GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Investigation of plant aldo-keto reductase involved in stemmadenine biosynthesis

Bachelor Thesis

Bachelor Biochemistry

submitted by

Marlen Siegmund

born in

Weimar

written at

Max-Planck-Institut for Chemical Ecology Jena

Faculty for Biology und Psychology

Georg-August-University of Göttingen

submission spring semester 2022

1. Examiner: Prof. Dr. Ivo Feußner

2. Examiner: Dr. rer. nat. Hornung

submitted: 05.07.2022

Abstract

Monoterpene indole alkaloids (MIAs) are a diverse and important class of plant natural products that include a number of medicinally significant compounds, often present at low concentrations within their native plant species. The early part of the vinblastine pathway results in stemmadenine, a precursor to many important MIA scaffolds. The final step of stemmadenine synthesis is catalysed by the RedOx2 enzyme. Here the RedOx2 crystal structure of *Catharanthus roseus* is presented with a resolution of 1.5 Å. Structural comparison to a relative protein in *Tabernanthe iboga* lead to key residues which were analysed using site directed mutagenesis and enzymatic assays. Two key residues were identified having an influence on the product formation and product ratio. These findings build a base to prevent the formation of side products and help to increase the stemmadenine production by RedOx enzyme.

Monoterpenoid inodol alkaloide (MIAs) sind eine Klasse von vielseitigen, chemisch komplexen Metaboliten, welche auch in der Medizin Anwendnung finden. Oft können diese jedoch nur in geringer Konzentation aus Pflanzenextrakten gewonnen werden. Der frühe Teil des Vinblastin Stoffwechselweges endet mit Stemmadenine. Dies ist ein Zwischenprodukt für viele wichtige MIAs. Die Synthese von Stemmadenine wird durch das RedOx2 Enzym katalysiert. In dieser Arbeit wird die Struktur von RedOx2 aus *Catharanthus roseus* mit einer Auflösung von 1.5 Å präsentiert. Mittels eines eng verwandten Enzymes aus *Tabernanthe iboga* wurde eine Strukturanalyse durchgeführt. Dafür wurden sowohl zielgerichtete Mutation, als auch enzymatische Untersuchungen genutzt. Dadurch wurden zwei Schlüssel Aminosäuren identifiziert, welche sowohl auf die Produkmenge, als auch das Produktverhältnis einen Einfluss haben. Die hier präsentierten Ergebnisse können die Grundlage für eine Vermeidung von unerwünschten Nebenprodukten, als auch eine erhöhte Stemmadenine-Produktion bilden.

Acknowledgements

I want to thank everyone from the Sarah O'Connor lab for the warm welcome during peak CoVid19 times. Over the months working in the lab I got to know so many new things, not only about science, but about working together and staying strong together during a period of crisis. It's been a honour working with everyone of you and I especially want to thank the following people.
Sarah O'Connor for giving me the opportunity to be part of your great team and your stunning research. Thanks to Lira Palmer for every motivation, each coffee chat and your absolute catching enthusiasm for science. I want to thank Mohamed Omar Kamileen for all the knowledge you passed to me, the time you spend teaching me and the fundamental basement I gained how to become a good scientist. Chloe Langley for all the talks and your help with the crystals, the docking and the protein mutants. And Nestor J. Hernandez Lozada for your help and good hints with dealing with protein purification and PyMol. A warm thank you also to the MS team Delia Ayled Serna Guerrero, Maritta Kunert and Sarah Heinicke for method developement, handling and taking care about all the instruments. And finally Jerome Basquin from MPI Biochemistry (Munich, Martinsried) for X-Ray crystallography and structure refinement.

Contents

Abbreviations

AKR aldo-keto reductase

BIA benzylisoquinoline alkaloid

Carb carbenicillin

cDNA complementary DNA

CHR chalcone reductase

CD circular dichroism

COR codeinone reductase

Cr *Catharanthus roseus*

CrRedOx2 oxidoreductase 2 of *Catharanthus roseus*

DNA deoxyribunucleic acid

ESI electrospray ionisation mode

FPLC fast protein liquid chromatography

GO geissoschizine oxidase

HEPES 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HF high fertility

HPLC high pressure liquid chromatography

IPTG isopropyl β-D-1-thiogalactopyranoside 2

LB-media lysogeny broth media

LC-MS liquid chromatography–mass spectrometry

MEP methylerythritol 4-phosphate

MIA monoterpenoid indole alkaloid

Ms *Medicago sativa*

MS	mass spectrometry
m/z	mass-to-charge ratio
NADP	nicotinamide adenine dinucleotide phosphate
NMR	nuclear magnetic resonance
OD	optical denisity
PDB	protein data base
Ps	Papaver somniferum
PCR	polymerase chain reaction
Q-ToF	quadrople time of flight
RedOx1	oxidoreductase 1
RedOx2	oxidoreductase 2
RMSD	root mean square deviation
RNA	ribonucleic acid
SDS-PAGE	sodium dodecyl sulphate – polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SGD	strictosidine glucosidase
STR	strictosidine synthase
Ti	*Tabernanthe iboga*
TIC	total ion chromatogram
TIM	triose phosphate isomerase
TiRedOx2	oxidoreductase 2 of *Tabernanthe iboga*
WT	wild type

List of Figures

List of Tables

1 Introduction

1.1 Biological background

1.1.1 Enzymes

When life evolved on earth organisms were in need to assimilate metabolites from their environment and convert them into useful scaffolds for their own purpose. Chemical machinery, called enzymes, were evolved to maintain this function and provide all necessary substances. All living systems on earth are dependent on enzymatic activity [39]. Their primary role is to lower the activation energy of various chemical reactions by acting as catalysts. Therefore, enzymes use various systems such as destabilising a compound by creating an altered environment in their active site. The active site consists of a binding and catalytic site for the desired compound. The compound that is taken up by the enzyme is called substrate. The modified scaffold that leaves the enzyme is called product. All states in between are called intermediates and have to be stabilised by the enzyme. Endless varying molecules can be used by different enzymes. Evolution lead to more substrate-specialised enzymes with higher efficiency at expense of promiscuity [51]. Promiscuity is the ability of one enzyme to take up more than one substrate and therefore catalyse different reactions. This property pushes chemical evolution forward. Some enzymes kept their promiscuity as a part of a stable system [48]. As a result of enzyme evolution we can see so called enzyme super families that share a common backbone of structure. The backbone can consist of a strongly conserved constitution of alpha helices and beta sheets which proved itself. Helices and sheets are called secondary structure of enzymes. These structures arise from the primary structure, the amino acid sequence of an enzyme. The tertiary structure then describes the combination and interaction of secondary segments [41]. Another way of structure preservation can be the arrangement of chemical properties or amino acid residues in the active site. Those residues preserve the ability of enzymes to catalyse certain reactions. Structures like this, alone or in combination, define the specification of families in today's enzymology [36].

1.1.2 Plant metabolism

A biosynthetic pathway consists of multiple steps usually catalysed by enzymes. The product of the first enzyme is used as substrate for the following enzyme. These systems enable construction of more complicated scaffolds. Metabolite networks arose among the first biological systems. Membranes separated individual organisms and selective pressure arose. As consequence, more effective pathways and development of new metabolites became essential [39]. Pathways increasing the biological fitness were established, enzymes became more efficient and the genes involved became more robust to activity-killing mutations [33]. The benefits of new substances influenced the structural and chemical evolution so that organisms became dependent on certain pathways, which are now summarised as primary metabolism. In plants, one common example is photosynthesis which converts the energy from sunlight into chemical energy, as well as provides the only carbon source of terrestrial plants. Therefore, carotenoids are needed as scaffolds that can absorb the light which arise from the methylerythritol 4-phosphate (MEP) pathway.
In plant biosynthesis, enzymes can lead to scaffold formations, and modifications to backbones of chemical scaffolds (tailoring)[50]. Due to evolution, the primary metabolism, which provides essential scaffolds, became a complicated network of pathways which is not yet completely understood.

1.1.3 Plant specialised metabolism and evolution

Pathways are typically classified into primary and secondary metabolism. The latter are defined as not being involved in central metabolism, or which the organism would die without [17]. Secondary metabolism is inter alia now referred to as specialised metabolism. In order to gain an advantage over the competition, the plant has to invest resources and energy. Since primary (central) metabolism is essential, using those enzymes for development directly bears a risk of losing them. A gene duplication provides a copy that can be mutated and adjusted without serious consequences. Often the gene undergoes a few mutations that can either broaden the enzyme specificity again or lead to completely new activities. There are several ways to explore new compounds through new enzymatic activity. (1) The enzyme can broaden its substrate specificity and apply the original chemical changes to different compounds. (2) An

established intermediate can be tailored in a distinct way by displaying it differently to the active site. (3) Hybrid enzymes can be created by adding an abiotic catalyst, such as transition metals, into the active site and change the chemistry applied to an intermediate [30].

1.1.4 Alkaloids in plants

One of the largest groups of specialised metabolites in plants are alkaloids. More than 12,000 different compounds belong to the group of structures with a heterocyclic nitrogen derived from an amino acid [42]. Alkaloids are known and used for their medicinal properties since ancient times. Despite size and versatility of the available and known alkaloids, morphine from opium poppy (*Papaver somniferum*, Papaveraceae) is most probably the most notable member. It was first isolated 1803 by Friedrich Sertürner and belongs to a group of poorly extractable products, the benzylisoquinoline alkaloids (BIAs) [45]. Those alkaloids derive from a common intermediate called (*S*)-reticuline. Morphine, and the earlier intermediate codeine, are compounds with high economic importance (figure 1).

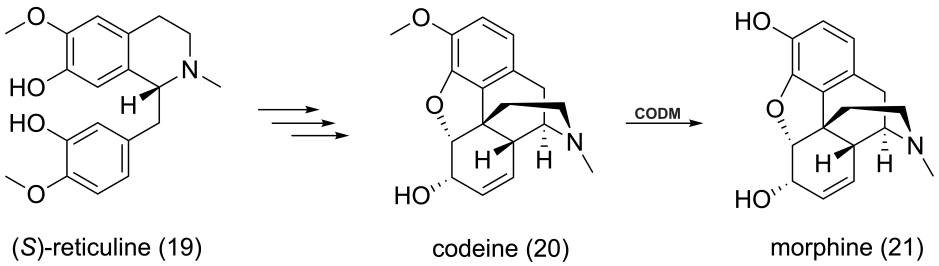

Figure 1: BIA scaffolds origin from (*S*)-reticuline. Codeine as important compounds for itself is also a precurser of morphine. The reaction is catalysed by an enzyme called codeine *O*-demethylase (CODM).

Another large group are monoterpenoid indole alkaloids (MIAs). The precursor to all MIAs is strictosidine (figure 2). Vinblastine and vincristine are two well known MIAs with anticancer activity produced by the plant *Catharanthus roseus* (Madagascar periwinkle).

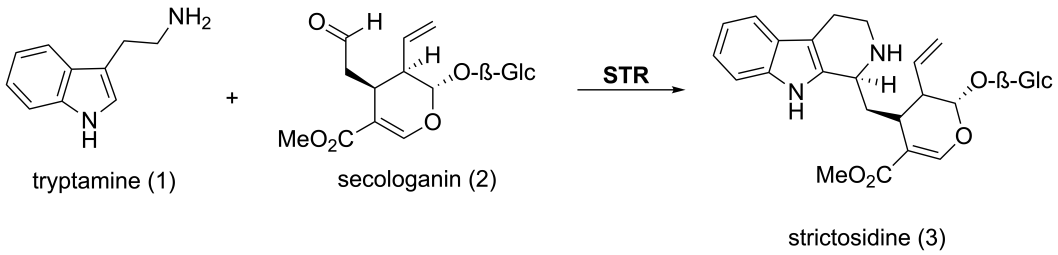

Figure 2: Formation of strictosidine by secologanin and tryptamine. The reaction is catalysed by strictosidine synthase (STR).

Apocynacea plant family Plants of the Apocynaceae family produce a majority of the known MIAs with various scaffolds utilising the precursor strictosidine [3]. The plant family is one of the largest group of plant families. It contains around 4600 species divided in five subfamilies (figure 3). The origin of the family lies in south east asia mainly and many species are reported to produce alkaloids. Various types are reported including indole, iboga and vinca alkaloids. Therefore, those plants were early and broadly used in traditional medicine. Today, an increase of production is attempted through engineering parts of important pathways into heterologous hosts, for instance strictosidine [9].

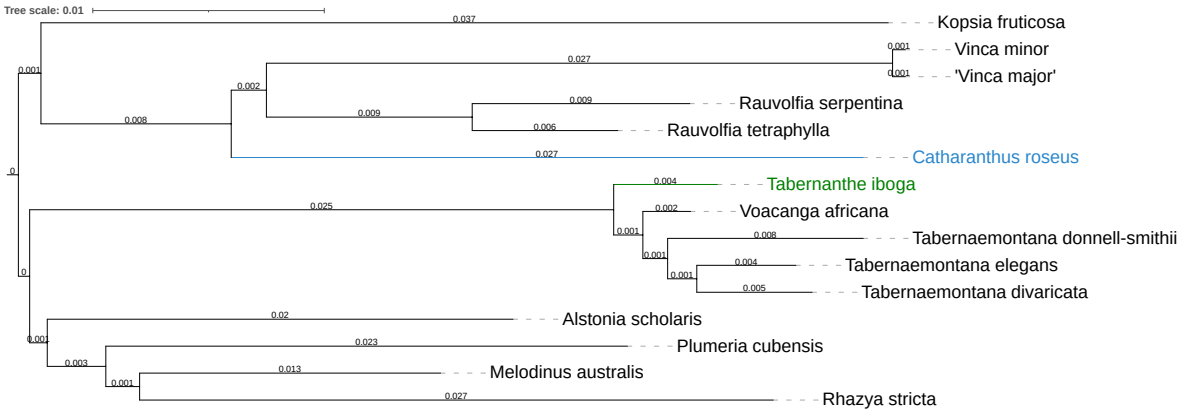

Figure 3: A subset of species of the Apocynaceae plant family where *Tabernanthe iboga* (green) and *Catharanthus roseus* (blue) belong to. Tree was made with iTOL [29].

1.2 Vinblastine pathway

Vinblastine and vincristine are important anti-cancer MIAs used for different forms of cancer like leukaemia and lymphoma types. The pathway shown in the following figure 4 consists of at least 20 enzymes, where most of them are known [10]. Parts of the pathway

are also reported from various species while vinblastine and vincristine themselves are only produced by the plant *Catharanthus roseus*. High efforts were made to fully reveal and understand the pathway. The segment from strictosidine up to stemmadenine can be simplified as the early MIA pathway in *C. roseus* (figure 4 grey). Beginning from tryptamine and secologanin, six enzymes are involved. Stemmadenine serves as an important, stable precursor to key MIA scaffolds, aspidosperma-type and iboga-type. Stemmadenine has to be produced by plants and isolated from plant extracts with a very low yield. Due to the novelty and rarity of this plant natural product, an increase of production is desired.

1.2.1 Stemmadenine pathway

Strictosidine is formed by tryptamine and secologanin (figure 4). Tryptamine provides the indole ring structure (indole alkaloid) while secologanin is a seco-iridoid (oxidized monoterpene). The reaction is called Picet-Spengler-condensation [19]. Strictosidine so derives from iridoids, which again derive from the isoprenoid pathway. The SGD (strictosidine glucosidase)[18] deglycosylates strictosidine resulting in dehydrogeissoschizine, which will used by GS (geissoschizine synthase), followed by GO (geissoschizine oxidase)[47]. The product is used then by RedOx1 [38]. Both, substrate and product of RedOx1 remain structurally undefined due to their unstable nature. This attributes of some intermediates can be explained by their high reactivity, which makes them difficult to characterise by isolation. Here, unstable intermediates spontaneously deformulate and accumulate as sink (or shunt) products. They convert immediately into a more stable scaffolds, for instance instead of preakuammicine figure 4 (**8**), akuammicine (**9**) can be found, which is not a part of this pathway. The RedOx2 enzyme then stabilises the product of RedOx1 by forming stemmadenine in an aldehyde reduction [38].

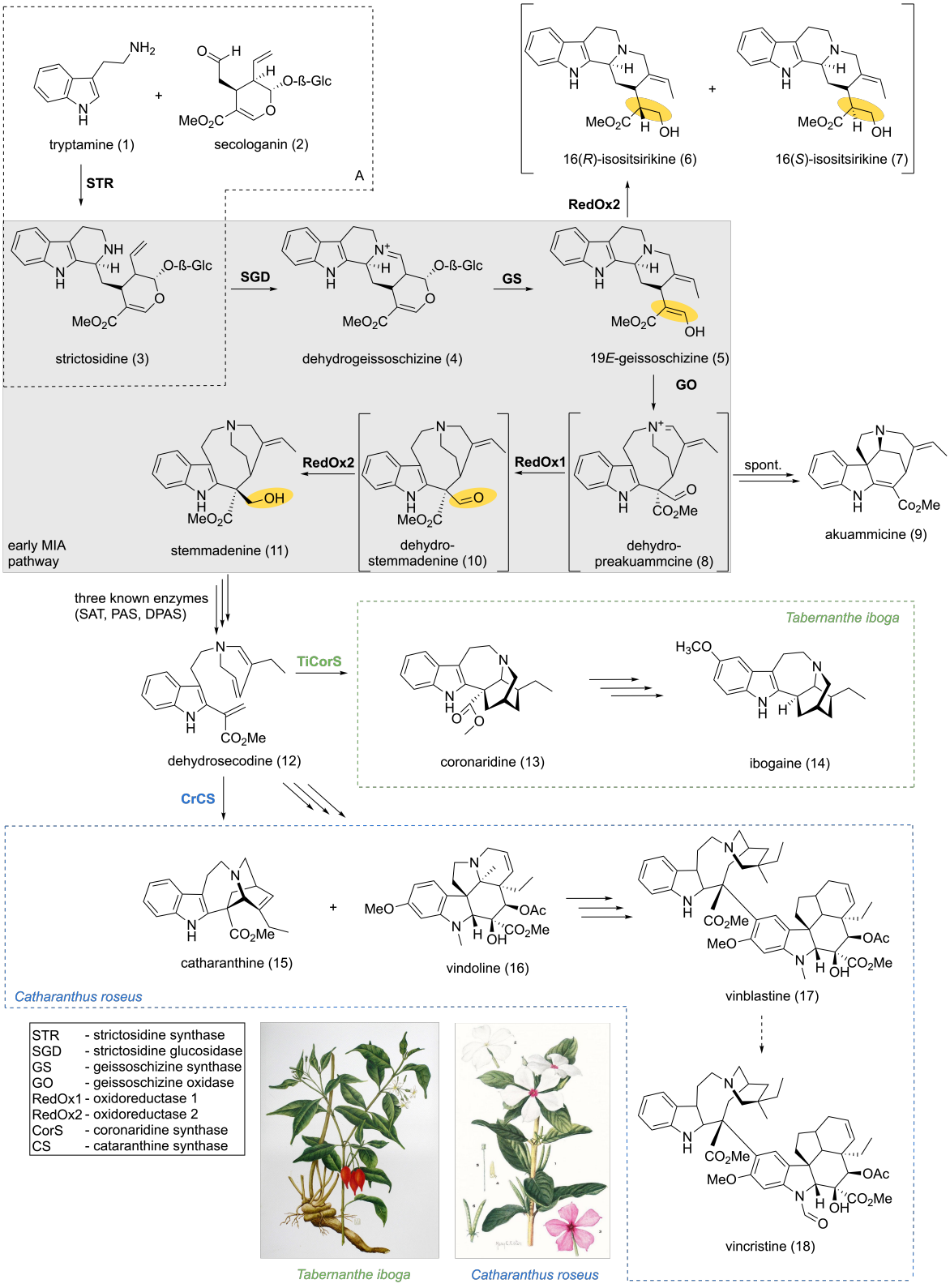

Figure 4: Pathway beginning with tryptamine **1** and secologanin **2** forming strictosidine **3** in a Picet-Spengler condensation **A**. The early MIA pathway is coloured in grey. The pathway divides at dehydrosecondine **12** in *T. iboga* products (green) coronaridine **13** and ibogaine **14**, and in *C. roseus* products (blue) cataranthine **15** and vindoline **16**. Yellow circles show where RedOx2 enzyme catalyses reactions. Multiple arrows indicate multiple transformations to yield the intermediate. Dashed arrows indicate transformations that are still unknown. Plant figures from [15],[32]. Figure made with ChemDraw [34].

1.2.2 Alkaloids in *Catharanthus roseus*

Catharanthus roseus (Madagascar periwinkle) is a well-studied plant species and is considered as a model plant in alkaloid biosynthesis. Because of its therapeutic value in producing the MIA anticancer compounds vinblastine and vincristine, elucidating the biosynthetic step remained a key challenge, inspiring many chemists, biochemists and plant scientists to discover ground breaking techniques in the field of plant natural product biosynthesis. Most of the enzymes involved in MIA biosynthesis were hence discovered, as a result of characterisation of vinblastine biosynthesis in *C. roseus* over the past 20 years. O'Connor, DeLuca and others pioneered this work [10],[38]. The pathway still remains to be fully elucidated and is one of the most complicated pathways in specialised metabolism. Technological progress in next generation sequencing made it possible to accelerate the gene discovery process and identify enzymes. The vinblastine pathway has a complex architecture because of its length and the compartmentalisation in the cell [12].

1.2.3 Alkaloids in *Tabernanthe iboga*

Another plant known for its potential medical use is *Tabernanthe iboga* (Iboga, Apocynacea). The plant primarily produces the hallucinogenic MIA alkaloid ibogaine accumulated in the root bark. In recent years ibogaine has shown promising effects in alleviating opiate withdrawal symptoms [52]. Transcriptomic analysis led to the discovery and characterisation of enzymes involved in ibogaine biosynthetic pathway from stemmadenine, an intermediate shared with vinblastine biosynthesis in *C. roseus* [16]. Recent studies by O'Connor (unpublished) showed that iboga shared gene candidates with high sequence similarity to the *C. roseus* early MIA enzymes leading to stemmadenine. Hence, functional characterisation of the iboga gene candidates was started. In some cases, like oxidoreductase 2 (RedOx2), the sequence similarity is 50% which gives reason to investigate the impact of these differences in the pathway.

1.3 RedOx2 enzyme

1.3.1 AKR super family

The RedOx2 enzyme belongs to the aldo-keto reductase (AKR) superfamily. The superfamily contains at least 190 members described so far [21]. Those enzymes originate from all known species. Particularly of great interest were the human 20α-hydroxysteroid dehydrogenase (h20α-HSD) involved in the binding of C21-steroids [13],[37]. In plants, the AKR enzymes of the model plant *Arabidopsis thaliana* were of great interest since they are involved in detoxification in stress situations [46]. AKR enzymes are reported to full fill numerous functions in plants by accepting a broad range of substrates [44]. Especially those involved in biotic and abiotic stress defence show a broad substrate specificity, which is helpful for detoxification of compounds for instance. However, a broad substrate range makes the classification into metabolism and physiological function of new members difficult. Meanwhile enzymes found to belong to specialised metabolism narrowed their substrate range rigorously [14]. The characteristics defining an AKR enzyme are of structural and also catalytic nature. Their cofactor is either NADPH or NADP+, catalysing either a reduction or oxidation reaction. The enzymes are about 37 kDa size and have an $(\alpha/\beta)_8$-barrel motif [24] (figure 5). Some also possess additional alpha-helices on top of the substrate binding pocket (figure 5 **A** A1, A2) and / or a structure closing the barrel motif at the other side of the enzyme (figure 5 **A** B1, B2). While the structures coloured in figure 5 **A** remain highly preserved, the loops coloured in figure 5 **B** are flexible in both structure and amino acid sequence. They are responsible for substrate recognition and uptake. This structure was first observed in the triose phosphate isomerase and is therefore referred as TIM-barrel [4]. The catalytic characteristics of the enzymes contain a catalytic tetrad of four amino acids: tyrosine, aspartic acid, lysine and histidine in a strongly conserved composition. Those sit deeply in the catalytic pocket. In general the AKR enzymes catalyse their reactions through a Bi-Bi kinetic mechanism, which means that the cofactor binds first and is released last when the reaction is complete [22]. The superfamily is again divided in seven sub-families [23]. Superfamilies in general show a sequence identity less than 40% to other superfamilies. A sequence identity of more than 60% is needed for clading in the same family. In the AKR superfamily the different families, in a numbered system, are again divided in subfamilies in an

alphabetic system. The majority of plant enzymes can be found in the AKR4 clade.

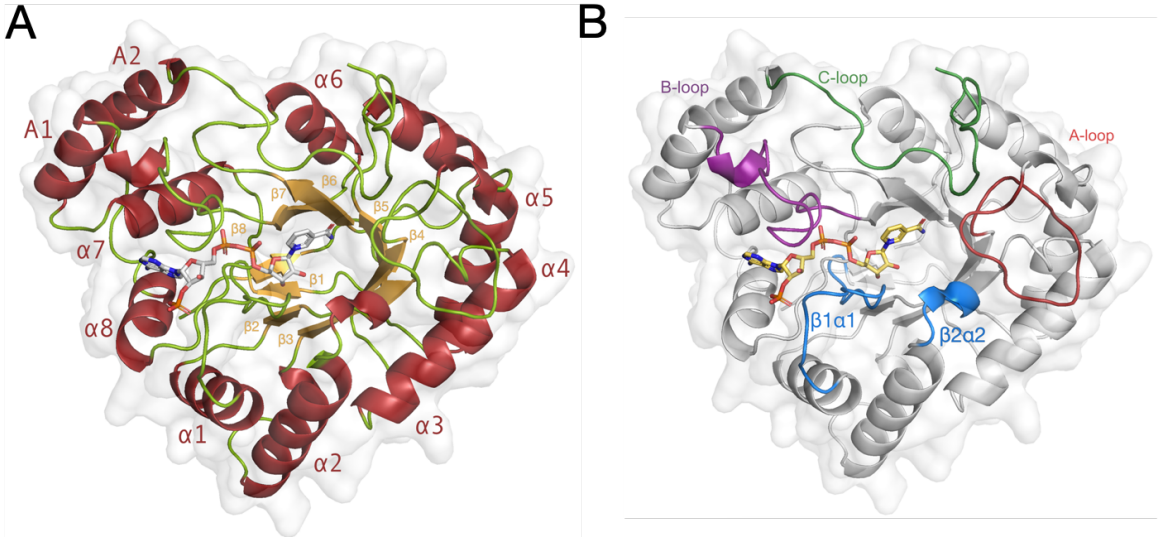

Figure 5: Secondary structure of chalcone reductase (CHR) with NADP+ bound [8] as example for an AKR protein structure. **A** alpha helices (red), beta sheets (yellow) loop structures (green) NADP+ (grey). Alpha helices and beta sheets are aligned in a conserved composition referred as TIM-barrel. They bear the catalytic tetrad responsible for the characteristic catalysis of AKR enzymes. **B** loop structures involved in substrate pocket are much less conserved and form the actual catalytic pocket, NADP+ (yellow). Figure made with PyMol [43].

1.3.2 AKR4 clade

The enzymes of the AKR4 clade are described to belong to the biotic and abiotic stress defence, as well as production of commercially important secondary metabolites such as morphine. The AKR4 family is again divided in three reported subfamilies: A, B and C. The family is mainly known for the chalcone reductase and codeine reductase of different species belonging to it [37]. All four enzymes known from the model plant *Arabidopsis thaliana* belong to the AKR4 family and are well studied in their substrate activity [44],[46].

1.3.3 Chalcone reductase of *Medicago sativa*

Chalcone reductase (CHR) is a NADPH dependent AKR4 involved in flavonoids and isoflavonoids biosynthesis and depends on NADPH. The enzyme was detected in numerous species and is involved in the formation of precursors for important flavonoids and isoflavonoids. There among others CHR interacts closely with chalcone synthase

(CHS) forming deoxychalcone. Despite its naming, chalcone reductase does not reduce chalcone [35] but rather intermediates of a multi-step reaction resulting in chalcone (27) and deoxychalcone (28) (figure 6 **D**). A detailed description of CHR of *Medicago sativa* (alfalfa) was published along with it's crystal structure in 2005 [8]. This structure was the first of the AKR4 family to be published and therefore served as a template for most plant AKRs. The structure diffracted to 1.7 Å as a dimer. $NADP^+$ is bound as cofactor (figure 6 **B**).

1.3.4 Codeinone reductase of *Papaver somniferum*

Another example is codeinone reductase (COR), involved in the formation of morphine, one of the most important alkaloids. As seen in figure 6 **C** the enzyme catalyses not only two different reactions but also the backward reactions of them. The promiscuity of the enzyme, using two different substrates, where both can lead to the production of morphine, is an example for desirable promiscuity. Those attributes make the enzyme an attractive target for protein engineering. For instance, a higher outcome could be achieved by inhibiting the protein to perform the backwards reaction using $NADP^+$ as a substrate. The crystal structure of COR1.3 of *Papaver somniferum* was published additional to a very detailed analysis including a residues-function relation [11]. The final structure was solved using molecular replacement and refined to a resolution of 2.4 Å. The structure diffracted as a hexamer with slight differences in each unit (figure 6 **A**). One loop is missing in the structure (figure 6 **A** light green). There is no ligand bound to the protein structure.

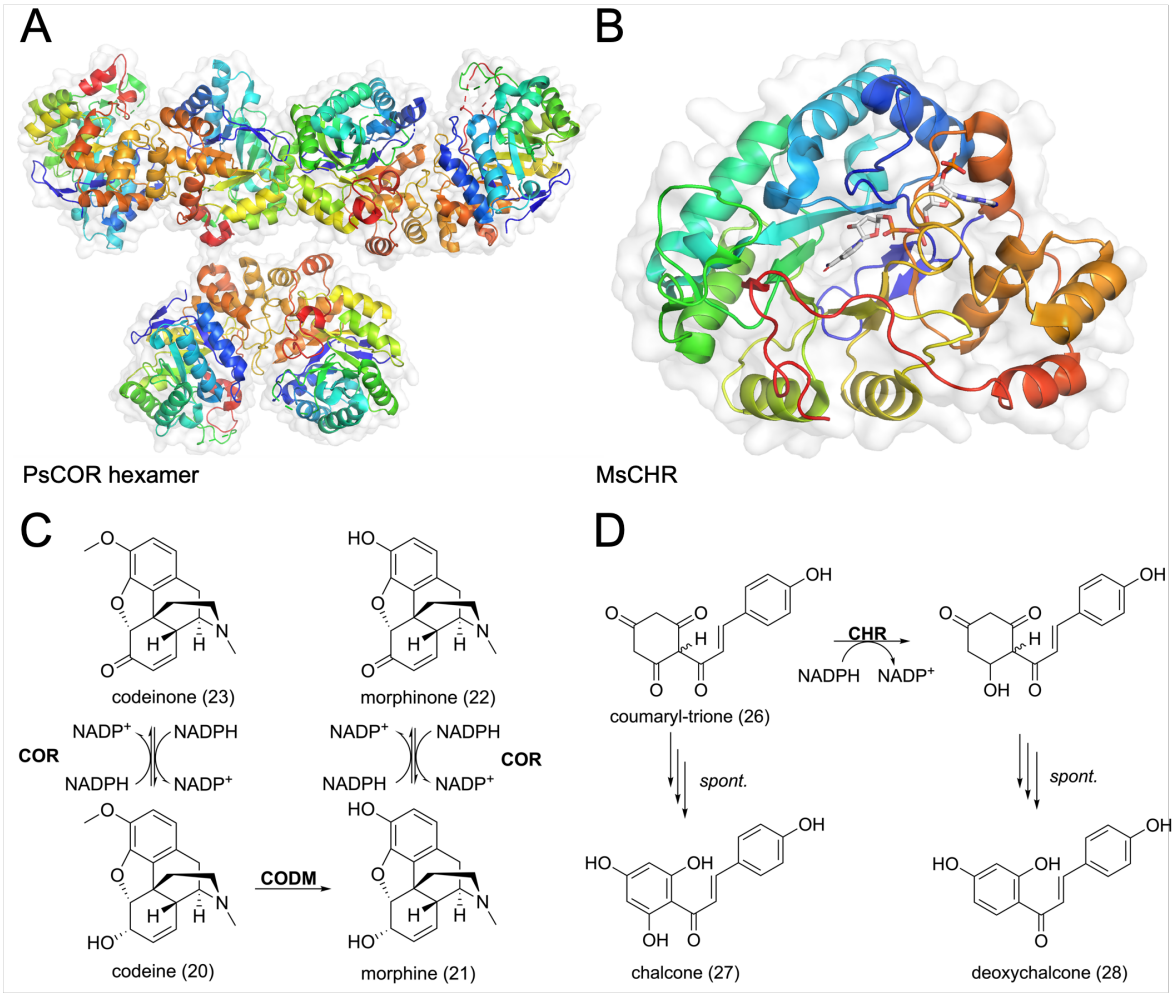

Figure 6: **A** Hexamer of PsCOR protein, each chain is coloured from N-terminus (blue) to C-terminus (red). **B** MsCHR protein structure with NADP+ (grey) bound in rainbow colouring sheme starting from N-terminus (blue) to C-terminus (red). **C** reaction catalysed by COR and CODM resulting in morpphine **21**, **D** proposed reaction catalysed by CHR resulting in chalcone **27** and deoxychalcone **28**. Figures made with PyMol and ChemDraw [43],[34].

1.3.5 RedOx2 enzyme

An example for undesirable promiscuity is the enzyme RedOx2, involved in the production of the very important intermediate stemmadenine [38] and belonging to the vinblastine pathway. Same as CHR and COR, RedOx2 is an oxidoreductase belonging the enzyme superfamily of aldo-keto reductase enzymes. As seen in figure 7 RedOx2 catalyses two different reactions in the pathway of stemmadenine. That means the enzyme is not as specialised as it could be. The first reaction catalysed by the enzyme is a reduction of an aldehyde group of dehydrostemmadenine. The reaction causing undesired sink products, in interest of producing stemmadenine, is the reduction of

geissoschizine which results in the isomers 16-(R/S)-isositsirikines. There, RedOx2 takes up an earlier intermediate of the pathway, 19E-geissoschizine. So far stemmadenine as product has been gained from plants extracts only. With investigating RedOx2 enzyme, in vitro production of stemmadenine could be possible and efficiently used.

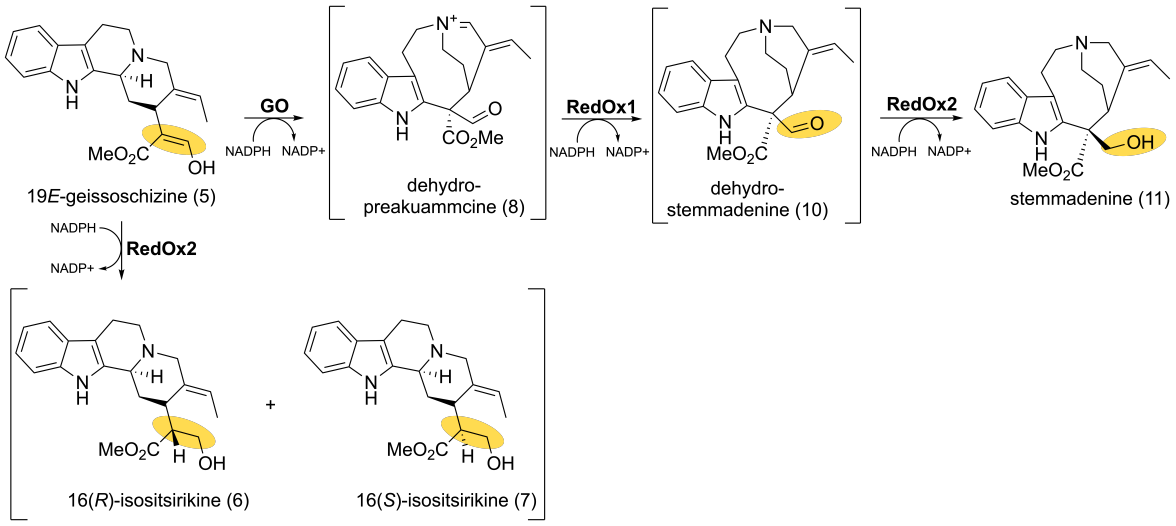

Figure 7: Part of the stemmadenine pathway beginning at 19E-geissoschizine and resulting in both stemmadenine and 16(R/S)-isositsirikines. The substrate and product of RedOx1 remain structurally unknown due to their instable nature and are proposed as shown here [38]. Reactions catalysed by RedOx2 are highlighted in yellow.

AlphaFold In 2021 AlphaFold [26] published a predicted structure of CrRedox2 (UniProt A0A2P1GIY9). AlphaFold is a protein structure database by DeepMind and EMBL-EBI which focuses on predicting as many precise protein structures as possible. The AlphaFold algorithm is an artificial intelligence system that was trained with already published structures. Those predictions turn out to be highly precise when compared to a actual crystal structure. However, the catalytic pocket of RedOx2 is built of flexible loops with no systematic secondary structure. The predicted structure lacks of model confidence in those areas. Furthermore no cofactors or substrates can affect the prediction, while the protein certainly adjusts to those factors. Summarised, these structures are a helpful tool but do not replace the need for protein crystallisation. The quality of the similarity of superimposed proteins can be measured with the root mean square deviation (RMSD) value, which describes the average distance between the backbone atoms.

2 Aims of this work

In this thesis I aim to investigate the enzyme RedOx2. Therefore, I will compare the activity of RedOx2 enzyme of *Catharanthus roseus* (named CrRedox2) with three homologues from *Tabernanthe iboga* (named TiRedox2). There are three TiRedox2 candidates present, where the first one and second one show a sequence identity of 83 % to CrRedox2. The third candidate shows a sequence identity of 50 % only.

Using crystallisation and site directed mutagenesis I will investigate the structure and the activity of the named enzymes on two reactions, namely the production of stemmadenine and the production of the undesirable sink product 16(R/S)-isositsirikines using 19E-geissoschizine as a substrate. I will optimise the reaction conditions for the enzymes and test the activity of them in different pH conditions.

In order to further characterise and analyse the enzyme, I aim to crystallise a RedOx2 enzyme and analyse it by using X-Ray crystallography. Structural comparison and *in silico* analysis of the gained protein structure and homology models will help me to understand the amino acid composition of the substrate binding pocket.

Based on structural analysis I will create mutants of the available enzymes using site directed mutagenesis and review the influence of these mutations on the catalytic activity. By doing so I will investigate the role of individual residues in the two reactions carried out by the enzyme RedOx2.

3 Methods

3.1 Analysis of the Transcriptome

The transcriptome was available from *Tabernanthe Iboga* [16] and *Catharanthus roseus* (Medicinal Plant Genomics Resource (MPGR) consortium (medicinalplantgenomics. msu. edu)) [27]. RedOx2 of *C. roseus* was BLAST [2] against the *T. iboga* transcriptome to receive the homologs of RedOx2 present in *T. iboga*. Primers were designed with Geneious Prime [6] and used to extract the gene from the cDNA. A phylogenetic tree was build using Geneious Prime tree builder (standard settings) and iTol [29].

3.2 Gene purification and primer design

The gene from the transcriptome was isolated using the specific primers of each gene. The gene of interest, amplified from cDNA, was cloned into pOPINF vector using In-Fusion cloning [5]. In-Fusion cloning is a gibson assembly method. The principle is shown in figure 8. pOPINF is a plasmid designed for errorless cloning and expression of proteins. The vector has multiple restriction sites where the gene of interest can be inserted. Here KpnI and HindIII were used as restriction enzymes for all described genes. Also the pOPINF vector has a antibiotic resistance of carbencillin (Carb). Using a media with antibiotics added, bacteria with unsuccessful infusion are sorted out.

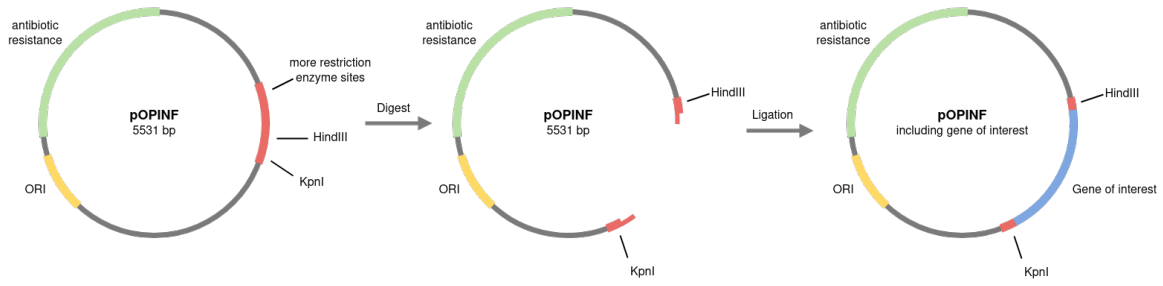

Figure 8: Cloning a certain gene of interest from plant transcriptome into pOPINF vector using KpnI-HF and HindIII-HF restriction enzymes.

3.3 Cloning

The first step into the experimental work was to clone the defined genes into a cloning vector. Therefore the genes were cloned into pOPINF cloning vector [5]. The infusion

reaction consists of 50 ng pOPINF, 100 ng template, 1 µL In-Fusion enzyme (Takara) and dH_2O up to a total reaction volume of 5 µL. The reaction was incubated at 50 °C for 15 min and then cooled on ice. The vector was transferred into *Escherichia coli* Top10 (*E. coli* Top10) during transformation. Therefore 2 µL of the In-Fusion reaction were added to 40 µL bacteria and incubated on ice for 20 min. Then the bacteria were heat shocked at 42 °C for exactly 40 s and cooled down on ice again for 5 min. 500 µL of LB-Media was added and the bacteria incubated at 37 °C for one hour. A volume of 100 µL was plated on a LB-Agar + Carb plate. Plates were stored over night in a 37 °C incubator. Next day a colony PCR using Phire enzyme approved the successful transformation of pOPINF into bacteria. For plasmid extraction, 10 mL LB-Media + Carb were set up for positive cultures and kept growing in a 37 °C shaker over night. Extraction was performed using the Promega-DNA-Purification-System kit. 250 ng DNA and 25 µM pOPINF sequencing primers (forward and reverse individual) were send for sequencing (by GENEWIZ, from Azenta Life Sciences). The sequencing results were used to identify the correct construct in frame of the cloning cassette. Therefore the forward and reverse sequencing of the same construct were assembled using the Geneious software. The sequenced construct was aligned to the original sequence and checked for errors. In order to express and purify the protein, the construct had to be transformed into *E. coli* BL21. (Method based on Invitrogen™ OneShot BL21(DE3) chemically competent *E. coli* user manual) This bacteria strain is designed for gene expression particularly since it can be induced with Isopropyl β-D-1-thiogalactopyranoside 2 (ITPG). Also the strain possesses the T7 RNA polymerase which is faster than the natural *E. coli* polymerase. Therefore the extracted construct, which was send for sequencing, was used instead of the infusion-reaction and the transformation was done as described before.

3.4 Site directed mutagenesis

Overlap extension PCR was used to introduce the point mutation into genes of interest [20]. Point mutations were introduced into CrRedox2 and TiRedox2.3 gene. Primers were designed using Geneious Prime. The reverse primer started at and included the point mutation. The annealing temperature of the primers was aimed for 55 °C Together with the outside primers of the enzyme, fragment A and B were first generated in a separate polymerase chain reaction (PCR). The reaction consisted of four steps and

consisted of 0.5 ng template, 0.5 µM each primer forward and reverse, 12.5 µQ5-hot start HF polymerase and water up to a volume of 25 µL. The initial denaturation of the cDNA was achieved by a temperature of 98 °C for 30 s plus 12 s. The now single stranded DNA was kept at the annealing temperature of 60 °C for 25 s so that primers could bind. The extending and replicating of the new fragment was done at the specific extending temperature of the polymerase, here 72 °C, for 1 min. The cycle of short (12 s) denaturation, annealing and extending was repeated 35 times. Last, a final extending is reached by 2 min at 72 °C. The reactions is cooled down to 4 degrees before continuing. In a third PCR the products of the A and B PCR were assembled. Therefore the PCR conditions were adjusted as following: 1 ng of both fragments, A and B, were added to and 12.5 µL Q5 hot start polymerase and dH_2O up to a reaction volume of 25 µL. The reaction was heat to 98 °C for 30 s. Then a cycle of 98 °C for 12 s, 60 °C for 30 s, 72 °C for 60 s, was repeated five times. The sample was cooled down to 25 °C for 5 min so that the outside primers could be added (1.25 µL each forward and reverse). Then the PCR could be continued with a 30 times cycle of 98 °C for 12 s, 60 °C for 30 s, 72 °C for 60 s and finalised with 72 °C for 2 minutes following by a cool down to 10 °C. The sample was then run through a 1 % agarose gel at 120 V, 400 A for 30 min and then extracted from the gel using Zymoclean Gel DNA recovery kit. The DNA concentration was measured with a NanoDrop and the final DNA product was transformed into bacteria as described before (first *E. coli* Top10, sequencing, then *E. coli* BL21). Bacteria with the correct constructs were stored at -80 °C as stock cultures. The stock was made from 500 µL of 10 mL overnight cultures and adding 1 mL of 50 % glyercol.

3.5 Protein purification

3.5.1 Bench top purification

The proteins were first purified in a small scale purification [10]. This way there activity and quality could be tested before doing a larger scale purification. First the bacteria, harbour the gene of interest, were taken from the glycerol stock and grown over night in 10 mL LB-Media + Carb. Then 1 mL of the culture was added to 100 mL 2xYT + Carb and incubated at 37 °C for about 6 hours until the cell density reached OD = 0.6. Protein expression was induced by adding IPTG to a final concentration of 0.3 µM.

The cultures were incubated at 18 °C for 17 hours. Then they were spinned in the centrifuge at 4000 x g for 15 min. The centrifuge was precooled to 4 °C. The supernatant was discarded and the cells resuspended with 10 mL A1+ buffer in total. Using the ultrasonic bath (2 s on, 3 s off, 4 min in total) the cells were lysed and then again spinned down at 35000 x g for 30 min. 400 µL of washed nickle beats sitting in A1 were added (500 µL) to the pure supernatant, containing the proteins. The tubes kept orbital shaking for 90 min at 4 °C. Then the tubes were spinned for 2 min at 1000 x g. The A1 containing all unbind proteins was discarded. 10 mL of A1 Buffer was added, mixed properly and spinned again for 2 min at 1000 x g. This washing step was repeated three times. The nickle beats were then incubated in 300 µL B1 buffer and spinned for 3 min at 1250 x g. This step was repeated once and each time 300 µL were pipetted into a new eppendorf tube. The proteins sitting in B1 were then buffer exchanged to A4 and concentrated using a 10 k cut-off filter tube. Last the enzyme absorbance A (1 A = 1 mg/mL) was measured using the NanoDrop and set of against the extinction coefficient ε using the Lambert-Beer-Law (equation 1) (lenght (l) in cm, extinction coefficient ε in L$\cdot$ mol$^{-1}\cdot$ cm^{-1}, concentration c in mol/L)

$$A = \epsilon \cdot c \cdot l \tag{1}$$

3.5.2 Fast protein liquid chromatography

Fast protein liquid chromatography (FPLC) is a technique to purify large amounts of clean protein. This high purity is needed for crystallisation. Using the ÄKTA purification system cultures could be purified in a one or two step purification with adjusted methods [10]. Therefore 1 L culture in 2xYT + carbicillin were started as and handled as described before with small cultures. A method was programmed at the local software and reused for all protein purification. Therefore especially the line pressure and the elution cut off had to be defined very precisely to adjust to the columns. The proteins were purified by affinity chromatography using a HisTrap first and with a size exclusion chromatography (SEC) column second. With this two step purification method, it was made sure that only for one thing proteins with His6-tag and for another thing proteins at the correct size were eluded into the collection plate. Again the protein absorbance was measured with the NanoDrop and set of against the extinction

coefficient to determine the concentration of the protein using the Lambert-Beer-Law.

3.6 SDS-PAGE

The purity and size of a protein was verified by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE). 5 µM protein was added to 5 µL TruPAGE buffer and water was added up to a total volume of 20 µL. The reaction was boiled for 10 min at 95 °C so that the secondary and tertiary structure of the protein was dissolved and the sodium dodecyl sulfate (SDS) would bind to the amino acid strain. After cooling down, 20 µL of the protein and 3 µL of the ladder were added to the electrophoresis gel. The gel run at 150 V, 400 A for one hour. Then buffer containing coomassie brilliant blue was added to a petri dish and the gel was kept shaking until the strong bands could be seen (around 1-2 hours). The gel was washed a few times with water.

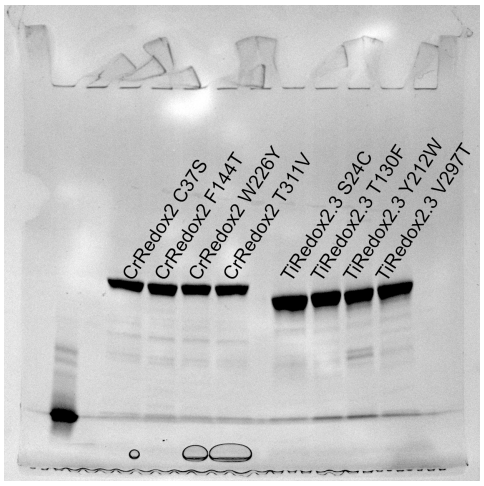

Figure 9: SDS page of all single mutants of CrRedox2 and TiRedox2.3 vice versa. The full ladder is missing.

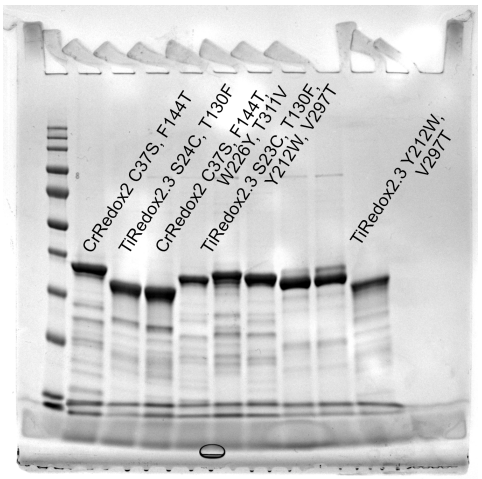

Figure 10: SDS page of remaining mutants of CrRedox2 and TiRedox2.3 additional proteins not used in this study.

3.7 Enzymatic assays

3.7.1 Conditions

One step reactions Different buffers were used for the same reaction at different pH: MES for pH 6.5, HEPES for pH 7.5, CHES for pH 8.5 and pH 9.5. As negative control, the reaction was set up with water instead of protein. First the enzymes were diluted to a 20 µM concentration in A4 buffer. The reaction then consisted of water, 50 mM buffer,

1 mM NADPH as cofactor, 1 µM enzyme and 20 µM substrate 19*E*-geissoschizine. The reaction had a total reaction volume of 100 µL. Reactions were incubated for one hour at 30 °C and then quenched with equal volume of 70 % methanol + 1 % formic acid (100 µL). And spinned for 10 min at 15000 rpm at room temperature. 50 µL were taken from the top layer and added into a glass vial with insert and analysed via LC-MS.

Coupled reactions Here 19*E*-geissochizine was used as a substrate again but two more enzymes were added: geissoschizine oxidase (GO) as well as RedOx1. This way three steps of the pathway were completed and started with a stable intermediate, since the actual substrate for RedOx2 to produce stemmadenine is not known. Again the reaction consisted of water, 50 mM buffer, 1 mM NADPH and 20 µM substrate. Three enzymes were added each 1 µM. The soluble enzymes RedOx1 and RedOx2 were added first, GO was added after. The substrate was added last. The total reaction volume was 100 µL and then treated as described before. As positive control, to check GO activity, another reaction without RedOx1 and RedOx2 was set up.

3.7.2 Liquid chromatography–mass spectrometry

Enzyme assays were analysed using the UHPLC (Ultra High Pressure Liquid Chromatography) system. A method specialised for MIA compounds was used as basis [16]. The machine was connected to Impact II UHR-Q-ToF (Ultra High Resolution Quadrople-Time-of-Flight) mass spectrometer (by Bruker Daltonics). The MS acquisition was performed in positive electrospray ionisation mode (+ESI). The capillary voltage was set to 4500 V and the end plate set to 500 V. Nitrogen was used as drying gas with a flow of 11 L/min at 250 °C and a nebulizer pressure of 2.5 bar. Acquisition was done at 12 Hz following a mass range from 80 to 1000 m/z where the sample is expected to elute. The acquisition had a data dependent MS/MS and an active exclusion window of 0.2 min as well as a reconsideration threshold of 1.8-fold change. Fragmentation of the scaffold was triggered on an absolute threshold of 400 and limited to a total cycle time range of 0.5 s. The stepping model was used for collision energy at a range of 20 to 50 eV. At the beginning of each sample run, a sodium formate-isopropanol calibration solution was infused into the source at 0.18 ml/h using a syringe pump. That way the MS spectra could be calibrated.

For the UHPLC two different methods were used. One in acidic and one in basic gradient. Acidic conditions were used first to check if the enzymes were active. Geissoschizine can not be seen as a sharp peak in the spectrum of acidic HPLC. Therefore basic condition were needed to elute geissoschizine and evaluate the substrate uptake of each protein. Data were analysed and chromatochrams generated using the Bruker Compass DataAnalysis software.

Acid gradient 2 µL of each sample was injected to an UHPC Phenomenex Kinetex XB-C18 column at 40 °C. Analytes were eluted with a gradient of running buffer. Solvent A was acidified water (0.1 % formic acid) and solvent B (mobile phase) was 100 % acetonitrile. The flow rate was 0.6 mL/min. The gradient started at 10 % of solvent B and over a time of 6 min linearly to 30 % solvent B. Followed by 1.5 min coloumn wash with 100 % solvent B and re-equilibration to 10 % B for 2.5 min.

Basic gradient Again 2 µL of each sample was injected onto an UHPC Waters BEH C18 column at 50 °C. Analytes were eluted with a gradient of running buffer. Solvent A was basified water (0.5 % v/v of 25 % NH_4OH and solvent B (mobile phase) was 100 % acetonitrile. The flow rate was 0.6 mL/min. The gradient started at 5 % of solvent B and over a time of 18 min linearly to 65 % solvent B. Followed by 2 min of coloumn wash at 95 % solvent B and a re-equilibration to 5 % B for 3 min.

3.8 Crystallisation

Both TiRedox2 and CrRedox2 were attempted to be crystallised. The Oryx8, protein crystallisation robot, was used for the vapor diffusion sitting drop method. First the different screening conditions were transferred from the JBScreen PEG (and others as named below) into the reservoir wells of the plate. Then the purified protein was concentrated to 8 mg/mL with a 10K-filter. The crystallisation robot pipetted the buffer from the reservoir into the drop chamber. Then enzyme and cofactor (1 mM NADP+) were added separately. The first drop chamber had a ratio of 1:1 of buffer and enzyme. The second chamber had a 2:1 ratio. The plate was sealed after dispersion and incubated in a non-vibration fridge at 20 °C. After one week a few crystals could be seen. The specific well conditions they grew in were used for a optimisation plate.

The original plate was kept. The crystals at the optimisation screen didn't have high quality. However, more crystals grew at the original screen after one month which could be harvested and sent to the Max-Planck-Institute of Biochemistry, Martinsried. There, the crystals were analysed using X-Ray diffraction method. The principle of the whole process is shown in figure 11. The crystal diffracting in best quality grew in conditions: 20 % PEG 3350, 200 mM lithium chloride. 30 % ethylene glycol was used as cryoprotectant.

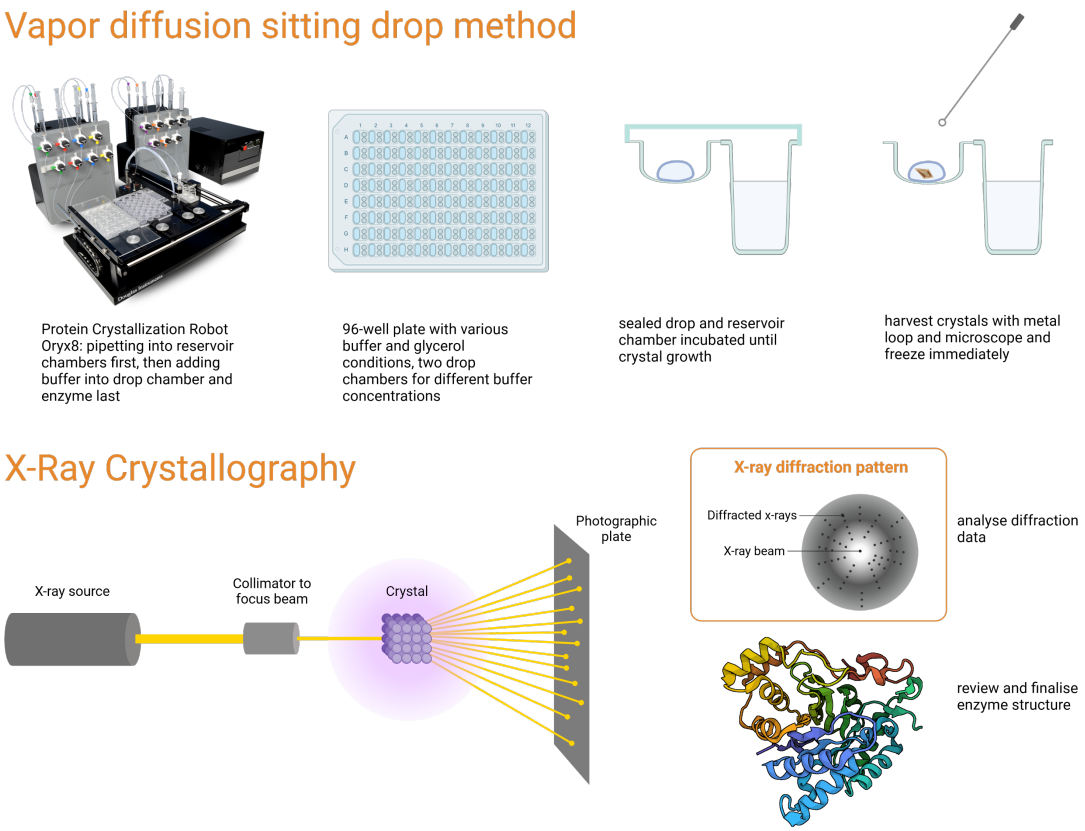

Figure 11: The process of enzyme crystallisation using the vapor diffusion sitting drop method. X-Ray diffraction and structure refinement was done by Jerome Basquin at Max-Planck-Institute for Biochemistry Muinch-Martinsried. Figure based on [7].

At the following two pictures (Figure 12, 13) crystals in the drop chamber can be seen. Clearly visible they are very different in several kinds. A coloured crystal is an indication that it's a protein and not a salt crystal. Furthermore the crystal should have a clear, intact (not scratched) surface and sharp edges. The crystal in figure 13 pointing horizontally is a good example four a rough surface. If the crystal is too thin, too small or consists of several stacked crystals it's not suitable for X-Ray diffraction.

It will break while harvesting or have a very bad resolution. For example the crystal in figure 12 could be identified as salt crystal via UV light. Protein crystals fluorescence under UV due to their aromatic amino acids. The crystals seen next to the coloured crystal are way to small for harvesting. After more time, bigger crystals may grow out of a cluster of smaller crystals. The smallest crystals can appear as white precipitant. Brown precipitant is an indication of the protein dying. The one crystal (vertically) in figure 13 could be harvested successfully and send into X-Ray analysis.

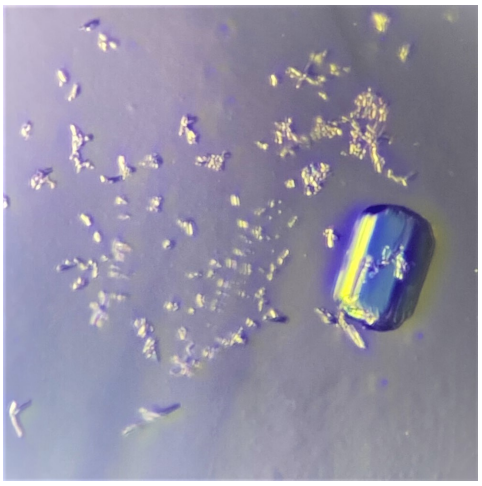

Figure 12: Salt crystal of a CrRedox2 screen. The crystal has a slightly scratched surface and is almost to large to pick up. UV-inspection indicated the crystal to contain no protein.

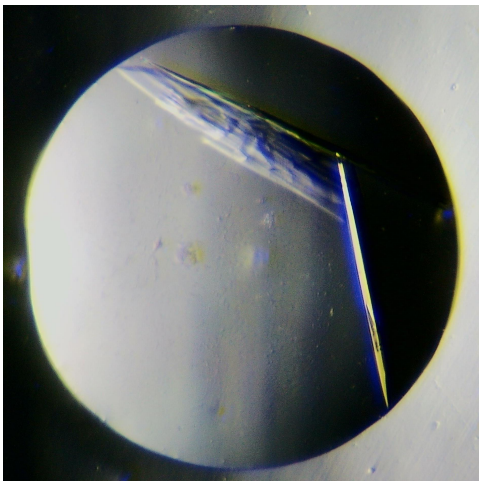

Figure 13: Protein crystal of CrRedox2 screen. The vertical crystal has a mainly clear surface. The horizontal crystal is layered and therefore not usable.

3.9 Homology Model

By aligning the amino acid sequence to the enzymes with crystal structure available, a homology model can be created. Therefore the algorithm from Swiss Model. On the website of SwissModel [49] the "User template mode" was used. The pdb file of CrRedox2 and the amino acid sequence of TiRedox2.3 was added. The predicted structure was downloaded as pdb file and opened in PyMol. PyMol is a software based on python to analyse enzyme structures visually [43].

3.10 Docking

Webina was used, a web browser integration of AutoDockVina [28]. The crystal structure and the substrate had to be uploaded as pdb. In the web application a rough

positioning of the substrate binding site could be given. The settings were adjusted to exhaustiveness: 50, number of modes: 20 and energy range: 20 (recommended by Chloe Langley). All predicted options were downloaded and opened into PyMol together with the crystal structure. Then a manual evaluation decided which docking position would make sense best. For every position the distance between the NADPH and the reaction site on the molecule was measured. All above 5 Å were discarded.

4 Results

This is the first report of enzymatic characterisation of RedOx2 homologues enzymes found in *T. iboga*. The enzymes were identified by homology based gene discovery in the published transcriptome of Iboga [16]. The enzyme CrRedox2 from *C. roseus* was used as bait to retrieve gene candidates from the Iboga transcriptome. There were three homologues found in *T. iboga* named as TiRedox2.1, TiRedox2.2 and TiRedox2.3. Their amino acid identity is shown in table 1.

Table 1: Amino acid sequence identity of all RedOx2 homologues in *C. roseus* and *T. iboga*.

Enzymes	CrRedox2	TiRedox2.1	TiRedox2.2	TiRedox2.3
CrRedox2				
TiRedox2.1	83 %			
TiRedox2.2	82 %	99 %		
TiRedox2.3	50 %	49 %	49 %	

Phylogenetic tree of the AKRs investigated in this study along with known plant AKRs are presented in figure 14. Two RedOx2 enzymes of *T. iboga* clade together with CrRedox2 (table 1). They share up to 83 % sequence identity with CrRedox2 (table 1). Meanwhile TiRedox2.3 is of interest since it shares only 50 % sequence identity, but similar sequence similarity and phylogenetic relations to plant chalcone reductases (CHRs) and codeinone reductase (COR) (figure 14, table 2). The closest enzyme to this homologue, found through enzyme alignment of AKR enzymes [25], is called AKR4B4 and known as a D-galacturonic acid reductase from *Fragaria x ananassa* (cultivated strawberry). Unfortunately no structural or catalytic information are known about this enzyme [1] which shows a sequence identity of 67% to TiRedox2.3. The amino acid sequence identity to reference enzymes is shown in table 2.

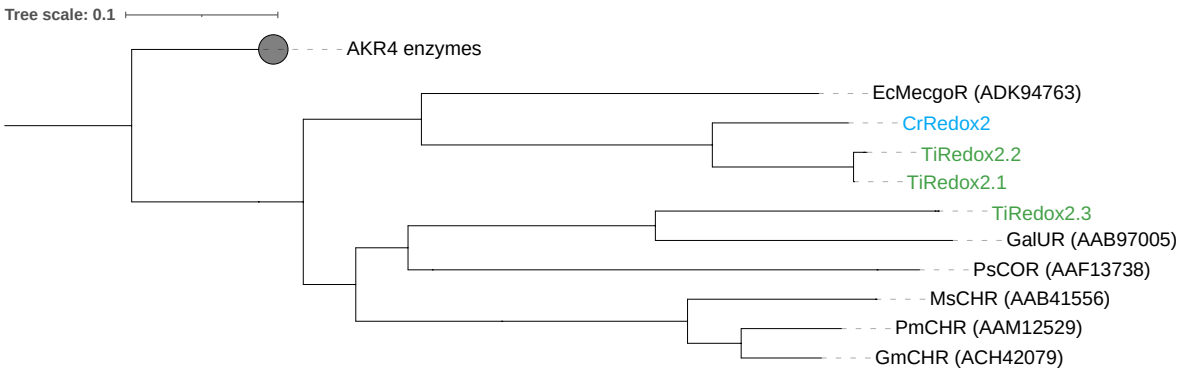

Figure 14: Phylogenetic relationship of known plant AKR enzymes with Redox2. Characterised plant AKR enzymes are presented with CrRedox2 (blue) and *T. iboga* RedOx2 homologes (green) are presented to depict the phylogenetic relationships. Genbank IDs of characterised AKRs are presented next to each enzyme. Tree was made with Geneious Prime. AKR sequences were downloaded from [25].

Table 2: Amino acid sequence identity of RedOx2 homologues and reference enzymes.

Enzymes	CrRedox2	TiRedox2.1	TiRedox2.3	MsCHR	PsCOR1.3
CrRedox2					
TiRedox2.1	83 %				
TiRedox2.3	50 %	49 %			
MsCHR	45 %	47 %	51 %		
PsCOR1.3	48 %	48 %	52 %	50 %	

4.1 Enzymatic assays

Heterologously expressed and purified enzymes were tested *in vitro* (figure 15). 19E-geissoschizine and NADPH were added to the each enzyme in each four different pH conditions starting from pH 6.5 to 9.5 in 0.5 increments. The assays were incubated for 1 hour and then analysed using LC-MS. The enzymes performed the reaction seen in figure 15 **A** resulting in 16(R/S)-isositsirikine isomers. This way the pH optimum of each enzyme should be determined. Figure 15 **B - E** shows exemplary LC-MS graphs for each enzyme in each pH condition. The substrate 19E-geissoschizine, does not chromatographically resolve in acidic mobile phase. The peak areas of the 16(R/S)-isositsirikine isomers were evaluated in figure 16. A trendline was added according to the values of 16(R)-isositsirikine. Based on the peak of the trend line the pH optimum of each enzyme could be estimated. CrRedox2, TiRedox2.1 and TiRedox2.2 show an optimum with pH 7.5 while TiRedox2.3 works most efficient with pH 8.5. Those results can give hints which residues are involved in catalysis.

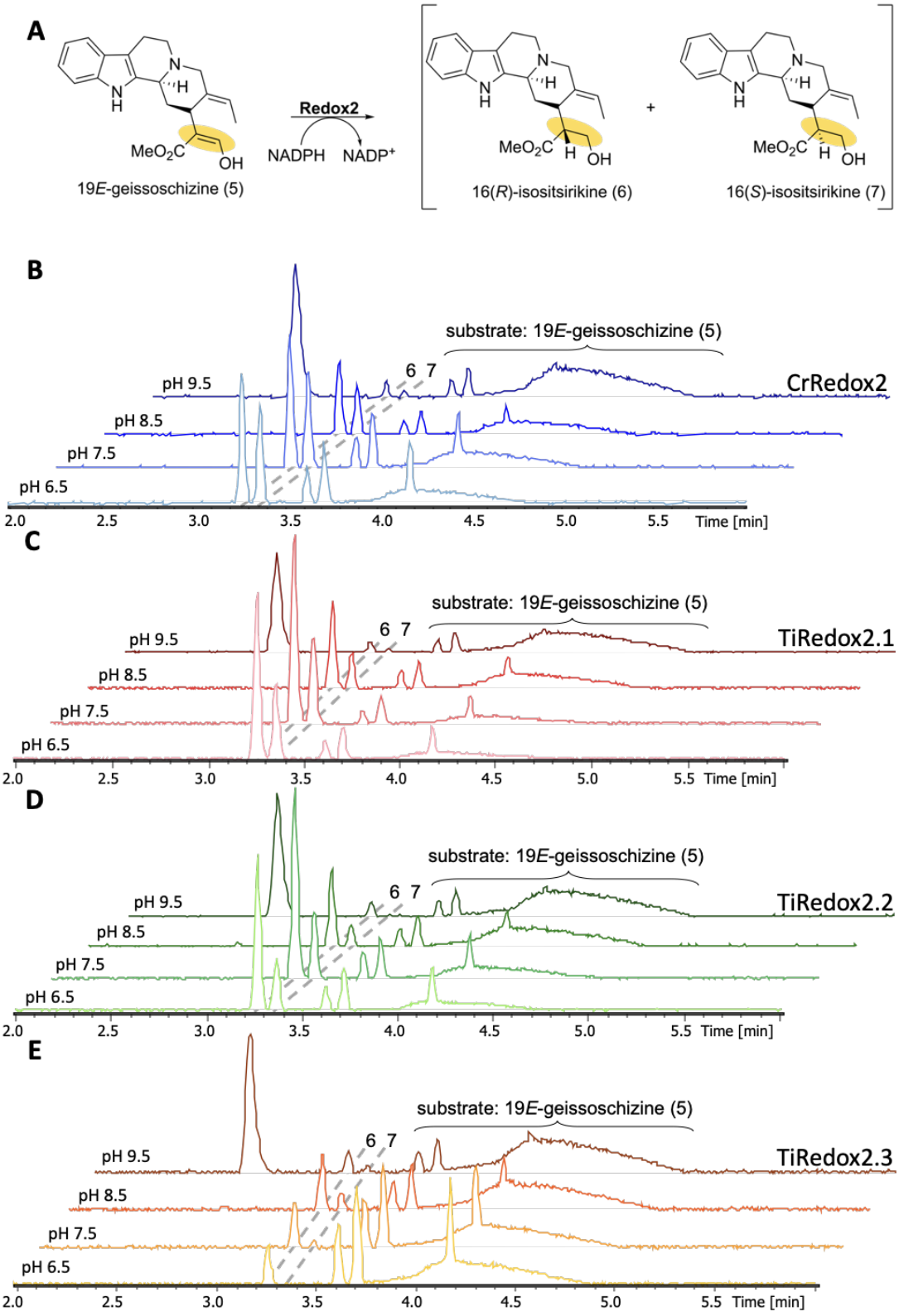

Figure 15: **A** Reaction catalysed by RedOx2 enzyme using 19E-geissoschizine as substrate and producing 16(R/S)-isositsirikine isomers. **B to E** Enzyme assays in pH conditions from 6.5 to 9.5 in 0.5 increments, analysis in acidic conditions, exemplary TIC LC-MS graphs are shown for **B** CrRedox2 **C** TiRedox2.1 **D** TiRedox2.2 **E** TiRedox2.3 where is **5** 19E-geissoschizine **6,7** 16(R/S)-isositsirikines.

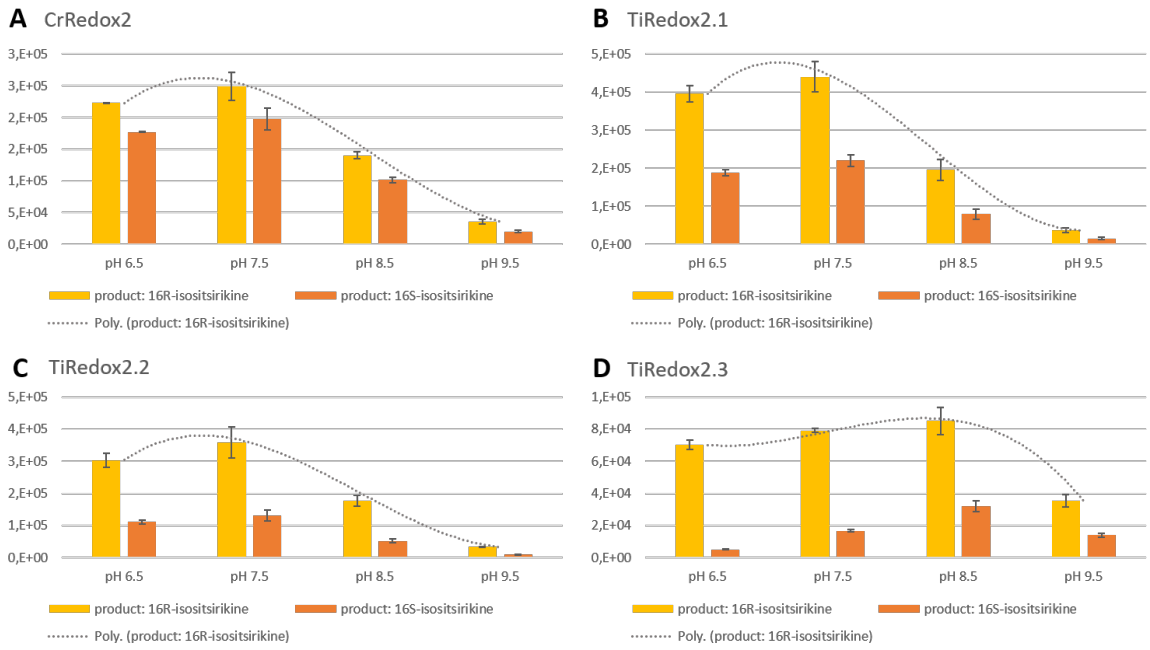

Figure 16: Product 16(R/S)-isositsirikine isomer ratio for all WT enzymes **A** CrRedox2 **B** TiRedox2.1 **C** TiRedox2.2 **D** TiRedox2.3 in different pH conditions (from pH 6.5 to 9.5 in 0.5 increments). Trend lines above 16(R)-isositsirikine described with $R^2 = 1$, 3rd degree polynomial, (n=3), error bars denote standard deviation.

4.2 Crystallisation

In order to crystallise RedOx2 enzyme, two enzymes were chosen for trial: CrRedox2 and TiRedox2.1. Commercially available screens were used where CrRedOx2 showed promising crystals. Based on these crystals, screen conditions were further optimised using individual composition of ingredients. However, more crystals grew at the commercial screen after one month, which showed good quality for X-Ray diffraction (chapter 3 figure 13). The screen was transported to collaborators at the Max-Planck-Institute for Biochemistry in Munich, Martinsried for crystal harvesting and remote data collection. Several crystals were harvested. Among the crystals picked, a promising data set was acquired to successfully solve the crystal structure. CrRedox2 crystal was solved as a dimer (figure 17 **A**), at 1.5 Å resolution bound to its inactive cofactor NADP+ as ligand figure 17 **B**) for the first time. Using this newly available CrRedox2 crystal structure, structural comparisons of reported plant AKR proteins were performed (table 3). AlphaFold was used to compare the AI based structure prediction with the solved CrRedox2 obtained experimentally by X-Ray crystallography. The RMSD value for the CrRedox2 structure predicted by AlphaFold and the actual crystal structure is 0.443 Å.

Table 3: Comparison of plant AKR protein structures and their attributes.

Protein	Species	Ligand	Chains	Resolution	PDB File	Publication	Citation
RedOx2	*Catharanthus roseus*	NADP+	Dimer	1.51 Å		2022	
CHR	*Medicago sativa*	NADP+	Dimer	1.70 Å	1ZGD	2005	[11]
COR	*Papaver somniferum*	none	Hexamer	2.40 Å	7MBF	2021	[8]

4.3 Structural comparison

Both reference enzymes, CHR [8] and COR [11], were published together with a residue-function analysis. Individual residues of the catalytic pocket were evaluated according to their impact on the carried out reaction. In the following analysis these were used as a base. For instance, no mutant was created regarding the catalytic tetrad since it was proved already, that these mutations kill enzymatic activity completely [11].

4.4 In silico analysis

4.4.1 Homology modeling

Homology modeling was performed using the online server from SwissModel [49]. Structural coordinates from CrRedox2 crystal structure were used to build homology models of TiRedox2.1, TiRedox2.2 and TiRedox2.3. The predicted structures were visually compared using PyMol [43]. The alignment gave a root-mean-square deviation (RMSD) value for every two compared structures (table 4).

Table 4: RMSD values for CrRedox2 crystal structure, CrRedox2 AlphaFold prediction and homology models of TiRedox2.1, TiRedox2.2, TiRedox2.3 based on CrRedox2 crystal structure. RMSD values were determined with PyMol.

RMSD	CrRedox2 crystal	CrRedox2 AlphaFold	TiRedox2.1	TiRedox2.2	TiRedox2.3
CrRedox2 crystal					
CrRedox2 AlphaFold	0.443				
TiRedox2.1	0.054	0.451			
TiRedox2.2	0.055	0.452	0.000		
TiRedox2.3	0.057	0.444	0.039	0.040	

4.4.2 Docking

Very rarely a protein crystallises with a ligand (substrate or cofactor) bound in the pocket of its active site. A theoretical alternative is docking. That's the prediction how a molecule could interact with the protein. Webina, a web browser integration of AutoDockVina [28] was used for docking both possible substrates, stemmadenine and

19*E*-geissoschizine, of RedOx2 into the structure of CrRedox2. All predicted docking options giving by Webina were downloaded and opened into PyMol together with the crystal structure. Then a manual evaluation decided which position would make sense best. For instance, it was not possible to define the position, where the catalysis takes place at the ligand. As result the substrate could be docked into the structure wrong way around. As optimal reaction distance 1 Å was aimed. The final positioning of the ligands into CrRedox2 can be seen in figure 17 **C** for stemmadenine and figure 17 **D** for 19*E*-geissoschizine. Based on the docking, residues could be identified that would be in close proximity to the substrate and therefore being involved in shaping the substrate pocket.

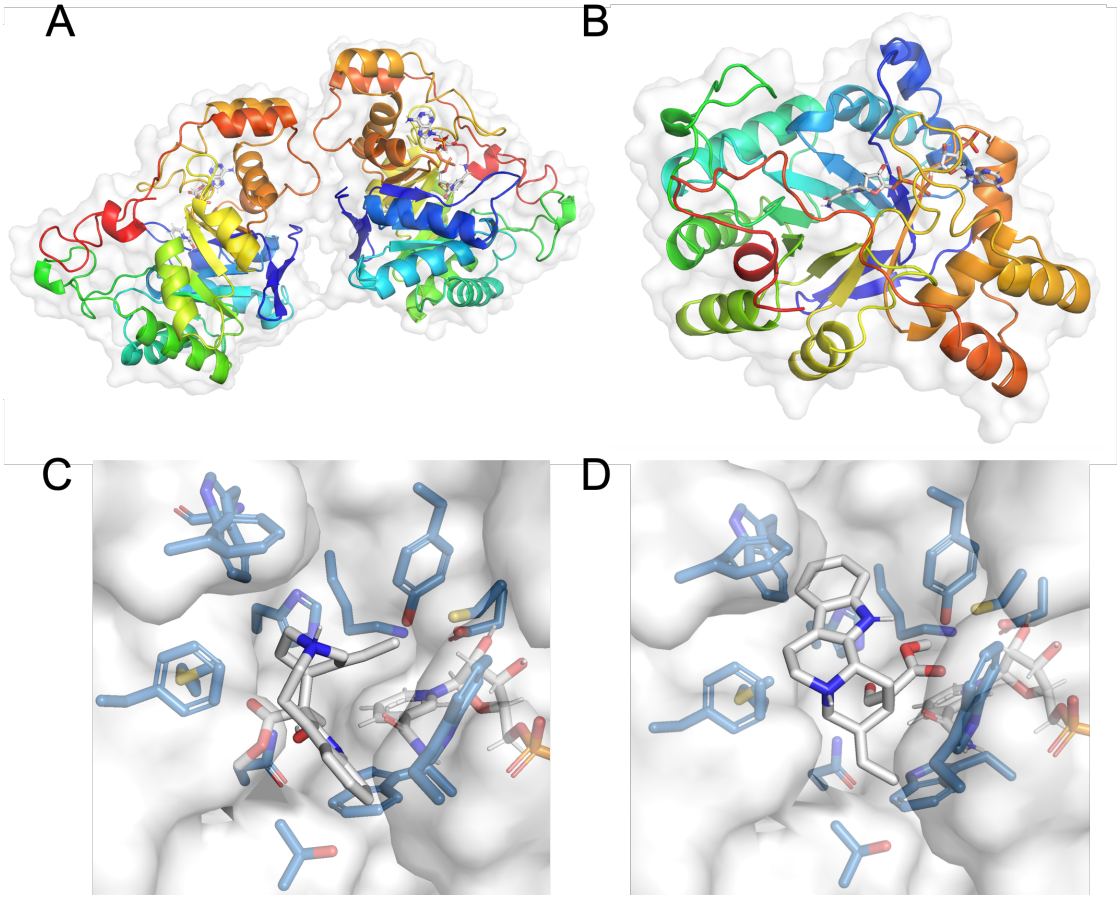

Figure 17: The crystal structure of CrRedox2 was solved using X-ray crystallography. The protein crystallised as a dimer **A** with NADP+ (grey) bound as inactive cofactor **B**. Docking was done using Webina [28] and resulted in positions, how the substrate could sit interact with the protein. Both stemmadenine **C** and 19*E*-geissoschzine **D** were docked into CrRedox2.

4.4.3 Substrate binding residues

Structural analysis information from CHR [8] and COR [46] helped analysing the structure with PyMol and identifying the residues involved in the substrate binding. The catalytic residues, building the catalytic tetrade common to all AKRs, were set as the base of the structural view as orientation. Before two different homology models were created for TiRedox2.3. One based on PsCOR structure and the other one based on CrRedox2. The models have a RMSD of 0.648 which is in the same dimension as the CrRedox2 structure and the CrRedox2 AlphaFold prediction. The model based on CrRedox2 was used. A homology model for each TiRedox2.1 and TiRedox2.2 based on the CrRedox2 structure was created and aligned but due to the fact that they were so similar to CrRedox2, an alignment could not give new information. So only the alignment of TiRedox2.3 and CrRedox2 was used. Differences in the substrate pocket were identified. In figure 21 the alignment of CrRedox2 and TiRedox2.3 can be seen. Using the structural information of CHR and COR1.3 analysis, some residues were not shortlisted for mutations since their influence on the catalysis seemed to be marginal. Also non of the catalytic tetrad residues were mutated since it has been proved before already, that these mutations result in a inactive enzyme [11].

Specifying mutations and description of positions A sequence alignment of all RedOx2 homologues, COR and CHR was made (figure 21 **C**). There, strongly conserved residues were identified with the help of the structural alignment. These were not only of catalytic but also structural nature. Most residues were conserved in CrRedox, TiRedox2.1 and TiRedox2.2. Four residues were chosen based not only on sequence alignment, but also based on structural alignment of CrRedox2 and TiRedox2.3. The first both residues sit on top or close to the entrance of the substrate pocket, while the third and fourth mutations take place deep inside the pocket with possible contact to the cofactor NADP(H). These could then have an influence on cofactor binding too. All positions in the following are labelled by amino acid structure of each protein.

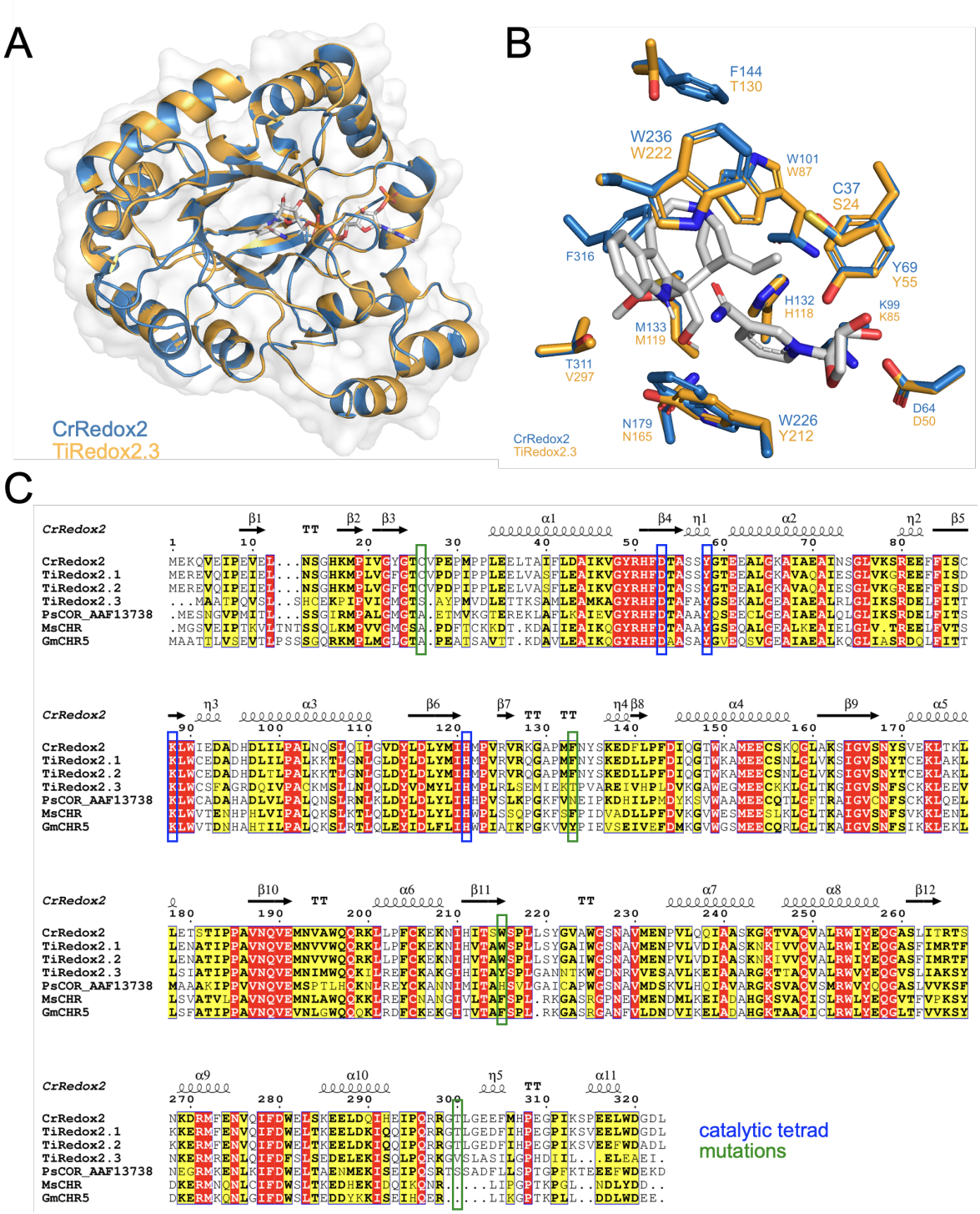

Figure 18: **A** CrRedox2 (blue) was aligned to homology model of TiRedox2.3 (orange) structure **B** Focus on the residues involved in the substrate binding based on the earlier docking of stemmadenine (grey) into CrRedox2. **C** Amino acid sequence alignment of all RedOx2 enzymes additionally to reference enzymes CHR and COR. Catalytic tetrad of AKR enzymes (blue) and chosen mutations (green) based on structural and sequence alignment. Figures made with PyMol [43] and ESPript [40].

CrRedox2 C37S The first mutation is sitting on top of the substrate pocket. In CHR there is a proline (P29), while in COR there is a glutamic acid (E28) In CrRedox2 there is a cysteine (C37) while in TiRedox2.3 there is a serine (S24). The position is described as one of the residues lining out the entrance of the catalytic pocket. The combination of M26 and E28 in COR is described to play an important role in defining the binding site. Both amino acids sit on the loop named there $\beta\,1\alpha\,1$ loop.

CrRedox2 F144T The second residue is a phenylalanine in all enzymes besides TiRedox2.3, where a threonine is found (T137). The positioning varies with the enzymes: CHR132, COR129 and CrRedox144. These residues point into the substrate-binding pocket whereas it has to be mentioned, that in COR the conformation of $\beta\,1\alpha\,1$ loop hinders the Phenylalanin from pointing to much into the pocket. Mutating this residue in COR affects the uptake of an alternative substrate [14].

CrRedox2 W226Y The third position shows an aromatic residue in every enzyme. Therefore it's likely that the position is important for shaping the substrate pocket since aromatic residues take up much more space than others. In this case the aromatic ring stacks with the nicotinamide unit of NADP+ where the reaction takes place. In CHR there is a phenylalanine (F214), while in CrRedox2 there is a tryptophan (W226) which is the largest of all amino acids. In COR there is a histidine (H213) which is the only charged option and a possible proton donor or acceptor in catalysis. In TiRedox2.3 there is a tyrosine (Y212).

CrRedox2 T311V The fourth and last residue to be mutated is continuously small in all enzymes such as an isoleucine in CHR (I298), a serine in COR (S298), a threonine in CrRedox2 (T311) and a valine in TiRedox2.3 (V297). Serine is the smallest amino acid of those mentioned. Based on the docking model, the amino acid has to be that small in order to fit in the substrate.

All mutations were created using site directed mutagenesis. First all four residues were mutated in both TiRedox2.3 and CrRedox2 to their corresponding residue. As a result there were 8 enzymes with each, one single mutation included. Second all four mutations were merged in one single enzyme with four mutations for both TiRedox2.3

and CrRedox2. Last two double mutants were created. They combined the first and second, and also the third and fourth mutation. In total 14 mutated enzymes were created.

4.5 Analysing mutations

First the influence of the mutations on the geissoschizine substrate uptake was tested. Therefore enzymes assays were screened using the same assay method and basic method for LC/MS analysis. The results are shown in figure 19. Not all mutated enzymes were available for this assay.

CrRedox2 mutants

C37S In comparison to the CrRedox2 WT, the first mutation makes no difference in catalysis.

F144T The second mutation has a big impact on catalysis. The enzymes takes up only half of the substrate and converts it. The double mutant of mutation one and two (C37S, F144T) shows the same pattern as the single mutation.

W226Y No difference in substrate uptake can be detected with this single mutation.

T311V Again there is no difference to the CrRedox2 WT enzyme regarding substrate uptake. The enzyme with a double mutation of W226Y, T311V is missing. The quadruple mutated enzyme shows the same pattern as the single mutant of F144T.

TiRedox2.3 mutants

S24C In comparison to the WT enzyme, there is no difference substrate uptake.

T130F The big error bars of the WT enzyme reaction make a comparison difficult. Still in comparison to the first mutation, the enzyme takes up less substrate. This result can not be carried forward to the double mutation of S24C, T130F.

Y212W With this single mutation the enzyme shows no activity anymore.

V297T Again the enzyme shows no activity anymore. While the double mutation enzyme is still inactive, the quadruple mutation is active again.

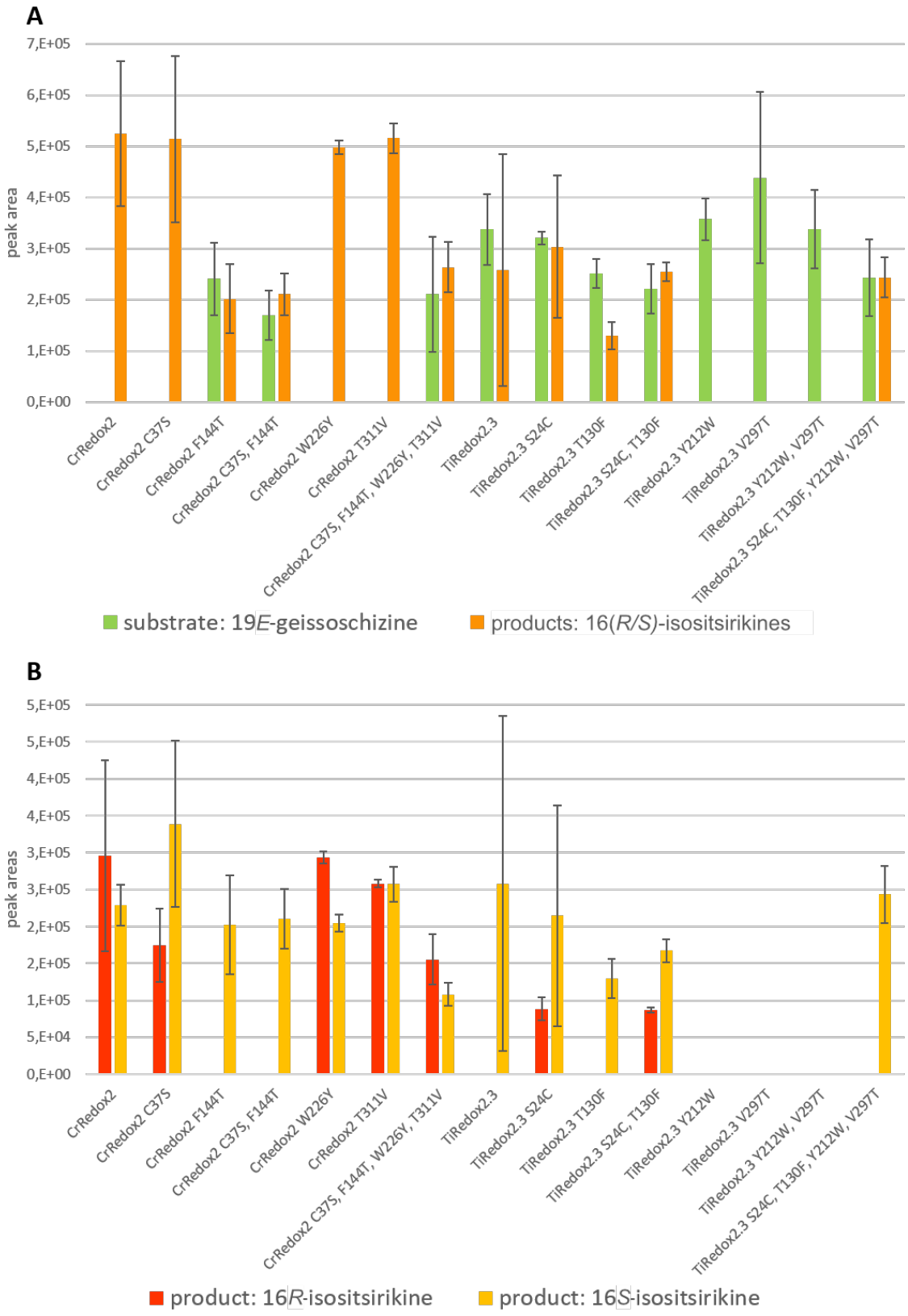

Figure 19: LC-MS peak area analysis for all mutants with 19*E*-geissoschizine as substrate, error bars denote standard deviation, (n=3). **A** substrate uptake and product formation, **B** product isomer ratio.

Table 5: Influence of CrRedox2 and TiRedox2.3 vice versa single mutations on total substrate uptake and isomer ratio based on figure 19.

Mutation CrRedox2	Influence	Mutation TiRedox2.3	Influence
C37S	none, inverted isomer ratio	S23C	none, sudden 16R-isositsirikine production
F144T	less substrate uptake, no 16R-isositsirikine production	T130F	none
W226Y	none	Y212W	inactive
T311V	none	V297T	inactive

After the mutants were tested in a coupled assay including the upstream enzymes CrGO and CrRedox1 (figure 20 **A**). This method was used since the substrate for RedOx2, dehydro-stemmadenine is instable and can not be isolated. Here 19E-geissoschizine was used as substrate again. Geissoschizine is the substrate of geissoschizine oxidase (GO). This assay has the disadvantage that GO is a membrane bound enzyme, which makes it difficult to determine the actual concentration. On the other hand RedOx1 is a p450 enzyme and easy to purify using batch purification. GO and RedOx2 compete for the substrate. As result not only the main products but also sink products can be seen in analysis. The coupled reactions were screened in acidic conditions. In figures 20 the graphs can be seen. In figure 20 **B** the WT enzymes are compared to each other. Clearly CrRedox2 and TiRedox2.1 produce a much higher amount of stemmadenine than TiRedox2.3. The third homologue shows less stemmadenine production and no isositsirikine production at all. Figure 20 **C** shows exemplary LC-MS graphs for CrRedox2 and it's mutants while figure 20 **D** shows the same for TiRedox2.3 and it's mutants.

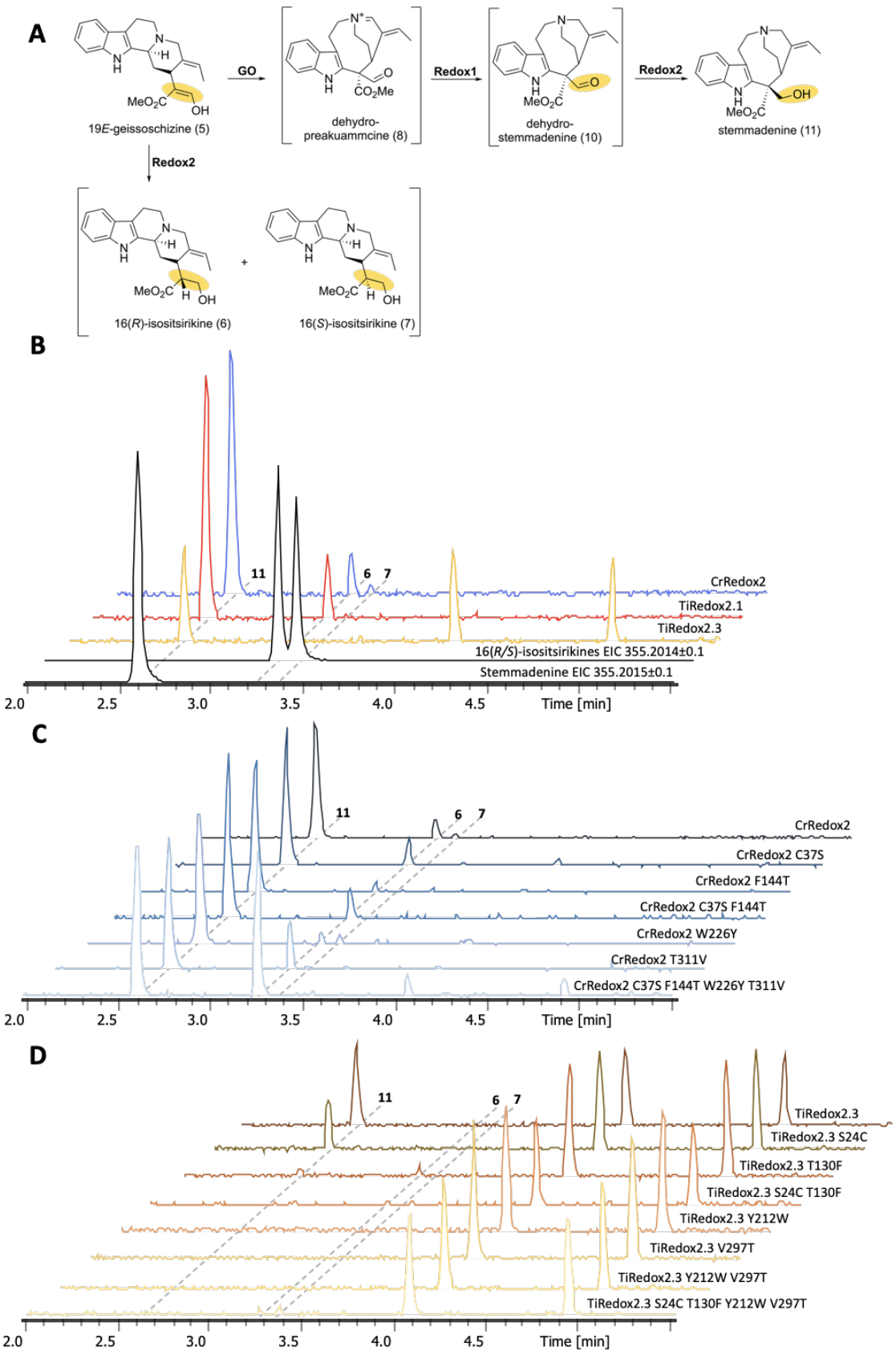

Figure 20: Coupled reaction including GO, RedOx1 and RedOx2 beginning with 19*E*-geissoschizine and resulting in stemmadenine, as well as sink products 16(*R/S*)-isositsirikine **A**. **B** showing stemmadenine and 16(*R/S*)-isositsirikine standards together with RedOx2 WT enzymes. **C** CrRedox2 WT and CrRedox2 mutants **D** TiRedox2.3 WT and TiRedox2.3 mutants with **11** 355 m/z Stemmadenine, **6,7** 355 m/z 16(*R/S*)-isositsirikines.

5 Discussion

High efforts were made revealing all catalytic steps and associated enzymes of the vinblastine pathway over a period of many years. The relatively long pathway depends on the efficiency of each step in order to produce a proper amount of the end product. Parts of the intermediates get lost as sink products. Stemmadenine is a key intermediate for important MIA scaffolds, representing the end of the so called early MIA pathway. The final step is catalysed by the enzyme RedOx2.

In this thesis the RedOx2 enzyme and its catalytic activity was addressed. Uncharacterised RedOx2 homologues from *T. iboga* were compared with a characterised RedOx2 enzyme from a closely related plant species *C. roseus*. This way the function of the true RedOx2 of *T. iboga* should be determined. The results stated here indicate, that the enzymes TiRedox2.1 and TiRedox2.2. represent the actual RedOx2 homologues. Meanwhile TiRedox2.3, which shares only 50 percent amino acid sequence identity, represents an enzyme, whose actual reaction remains unknown.

The RedOx2 enzymes of *C. roseus* and *T. iboga* were compared to each other by testing their ability to form isositsirikines with the substrate 19E-geissoschizine at different pH conditions. As result the differences between TiRedox2.3 and the other RedOx2 enzymes stand out. In enzymatic assays CrRedox2, TiRedox2.1 and TiRedox2.2 show similar peak heights for isositsirikines in LC-MS diagrams, while the peaks of TiRedox2.3 isositsirikines remain much lower (figure 16).

Protein crystals were send for X-Ray diffraction and revealed the structure of CrRedox2 bound with NADP+ at a resolution of 1.5 Å. The structure allowed the creation of precise homology models of TiRedox2 enzymes. Structural alignment, as well as docking the ligands stemmadenine and 19E-geissoschizine, were used to explore and review changes in the active site between RedOx2 enzymes. Also, CHR and COR structure, well studied plant AKR's, were used as reference to hypothesise the function of individual residues of the active site. For instance, based on this previous studies, non of the catalytic tetrad residues were targeted, since their mutation leading to inactive enzymes has been proved already [11]. This *in silico* analysis lead to four residues involved in

substrate binding, which were then targeted for site directed mutagenesis by creating mutations of CrRedox2 and TiRedox2.3 vice versa (figure 21).

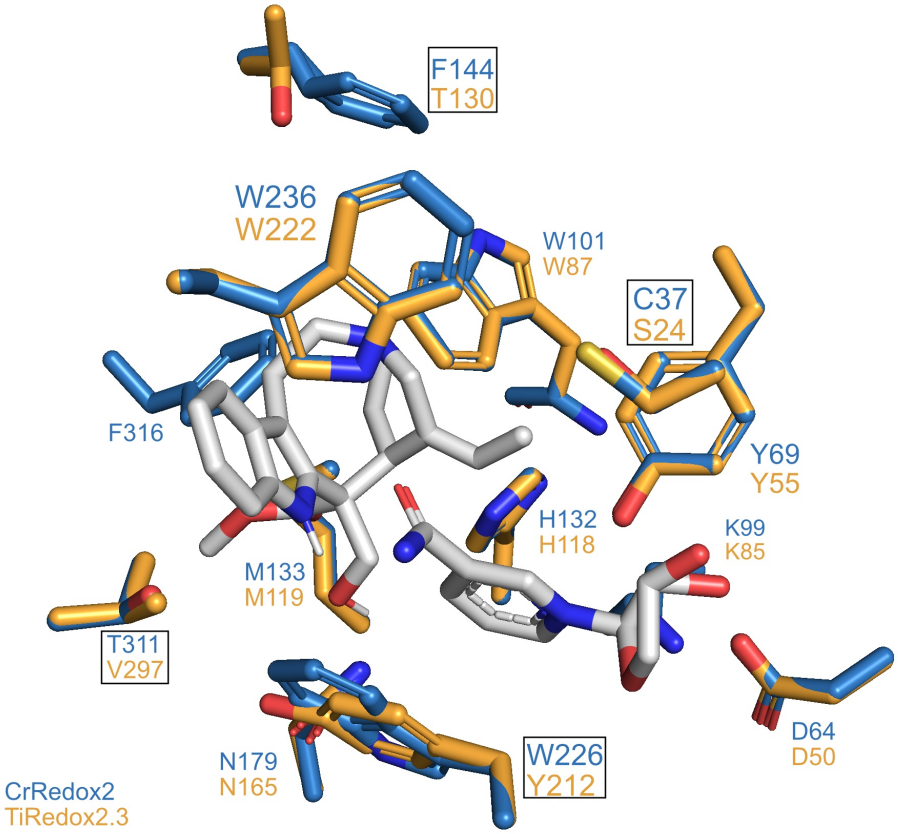

Figure 21: Residues involved in substrate binding inside the RedOx2 enzyme. Mutations in CrRedox2 (blue) and TiRedox2.3 (orange) enzyme vice versa in black boxes. Figure created with PyMol [43].

Mutated enzymes were purified and then again tested on 19E-geissoschizine substrate resulting in the production of 16(R/S)-isositsirikines. Differences appearing here, lead to the influence of single mutations on catalysis.

CrRedox2 F114T decreased the substrate uptake noticeably and hindered the production of 16(R)-isositsirikine. This result could be proven by double and quadruple mutants showing the same pattern. Since Phenylalanine is considerably larger amino acid than threonine, big aromatic chain of phenylalanine could hold 19E-geissoschizine in place while catalysis. The large space offered by the mutation could lead to an instable positioning of the substrate. Meanwhile the TiRedox2.3 T130F mutation did not show an increase as expected, but a slight decrease of product formation too. Here the larger residue added could have prevent the substrate from entering the active site.

The second mutation did not have an influence on the substrate uptake but only on isomer ratio. The exchange of cysteine to serine in CrRedox2 (C37S) results in a higher production of 16(S)-isositsirikine. While cysteine offers a thiol group, serine contains a hydroxy group. Oxygen is more electronegative and can act as stronger H-bond donor [31] which could stabilise the substrate stronger and force the production of 16(S)-isositsirikine. The double mutant of TiRedox2.3 S24C, T13F showed the same behaviour. However, in the double mutant of CrRedox2, influence of F144T mutation was more decisive so that the same pattern as CrRedox2 C37S could not be observed. The results of the first both mutations lead to the assumption that both residues are involved in substrate binding and orientation inside the catalytic pocket.

The third and fourth single mutations, siting inside the substrate pocket, namely CrRedox2 W226Y, T311V and TiRedox2.3 Y212W, V297T showed either no change in activity (CrRedox2) or total loss of activity (TiRedox2.3) for the production of 16(R/S)-isositsirikines. Even though it has to be mentioned that no circular dichroism (CD) spectrum was done with those mutants. Therefore, it cannot be ruled out that the results obtained from the coupled reactions with the mutated enzymes was due to the inactive Redox2 because of missfolding or the effect of the mutation causing loss of specificity for Redox1 product. Another explanation for the inactiveness of the enzymes could be steric hinderance due to the architecture of the catalytic pocket of TiRedox2.3. This explanation could fit to the fact, that the combination of four mutations based on the CrRedox2 pocket then again create an enzyme, that is active again. More detail information explaining those data could be achieved by investigating the interaction of residues themselves, which was not done in this study.

Then all available enzymes were tested in combination with GO and RedOx1 enzyme, using 19E-geissoschizine as substrate again. RedOx1 provides the unstable intermediate which acts as substrate for RedOx2 enzyme to convert the final product stemmadenine, a stable intermediate in MIA biosynthesis. Unfortunately no qualitative increase of stemmadenine production could be observed. Though the quadruple mutant of CrRedox2 lead to a large increase in the sink product 16(R)-isositsirikine. The mutants of

TiRedox2.3 did not show any stemmadenine production besides TiRedox2.3 S24C. Peaks with a 325 m/z (most probably belonging to 16(R/S)-deshydroxymethyl stemmadenine) were found instead, which most probably originated through the spontaneous conversion of the instable RedOx1 product [38]. Since the experiment did not give the expected result of various amount of stemmadenine production, it can not be concluded if and how single residue mutations influence the production of stemmmadenine.

The TiRedox2.3 model, structural analysis based on, was a homology model based on CrRedox2 structure. It might be that the amino acid sequence identity of 50 % allows enough opportunities for a slightly different architecture and orientation of the catalytic loops in TiRedox2.3. The homology model of TiRedox2.3 was missing the first 18 residues at the N-terminus. Based on the structure of CrRedox2, there is at least one more residue (CrRedox2 F316) which could have an influence on the catalytic pocket. Even so, without a proper homology model the corresponding residues could not be identified through sequence alignment. This is influenced by the fact that TiRedox2.3 contains mainly smaller amino acids (e.g. isoleucine, glycine) while in CrRedox2 there are bigger ones (e.g. phenylalanine, histidine) in the region of the N-terminus, which shift the whole loop in the tertiary structure of the enzyme.

5.1 Further research

In order to further validate and confirm the results stated here, more experiments are necessary and can help to clarify discrepancies in the results obtained. Since one of the last datasets gave unexpected results, the intactness of the used proteins should be proved (regarding the possibility of miss-folding) and the experiment repeated. Additionally the most important experiment to follow is the analysis of TiRedox2.3. Therefore the enzyme assays should be repeated in higher pH conditions. Together with NADP+ as cofactor and isositsirikines as substrate, it could be that the enzyme catalyses the backwards reaction. The ability of AKR enzymes to use both NADPH and NADP+ as cofactor has been shown before already [11]. Repeating the same process for coupled reaction including GO and RedOx1, a significant increase of stemmadenine production could possibly be seen.

To investigate the activity of RedOx2 enzyme *in vivo*, a study could be designed to transiently introduce these enzymes in a heterologous host and test their activity towards the production of stemmadenine. To complete knowledge about the RedOx2 enzyme, further investigation is needed to reveal the structure of the RedOx1 product, which is the RedOx2 substrate. Another interesting perspective would be the investigation of the phylogenetic relationship of RedOx2 to COR and CHR. Eventually it could be reproduced how and where from RedOx2 evolved, giving more information how the pathway of stemmadenine itself was formed.

5.2 Conclusion

By the site specific enzyme engineering of RedOx2 enzyme, the undesired affinity towards 19E-geissoschizine could be modified, so that the intermediate will be almost exclusive for the enzyme GO. The GO enzyme will then lead the pathway towards the production of stemmadenine. This achievement will help to engineer the overproduction of stemmadenine in heterologous hosts and make it a more easily accessible intermediate. The residues involved in substrate recognition could be identified as CrRedox2 C37 and F144. As described earlier with CHR and COR enzyme [8],[11], and now also in case of RedOx2 residues sitting on top of the catalytic pocket are very important for substrate recognition. Though, non of the of the mutations resulted in a significant difference in both sink product formation or stemmadenine production. The structure of CrRedox2 will line up to the structures of chalcone reductase (CHR) and codeinone reductase (COR) as representative of important plant aldo keto reductases. However, this is the first reported RedOx2 crystal structure in stemmadenine synthesis.

References

[1] F. Agius, R. González-Lamothe, J. L. Caballero, J. Muñoz-Blanco, M. A. Botella, and V. Valpuesta. Engineering increased vitamin c levels in plants by overexpression of a d-galacturonic acid reductase. *Nature Biotechnology*, 21:177–181, 2003.

[2] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. Basic local alignment search tool. *Journal of Molecular Biology*, 215:403–410, 1990.

[3] U. Anand, S. Nandy, A. Mundhra, N. Das, D. K. Pandey, and A. Dey. A review on antimicrobial botanicals, phytochemicals and natural resistance modifying agents from apocynaceae family: Possible therapeutic approaches against multidrug resistance in pathogenic microorganisms. *Drug Resistance Updates*, page 100695, 2020.

[4] D. Banner, A. Bloomer, G. Petsko, D. Phillips, and I. Wilson. Atomic coordinates for triose phosphate isomerase from chicken muscle. *Biochemical and Biophysical Research Communications*, 72:146–155, 1976.

[5] N. S. Berrow, D. Alderton, S. Sainsbury, J. Nettleship, R. Assenberg, N. Rahman, D. I. Stuart, and R. J. Owens. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Research*, 35:e45, 2007.

[6] Biomatters. Geneious prime, 2021. Available at https://www.geneious.com, version 2021.1.1.

[7] BioRender. X-ray crystallography, 2022. Available at https://app.biorender.com/biorender-templates/figures/5c95180bc2753f33003dacd3/t-60c774bef689dc00aad87e7e-x-ray-crystallography.

[8] E. K. Bomati, M. B. Austin, M. E. Bowman, R. A. Dixon, and J. P. Noel. Structural elucidation of chalcone reductase and implications for deoxychalcone biosynthesis. *Journal of Biological Chemistry*, 280:30496–30503, 2005.

[9] S. Brown, M. Clastre, V. Courdavault, and S. E. O'Connor. De novo production of the plant-derived alkaloid strictosidine in yeast. *Proceedings of the National Academy of Sciences*, 112:3205–3210, 2015.

[10] L. Caputi, J. Franke, S. C. Farrow, K. Chung, R. M. Payne, T.-D. Nguyen, T.-T. T. Dang, I. S. T. Carqueijeiro, K. Koudounas, T. D. de Bernonville, et al. Missing enzymes in the biosynthesis of the anticancer drug vinblastine in madagascar periwinkle. *Science*, 360:1235–1239, 2018.

[11] S. C. Carr, M. A. Torres, J. S. Morris, P. J. Facchini, and K. K. Ng. Structural studies of codeinone reductase reveal novel insights into aldo-keto reductase function in benzylisoquinoline alkaloid biosynthesis. *Journal of Biological Chemistry*, 297, 2021.

[12] V. Courdavault, N. Papon, M. Clastre, N. Giglioli-Guivarc'h, B. St-Pierre, and V. Burlat. A look inside an alkaloid multisite plant: the catharanthus logistics. *Current Opinion in Plant Biology*, 19:43–50, 2014.

[13] J.-F. Couture, P. Legrand, L. Cantin, V. Luu-The, F. Labrie, and R. Breton. Human 20α–hydroxysteroid dehydrogenase: Crystallographic and site-directed mutagenesis studies lead to the identification of an alternative binding site for c21-steroids. *Journal of Molecular Biology*, 331:593–604, 2003.

[14] M. Dastmalchi, L. Chang, M. A. Torres, K. K. Ng, and P. J. Facchini. Codeinone reductase isoforms with differential stability, efficiency and product selectivity in opium poppy. *The Plant Journal*, 95:631–647, 2018.

[15] M. E. Eaton and J. K. Small. Addisonia: colored illustrations and popular descriptions of plants. *The Science Press Printing Company*, 11:27–28, 1926.

[16] S. C. Farrow, M. O. Kamileen, J. Meades, B. Ameyaw, Y. Xiao, and S. E. O'Connor. Cytochrome p450 and o-methyltransferase catalyze the final steps in the biosynthesis of the anti-addictive alkaloid ibogaine from tabernanthe iboga. *Journal of Biological Chemistry*, 293:13821–13833, 2018.

[17] G. S. Fraenkel. The raison d'etre of secondary plant substances: These odd

chemicals arose as a means of protecting plants from insects and now guide insects to food. *Science*, 129:1466–1470, 1959.

[18] A. Geerlings, M. M.-L. Ibanez, J. Memelink, R. van der Heijden, and R. Verpoorte. Molecular cloning and analysis of strictosidine β-d-glucosidase, an enzyme in terpenoid indole alkaloid biosynthesis in catharanthus roseus. *Journal of Biological Chemistry*, 275:3051–3056, 2000.

[19] M. A. Hicks, A. E. Barber, L.-A. Giddings, J. Caldwell, S. E. O'Connor, and P. C. Babbitt. The evolution of function in strictosidine synthase-like proteins. *Proteins: Structure, Function, and Bioinformatics*, 79:3082–3098, 2011.

[20] S. N. Ho, H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, 77:51–59, 1989.

[21] D. Hyndman, D. R. Bauman, V. V. Heredia, and T. M. Penning. The aldo-keto reductase superfamily homepage. *Chemico-Biological Interactions*, 143:621–631, 2003.

[22] J. M. Jez, M. J. Bennett, B. P. Schlegel, M. Lewis, and T. M. Penning. Comparative anatomy of the aldo–keto reductase superfamily. *Biochemical Journal*, 326:625–636, 1997.

[23] J. M. Jez, T. G. Flynn, and T. M. Penning. A new nomenclature for the aldo-keto reductase superfamily. *Biochemical Pharmacology*, 54:639–647, 1997.

[24] J. M. Jez and T. M. Penning. The aldo-keto reductase (akr) superfamily: an update. *Chemico-Biological Interactions*, 130:499–525, 2001.

[25] J. Joo, B. Himes, and T. Penning. Aldo-keto reductase (akr) superfamily database, 2021. Available at https://akrsuperfamily.org.

[26] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596:583–589, 2021.

[27] F. Kellner, J. Kim, B. J. Clavijo, J. P. Hamilton, K. L. Childs, B. Vaillancourt, J. Cepela, M. Habermann, B. Steuernagel, L. Clissold, et al. Genome-guided investigation of plant natural product biosynthesis. *The Plant Journal*, 82:680–692, 2015.

[28] Y. Kochnev, E. Hellemann, K. C. Cassidy, and J. D. Durrant. Webina: an open-source library and web app that runs autodock vina entirely in the web browser. *Bioinformatics*, 36:4513–4515, 2020.

[29] I. Letunic and P. Bork. Interactive tree of life (itol) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, 49:W293–W296, 2021. version 6.5.8.

[30] R. B. Leveson-Gower, C. Mayer, and G. Roelfes. The importance of catalytic promiscuity for enzyme design and evolution. *Nature Reviews Chemistry*, 3:687–705, 2019.

[31] I. A. Lobo, P. A. Robertson, L. Villani, D. J. Wilson, and E. G. Robertson. Thiols as hydrogen bond acceptors and donors: spectroscopy of 2-phenylethanethiol complexes. *The Journal of Physical Chemistry A*, 122:7171–7180, 2018.

[32] V. Logar. Iboga, 2010. Available at https://render.fineartamerica.com/images/rendered/share/780912&domainId=12.

[33] H. A. Maeda and A. R. Fernie. Evolutionary history of plant metabolism. *Annual Review of Plant Biology*, 72:185–216, 2021.

[34] N. Mills. Chemdraw ultra 10.0 cambridgesoft, 100 cambridgepark drive, cambridge, ma 02140. www.cambridgesoft.com., 2006. version 20.1.0.112.

[35] S. Oguro, T. Akashi, S.-i. Ayabe, H. Noguchi, and I. Abe. Probing biosynthesis of plant polyketides with synthetic n-acetylcysteamine thioesters. *Biochemical and Biophysical Research Communications*, 325:561–567, 2004.

[36] C. A. Orengo and J. M. Thornton. Protein families and their evolution—a structural perspective. *Annual Review of Biochemistry*, 74:867–900, 2005.

[37] T. M. Penning. The aldo-keto reductases (akrs): Overview. *Chemico-Biological Interactions*, 234:236–246, 2015.

[38] Y. Qu, M. E. Easson, R. Simionescu, J. Hajicek, A. M. Thamm, V. Salim, and V. De Luca. Solution of the multistep pathway for assembly of corynanthean, strychnos, iboga, and aspidosperma monoterpenoid indole alkaloids from 19e-geissoschizine. *Proceedings of the National Academy of Sciences*, 115:3180–3185, 2018.

[39] M. Ralser. The rna world and the origin of metabolic enzymes. *Biochemical Society Transactions*, 42:985–988, 2014.

[40] X. Robert and P. Gouet. Deciphering key features in protein structures with the new endscript server. *Nucleic Acids Research*, 42:W320–W324, 2014.

[41] M. G. Rossmann and P. Argos. Protein folding. *Annual Review of Biochemistry*, pages 497–532, 1981.

[42] S. Schläger and B. Dräger. Exploiting plant alkaloids. *Current Opinion in Biotechnology*, 37:155–164, 2016.

[43] L. Schrödinger and W. DeLano. Pymol, 2020. Available at http://www.pymol.org/pymol, version 2.4.0.

[44] D. Sengupta, D. Naik, and A. R. Reddy. Plant aldo-keto reductases (akrs) as multi-tasking soldiers involved in diverse plant metabolic processes and stress defense: A structure-function update. *Journal of Plant Physiology*, 179:40–55, 2015.

[45] F. Sertuerner. Ueber das Morphium, eine neue salzfähige Grundlage, und die Mekonsäure, als Hauptbestandtheile des Opiums, 1817. https://doi.org/10.1002/andp.18170550104.

[46] P. J. Simpson, C. Tantitadapitak, A. M. Reed, O. C. Mather, C. M. Bunce, S. A. White, and J. P. Ride. Characterization of two novel aldo–keto reductases from arabidopsis: expression patterns, broad substrate specificity, and an open active-site structure suggest a role in toxicant metabolism following stress. *Journal of Molecular Biology*, 392:465–480, 2009.

[47] E. C. Tatsis, I. Carqueijeiro, T. D. de Bernonville, J. Franke, T.-T. T. Dang, A. Oudin, A. Lanoue, F. Lafontaine, A. K. Stavrinides, M. Clastre, et al. A three enzyme system to generate the strychnos alkaloid scaffold from a central biosynthetic intermediate. *Nature Communications*, 8:1–10, 2017.

[48] O. K. Tawfik and D. S. Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annual Review of Biochemistry*, 79:471–505, 2010.

[49] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, et al. Swiss-model: homology modelling of protein structures and complexes. *Nucleic Acids Research*, 46:W296–W303, 2018.

[50] J.-K. Weng, J. H. Lynch, J. O. Matos, and N. Dudareva. Adaptive mechanisms of plant specialized metabolism connecting chemistry to function. *Nature Chemical Biology*, 17:1037–1045, 2021.

[51] J.-K. Weng, R. N. Philippe, and J. P. Noel. The rise of chemodiversity in plants. *Science*, 336:1667–1670, 2012.

[52] M. Yamasaki and E. Leete. Biosynthesis of the iboga alkaloids: the incorporation of tryptophan-3-c14 into ibogaine. *Tetrahedron Letters*, 5:1499–1501, 1964.

Declaration of independent work

I hereby declare that I have produced this work independently and without outside assistance, and have used only the sources and tools stated.
I have clearly identified the sources of any sections from other works that I have quoted or given in essence.
I have complied with the guidelines on good academic practice at the University of Göttingen.
If a digital version has been submitted, it is identical to the written one.
I am aware that failure to comply with these principles will result in the examination being graded "Nicht bestanden", i.e. failed.

Weimar 05.07.2022 M. Sieg

Place, Date Signature