6.1), thus contributing to the improved overall performance (fig. 6.1). However, despite poorer coupling efficiency, the non-fused redox partners proved generally more effective than YR-P5 in supporting P450 catalysis (table 6.3, figs. 6.1, 6.6 and 6.7), which is likely governed by faster overall electron transfer (accelerated NADPH oxidation as well as P450 reduction). The higher amount of progesterone overoxidation products formed by CYP106A2 when the non-fused redox partners are used as opposed to YR-P5, is also consistent with this notion (table 6.3).

Examination of the linker properties using MD simulations revealed that the $(GGGGS)_n$ linkers are highly flexible and tend to adopt compact random structures, whereas the $([E/L]PPPP)_n$ linkers maintain their linear conformation and are

therefore structurally more rigid (Fig. 6.4). These results are consistent with the notion that glycine-rich and proline-rich amino acid sequences are flexible and rigid in nature, respectively [170, 171, 172]. Moreover, the high mobility of glycine-rich linkers is apparent from their lack of structural resolution in X-ray structures of a multitude of artificial fusion proteins [173]. Due to their flexibility, glycine-rich linkers are unstructured and tend to provide limited domain separation [174, 175], whereas structurally rigid linkers, such as proline-rich linkers are more likely to separate the fusion partners [171, 176].

For flavodoxins the binding areas for the electron donating reductase and the terminal electron acceptor (P450) are partially overlapping, which precludes the formation of a ternary protein complex [177, 178]. Considering the fusion constructs, an optimal linker should therefore facilitate the formation of an electron transfer complex between Fpr and YkuN as well as between YkuN and the P450. In view of the superior performance of the YR-P5 construct these criteria are apparently best met by the $([E/L]PPPP)_5$ linker. Indeed, both the activity of Fpr (NADPH oxidation rate) and YkuN (heme-iron reduction) are reasonably well preserved in the YR-P5 construct (tables 6.1 and 6.2). Herein, the rigid P5 linker may restrict the degrees of freedom of the fused redox partners, while allowing them to mutually interact in an effective manner, thus promoting intramolecular electron transfer. On the other hand, the extended rigid P5 linker (25 residues) may increase the distance between the redox centres of the fusion partners, thereby disfavouring intramolecular electron transfer and facilitating electron transfer between fusion proteins (i.e. intermolecularly). The available experimental data for YR-P5 seems to be more consistent with a predominantly intermolecular electron transfer pathway. YR-P5 lacks a strict functional coupling between the Fpr and YkuN domain and its Fpr domain is readily accessible for external YkuN (fig. 6.4), which is likely facilitated by a higher degree of separation of the fusion partners in case of the P5 linker. Moreover, P450 mediated catalysis exhibited a higher order (sigmoidal) dependence on fusion enzyme concentration suggesting an intermolecular contribution, whereas a linear dependence would have been indicative of exclusive intramolecular electron transfer (fig. 6.1(b)). At elevated redox partner concentration, which increases the collision frequency between fusion proteins, apparent limitations in electron transfer were (partially) overcome and the performance of YR-P5 was similar to that of the non-fused Fpr/YkuN (fig. 6.1(b)).

Finally, the superior construct YR-P5 exhibits versatility as it effectively supported monooxygenase activity of functionally diverse P450s, including *B. subtilis* CYP109B1 (fig. 6.1), *T. fusca YX* CYP154E1 (fig. 6.6), *B. megaterium* CYP106A2 (table 6.3) and bovine CYP21A2 (fig. 6.7). Thus, functional promiscuity of YR-P5 is not limited to bacterial P450s.

Considering the feasibility of the Fpr/YkuN system and its fused derivative YR-P5 for applied biocatalysis, the TTN presented here are rather modest. In case of CYP154E1 with β-ionone as substrate TTN were 1,582 with YR-P5 and 1,948

with the non-fused Fpr/YkuN. On the other hand, we recently observed that a Fpr/YkuN/CYP154E1 system can achieve TTN of up to 20,000 with stilbene as substrate [179]. Although TTN in P450-based multi-component systems depend on many factors, the P450-substrate match seems to play a prominent role. For example, CYP154C5 catalysis supported by PdR/Pdx achieved TTN of 2,440 with progesterone and 3,341 with androstenedione under optimized conditions [180]. For the natural fusion P450 BM3 from *Bacillus megaterium* (and mutants thereof), TTN values range from 890 with omeprazole [181], 2,200 with n-octane [182], to 6,195 with anisole [183], and even TTN as high as 24,363 or 45,000 have been reported in case of cyclooctane [184] and propane [185], respectively.

Taken together, the intrinsic dependence on redox partner(s) represents an important limiting factor in the utilization of P450s as biocatalysts. Availability of suitable redox partner(s) is therefore a prerequisite to successfully explore the catalytic potential of P450s [186]. By covalent fusion the complexity of the bacterial Fpr/YkuN electron transfer chain was effectively reduced, allowing easy protein production, purification and handling. Moreover, a stabilising effect on Fpr was noted upon fusion to YkuN. Through linker engineering the activity of the fusion construct could be tuned such that the high activity of the individual redox partners and the promiscuity for different P450s was largely preserved, with overall best performance of the YR-P5 construct. The versatile YR-P5 may serve as an effective surrogate electron transfer system for exploitation of the catalytic potential of (orphan) P450s.

Chapter 7

Conclusion

During my work I contributed ideas and theoretical calculations that provide ground for further studies in the biotechnological application of redox systems depending on *adrenodoxin* and *Electron transfer protein 1*, *CYP106A2* together with *adrenodoxin reductase*.
In a first collaboration with Prof. Bernhardts lab we took part in engineering CYP106A2 towards a 9-α- and 6-β-hydroxylation. Our suggestions for CYP106A2 mutations towards 6-β-hydroxylations were exchanges at positions 247 (change from threonine to alanine), 243 (alanine to serine) and 173 (phenylalanine to alanine). Combining those replacements lead to seven new proteins. The exchange of alanine to serine at position 243 increased the production of 6-β-hydroxylation of progesterone by more than ten-fold which was due to an additional hydrogen-bond to the carbonyl oxygen in position 21 that could be formed by the substrate progesterone. Subsequent MD simulations could explain this observation. The distances to the 6-β position to the central heme iron were shorter than those to the 15-β throughout the simulation.
In another collaboration we were looking for an optimal fitting redox system for CYP106A2. Here the combination of ferredoxins Adx, Etp1^{fd}, Fdx2, FldA and YkuN with reductases Arh1, BmCPR and Fpr resulting in 11 different complexes was established (chapter 3). Again progesterone was used since it is a well studied substrate of CYP106A2. The preferred redox system is Adx together with AdR. The inspected redox systems all were able to deliver electrons from NAD(P)H to CYP106A2, which was determined experimentally in Prof. Bernhardts lab. The differences appeared in variable overall turnover of progesterone (e.g. Etp1(516-618)/AdR had 100% turnover, Etp1(516-618)/BmCPR had 79%) as well as in product variability including mono- and polyhydroxylated progesterone. When using Adx(4-108) together with AdR after 30 minutes of conversion, 65% of progesterone is polyhydroxylated, 27% is 15β-OH, 5% are other monohydroxylated products and the remaining part is unhydroxylated substrate. Exchanging Adx(4-108) with Etp1(516-618) shifted the outcome to 18% poly-OH, 78% 15β-OH, 3% other

monohydroxylated progesterones, and no remaining substrate. Similar results were seen when investigating Adx(1-128) and AdR or Fdx2 together with Arh1. From these redox systems, Etp1 and Adx were investigated using theoretical approaches. Protein-protein docking of Etp1(516-618) and Adx(4-108) to the binding interface of CYP106A2 revealed the possible reason for this product spectrum. Phe592 (which is at the corresponding position of Tyr82 in Adx) of Etp1(516-618) binds to Phe107, Pro359, and Leu356 of CYP106A2 in a sandwich like manner. This allows the iron-sulfur cluster to approach the heme center of CYP106A2 as close as 19 Å in its closest conformation found. Subsequent electron transfer should therefore be faster as for Adx(4-108). Here the distance of iron-sulfur cluster and heme center is 24 Å. Also position Tyr82 of Adx(4-108) cannot interact with Phe107, Pro359, and Leu356 of the CYP1062 because it is sterically hindered by Arg14 and Asp41. But due to larger electrostatic contributions a slower release of Adx(4-108) is conceivable. Those findings suggest that Adx has a slower electron transfer rate than Etp(516-618) and its interactions are stronger so that the substrate dwells longer in the active site of the CYP106A2 ending in more polyhydroxylated products.
In chapter 6 the focus lied on the establishment of fused redox systems delivering electrons to different CYP P450s, *B. subtilis* CYP109B1, *Thermobifida fusca* CYP154E1, *Bacillus megaterium* CYP106A2 and bovine microsomal CYP21A2, respectively. The redox systems consisted of *E. coli*'s Flavodoxin reductase Fpr and Flavodoxin YkuN of *B. subtilis* (complex to be named YR). They have been fused together with flexible glycine $(GGGGS)_n$ and stiffer proline linkers $([E/L]PPPP)_n$ of different lengths (with n = 1 - 5) in Prof. Urlachers lab. Fusing the electron transfer proteins should lower the complexity of the redox chain as well as improve the catalytic properties. The best result was seen for the YR-P5 (proline linker with length of n = 5). Here a 1.7 fold increase of coupling efficiency could be seen compared to the separated proteins (49% vs. 28%). Nevertheless the oxidation rate of NADPH as well as the CYP reduction rate decreased by 2.3 and 1.7, respectively. Theoretical investigation of the linker constructs through MD simulation did provide proof why in general the proline-rich linkers outperform the glycine-rich ones. We reconstructed each linker in an elongated conformation with the Hyperchem program and performed MD simulations for 1 ns with prior energy minimizations as well as equilibration. We were able to show that the glycine-rich linkers experienced a hydrophobic collapse after a short period of time. The proline-rich linker stayed stable throughout the simulation and showed only slight bending. The space provided by the proline-rich linkers allowed the required accessibility of the actual redox centers, which obviously limits the local turn-over.
Chapter 5 formulates our attempts to systematically investigate amino acid exchanges or deletions in bovine Adx regarding the change of its redox potential. First we analysed the effect of deletion of threonine at position 49. This dramatically distorts the arrangement of its iron-sulfur cluster since it has a direct effect on cysteine at position 46, which ligates one of the iron atoms. Our computations show that the sterical effect of this deletion leads to a contribution of at least 50mV to the lower redox potential due to electronic effects on the iron-sulfur cluster. Therefore,

we are able to estimate this contribution from that of the remaining protein. For our research aiming at a systemic evaluation of amino acid exchange influences on the redox potential of Adx we came across a publication of Zhou [56]. They used a simplified sphere model for protein and introduced the idea of setting dipole directions of amino acids in relation to the redox potential. This approximation is valid as long as the contributions from the iron-sulfur cluster remain largely unchanged upon mutation of the protein. The influence on the redox potential is most distinct when a buried amino acid in close distance to the iron-sulfur cluster is changed. Also the direction of its electric dipole plays a role. An increase of the redox potential should be seen if the dipole vector points towards the iron-sulfur cluster, whereas a decrease is expected if it points away. Perpendicular dipole vectors should only show marginal effects. Based on this we suggested two mutations at position 43 with opposite effects that still awaits experimental validation at present.
Before that I tried an attempt to correlate experimentally derived redox potential with computed ones. The goal was (like in chapter 5) to implement a systematic evaluation that could estimate the influence of exchanging amino acids in bovine Adx(4-108) WT molecules on the reduction potential. The influence could be measured in deviation from a normalized computed redox potential in relation to the real experimental values. Lowering the redox potential would improve the performance of Adx in terms of e.g. electron transfer rate. Experimentally derived values for reduction potentials were collected from the excellent review from Ewen et al. in 2011 [113]. Table 7.1 lists the mutations we took to computationally derive redox potentials. Here the terms in brackets account for length of molecule and positions of mutations (e.g. Adx(1-128/D76E) denotes full length Adx with an amino acid exchange at position 76 where D (aspartic acid) is exchanged by E (glutamic acid)). Starting from the crystal structure of wild type bovine Adx (pdb entry 1AYF) we

Adx(1-112)	Adx(1-128/H56T)	Adx(1-128/T49Y)	Adx(4-108)
Adx(1-112/S112W)	Adx(1-128/K6E)	Adx(1-128/T54A)	Adx(4-108/R14A)
Adx(1-112/Y82F/S112W)	Adx(1-128/K6E/K22Q/K24Q)	Adx(1-128/T54S)	Adx(4-108/R14E)
Adx(1-112/Y82L/S112W)	Adx(1-128/K6E/K22Q/K24Q/K98E)	Adx(1-128/T71E)	Adx(4-108A)
Adx(1-112/Y82S/S112W)	Adx(1-128/K6E/K98E)	Adx(1-128/T71V)	Adx(4-108K)
Adx(1-128)	Adx(1-128/K98E)	Adx(1-128/Y82F)	Adx(4-108S)
Adx(1-128/D76E)	Adx(1-128/T49A)	Adx(1-128/Y82L)	Adx(4-108W)
Adx(1-128/H56Q)	Adx(1-128/T49L)	Adx(1-128/Y82S)	Adx(4-114)
Adx(1-128/H56R)	Adx(1-128/T49S)	Adx(4-107)	Adx(4-128)

Table 7.1: Overview of the different modeled Adx mutations. Terms in brackets account for length of molecule and positions of mutations (e.g. Adx(1-128/D76E) denotes full length Adx with an amino acid exchange at position 76 where D (aspartic acid) is exchanged by E (glutamic acid)).

built homology models with the SWISS-MODEL [119, 187, 120, 188] homology modeling server according to the mutations listed in table 7.1. To relax tensions and get

rid of other sterical discrepancies we simulated each of the Adx mutant in a water bath at constant temperature and pressure for about 1 ns in TIP3P water [189] (relaxation). Charges for the iron-sulfur clusters were taken from a prior publication from Shakya et al [190]. The subsequent computation of the redox potential with the APBS program, however, did not produce reasonable values for further use. Attempts to change the order of relaxation and building of homology models did not have an influence on the quality of the results. The two approaches with changing order were as follows. First relax the wild type, mutate it and do an energy minimization and calculate the redox potential. Second, mutate the wild type and relax it afterwards. The failure may be due to the lack of crystal structures for the different mutants. Also the missing information on the conformational change of the [Fe2S2] center upon reduction is probably a big issue. Poisson-Boltzmann theory requires the accurate three-dimensional structure representation of both the oxidized as well as the reduced Adx in solution, along with appropriate assignment of atomic charges and protonation states for each amino acid [111, 115]. The problem with measuring the effect of conformational changes needs to be computed by quantum mechanical methods like DFT, which are very time demanding and inappropriate for the size as well as the amount of molecules. A detailed analysis of the set of Adx mutants (table 7.1) revealed that the redox potential strongly correlates with the accessible surface area. This means that exchanging a big amino acid with a small one should change the redox potential massively. But only on certain positions in the Adx which then are the important contact interfaces (e.g.Y82S, position 82 plays an important role). Since there was seen a big decrease in redox potential for the exchange of proline to lysine at position 108 another exchange was testet besides the ones at position 82. The proposed candidates for new mutations for lowering the redox potential are given in table 7.2. The terms in brackets account for length of the molecule and positions of mutations (e.g. Adx(4-108/P108R) denotes a trimmed version of Adx (from amino acid at position 4 to 108) with an amino acid exchange at position 108 where P (proline) is exchanged by R (Arginine)). The redox potentials are given in mV together with the standard deviation, which is rather high. Exchange of tyrosine at position 82 with serine should lower the redox potential due to changes in terms of size. Exchange with asparagine could lead into an internal salt bridge with the proposed arginine at position 108. Both changes at position 82 were conjugated with changes at position 108 to arginine and lysine. However, results show that the changes only have minor influences and our propositions did not have the desired effects. In addition to that, the constructs of these mutants in the wetlab were rather unstable. proline at position 108 is almost indispensable as stabilizing residue.

This work dealt with the optimization of protein-protein and protein-ligand interactions trying to add an other piece to the puzzle of decrypting the way for the biotechnological application of CYP P450 redox systems. For the case of Adx my focus shifted more and more in the direction of fundamental research. Meaning to investigate why and how the redox potential changes upon introducing other amino acids into the structure. Two attempts were undertaken that failed to explain this.

Mutation	Avg redox potential[mV]	SD
Adx(4-108)	-324.62	2.45
Adx(4-108/P108R)	-319.03	21.83
Adx(4-108/Y82S/P108R)	-314.36	2.63
Adx(4-108/Y82D/P108R)	-372.99	27.24
Adx(4-108/Y82S/P108K)	-336.44	12.36
Adx(4-108/Y82D/P108K)	-341.88	24.04

Table 7.2: Average redox potentials from newly proposed Adx mutations. In relation to the shorter wild type Adx4-108 all others are not noteworthy better or worse than others from [113]. Also the standard deviation is very high. Terms in brackets account for length of molecule and positions of mutations (e.g. Adx(4-108/P108R) denotes a trimmed version of Adx (from amino acid at position 4 to 108) with an amino acid exchange at position 108 where P (proline) is exchanged by R (arginine)).

The validation of another theory that could explain this is still ongoing. Theoretical groundwork has been done and suffices experimental validation (see chapter 5). Building up on this one could be able to predict results of exchanges of amino acids. Furthermore this could be extended to other ferre- or flavodoxins affecting other CYP P450 redox systems. Besides the focus on Adx and the improvement of interaction to CYP P450s a proof for efficiency could be delivered for a fusion of a flavodoxin and a reductase (chapter 6). The kind of linker construct was important for keeping the two proteins separate over a certain distance that allowed for regenerating the system for following oxidations and reductions of the heme centers and electron donors, NAD(P)H respectively. Since there were no crystal structures of fused systems from nature available, a complete picture about their behaviour in solution (with e.g. MD simulations) cannot be delivered yet. However, this concept also could be applied to Adx or Etp1 with a given reductase for further tuning the electron transport efficiency. Protein engineering of CYP106A2 lead to a substantial success in steering the synthesis of progesterone to alternatively hydroxylated progesterone products, that are important for diverse pharmaceutical applications. In particular, using docking of substrates into the available crystal structure and likewise into homology models of rationally designed mutants, proved to be a valuable tool to complement experimental work. Furthermore, molecular dynamics simulations revealed that reorientation of primary hydroxylation products inside the binding pocket occurs on a fast time scale. This prompted further experimental studies regarding the electron transport towards the cytochrome, which are still ongoing.

List of Figures

List of Tables

Bibliography

[1] H. S. Mason, W. L. Fowlks, and E. Peterson. Oxygen transfer and electron transport by the phenolase complex1. *Journal of the American Chemical Society*, 77(10):2914–2915, 1955.

[2] Osamu Hayaishi, Masayuki Katagiri, and Simon Rothberg. Mechanism of the pyrocatechase reaction. *Journal of the American Chemical Society*, 77(20):5450–5451, 1955.

[3] M. KLINGENBERG. Pigments of rat liver microsomes. *Arch. Biochem. Biophys.*, 75(2):376–386, Jun 1958.

[4] D. GARFINKEL. Studies on pig liver microsomes. I. Enzymic and pigment composition of different microsomal fractions. *Arch. Biochem. Biophys.*, 77(2):493–509, Oct 1958.

[5] T. OMURA and R. SATO. THE CARBON MONOXIDE-BINDING PIGMENT OF LIVER MICROSOMES. I. EVIDENCE FOR ITS HEMOPROTEIN NATURE. *J. Biol. Chem.*, 239:2370–2378, Jul 1964.

[6] R. W. ESTABROOK, D. Y. COOPER, and O. ROSENTHAL. THE LIGHT REVERSIBLE CARBON MONOXIDE INHIBITION OF THE STEROID C21-HYDROXYLASE SYSTEM OF THE ADRENAL CORTEX. *Biochem Z*, 338:741–755, 1963.

[7] I. C. Gunsalus, B. N. Ganguli, M. Katagiri, J. C. Tsibris, P. Debrunner, and H. Frauenfelder. Oxygenation: a specific soluble cytochrome p-450 coupled enzyme complex. *Science*, 160(3826):438–439, Apr 1968.

[8] D. W. Nebert and F. J. Gonzalez. P450 genes: structure, evolution, and regulation. *Annu. Rev. Biochem.*, 56:945–993, 1987.

[9] D. R. Nelson. The cytochrome p450 homepage. *Hum. Genomics*, 4(1):59–65, Oct 2009.

[10] D. Werck-Reichhart and R. Feyereisen. Cytochromes P450: a success story. *Genome Biol.*, 1(6):REVIEWS3003, 2000.

[11] Rita Bernhardt and Vlada B Urlacher. Cytochromes p450 as promising catalysts for biotechnological application: chances and limitations. *Applied microbiology and biotechnology*, 98(14):6185–6203, 2014.

[12] F. Hannemann, A. Bichet, K. M. Ewen, and R. Bernhardt. Cytochrome P450 systems–biological variations of electron transport chains. *Biochim. Biophys. Acta*, 1770(3):330–344, Mar 2007.

[13] Paul R. Ortiz de Montellano (eds.). *Cytochrome P450: Structure, Mechanism, and Biochemistry*. Springer International Publishing, 4 edition, 2015.

[14] Linda Owers Narhi and Armand J Fulco. Characterization of a catalytically self-sufficient 119,000-dalton cytochrome p-450 monooxygenase induced by barbiturates in bacillus megaterium. *Journal of Biological Chemistry*, 261(16):7160–7169, 1986.

[15] Linda O Narhi and Armand J Fulco. Identification and characterization of two functional domains in cytochrome p-450bm-3, a catalytically self-sufficient monooxygenase induced by barbiturates in bacillus megaterium. *Journal of Biological Chemistry*, 262(14):6683–6690, 1987.

[16] J. Meyer. Ferredoxins of the third kind. *FEBS Lett.*, 509(1):1–5, Nov 2001.

[17] WJ McAleer, TA Jacob, LB Turnbull, EF Schoenewaldt, and TH Stoudt. Hydroxylation of progesterone by bacillus cereus and bacillus megaterium. *Archives of biochemistry and biophysics*, 73(1):127–130, 1958.

[18] Anders Berg, Kjell Carlström, Jan-Åke Gustafsson, and Magnus Ingelman-Sundberg. Demonstration of a cytochrome p-450-dependent steroid 15β-hydroxylase in bacillus megaterium. *Biochemical and biophysical research communications*, 66(4):1414–1423, 1975.

[19] ANDERS Berg, Jan-Ake Gustafsson, and MAGNUS Ingelman-Sundberg. Characterization of a cytochrome p-450-dependent steroid hydroxylase system present in bacillus megaterium. *Journal of Biological Chemistry*, 251(9):2831–2838, 1976.

[20] Anders Berg, Magnus Ingelman-Sundberg, and Jan-&e Gustafsson. Purification and characterization of cytochrome p-450. *J. Biol. Chem*, 254:5264–5271, 1979.

[21] A Berg, M Ingelman-Sundberg, and Jan-Ake Gustafsson. Isolation and characterization of cytochrome p-450meg. *Acta biologica et medica Germanica*, 38(2-3):333–344, 1978.

[22] A Berg and Joseph J Rafter. Studies on the substrate specificity and inducibility of cytochrome p-450meg. *Biochemical Journal*, 196(3):781–786, 1981.

[23] Anders Berg. Characterization of the ferredoxin component of the steroid 15β-hydroxylase system from bacillus megaterium. *Biochemical and biophysical research communications*, 105(1):303–311, 1982.

[24] Birgit Simgen, Jörg Contzen, Rolf Schwarzer, Rita Bernhardt, and Christiane Jung. Substrate binding to 15β-hydroxylase (cyp106a2) probed by ft infrared spectroscopic studies of the iron ligand co stretch vibration. *Biochemical and biophysical research communications*, 269(3):737–742, 2000.

[25] Reimund Rauschenbach, Marina Isernhagen, Christiane Noeske-Jungblut, Werner Boidol, and Gerhard Siewert. Cloning sequencing and expression of the gene for cytochrome p450meg, the steroid-15β-monooxygenase from bacillus megaterium atcc 13368. *Molecular and General Genetics MGG*, 241(1):170–176, 1993.

[26] Hitosi Agematu, Naoki Matsumoto, Yoshikazu Fujii, Hiroki Kabumoto, Satoru Doi, Kazuhiro Machida, Jun Ishikawa, and Akira Arisawa. Hydroxylation of testosterone by bacterial cytochromes p450 using the escherichia coli expression system. *Bioscience, biotechnology, and biochemistry*, 70(1):307–311, 2006.

[27] Sabrina Bleif, Frank Hannemann, Michael Lisurek, Jens Peter von Kries, Josef Zapp, Matthias Dietzen, Iris Antes, and Rita Bernhardt. Identification of cyp106a2 as a regioselective allylic bacterial diterpene hydroxylase. *ChemBioChem*, 12(4):576–582, 2011.

[28] Sabrina Bleif, Frank Hannemann, Josef Zapp, David Hartmann, Johann Jauch, and Rita Bernhardt. A new bacillus megaterium whole-cell catalyst for the hydroxylation of the pentacyclic triterpene 11-keto-β-boswellic acid (kba) based on a recombinant cytochrome p450 system. *Applied microbiology and biotechnology*, 93(3):1135–1146, 2012.

[29] Daniela Schmitz, Josef Zapp, and Rita Bernhardt. Hydroxylation of the triterpenoid dipterocarpol with cyp106a2 from bacillus megaterium. *The FEBS journal*, 279(9):1663–1674, 2012.

[30] Simon Janocha, Yvonne Carius, Michael Hutter, C Roy D Lancaster, and Rita Bernhardt. Crystal structure of cyp106a2 in substrate-free and substrate-bound form. *ChemBioChem*, 17(9):852–860, 2016.

[31] K. SUZUKI and T. KIMURA. AN IRON PROTEIN AS A COMPONENT OF STEROID 11-BETA-HYDROXYLASE COMPLEX. *Biochem. Biophys. Res. Commun.*, 19:340–345, Apr 1965.

[32] T. Omura, E. Sanders, R.W. Estabrook, D.Y. Cooper, and O. Rosenthal. Isolation from adrenal cortex of a nonheme iron protein and a flavoprotein functional as a reduced triphosphopyridine nucleotide-cytochrome p-450 reductase. *Archives of Biochemistry and Biophysics*, 117(3):660 – 673, 1966.

[33] T. Kimura and K. Suzuki. Components of the electron transport system in adrenal steroid hydroxylase. Isolation and properties of non-heme iron protein (adrenodoxin). *J. Biol. Chem.*, 242(3):485–491, Feb 1967.

[34] Rita Bernhardt. The role of adrenodoxin in adrenal steroidogenesis. *Current Opinion in Endocrinology, Diabetes and Obesity*, 7(3):109–115, 2000.

[35] Rita Bernhardt and Michael R Waterman. Cytochrome p450 and steroid hormone biosynthesis. *The ubiquitous roles of cytochrome P450 proteins*, 3:361–396, 2007.

[36] W Miller. Steroidogenic enzymes. In *Disorders of the human Adrenal cortex*, volume 13, pages 1–18. Karger Publishers, 2008.

[37] Natallia Strushkevich, Farrell MacKenzie, Tatyana Cherkesova, Irina Grabovec, Sergey Usanov, and Hee-Won Park. Structural basis for pregnenolone biosynthesis by the mitochondrial monooxygenase system. *Proceedings of the National Academy of Sciences*, 108(25):10139–10143, 2011.

[38] Dirk Beilke, Roland Weiss, Frank Löhr, Primož Pristovšek, Frank Hannemann, Rita Bernhardt, and Heinz Rüterjans. A new electron transport mechanism in mitochondrial steroid hydroxylase systems based on structural changes upon the reduction of adrenodoxin. *Biochemistry*, 41(25):7969–7978, 2002.

[39] Matthias Bureik, Burkhard Schiffler, Yasushi Hiraoka, Frank Vogel, and Rita Bernhardt. Functional expression of human mitochondrial cyp11b2 in fission yeast and identification of a new internal electron transfer protein, etp1. *Biochemistry*, 41(7):2311–2321, 2002.

[40] Jürgen J Müller, Frank Hannemann, Burkhard Schiffler, Kerstin M Ewen, Reinhard Kappl, Udo Heinemann, and Rita Bernhardt. Structural and thermodynamic characterization of the adrenodoxin-like domain of the electron-transfer protein etp1 from schizosaccharomyces pombe. *Journal of inorganic biochemistry*, 105(7):957–965, 2011.

[41] Burkhard Schiffler and Rita Bernhardt. Bacterial (cyp101) and mitochondrial p450 systems - how comparable are they? *Biochemical and biophysical research communications*, 312(1):223–228, 2003.

[42] Heike Uhlmann and Rita Bernhardt. The role of threonine 54 in adrenodoxin for the properties of its iron-sulfur cluster and its electron transfer function. *Journal of Biological Chemistry*, 270(50):29959–29966, 1995.

[43] Asya V Grinberg, Frank Hannemann, Burkhard Schiffler, Jürgen Müller, Udo Heinemann, and Rita Bernhardt. Adrenodoxin: structure, stability, and electron transfer properties. *Proteins: Structure, Function, and Bioinformatics*, 40(4):590–612, 2000.

[44] Burkhard Schiffler, Martin Kiefer, Andreas Wilken, Frank Hannemann, Hans Werner Adolph, and Rita Bernhardt. The interaction of bovine adrenodoxin with cyp11a1 (cytochrome p450scc) and cyp11b1 (cytochrome p45011β) acceleration of reduction and substrate conversion by site-directed mutagenesis of adrenodoxin. *Journal of Biological Chemistry*, 276(39):36225–36232, 2001.

[45] Heike Uhlmann, Stefania Iametti, Giuseppe Vecchio, Francesco Bonomi, and Rita Bernhardt. Pro108 is important for folding and stabilization of adrenal ferredoxin, but does not influence the functional properties of the protein. *The FEBS Journal*, 248(3):897–902, 1997.

[46] Burkhard Schiffler, Matthias Bureik, Wolfgang Reinle, Eva-Christina Müller, Frank Hannemann, and Rita Bernhardt. The adrenodoxin-like ferredoxin of schizosaccharomyces pombe mitochondria. *Journal of inorganic biochemistry*, 98(7):1229–1237, 2004.

[47] Hans-Emil KNOELL and Joachim KNAPPE. Escherichia coli ferredoxin, an iron-sulfur protein of the adrenodoxin type. *European journal of biochemistry*, 50(1):245–252, 1974.

[48] Germaine Sainz, Jean Jakoncic, Larry C Sieker, Vivian Stojanoff, Nukri Sanishvili, Marcel Asso, Patrick Bertrand, Jean Armengaud, and Yves Jouanneau. Structure of a [2fe–2s] ferredoxin from rhodobacter capsulatus likely involved in fe–s cluster biogenesis and conformational changes observed upon reduction. *JBIC Journal of Biological Inorganic Chemistry*, 11(2):235–246, 2006.

[49] Vytas Reipa, Marcia J Holden, Martin P Mayhew, and Vincent L Vilker. Temperature dependence of the formal reduction potential of putidaredoxin. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 1459(1):1–9, 2000.

[50] Marina V Donova and Olga V Egorova. Microbial steroid transformations: current state and prospects. *Applied microbiology and biotechnology*, 94(6):1423–1447, 2012.

[51] P Fernandes, Aurelio Cruz, Blaga Angelova, HM Pinheiro, and JMS Cabral. Microbial conversion of steroid compounds: recent developments. *Enzyme and microbial technology*, 32(6):688–705, 2003.

[52] Lina Maria Klara Renate Schiffer. *Human steroidogenic cytochromes P450: biotransformation of drugs and biotechnological application.* PhD thesis, 2016.

[53] Matthias Bureik and Rita Bernhardt. Steroid hydroxylation: microbial steroid biotransformations using cytochrome p450 enzymes. *Modern Biooxidation: Enzymes, Reactions, and Applications*, pages 155–176, 2007.

[54] Wang-Yu Tong and Xiang Dong. Microbial biotransformation: recent developments on steroid drugs. *Recent patents on biotechnology*, 3(2):141–153, 2009.

[55] Julia Nikolaus, Kim Thoa Nguyen, Cornelia Virus, Jan L Riehm, Michael Hutter, and Rita Bernhardt. Engineering of cyp106a2 for steroid 9α-and 6β-hydroxylation. *Steroids*, 120:41–48, 2017.

[56] Huan-Xiang Zhou. Control of reduction potential by protein matrix: lesson from a spherical protein model. *Journal of Biological Inorganic Chemistry*, 2(1):109–113, 1997.

[57] Nathan A Baker, David Sept, Simpson Joseph, Michael J Holst, and J Andrew McCammon. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences*, 98(18):10037–10041, 2001.

[58] Loup Verlet. Computer" experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Physical review*, 159(1):98, 1967.

[59] RW Hockney, SP Goel, and JW Eastwood. Quiet high-resolution computer models of a plasma. *Journal of Computational Physics*, 14(2):148–158, 1974.

[60] Helen M. Berman, John Westbrook, Zukang Feng, Gary Gilliland, T. N. Bhat, Helge Weissig, Ilya N. Shindyalov, and Philip E. Bourne. The protein data bank. *Nucleic Acids Research*, 28(1):235, 2000.

[61] Herman JC Berendsen, David van der Spoel, and Rudi van Drunen. Gromacs: a message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, 91(1-3):43–56, 1995.

[62] Wendy D Cornell, Piotr Cieplak, Christopher I Bayly, Ian R Gould, Kenneth M Merz, David M Ferguson, David C Spellmeyer, Thomas Fox, James W Caldwell, and Peter A Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules j. am. chem. soc. 1995, 117, 5179- 5197. *Journal of the American Chemical Society*, 118(9):2309–2309, 1996.

[63] Alexander D MacKerell Jr, Donald Bashford, MLDR Bellott, Roland Leslie Dunbrack Jr, Jeffrey D Evanseck, Martin J Field, Stefan Fischer, Jiali Gao, H Guo, Sookhee Ha, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The journal of physical chemistry B*, 102(18):3586–3616, 1998.

[64] Lukas D Schuler, Xavier Daura, and Wilfred F Van Gunsteren. An improved gromos96 force field for aliphatic hydrocarbons in the condensed phase. *Journal of Computational Chemistry*, 22(11):1205–1218, 2001.

[65] William L Jorgensen, David S Maxwell, and Julian Tirado-Rives. Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118(45):11225–11236, 1996.

[66] Herman JC Berendsen, JPM van Postma, Wilfred F van Gunsteren, ARHJ DiNola, and JR Haak. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690, 1984.

[67] Michele Parrinello and Aneesur Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12):7182–7190, 1981.

[68] Douglas B Kitchen, Hélène Decornez, John R Furr, and Jürgen Bajorath. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature reviews Drug discovery*, 3(11):935–949, 2004.

[69] R. Wang, L. Lai, and S. Wang. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput. Aided Mol. Des.*, 16(1):11–26, Jan 2002.

[70] Sergio Filipe Sousa, Pedro Alexandrino Fernandes, and Maria Joao Ramos. Protein–ligand docking: current status and future challenges. *Proteins: Structure, Function, and Bioinformatics*, 65(1):15–26, 2006.

[71] Irwin D Kuntz, Jeffrey M Blaney, Stuart J Oatley, Robert Langridge, and Thomas E Ferrin. A geometric approach to macromolecule-ligand interactions. *Journal of molecular biology*, 161(2):269–288, 1982.

[72] Garrett M Morris, Ruth Huey, William Lindstrom, Michel F Sanner, Richard K Belew, David S Goodsell, and Arthur J Olson. Autodock4 and autodocktools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry*, 30(16):2785–2791, 2009.

[73] Garrett M Morris, David S Goodsell, Robert S Halliday, Ruth Huey, William E Hart, Richard K Belew, Arthur J Olson, et al. Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function. *Journal of computational chemistry*, 19(14):1639–1662, 1998.

[74] Steven Demers. *Advanced density functional theory methods for materials science*. PhD thesis, Citeseer, 2013.

[75] David Sholl and Janice A Steckel. *Density functional theory: a practical introduction*. John Wiley & Sons, 2011.

[76] Max Born and Robert Oppenheimer. Zur quantentheorie der molekeln. *Annalen der Physik*, 389(20):457–484, 1927.

[77] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Physical review*, 136(3B):B864, 1964.

[78] Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965.

[79] Todd J. Dolinsky, Paul Czodrowski, Hui Li, Jens E. Nielsen, Jan H. Jensen, Gerhard Klebe, and Nathan A. Baker. Pdb2pqr: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Research*, 35(suppl_2):W522, 2007.

[80] Todd J. Dolinsky, Jens E. Nielsen, J. Andrew McCammon, and Nathan A. Baker. Pdb2pqr: an automated pipeline for the setup of poisson boltzmann electrostatics calculations. *Nucleic Acids Research*, 32(suppl_2):W665, 2004.

[81] T. Sagadin, J.L. Riehm, M.C. Hutter, and R. Bernhardt. Towards an optimal redox chain for cyp106a2 from bacillus megaterium atcc 13368. submitted.

[82] Michael Lisurek, Min-Jung Kang, Rolf W Hartmann, and Rita Bernhardt. Identification of monohydroxy progesterones produced by cyp106a2 using comparative hplc and electrospray ionisation collision-induced dissociation mass spectrometry. *Biochemical and biophysical research communications*, 319(2):677–682, 2004.

[83] Min-Jung Kang, Michael Lisurek, Rita Bernhardt, and Rolf W Hartmann. Use of high-performance liquid chromatography/electrospray ionization collision-induced dissociation mass spectrometry for structural identification of monohydroxylated progesterones. *Rapid communications in mass spectrometry*, 18(23):2795–2800, 2004.

[84] Kerstin M Ewen, Burkhard Schiffler, Heike Uhlmann-Schiffler, Rita Bernhardt, and Frank Hannemann. The endogenous adrenodoxin reductase-like flavoprotein arh1 supports heterologous cytochrome p450-dependent substrate conversions in schizosaccharomyces pombe. *FEMS yeast research*, 8(3):432–441, 2008.

[85] Matthias Bureik, Katja Hübel, Calin-Aurel Dragan, Jochen Scher, Hans Becker, Natalie Lenz, and Rita Bernhardt. Development of test systems for the discovery of selective human aldosterone synthase (cyp11b2) and 11β-hydroxylase (cyp11b1) inhibitors.: Discovery of a new lead compound for the therapy of congestive heart failure, myocardial fibrosis and hypertension. *Molecular and cellular endocrinology*, 217(1):249–254, 2004.

[86] Michael Kleser, Frank Hannemann, Michael Hutter, Josef Zapp, and Rita Bernhardt. Cyp105a1 mediated 3-hydroxylation of glimepiride and glibenclamide using a recombinant bacillus megaterium whole-cell catalyst. *Journal of biotechnology*, 157(3):405–412, 2012.

[87] Anna Hobler, Norio Kagawa, Michael C Hutter, Michaela F Hartmann, Stefan A Wudy, Frank Hannemann, and Rita Bernhardt. Human aldosterone synthase: recombinant expression in e. coli and purification enables a detailed biochemical analysis of the protein on the molecular level. *The Journal of steroid biochemistry and molecular biology*, 132(1):57–65, 2012.

[88] Elisa Brill, Frank Hannemann, Josef Zapp, Gerit Brüning, Johann Jauch, and Rita Bernhardt. A new cytochrome p450 system from bacillus megaterium dsm319 for the hydroxylation of 11-keto-β-boswellic acid (kba). *Applied microbiology and biotechnology*, 98(4):1703–1717, 2014.

[89] Rachel J Lawson, Claes von Wachenfeldt, Ihtshamul Haq, John Perkins, and Andrew W Munro. Expression and characterization of the two flavodoxin proteins of bacillus subtilis, ykun and ykup: biophysical properties and interactions with cytochrome p450 bioi. *Biochemistry*, 43(39):12390–12409, 2004.

[90] Mohammed Milhim, Adrian Gerber, Jens Neunzig, Frank Hannemann, and Rita Bernhardt. A novel nadph-dependent flavoprotein reductase from bacillus megaterium acts as an efficient cytochrome p450 reductase. *Journal of biotechnology*, 231:83–94, 2016.

[91] Patrick J Bakkes, Stefan Biemann, Ansgar Bokel, Marc Eickholt, Marco Girhard, and Vlada B Urlacher. Design and improvement of artificial redox modules by molecular fusion of flavodoxin and flavodoxin reductase from escherichia coli. *Scientific reports*, 5:12158, 2015.

[92] Marco Girhard, Tobias Klaus, Yogan Khatri, Rita Bernhardt, and Vlada B Urlacher. Characterization of the versatile monooxygenase cyp109b1 from bacillus subtilis. *Applied microbiology and biotechnology*, 87(2):595–607, 2010.

[93] Michael Ringle, Yogan Khatri, Josef Zapp, Frank Hannemann, and Rita Bernhardt. Application of a new versatile electron transfer system for cytochrome p450-based escherichia coli whole-cell bioconversions. *Applied microbiology and biotechnology*, 97(17):7741–7754, 2013.

[94] Irina A Pikuleva, Carolyn Cao, and Michael R Waterman. An additional electrostatic interaction between adrenodoxin and p450c27 (cyp27a1) results in tighter binding than between adrenodoxin and p450scc (cyp11a1). *Journal of Biological Chemistry*, 274(4):2045–2052, 1999.

[95] Cornelia Virus. Etablierung und evaluierung eines screeningsystems für die molekulare evolution der steroid-15beta-hydroxylase (cyp106a2) aus bacillus megaterium atcc 13368. 2006.

[96] Alexander Müller, Jürgen J Müller, Yves A Muller, Heike Uhlmann, Rita Bernhardt, and Udo Heinemann. New aspects of electron transfer revealed by the crystal structure of a truncated bovine adrenodoxin, adx (4–108). *Structure*, 6(3):269–280, 1998.

[97] Michael Lisurek, Birgit Simgen, Iris Antes, and Rita Bernhardt. Theoretical and experimental evaluation of a cyp106a2 low homology model and production of mutants with changed activity and selectivity of hydroxylation. *Chembiochem*, 9(9):1439–1449, 2008.

[98] Kim Thoa Nguyen, Cornelia Virus, Nils Günnewich, Frank Hannemann, and Rita Bernhardt. Changing the regioselectivity of a p450 from c15 to c11 hydroxylation of progesterone. *ChemBioChem*, 13(8):1161–1166, 2012.

[99] J Tamm, M Seckelmann, U Volkwein, and E Ludwig. The effect of the antiandrogen iiα-hydroxyprogesterone on sebum production and cholesterol concentration of sebum. *British Journal of Dermatology*, 107(1):63–70, 1982.

[100] AH Van der Willigen, JD Peereboom-Wynia, Th Van Joost, and E Stolz. A preliminary study of the effect of 11a-hydroxyprogesterone on the hair growth in men suffering from androgenetic alopecia. *Acta dermato-venereologica*, 67(1):82–85, 1986.

[101] HIROYUKI Morita, MINGYI Zhou, MARK F Foecking, ELISE P Gomez-Sanchez, EDUARDO N Cozza, and CELSO E Gomez-Sanchez. 11 beta-hydroxysteroid dehydrogenase type 2 complementary deoxyribonucleic acid stably transfected into chinese hamster ovary cells: specific inhibition by 11 alpha-hydroxyprogesterone. *Endocrinology*, 137(6):2308–2314, 1996.

[102] Graham W Souness and David J Morris. 11α-and 11β-hydroxyprogesterone, potent inhibitors of 11β-hydroxysteroid dehydrogenase, possess hypertensinogenic activity in the rat. *Hypertension*, 27(3):421–425, 1996.

[103] JA Hogg, PF Beal, AH Nathan, FH Lincoln, WP Schneider, BJ Magerlein, AR Hanze, and RW Jackson. The adrenal hormones and related compounds. i. a direct synthesis of hydrocortisone acetate and cortisone acetate from 11α-hydroxyprogesterone. *Journal of the American Chemical Society*, 77(16):4436–4438, 1955.

[104] Josef Fried, Richard W Thoma, John R Gerke, Josef E Herz, Milton N Donin, and D Perlman. Oxidation of steroids by microörganisms. ii. 1 hydroxylation in position 11 and synthesis of cortisone from reichstein's compound s. *Journal of the American Chemical Society*, 74(15):3962–3963, 1952.

[105] H Yoshioka, S Asada, and S Fujita. Process for production of 6β, 14α-hydroxyandrost-4-ene-3, 17-dione. *European Patent*, 5(99,658), 1994.

[106] Akihiko Yamada, Morio Yamada, Yukihisa Fujita, Takashi Nishigami, Keiji Nakasho, and Kunio Uematsu. Self-augmentation effect of male-specific products on sexually differentiated progesterone metabolism in adult male rat liver microsomes. *Journal of Biological Chemistry*, 276(7):4604–4610, 2001.

[107] Perlman David, James D Dutcher, and Elwood O Titus. 9alpha-hydroxyprogesterone, June 24 1958. US Patent 2,840,578.

[108] Cornelia Virus and Rita Bernhardt. Molecular evolution of a steroid hydroxylating cytochrome p450 using a versatile steroid detection system for screening. *Lipids*, 43(12):1133–1141, 2008.

[109] Reimund Rauschenbach. *Klonierung, Expression und Mutagenese des Gens für die Steroid-15 [beta]-Hydroxylase aus Bacillus megaterium ATCC 13368.* 1993.

[110] Flora Marta Kiss, Daniela Schmitz, Josef Zapp, Tobias KF Dier, Dietrich A Volmer, and Rita Bernhardt. Comparison of cyp106a1 and cyp106a2 from bacillus megaterium–identification of a novel 11-oxidase activity. *Applied microbiology and biotechnology*, 99(20):8495–8514, 2015.

[111] Jean-Marie Mouesca, Jun L Chen, Louis Noodleman, Donald Bashford, David A Case, et al. Density-functional poisson-boltzmann calculations of redox potentials for iron-sulfur clusters. *Journal of the American Chemical Society*, 116(26):11898–11914, 1994.

[112] Mitsuo Shoji, Kenichi Koizumi, Takeshi Taniguchi, Yasutaka Kitagawa, Shusuke Yamanaka, Mitsutaka Okumura, and Kizashi Yamaguchi. Theory of chemical bonds in metalloenzymes iii: Full geometry optimization and vibration analysis of ferredoxin-type [2fe–2s] cluster. *International Journal of Quantum Chemistry*, 107(1):116–133, 2007.

[113] Kerstin Maria Ewen, Michael Kleser, and Rita Bernhardt. Adrenodoxin: The archetype of vertebrate-type [2fe2s] cluster ferredoxins. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1814(1):111 – 125, 2011. Cytochrome P450: Structure, biodiversity and potential for application.

[114] Andy Zöllner, Frank Hannemann, Michael Lisurek, and Rita Bernhardt. Deletions in the loop surrounding the iron–sulfur cluster of adrenodoxin severely affect the interactions with its native redox partners adrenodoxin reductase and cytochrome p450 scc (cyp11a1). *Journal of inorganic biochemistry*, 91(4):644–654, 2002.

[115] Matthias G Ullmann, Louis Noodleman, and David A Case. Density functional calculation of pk a values and redox potentials in the bovine rieske iron-sulfur protein. *Journal of Biological Inorganic Chemistry*, 7(6):632–639, 2002.

[116] Bradley Scott Perrin, Shuqiang Niu, and Toshiko Ichiye. Calculating standard reduction potentials of [4fe–4s] proteins. *Journal of computational chemistry*, 34(7):576–582, 2013.

[117] AK Churg and A Warshel. Control of the redox potential of cytochrome c and microscopic dielectric effects in proteins. *Biochemistry*, 25(7):1675–1681, 1986.

[118] Paul D Swartz, Brian W Beck, and Toshiko Ichiye. Structural origins of redox potentials in fe-s proteins: electrostatic potentials of crystal structures. *Biophysical journal*, 71(6):2958–2969, 1996.

[119] Marco Biasini, Stefan Bienert, Andrew Waterhouse, Konstantin Arnold, Gabriel Studer, Tobias Schmidt, Florian Kiefer, Tiziano Gallo Cassarino, Mar-

tino Bertoni, Lorenza Bordoli, and Torsten Schwede. Swiss-model: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*, 42(W1):W252, 2014.

[120] Konstantin Arnold, Lorenza Bordoli, JÃ¼rgen Kopp, and Torsten Schwede. The swiss-model workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, 22(2):195, 2006.

[121] Carlo Adamo and Vincenzo Barone. Toward reliable density functional methods without adjustable parameters: The pbe0 model. *The Journal of chemical physics*, 110(13):6158–6170, 1999.

[122] Florian Weigend and Reinhart Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for h to rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics*, 7(18):3297–3305, 2005.

[123] Marat Valiev, Eric J Bylaska, Niranjan Govind, Karol Kowalski, Tjerk P Straatsma, Hubertus JJ Van Dam, Dunyou Wang, Jarek Nieplocha, Edoardo Apra, Theresa L Windus, et al. Nwchem: a comprehensive and scalable open-source solution for large scale molecular simulations. *Computer Physics Communications*, 181(9):1477–1489, 2010.

[124] Louis Noodleman. Valence bond description of antiferromagnetic coupling in transition metal dimers. *The Journal of Chemical Physics*, 74(10):5737–5743, 1981.

[125] Axel D Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical review A*, 38(6):3098, 1988.

[126] John P Perdew, John A Chevary, Sy H Vosko, Koblar A Jackson, Mark R Pederson, Dig J Singh, and Carlos Fiolhais. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical Review B*, 46(11):6671, 1992.

[127] Axel D Becke. Density-functional thermochemistry. iii. the role of exact exchange. *The Journal of chemical physics*, 98(7):5648–5652, 1993.

[128] Chengteh Lee, Weitao Yang, and Robert G Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical review B*, 37(2):785, 1988.

[129] Yan Zhao and Donald G Truhlar. The m06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four m06-class functionals and 12 other functionals. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 120(1):215–241, 2008.

[130] Robert K Szilagyi and Mark A Winslow. On the accuracy of density functional theory for iron-sulfur clusters. *Journal of computational chemistry*, 27(12):1385–1397, 2006.

[131] Travis V Harris and Robert K Szilagyi. Iron–sulfur bond covalency from electronic structure calculations for classical iron–sulfur clusters. *Journal of computational chemistry*, 35(7):540–552, 2014.

[132] Patrick J Bakkes, Jan L Riehm, Tanja Sagadin, Ansgar Rühlmann, Peter Schubert, Stefan Biemann, Marco Girhard, Michael C Hutter, Rita Bernhardt, and Vlada B Urlacher. Engineering of versatile redox partner fusions that support monooxygenase activity of functionally diverse cytochrome p450s. *Scientific reports*, 7(1):9570, 2017.

[133] Kerstin Maria Ewen, Frank Hannemann, Stefania Iametti, Anna Morleo, and Rita Bernhardt. Functional characterization of fdx1: evidence for an evolutionary relationship between p450-type and isc-type ferredoxins. *Journal of molecular biology*, 413(5):940–951, 2011.

[134] Peter Hlavica. Assembly of non-natural electron transfer conduits in the cytochrome p450 system: a critical assessment and update of artificial redox constructs amenable to exploitation in biotechnological areas. *Biotechnology advances*, 27(2):103–121, 2009.

[135] Pei-rang Cao, Hannes Bülow, Bruno Dumas, and Rita Bernhardt. Construction and characterization of a catalytic fusion protein system: P-450 11β-adrenodoxin reductase-adrenodoxin. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, 1476(2):253–264, 2000.

[136] Sheila J Sadeghi, Yergalem T Meharenna, Andrea Fantuzzi, Francesca Valetti, and Gianfranco Gilardi. Engineering artificial redox chains by molecular "lego". *Faraday discussions*, 116:135–153, 2000.

[137] Shengying Li, Mani Raj Chaulagain, Allison R Knauff, Larissa M Podust, John Montgomery, and David H Sherman. Selective oxidation of carbolide c–h bonds by an engineered macrolide p450 mono-oxygenase. *Proceedings of the National Academy of Sciences*, 106(44):18463–18468, 2009.

[138] Shengying Li, Larissa M Podust, and David H Sherman. Engineering and analysis of a self-sufficient biosynthetic cytochrome p450 pikc fused to the rhfred reductase domain. *Journal of the American Chemical Society*, 129(43):12940, 2007.

[139] Kirsty J McLean, Hazel M Girvan, and Andrew W Munro. Cytochrome p450/redox partner fusion enzymes: biotechnological and toxicological prospects. *Expert opinion on drug metabolism & toxicology*, 3(6):847–863, 2007.

[140] Clemens von Bühler, Priska Le-Huu, and Vlada B Urlacher. Cluster screening: an effective approach for probing the substrate space of uncharacterized cytochrome p450s. *ChemBioChem*, 14(16):2189–2198, 2013.

[141] Matthew M Purdy, Laura S Koo, Paul R Ortiz de Montellano, and Judith P Klinman. Steady-state kinetic investigation of cytochrome p450cam: interaction with redox partners and reaction with molecular oxygen. *Biochemistry*, 43(1):271–281, 2004.

[142] Christopher M Jenkins and Michael R Waterman. Nadph- flavodoxin reductase and flavodoxin from escherichia coli: characteristics as a soluble microsomal p450 reductase. *Biochemistry*, 37(17):6106–6113, 1998.

[143] Zhan Deng, Alessandro Aliverti, Giuliana Zanetti, Adrián K Arakaki, Jorgelina Ottado, Elena G Orellano, Nora B Calcaterra, Eduardo A Ceccarelli, Néstor Carrillo, and P Andrew Karplus. A productive nadp+ binding mode of ferredoxin–nadp+ reductase revealed by protein engineering and crystallographic studies. *Nature Structural & Molecular Biology*, 6(9):847–853, 1999.

[144] Luciano Piubelli, Alessandro Aliverti, Adrián K Arakaki, Néstor Carrillo, Eduardo A Ceccarelli, P Andrew Karplus, and Giuliana Zanetti. Competition between c-terminal tyrosine and nicotinamide modulates pyridine nucleotide affinity and specificity in plant ferredoxin-nadp+ reductase. *Journal of Biological Chemistry*, 275(14):10472–10476, 2000.

[145] Isabel Nogués, Jesús Tejero, John K Hurley, Darío Paladini, Susana Frago, Gordon Tollin, Stephen G Mayhew, Carlos Gómez-Moreno, Eduardo A Ceccarelli, Néstor Carrillo, et al. Role of the c-terminal tyrosine of ferredoxin-nicotinamide adenine dinucleotide phosphate reductase in the electron transfer processes with its protein partners ferredoxin and flavodoxin. *Biochemistry*, 43(20):6127–6137, 2004.

[146] Kirsty J McLean, Dominika Luciakova, James Belcher, Kang Lan Tee, and Andrew W Munro. Biological diversity of cytochrome p450 redox partner systems. In *Monooxygenase, Peroxidase and Peroxygenase Properties and Mechanisms of Cytochrome P450*, pages 299–317. Springer, 2015.

[147] Christopher C Moser, P Leslie Dutton, et al. Nature of biological electron transfer. *Nature*, 355(6363):796, 1992.

[148] Tsuneo Omura and Ryo Sato. The carbon monoxide-binding pigment of liver microsomes ii. solubilization, purification, and properties. *Journal of Biological Chemistry*, 239(7):2379–2385, 1964.

[149] Marco Girhard, Florian Tieves, Evelyne Weber, Martha Sophia Smit, and Vlada B Urlacher. Cytochrome p450 reductase from candida apicola: versatile redox partner for bacterial p450s. *Applied microbiology and biotechnology*, 97(4):1625–1635, 2013.

[150] Anna M Bogazkaya, Clemens J von Bühler, Sebastian Kriening, Alexandrine Busch, Alexander Seifert, Jürgen Pleiss, Sabine Laschat, and Vlada B Urlacher. Selective allylic hydroxylation of acyclic terpenoids by cyp154e1 from thermobifida fusca yx. *Beilstein journal of organic chemistry*, 10(1):1347–1353, 2014.

[151] Elisabetta Brenna, Claudio Fuganti, Stefano Serra, and Philip Kraft. Optically active ionones and derivatives: preparation and olfactory properties. *European Journal of Organic Chemistry*, 2002(6):967–978, 2002.

[152] J Lalko, A Lapczynski, D McGinty, S Bhatia, CS Letizia, and AM Api. Fragrance material review on β-ionone. *Food and chemical toxicology*, 45(1):S241–S247, 2007.

[153] Hideaki Kakeya, Takeshi Sugai, and Hiromichi Ohta. Biochemical preparation of optically active 4-hydroxy-β-ionone and its transformation to (s)-6-hydroxy-α-ionone. *Agricultural and biological chemistry*, 55(7):1873–1876, 1991.

[154] Walter Eschenmoser, Peter Uebelhart, and Conrad Hans Eugster. Synthese und chiralität der enantiomeren 6-hydroxy-α-ionone sowie von cis- und trans-5,6-dihydroxy-5,6-dihydroxy-β-iononen. *Helvetica Chimica Acta*, 64(8):2681–2690.

[155] Vlada B Urlacher, Akhmadjan Makhsumkhanov, and Rolf D Schmid. Biotransformation of β-ionone by engineered cytochrome p450 bm-3. *Applied microbiology and biotechnology*, 70(1):53–59, 2006.

[156] Daniela Schmitz, Josef Zapp, and Rita Bernhardt. Steroid conversion with cyp106a2–production of pharmaceutically interesting dhea metabolites. *Microbial cell factories*, 13(1):81, 2014.

[157] Perrin C White, Maria-Teresa Tusie-Luna, Maria I New, and Phyllis W Speiser. Mutations in steroid 21-hydroxylase (cyp21). *Human mutation*, 3(4):373–378, 1994.

[158] Simone Brixius-Anderko, Lina Schiffer, Frank Hannemann, Bernd Janocha, and Rita Bernhardt. A cyp21a2 based whole-cell system in escherichia coli for the biotechnological production of premedrol. *Microbial Cell Factories*, 14(1):135, Sep 2015.

[159] Hans P BLASCHKOWSKI, Joachim KNAPPE, Monika LUDWIG-FESTL, and Gabriele NEUER. Routes of flavodoxin and ferredoxin reduction in escherichia coli. *European Journal of Biochemistry*, 123(3):563–569, 1982.

[160] V. Bianchi, R. Eliasson, M. Fontecave, E. Mulliez, D.M. Hoover, R.G. Matthews, and P. Reichard. Flavodoxin is required for the activation of the anaerobic ribonucleotide reductase. *Biochemical and Biophysical Research Communications*, 197(2):792 – 797, 1993.

[161] P Razquin, S Schmitz, ML Peleato, MF Fillat, C Gómez-Moreno, and H Böhme. Differential activities of heterocyst ferredoxin, vegetative cell ferredoxin, and flavodoxin as electron carriers in nitrogen fixation and photosynthesis in anabaena sp. *Photosynthesis research*, 43(1):35–40, 1995.

[162] D Nieva-Gomez, Gary P Roberts, S Klevickis, and Winston J Brill. Electron transport to nitrogenase in klebsiella pneumoniae. *Proceedings of the National Academy of Sciences*, 77(5):2555–2558, 1980.

[163] Katsuhiko Fujii and FM Huennekens. Activation of methionine synthetase by a reduced triphosphopyridine nucleotide-dependent flavoprotein system. *Journal of Biological Chemistry*, 249(21):6745–6753, 1974.

[164] Olwen M Birch, Martin Fuhrmann, and Nicholas M Shaw. Biotin synthase from escherichiacoli, an investigation of the low molecular weight and protein components required for activity invitro. *Journal of Biological Chemistry*, 270(32):19158–19165, 1995.

[165] Mi-Sook Dong, Hiroshi Yamazaki, Zuyu Guo, and F Peter Guengerich. Recombinant human cytochrome p450 1a2 and an n-terminal-truncated form: construction, purification, aggregation properties, and interactions with flavodoxin, ferredoxin, and nadph-cytochrome p450 reductase. *Archives of Biochemistry and Biophysics*, 327(1):11–19, 1996.

[166] Lisa McIver, Claire Leadbeater, Dominic J Campopiano, Robert L Baxter, Simon N Daff, Stephen K Chapman, and Andrew W Munro. Characterisation of flavodoxin nadp+ oxidoreductase and flavodoxin; key components of electron transfer in escherichia coli. *European journal of biochemistry*, 257(3):577–585, 1998.

[167] Federico Sabbadin, Gideon Grogan, and Neil C. Bruce. *LICRED: A Versatile Drop-In Vector for Rapid Generation of Redox-Self-Sufficient Cytochromes P450*, pages 239–249. Humana Press, Totowa, NJ, 2013.

[168] Hidehiko Hirakawa and Teruyuki Nagamune. *Use of Sulfolobus solfataricus PCNA Subunit Proteins to Direct the Assembly of Multimeric Enzyme Complexes*, pages 149–163. Humana Press, Totowa, NJ, 2013.

[169] Sang Taek Jung, Ryan Lauchli, and Frances H Arnold. Cytochrome p450: taming a wild type enzyme. *Current opinion in biotechnology*, 22(6):809–817, 2011.

[170] Patrick Argos. An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. *Journal of molecular biology*, 211(4):943–958, 1990.

[171] Richard A George and Jaap Heringa. An analysis of protein domain linkers: their classification and role in protein folding. *Protein engineering*, 15(11):871–879, 2002.

[172] Takanori Tanaka, Shigeyuki Yokoyama, and Yutaka Kuroda. Improvement of domain linker prediction by incorporating loop-length-dependent characteristics. *Peptide Science*, 84(2):161–168, 2006.

[173] Vishnu Priyanka Reddy Chichili, Veerendra Kumar, and J Sivaraman. Linkers in the structural biology of protein–protein interactions. *Protein Science*, 22(2):153–167, 2013.

[174] Toon H Evers, Elisabeth MWM van Dongen, Alex C Faesen, EW Meijer, and Maarten Merkx. Quantitative understanding of the energy transfer between fluorescent proteins connected via flexible peptide linkers. *Biochemistry*, 45(44):13183–13192, 2006.

[175] Ryoichi Arai, Willy Wriggers, Yukihiro Nishikawa, Teruyuki Nagamune, and Tetsuro Fujisawa. Conformations of variably linked chimeric proteins evaluated by synchrotron x-ray small-angle scattering. *PROTEINS: Structure, Function, and Bioinformatics*, 57(4):829–838, 2004.

[176] Joshua S Klein, Siduo Jiang, Rachel P Galimidi, Jennifer R Keeffe, and Pamela J Bjorkman. Design and characterization of structured protein linkers with differing flexibilities. *Protein Engineering Design and Selection*, 27(10):325–330, 2014.

[177] Diane A Hall, Craig W Vander Kooi, Chad N Stasik, Shawn Y Stevens, Erik RP Zuiderweg, and Rowena G Matthews. Mapping the interactions between flavodoxin and its physiological partners flavodoxin reductase and cobalamin-dependent methionine synthase. *Proceedings of the National Academy of Sciences*, 98(17):9521–9526, 2001.

[178] Peter Hlavica. Mechanistic basis of electron transfer to cytochromes p450 by natural redox partners and artificial donor constructs. In *Monooxygenase, Peroxidase and Peroxygenase Properties and Mechanisms of Cytochrome P450*, pages 247–297. Springer, 2015.

[179] Ansgar Rühlmann, Dragutin Antovic, Thomas J. J. Müller, and Vlada B. Urlacher. Regioselective hydroxylation of stilbenes by engineered cytochrome p450 from thermobifida fusca yx. *Advanced Synthesis & Catalysis*, 359(6):984–994.

[180] Paula Bracco, Dick B. Janssen, and Anett Schallmey. Selective steroid oxyfunctionalisation by cyp154c5, a bacterial cytochrome p450. *Microbial Cell Factories*, 12(1):95, Oct 2013.

[181] Hyun-Hee Jang, Sang-Hoon Ryu, Thien-Kim Le, Tiep Thi My Doan, Thi Huong Ha Nguyen, Ki Deok Park, Da-Eun Yim, Dong-Hyun Kim, Choong-Kyung Kang, Taeho Ahn, Hyung-Sik Kang, and Chul-Ho Yun. Regioselective c-h hydroxylation of omeprazole sulfide by bacillus megaterium cyp102a1 to produce a human metabolite. *Biotechnology Letters*, 39(1):105–112, Jan 2017.

[182] Steffen C. Maurer, Katja Kühnel, Leonard A. Kaysser, Sabine Eiben, Rolf D. Schmid, and Vlada B. Urlacher. Catalytic hydroxylation in biphasic systems using cyp102a1 mutants. *Advanced Synthesis & Catalysis*, 347(7-8):1090–1098.

[183] Alexander Dennig, Nina Lülsdorf, Haifeng Liu, and Ulrich Schwaneberg. Regioselective o-hydroxylation of monosubstituted benzenes by p450 bm3. *Angewandte Chemie International Edition*, 52(32):8459–8462.

[184] C. A. Müller, A. M. Weingartner, A. Dennig, A. J. Ruff, H. Gröger, and Ulrich Schwaneberg. A whole cell biocatalyst for double oxidation of cyclooctane. *Journal of Industrial Microbiology & Biotechnology*, 43(12):1641–1646, Dec 2016.

[185] Rudi Fasan, Mike M. Chen, Nathan C. Crook, and Frances H. Arnold. Engineered alkane-hydroxylating cytochrome p450bm3 exhibiting nativelike catalytic properties. *Angewandte Chemie International Edition*, 46(44):8414–8418.

[186] Vlada B. Urlacher and Marco Girhard. Cytochrome p450 monooxygenases: an update on perspectives for synthetic application. *Trends in Biotechnology*, 30(1):26–36, 2018/03/22.

[187] Florian Kiefer, Konstantin Arnold, Michael Künzli, Lorenza Bordoli, and Torsten Schwede. The swiss-model repository and associated resources. *Nucleic Acids Research*, 37(suppl_1):D387, 2009.

[188] Nicolas Guex, Manuel C. Peitsch, and Torsten Schwede. Automated comparative protein structure modeling with swiss-model and swiss-pdbviewer: A historical perspective. *ELECTROPHORESIS*, 30(S1):S162–S173, 2009.

[189] William L Jorgensen and Jeffry D Madura. Quantum and statistical mechanical studies of liquids. 25. solvation and conformation of methanol in water. *Journal of the American Chemical Society*, 105(6):1407–1413, 1983.

[190] Saurabh Kumar Shakya, Wei Gu, and Volkhard Helms. Molecular dynamics simulation of truncated bovine adrenodoxin. *Biopolymers*, 78(1):9–20, 2005.