Mechanistic Studies of Sulfur-Carbon Bond Formation by Metal-Dependent Enzymes

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Sebastian Flückiger

aus Münchenbuchsee BE

Basel, 2019

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von
Prof. Dr. Florian P. Seebeck
Prof. Dr. Thomas R. Ward

Basel, den 27.02.2018

Prof. Dr. Martin Spiess
Dekan der Philosophisch-
Naturwissenschaftlichen Fakultät

People keep saying,

"Science doesn't know everything!"

Well, science *knows* it doesn't know everything;

otherwise, it would stop."

- Dara Ó Briain

Table of Contents

Abbreviations

2A-DMH	2-amino N^{α},N^{α}-dimethyl histidine
ACV	δ-(*L*-α-aminoadipyl)-*L*-cystenyl-D-valine
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
BMC	4-bromomethyl-7-methoxycoumarin
CAII	Carbonic anhydrase II
CDNB	Chlorodinitribenzyl
CDNB	Chlorodinitrobenzyl
CDO	Cysteine dioxygenase
CoA	Coenzyme A
CSA	Cysteine sulfinic acid
CSS	Complex formation significance score
CT	Charge transfer
Cy_1	Cyclization domain
DFT	Density functional theory
DHAP	Dihydroxyacetone-phosphate
DMH	Dimethylhistidine
DMSO	Dimethylsulfoxide
DTB	Dethiobiotin
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EDTA	Ethylenediaminetetraacetic acid
EGT	Ergothioneine
Egt1	EgtB from *Neurospora Crassa*
$EgtB_{fungi\_C}$	EgtB from *Chaetomium thermophilum*
$EgtB_{fungi\_D}$	EgtB from *Dichomitus Squalens*
$EgtB_{fungi\_T}$	EgtB from *Thermothelomyces thermophilia*
$EgtB_{thermo1}$	EgtB from *Mycobacterium thermoresistibile*
$EgtB_{thermo2}$	EgtB from *Candidatus chloracidobacterium thermophilum*
$EgtB_{Vapar}$	EgtB homologue from *Variovorax paradoxus*
fCDO	Cysteine dioxygenase from *Dichomitus squalens*

GAP	Glyceraldehyde-3-phosphate
GDP	Guanine diphosphate
GSH	Glutathione
GST	Glutathione-*S*-Transferase
GT	Glycosyltransferase
GT	Glycosyltransferase
HEPES	2-(4-(2-hydroxyethyl)piperazin-1-yl)ethane-1-sulfonic acid
HOMO	Highest occupied molecular orbital
HRMS	High-resolution mass spectrometry
HRMS	High Resolution Mass Spectroscopy
IPNS	Isopenicillin N synthase
ITC	Isothermal titration calorimetry
KIE	Kinetic isotope effect
KISE	Kinetic solvent isotope effect
KSIE	Kinetic solvent isotope effect
MAPEG	Membrane-associated proteins in eicosanoid and glutathione metabolism
mBCl	Monochlorobimane
MG	Methylglyoxal
MO	Molecule orbitals
MS	Mass spectroscopy
MSH	Mycothiol
NMR	Nuclear magnetic resonance
NRPS	Nonribosomal Peptide Synthetases
RMSD	Root-mean-square-deviation
ROS	Reactive oxygen species
rSAM	Radical *S*-Adenosylmethionine
SAM	*S*-Adenosyl-L-methionine
SSN	Sequence similarity network
TCEP	Tris(2-carboxyethyl)phosphine
Tm	Melting temperature
TS	Transition state
UDP	Uridine diphosphate
γ-GC	γ-glutamylcysteine

Abstract

Sulfur is an essential element for all living organisms. In a large variety of relevant compounds like amino acids, cofactors and other natural products carbon-sulfur bonds are found. The understanding of the mechanistic details of carbon sulfur bond formation catalyzed by enzymes is from particular interest. The gained knowledge can be used for the discovery of new biosynthetic pathways of sulfur containing natural products as well as for the design of artificial enzymes able to perform carbon-sulfur bond formation.

In this thesis we investigate bacterial and fungal enzymes that catalyze carbon-sulfur bond formation. First, we characterized the catalytic activity of an unusual class of glutathione-S-transferases. These enzymes require a bivalent metal ion to activate their thiol substrate for nucleophilic attack onto carbon electrophiles. In vitro reconstitution of these enzymes, variation of the electrophile, the catalytic metal and the reaction conditions, compounded with structural studies revealed important insight into the catalytic mechanism of enzyme-catalyzed metal-dependent C-S bond formation.

Secondly, we examined a distantly related class of metalloenzymes that occur in bacteria and fungi. These metalloenzymes are oxygen-dependent and catalyze oxidative coupling of histidine derivatives with cysteine derivatives by forming a new C-S bond. This reaction is a key step in ergothioneine biosynthesis. Kinetic and structural examination of these enzymes were used to study the subtle but functionally significant differences between fungal and bacterial enzymes. Overall, we have broadened our knowledge of enzymes forming carbon-sulfur bonds and containing a DinB domain, along with discovering a new class of fungal cysteine dioxygenase.

1. Introduction

1.1 Sulfur

Carbon, hydrogen, oxygen and nitrogen are the elements which come to mind immediately if we think about the non-metal atomic composition of living organism. We would most likely complement the list with phosphorus as a main component of DNA. Although sulfur is often neglected, it is an essential element for all organisms and ubiquitous in Nature. The essential role of sulfur is based on the unique chemical flexibility of this element as will be illustrated in the following sections.

1.2 Sulfur Chemistry in Nature

Sulfur is found as the 16th element in the periodic table in the group containing the chalcogens and contains six valence electrons. One of the most unique properties of sulfur is the multiple different oxidation states it can adopt, from hydrogen sulfide with an oxidation state of -2 to sulfate with an oxidation state of +6. All oxidation states are found in Nature (Figure 1).

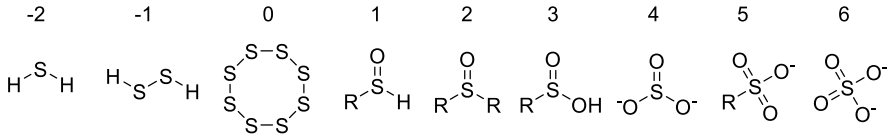


Figure 1. Examples of the different oxidation states of sulfur. Hydrogen sulfide (-2), hydrogen disulfide (-1), elemental sulfur (0), sulfenic acid (+1), sulfoxide (+2), sulfinic acid (+3), sulfite (+4), sulfonic acid (+5) and sulfate (+6).

Sulfur is found in primary metabolites such as the amino acids cysteine, methionine and homocysteine (Figure 2). Cysteine and methionine are both used as building blocks for proteins and contribute to the diverse properties of proteins.[1]

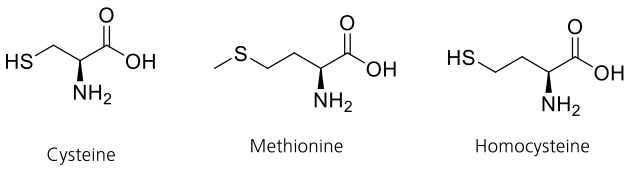


Figure 2. Chemical structures of the sulfur-containing amino acids cysteine, methionine and homocysteine.

The side chain of cysteine contains a thiol group, which can be used as a metal ligand. Cysteine metal ligands are observed, for example, in the copper-dependent formylglycine-generating enzyme or the

zinc-dependent DNA repair protein Ada.[2,3] The thiol group of cysteine can further form disulfide bridges to stabilize the secondary structure of proteins, such as in the structural skin protein collagen or the fibrous proteins of keratin, which making up hair, horns and nails.[4,5] The pK_a of the cysteine thiol found in proteins in a the range of 3.3, such as in papain to 9.8, such as in the acetyl-coenzyme A binding protein (T17C).[6–8] Most of the time and under physiological conditions, the thiol group inside of proteins can be found in its deprotonated state. This expands the functionality of thiols to act as bases, acids or nucleophiles. In an example of cysteine protease, the nucleophilic attack of a deprotonated cysteine is responsible for the cleavage of a peptide bond.[9,10] In the amino acid methionine, the sulfur is found as a thioether, which is a rather non-polar functional group. Nevertheless, thioethers are found as metal ligands as seen in the crystal structure of auracyanin A, a blue copper protein from *Chloroflexus aurantiacus*.[11] Furthermore in the example of the biosynthesis of the cofactor *S*-adenosyl-methionine, the thioether can act as a nucleophile.[12] An additional chemical function of thiols and thioethers is their ability to form sulfur center hydrogen bonds.[13] These interactions proved to be stronger than expected and are involved in protein structure stabilization and in physicochemical processes.[13–16]

Other sulfur-containing compounds in Nature include a variety of enzymatic cofactors (Figure 3). For example biotin (Vitamin B7), a cofactor is involved in enzymatic carboxylation, decarboxylation and trans carboxylation reactions.[17] Biotin is essential for all organisms but only produced by microorganisms.[18] Another co-factor, *R*-lipoic acid, is involved in acyl transfer reactions in mitochondria.[19,20] A further central co-factor is *S*-adenosyl-*L*-methionine (SAM) which is involved in methyl transfer reactions, as well as functioning as an initiator in radical reactions.[17,19,21]

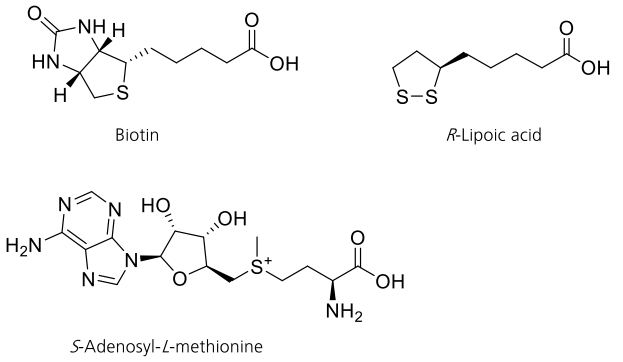


Figure 3. Chemical structures of a selection of important sulfur-containing co-factors.

Carbon-sulfur (C-S) bonds are also found in many bioactive natural products. Gliotoxin is a virulence factor of the human pathogen, *Aspergillus fumigatus*.[22] Lincomycin A, found in *Streptomyces lincolnensis*, and penicillin N, from *Penicillium* fungi, are two sulfur-containing molecules with antibiotic properties.[23–25] Sulforaphane and allicin are natural products from crucifers and garlic which are involved in plant defense mechanisms.[26,27] The two thiohistidines ergothioneine, isolated from the ergot fungus and ovothiol, isolated from sea urchin eggs (*Paracentrotu lividus*) are thiol compounds which are presumably involved in cellular redox homeostasis.[28–30]

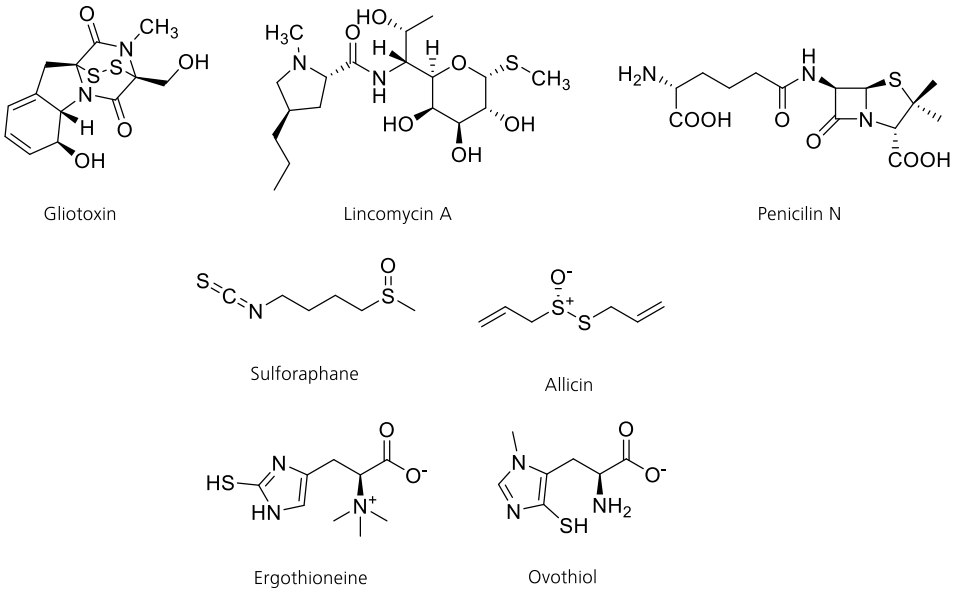
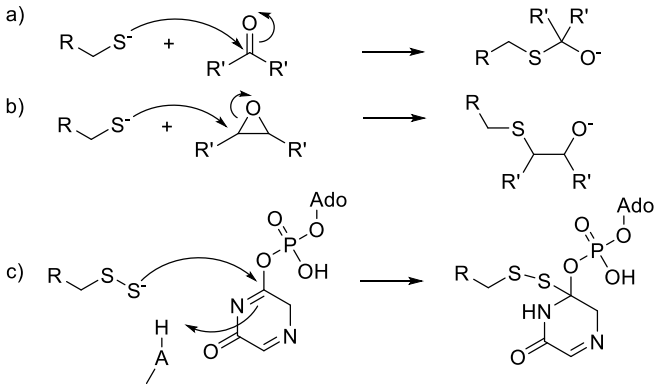


Figure 4. Examples of natural products which contain C-S bonds.

The versatility of sulfur is also seen in the formation of C-S bonds in natural products. There are several approaches by which a C-S bond can be formed each exploiting alternative reactivities of sulfur. The three main classes of C-S bond forming reactions are electrophilic reactions, nucleophilic reactions as well as radical reactions, each of which will be subsequently discussed.

1.3 Nucleophilic Carbon-Sulfur Formation

The first and perhaps the most straight forward approach for the formation of a C-S bond is the nucleophilic attack of a thiol on an electrophile (Scheme 1). By forming a thiolate-ion the nucleophilicity of the sulfur atom can be increased, which further promotes the reaction.[31] Multiple classes of enzymes utilize this strategy to form C-S bonds *via* a nucleophilic addition or substitution reaction. Two important classes of enzymes which catalyze such nucleophilic reactions are the *S*-transferases and the C-S cyclases.[25] Another strategy involves the use of persulfide as a nucleophile. Persulfide-containing enzymes are often involved in the formation of primary metabolites.[32]



Scheme 1. Nucleophilic reaction of a thiolate or persulfide to an electrophile. The thiolate attack at a carbonyl **a)** or at a peroxide **b)**. **c)** Persulfide attack on a the C4 atom of an activated uridine.

1.3.1 *S*-Transferases

Many sulfur-containing natural products are formed by *S*-transferases, which catalyze the attack of a thiol group on an activated carbon. Different families of *S*-transferases exist, such as the *S*-methyltransferases, the *S*-glycosyltransferases, and the glutathione-*S*-transferases.[25]

S-Methyltransferase, for example, are involved in the stepwise methylation of sulfide to dimethylsulfide, as well as in the detoxification process of gliotoxin, a mycotoxin found in *Aspergillus fumigatus* (Figure *5*).[33,34]

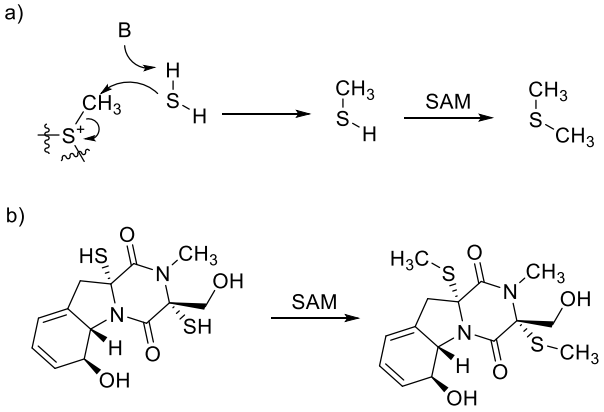


Figure 5. **a)** Stepwise methylation of hydrogen sulfide by a *S*-methyltransferase. **b)** Deactivation of the mycotoxin gliotoxine, by SAM-dependent methylation of the two available thiol groups.[33,34]

While *S*-glycosyltransferases are involved in the *S*-glycosylation of *S*-linked glycopeptides and glucosinolates the formation of *S*-linked sugar molecules is relatively rare in natural products. Nevertheless an *S*-linked glycopeptide with antimicrobial properties, called sublancin was found in *Bacillus subtilis* (Figure 6 a) & b)). Another example of a thio sugar is glucosinolate, a natural product class found as a plant defense in crucifers (Figure 6c).[27,35]

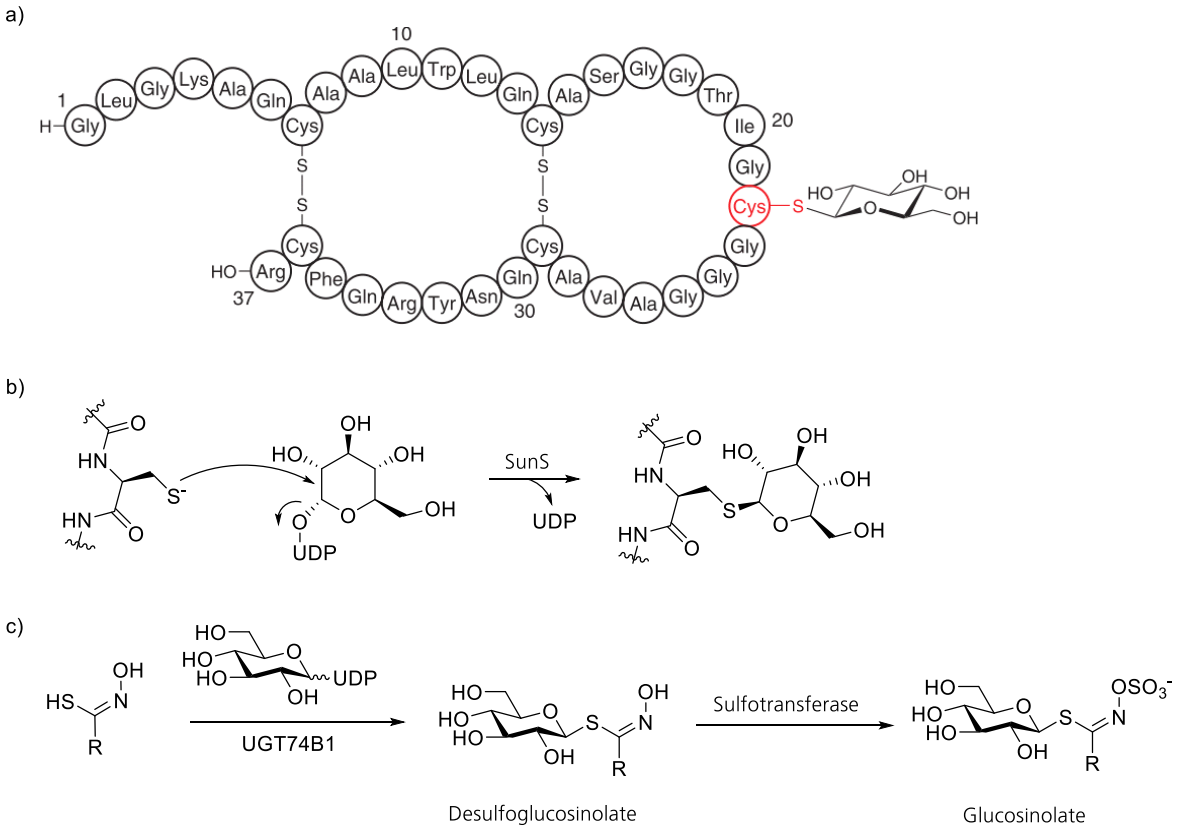


Figure 6. **a)** Structure of sublancin, an *S*-linked glycopeptide.[35] **b)** The proposed reaction mechanism to form a glycosylated product from a cysteine thiolate and UDP-glucose by the glycosyltransferase SunS.[36] **c)** *In vitro* analysis of the enzyme UGT74B1 from *Arabidopsis thaliana* showed *S*-glycosylation activity with a thiohydroximix group and UDP-glucose to yield desulfoglucosinolate. This is further used as a sulfotransferase substrate to form glucosinolate.[27]

An example of a nucleophilic reaction where a glutathione-*S*-transferase is involved, is the biosynthesis of Leukotriene C4, an inflammatory mediator produced in leukocytes.[37] Glutathione-*S*-transferase catalyzes the reaction, in which the epoxide ring of Leukotriene A4 is attacked by the thiol group of glutathione to form Leukotriene C4 (Figure 7).[37]

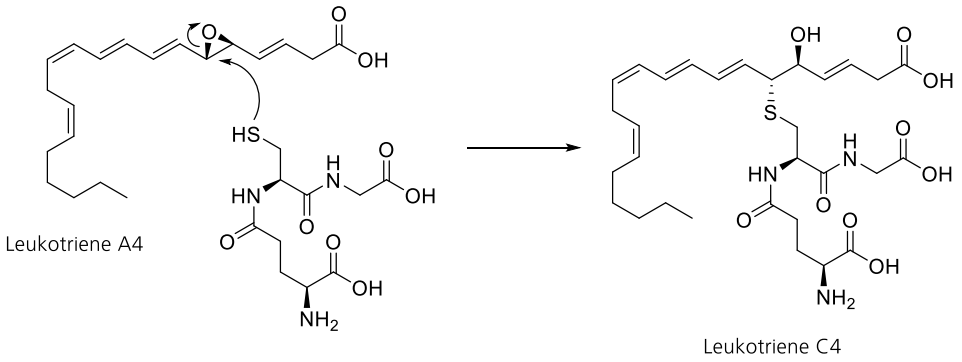


Figure 7: The formation of Leukotriene C4 is catalyzed by a glutathione-*S*-transferase. The thiolate attacks one C-atom of the epoxide ring as a nucleophile.[25,37]

1.3.2 C-S Cyclases

Intramolecular nucleophilic C-S bond formations are catalyzed by cyclases. For example several examples are involved in the formation of thiazoline rings on nonribosomal peptide synthetase (NRPS) assembly lines.[38,39] Other cyclases have been reported to be involved in the biosynthesis of lanthipeptides: the cyclases catalyze the conjugate addition, which forms the characteristic lanthionine rings.[40]

Two examples of natural products which contain a thiazoline ring formed by a C-S cyclase are the two high-affinity iron-chelating molecules, siderophore, yersiniabactin from *Yersinia pestis* and pyochelin from *Pseudomonas aeruginosa* (Figure 8).[41,42] The mechanism of the thiazoline ring formation is in both cases a cyclocondensation, which is catalyzed by domains typically enclosed in the NRPS module. The cyclisation domains catalyze both the condensation reaction and the heterocycle formation.[25]

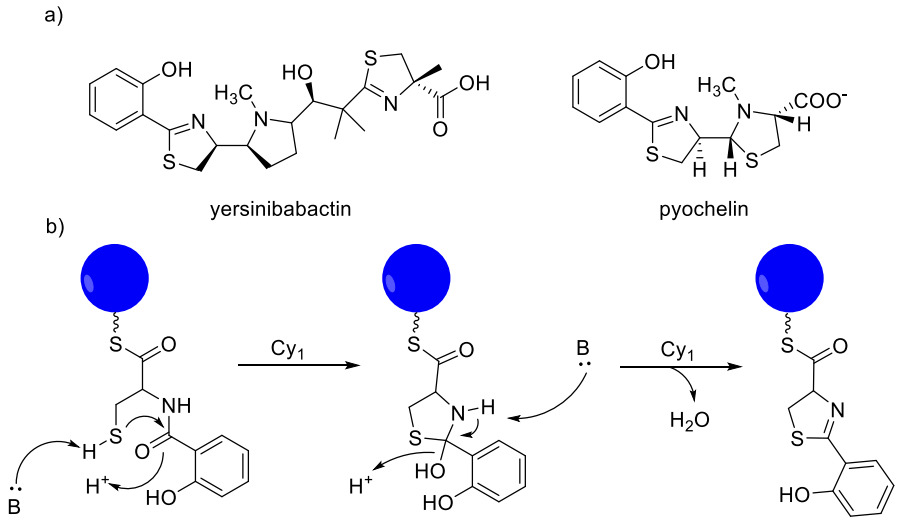


Figure 8. **a)** Structures of the siderophore yersiniabactin and pyochelin.[41,42] **b)** Reaction mechanism of the cyclization domain (Cy_1) of the NRPS module. The first step involves a cyclisation, followed by a condensation reaction to form the thiazoline ring.[25]

An example of a lanthipeptide which is formed by a Michael-like addition mechanism is Nisin A, an antimictobial.[40,43] This peptide is ribosomally synthesized and post-translationally modified by the dehydratase NisB, the cyclase NisC and the protease NisP.[43] The biosynthesis of the lanthionine ring starts with the dehydration of a serine or threonine residue to form a Michael acceptor. In the next step a cyclisation occurs by the attack of the thiol group of a cysteine residue to the unsaturated amino acids, resulting in the final product (Figure 9).[43]

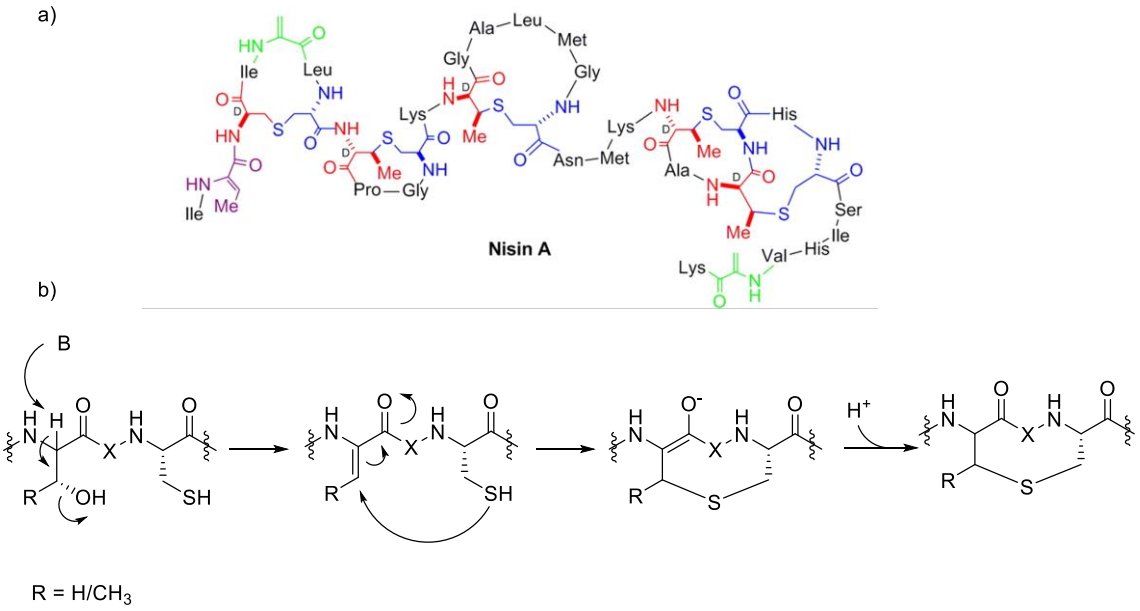
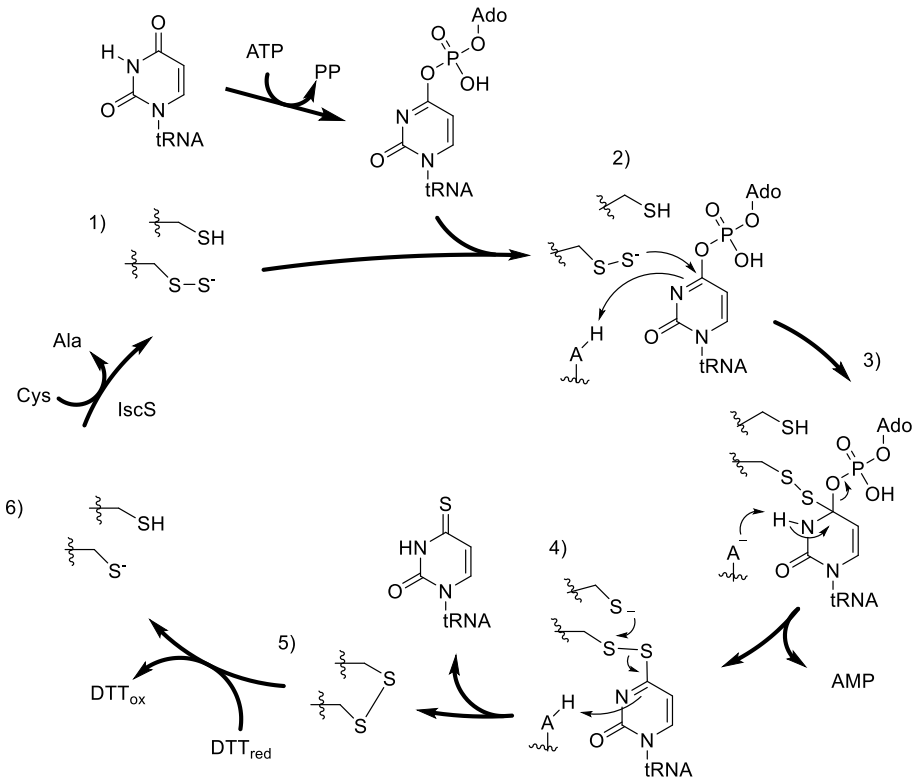


Figure 9. **a)** Structure of the antimicrobial peptide Nisin A.[43] **b)** Reaction mechanism for the formation of the lanthionine ring. A dehydration reaction leads to the formation of a Michael acceptor, which is then attacked by the thiol group of a cysteine residue to form the cyclic product.[43]

1.3.3 Persulfide as a Nucleophile

Persulfides (R-S-S-H) are involved in iron-sulfur cluster biosynthesis, as well as in the biosynthesis of various other primary metabolites, such as the cofactor thiamin or 4-thiouridine. 4-Thiouridine is found at position 8 of tRNA and is used by some bacteria as a photosensor, which regulates bacterial growth during UV-light exposure.[32,44] In the biosynthesis of 4-thiouridine in *E. coli*. two enzymes ThiI and IscS are involved.[45,46]

It was proposed that the persulfide acted as a nucleophile (Scheme 2). in the first step the sulfurtransferase IscS transfers the sulfur of a free cysteine to a cysteine residue in the active site of the enzyme ThiI, thereby forming a persulfide and an alanine(**1**). This enzyme-bound persulfide can attack an ATP-activated tRNA to form the disulfide bond that links the enzyme ThiI to the tRNA (**2-3**). In the next step a second active site cysteine residue, which was identified through point mutation studies, liberates the thiouridine residues by forming a disulfide bond (**4**).[47] This disulfide bond can be reduced to form the two cysteine residues in the active site and regenerate the ThiI enzyme (**5-6**).[32]



Scheme 2. The proposed reaction mechanism for the generation of 4-thiouridine catalyzed by the enzyme ThiI. The persulfide in the active site, which is generated by the sulfurtransferase IscS, can attack an activated uridine (**1-3**). The newly formed disulfide bond that links the tRNA to ThiI can be liberated by a cysteine which is present in the active site (**4**). The reduction of the disulfide bond regenerates the enzyme (**5**).[32]

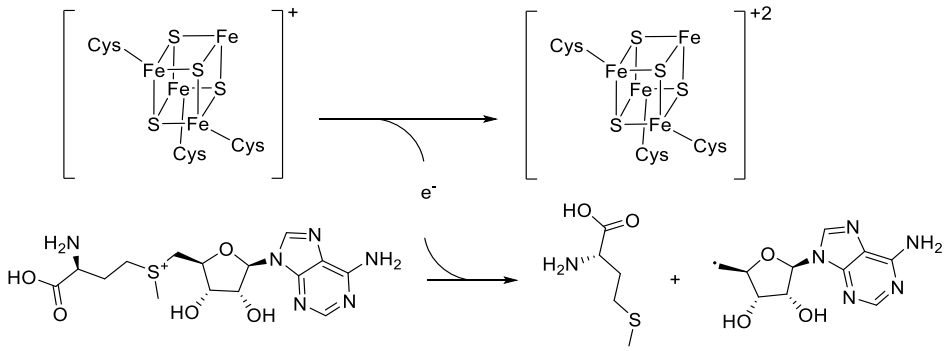
1.4 Radical Sulfur Reactions

Another reaction mechanism that is used by enzymes for C-S bond formation is the radical reaction. In Nature, this strategy to form C-S bonds is found in radical *S*-adenosylmethionine enzymes and various oxidases and oxygenases.[48,49]

1.4.1 Radical *S*-Adenosylmethionine Enzymes

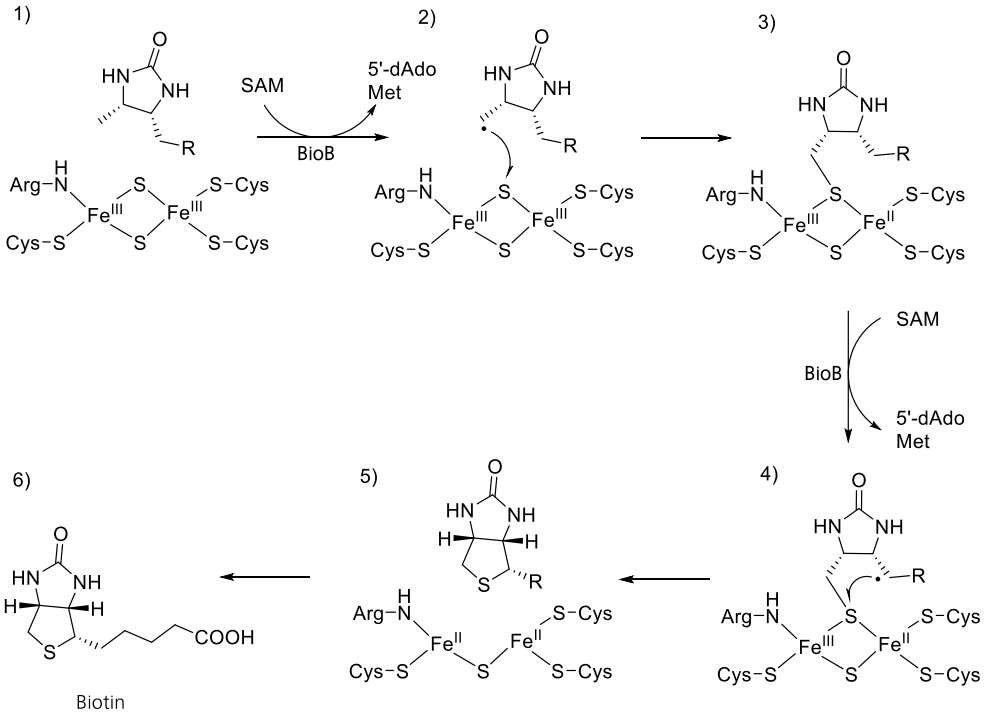
Radical *S*-adenosylmethionine (rSAM) enzymes use an [4Fe-4S] cluster to reductively cleave an SAM molecule to form methionine and the highly reactive 5'-deoxyadenosyl radical (Scheme 3). This radical is then able cleave unreactive C-H bonds in order to perform chemically challenging transformations. This strategy is involved in the biosynthesis of many different natural products.[17,48,50] In the diverse family of rSAM enzymes, only a few enzymes were found to be involved in C-S bond formation. rSAM enzymes catalyze the C-S bond formation in the biosynthesis of the cofactors biotin and lipoic acid. Other recently

reported examples which use rSAM enzymes in their biosynthesis include two natural product classes of albomycin, an antibiotic from *Streptomyces*, and sactipeptides, a subclass of RiPP products.[25]



Scheme 3. Reduction of SAM by a [4Fe-4S] cluster to form methionine and the reactive 5'-deoxyadenosyl radical.[48]

In the biosynthesis of the cofactor biotin, which is involved in carboxylation and decarboxylation processes, the following mechanism was proposed (Scheme 4). In a first step of the reaction, SAM is reduced by a [4Fe-4S] sulfur cluster to form a 5'-deoxyadenosyl radical (**1**). This radical abstracts the hydrogen of the substrate dethiobiotin (DTB) at the position C9 (**2**). The DTB radical then reacts with a second iron-sulfur cluster [2Fe-2S] to form the intermediate 9-mercaptodethiobiotin (**3**). Another SAM molecule is reduced to generate a second radical which abstracts a hydrogen at position C6 (**4**). Finally, the thiophane ring is closed and the product biotin is released (**5-6**). This mechanism is based on the dimeric crystal structure of the enzyme biotin synthase which was co-crystalized with the substrate DTB, with SAM and both iron-sulfur clusters.[51] The order of hydrogen abstraction was identified by chemically synthesizing the 9-mercaptodethiobiotin, which was transformed by biotin synthase to the product biotin.[52] The rate limiting step was proposed to be the hydrogen abstraction of the C9 methyl group. The labeled ($^2H_3$-9-methyl)-DTB substrate had almost no reactivity and the primary kinetic isotope effect was estimated to be more than 40.[17]

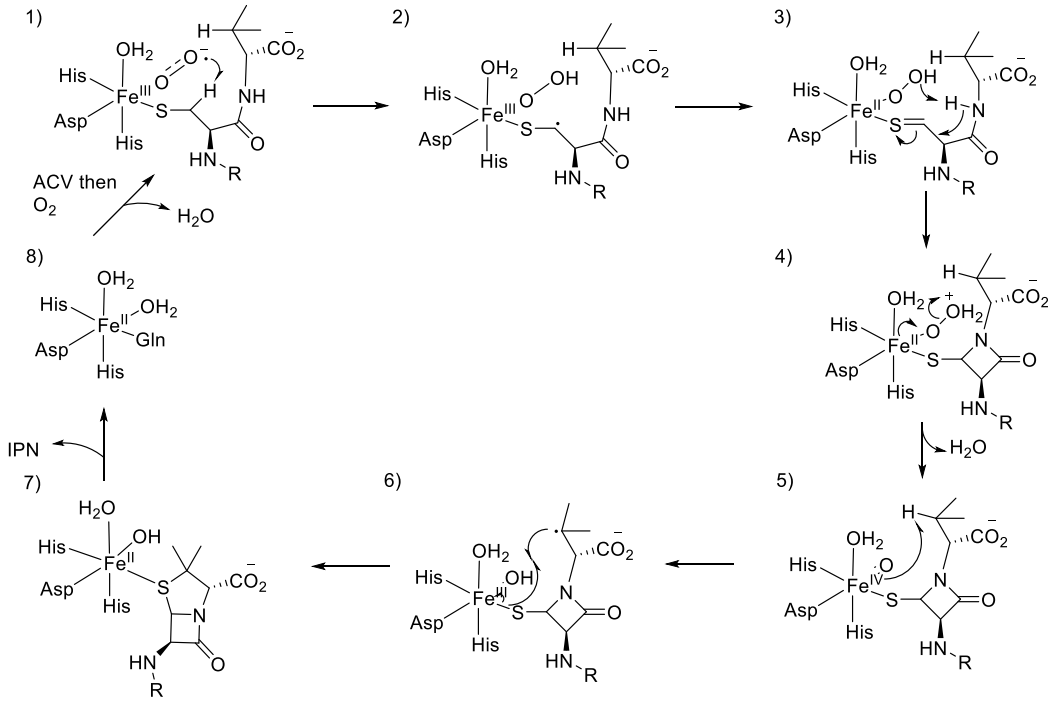


Scheme 4. Proposed reaction mechanism of the last step of biotin biosynthesis. The reduction of SAM forms a 5'-deoxyadenosyl radical which abstracts the hydrogen of C9 of DTB (1). The newly formed C9 radical can attack the sulfur of the [2Fe-2S] cluster to form the 9-mercaptodethiobiotin intermediate (2). In a third step the hydrogen of the C6 is abstracted by a SAM induced radical (3) and forms the thiophane ring of biotin (4).

1.4.2 Isopenicillin N synthase (IPNS)

The first antibiotic substance penicillin was discovered in 1929 by Alexander Fleming.[53] Penicillin contains a β-lactam and a thiazolidine ring and irreversibly inhibits the biosynthesis of peptidoglycan components, which are required in the cell wall of gram-positive bacteria.[54] In the last step of the biosynthesis of penicillin N the enzyme isopenicillin N synthase (IPNS) catalyzes the ring formation of both the β-lactam and the thiazolidine ring. The substrate of IPNS was identified to be the linear peptide δ-(*L*-α-aminoadipyl)-*L*-cystenyl-D-valine (ACV).[55] In addition, it was demonstrated that ascorbate and Fe^{2+} increased the reaction rate of IPNS and that oxygen is required as electron acceptor.[56] It was proposed that IPNS is a non-heme iron oxidase with a radical reaction mechanism. The crystal structure with the substrate ACV, iron, and the dioxygen analogue NO was solved under anaerobic conditions.[57] This structure revealed the octahedral binding motive for the iron atom, which consisted of two histidine residues, one aspartate, the thiol group of ACV, a water molecule and NO. Based on crystal structures, mechanistic studies and computational studies, the following mechanism was proposed (Scheme 5).[49,57–59] In the first phase, the *β*-lactam ring is formed in a two-electron oxidative cyclization. After the binding of the substrates, ACV and O_2 to the active site (**1**), the hydrogen of the beta carbon of cysteine in ACV is cleaved (**2**). Afterwards, a thioaldehyde can

be formed by the coordination of the formed thioalkyl radical to the Fe(III) (**3**). Cleavage of the O-O bond coupled to the attack of the amide on the thioaldehyde (**4**) and water elimination led to the formation of the β-lactam ring and an iron(IV) ferryl-oxo complex (**5**). In the second phase, another two-electron oxidative cyclization occurs whereby the thiazolidine ring is formed. The ferryl-oxo species abstracts a hydrogen atom from the beta position of valine (**5**). The formed valinyl radical then reacts with the coordinated sulfur (**6**) to yield the product isopenicillin N (**7**). In recent studies, it was possible to observe both reactive iron intermediates, the ferric superoxide and the iron(IV) ferryl-oxo species by rapid-freeze-quench Mössbauer spectroscopy and UV-Vis absorption spectroscopy.[59]



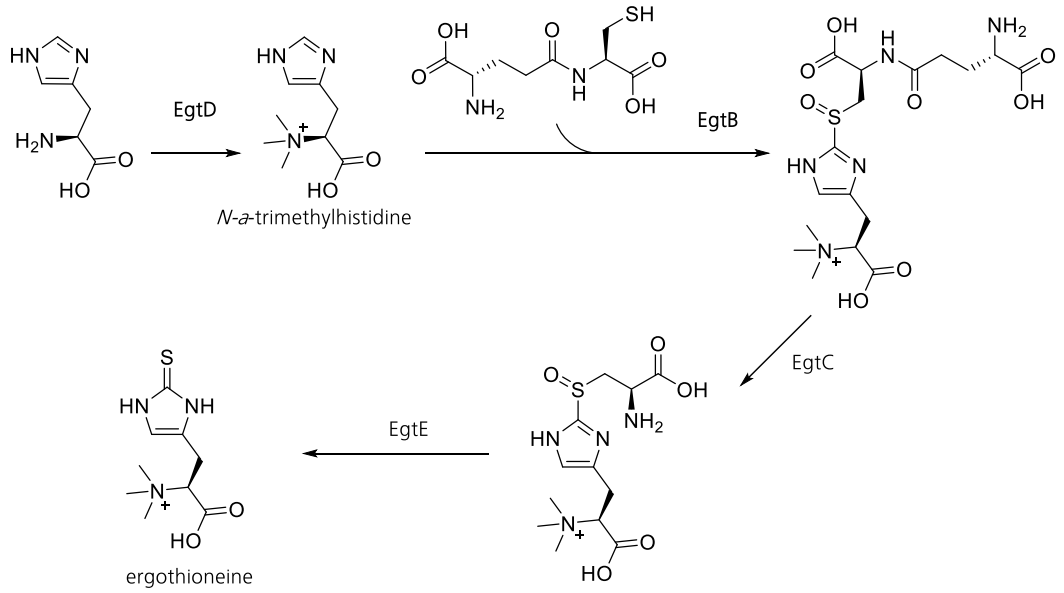
Scheme 5. Proposed reaction mechanism of IPNS. **1)** After the binding of oxygen and ACV, a proton is abstracted to form the superoxo species. **2)** The alkyl radical can form a thioaldehyde **(3)** which is attacked by the deprotonated amine group to form the β-lactam ring. **4)** The elimination of water leads to the formation of a ferryl-oxo species **(5)** which is able to abstract a hydrogen form the beta C of valine. **6)** The valinyl radical then react with the coordinated sulfur to yield the final product Isopenicillin N (**7**).[49,59]

1.4.3 Ergothioneine

A unique sulfur-containing natural product which contains a C-S bond formed in a radical reaction is Ergothioneine (EGT).[60] EGT is a trimethylbetaine of *2*-thiol-*L*-histidine and is produced in non-yeast fungi actinomycetes and cyanobacteria.[61] The interest in this thiol increased tremendously after the discovery of increased EGT concentrations in blood and tissues of higher organisms such as rats, pigs and humans as well as in higher plant materials.[62–64] No evidence has been found that those higher organism are able to

produce EGT by themselves. Therefore, the molecule must be taken up by nutrition. The discovery of the membrane protein OCTN1, which is a EGT transporter in humans, underlined the importance of the uptake and raised questions about the biological function of EGT.[65] In bacteria and fungal organisms the *in vivo* function of EGT- antioxidant activity, protection against oxidative stress, defense against heat shock and UV radiation and enhanced life- were found.[66–70]

The complete biosynthetic pathway of EGT was discovered in *Mycobacterium smegmatis*.[71] The bacterial enzymes, involved in the biosynthesis are located in a five gene cluster and consist of EgtA, a γ- glutamyl cysteine synthetase, EgtB a non-heme iron-dependent sulfoxide synthase, EgtC a glutamine amidotransferase, EgtD a SAM-dependent methyltransferase and EgtE a PLP-dependent liase.[71] The biosynthesis of EGT starts from *L*-histidine, which is trimethylated by EgtD to form *N,N,N*-α-trimethyl histidine (TMH). Subsequently, the sulfoxide synthase EgtB forms the C-S bond between the TMH and γ-glutamyl cysteine. The glutamine residue is then removed by EgtC to form the intermediate hercynylcysteine sulfoxide which is then transformed in the last step by EgtE to EGT (Scheme 6).



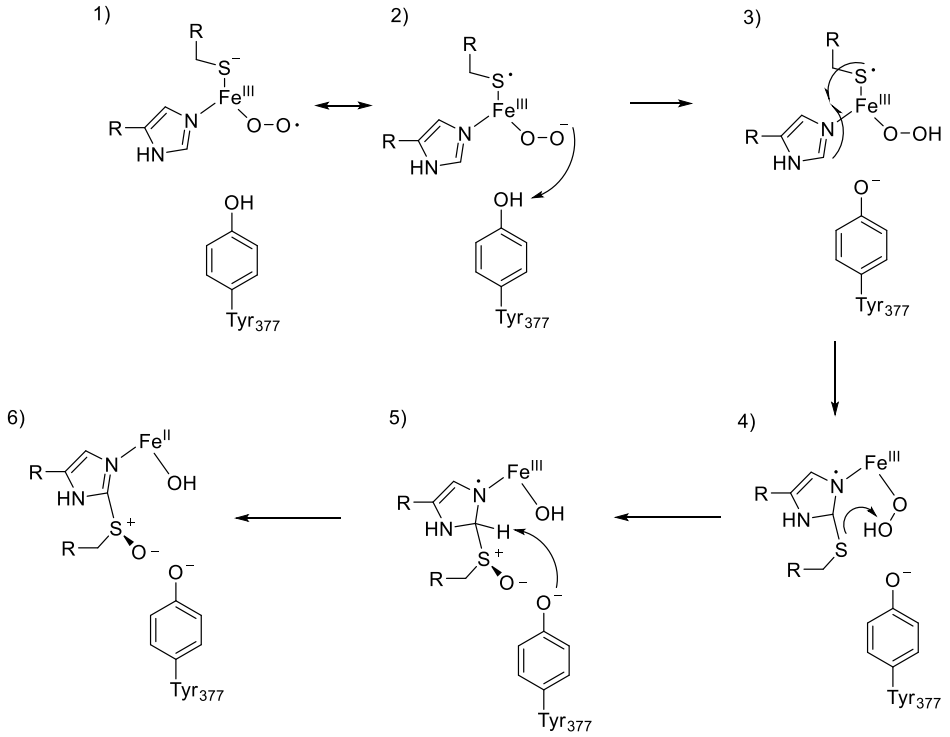
Scheme 6: The biosynthesis of EGT in mycobacteria. Histidine is methylated by EgtD, followed by a sulfoxidation with γ-glutamyl cysteine catalyzed by EgtB. EgtC cleaves the glutamyl skeleton form the sulfoxide and finally EgtE catalyses the formation of the final product EGT.[71]

1.4.4 Sulfoxide Synthase EgtB

The radical-induced C-S bond formation between the imidazole ring of TMH and a sulfur source in the biosynthesis of EGT is catalyzed by the enzyme EgtB.[71] EgtB is a non-heme iron-dependent sulfoxide

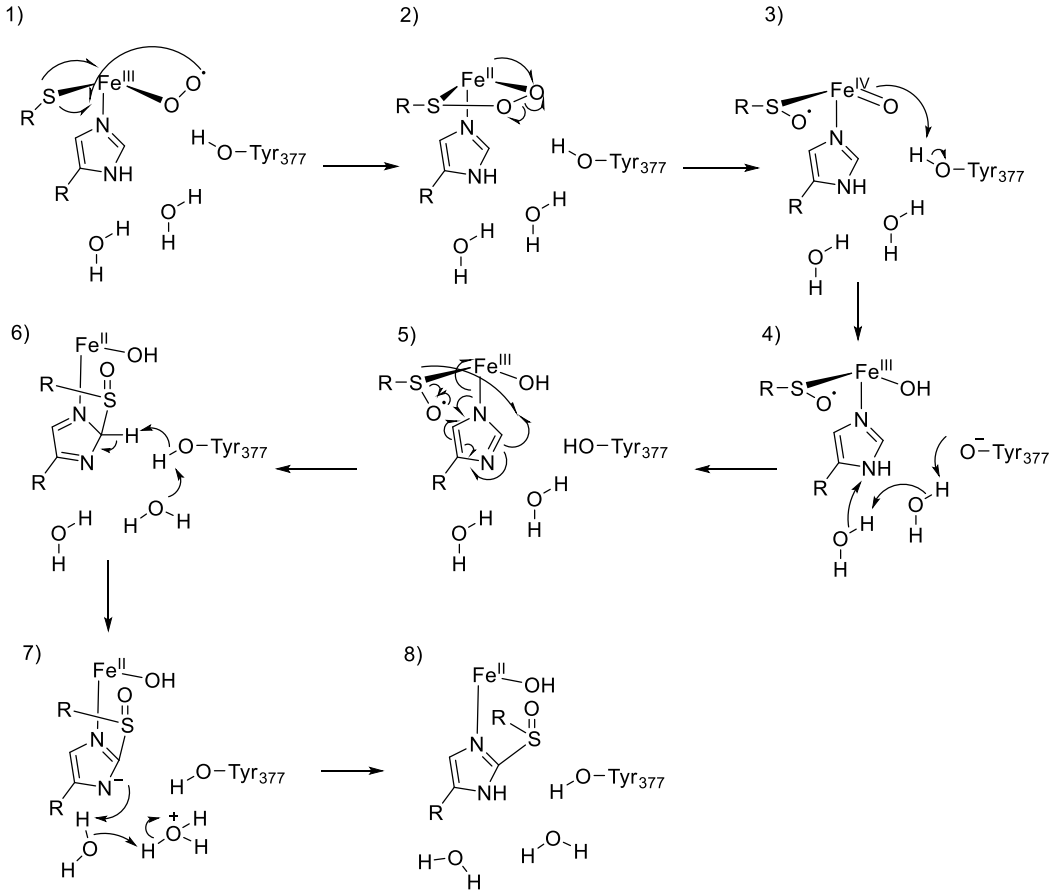
synthase. To date at least two possible reaction mechanisms are under discussion for EgtB in Mycobacteria.[60,72] The main difference between these proposed mechanism is whether the S-O or the C-S bond is formed first.

The first mechanism of EgtB was proposed according to kinetic studies and the crystal structure of the bacterial EgtB from *Mycobacteria thermoresistibile* ($EgtB_{thermo1}$).[60] The oxidative sulfurization catalyzed by EgtB requires an iron cofactor and oxygen in addition to its substrates γ-glutamylcysteine and TMH. In the absence of ascorbate, the number of turnovers was limited in the *in vitro* assay for $EgtB_{thermo1}$.[60] This observation indicates that the starting oxidation state of the iron is required to be in its reduced ferrous state. Upon incubation with γ-glutamylcysteine, a deactivation and an absorption band at 565 nm were observed. This absorption band vanished after the addition of ascorbate.[60] A plausible explanation for this result is a sulfur to iron(III) charge transfer, which was also observed in iron-sulfur(III) enzyme complexes of cysteine dioxygenases.[73] In the $EgtB_{thermo1}$ crystal structure the enzyme's iron binding site was identified as a three-histidine facial triad (H51, H134 and H138). Together with the two substrates bound (γ-glutamylcysteine, dimethylhistidine) and a water molecule, an octahedral metal binding sphere could be observed (PDB: 4x8d, 1.98 Å). Based on this observation, the following mechanism was proposed (Scheme 7).[60,74] In the first step the substrates and the oxygen bind to the iron(II) complex. It well accepted that complexation of oxygen to iron forms a ferric superoxo species (**1**). This iron-oxygen bound species may be in an equilibrium with a peroxide anion (O_2^{2-}) and a thiyl radical. A proton transfer from Tyr_{377} is proposed to stabilize this peroxide anion (**2**). This stabilization is coupled to the attack of the thiyl radical on to the imidazole ring to form the C-S bond and produce an iminyl radical on TMH (**3**). The formed thioether can then attack the distal oxygen (**4**). After the oxidation of the thioether, deprotonation of the C5 position of imidazole takes place resulting in the rearomatization and ligand-to-metal electron transfer to regenerate the ferrous cofactor (**5**).[60,74] In the final step of the reaction the sulfoxide product is released from the enzyme (**6**).



Scheme 7: Proposed mechanism of EgtB, the superoxo iron species forms the peroxide anion and the thiyl radical. The protonation of the peroxide anion is coupled to the thiyl radical attack of the imidazole. The thioether is oxidized by the peroxide species and finally deprotonation and ligand-to-metal electron transfer leads to rearomatization and to the formation of the final product.[60,74]

The second proposed mechanism by Wei *et al.*, was based on theoretical studies with DFT calculations, where the first step remains the same (Scheme 8).[72] The substrates bind to the iron-metal center and a ferric superoxo complex and formed (**1**). Afterwards the S-O bond is generated by the attack of the superoxide complex and resulting in a ferrous thiyl peroxide intermediate (**2**). In the next step the O-O bond is homolytically cleaved to form an iron(IV) ferryl-oxo species and an RSO-species (**3**). The deprotonation of Tyr377 by the iron(IV) ferryl-oxo species followed by the deprotonation of the imidazole ring by Tyr377 over two water molecules results in an anionic imidazole through a partial radical reaction (**4**). The C-S bond is then formed by a radical coupling reaction (**5**). Finally, the imidazole ring is rearomatized by a two-step proton transfer resulting in the final product (**6-8**). The calculated rate limiting step was the proton abstraction from the imidazole carbon by Tyr377 and the calculated deuterium kinetic isotope effect (KIE) for the TMH imidazole $C_\varepsilon H/C_\varepsilon D$ had a value of 5.7. This reaction model is not supported by kinetic data from our laboratory as no KIE with C-2 deuterium-labelled TMH was observed.[74]



Scheme 8. Mechanism of Wei *et al.* based on density function theory (DFT) calculations. 1) The sulfur reacts with the superoxo specie to form a ferrous thiyl peroxide intermediate. 2) The homolytic cleavage if the O-O bond forms the ferryl oxo species. 3) This iron(IV) ferryl-oxo species deprotonates the hydroxide group of Tyr377. 4) The hydroxide ion of Tyr377 is able to deprotonate the imidazole ring over two water molecules. 5) The C-S bond is formed by radical coupling. 6-7) A two step proton transfer rearomatizes the imidazol ring and forming the final product (8).[72]

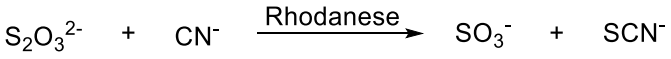
1.5 Electrophilic Sulfur Reaction

In addition to the more common nucleophilic and radical reactions, sulfur can also participate in reactions in which it reacts as an electrophile. Electrophilic sulfurs are mainly found in persulfides and disulfites, which can be attacked by nucleophiles. One of the best described examples is the reaction of cyanide with disulfite, which is catalyzed by the enzyme class, rhodanese.

1.5.1 Rhodanese

Rhodaneses are involved in the catalytic sulfur transfer reactions to the cytotoxic nucleophilic cyanide to form thiocyanate.[75] The cyanide anion or protonated hydrogen cyanide form (pK_a = 9.21 at 25°C) can inhibit key metabolic enzymes and is therefore of is toxic to most living organisms. A prominent example of toxicity

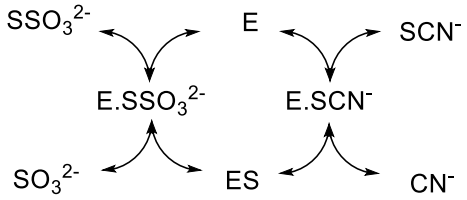
is the inhibition properties of cyanide on cytochrome c oxidases, thereby blocking aerobic cellular respiration.[76] The sulfur source of this reaction is the sulfite anion (Scheme 9).[75,76]



Scheme 9. The reaction scheme of the detoxification reaction of cyanide catalyzed by rhodanese.[75,76]

The function of rhodanese and rhodanese-like enzyme is not limited to the detoxification of cyanide. Other functions such as thioredoxin oxidase activity or the involvement in the selenium metabolism have been characterized.[77,78] Nevertheless, the best characterized reaction system is the bovine liver rhodanese, which catalyzes the *in vitro* detoxification reaction between thiosulfate and the cyanide to form the less toxic thiocyanate.[75,77,79]

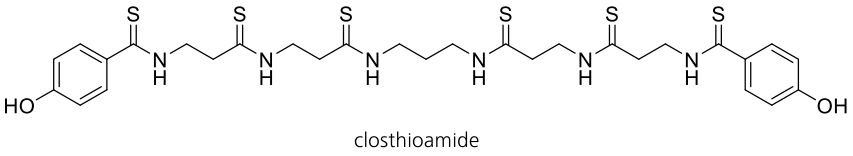
The transfer reaction of the sulfur atom is best described by a double displacement mechanism (Scheme 10). In the first step, the thiosulfate is taken up by the enzyme and coordinated to a metal ion inside the active site. A conformational change of the enzyme stretches and weakens the sulfur-sulfur bond of thiosulfate. The breaking of the sulfur-sulfur bond and the formation of the persulfide intermediate with a cysteine residue in the active site is the rate limiting step of this reaction. The persulfide intermediate of the bovine rhodanese was identified by protein crystallography (PDB: 1rhs, 1.36 Å).[80] The sulfur loaded-enzyme then reacts in the second step with the nucleophilic cyanide to form the final product thiocyanate.



Scheme 10. The double displacement reaction of bovine rhodanese. In the first step a sulfur atom of thiosulfate is covalently attached to the enzyme by forming a persulfide. In the second step the nucleophilic cyanide abstracts the sulfur atom from the enzyme to form the final product.[75]

1.6 Objectives

The mechanistic possibilities of enzymes to form C-S bonds in Nature is highly fascinating. Although the mechanistic understanding of C-S bond formation in enzymes has increased over the last few years, there are still numerous sulfur-containing natural products, for which the biosynthetic routes are unknown.[25] One such example is closthioamide a compound with antibiotic properties found in the anaerobic bacterium *Clostridium cellulolyticum* (Figure 10).[81] Closthioamide has numerous C-S bonds, yet the enzyme or enzymes that form these bonds are as of yet unknown.



closthioamide

Figure 10. The chemical structure of closthioamide with an unknown biosynthesis.

In other pathways, the responsible enzyme had most often been identified, however the mechanism of C-S bond formation is still not completely understood. For example, the proposed mechanism of the catalytic oxidative C-S bond formation by the sulfoxide synthase EgtB is still under discussion.[60,72,82] The focus of the mechanistic research on EgtB was based on bacterial homologues of this enzyme, with an available crystal structure.[60,74,83] Furthermore the evolution of enzymes which catalyze this distinct oxidative C-S bond formation has not yet elucidated. By the expansion of the mechanistic research to enzymatic homologues of different organisms a platform of comparative enzymology can be established. This helps to identify the core catalytic features which are required for the reaction and can support proposed mechanisms.

This thesis aims to investigate four different enzymes which are involved in metal induced sulfur activation. In the first chapter, a glutathione *S*-transferase with a distinct metal binding site is described. We shed light on the metal-induced sulfur activation of this simple enzymatic system with a kinetic model system and presen solved crystal structures of the enzyme.

In the second chapter, we expand our knowledge about EgtB enzymes by investigating EgtB homologues found in eukaryotes. This investigation led to the discovery of a novel type of cysteine dioxygenase in fungi.

In the last chapter, we investigated a EgtB homologues with unknown function and aimed to discover the biosynthesis of selenoneine in bacteria.

This thesis complements the understanding of C-S bond formation in metal-dependent enzymes. This is a further step to discovering unknown and understanding known related sulfur activating enzymes. Furthermore, the increased knowledge of C-S bond formation in enzymes can be used for the construction of artificial enzymes.

2 Glutathione *S*-Transferases

2.1 Introduction

2.1.1 Glutathione

It has been proposed that molecular oxygen was formed in two periods - between 2.5-2.2 and 0.8-0.5 billion years ago - and these events significantly changed the atmosphere and the living conditions for ancestral organisms.[84;85] The newly formed oxidative environment led to oxidative stress and increase in the amount of reactive oxygen species (ROS) in early cells.[86;87] Elementary oxygen can be activated by flavin chemistry and other electron transfer processes, for example is activation by metal ions to form superoxide anions, hydrogen peroxide and hydroxyl radicals.[88] These oxygen species can then further react with metabolites, DNA and proteins, leading to irreparable damages to cells.

In addition to oxidative radical species, living organisms were also exposed to toxic electrophiles. These electrophiles can either be formed as side products of metabolic processes or taken up from the environment. One example of a metabolic side product is methylglyoxal formed during glycolysis. Glycolysis is an ancestral pathway found in bacteria, archaea and eukaryotes, where glucose is converted to pyruvate and the energy carrier ATP is generated. During glycolysis, a side reaction can occur where 2-oxoaldehyde can be formed. Phosphate can eliminate from glyceraldehyde-3-phosphate (GAP) and dihydroxyacetone-phosphate (DHAP) to form methylglyoxal (MG), a strong electrophile (Figure 11).[89] This electrophile can react spontaneously with nucleophiles from biomolecules such as proteins, lipids and nucleic acids and form cytotoxic and mutagenic advanced glycation end products.[88–90]

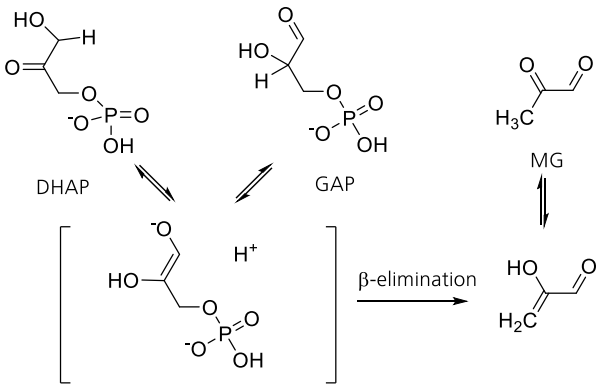


Figure 11. Formation of methylglyoxal from GAP or DHAP. After the formation of the enol form, spontaneous β-elimination of the phosphate can occur and methylglyoxal is formed.

One strategy Nature has evolved to deal with both oxidative and electrophilic toxic reactants is using low molecular weight thiols as scavengers. Enzymes co-evolved to improve the efficiency of such scavengers. The most well-known and best studied low molecular weight thiol, which is present in most aerobic organisms, is glutathione, the tripeptide of γ-*L*-glutamyl-*L*-cysteineglycine (Figure 12a).[88] Glutathione is synthesized in two steps from the amino acids *L*-glutamate, *L*-cysteine and *L*-glycine catalyzed by the two enzymes γ-glutamylcysteine synthetase and glutathione synthase (Figure 12b).[91]

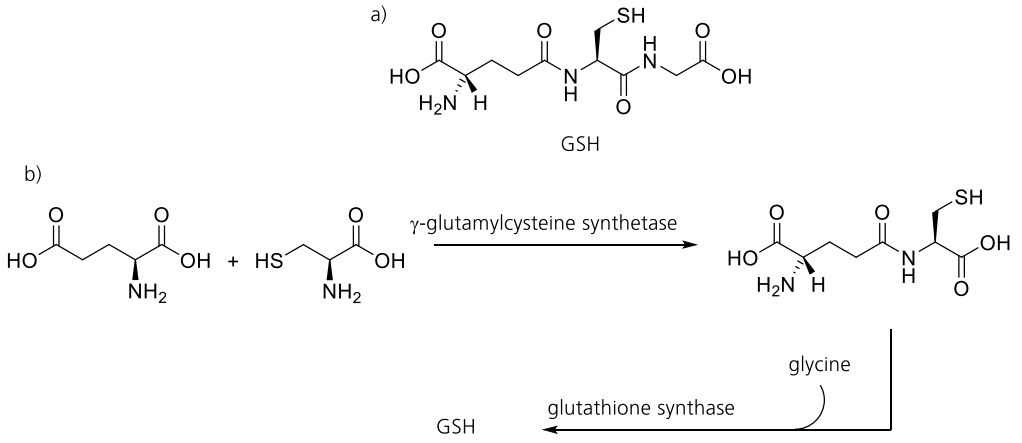
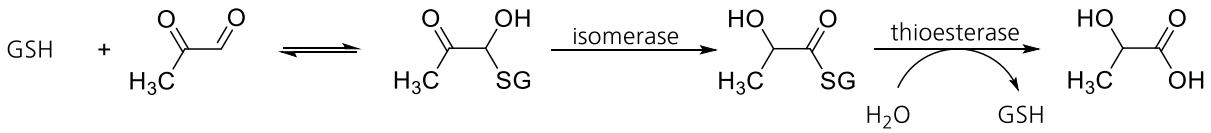


Figure 12. **a)** The structure of glutathione (γ-*L*-glutamyl-*L*-cysteineglycine). **b**) Biosynthesis of glutathione. In the first step γ-glutamylcysteine synthetase catalyzes the formation of γ-glutamylcysteine. In the second step the peptide bond of the γ-glutamylcysteine and glycine is formed by the enzyme glutathione synthase.

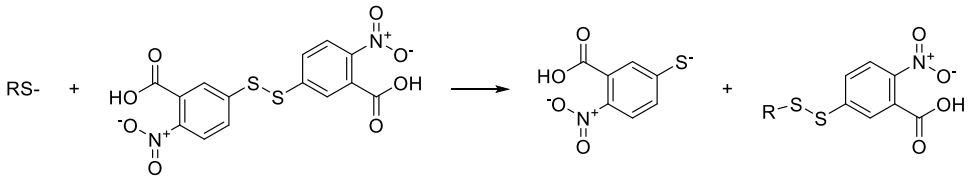
2.1.2 Biochemistry of Glutathione

The chemical properties of small molecular weight thiols are suitable for their function as electrophile scavengers. This can be illustrated on the example of glutathione and methylglyoxal. After the nucleophilic attack of the thiol group of glutathione to the aldehyde, an isomerization and a thioester transfer takes place to form the more stable and less reactive 2-hydroxycarboxylic acid (Scheme 11).[88]



Scheme 11. The GSH-dependent conversion of methylglyoxal to 2-hydroxycarboxylic acid. After the addition of glutathione, an isomerase and a thioesterase catalyze the transformation into the final product.

The pK_a values of alkylthiol groups are close to physiological pH, which means that they can be found in the deprotonated form. This is significant since the nucleophilicity, and therefore the reactivity, of thiolate is much higher than of thiols.[92] Whitesides and coworkers examined the reaction rate between thiols and Ellman's reagent (Scheme 12). They showed that the maximum rate of thiols to react as nucleophilies with the Ellman's reagent is when the pK_a of the thiols is close to the pH of the solution.[31] The pK_a value for glutathione is 9 and is therefore found predominantly protonated at physiological pH (>98%).[88] However, this pK_a is can easily be modified inside of the active sites of enzymes (see chapter 2.1.4).



Scheme 12. Reaction scheme between a thiolate residue and Ellman's reagent.

The use of cysteine derivatives as scavengers, instead of the free amino acid cysteine, has several advantages. Firstly, cysteine is a primary metabolite which is used in multiple pathways, for example in protein production. To have a separation of detoxification processes from cysteine-dependent anabolic processes is favorable for an organism.[88] Secondly, free cysteine is sensitive towards auto-oxidation to cystine.[30,93] The free amino group and the thiolate group can bind free metal ions (such as iron or copper) and catalyze electron transfer reactions, forming thiyl radicals, as well as ROS.[94] For example hydrogen peroxide can be formed, which can further react *via* the Fenton reaction to hydroxyl radicals and damage cell components.[93–95] This is also a reason why the intracellular cysteine concentration is low compared to other amino acids.[96,97] A solution to minimize the autooxidation properties of cysteine derivatives is to protect the amino group of cysteine (Figure 13).[88,98] In the case of glutathione, which is found in eukaryotes and most gram-negative bacteria, the carboxylic acid is protected by a glycine moiety and the amine group by a glutamate moiety.[99] Another low molecular weight thiol is mycothiol, which is found in Actinobacteria, and contains an acetylated amine and carboxylic acid group connected to a glucosamine-inositol dimer.[100] In the case of γ-glutamylcysteine which is found in halobacteria, the amine group is protected by a glutamate residue.[101]

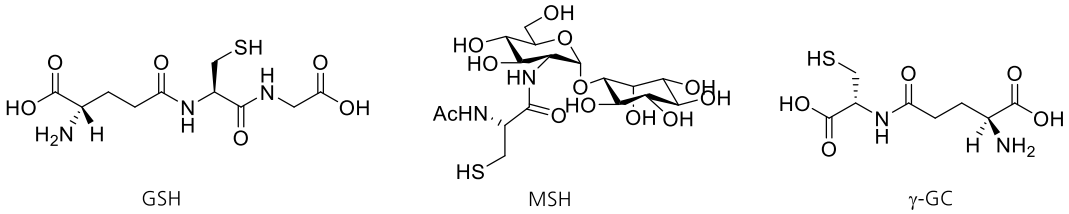
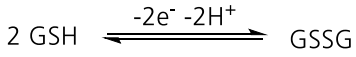


Figure 13. Structure of the cysteine derivatives glutathione (GSH), mycothiol (MSH) and γ-glutamylcysteine (γ-GC)

A further central chemical function of thiols is the ability to undergo redox chemistry. The oxidation of glutathione to glutathione disulfide is a reversible process (Scheme 13). The GSH/GSSG couple has an electrode potential of 0.26 V under physiological conditions and is used as a central redox buffer in the cell.[102]



Scheme 13: Redox reaction of glutathione

To activate glutathione as a scavenger, and to keep the redox potential constant, an efficient glutathione system has evolved.[88] A plethora of different enzymes are involved in the glutathione system, such as glutaredoxins, glutathione peroxidases, glutathione reductases and glutathione-*S*-transferases (Figure 14). Glutaredoxins are responsible for the reduction of oxidized molecules inside of the cell, for example unfavored disulfide bridges inside of proteins. Glutathione peroxidases, catalyze the reduction of free hydroperoxide to alcohol or, in the case of H_2O_2, to water. To maintain glutathione concentration inside of the cell the enzyme glutathione reductase is responsible for reducing glutathione disulfides. And glutathione-*S*-transferases are responsible for activating glutathione to catalyze the reactions with different electrophiles.

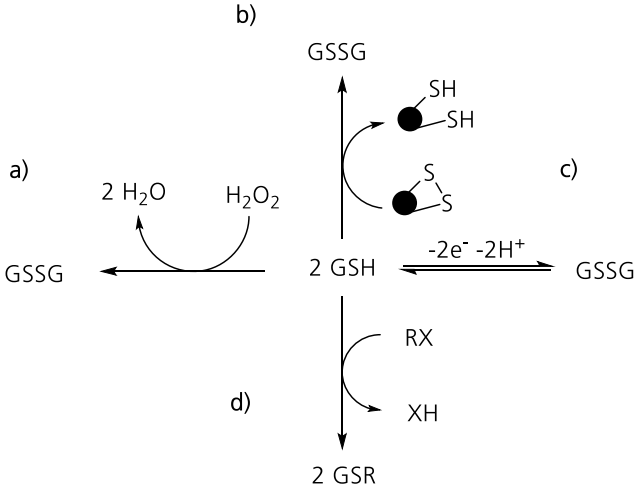


Figure 14. Four different reactions of the glutathione system a) H_2O_2 can be reduced to H_2O by glutathione peroxidases. b) Glutaredoxin catalyzes the reduction of sulfur disulfide bonds. c) The redox conversion from GSH to GSSG is catalyzed by glutathione reductases. d) The conjugation reaction of electrophiles with glutathione is catalyzed by glutathione S-transferases and MAPEG enzymes.

2.1.3 Glutathione *S*-Transferases

Glutathione *S*-transferases (GST) are involved in cellular detoxification processes catalyzing the conjugation of electrophilic substrates to glutathione. Further examples of GST with isomerization activity as well as peroxidase activity have been described previously.[88,103,104]

The first GST activity was discovered in 1961 in rat liver. Combes and Stakelum discovered a glutathione bromosulfophthalein adduct in presence of liver lysate.[105] In the same year, Booth *et al.* also observed the enzymatic formation of glutathione adducts with chloronitrobenzenes.[106] These observations, were followed by the discovery of multiple GSTs in mammals, plants, fungi and bacteria.[88,107] The GST, are separated into two entirely distinct enzyme families, the cytosolic soluble GST and the hydrophobic membrane bound GST.

The soluble GST family is categorized on the basis of their degree of sequence identity into eight classes; alpha, mu, pi, theta, zeta, omega, kappa and sigma.[108] The general structure of a GST contains an *N*-terminal thioredoxin fold, a four stranded antiparallel β-sheet sandwiched between three α-helices, while the *C*-terminus is variable. Glutathione binds within the *N*-terminal thioredoxin fold at a position denoted the G-site (Figure 15, left).[109]

The family of the membrane bound GST, which are also called MAPEG (membrane-associated proteins in eicosanoid and glutathione metabolism), is a distinct family of enzymes which are integrated into membrane components.[110] They are structurally unrelated to soluble GSH but show similar activity. The architecture of MAPEG enzymes consist of a homotrimer, where each monomer is composed of a four α-helices transmembrane segment (Figure 15, right).[111]

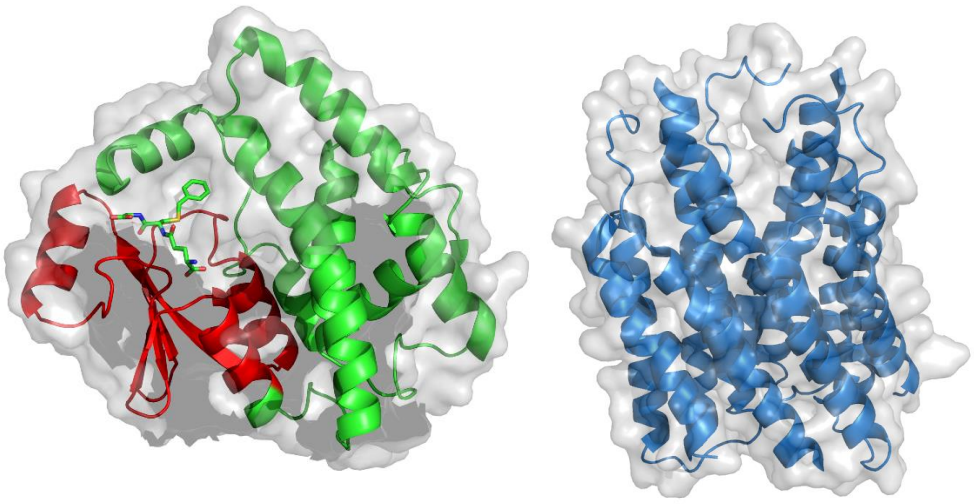


Figure 15. **Left:** Crystal structure of the human alpha class glutathione transferase A1-1 in complex with the ligand *S*-benzyl-glutathione. The thioredoxin fold is highlighted in red (PDB: 1guh, 2.6 Å).[109] **Right:** A microsomal rat glutathione transferase 1 of the MAPEG superfamily (PDB: 2h8a, 3.2 Å).[111]

2.1.4 Glutathione Activation in *S*-transferases

The glutathione-*S*-transferases described in the previous section make use of different catalytic mechanisms to increase the reactivity between glutathione and the target electrophiles and thereby increase the reaction rate. A common strategy is to activate the glutathione substrate is by lowering the pK_a of the thiol group to form the more reactive thiolate inside of the active site.[88,104,110] The thiolate can be stabilized by a hydrogen bond between the negatively charged thiolate and in most cases a nearby serine or tyrosine residue. In the rat glutathione *S*-Transferase M1-1, Tyr6 was determined to be the hydrogen bond donor within the active site.[112] In a mutational study, Tyr6 (pK_a of 10) was replaced with a 3-fluorotyrosine (pK_a of 7.5).[112,113] Due to the lower pK_a, 3-fluorotyrosine was expected to be deprotonated at physiological pH and should act as a general base instead of electrophilically stabilizing the thiolate anion. This hypothesis was supported by the large inverse kinetic solvent isotope effect (KSIE) of the mutant (KSIE = 0.5) compared to the wildtype (KSIE = 0.9) in a reaction with GSH and chlorodinitrobenzyl (CDNB). These observations point to a proton transfer in the mutant while the modest isotope effect in the wildtype points to a hydrogen bond stabilization rather than proton transfer (Figure 16a).[113]

Another method to stabilize the thiolate form is the position of a positively charged electrostatic field inside of the active site next to the thiol group of glutathione. If, for example, a positively charged arginine residue is located next to the thiol group, the deprotonated thiolate is stabilized and therefore favored (Figure 16b). This was observed in the human MAPEG enzyme LTC4 synthase which conjugates glutathione to the unstable epoxide leukotriene A_4 to form leukotriene C_4.[114] In the crystal structure, R104 was observed to be pointing directly towards the thiol group in the co-crystallized glutathione at a short distance of 3.3 Å (PDB: 4ntb, 2.7 Å).[114]

An alternative strategy to increase the reactivity of the nucleophile, is to instead activate the electrophilic substrate, for example with a Lewis acid. Such a strategy has been observed in the fosfomycin resistance protein FosA, where a Mn^{2+} ion coordinates the epoxide ring of fosfomycin, an antibiotic compound produced by *Streptomyces*.[115,116] The coordination of the Mn^{2+} ion accepts electron density and thereby makes the carbons of the epoxide ring more electrophilic. This facilitates the attack of the nucleophilic thiol group of glutathione (Figure 16c).[116]

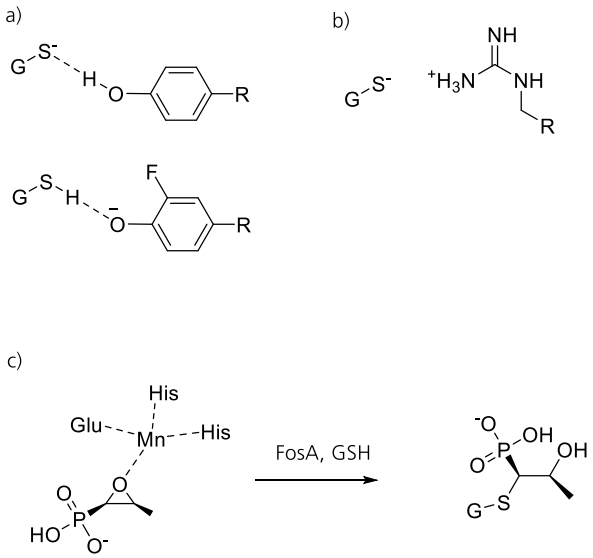


Figure 16. **a)** Interaction of tyrosine and 3-fluorotyrosine in the active site of Rat glutathione *S*transferase M1-1, showing two mechanisms. The electrostatic stabilization of the wildtype and the base induced deprotonation with the unnatural amino acid.[112] **b)** The positive charge of Arg105 forms a electrostatic field in the active site of LTC4 synthase to stabilize the thiolate form.[114] **c)** The epoxide ring of fosfomycin is activated by a manganese and thereby facilitates the nucleophilic attack of glutathione in the enzyme FosA.[116]

The activities of many glutathione *S*-transferase were identified with use of the artificial aromatic electrophile CDNB. The catalytic mechanism for this aromatic sp^2-hybridized electrophile can be described as a nucleophilic aromatic substitution S_{NAr} which involves the formation of a σ-complex intermediate. Even though the binding mechanism has not been described for every glutathione *S*-transferase, a possible mechanism has been proposed (Figure 17).[88] In the first step glutathione binds to the glutathione binding site of the enzyme (G-site). The hydrogen bonding network inside of the active site lowers the pK_a of glutathione from 9 to 5.2-6.8 thereby favoring the thiolate form of the substrate.[88] Afterwards, the electrophiless enter the active site into a second binding site, called the H-site. The nucleophilic addition of the activated glutathione to the electrophile CDNB, leads to a σ-complex (cyclohexadienyl cation) intermediate. This intermediate dissociates by an elimination reaction to form the final product. In the last step the glutathione-electrophile adduct is released from the binding site.

Another commonly used electrophile is methyl iodide. For this sp^3-hybridized substrate the proposed mechanism is an S_N2 reaction. The attack of the electrophile by glutathione and the release of the leaving group is concerted (Figure 17).[88]

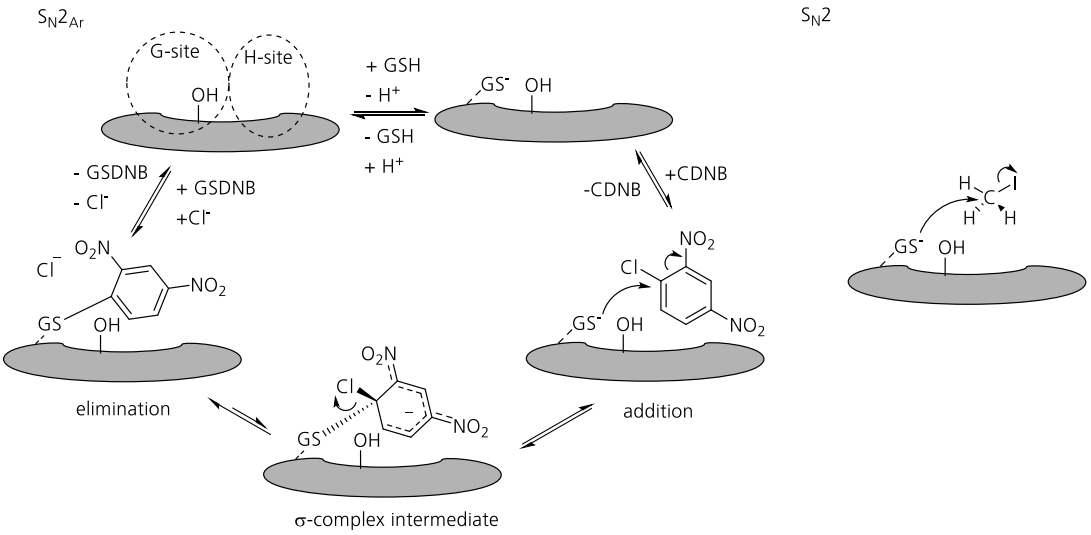


Figure 17. **Left:** Model for an S_{NAr} mechanism of a GST with CDNB. The mechanism has a separated addition and elimination step with the σ-complex intermediate which is formed. **Right:** Model for an S_N2-like mechanism with methyl iodide.[88]

2.1.5 DinB2: A New Class of Glutathione *S*-Transferase

In 2011 Newton *et al.* discovered a glutathione *S*-transferase EF_3021 from *Enterococcus faecalis*.[117] This enzyme was discovered because of its structural similarity to a newly found mycothiol *S*-transferase.[117,118] The biological function of EF_3021 is unknown, nevertheless the glutathione specificity could be established with monochlorobimane (mBCl) a fluorophoric electrophilic reagent, which is used for detection of *S*-transferase activity.[117] The apo crystal structure (PDB: 3cex, Figure 18), was solved in a structural genomic program, and the amino acid sequence places the enzyme in the DinB-like superfamily. The DinB superfamily has it name from the DNA-damage induced gene DinB found in *Bacillus subtilis* and consists of eight subfamilies (DinB, DinB_2, DUF1569, DUF1572; DUF1706, DUF1993, DUF664, MDMPI_N). The most characteristic features of the DinB enzymes are the four-helix up-down-down-up bundle structure and a predicted metal binding in the active site.[118] The metal binding site is often a conserved histidine-triad motif. Nevertheless, in some cases histidines have been replaced by aspartates and glutamates. Furthermore, the DinB enzymes form dimeric structures in most cases.[118]

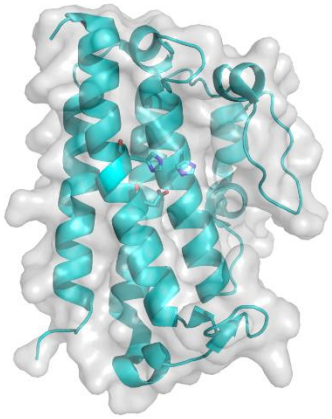


Figure 18. The apo crystal structure of EF_3021 with the His-His-Asp metal binding triad highlighted (PDB 3cex, 2.0 Å).

This new discovered glutathione *S*-transferase EF_3021 has remarkable differences to previously described glutathione *S*-transferases and MAPEG enzymes. EF_3021 has an unrelated protein fold compared to other glutathione *S*-transferases and it has a metal binding site.[117] The activation of the sulfur of a substrate with a metal is a completely different catalytic strategy than those previously described for glutathione *S*-transferases.

Enzymes of the DinB-like superfamily are involved in the formation of different C-S bonds with low molecular weight thiols (Figure 19). Next to the GST EF_3021, several other examples with known and unknown function have been reported.[117,118,60,119] A mycothiol-dependent maleylpyruvate isomerase with a DinB domain was discovered in *Corynebacterium glutamicum*.[120] The DinB2 protein LmbV is involved in the biosynthesis of lincomycin A and catalyzes a thiol exchange of the lincosamide unit, in which mycothiol is replaced by ergothioneine.[119] Another enzyme which has been extensively studied in our research group is the EgtB enzyme, a sulfoxide synthase which is involved in the ergothioneine biosynthesis. EgtB catalyzes the radical induced formation of a sulfoxide out of the substrates trimethylhistidine and γ-glutamylcysteine. It consist of two domains, a FGE-sulfatase like and a DinB domain.[60]

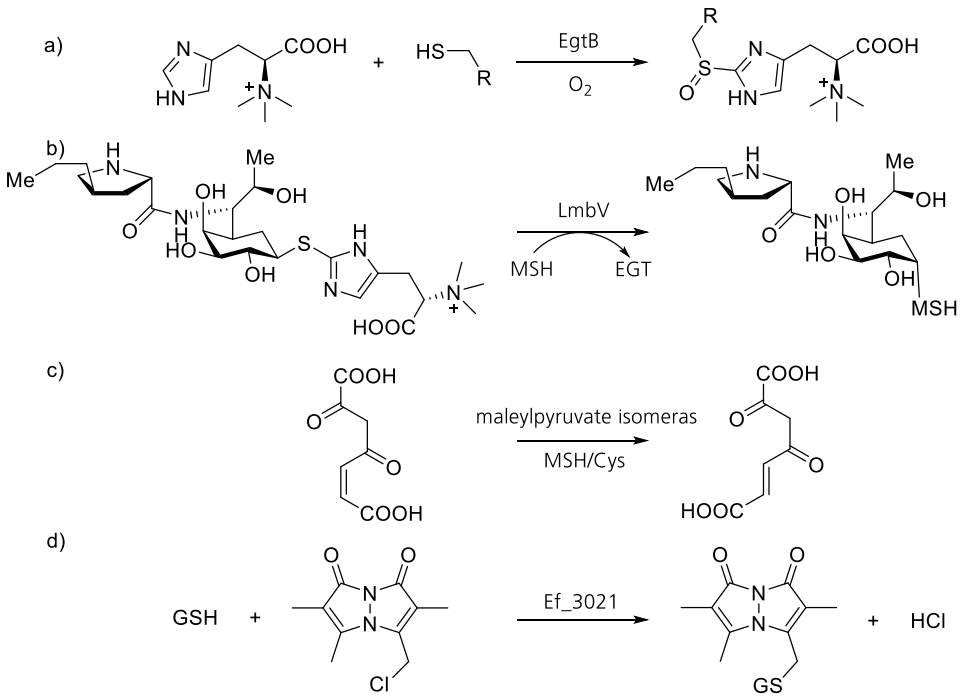


Figure 19. Examples of enzymes with known functions that contain a DinB domain. **a)** In ergothioneine biosynthesis, the iron-dependent EgtB catalyzes the formation of a sulfoxide from trimethylhistidine, γ-glutamylcysteine and molecular oxygen.[60] **b)** In lincomycin A biosynthesis, LmbV catalyzes the replacement of an ergothioneine moiety with a mycothiol moiety.[119] **c)** A cysteine-dependent as well as a mycothiol-dependent maleylpyruvate isomerase containing a DinB domain were identified.[120,121] **d)** The glutathione specificity of EF_3021 was established with monochlorobimane.[117]

A remarkable feature of the EF_3021 enzyme is the fold similarity to the iron-dependent EgtB. By comparing the two DinB domains a high structural accordance to each other is visible (Figure 20). The root-mean-square deviation (RMSD) was calculated by the program PyMOL to be 2.252 Å. This raises the question of the functional and evolutionary relationship between the two enzymes.

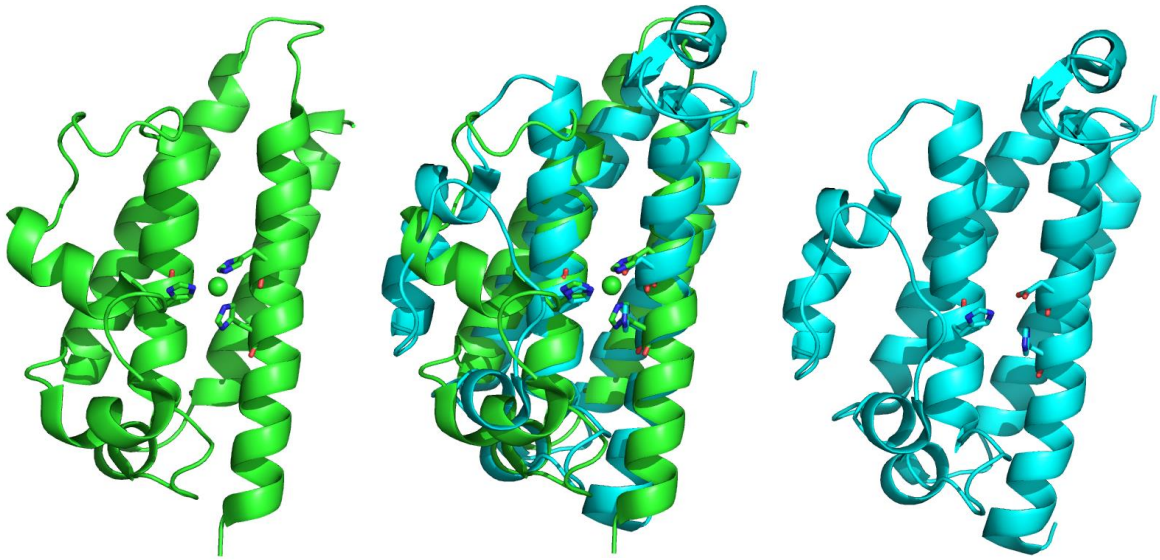


Figure 20. **Left** DinB domain (green) of EgtB from *Mycobacterium thermoresistibile* (PDB: 4x8d, 1.98 Å) **Middle**: Superposition of both domains show high structural similarity **Right**: DinB domain (cyan) of EF_3021 (PDB: 3cex, 2.0 Å).

2.1.6 Activation of Thiols with Metal Ions

The activation of sulfur-metal bonds has been observed in a variety of proteins and enzymes. The activation is directly influenced by multiple factors including the ligand set, the charge of the metal site, hydrogen bonding and the metal identity.

A historical example of one of the first and best studied sulfur activation systems is the Ada protein, a protein which is responsible for the reparation of DNA alkylation damages. It contains a complex of four cysteines bound to a zinc ion in the active site.[2]

The Ada enzyme of *E. coli* consist of two domains, a 20-kDa *N*-terminal domain and a 19 kDa *C*-terminal domain. The *N*-terminal domain contains four cysteine residues which coordinate to a zinc-ion (Figure 21). This domain is responsible for the repair of methylated phosphotriesters by irreversibly transferring the methyl group to a cysteine from the active site (Figure 21).[2] The mechanism of this protein was studied by Lippard *et al.* with a simplified model system. Lippard and coworkers used the $[(CH_3)_4N]_2[Zn(SC_6H_5)_4]$ complexes to mimic the protein and $(CH_3O)_3PO$ as a substrates (Scheme 14).[122] The main mechanistic question was whether the thiolate, which attacks the phosphotriester was bound to the metal (associative mechanism) or whether the thiolate first dissociates from the metal and then attacks the electrophile (dissociative mechanism). The ion pairing, ligand dissociation and the combination of both effects were measured using ^{1}H NMR spectroscopy in deuterated DMSO. A high rate of ligand exchange in solution was observed. Furthermore, the metal-bound thiolate moiety shows decreased nucleophilicity relative to that of free thiolates. Therefore, they proposed that the active nucleophiles in their system are dissociated thiolate

ligands. Based on these studies, they proposed a dissociative reaction mechanism in the model system, as well as in the Ada protein.

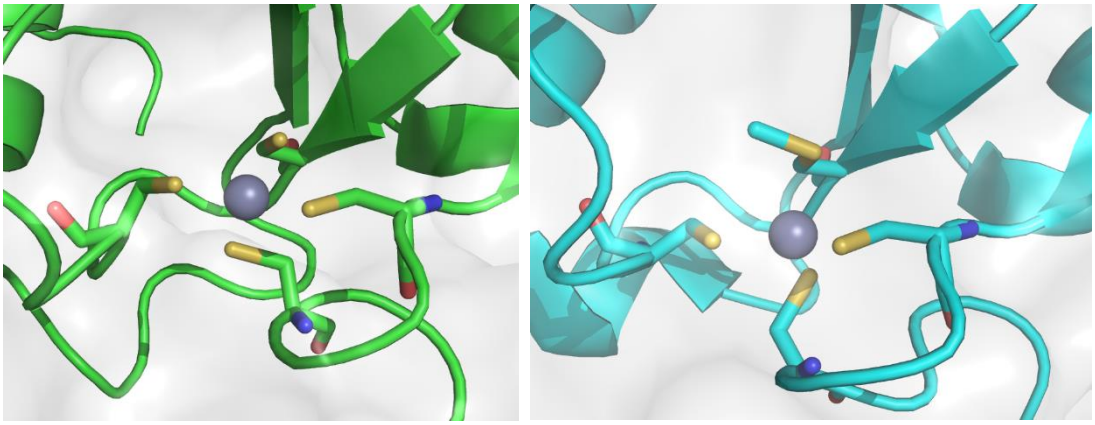
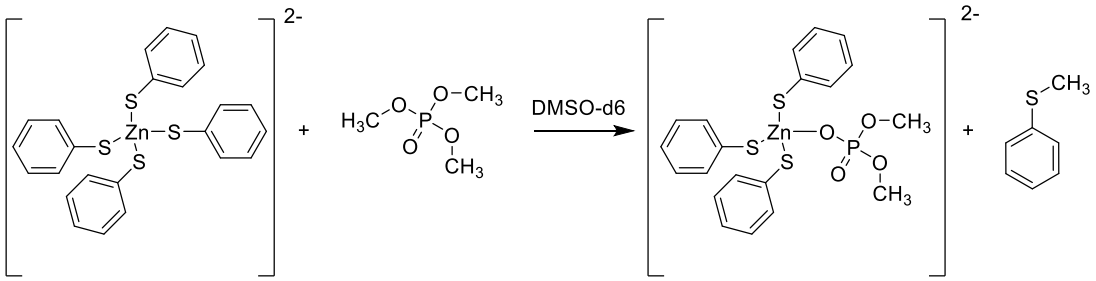


Figure 21. **Left:** View of the active site of Ada from *E. coli* with the zinc atom and the four cysteine ligands before alkylation (PDB 1eyf). **Right:** Active site of the same protein after alkylation reaction (PDB: 1u8b 2.1 Å).



Scheme 14. Model system for the study of the alkylation reaction of the Ada enzyme. $[(CH_3)_4N]_2[Zn(SC_6H_5)_4$ reacts in deuterated DMSO with $(CH_3O)_3PO$ to form the alkylated product.

In metalloenzymes, the metal ion is often coordinated to a ligand amino acid, such as histidine, cysteine and aspartate/glutamate. The influence of the ligand set was measured in a simplified model system of S-alkylation reactions. The rate of the reaction between a tetrahedral zinc-thiolate complex with methyl iodide was measured. By changing one of the zinc ligands from O to N to S.[123] In a second study, a similar alkylation reaction between methyl iodide and a zinc complex containing different numbers of S-donors $[ZnN_3S]^0$, $[ZnN_2S_2]^0$, $[ZnNS_3]^0$, $[ZnS_4]^0$ was measured. It was shown that the second order rate constant increased proportionally to the number of S donor ligands. An explanation could be that the electron density at the zinc metal ion increases with the sulfur content of the ligand. This lead to a more ionic and therefore more reactive nature of the thiolate ligands.[124]

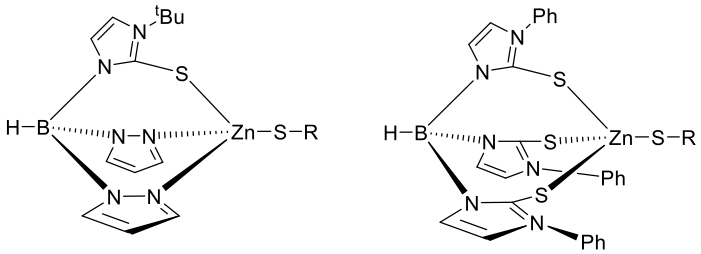


Figure 22. The chemical structure of the complex $[ZnN_2S_2]^0$ (left) and $[ZnS_4]^0$ (right).

Another factor which influences the activity of the thiolate–metal bond is the charge of the complex. In the ligand series $[ZnS_4]^{2-}$, $[ZnS_3N]^-$ and $[ZNS_2N_2]^0$ the highest alkylation activity was measured in the most negatively charged complex and the lowest in the neutral complex. The negative charge of the complex is increases the electron density of the thiolate resulting in a higher nucleophilicity.[125]

The reactivity of a metal-bound thiolate complex can be further altered with hydrogen bonds. It is known that in the active site of enzymes a complex hydrogen-bond network can exist, therefore it is important to understand the effects of hydrogen bonding. To probe the influence of a hydrogen bond to zinc-thiolate complexes Chiou *et al.* characterized the reaction rate of two different mononuclear zinc-thiolate complexes (Figure 23). The alkylation reaction of the complexes with iodomethane or α-bromo-toluene was monitored by ^{1}H-NMR. The obtained results showed a decrease of reactivity by more than an order of magnitude due to the additional amide N-H-S hydrogen bond.[126] These results are in agreement with similar studies and show that the nucleophilicity is directly influenced by hydrogen bonds.[127–129]

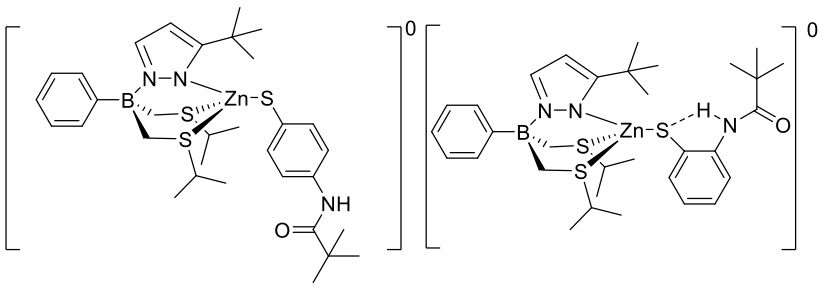


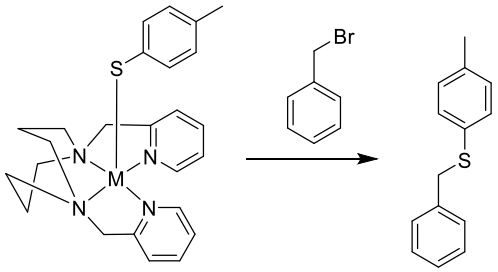
Figure 23. Zinc complexes were used to measure the influence of hydrogen bonding towards the thiolate. **Left:** Zinc-complex without a hydrogen bond. **Right:** A zinc complex with a hydrogen bond to a backbone amide group.[126]

Even with the limited presented studies of different model systems it is possible to recognize the numerous possibilities an enzyme has to modify and tune the active metal site for a certain reaction. Nevertheless, the most important aspect has not yet been discussed, the identity of the metal ion.

2.1.7 The Influence of the Central-Metals in Thiol Alkylation

A problem with the binding of a thiolate to a transition metal is the decrease of charge density at the sulfur atom which should lead to a decreased nucleophilicity. This can be explained by the Lewis acid/Lewis base concept, where the filled thiolate orbital can interact with the empty d orbitals of the metal. Since the Lewis acidity of each metal differs, we expect the metal identity to have a large impact on thiolate activation.

In 2004, Fox *et al.* carried out a study with a model system to investigate the influence of different divalent transition metals towards sulfur-activation in alkylation reactions.[130] A 1,5-bis(2-pyridylmethyl)-1,4-diazacyclooctane (L^8py_2) ligand, *p*-toluenethiolate and benzyl bromide were used for the model system as shown in Scheme 15.



M = Fe^{2+}, Co^{2+}, Ni^{2+}, Zn^{2+}

Scheme 15. The model reaction used by Fox *et al.* with the 1,5-bis(2-pyridylmethyl)-1,4-diazacyclooctane (L^8py_2) metal ligand, *p*-toluenethiolate and the second adduct benzyl bromide which form together the thioether alkylation product.

The metal complexes were synthesized using the tetradentate ligand L^8py_2 and hydrated metal(II) perchlorate salts. These precursors were then mixed with sodium *p*-toluenethiolate to result in the metal(II)-thiolate complexes. The final complexes were crystalized and x-ray crystal structures were obtained. All metal complexes showed a pentacoordinate metal ion in a distorted square-pyramidal geometry. In the complexes with Ni and Co a thiolate-to-metal charge transfer (CT) could be observed with an absorption shift to a lower energy.

The second order kinetics of the reaction with the metal(II)-thiolate complex and benzyl bromide was measured at 30°C in acetonitrile. The observed second order constants (k_2) are shown in Table 1. Interestingly, the reaction rate k_2 for M = Fe(II) and Co(II) are one order of magnitude smaller, and therefore slower, than k_2 for M = Ni(II) and Zn(II). For the metal(II)-thiol complex an associative mechanism was proposed, since there was no ligand exchange visible in the ^{1}H-NMR. In the NMR experiment a clear complex signal is visible if the ligand exchange is slow, while a brought signal is observed if the ligand exchange of a complex is fast.

Table 1: Second order rate-constants of the model reaction with different metals.

$[L^8py_2M(SAr)]^+$	k_2, $M^{-1}s^{-1}$
M = Fe	5×10^{-3}
M = Co	3×10^{-3}
M = Ni	3×10^{-2}
M = Zn	3×10^{-2}
NaSAr	$>>10^{-2}$

In order to explain the observation that the alkylation reaction with zinc and nickel is ten times faster than the alkylation reaction with cobalt and iron, density functional theory (DFT) calculations were performed for each metal complex. The highest-occupied molecular orbitals (HOMO) which are relevant for the calculation are the two 3p orbitals of the thiolate that interact with the metal d orbitals. The 3p orbitals of the sulfur atom in *p*-toluenethiolate, also labeled as π orbitals, are oriented in the plane of the aromatic ring ($\pi_{in\ plane} = \pi_{ip}$) or perpendicular to this plane ($\pi_{orthogonal} = \pi_{op}$) (Figure 12). The π_{ip} interacts with the d_{yz} orbital of the metal while the π_{op} ineracts with the d_{z2} orbital of the metal.

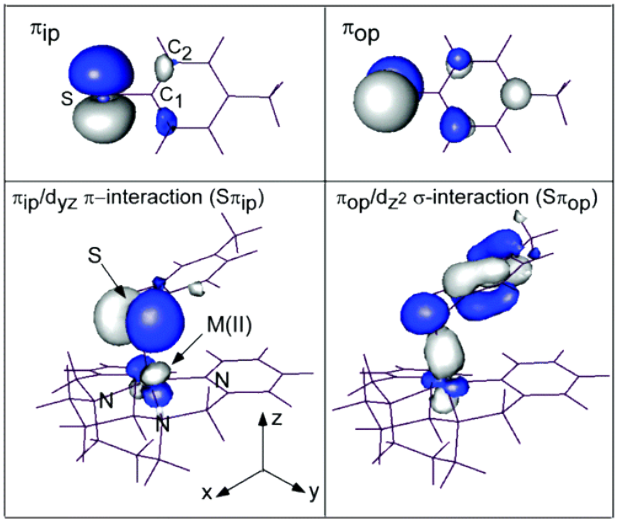


Figure 24. the π_{ip} and π_{op} orbital of the thiolate and the interaction with those orbitals and the d_{yz} and d_{z2} orbital of the metal.[130]

There are three electronic effects which influence the sulfur-metal bond in a way that could explain the different rates of the alkylation reaction with different metals. The first effect is due to the change in the charge of the metal nucleus. In the periodic table, from left to right, the molecule orbitals (MOs) of the

metal d orbitals decrease in energy relative to the thiolate π-based MO (Figure 25). This means that the gap between thiolate π-based MO and metal d-based MO is diminishing. This energy gap defines the covalency of the bond. This is in good accordance with the observed thiolate-to-metal charge transfer for Co and Ni complexes. In the case of the Zn(II) complex, the situation is different since the HOMOs have thiolate character and the metal d-based MOs are much lower in energy.

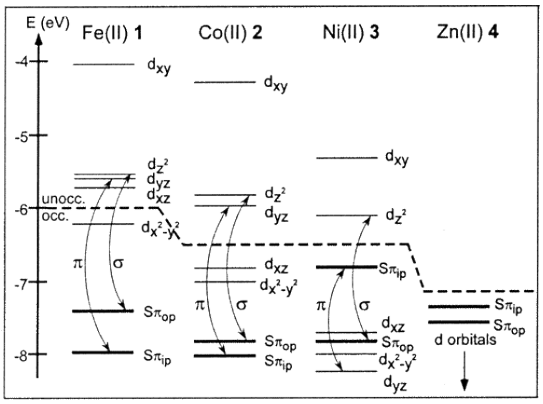


Figure 25. DFT calculated orbital energies of the metal-thiolate complexes. The pairs of the MO between the metal σ and sulfur π orbital which show dominant sulfur-metal interaction are connected by arrows.[130]

The second effect is based on the number of electrons which are involved in the interaction between ligand orbitals and metal d orbitals. If the interaction of the filled ligand orbital is with a half-filled metal d orbital, a transfer of electron density onto the metal occurs, which has a stabilizing effect towards the ligand-based orbitals (Figure 26A). However, if in addition to the ligand orbital the metal d orbitals are also doubly occupied, no net transfer of electron density occurs (Figure 26B). In addition, if the ligand orbital is higher in energy than the metal d orbitals, the ligand orbitals are destabilized. In this case the nucleophilicity of the complex is increased by the so called "filled/filled" interactions (Figure 26C). For the Fe(II) complex with (d^6 configuration) and for the Co(II) complex (d^7 configuration), the thiolate π MO are stabilized by the half-filled d orbitals of the metals. In the Ni(II) complex (d^8 configuration) both the ligand and the metal orbitals are fully occupied which leads to a filled/filled interaction. This destabilization of the thiolate orbitals leads to a higher nucleophilicity of the ligand and therefore to a higher reactivity for alkylation.

In the complex with Zn(II) (d^{10} configuration) the completely filled metal orbitals are much lower in energy than the sulfur-based HOMO (Figure 26D). Therefore, almost no interaction is taking place and the Zn-S bond has dominant ionic character. This ionic bond increases the negative charge on the sulfur and therefore increases the nucleophilicity.

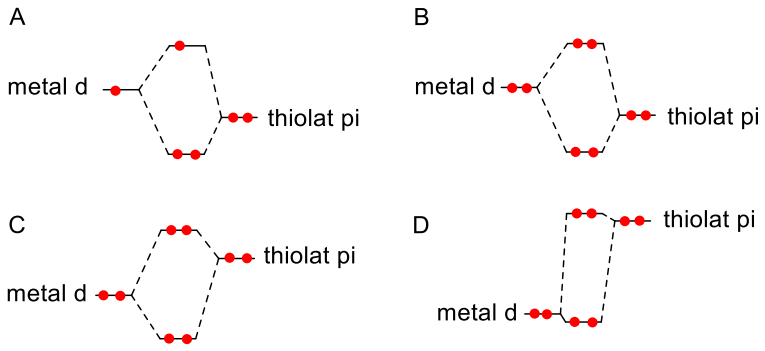


Figure 26. Transition metal-thiolate orbital interaction. A) Normal metal-ligand interactions. B) Filled/filled interaction with a destabilized metal-based molecular orbital. C) Filled/filled interaction with a destabilized thiolate-based moleculare orbital. D) Minimal interaction of the metal based closed shell d orbitals and the ligand based π orbitals.

The third effect is the spin of the complex which changes across the transition metal series. It is difficult to estimate the effect of the spin towards the reaction rate. In the results of Fox *et al.,* a correlation between spin density and reactivity was not observed.[130]

2.1.8 Mechanism of Thiol-Metal Alkylation

To address the question of the mechanism for metal activated sulfur alkylation, three possible mechanisms were proposed for the reaction with the electrophile methyl iodide.[129] The first mechanism is an S_N2 dissociative pathway. The thiolate dissociates from the metal complex and reacts then as the reactive species with the alkylation agent CH_3I. This reaction takes places via the S_N2 like transition state 1 (TS 1, Figure 27A). In an associative mechanism, the alkylation occurs directly between the thiolate-metal and the electrophile (TS 2, Figure 27B). This leads to the zinc-bound thioether which then dissociates to the final alkylation product. The third possible pathway is the formation of a four-center transition state in which the thiolate alkylation and the zinc thiolate dissociation occur simultaneously (TS 3, Figure 27C). This associative reaction is called σ-bond metathesis.[129,131]

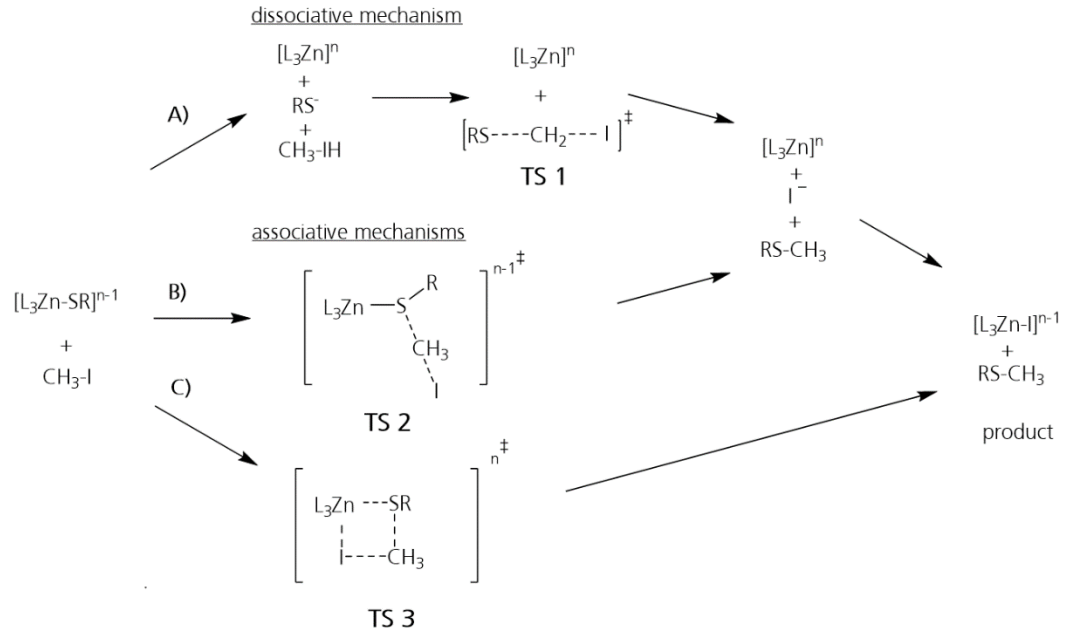


Figure 27. The three possible mechanisms. **A)** Dissociative mechanism in which the first step is the dissociation of the thiolate from the metal complex. The thiolate further reacts with the methyliodide to form the final alkylation product. **B)** The associative mechanism where nucleophile substitution occurs between the electrophile and the thiolate which is bound to the metal. **C)** An associative mechanism over a four-center transition state (σ-bond metathesis).

In the literature, there are many examples are found which suggest alkylation of thiolate complexes is by an associative mechanism.[129,130,132] In addition to different model systems, mechanistic studies with enzymes like farnesyltransferase or Ncp7 zinc finger were also in agreement with this mechanism.[132,133] In contrast to these findings, Wilker and Lippard showed by ^{1}H-NMR experiments that the double negatively charged complex $Zn(SPh)_4^{2-}$ undergoes alkylation as a dissociated thiolate.[122] Furthermore, the Ada DNA repair protein processes the catalyzed thiol alkylation *via* a dissociative mechanism to form a free thiolate as reactive species.[125,134]

To overcome this inconsistent reported data and to make a link between the structure of the zinc complexes, their reactivity and their alkylation mechanism, Picot *et al.* performed calculation studies with a thiolate alkylation model system. In their computer model, they let the complexes shown in Figure 28 react with methyl iodide and calculated the transition state energies of a possible associative and dissociative transition state. In their calculated models, they could exclude a σ-bond metathesis mechanism because this pathway is by far the least favored mechanism in their gas, as well as their aqueous model.

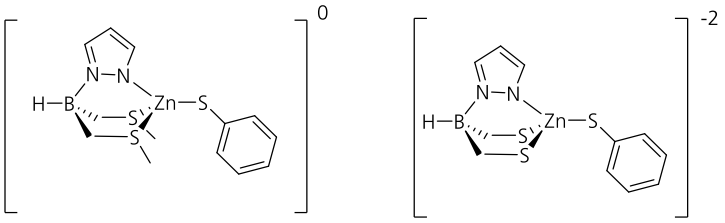


Figure 28. The neutral and the anionic thiolate-zinc complex used for the mechanistic alkylation studies of Picot *et al.*[129]

The calculated free energy of the dissociative intermediates and the transition state of the dissociative and the associative complexes are displayed in Figure 29. The calculations were performed in an aqueous solution model and show that the alkylation of zinc-bound thiolate reactions are exothermic. The calculations indicate that in the neutral complex the associative mechanism is most competitive-, while in the double negatively charged complex the reaction follows the dissociative mechanism. These calculations can be explained with the dissociation energy. The dissociation energy for a neutral core is much higher than for a negatively charged core. Even though the free thiolate would be more reactive than the zinc-thiolate complex, the negative effect for the dissociation of the thiolate favors the associative mechanism of the neutral complex. In the electron rich complex, the zinc-thiolate interaction is weaker and favors the dissociation of the more reactive thiolate.

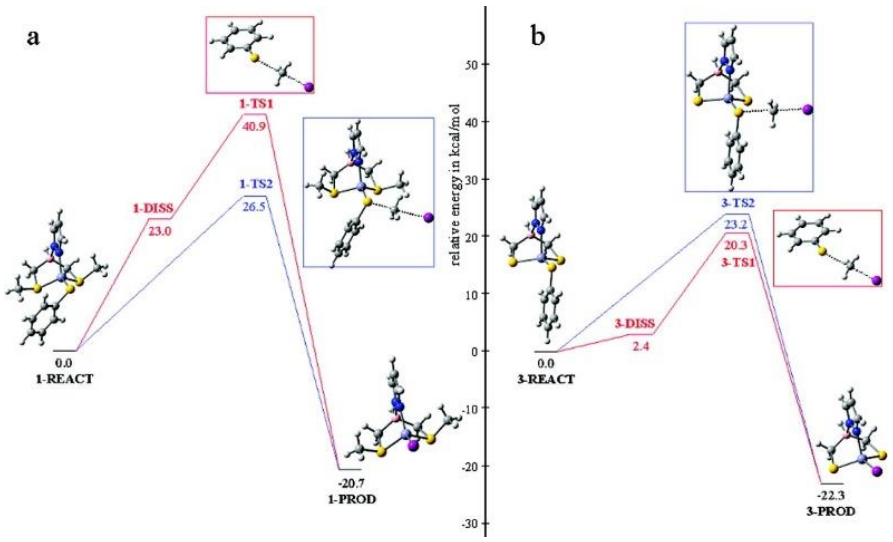


Figure 29. **a)** Free-energy profiles in an aqueous solution for the model neutral complex. The associative mechanism (blue) is lower in energy and therefore favored over the dissociative mechanism(red). **b)** In the negatively charged complex the dissociation step has only a low energy barrier and therefore the whole mechanism is energetically favored over the associative mechanism.[129]

2.2 Aim of the Chapter

The EF_3021 of *Enterococcus faecalis* is a metal-dependent glutathione *S*-transferase comprised of a DinB2 domain. This domain is structurally related to the DinB domain of the sulfoxide synthase EgtB. Both enzymes are involved in the formation of a new carbon-sulfur (C-S) bond. This bond formation is catalyzed in both enzymes by a metal-thiol activation. However, the reactions follow different chemical pathways. EgtB from *Mycobacteria thermoresistibile* is a non-heme iron-dependent enzyme and catalyzes C-S bond formation, involving a 4-electron oxidation reaction, through a radical mechanism- while EF_3021 uses metals to form an active thiolate which most likely reacts as a nucleophile in an S_N2 reaction with an electrophilic co-substrate. [60] The protein EF_3021 only consist of the DinB2 domain while the protein EgtB is a fusion between a DinB and an FGE-like domain. This raises questions of evolutionary relationships between the simpler GST and the more specific EgtB.

With the recombinant gene of EF_3021 in hand, we produced the enzyme in *E. coli*, functionally characterized it *in vitro* and probed the influence of different metals. To underline our kinetic findings, we wanted to structurally characterize the enzyme with x-ray crystallography. We aimed to identify the glutathione binding residues, as well as define the geometry of the metal binding site, with the crystal structures of the enzyme in the presence of different metals and the native substrate glutathione.

2.3 Results and Discussion

2.3.1 Development of a Study System for EF_3021

The codon optimized gene of EF_3021 was purchased from Genscript and cloned into a pET19 production vector. The enzyme EF_3021 from *Enterococcus faecalis* was produced in *Escherichia coli* BL21 cells yielding 5-7 mg/L of purified enzyme.

The native substrate of EF_3021 next to glutathione is unknown (see below 2.3.8). The glutathione activity was discovered with the generic electrophile monochlorobimane.[117] We were interested in the kinetic characterization of EF_3021 and therefore decided to replace the rather expensive monochlorobimane with a low-cost and UV-detectable benzyl halogen (Figure 30).

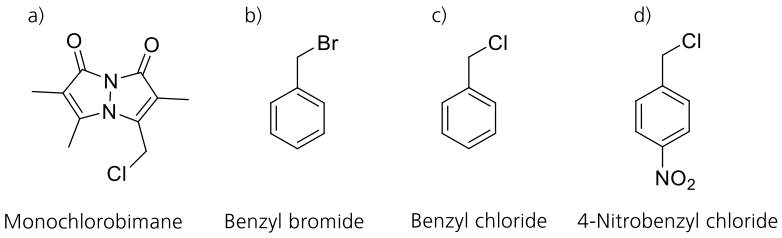
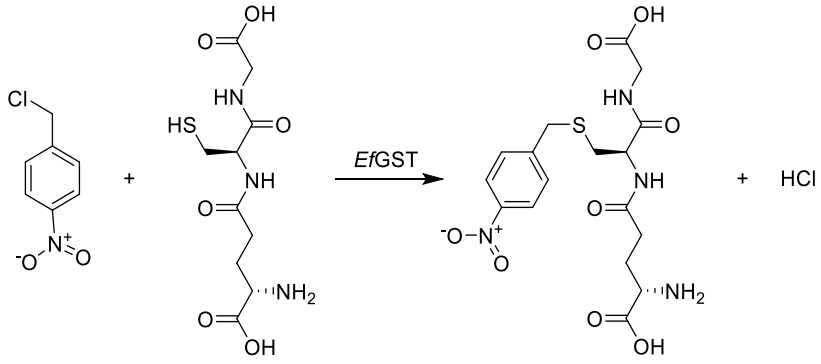


Figure 30. The structure of monochlorobimane and halogen benzyl compounds, which were tested as electrophiles. **a)** Monochlorobimane, **b)** Benzyl bromide, **c)** Benzyl chloride, **d)** 4-Nitrobenzyl chloride.

First, we chose the compound benzyl bromide. The electrophilicity of this compound was so strong that we observed a fast-non-enzymatic spontaneous reaction with glutathione. Therefore, we decided to lower the activity of the electrophile by changing to the compound benzyl chloride. The low solubility (0.5 g/l at 20°C) in aqueous solution proved to be unsuitable for the designed HPLC assay in a buffer system containing only 10% acetonitrile.[135] Therefore we changed the electrophile to the more soluble 4-nitrobenzylchloride (Scheme 16). 4-Nitrobenzylchloride has a UV-Vis absorption maximum of 272 nm and a measured extinction coefficient of $\varepsilon = 11.5 \pm 2.0 \times 10^3$ $cm^{-1}M^{-1}$. The product formation was monitored in an HPLC assay at an absorbance of 260 nm.



Scheme 16. Model reaction of EF_3021 with 4-nitrobenzylchloride and glutathione.

Newton *et al.* described EF_3021 as a glutathione transferase with selective glutathione activity. No activity with other low molecular weight sulfur molecules, like mycothiol and bacillithiol was observed.[117] To screen the sulfur source selectivity of this enzyme the following thiol containing compounds were tested: glutathione, *L*-cysteine, γ-glutamylcysteine, *N*-acethyl cysteine and *D/L*-homocysteine. The enzyme catalyzed only the formation of glutathione addition product as was observed in these measurements.

In the first kinetic characterization we used the recombinant produced EF_3021 containing a His_6-tag used for the Ni-affinity column in the purification step. With this enzyme and the substrates glutathione and 4-nitrobenzyl chloride a maximal reaction rate of 0.0078 ± 0.0002 min^{-1} was measured. By removing the His_6-tag of the enzyme with the recombinant TEV-protease the reaction rate of the enzyme increased 72-fold to 0.56 ± 0.01 min^{-1}. Consequently, all further experiments were measured with His_6-tag free enzymes. A possible explanation for the decreased activity of the enzyme with the His_6-tag could be that the quaternary structure of the enzyme, which is relevant for catalysis is disturbed.

The quaternary structure of the enzyme was analyzed by size-exclusion chromatography (Figure 31). The retention volume showed that the enzyme is found in solution as a dimer. The enzyme eluted also as a dimer after treatment with the reducing reactant dithiothreitol (DTT). This indicates that the dimer-interaction is not mediated by an intramolecular disulfide bond. The observation of a dimeric structure supports the possibility that additional amino acid residues like a His_6-tag can have negative effects towards the stability.

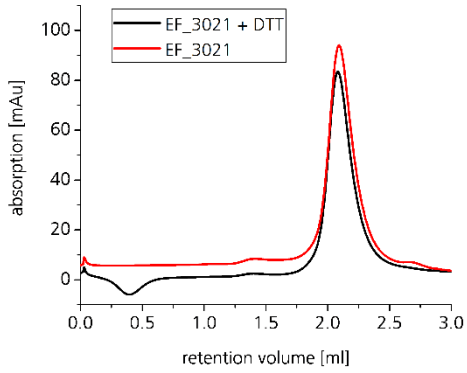
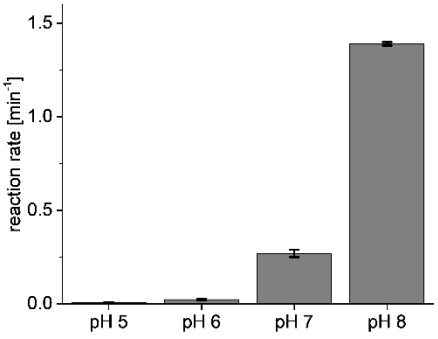


Figure 31. Analytical size exclusion chromatography of EF_3021 with and without the reducing agent DTT.

The assay conditions were similar to the condition used in the monochlorobimane measurements.[117] We used 50 mM NaCl and 50 mM 2-(4-(2-hydroxyethyl)piperazin-1-yl)ethane-1-sulfonic acid (HEPES) buffer at a pH of 7. The p*K*a of glutathione in aqueous solution is 9 which means that the majority (ca. 99%) of the molecules are protonated.[136] Nevertheless, a background reaction of glutathione with the electrophile 4-nitro-benzylchloride could be observed at pH 7. This background reaction has a metal independent rate constant of 6.9 ± 0.4 x 10^{-5} $min^{-1}uM^{-1}$ and was always subtracted from the measured reaction rates. The

observation that this background reaction is highly pH dependent lead to the assumption that the background reaction occurs predominant between the thiolate form of glutathione and the electrophile (Figure 32).



pH	Background reaction [min^{-1}]
5	(8.3 ± 0.9) x 10^{-3}
6	(2.3 ± 0.5) x 10^{-2}
7	0.27 ± 0.02
8	1.39 ± 0.01

Figure 32. Non-enzymatic background reaction of glutathione at different pH values. The concentration of glutathione and 4-nitrobenzylchloride was 1 mM in every entry. The Britton-Robinson buffer system was used for this reaction.[137]

The Michaelis-Menten kinetics with the EF_3021 loaded with Zn^{2+} was measured at a pH of 7.0. While glutathione proved to be a good substrate we were not able to saturate the enzyme with the artificial substrate 4-nitrobenzylchloride. Therefore, we kept the electrophile concentration in all our kinetic measurements constant at a concentration of 500 µM. Under these conditions the enzyme EF_3021 showed a high glutathione affinity with a Michaelis-Menten parameter K_M of 1.6 ± 0.3 µM, a k_{cat} of 0.56 ± 0.01 min^{-1} and a catalytic efficiency of 5.8 ± 1.2 x10^3 $s^{-1}M^{-1}$.

We measured the kinetic parameter of our enzyme also at a lower pH of 5 with a reduced background reaction. Interestingly the enzyme remained 100% active in the lower pH range. The Michaelis-Menten kinetics with Zn^{2+} were measured at pH 5 and showed a similar k_{cat} of 0.59 and a 10 times higher K_M of 16.9 µM (Table 2). Due to the higher biological relevance of a system at pH close to the physiological pH of 7.4 further reaction were carried out at pH 7.

Table 2. Michaelis-Menten parameters for the enzyme reaction EF_3021 with a constant concentration of 4-nitrobenzylchloride (500 µM) at pH 5 and pH 7.

Metal cations	K_M [µM]	k_{cat} [min^{-1}]	k_{cat}/K_M [$s^{-1}M^{-1}$]
Zn^{2+} pH 7	1.6 ± 0.3	0.56 ± 0.01	5.8 ± 1.2 x 10^3
Zn^{2+} pH 5	16.8 ± 0.5	0.59 ± 0.01	0.59 ± 0.03 x 10^3

Furthermore, we probed the nucleophilicity of glutathione with different electrophiles. The electrophilicity of the benzylic position of benzyl chloride is directly influenced by the functional groups in the *para*-position of the benzyl ring. An electron-withdrawing group such as nitro, nitrile or trihalide increases the electrophilicity, while electron-donating groups like ethers, amines or phenol groups decrease the electrophilicity of the benzylic position.

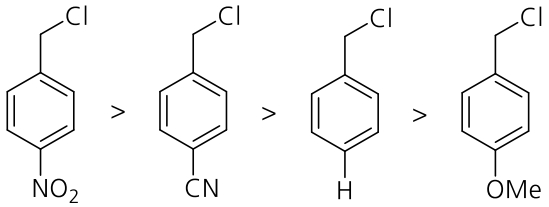


Figure 33. Benzylchloride analogues ordered from the most electrophilic 4-nitrobenzylchloride to the least electrophilic methoxybenzylchloride.

To test how nucleophilic the enzyme-activated glutathione for the alkylation reaction with artificial substrates is, we changed the electrophile from the strong electron-withdrawing group nitro to the less withdrawing nitrile group. We hence replaced the substrate 4-nitrobenzylchloride with 4-(chloromethyl)-benzonitrile. What we observed was an expected reduction of the reaction rate. The observed reaction rate decreased by a factor of 2.2 from 0.56 ± 0.01 min^{-1} to 0.251 ± 0.004 min^{-1} in the Zn^{2+}-loaded enzyme. Afterwards we changed the substrate to 4-methoxybenzylchloride, which contains an electron-donating group instead of an electron-withdrawing group. In this reaction setup we were not able to measure any product formation.

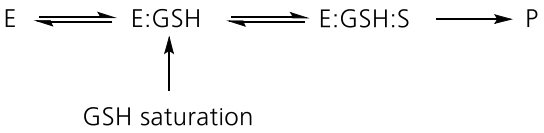
Table 3. Michaelis-Menten kinetic parameters for different substrate with the Zn^{2+}-loaded enzyme.

Metal cations	Substrate	K_M [µM]	k_{cat} [min^{-1}]	k_{cat}/K_M [$s^{-1}M^{-1}$]
Zn^{2+}	4-nitrobenzylchloride	1.6 ± 0.3	0.56 ± 0.01	5.8 ± 1.2 x 10^3
Zn^{2+}	4-(chloromethyl)benzonitrile	1.9 ± 0.3	0.251 ± 0.004	2.2 ± 0.4 x 10^3
Zn^{2+}	4-methoxybenzyl chloride	n.a.	n.a.	n.a.

The proposed reaction mechanism of EF_3021 with sp^3 hybridized benzyl chloride derivatives is an associative S_N2 reaction. In the active site of EF_3021 the Zn^{2+}-complex is formed by the two neutral histidine ligands, the negative ligand aspartic acid and the deprotonated, negatively charged ligand glutathione. Together with the Zn^{2+} ion the net charge of the complex is zero. Following the arguments and model of Picot *et al.* (2.1.7), we can propose an association mechanism for the glutathione-*S*-transferase. Also, the kinetic observation of the sustained reactivity at lower pH supports the proposed association mechanism.

2.3.2 Kinetic Characterization of EF_3021 with Different Metals

The enzyme reaction with the two substrates is divided into two steps. First the enzyme binds glutathione to form the enzyme substrate complex. Secondly this complex reacts with the electrophilic substrate to the product. Since our enzyme has a high affinity to glutathione it is easy to saturate and form the predominant E:GSH complex. With this predominant E:GSH complex and the electrophile we can describe the system as a bimolecular reaction.



Scheme 17: The enzymatic reaction of the two-step reaction scheme. In the first step the enzyme-GSH complex is formed, while in the second step the E:GSH complex reacts with the electrophilic second substrate to form the product. By increasing the GSH concentration the enzyme can be completely saturated and the reaction can then be described as a bimolecular reaction between E:GSH and the second substrate.

This bimolecular reaction can be compared to the model system used by Fox *et al.* in their thiol alkylation study with different metals (see 2.1.6).[130] The difference is that in our case instead of a simplified model system we have a glutathione-dependent metalloenzyme to examine the nucleophilicity of the ternary complex. Therefore, we wanted to measure the enzyme reaction with different loaded metals. The enzyme was first treated with ethylenediaminetetraacetic acid (EDTA) after purification to remove all metals from the enzyme and then different divalent metal ions were added to the reaction mixtures during the kinetic assay. With this method we first measured the rate of the enzyme with the metals Zn^{2+}, Ni^{2+}, Mn^{2+}, Fe^{2+} and Co^{2+} at a pH of 7 at high substrate concentrations (Figure 34).

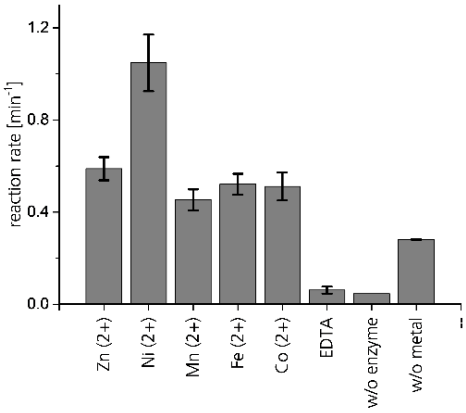


Figure 34. Reaction rates of the EF_3021 catalyzed reaction with glutathione (500 µM), 4-Nitrobenzylchloride (500 µM) and different metals.

In our measurements, we observed that the reaction rates of the enzymes with different metal ions were quite close to each other and in the same order of magnitude. These surprising results let us make a detailed enzyme kinetic characterization of EF_3021 with the divalent ions Zn^{2+}, Ni^{2+}, Mn^{2+}, Fe^{2+} and Co^{2+} (Table 4). We did not include copper in our measurements due to its strong redox activity. The measurements with Fe^{2+} were performed with and without the addition of ascorbate to ensure that all iron is in the ferrous state.

Table 4: Michaelis-Menten parameters at pH 7 with different metal ions.

Metal cations	K_M [µM]	k_{cat} [min^{-1}]	k_{cat}/K_M [$s^{-1}M^{-1}$]
Zn^{2+}	1.6 ± 0.3	0.56 ± 0.01	5.8 ± 1.2 x 10^3
Co^{2+}	26 ± 5	0.59 ± 0.03	0.3 ± 0.1 x 10^3
Mn^{2+}	4.1 ± 1	0.25 ± 0.01	(1.0 ± 0.3) x 10^3
Ni^{2+}	(6.5 ± 1.8) x 10^3	22 ± 6	(0.05 ± 0.03) x 10^3
Fe^{2+}	(1.1 ± 0.4) x 10^3	1.3 ± 0.1	(0.02 ± 0.01) x 10^3
Fe^{2+} + Ascorbate	(1.5 ± 0.5) x 10^3	1.0 ± 0.1	(0.011 ± 0.005) x 10^3

The Michaelis-Menten constant K_M can be interpreted as the affinity of the enzyme towards the substrate. In our measurements we observe significant differences between the glutathione binding to the enzyme containing different metals. While the K_M value was low in the range of 1-26 µM for the metals Zn^{2+}, Co^{2+}, and Mn^{2+} it was much higher for Fe^{2+} (1100 µM) and Ni^{2+} (6500 µM). In our case we were more interested in examining the nucleophilicity of the ternary complex and therefore the differences of the single K_M values are less important.

In the thiol alkylation model system of Fox *et al.* they observed an increase of the reactivity by 10-fold in Zn^{2+} and Ni^{2+} containing complexes.[130] In our results we observed significantly different results. The maximal rate was observed in the reaction with Ni^{2+} loaded enzyme (22 min^{-1}). This reaction rate was 38-fold faster than the Zn^{2+} loaded enzyme (0.56 min^{-1}). Interestingly the kinetic measurements with Co^{2+} showed very similar reaction rates of 0.59 min^{-1}. With manganese the reaction rate dropped by two-fold compared to the Zn^{2+} enzyme (0.25 min^{-1}) and the Fe^{2+} loaded enzyme was twice as fast as Zn^{2+} 1.0 min^{-1}. In our measured results we observed that only the Ni^{2+} loaded enzyme shows a significant difference to the other metals while the rates of Mn^{2+}, Co^{2+}, Zn^{2+} and Fe^{2+} were all similar to each other.

The reaction rate in the enzyme reaction with Ni^{2+} is much higher than with other metals. Especially the observation with Zn^{2+} is in discrepancy to the model systems of Fox *et al.* and Gennari *et al.*[130,138] They showed in their system a similar nucleophilicity and reactivity of nickel and zinc complexes in computer models as well as with kinetic data.[130,138] This fast catalytic activity could point toward a nickel-dependent enzyme. Nevertheless, the GST EF_3021 is presumably not nickel-dependent because there are strong arguments against this theory. The K_M value of glutathione with Ni^{2+} ions of 6 mM is very high and even in cells containing glutathione in concentrations in the millimolar range this K_M seems to be very inefficient. In addition, the catalytic efficiency which is defined as k_{cat}/K_M, is 50 ± 30 [$s^{-1}M^{-1}$], is 100 times lower than in the Zn^{2+} measurements and does not point to a natural nickel-dependent enzyme. In addition, Ni^{2+} cations are potentially toxic and cause multiple cell damages, they are present in bacterial cells only in the nanomolar range and the uptake and distribution is complicated and well regulated.[139] Most Ni^{2+} enzymes are combined with accessory proteins, which are involved in the delivery of Ni to the apoenzyme, for example in ureases.[139] In our enzyme EF_3021 no Ni^{2+} carrier proteins are found in the genome.

The measured kinetic data shows that the scaffold of the protein can accept many different metal ions for catalytic activity. The main question is how the enzyme selects one particular metal? Is it a random occupation or might there be other factors responsible for the metal selectivity, such as ligand-induced selectivity?

Some effects which might influence the metal binding in the cell are the absolute metal concentration in the organism, the stability of the complex as well as the architecture of the active site. The free metal ion concentration is highly regulated inside the cytosol. The concentration of Mn^{2+} and Fe^{2+} are in the µM range, while the Co^{2+} and Ni^{2+} concentration are in the nM range and the Zn^{2+} ions are even lower in the picomolar range.[140] In addition to the concentration, the stability of the complexes formed by divalent metals has a high impact. This stability of the ligand-metal complexes can be described by the Irving-Williams series, with the following order $Mg^{2+} < Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+}$.[141,142] While the concentration of Zn^{2+} is lower than the concentration of other metal ions it still can form very stable complexes and therefore replace other incorporated metal ions. An additional factor is the geometry and coordination mode of the enzymatic metal binding site, which can influence the metal selectivity, and which is discussed below.

Another possibility could be that EF_3021 can us any divalent transition metal from the first row. The concentration of glutathione in cells can be in the millimolar range, therefore EF_3021 loaded with a random divalent transition metal from the first row of the periodic table could be a sufficient catalyst. Especially if the glutathione *S*-transferase is involved in detoxification processes the ability to accept different central metals could be an advantage.

2.3.3 Crystal Structure

The enzyme loaded with different metals showed very similar *S*-transferase activity. We wondered if the metal changes the structure of the enzymes. To examine the structure of the nucleophilic complex we decided to crystallize the protein with glutathione and different metals.

The crystal structure of EF_3021 (PDB: 3cex) was already described in the literature. Nevertheless the reproducibility to obtain crystals proved to be a challenging task. At first and in collaboration with Marcel Meury, we were not able to obtain any crystals in multiple commercial and optimized 96-well screens, even though the protein was produced in good yield and in high purity. However, we found a solution to obtain crystals by replacing the reservoir solution in the 3 drop 96 well plate with a 1.5 M NaCl solution. This method was described in the crystal structure of a structurally related but unknown enzyme from *Geobacillus stearothermophilus*.[118] With the newly found crystallization conditions we could co-crystallize the enzyme with the substrate glutathione and with different metals.

Marcel Meury solved the structures of EF_3021 with the central metal ion Zn^{2+} and glutathione (1.5 Å resolution), Co^{2+} and Glutathione (1.6 Å resolution), Mn^{2+} without glutathione (1.6 Å resolution) and Ni without glutathione (1.6 Å resolution). All solved crystal structures have an asymmetric unit which contains two molecules arranged as dimers (Figure 35). The crystal structures with zinc and cobalt belong to the space group P121, while the crystal structure with manganese and nickel belong to the space group P1211 the cell constants are shown in Table 5.

Table 5. Unit cells parameters, space groups and resolution of the solved crystal structures and structure from literature.

Structure	Resolution	a	b	c	α	B	γ	Space group
Zn^{2+} + GSH	1.5 Å	56.133 Å	51.845 Å	62.507 Å	90°	112.558°	90°	P121
Co^{2+} + GSH	1.6 Å	56.312 Å	52.839 Å	62.412 Å	90°	112.664°	90°	P121
Mn^{2+}	1.6 Å	55.334 Å	52.635 Å	62.015 Å	90°	111.76o°	90°	P1211
Ni^{2+}	1.7 Å	55.629 Å	52.552 Å	62.108 Å	90°	111.95°	90°	P1211
3cex	2.0 Å	55.101 Å	55.137 Å	62.429 Å	90°	108.20°	90°	P121

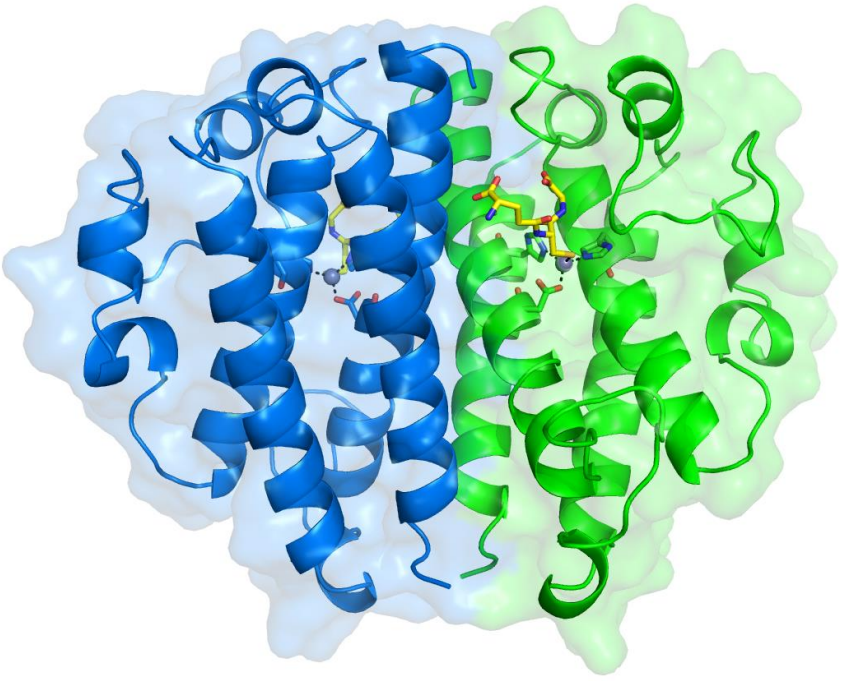


Figure 35: Solved crystal structure of the dimeric enzyme EF-3021 in a 1.5 Å resolution with Zn^{2+}, the ligands glutathione and two glycerol molecules.

The asymmetric unit of the crystal structures contained two molecules, forming a dimer. The obtained crystal structures are very similar to each other. The root-mean-square deviation (RMSD) was calculated with PyMOL for the crystal structures. The RMSD is a value which describes the average distance between two superimposed proteins. We aligned all observed structure to the Zn^{2+} structure containing the highest resolution of 1.5 Å.

Table 6. The RMSD values calculate by PyMOL between the Zn^{2+} structure and the crystal structures. The atoms which were used for the calculation are shown in the second column.

Crystal Structure	RMSD [Å]	Atoms
Co^{2+}	0.141	151 to 151
Ni^{2+}	0.165	146 to 146
Mn^{2+}	0.161	138 to138
3CEX	0.230	150 to 150

The RMSD values were below 0.2 Å therefore the detailed structural discussion is based on the structure of the enzyme with Zn^{2+} and glutathione. In this structure, the complete peptide chain from residue 1 to 169 is modelled. The monomer contains the four α-helix bundles of the DinB2 superfamily. Helix one (4-24) is linked to helix two by a 16 amino acid linker. This linker contains a small α-helix structure containing five residues. Helix two (41-59) is linked by a long linker which contains 40 amino acids to helix three (101-121). In this longer linker, three small α-helices and a random coiled section (76-89) are present. Between the third and the fourth helix (141-167) the linker containing 19 amino acid residues contain a short α-helix and another random coil (128-140).

2.3.4 Crystallographic Interface

Each monomer consists of a four-α-helix bundle DinB structure. Residues located on the first and fourth α-helix of each monomer are involved in the interface interaction of the dimer. In the crystal structure with Zn^{2+} (1.5 Å) we identified 38 residues, which are involved in the interface formation. Hydrogen bonds are formed by 17 residues and salt bridges are formed by 12 residues. The residues which forms hydrogen bonds between each monomer were identified as follows: Gln162 with Glu29, Arg184 with Leu57, Met 21 with Glu144, Arg33 with Val163, Arg37 with Asp170, Gln178 with Gly177 and Lys59 with Tyr 181. The residue forming salt bridges were identified as Met21 to Glu144, Arg37 to Asp170 (Figure 36).

As described above the quaternary structure of EF_3021 in solution was identified to be dimeric. Furthermore, the active site is located in the interface of the two monomers (see below). We therefore propose that the observed interactions are strong enough to provide a stable cohesion between the single monomers.

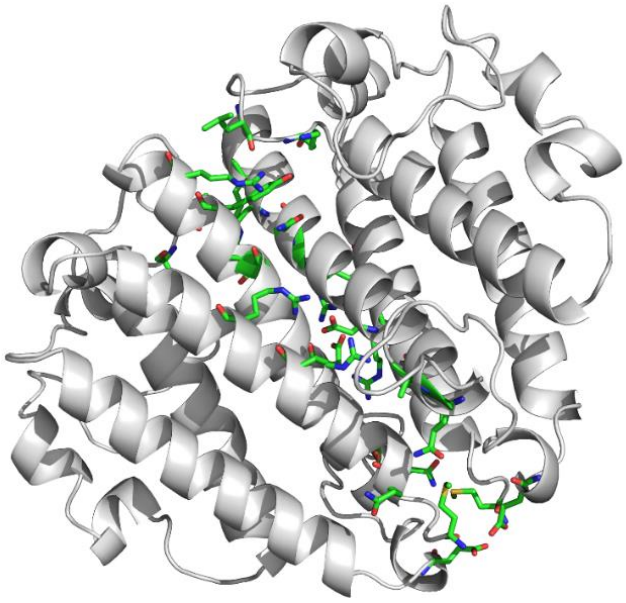


Figure 36. Overview of the residues which are involved with hydrogen bonds and salt bridges are highlighted in the interface between the two monomers of EF_3021.

Another method to analyze the interface interactions is to measure the complex formation significance score (CSS) by the web application PDBePisa (http://www.ebi.ac.uk/pdbe/pisa/).[143–145] The CSS is calculated based on the crystal structure and displays the relevance to form a complex. The CSS score can range between 0 and 1, where 0 is no and 1 is an essential relevance for the complex formation. In the Zn^{2+} structure with the glutathione ligand the CSS was calculated to be 0.206. This is a low score, which indicates that the interface plays an auxiliary role in the complex formation. We used the same method to analyze the CSS value of the crystal structure from the literature (PDB: 3cex, 2 Å). The score was calculated to be 0.944, which indicates an essential role for the complex formation. This rather big difference between the EF_3021 Zn^{2+} structure and the structure from the literature is surprising. One explanation could be the substrate binding in our Zn-structure. The glutathione is found in the interface of the two monomers and might reduce the number of hydrogen bonds and salt bridge formation. This could lower the CSS score calculated from the PDBePisa analysis.

2.3.5 The Active Site of EF_3021

The active site of the enzyme EF_3021 is located in a small notch close to the surface. The active site consists of residues from both monomers, which means the active site is only formed in the dimer of the protein. The protein is symmetrical and therefore contains two active sites per dimer. The metal binding motive is located between the second and fourth helix and contains two histidine and one aspartate ligand. This 2-His-1-carboxylate facial triad is a common metal ligand site for different metals like Zn^{2+} or Fe^{2+}.[146,147] The

first histidine is located on the second helix at position His47. The aspartic acid and the second histidine are located on the fourth helix at positions Asp151 and His155. The metal is bound to both histidine residues by the N^{τ} of the imidazole ring at a distance of 2.1 Å (His47) and 2.2 Å (His155). The carboxylic acid is bound to the metal ions at a distance of 2.1 Å (Asp151). The thiol group of glutathione completes the tetrahedral metal binding motive with a distance of 2.2 Å.

The metal ion is coordinated by four ligands in a distorted tetrahedral geometry (Figure 37). A symmetrical tetrahedron contains four ligands and the tetrahedral angle of 109.5°, which is defined as the angle between the corners looking from the middle of the tetrahedron. In the crystal structure with the bound ligand glutathione and the central metal Zn^{2+} we observe a distortion of this optimal angle. The longest angle is 124.1° a difference of 14.6° and the shortest angle is 100.5° a difference of 9°. A reason for this distortion could be that the enzyme activity is increased if the thiolate of glutathione is slightly shifted towards the second substrate binding pocket as seen in the crystal structure. A summary of the distances and angles of all solved crystal structures is displayed in Table 7.

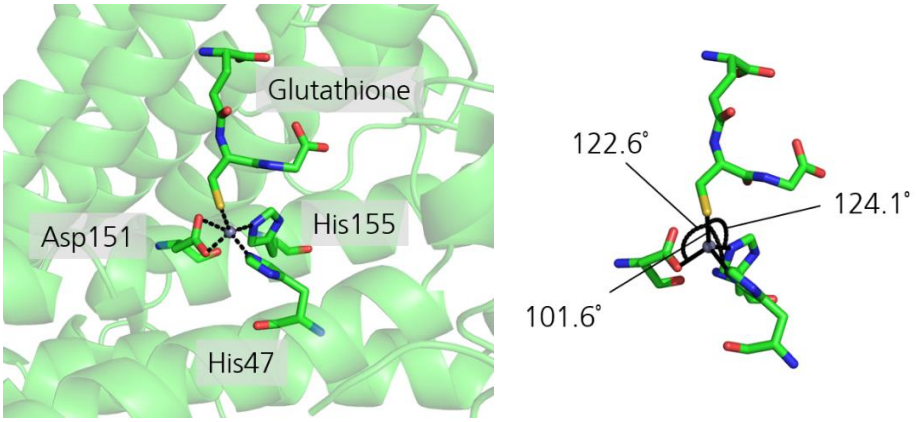


Figure 37. **Left:** Tetrahedral Zn^{2+} binding site, located between the second and fourth helix. Ligands His47, Asp151, His155 and glutathione forming a distorted tetrahedron. **Right:** Three of the six tetrahedral angles are shown, and the discrepancy to the optimal tetrahedral angel of109.5°.

Table 7. Distances and angles between the metal and the ligands in all solved structures. In the structure of Mn and Ni instead of the bound glutathione two water molecules were observed, which are coordinating to the central metal. The resolution of the structures are 1.5 Å for Zn and 1.6 Å for the Co, Mn and Ni structure.

	M = Zn	M = Co	M = Mn	M = Ni
Resolution [Å]	1.5	1.6	1.6	1.6
Distances [Å]				
M-His47	2.1	2.1	2.1	2.1
M-His155	2.2	2.0	2.2	2.0
M-Asp151	2.1	2.0	2.2	2.1
M-S	2.2	2.2	0	
H_2O236	0	0	2.4	
H_2O238	0	0	2.2	
H_2O64				2.2
H_2O260				2.2
Angels [d				
His47-M-His155	101.7	99.1	96.1	96.0
His47-M-Asp151	102.1	104.1	94.1	103.0
His47-M-S	100.5	95.7		
His155-M-Asp151	101.6	99.1	90.7	95.7
His155-M-S	124.1	128.3		
Asp151-M-S	122.6	124.5		
His47-M-H_2O236			89.4	
His47-M-H_2O238			115.0	
His155-M-H_2O236			169.5	
His155-M-H_2O238			91.4	
Asp151-M-H_2O236			97.8	
Asp151-M-H_2O238			150.4	
H_2O236-M-H_2O238			78.2	
His47-M-H_2O64				90.9
His47-M-H_2O260				101.2
His155-M-H_2O64				172.7
His155-M-H_2O260				92.1
Asp151-M-H_2O64				85.1
Asp151-M-H_2O260				153.6
H_2O64-M-H_2O260				84.1

In the second crystal structure with the ligand glutathione and the central metal cobalt, we obtained a very similar structural symmetry. The distances between the metal ligands and the metal are similar to the Zn^{2+} structure and are between 2.0-2.2 Å. Again, a distorted tetrahedron is formed with the two histidines, the aspartic acid and glutathione. The direct overlay between the $Zn^{2+}$and the Co^{2+} structure shows only marginal differences. If we compare the angle between these structures, we can see that in the case of the Co^{2+} structure the distortion is slightly bigger than in the Zn^{2+} structure. The biggest tetrahedral angle of 128.3˚ deviates by 18.8˚ from the ideal symmetric angle 109.5˚and the smallest 95.7˚ by 13.8˚. This slightly higher distortion in the cobalt complex does not directly have negative consequences on the reaction rate of our model reaction with the substrate 4-nitrobenzylchloride.

In conclusion we are not able to observe major differences between the nucleophilic complexes of Zn^{2+} and Ni^{2+}, they are exceedingly similar to each other.

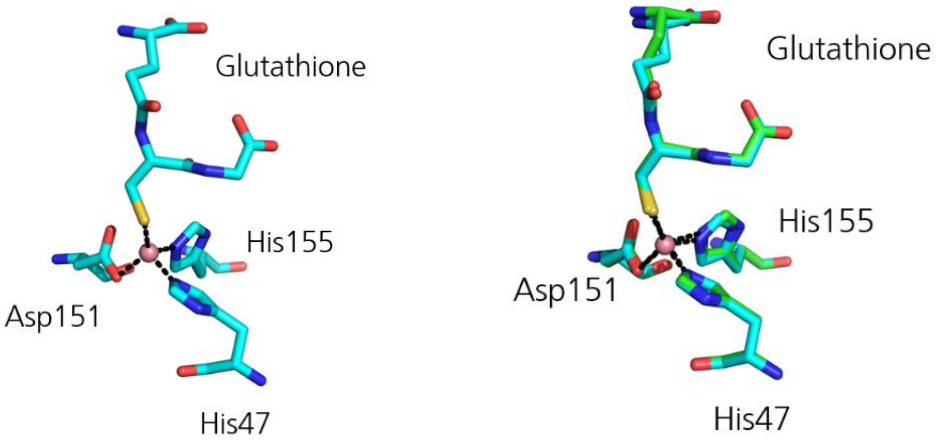


Figure 38. **Left:** tetrahedral binding site of the cobalt complex with the ligands His47, Asp151, His155 and glutathione forming a distorted tetrahedron. **Right:** Alignment of the cobalt and the Zn^{2+} metal binding site.

In the crystal structure containing Mn^{2+} and Ni^{2+}, the substrate glutathione is missing. In the Ni^{2+} structure this is not surprising since the binding affinity of glutathione with the Ni^{2+} structure is low (K_M value of 6 mM). In the Mn^{2+} structure this argument is not valid (K_M = 4 µM), and it is not clear why the glutathione is missing in the crystal structure. In the two structures, instead of the glutathione coordinating the metal, two coordinating water molecules were observed. In the Mn^{2+} structure the water molecules H_2O_{236} and H_2O_{238}, and in the Ni^{2+} structure the two water molecules H_2O_{64} and H_2O_{260} were bound to the metal at distances between 2.2-2.4 Å. This additional ligand leads to the formation of a distorted trigonal bipyramidal geometry. The two ligands which correspond to the axial ligands in the bipyramidal geometry

are His_{155} and H_2O_{236} in the manganese structure and His_{155} and H_2O_{64} in the nickel structure. The observed angles (Mn: 169.5˚, Ni: 172.7˚) correspond quite well to the theoretical 180˚C angles of the axial ligands. The other ligands in equatorial position are at an angle close to 90˚ towards the axial position. Nevertheless, the distortion between the three equatorial ligands with a theoretical angle of 120˚ is high. The two structures (Mn^{2+} and Ni^{2+}) are very similar to each other and show similar angles between each ligand. The conformational change of the metal ligands between the glutathione bound and glutathione unbound structure is minimal. This can be seen in the superposition of the Mn^{2+} and Ni^{2+} structure with the Zn^{2+} structure (Figure 39).

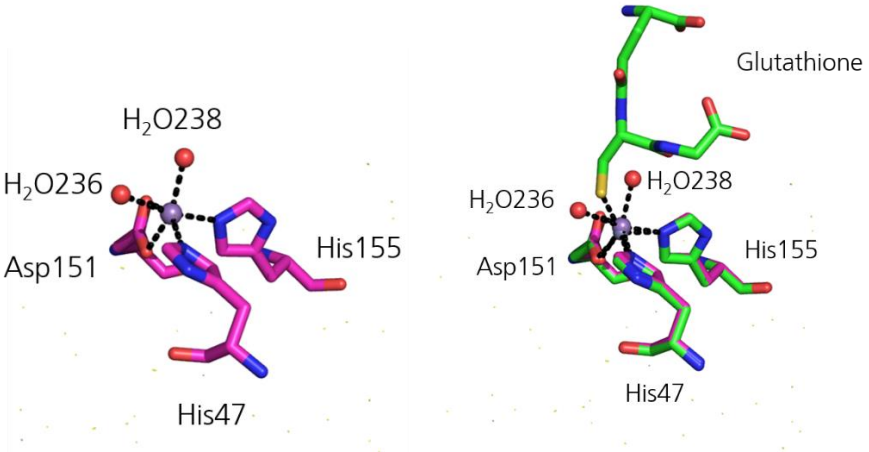


Figure 39. **Left:** distorted metal binding site in the Mn^{2+} structure with the two water molecules H_2O_{236} and H_2O_{238}. **Right:** Alignment of the Mn^{2+} and the Zn^{2+}metal binding site.

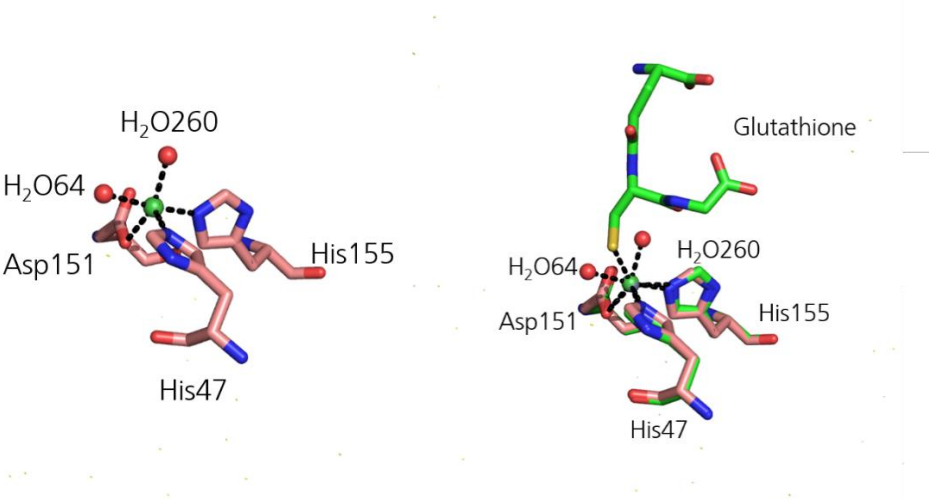


Figure 40. **Left:** distorted metal binding site in the Ni^{2+} structure with the two water molecules H_2O_{64} and H_2O_{260}. **Right:** Alignment of the Ni^{2+} and the Zn^{2+} metal binding site.

2.3.6 Metal-Binding of EF_3021

The obtained crystal structure with Zn^{2+} and bound glutathione shows a tetrahedral binding symmetry. Different metal ions usually prefer specific coordination geometries. It has been shown that Co^{2+}, Mn^{2+}, and Fe^{2+} prefer an octahedral coordination with six ligands while Zn^{2+} prefers a tetrahedral binding motif with 4 ligands. For Ni^{2+} metal ions, octahedral and tetrahedral coordination are equally frequent.[140] The properties of a metal-binding site are influenced not only by the ligands which are directly interacting with the metal ions but also by the ligands which are in the second shell. These "indirect ligands" form a hydrogen network with the metal ligands. This hydrogen bonding network has direct influence on the metal center by changing the metal affinity, the orientation of the ligands for optimal metal binding and enhancing the electrostatic interaction between ligand and metal ion.[148,149]

One example of an extensively studied hydrogen network is in the carbonic anhydrase II (CAII).[148] CAII is a Zn^{2+}-dependent enzyme which contains three histidine metal ligands and catalyzes the reversible hydration of CO_2 to H_2CO_3. The three histidines are coordinated by hydrogen bonds to a glutamic acid, a glutamine and a main chain carbonyl of an asparagine. By making single point mutations of those residues Kiefer *et al.* measured the dissociation constant K_D of the Zn^{2+} and the energy of the single hydrogen bonds.[148] The removal of a hydrogen bond from an indirect ligand reduces the Zn^{2+} affinity by 0.9 kcal/mol. This effect is additive and it was estimated, that the indirect ligands together increase the affinity of Zn^{2+} by 5 kcal/mol. In the mutants of these indirect ligands the Zn^{2+} affinity decreased by up to 40 fold and the activity of the enzyme up to 60 fold compared to the wild type.[148] These observations underline that in the description of metalloenzymes indirect ligands should be included.

In the Zn^{2+} structure, we also observe second shell ligands. The $N^{\delta 1}$ of His47 (the second nitrogen of the imidazole ring which is not bound to the ligand) forms a hydrogen bond to the amino acid Asp85 (2.6 Å). Furthermore, the $N^{\delta 1}$ of the second His155 is coordinated to the carbonyl group of Gln158 (2.7 Å) (Figure 41).

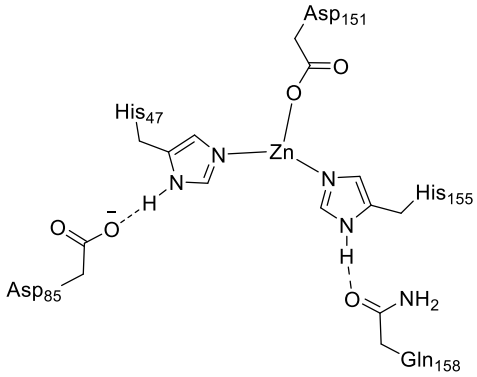


Figure 41: Second-shell ligands Asp85 and Gln158 which are coordinating the two histidine of the metal binding site.

These second shell ligands point towards a less flexible metal binding site with defined function. We therefore speculate that despite the measured metal versatility the enzyme is Zn^{2+}-dependent in vivo. The low K_M value for glutathione, the predominant ionic character of the Zn^{2+}-metal binding and the tetrahedral metal geometry support this assumption. Furthermore the product release of a formed thioether in a glutathione *S*-transferase mechanism is not limiting with a Zn^{2+} metal ion since the bond strength of Zn^{2+}-thioether is much lower than the Zn^{2+}-thiolate.[138] This is a difference to Ni^{2+}-complexes which are able to bind thioether and can decrease the rate of product release.[130,138] To test the binding of the thioether product to the nickel of the active center we measured the reaction rate in presence of the previously formed glutathione nitro benzyl adduct. In the reaction with the nickel-loaded enzyme we observed a decrease of the activity by three-fold after addition of 200 µM product. In contrast the Zn^{2+} reaction showed no change in reactivity (Figure 42). This rather small product inhibition effect with nickel shows that the product can weakly bind to the enzyme, which is in agreement with the literature.

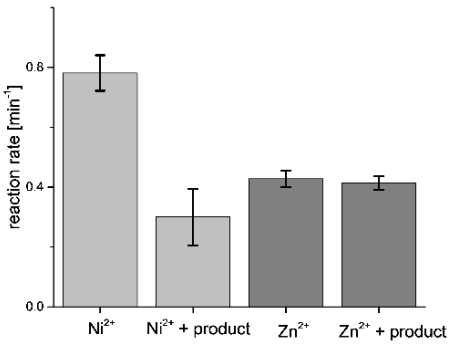
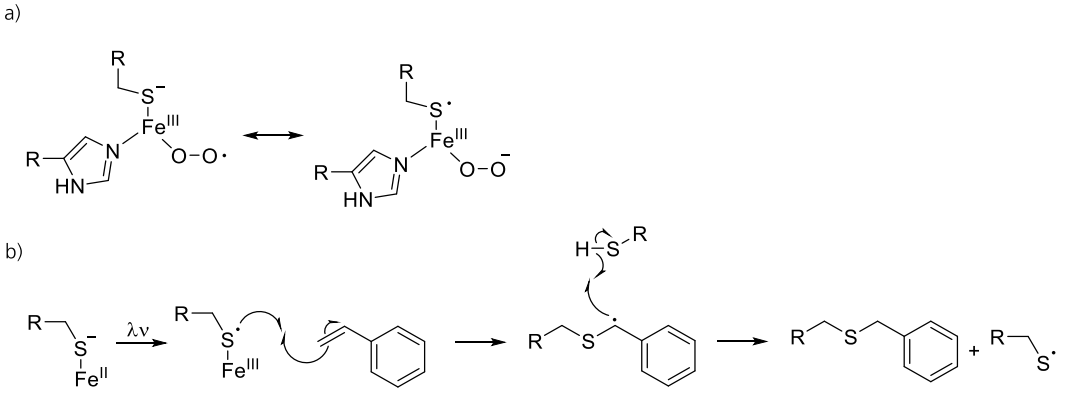


Figure 42. Product inhibition reactions. Light gray Ni^{2+}, without and with 200 µM product. Dark gray Zn^{2+} reaction without and with 200 µM product.

2.3.7 Possible Radical Mechanism

Our enzyme EF_3021 show a high similarity in the structural fold of the metal binding to EgtB. In the iron-dependent sulfoxide synthase EgtB an elementary oxygen binds to the iron and forms a superoxide species. This superoxide species is stabilized by the iron and reacts further in the EgtB radical mechanism.[60] We were able to load our enzyme EF_3021 with Fe^{2+}. Therefore, we were wondering if we could observe a radical reaction with glutathione and find a mechanistic relation between the two DinB enzymes. A mechanistic connection could reveal a possible evolutionary path, showing how the more complex EgtB enzyme containing two different domains, could have evolved out of the simpler glutathione *S*-transferase.

If the Fe^{2+} metal ion would be able to stabilize a glutathione thiyl radical in EF_3021, this thiyl-radical could then further react in a radical addition to a radical acceptor like styrene. The radical addition product could abstract a hydrogen atom of another glutathione and sustain the chain reaction (Scheme 18). To generate a glutathione thiyl radical we tried to irradiate the glutathione-enzyme reaction with UV light (366 nm). The identification of the addition product of styrene with glutathione in the enzyme reaction was not successful, we were only able to observe the polymerization of styrene.



Scheme 18. **a)** The supeoxo radical formed in the first step of the sulfoxide synthase mechanism of EgtB. **b)** a possible radical addition between glutathione and styrene.

2.3.8 Second Native Substrate

All kinetic experiments have been measured with artificial electrophiles which have little biological relevance, because we do not know what the second native substrate of EF_3021 is. There are leads which can point to some possible native substrates. In the binding site next to the glutathione we observed two bound glycerol molecules (Figure 43). The binding is rather unspecific over a few hydrogen bonds. The binding of two glycerol molecules could indicate that the second substrate is a polar molecule for example a sugar analogue.

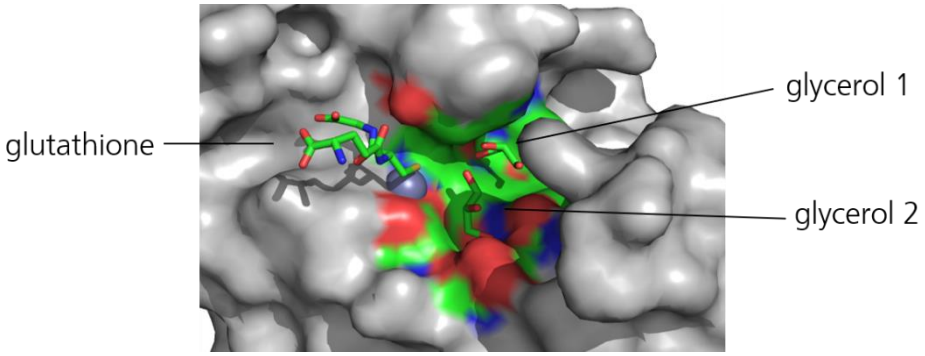


Figure 43. The two glycerol molecules are located in the active site close to the glutathione.

Furthermore, it is common in bacteria that enzymes which are used in biosynthesis pathways are located in gene clusters.[150] Therefore, it is worthwhile to analyze gens which are located close to our GST EF_3021 and if see whether these genes are conserved in other organisms containing a GST homologue. In the genome map of *Enterococcus faecalis* we identified next to the GST gene a polysaccharide lyase (EF_3023). A possible function of EF_3021 might be the isomerization of a sugar degradation product. Nevertheless, the polysaccharide lyase is not in a highly conserved gen cluster in other organism which produce EF_3021 homologues.

Figure 44. Genome map of *Enterococcus faecalis* with the glutathione *S*-transferase EF_3021, and a close encoded polysaccharide lyase (EF_3023).

2.3.9 Modification of the Active Site

The active site revealed a tryptophan W86 which is located exactly above the metal binding site and might be involved in the binding of the second substrate (Figure 45). This amino acid residue could block the direct access of the artificial electrophiles to the metal bound thiolate and thereby forcing the glutathione to dissociate first before reacting. This would be an explanation why we do not observe metal-depending kinetic differences.

To improve the accessibility of the artificial substrate to the metal-bound glutathione we decided to replace the bulky W86 residue with alanine or glycine. We chose these residues because they are often found in α-helices and should therefore have a minimal impact on the stability of the protein. The mutants EF_3021$_{W86A}$ and EF_3021$_{W86G}$ were produced and showed no structural differences to the wild type.

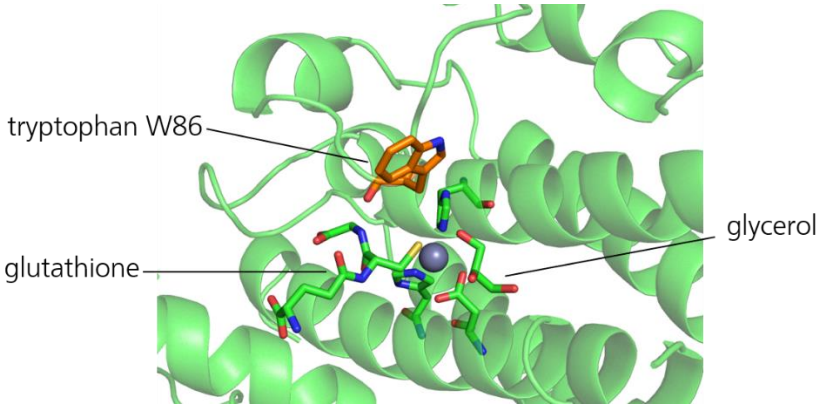


Figure 45: The tryptophan 86 (higlighted in orange) which is located directly above the active site with the metal Zn^{2+}.

In the next step we repeated the kinetic experiment with both mutants at pH 7 and pH 5. The results showed a large decrease of activity displayed in Table 8**Fehler! Verweisquelle konnte nicht gefunden werden.**. While the mutant EF_3021$_{W86A}$ showed a 25-fold reduced activity the activity of EF_3021$_{W86G}$ was below our detection limit. These results indicate that the formed mutation cannot stabilize the artificial substrate or is not able to activate glutathione anymore.

Table 8. Reaction rate of the wildtype enzyme and the mutants, with 500 µM 4-Nitrobenzylchloride, 1000 µM glutathione,1 µM enzyme and 10 µM $ZnCl_2$.

Enzyme	Reaction rate pH 5 [min^{-1}]	Reaction rate pH 7 [min^{-1}]
Ef_3021_{wt}	0.536	0.4107
Ef_3021_{W86A}	0.034	0.0158
Ef_3021_{W86G}	0.001	0.0007

To exclude the possibility that glutathione is not bond anymore, the UV-vis experiment with Co^{2+} was repeated. It could be shown that the cobalt-to-sulfur charge transfer still occurs in the mutant W86A which indicates that the glutathione is still able to bind to the enzyme (Figure 46). The extinction coefficient is 95 $M^{-1}cm^{-1}$ which is close to the extinction coefficient of the wildtype enzyme of ε_{512} = 166 $M^{-1}cm^{-1}$.

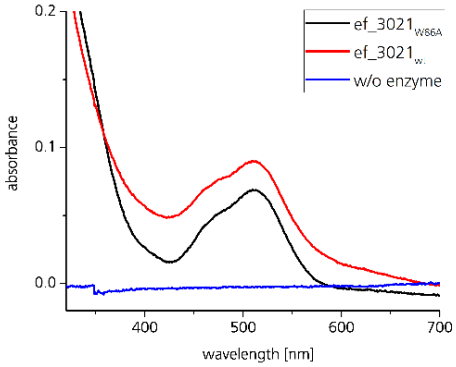


Figure 46. UV-Vis chromatogram with the Co loaded enzyme and different glutathione concentrations. The sulfur to cobalt charge transfer is visible at a wavelength of 512 nm.

2.3.10 The Distribution of Glutathione *S*-Transferase in Bacteria

The native substrate glutathione interacts with a number of residues within the active site. Two arginines Arg163 (2.9 Å) and Arg164 (2.9 Å), coordinate the two-carboxylic groups of glutathione. These residues are located on the fourth helix of the monomer (Figure 47). In addition, a coordination between Lys39 and the carboxylic acid of glycine from glutathione is visible. Another hydrogen bond is formed between the amide carbonyl of Trp86 and the amide bond of glycine from glutathione. Water molecules are also involved in the binding network. A water-mediated connection between the amine group of Gln158 from monomer B and the amide carbonyl of the glycine from glutathione (5.7 Å), as well as a water-mediated connection between the amide carbonyl of His88 and the amide bond of glycine from glutathione is formed. It is important to note that the residues Gln24, Arg163, Arg164 are not located on the same protein chain as the metal binding site or the other residues involved in the glutathione binding.

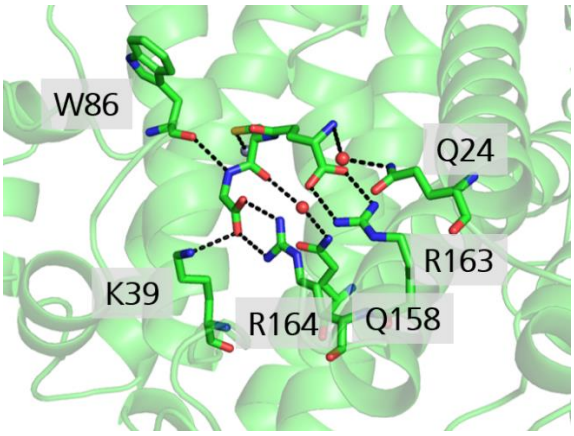


Figure 47. Glutathione binding pocket of EF_3021 containing the residues which directly coordinate the substrate (K39, W86, R163 and R164). In addition, the two amino acids which form indirect water network Q24 and Q158 are displayed.

In the structure of EF_3021 the two residues Arg163 and Arg164 were identified to be forming salt bridges to two carboxylate groups of glutathione. By blasting the sequence of EF_3021 with UniProt.org, we could identify organisms which contain a related enzyme with conserved Arg163 and Arg164 (Figure 48). The organisms were found in the gram-positive bacterial phylum of Firmicutes and Actinobacteria and in the gram-negative phylum of Proteobacteria. The presence of EF_3021 GST homologues indicates that those organisms presumably produce glutathione in their cells. Gram-positive bacteria are known to use mostly other small molecular weight thiol, such as mycothiol, in their cells.[99] The identification of GSH-dependent Actinobacteria is rather unusual, since they are known to predominantly produce mycothiol.[151]

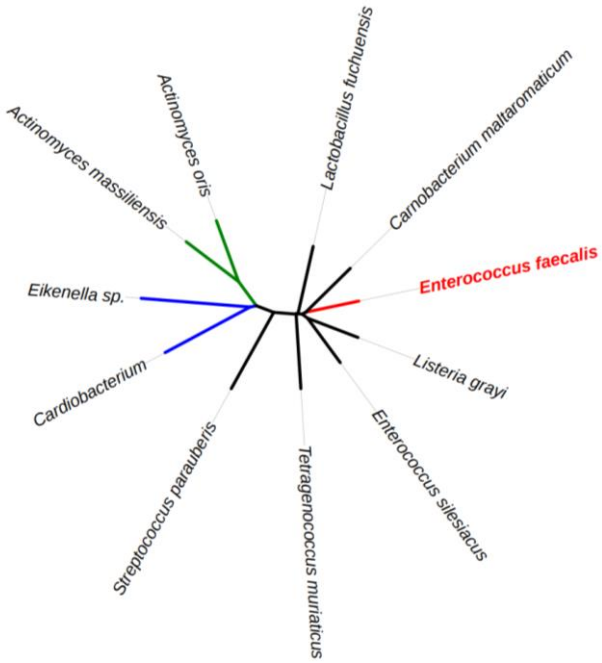
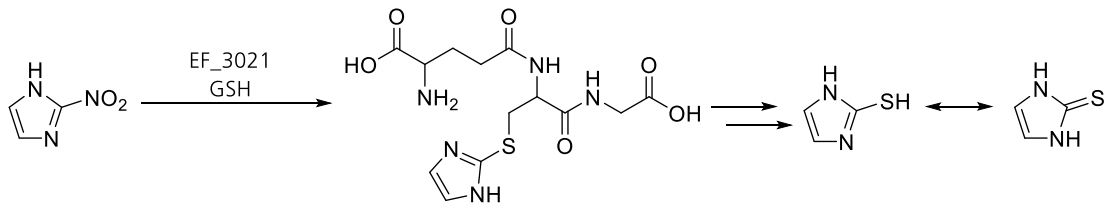


Figure 48. Phylogenetic tree showing selected GST homologues of EF_3021 (red). The firmicutes are shown in black, the actinobacteria in green and the proteobacteria in blue.

2.3.11 Hypothesis of an Alternative Ergothioneine Biosynthesis

The sulfoxide synthase EgtB catalyses the main step in ergothioneine biosynthesis, forming the C-S bond between histidine and a cysteine derivative.[71] The structural homology of the GST EF_3021 allows us to speculate, that there could be an alternative route to produce ergothioneine. The C-S bond could also be formed by a nucleophilic attack of a glutathione molecule to an activated imidazole ring of a histidine derivative. After successful addition, the glutathione moiety could be degraded by a γ-glutamyltransferase and an EgtE/OvoB homologue could perform the transformation to give the final thiol compound, ergothioneine.

It has been reported, that UV-light induced damage in nucleobases leads to nitro group formation and that 8-nitroguanine can be formed.[152] Nitro groups are a potential leaving group. Therefore, we tested the activated imidazole *4*- and *2*- nitroimidazole for a possible reaction with glutathione, catalyzed by EF_3021 (Scheme 19). Our hypothesis could not be confirmed, and we were not able to detect any formation of the glutathione-imidazole addition product. Nevertheless, the questions about the evolutionary relationship between EgtB and EF_3021 and the evolution of the DinB domain remain.



Scheme 19. A hypothetical and alternative pathway to form a C-S bond between an imidazole ring and the glutathione. The addition product can then further be transferred by a γ-glutamyltransferase and an EgtE/OvoB homologe to the 1,3-dihydro-imidazole-2-thione.

2.4 Conclusion

We were able to characterize the metal-dependent glutathione *S*-transferase EF_3021 with different artificial substrates and different binding metals. The observation that the EF_3021 enzyme scaffold accepts different divalent transition metal ions with similar catalytic efficiency was counter our expectations and different from previously reported systems in the literature.[130] Furthermore, we obtained crystal structures containing the substrate glutathione and different metals showing a distorted tetrahedral active site with bound glutathione and a distorted trigonal bipyramidal geometry in the glutathione-free structures. To be able to compare our kinetic data directly to results from similar systems reported in the literature density functional theory calculations are needed from our obtained crystal structures.

Bioinformatics show that this glutathione *S*-transferase is found in different bacteria phyla. The system of EF_3021, which is formed by the dimerization of two DinB domains, is much simpler system than in fusion proteins which contain a second unrelated domain. The fusion of the DinB structure with other domains such as the FGE-like domain in sulfoxide synthase EgtB, or the MDMPI domain in mycothiol maleylpyruvate isomerase enzymes, show that the simple DinB structure evolved over time to catalyze more complex biochemical reactions. We attempted to find evidence for a direct connection between the simple GST and a possible ancestral ergothioneine pathway, however with the use of 2-nitro and 4-nitro imidazole no link was found. Nevertheless, this approach should be repeated with 4-nitrohistidine and other more electrophilic molecules.

2.5 Experimental

Plasmid Construction:

The gene of the protein *Ef*GST (WP_002375627) was ordered codon optimized for production in *Escherichia coli* from GenScript® and delivered in a pUC57 plasmid. The gene was ligated into the expression vector pET19 using the restriction enzymes NdeI/XhoI. The sequence of the plasmid was verified by sequence analysis of Microsynth AG.

Sequence of pET19 EF_3021 wt

MGHHHHHHAENLYFQGHMKVTQLSSETLDRAHERFEETLAQMTVAEANTMPAPLIKSVTWLMWHTARELDLQISAL
NHSDPLWLSQHWTEKFALDLPDETEDWHHTPEEAAKVVVAEKQLLSDYLAASVALTKSYLDQIKEEQLSDVIDKNWT
PPVTRQVRLVSAIDDAVMHSGQAVYTRRLVIGK-

EF_3021 Variants

The two EF_3021 variants EF_3021$_{W86G}$ and EF_3021$_{W86A}$ were constructed using site directed mutagenesis using the pET19-EF_3021 as a template and synthetic primers from Microsynth AG. The PCR fragments as well as the target pET19 vector were digested with the restriction enzymes XhoI and NdeI. The fragments were ligated into the target vector using T4 Ligase and the final construct was confirmed by sequencing analysis (Microsynth AG).

Sequence of pET19 EF_3021$_{W86A}$

MGHHHHHHAENLYFQGHMKVTQLSSETLDRAHERFEETLAQMTVAEANTMPAPLIKSVTWLMWHTARELDLQISAL
NHSDPLWLSQHWTEKFALDLPDETEDAHHTPEEAAKVVVAEKQLLSDYLAASVALTKSYLDQIKEEQLSDVIDKNWTP
PVTRQVRLVSAIDDAVMHSGQAVYTRRLVIGK-

Sequence of EF_3021$_{W86G}$

MGHHHHHHAENLYFQGHMKVTQLSSETLDRAHERFEETLAQMTVAEANTMPAPLIKSVTWLMWHTARELDLQISAL
NHSDPLWLSQHWTEKFALDLPDETEDGHHTPEEAAKVVVAEKQLLSDYLAASVALTKSYLDQIKEEQLSDVIDKNWTP
PVTRQVRLVSAIDDAVMHSGQAVYTRRLVIGK-

Table 9. Sequence of the used primers for the mutation EF3021 W86A and W86G.

Nr.	Name	Sequence
28	EF3021 W86As	AAA CCG AAG ATG CGC ATC ACA
29	EF3021 W86Aa	TGT GAT GCG CAT CTT CGG TTT
30	EF3021 W86Gs	AAA CCG AAG ATG GCC ATC ACA
31	EF3021 W86Ga	TGT GAT GGC CAT CTT CGG TTT

Protein production

The expression vector pET19 containing *Ef*GST was transferred into BL21 (DE3) pLys. The cells were grown in lysogeny broth containing ampicillin (100 mg/l) and chloramphenicol (34 mg/l) at 37°C. After the OD_{600} of the cell culture reached 0.8 the cultures were induced by adding 0.1 mg/ml isopropyl-β-D-thiogalactopyranosid (IPTG). Protein was expressed for another three hours, harvested by centrifugation (7000 rpm) and stored at -20°C until protein purification. For the purification the cells pellet was resuspended into lysis buffer (300 mM NaCl , 50 mM Na_2HPO_4) and lysed by sonication. To remove the insoluble cell components the mixture was centrifuged for one hour at 4000 rpm at 4°C. Ni^{2+} NTA agarose beads (Qiagen) were added and after binding of the enzyme for 20 minutes at 4°C and washing with 20 mM imidazole, the purified protein was eluted with 250 mM imidazole. The N-terminal His_6-tag of the purified protein was then removed with the recombinant tabacco etch virus (TEV) protease in 50 mM Tris pH 8.0, 50 mM NaCl, 0.5 mM EDTA and 1 mM DTT. The cleavage was monitored with SDS-page and finished after 36 hours at 4°C. The protein mixture was then filtered over Ni^{2+} NTA agarose, incubated for 10 minutes with 10 mM EDTA and purified by size exclusion chromatography.

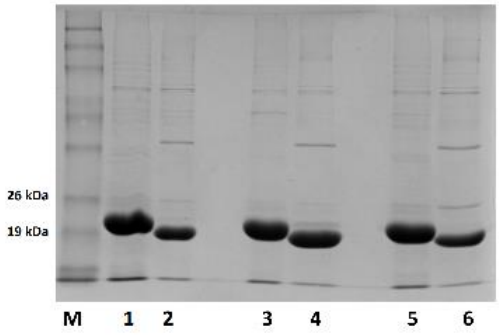


Figure 49. SDS-page picture of EfGSTwt, $EfGST_{W105A}$, $EfGST_{W105G}$. M: Molecular weight marker; 1 EfGSTwt with His_6-tag, 2 EfGSTwt without His_6-tag, 3 $EfGST_{W105A}$ with His_6-tag, 4 $EfGST_{W105A}$ without His_6-tag, 5 $EfGST_{W105G}$ with His_6-tag, 6 $EfGST_{W105G}$ without His_6-tag.

Table 10. HRMS measurements of the purified proteins with the calculated mass.

Enzyme	His-Tag	Calculated Mass	Measured Mass
EF_3021 wt	+	21199	21198
EF_3021 wt	-	19453	19452
EF_3021 W105A	+	21084	21083
EF_3021 W105A	-	19338	19337
EF_3021 W105G	+	21070	21068
EF_3021 W105G	-	19324	19323

FPLC analysis

The quaternary structures were analyzed by size exclusion (Äkta FPLC, GE Healthcare) using a Superdex 200 5/150 GL column. Protein samples (0.1-0.5 mg) were injected and eluted in a constant flow of 0.2 ml/min in degassed FPLC buffer (200 mM NaCl, 50 mM Tris/HCl pH 8) at room temperature.

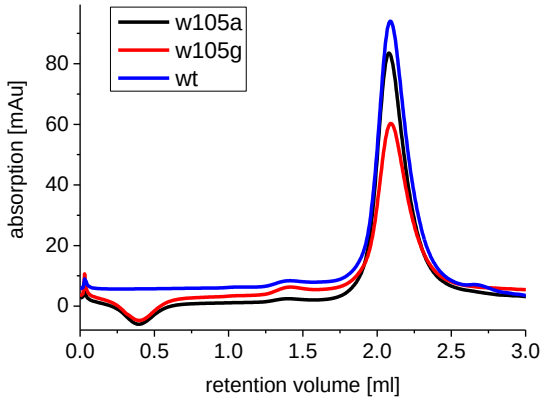


Figure 50. FPLC traces of EF_3021, and the two mutants W105A and W105G in degased FPLC buffer (200 mM NaCl and 50 mM Tris/HCl pH 8).

HPLC Assay

The samples were analyzed by RP-HPLC (Gemini-NX, 5 µM C18 150 x 4.6 mm column) with the gradient shown in Table *11*. The product concentration was determined with a calibration curve of the substrates since there is no expected change of the extinctions coefficient between the substrate and the product. The recorded chromatograms were at 265 nm for 4-Nitrobenuylchloride, 230 nm for 4-chloromethyl benzonitrile and

Table 11. HPLC gradient used for the measurements of reaction rates. Buffer A: H2O, 1 % MeCN, 0.1 % TFA. Buffer B: MeCN, 0.085% TFA. Flow rate 0.75 ml/min.

Time [min]	Buffer B [%]
0	0
12	95
13.5	95
14	0
16	0

HPLC assay: Analysis of reaction rates and Michaelis-Menten kinetic

The enzyme EF_3021 was diluted to a final concentration of 1 µM in a reaction condition containing 50 mM HEPES buffer pH 7 and 50 mM NaCl. The different metals were supplemented as corresponding salts ($ZnCl_2$, $NiCl_2$, $CoCl_2$*6 H_20, $FeSO_4$, $MnCl_2$) to a final concentration of 10 µM. In the reaction with $FeSO_4$, measurements with and without 0.5 mM ascorbate were recorded. The artificial substrates benzylchlorides (*4*-nitrobenzylchloride, *4*-(chloromethyl)benzonitrile or *4*-metoxy benzylchloride) were added to a final concentration of 500 µM. The glutathione concentrations were varied between 5-10000 µM. The reaction

was incubated at a constant temperature of 25°C and time point samples were quenched with 2% TFA. Samples were analyzed by RP-HPLC as described above. The reaction at pH 5 were performed with a acetic acid buffer instead of a HEPES buffer.

Michaelis Menten plots:

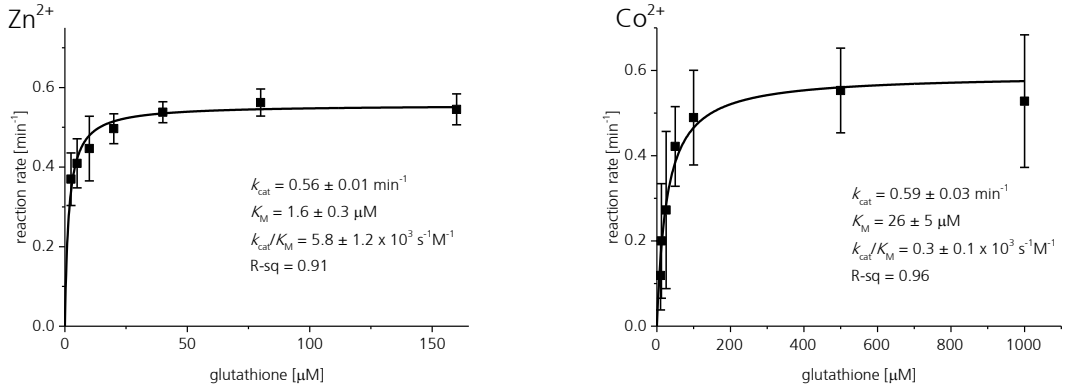


Figure 51. **Left**: Michaelis-Menten plot with 10 µM $ZnCl_2$, 1 µM enzyme, 0.5 mM *4*-nitrobenzylchloride and changing glutathione concentrations. **Right**: Michaelis-Menten plot with 10 µM $CoCl_2$, 1 µM enzyme, 0.5 mM *4*-nitrobenzylchloride and changing glutathione concentrations.

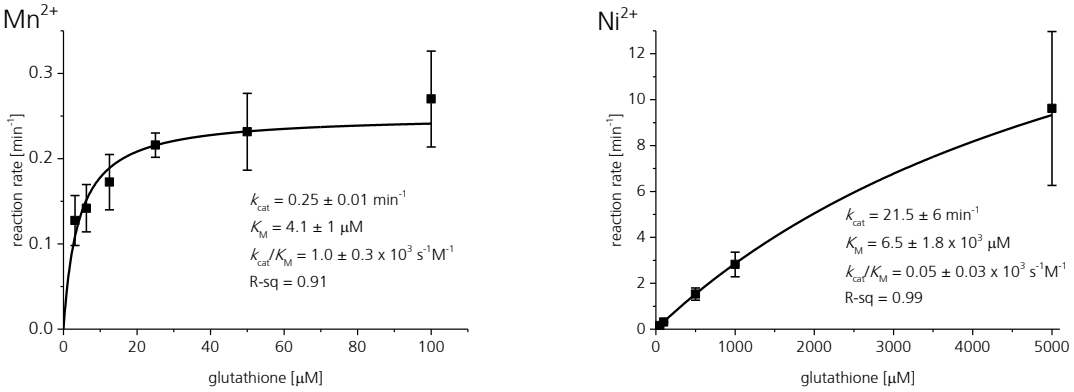


Figure 52. **Left**: Michaelis-Menten plot with 10 µM $MnCl_2$, 1 µM enzyme, 0.5 mM *4*-nitrobenzylchloride and changing glutathione concentrations. **Right**: Michaelis-Menten plot with 10 µM $NiCl_2$, 1 µM enzyme, 0.5 mM *4*-nitrobenzylchloride and changing glutathione concentrations.

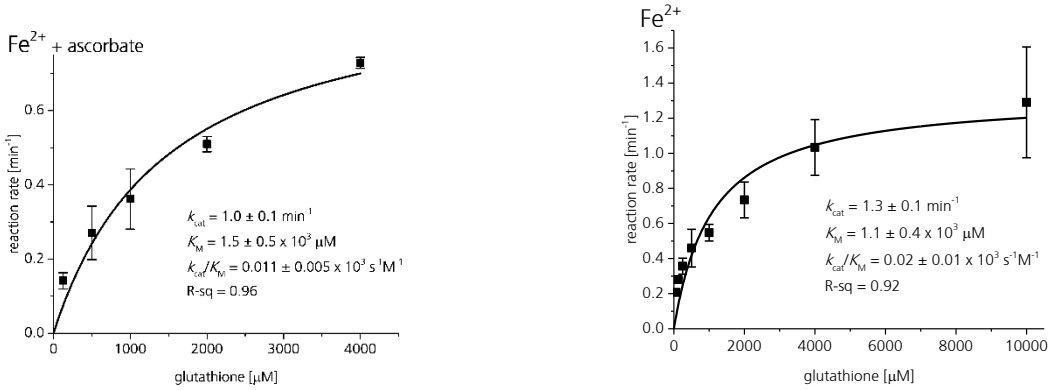


Figure 53. **Left**: Michaelis-Menten plot with 10 µM $FeSO_4$, 0.1 mM ascorbate, 1 µM enzyme, 0.5 mM *4*-nitrobenzylchloride and changing glutathione concentrations. **Right**: Michaelis-Menten plot with 10 µM $FeSO_4$, 1 µM enzyme, 0.5 mM *4*-nitrobenzylchloride and changing glutathione concentrations.

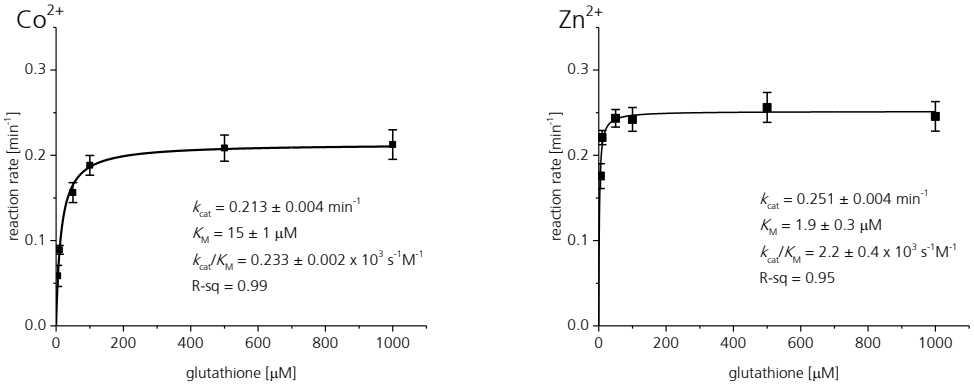


Figure 54. **Left**: Michaelis-Menten plot with 10 µM $CoCl_2$, 1 µM enzyme, 0.5 mM *4*-(chloromethyl)benzonitrile and changing glutathione concentrations. **Right**: Michaelis-Menten plot with 10 µM $ZnCl_2$, 1 µM enzyme, 0.5 mM *4*-(chloromethyl)benzonitrile and changing glutathione concentrations

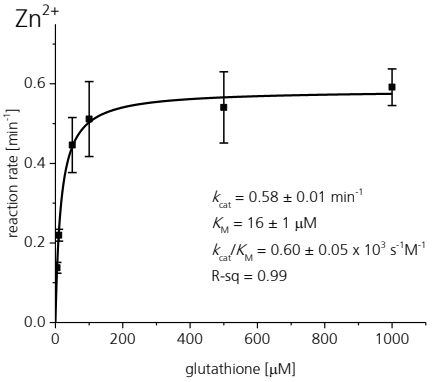


Figure 55. Michaelis-Menten plot with 10 µM $ZnCl_2$, 1 µM enzyme, 0.5 mM *4*-nitrobenzylchloride and changing glutathione concentrations at a pH of 5.

Reaction with Styrol

The enzyme EF_3021 was diluted to a final concentration of 2 µM in a reaction condition containing 100 mM HEPES buffer pH 7 and 40 mM NaCl. $FeSO_4$, was added to a final concentration of 10 µM with and without ascorbate. The reaction was incubated at room temperature and was measured before and after two times 30 min UV light (366 nm) radiation.

Reaction with 2-Nitroimidazole and 4-Nitroimidazole

The enzyme EF_3021 was diluted to a final concentration of 2 µM in a reaction condition containing 100 mM HEPES buffer pH 7 and 40 mM NaCl. $ZnCl_2$ 10 µM and glutathione 1 mM were added, and the reaction was started with 1 mM of the imidazole derivate. The reaction was monitored over-night and analyzed according to the HPLC method described above.

UV-Vis

To a diluted enzyme concentration (500 µM) in HEPES buffer (50 mM) and NaCl (50 mM), $CoCl_2$ (1 mM) and glutathione (2 mM) were added. In the control experiment only $CoCl_2$ and glutathione were mixed in the same buffer system to exclude non-enzymatic sulfur-metal charge transfer.

Crystallization:

For the crystallization experiments the His_6-tag cleaved and FPLC purified protein was loaded with different metal cations. For the Ni^{2+} and Zn^{2+} cations the protein was dialyzed against a low metal concentration ($NiCl_2$ or $ZnCl_2$) over night to avoid protein precipitation. Manganese and cobalt chloride were added directly to the enzyme solution (2 mM end concentration). The enzyme was then concentrated using centrifuge concentrators (Millipore) to a final concentration of 40 mg/ml and flash frozen in liquid nitrogen. The protein was screened for crystallization with commercial available crystal screens from Molecular Dimensions using 96-well 3-drop Swissci plates. Each drop contained 0.2 µl protein solution and 0.2 µl crystallization solution at a constant temperature of 20°C. For a successful crystallization, the reservoir solution needed to be changed to 1.5 M NaCl solution as reported for the enzyme GsDinB of *Geobacillus stearothermophilus*.[118] Single crystals were harvested out of the optimized JSCG+ position H3 screen, which contained 0.1 M Bis-Tris pH5.5, 15-30% of the precipitant PEG 3350 and reservoir drops containing 1-2 M NaCl. Typical observed rectangular crystals appeared after 2 days (figure x).

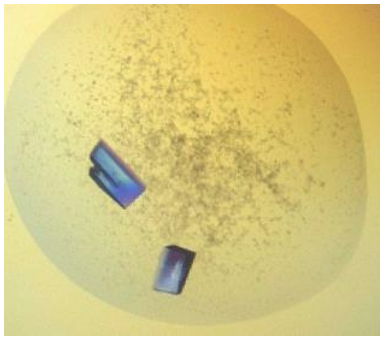


Figure 56. Observed crystals of EF_3021 in the iotimized crystal screen.

The crystals were fished, flash frozen in liquid nitrogen and stored until data collection. The data were collected by Marcel Meury at the Swiss Light Source (SLS), Villigen, Switzerland using a Pilatus 6M detector. The structure were solved using the template (PDB 3Cex)

3 EgtB Homologues in Eukaryotes

3.1 Introduction $EgtB_{fungi}$

3.1.1 EgtB a DinB Protein

In the previous chapter the mechanism of a C-S bond forming DinB metalloenzyme was described. Glutathione *S*-transferase was active with different metals forming the C-S bond *via* an ionic mechanism. Another enzyme, which contains a DinB domain, is EgtB, and is involved in the biosynthesis of ergothioneine.[71] The crystal structure showed that EgtB consists of two domains (Figure 57). The DinB domain, a four-helical bundle fused to a formylglycine (FGE)-sulfatase-like domain, is comprised of a C-type lectin fold and an extended parallel beta sheet. The active site of the enzyme is located between the domains and contains a metal binding site made up of a facial histidine triad.[60] The architecture of this metal binding site is fairly conserved in the DinB protein superfamily as seen in EgtB and the glutathione-*S*-transferase EF_3021. Despite a similar metal binding site, the mechanisms by which EgtB and EF_3021 catalyzes C-S bonds differs. While EgtB catalyzes an O_2-dependent oxidative C-S bond formation, EF_3021 catalzyzes an alkyl transfer in an ionic mechanism. Our studies on the bacterial EgtB revealed a redox active iron ion, which activates molecular oxygen to form the new C-S bond in a 4 e^- oxidation reaction.[60,74] This reaction type is called oxidative sulfurization.[71]

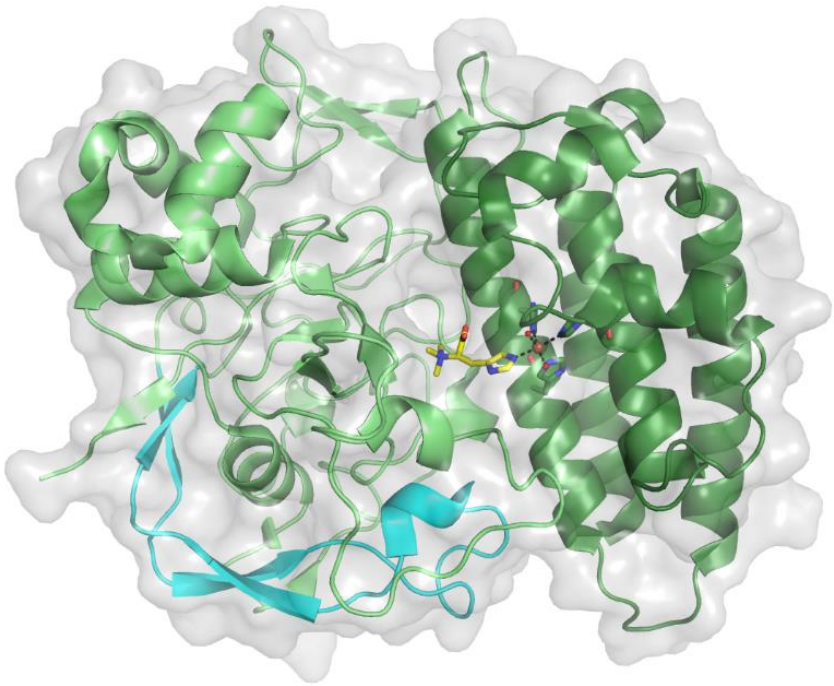
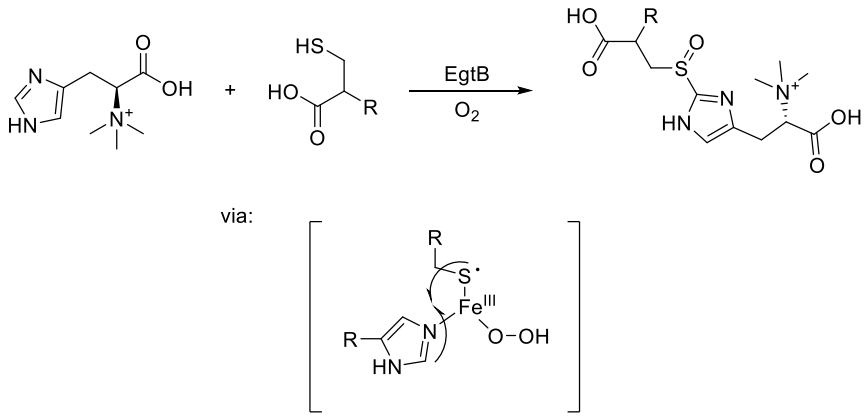


Figure 57. Crystal structure of the bacterial EgtB (*M. thermoresistibile*) with the ligands, γ-glutamylcysteine and *N,N*-dimethyl-*L*-histidine, the DinB domain in highlighted in blue (PDB: 4x8e, 1.6 Å).[60]

Thus far only bacterial EgtB enzyme systems have been studied in our group. However, ergothioneine biosynthesis is also found in eukaryotic species. In this chapter we have a closer look to the EgtB homologue found in the eukaryotic EgtB biosynthesis. We examine the similarities and differences to the bacterial system and contribute understanding of the reaction mechanism of the EgtB enzymes.

3.1.2 The Variety of EgtB enzymes

In the biosynthesis of ergothioneine the enzyme EgtB catalyzes the reaction between *N,N,N*-α-trimethylhistidine (TMH), γ-glutamylcysteine and dioxygen to form the sulfoxide product. In the reaction mechanism of EgtB (Scheme 20, described in short below) the oxygen is activated *via* an iron(III)superoxo species to form a thiyl radical, which proceeds to form the new C-S bond. Despite the considerable complexity of the EgtB reaction a surprisingly low conservation in the enzyme architecture is observed in Nature.



Scheme 20. Reaction catalyseld by EgtB. The formation of the C-S bond goes *via* a thiyl radical.[60]

At least six different EgtB related enzyme types have been identified so far (Figure 58). The EgtB enzyme from *Mycobacteria thermoresistibile* (EgtB type M) has been characterized in details.[60,71,74,83] The enzyme size is 445 amino acids and the substrates were identified to be γ-glutamylcysteine (γ-GluCys) and trimethylhistidine (TMH).[71,72] In the active site of the enzyme a mechanistically relevant tyrosine residue, which acts as catalytic acid, has been identified.[60,83]

In the fungal *N. crassa* another distinct EgtB type (EgtB fungi) was found. The size of this eukaryotic EgtB is 876 amino acids, which contains next to a EgtB domain a fused EgtD domain.[67] Furthermore the substrate γ-glutamylcysteine is changed to cysteine.[153]

Recently, we discovered a second eukaryotic EgtB type in the organism of *Dichomitus squalens*, which differs from the other fungal EgtB in size and misses the catalytic tyrosine residue. This EgtB enzyme (fCDO) was identified to be a cysteine dioxygenase and will be discussed in the second part of this chapter.

A EgtB homologue from the organism of *Variovorax paradoxus*, which has significant sequence differences to the bacterial EgtB enzymes. This EgtB homologue with unknown functions will be described in chapter 4.

A EgtB related enzyme type are the OvoA enzymes, which are involved in the biosynthesis of ovothiol A and differs in the regioselectivity of the C-S bond formation.[154]

In 2017 a short OvoA-type enzyme ($EgtB_{ovo}$) was reported from the cyanobacteria *Microcystis aeruginosa*.[155] *In vitro* characterization showed that this enzyme is able to produce out of TMH and cysteine the C2-sulfoxide product found as an intermediate in the ergothioneine biosynthesis. This example of a OvoA homologues with EgtB functionality is an example of convergent evolution and a distinct catalytic EgtB-type.[155]

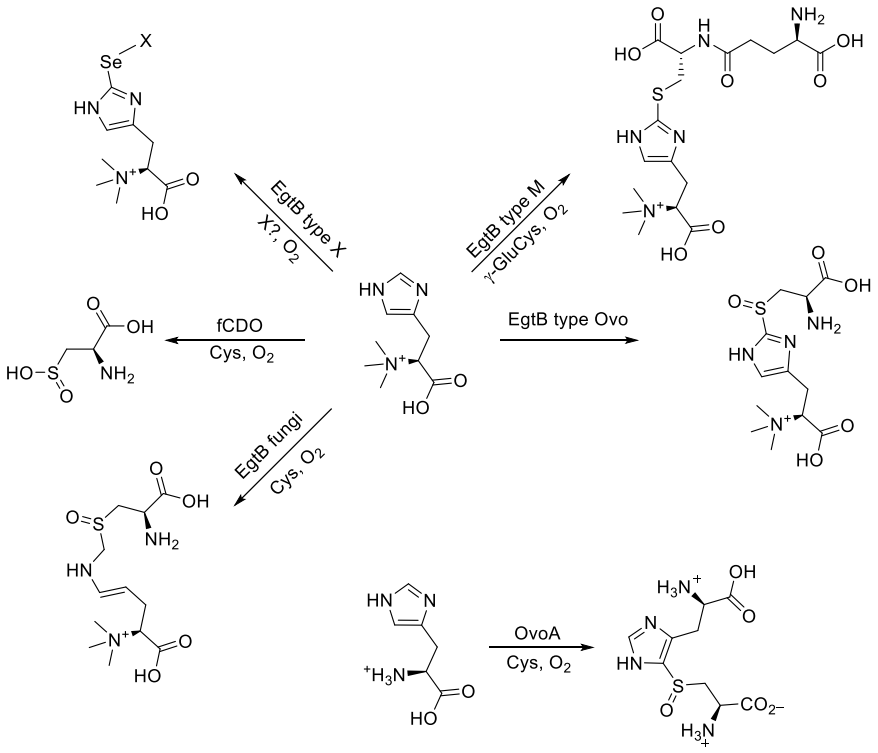
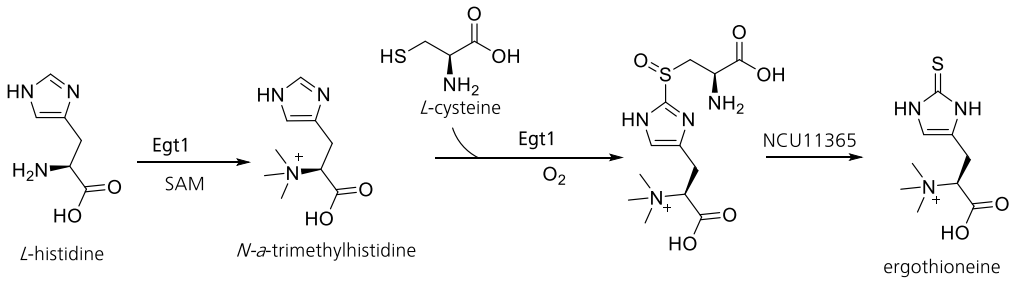


Figure 58. Six different types of EgtBs, which are found in Nature. EgtB type M from Mycobacteria, EgtB type Ovo found in cyanobacteria.[60,155] The EgtB fungi is found in *ascomycota* and *basidiomycota*, while the fCDO type is only found in *basidiomycota*.[67] The unknown EgtB type is found in proteobacteria (unpublished data). The related OvoA type which is involved in the ovothiol biosynthesis rather than the ergothioneine biosynthesis is found in enterobacteria.[154]

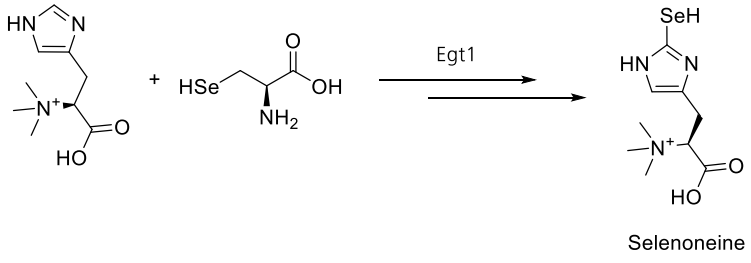
3.1.3 The Eukaryotic Ergothioneine Biosynthesis

Ergothioneine biosynthesis in eukaryotes differs from the bacterial pathway described in *M. smegmatis* (Chapter 1) in numerous aspects (Scheme 21).[67,153] Firstly, in eukaryotes the enzymes EgtD and EgtB are found solely as a single fusion protein Egt1, which can catalyze both the SAM-dependent trimethylation of histidine as well as the following oxidative sulfurization.[69,153,156] In contrast, the enzymes in the bacterial pathway are encoded by separate polypeptides for each respective function. Moreover, the EgtB domain of EgtB uses cysteine as a substrate instead of γ-glutamylcysteine. Therefore, the γ-glutamyl cysteine synthetase EgtA and the glutamine amidotransferase EgtC, which are present in *M. smegmatis* no longer required. The eukaryotic biosynthestic route is also now much shorter, with the fusion protein and the use of cysteine as a substrate, EgtA and EgtC are effectively redundant, and only two enzymes are required for the ergothioneine biosynthesis. In summary, the eukaryotic ergothioneine biosynthetic pathway consists of the EgtD/EgtB fusion protein, Egt1, and a C-S lyase (NCU11365 in *Neurospora crassa*).[153]
In addition, a sequence alignment of the EgtB of the bacterial enzyme and the eukaryotic enzymes (see discussion) revealed differences in the binding mode of the sulfur-containing substrate.



Scheme 21. The three steps of ergothioneine biosynthetic pathway in the fungus *N. crassa*, catalyzed by two enzymes Egt1 and the C-S lyase NCU11365.[153]

Another difference of Egt1 to bacterial EgtB is the proposed connection of Egt1 to the biosynthesis of selenoneine, the selenium containing analogue of ergothioneine (Scheme 22). *In vivo* studies with overexpressed Egt1 enzyme in the fungi *Schizosaccharomyces pombe* selenoneine formation was observed.[156]



Scheme 22. Proposed biosynthesis of selenoneine in fission yeast *S. pombe*.[156]

These differences between the bacterial and the eukaryotic biosynthesis of ergothioneine prompted us to study the eukaryotic EgtBs from the organisms *Chaetomium thermophilum* and *Thermothelomyces thermophilia*. By sequence alignment and homology studies these homologues lack the cysteine binding motif conserved in bacteria, which must thereby bind cysteine using an alternative binding site to that used in bacteria. The kinetic characterization showed that the activities of the fungal EgtBs were in a similar range to the bacterial enzymes. While their *in vitro* activity with selenocysteine was greater than that of bacterial EgtBs. With the identification of an alternative substrate selenocysteine, we were able to identify significant differences in the oxygen activation between the bacterial and fungal EgtB enzymes. In the second part of the chapter we identified an EgtB homologue in *D. squalens* with cysteine dioxygenase activity. Studying similar enzymes from different organisms gives us the opportunity to establish a research platform for comparative enzymology, which in turn allows us to isolate the core catalytic principles.

3.2 Aim of the Chapter

To complement our knowledge on C-S catalysis by EgtB enzymes we aimed to characterize the fungal EgtB/EgtD fusion type *in vitro*. We wanted to probe the enzymatic reaction with selenocysteine and its possible role in eukaryotic selenoneine biosynthesis. A structural characterization would provide a strong platform for mechanistic hypothesis, therefore we attempted to obtain a protein crystal structure of the novel EgtB/EgtD fusion architecture.

In the second part of the chapter we aimed to identify the function of a unique EgtB homologue in basidiomycetes. In *Dichomitus squalens* we identified next to the EgtB/EgtD fusion protein a second EgtB homologue, which misses the catalytic acid Y377 identified in $EgtB_{thermo1}$. This let us speculate that this enzyme is a cysteine dioxygenase rather than a sulfoxide synthase. Next to the functional characterization we aimed to identify the distribution of these EgtB homologues in Nature.

3.3 Results and Discussion: $EgtB_{fungi\_C}$ and $EgtB_{fungi\_T}$

3.3.1 The Enzyme $EgtB_{fungi}$

To investigate the properties of the ergothioneine biosynthesis in eukaryotes, we chose two homologues of Egt1 from *Neurospora crassa,* $EgtB_{fungi\_C}$ from the thermophilic fungi *Chaetomium thermophilum* as well as $EgtB_{fungi\_T}$ from *Thermothelomyces thermophilia*.[153] We focused on the second step of the ergothioneine biosynthetic pathway, the C-S bond formation, which is catalyzed by the EgtB domain. In the first step we compared the enzyme sequences and a homology model of the $EgtB_{thermo\_C}$ with the EgtB from *M. thermoresistibile*. ($EgtB_{thermo1}$) to identify conserved residues which are presumably involved in the substrate binding.

3.3.2 TMH Binding

Both the fungal and bacterial enzymes catalyzing the formation of the sulfoxide intermediate use trimethylhistidine (TMH) as a substrate. The crystal structure of $EgtB_{thermo1}$ with TMH bound (PDB: 4x8e, 1.6 Å) allowed for the identification of the amino acid residues which are presumably involved in TMH binding. Tyr380 forms a water-mediated hydrogen bridge to the imidazole ring. Two *N*-α-methyl groups of TMH are pointing against the indole ring of Trp415. The third *N*-α-methyl group forms dipolar interactions with Asn414 and Gln137. We looked for conservation of these residues in sequence alignments and homology modelling of the fungal enzyme. Residues Tyr842, Trp874 and Asn873 were found to be conserved therefore were suggested as possible TMH-binding residues (Figure 59). Only one difference was observed in the fungal enzyme. Instead of a Gln137, which interacts with the third *N*-α-methyl group, a methionine residue Met532 was identified. The change from a polar residue Gln137 to the unpolar methionine in an active site could be problematic, because the polar interaction between the substrate and the enzyme disappeared. However, in the case of $EgtB_{fungi\_C}$ we expected no significant change in the TMH binding properties because all the other identified TMH-binding residues are conserved.

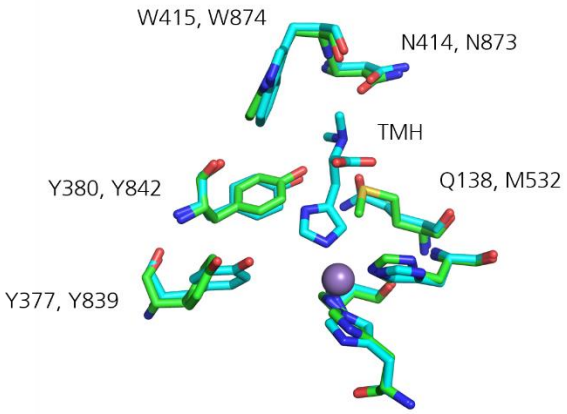


Figure 59. Superposition of the TMH binding residues of the crystalstructure (pdb: 4x8e, 1.6 Å, cyan) and a homology model of $EgtB_{fungi\_C}$ (green). The first number of the labeled amino acids corresponds to the residue position in bacterial EgtB, the second number to the fungal EgtB.

3.3.3 Cysteine Binding

The thiol-containing second substrate differs between the fungal and bacterial enzymes. $EgtB_{fungi\_C}$ uses cysteine, instead of γ-glutamylcysteine (γ-GluCys), as a substrate. We were interested in comparing the binding sites of the substrates to identify differences associated with the binding. For $EgtB_{thermo1}$, it was possible to solve the ternary complex with manganese, dimethylhistidine and γ- GluCys bound (PDB: 4x8d, 1.98 Å).[60] In this structure, the residues for γ-GluCys were identified. Arg87 and Arg90 forms salt bridges to the carboxylic group of the cysteine moiety of γ-GluCys. In addition, salt bridges are formed between the carboxylate group of the glutamyl residue and the arginines, Arg420 and Arg416. We expected to see a similar substrate binding architecture were the two arginine Arg87 and Arg90, which form a RXXR binding motif are conserved.

In the sequence of the fungal EgtB, neither of the two arginines were found to be conserved. Furthermore, the sequence alignment lacks any conservations (Figure 60, Figure 61). The absence of an RXXR motif allowed us to propose that the fungal-type enzymes might have an alternative cysteine binding mode.

In addition, we identified Arg465 in the fungal gene of $EgtB_{fungi\_C}$, which is highly conserved in all examined fungal sequences. It is tempting to presume that this arginine is involved in the binding of the carboxylate group of the substrate cysteine.

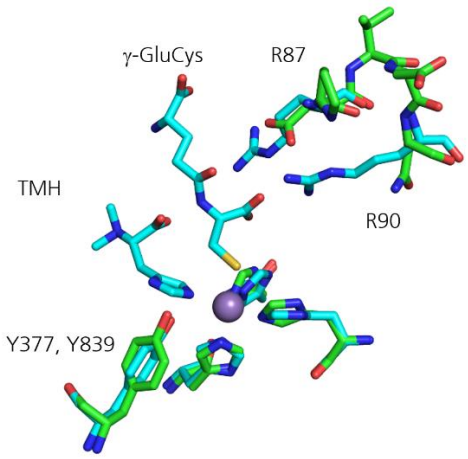


Figure 60. Superposition of the homology model of $EgtB_{fungi\_C}$ and the crystal structure of $EgtB_{thermo1}$ (PDB: 4x8d, 1.98 Å). The arginine residues of the bacterial enzymes which are involved in the cysteine binding are highlighted. In the homology model a glutamate is located at the position of the arginine.

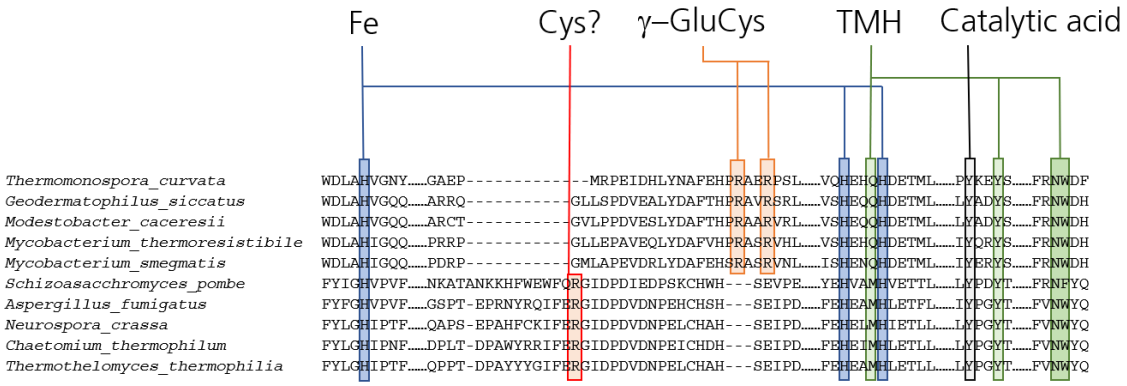
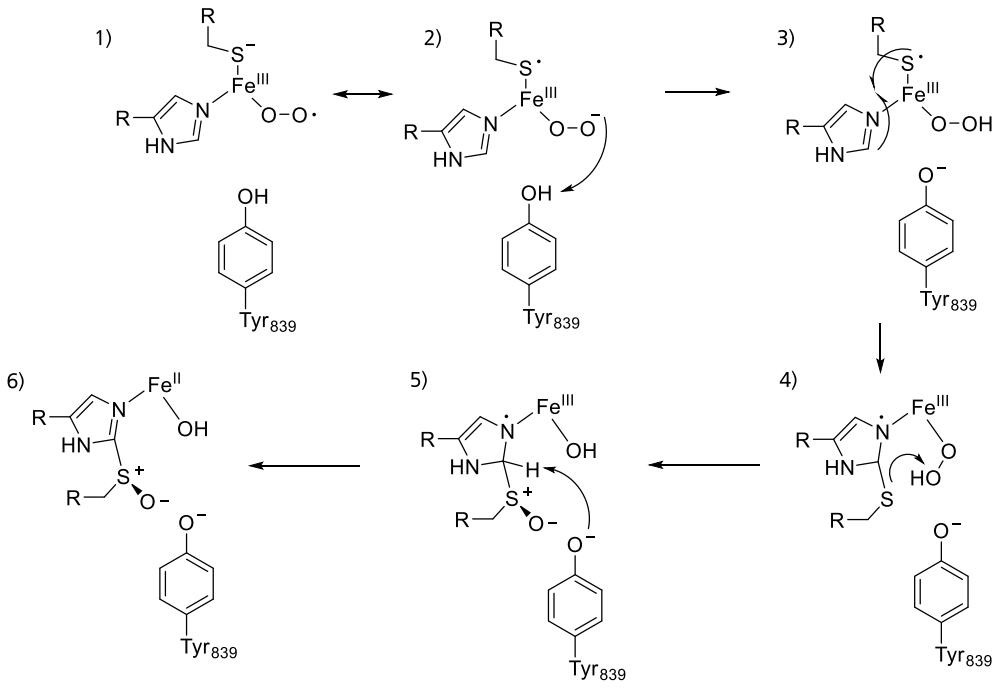


Figure 61. Sequence alignment of selected bacterial and eukaryotic sequences. The amino acid which were identified in EgtB from *M. thermoresistibile* are highlighted. Blue: iron binding; green: TMH binding; black: catalytic acid; orange: *γ-GluCys* binding; red: conserved arginine in the fungal enzyme hypothesized to bind cysteine.

3.3.4 Proposed Mechanism of $EgtB_{fungi}$

As described before we were able to identify a similar TMH binding site and a different cysteine binding site. The 3-His facial triad which is responsible for the iron binding was identified to be conserved (His436, His529 His533). Furthermore the catalytic relevant tyrosine was also found in the fungal enzyme. Therefore we proposed the following mechanism (Scheme 23). In the first step TMH, cysteine and oxygen bind to the ferrous iron (**1**), which is coordinated by the 3-His facial triad. While the binding order of γ-glutamylcysteine and TMH in the bacterial enzyme $EgtB_{thermo1}$ was identified as sequential (TMH binds first), it is not known for $EgtB_{fungi\_C}$.[74,159] The binding of oxygen forms a superoxo iron(III) species, which is in an equilibrium with a peroxide anion (O_2^{-2}) and a thiyl radical (**2**). Proton transfer to the peroxide anion can stabilize this form. In $EgtB_{thermo1}$ the tyrosine residue 377 was identified as proton donor.[60,83] In the sequence of $EgtB_{fungi\_C}$, we were able to identify the conserved catalytic tyrosine Tyr839. The stabilization of the peroxide anion enables the radical attack of the thiyl radical on the C5 position of the imidazole ring, forming the C-S bond (**3**), which is further oxidized to the sulfoxide (**4**). In the next step, the imidazole ring is deprotonated by the tyrosine residue (**5**), which results in the rearomatization of the imidazole ring and a ligand-to-metal electron transfer. Finally, the sulfoxide product is released from the enzyme (**6**).



Scheme 23. Proposed reaction mechanism of $EgtB_{fungi\_C}$ For simplification the coordinating 3-His facial triad is not shown **1)** The binding of the substrates cysteine, TMH and oxygen forms a the octahedral-coordinated superoxo iron(III) species; **2)** This species is in equilibrium with a thiyl radical and peroxide ion. The peroxide ion can be deprotonated by the Tyr839 residue to stabilize the thyil radical; **3)** The stabilized thiyl radical is able to attack the imidazole ring at the C5 position, forming the C-S bond; **4)** The selenoether is then oxidized by the peroxide to form the sulfoxide; **5)** Deprotonation of the imidazole ring leads to rearomatization; **6)** The formed product is released from the active site.

3.3.5 Kinetic Characterization

For the kinetic characterization of the eukaryotic EgtB we expected to observe a similar product formation rate as seen in the bacterial enzymes. Therefore we chose the same assay used for kinetically characterization of bacterial EgtB.[60]

This kinetic assay is based on the monitoring of the product formation with a cation-exchange HPLC. The expected product is a sulfoxide which has a specific maximal UV-absorption at 265 nm.[74] The reaction conditions contained 100 mM phosphate buffer, 100 mM NaCl, 2 mM ascorbate, 2 mM TCEP, 1 µM enzyme and 10 µM $FeSO_4$ and the reaction temperature was set at 20 °C. To quantify the kinetic data, we used steady state Michaelis-Menten kinetics. Under saturating conditions of one substrate, we changed the concentration of the second substrate and measured the reaction rate of the enzyme at each point. These data points were then fitted with the Michaelis-Menten equation using OriginPro (Eq. 1). With this method we measured the steady-state kinetic Michaelis-Menten parameters K_M and v_{max} of each substrate for both enzymes $EgtB_{fungi\_C}$ and $EgtB_{fungi\_T}$. In addition, the enzyme concentration independent reaction rate k_{cat} of the enzymes was calculated (Eq. 2).

$$v_0 = \frac{v_{max} \cdot [S]}{K_m + [S]} \qquad \text{(Eq. 1)}$$

$$k_{cat} = \frac{v_{max}}{[E] + [ES]} \qquad \text{(Eq. 2)}$$

The plasmid of the two thermophilic EgtB homologue enzymes were purchased from GenScript and successfully recombined expressed in BL21 cells of *E. coli* in a yield of 1-2 mg per liter of cell culture. The measured reaction rate k_{cat} for the two thermophilic fungal enzymes $EgtB_{fungi\_C}$ (0.07 s^{-1}) and $EgtB_{fungi\_T}$ (0.06 s^{-1}) were similar to each other. Compared to the bacterial enzymes from *Mycobacteria*, which were measured with the same assay at a slightly higher temperature (26 °C), these rates are much lower. EgtB from *Mycobacteria smegmatis* is 17 times and EgtB from *Mycobacteria thermoresistibile* ($EgtB_{thermo1}$) is 12 times faster than $EgtB_{fungi\_C}$ (Figure 62).[60] Compared to another bacterial EgtB from the organism *Candidatus chloracidobacterium thermophilum* ($EgtB_{thermo2}$), the difference is smaller. The reaction rate of $EgtB_{thermo2}$ is with 0.2 s^{-1} only 3 times faster than the fungal enzymes (Figure 62).[74]

The enzyme Egt1 from the fungi *N. crassa* was characterized by a different group with a different assay, where the oxygen consumption was measured with a NeoFox oxygen electrode.[153] In their measurements they observed a fast reaction rate of 2.3 s^{-1}. This rate is 33 times faster than the observed rates with our enzymes (Figure 62). These results are rather surprising, since the fungal enzymes are related to each other (66% sequence identity). A possible explanation is the different measurement methods used in the kinetic assay. It is possible that there are side reactions occurring during catalysis which might consume additional

oxygen and would have a high impact to the measured oxygen consumption. In addition, the Egt1 enzyme from *N. crassa* was purified under anaerobic conditions, this might also increase the activity of the enzyme since it is protected from oxidative stress and possible oxygen-induced damaging during the purification steps. However, we observed only a minimal decrease in activity if the fungal enzymes were exposed to the air for several hours.

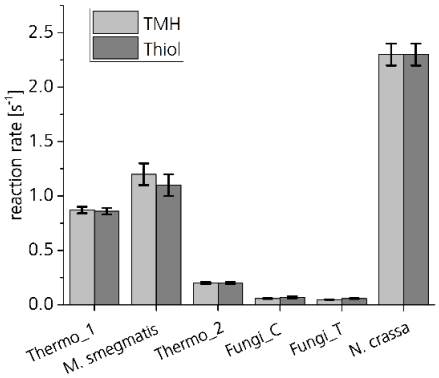


Figure 62. Reaction rate of the bacterial EgtB enzymes of *M. thermoresistibile* (thermo1) *M. smegmatis*, *C. thermophilum* (thermos 2) and the fungal enzymes $EgtB_{fungi\_C}$, $EgtB_{fungi\_T}$ and Egt1 from *N. crassa*. Light gray is k_{cat} with the saturated substrate TMH. Dark gray is k_{cat} with the saturated thiol substrate.

The K_M values for the $EgtB_{fungi}$ enzymes were low for the substrate TMH (11 µM). This indicates a strong affinity of TMH towards the active site of the EgtB-domain of the enzyme. The K_M of the second substrate cysteine is in the same range as in the bacterial enzymes, between 38 and 64 µM. The kinetic parameters of the fungal EgtB are listed in Table 12.

If we compare the K_M values of our enzymes to the values reported for Egt1, it is apparent that the K_M-values for $EgtB_{Fungi\_C}$ and $EgtB_{fungi\_T}$ are 10-40 times lower. The reported K_M values of Egt1 are rather high with 436 µM for TMH and 603 µM for cysteine and seem to be questionable, especially for the cytotoxic amino acid cysteine which is normally present in cells and only in very low concentrations of 10-100 µM.[97] The catalytic efficiency, which is defined by k_{cat}/K_M, for the fungal enzymes were within the same range between 900-5400 $s^{-1}M^{-1}$. The difference in the measured kinetic data might be influenced more strongly by the different measuring methods and not because of the differences in the amino acid sequence (sequence identity of 66%) between the enzymes.

Table 12. Michaelis-Menten parameters for the thermophilic eukaryotic EgtB. The reaction conditions for $EgtB_{fungi\_C}$ and $EgtB_{fungi\_T}$ were 100 mM phosphate buffer and 100 mM NaCl, 2 mM ascorbate, 2 mM TCEP, 1 µM enzyme and 10 µM $FeSO_4$. The substrate concentrations were changed between 5-1000 µM and the reaction temperature was set to 20 °C and 40 °C.

EgtB	Temp.	Substrate	k_{cat}, [s^{-1}]	K_M [µM]	k_{cat}/K_M [$s^{-1}M^{-1}$]

Chaetomium Thermophilum	*20°C*	TMH: Cys:	0.058 ± 0.002 0.07 ± 0.01	11 ± 2 38 ± 3	(5.4 ± 1.1) x 10^3 (1.8 ± 0.2) x 10^3
Thermothelomyces Thermophilia	*20°C*	TMH: Cys:	0.0485 ± 0.0004 0.058± 0.002	11.6 ± 0.4 64 ± 8	(4.2 ± 0.2) x 10^3 (0.9 ± 0.1) x 10^3
Chaetomium Thermophilum	*40°C*	TMH: Cys:	0.27 ±0.01 0.29 ± 0.01	10 ± 2 58 ± 6	(2.7 ± 0.6) x10^4 (5.0 ± 0.7) x 10^3

Chaetomium Thermophilum is a fungus found in compost and it has an optimal growth temperature of 50-55 °C.[160] Enzymes of thermophilic species normally have an optimal activity close to the optimal growth conditions.[161] Therefore, we measured the Michaelis-Menten kinetic parameters additionally at a higher temperature of 40 °C (Table 12). We observed a 4-fold increase in the velocity, demonstrating that the enzyme is more active at a higher temperature, as expected.

Furthermore, we analyzed the quaternary structure of $EgtB_{fungi\_C}$ with analytical size exclusion chromatography. We observed that the enzyme was present in solution as a dimer with a total mass of 205 kDa. With the addition of the reducing agent dithioreitol (DTT) we observed a more defined signal however a clear identification of a trimer or tetramer formed by a disulfide bridge was not possible (Figure 63).

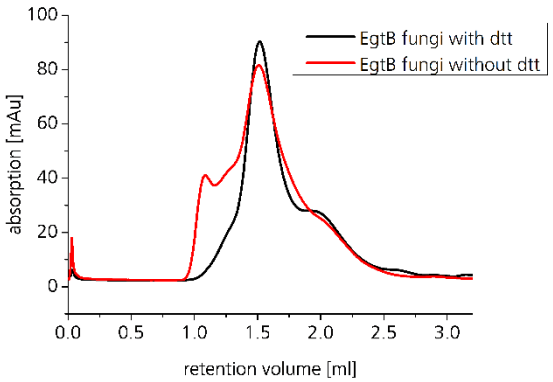
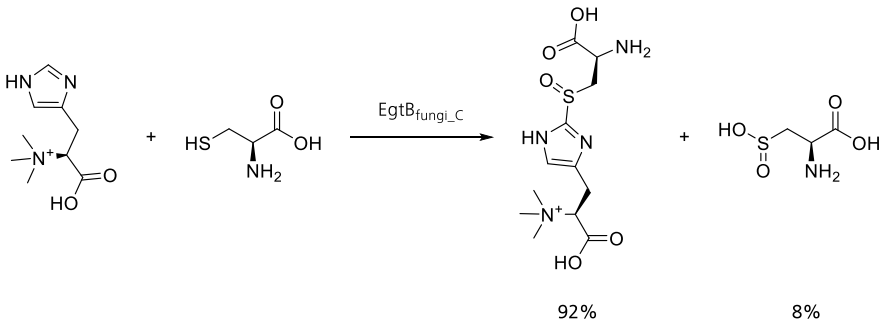


Figure 63. Analytical size exclusion of purified $EgtB_{fungi\_C}$ with and without addition of 10 mM TCEP.

3.3.6 Cysteine Sulfinic Acid as Side Product

In the $EgtB_{fungi\_C}$ reaction mechanism, described in 3.3.1 a crucial step involves the protonation of the superoxo iron(III) species. For this proton transfer from Tyr839 to the oxygen to occur, a precise positioning of the substrates and catalytic residues is required. If this protonation does not occur fast enough, an alternative pathway can occur which results in the formation of a second product, cysteine sulfinic acid. This second parallel pathway is seen in different EgtB and OvoA enzymes.[74,153,162] The proposed mechanism of this cysteine dioxygenase pathway is described in the second part of this chapter.

With a ^{1}H-NMR experiment, we were able to monitor the product distribution between the main sulfoxide product and the side product, cysteine sulfinic acid (CSA) after 12 hours of reaction. The sulfoxide was observed as the major (91.8 ± 1.5 %) and the CSA as a minor product (8.2 ± 1.5 %, Scheme 24). This side reaction was also observed in the enzymes Egt1 from *N. crassa,* where 8% CSA was formed.[153] In the bacterial enzyme EgtB, from *M. thermoresistibile*, the cysteine dioxygenase activity (< 1 %) was minimal.[74] Moreover in the OvoA enzyme (40%) and in EgtB from *Candidatus chloracidobacterium thermophilum* (20%), cysteine sulfinic acid was also observed as side-product (Figure 64).[74,162]



Scheme 24. The reaction catalyzed by the enzyme $EgtB_{fungi\_C}$. Alongside, sulfoxide formation (92%), the side product cysteine sulfinic acid (8%) is produced.

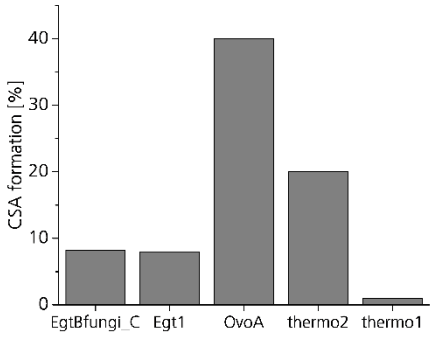


Figure 64. Cysteine sulfinic acid (CSA) formation (%) of the measured $EgtB_{fungi\_C}$ and the data reported from the literature for the enzymes Egt1, OvoA, $EgtB_{thermo1}$ and $EgtB_{thermo2}$.[74,153,162]

3.3.7 Binding Order of $EgtB_{fungi\_C}$

The binding order of an enzyme can provide information about the architecture of the binding site. Furthermore, without the knowledge of the binding order, the kinetic parameters cannot be interpreted since all measured K_M values are apparent K_M values.

If the binding order is sequential, the binding of the first substrate promotes the binding of the second substrate. The binding of the first substrate A can induces a conformational change in the enzyme which allows the second substrate B to bind. Another possibility is that the substrate A forms a portion of the binding site of substrate B. Alternatively, the binding site of substrate A is less accessible than the binding of the substrate B. In a random binding order both substrates can access the active site of the protein with equal likelihood (Figure 65).

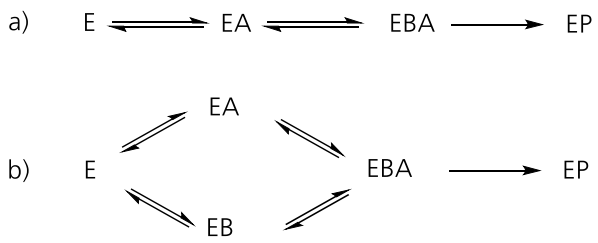


Figure 65. **a)** Sequential binding order of an enzyme system with two substrates. **b)** Random binding order of an enzyme system with two substrates.

In the bacterial EgtB enzymes from *Candidatus chloracidobacterium thermophilum* and *Mycobacteria thermoresistibile*, the binding order was identified to be sequential.[74,159] The TMH first binds to the enzyme, followed by cysteine or γ-glutamylcysteine, respectively. This observation is in agreement with the architecture of the published EgtB structure, where the γ-glutamylcysteine binding site is located on top of the TMH binding site.[60] Since the enzyme architecture of the cysteine binding enzyme is different in the eukaryotic enzyme compared to the bacterial enzyme, we were interested in identifying the binding order of the enzyme $EgtB_{fungi\_C}$.

In a rapid-equilibrium assumption for a bisubstrate mechanism system, the variation of one substrate, while maintaining the concentration of the second substrate, can be used to distinguish between a random or sequential mechanism.[163] In the initial binding order studies, we therefore varied the cysteine concentration at a fixed TMH concentration. We observed that the initial velocity of the reaction increased as the cysteine concentration was raised. The same trend was observed if the cysteine concentration was constant and the TMH concentration was varied, however, the K_M value of the reaction did not change significantly in either case.

In an ordered mechanism the substrate B, which binds second to the enzyme, influences the substrate affinity of the first bound substrate A. If the concentration of substrate B trend to infinite, no free enzyme exists since the second substrate B traps substrate A in the enzyme, forming the EAB complex. Therefore, the K_M value of substrate A should decrease by increasing the concentration of substrate B. The double reciprocal plot (which correlates for V_{max}/K_M) should go through zero at infinitive concentration of substrate B. This is not the case in our measurement and therefore a random ordered binding mechanism was tentatively proposed.

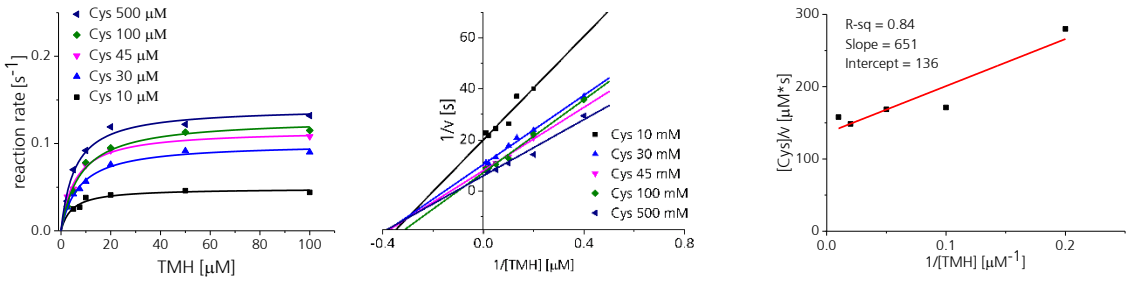


Figure 66. **Left**: Reaction rate of $EgtB_{fungi\_C}$ at a constant cysteine level and different TMH concentrations. **Middle**: A Lineweaver-Burk plot of different cysteine concentrations. **Right**: Slope of the line from the Lineweaver-Burk plot.

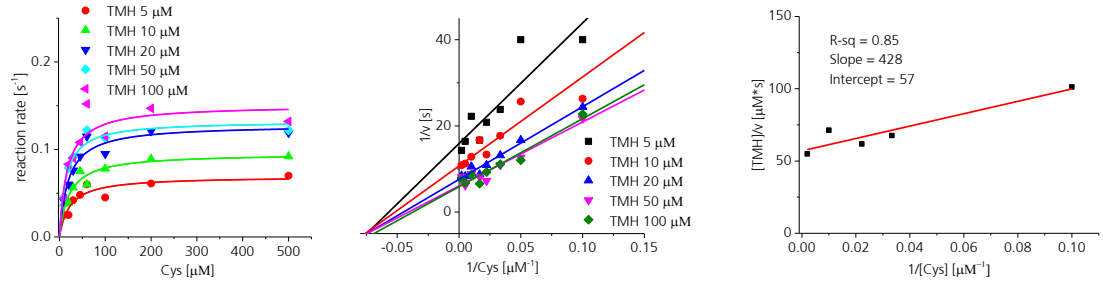


Figure 67. **Left**: Reaction rate of $EgtB_{fungi\_C}$ at a constant TMH level and different cysteine concentrations. **Middle**: A Lineweaver-Burk plot of different TMH concentrations. **Right**: Slope of the line from the Lineweaver-Burk plot.

However, this analysis is not precise because we neglected the third substrate, molecular oxygen. Oxygen is available in concentrations of ~0.25 mM at 20°C.[164] If oxygen concentration is saturating the binding order can look like random even though it might be an ordered substrate binding sequence. For a rigorous analysis of a three substrate system, the inclusion of varying concentrations of O_2 is required.[163]

Another possibility to determine the binding order of a protein is to use a competitive inhibitor of a substrate. If the inhibition pattern shows that the inhibitor is competitive to the first and uncompetitive to the second substrate, the binding order can be identified.

The binding order was determined in $EgtB_{thermo2}$ from *Candidatus chloracidobacterium thermophilum* with selenocysteine, which was shown to be a competitive inhibitor towards cysteine, and an uncompetitive towards TMH (Figure 68).[74] Therefore it was suggested that TMH binds first and then cysteine binds to the enzyme:TMH complex.[74] In EgtB from *Mycobacterium thermoresistibile* 2-amino N^{α},N^{α}-dimethyl histidine was used as an inhibitor.[159]

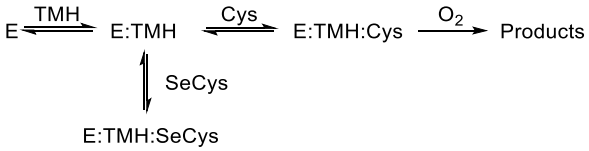


Figure 68: The binding mechanism of the bacterial EgtB.

We also attempted to measure the same kinetic data with selenocysteine as a competitive inhibitor to determine the binding mechanism of the eukaryotic enzyme. To our surprise we could observe, that selenocysteine was a poor inhibitor for the enzyme (Figure 69). Instead of inhibition, the selenoether product was formed during the reaction and it was not possible to measure the binding order (see next section).

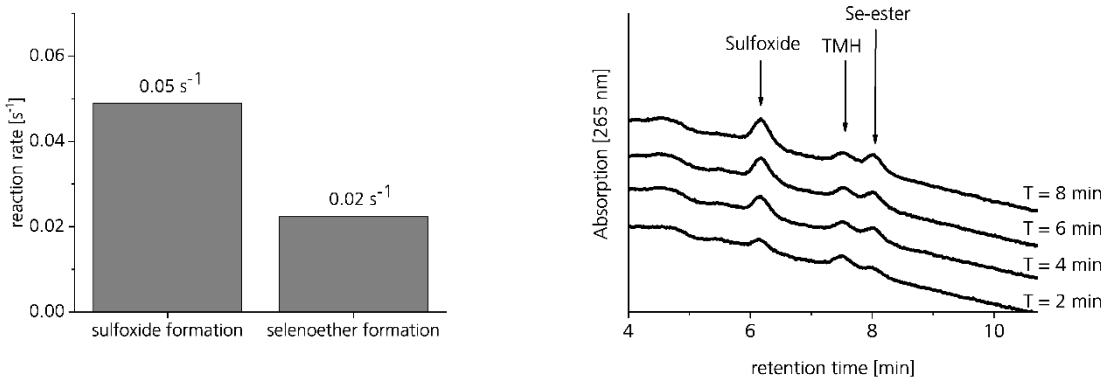


Figure 69. **Left:** The reaction rate of sulfoxide formation and selenoether formation in the same reaction mixture containing: 100 mM phosphate buffer pH 8, 100 mM NaCl, 2 mM ascorbate, 10 µM $FeSO_4$, 2 mM TCEP, 1 mM TMH, 60 µM cysteine and 200 µM selenocysteine, 0.1 µM $EgtB_{fungi\_C}$ in a total volume of 250 µL. The reaction temperature was 20°C and the samples of the reactions were quenched with 1 M phosphoric acid in a ratio 1:2. The samples were analyzed with the standard HPLC method described in the experimental section. **Right:** HPLC chromatogram of the same reaction monitored at four different time points.

We changed the inhibitor to 2-amino N^{α},N^{α}-dimethyl histidine (2A-DMH, Figure 71). This inhibitor imitates the substrate TMH. However, the amino group at the C2 position impedes the attack of the thiyl radical and therefore inhibits the overall reaction. In the reaction with the enzyme $EgtB_{thermo1}$, a low inhibitor constant K_I of 14 µM was calculated, which made 2A-DMH a suitable inhibitor.[159]

We tested 2A-DMH in a 10 fold excess towards the substrate TMH with $EgtB_{fungi\_C}$. In our measurement we could observe an inhibition of 63% (Figure 70). This indicates that this TMH analogues binds to the active site of $EgtB_{fungi\_C}$ with low affinity, and is therefore only a poor inhibitor.

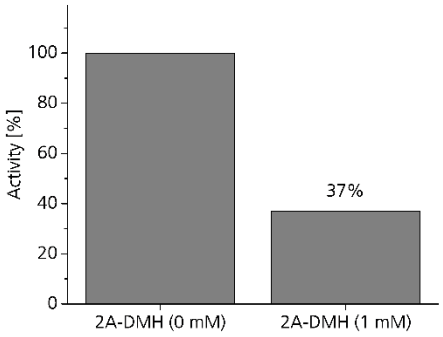


Figure 70. Inhibition reaction with and without 2A-DMH. The reaction conditions were: 100 mM phosphate buffer pH 8, 100 mM NaCl, 2 mM ascorbate, 10 µM $FeSO_4$, 2 mM TCEP, 0.1 mM TMH, 0.1 mM cysteine, 0.2 µM $EgtB_{fungi\_C}$ in a total volume of 250 µL. The inhibitor concentration was 0 mM or 1 mM and the reaction temperature was 20°C. The samples of the reactions were quenched with 1 M phosphoric acid in a ratio 1:2 and analyzed with the standard HPLC method described in the experimental section.

To identify a more suitable inhibitor with which to determine the binding order of $EgtB_{fungi\_C}$, further experiments are required. Possible inhibitors could be dimethylhistidine analogues, such as 2-methyl N^{α},N^{α}-dimethyl histidine, or 2-fluoro N^{α},N^{α}-dimethyl histidine which could be tested in future experiments.

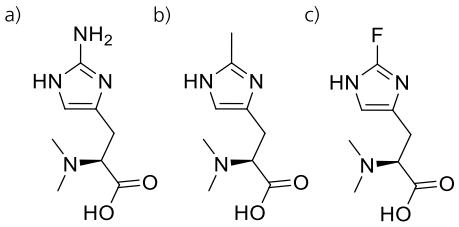
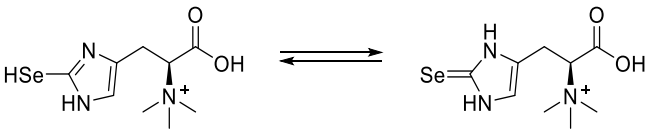


Figure 71. Possible inhibitors to determine the binding order of eukaryotic EgtB: **a)** 2-amino N^{α},N^{α}-dimethyl histidine **b)** 2-methyl N^{α},N^{α}-dimethyl histidine and **c)** 2-fluoro N^{α},N^{α}-dimethyl histidine.

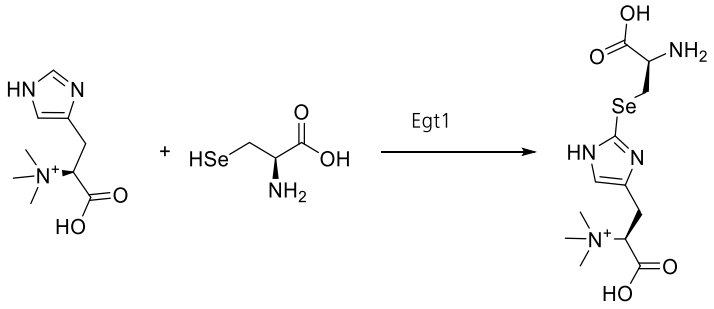
3.3.8 Selenoneine

Another natural product which was recently discovered is selenoneine, the selenium-containing analogue of ergothioneine (Scheme 25).[165] It was first detected in the blood of bluefin tuna, *Thunnus orientalis,* and was connected to mercury detoxification.[165,166] Selenoneine has been shown to have strong *in vitro* antioxidant properties.[166] Furthermore, it was discovered in human blood and in the methylated form (Se-methylselenonine) in urine.[167,168]



Scheme 25. Tautomeric structure of selenoneine with the shown selenol and selenoxo-form.[166]

In the fission yeast, *Schizosaccharomyces pombe* it was shown that the ergothioneine biosynthetic pathway of fungal species was also able to produce selenoneine *in vivo*. With the modified fungal strain with an Egt1 overexpressing system and supplementation of the growth medium sodium selenite, it was possible to observe selenoneine formation. As an intermediate, the selenoether, rather than the expected selenoxide was identified (Scheme 26).[156]



Scheme 26. Reaction of TMH and selenocysteine which results in the selenoether formation. The selenoether was identified in the fisson yeast *S. pombe*.[156]

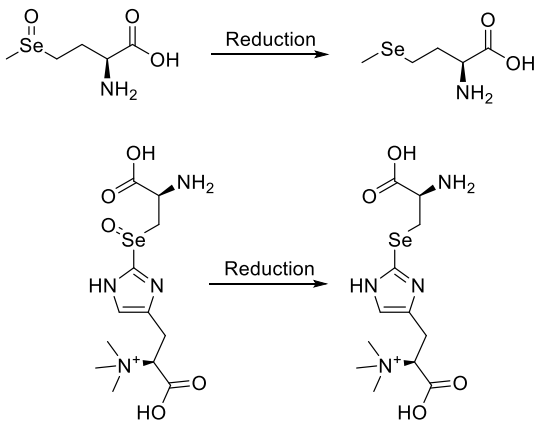
In our binding order studies with $EgtB_{fungi\_C}$, we observed that selenocysteine is a substrate, rather than an inhibitor, for the enzyme. These results are in agreement with the discovery of selenoneine formation in *S. pombe*. To verify this observation, we measured the Michaelis-Menten kinetic parameters with the same assay described previously for the sulfoxide measurements. Under saturating selenocysteine and TMH conditions, we were able to monitor the rate of selenoether formation with a k_{cat} of 0.04 s^{-1} at 20°C. This rate is only 1.7 times slower than the reaction rate measured using cysteine as a substrate. The apparent K_M of selenocysteine with a constant TMH concentration of 500 mM was calculated to be 33 µM. Under

saturating concentrations of TMH the binding affinity of selenocysteine is almost identical to the affinity of cysteine in $EgtB_{fungi\_C}$.

Table 13. Michaelis-Menten kinetic parameters of $EgtB_{fungi\_C}$ with cysteine and selenocysteine. The reaction conditions for $EgtB_{fungi\_C}$ and $EgtB_{fungi\_T}$ were 100 mM phosphate buffer and 100 mM NaCl, 2 mM ascorbate, 4 mM TCEP, 1 µM enzyme and 10 µM $FeSO_4$. The substrate concentrations were changed between 5-1000 µM and the reaction temperature was set to 20°C and 40°C.

EgtB	Temp.	Substrate	k_{cat}, [s^{-1}]	K_M [µM]	k_{cat}/K_M [$s^{-1}M^{-1}$]
C. Thermophilum	*20°C*	Cys:	0.07 ± 0.01	38 ± 3	(1.8 ± 0.2) x 10^3
C. Thermophilum	*20°C*	Se-Cys:	0.041± 0.002	33 ± 6	(1.2 ± 0.3) x 10^3

In the kinetic assay, we observed only the reduced selenoether and not the expected selenoxide product. These results are consistent with the observation of the product formed *in vivo* by Egt1 in *S. pombe* as well as *in vitro* the bacterial $EgtB_{thermo1}$ and $EgtB_{thermo2}$. Mechanistic studies with the bacterial enzymes showed no evidence for an alternative mechanism.[74] A plausible explanation is that the selenoxide is immediately reduced after the enzymatic formation and therefore only the selenoether is observed. The reduction of selenoxide has been shown, for example, in *L*-methionine selenoxide, which is reduced to seleno-*L*-methionine by methimazole, ascorbate, and endogenous thiols (Scheme 27).[169] In the kinetic assay, we have the reducing agents, ascorbate and TCEP, in high concentrations which efficiently reduce selenoxide.



Scheme 27. **Top**: *L*-Mmethioneine selenoxide is reduced to seleno-*L*-methionine with as ascorbate and endogenous thiols. **Bottom**: The presumably formed selenoxide is reduced in the reaction mixture by TCEP or ascorbate to the observed selenoether.

In the bacterial $EgtB_{thermo1}$ and $EgtB_{thermo2}$, under saturating substrate concentrations, selenoether formation was found to be 60-170 times slower than the sulfoxide formation (Table 14, Figure 72).[74] These differences

in reactivity towards selenium-containing substrates indicate a distinct catalytic difference between the bacterial enzyme types and the fungal EgtB.

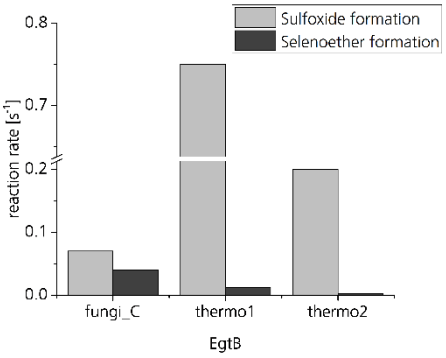


Figure 72. The reaction rates of the sulfoxide formation light gray and the selenoether formation in the enzymes $EgtB_{fungi\_C}$, $EgtB_{thermo1}$ and $EgtB_{thermo2}$.[74]

Table 14: Reaction rates of sulfoxide and selenoether formation of $EgtB_{fungi\_C}$ and two bacterial EgtBs.[74]

Sulfoxide synthase	$k_{obs,\ sulfoxide}$, [s^{-1}]	$k_{obs,\ Se\text{-}ether}$, [$s^{-1}$]	$k_{obs,\ sulfoxide} / k_{obs,\ Se\text{-}ether}$, [$s^{-1}$]
$EgtB_{fungi\_C}$	0.07	0.04	2
$EgtB_{M.\ thermoresistibile}$[74]	0.75	0.012	60
$EgtB_{C.\ thermophilum}$[74]	0.2	0.002	170

In the *in vitro* assay utilizing selenocysteine as the substrate we observed a fast decrease in enzyme activity after 10 minutes (Figure 73). Therefore, all kinetic measurements were performed within the first four minutes.

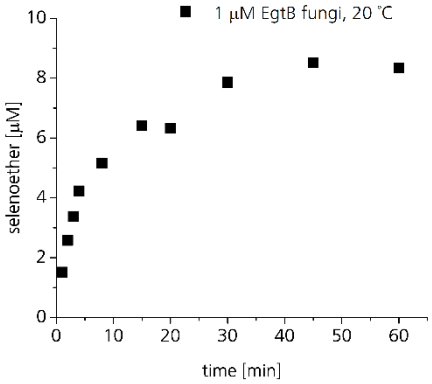


Figure 73. Selenoether formation over a time range of one hours. The condition for the assay were as follows: 100 mM phosphate buffer pH 8, 100 mM NaCl, 2 mM ascorbate, 10 µM $FeSO_4$, 10 mM TCEP, 2 mM selenocysteine, 2 mM TMH and 1 µM $EgtB_{fungi\_C}$ to a total volume of 0.5 mL. The reaction temperature was 20°C and the samples of the reactions were quenched with 1 M phosphoric acid in a ratio 1:2. The samples were analyzed with the standard HPLC method described in the experimental section.

Two possible explanations for this time-dependent loss of activity could be 1) the deactivation of the enzyme or 2 the fast consumption of oxygen in a selenocysteine-dependent side reaction, which depletes oxygen concentrations below saturating conditions. To test the later, oxygen consumption was tested in the following experiment. Two enzyme reactions were monitored over 2 hours in identical conditions. In one reaction fresh enzyme was subsequently added, while in the control reaction the same volume of pure H_2O was added. After an additional 30 minutes this was step was repeated. In both reactions we observed an increase of the initial reaction rate directly after addition. The reaction rate divided by the enzyme concentration showed identical activity for both reaction mixtures after 30 minutes and after 60 minutes. These results indicates that oxygen can be a limiting factor in the selenocysteine reaction.

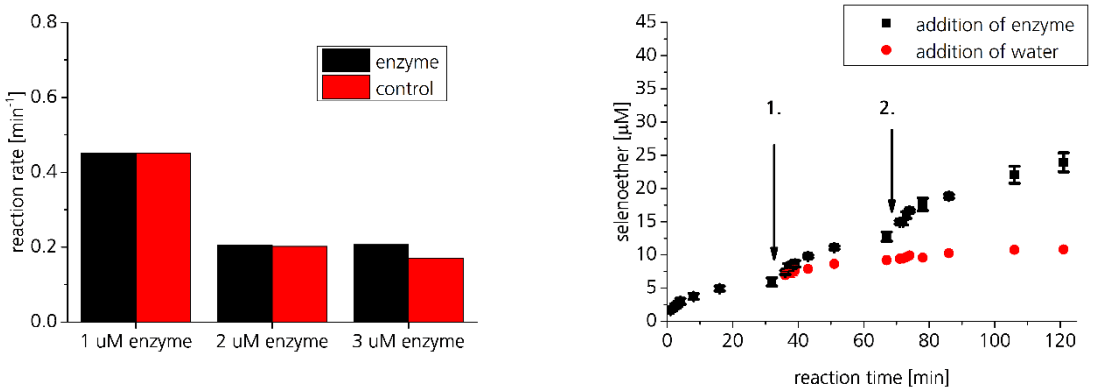


Figure 74. **Left:** Initial rate in the first four minutes after addition of enzyme or water was measured and divided by the amount of enzyme inside of the reaction mixture. **Right:** Product formation monited over 120 minutes with enzyme or water addition after 32 and 64 minutes. Reaction contained 0.5 mM selenocysteine, 1 mM TMH, 4 mM TCEP, 2 mM ascorbate, 10 uM Fe and 1 uM enzyme.

3.3.9 Oxygen Binding Site

In the eukaryotic EgtB, selenocysteine was identified as a substrate with a comparable activity to the substrate cysteine. In the bacterial enzymes the reaction rates with selenocysteine or seleno-γ-glutamylcysteine were much lower. A possible explanation for this observation could be a distinct difference in the oxygen activation between the EgtB enzymes.

In the enzyme cysteine dioxygenase, a non-heme iron enzyme the difference between selenocysteine and cysteine was investigated. In this enzyme, which has a similar active site architecture to EgtB, the oxidation of selenocysteine was not observed.[73] Multiple computational studies have been carried out to elucidate the mechanistic details, each with different outcome.[170–172] Che *et al.* proposed that the superoxo iron(III) species, which is an active oxidant in the reaction cannot be formed with the substrate selenocysteine. This is due to ability of selenium to donate electron density better to the iron atom than sulfur, which could

result in no change in the oxidation state of the iron center upon O_2 binding.[172] In contrast, Blaesi *et al.* proposed that the O_2 cannot bind when selenocysteine is bound in the active site. In their computational model, it was not possible to identify an accessible low energy intermediate for the selenocysteine pathway.[171]

We proposed that the binding pocket of the dioxygen could influence the activation of the iron(III)-superoxo specie. Therefore, we compared the crystal structure of $EgtB_{thermo1}$ with the homology model of $EgtB_{fungi\_C}$. In the ternary complex of $EgtB_{thermo1}$, we were able to identify a water molecule which was coordinated to the manganese ion in the position where the oxygen binding was expected. In $EgtB_{thermo1}$ the glutamine residue Q55 and the tyrosine Y377 are found close (distance ~ 3 Å) to this water molecule. In the homology model of $EgtB_{fungi\_C}$, the tyrosine Y839 is conserved while the glutamine is missing. Instead, a phenylalanine F440 is found. This observation would indicate that the enzymatic oxygen environment in the bacterial enzyme is much more polar than in the eukaryotic enzyme. However, the quality of the homology model is rather poor. Since the cysteine binding site of both enzymes, which is located close to the active site, is completely different, the model is not trustworthy. For a confirmation of the observed change in the polarity of the oxygen binding site, a crystal structure is required.

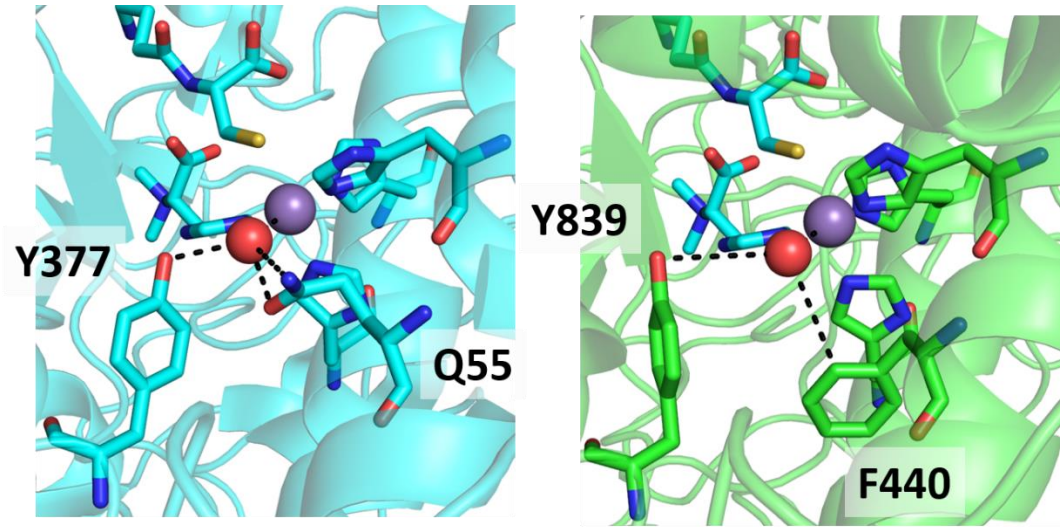


Figure 75. **Left:** crystal structure of the ternary $EgtB_{thermo1}$ crystal structure. The water molecule (red sphere) which is bond to the manganese (purple sphere) is in hydrogen bonding distance of the Q55 and Y377. **Right:** The homology model of $EgtB_{fungi\_C}$ contains Y839 and, instead of the glutamine, a phenylalanine is found in the surrounding of the water molecule.

3.3.10 Substrate Kinetic Isotope Effect

In the proposed EgtB mechanism a proton is abstracted from C2 of the imidazole ring after the formation of the C-S bond (Figure 76). By replacing the substrate TMH with C2-^{2}H-*N*$^{\alpha}$-*N*$^{\alpha}$-*N*$^{\alpha}$-trimethyl-*L*-histidine (D-TMH) we can measure the substrate kinetic isotope effect (KIE) and probe if this deprotonation step is rate limiting for the substrates cysteine and selenocysteine.

In previous studies with the bacterial $EgtB_{thermo1}$ and $EgtB_{thermo2}$ and D-TMH the measurements with cysteine showed a substrate KIE, which was close to one.[74] Interestingly, when the substrate was changed to selenocysteine, the protonation became the rate limiting step with a substrate KIE with D-TMH of 3 in $EgtB_{Thermo2}$.[74]

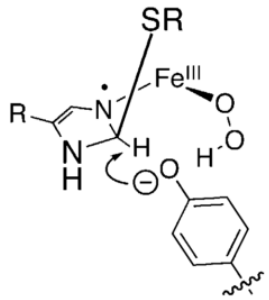


Figure 76: The protonation step in the proposed mechanism of EgtB.

In $EgtB_{fungi\_C}$, the substrate KIE with D-TMH and the substrates cysteine and selenocysteine were measured. We could observe a substrate KIE with D-TMH of 1.02 ± 0.09 for the reaction with cysteine and 1.15 ± 0.03 with selenocysteine. Both values are close to unity and shows that the deprotonation of the imidazole ring is neither the rate limiting step in the sulfoxide formation or in the selenoether formation. Especially the observation of the absence of a KIE with selenocysteine in the $EgtB_{fungi\_C}$ show that the energy profile of the eukaryotic reaction has changed towards the bacterial reaction. The C-H bond cleavage is not anymore the rate limiting step.

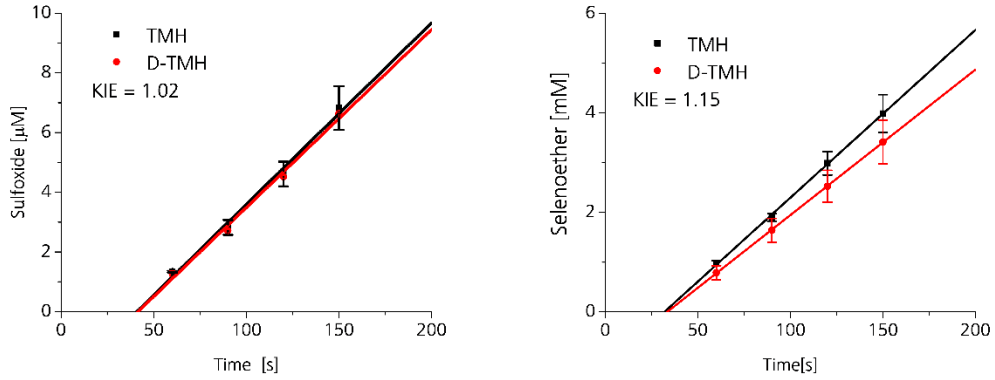


Figure 77. Left: The measured substrate isotope effect with D-TMH and cysteine measured for the sulfoxide formation. Right: The measured substrate isotope effect with D-TMH and selenocysteine measured for the selenoether formation. The condition for the assay were as follows: 100 mM phosphate buffer pH 8, 100 mM NaCl, 2 mM ascorbate, 10 µM $FeSO_4$, 2 mM TCEP for cysteine reactions, 3 mM TCEP for selenocysteine reactions, 500 µM D-TMH, 500 µM cysteine or 200 µM selenocysteine, 1 µM $EgtB_{fungi\_C}$ to a total volume of 250 µL. The reaction temperature was 20˚C and the samples of the reactions were quenched with 1 M phosphoric acid in a ratio 1:2. The samples were analyzed with the standard HPLC method described in the experimental section.

Table 15: Measured substrate KIE for $EgtB_{fungi\_C}$. compared to bacterial substrate KIE of bacterial enzymes.

Enzyme	Substrate	
$EgtB_{fungi\_C}$	Cysteine	1.02 ± 0.09
$EgtB_{fungi\_C}$	Selenocysteine	1.15 ± 0.03
$EgtB_{C.thermophilum}$[74]	Cysteine	0.9 ± 0.1
$EgtB_{C.thermophilum}$[74]	Selenocysteine	3.1 ± 0.1

3.3.11 Enzyme Stability

Recombinantly purified enzymes are constantly exposed to oxidative stress therefore a limited lifetime is expected in *in vitro* assays. Especially in oxygenases, which produce reactive oxygen species as intermediates, the probability of the enzyme damaging due to oxidation is increased. It is therefore not surprising that several iron-dependent enzymes, which activate molecular oxygen struggle with detrimental side reactions and loses their activity in a short amount of time. One example of an enzyme which has a fast deactivation rate, is the Fe^{2+}, O^{2-} and α-ketoglutarate-dependent oxygenase, clavaminate synthase.[173] By measuring radiolabeled CO_2 formation it was shown that in the clavaminate synthase from *Steptomyces clavuligerus*, catalyze a partial decarboxylation of α-ketoglutarate in the absence of the substrate.[174] These results indicate that the oxygen activation can occur without bond substrate and lead to the formation of reactive oxygen species such as peroxide or hydroxyl radicals.[174] A possible strategy to overcome this oxygen induced deactivation is to regulate the binding order of the substrates. If the oxygen binds as the last the probability

of damaging side reaction can be reduced.[175] This highly ordered oxygen binding is observed in most non-heme iron-dependent enzymes for example in the cysteine dioxygenase.[175,176] In cysteine dioxygenase this binding order was identified in EPR studies with the oxygen analogue nitrous oxide.[176]

In the non-heme iron EgtB enzymes of bacteria, a high stability was observed. In the kinetic measurements the activity was constant over hundreds of turnovers.[74] The binding order of oxygen was not measured in the bacterial EgtB, because the exact control of the oxygen concentration is challenging.[74] Nevertheless the assumption that oxygen binds only the ternary complex of the iron-TMH-γ-glutamyl cysteine complex, would support the observed stability of these enzymes.

To test the stability of the eukaryotic $EgtB_{fungi\_C}$ enzyme we monitored product formation over a time span of four hours. We observed a decrease in the enzyme activity after 50 minutes at 26°C as well as 40°C.

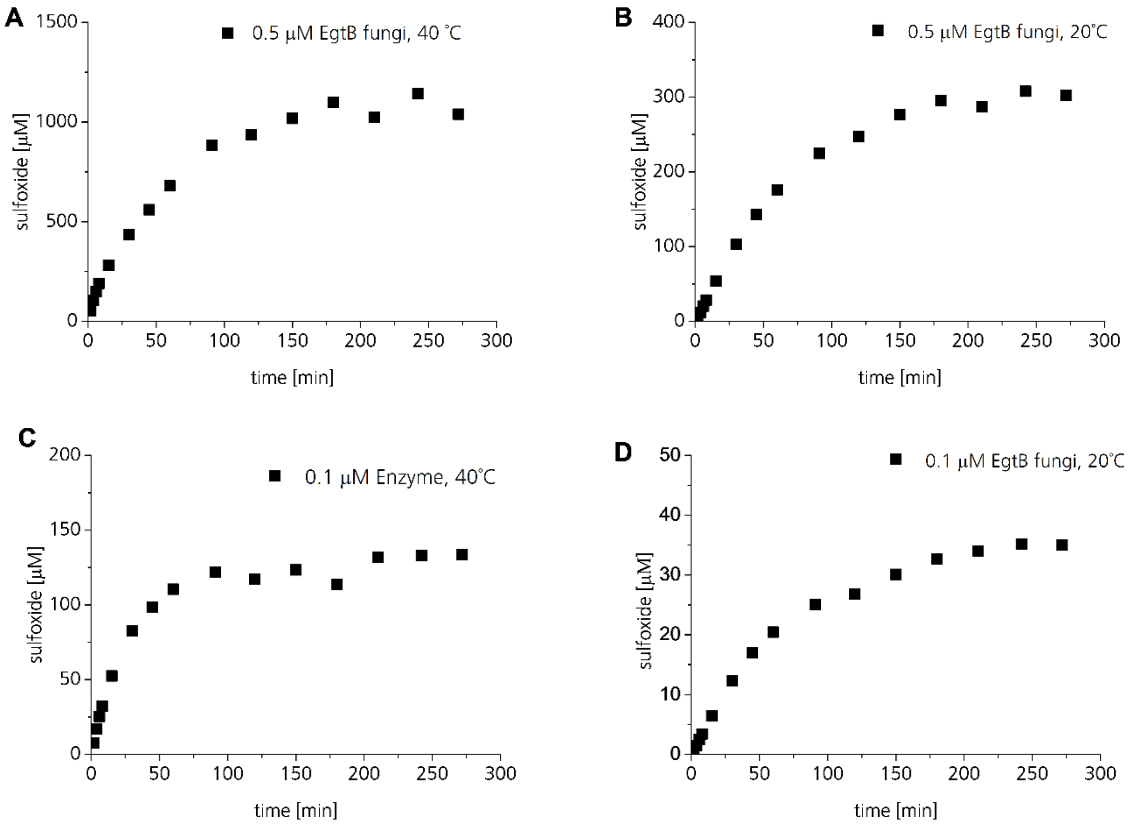


Figure 78 Sulfoxide formation over a time range of four hours. The condition for the assay were as follows: 100 mM phosphate buffer pH 8, 100 mM NaCl, 2 mM ascorbate, 2 µM $FeSO_4$, 10 mM TCEP, 2 mM cysteine, 2 mM TMH and 0.5 µM (**A-B**) or 0.1 µM (**C-D**) $EgtB_{fungi\_C}$ to a total volume of 0.5 mL. The reaction temperature was 40°C (**A, C**) or 20°C (**B, D**) and the samples of the reactions were quenched with 1 M phosphoric acid in a ratio 1:2. The samples were analyzed with the standard HPLC method described in the experimental section.

In the next experiment, we tested if product inhibition could be the reason for this observation. To the enzyme reaction, 0 mM, 100 mM, 200 mM and 400 mM of the product was added and the reaction rate was monitored over the first 10 minutes. We observed no difference in the activity and therefore exclude the deactivation due to product inhibition.

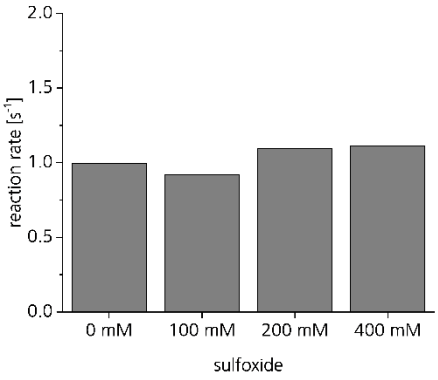


Figure 79. Product inhibition measurements. The reaction rates were measured with 0 mM, 100 mM, 200 mM and 400 mM premixed sulfoxide product.

Another possible explanation for the deactivation of the enzyme could be the oxidation of catalytic relevant residues during catalysis. This would lead to an oxidation signature inside of the enzyme, which might be observable with mass spectroscopy (MS) analysis. With a protein mass of more than 100 kDa, it was not possible to measure a full length HRMS of the enzyme. We therefore decided to analyze the enzyme fragments after tryptic digestion.

The enzyme reaction was monitored until the complete loss of $EgtB_{fung\_C}$ activity was observed. Afterwards the enzyme was digested with trypsin and analyzed by low resolution LC-MS. As a control experiment the enzyme was digested without any catalytic turnovers. With the low resolution MS, we were able to identify 51 of 55 (95%) of the protein fragments. However, it was not possible to identify a distinct difference between the deactivated enzyme and the control.

For a more precise analysis we repeated the experiment in the proteomics core facility at the Biozentrum of the University of Basel with HRMS fragment analysis. With the HRMS analysis a sequence coverage of 47-60% was obtained. Nevertheless, this experiment also did not show a statistically significant difference between the deactivated sample and the control. Since the sequence coverage was not 100%, it cannot be excluded that a possible distinct modification was present in a fragment which was not measured.

In HRMS analysis, it is common to observe methionine residues as the most oxidized amino acid (+16 Da mass), due to the sulfur-containing side chain which is susceptible to oxidation.[177] Interestingly, in our experiments we observed different results in both the catalytic deactivated enzyme and the control enzyme. The most oxidized amino acid was in either case tyrosine followed by phenylalanine (Figure 80). This oxidation pattern indicatef that the enzyme $EgtB_{fungi\_C}$ is in general exposed to a higher amount of reactive oxygen species. A direct connection of an oxidation pattern towards the catalysis was not observable.

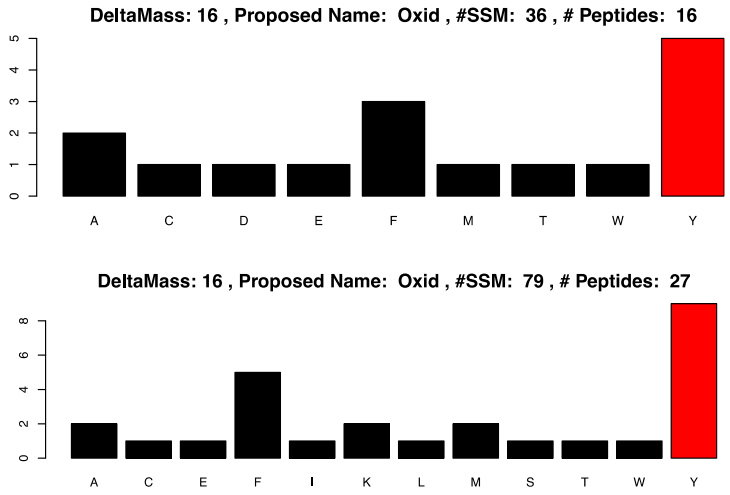


Figure 80. **Top**: Distribution of oxidation found in the peptide fragment of the deactivated enzyme. **Bottom**: Distribution of oxidation found in the peptide fragment of the control enzyme.

3.3.12 The EgtD-Domain

A main difference between the fungal EgtB enzymes and the bacterial EgtB enzymes is the protein size. The eukaryotic enzymes contain an additional fused domain next to the EgtB domain, which is a homologue to the EgtD methyltransferase found in *M. smegmatis*.

The SAM-dependent methyltransferase EgtD, which was characterized in *Mycobacteria smegmatis*, catalyzes the first step of the biosynthesis of ergothioneine by forming TMH from histidine.[21] The crystal structure of EgtD, co-crystalized with the intermediate dimethylhistidine and *S*-adenosyl-*L*-homocysteine (PDB: 4pio, 1.8 Å), revealed the amino acids which are involved in histidine binding.[21] A network of seven hydrogen bonds between the enzyme and dimethylhistidine were identified. The carboxylate group of histidine interacts with the side chains of Tyr206, Ser284, Tyr56 and Asn166. Furthermore, the amino acids Phe47 and Tyr39 defines the binding pocket of the imidazole ring.

A sequence alignment between the bacterial and the fungal EgtD domain revealed the presence, and conservation of, all amino acids which are connected to the binding of histidine. In the fungal enzyme the conserved residues for the histidine binding were identified as the amino acid residues Tyr77, Phe85, Tyr94, Asn208, Tyr248 and Ser340 (Figure 81).

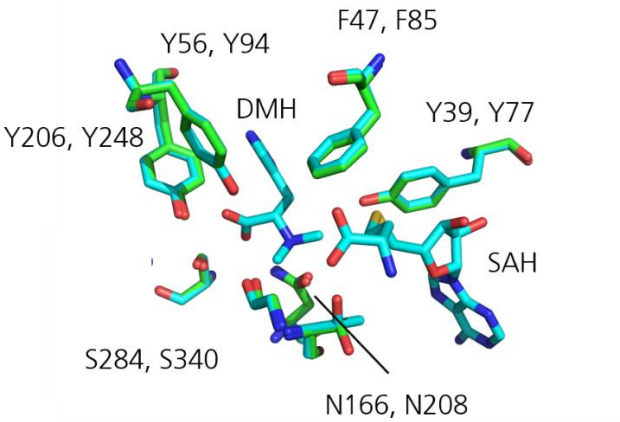


Figure 81. Supperposition of the EgtD crystal structure (cyan, pdb: 4pio, 1.8 Å) and the homology model of $EgtB_{fungi\_C}$ (green) shows no significant difference of the catalytic relevant residues.[21] The first number of the labeled amino acids corresponds to the residue position in bacterial EgtD, the second number to the fungal EgtD.

In the characterization of Egt1 from *N. crassa*, the methyltransferase activity was not measured. Furthermore, the reported sequence of the Egt1 examined by Hu *et al.* is a truncated version of Eg1. The sequence displayed in the supporting information is different to the sequence found for wild type enzyme.[153] The amino acid between the position Gly201 and Asn243 are not present in the reported sequence.[153] This truncated version of Egt1 misses central amino acids which are relevant for the methyltransferase activity and questions the overall reliability of the examined enzyme.

We were able to successfully measure for the first time the methyltransferase activity of a eukaryotic EgtD/EgtB fusion protein. With the previously published HPLC-kinetic assay a maximal histidine methylation rate of 2.2 min^{-1} at 40 °C was measured by Laëtitia Misson.[21] This value is 17 times slower than the reaction rate measured for EgtD in *Mycobacteria smegmatis* (34.2 min^{-1}) at room temperature.[21]

3.3.13 Protein Crystallization Attempts

To elucidate the differences in the architecture of eukaryotic EgtB and EgtB from *M. thermoresistibile*, we attempted to crystallize $EgtB_{fungi\_C}$. To reduce the flexibility of the protein and support the natural crystal packing we removed the His_6-tag with TEV protease. The protein was then purified with size exclusion chromatography to improve the purity for crystallization. In a first crystallization round, multiple

commercially available crystal screens were used, however we were only successful in obtaining protein crystals using one of these single conditions (G8 of Morpheus® HT screen,.37.5 %w/v M1K3350, 0.1 M MB2 7.5 pH, 10% MCA).

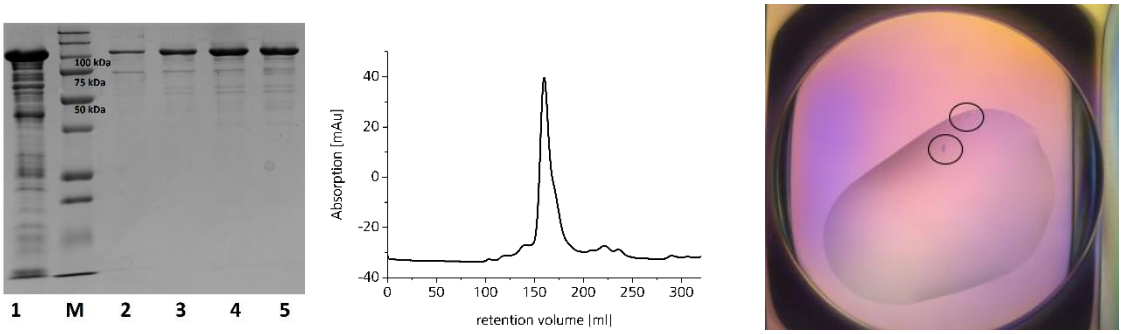


Figure 82. **Left:** SDS page with a standard marker (M) of EgtB fungi before (line 1) and after (line 2-5) FPLC purification. **Middle**: Size exclusion chromatogram of $EgtB_{fungi\_C}$. **Right**: First observed crystals in the Morpheus® HT crystal screen (position G8).

These conditions were used as an initial point for an optimized screen. The optimized screen contained different precipitant concentration (30-40% M1K3350), and different complex ingredients at two different pH values. With the optimized screen, we were able to obtain crystals under multiple conditions. During the harvesting of the crystals, we realized that the crystals were located under a polymerlike sticky layer. In addition, the crystals were very fragile and easily broke apart. The few crystals we were able to harvest diffracted only in a low resolution of >10 Å and therefore no data was collected.

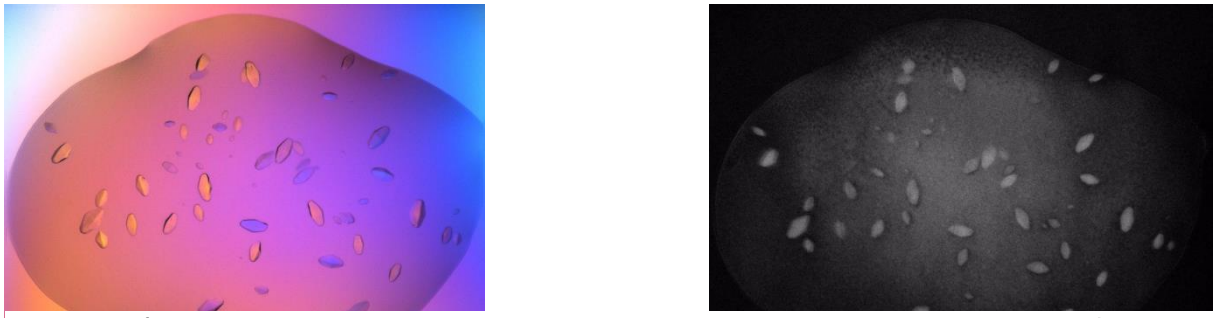


Figure 83. **Left:** Image of small EgtBfungi_C crystals observed in a drop in the optimized crystal screen. **Right:** Image of the same drop under 280 nm UV-light. Protein absorption is visible.

The polymerlike layer in the crystal drops and the fragile crystals could be caused by a polymerization reaction of the protein. One possibility how protein a can form polymers is the formation of disulfide bridges between surface cysteines. Therefore, we located the cysteine residues which were exposed to the surface of the protein according to the homology model made by EgtB from *M. thermoresistibile* and EgtD from *M. smegmatis*. The enzyme was found to contain eight cysteine residues. We were able to identify the

cysteine residues Cys238, Cys365, Cys477 and C754 in the homology models which were exposed to the surface of the protein and could form disulfide bridges.

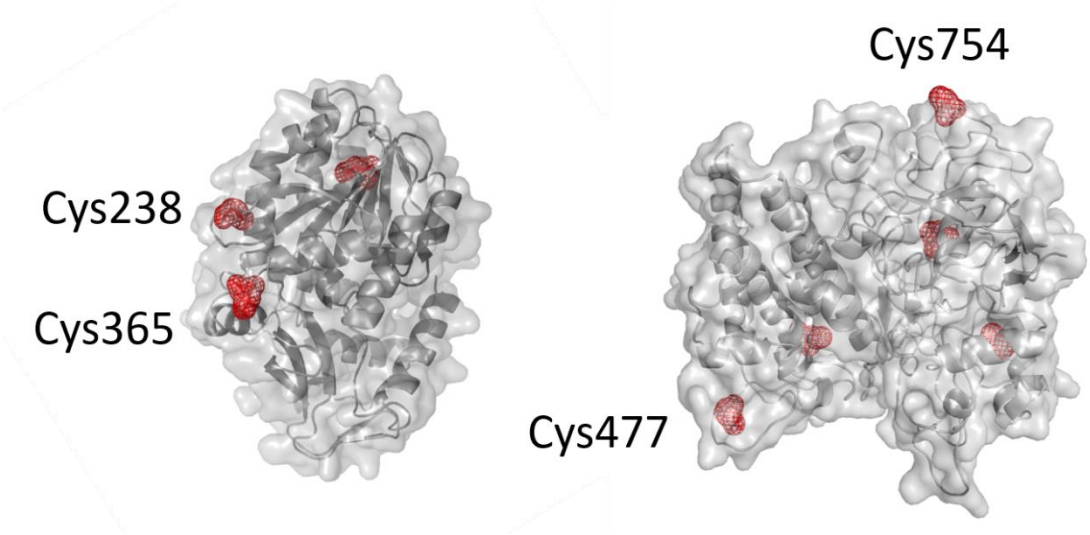
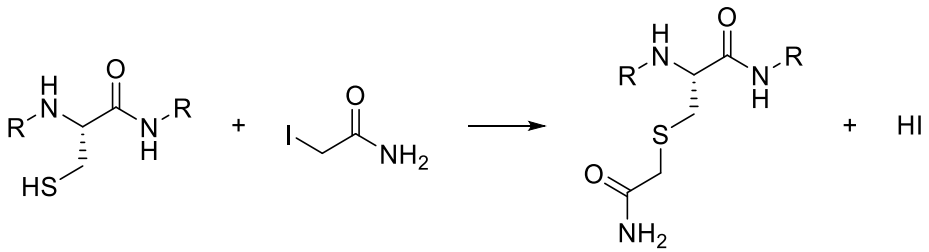


Figure 84. Cysteine residues are highlighted in red in the EgtD homology model (**left**) and in the EgtB homology model (**right**). The surface cysteines are labeled by the amino acid sequence position.

To deactivate these surface cysteines, we masked them by alkylation with iodoacetamide (Scheme 28). The protein was first reduced with TCEP and then iodoacetamide was added. Since the protein has a size of 100 kDa, we were not able to measure HRMS and therefore we were not able to monitor the reaction. However in the size exclusion purification, a high protein aggregation due to the chemical treatment could be observed. The purified enzyme was then used for different crystal screens, nevertheless we could not observe anymore crystal formation with the modified enzyme.



Scheme 28. Reaction of iodoacetamide and surface cysteines to form the deactivated alkylcysteines.

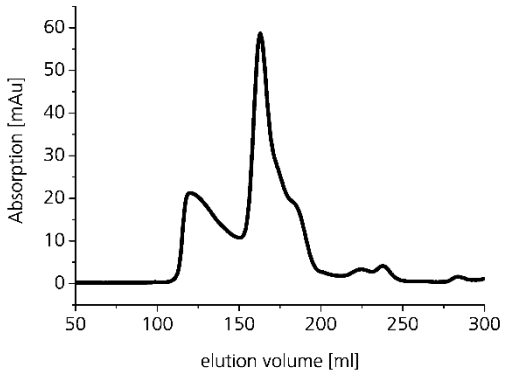


Figure 85. FPLC-purification of $EgtB_{fungi\_C}$ after iodoacetamide treatment.

In the next step we exchanged the enzyme $EgtB_{fungi\_C}$ to the homologue $EgtB_{fungi\_T}$. This new enzyme was then used in new crystallization attempts. Even though this enzyme comes from another organism and contains less surface cysteines we could only observe very fragile crystals and were not able to isolate them without destruction of the crystals.

3.4 Results and Discussion fCDO

3.4.1 Discovery of two EgtB Homologues in D. squalens

A BLAST search of the genome of the basidiomycete *Dichomitus Squalens*, revealed two different EgtB homologues. The sequence alignment showed that one of the homologue is a EgtB/EgtD fusion protein and related to $EgtB_{Fungi\_C}$ (35% sequence identity). The other EgtB homologue (33 % sequence identity to $EgtB_{fungi\_C}$) is much shorter in size and contains only the EgtB domain. In addition, the catalytic acid tyrosine had been replaced by a phenylalanine. This residue (Tyr377) was identified in the EgtB of *M. thermoresitibile* as crucial to protonate the superoxo iron(III) intermediate.[60,74,83] The $EgtB_{thermo1}$ mutant, Y377F, almost lost complete sulfoxide synthase activity. Instead, a cysteine dioxygenase reaction was observed. It was possible to show that a single point mutation of this residue could convert the sulfoxide synthase into a cysteine dioxygenase.[83] We therefore speculated that this second EgtB is a naturally evolved fungal cysteine dioxygenase (fCDO), instead of a sulfoxide synthase. The sequence alignment between the EgtBs of *M. thermoresistibile*, fungal *C. thermophilum* and the two EgtB versions found in *D. squalens* is displayed in Figure 86.

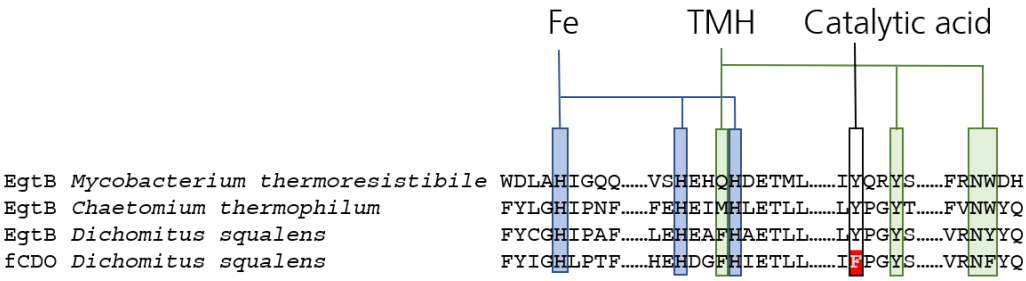


Figure 86: Sequence alignment between EgtB of *Mycobacteria thermoresistibile*, EgtB of the fungal *Chaetomium thermophilum* and the two EgtB versions found in *D. squalens*. The iron binding residues, the TMH binding residues and the catalytic acid residue are highlighted. The phenylalanine residue F442s in fCDO of *D. squalens* is highlighted in red.

3.4.2 Homology Model of fCDO

To illustrate the architecture of the fCDO, a three-dimensional structure was modeled by Swiss model (https://swissmodel.expasy.org) with the EgtB structure of *Mycobacteria thermoresistibile* (PDB: 4x8d) as a template.

Three histidine residues H100, H200, H204 were identified as the conserved iron binding residues. The residues Y445, F477, N476 and F203 were proposed to be involved in the TMH binding of fCDO. While Y445 and N476 are conserved in the TMH binding site of $EgtB_{thermo1}$, F477 and F203 are different. The replacement of a tryptophan for phenylalanine should have a minor effect for the architecture of the binding site. The change of Q138 to F203 is more unusual. Nevertheless, this single difference should not change the overall binding of TMH. The cysteine binding is different towards the bacterial enzyme and closely related to the fungal enzyme discussed before. Most interestingly the Y377 in the bacterial enzyme is replaced by a phenylalanine F442 (Figure 87).

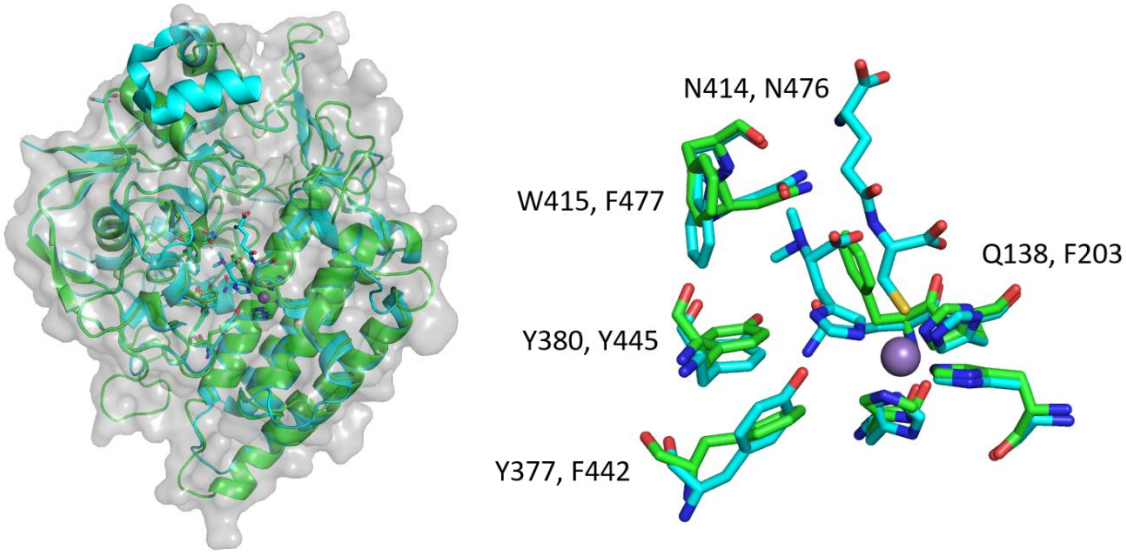
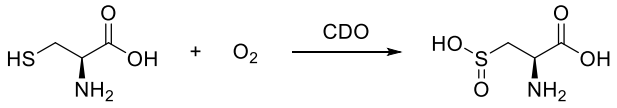


Figure 87. **Left:** Superposition of homology model $EgtB_{M.\ thermoresistibile}$ (PDB: 4x8d, 1.98 Å), with the homology model for fCDO. **Right**: superposition of the residues involed in the TMH binding. The cytalytic acid Y377 of the bacterial enzyme is replaced by F442 in the fCDO.

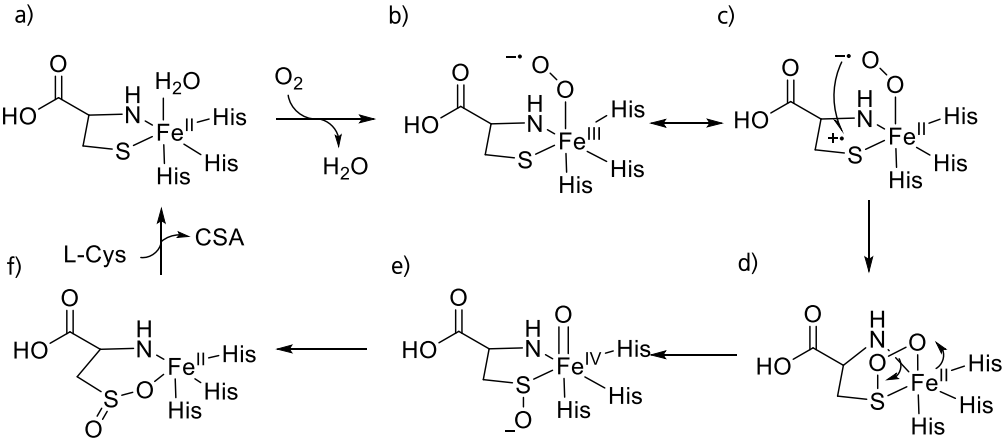
3.4.3 Proposed Mechanism of fCDOs

Cysteine dioxygenase (CDO) are mononuclear non-heme enzymes which convert cysteine to cysteine sulfinic acid (CSA) via the addition of molecular oxygen (Scheme 29). CDOs are found in mammalian cells, fungi and some bacteria.[1,178,179] One of the main function of these enzymes are to regulate the level of the cytotoxic amino acid cysteine. Furthermore they are involved in taurine biosynthesis and in sulfate production.[180] The first crystal structure of a mouse CDO was solved in 2006 and revealed a characteristic β-barrel fold of the cupin superfamily.[179] The active site revealed a three-histidine facial triad as an iron ligand.[181] Interestingly, in eukaryotes, a cross-link between a cysteine and a tyrosine in the active site was observed.[179] In enzymes found in prokaryotes, this crosslink is absent.[157]



Scheme 29. Cysteine 4 e^- oxidation catalyzed by CDO to form CSA.

The proposed mechanism starts with the binding of cysteine (**a**) followed by the second substrate dioxygen (**b**). The resulting superoxo iron(III) species can attack the radical cation resonance form of the thiolate (**c**) forming a four-membered ring between the sulfur iron and the two oxygen (**d**). Heterolytic cleavage of the dioxygen bond results in a Fe(IV)-oxo species (**e**). In the last step, an oxygen atom transfer results in the final product CSA (Scheme *30*).[179,182]

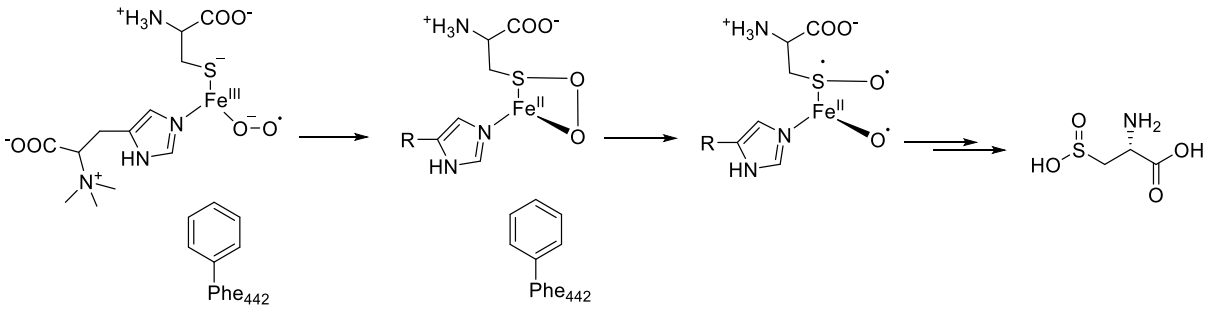


Scheme 30. Proposed reaction mechanism catalyzed by CDO to form CSA form cysteine.[179]

The new discovered fCDO in *D. squalens*, is a EgtB homologue and therefore an EgtB structure containing a DinB domain and FGE-like domain, as seen in $EgtB_{thermo1}$, was expected. This structure is evolutionary

unrelated to the β-barrel fold of the cupin superfamily found in CDO. However, the analysis of the active site architecture of an available crystal structures revealed significant similarities between CDO and EgtB. Both enzyme classes have a three-histidine facial triad and bind iron in the ferrous state. Furthermore, both enzyme classes form an octahedral complex with the bound substrates and, in the proposed mechanisms of both enzyme a cysteine bound superoxo iron(III) species is involved.

Based on the similarities of the active site architecture between CDO and EgtBs we proposed a similar catalytic mechanism for fCDO, as proposed for CDO. After the binding of cysteine and oxygen, and the binding of the activator TMH, the resulting superoxide can attack the sulfur atom to form an irreversible S-O bond. Homology O-O bond scission followed by radical rebound results in CSA (Figure 68).



Scheme 31. Proposed reaction mechanism of fCDO from *D. squalens*.

3.4.4 Cysteine Dioxygenase Activity

The genes of the enzyme fCDO and EgtB/EgtD fusion proteins from the organism *D. squalens* were ordered from GenScript® and cloned into a pET28 production vector of *E. coli*. The enzyme fCDO was produced in a poor yield of 1 mg protein per liter of culture using the plasmid pL1SL2 containing the gene for an *E. coli* chaperone. In the first experiment, the cysteine dioxygenase activity was established by ^{1}H-NMR. In the reaction with the two substrates, cysteine and TMH, ascorbate, iron and TCEP, we could observe the exclusively formation of cysteine sulfinic acid (CSA). No sulfoxide formation could be observed in the ^{1}H-NMR. As a control the second EgtB homologue of *D. squalens* $EgtB_{fungi\_D}$, which is a EgtB/EgtD fusion enzyme, was produced and the product distribution was measured via ^{1}H-NMR (Figure 88). In $EgtB_{fungi\_D}$ the sulfoxide formation was the major product (80%). These results supported our assumption, that this EgtB variant of *D. squalens* is indeed a cysteine dioxygenase.

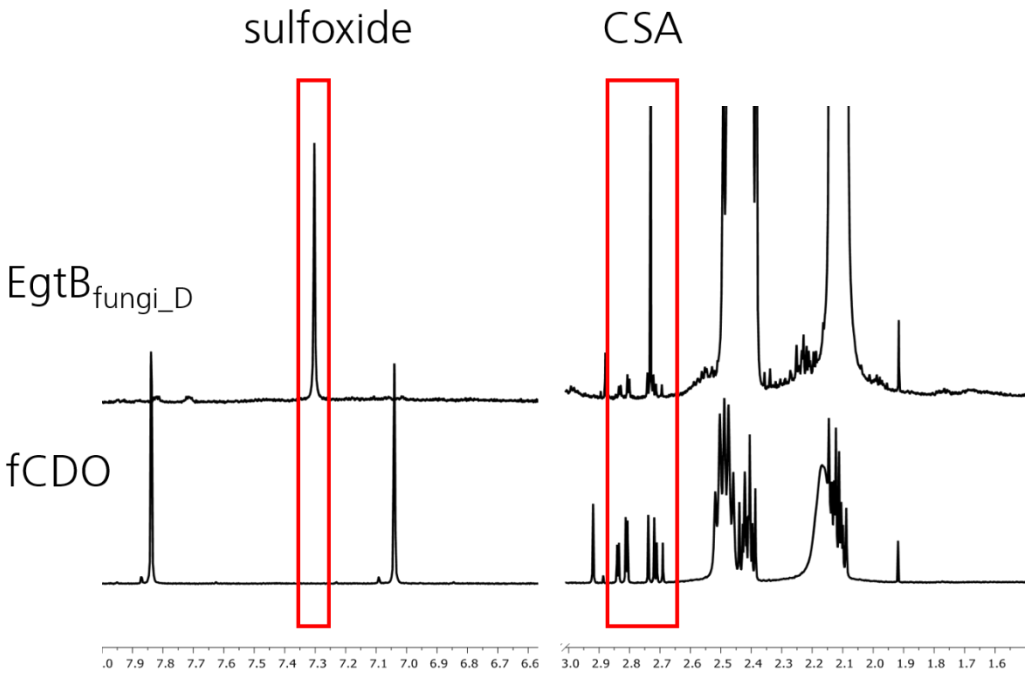


Figure 88. ^{1}H-NMR spectrum of the reaction of $EgtB_{fungi\_D}$ and fCDO with TMH and cysteine shows the formation of CSA and only in the enzyme $EgtB_{fungi\_D}$ the sulfoxide product.

For a closer kinetic characterization, we designed a project suitable for a Master thesis. Under my supervision, Nico Valerio Igareta performed the kinetic characterization, substrate specificity studies, molecular cloning and activator kinetic studies (unpublished data). Furthermore, we started crystallization attempts in collaboration with Florian Leisinger.

3.4.5 Evolutionary Analysis

The enzyme fCDO was found in *D. squalens,* which belongs to the white rot fungus. This fungal species is capable of degrading lignin, as well as other wood components, and is grouped into the division of basidiomycetes. After the discovery of the *D. squalens* cysteine dioxygenase variant, which is related to EgtB, we started to look for other fCDO homologues.

We collected the eukaryotic EgtBs sequences from the published sequence similarity network (SSN) of OvoAs and EgtBs.[155] Each sequence represented a collection of sequences which had 75% or higher similarity to each other. With those available sequences, a phylogenetic tree was created. The two EgtB homologue sequences of *D. squalens*, EgtB and fCDO were added to the phylogenetic tree to highlight their position (Figure 89).

Most of the identified sequences belong to the *Dikarya* a sub-kingdom of fungi containing the "higher fungi" divisions, *Ascomycota* and *Basidiomycota*. While the fusion protein EgtB/EgtD was found in both divisions, the identified fCDO homologues were only detected in the basidiomycetes.

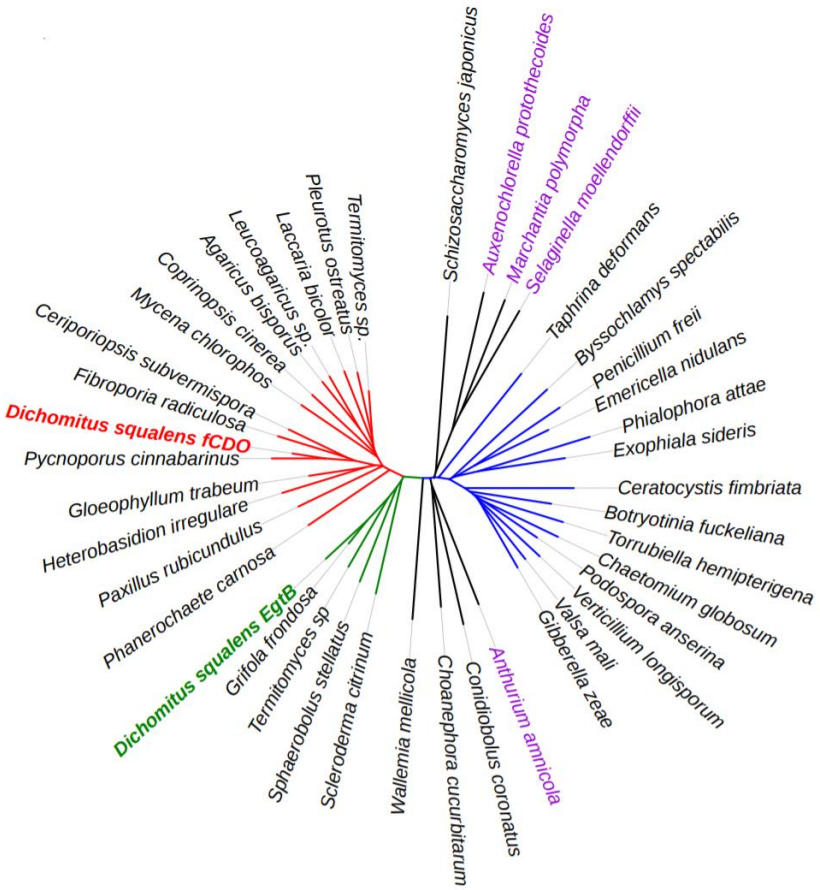


Figure 89: Phylogenetic tree of the sequences from the eukaryotic EgtB's from SSN. The EgtB from Ascomycota organisms are highlighted in blue. The EgtB from *basidiomycota* are highlighted in green and the fCDO from *basidiomycota* (EgtB homologue where the catalytic acid tyrosine is replaced by a phenylalanine) are highlighted in red. The identified plant organisms are highlighted in purple. The other sequences belong to fungi outside of the subkingdom Dikarya.

Next to the fungal EgtBs were sequences of EgtB which belong to organisms from the kingdom of Plantea. This is remarkable since there are no reported examples of ergothioneine producing plants in the literature. The sequence of these plant EgtB homologues contained the conserved catalytically relevant amino acids found in fungi enzyme. Therefore it can be assumed that these plant species are able to produce ergothioneine. In our phylogenetic tree EgtB homologues were found in the green algae *Auxenochlorella prothecoides*, in the liverwort *Marchantia polymorpha*, and in the lycophyte *Selaginella moellendorffii*. Another plant which contained an EgtB homologue was *Anthurium amnicola,* also known as the flamingo flower. Nevertheless, the data about this sequence of *Anthurium amnicola* are limited and the closest related sequence (77% sequence identity) belonged to the fungi *Rhizophagus irregularis* from the division

of *glomeromycota* a known symbiotic fungi living close to the roots of host plants.[183] Therefore, it could not be excluded that this sequence belongs to another mycorrhizal species, which was accidentally sequenced rather than to *Anthurium amnicola*. Because plants genomes have rather long genome sizes, sequencing is expensive and therefore the amount of solved plant genomes are limited.[184,185] It is likely that many more plants species are able to produce ergothioneine.

Another question revolves around the evolution of the eukaryotic EgtBs, wheter they have a common ancestor or, if in some organisms, the genes were taken up from bacteria by horizontal gene transfer. The main difference we could identify between the bacterial EgtB and the fungal EgtB is the sequence of the cysteine binding site. Therefore, this sequence, which is located between the second and the third α-helix of the DinB domain was analyzed (Figure 90). In all sequences from the SSN, we could identify the conserved cysteine binding which is present in the fungal EgtBs. Especially the arginine residue (Arg465 in $EgtB_{fungi\_C}$), which presumably is involved in the binding of the carboxylate group of cysteine, is completely conserved in all structures. This finding would suggest that the fungal, as well as the plant, have a common eukaryotic ancestor which contained this specific cysteine binding motif. This ancestor contained the EgtB gene presumably before the divergence between fungi and plant, which happened more than one billion years ago.[186]

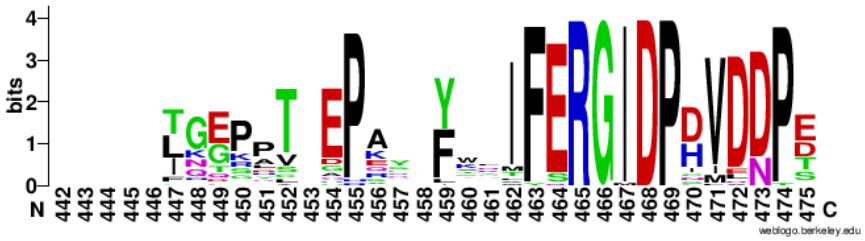
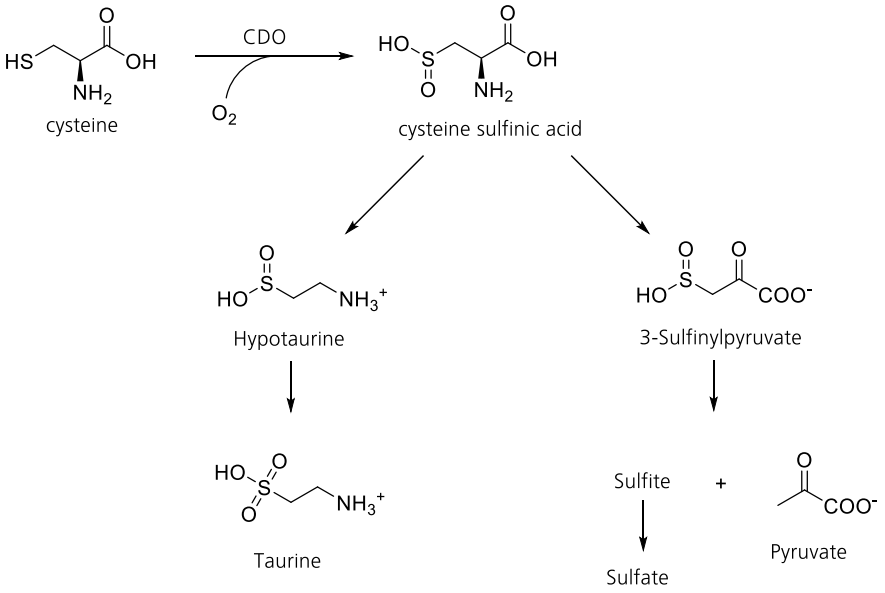


Figure 90. The highly conserved cysteine binding sequence found in the eukaryotic EgtB homologues taken from the SSN. The arginine (Arg 465 in $EgtB_{fungi\_C}$) is conserved in all organisms. The weblogo was generated with the website http://weblogo.berkeley.edu/logo.cgi.

3.4.6 Possible Role of Cysteine Sulfinic Acid

In the literature cysteine dioxygenases, which are found in bacteria, fungi and metazoan, are usually described as efficient regulator for cellular cysteine levels.[157,178,180,181,187] After the addition of molecular oxygen to cysteine, CSA is further used for taurine biosynthesis or the formation of sulfate (Scheme 32).[1]

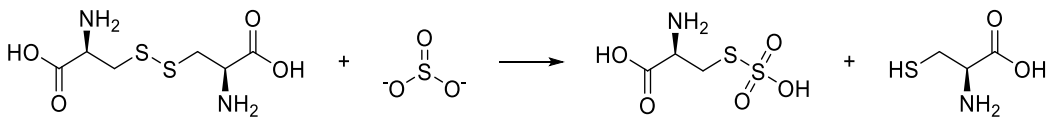


Scheme 32: Dioxygenation of cysteine by CDOs and the two degrading pathways for CSA. CSA can be transformed to hypotaurine by a cysteine sulfonate decarboxylase, which further is processed to taurine. In the alternative pathway CSA is metabolized to 3-sulfinicpyruvate by aspartate aminotransferase, which decomposes to sulfite and pyruvate.[157,181]

The cysteine dioxygenase from *D. squalens* differs not only in the amino acid sequence and structure from previously found CDOs, but also requires an activator for efficient catalysis. This activator is TMH, which means that this enzyme is only active if a SAM-dependent histidine methyltransferase is available in parallel. This would make a connection between ergothioneine biosynthesis and cysteine sulfinic acid.

Recently another function of CSA was discovered, CSA involved in the degradation of keratin by dermathophytes.[187]

In 1972, Kunert showed that sulfite is able to irreversibly reduce cysteine disulfide bridges by forming *S*-sulfocysteine (Scheme 33).[188] Furthermore, a sulfite efflux pump Ssu1 was found in Baker's yeast.[189] The discovery of the gene for both the cysteine dioxygenase Cdo1 and the sulfite efflux pump Ssu1 in dermatophytes, led to a new model for keratin degradation.[187] In 2013, it was shown in a knock-out study in the fungus *Arthroderma benhamiae,* that both enzymes, Ssu1 and Cdo1 are required for successful growth of keratin-containing tissue, such as hair or nails.[187]



Scheme 33. The reduction reaction of sulfite towards cystine disulfide bonds.

Hard keratin is composed of peptide chains which are tightly linked and contain multiple cysteine disulfide bridges. These disulfide bridges make the material completely water insoluble and form brick-like layers. For an efficient degradation of keratin, the disulfide bridges need to be reduced first. In the dermatophyte *A. benhamiae*, sulfite which is formed over the intermediate CSA is secreted out of the cell by the sulfite efflux pump Ssu1. The sulfite then reduces the disulfide bridges of keratin and make the single polypeptide chain more accessible for the proteases which are secreted as well. The proteolytically digested amino acids are then taken up by the organism. The cysteines which are taken up can further be oxidized by CDO to form new sulfite molecules (Figure 91).

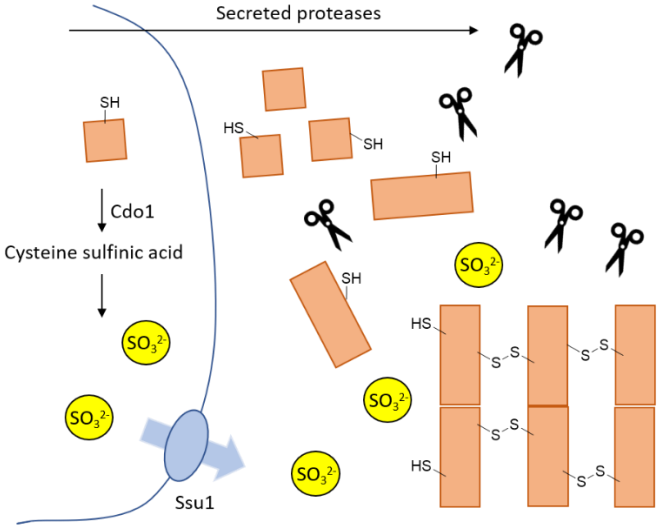


Figure 91. Keratin degradation system in *A. benhamiae*. Cysteine is oxidized by Cdo1 to CSA, which is further metabolized to sulfite. Sulfite is then segregated by Ssu1 to reduce the disulfide bridges of extracellular keratin. Keratin can therefore better by digested by secreted protease. The digested amino acids are then taken up by the fungi.[187]

A possible function of the fCDO could be the formation of sulfite as seen in the dermatophyte *A. benhamiae*. Especially for the wood degrading fungus, *D. squalens,* an extracellular reducing agent could be an advantage to support proteolytic digestion of nutrients which are difficult to degrade. To find more evidence for this hypothesis, the genome of *D. squalens* was screened for either an alternative cysteine dioxygenase Cdo1 homologue and/or for a sulfite efflux pump Ssu1 homologue. It was not possible to find a Cdo1 (*A. benhamiae)* homologue in the organism *D. squalens* (taxid:114155) by a BLAST search on the NCBI website, however, it was possible to identify three Ssu1-related enzymes with a similar size and sequence identities of 30-33%. The observation that the organism *D. squalens* has no other characterized cysteine dioxygenase enzyme and contains sulfite efflux related enzymes could support a similar sulfite mechanism as seen in *A. benhamiae*. However, the alternative function of fCDO to regulate the cysteine concentration inside of the cell is also likely.

Since the fCDO was only observed in basidiomycetes, we searched for other cysteine dioxygenases in this division. By a BLAST search with the sequence of Cdo1 from *A. benhamiae*, it was only possible to identify 14 sequences within the whole division of basidiomycetes. This result of a scarce distribution of cysteine dioxygenase could explain the evolutionary pressure for the emergence of an alternative cysteine dioxygenase enzyme in basidiomycetes.

3.5 Conclusion

It was possible to produce and kinetically characterize active forms of the fungal EgtB/EgtD fusion protein $EgtB_{fungi\_C}$ and $EgtB_{fungi\_T}$. The *in vitro* sulfoxide synthase as well as SAM-dependent methyltransferase activity were established. A different cysteine binding mode was identified with sequence alignment and a homology model with the published EgtB structure.[60] The sulfoxide formation reaction rate of the enzyme was slightly reduced compared to bacterial enzymes.

Efficient selenoether formation was observed with the alternative substrate selenocysteine using the EgtB/EgtD fusion enzymes of eukaryotes. This observation stands in contrast to the slow reaction of selenocysteine with bacterial EgtBs. This result might be due to a more efficient oxygen binding or increased oxygen activation abilities of the fungal enzymes. In addition, the measured substrate KIE for cysteine, as well as selenocysteine, with D-TMH was close to unity. In bacterial EgtBs the hydrogen abstraction was the rate limiting step with selenium containing substrates. These results indicate that in the fungal enzyme the energy profile of the reaction with selenocysteine has changed.

The $EgtB_{fungi\_C}$ enzyme showed self-deactivation properties during catalysis, which was not observed in bacterial enzymes. Identified unspecific oxidations could point to an increased formation of uncontrolled reactive oxygen species inside of the EgtB active site.

It was possible to find conditions to obtain protein crystals of $EgtB_{fungi\_C}$., however, it was not possible to collect diffraction data with a resolution below 10 Å. Also, the crystallization attempts with the homologue $EgtB_{fungi\_T}$ failed.

In the second part of this chapter, a new fungal cysteine dioxygenase (fCDO) was identified in the organism *D. squalens*. The *in vitro* reaction was observable with ^{1}H-NMR. This result highlights the mechanistic relationship between the two non-heme iron-dependent enzymes EgtB and CDO. The fCDO is related to the eukaryotic EgtB/EgtD fusion protein and only found in the division of basidiomycetes. The scarce distribution of CDO in basidiomycetes could be a reason for the evolution of this alternative CDO. The requirement of the activator TMH connects this enzyme to the ergothioneine biosynthesis.
A possible function, next to the regulation of the cysteine level inside of the organism could be the formation of sulfite. Sulfite could be used as an extracellular reducing agent for the support of nutrient degradation. CDO induced sulfite production linked to the keratin degradation has been observed in the dermatophyte *A. benhamiae*.[187]

3.6 Experimental

The gene of the protein $EgtB_{fungi\_C}$, $EgtB_{fungi\_T}$, fCDO (WP_002375627) was codon optimized ordered for production in *Escherichia coli* from GenScript® and delivered in a pUC57 plasmid. The gene was ligated into modified expression vector pET28 using the restriction enzymes NdeI/XhoI. The modified pET28 vector contained a TEV-cleavage site after the His_6-tag. The sequence of the plasmid was verified by sequence analysis of Microsynth AG

Sequence of $EgtB_{fungi\_C}$

MGHHHHHHAENLYFQGHMPGLENPVLASQTGQGRLLAIKEKKRLPDVRVKIGEKASFDIIDIRQGSVEMNLKVEI
LSMFLTKNGPRKLPTLLLYDERGLQLFEKITYLEEYYLTNDEIEVLQKYSADIAKLIPEGAMLIELGSGNLRKVN
LLLRAFEDAGKSIDYYALDLSKQELERTLAQLPHYQYVRAHGLLGTYDDGRAWLKHPSRASRQKCILSLGSSVGN
FDRADAAAFLKTFADILGPGDTMLIGLDACNDPARVYHAYNDKEGVTHELASQAGSGDSADESIHRFILNGLRHA
NKILGETVFVEAEWRVIGEYVYDGQGGRHQAFYAPLHDTTVLGQLIRPHDRIQVEQSLKYSPAEAELLWKRAGME
EIGHWRCRDEYGVHMLSKPKMAFGLIPSVYARSALPSLEEWESLWAAWDTLTQEMLPPEELLSKPIKLRNACIFY
LGHIPNFLDVQLSKVTTDPLTDPAWYRRIFERGIDPDVDNPEICHDHSEIPDEWPPADEILEYQTRVRARLRKYY
ENGVENIPRHIGRAIWVGFEHEIMHLETLLYMMLQSDKTRPPPNVPVPDWEKLAAKARSERVPNEWFDIPEQEIT
IGLDDPEDETDPNVHYGWDNEKPVRRAKVHAFQAKGRPITNEEYATYLYNTHGSQIPASWAYTKEKDPQNGVSGT
NGHSTIANGTAPLPESFLEDKAVKTVFGLVPLKYALDWPVFASYNELAACAAWMGGRIPTFEEVRSIYAHVEARK
KQKEAQKHLAQTVPAVNAHLCNNGVEISPPATPPAGTAAATAEGDESENSLFIDLDGANVGFQHWHPVPVTNRGG
ELAGQAECGGVWEWTSSVLRPWDGFKPMTLYPGYTADFFDEKHNIVLGGSWATHPRIAGRKSFVNWYQRNYPYAW
VGARLVRDLP

Sequence of $EgtB_{fungi\_T}$

MGHHHHHHAENLYFQGHMPIAQETTVVPQGATPRLGAIKEKKGVPPLRTAAATGPYIIDIRHAVVEINLKAEVLA
QFRARDGPRKLPTLLLYDENGLQLFEKITYLEEYYLTNDEIAVLKSYAADIVKYIPSGAMVIELGSGNLRKVNLL
LQALEDAAKDIDYYALDLSQQELERTLAQLPPYKHVRAHGLLGTYDDGRLWLKDPTIATRQKCILSLGSSVGNFD
RREAGAFLKSFADVLGPDDTMLIGLDACDDPAKVYHAYNDKEGITHEFILNGLRHANRILGENAFVEKDWRVIGQ
YVYDAEGGRHQAFYAPVRETIVMGERIRPHDRIQVEQSLKYSAAEAEELWRRAGMTEIAQWRHLKEYGLHMLARP
RMAFSLTPSVYARTALPSMRDWEGLWAAWDVVTRDMLPPEEILEKPIKLRNACIFYLGHIPTFLDIQLTKTTKQP
PTDPAYYYGIFERGIDPDVDNPELCHAHSEIPDEWPPVDEIRAYQGRVRARLQSLYAAGIDAIPRHVGRAIWVGF
EHEAMHLETLLYMMLQSDRTRPPPRIPAPDFESLAAKARSERVPNQWFDIPEQEVVIGLDDPEDGTDPHAPYGWD
NEKPLRRVKVHAFQAQGRPISNEEYARYLYNTRTTKIPASWAQIPGGTVKTSDETAAAGGGTDSAEANNNNNENG
RTNSHSLANGVALPESFLDDKAVRTVYGLVPLKHALDWPVFASYDELAGCAAWMGGRIPTFEEVRSIYKHAEALK
KEQAENQLSQTVPAVNGHLTNNGVEISPPATPPGSTDAGSEADSRDRRLSQEDLFIDLNGANVSFSHWHPTPVTS
RGNQLAGQSDAGGVWEWTSSVLRPWDEFQAMSLYPGYTADFFDEKHNMVLGGSWATHPRLAGRKSFVNWYQRNYP
YAWIGARLVRDVE

Sequence of fCDO

MGHHHHHHAENLYFQGHMLNLFCEAGLRSIIGLPSRSANTTTTLYILERPPLTFPPFSSTPRSAPNPIANPYTLP
SLAEFTRAWAIWDLITLGMIPSELLHSKPIDLRHKPLFYIGHLPTFANILLSRVIGEREVGPRHYLTTFERGIDP
SVDDPERCHSHSEVPERDEDWPVIGDVLAYRDEVREKVIKRIFAEVESGERALTRRLARTMVMVHEHDGFHIETL
LYMLIQRAGTGMLPPPGFAPPPWPALAAQWDAIPAPTTPTVTLGPATITMGHDDQEPDDLLPALEHDVGAHEFGW
DNESPARAVHVGAFRVEWRPVTNGEFLAFWQGPGKDVVDMPASWAQTEDGEVRVRTLYGPVPMAHAKHWPVLTAY
DDLAKYAAHKGGRIPTEPELRLFLDAYQVGYEEGANTGFRHWHPLPATAGLQEIDGGRGSNGGVWEWTATALDAH
PGFVGTGIFPGYSSDFFDGKHQVVLGASYATIPRLGDRRTVRNFYQHNYPYPWVGARVAYDV

Sequence of $EgtB_{fungi\_D}$

MASKSANAIQILDIHTRVSPPSGASIRDQIVSGLSQPVGGKTLPTLLLYDERGLRLYDEITTDAAEYYLFPAEEE
ILKNKADGIVRVMHSGIANGELVDEVVVELGAGALRKTSLILGALARLVPNPAPTPPISYHALDLEKRELERTLM
ELNSSHVGVELKGKVATAGLCGTYDDGLKFIVEGGLQDRNALERIDTAFGSQYPVERVGRDASPSSAASSRGRTD
RTEATPPSTPGAQQPLHILFLGSSLGNFARGEDAAFLKSLPLRPGSGDTLLLGMDHGNDARRIEAAYNDSKGVTR
KFIMNGLVSAGRALGDESLFAQDKWEYVGKYNEESRRHEAYYRSACDQTVVDPETKALFPFVKDELIRVEVSHKY
SERDAWTVFTEANLCPVHRWTDRSSQYSLWLLERPKFSFPLLKWPSSSDAKTVARSPFSLPTVDEWRDMWAAWDF
ITRQMIPPSMLFQKPIDLRHICLFYCGHIPAFLSIHLSKLLQEPDTEPVEYKYIFERGIDPIVDDPTQCHPHSEV
PQKDEDWPSLPSILQFQSRVRERVINLYRDIDSGKVTLTRKIARVMQMTLEHEAFHAETLLYMLLQRAGTGTIPP
SGFIPPPWEVLAESWDKQPLPAAETVTLGPEEISLGHDDDESQDDSTEGVLEHSFGWDNEHPKRTVKVEEFKIEW
RPVTNGQFYEFYTGPGKDKVQFPKSWVELDGEVFVRTLYGPVPIKIAWHWPIKTSYDNLSVYANVKGGRIPTEPE
LRLFLDKFECGYEGGANVGFRNWHPVPATTGGPADGGKGHNGGVWEWTSTVLDKYEGFVPSKLYPGYSTDFFDTH
HQVVIGGSYTTIPRLAERRTVRNYYQHNYPYAWVGARVAYDV

Protein Production

The fungal genes were purchased by Genscript® and cloned into a modified pET28 production vector, which contained a TEV-cleavage site after the N-terminal His_6-tag. The modified expression vector pET28 containing a eukaryotic enzyme was transferred into BL21 (DE3) pLys containing an additional plasmid pL1SL2, with the encoded *E. coli* chaperon DnaK. The cells were grown in lysogeny broth containing ampicillin (100 mg/L) and chloramphenicol (34 mg/L) and kanamycin (50 mg/mL) at 37˚C. After the OD_{600} of the cell culture reached 0.6-0.8 the cultures were cooled down to 18˚C and then induced by adding 0.1 mg/mL isopropyl-β-D-thiogalactopyranoside (IPTG). The protein was expressed for16 hours at 18˚C, harvested by centrifugation (7000 rpm) and stored at -20 ˚C until protein purification. For the purification, the cell pellet was resuspended into lysis buffer (300 mM NaCl, 50 mM Na_2HPO_4) and lysed by an Avestin EmulsiFelx-C3 homogenizer. To remove the insoluble cell components, the mixture was centrifuged for one hour at 4000 rpm at 4˚C. Ni^{2+} NTA agarose beads (Qiagen) were added and after binding of the enzyme for 20 minutes at 4˚C. The Ni^{2+} column was washed with 20 mM imidazole and afterwards the protein was eluted with 250 mM imidazole. To remove Ni^{2+} ions from the purification step, the protein was incubated for 20 minutes at 0˚C with 10 mM EDTA. After two rounds (minimum 4 hours) of dialysis in 50 mM Tris/HCl pH 8 and 50 mM NaCl buffer the protein was flash frozen and stored for further use.

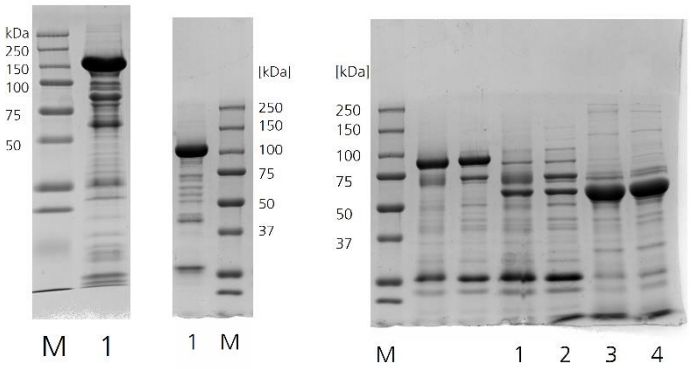


Figure 92: SDS pages after Ni-NTA affinity chromatography. Left: M: molecular weight marker, 1: $EgtB_{fungi\_C}$ Middle: E: 1 $EgtB_{fungi\_T}$, M: molecular weight marker, right: M: molecular weight marker, 1-2: fCDO produced without pL1SL2; 3-4 fCDO produced with pL1SL2.

Size exclusion chromatography.

The size exclusion protein purification was performed on an Äkta Prime with a HiLoad Superdex 26/600 column. The flow rate was 2 mL/min in an elution buffer containing 50 mM HCl/Tris pH 8 and 200 mM NaCl.

HPLC assay

A cation-exchange HPLC containing a Luna 5u SCX column (100 Å, 150 x 4 mm, Phenomenex) was used for the reaction mixture analysis. The mobile phase (solvent A) was 20 mM phosphoric acid pH 2 and the product were eluted with a gradient of solvent B (1 M NaCl, 20 mM phosphoric acid, pH 2). The gradient was defined according to the following method and the product formation was monitored at an absorption of 265 nm.

Table 16: Cation-exchange standard HPLC method used for the EgtB reactions.

Time (min)	Solvent A (%)	Solvent B (%)
0.01	85	15
2	80	20
8	1	99
9	1	99
11	85	15

Michaelis-Menten plots/Enzyme assay.

The eukaryotic enzymes were measured with the standard condition containing 100 mM phosphate buffer and 100 mM NaCl, 2 mM ascorbate, 2 mM TCEP, 1 µM enzyme and 10 µM $FeSO_4$.[60] The substrate concentrations were changed between 5-1000 µM and the reaction temperature was set to 20°C. The

reaction mixture was quenched in a ratio of 2:1 with 1 M phosphoric acid and then analyzed by cation exchange HPLC with the method described before.

For the selenocysteine measurements the TCEP concentration was increased to 5 mM.

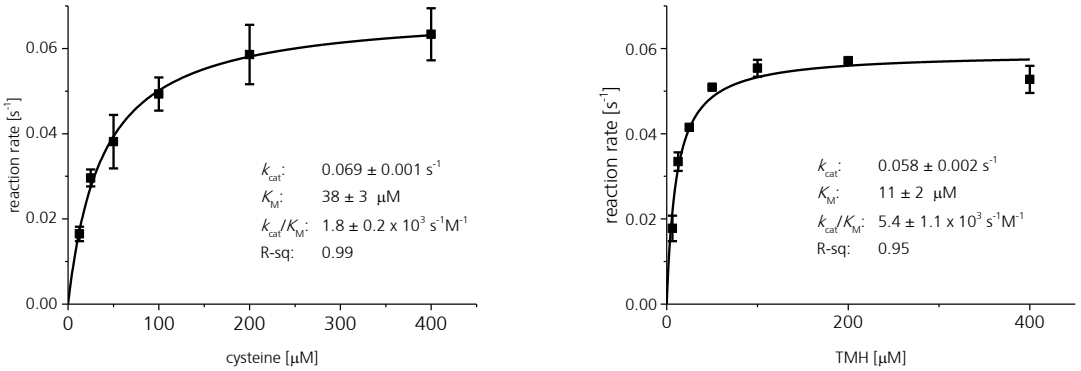


Figure 93. **Left:** Michalis-Menten plot of $EgtB_{fungi\_C}$ at 20°C with different cysteine concentrations and 0.5 mM TMH. **Right:** Michalis-Menten plot of $EgtB_{fungi\_C}$ at 20°C with different TMH concentrations and 0.5 mM cysteine.

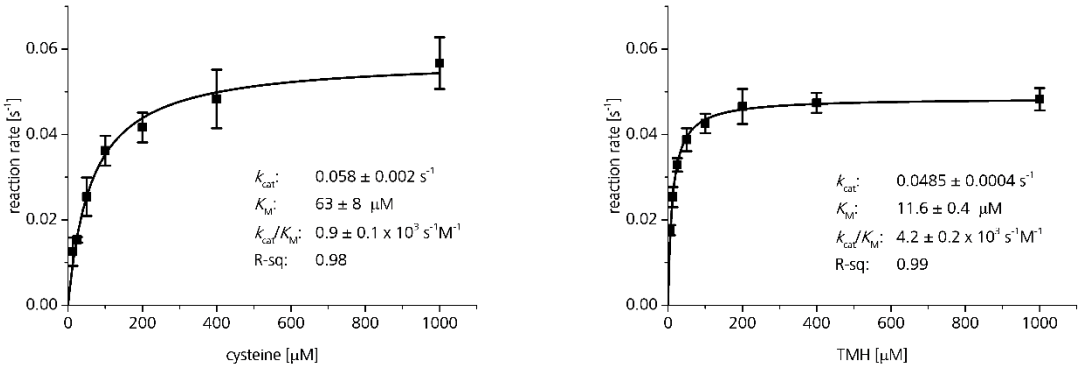


Figure 94. **Left:** Michalis-Menten plot of $EgtB_{fungi\_T}$ at 20°C with different cysteine concentrations and 0.5 mM TMH. **Right:** Michalis-Menten plot of $EgtB_{fungi\_T}$ at 20°C with different TMH concentrations and 0.5 mM cysteine.

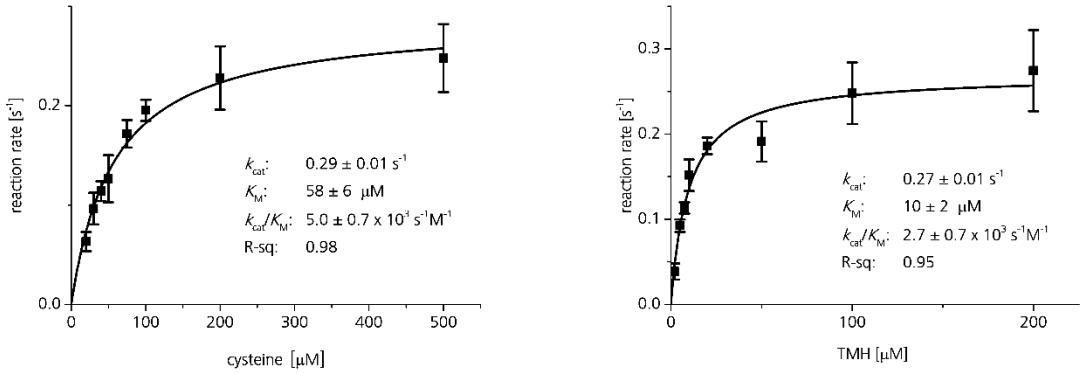


Figure 95. **Left:** Michalis-Menten plot of $EgtB_{fungi\_C}$ at 40˚C with different cysteine concentrations and 0.5 mM TMH. **Right:** Michalis-Menten plot of $EgtB_{fungi\_C}$ at 40˚C with different TMH concentrations and 0.5 mM cysteine.

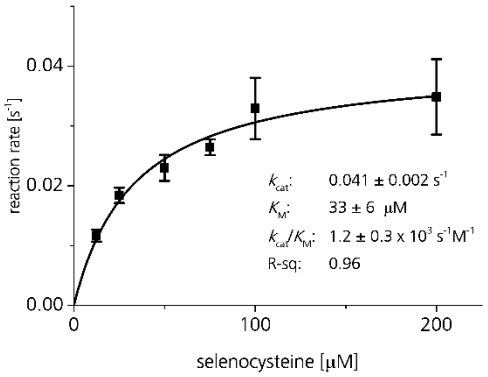


Figure 96. Michalis-Menten plot of $EgtB_{fungi\_C}$ at 20˚C with different selenocysteine concentrations. The TCEP concentration was increased to 5 mM.

Kinetic Isotope effect:

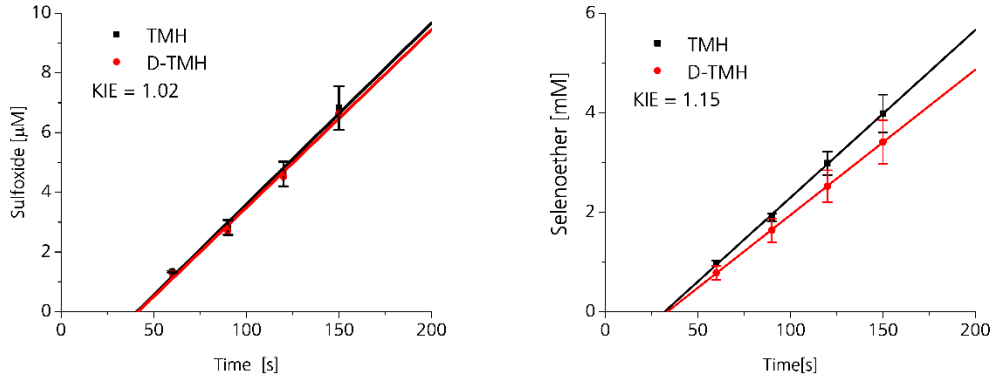


Figure 97. **Left**: The measured substrate isotope effect with D-TMH and cysteine measured for the sulfoxide formation. **Right**: The measured substrate isotope effect with D-TMH and selenocysteine measured for the selenoether formation. The condition for the assay were as follows: 100 mM phosphate buffer pH 8, 100 mM NaCl, 2 mM ascorbate, 10 µM $FeSO_4$, 2 mM TCEP for cysteine reactions, 3 mM TCEP for selenocysteine reactions, 500 µM D-TMH, 500 µM cysteine or 200 µM selenocysteine, 1 µM $EgtB_{fungi\_C}$ to a total volume of 250 µLl. The reaction temperature was 20˚C and the

samples of the reactions were quenched with 1 M phosphoric acid in a ratio 1:2. The samples were analyzed with the standard HPLC method described before.

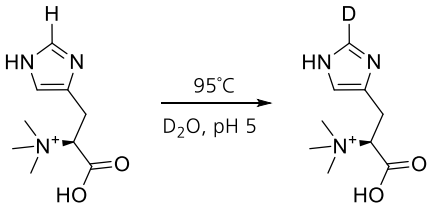


Figure 98. Synthesis of C2-deuterated TMH: TMH was dissolved into a D_2O solution pH 5 (acidified by DCl) to a final concentration of 5 mM. The reaction was heated up to 90°C for 72 hours. The deuteration was controlled by ^{1}H-NMR (400 MHz, D_2O) δ ppm 7.26 (s, 1H), 3.88 (dd, J = 12.0, 3.9 Hz, 1H), 3.44-3.24 (m, 2H), 3.23 (s, 9H).

Binding order kinetic

For the binding order kinetics Michaelis Menten curves were measured at different constant levels of one substrate. This was then repeated for the second substrate. The constant cysteine levels were (10, 30, 45, 100 and 400 µM) and the constant TMH levels were (5, 10, 20, 50 and 100 µM).The reaction mixtures contained: 100 mM phosphate buffer pH 8, 100 mM NaCl, 2 mM ascorbate, 10 µM $FeSO_4$, 2 mM TCEP and 0.1 µM enzyme. Time points were taken between 2 and 8 minutes and analyzed with the HPLC assay described before.

Selenocysteine inhibition reaction

The reaction rate of sulfoxide formation and selenoether formation in the same reaction mixture containing: 100 mM phosphate buffer pH 8, 100 mM NaCl, 2 mM ascorbate, 10 µM $FeSO_4$, 2 mM TCEP, 1 mM TMH, 60 µM cysteine and 200 µM selenocysteine, 0.1 µM $EgtB_{fungi\_C}$ in a total volume of 250 µL. The reaction temperature was 20°C and the samples of the reactions were quenched with 1 M phosphoric acid in a ratio 1:2. The samples were analyzed with the standard HPLC method described in the experimental section

2A-DMH inhibition reaction

The 2-amino N α -dimethyl-L-histidine was prepared in our labrarotory with the previously reported synthesis route.[159] Inhibition reaction with and without 2A-DMH were performed under the following conditions: 100 mM phosphate buffer pH 8, 100 mM NaCl, 2 mM ascorbate, 10 µM $FeSO_4$, 2 mM TCEP, 0.1 mM TMH, 0.1 mM cysteine, 0.2 µM $EgtB_{fungi\_C}$ in a total volume of 250 µL. The inhibitor concentration was 0 mM or 1 mM and the reaction temperature was 20°C. The samples of the reactions were quenched with 1 M phosphoric acid in a ratio 1:2 and analyzed with the standard HPLC method described in the experimental section.

NMR experiment:

A reaction mixture containing 2 mL of the following ingredients was prepared: 100 mM phosphate buffer pH 8, 100 mM NaCl, 2 mM ascorbate, 10 µM $FeSO_4$, 2 mM TCEP, 1000 µM cysteine, 4000 µM TMH, 2 µM $EgtB_{fungi\_C}$, 2 µM $EgtB_{fungi\_D}$ or 2 µM fCDO. After the addition of the enzyme the reactions were incubated for 12 h at room temperature. In the next morning the samples were lyophilized and dissolved in 600 µl D_2O. A ^1H NMR (600 MHz) spectrum for both enzymes were obtained.

Tryptic digestion:

The enzyme was catalyzing the sulfoxide or the selnoether formation until complete deactivation. The deactivation was monitored by HPLC. The reaction conditions were as follows: 100 mM phosphate buffer pH 8, 100 mM NaCl, 2 mM ascorbate, 10 µM $FeSO_4$, 2 mM TCEP, 2 mM cysteine reactions, 2 mM TCEP, 2000 µM cysteine, 2000 µM TMH, 4 µM $EgtB_{fungi\_C}$ at room temperature in a total volume of 10 mL. The protein sample were concentrated to 100 µL and digested with 0.1 µg/µL trypsin (purchased from Roth Chemie GmbH) for 2 hours. The tryptic digestion was monitored with SDS-page. After the digestion 1 % TFA and 2 mM DTT were added and the sample were injected into the UPLC-MS.

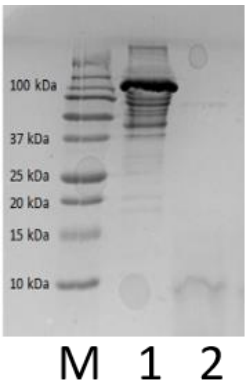


Figure 99. SDS-page of the trypsin digestion. M: molecular weight marker, 1: $EgtB_{fungi\_C}$ before tryptic digestion. 2: $EgtB_{fungi\_C}$ after 2 hours of tryptic digestion (trypsin concentration: 0.1 µg/µL).

Crystallization:

The *N*-terminal His_6-tag of the purified protein $EgtB_{fungi\_C}$ or $EgtB_{fungi\_T}$ was then removed with the recombinant tobacco etch virus (TEV) protease in 50 mM Tris pH 8.0, 50 mM NaCl, 0.5 mM EDTA and 1 mM DTTfor 24 hours at 4°C. The protein mixture was then filtered over Ni^{2+} NTA agarose, incubated for 10 minutes with 10 mM EDTA and purified by size exclusion. The enzyme was then concentrated using centrifuge concentrators (Millipore) to a final concentration of 15 mg/mL and flash frozen in liquid nitrogen. The protein solution for crystallization contained next to the enzyme (10 mg/mL), 2 mM TCEP, 1 mM SAH, 1 mM DMH, 2 mM ascorbate and 20 µM $FeSO_4$.

The protein was screened for crystallization with commercial available crystal screens from Molecular Dimensions using 96-well 3-drop Swissci plates. Each drop contained 0.2 µL protein solution and 0.2 µL crystallization solution at a constant temperature of 20 ˚C. Protein crystals were observed only in position G8 from the commercially available Morpheus HT screen containing the conditions: 0.1 M carboxylic acid (0.2 M Sodium formate; 0.2 M Ammonium acetate; 0.2 M Sodium citrate tribasic dihydrate; 0.2 M Sodium potassium tartrate tetrahydrate; 0.2 M Sodium oxamate), 1 M Buffer at pH 7.5 (Sodium HEPES; MOPS (acid)) and 50 % v/v precipitant mixture (25% v/v MPD; 25% PEG 1000; 25% w/v PEG 3350). This condition was used for further optimization. The optimization screen contained in the x-axis complex ingredients from 4-19 %. In the y-axis precipitant concentrations from 31.5-42 %w/v and this at two different pH (pH 7 and pH 7.5). The crystals, which appeared after 36 hours were flashfrozen in liquid nitrogen, stored and then further used for data collection at the Swiss light Source (SLS), Villigen, Switzerland. The crystals did not diffract or diffract in a low resolution of more than 10 Å.

Alkylation of surface cysteine

$EgtB_{fungi\_C}$ was diluted to a final concentration of 1 mg/mL in a reaction mixture containing 2 mM TCEP, 50 mM HCl/Tris and 50 mM NaCl. The reaction was incubated at RT for 30 min. To the reaction mixture 2-iodo acetamide was added to a final concentration of 5 mM. The reaction was incubated for another 120 mins in absence of light. After 120 mins 5 mM DTT were added and the protein mixture was purified with size exclusion chromatography.

Oxygen deactivation:

The reaction conditions were as follows: 100 mM phosphate buffer pH 8, 100 mM NaCl, 2 mM ascorbate, 2 µM $FeSO_4$, 10 mM TCEP, 2 mM cysteine or 2mM selenocysteine, 2000 µM cysteine, 2000 µM TMH, 0.1-1 µM $EgtB_{fungi\_C}$ in a total volume of 500 µL. The reaction temperature was maintained at either 20 ˚C or 40 ˚C.

4 $EgtB_{Vapar}$

4.1 Introduction

4.1.1 A novel class of EgtB Analogues

In the previous chapter, the plasticity of the EgtB-related enzymes were described. We found two different EgtB homologues in Firmicutes, a fusion protein of a methyltransferase, a sulfoxide synthase as well as a cysteine dioxygenase. We also discovered the existence of at least two different EgtB-homologue enzymes in the organism, *Variovorax paradoxus*, a gram negative proteobacterium. The first homologue is related to the EgtB from *Chloracidobacterium thermophilum* with a sequence identity of 55%. The second however shows only a sequence identity of 32% towards the same EgtB. The existence of two different EgtB homologues in the same organism points towards two different function. The function of the second EgtB homologue $EgtB_{Vapar}$ which contains a DinB domain as well as a FGE-like domain is not known. In the sequence similarity network (SSN) of EgtB enzymes a distinct class of enzymes with the unknown $EgtB_{Vapar}$ and homologues were discovered.[155]

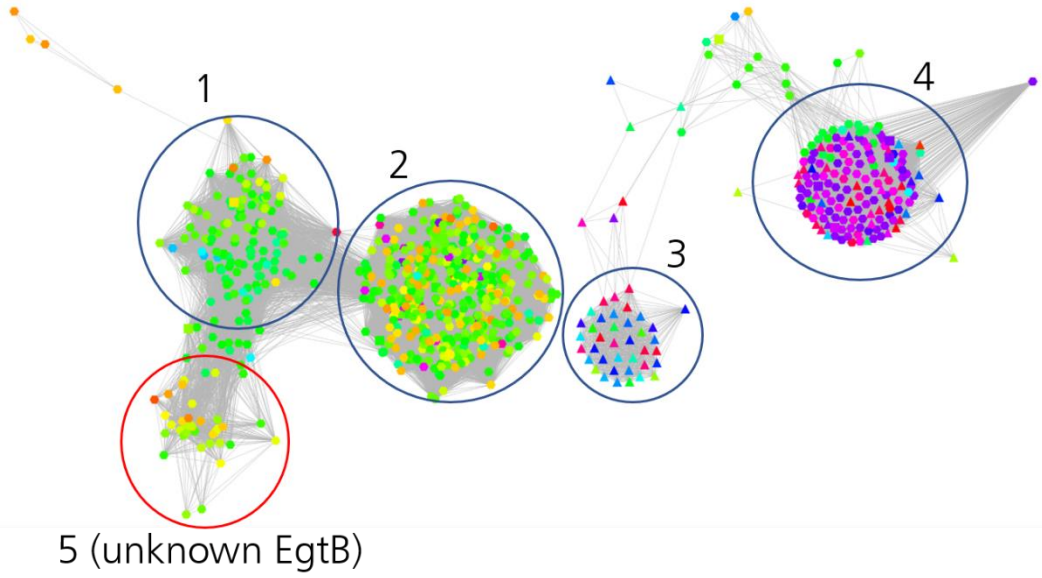
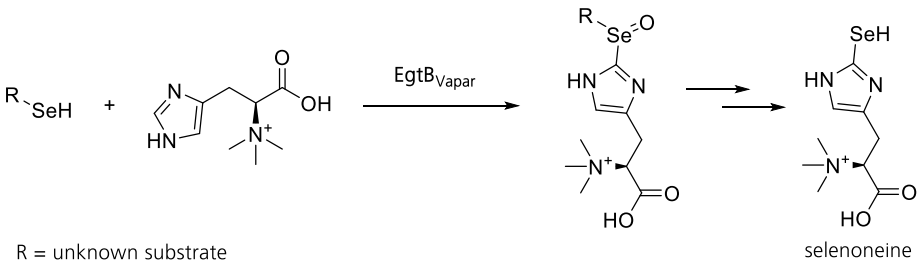


Figure 100. Sequence similarity network of EgtB and OvoA homologues. Sequences with a similarity higher than 75% are represented by a node. The group 1 and 2 belong to the bacterial EgtBs, group 3 to the eukaryotic EgtBs and group 4 to the OvoAs. Group 5, where $EgtB_{Vapar}$ was located in, are bacterial EgtB with unknown substrates.[155]

4.1.2 Biosynthesis of Selenoneine

Selenoneine is the selenium-containing analogue of ergothioneine and was discovered in tuna fish (see also Chapter 3).[165] We discovered *in vitro* assays that the eukaryotic EgtBs are able to form the carbon-selenium bond quite efficiently. These results are supported by the observation of selenoneine production in *in vivo* studies in the fisson yeast *S. pombe*.[156] However, the bacterial EgtB enzymes of *Myocobacteria thermoresistibile* and *Chloracidobacterium thermophilum* were found to be poor catalysts for the formation of the carbon-selenium bond.[74] We speculated that selenoneine is also found in prokaryotic species and therefore an alternative biosynthetic pathway in bacterial organisms needs to exist for the formation of selenoneine. A possible alternative route for selenoneine formation could be the $EgtB_{Vapar}$ with unknown substrates and a different active site architecture.



Scheme 34. A proposed biosynthetic route of an unknown selenosubstrate with trimethylhistidine to form selenoneine in bacteria.

4.1.3 Methods for Monitoring the Substrate Binding

In the enzyme, $EgtB_{Vapar}$ the substrates are unknown. A strategy to identify the right substrates could be to analyze possible ligand-enzyme interactions. Substrates usually have strong binding properties to enzymes. There are multiple methods to investigate this binding interaction between ligands and enzymes. Two useful and easily accessible methods are fluorescence based thermal shift assay and the isothermal titration calorimetry.[190,191]

In the fluorescence based thermal shift assay the melting curve of a protein is measured with and without ligand in the presence of a fluorescent dye. The difference of the melting temperatures can be used to estimate the ligand binding.

The fluorescent dye needs to be highly fluorescent in a non-polar environment while, in polar solution, the fluoresce needs to be quenched. If the protein denaturates, the fluorescent dye can bind to the hydrophobic residues and thereby start being fluorescent. By monitoring the fluorescence signal, the melting curve and therefore the stability of the protein can be measured.[191]

After the addition of a binding ligand, the melting curve increases in most of the cases. This observation is based on the free energy contribution of a binding ligand which increases the temperature-dependent Gibbs free energy of unfolding ΔG_u and results in an increased melting temperature (T_m).[191,192] The stabilizing effect of a ligand is proportional to its concentration and the affinity for the binding site.[193] This method is ideal to get an overview of compounds, which might bind to the enzyme and influence the T_m of the protein. A big advantage is that up to 30 potential ligands can be tested in the same experiment with rather low enzyme concentrations. Nevertheless, thermodynamic interpretations are difficult because ΔT_m is not always a reflection of their relative affinity and misinterpretation can occur.[191]

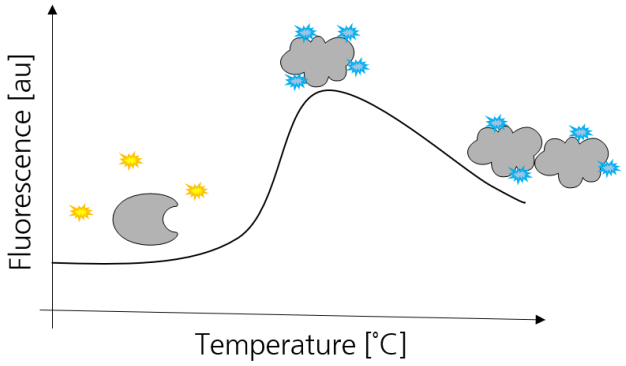


Figure 101. The unfolding curve monitored in the presence of a fluorescene dye (e.g. SYPRO orange). The fluorescence is quenched in the aqueous solution. After the protein is unfolded the dye can aggregate to the hydrophobic parts of the enzyme and lead to a strong fluorescence. The melting temperature of the protein can be determined by forming the first derivative of the displayed curve.[191]

A second more precise method to measure protein-ligand interaction is isothermal titration calorimetry (ITC). ITC is based on the measurement of heat effects, which are caused by biological binding interactions between a ligand and an enzyme. This method has broad applications and it has been used to measure macromolecule interaction with small ligands, other proteins, nucleic acids or drugs.[194–197]

The working principle of ITC is simple. The ligand is added in small fraction to a sample cell containing the protein solution. After each addition of the ligand, the heat difference between the sample cell and the reference cell is measured and adjusted. The required heat to adjust both cells is monitored in a titration thermogram. This data can be further processed to determine thermodynamic parameters.

With a single ITC measurement, it is possible to receive a full thermodynamic description of the interacting system including the binding enthalpy ΔH, the binding entropy ΔS, as well as the binding constant K. A further advantage of ITC is that no chromophore or fluorophore is required.[190,198,199] Nevertheless, a limiting factor is the high concentration of protein, which is required in each measurement.

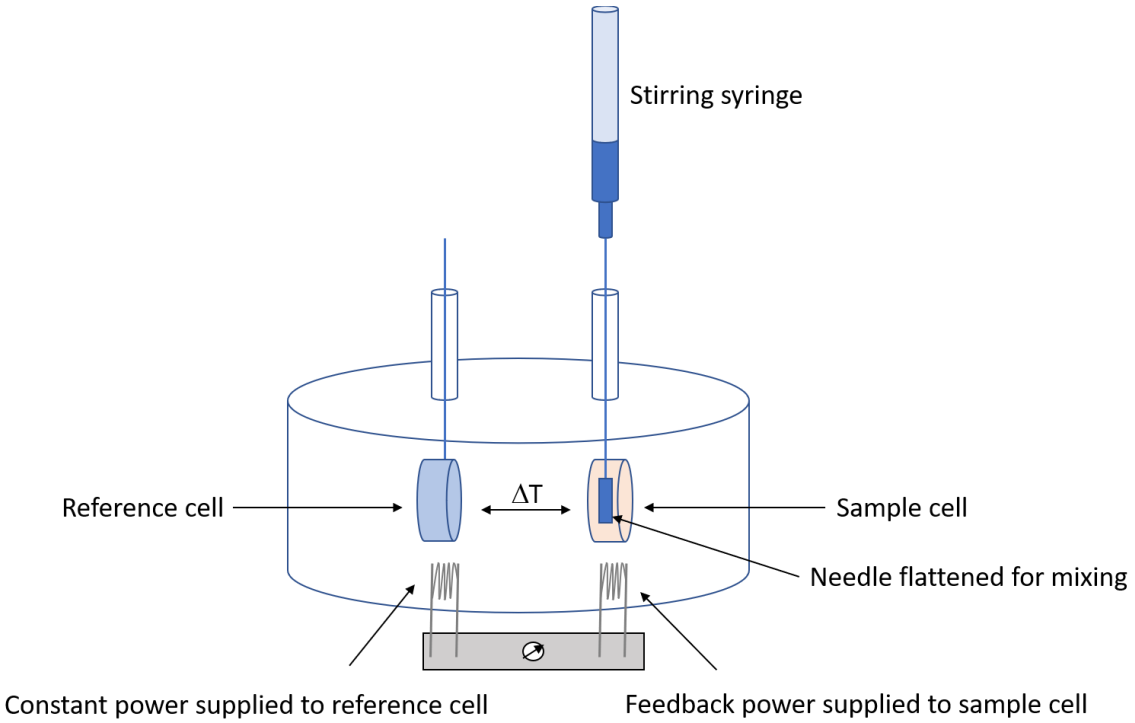


Figure 102. Schematic composition of a power compensation ITC. The sample cell and the reference cell are located in a temperature isolated shield. The ligand sample is added with a syringe to direct to the sample cell. The feedback power supply which is proportional to ΔT is used to determine the amount of energy which is used to maintain a constant temperature between the reference and sample cell.[190]

4.2 Aim of the Chapter

The discovery that two different EgtB analogue are present in the proteobacteria, *Variovorax paradoxus*, raises questions about the function of both these enzymes. While one enzyme was annotated as a cysteine-dependent sulfoxide synthase, the other protein belongs to a distinct enzyme cluster with an unknown function. Our goal was to analyze the characteristics of this unknown EgtB variant by analyzing the amino acid sequence and creating a homology model. Furthermore, the identification of the native substrates as well as the products, were of central interest. In addition, we were interested to obtain a crystal structure to get further hints on the function of this EgtB and to identify similarities and differences to other identified EgtB-related enzymes.

4.3 Results and Discussion

4.3.1 Sequence alignment

The enzyme EgtB$_{Vapar}$ consists of a DinB and a FGE-like domain. The sequence alignment between the enzyme EgtB from *Mycobacteria thermoresistibile* (EgtB$_{thermo1}$) and EgtB$_{Vapar}$ showed a sequence identity of 30%. The 3-His facial triad metal binding site is conserved in EgtB$_{Vapar}$ and the residues His71, His167 and H171 were identified as iron ligands. Furthermore, the tyrosine proposed as a catalytic acid in the mechanism of EgtB$_{thermo1}$ is conserved in EgtB$_{Vapar}$ (Tyr365).

4.3.2 Trimethylhistidine

The homology model and the sequence alignment showed that the amino acid residues of the trimethylhistidine binding site of EgtB$_{thermo1}$ are similar to the residues identified in EgtB$_{Vapar}$. The residue Asn395 and Y366 are conserved. The replacement of a glutamine to methionine in EgtB$_{Vapar}$ (Met170) was also observed in the eukaryotic enzymes, where TMH was identified as a native substrate. A second difference of EgtB$_{Vapar}$ is the replacement of a tryptophan residue to a phenylalanine (F396). The change of the aromatic amino tryptophan to the slightly smaller aromatic amino acid phenylalanine should not have a major impact to the binding affinity of TMH. The similarity of the binding residue in the homology model (Figure 103) as well as in the sequence alignment points towards TMH as one of the native substrates for EgtB$_{Vapar}$.

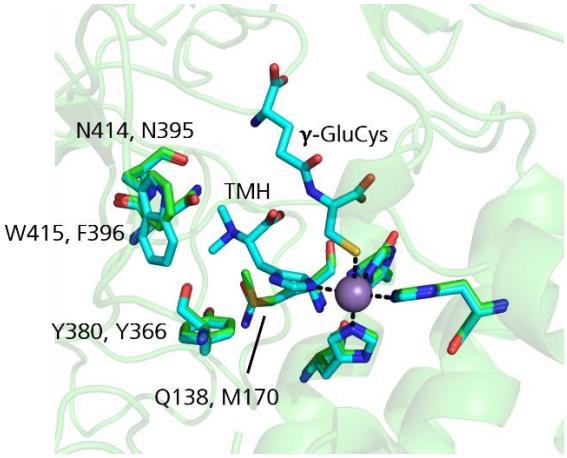


Figure 103. Superposition of the crystal structure of EgtB$_{thermo1}$ (cyan, pdb: 4x8d, 1.98 Å) and a EgtB$_{Vapar}$ homology model (green). The first number of the labeled amino acids corresponds to EgtB$_{thermo1}$ the second number to EgtB$_{Vapar}$. The residue Met170, Tyr366, Asn395 and Phe396 were tentatively identified to be involved in TMH binding.

4.3.3 Trimethylhistidine binding

The gene of $EgtB_{Vapar}$ was received from GenScript® and produced in *E. coli* BL21 cells with a good yield of 5-7 mg/L culture and good purity. With the isolated protein in hand, we were able to measure the TMH binding properties of this enzyme. In the first experiment, we performed a fluorometric thermal shift assay, where the melting temperature of the protein T_m was monitored with the suitable fluorescence dye SYPRO Orange.[191] With this method, we measured the binding properties of the enzyme without ligand, with histidine, dimethylhistidine and trimethylhistidine. The melting curves were measured in a triplicate and the first derivatives of these measurements were calculated to obtain the melting temperatures (Figure 104).

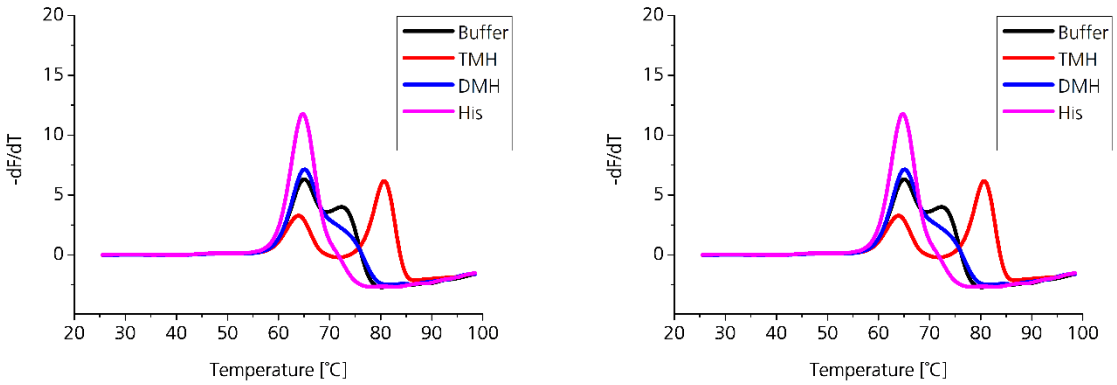


Figure 104. **Left:** Average of a triplicate measurements of the melting curve of the enzyme $EgtB_{Vapar}$ with different ligands. The enzyme concentration was 2.5 µM in 50 mM HCl/Tris buffer, 50 mM NaCl and a ligand concentration of 5 mM. **Right:** The first derivative curve of the melting curves.

In the measurements of the melting temperature we observe two melting temperature maxima at different temperatures. A possible explanation for this observation is that the enzyme $EgtB_{Vapar}$ consists of two different domains, which can have different melting temperatures. Furthermore, we observed that the ligand TMH significantly changed the second melting temperature. A high temperature difference of + 9 °C was observed, which indicates that TMH can interact with the protein and lead to its stabilization. The other ligands DMH and His did not have any stabilizing properties towards the enzyme. Interestingly, it seems that DMH and His have a negative effect to the stability of the second domain. The high stabilization effect of TMH supports the assumption that TMH is a native substrate of the enzyme.

Table 17: The melting temperatures measured with and without ligand and the change towards the control (Buffer).

Ligand	Melting temp. 1	Melting temp. 2	Δ Melting temp. 1	Δ Melting temp. 2
Buffer	65 °C	72 °C	-	-
TMH	64 °C	81 °C	-1 °C	9 °C
DMH	65 °C	-	0 °C	-
His	65 °C	-	0 °C	-

Another method to measure the binding properties of a potential substrate is the isothermal titration calorimetry (ITC). By adding a ligand in small portion to the enzyme solution, it is possible with the ITC to measure the thermodynamic parameters of interactions in solution.[190] With this method, we were able to measure the association rate K_a, the enthalpy ΔH and the entropy ΔS of the enzyme with the ligand TMH.

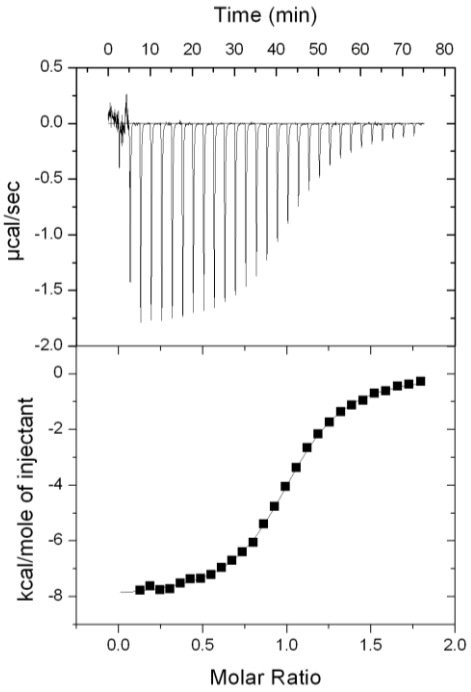


Figure 105: Final Figure of the ITC measurements measured with 86 µM enzyme in 50 mM HCl/Tris and 100 mM NaCl buffer. The ligand solution (1 mM TMH) was added in 10 µL portions.

The K_D value, which corresponds to the amount of ligand which is required to saturate half of the enzyme was calculated with a 1:1 fitting model. It was calculated to be 2.5 µM. This is a low value in the range of the K_D values measured for the EgtD enzyme of *M. smegmatis*.[21] This provides further evidence for TMH being the native substrate of $EgtB_{Vapar}$. The binding of the ligand is an exothermic reaction with an enthalpy value of – 8.1 kcal/mol. The entropy was calculated to be -1.5 cal/molK (Table 18).

Table 18. Calculated ISC parameters with 0.086 mM enzyme and 1 mM TMH at 25 °C.

Enzyme	K_A [M^{-1}]	K_D[M]	ΔH [kcal/mol]	ΔS [cal $mol^{-1}K^{-1}$]
$EgtB_{Vapar}$ + TMH	(4.0 ± 0.1) x 10^5	(2.5 ± 0.1) x 10^{-6}	- 8.1	- 1.5

4.3.4 Second substrate

Analysis of the $EgtB_{Vapar}$ sequence, combined with a homology model, fluorescence and ICT are consistent and strongly suggest that TMH is the first substrate for $EgtB_{Vapar}$. However, we do not know what the second substrate of $EgtB_{Vapar}$ might be. The sequence alignment shows that one arginine residue, which is involved in the carboxylic group binding of cysteine in $EgtB_{thermo1}$ is conserved. The other arginine is replaced by a histidine in the $EgtB_{Vapar}$ sequence. In addition, we observed that an alanine residue which is located close to the active site in the $EgtB_{thermo1}$ is changed in $EgtB_{Vapar}$ to a serine residue (Figure 106). In the homology model the Arg120 is congruent with the Arg90 of $EgtB_{thermo1}$, while the rest of the binding pocket looks completely different (Figure 107).

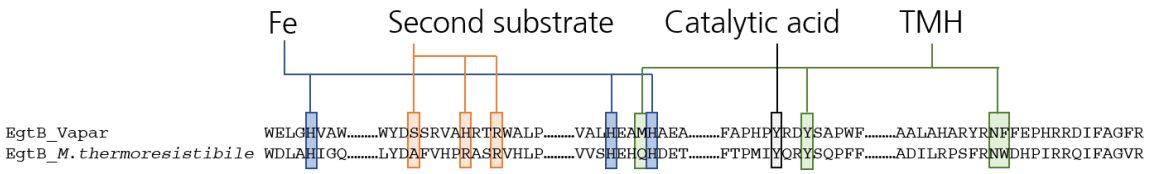


Figure 106. Sequence alignment between the two EgtB sequences $EgtB_{thermo1}$ and $EgtB_{Vapar}$.

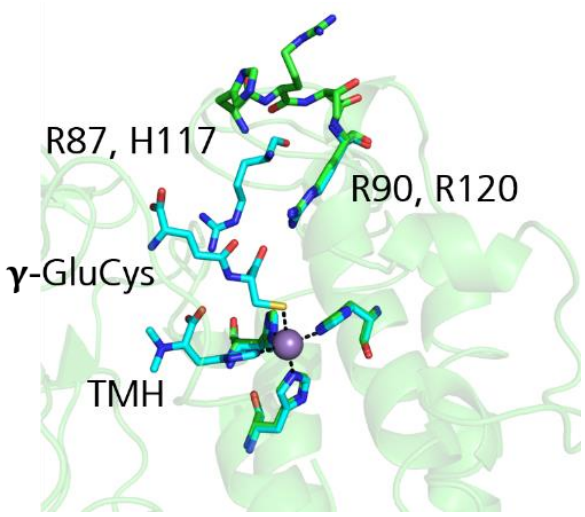


Figure 107. Highlighted catalytic residues involved in the binding of the second substrate. Only one arginine residue of the bacterial EgtB is conserved. First numbered residue refers to the crystal structure (cyan, PDB: 4x8d, 1.98 Å) second number to the homology model (green).

4.3.5 Three-Gene Cluster

One possibility to identify a potential substrate is to look at the genomic map of the organism and identify genes, which are located in gene clusters around the target enzyme. Especially in bacteria, numerous examples of genes involved in the same biochemical pathways are located in clusters.[150] Examples for gene clusters are the ergothioneine biosynthesis in *Mycobacteria smegmatis* or the biosynthesis of tryptophan and histidine in *Salmonella thyphimurium*.[71,200] In the gene map of *Variovorax paradoxus* we discovered that next to the $EgtB_{Vapar}$ gene, genes for a glycosyltransferase and a selenokinase are located (Figure *108*).

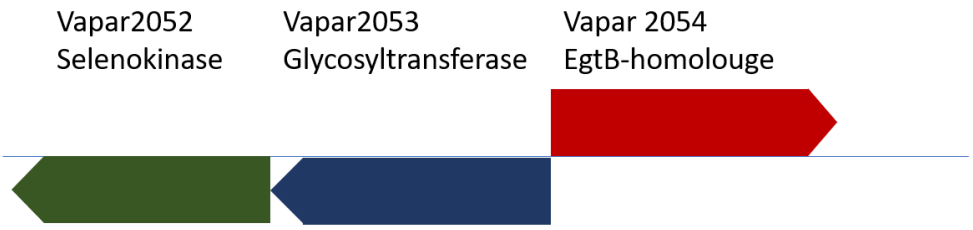


Figure 108. The gene cluster found in the organism *Variovorax paradoxus* containing the three enzymes, Vapar2052 a selenokinase, Vapar2053 a glycosyltransferase and Vapar2054 a EgtB homologue.

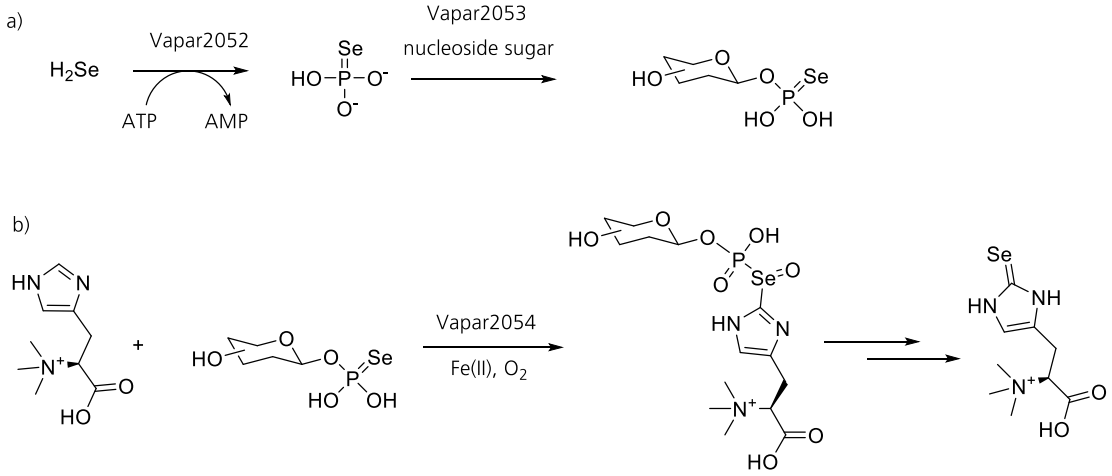
In the next step, we analyzed the gene map of the sequences which are in the same group as $EgtB_{Vapar}$ in the sequence similarity network (SSN).[155] We analyzed the available 47 sequences from the SSN and we were able to identify in 30 (65%) of those sequences a gene cluster containing the EgtB sequence, a glycosyltransferase and a selenokinase (Table 19). This conserved and ordered gene clusters point to a dependent biological relevance. There are different reasons why 100 % conservation of this gene cluster was not observed. First, we also identified EgtB homologues which are in this group but contain significant differences in the amino acid sequence towards $EgtB_{Vapar}$. Second also the genomes of bacteria are not only 2-dimensional and therefore a gene cluster can be split in fragments in certain organisms. In summary, we can state that the conservation of those three genes is probably more than just a coincidence.

Table 19: List of sequences from the SSN. The accession number of the NCBI website is listed. For every sequence the gene region around the EgtB homologue was searched for glycosyltransferase and selenokinase. The identified genes are displayed.

		EgtB	Glycosyltransferase	Selenokinase
1	*Limnobacter_sp.*	WP_008247400	n.a.	n.a.
2	*Sulfurifutis variabilis*	WP_096459807	WP_096459807	WP_096459807
3	*Anaeromyxobacter_sp.*	WP_011984884	WP_011984884	n. a.
4	*Methylobacterium_oryzae*	WP_043345642	WP_043345642	WP_043345642
5	*Halomonas_chromatireducens*	WP_066452119.1	WP_066452434.1	WP_066452120.1
6	*Streptomyces_luteus*	WP_043376033.1	WP_043376035.1	n.a.
7	*Thiobacillus_denitrificans*	WP_011311534.1	n.a.	WP_011311534.1
8	*Rhodoferax_ferrireducens*	WP_049765633.1	n.a.	n.a.
9	*Cupriavidus_necator*	WP_013952020.1	WP_013952021.1	WP_041228525.1
10	*Janthinobacterium_sp.*	WP_046684819.1	n.a.	n.a.
11	*Ralstonia_pickettii**	WP_042886710.1	n.a.	n.a.
12	*Betaproteobacteria_bacterium*	WP_023471952.1	WP_023471951.1	WP_023471950.1
13	*Methyloversatilis_sp.*	WP_069039889.1	WP_069039890.1	WP_069040729.1
14	*Variovorax_paradoxus*	WP_047785598.1	n.a.	n.a.
15	*Thiobacillus_denitrificans*	WP_011311534.1	n.a.	WP_011311533.1
16	*uncultured_beta_proteobacterium**	WP_059756770.1	n.a.	WP_059756773.1
17	*Rhizobacter_sp.*	WP_082548859.1	WP_056464217.1	WP_056464215.1
18	*Massilia_sp.*	WP_056128449.1	WP_056127872.1	WP_056127869.1
19	*Massilia_sp._2*	WP_056405421.1	n.a.	WP_056404770.1
20	*Cupriavidus_basilensis*	WP_043346234.1	WP_043346232.1	n.a.
21	*Janthinobacterium_sp.*	WP_054264600.1	WP_054263186.1	WP_054263187.1
22	*Rhodoferax_ferrireducens*	WP_049765633.1	WP_011463913.1	WP_011463912.1
23	*Sulfuritalea_hydrogenivorans*	WP_084207358.1	WP_052473550.1	WP_041099245.1
24	*Massilia_sp._3*	WP_047826502.1	WP_047826493.1	WP_047826492.1
25	*Duganella_sp.*	WP_057263203.1	WP_057263204.1	WP_057263205.1
26	*Cupriavidus_sp.*	WP_035827648.1	WP_035827650.1	WP_035827652.1
27	*Massilia_sp._4*	WP_081925800.1	n.a.	n.a.
28	*Variovorax_sp.*	WP_068674688.1	WP_068674895.1	WP_068674686.1
29	*Methyloversatilis_sp.*	WP_069039889.1	WP_069039890.1	WP_069040729.1
30	*beta_proteobacterium*	WP_054125751.1	WP_082374809.1	WP_054125670.1
31	*Burkholderiales_bacterium*	WP_009551713.1	WP_009551714.1	WP_009551714.1
32	*Lautropia_sp.**	WP_054139983.1	WP_082354828.1	WP_054139982.1
33	*Leptothrix_cholodnii*	WP_012346753.1	WP_012346754.1	WP_012346755.1
34	*Hydrogenophaga_crassostreae*	WP_066085507.1	WP_066085510.1	WP_066085513.1
35	*Candidatus_Gallionella_acididurans**	WP_008447177.1	WP_008447176.1	WP_008447175.1
36	*Methylibium_sp.*	WP_047494294.1	n.a.	n.a.
37	*Variovorax_paradoxus*	WP_021006821.1	WP_021006820.1	WP_021006819.1
38	*Betaproteobacteria_bacterium*	WP_023471952.1	WP_023471951.1	WP_023471950.1
39	*Variovorax_sp.*	WP_068629109.1	WP_068629107.1	WP_068629105.1
40	*Candidatus_Accumulibacter_sp.*	WP_075148898.1	WP_075148899.1	n.a.
41	*Candidatus_Accumulibacter_sp.*	EXI86331	n.a.	n.a.
42	*Lautropia_sp.*	ODT37043.1	n.a.	n.a.
43	*Anaeromyxobacter_sp.*	WP_059437016.1	WP_059437017.1	n.a.
44	*Noviherbaspirillum_autotrophicum*	WP_052452189.1	n.a.	n.a.
45	*Noviherbaspirillum_autotrophicum*	WP_040042554	n.a.	n.a.
46	*Thiobacillus_denitrificans*	WP_059756770.1	n.a.	WP_059756773.1

4.3.6 Proposed Second Substrate

The most obvious reason for the conservation of the three-gene cluster containing an EgtB analogue, a glycosyltransferase and a selenokinase would be that all of them are directly involved in the same biosynthesis. The presence of a selenokinase would suggest that seleno phosphate is involved in the biosynthesis and therefore that the final product presumably contains a selenium residue. We know that TMH is presumably a substrate for this enzyme, we expect a final product, which contains a trimethylhistidine backbone as well as selenium. A logical product would be selenoneine, the ergothioneine analogue described in chapter 3.1.4. In this pathway a possible function for the glycosyltransferase could be the formation of a seleno phosphate carrier molecule based on a sugar molecule, which enables the carbon-selenium bond formation. A possible biosynthesis of selenoneine could start with the formation of selenophosphate from selenite and ATP catalyzed by Vapar2052. The selenophosphate could then be transferred to a sugar molecule by the glycosyltransferase Vapar2053. In the final step the $EgtB_{Vapar}$ (Vapar2054) could form the selenoxide which could presumably get reduced and decompose spontaneous inside of the cell to the final product selenoneine (Scheme 35).



Scheme 35. Possible biosynthesis of selenoneine with the Vapar gen-cluster. **a)** In the first reaction selenide is phosphorylated with ATP catalyzed by the selenokinase Vapar2053. The product selenophosphate is then transferred to a sugar moiety by the glycosyltransferase Vapar2053. **b)** The thiosugar reacts then with TMH to form the selenoxide, which is then reduced and hydrolyzed to give the final product selenoneine.

The sequence identity of the selenokinase was 49 % towards the structurally and functionally characterized selenophosphate synthetase from *E. coli* (PDB: 3u0o).[201] A sequence alignment showed that the catalytic residues inside of the catalytic center are conserved, so we were very convinced that this enzyme is catalyzing the formation of selenophosphate out of selenite and ATP.

4.3.7 Glycosyltransferase Vapar2053

We were not able to find functional information about the conserved glycosyltransferase. Glycosyltransferase (GT) are classified into over 90 different distinct GT-families.[202,203] Each family is defined by amino acid sequence similarity, and is expected to have a related 3D-structure and a similar stereoselectivities. The different families are summarized in the CAZy database (www.cazy.org). Despite the functional and sequence diversity, the 3D structure of GTs are limited to three different main forms GT-A, GT-B and GT-C, which are found in Nature.[202–204] The GT-A structure contains a conserved DXD motif which binds to a divalent metal ion. The GT-B structure contains two Rossmann-like folds, where the binding site is located normally in the interdomain region of the two folds. The third structure is GT-C which contains a *N*-terminal transmembrane domain as well as a *C*-terminal globular domain.[203]

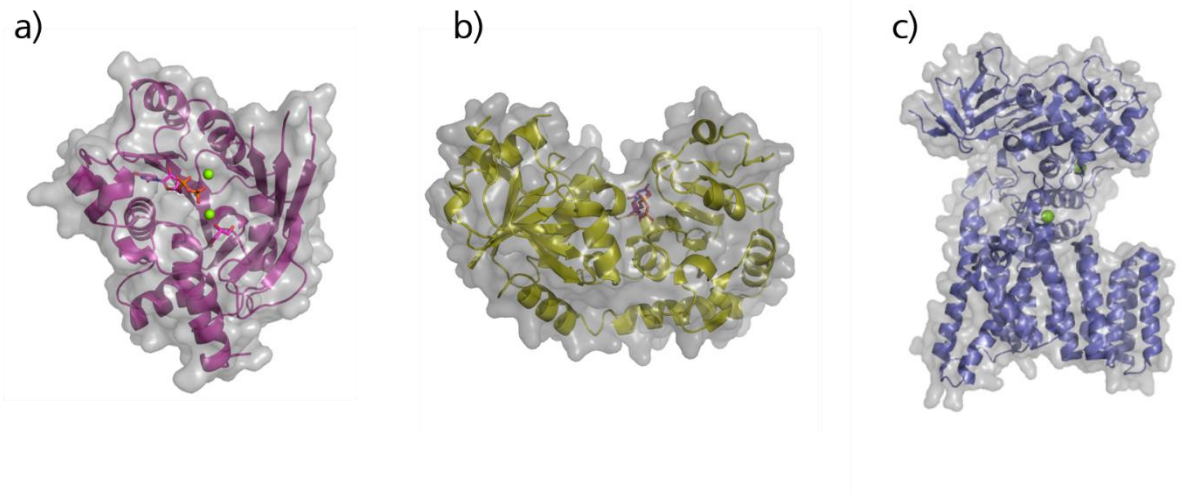
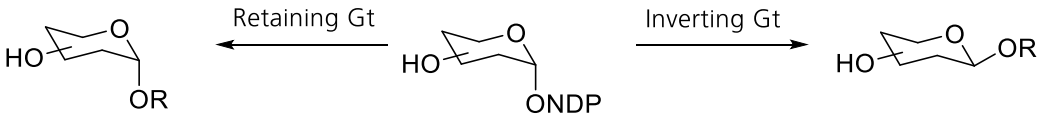


Figure 109. **a)** The GT SpsA from *Bacillus subtilis* with a GT-A fold with the ligands UDP and glycerol and two magnesium ions (PDB: 1qgs, 2.0 Å).[205] **b)** Bacteriophage T4 β-glucosyltransferase with the GT-B fold and UDP as a ligand (PDB: 2bgu, 2.2 Å).[206] **c)** Bacterial oligosaccharyltransferase of *Campylobacter lari* with a GT-C-fold and magnesium ions (PDB: 3RCE, 3.4 Å).[207]

In GTs the transfer of the glycosyl group can be grouped into two basic reaction mechanisms (Scheme 36). In the inverting mechanism the anomeric position of the sugar change from α in the substrate to β in the product. In the retaining mechanism the stereochemistry between the substrate and the product is not changed (Scheme 36).[202,203]



α-glycoside NDP-a-sugar β-glycoside

Scheme 36. The two-possible mechanisms of glycosyltransferases. In the retaining mechanism the stereochemistry of the anomeric position of the substrate is not changed, while in the inverting mechanism the product has a β configuration.

Stereoselectivity in glycosyltransferase families are mostly conserved. Only a few exceptions of glycosyl hydrolase have been found, where the stereoselectivity was not consistent within a certain family.[202] This observation lead to the classification of GT family into specific clans according to their fold and stereoselectivity (Figure 110).[208] This classification model is limited to the most common GT-A and GT-B folds.

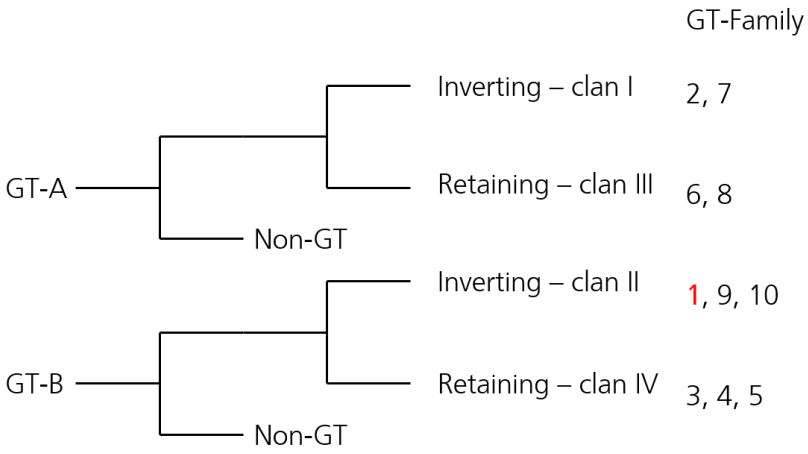


Figure 110: The glycosyltransferase classification system involving different clans. The GT-Family 1-10 in the corresponding clans are shown. GT-Family 1, which is classified in the clan II with inverting chemistry and where Vapar2053 belongs to is highlighted in red.[202,208]

The GT enzyme Vapar2053A, which is found in the *V. paradoxus* gene cluster is classified into the glycosyl transferase family 1 and therefore belongs to clan II. This means that Vapar2053 has a GT-B 3D structure with an inverting mechanistic chemistry.

4.3.8 An Alternative Glycosyltransferase

Because of solubility problems, the glycosyltransferase Vapar2053 of *Variovorax paradoxus* was not producible. Therefore, the studied glycosyltransferase was changed to the homologue of *Sulfurifustis variabilis* GT_{SVA}, which contained a sequence identity of 40%. *Sulfurifustis variabilis* is a gram-negative bacterium which belongs to the family of *Acidiferrobacteraceae* and has an optimal growth temperature of 42-45°C.[209] This protein was producible in *E. coli*. Nevertheless, during the dialysis step of the protein purification precipitation occurred. In addition, an absorption band at 260 nm was measured, which could indicate that a nucleoside building block is bound to the protein. Without the dialysis step, it was possible to purify the enzyme in a yield of 4 mg/L dissolved in an elution buffer containing 250 mM imidazole.

In general, the sugar molecule which is transferred by the glycosyltransferase comes from a sugar diphosphate nucleotide. In the literature, different nucleotides are reported as substrates for glycosyltransferase from family 1. Examples are the GDP-dependent phosphatidyl mannosyltransferase form *Mycobacteria tuberculosis* and the UDP-dependent HepE (heterocyst envelope polysaccharide) in *Cyanobacteria*.[210,211]

One method to identify which nucleotide is used in the GT_{SVA} is by determining the binding affinity of the enzyme. The commercially available adenosine diphosphate (ADP), guanine diphosphate (GDP) and uridine diphosphate (UDP) were used as substrates for a thermal shift experiment (Figure 111). When ADP or glucose were used, no shift of the melting temperature of GT_{SVA} was observed. With the substrate GDP, a shift of 4 °C was observed while with UDP a shift of 9 °C was observed. The native substrate for the glycosyltransferase GT_{SVA}, as well as for GT_{Vapar}, is probably a uridine diphosphate analogue. In addition, the binding properties of the D-glucose molecule was tested. No change in the melting temperature of the enzyme during the thermal shift experiment was observed. This result was expected since the substrate is a sugar bound nucleotide and not a single sugar molecule.

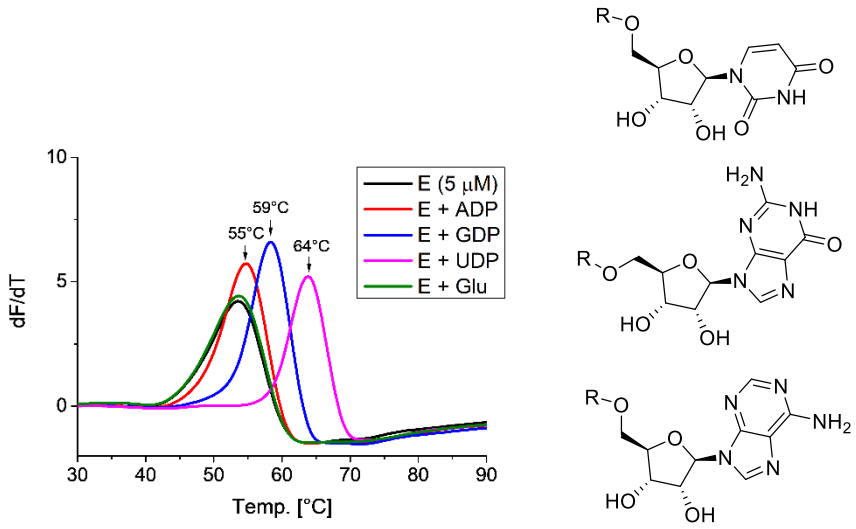


Figure 111. **Left:** Thermal shift assay with 5 µM GT_{SVA} and different substrates (black: control; red: ADP; blue GDP; purple: UDP; green: Glucose). **Right:** Chemical structures of the used substrates ADP, GDP and UDP.

Table 20: Melting temperature of GT_{SVA} with the different substrates.

Ligand	Melting temp.	Δ Melting temp.
Buffer	55°C	0°C
ADP	55°C	0°C
GDP	59°C	4°C
UDP	64°C	9°C
Glucose	55°C	0°C

4.3.9 Reaction with UDP-glucose and UDP-galactose

The determination which sugar molecule bound to the UDP nucleotide is used in the enzyme reaction is much more difficult. Since the nucleoside UDP had the best stabilizing effects towards the enzyme we purchased the two commercially available UDP-sugars, UDP-glucose and UDP-galactose. We repeated the thermal shift experiment and could observe that both UDP-sugars had a positive effect towards the enzyme melting temperature. In the case of UDP-glucose, a shift of the melting temperature of 8.4 °C was observed while the UDP-galactose had a shift of 6.4°. However, the UDP alone had a more stabilizing effect to the enzyme (T_m = +11 °C). These results were surprising since we would expect that the substrate of the glycosyltransferase would stabilize it better than the side product which is formed during the reaction. Nevertheless, in the bacteria *E. coli* the concentration of free UDP are almost two times lower than UDP-

glucose.[212] The higher accessibility of UDP-glucose inside of the cell could still compensate the slightly lower binding properties. UDP-glucose could be the right native substrate of GT_{SVA} and Vapar2053.

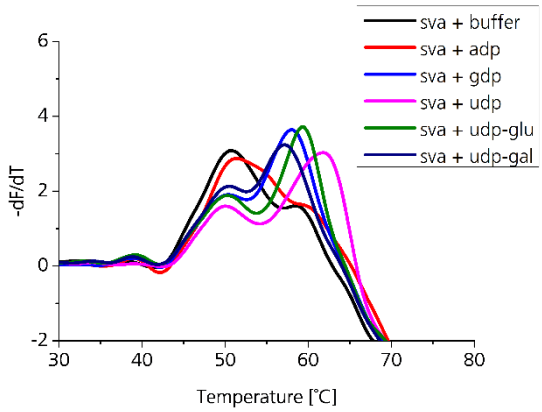


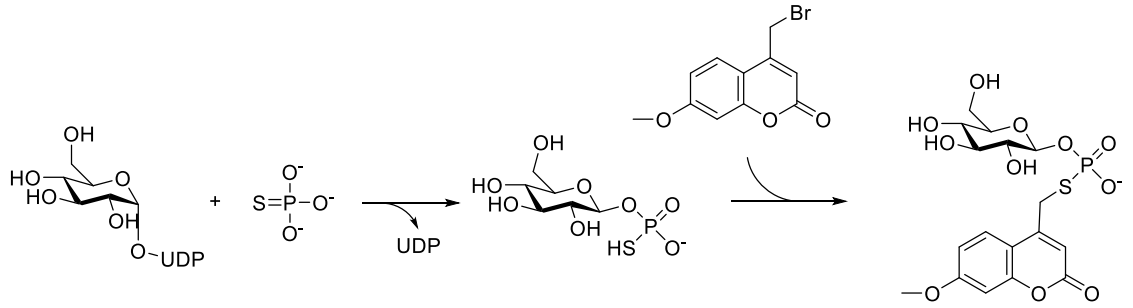
Figure 112. Thermal shift assay with 5 µM GT_{SVA} and different substrates (black: control; red: ADP; blue GDP; purple: UDP; green: UDP-glucose; dark blue: UDP-galactose).

Table 21. Melting temperature of GT_{SVA} with different substrates.

Ligand	Melting temp.	Δ Melting temp.
Buffer	50.8 °C	-
ADP	51.5 °C	0.7 °C
GDP	58.2 °C	7.4 °C
UDP	61.8 °C	11 °C
UDP-Glucose	59.2 °C	8.4 °C
UDP-Galactose	57.2 °C	6.4 °C

Based on the thermal shift experiment, UDP-glucose might be the right nucleoside-sugar substrate for the glycosyltransferase. The proposed second substrate is selenophosphate, which is a very reactive and air sensitive compound.[213] Therefore, the selenophosphate was replaced to the less reactive and commercialy available thiophosphate in the glycosyltransferase reaction. Since the thiosugar product is difficult to monitor it was labeled with the chromophore 4-bromomethyl-7-methoxycoumarin (BMC). This chromophore has been reported to efficiently label thiol groups and the adducts can be monitored at 330 nm (Scheme 37).[83] The reaction mixture containing the enzyme GT_{SVA}, UDP-glucose (1 mM) in 20 mM thiophosphate buffer (pH 7) was allowed to react at 25°C. This reaction mixture was quenched after 1, 2 and 12 hours by the addition of 40 µL acetonitrile and 10 µL of 20 mM BMC in DMSO. This reaction mixture was incubated for 30°C and then diluted in a 1:1 ratio with 0.1 %TFA. The coumarin adducts were analyzed by reverse phase HPLC at an absorbance of 330 nm. With this assay, it was not possible to identify a distinct product which was formed during the reaction. There are different elements of uncertainty which could

explain these results. First, it is not clear if the enzyme is active in the presence of 250 mM imidazole. Secondly, we only had evidence that UDP-glucose is the right substrate. And thirdly it is difficult to estimate how well the thio-sugar can be labeled with BMC. Furthermore, we proposed a seleno-compound as a substrate but, used a thiol-sugar in our assay.

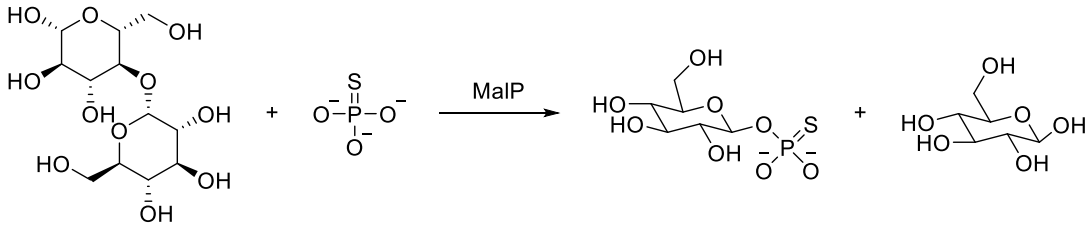


Scheme 37. The schematic attempt to observe thiosugar formation with the BMC-assay. At pre-defined time points the reaction mixture was quenched with a solution containing 5 mM BMC in a acetonitrile DMSO mixture (1:4). After an incubation time of 30 minutes, the samples were diluted 1:1 in 0.1% TFA and analyzed by reverse phase HPLC by an absorbance of 330 nm.

4.3.10 An Alternative Route to Form Thiophosphate-Glucose

Even though the right thiophosphate-glucose could not be observed as product with the glycosyltransferase, it could not be excluded as the right substrate for the $EgtB_{Vapar}$ enzyme. Therefore we searched for an alternative way to produce the thiosugar. GT_{SVA} belongs to the 1 family of GT, therefore it can be assumed that it operates by an inverting mechanism which results in a β-glycoside.

It has been shown that maltose phosphorylase (MalP) is able to form β-glucose-1-phosphate from mannose and phosphate.[214] Therefore, this enzyme might also be able to form the thiosugar if phosphate is replaced with thiophosphate (Scheme 38). The desired thiosugar could be monitored with ^{1}H-NMR and ^{31}P-NMR. It was then isolated by acetone induced enzyme precipitation followed by lyophilization.



Scheme 38. The formation of β-glucose-1-phosphate out of mannose and thiophosphate catalyzed by the enzyme MalP.

4.3.11 Substrate Binding Studies

To identify the second substrate for $EgtB_{Vapar}$, the binding affinity of possible ligands was measured with the thermal shift experiment. The thermal shift experiment was performed with *β*-glucose-1-thiophospha te, thiophosphate, phosphate and TMH as ligands. In addition, the combination of TMH with thiosugar, thiophosphate and phosphate was measured (Figure 113).

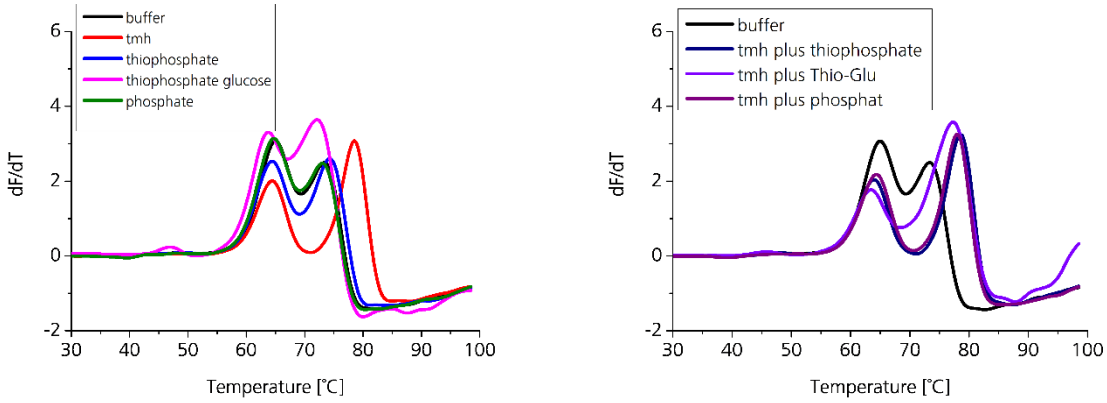


Figure 113. Melting curves of $EgtB_{Vapar}$ (2.5 µM) with buffer and different ligands (5 mM). For simplification the results of the same experiment is displayed in two figures. **Left:** The protein with one ligand. (black: control; red: TMH; blue thiophosphate; purple: thiophosphate glucose; green: phosphate) **Right:** The protein with a combination of TMH and a second ligand (black: control; dark blue: TMH plus thiophosphate; violet: TMH plus thiophosphate glucose; dark purple: TMH plus phosphate).

In this experiment, two maxima were observed after the derivative of the melting curve of $EgtB_{Vapar}$ with and without different ligands. While the first maxima of the measured curves was quite constant at a temperature around 64 °C, the second maxima was affected by the different ligands. A strong binding of TMH could be observed, which results in a stabilization of the melting temperature by 5 °C. The β-glucose-1-thiophosphate had no stabilizating effect. The ligand thiophosphate had a small shift of 0.9 °C. The combination of thiophosphate or *β*-glucose-1-thiophosphate with TMH could not increase the stability of the protein more than the ligand TMH by itself (Table 22). The control reaction with phosphate had only a minimal effect to the stability of $EgtB_{Vapar}$.

The observation that β-glucose-1-phosphate did not stabilize the enzyme would indicate that it is not a binding substrate. This result contrasts our proposed reaction pathway for the enzyme reaction of $EgtB_{Vapar}$, where a seleno-sugar molecule was expected to react with TMH. The change of the selenophosphate group to a thiophosphate group should have only a secondary effect to the binding properties of the ligand. Another explanation could be that instead of *β*-glucose-1-selenophospha te, another sugar molecule is the right substrate.

Interestingly, a small stabilizing shift with the ligand thiophosphate could be observed. This shift of 0.9 °C would indicate that an interaction of the ligand with the enzyme is taking place. This observation could point to an alternative route which is described in the next paragraph where selenophosphate is bound first and the glycosyltransferase is not involved in the carbon-selenium bond formation reaction of $EgtB_{Vapar}$. In combination with the second ligand TMH, the stabilizing effect is not additive. A possible explanation could be that after the binding of TMH, the enzyme $EgtB_{Vapar}$ is in the energetically most stable conformation and that the binding of the second ligand no longer a positive effect towards the stability.

Table 22. Melting temperatures of the $EgtB_{Vapar}$ with different ligands.

Ligand	Melting temp. 1	Melting temp. 2	Δ Melting temp. 1	Δ Melting temp. 2
Buffer	64.8 °C	73.5 °C	0 °C	0 °C
TMH	64.5 °C	78.5 °C	-0.3 °C	5 °C
Thiophosphate	64.5 °C	74.4 °C	-0.3 °C	0.9 °C
β -glucose-1-phosphate	63.8 °C	72.1 °C	-1 °C	-1.4 °C
Phosphate	64.8 °C	73.1 °C	0 °C	-0.4 °C
TMH + Thiophosphate	63.8 °C	78.5 °C	-1 °C	5 °C
TMH + β-glucose-1-phosphate	63.5 °C	77.2 °C	-1.3 °C	3.7 °C
TMH + Phosphate	64.5 °C	77.8 °C	-0.3 °C	4.3 °C

4.3.12 ICT Measurements with Thiophosphate

The observation of the small shift observed in the thermal shift assay with thiophosphate promoted us to repeat the ICT assay using thiophosphate as a ligand (Figure 114). In the first experiment, 4 mM thiophosphate was added to the buffer solution (50 mM NaCl and 50 mM Tris/HCl) containing 0.07 mM $EgtB_{Vapar}$. In a second experiment, we added 4 mM thiophosphate to an enzyme solution containing 1 mM of the proposed second substrate TMH. In a third experiment, we measured the addition of thiophosphate to a 1 mM TMH buffer solution without the enzyme. In the ITC measurements, we observed a strong background reaction, which lowers the informative values. Nevertheless, we were able to observe a distinct binding of thiophosphate with the enzyme in the first experiment. The calculated K_D value was rather low at 15 µM and is similar to the observed binding constants for substrates in other EgtBs. In the second experiment with enzyme and the proposed substrate TMH, any binding was barely observable, because of the high background. The calculated K_D increased to a value of 63 µM. In the third experiment, we did not see any binding and only dilution of the sample is visible. The thermodynamic parameters are summarized in Table 23. These measurements show a distinct binding of thiophosphate to the enzyme and let us propose an alternative route for the biosynthesis of selenoneine.

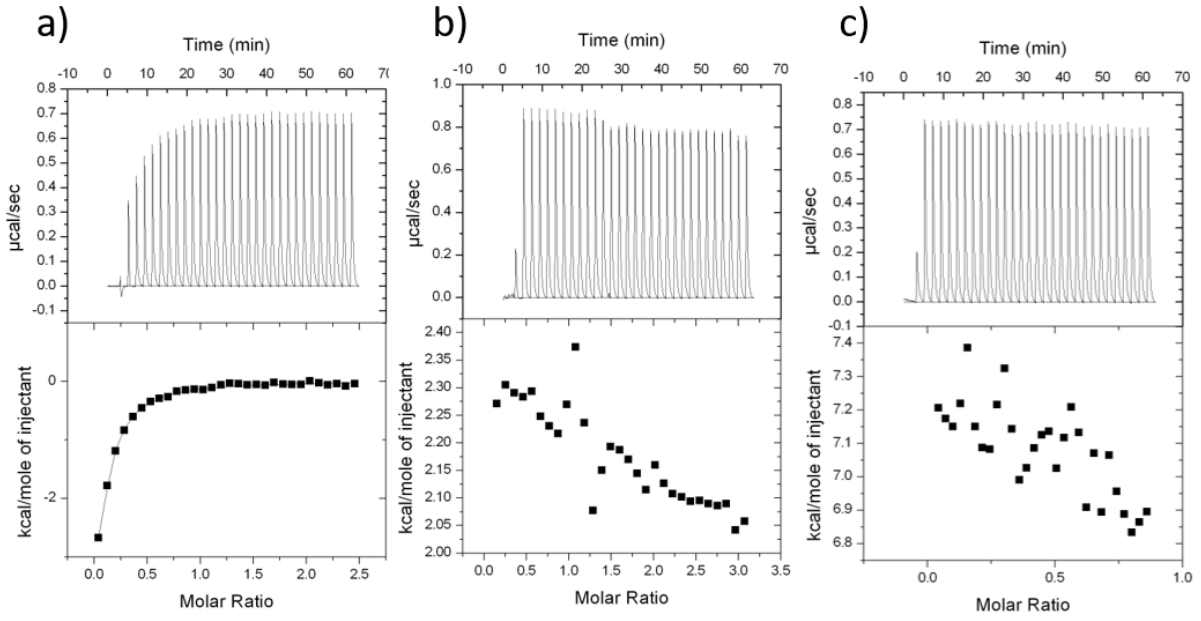


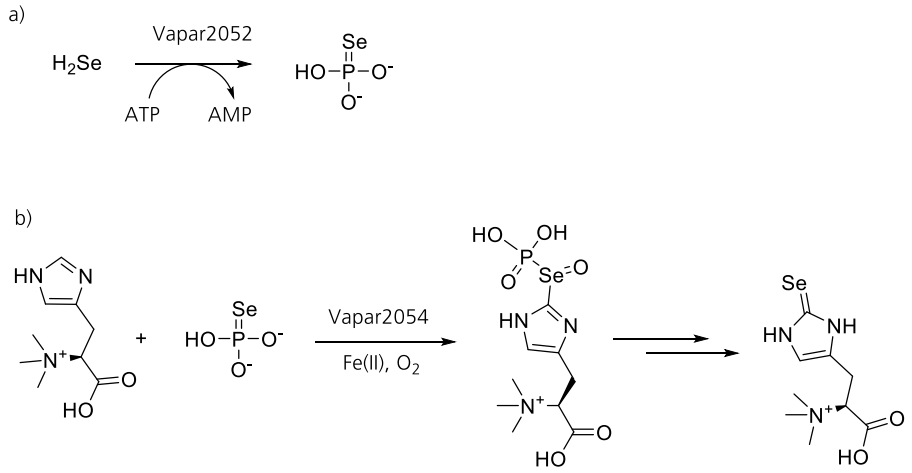
Figure 114. ITC measurements with EgtBvapar and Thiophosphate **a)** To70 µM enzyme in 50 mM HCl/Tris and 50 mM NaCl buffer the ligand solution (4 mM thiophosphate) was added in 2 µl portions. **b)** The enzyme solution contained an additional 1 mM TMH. The ligand (4 mM thiophosphate) was added in 2 µl portions. **c)** Control experiment with 1 mM TMH and 4 mM thiophosphate without enzyme.

Table 23. Thermodynamic parameters of the ITC experiments with the ligand thiophosphate (4 mM) added to a buffer solution containing enzyme (0.07 mM), enzyme (0.07 mM) plus TMH (1 mM) and only TMH (1 mM).

Enzyme	Ligand	K_A [M^{-1}]	K_D[M]	ΔH [kcal/mol]	ΔS [cal mol-1**K-1**)
$EgtB_{Vapar}$	Thiophosphate	(6.5 ± 0.2) x 10^4	15 ± 1	- 7.5	- 3.3
$EgtB_{Vapar}$ + 1 mM TMH	Thiophosphate	(1.6 ± 1.2) x 10^4	63 ± 2	+ 0.5	+ 21
1 mM TMH	Thiophosphate	n.a.	n.a.	n.a.	n.a.

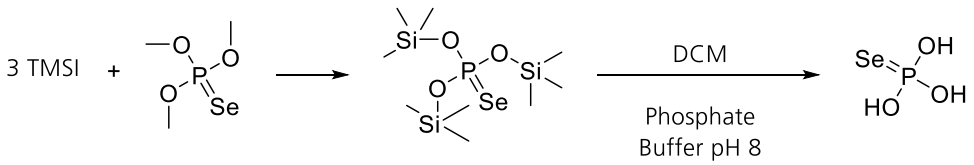
4.3.13 Alternative Routes

Since the proposed β-glucose-1-thiophosphate had no positive binding effect in the thermal shift assay to the $EgtB_{Vapar}$ but thiophosphate did, an alternative reaction pathway was proposed. A possible alternative route for the enzyme reaction of $EgtB_{Vapar}$ could be to form directly the selenoxide from the substrates trimethylhistidine and selenophosphate. The selenoxide might then spontaneously decompose to the final product selenoneine (Scheme 39).



Scheme 39. An alternative proposed biosynthesis catalyzed by the selenokinase Vapar2052 and the $EgtB_{Vapar}$ for the formation of selenoneine.

We decided to test this reaction under the standard EgtB reaction conditions containing an increased amount of reducing agent (5 mM TCEP) and increased enzyme concentration (10 µM enzyme). We increased the concentration of reducing agent because the possible substrate selenophosphate is sensitive towards oxidation.[213] The substrate concentrations were 1 mM TMH and 1 mM freshly prepared selenophosphate. Selenophosphate was prepared in our laboratory by hydrolysis of the stable tris(trimethylsilyl) selenophosphate an under oxygen-free atmosphere (Scheme 40).[213] The selenophosphate was then directly used in the enzyme assay. The reaction was quenched after 15 minutes and after 12 hours with 1 M phosphoric acid and analyzed with the cation exchange HPLC.



Scheme 40. The last two steps of the synthesis of monoselenophosphate. In the first step, the protecting groups of trimethyl-selenophosphate was replaced by trimethylsilane. This stable intermediate can easily be stored and freshly hydrolysed under oxygen-free atmosphere to form selenophosphate.[213]

In all measurements analyzed on the cation exchange HPLC we were not able to identify any new peak formation or TMH consumption. It was also not possible to identify a distinct selenoneine formation inside the reaction. The same reaction was repeated in addition with thiophosphate (1 mM), TMH (1 mM) and enzyme. Also, in this measurements we were not able to observe any sulfoxide formation.

4.3.14 Crystallization

To structurally characterize $EgtB_{Vapar}$ and to elucidate the architecture of the binding site of the second substrate, we started crystallization trials. The protein $EgtB_{Vapar}$ was easily producible in a good purity. After the cleavage of the His_6-Tag, the protein was purified with size exclusion chromatography. In the next step conditions for protein crystallization were determined. 96-Well 3 drop Swissic plates containing commercially available crystal screens were used with a protein concentration of 10 mg/mL. Every screen contained 96 different conditions with varying pH, precipitant and additives. For every condition, three different protein to crystal screen ratios were used. The protein was used freshly produced as well as previously shock frozen, and stored by -80 °C. Under the multiple conditions used in the primary trials, no crystals could be observed, and only protein precipitation and random protein aggregation was detected.
A strategy to improve crystallization is to analyze the secondary structure of a protein and remove disordered amino acid segments. These disorder amino acid regions could prevent the crystallization of the protein due to the high flexibility and poor ability to form an ordered alignment. Therefore, the secondary structure of $EgtB_{Vapar}$ was analyzed with Phyre[2] an online protein fold recognition server (http://www.sbg.bio.ic.ac.uk/servers/phyre2/html/page.cgi?id=index). As shown in Figure 115, the first 23 amino acids could be identified as disordered and a new protein construct was designed. For this construct, the amino acids 2-21 were deleted and a truncated protein was formed. The new plasmid was then transformed, and the truncated protein was produced *E. coli* in BL21 cells. The production of this new construct was not successful, and it was not possible to isolate a pure protein. The solubility of the new enzyme construct decreased tremendously.

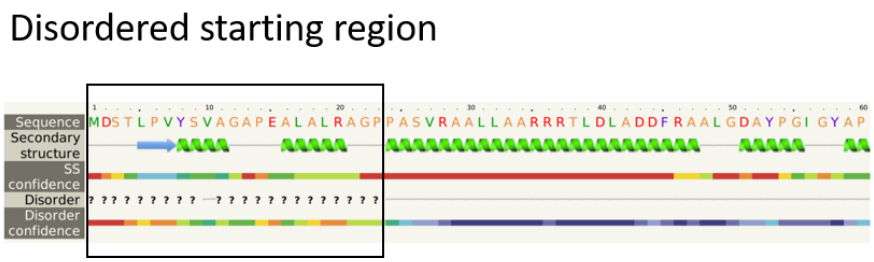


Figure 115. The analysis of the starting region of the enzyme $EgtB_{Vapar}$ by the online server Phyre2 showing that the first 23 amino acid are disordered.

Since the attempt to use the truncated enzyme did not work because the enzyme was not soluble anymore we changed back to the full length enzyme. In the next trial the temperature during the crystallization was changed from 20 ˚C to 4 ˚C. In addition, the protein concentration was increased from 10 mg/mL to 20 mg/mL. Again, only precipitation and aggregation occured under the used conditions.

Finally, in the last trial, the protein concentration was increased to 31 mg/mL and changed back to 20 ˚C. Under this condition, in certain drops crystal formation could be observed in the MorpheusHT screen, as well as in the PEG Ion Screen (Figure 116). The crystal from Morpheus D12 (37.5% m/v M1K3350 (Precipitant), 0.1 M MB3 8.5 pH (buffer) and 10% alcohols (complex ingredients) was flash frozen in liquid nitrogen and stored until data collection. The crystal was measured at the Swiss Light Source (SLS), Villigen, Switzerland and the crystal structure was successfully solved by Anja Stampfli. The resolution of the apo crystal structure from $EgtB_{Vapar}$ was 2.7 Å.

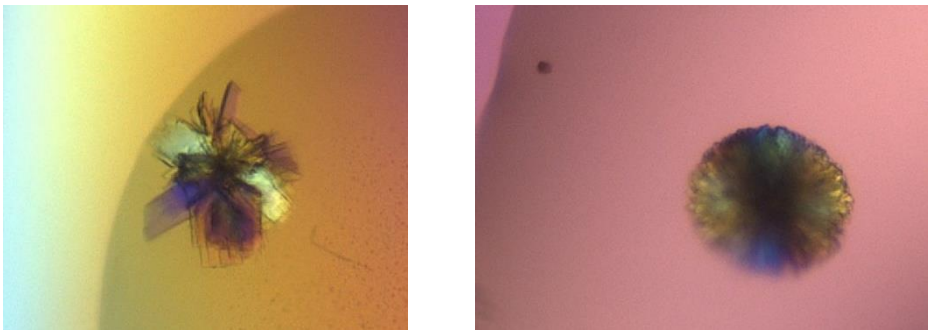


Figure 116. Protein crystals observed in the Morpheus HT screen A12 and E12.

4.3.15 The $EgtB_{VaparS112A}$ Mutant

In the crystal structure of the $EgtB_{Vapar}$, we identified the S112 residue, which is located in the active site pointing towards the binding site of a potential substrate. In the crystal structure of $EgtB_{thermo1}$ the A82 residue is found at a very similar position. We assumed that the purpose of S112 is to form a hydrogen bond towards the unknown substrate.

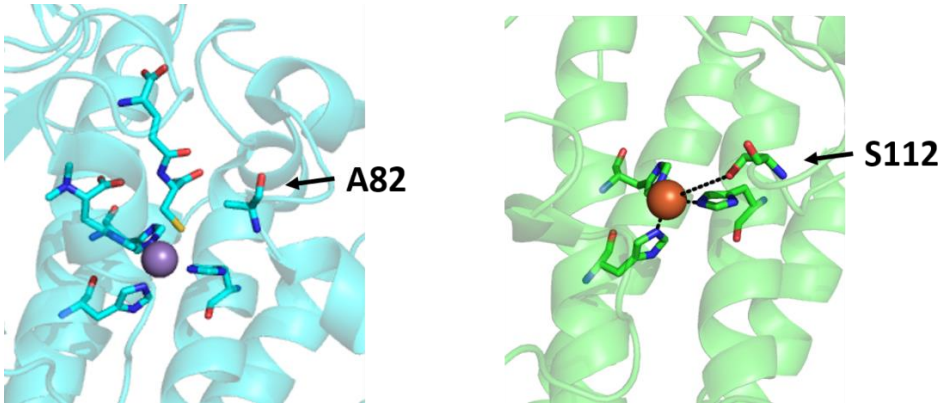


Figure 117. **Left:** Crystal structure of $EgtB_{M.thermoresistibile}$ with the highlighted alanine A82 residue (PDB: 4x8d, x) **Right:** Crystal structure of $EgtB_{Vapar}$ with the highlighted serine S112 residue (unpublished, 2.7 Å).

The effects of hydrogen bonds to iron-sulfur complexes have been studied in many enzymes and model systems. Hydrogen bonds change the redox potential of Fe-S clusters in proteins as well as artificial complexes.[215] In the example of an artificial iron-sulfur cubane complexes [4Fe-4S] the addition of an OH···S bond raised the oxidation potential of the complex, which was measured with photoelectron spectroscopy by ~130 mV.[216] In the enzyme superoxide reductase from *Desulfoarculus baarsii*, it was shown that hydrogen bonds towards the cysteine ligand are involved in tuning the strength of the S-Fe bond.[217] Another example where the effects of hydrogen bonds towards S-Fe bonds was investigated in P450 model complexes. With sulfur-K-edge X-ray absorption spectroscopy (XAS) and DFT calculation it was shown that the covalency of the metal-ligand was reduced by a H-bonding to the sulfur atom. In addition an increased reduction potential was observed.[16]

To test the influence of the hydrogen bond in the EgtB enzyme the residue A82 of $EgtB_{thermo1}$ was replaced with a serine. It could be shown in our laboratory that a hydrogen bound between the iron-bond γ-glutamyl cysteine and this serine is most likely formed.[74] The sulfoxide formation activity of this new enzyme mutant was reduced by 5 fold. A possible explanation for this observation could be that the hydrogen bond reduces the F-Se covalency thereby stabilizing the negative charge on the sulfur atom. This stabilization then hindered the formation of the proposed thiyl radical which is a central step in the proposed mechanism.[60,74] Nevertheless, if the substrate was changed in $EgtB_{thermo1}$ from γ-glutamylcysteine to γ-glutamylselenocysteine, no decrease of the activity was observed. For seleno-compounds, this hydrogen bond might have positive effects for activation, by delocalizing the increased electron density of selenium and thereby make it more similar to sulfur.

In the search of a potential substrate, we predominantly used thiol compounds, instead of seleno compounds, since they are less sensitive towards autooxidation.[213] As seen in the enzyme $EgtB_{thermo1}$, a

hydrogen bond decreases the activity with the thiol substrate.[74] Therefore, we speculated that the removal of the hydrogen bond formed by S112 could favor the reaction with thio-compounds like thiophosphate. The mutant $EgtB_{Vapar\ S112A}$ was constructed and produced.

With the mutant $EgtB_{Vapar\ S112A}$, we tested the reaction with TMH and thiophosphate. The reaction contained the standard conditions with 2 mM ascorbate, 2 mM TCEP, 10 µM $FeSO_4$ and 2 µM enzyme. The reaction was monitored by cation exchange HPLC method used for the eukaryotic EgtB characterization in chapter 3. We were unable to detect product formation. Although we chose the wrong substrate, we cannot exclude that this serine residue has other functionality and that the mutant of this enzyme might be not active anymore.

4.4 Conclusion and Outlook

A new EgtB analogue in the organism *Variovorax paradoxus* was identified which belongs to a distinct group of EgtB homologues. The identification of a selenokinase and a glycosyltransferase in a conserved gene cluster and the unknown biosynthesis of the nature product selenoneine in bacteria encouraged us to speculate that this group of enzymes might be involved in its biosynthesis.

Supported by amino acid sequence analysis, thermal shift assays and ITC measurements we were able to provide strong evidence that TMH is one of the native substrates. The low measured K_D value of 2.5 µM for TMH is a further argument for this finding. The second substrate, which might be a selenophosphate or a selenophosphate coupled to a sugar moiety, could not be identified. An explanation for the unsuccessful identification of the second substrate could be that a different UDP-nucleotide is required. We only examined the nucleotide UDP-glucose and UDP-galactose. Another possibility could be UDP-N-acetylglucosamine or UDP-glucuronate which are also found as metabolites in bacteria.[218]

The observation of that thiophosphate binds as ligand in ITC measurements let us propose an alternative biosynthetic route for selenoneine. Nevertheless, it was not possible to observe any product formation with either selenophosphate nor thiophosphate.

In addition to attempts to functionally characterize $EgtB_{Vapar}$ we were able to structurally characterize it, by obtaining an apo crystal structure of the enzyme at a resolution of 2.7 Å.

For future work, it would be interesting to co-crystalize the enzyme with TMH, thiophosphate and selenophosphate to determine if it is possible to obtain a ligand-enzyme crystal complex. Furthermore, the identification of the right substrate for the conserved glycosyltransferase would help to elucidate the biosynthetic pathway in which $EgtB_{Vapar}$ is involved. A further strategy to find out the second substrate might be to use molecular modeling approaches with the solved crystal structure.

4.5 Experimental

Plasmid Construction:

The gene for protein $EgtB_{Vapar}$, and GT_{SVA} were ordered codon optimized for production in *Escherichia coli* from GenScript® and delivered in a pUC57 plasmid. The genes were ligated into the expression vector pET19 and pEt28 using the restriction enzymes NdeI/XhoI. The sequence of the plasmid was verified by sequence analysis of Microsynth AG. MalP was ordered by General Biosystems and delivered in a pET28a vector, which was directly used for protein production.

$EgtB_{Vapar}$

MGHHHHHHAENLYFQGHMDSTLPVYSVAGAPEALALRAGPPASVRAALLAARRRTLDLADDFRAALGDAYPGIGY
APELNPPLWELGHVAWFQEWWIGRNRQRARGVACEPDHAREPSLLPQADAWYDSSRVAHRTRWALPLPDAEATRG
YLERTFAQTLALLDELPPDAHDDALYFFRLVALHEAMHAEAAAYMAEGLGISLREGGPGPQLAEDAELELPARRF
RIGSEAGAGFAFDNELLPHDVAIGPLRIDAQAVSWARFLPFVEAGGYENPAWWSDAGRGWLARQPLRQPACLRAA
GTGWQQQRGGRWLPLGPQAAAVHLNAHEAEAWCRWAGRRLPTEAEWECAALTLPGFAWGRVWEWTSSPFEPYPGF
APHPYRDYSAPWFGTRRVLRGACHATSAALAHARYRNFFEPHRRDIFAGFRSCRAAGG

$EgtB_{Vapar\_S112A}$

MGHHHHHHAENLYFQGHMDSTLPVYSVAGAPEALALRAGPPASVRAALLAARRRTLDLADDFRAALGDAYPGIGY
APELNPPLWELGHVAWFQEWWIGRNRQRARGVACEPDHAREPSLLPQADAWYD**A**SRVAHRTRWALPLPDAEATRG
YLERTFAQTLALLDELPPDAHDDALYFFRLVALHEAMHAEAAAYMAEGLGISLREGGPGPQLAEDAELELPARRF
RIGSEAGAGFAFDNELLPHDVAIGPLRIDAQAVSWARFLPFVEAGGYENPAWWSDAGRGWLARQPLRQPACLRAA
GTGWQQQRGGRWLPLGPQAAAVHLNAHEAEAWCRWAGRRLPTEAEWECAALTLPGFAWGRVWEWTSSPFEPYPGF
APHPYRDYSAPWFGTRRVLRGACHATSAALAHARYRNFFEPHRRDIFAGFRSCRAAGG

$EgtB_{Vapar\_truncated}$

MGHHHHHXAENLYFQGHMGPPASVRAALLAARRRTLDLADDFRAALGDAYPGIGYAPELNPPLWELGHVAWFQEW
WIGRNRQRARGVACEPDHAREPSLLPQADAWYDSSRVAHRTRWALPLPDAEATRGYLERTFAQTLALLDELPPDA
HDDALYFFRLVALHEAMHAEAAAYMAEGLGISLREGGPGPQLAEDAELELPARRFRIGSEAGAGFAFDNELLPHD
VAIGPLRIDAQAVSWARFLPFVEAGGYENPAWWSDAGRGWLARQPLRQPACLRAAGTGWQQQRGGRWLPLGPQAA
AVHLNAHEAEAWCRWAGRRLPTEAEWECAALTLPGFAWGRVWEWTSSPFEPYPGFAPHPYRDYSAPWFGTRRVLR
GACHATSAALAHARYRNFFEPHRRDIFAGFRSCRAAGG

GT_{SVA}

MGSSHHHHHHSSGLVPRGSHMRIVLITPAKRGSKAGNRATAVRWARHLRALGHRIDVAVDYDGRPADLMVALHAW
RSAAAVERFRERYPGRPLILALTGTDLYRFLATDPEVTVRSMALADRLVVLHDLAHEALDPSLRHKVRVIHQSAL
PLARHRRPPTRYFDVCVIGHLREEKDPFRAAYAARGLPAGSRIRIVHLGKGHGPEWEDEARDEMRRNPRYLWRGE
VPFARVRETMARSRLMVISSRMEGGANVVSEACVARLPVIASDIPGNVGLLGPAYPGYYPLEDSAALATLLLRAE
RDPAFYRALKQACTARAPLFTPARERRAWKSLLAEVTARESEDLEPPGRQGRQGVTEIPDP

MalP

MKRIFEIDPWKVITHKFDPKDKRLQESMTAIGNDYMGMRGNFEEGYSGDSLQGTYLAGVWFPDKTVVGWWKNGYP
KYFGKTPNAPSFIGIGINVNGEKVDLAKVKFSDFELSLDMHQGLLSRSFIYEGKDVKVKLEFERFLHIVQKEAAL
IKVKATVLEGHAKIDFDSTLDGTVVNEDSNYGDRFWIPLGEDKDEKTIQVKTKKNPYDVPQFTVLLKEALRNNGV
AVNGEVTTEDAKLSERFSVELDEGQSYELEKDVIVVTSRDVEEKDQAAVANNLMSKLQTKSFEDNLADHTEAWKK
RWETSDVEISGDDAAQQGIRFNICQLFMTYYGEDKRLNVGPKGFTGEKYGGATYWDTEAFIVPMYLAVTKPSVTR

ALLQYRHDQLPGAYHNAKEQGLPGALFPMVTFNGIECHNEWEITFEEIHRNADIPHAIAMYTDYTGDDSYVKNEG
MDVLVGTARFWAARVHWSKMRNKYVMHGVTGPNEYENNVNNNWFTNTMARWLLKYTLERLPLATKEAQERVRVTD
EEKAKWQDIVDNMYLPEDEDLGIFLQQDDFLDKDIRPVTEIEDQRPINQHWSWDKILRSPFIKQADVLQGIYFFD
DQYTMDQKEKNFDFYEPLTVHESSLSPCIYSIMAAELGKKEKAVELYQRTARLDLDNYNNDTVDGLHITSMSGSW
LAIVQGFAGMRYDHDQLKFNPFVPDGWDHYSFKINYRGRLIEVYVDHDECKITLLSGDDLEVMVHDNKLDLKEGK
TKCLKA

The two $EgtB_{Vapar}$ variants $EgtB_{Vapar\_S112A}$ and $EgtB_{Vapar\_truncated}$ were constructed using site directed mutagenesis using the pET19- $EgtB_{Vapar}$ as a template and synthetic primers from Microsynth AG. The PCR fragments as well as the target pET19 vector were digested with the restriction enzymes XhoI and NdeI. The fragments were ligated into the target vector using T4 Ligase and the final construct was confirmed by sequencing analysis (Microsynth AG).

Table 24: Primer list used for the cloning of the $EgtB_{Vapar}$ variants.

Vapar G22	5'-tatatcatatgGGTCCGCCGGCGTCAG-3'
Vapar2054 reverse	5'-TaattCTCGAGTTAGCCACCTGCGGCACG-3'
Vapar S112Aa	5'-CAC ACG AGA GGC GTC ATA CCA-3'
Vapar S112As	5'-CCT GGT ATG ACG CCT CTC GTG-3'

Protein Production

EgtBVapar/MalP

The expression vector pET19 containing $EgtB_{Vapar}$ or the expression vector pET28 containing MalP enzyme were transferred into BL21 (DE3) pLys. The cells were grown in lysogeny broth containing ampicillin (100 mg/L) and chloramphenicol (34 mg/L) at 37°C for pET19 vector. For the pET28 vector the antibiotics chloramphenicol (34 mg/ml) and kanamycin (50 mg/mL) were used. After the OD_{600} of the cell culture reached 0.8 the cultures were cooled down to 18°C and then induced by adding 0.1 mg/mL isopropyl-β-D-thiogalactopyranosid (IPTG). The protein was expressed for 16 hours at 18 °C, harvested by centrifugation (7000 rpm) and stored at -20 °C until protein purification. For the purification, the cells pellet was resuspended into lysis buffer (300 mM NaCl, 50 mM Na_2HPO_4) and lysed by an Avestin EmulsiFelx-C3 homogenizer. To remove the insoluble cell components the mixture was centrifuged for one hour at 4000 rpm at 4 °C. Ni^{2+} NTA agarose beads (Qiagen) were added and after binding of the enzyme for 20 minutes at 4 °C. The Ni^{2+} column was washed with 20 mM imidazole and afterwards the protein was eluted with 250 mM imidazole. To remove Ni^{2+} ions from the purification step the protein was incubated for 20 minutes at 0 °C with 10 mM EDTA. After two round (minimum 4 hours) of dialysis in 50 mM Tris/HCl pH 8 and 50 mM NaCl buffer the protein was flash frozen and stored for further use.

The GT_{SVA} was produced with the expression vector pET28 in BL21 containing additionally to the pLys plasmid the chaperon pL1SL2 plasmid. The cells were grown in lysogeny broth containing ampicillin (100 mg/L) and chloramphenicol (34 mg/L) and kanamycin (50 mg/mL) at 37 °C. The production and purification were performed with the protocol described before. After purification, the dialysis steps were skipped, and the protein was stored in the fridge for further use.

Table 25. Calculated and observed HRMS mass of the different enzymes.

Protein	Calculated Mass [Da]	Found Mass [Da]
$EgtB_{Vapar}$: pET28 pET19-$_6$His-Tag	48096.0 46258.1	48094.2 46257.0
$EgtB_{Vapar\_truncated}$	46020.6	n.a.
$EgtB_{Vapar,\ S112A}$	47987.9	47987.1
MalP	89155.2	n.a.
GT_{SVA}	40219.1	40218.5

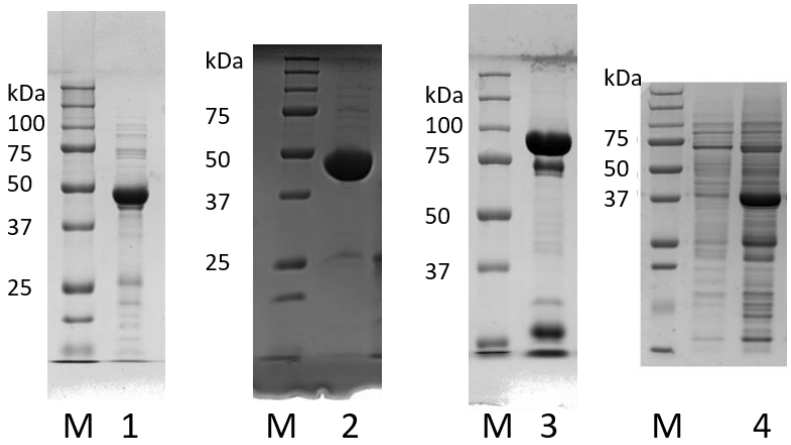


Figure 118: SDS page of the purified enzymes: M: molecular weight marker, 1: $EgtB_{Vapar}$, 2. $EgtB_{Vapar\ S112A}$, 3 MalP and 4 GT_{SVA}.

MalP NMR

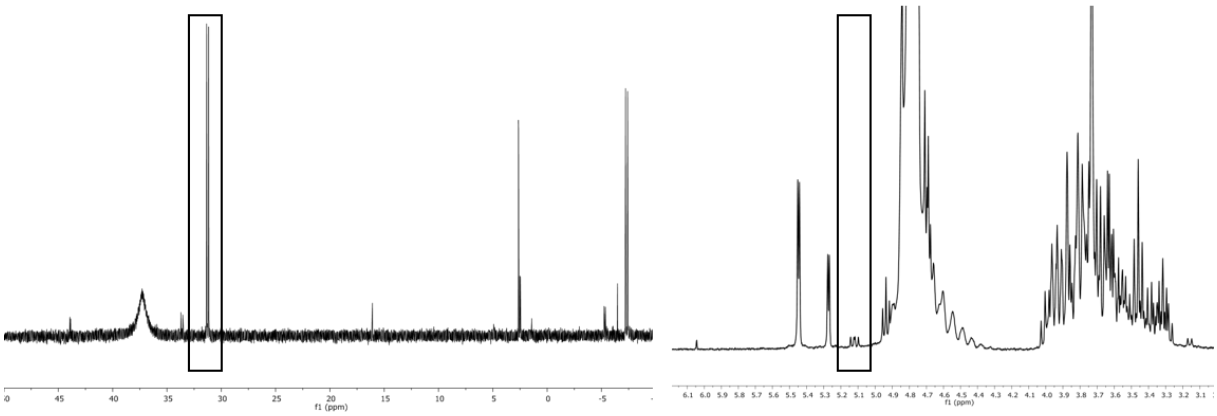


Figure 119. NMR oft the crude mixture of β-glucose-1-thiophosphate produced by the reaction of thiophosphate and mannose catalyzed with MalP. Left: ^{31}P-NMR the new formed phosphate peak is highlighted in the box. Right: ^{1}H-NMR theβ-protons of a new formed glucose species is highlighted in the box.

Isothermal titration calorimetry (ITC): TMH measurement

TMH was dissolved in Tris/HCl (50 mM, pH 8) and NaCl (100 mM) to a final concentration of 1 mM. Frozen $EgtB_{Vapar}$ protein was thawed and diluted in the same buffer (Tris/HCl (50 mM, pH 8) and NaCl (100 mM)) to a final concentration of 0.086 mM. The sample cell of an isothermal titration calorimeter (VP200-ITC system, MicroCal LLC) was then loaded with the protein solution. The TMH solution was added to the protein solution in 2 mL steps by syringe/stirrer every 2 min. The measured data were analyzed by the program Origin7.

Fluorescence based thermal Shift Assay:

Sample preparation: The protein stock solution was mixed with the SYPRO Orange dye (5000x) and buffer (50 mM HCl/Tris buffer (pH8) supplemented with 50 mM NaCl) to a concentration of 5-10 µM enzyme and 2x SYPRO Orange dye. The solution was mixed 1:1 with buffer (control) or ligand solution (10 mM). The final concentration of the reaction mixtures was 2.5-5 µM enzyme and 1x SYPRO Orange and 5 mM ligand. All reactions were performed as triplicate.

The thermal shift fluorescence measurements were performed on a Rotor-Gene Q, real time PCR from QIAGEN with a constant temperature gradient (30 to 100°C at 5°C/min).

Thiosugar/BMC-Assay

The reaction mixture containing 1-5 µM GT_{SVA}, 1 mM UDP-Glu and in a 1 mM thiophosphate in 50 mM Tris/HCl buffer and 20 mM NaCl was incubated overnight at room temperature. In the next morning 40 µL of the reaction mixture was added to a 50 µL of a 2 mM BMC solution (DMSO:MeCN 1:4). After 30 minutes the samples were diluted with 90 µL 0.1% TFA and quantified by RP-HPLC (at 330 nm).

HPLC-Assay

A cation-exchange HPLC containing a Luna 5u SCX column (100 Å, 150 x 4 mm, Phenomenex) was used for the reaction mixture analysis. The mobile phase (solvent A) was 20 mM phosphoric acid pH 2 and the product were eluted with a gradient of solvent B (1 M NaCl, 20 mM phosphoric acid, pH 2). The gradient was defined according to the following method and the product formation was monitored at an absorption of 265 nm.

Table 26: cation-exchange standard HPLC method used for the EgtB reactions. Solvent A (20 mM phosphoric acid pH 2) Solvent B (1 M NaCl, 20 mM phosphoric acid, pH 2).

Time (min)	Solvent A (%)	Solvent B (%)
0.01	85	15
2	80	20
8	1	99
9	1	99
11	85	15

RP-HPLC

Table 27: reversed-phase HPLC method used for the analysis of the BMC assay. Solvent A (H_2O, Acetonitrile 1 %, TFA 0.1%). Solvent B (Acetonitrile, TFA 1 %).

Time (min)	Solvent A (%)	Solvent B (%)
0.01	5	95
10	35	65
25	95	5
28	95	5
35	5	95

FPLC-Purification

The $EgtB_{Vapar}$ were purified with size exclusion chromatography for the crystallization. The size exclusion protein purification was performed on an Äkta Prime with a HiLoad Superdex 26/600 column. The flow rate was 2 ml/min in an elution buffer containing 50 mM HCl/Tris pH 8 and 200 mM NaCl.

Crystallization

The crystallization condition for $EgtB_{Vapar}$ were determined with the sitting drop 96-well plates. The dispensing robot (gryphon) was set up the crystallization drops containing 0.2 µl enzyme solution and 0.2 µl EgtBVapar (31 mg/mL). The screens were analyzed with the automatic system Rockimager at a constant temperature of 20°C. The final conditions for the crystal structure were the Morpheus HT screen A12 and E12. The harvesting of the crystallization, the data collection and solving of the crystal structure were made by Anja Stampfli.

5 Final Conclusion and Outlook

This thesis describes the characterization of four different enzymes, all of which are involved in metal-dependent sulfur activation reactions.

The first examined enzyme, EF_3021, is a glutathione-*S*-transferase from the pathogenic organism *Enterococcus faecalis*. The development of a model system with 4-Nitrobenzylchloride enabled the kinetic analysis of this enzyme. It was shown that this enzyme is active with a variety of different divalent transition metals. The reaction rates are only minimally dependent on the central metal atom. The exception being Ni^{2+}, which showed higher activity. These results contrast previously reported alkylation studies with artificial sulfur-metal complexes.[1] Crystal structures with various metals and glutathione were obtained and the architecture of the ternary complex with glutathione was identified to be tetrahedral. A combination of structural data with the measured kinetic results and density functional theory calculations would be required for a complete description of the sulfur activation in EF_3021. This would also make it possible to compare the enzymatic metal-thiol complex with other reported data.

The second characterized enzyme was the EgtB/EgtD fusion protein from the thermophilic fungi *Chaetomium thermophilum*. This enzyme is involved in the first and second step of eukaryotic ergothioneine biosynthesis, catalyzing the SAM-dependent trimethylation of histidine as well as the sulfoxide formation between trimethylhistidine and cysteine. The kinetic characterization of the C-S bond formation step showed that the activity and enzymatic efficiency of those enzymes are similar to these for bacterial EgtBs. The sequence alignment and homology model for the fungal EgtB did however show that the bacterial cysteine binding site is not conserved, therefore an alternative binding mode must be used. An additional contrast to bacterial enzymes is that selenocysteine is a substrate for the *in vitro* assay. A possible explanation could be a different oxygen activation mechanism. For further studies to probe this oxygen activation a crystal structure would be useful.

In the *Bacidomycetes Dichomitus squalens* a second distinct eukaryotic EgtB homologue was identified. The lack of the catalytically relevant tyrosine indicates a different function of this enzyme. It was possible to identify cysteine dioxygenase activity as previously observed in a EgtB mutant lacking the same tyrosine residue.[2] With the discovery of this EgtB homologue a new cysteine dioxygenase class was identified. The requirement of TMH as an activator connects the ergothioneine biosynthesis to the production of cysteine sulfinic acid. This connection could be a lead to a yet undiscovered function of the cysteine sulfinic acid in fungal organisms. Further experiments are needed for the kinetic characterization of this fungal cysteine dioxygenase.

Finally, an unknown EgtB homologue $EgtB_{Vapar}$ from the organism *Variovorax paradoxus* was examined. The observation of a conserved gene cluster containing a glycosyltransferase and a selenokinase was connected to a possible selenoneine biosynthesis pathway in bacteria. One of the substrates was identified to be TMH. ITC experiments with thiophosphate showed an active binding to the enzyme and thus selenophosphate was hypothesized to be a possible second substrate. Nevertheless, it was not possible to observe product formation with TMH and selenophosphate in the *in vitro* assay. In addition, crystallization conditions for a successfully solved crystal structure were discovered and optimized. Future experiments would be to co-crystalize the enzyme with thiophosphate and trimethylhistidine and to identify wheter the thiophosphate binds to the central metal ion.

The common denominator in all of the four described enzymes is the metal binding DinB domain. This domain provides a platform for the sulfur activation *via* a metal ion, whether it is an ionic alkyltransfer reactions or a more complex radical mechanism. The dimerization or fusion to another fold of this helical domain diversifies the chemistry that can be catalyzed. The chemical versatility of the activated sulfur and the immense possibility for combinations of the DinB domain with a second domain makes the DinB structure an extremely and powerful tool in enzymatic sulfur chemistry and especially in the formation of C-S bonds.

6 References

1. Stipanuk, M. H. SULFUR AMINO ACID METABOLISM: Pathways for Production and Removal of Homocysteine and Cysteine. *Annu. Rev. Nutr.* **24,** 539–577 (2004).
2. Myers, L. C., Verdine, G. L. & Wagner, G. Solution Structure of the DNA Methyl Phosphotriester Repair Domain of Escherichia coli Ada. *Biochemistry* **32,** 14089–14094 (1993).
3. Meury, M., Knop, M. & Seebeck, F. P. Structural Basis for Copper–Oxygen Mediated C–H Bond Activation by the Formylglycine-Generating Enzyme. *Angew. Chemie - Int. Ed.* **56,** 8115–8119 (2017).
4. Bragulla, H. H. & Homberger, D. G. Structure and functions of keratin proteins in simple, stratified, keratinized and cornified epithelia. *J. Anat.* **214,** 516–559 (2009).
5. Shoulders, M. D. & Raines, R. T. Collagen Structure and Stability. *Annu. Rev. Biochem.* **78,** 929–958 (2009).
6. Awoonor-Williams, E. & Rowley, C. N. Evaluation of Methods for the Calculation of the pK_a of Cysteine Residues in Proteins. *J. Chem. Theory Comput.* **12,** 4662–4673 (2016).
7. Pinitglang, S. Watts, A. B. Patel, M. & Brocklehurst K.. A classical enzyme active center motif lacks catalytic competence until modulated electrostatically. *Biochemistry* **36,** 9968–9982 (1997).
8. Jensen, K. S., Pedersen, J. T., Winther, J. R. & Teilum, K. The pKa value and accessibility of cysteine residues are key determinants for protein substrate discrimination by glutaredoxin. *Biochemistry* **53,** 2533–2540 (2014).
9. McGrath, M. E. The lysosomal cysteine proteases. *Annu. Rev. Biophys. Biomol. Struct.* **28,** 181–204 (1999).
10. Verma, S., Dixit, R. & Pandey, K. C. Cysteine proteases: Modes of activation and future prospects as pharmacological targets. *Front. Pharmacol.* **7,** 1–12 (2016).
11. Lee, M. del Rosario, M. C. Harris, H. H. Blankenship, R. E. Guss, J. M. & Freeman, H. C. The crystal structure of auracyanin A at 1.85 Å resolution: The structures and functions of auracyanins A and B, two almost identical 'blue' copper proteins, in the photosynthetic bacterium Chloroflexus aurantiacus. *J. Biol. Inorg. Chem.* **14,** 329–345 (2009).
12. Lu, S. C. S-Adenosylmethionine. *Int. J. Biochem. Cell Biol.* **32,** 391–395 (2000).
13. Scheiner, S. *Noncovalent Forces*. *Noncovalent Forces* (Springer International Publishing, 2015).
14. Valdés-Martínez, J. *et al.* Study of the sulfur atom as hydrogen bond acceptor in N(2)-pyridylmethyl-N′-arylthioureas. *J. Chem. Crystallogr.* **34,** 533–540 (2004).
15. Zhou, P., Tian, F., Lv, F. & Shang, Z. Geometric characteristics of hydrogen bonds involving sulfur atoms in proteins. *Proteins Struct. Funct. Bioinforma.* **76,** 151–163 (2009).

16. Dey, A. Okamura, T.-a. Ueyama, N. Hedman, B. Hodgson, K. O. Solomon & E. I. Sulfur K-edge XAS and DFT calculations on P450 model complexes: Effects of hydrogen bonding on electronic structure and redox potentials. *J. Am. Chem. Soc.* **127,** 12046–12053 (2005).
17. Fugate, C. J. & Jarrett, J. T. Biotin synthase%: Insights into radical-mediated carbon – sulfur bond formation. *BBA - Proteins Proteomics* **1824,** 1213–1222 (2012).
18. Hamid M., S. *Biotin: Biochemical, Physiological and Clinical Aspects*. *Water Soluble Vitamins* (2012). doi:10.1007/978-94-007-2199-9
19. Richter, M. Functional diversity of organic molecule enzyme cofactors. *Nat. Prod. Rep.* **30,** 1324 (2013).
20. Hoja, U. *et al.* HFA1 encoding an organelle-specific acetyl-CoA carboxylase controls mitochondrial fatty acid synthesis in Saccharomyces cerevisiae. *J. Biol. Chem.* **279,** 21779–21786 (2004).
21. Vit, A., Misson, L., Blankenfeldt, W. & Seebeck, F. P. Ergothioneine biosynthetic methyltransferase EgtD reveals the structural basis of aromatic amino acid betaine biosynthesis. *ChemBioChem* **16,** 119–125 (2015).
22. Disulfide, T. *et al.* Supplemental Material for and its Role for Self-Resistance of the Human Pathogen Aspergillus fumigatus. 1–6 (2010).
23. MacLeod, A. J., Ross, H. B., Ozere, R. L., Digout, G. & van Rooyen, C. E. Lincomycin: A New Antibiotiv Active Against Staphylococci and Other Gram-Positive Cocci. *Canad. Med. Ass. J.* **91,** 1056–1060 (1964).
24. Waxman, D. J. & Strominger, J. L. Penicillin-Binding Proteins and the Mechanism of Action of Beta-Lactam Antibiotics1. *Annu. Rev. Biochem.* **52,** 825–869 (1983).
25. Dunbar, K. L., Scharf, D. H., Litomska, A. & Hertweck, C. Enzymatic Carbon-Sulfur Bond Formation in Natural Product Biosynthesis. *Chem. Rev.* **117,** 5521–5577 (2017).
26. Borlinghaus, J., Albrecht, F., Gruhlke, M. C. H., Nwachukwu, I. D. & Slusarenko, A. J. Allicin: Chemistry and biological properties. *Molecules* **19,** 12591–12618 (2014).
27. Halkier, B. A. & Gershenzon, J. BIOLOGY AND BIOCHEMISTRY OF GLUCOSINOLATES. *Annu. Rev. Plant Biol.* **57,** 303–333 (2006).
28. Akanmu, D., Cecchini, R., Aruoma, O. I. & Halliwell, B. The antioxidant action of ergothioneine. *Arch. Biochem. Biophys.* **288,** 10–16 (1991).
29. Palumbo, A. d'Ischia, M. Misuraca G. & Prota G. Isolation and structure of a new sulphur-conatining aminoacid from sea urchin eggs. *Tetrahedron Letters* **23,** 3207–3208 (1982).
30. Fahey, R. C. Glutathione analogs in prokaryotes. *Biochim. Biophys. Acta - Gen. Subj.* **1830,** 3182–3198 (2013).
31. Whitesides, G. M., Liburn, E. J. & Szajewski, P. Rates of Thiol-Disulfide Interchange Reactions between Mono- and Dithiols and Ellman's Reagentl George. *J. Org. Chem.* **42,** 332–338 (1977).
32. Mueller, E. G. Trafficking in persulfides: Delivering sulfur in biosynthetic pathways. *Nat. Chem. Biol.*

2, 185–194 (2006).

33. Drotar, A., Fall, L. R., Mishalanie, E. A., Tavernier, J. E. & Fall, R. A. Y. Enzymatic Methylation of Sulfide , Selenide , and Organic Thiols by Tetrahymena thermophila. *Appl. Environ. Microbiol.* **53,** 2111–2118 (1987).
34. Li, B., Forseth, R. R., Bowers, A. A., Schroeder, F. C. & Walsh, C. T. A Backup Plan for Self-Protection: *S*-Methylation of Holomycin Biosynthetic Intermediates in Streptomyces clavuligerus. *ChemBioChem* **13,** 2521–2526 (2012).
35. Oman, T. J., Boettcher, J. M., Wang, H., Okalibe, X. N. & Van Der Donk, W. A. Sublancin is not a lantibiotic but an *S*-linked glycopeptide. *Nat. Chem. Biol.* **7,** 78–80 (2011).
36. Wang, H. & Van Der Donk, W. A. Substrate selectivity of the sublancin *S*-glycosyltransferase. *J. Am. Chem. Soc.* **133,** 16394–16397 (2011).
37. Samuelsson, B. DETECTION OF LEUKOTRIENE A4 AS AN INTERMEDIATE IN THE BIOSYNTHESIS OF LEUKOTRIENES C4 and D4. *FEBS Lett.* **122,** 83–86 (1980).
38. Crosa, J. H. & Walsh, C. T. Genetics and Assembly Line Enzymology of Siderophore Biosynthesis in Bacteria. *Microbiol. Mol. Biol. Rev.* **66,** 223–249 (2002).
39. Roy, R. S., Gehring, a M., Milne, J. C., Belshaw, P. J. & Walsh, C. T. Thiazole and oxazole peptides: biosynthesis and molecular machinery. *Nat. Prod. Rep.* **16,** 249–263 (1999).
40. Knerr, P. J. & van der Donk, W. A. Discovery, Biosynthesis, and Engineering of Lantipeptides. *Annu. Rev. Biochem.* **81,** 479–505 (2012).
41. Brandel, J. Humbert, N. Elhabiri, M. Schalk, I. J. Mislin, G. L. A. & Albrecht-Gary, A-M. Pyochelin, a siderophore of Pseudomonas aeruginosa: Physicochemical characterization of the iron(iii), copper(ii) and zinc(ii) complexes. *Dalt. Trans.* **41,** 2820 (2012).
42. Drechsel, H. Stephan, H. Lotz, R. Haag, H. Zähner, H. Hantke, K. & Jung, G. Structure elucidation of yersiniabactin, a siderophore from highly virulent Yersinia strains. *Liebigs Ann.* **1995,** 1727–1733 (1995).
43. Li, B. & Van Der Donk, W. A. Identification of essential catalytic residues of the cyclase NisC involved in the biosynthesis of nisin. *J. Biol. Chem.* **282,** 21169–21175 (2007).
44. Favre, A., Saintomk, C., Clivio, P. & Laugâa, P. New Trends in Photobiology Thionucleobases as intrinsic photoaffinity probes of nucleic acid structure and nucleic acid-protein interactions. *J. Photochem. Photobiol.* **42,** 109–124 (1998).
45. Mueller, E. G., Buck, C. J., Palenchar, P. M., Barnhart, L. E. & Paulson, J. L. Identification of a gene involved in the generation of 4-thiouridine in tRNA. *Nucleic Acids Res.* **26,** 2606–2610 (1998).
46. Lauhon, C. T. & Kambampati, R. The iscS gene in Escherichia coli is required for the biosynthesis of 4- thiouridine, thiamin, and NAD. *J. Biol. Chem.* **275,** 20096–20103 (2000).
47. Mueller, E. G., Palenchar, P. M. & Buck, C. J. The Role of the Cysteine Residues of ThiI in the

Generation of 4-Thiouridine in tRNA *. *J. Biol. Chem.* **276,** 33588–33595 (2001).

48. Broderick, J. B., Du, B. R., Duschene, K. S. & Shepard, E. M. Radical S - Adenosylmethionine Enzymes. *Chem. Rev.* **114,** 4229–4317 (2014).
49. Peck, S. C. & van der Donk, W. A. Go it alone: four-electron oxidations by mononuclear non-heme iron enzymes. *J. Biol. Inorg. Chem.* **22,** 381–394 (2017).
50. Bridwell-Rabb, J., Zhong, A., Sun, H. G., Drennan, C. L. & Liu, H. A B12-dependent radical SAM enzyme involved in oxetanocin A biosynthesis. *Nature* **544,** 322–326 (2017).
51. Berkovitch, F., Nicolet, Y., Wan, J. T., Jarrett, J. T. & Drennan, C. L. Crystal Structure of Biotin Synthase , an S -Adenosylmethionine – Dependent Radical Enzyme. *Science (80-.).* **303,** 76–80 (2004).
52. Tse Sum Bui, B Lotierzo, M. Escalettes, F. Florentin, D. & Marquet A. Further Investigation on the Turnover of Escherichia coli Biotin Synthase with Dethiobiotin and 9-Mercaptodethiobiotin as Substrates †. *Biochemistry* **43,** 16432–16441 (2004).
53. Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929. *Bull. World Health Organ.* **79,** 780–790 (2001).
54. Tipper, D. J. & Strominger, J. L. Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proc. Natl. Acad. Sci. U. S. A.* **54,** 1133–1141 (1965).
55. Isomers, O. & Cystinylvaline, O. F. The Utilization of L-Cystinyl-L-Valine for Penicillin Biosynthesis. *Biochem. J* **76,** 323–327 (1958).
56. Pang, C. P. *et al.* Purification of isopenicillin N synthetase. *Biochem. J.* **222,** 789–95 (1984).
57. Roach, P. L. *et al.* Structure of Isopenecillin-N-synthase Complexed with Substrate and the Mechanism of Penecillin Formation. *Nature* **387,** 827–830 (1997).
58. Baldwin, J. E. & Bradley, M. Isopenicillin N Synthase: Mechanistic Studies. *Chem. Rev.* **90,** 1079–1088 (1990).
59. Tamanaha, E. *et al.* Spectroscopic Evidence for the Two C − H-Cleaving Intermediates of. *J. Am. Chem. Soc.* **138,** 8862–8874 (2016).
60. Goncharenko, K. V., Vit, A., Blankenfeldt, W. & Seebeck, F. P. Structure of the sulfoxide synthase EgtB from the ergothioneine biosynthetic pathway. *Angew. Chemie - Int. Ed.* **54,** 2821–2824 (2015).
61. Cheah, I. K. & Halliwell, B. Ergothioneine; antioxidant potential, physiological function and role in disease. *Biochim. Biophys. Acta* **1822,** 784–93 (2012).
62. Eagles, B. A. & Harry, M. V. The Physiology of ergothioneine. *J. Biol. Chem.* **80,** 615–622 (1928).
63. MELVILLE, D. B., HORNER, W. H. & LUBSCHEZ, R. Tissue ergothioneine. *J. Biol. Chem.* **206,** 221–228 (1954).
64. MELVILLE, D. B. & Eich, S. THE OCCURRENCE OF ERGOTHIONEINE IN PLANT MATERIAL. *J. Biol.*

Chem. **218,** 647 (1956).

65. Gründemann, D. *et al.* Discovery of the ergothioneine transporter. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 5256–5261 (2005).
66. Alamgir, K. M., Masuda, S., Fujitani, Y., Fukuda, F. & Tani, A. Production of ergothioneine by Methylobacterium species. *Front. Microbiol.* **6,** 1–12 (2015).
67. Bello, M. H., Barrera-Perez, V., Morin, D. & Epstein, L. The Neurospora crassa mutant Nc??Egt-1 identifies an ergothioneine biosynthetic gene and demonstrates that ergothioneine enhances conidial survival and protects against peroxide toxicity during conidial germination. *Fungal Genet. Biol.* **49,** 160–172 (2012).
68. Saini, V. *et al.* Ergothioneine Maintains Redox and Bioenergetic Homeostasis Essential for Drug Susceptibility and Virulence of Mycobacterium tuberculosis. *Cell Rep.* **14,** 572–585 (2016).
69. Sheridan, K. J. *et al.* Ergothioneine Biosynthesis and Functionality in the Opportunistic Fungal Pathogen, Aspergillus fumigatus. *Sci. Rep.* **6,** 35306 (2016).
70. Cumming, B. M., Chinta, K. C., Reddy, V. P. & Steyn, A. J. C. Role of Ergothioneine in Microbial Physiology and Pathogenesis. *Antioxid. Redox Signal.* **0,** ars.2017.7300 (2017).
71. Seebeck, F. P. In vitro reconstitution of Mycobacterial ergothioneine biosynthesis. *J. Am. Chem. Soc.* **132,** 6632–3 (2010).
72. Wei, W. J., Siegbahn, P. E. M. & Liao, R. Z. Theoretical Study of the Mechanism of the Nonheme Iron Enzyme EgtB. *Inorg. Chem.* **56,** 3589–3599 (2017).
73. Gardner, J. D., Pierce, B. S., Fox, B. G. & Brunold, T. C. Spectroscopic and computational characterization of substrate-bound mouse cysteine dioxygenase: Nature of the ferrous and ferric cysteine adducts and mechanistic implications. *Biochemistry* **49,** 6033–6041 (2010).
74. Goncharenko, K. The Catalytic Mechanism of the Iron-Dependent Sulfoxide Synthase EgtB. (2017).
75. Parameters, T. The Mechanism of the Rhodanese-catalyzed Thiosulf ate-Cyanide Reaction. **243,** 1892–1899 (1968).
76. Cipollone, R., Ascenzi, P., Tomao, P., Imperi, F. & Visca, P. Enzymatic Detoxification of Cyanide%: Clues from Pseudomonas aeruginosa. *J. Mol. Microbiol. Biotechnol.* 199–211 (2008). doi:10.1159/000121331
77. Bordo, D. & Bork, P. The rhodanese/Cdc25 phosphatase superfamily. Sequence – structure – function relations. *EMBO Rep.* **3,** 741–746 (2002).
78. Nandi, D. L., Horowitz, P. M. & Westley, J. Rhodanese as a thioredoxin oxidase. *Int. J. Biochem. Cell Biol.* **32,** 465–473 (2000).
79. Schlesinger, P. & Westley, J. An expanded mechanism for rhodanese catalysis. *J. Biol. Chem.* **249,** 780–788 (1974).
80. Gliubich, F., Berni, R., Colapietro, M., Barba, L. & Zanotti, G. Structure of sulfur-substituted

rhodanese at 1.36 Å resolution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 481–486 (1998).

81. Lincke, T., Behnken, S., Ishida, K., Roth, M. & Hertweck, C. Closthioamide: An unprecedented polythioamide antibiotic from the strictly anaerobic bacterium Clostridium cellulolyticum. *Angew. Chemie - Int. Ed.* **49,** 2011–2013 (2010).
82. Faponle, A. S., Seebeck, F. P. & De Visser, S. P. Sulfoxide Synthase versus Cysteine Dioxygenase Reactivity in a Nonheme Iron Enzyme. *J. Am. Chem. Soc.* **139,** 9259–9270 (2017).
83. Goncharenko, K. V. & Seebeck, F. P. Conversion of a non-heme iron-dependent sulfoxide synthase into a thiol dioxygenase by a single point mutation. *Chem. Commun.* **52,** 1945–1948 (2016).
84. Anbar, A. D. OCEANS: Elements and Evolution. *Science (80-.).* **322,** 1481–1483 (2008).
85. Bekker, A. *et al.* Dating the rise of atmospheric oxygen. *Nature* **427,** 117–120 (2004).
86. Thannickal, V. J. Oxygen in the evolution of complex life and the price we pay. *Am. J. Respir. Cell Mol. Biol.* **40,** 507–510 (2009).
87. Falkowski, P. G. Tracing Oxygen ' s Imprint on Earth ' s Metabolic Evolution. *Science (80-.).* **1724,** 1724–1725 (2006).
88. Deponte, M. Glutathione catalysis and the reaction mechanisms of glutathione-dependent enzymes. *Biochim. Biophys. Acta - Gen. Subj.* **1830,** 3217–3266 (2013).
89. Kalapos, M. P. Methylglyoxal in living organisms - Chemistry, biochemistry, toxicology and biological implications. *Toxicol. Lett.* **110,** 145–175 (1999).
90. Thornalley, P. J. & Rabbani, N. Glyoxalase in tumourigenesis and multidrug resistance. *Semin. Cell Dev. Biol.* **22,** 318–325 (2011).
91. Meister, A. & Anderson, M. E. GLUTATHIONE. *Ann. Rev. Biochem.* **52,** 711–760 (1983).
92. Roberts, D. D., Lewis, S. D., Ballou, D. P., Olson, S. T. & Shafer, J. A. Reactivity of Small Thiolate Anions and Cysteine-25 in Papain toward Methyl Methanethiosulfonate. *Biochemistry* **25,** 5595–5601 (1986).
93. Nath, K. a & Salahudeen, a K. Autoxidation of cysteine generates hydrogen peroxide: cytotoxicity and attenuation by pyruvate. *Am. J. Physiol.* **264,** F306–F314 (1993).
94. Held, K. D. & Biaglow, J. E. Mechanisms for the oxygen radical-mediated toxicity of various thiol-containing compounds in cultured mammalian cells. *Radiat. Res.* **139,** 15–23 (1994).
95. Winterbourn, C. C. Toxicity of iron and hydrogen peroxide: the Fenton reaction. *Toxicol. Lett.* **82–83,** 969–974 (1995).
96. Gaitonde, M. K. A spectrophotometric method for the direct determination of cysteine in the presence of other naturally occurring amino acids. *Biochem. J.* **104,** 627–633 (1967).
97. Cooper, A. J. L. Biochemistry of sulfur- containing amino acids. *Ann. Rev. Biochem.* **52,** 187–222 (1983).
98. Munday, R. Toxicity of thiols and disulphides: Involvement of free-radical species. *Free Radic. Biol.*

Med. **7,** 659–673 (1989).

99. Fahey, R. C., Brown, W. C., Adams, W. B. & Worsham, M. B. Occurrence of glutathione in bacteria. *J. Bacteriol.* **133,** 1126–1129 (1978).

100. Newton, G. L., Buchmeier, N. & Fahey, R. C. Biosynthesis and Functions of Mycothiol, the Unique Protective Thiol of Actinobacteria. *Microbiol. Mol. Biol. Rev.* **72,** 471–494 (2008).

101. Newton, G. L. & Javor, B. γ-Glutamylcysteine and thiosulfate are the major low-molecular-weight thiols in halobacteria. *J. Bacteriol.* **161,** 438–441 (1985).

102. Jones, D. P. & Liang, Y. Measuring the poise of thiol/disulfide couples in vivo. *Free Radic. Biol. Med.* **47,** 1329–1338 (2009).

103. Johansson, A. S. & Mannervik, B. Human Glutathione Transferase A3-3, a Highly Efficient Catalyst of Double-bond Isomerization in the Biosynthetic Pathway of Steroid Hormones. *J. Biol. Chem.* **276,** 33061–33065 (2001).

104. Hayes, J. D., Flanagan, J. U. & Jowsey, I. R. Glutathione Transferases. *Annu. Rev. Pharmacol. Toxicol.* **45,** 51–88 (2005).

105. Combes, B. & Stakelum, G. S. A liver enzyme that conjugates sulfobromophthalein sodium with glutathione. *J. Clin. Invest.* **40,** 981–988 (1961).

106. Booth, J., Boyland, E. & Sims, P. An enzyme from rat liver catalysing conjugations with glutathione. *Biochem. J.* **79,** 516–24 (1961).

107. Mannervik, B., Board, P. G., Hayes, J. D., Listowsky, I. & Pearson, W. R. Nomenclature for mammalian soluble glutathione transferases. *Methods Enzymol.* **401,** 1–8 (2005).

108. Sheehan, D., Meade, G., Foley, V. M. & Dowd, C. A. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochem. J.* **360,** 1–16 (2001).

109. Sinning, I. *et al.* Structure determination and refinement of human alpha class glutathione transferase A1-1, and a comparison with the Mu and Pi class enzymes. *Journal of Molecular Biology* **232,** 192–212 (1993).

110. Dixon, D. P. & Edwards, R. Glutathione Transferases☆. *Arab. B.* 1–13 (2010). doi:10.1016/B978-0-12-801238-3.04351-8

111. Holm, P. J. *et al.* Structural Basis for Detoxification and Oxidative Stress Protection in Membranes. *J. Mol. Biol.* **360,** 934–945 (2006).

112. Xiao, G., Parsons, J. F., Tesh, K., Armstrong, R. N. & Gilliland, G. L. Conformational changes in the crystal structure of rat glutathione transferase M1-1 with global substitution of 3-fluorotyrosine for tyrosine. *J. Mol. Biol.* **281,** 323–39 (1998).

113. Parsons, J. F. & Armstrong, R. N. Proton configuration in the ground state and transition state of a glutathione transferase-catalyzed reaction inferred from the properties of tetradeca(3-

fluorotyrosyl)glutathione transferase. *J. Am. Chem. Soc.* **118,** 2295–2296 (1996).

114. Molina, D. M. *et al.* Structural basis for synthesis of inflammatory mediators by human leukotriene C4 synthase. *Nature* **448,** 613–616 (2007).

115. Roberts, A. A. *et al.* Mechanistic studies of FosB: a divalent-metal-dependent bacillithiol-*S*-transferase that mediates fosfomycin resistance in Staphylococcus aureus. *Biochem. J* **451,** 69–79 (2013).

116. Rife, C. L., Pharris, R. E., Newcomer, M. E. & Armstrong, R. N. Crystal structure of a genomically encoded fosfomycin resistance protein (FosA) at 1.19%?? resolution by MAD phasing off the L-III edge of TI+. *J. Am. Chem. Soc.* **124,** 11001–11003 (2002).

117. Newton, G. L., Leung, S. S., Wakabayashi, J. I., Rawat, M. & Fahey, R. C. The DinB superfamily includes novel mycothiol, bacillithiol, and glutathione *S*-transferases. *Biochemistry* **50,** 10751–10760 (2011).

118. Cooper, D. R., Grelewska, K., Kim, C.-Y., Joachimiak, A. & Derewenda, Z. S. The structure of DinB from Geobacillus stearothermophilus: a representative of a unique four-helix-bundle superfamily. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* **66,** 219–24 (2010).

119. Zhao, Q., Wang, M., Xu, D., Zhang, Q. & Liu, W. Metabolic coupling of two small-molecule thiols programs the biosynthesis of lincomycin A. *Nature* **518,** 115–119 (2015).

120. Feng, J. *et al.* The gene ncgl2918 encodes a novel maleylpyruvate isomerase that needs mycothiol as cofactor and links mycothiol biosynthesis and gentisate assimilation in Corynebacterium glutamicum. *J. Biol. Chem.* **281,** 10778–10785 (2006).

121. Liu, T.-T. & Zhou, N.-Y. Novel L-Cysteine-Dependent Maleylpyruvate Isomerase in the Gentisate Pathway of Paenibacillus sp. Strain NyZ101. *J. Bacteriol.* **194,** 3987–3994 (2012).

122. Wilker, J. J. & Lippard, S. J. Alkyl Transfer to Metal Thiolates: Kinetics, Active Species Identification, and Relevance to the DNA Methyl Phosphotriester Repair Center of Escherichia coli Ada. *Inorg. Chem.* **36,** 969–978 (1997).

123. Warthen, C. R., Hammes, B. S., Carrano, C. J. & Crans, D. C. Methylation of neutral pseudotetrahedral zinc thiolate complexes: Model reactions for alkyl group transfer to sulfur by zinc-containing enzymes. *J. Biol. Inorg. Chem.* **6,** 82–90 (2001).

124. Rombach, M. *et al.* Thiolate alkylation in tripod zinc complexes: A comparative kinetic study. *Inorg. Chem.* **45,** 4571–4575 (2006).

125. Wilker, J. J. & Lippard, S. J. Modeling the DNA Methylphosphotriester Repair Site in Escherichia coli Ada. Why Zinc and Four Cysteines? *J. Am. Chem. Soc.* **117,** 8682–8683 (1995).

126. Chiou, S.-J., Riordan, C. G. & Rheingold, A. L. Synthetic modeling of zinc thiolates: Quantitative assessment of hydrogen bonding in modulating sulfur alkylation rates. *Proc. Natl. Acad. Sci.* **100,** 3695–3700 (2003).

127. Smith, J. N., Shirin, Z. & Carrano, C. J. Control of thiolate nucleophilicity and specificity in zinc metalloproteins by hydrogen bonding: Lessons from model compound studies. *J. Am. Chem. Soc.* **125,** 868–869 (2003).

128. Smith, J. N., Hoffman, J. T., Shirin, Z. & Carrano, C. J. H-bonding interactions and control of thiolate nucleophilicity and specificity in model complexes of zinc metalloproteins. *Inorg. Chem.* **44,** 2012–2017 (2005).

129. Picot, D., Ohanessian, G. & Frison, G. The alkylation mechanism of zinc-bound thiolates depends upon the zinc ligands. *Inorg. Chem.* **47,** 8167–8178 (2008).

130. Fox, D. C., Fiedler, A. T., Halfen, H. L., Brunold, T. C. & Halfen, J. a. Electronic structure control of the nucleophilicity of transition metal-thiolate complexes: An experimental and theoretical study. *J. Am. Chem. Soc.* **126,** 7627–7638 (2004).

131. Penner-Hahn, J. Zinc-promoted alkyl transfer: a new role for zinc. *Curr. Opin. Chem. Biol.* **11,** 166–171 (2007).

132. Huang, C. C., Hightower, K. E. & Fierke, C. A. Mechanistic studies of rat protein farnesyltransferase indicate an associative transition state. *Biochemistry* **39,** 2593–2602 (2000).

133. Miller Jenkins, L. M. *et al.* Studies on the Mechanism of Inactivation of the HIV-1 Nucleocapsid Protein NCp7 with 2-Mercaptobenzamide Thioesters. *J. Med. Chem.* **48,** 2847–2858 (2005).

134. Mishina, Y., Duguid, E. M. & He, C. Direct reversal of DNA alkylation damage. *Chem. Rev.* **106,** 215–232 (2006).

135. Ohnishi, R. & Tanabe, K. A New Method of Solubility Determination of Hydrolyzing Solute-Solubility of Benzyl chloride in Water. *Bull. Chem. Soc. Jpn.* **41,** 2647–2469 (1971).

136. Tang, S. S. & Chang, G. G. Kinetic characterization of the endogenous glutathione transferase activity of octopus lens *S*-crystallin. *J. Biochem.* **119,** 1182–8 (1996).

137. Britton, H. T. K. & Robinson, R. A. Universal Buffer Solutions and the Dis- sociation Constant of Verona. *J. Chem. Soc.* 1456–1462 (1931).

138. Gennari, M. *et al.* Experimental and computational investigation of thiolate alkylation in Ni II and Zn II complexes: Role of the metal on the sulfur nucleophilicity. *Inorg. Chem.* **50,** 10047–10055 (2011).

139. Li, Y. & Zamble, D. B. Nickel homeostasis and nickel regulation: An overview. *Chem. Rev.* **109,** 4617–4643 (2009).

140. Dudev, T. & Lim, C. Competition among metal ions for protein binding sites: Determinants of metal ion selectivity in proteins. *Chem. Rev.* **114,** 538–556 (2014).

141. Irving, H. & Williams, R. J. P. Order of Stability of Metal Complexes. *Nature* **162,** 746–747 (1948).

142. Williams, R. J. P. Chemical selection of elements by cells. *Coord. Chem. Rev.* **216–217,** 583–595 (2001).

143. Krissinel, E. & Henrick, K. Detection of Protein Assemblies in Crystals. *CompLife* 163–174 (2005).

144. Krissinel, E. Crystal Contacts as Nature ' s Docking Solutions. *J. Comput. Chem.* **31,** 133–143 (2009).
145. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372,** 774–797 (2007).
146. Koehntop, K. D., Emerson, J. P. & Que, L. The 2-His-1-carboxylate facial triad: A versatile platform for dioxygen activation by mononuclear non-heme iron(II) enzymes. *J. Biol. Inorg. Chem.* **10,** 87–93 (2005).
147. Alberts, I. L., Nadassy, K. & Wodak, S. J. Analysis of zinc binding sites in protein crystal structures. *Protein Sci.* **7,** 1700–1716 (1998).
148. Kiefer, L. L., Paterno, S. A. & Fierke, C. A. Hydrogen Bond Network in the Metal Binding Site of Carbonic Anhydrase Enhances Zinc Affinity and Catalytic Efficiency. *J. Am. Chem. Soc.* **117,** 6831–6837 (1995).
149. Maret, W. & Li, Y. Coordination dynamics of zinc in proteins. *Chem. Rev.* **109,** 4682–4707 (2009).
150. Ballouz, S., Francis, A. R., Lan, R. & Tanaka, M. M. Conditions for the evolution of gene clusters in bacterial genomes. *PLoS Comput. Biol.* **6,** (2010).
151. Johnson, T., Newton, G. L., Fahey, R. C. & Rawat, M. Unusual production of glutathione in Actinobacteria. *Arch. Microbiol.* **191,** 89–93 (2009).
152. Oshima, H., Sawa, T. & Akaike, T. 8-Nitroguanine, a Product of Nitrative DNA Damage Caused by Reactive Nitrogen Species: Formation, Occurrence, and Implications in Inflammation and Carcinogenesis. *Antioxid. Redox Signal.* **8,** 1033–1045 (2006).
153. Hu, W. *et al.* Bioinformatic and biochemical characterizations of C-S bond formation and cleavage enzymes in the fungus neurospora crassa ergothioneine biosynthetic pathway. *Org. Lett.* **16,** 5382–5385 (2014).
154. Mashabela, G. T. M. & Seebeck, F. P. Substrate specificity of an oxygen dependent sulfoxide synthase in ovothiol biosynthesis. *Chem. Commun.* **49,** 7714 (2013).
155. Liao, C. & Seebeck, F. P. Convergent Evolution of Ergothioneine Biosynthesis in Cyanobacteria. *ChemBioChem* 1–5 (2017). doi:10.1002/cbic.201700354
156. Pluskal, T., Ueno, M. & Yanagida, M. Genetic and metabolomic dissection of the ergothioneine and selenoneine biosynthetic pathway in the fission yeast, S. pombe, and Construction of an overproduction system. *PLoS One* **9,** (2014).
157. Dominy, J. E., Simmons, C. R., Karplus, P. A., Gehring, A. M. & Stipanuk, M. H. Identification and characterization of bacterial cysteine dioxygenases: A new route of cysteine degradation for eubacteria. *J. Bacteriol.* **188,** 5561–5569 (2006).
158. Kasperova, A. *et al.* Isolation of recombinant cysteine dioxygenase protein from Trichophyton mentagrophytes. *Mycoses* **54,** 456–462 (2011).
159. Engi, P. Chemical Probes for Mechanistic Enzymology. (2017).

160. Bock, T. *et al.* An integrated approach for genome annotation of the eukaryotic thermophile Chaetomium thermophilum. *Nucleic Acids Res.* **42,** 13525–13533 (2014).

161. Cava, F., Hidalgo, A. & Berenguer, J. Thermus thermophilus as biological model. *Extremophiles* **13,** 213–231 (2009).

162. Song, H. *et al.* Cysteine oxidation reactions catalyzed by a mononuclear non-heme iron enzyme (OvoA) in ovothiol biosynthesis. *Org. Lett.* **16,** 2122–2125 (2014).

163. Purich, D. L. *Enzyme Kinetics*. (2010).

164. Fondriest Environmental, Inc. 'Dissolved Oxygen.' Fundamentals of Environmental Measurements. 19 Nov. 2013. Web. < http://www.fondriest.com/environmental-measurements/parameters/water-quality/dissolved-oxygen/ >.

165. Yamashita, Y. & Yamashita, M. Identification of a novel selenium-containing compound, selenoneine, as the predominant chemical form of organic selenium in the blood of bluefin tuna. *J. Biol. Chem.* **285,** 18134–18138 (2010).

166. Yamashita, Y., Yabu, T. & Yamashita, M. Discovery of the strong antioxidant selenoneine in tuna and selenium redox metabolism. *World J. Biol. Chem.* **1,** 144–150 (2010).

167. Yamashita, M., Yamashita, Y., Ando, T., Wakamiya, J. & Akiba, S. Identification and determination of selenoneine, 2-Selenyl-N α, N α, N α -Trimethyl-l-histidine, as the major organic selenium in blood cells in a fish-eating population on remote Japanese Islands. *Biol. Trace Elem. Res.* **156,** 36–44 (2013).

168. Klein, M., Ouerdane, L., Bueno, M. & Pannier, F. Identification in human urine and blood of a novel selenium metabolite, Se-methylselenoneine, a potential biomarker of metabolization in mammals of the naturally occurring selenoneine, by HPLC coupled to electrospray hybrid linear ion trap-orbital ion tra. *Metallomics* **3,** 513 (2011).

169. Krause, R. J. & Elfarra, A. A. Reduction of l-methionine selenoxide to seleno-l-methionine by endogenous thiols, ascorbic acid, or methimazole. *Biochem. Pharmacol.* **77,** 134–140 (2009).

170. Kumar, D., Thiel, W. & De Visser, S. P. Theoretical study on the mechanism of the oxygen activation process in cysteine dioxygenase enzymes. *J. Am. Chem. Soc.* **133,** 3869–3882 (2011).

171. Blaesi, E. J., Gardner, J. D., Fox, B. G. & Brunold, T. C. Spectroscopic and computational characterization of the NO adduct of substrate-bound Fe(II) cysteine dioxygenase: Insights into the mechanism of O2 activation. *Biochemistry* **52,** 6040–6051 (2013).

172. Che, X., Gao, J., Liu, Y. & Liu, C. Metal vs. chalcogen competition in the catalytic mechanism of cysteine dioxygenase. *J. Inorg. Biochem.* **122,** 1–7 (2013).

173. Busby, R. W. Expression and Purification of two Isozymes of Clavaminate Synthase and Initial Characterization of the Iron Binding Site. *The Journal of biological chemistry* **270,** 4262–4269 (1995).

174. Salowe, S. P., Neil Marsh, E. & Townsend, C. A. Purification and Characterization of Clavaminate Synthase from Streptomyces clavuligerus: An Unusual Oxidative Enzyme in Natural Product Biosynthesis. *Biochemistry* **29,** 6499–6508 (1990).

175. Solomon, E. I., Goudarzi, S. & Sutherlin, K. D. O 2 Activation by Non-Heme Iron Enzymes. *Biochemistry* **55,** 6363–6374 (2016).

176. Pierce, B. S., Gardner, J. D., Bailey, L. J., Brunold, T. C. & Fox, B. G. Characterization of the nitrosyl adduct of substrate-bound mouse cysteine dioxygenase by electron paramagnetic resonance: Electronic structure of the active site and mechanistic implications. *Biochemistry* **46,** 8569–8578 (2007).

177. Guan, Z., Yates, N. A. & Bakhtiar, R. Detection and characterization of methionine oxidation in peptides by collision-induced dissociation and electron capture dissociation. *J. Am. Soc. Mass Spectrom.* **14,** 605–613 (2003).

178. Kumar, V. *et al.* Purification and Characterization of a Cysteine Dioxygenase from the Yeast Phase of Histoplasma capsulatum. *Biochemistry* **22,** 762–768 (1983).

179. McCoy, J. G. *et al.* Structure and mechanism of mouse cysteine dioxygenase. *Proc. Natl. Acad. Sci.* **103,** 3084–3089 (2006).

180. Stipanuk, M. H., Ueki, I., Dominy, J. E., Simmons, C. R. & Hirschberger, L. L. Cysteine dioxygenase: A robust system for regulation of cellular cysteine levels. *Amino Acids* **37,** 55–63 (2009).

181. Joseph, C. A. & Maroney, M. J. Cysteine dioxygenase: structure and mechanism. *Chem. Commun.* 3338 (2007). doi:10.1039/b702158e

182. Tchesnokov, E. P. *et al.* An iron–oxygen intermediate formed during the catalytic cycle of cysteine dioxygenase. *Chem. Commun.* **52,** 8814–8817 (2016).

183. Balog, A. *et al.* The arbuscular mycorrhizal fungus Rhizophagus irregularis affects arthropod colonization on sweet pepper in both the field and greenhouse. *J. Pest Sci. (2004).* **90,** 935–946 (2017).

184. The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* **408,** 796–815 (2000).

185. Hamilton, J. P. & Robin Buell, C. Advances in plant genome sequencing. *Plant J.* **70,** 177–190 (2012).

186. Wang, D. Y.-C., Kumar, S. & Hedges, S. B. Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. *Proc. R. Soc. B Biol. Sci.* **266,** 163–171 (1999).

187. Grumbt, M. *et al.* Keratin degradation by dermatophytes relies on cysteine dioxygenase and a sulfite efflux pump. *J. Invest. Dermatol.* **133,** 1550–1555 (2013).

188. Kunert, J. Keratin decomposition by dermatophytes: Evidence of the sulphitolysis of the protein. *Experientia* **28,** 1025–1026 (1972).

189. Park, H. & Bakalinsky, A. T. SSUI mediates sulphite efflux in Saccharomyces cerevisiae. *Yeast* **16,**

881–888 (2000).

190. Freyer, M. W. & Lewis, E. A. Isothermal Titration Calorimetry: Experimental Design, Data Analysis, and Probing Macromolecule/Ligand Binding and Kinetic Interactions. *Methods Cell Biol.* **84,** 79–113 (2008).

191. Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2,** 2212–2221 (2007).

192. Schellman, J. A. Temperature, stability, and the hydrophobic interaction. *Biophys. J.* **73,** 2960–4 (1997).

193. Vedadi, M. *et al.* Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 15835–40 (2006).

194. Velazquez-Campoy, A. & Freire, E. Isothermal titration calorimetry to determine association constants for high-affinity ligands. *Nat. Protoc.* **1,** 186–191 (2006).

195. Pierce, M. M., Raman, C. S. & Nall, B. T. Isothermal Titration Calorimetry of Protein – Protein Interactions. *Methods* **221,** 213–221 (1999).

196. Matulis, D., Rouzina, I. & Bloomfield, V. A. Thermodynamics of DNA binding and condensation: Isothermal titration calorimetry and electrostatic mechanism. *J. Mol. Biol.* **296,** 1053–1063 (2000).

197. Boonsongrit, Y., Mueller, B. W. & Mitrevej, A. Characterization of drug-chitosan interaction by 1H NMR, FTIR and isothermal titration calorimetry. *Eur. J. Pharm. Biopharm.* **69,** 388–395 (2008).

198. Song, C., Zhang, S. & Huang, H. Choosing a suitable method for the identification of replication origins in microbial genomes. *Front. Microbiol.* **6,** (2015).

199. Freire, E., Mayorga, O. L. & Straume, M. Isothermal Titration. *Anal. Chem.* **62,** 950A–959A (1990).

200. Of, C., Related, F., In, G., Urium, S. T. & Demerec, B. Y. M. CLUSTERING OF FUNCTIONALLY RELATED GENES IN SALMONELLA TYPHIMURIUM. *Genetics* **51,** 1057–1060 (1964).

201. Noinaj, N. *et al.* Structural insights into the catalytic mechanism of Escherichia coli selenophosphate synthetase. *J. Bacteriol.* **194,** 499–508 (2012).

202. Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu. Rev. Biochem.* **77,** 521–555 (2008).

203. Liang, D.-M. *et al.* Glycosyltransferases: mechanisms and applications in natural product development. *Chem. Soc. Rev.* **44,** 8350–8374 (2015).

204. Breton, C., Šnajdrová, L., Jeanneau, C., Koča, J. & Imberty, A. Structures and mechanisms of glycosyltransferases. *Glycobiology* **16,** 29–37 (2006).

205. Charnock, S. J. & Davies, G. J. Structure of the nucleotide-diphospho-sugar transferase, SpsA from Bacillus subtilis, in native and nucleotide-complexed forms. *Biochemistry* **38,** 6380–6385 (1999).

206. Vrielink, A., Rüger, W., Driessen, H. P. & Freemont, P. S. Crystal structure of the DNA modifying

enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. *EMBO J.* **13,** 3413–22 (1994).

207. Lizak, C., Gerber, S., Numao, S., Aebi, M. & Locher, K. P. X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **474,** 350–356 (2011).
208. Coutinho, P. M., Deleury, E., Davies, G. J. & Henrissat, B. An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* **328,** 307–317 (2003).
209. Kojima, H., Shinohara, A. & Fukui, M. Sulfurifustis variabilis gen. nov., sp. nov., a sulfur oxidizer isolated from a lake, and proposal of Acidiferrobacteraceae fam. nov. and Acidiferrobacterales ord. nov. *Int. J. Syst. Evol. Microbiol.* **65,** 3709–3713 (2015).
210. Wang, X. P. *et al.* Structural and enzymatic analyses of a glucosyltransferase Alr3699/HepE involved in Anabaena heterocyst envelop polysaccharide biosynthesis. *Glycobiology* **26,** 520–531 (2016).
211. Giganti, D. *et al.* Secondary structure reshuffling modulates glycosyltransferase function at the membrane. *Nat. Chem. Biol.* **11,** 16–18 (2015).
212. Bennett, B., Kimball, E. & Gao, M. Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nat. Chem. …* **5,** 593–599 (2009).
213. Glass, R. S. *et al.* Monoselenophosphate: Synthesis, Characterization, and Identity with the Prokaryotic Biological Selenium Donor, Compound SePX. *Biochemistry* **32,** 12555–12559 (1993).
214. Nakai, H. *et al.* The maltodextrin transport system and metabolism in Lactobacillus acidophilus NCFM and production of novel α-glucosides through reverse phosphorolysis by maltose phosphorylase. *FEBS J.* **276,** 7353–7365 (2009).
215. Langen, R., Jensen, G. M., Jacob, U., Stephens, P. J. & Warshel, A. Protein control of iron-sulfur cluster redox potentials. *J. Biol. Chem.* **267,** 25625–25627 (1992).
216. Yang, X., Niu, S., Ichiye, T. & Wang, L.-S. Direct Measurement of the Hydrogen-Bonding Effect on the Intrinsic Redox Potentials of [4Fe–4S] Cubane Complexes. *J. Am. Chem. Soc.* **126,** 15790–15794 (2004).
217. Mathé, C. Weill, C. O. Mattioli, T. A. Berthomieu, C. Houée-Levin, C. Tremey, E. & Nivière, V. Assessing the role of the active-site cysteine ligand in the superoxide reductase from Desulfoarculus baarsii. *J. Biol. Chem.* **282,** 22207–22216 (2007).
218. Bennett, B., Kimball, E. & Gao, M. Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nat. Chem.* **5,** 593–599 (2009).

7 Appendix

7.1 General Experimental

The solvents used in reaction and assay were purchased in HPLC-grade und directly used. The reagents used in reaction, in assays and as substrates were purchased from commercial sources. The UHPLC/MS-measurements were measured on an Agilent 1290 Infinity system. The column was a Zorbax Eclipse Plus C-18-column (2.1 x 50 mm; particle size 1.8 µm; Agilent), with a flow rate of 0.4 ml/min. The coupled mass spectrometer was an Agilent 6130 quadrupol-MS. Eluents were A: acetonitrile and B: water/acetonitrile/TFA (1000:10:1). RP-HPLC was used with a Gemini-NX, 5 µM C18 150 x 4.6 mm column column and the solvent system A =acetonitrile/TFA 1000:1, B = water/acetonitrile/TFA (1000:10:1). Compounds were detected at 260 nm Cation exchange HPLC was used with a Luna 5u SCX column (100 Å, 150 x 4 mm, Phenomenex). Compounds were detected at 265 nm. The elution buffer was 20 mM phosphate and 1 M NaCl at pH 2. A Seven Easy pH meter from Mettler Toledo was used for pH measurements. A Nanodrop 2000 Spectrophotometer from Thermo Scientific was used for absorbance measurements. NMR spectra were measured on a Bruker 400 MHz or a Bruker 500 MHz instrument. Size exclusion was performed with an Äkta FPLC, GE Healthcare using a Superdex 200 5/150 GL column. In a constant flow of 0.2 ml/min in degassed FPLC buffer (200 mM NaCl, 50 mM Tris/HCl pH 8).

Acknowledgments:

I would like to thank Prof. Dr. Florian Seebeck for giving me the opportunity to work in his group. I am very grateful that he gave me the chance to change my research field from organic synthesis to my preferred topic biochemistry. Furthermore, I would like to thank him for his support and patience during my PhD.
I would like to thank Prof. Dr. Thomas Ward for accepting the co-examination and taking the time to read my thesis.
I would like to thank Prof. Christof Sparr for chairing my PhD defense and for the good working atmosphere during the teaching courses.

Some special thanks go to the current and former members of the Seebeck group. I enjoyed the great atmosphere and the unique team spirit of our group and was glad to be a part of. Thanks to Dr. Gabriel Mashabela, Dr. Sebastien Coyne, Dr. Roxana Lemnaru, Dr. Kristina Goncharenko, Reto Burn, Florian Leisinger, Dzmitry Miarzlou, Alice Maurer, Julia Hildesheim, Alma Idrizovic, Thanh Dang and Tom Tran. I would further thank Dr. Cangsong Liao and Dr. Matthias Knop not only for the great time in the Lab304 but also for all the support and the enormous amount of help and shared knowledge. Furthermore, I would like to thank Maria Belyaeva, for helping me with the ITC-experiments and Dr. Marcel Meury for helping me with my crystallization attempts and for solving the crystal structures of EF_3021. Special thanks go to Nico Valerio Igaereta, my master student, for all the work he did.

Without the help of my thesis correction-, formatting-, input-, feedback-, wizard-team I would have never been able to finish my thesis. I am very grateful and would like to thank Dr. David Lim, Dr. Pascal Engi, Tobias Brunner, and especially Anja Stampfli.
I also want to thank to my current and former members of the Wettstein-WG, Philipp, Oli, Patricia, Maria and Sara. Even though I was not too often in Basel and a bit stressed in the last weeks I felt always welcomed and home in the WG.

I would like to thank my family for all the support over the last four years. A special thanks to my two best brothers Matthias and Samuel which were always ready to help me to get some distraction with MTG and board games. Furthermore I am grateful to my parents to give me all the opportunities I had in live, supported me in every moment and thought me to never give up.

The last weeks/month of my PhD were very stressful and uncomfortable for all involved parties. Nevertheless, one person stood at my side and supported me the whole time with her love, patience and the endless believe in me. Therefore, my last and most important thank goes to Anja Schmutz.