

# **Identification and Characterization of a Novel Copper Dependent Enzyme**

**Inauguraldissertation**

zur  
Erlangung der Würde eines Doktors der Philosophie  
vorgelegt der  
Philosophisch-Naturwissenschaftlichen Fakultät  
der Universität Basel

Von

**Matthias Knop**

aus Ribnitz-Damgarten, Deutschland

Basel, 2017

*Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel  
edoc.unibas.ch*

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

Auf Antrag von

Prof. Dr. Florian P. Seebeck

Prof. Dr. Thomas R. Ward

Basel, den 21.02.2017

Prof. Dr. Martin Spiess

Dekan der Philosophisch-  
Naturwissenschaftlichen Fakultät





---

- Do. Or do not. There is no try. -

---

---

## Abstract

The Formylglycine generating enzyme (FGE) catalyses the oxygen dependent oxidation of a highly conserved cysteine residue in sulfatases and phosphatases to formylglycine (fGly).<sup>[3]</sup> In sulfatases the hydrated form of fGly serves as the catalytic residue to cleave sulfate esters of their substrates.<sup>[46][47]</sup> After the first identification of FGE as the key enzyme in sulfatase activation, it has been quickly associated with multiple sulfatase deficiency (MSD), a rare but fatal lysosomal storage disease. Various mutations in the FGE encoding *SUMF1* gene have been identified in patients suffering from MSD.<sup>[48]</sup>

Alongside the medicinal importance, FGE has gained significant attention due to the fact, that it has been identified as a cofactor-independent oxidase with an unknown mechanism of oxygen activation. The absence of metals and cofactors in published crystal structures and enzyme preparations raised the question of how oxygen activation can occur.

We hereby show that FGE is a copper dependent oxidase. Recombinant FGE from *Thermomonospora curvata* contains a disulfide bond but is readily reduced in the presence of reducing agents such as DTT or cysteamine. Reduced FGE shows a near atto molar affinity to copper(I) and binds copper throughout multiple turnovers. Copper binds to two highly conserved cysteine residues in the active site of FGE where it is involved in oxygen activation. While copper can be replaced by other metals such as silver, only copper facilitates the reaction, which supports its participation in redox chemistry.

Furthermore, we describe the dependence of FGE on a peptidyl reducing agent containing a thioredoxin-like CxC amino acid motif. FGE from *Homo sapiens* is linked to an N-terminal domain bearing the CxC sequence. This peptide tag is believed to be involved in retention of FGE to the ER. We additionally suggest that this domain serves as the immediate reducing agent of FGE and is then reduced by other pathways. Sequence similarities of this domain to the substrate polypeptide indicate that the N-terminal domain binds to the substrate binding groove of FGE. This suggests a ping-pong mechanism of product formation and subsequent reduction of the oxidized enzyme intermediate.

In addition to the mechanistic interest, FGE has been used as a tool to site-specifically introduce aldehyde functionalities into recombinant proteins for further bioconjugation. We now identify a set of five mutations which increases the catalytic efficiency of FGE by 38-fold. The resulting variant completely oxidizes a specific cysteine residue in a target protein within 90 min with only 0.3 % of catalyst loading at room temperature. This high reactivity is maintained even when the recognition motif is installed into an internal loop region of the target protein.

We envision that these findings will not only drive the mechanistic discussion about the previously mysterious oxygen activation in FGE but will also support the usage of fGly in bioconjugation research.

---



---

## Abbreviations

ABDT	2-amino butanedithiole
BCS	Bathocuproinedisulfonic Acid
CAO	Copper-Dependent Amine Oxidase
CD	Circular Dichroism
DTT	Dithiothreitol
EPR	Electron Paramagnetic Resonance
ER	Endoplasmic Reticulum
ERp44	ER Resident Protein
FAD	Flavine Adenine Dinucleotide
FGE	Formylglycine generating enzyme
FGE <sub>AC</sub>	FGE <sub>C187A,C231A,C284S,C298A</sub> from <i>T. curvata</i>
FGE <sub>smegmatis</sub>	FGE from <i>M. smegmatis</i>
FGE <sub>WT</sub>	Wild Type of FGE from <i>T. curvata</i>
fGly	Formylglycine
GFP	Green Fluorescent Protein
GSH	Glutathione
HRMS	High Resolution Mass Spectroscopy
IAA	2-Iodoacetamide
ICP	Inductively Coupled Plasma
KIE	Kinetic Isotope Effect
MSD	multiple sulfatase deficiency
PDI	Protein Disulfide Isomerase
RLuc	Luciferase from <i>Renilla reniformis</i>
SHE	Standard Hydrogen Electrode
SUMF1	Sulfatase Modifying Factor 1
TCEP	Tris(2-carboxyethyl)phosphine
TPQ	2,4,5-trihydroxy-phenylalanine
Trx	Thioredoxin
UOX	Urate Oxidase



---

## Outline

1.	Understanding the impossible.....	1
2.	Introduction - Oxygen and Metals in Enzyme Mechanisms .....	3
2.1	Oxygen Using Enzymes - From Oxydases and Oxygenases.....	3
2.1.1	Cofactor-Independent Oxidases.....	5
2.1.2	Cofactor-Independent Oxygenases.....	6
2.2	Metals in Enzyme Active Sites.....	7
2.2.1	Iron enzymes .....	7
2.2.2	Copper oxidases and oxygenases.....	8
2.2.3	Copper(I) Enzymes.....	10
2.2.4	Lytic polysaccharide monooxygenases .....	10
2.3	Formylglycine Generating Enzyme.....	12
2.3.1	Sulfatases in <i>Homo sapiens</i> .....	12
2.3.2	FGE in human cells .....	13
2.3.3	Structure of FGE .....	14
2.3.4	FGE mechanism .....	15
2.4	Bioconjugation.....	18
3.	Aim of this Thesis.....	21
4.	Formylglycine Formation – A Copper Dependent Reaction .....	23
4.1	In vitro Reconstitution of FGE with Copper(I).....	23
4.1.1	Establishing a robust study system .....	23
4.1.2	Analysis of the FGE catalyzed reaction.....	25
4.1.3	Metal dependent activity of FGE.....	27
4.2	Conclusion .....	32
4.3	Experimental .....	33
4.4	Appendix .....	37
5.	Copper Binding in Formylglycine Generating Enzyme .....	41
5.1	Identification of metal binding site in FGE from <i>T. curvata</i> .....	41
5.2	Conclusion .....	49
5.3	Experimental .....	51
5.4	Appendix .....	54
6.	Identification of Reducing Agents for FGE .....	57

---

6.1	FGE reduction <i>in vitro</i> .....	57
6.1.1	FGE activity in the absence of reducing agent .....	57
6.1.2	Scope of reducing agents compatible with FGE.....	61
6.1.3	Reducing agent <i>in vivo</i> .....	63
6.2	Conclusion .....	67
6.3	Experimental .....	68
6.4	Appendix.....	71
7.	Optimization of FGE driven aldehyde formation <i>in vitro</i> .....	73
7.1	Identification of <i>in vitro</i> gain of function cysteine mutation .....	74
7.2	Mutation of active site Tyr273 .....	77
7.3	Conclusion .....	82
7.4	Experimental .....	83
7.5	Appendix.....	86
8.	Final Conclusion and Outlook .....	89
9.	References .....	91
	Acknowledgement.....	99

---

## 1. Understanding the impossible

Enzymes belong to the most fascinating creations of nature. To fully understand how these protein catalysts can perform the most difficult reactions with exceptional specificity or simply how they form stable constructs is still one of the main quests of scientists all around the globe. The first enzyme to be discovered appeared to be a diastase in the 1830s.<sup>[1]</sup> Not until the early 19<sup>th</sup> century, however, the very first enzyme was purified and crystallized. This groundbreaking work was then little later acknowledged with the Nobel Prize in Chemistry for Sumner, Northrop and Stanley in 1946.<sup>[2]</sup> Curiosity and the eager for knowledge have thriven many scientists and researchers to investigate the complex mechanisms behind enzymes ever since.

The formylglycine generating enzyme (FGE) is such an enzyme that captivates the scientific community in recent years. It catalyzes the site specific oxidation of a conserved cysteine residue within sulfatases.<sup>[3]</sup> Dysfunction of sulfatases causes a rare but lethal disease, the multiple sulfatase deficiency (MSD).<sup>[4]</sup> In pursuit of investigating possible treatments of MSD, understanding the details of FGE catalysis is key. In addition to the medicinal interest, FGE has gained considerable attention due to widespread applications opening up in the field of covalent protein labeling and immobilization.<sup>[5][6]</sup>

However, research in FGE has been driven mainly due to another key characteristic: the absence of metals and cofactors while performing an oxidase type reaction. Oxygen dependent oxidation of organic matter is subject to the rules of quantum chemistry, meaning that molecular oxygen in its triplet state, as present in the atmosphere, cannot react with unactivated organic matter (see chapter 2). Dioxygen is a diradical in the triplet state and does not spontaneously react with other molecules in the singlet state. However, enzymes exist which perform oxygen activation and do not rely on conventional activation mechanisms such as metal ions or organic cofactors. To some extent this observation might appear impossible. Careful evaluation, however, shows that these enzymes often perform reactions by using 'cofactor-like' substrates.<sup>[8][9]</sup>

The substrate of FGE is a simple polypeptide consisting of only proteogenic amino acids. The mechanisms of oxygen activation by FGE in absence of cofactor or metal can not be understood. It has therefore been a main target to elucidate the questionable mechanism and to understand the FGE-catalyzed oxygen activation.

This work describes the investigation of the FGE mechanism and the resulting identification of FGE as being a copper(I) dependent oxidase.

---

---

---

## 2. Introduction - Oxygen and Metals in Enzyme Mechanisms

About 2.45 billion years ago the earth was a different place to the world we live in today. Geological and geochemical evidence shows, that the level of atmospheric oxygen was significantly lower than it is today.<sup>[11]</sup> The amount of available oxygen increased to considerable levels by 2.2 billion years ago when the first organisms, such as cyanobacteria, released molecular oxygen into the atmosphere.<sup>[12]</sup> Higher organisms evolved, by using the increased oxygen levels to produce energy for their living. Most 'recently' the *Homo sapiens* populated the now oxygen rich world, only thanks to the first appearance of O<sub>2</sub> billions of years in the past. Even though anaerobic organisms exist, to which oxygen in fact is toxic<sup>[13]</sup>, the vast majority of today's animal kingdom uses oxygen for their living.<sup>[14]</sup>

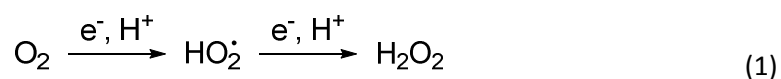
Long before oxygen found its way into the biochemistry of low level organisms, transition metals were present to play central roles in catalysis. The oceans of the ancient world were filled with water soluble metal salts that could be readily incorporated into the simple single cellular organisms of that time.

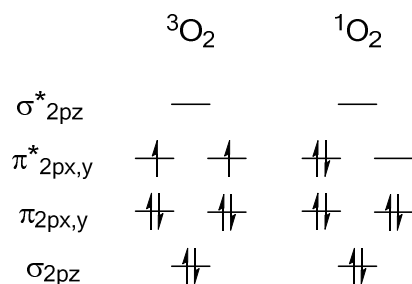
When the concentration of molecular oxygen in the atmosphere increased, new enzymatic metal centers evolved that could make use of both, redox properties of the transition metals and the oxidation potential of molecular oxygen. A new world of chemistry has been opened up for many new organisms to develop unique biochemical and biosynthetic pathways and to create the world in which we happen to live to date.

### 2.1 Oxygen Using Enzymes - From Oxydases and Oxygenases

When oxygen became available for an increasing number of organisms, new chemical pathways were developed. However, all of them had to face one key property of molecular oxygen - unreactivity towards singlet molecules.<sup>[7]</sup> Oxygen is a molecule in a triplet state (two unpaired electrons). Organic matter is mostly present in a singlet state (only paired electrons). Organisms, more particularly enzymes within the organisms, where therefore faced with one major problem: Oxygen activation. For a reaction to occur, oxygen in the triplet state has to be excited or activated to the singlet state in order to overcome the spin restriction (see Fig. 2-1 and Eq. 1). While the formation of the peroxy radical is endothermic, hence requires an energy input ( $E_0'$  for O<sub>2</sub>/O<sub>2</sub><sup>-</sup> = -0.33 V), the following hydrogen peroxide formation can occur spontaneously.

Many enzymes have evolved which can overcome the initial energy barrier and in the following reaction can harness the oxidizing capability of molecular oxygen. In order to make use of the strongest available oxidant in biology, enzymes such as oxygenases often contain redox-active metal ions that are capable of transferring a single electron to dioxygen (i.e. iron(II), copper (I)). Others use organic cofactors to act as an electron donor (i.e. flavins).<sup>[15]</sup> However, some rather rare examples exist in which enzymes are able to activate oxygen without the requirement of external reagents.<sup>[16]</sup> In general, enzymes that use dioxygen can be divided into two groups, oxidases, which use oxygen as an electron acceptor and reduce it to water or hydrogen peroxide,





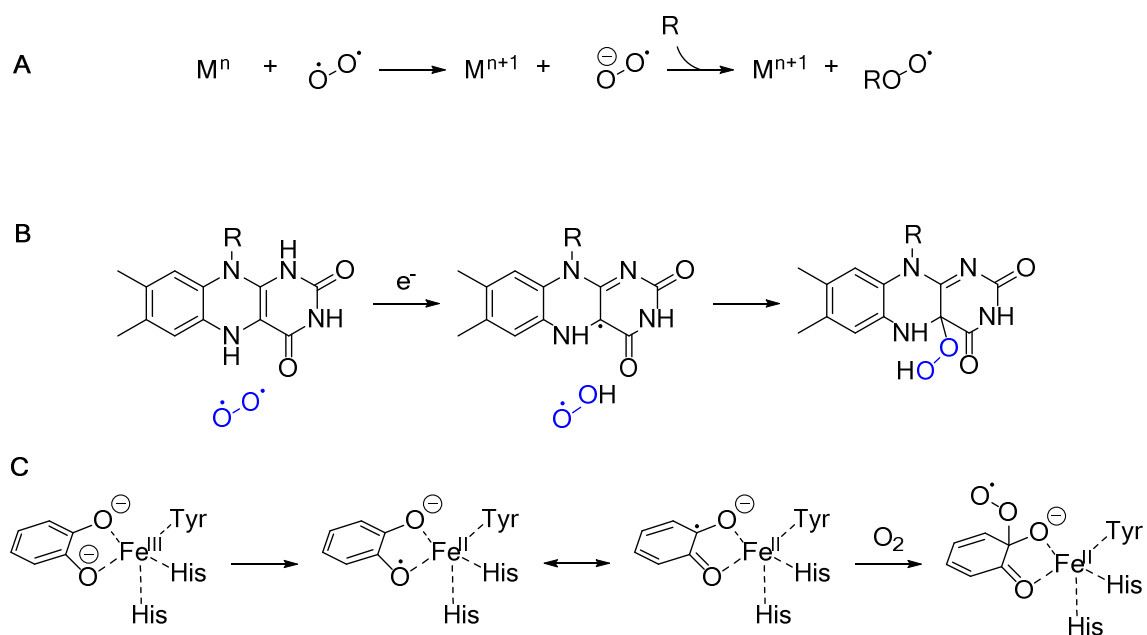
**Fig. 2-1:** Molecular orbital diagrams for triplet ( ${}^3\text{O}_2$ ) and singlet oxygen ( ${}^1\text{O}_2$ ).

and oxygenases, which incorporate the oxygen atoms into the substrate and product molecules. This group again divides in two sub-groups. Mono-oxygenases only incorporate a single oxygen atom whereas di-oxygenases transfer both atoms of dioxygen to the substrate. Most dioxygenases use metal ions in their active centres. These transition metal containing enzymes can use mainly three strategies to activate oxygen.<sup>[15]</sup> Firstly, orbitals with unpaired 3d electrons of the transition metal can overlap with the  $\pi^*$  orbitals of the oxygen molecule, containing the unpaired electrons (see Fig. 2-1). The resulting complex of metal and oxygen can then undergo reactions with singlet molecules. Secondly, single electron transfer from the metal to oxygen can occur. Transition metals in enzymes can often be found in two distinct oxidation states (i.e. Cu(I)/Cu(II) or Fe(II)/Fe(III)). These metals can transfer a single electron to dioxygen, which generates superoxide (see Scheme 2-1, A). The following reaction of superoxide (a singlet molecule) and organic matter in the singlet state is then allowed. The oxidized metal ions consequently need to be reduced by a suitable reducing agent to be able to perform the next electron transfer reaction.

A third method of oxygen activation is single electron transfer from organic molecules that can stabilize radicals. These cofactors, such as flavins or NADH, can similarly directly react with dioxygen (see Scheme 2-1, B). A hydrogen atom is transferred from flavin to triplet oxygen, generating two radical molecules. Upon spin inversion the reaction partners collapse into flavin hydroperoxide.<sup>[17]</sup> A last possible pathway to allow for reaction with oxygen is activation of the substrate. While the reaction of triplet oxygen with organic material, that contains only paired electrons, is quantum chemically forbidden, the reaction with radicals is a spin-allowed process. Such reactions occur when substrate molecules can stabilize radicals, such as in the case of intradiol catechol dioxygenase (see Scheme 2-1 C).<sup>[18]</sup> The required substrate activation has been studied using artificial catechol iron complexes. It is believed that the Lewis acidity of the iron atom increases the covalent character of the iron-catechol interactions, which in turn enhances the semiquinone character of the bound ligand and makes the catechol ligand more prone to oxygen attack.

While there are many examples in which oxygen activation can be performed under participation of metal ions (even earth abundant transition metals such as molybdenum, as in the case of xanthine oxidoreductase<sup>[19]</sup>) or organic cofactors, only few enzymes catalyze cofactor and metal-free oxygen activation.

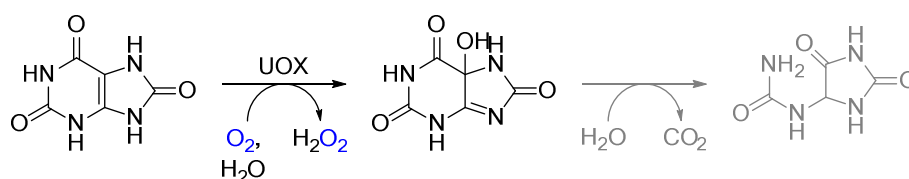




**Scheme 2-1:** Enzymatic approaches to oxygen activation. **A:** Electron transfer from a metal to triplet oxygen to generate superoxide, which can undergo reactions with organic material. **B:** Hydrogen atom transfer from flavin to molecular oxygen, followed by recombination of superoxide with the resulting flavin radical. **C:** Metal assisted radical formation on the substrate, with subsequent recombination of substrate radical and  $^3O_2$ .

### 2.1.1 Cofactor-Independent Oxidases

One particular and well studied example of cofactor-free oxidases is the urate oxidase (UOX). The enzyme belongs to the purine degradation cascade and more specifically catalyzes the hydroxylation of uric acid (see Scheme 2-2). By the use of EPR spectroscopy, a number of radicals are readily observable upon simple mixing of UOX and uric acid.<sup>[21]</sup> These different radical species were assigned to be of the same origin, since the structure of uric acid allows for easy delocalization of the unpaired electron throughout the molecule. Different radical traps then are capable of trapping different products depending on the delocalization of the radical. How UOX does activate oxygen or the substrate remains unclear. Typically, oxidases that do not use cofactors or metals activate oxygen first, which then reacts as superoxide radical anion ( $\cdot O_2^-$ ) leading to peroxidation of the substrate. However, no superoxide or hydroxyl radicals were observed in case of UOX, indicating that these either do not form or do not leave the active site.<sup>[21]</sup>

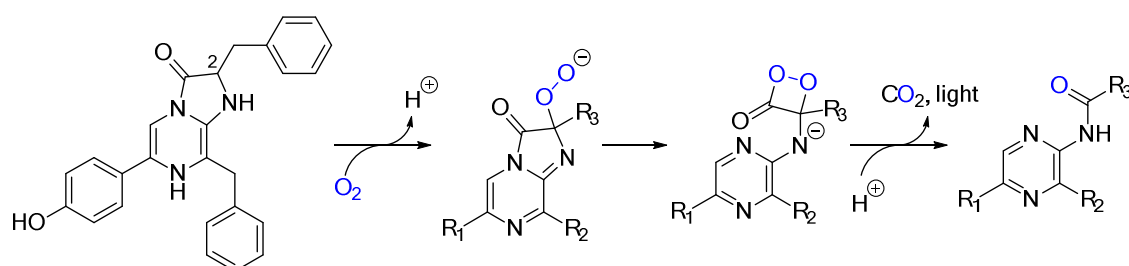


**Scheme 2-2:** Reaction catalyzed by UOX under participation of molecular oxygen and the follow-up reaction to yield allantoin, the final degradation product.

---

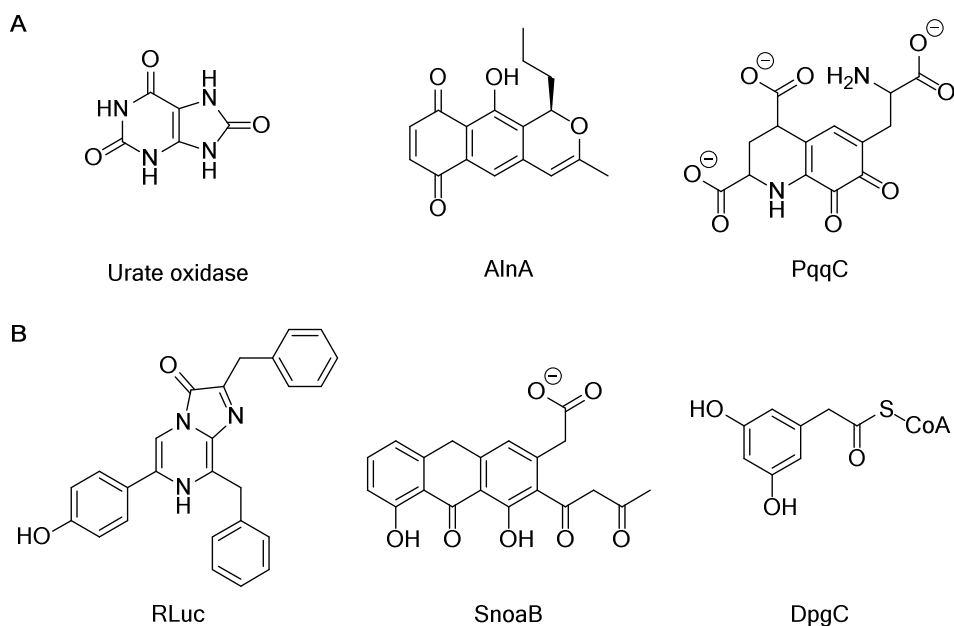
### 2.1.2 Cofactor-Independent Oxygenases

Similarly to cofactor-free oxidases, oxygenases exist that do not require an external electron donor for the reaction to be catalyzed. In biotechnology the green fluorescent protein (GFP) from marine organisms such as *Renilla reniformis* (a sea pansy), has grown to a laboratory standard within the last decades. This was mainly due to its ease to be produced and the characteristic fluorescent emission upon irradiation with light. Luciferase (RLuc) is an enzyme also found in *R. reniformis*. When the sea pansy is disturbed, a cascade of nervous system responses triggers the release of  $\text{Ca}^+$  ions which in turn causes the release of coelenterazine (the substrate of RLuc) from the coelenterazine binding protein. RLuc then catalyzes the monooxygenation of the substrate and releases the product as well as carbon dioxide while emitting blue light in the process. In the organism, resonance energy transfer occurs, transmitting the energy to GFP which then emits its characteristic green light ( $\lambda_{\text{em,GFP}} = 509 \text{ nm}$ ). The luciferase reaction is believed to proceed via an attack of dioxygen to C2 of the substrates imidazole ring (see Scheme 2-3). Intramolecular attack could then lead to the release of carbon dioxide and light. Imidazolopyrazinones undergo easy dioxygenolysis in aprotic solvents, when a suitable base is also present, making luminescence visible even in the absence of a catalyst.<sup>[20]</sup> Hence, the luciferase ‘simply’ has to do little more than providing a suitable binding site as well as a catalytic base.



**Scheme 2-3:** Proposed reaction pathway of luciferase from *R. reniformis*.<sup>[8]</sup>

The examples of cofactor-free oxidases and oxygenases shown above share one common feature: the basic structure of the substrate is aromatic and rich in carbonyls. With regards to other cofactor-free oxidases and oxygenases (see Scheme 2-4) this trend continues on. Without any doubt, analogies to cofactors such as flavin adenine dinucleotide (FAD, see also Scheme 2-1) exist. One can interpret these substrates as cofactor-like molecules, which perform similar tasks and can consequently act as electron donor itself. The requirement for external cofactors is thereby omitted.



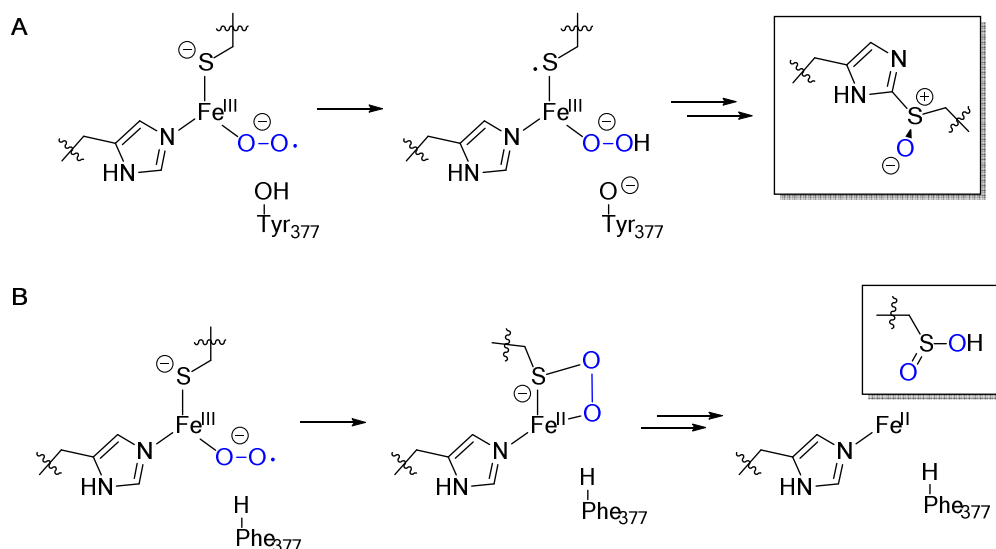
**Scheme 2-4:** Substrate scope of cofactor-independent oxidases (**A**) and oxygenases (**B**) for which substrate assisted oxygen activation has been observed.<sup>[21]-[25]</sup>

## 2.2 Metals in Enzyme Active Sites

The cellular machinery is tightly connected to the presence of metal cations catalyzing a variety of reactions within numerous metalloenzymes. Examples range from ‘simple’ tasks such as protein structure stabilisation, signalling, Lewis acid functionalities but also more complex reactions such as electron transfer and oxygen activation. These metal-cofactors are very simple but at the same time highly versatile cofactors. Based on their specific chemical and physiological function, a large variety of metals have found their way into the cells of living organisms.<sup>[26][27]</sup>

### 2.2.1 Iron enzymes

Among the many metal cations found in nature iron plays a prominent role. Heme coordinated iron is the central cation in the oxygen transport system of eukaryotic organisms. It therefore most likely presents the best known metal in the human body even in the non-scientific community. While there are numerous enzymes using heme-coordinated iron (i.e. P450 monooxygenases) an array of protein catalysts can tightly bind single iron atoms in their active centres. These enzymes similarly perform a wide range of reactions involving molecular oxygen. Non-heme iron binding sites typically contain histidine and carboxylate residues but also tyrosines and other hydroxy groups are often involved.<sup>[14]</sup> One example of such iron binding enzymes is the sulfoxide synthase EgtB, which is an essential enzyme in the biosynthesis of ergothionine, a 2-thio-histidine.<sup>[28]</sup> As with many metal enzymes, oxygen activation is achieved by electron transfer from the iron center towards O<sub>2</sub>. Oxygen is then reduced by electron transfer from  $\gamma$ -glutamyl cysteine, yielding a thiyl radical and an iron coordinated peroxide derivative (see Scheme 2-5, A). The sulfur radical then further proceeds to the product by C-S bond formation with trimethyl histidine. Quite interestingly, upon mutation of a single active site residue



**Scheme 2-5:** Proposed mechanism of EgtB from *Mycobacterium thermoresistibile*. Oxygen activation occurs upon binding to the iron center where one electron is transferred to generate the peroxyl radical anion.

(Tyr377Phe) this reactivity can be altered to generate a cysteine dioxygenase activity (see Scheme 2-5, B). Removal of the hydrogen bonding tyrosine residue changes the protonation state of the iron bound oxygen. This then weakens the oxygen-oxygen bond, which in turn yields a favourable O-O bond cleavage. One oxygen atom is then transferred to the sulfur atom of  $\gamma$ -glutamyl cysteine to generate the oxidized product.

This example shows the wide array of oxidation reactions that can be catalyzed by iron center enzymes even if only small changes are introduced into the active site.

### 2.2.2 Copper oxidases and oxygenases

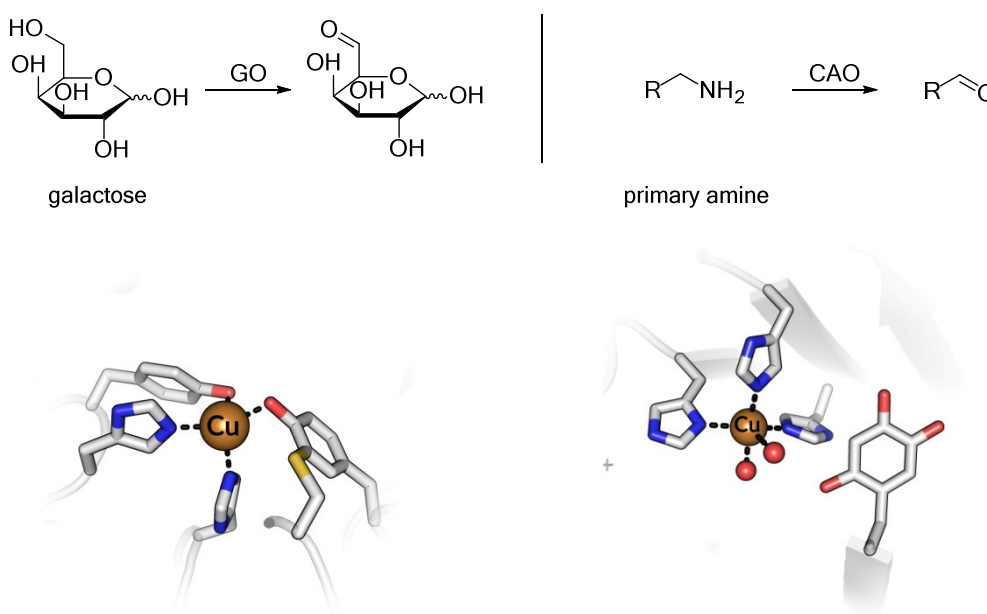
Aside from iron, another very prominent metal in enzyme active sites is copper. Copper plays an important role in living systems. Although being present in only the parts per million concentration range in the living cell, it is a key element in a variety of biological oxidation and reduction reactions.<sup>[30][31]</sup>

Copper in biological systems can most often be found in two distinct redox states. The vast majority of proteins display a copper(II) active site. Noteworthy and well studied examples for this class of copper enzymes are galactose oxidase and copper amine oxidases.<sup>[32][33]</sup> Galactose oxidase has been known since the late 1950s and hence is one of the best known copper containing enzymes. Nevertheless it bears quite an interesting active site (see Fig. 2-2 left). The copper(II) ion is coordinated in a square pyramidal geometry by a number of histidine and tyrosine residues. More interestingly, one of the tyrosine residues has undergone post translational modification to afford a cofactor which contains a C-S bond linked structure of a tyrosine and a cysteine residue. It is believed, that the *ortho* configuration of the carbon-sulfur bond favours the delocalization of the  $\pi$ -system towards the sulfur, which can stabilize a radical on the tyrosine hydroxyl group. A tryptophane residue right in place on top of the cysteine-tyrosine cofactor further stabilizes the delocalized system.<sup>[34]</sup> This unique architecture then reduces the redox potential of said tyrosine residue from  $\approx 0.9$  V to  $\approx 0.4$  V.

Copper-dependent amine oxidases (CAOs) represent a different example of copper dependent enzymes that additionally contains an enzyme originated cofactor. CAOs catalyze the oxidation of

a primary alcohol to an aldehyde functionality. They display an ubiquitous group of enzymes that can be found in various domains of life such as bacteria, yeast, plants and mammals.<sup>[35]</sup> The active site contains copper(II) which is held in place by coordination to three histidine residues and a number of water molecules. In addition to the copper ion the active site contains a tyrosine derived cofactor, namely 2,4,5-trihydroxy-phenylalanine (topaquinone, TPQ). Interestingly copper is only required for the cofactor genesis, but does not participate in the actual turnover, where copper coordinated water provides hydrogen bonding towards the deprotonated cofactor. The substrate amino group forms a Schiff base with C3 of TPQ which hydrolyses to form the aldehyde (similar to the transamination mechanism). The amino group resides on the cofactor and under participation of oxygen and water is again hydrolysed to the aldehyde containing recycled TPQ.

The cofactor alone is believed to adopt a rather flexible orientation. When TPQ is coordinated to copper (adopting the “on-copper” conformation) the enzyme is in fact inactive. Upon binding of C4-oxygen to the copper ion the orientation is fixed, which causes the cofactor to face away from the substrate entry channel and hence inactivates the complex.<sup>[36]</sup> This being said, the multifarious traits of copper in enzymatic active sites are clearly shown.

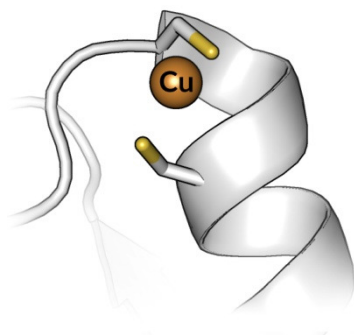


**Fig. 2-2: Top:** Reactions catalyzed by galactose oxidase (GO) and copper-dependent amine oxidases (CAO) **Bottom:** Active site view of galactose oxidase (left, PDB code 2EIC) with the tyrosine-cysteine cofactor and copper dependent amine oxidase (right, PDB code 1IU7) with TPQ.<sup>[37][38]</sup>

---

### 2.2.3 Copper(I) Enzymes

Copper binding motifs in Cu<sup>2+</sup>-enzymes display related active site structures with mainly nitrogen and oxygen based ligands. Different protein structures, however, have evolved that specifically bind Cu<sup>+</sup> ions. Although examples exist that bind copper(I) in a histidine rich active site (i.e. Hemocyanin, contains two copper(I) atoms that upon irreversible oxygen binding are oxidized to Cu<sup>2+</sup> [39][40]), most copper(I) binding sites typically contain an increased number of cysteine residues and show high affinities to the metal cation. Further differences arise with regards to the function of these enzymes. Copper(II) containing proteins often catalyze redox reactions which involve electron transfer from copper to substrates and cofactors.<sup>[31]</sup> Quite differently, the function of many Cu(I) proteins is copper transport in which the redox state of the metal ion is not affected. A typical example of these chaperon-like proteins is the Cu(I)-binding protein Atx1.<sup>[41]</sup> Being involved in copper transport to a second copper-containing protein, the P-Type ATPase, it requires to tightly bind copper but also efficiently release the metal at a later stage. Copper is delivered across the membrane to the *trans*-Golgi network where it will be further distributed to other metal proteins.<sup>[42]</sup> The metal binding site of Atx1 is dominated by a conserved MxCxC motif located on a solvent accessible and flexible loop (see Fig. 2-3). Similar to other copper chaperons like Ccs1 (a copper transporter for superoxide dismutases<sup>[43]</sup>) and Cox17 (delivering copper ions to cytochrome c oxidase<sup>[44]</sup>) this cysteine rich motif ensures tight binding of copper(I) with a  $K_D$  of  $\approx 10^{-17}$  M.

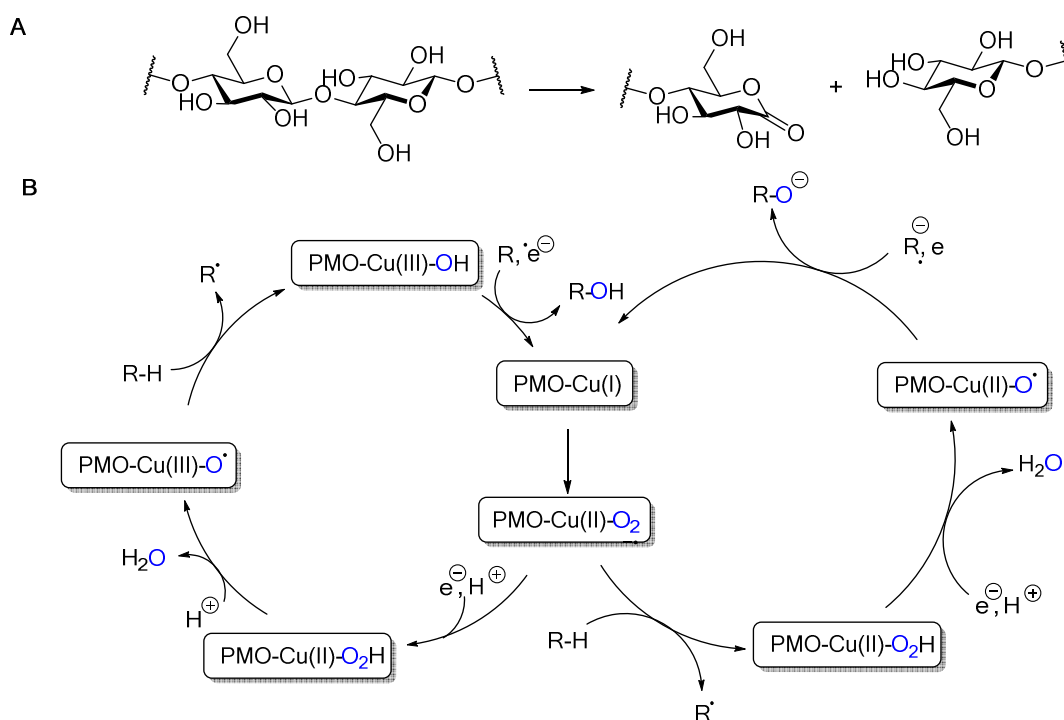


**Fig. 2-3:** Active site view of human Atx1 (PDB code 1FEE) with Cys12 and Cys15 being involved in copper binding.

### 2.2.4 Lytic polysaccharide monooxygenases

In the last years copper dependent lytic polysaccharide monooxygenases have gained reasonable interest due to their capability of breaking glycosidic bonds to produce biofuel from organic materials. They are found in many different organisms such as fungi and bacteria, but also viruses.<sup>[67]</sup> They do not share remarkable sequence homology, however, the overall fold as well as the structure of the active site appears to be similar.

The central feature of LPMO catalysis, is the copper atom at the surface of the reactive pocket. It is coordinated by two histidine residues (one of which is additionally methylated at the  $\delta$ -nitrogen in fungal LPMOs) as well as water molecules. These will be replaced by the hydroxy groups of the sugar upon substrate binding.<sup>[66]</sup>



**Scheme 2-6:** Overall reaction catalyzed by LPMOs (A) with two proposed mechanisms (B).<sup>[68]</sup>

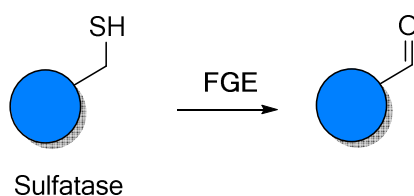
The mechanistic details of LPMOs have been subject to various research projects with different outcomes. In general, two distinct mechanisms are proposed (see Scheme 2-6) with different orders of action.<sup>[68]</sup> In both cases the catalytic cycle is initiated by oxygen activation via binding of molecular oxygen to the Cu(I) species in the resting state. The resulting superoxide radical anion species can then undergo direct reaction with the substrate (Scheme 2-6, B right pathway) to generate the substrate radical as well as the peroxide anion coordinated copper center. The latter is then reduced by a suitable reducing agent to generate a Cu(II)-oxyl radical which upon recombination with the radical substrate is transferred into the product.

The second suggested mechanisms involves the formation of a rather unusual Cu(III) species (Scheme 2-6, B left pathway) which is generated by heterolytic O-O bond cleavage and release of water after the initial superoxide radical anion species is reduced to the cupric peroxy species. The resulting Cu(III) species has been hypothesized in analogy to Compound I in P450s<sup>[69]</sup> and can be stabilized by a close by tyrosine residue in the active site. Model complexes have been synthesized that resemble the nature of this intermediate.<sup>[70]</sup> Hydrogen atom abstraction from the substrate is followed by radical recombination to yield the product as well as the cupric enzyme which can be reduced to the cuprous species, upon which the catalytic cycle is then closed.

## 2.3 Formylglycine Generating Enzyme

Nature has evolved a number of setups in order to tame the janus-faced temper of molecular oxygen. Examples are cofactor participation or redox active transition metals. However, other proteins exist, that do not require such assistance, but at the same time are capable of oxidizing their respective substrates in the absence of external support ((see chapter 2.1 and 2.2). These specialized enzymes however typically perform reactions on substrates with extended  $\pi$ -systems in polycyclic aromatic substrates. The oxidation of less electron rich substrates, though, still remains a challenge.

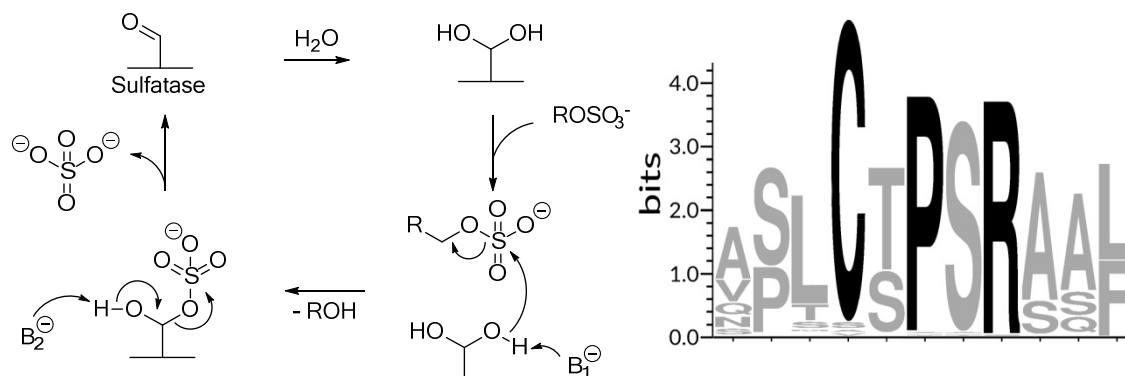
The Formylglycine generating enzyme (FGE) has been described as being capable of oxidizing a specific cysteine residue within a defined motif of sulfatases during translation (see Scheme 2-7).<sup>[3][45]</sup>



**Scheme 2-7:** Oxidation of a sulfatase cysteine catalyzed by the formylglycine generating enzyme.

### 2.3.1 Sulfatases in *Homo sapiens*

Sulfatases catalyze the hydrolysis of sulfate groups on various different substrates. The family of sulfatases contain a number of subgroups with different function, however they share a common active site residue. The catalytic activity requires one of the active site cysteines (in some sulfatases this can be a serine ) to be oxidized to formylglycine (fGly) which hydrolyses to the diol form in the mature sulfatase (see Fig. 2-4, left).<sup>[46][47]</sup> A base mediated nucleophilic attack towards the sulfate group of the substrate yields the product as well as the sulfated diol. Deprotonation of this then results in the release of sulfate as well as the aldehyde containing enzyme. The highly conserved cysteine residue is located in an CxPxR motif (see Fig. 2-4, right) that triggers substrate



**Fig. 2-4:** **Left:** Mechanism of aldehyde catalyzed sulfate ester cleavage by sulfatases. **Right:** Sequence logo showing the conservation of the active site motif CxPxR in 100 sulfatases homologues (after BLAST search) in *H. sapiens*.



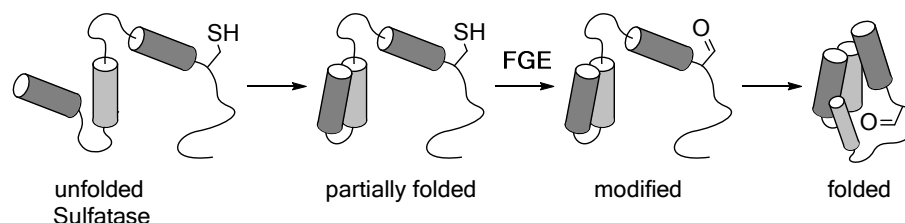
recognition by FGE. The cleavage of sulfate esters is therefore dependent on the presence of the fGly residue in the active site. Insufficient sulfatase activity leads to a rare but severe disorder called multiple sulfatase deficiency (MSD).<sup>[48]</sup> Patients suffering from MSD show very low activity of all sulfatases which has been traced back to dysfunction of FGE. It was shown that FGE not efficiently introduces the active site aldehyde in these patients. Gene sequencing revealed a number of mutations on the sulfatase modifying factor 1 gene (*SUMF1*) which expresses to the formylglycine generating enzyme.<sup>[49]</sup> These mutations result in deletions of sequences within FGE (loss of residue 327-366 upon deletion of G661 in *SUMF1*) or altered conserved residues throughout the enzyme scaffold, but mostly in the C-terminal subdomain of FGE (i.e. R349W/Q or C336R). While the individual gene defects were of different nature, seven out of seven patients showed mutations or deletions in subdomain three of FGE (Arg327-Ala366, C-terminal domain), suggesting the particular importance of this domain.

### 2.3.2 FGE in human cells

In humans, after expression of *SUMF1* in the endoplasmic reticulum (ER), FGE remains localized in the ER where it ultimately performs catalysis on the respective sulfatases. FGE does not contain a signaling peptide sequence of the KDEL type, suggesting that interactions with other proteins cause FGE to remain in the ER.<sup>[49]</sup> It interacts with protein disulfide isomerase (PDI) and a ER resident protein (ERp44), both of which are members of the thioredoxin family as well as ERGIC-53, a lectin, moving between the golgi and the ER.<sup>[50]</sup> The interaction with ERp44 occurs through a N-terminal domain of FGE (residue 34-68 of the human variant) through disulfide bond formation as well as non-covalent interactions.<sup>[51]</sup> This N-terminal domain has been identified as being crucial for retention in the ER and hence for *in vivo* activity, but appears to be dispensible for *in vitro* reactivity. It contains two cysteine residues (Cys50 and Cys52) that are involved in binding to ERp44 but also lead to homodimer or intramolecular disulfide bond formation.<sup>[45]</sup> The homodimer, however, appears to be labile and is readily reduced to the monomers in the presence of reducing agent. The *in vivo* activity only depends on the presence of Cys52 but not Cys50.<sup>[51]</sup>

The enzyme additionally contains a N-glycosylation at Asn141, which however is not essential for the function.<sup>[52]</sup> A part of FGE is secreted from the ER and upon cleavage of the N-terminal domain (residue 34-72) the majority is inactivated. However, 20-30 % are not truncated and remain active.<sup>[53]</sup> The glycosylation at Asn141, though, does not influence secretion.<sup>[52]</sup>

Modification of the sulfatase substrate occurs inside the ER where FGE catalyzes the oxidation co-translationally with expression of the substrate gene.

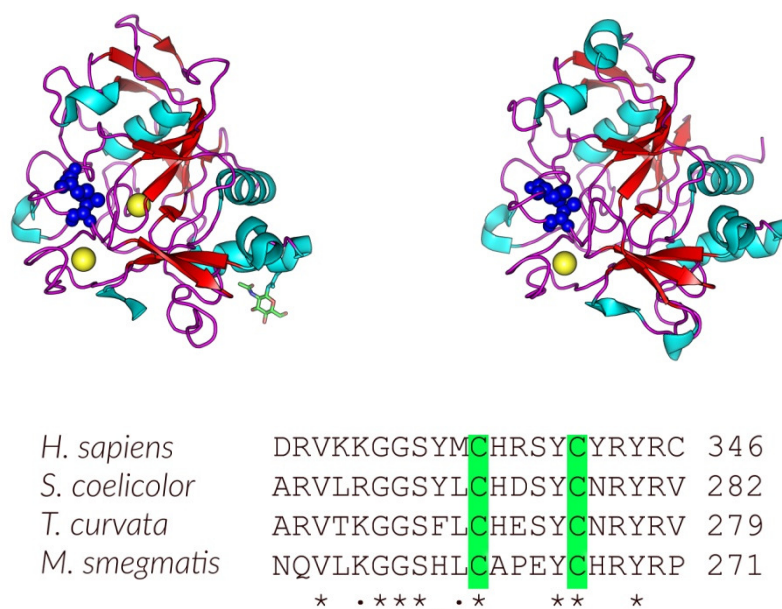


**Fig. 2-5:** FGE catalysis on sulfatases occurs co-translationally in the endoplasmic reticulum.

### 2.3.3 Structure of FGE

Crystal structures of FGE have been published for enzymes (*Streptomyces coelicolor*) from bacterial and eukaryotic (*Homo sapiens*) organisms. The structures from the different sources share the same fold with a rmsd of 0.65 Å and provide excellent tools for researchers investigating FGE.<sup>[54]</sup> The overall fold of the enzyme is quite unique and has hence been described as the FGE-fold with 64 % adopting a random coiled secondary structure (see Fig. 2-6, top). This fold has been identified as a subclass of the C-type lectin fold.<sup>[55]</sup> While human FGE contains two Calcium atoms, the bacterial homologue only binds one. Asparagine 141 in human FGE is N-glycosylated with a N-acetyl glucosamine.

Both, human and bacterial, FGE contain a number of conserved cysteine residues. The eukaryotic variant contains six cysteines that are oxidized to cystines, while the prokaryotic protein only contains five, with one disulfide bond between Cys272 and Cys277. Three out of these cysteines are highly conserved, with two of them being positioned in the active site of the enzyme (see Fig. 2-6, bottom). Additional crystal structures of human FGE show these two residues under various oxidizing conditions. Sulfenic and sulfonic acids have been observed. A cysteine mutant of human FGE (Cys336Ser) has additionally been cocrystallized with a substrate peptide (sequence: LCTPSRA), showing a disulfide bond between the remaining Cys341 and the substrate cysteine. This furthermore led to the identification of the substrate binding groove on the surface of FGE. In addition to contacts to water molecule and the peptide backbone, other key interactions between the substrates conserved CxPSR motif and FGE are hydrogen bonding of the serine hydroxy group with Asn352 and of the arginine with Asp154, Asp355 and Ser357. With a total number of six van-der-Waals contacts, the terminal arginine residue appears to be of high importance (see also sequence conservation in sulfatases, chapter 2.3.1).



**Fig. 2-6: Top:** X-ray structure of FGE from *Homo sapiens* (left, PDB 1Y1E<sup>[48]</sup>) and *Streptomyces coelicolor* (right, PDB 2Q17<sup>[54]</sup>) colored according to secondary structure (helices: cyan;  $\beta$ -sheet: red; random coiled: magenta). Cysteine336 and Cys341 are colored in blue. Calcium atoms are shown in yellow. **Bottom:** Sequence alignment of FGE from *Homo sapiens*, *Streptomyces coelicolor*, *Thermomonospora curvata* and *Mycobacterium smegmatis*, with conserved active site cysteines highlighted in green.

---

The availability of FGE crystal structures has furthermore driven mechanistic investigations. The covalently linked substrate in the active site being one main feature of the proposals, it remains unclear how oxygen activation can occur in the absence of metals or cofactors.

#### 2.3.4 FGE mechanism

The FGE catalyzed reaction is one of the most remarkable oxidation reactions known. The oxidation of the thiol group without participation of metals or cofactors is complex and a complete understanding of the underlying mechanisms still lacks. After the first identification of FGE it has been shown to be dependent on oxygen, most likely as terminal electron acceptor.<sup>[45]</sup> Additionally, the dependence on reducing agent has been demonstrated. Cell extracts containing FGE were assayed in the presence of GSH and DTT, upon which an elevated aldehyde formation was observed.<sup>[3]</sup> In search of answers regarding the basis for oxygen activation various crystal structures have been reported including bacterial (*Streptomyces coelicolor*) as well as human enzymes (see above page 14).<sup>[48][54]</sup> The human FGE contains six cysteine residues. Two of these cysteines have been identified as being highly important for activity. In fact, upon co-crystallization of FGE with a short peptide representing the substrate binding motif, a mixed disulfide between the substrate and the enzyme is observed.<sup>[56]</sup> Additionally, the importance of the two highly conserved cysteine residues Cys336 and Cys341 (numbering based on human variant) has been shown in serine mutants of the respective amino acid which show only residual activity *in vivo*. Several mechanisms have since been proposed, most of which include disulfide bond formation between the substrate and the enzyme (see Fig. 2-7) which is the first step of the proposal. This mixed disulfide is then subject to the formal addition of oxygen, being the most questionable step as oxygen activation is required for this reaction to occur. However, it follows that the sulfur-peroxyde intermediate is reduced by an appropriate reducing agent which leads to the formation of a first equivalent of water as well as the sulfenic acid on the enzymes cysteine. Subsequent disulfide bond shuffling releases the enzyme in its resting state as well as the substrate sulfenic acid, which contains the necessary oxidation equivalent to decompose to another equivalent of water and the thioaldehyde. Hydrolysis of this, finally yields the aldehyde containing product.

The direct reaction of oxygen with the FGE cysteine, however, is thermodynamically not

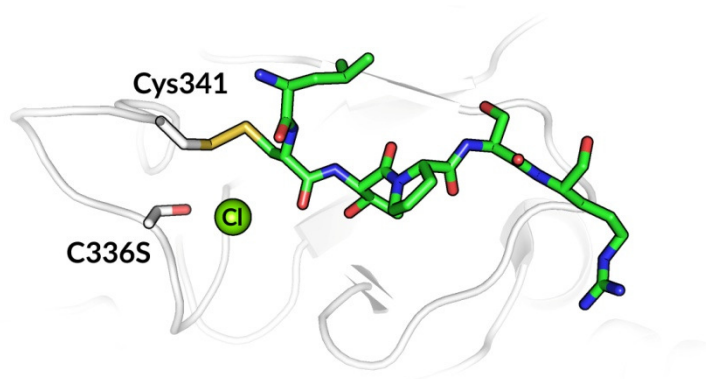
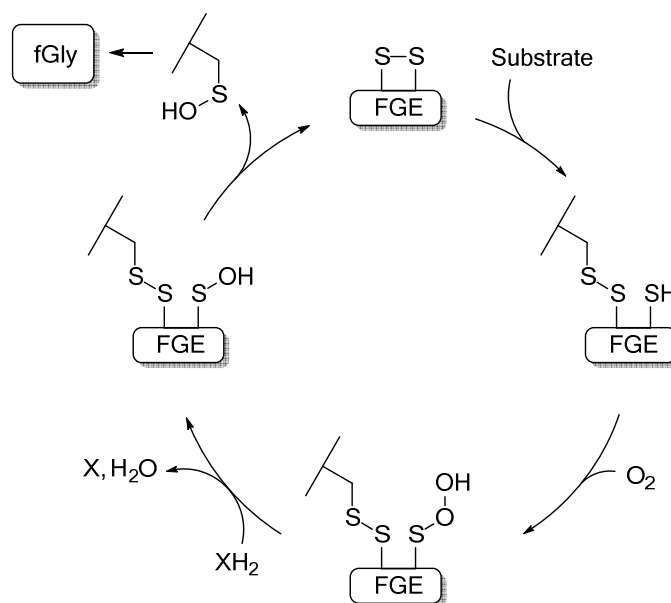


Fig. 2-7: Active site view of human FGE (PDB code 2AIK) with Cys341 and Cys15 being involved in copper binding.

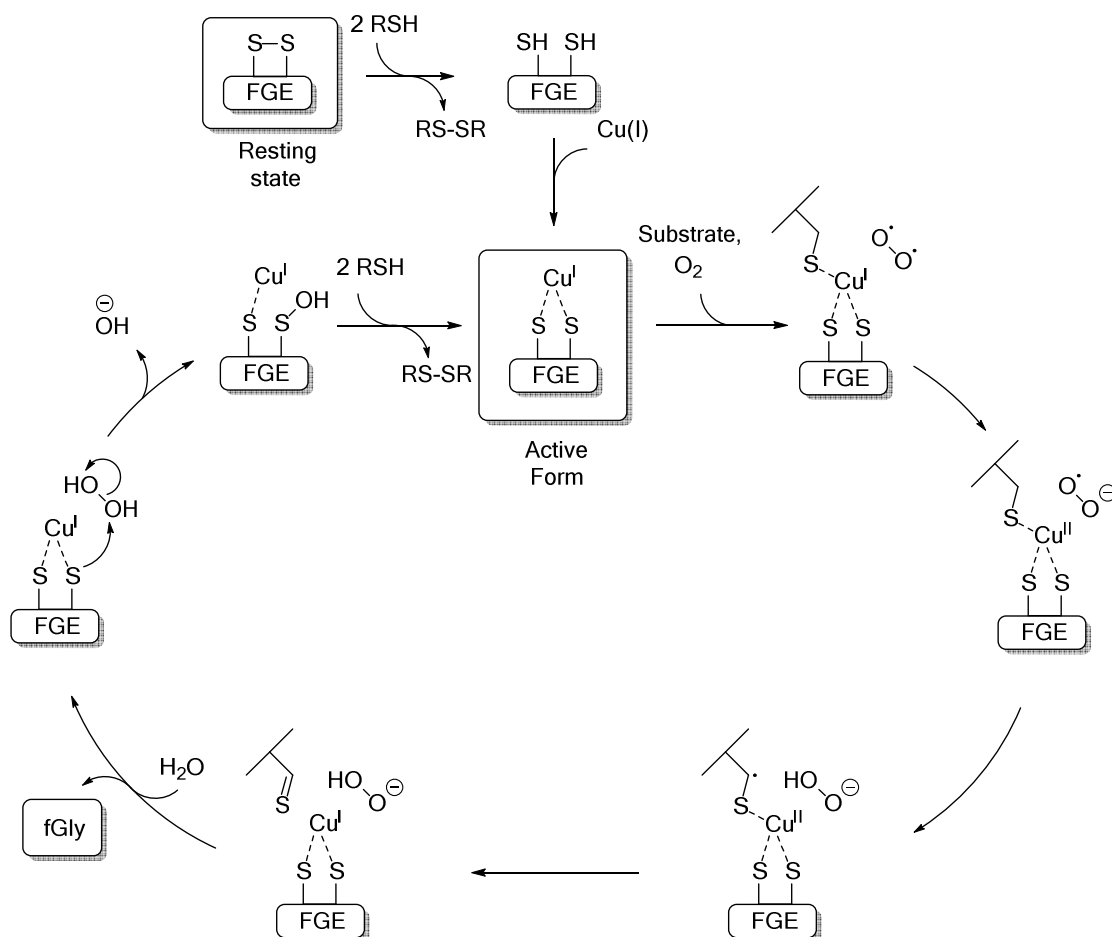


**Scheme 2-8:** Proposed mechanism of fGly formation without participation of cofactors. The proposed disulfide bond intermediate was observed in crystal structures.

favorable.<sup>[57]</sup> Single electron transfer from sulfur to oxygen leads to the generation of two unstable products (superoxide ion and HS $\cdot$ ). A two electron transfer would be thermodynamically possible, but the presence of two unpaired electrons in the highest occupied orbitals of  $^3\text{O}_2$  (see Fig. 2-1) that are similar in energy disfavours this process. The mechanisms hence required additional investigation to solve the riddle about oxygen activation.

The participation of metal ions has been excluded due to absence of iron, copper or zinc in purified protein preparations.<sup>[54]</sup> Even though traces of copper were found to be present in some cases (0.09 mol/mol FGE) enzyme samples which have been supplied with additional copper proved to be inactive.<sup>[45]</sup> However, according to the previous remarks, oxygen activation can only be achieved if a suitable electron donor is present. We therefore anticipated, that the mechanistic proposals do not reflect the complete truth of FGE catalysis.

Upon reinvestigation of the metal dependency of FGE we showed, that reactivity in the presence of copper largely exceeds the previously observed reaction rates.<sup>[64]</sup> The mechanistic proposal was then extended with the presence of copper (see Scheme 2-9). The resting state of the enzyme shows a disulfide bond in the active site, which in a first step will be reduced by a suitable reducing agent (*in vitro* often DTT). Reduced FGE is then able to bind copper(I), which afterwards has an open coordination site for oxygen to coordinate. This copper complex then transfers a single electron to dioxygen, generating the superoxide radical anion as well as Cu(II).  $\text{O}_2^-$  will then abstract an hydrogen atom from the substrate (similarly to other iron dependent FGEs<sup>[62]</sup>), which represents the rate limiting step. The substrate radical transfers an electron back to copper(II) yielding the thioaldehyde, which subsequently will hydrolyze to the fGly product. Hydrogen peroxide is not detectable in the reaction mixture, suggesting that oxygen is fully reduced to water.<sup>[65]</sup> One of the active site cysteines, attacks the peroxide anion to release a hydroxide and forms a sulfenic acid. This active form of FGE is then regenerated by reduction of the oxidized cysteine by the reducing agent present in solution.



**Scheme 2-9:** Mechanistic proposal of fGly formation under participation of copper as a cofactor. Binding of copper was shown in kinetic analyses.

It has to be noted, that a second class of SAM dependent FGEs has been identified. While the substrate recognition motif is similar (only cysteine is replaced by serine), these FGEs are dependent on an iron sulfur cluster as electron donor as well as SAM for hydrogen atom abstraction of the substrate, but catalyze the fGly formation in the absence of oxygen. Hence the mechanisms happen to be unrelated to the one discussed here.<sup>[62][63]</sup>

The proposed reaction cycle shows various similarities to the LPMO-catalyzed oxidation. Firstly, oxygen activation occurs in an initial step, by electron transfer from copper(I) to oxygen and generation of the cuprous intermediate. Also, hydrogen atom abstraction has been shown to be the key step in the reaction and last but not least, the requirement of an external reducing agent. The reducing agent in the catalytic cycle of LPMOs was proposed to be a complex electron donor system.<sup>[71]</sup> In contrast, the reducing agent which is involved in the FGE reactivation *in vivo* is yet to be identified.

While there are differences in the mechanistic details, further comparisons between these two very distinct classes of oxygenases can most likely be drawn in the future, which will be of advantage for researchers working in both fields.

Again, LPMOs and FGE share similarities not only on the mechanistic level. The importance of LPMO for the biorefinery has been outlined before. Likewise, FGE has gained compelling interest in biotechnology as it opened up new possibilities to site-specifically introduce aldehyde

---

functionalities into virtually any recombinant protein. The defined recognition motif has been genetically introduced into recombinant proteins of interest upon which FGE catalyzed aldehyde formation is followed by chemical ligation techniques to introduce virtually any chemical label.<sup>[5][58]</sup>

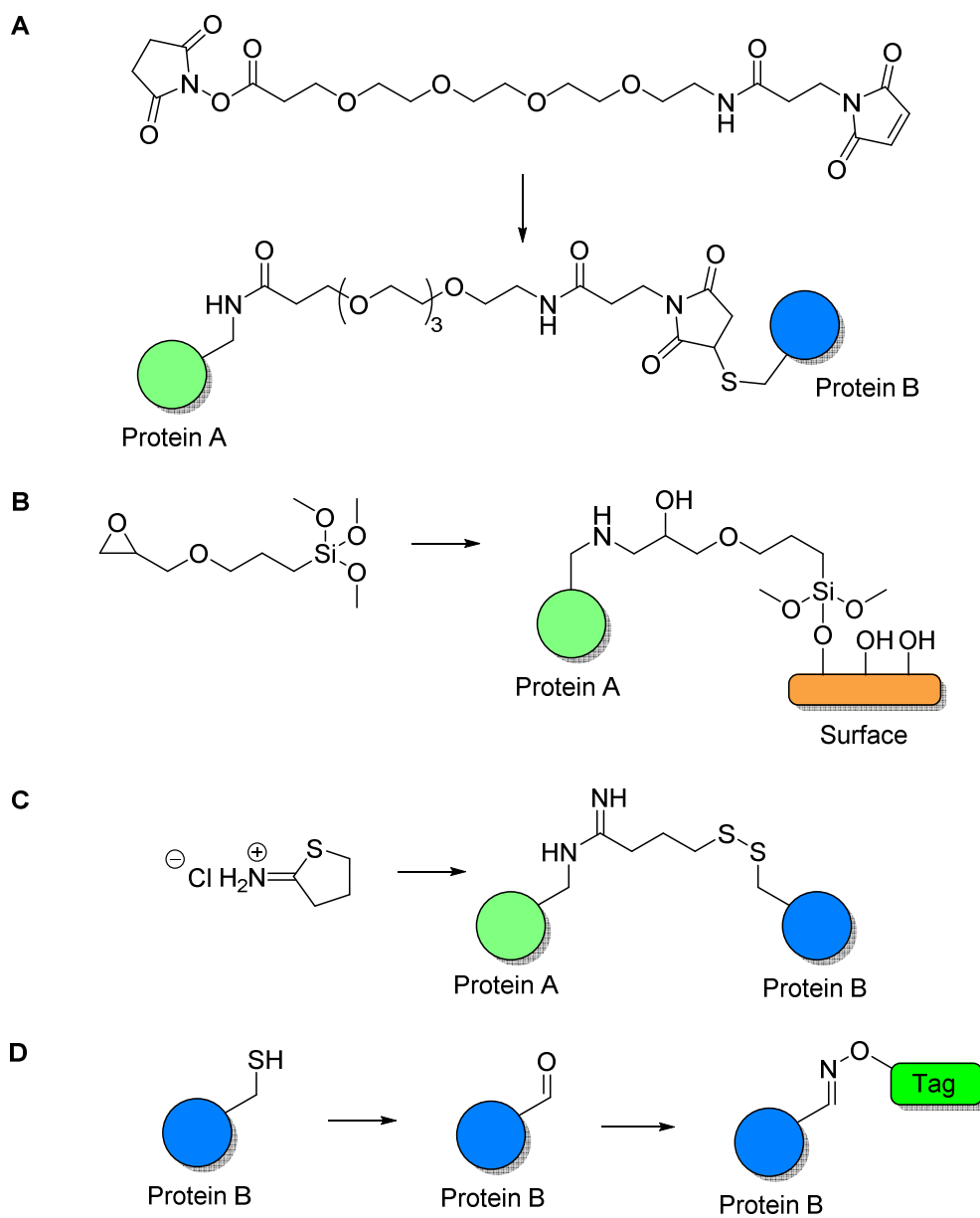
## 2.4 Bioconjugation

Ever since enzymes were first identified and isolated researchers have reached out to utilize the potential of biocatalysts to facilitate or simplify industrial processes. Shortly after the discovery of diastase, a starch hydrolysing enzyme discovered in the 1830s, first applications have emerged mostly to produce dextrin.<sup>[72]</sup> With the industrial revolution and medicinal and biochemical science fastly expanding the field of enzymology, other, more complex, applications emerged. The first enzymatic synthesis of isomaltose was developed in 1898 by Croft-Hill in which he reacted a yeast extract with a glucose solution.<sup>[73]</sup> Today enzymes are used in organic synthesis, in fermentations and to generate new functionalized materials.<sup>[74][75][80]</sup>

Bioconjugation is the 'simple' attachment of one molecule to another via a new covalent bond in which one of the two conjugates originates from a biological source or displays a biological function. Examples range from the attachment of fluorescent labels to antibodies, which in turn bind their specific antigen and hence enable visual input of their binding locus, to surface localized affinity ligands or even enzymes, which can perform catalysis on a solid surface.<sup>[76]</sup> The newly formed bond typically is formed by the use of a linker moiety which displays two distinct reactive groups on either side, enabling different chemistry to form covalent bonds (see Fig. 2-8). Examples for covalent attachment are the carbon-sulfur bond between maleimid and protein cysteins or esters of N-hydroxysuccinimid groups which can be displaced by lysine residues. While the designs shown in Fig. 2-8 provide good methods for interlinking proteins or labels, some do require extensive synthesis. Others are less demanding with regards to synthesis, but show other disadvantages like the lack of regio specificity.

In search of an easy to install and selective linker to combine any protein of interest with virtually any tag, researchers have reached out to utilize the remarkable reactivity of the formylglycine generating enzyme to generate a non-proteinogenic and thereby unique amino acid on the target protein, which can be covalently linked in a subsequent step. A sequence comparison of sulfatases from different organisms revealed that FGE recognizes a defined amino acid motif (with CxPxR being highly conserved).<sup>[78]</sup> However, little later it was shown, that replacement of proline or arginine by alanine, likewise leads to the formylglycine containing protein, indicating that the recognition motif is more flexible (at least when using FGE from *E.coli*).<sup>[79]</sup> The amino acid sequence can genetically be introduced into the structure of recombinant proteins, which will then be recognized by FGE. Upon coexpression with FGE in *E.Coli* the tagged proteins can be isolated with the aldehyde being installed at high conversion rates. After protein purification the aldehyde containing protein is reacted with hydroxylamines or hydrazines to yield the labeled protein. While initial setups were limited to C- or N-terminal modification, other designs with slightly extended recognition sites yield the formylglycine insertion into internal loop regions of for example GFP or even antibodies.<sup>[80][81]</sup> In addition to protein labeling with fluorophores others have demonstrated that formylglycine containing proteins can efficiently be immobilized on

nanocellulose or agarose.<sup>[80][82]</sup> With no doubt, FGE catalyzed aldehyde formation has found its way into current biotechnological methodologies.



**Fig. 2-8:** Different approaches to covalently link proteins or enzymes to other proteins, surfaces or functionalization tags. **A:** The reaction of maleimides with surface exposed cysteines and the replacement of N-hydroxysuccinimids by lysine residues. **B:** Silane decorated epoxide building blocks are immobilized on surfaces and subsequently attacked by lysine residues of the target protein.<sup>[77]</sup> **C:** 2-iminothiolane is used to covalently link lysines and cysteines. **D:** Selective oxidation of a defined cysteine residue to an aldehyde functionality is followed by chemical ligation with hydroxylamines.

---

---



---

### 3. Aim of this Thesis

Oxygen activation in FGE has been unexplained since the first discovery of the enzyme in 1995.<sup>[4]</sup> The most important residues in the active site have been identified as two conserved cysteine residues.<sup>[48]</sup> While initial mechanistic pathways were proposed based on structural observation, detailed biochemical investigations of the reaction mechanism were not available.

The identification and quantification of side products is vital for the understanding of the mechanism of a reaction. One first goal of this work was therefore to provide detailed investigations of the reaction stoichiometries.

Furthermore, the direct electron transfer of sulfur to oxygen as discussed before is not favourable. Consequently, one aim of this work was to elucidate how oxygen activation can occur in FGE even though no cofactors were believed to participate in the reaction. The identification of potentially overlooked cofactors and the investigation of their method of action is an essential part of this assessment.

FGE from *T. curvata* shows low activity at standard conditions *in vitro*. In an attempt to improve biotechnological applications, we furthermore assay possibilities to improve the utility of FGE for *in vitro* labeling reactions of recombinant proteins.

---

---

---

## 4. Formylglycine Formation – A Copper Dependent Reaction

Oxygen activation requires the transition of oxygen from a triplet to singlet state (see chapter 2.1). While enzymes such as cytochrome P450s or polysaccharide monooxygenases use metal centres to achieve electron transfer towards molecular oxygen, others (i.e. flavoproteins) have evolved cofactor dependent mechanisms, in which radicals stabilized on organic cofactors can then react with molecular oxygen.

The formylglycine generating enzyme has been described as a metal- and cofactor-independent monooxygenase. While other cofactor-independent enzymes that use oxygen as a terminal electron acceptor are known, FGE is a special case since its active site contains only proteogenic amino acid functional groups and lacks any polyaromatic moieties that are typically involved in cofactor independent oxidases and oxygenases. The two most important residues in FGE catalysis have been identified as two highly conserved cysteine residues at the end of a substrate binding groove. While spontaneous reaction of thiols with oxygen can be observed in many biochemical reactions, this type of oxidation typically leads to oxidized sulfur species such as disulfides or sulfenic acids.

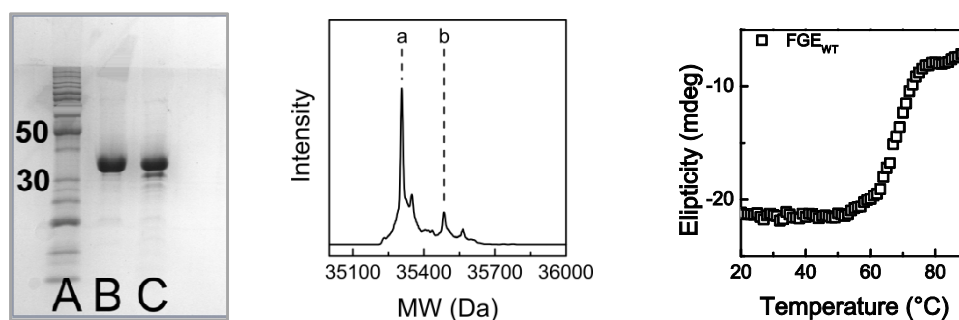
Mechanistic proposals for how the FGE catalyzed aldehyde formation occurs previously contained the formation of a mixed disulfide with the substrate. This was followed by the formal addition of O<sub>2</sub> and subsequent reduction to the sulfenic acid. The details of how sulfur mediated oxygen activation should occur, however, have remained unsolved. Direct reaction of a cysteine residue with molecular oxygen is a spin forbidden process. In contrast, thiols often show high affinities to metal ions, such as zinc or copper. We therefore propose that the mechanistic hypotheses outlined prior to this work did not resemble the complete truth of FGE catalysis and might lack one key element that can explain the observed oxidation reaction with the substrate.

In the following chapter we show that FGE catalysis requires equimolar amounts of reducing agents and that hydrogen sulfide is released in quantitative amounts. Furthermore the dependency of FGE catalysis on copper(I) is demonstrated.

### 4.1 *In vitro* Reconstitution of FGE with Copper(I)

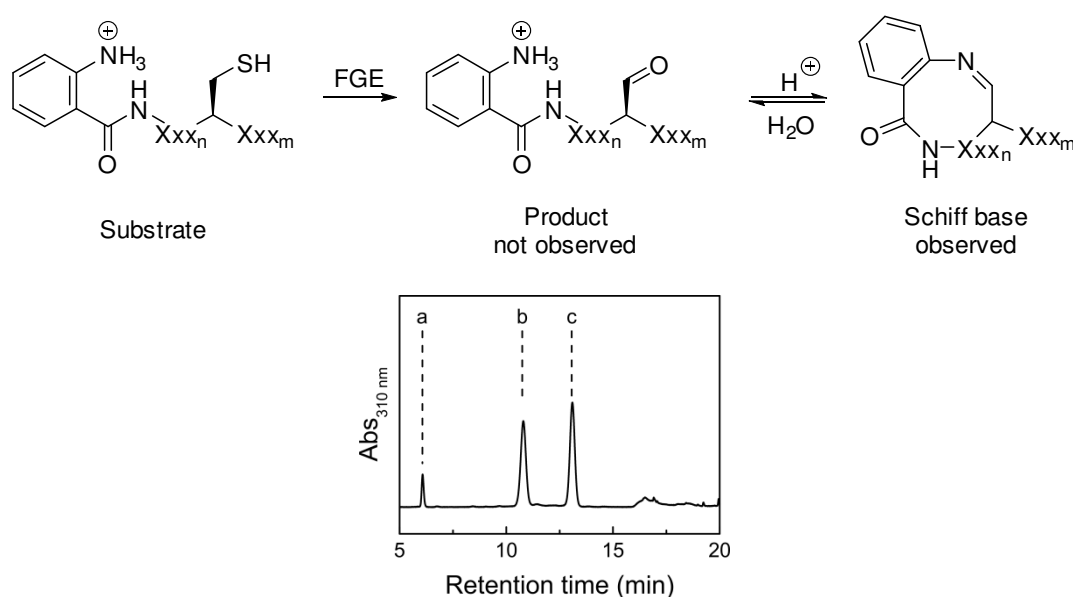
#### 4.1.1 Establishing a robust study system

To further advance FGE research, we selected an FGE variant from *Thermonomospora curvata*, an organism first isolated from stable manure, and known to prefer elevated temperatures.<sup>[86]</sup> The selected variant shares 58 % sequence homology with the human variant (FGE<sub>human</sub>). Conclusions made for FGE from *T. curvata* should thereby allow for comparisons with the FGE<sub>human</sub>. Similarly, published data obtained for FGE from *H. sapiens* and *S. coelicolor* should also be applicable to our research (sequence homology of FGE from *curvata* to the streptomycete protein is 64 %). Working with proteins from thermostable organisms can have the advantage of potentially higher production and isolation yields and higher resistance towards thermal stress. This in turn can broaden the spectrum of possible applications for later use in biotechnology. Indeed, *E. coli* cell cultures in shaking flasks yield good amounts of 5 – 10 mg/L purified protein sample



**Fig. 4-1:** **Left:** SDS-PAGE picture of FGE<sub>WT</sub> after purification using Ni<sup>2+</sup>-NTA affinity chromatography. A: Molecular weight marker; B: FGE<sub>WT</sub>; C: FGE<sub>WT</sub> + DTT. **Middle:** Mass spectrum of FGE<sub>WT</sub> after protein purification showing the protein peak (a) as well as a glycosylation adduct on the His<sub>6</sub>-tag (b)<sup>[88]</sup>. MW<sub>calc,WT</sub> = 35308.1 Da; MW<sub>obs,WT</sub> = 35306.5 Da. **Right:** Determination of temperature of thermal unfolding by circular dichroism. FGE was diluted to a concentration of c<sub>FGE</sub> = 10 μM in a buffer containing 2 mM DTT and ellipticity was measured at λ<sub>CD</sub> = 222 nm while heating at 1 °C/min.

(see Fig. 4-1) after purification by Ni-affinity chromatography. This can be increased up to 20 - 30 mg/L of medium when performing a controlled fermentation production. The HRMS spectra of FGE from *T. curvata* (FGE<sub>WT</sub>) after overexpression and isolation from BL-21 *E. coli* cells shows the expected mass after cleavage of the initial methionine (MW<sub>calc,WT</sub> = 35308.1 Da; MW<sub>obs,WT</sub> = 35306.5 Da). The observed mass difference of ΔMW = 2 Da can be attributed to the formation of a disulfide bond within the protein structure. During *in vitro* experiments this protein preparation is active in the presence of reaction buffer supplied with substrate and DTT as a reducing agent.<sup>[3]</sup> Moreover, the protein is stable up to a temperature of T<sub>unfold</sub> = 55 °C with no sign of thermal unfolding and an activity maximum at a temperature of T<sub>active</sub> = 45 °C<sup>[87]</sup> and a pH optimum of around 10-11 (see appendix Fig. 4-11). The reaction rates can readily be determined using a synthetic peptide



**Fig. 4-2:** Product identification by HPLC using the specific absorbance of the Abz group at λ<sub>Abz</sub> = 310 nm. **Top:** FGE catalyzed aldehyde formation is followed by acid catalyzed schiff base formation. The aldehyde product is not observed. **Bottom:** Representative HPLC chromatogram showing the oxidized DTT (a), the substrate peak (b) and the product peak (Schiff base, c, MW<sub>calc,schiff base</sub> = 987.5 Da, MW<sub>obs, product</sub> = 987.5 Da).

encoding the recognition motif CxPxR (peptide sequence: Abz-SALCSPTRA-NH<sub>2</sub>). After aldehyde formation, the product undergoes acid catalyzed Schiff base formation with the ortho amino group of the terminal Abz group further facilitating peak separation (see Fig. 3-2).

#### 4.1.2 Analysis of the FGE catalyzed reaction

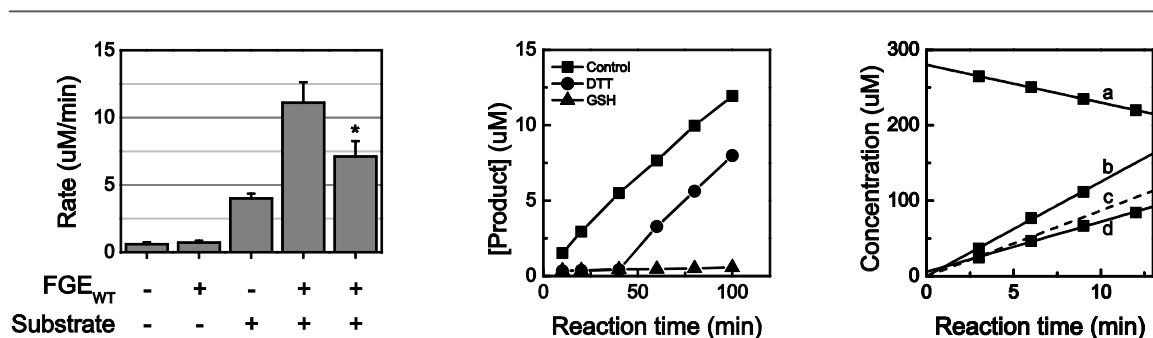
The main consensus about the mechanism of FGE is that the two highly conserved active site cysteines are involved. The reactivity of protein sulfur groups is largely determined by their redox state, hence our concerns lied with the oxidation state of the two active site cysteines. Crystal structures of the FGE from *H. sapiens* and *S. coelicolor* contain a disulfide bond between the two active site cysteines. In order to simplify the FGE system and to avoid disulfide shuffling side effects we constructed a FGE variant in which all cysteines, aside from the active site residues, are mutated to either alanine or seine (FGE<sub>C187A,C231A,C284S,C298A</sub> = FGE<sub>4C</sub>). The variant is readily producible in BL-21 *E. Coli* cells. The protein yields however, happen to be lower ( $m_{4C} = 2$  mg/L of culture vs.  $m_{WT} = 6$  mg/L of culture) which is to be expected after introduction of four mutations into the protein structure. Regardless of the lower yields, the isolated protein sample gave the expected mass in HRMS measurements ( $MW_{calc,4C} = 35195.8$  Da;  $MW_{obs,4C} = 35193.8$  Da), again indicating the absence of a cofactor or bound metals. With regards to the activity of the new variant we did observe an unexpected increase in reactivity under the standard conditions. The catalytic efficiency increased to  $k_{cat}/K_{M,4C} = 1000$  M<sup>-1</sup>min<sup>-1</sup> (see  $k_{cat}/K_{M,WT} = 100$  M<sup>-1</sup>min<sup>-1</sup>).

With the new variant FGE<sub>4C</sub> containing only the two cysteines residues involved in the reaction, we were now able to selectively address their redox state by chemical means. We used two published assays to quantify the amount of free thiols and disulfide bonds in FGE from *T. curvata* (see Tab. 4-1). The active site cysteines were found to be oxidized which is in good agreement with published crystal structures<sup>[56][54]</sup> as well as the mass increment of  $\Delta MW = -2$  Da in HRMS measurements. Alkylation attempts of FGE<sub>4C</sub> using 2-iodoacetamide (IAA) did not result in the labeled active site cysteines unless these were previously reduced by addition of DTT. This observation provides further evidence that the cysteines exist in an oxidized state (see Tab. 4-1).

Researchers have previously observed the dependence of FGE activity upon the addition of thiols *in vitro*.<sup>[3]</sup> In order to gain a complete understanding of the details of FGE catalysis, we quantified this correlation. Even in absence of FGE the oxidation of DTT is significant (see Fig. 4-3). However, during turnover the consumption of reduced DTT is enhanced. When comparing the rates of product formation, consumption of substrate and DTT oxidation, we

**Tab. 4-1:** Thiol and disulfide content of FGE from *T. curvata* determined using Ellmann's reagent<sup>[92][93]</sup> and the degree of alkylation after treatment of FGE<sub>4C</sub> with 2-iodo acetamide. <sup>[a]</sup>not determined

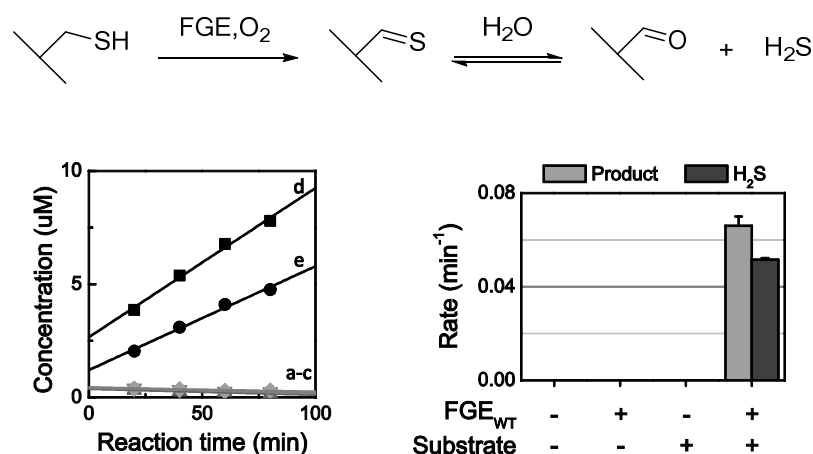
	SH/FGE	(SH + S-S)/FGE	IAA adducts
FGE <sub>WT</sub>	4.3 ± 0.2	3.8 ± 0.2	- <sup>[a]</sup>
FGE <sub>4C</sub>	0.4 ± 0.2	1.3 ± 0.1	0
FGE <sub>4C</sub> + DTT	- <sup>[a]</sup>	- <sup>[a]</sup>	2



**Fig. 4-3: Left:** Rates of spontaneous or FGE catalyzed DTT oxidation. In the absence of substrate the DTT oxidation rate is low. Only in the presence of substrate and even more under steady state conditions during conversion, DTT consumption is substantial. The asterisk (column 5) indicates the DTT oxidation rate during turnover corrected by subtraction of the background oxidation determined in the absence of FGE. **Middle:** Requirement of reducing agent. Product formation during the FGE<sub>4C</sub> catalyzed reaction in the absence of reducing agent is slow. Only upon addition of reducing agent the reaction rate increases, with GSH not having an effect. ( $c_{\text{FGE}} = 0.5 \mu\text{M}$ ,  $c_{\text{Cu(I)}} = 0.5 \mu\text{M}$ ,  $c_{\text{substrate}} = 50 \mu\text{M}$ ). **Right:** Rates of substrate consumption (a) and product formation (d) as well as the rate of DTT oxidation without (b) and with background correction (c, dashed).

observe a correlation between these three rates (Fig. 4-3, right), indicating turnover dependent oxidation of reducing agent. Indeed, in the absence of reducing agent FGE activity is reduced by 50 fold. Upon addition of reducing agents like DTT and cysteamine (2-aminoethane-1-thiol) to this reaction the reactivity is restored to full reactivity (see Fig. 4-3, middle). When using larger thiols (i.e. glutathione) or a non-thiol reducing agent (i.e. TCEP) the product formation remains slow (see also Appendix Fig. 4-14).

In order to further analyze the FGE catalyzed reaction, we then tried to identify other products aside the formylglycine residue. It has been previously stated that oxidation of the cysteine residue to the thioaldehyde is followed by hydrolysis to fGly. However, direct evidence for the detection of hydrogen sulfide ( $\text{H}_2\text{S}$ ) is missing. We therefore used a published assay to quantify the amount of  $\text{H}_2\text{S}$  released during turnover. We show the stoichiometric formation of  $\text{H}_2\text{S}$  and fGly (Fig. 4-4 bottom left, e and d) in the course of the FGE catalyzed reaction. This further



**Fig. 4-4: Top:** FGE catalyzed oxidation of the substrate cysteine to a thiol aldehyde and subsequent spontaneous hydrolysis to the aldehyde moiety and hydrogen sulfide. **Bottom left:** Quantification of  $\text{H}_2\text{S}$  formation. FGE catalyzed product formation is coupled to  $\text{H}_2\text{S}$  formation. In the absence of or presence of only substrate or FGE no hydrogen sulfide is formed (a-c). Upon mixing of substrate and FGE in the presence of DTT the product formation (d) and release of hydrogen sulfide (e) are correlated ( $c_{\text{FGE}} = 1 \mu\text{M}$ ,  $c_{\text{substrate}} = 100 \mu\text{M}$ ). **Bottom right:** Rates of product formation and release of hydrogen sulfide under various conditions.

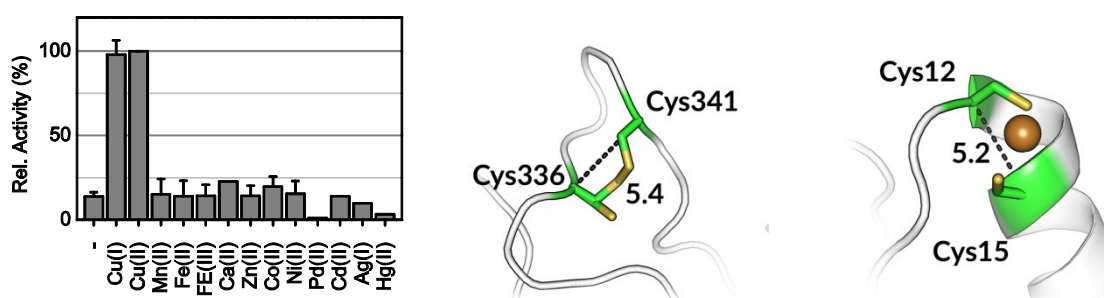
indicates that the sulfur atom of the substrate is not irreversibly oxidized.

The reduction of oxygen to water requires four electrons. Two of these electrons are provided by the substrate which gets oxidized. The remaining two are provided by the reducing agent which is stoichiometrically consumed. In agreement with these observations,  $H_2O_2$  (a theoretically possible product derived from molecular oxygen) is not detectable.<sup>[65]</sup>

#### 4.1.3 Metal dependent activity of FGE

After having quantified the side-products of the FGE-catalyzed reaction, the rationale of how oxygen activation can occur by the active site cysteines, still remains unclear. One possible explanation is the absence of a redox-active cofactor. Investigations of the metal content of isolated FGE were performed previously. Analysis via ICP-MS showed the presence of calcium as well as potassium (1.65 and 0.69 mol/mol FGE, respectively) in samples of human FGE.<sup>[45]</sup> A different study analyzing bacterial FGE from *Streptomyces coelicolor* and *Mycobacterium tuberculosis* identifies one  $Ca^{2+}$  in the *S. coelicolor* variant, while FGE from *M. tuberculosis* did not contain significant amount of metals.<sup>[54]</sup> The residues involved in calcium binding, however, are highly conserved within FGEs. While such alkali and alkaline earth metals do not suffice to catalyze redox chemistry, they might contribute to protein stability and folding or serve as a Lewis acids. However, the absence of such metals in (still active) FGE from *M. tuberculosis* leads to the conclusion that these metals are not essential for *in vitro* turnover. Interestingly, human enzyme preparations were found to contain 0.09 mole of copper per mole of FGE. Upon supplementation of the reaction buffer with excess of  $Cu(II)Cl_2$  (Cu : FGE 1000 : 1) FGE was strongly inhibited ( $IC_{50} \approx 5 \mu M$ ).<sup>[45]</sup> The effect of heavy metals on protein folding and stability has been subject to many research projects and it is well recognized that such metals can be toxic towards many organisms.<sup>[89][90]</sup> In fact, protein aggregation and inactivation can be caused by large amounts of transition metals in the medium.

With this in mind, we intended to reinvestigate the previously published data by assaying the rate of product formation in the presence of various metal salts. Interestingly, the conversion rates were found to be significantly higher in cases in which the reaction buffer was supplemented with one mole of copper per mole of FGE (reaction rate enhancement of 10 fold,

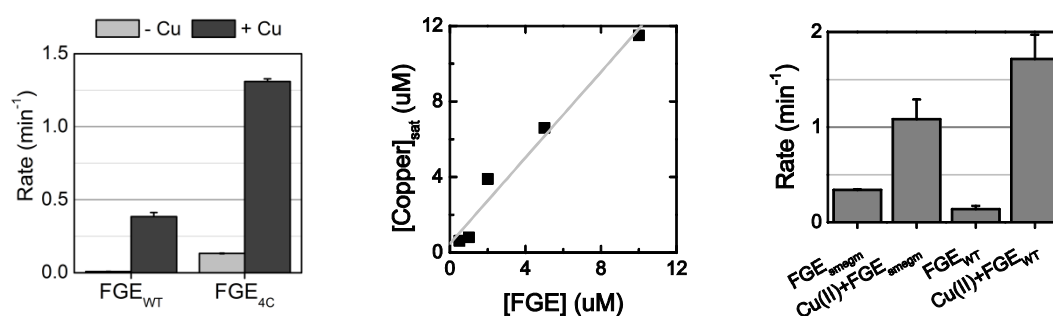


**Fig. 4-5:** FGE is a copper dependent oxidase. **Left:** Relative reaction rates in the presence of different metal salts.  $FGE_{WT}$ . Only copper salts accelerated the FGE catalyzed reaction, while others have no effect on the reaction or have an inhibitory effect. **Middle:** Crystal structure of FGE from *H. sapiens* (PDB 1Y1E, 1.7 Å). Shown are the two cysteines that have been identified as being essential for catalysis and the distance between their alpha carbon. **Right:** Crystal structure of Atx1 from *H. sapiens* (PDB 1FEE, 1.8 Å). Shown are the two cysteines involved in copper binding and the distance between their  $\alpha C$  as well as a copper atom.

see Fig. 4-5, left). Importantly, none of the other metals showed a comparable effect, which suggests a potential role for copper as a redox active metal rather than serving as a Lewis acid or providing structural support. Not only does the presence of copper increase the  $k_{\text{cat}}$  of FGE<sub>WT</sub>, it was also found to lower the  $K_M$  value. While we were not able to saturate the enzyme with substrate in the absence of copper (up to a substrate concentration of [Substrate] = 1000  $\mu\text{M}$ ), addition of copper decreased the  $K_M$  to  $K_{M,\text{WT}} = 580 \pm 40 \mu\text{M}$ . When analyzing FGE<sub>4C</sub>, copper dependent increase in reactivity was observed similarly to the wild type reaction. Again, after the addition of copper the reaction rate increased by 10-fold (see Fig. 4-6) and the catalytic efficiency was increased by 20-fold ( $k_{\text{cat}}/K_{M,4\text{C-Cu}} = 1000 \text{ M}^{-1}\text{min}^{-1}$  vs.  $k_{\text{cat}}/K_{M,4\text{C}} = 20000 \text{ M}^{-1}\text{min}^{-1}$ ). This is attributed to the lower  $K_M$  value which is, similar to FGE<sub>WT</sub>, also observed for FGE<sub>4C</sub> (no copper:  $K_{M,4\text{C}} = 600 \mu\text{M}$ ; with copper:  $K_{M,4\text{C}} = 200 \mu\text{M}$ ). The higher catalytic efficiency is thereby combined result of the higher  $k_{\text{cat}}$  and the lower  $K_M$ .

In addition to the metal, enzyme and substrate the reaction buffer contained EDTA as a chelating agent for any excess metal ions, as well as DTT as a reducing agent. EDTA mostly binds  $\text{Cu}^{2+}$  with a stability constant of  $K_D = 10^{-18} \text{ M}$  but does not bind  $\text{Cu}^+$ .<sup>[91]</sup> In contrast, DTT with its low redox potential of -0.33 V (against SHE) will quickly reduce any  $\text{Cu}^{2+}$  to  $\text{Cu}^+$  (+0.35 V). We therefore believe, that the active species in solution is best described as being Cu(I) and the elevated reaction rate after addition of Cu(II) is a result of the immediate reduction of Cu(II) to Cu(I) which will then interact with FGE.

Copper(I) and copper(II) are both often found in redox active enzymes (see also chapter 2.2), but require different binding sites. In fact, when analyzing the FGE active site for possible copper binding sites, one can identify similarities between FGE and Cu(I) binding proteins such as Atx1 (see Fig. 4-5 and chapter 2.2). Analogous to Atx1 FGE also provides two cysteine residues in the active site. When taking the carbonyl carbon as a reference, the two appear to be 5.4 Å (PDB ID: 1Y1E, 1.7 Å resolution) apart while the two cysteines in the copper chaperon are located within a distance of 5.2 Å (PDB ID: 1FEE, 1.8 Å resolution). Since the cysteines in Atx1 provide excellent copper binding capabilities, these structural similarities might explain copper binding also in FGE. Arguably, the disulfide bond in the human crystal structure might force the cysteines into this tight conformation, however in the structure of FGE<sub>C336S</sub> with bound substrate, the introduced



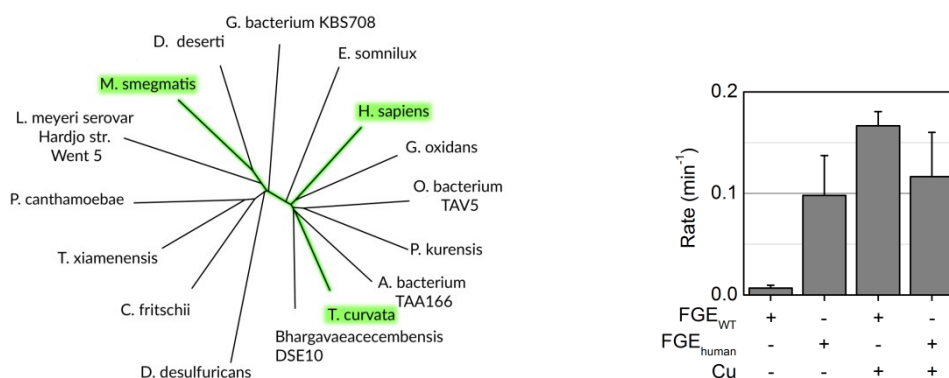
**Fig. 4-6:** Left: Rates of product formation using different FGE variants. FGE<sub>WT</sub> as well as FGE<sub>4C</sub> are likewise activated by addition of copper. Middle: Copper concentration required to saturate increasing amounts of FGE<sub>4C</sub>. The reaction rates at different concentrations of FGE were assayed each at increasing copper concentrations. The lowest concentration of copper yielding the highest reaction rate was plotted against the respective FGE concentration. The data was fitted with a linear plot. The slope of which is close to 1  $\mu\text{M Cu}/\mu\text{M FGE}$ , indicating the 1 : 1 relationship between metal and enzyme. Right: Increase in reaction rates of FGE from *Mycobacterium smegmatis* and *Thermomonospora curvata* upon copper treatment.



and remaining cysteine (now engaged in a disulfide bond with the substrate) are separated by a distance of 5.5 Å (PDB ID: 2AIK, 1.73 Å).

We then further investigated whether copper forms a complex with FGE or rather provides a source of reactive oxygen species in solution which could then interact with the enzyme. To do so, we analyzed the reaction rate of FGE<sub>AC</sub> in the presence of increasing amounts of copper and determined the amount of copper required for maximal activity (see also Appendix Fig. 4-13). After doing so for various concentrations of FGE we observed a correlation between copper concentration required for saturation and the enzyme concentration (see Fig. 4-6, middle). The slope of the linear fit is  $m = 1$ , showing that one copper ion is required to activate one molecule of FGE. This favours the hypothesis, that copper forms a complex with FGE, and excess of copper does not further contribute to the reaction after FGE saturation with copper. We therefore conclude that the observed effect is not a results of an increased concentration of ROS in solution (by copper catalyzed reduction), but of the formation of a complex between FGE and the metal.

To verify that copper dependent rate acceleration is not just an erratic feature of this particular homologue, we cloned a FGE variant from *Mycobacterium smegmatis* (FGE<sub>smegmatis</sub>) and analyzed the reaction rate in the presence of copper(I). The two proteins both originate from actinobacteria but share a sequence homology of only 46 %. Despite this low homology, both proteins show higher reaction rates in the presence of one equivalent of copper (see Fig. 4-6, right). In fact, in an unrooted phylogenetic tree (see Fig. 4-7, left), the two proteins are located on opposite branches, with the human variant (FGE<sub>human</sub>) being more closely related to FGE<sub>WT</sub> from *T. curvata* (sequence homology 58 %). The fact, that the copper dependent rate acceleration is observable for two distantly related homologues, such as FGE<sub>WT</sub> and FGE<sub>smegmatis</sub>, suggests that copper dependent rate acceleration should also be observable for other FGEs. In fact, the copper activation has additionally been shown for FGE from *Streptomyces coelicolor* and *Homo sapiens*, further supporting the relevance of copper mediated activation of FGE.<sup>[94]</sup> However when similar



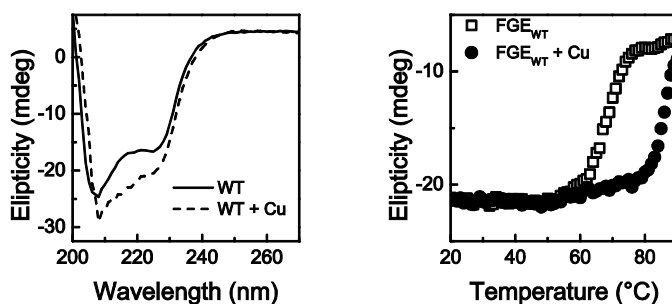
**Fig. 4-7: Left:** Unrooted phylogenetic tree showing the relation between FGEs from different organisms. Selected were the closest FGE homologues in 15 bacterial phyla and *H. sapiens*. FGEs from the following organisms have been considered: Gemmatimonadetes bacterium KBS708, Entomoplasma somnilux, *Homo sapiens*, *Gluconobacter oxydans*, *Opiritaceae bacterium TAV5*, *Physcisphaeramikurensis*, *Acidobacteriaceae bacterium TAA166*, *thermomonospora curvata*, *Bhargavaeacecembensis DSE10*, *Deferribacter desulfuricans*, *Chlorogloeopsis fritschii*, *Thermophagus xiamenensis*, *Parachlamydiaa canthamoebae*, *Leptospira meyeri serovar Hardjo str. Went 5*, *Mycobacterium smegmatis*, *Deinococcus deserti*. **Right:** Reaction rate of FGE<sub>WT</sub> and FGE<sub>human</sub> in the presence of copper. Increased reactivity is observed only for FGE<sub>WT</sub> while FGE<sub>human</sub> shows comparable activity to the copper free reaction. The active site of FGE<sub>human</sub> might be preloaded with copper from the culture medium.

reactions with FGE<sub>human</sub> were performed in our laboratory, we did not observe an enhanced reaction rate in the presence of copper (see Fig. 4-7). We concluded that the enzyme might be preloaded with copper from the expression medium and is therefore not activated by additional copper.

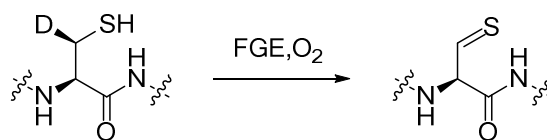
The identification of copper as an important cofactor for FGE catalysis, we then further investigated copper induced structural effects. While the structure consists largely of random coil (see chapter 2.3.3), analysis of the CD spectrum before and after addition of copper shows an increase in the alpha helical content (see Fig. 4-8, left). In the presence of copper, the point of thermal unfolding is significantly shifted towards higher temperatures ( $\Delta T > 15^\circ\text{C}$ , see Fig. 4-8, left). The graph shows two transition points. A minor one, only contributing to  $\approx 10\%$  of the unfolding (based on the overall change in ellipticity) and the global unfolding, responsible for the major loss of CD signal. The first minor transition is initiated at temperature comparable to those of the global transition in the copper free sample ( $60 - 70^\circ\text{C}$ ). The second, major transition in the FGE:Cu complex however occurs at significantly higher temperatures. We hypothesize, that a smaller, less stable loop is prone to denature at lower temperatures ( $T < 80^\circ\text{C}$ ) whereas the core of the protein is highly stabilized upon addition of copper with global denaturing occurring only at temperatures higher than  $80^\circ\text{C}$ . A similar copper induced stabilization has been observed for metallo chaperons CopZ from *Bacillus subtilis* and Atox1 from *Homo sapiens*.<sup>[95]</sup> Both proteins share the copper binding motif MxCxC and the denaturation temperature in the presence of copper increases by  $\Delta T = 3 - 10^\circ\text{C}$ . While the copper binding site is located on a surface exposed loop region, it has been suggested that copper binding induces rearrangements throughout the protein structure which enhances the core stability. In contrast, other examples exist in which copper was shown to enhance the rate of proteolysis of CopZ from *Enterococcus hirae*, which in turn displays the toxicity of elevated copper concentrations *in vivo*.<sup>[96]</sup>

The enhanced reactivity of FGE in the presence of copper not only enables researchers to improve existing protein labeling methods, but also proves to be of particular importance in the understanding of the FGE mechanism. Previous studies have suggested a potential mechanism, but to some extent rely on assumptions and do not specify on the way of oxygen activation. The finding, that FGE activity relies on a redox active metal centre explains for the first time how oxygen activation can occur.

To further investigate the mechanism of action of FGE we determined the rate limiting step of the reaction. We therefore synthesized a deuterated substrate peptide with a deuterium atom in position of the (*R*)- $\beta$ -hydrogen (see Scheme 4-1). When supplying the reaction with the

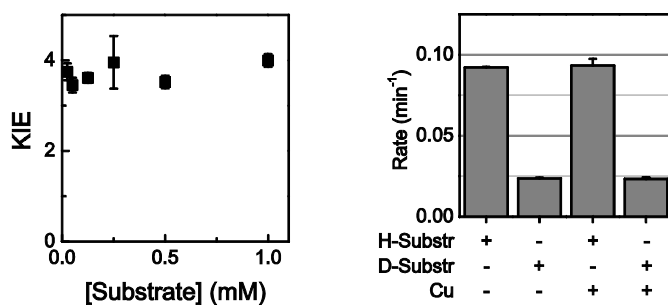


**Fig. 4-8:** Circular dichroism analysis of FGE<sub>WT</sub> in the presence of copper. Cu(I) binding induces an increase in helical content (left) as well as an increased resistance towards thermal stress (right).



**Scheme 4-1:** Reaction scheme showing the selective removal of the (*R*)- $\beta$ -hydrogen of the substrate cysteine in the course of the FGE catalyzed reaction.

deuterated peptide analog the reaction rate of FGE<sub>WT</sub> and FGE<sub>4C</sub> in the absence of copper is not affected ( $k_H/k_D = 0.8 - 1.2$ ). In the presence of copper, however, the rate of product formation shows a substrate kinetic isotope effect (KIE) of  $KIE_{WT} = 3.7 \pm 0.2$  and  $KIE_{4C} = 3.0 \pm 0.2$ . The products of both, the non-deuterated and the deuterated substrate, share the same mass as well as isotope pattern (see Appendix Fig. 4-15), indicating loss of the deuterium during turnover. The isotope effect proved to be independent of the substrate concentration (see Fig. 4-9). We therefore conclude that the presence of copper increases the extent to which hydrogen abstraction from the substrate is rate limiting. To further support this finding we additionally tested for isotope effects when using human FGE. In agreement with the data obtained with FGE<sub>WT</sub> and FGE<sub>4C</sub>, copper-supplemented FGE<sub>human</sub> shows an isotope effect of  $KIE_{human} = 4.0 \pm 0.1$ . In contrast to the previous observations with FGE<sub>WT</sub>, FGE<sub>human</sub> shows an isotope effect of  $KIE_{human,-Cu} = 3.9 \pm 0.1$  even in the absence of copper. In agreement with our earlier observation that copper addition does not positively influence the reaction rate, one can interpret the human variant as being preloaded with copper, already after isolation from the expression system. The isotope effect without the presence of copper can hence not be determined without removal of the metal. Efficient methods to remove the metal and sustain the activity are yet to be developed. However, what remains is the observation that both FGE<sub>WT</sub> as well as FGE<sub>human</sub> share the same rate limiting step, indicating that both variants follow the same reaction scheme and that mechanistic findings of FGE<sub>WT</sub> will hold true also for the human enzyme.



**Fig. 4-9:** **Left:** The substrate kinetic isotope effect at increasing substrate concentration. The reaction mixture was supplied with FGE<sub>WT</sub> and one molar equivalent of copper. **Right:** The isotope effect of FGE<sub>human</sub> in the presence and absence of copper is indistinguishable as FGE<sub>human</sub> is already preloaded with copper.

---

## 4.2 Conclusion

Copper dependent oxidation reactions have been known for many years. The participation of copper ions in the FGE mechanism, however, has been excluded until recently due to the absence of metals in protein preparations and accordingly crystal structures.<sup>[45]</sup> While residual copper (0.09 Cu/FGE) was previously identified in some FGE samples isolated from eukaryotic cell lines, the reactivity of such enzymes in the presence of copper decreased. In contrast our present results show the rate enhancement of FGE upon incubation with one molar equivalent of Cu(I). Interestingly, the reaction rate of FGE<sub>WT</sub> in the absence of additional metal is 6 % of the maximum reactivity that we observe under standard conditions with copper present ( $v_{\text{WT-Cu}} = 0.01 \text{ min}^{-1}$  vs.  $v_{\text{WT+Cu}} = 0.17 \text{ min}^{-1}$ ). The low activity might therefore be a representation of the low copper loading, with metal independent turnover not being possible. The inhibitory effect of copper previously reported by other researchers could be attributed to the large excess (1000 fold) of copper during the setup of the reaction.<sup>[45]</sup> The toxic effects of copper as well as metal catalyzed formation of reactive oxygen species on cells and proteins is well established.<sup>[7][42]</sup> We therefore demonstrated the requirement of a 1:1 ratio between copper and FGE to yield the highest activity. The requirement of copper for catalysis was additionally demonstrated for a second bacterial protein (*M. Smegmatis*) showing the generality of this observation. The work of other research groups, who likewise observed the copper mediated rate enhancement *in vitro* and *in vivo* further support our findings.<sup>[94][97]</sup>

In addition to the participation of copper(I) in FGE catalysis we demonstrated the stoichiometric release of hydrogen sulfide as well as the first quantitative analysis of reducing agent consumption during turnover. Although, the requirement of reducing agents has been reported by others prior to this work<sup>[3]</sup> quantification has lacked.

For the first time, the equimolar release of H<sub>2</sub>S and the quantitative consumption of reducing agent have been demonstrated. We therefore suggest that the oxidation equivalent of molecular oxygen is transferred to the substrate and the reducing agent, rather than being released in hydrogen peroxide. In agreement with this observation, H<sub>2</sub>O<sub>2</sub> has not been detected during turnover.<sup>[65]</sup>

The finding that FGE requires copper for catalysis has further enriched the mechanistic discussion and resulted in the identification of the C-H bond cleavage as the rate limiting step. It also stimulates the development of FGE based bioconjugation applications, since *in vitro* and *in vivo* reactivity of FGE is significantly increased.<sup>[97]</sup> Together with the more stable protein fold this will most certainly support the development of new methodologies in biotechnology as it helps to overcome current limitations regarding low reactivity.

The lysosomal storage disorder MSD has been attributed to insufficient FGE activity. Different gene mutations and deletions in the sulfatase modifying factor 1 gene have been identified.<sup>[48][98]</sup>

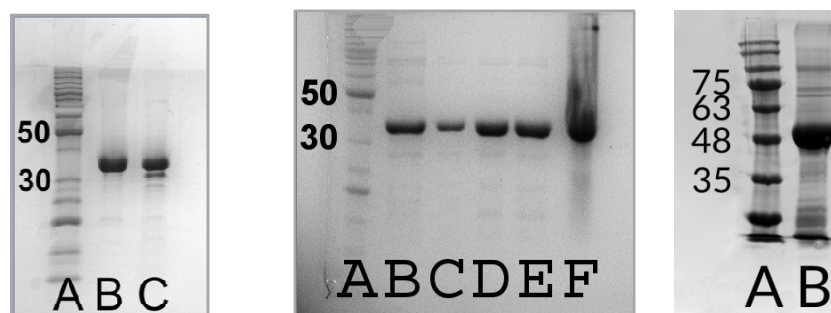
The question might be raised, whether incomplete copper loading of FGE causes MSD or whether mutations on the FGE gene lead to insufficient metal binding. Similarly, proteins involved in copper delivery and transport such as the metallo chaperon Atx1 may contribute to the disfunction of sulfatases. In this light, the identification of copper as an essential contributor to FGE activity will further drive researchers to investigate the role of formylglycine related diseases such as MSD.

### 4.3 Experimental

**Plasmid Construction.** Genes encoding FGE<sub>WT</sub>, FGE<sub>4C</sub> and FGE<sub>smegmatis</sub> were ordered from genscript in a pUC57 vector with a N-terminal His<sub>6</sub>-tag. Genes were digested using NdeI and XhoI restriction enzymes and subsequently ligated into similarly treated pET28 or pET19 expression vectors using T4 DNA ligase. The identity of final plasmids was confirmed by sequencing analysis (Microsynth).

The gene for FGE<sub>smegmatis</sub> was amplified from *Mycobacterium smegmatis* genomic DNA (DSM # 43465). This fragment was cloned into a pOPIN-expression vector using restriction free protocols. The resulting vector (pOPIN\_FGE<sub>smegmatis</sub>) encodes FGE<sub>smegmatis</sub> as a fusion with an N-terminal sumo domain (sumo: small ubiquitin-like modifier) to improve stability and solubility.

**Protein Expression and Purification.** Expression plasmids of FGE<sub>WT</sub>, FGE<sub>4C</sub> and FGE<sub>smegmatis</sub> were transferred into BL-21 pLys *E. coli* cells. Pre-cultures were inoculated with one single bacterial colony equipped with the respective plasmid. After incubation at T<sub>inc</sub> = 37 °C overnight expression cultures of 1 L lysogeny broth supplemented with kanamycine and chloramphenicol at concentrations of 50 mg/L and 34 mg/L, respectively, or ampicilin and chloramphenicol (100 mg/L and 34 mg/L respectively) were inoculated by addition of 1 mL preculture. After incubation at T<sub>inc</sub> = 37 °C until the optical density reached OD = 0.6 - 0.8, gene expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM. Cultures were continuously incubated at T<sub>express</sub> = 37 °C for additional t<sub>express</sub> = 3 – 4 hrs when cells were harvested by centrifugation at 8000 rpm. Cells were stored at T<sub>store</sub> = -20 °C until purification. The cell pellet was then thawed and resuspended in lysis buffer containing 50 mM phosphate buffer pH 8 and 200 mM NaCl. Cell lysis was induced by sonication and the insoluble fraction was removed by ultra-centrifugation at 9000 rpm at T = 4 °C. The lysate was treated with Ni-NTA agarose beads and incubated for 20 min at T<sub>bind</sub> = 4 °C. The resin was loaded onto a gravity flow column where it was washed with excess lysis buffer containing increasing concentration of imidazole from c<sub>imid</sub> = 0 – 20 mM. FGE was fractionated by elution from the resin by addition of 250 mM imidazole in lysis buffer. Collected fractions were dialyzed against a buffer containing 50 mM Tris-HCl pH 8 and 50 mM NaCl twice for t<sub>dial</sub> ≥ 4 hrs at T<sub>dial</sub> = 4 °C. The final protein samples were flash frozen in liquid N<sub>2</sub> and stored at T<sub>store</sub> = -80 °C until needed.



**Fig. 4-10:** SDS-PAGE pictures of FGE<sub>WT</sub> (left) and FGE<sub>4C</sub> (right) after purification by Ni-NTA affinity chromatography. **Left:** A: Molecular weight marker; B: FGE<sub>WT</sub>; C: FGE<sub>WT</sub> + 5 mM DTT. **Middle:** A: Molecular weight marker; B-E: collected and pooled fractions of FGE<sub>4C</sub> after affinity chromatography; F: insoluble fraction after cell lysis of FGE<sub>4C</sub> cell pellet. **Right:** A: Molecular weight marker; B: FGE<sub>smegmatis</sub>.

---

Protein identity was confirmed by SDS-PAGE and HRMS (see Fig. 4-10 and Tab. 4-2). Samples of FGE<sub>human</sub> were kindly supplied by Prof. Dr. Thomas Dierks, University of Bielefeld.

**Tab. 4-2:** Calculated and observed molecular weights of proteins described in chapter 4 with comments on the mass difference.

Protein	MW <sub>calc</sub> [Da]	MW <sub>obs</sub> [Da]	Comment
FGE <sub>WT</sub>	35308.1	35306.5	Disulfide bond
FGE <sub>4C</sub>	35195.8	35193.8	Disulfide bond
FGE <sub>smegmatis</sub>	44637.8	44633.4	Disulfide bonds

**Determination of thermal unfolding by circular dichroism.** FGE was diluted to a final concentration of  $c_{\text{FGE}} = 10 \mu\text{M}$  in a buffer containing 50 mM Tris, 50 mM NaCl and 2 mM DTT. Thermal unfolding was monitored at  $\lambda_{\text{CD}} = 222 \text{ nm}$  in a quartz cuvette of 1 mm pathlength while heating at  $1 \text{ }^\circ\text{C}/\text{min}$  in steps of  $1 \text{ }^\circ\text{C}$ . Data points were fitted with a boltzman distribution curve (see Eq. 2) with  $x_0$  being the midpoint of thermal transition.

$$y = A_1 + \frac{A_1 - A_2}{1 + \exp\left(\frac{x - x_0}{dx}\right)} \quad (2)$$

**Peptide synthesis.** Peptide substrates were synthesized by Fmoc solid phase peptide synthesis following standard protocols on a peptide synthesizer. Shortly, the most C-terminal amino acid was coupled to Rink amide resin. Deprotection of the primary amine is accomplished by incubation with 40 % piperidine in DMF. The subsequent amino acid is added and activated by HCTU in DMF upon which coupling occurs. Unreacted peptide is quenched by addition of acetic anhydride DMF. The peptide is sequentially synthesized with the last step being coupling of 2-amino benzoic acid. Cleavage from the resin and deprotection of the amino acid side chains is simultaneously achieved by incubation with the cleavage cocktail (TFA:TIPS:DCM 96.5 : 2.5 : 1) for 3 hrs. Impurities have been removed by RP-HPLC after which lyophilization yields the final peptide substrate. The identity of the peptide was confirmed by ESI-MS (sequence: Abz-SALCSPTRA-NH<sub>2</sub>; MW<sub>calc</sub> = 1023.5 Da,  $[\text{M}+\text{H}]^+_{\text{obs}} = 1024.5 \text{ Da}$ ,  $[\text{M}+2\text{H}]^{2+}_{\text{obs}} = 513.0 \text{ Da}$ ).

The deuterated peptide analog was kindly supplied by Pascal Engi.

**Analysis of reaction rates.** Reaction rates were determined using a discontinuous HPLC assay. If not stated otherwise, FGE was diluted to a final concentration of  $c_{\text{FGE}} = 2 \mu\text{M}$  in a reaction buffer containing 200  $\mu\text{M}$  substrate peptide, 2 mM DTT, 50 mM EDTA, 50 mM NaCl and 50 mM Tris-HCl buffer pH 8. Samples containing additional metal salts were supplemented with the respective metal ion at a final concentration of  $c_{\text{metal}} = 2 \mu\text{M}$ . The reaction was incubated at  $T_{\text{react}} = 25 \text{ }^\circ\text{C}$ . Throughout the reaction time, samples were quenched by addition of one volume equivalent of 4 M urea in 2 % TFA, which were subsequently analyzed by RP-HPLC (Gemini-NX, 5 $\mu\text{m}$  C18 250 x 4.6 mm column, gradient see Tab. 4-3,  $\text{RT}_{\text{substrate}} = 10.8 \text{ min}$ ,  $\text{RT}_{\text{product}} = 13.1 \text{ min}$ ). The substrate concentration was determined by comparison to a calibration curve established with free 2-amino benzoic acid. The product concentration was determined by comparison to a calibration curve established with samples of fixed product concentration (known substrate concentration after complete turnover).

**Tab. 4-3:** HPLC gradient used for the separation of the substrate peptide and the product peptide of the FGE catalyzed reaction. Buffer A: H<sub>2</sub>O, 1 % MeCN, 0.1 % TFA. Buffer B: MeCN, 0.085 % TFA. Flow rate 1 mL/min.

Time [min]	Buffer B [%]
0	15
0.5	15
12.0	20
17.0	95
20.7	95
22.0	15
25.0	Stop

**Michaelis Menten Kinetics.** Reaction rates at increasing amounts of substrate were determined using a discontinuous HPLC assay. FGE was diluted to a final concentration of  $c_{\text{FGE}} = 2 \mu\text{M}$  in a reaction buffer containing 0 - 1000  $\mu\text{M}$  substrate peptide, 5 mM DTT, 50 mM EDTA, 50 mM NaCl and 50 mM Tris-HCl buffer pH 8. Samples containing additional copper were supplemented with CuSO<sub>4</sub> at a final concentration of  $c_{\text{Cu}} = 2 \mu\text{M}$ . The reaction was incubated at  $T_{\text{react}} = 25 \text{ }^\circ\text{C}$ . Throughout the reaction time, samples were quenched by addition of one volume equivalent of 4 M urea in 2 % TFA, which were subsequently analyzed by RP-HPLC (see above).

**Analysis of thiol and disulfide content.** The concentration of free thiols in samples of FGE was determined by following a known standard procedure.<sup>[92]</sup> Shortly, proteins were diluted in 8 M guanidine-HCl to a final concentration of  $c_{\text{FGE}} \leq 40 \mu\text{M}$  upon which 0.2 volume equivalent of "buffer A" (100 mM boric acid, 0.2 mM EDTA) were added. Subsequently, 0.02 volume equivalents of a solution of 5,5'-dithiobis-2-nitrobenzoic acid (DTNB,  $c_{\text{DTNB}} = 10 \text{ mM}$ ) were added. After an incubation of  $t_{\text{inc}} = 5 \text{ min}$  the absorbance at  $\lambda = 412 \text{ nm}$  was determined against a blank containing only buffer A and guanidine-HCl ( $\epsilon_{\text{DTNB},412} = 14150 \text{ M}^{-1}\text{cm}^{-1}$ ).

The concentration of free thiols was determined by comparison to a calibration curve established with samples of increasing amounts of L-cysteine. The concentration of disulfide bonds was determined following a known standard procedure.<sup>[93]</sup> Disodium 2-nitro-5-thiosulfobenzoate (NTSB) stock solution was prepared by dissolving 70 mg of DTNB in an aqueous 10 M Na<sub>2</sub>SO<sub>3</sub> solution. Air was bubbled through the mixture until a pale yellow color was observed. The end point of the reaction was determined by absorbance measurement at  $\lambda = 412 \text{ nm}$  ( $\epsilon = 13900 \text{ M}^{-1}\text{cm}^{-1}$ ). NTSB assay solution was prepared by diluting NTSB stock solution by 100 fold in "NTSB buffer" solution (Trizma base 2.228 g, guanidine-HCl 28.659 g, EDTA (500 mM) 0.6 mL, Na<sub>2</sub>SO<sub>3</sub> 1.2604 g, degassed H<sub>2</sub>O 100 mL, pH 9.5). Protein samples were prepared by dilution in NTSB assay solution and were kept in the dark. After incubation at RT for 5 min absorbance at  $\lambda = 412 \text{ nm}$  was measured against a blank containing only NTSB assay solution. The concentration of disulfide bonds is determined by comparison to a calibration curve established with samples of increasing amounts of L-cystine.

**Quantification of H<sub>2</sub>S formation.** Hydrogen sulfide concentration was determined following a known procedure.<sup>[99]</sup> FGE was diluted to a final concentration of  $c_{\text{FGE}} = 1 \mu\text{M}$  in a reaction mixture containing 100  $\mu\text{M}$  substrate, 2 mM DTT, 50 mM NaCl and 50 mM Tris-HCl pH 8.0. After various time points, samples for sulfide analysis were quenched by addition of 0.08 volume equivalent of sulfide assay solution (10.8 mM N,N-dimethyl-p-phenylenediamine, 37 mM FeCl<sub>3</sub> in 6 M HCl).

---

Sulfide formation was monitored by UV/vis spectroscopy after an incubation time of 20 min at a wavelength of  $\lambda = 667$  nm. The concentration of  $\text{H}_2\text{S}$  is determined by comparison to a calibration curve established with samples of increasing amounts of  $\text{Na}_2\text{S}$ .

**Labeling with 2-iodo acetamide.** FGE was diluted to a final concentration of  $c_{\text{FGE}} = 45 \mu\text{M}$  in a reaction mixture containing 2 mM DTT, 50 mM EDTA, 50 mM NaCl and 50 mM Tris-HCl pH 8. The reaction was incubated at RT for 30 min, after which 2-iodo acetamide was added to a final concentration of  $c_{\text{IAA}} = 5$  mM. The reaction was then incubated at RT for 120 min. Protein modification was subsequently analyzed by LC/HRMS.

**Copper titration.** The copper(I) saturation was determined by diluting  $\text{FGE}_{4\text{C}}$  to a final concentration of  $c_{\text{FGE}} = 0.5 - 10 \mu\text{M}$  in a reaction buffer containing 0 – 5 molar equivalents of  $\text{CuSO}_4$ , 200  $\mu\text{M}$  substrate, 2 mM DTT, 50 mM EDTA, 50 mM NaCl and 50 mM Tris-HCl pH 8. Reactions were then incubated at  $T_{\text{react}} = 25$  °C. The reaction rates were determined as described previously (see above).



#### 4.4 Appendix

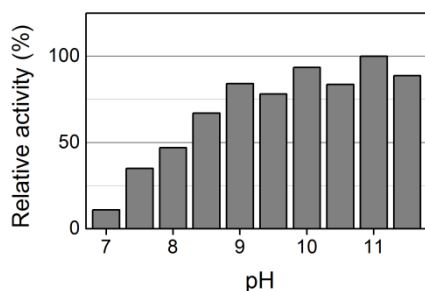


Fig. 4-11: Relative rate of product formation at increasing pH. FGE<sub>WT</sub> shows a pH optimum of pH 10 – 11.

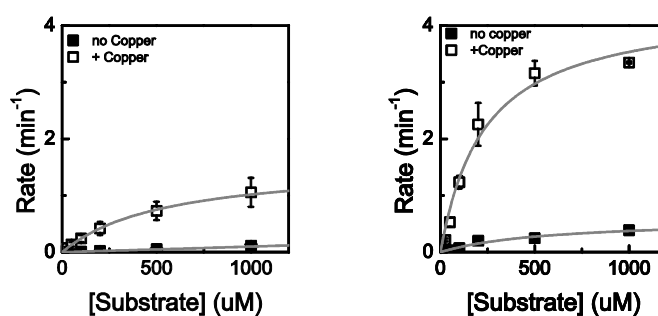


Fig. 4-12: Reaction rates of FGE<sub>WT</sub> (left) and FGE<sub>4C</sub> (right) at different substrate concentrations in the absence (solid square) and in the presence (empty square) of one molar equivalent of Cu-EDTA. FGE was diluted to a final concentration of  $c_{\text{FGE}} = 2 \mu\text{M}$  in a reaction mixture containing increasing amounts of substrate, 5 mM DTT, 50 mM EDTA, 50 mM NaCl and 50 mM Tris pH 8.0. Reactions in the presence of copper additionally contained  $2 \mu\text{M}$  of  $\text{CuSO}_4$ . Reaction was then incubated at  $T_{\text{inc}} = 25 \text{ }^\circ\text{C}$ . After various time points the reaction was quenched for product analysis by addition of 1 volume equivalent of 1% TFA in 4 M urea. Product formation was analyzed by RP-HPLC from which the reaction rate has been calculated.

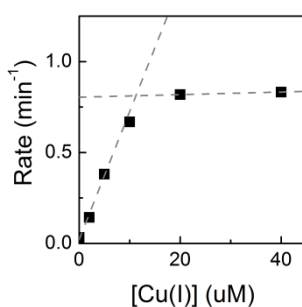
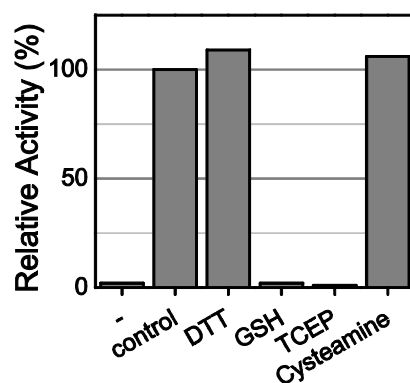
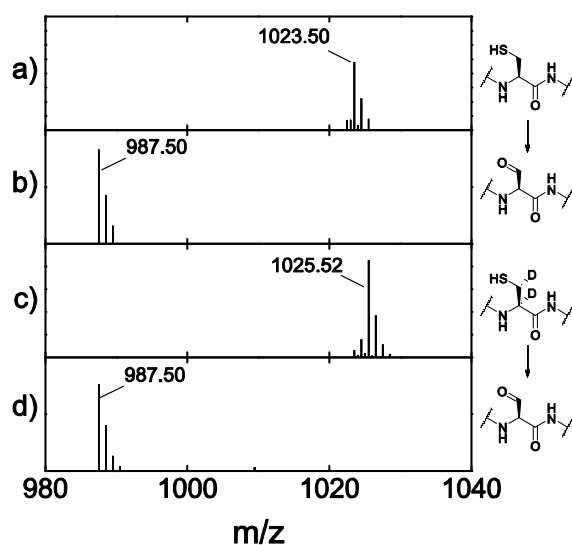


Fig. 4-13: Reaction rates of FGE<sub>4C</sub> at different concentrations of copper. FGE was diluted to a final concentration of  $c_{\text{FGE}} = 10 \mu\text{M}$  in a reaction mixture containing increasing amounts of copper,  $200 \mu\text{M}$  substrate, 2 mM DTT, 50 mM EDTA, 50 mM NaCl and 50 mM Tris pH 8.0. Reactions in the presence of copper additionally contained  $2 \mu\text{M}$  of  $\text{CuSO}_4$ . Reaction was then incubated at  $T_{\text{inc}} = 25 \text{ }^\circ\text{C}$ . After various time points the reaction was quenched for product analysis by addition of 1 volume equivalent of 1% TFA in 4 M urea. Product formation was analyzed by RP-HPLC from which the reaction rate has been calculated.



**Fig. 4-14:** Relative activity of FGE in the presence of different reducing agents. FGE<sub>4C</sub> was diluted to a final concentration of  $C_{FGE} = 0.5 \mu\text{M}$  in a reaction mixture containing  $50 \mu\text{M}$  substrate,  $0.5 \mu\text{M}$  CuSO<sub>4</sub>, 50 mM EDTA, 50 mM NaCl and 50 mM Tris pH 8.0. Reducing agent was added after an incubation time of  $t_{inc} = 40$  min to a final concentration of  $C_{reduce} = 2$  mM. Activity in the absence of reducing agent was calculated as average of the rates before addition of reducing agent. Positive controls were supplied with DTT at time  $t = 0$  min. After various time points the reaction has been quenched for product analysis by addition of 1 volume equivalent of 1 % TFA in 4 M urea. Product formation has been analyzed by RP-HPLC. The relative rate has then been calculated, with respect to the control reaction which contained DTT from the beginning.



**Fig. 4-15:** High resolution mass spectrometry of the non-deuterated peptide substrate (a) and the selectively deuterated peptide substrate (c) and their respective products (b & d). Products were isolated by RP-HPLC after incubation with FGE overnight.

**Tab. 4-4:** Amino acid sequences of FGEs from different organisms described in chapter 4.

Source Organism	Sequence
<i>Thermomonospora curvata</i>	GSSHHHHHSSGLVPRGSHMPSFDFDIPRRSPQEIAKGMVAIPGGTFRMGGEDPDAFPE DGEGPVRTVRLSPFLIDRYAVSNRQFAAFVKATGYVTD AERYGWSFVFH AHVAPGTPVM DAVVPEAPWWVAVPGAYWKAPEGPGSSITDRPNHPVVHVS WND AVAYATWAGKRLPT EAEWEMAARGGLDQARYPWGNELTPRGRHRCNIWQGT F PVHDTGEDGYTGTAPVNAF APNGYGLYNVAGNVWEWCADWW SADWHATESPATRIDPRGPETGTARVTKGG SFLCH ESYCNRYRVAARTCNTPDSSAAHTGFRCAADPL
<i>Mycobacterium smegmatis</i>	GSSHHHHHGGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAF A KRQ GKEMDSLRF LYDGIRIQADQTPEDLDMEDNDIIEAHREQIGGLTELVEVPGGSFRMG STSFYPEEAPVHTATVGDFAIERHPVTNAQFAEFVEATGYVTVAERPLDPKLYPGVPEADLV PGSLVFRPTSGPVDLRDWRQWWDWAPGACWHHPFGPRREFCRPDHPVVQVAYPDAV AYATWAGRRLPTEAEWEYAARGGPKGGVGFLYAWGDEVCPDQQLMANTWQ GKFPYR NDGALGWKGTSPVGT FPPNRLGLVDMIGNVWEWTATKFS AHRPGDESHVTCCPPPSG GDPGVNQVLKGGSHLCAPEYCHRYRPAARSPQSQDSSTTHIGFRCVVS
<i>Homo sapiens</i>	LAHSMVPIPAGVFTMGTD DDPQIKQDGEAPARRVTIDAFYMDAYEVSNT EFEK FVNSTGY LTEAEKFGDSFVFEGMLSEQVKTNIQQAVAAAPWWLPVKGANWRHPEGPDSTILHRPD HPVLHVS WND AVAYCTWAGKRLPTEAEWEYSCRGGLHNRLFPWGNKLQPKGQHYANI WQGEFPVTNTGEDGFQGTAPVDAFPNGYGLYNIVGNAWEWTS DWWTVHHSVEETL NPKGPPSGKDRVKKGGSYMCHRSYCYRYRCAARSQNTPDSSASNLGFRCAADRLP

---

---

---

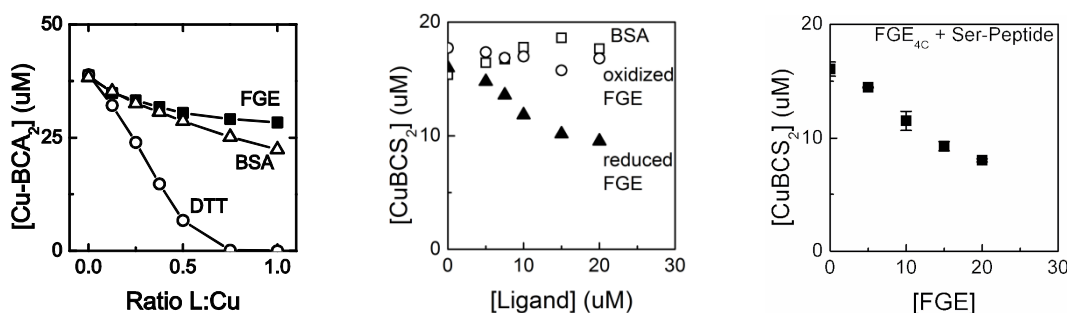
## 5. Copper Binding in Formylglycine Generating Enzyme

FGE has previously been identified as a cofactor-free oxidase. This classification raised significant interest in FGE research and was primarily based on the absence of cofactors or metals in protein preparations and existing crystal structures. Additionally, no distinct metal recognition sites had been identified. With the identification of copper as a crucial electron donor in FGE catalysis, the identification of the metal binding site is an important target. Many copper containing proteins and enzymes use thiolates or thiol ethers to coordinate metals in active or binding sites. Notable examples such as Atx1 have been discussed earlier (see also section 2.2). Similarly, the FGE active site contains two highly conserved cysteine residues which are believed to be involved in disulfide formation with the substrate.<sup>[56]</sup> The previous chapter showed the requirement of a 1:1 ratio of copper to enzyme, suggesting formation of a protein : metal complex.

In the following chapter we show the participation of the two conserved cysteine residues in metal binding. Since only reduced thiols efficiently bind copper, we hypothesize that the disulfide between FGE and the substrate, which is observed in crystal structures, is an artefact of the experimental conditions and that the substrate coordinates to copper via its cysteine residue. Additionally, copper remains in the active site over the course of multiple turnover. From these findings we propose that FGE should be categorized as a copper-dependent oxidase.

### 5.1 Identification of metal binding site in FGE from *T. curvata*

The effect of copper on FGE reactivity is significant. However, none of the crystal structures published to date show a transition metal present in the protein structure. Many metal containing proteins can be isolated in the metal bound form and in a subsequent step require removal of the metal to yield the apo-protein. While the copper content of the respective culture medium could be a limiting factor, other metal containing proteins have been isolated under similar conditions (human cancer cell line or *E.Coli*). The absence of copper from FGE preparations requires alternative explanations. One possible reason is the failure of the protein to efficiently bind copper. To test this hypothesis we adopted a previously published method to determine copper(I) binding affinities.<sup>[100]</sup> This assay relies on the equilibrium between a colored complex of copper(I) and bathocuproinedisulfonic acid (BCS) and the protein (or other ligand) of interest. We found, that recombinantly produced FGE does not bind copper significantly stronger than bovine serum albumin (BSA) and far less strongly than DTT, which is always present in the reaction mixture (see Fig. 5-1, left). It has been reported that for the copper binding protein Atx1, the oxidized structure containing a disulfide bond in the active site, is absent of metal ions and is therefore not biologically relevant.<sup>[101]</sup> Knowing that FGE from *T. curvata* contains a disulfide bond in the active site upon isolation from the host organism (see chapter 4.1), we propose, that only reduced FGE is able to bind copper. This might explain the lack of metal ions in published crystal structures since FGE requires to be actively reduced in order to bind copper. Indeed, after reduction of FGE<sub>4C</sub> with DTT, the affinity for copper(I) increases significantly



**Fig. 5-1:** Copper binding of different ligands in competition with bathocuproinedisulfonic acid (BCS). **Left:** FGE in its resting state does not bind copper stronger than a random control protein with no known copper affinity (BSA) and far less strongly than DTT, which is known to be a strong copper(I) ligand. **Right:** Direct comparison of reduced and oxidized FGE<sub>4C</sub> as a ligand of copper(I).

(see Fig. 5-1, middle) with dissociation constants comparable to other copper(I) binding proteins ( $K_{D,WT} = 10^{-17.1} \text{ M}$ ;  $K_{D,4C} = 10^{-17.1} \text{ M}$ ;  $K_{D,Atx1} = 10^{-17.4} \text{ M}^{[100]}$ ). Importantly, the copper affinities of FGE<sub>WT</sub> and FGE<sub>4C</sub> appear to be very similar, leading us to conclude, that none of the cysteine residues outside the active site contribute to copper binding. To support this notion about the location of copper binding and the crucial role of the cysteine residues in this, we analyzed additional two variants of FGE that lack either of the two active site cysteines (FGE<sub>C269S</sub> and FGE<sub>C274S</sub> see Tab. 5-1). To ensure a completely reduced state of cysteines within the enzyme we performed the BCS in the presence of 2 cysteamine as a reducing agent. This small molecular weight thiol does not interfere with the assay up to concentration of 2 mM but readily reduces the enzyme (see appendix Fig. 5-9). In agreement with previous reports, FGE<sub>C269S</sub> and FGE<sub>C274S</sub> show little catalytic activity, with regards to the wild type enzyme (see Tab. 5-1).<sup>[48]</sup> Additionally, they do not show any sign of competition for Cu(I) during the copper affinity assay (see also Appendix Fig. 5-8) supporting the hypothesis that copper binding is only achieved under participation of both cysteine residues in the active site.

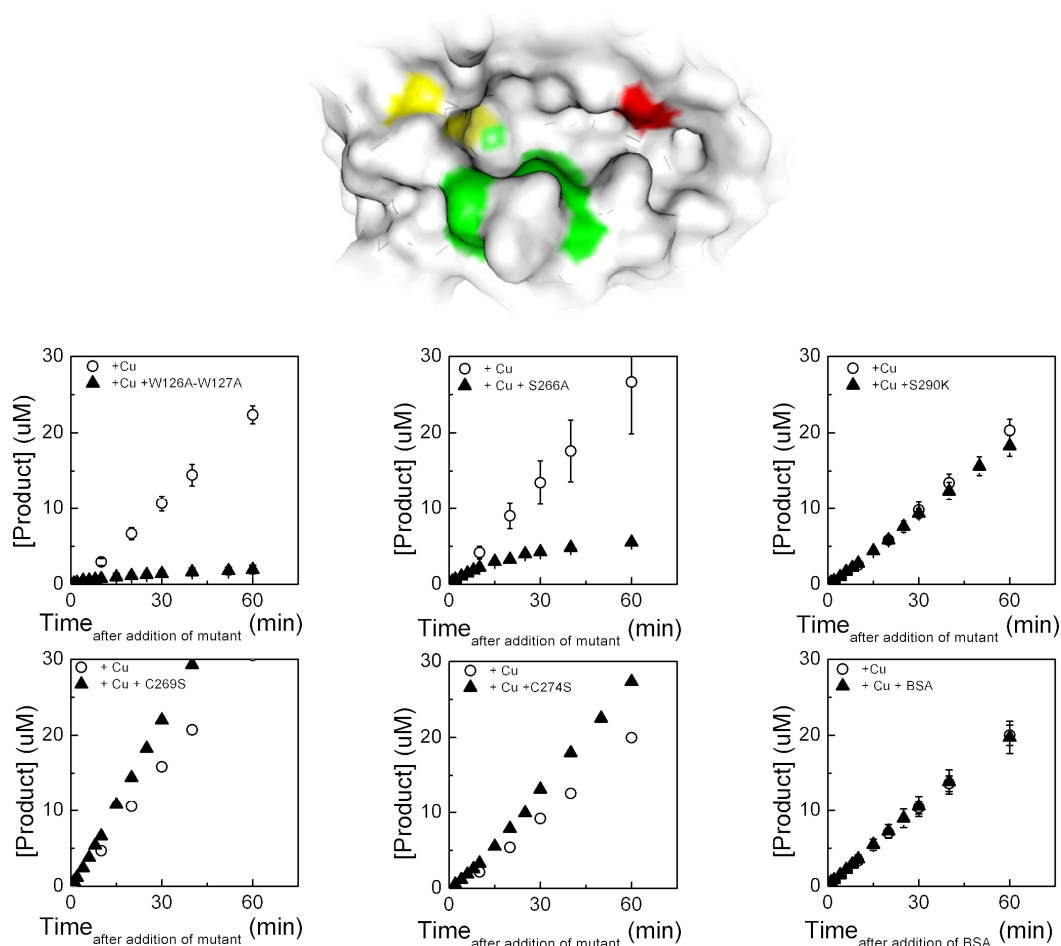
DTT has an affinity two orders of magnitude below that of FGE<sub>WT</sub> and FGE<sub>4C</sub> (see also Tab. 5-1). Under our *in vitro* reaction conditions, however, the concentration of reducing agent is three orders of magnitude above that of FGE ( $C_{FGE} = 2 \mu\text{M}$  vs.  $C_{DTT} = 2 \text{ mM}$ ). The rational

**Tab. 5-1:** Copper(I) binding affinities of different FGE variants and DTT determined by using a UV based competition assay with BCS. The respective reaction rates and the rates of copper release for selected mutants are shown.

Ligand	pK <sub>D</sub>	k <sub>react</sub> [min <sup>-1</sup> ]	k <sub>off,Cu</sub> [min <sup>-1</sup> ]
FGE <sub>WT</sub>	17.1	0.42	-
FGE <sub>4C</sub>	17.1	1.3	-
FGE <sub>C269S</sub>	n.a.	0.3 · 10 <sup>-3</sup>	-
FGE <sub>C274S</sub>	n.a.	0.01 · 10 <sup>-3</sup>	-
FGE <sub>S266A</sub>	17.7	6.3 · 10 <sup>-3</sup>	0.06
FGE <sub>S290K</sub>	16.3	1.2 · 10 <sup>-3</sup>	-
FGE <sub>W108A,W109A</sub>	16.2	0.3 · 10 <sup>-3</sup>	0.09
DTT	15.1	-	-

behind the competition of FGE for copper in such a buffer remains unclear. One explanation might be an increased affinity for FGE under the reaction conditions. We therefore analyzed the copper binding affinities of the substrate-enzyme complex. The cysteine containing peptide however, is a strong copper binder itself (see Appendix Fig. 5-10) this renders saturating concentrations of the substrate incompatible with the BCS assay. To circumvent this issue, we synthesized a peptide with a serine instead of the substrate cysteine (peptide sequence: Abz-SAL $\underline{S}$ SPTRA-NH $_2$ ). Investigations of the copper binding affinity in the presence of this serine containing substrate did yield similar affinities as the substrate free samples ( $K_{D,4C} = 10^{-17.1}$  M vs.  $K_{D,4C+Peptide} = 10^{-17.4}$  M, see also Fig. 5-1, right). However, the serine containing peptide was also shown to have only weak interactions with FGE $_{4C}$  in an inhibition assay (see Appendix Fig. 5-10) which questions the viability of this enzyme-substrate complex.

To further advance our understanding of copper binding, we generated additional inactive mutants. The variants FGE $_{S266A}$ , FGE $_{S290K}$  and FGE $_{W108A,W109A}$  prove to be many fold less active than FGE $_{WT}$  but still show high copper affinities (see Tab. 5-1). Trp126 and Trp127 as well as Ser266 (green in Fig. 5-2) are located in close vicinity to the active site cysteines (yellow in Fig. 5-2,  $d < 14$  Å), whereas Ser290 (red in Fig. 5-2) is positioned at the very end of the substrate binding cavity

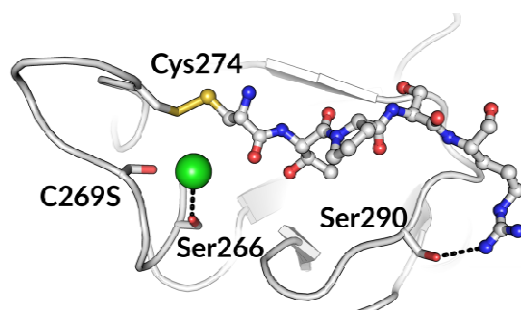


**Fig. 5-2:** Capability of different inactive FGE mutants to compete for copper(I) in direct comparison with FGE $_{4C}$ . **Top:** Surface representation of the substrate binding groove of FGE $_{human}$  (PDB 1Y1E $^{[56]}$ ) with locations of the introduced mutations. Yellow: Active site cysteines Cys269 and Cys274; Green: Trp126, Trp127 and Ser266. Red: Ser290. **Bottom:** Product production of FGE $_{4C}$  in the presence of a second protein. Only FGE variants that do contain the active site cysteines and bind the substrate can efficiently compete for Cu(I).

(distance to Cys269 = 18 Å). We introduced this specific mutation with the intent to destroy hydrogen bonding interactions with the substrate and to thereby generate a FGE variant that does not produce the product yet still contains an intact active site. In addition to the lost hydrogen bonding interaction, the protonated lysine residue would introduce a positive charge and thereby generate repulsive interactions with the positively charged arginine of the substrate. This would then further weaken substrate binding. By doing so we generated a mutant that shows a catalytic efficiency that is 440-fold lower than that of the wild type ( $k_{\text{cat}}/K_{\text{M,S290K}} = 6.6 \pm 0.2 \text{ M}^{-1}\text{min}^{-1}$  vs.  $k_{\text{cat}}/K_{\text{M,WT}} = 2900 \pm 50 \text{ M}^{-1}\text{min}^{-1}$ ) despite the introduced mutation being positioned far away from the reactive centre of FGE (see also Fig. 5-2, left and Tab. 5-1). In contrast, the mutated tryptophane residues are in close vicinity to the active site and substitution of these with alanine residues most likely strikes a hole into the structure, which then significantly disturbs the active site architecture. In agreement with these remarks, the resulting protein variant has a catalytic efficiency far below that of the wild type enzyme ( $k_{\text{cat}}/K_{\text{M,W108A,W109A}} = 1.7 \pm 0.3 \text{ M}^{-1}\text{min}^{-1}$  vs.  $k_{\text{cat}}/K_{\text{M,WT}} = 2900 \pm 50 \text{ M}^{-1}\text{min}^{-1}$ ). However, it cannot be saturated with substrate, indicating some loss of protein-peptide interactions. The effect of the serine to alanine mutation, in contrast, is more subtle. The serine residue appears perfectly positioned to undergo van der waals interactions with a nearby chloride ion (see Fig. 5-3), most likely contributing to the charge composition of the active site. Removal of these interactions could consequently interfere with local charge distributions and thereby also influence the  $pK_{\text{a}}$ 's of other nearby active site residues (i.e. Cys269 or Cys274), rendering FGE<sub>S266A</sub> 60-fold lower in activity than FGE<sub>WT</sub> ( $k_{\text{cat}}/K_{\text{M,S266A}} = 49 \pm 8 \text{ M}^{-1}\text{min}^{-1}$  vs.  $k_{\text{cat}}/K_{\text{M,WT}} = 2900 \pm 50 \text{ M}^{-1}\text{min}^{-1}$ ).

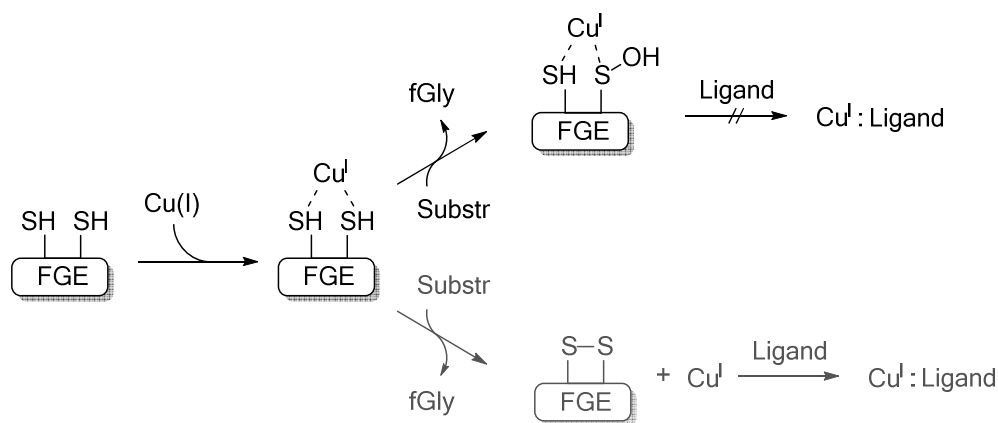
The metal binding capability of the inactive mutant FGE<sub>W108A,W109A</sub> is only slightly lower than for the active mutants FGE<sub>WT</sub> and FGE<sub>4C</sub> (see Tab. 5-1). The effect of second-coordination residues on the  $pK_{\text{a}}$  of cysteine residues has been evaluated previously for the copper chaperon Atx1 and might provide an explanation for the decreased affinities of the variants with mutations close to Cys269 or Cys274.<sup>[102]</sup>

With the copper binding in the active site of FGE being shown, the question arises, whether Cu(I) acts as in complex with the enzyme and hence can be categorized as a cofactor or whether the metal is released after each turnover and is then reduced in solution rather than bound to the enzyme (see Scheme 5-1). One way to test this is to provide a ligand that can efficiently remove copper ions from solution and hence make them unavailable for FGE as soon as the metal leaves



**Fig. 5-3:** Crystal structure of FGE<sub>human</sub> with disulfide bound substrate (PDB 2AIK<sup>[56]</sup>, 1.73 Å) showing the location of inactivating mutations S266A and S290K (numbering based on FGE<sub>WT</sub>). Ser266 is located in van der waals distance to the chloride ion (green sphere,  $d_{\text{Ser266-Cl}} = 3 \text{ Å}$ ). Ser290 is located at the end of the substrate binding groove, where it provides hydrogen bonding to the arginine residue of the substrate ( $d_{\text{Ser290-Arg}} = 3 \text{ Å}$ , side chain oxygen of serine to  $\eta$ -nitrogen of arginine).





**Scheme 5-1:** Possible pathways for copper binding and unbinding. Copper(I) can either remain in the active site as an integral part of the enzyme, which then does not allow for an external ligand to remove the metal and hence making it unavailable for FGE. Secondly, if copper is released into the solution after each turnover, external copper binder should be able to bind copper and consequently reduce the effective copper concentration available for FGE.

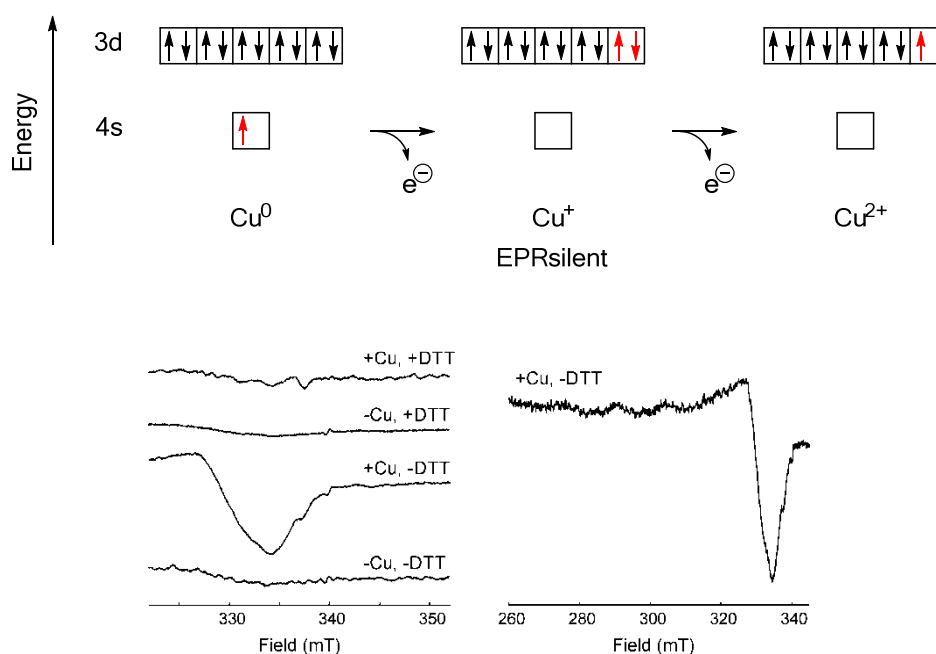
the enzyme complex. To our mind the most suitable ligand with a similar affinity for copper(I) is FGE itself. We prepared an active Cu:FGE<sub>4C</sub> complex by incubating the enzyme with substrate and reducing agent in the presence of copper. One minute after initiation, the reaction was additionally supplemented with 10 molar equivalents of inactive mutants (with regards to FGE<sub>4C</sub>) and the change in product concentration was monitored (see Fig. 5-2). Control experiments with BSA, FGE<sub>C269S</sub> or FGE<sub>C274S</sub>, did not influence the reaction rate of the Cu:FGE<sub>4C</sub>, which supports the previous data from the BCS copper(I) competition assay. Interestingly, only FGE<sub>S266A</sub> and FGE<sub>W108A,W109A</sub> show an inhibitory effect, with the final rate after equilibration of the system being at least 15 fold lower than the initial rate in the absence of competing ligand. This is in good agreement with the copper being redistributed between 0.5  $\mu$ M FGE<sub>4C</sub> and 4.5  $\mu$ M of either FGE<sub>S266A</sub> or FGE<sub>W108A,W109A</sub>. The third variant, FGE<sub>S290K</sub>, does not influence the reaction rate. Surprisingly, if only incubated with BCS and Cu(I) FGE<sub>S290K</sub> shows copper binding affinities comparable to that of other mutants and only 1-fold lower than wild type. Despite this observation, the mutant shows no copper binding affinity in direct competition with FGE<sub>4C</sub> under turnover conditions. We hereby conclude, that the copper binding affinity is enhanced in the presence of substrate, since only substrate binding should be perturbed in FGE<sub>S290K</sub>.

When looking at the removal of copper from FGE, the rate at which the product concentration levels off in the presence of FGE<sub>S266A</sub> and FGE<sub>W108A,W109A</sub> is of particular interest. A combination of an exponential and linear fit (see Eq. (3)) gives the rates at which copper leaves the active site of FGE<sub>4C</sub> as  $k_{off, Cu(I)}$  (see also Tab. 5-1). When supplying the reaction mixture with inactive mutants, copper is removed at a rate of 0.06 and 0.09  $\text{min}^{-1}$  with FGE<sub>S266A</sub> and FGE<sub>W108A,W109A</sub>, respectively. This rate is 10 fold lower than the reaction rate of Cu : FGE<sub>4C</sub> in the absence of competing ligand ( $v_{4C} = 0.82 \pm 0.10 \text{ min}^{-1}$ ), giving reasons to conclude that copper leaves the active site of the active enzyme every 10<sup>th</sup> turnover. This observation is a clear indication that copper release and subsequent re-binding cannot be part of the catalytic cycle and that FGE can hence be categorized as a copper-containing oxidase.

$$[P] = A \cdot (1 - \exp(k_{off, Cu(I)} \cdot t)) + m \cdot t \quad (3)$$

For catalysis to occur in copper containing oxidases, the oxidation state of the metal is vital. Many proteins use single or multiple copper atoms as the active centre. Numerous examples exist of such oxidoreductases that keep the metal in a +1 or +2 oxidation state.<sup>[31][104]</sup> We previously established an optimal copper to protein ratio of 1:1 to reach maximal reaction rates, indicating a mononuclear copper center (see chapter 4.1). What remains to be untangled is the oxidation state of the metal centre. Typically, enzyme redox centres are determined by electron paramagnetic resonance (EPR), a method for the detection of unpaired electrons.<sup>[105]</sup> While metal centres with only paired electrons such as Fe(II) or Cu(I) do not show EPR signals, other oxidation states generate distinct signals that can be interpreted. Careful evaluation of these patterns may provide information about the type of radical or the environment.

In the case of FGE, our intention was to determine the redox state of the copper atom. Upon incubation of FGE with substrate in the presence and absence of copper and DTT we observed different EPR spectra after freeze quenching the reaction (see Fig. 5-4). FGE catalysis is slow in the absence of metal and DTT (see also chapter 4.1). With no metals being present and no turnover occurring, FGE consequently appears EPR silent. We interpret this as the absence of free radicals or unpaired electrons. Upon addition of  $\text{CuSO}_4$  to the reaction mixture, the EPR spectrum shows a strong signal for copper(II). Due to the absence of reducing agents the metal remains in its +2-oxidation state. The enzyme, likewise, remains in its oxidized state with a disulfide bond present in the active site and is unable to bind copper (see also page 42). Under turnover conditions, i.e. in the presence of copper and DTT, the reducing agent generates the free thiols in the active site as well as Cu(I), which FGE will then tightly bind. Addition of substrate initiates the reaction, during which the metal remains in the oxidation state +1. The absence of any EPR signal during catalysis can be interpreted as the absence of Cu(II) (or Cu(0)). However, a +2 species



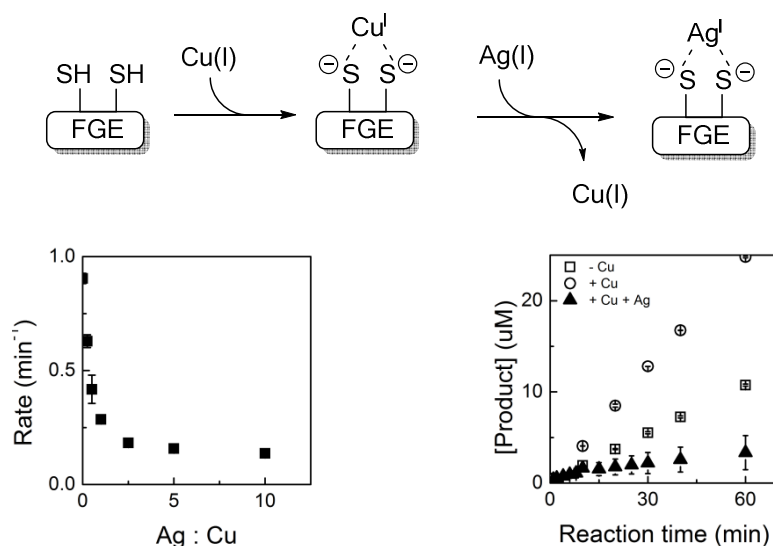
**Fig. 5-4: Top:** Electron configuration of copper in the ground state and in the first and second oxidative state. **Bottom:** X-band ( $\sim 9.52$  GHz) continuous-wave EPR spectra of  $50 \mu\text{M}$  FGE in the presence of  $500 \mu\text{M}$  substrate,  $0 - 50 \mu\text{M}$  copper and  $0 - 5$  mM DTT at  $T_{\text{EPR}} = 150$  K. Only samples containing copper(II) without addition of reducing agent show an EPR signal, whereas reaction mixtures that have been supplied with copper and DTT show the metal as being EPR silent (i.e.  $\text{Cu}^+$ ). Zoom of the region with strongest Cu(II) signal component (left) and full spectrum of the Cu(II) species (right).

---

could be short living, hence low in concentration which would render the intermediate invisible with this method.

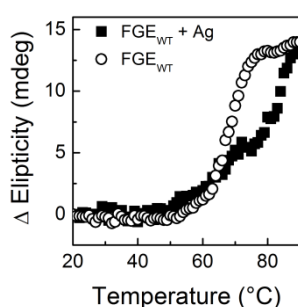
Knowing that FGE contains  $3d^{10} 4s^0$  copper allows us to search for other metals that could adopt such a configuration. These metals could then possibly replace copper in the active site. Other Cu(I) binding proteins such as metallochaperons Atx1 and P-type ATPases have been shown to bind Cd, Hg, and Ag as well. Silver, in particular, has been a well studied example of a replacing metal for copper proteins.<sup>[106]</sup> Our previous investigations have shown that none of the three metals are capable of supporting FGE catalysis (see chapter 4.1), which can be explained by their different redox potential ( $E_{Cu}^0 = 0.35$  V vs.  $E_{Ag}^0 = 0.80$  V). However, binding of the metal might still be possible. Our interest was therefore to investigate the possible replacement of active site copper by silver ions. Similarly to the perviously shown removal of copper by inactive FGE mutants, we then went on to analyze copper replacement by silver. The reaction rate in the presence of silver being significantly lower than with copper ( $k_{Ag}/k_{Cu} \approx 0.1$ , see also Fig. 4-5 section 4.1.3) one can again simply monitor the product formation. If silver would be capable of replacing copper from the active site, the following inactive Ag:FGE complex should lead to a lower rate of product formation. Indeed, when doing so, one can clearly observe a reduced rate of product release upon the addition of silver(I) (see Fig. 5-5). Reaching the half maximal rate at 0.5 equivalents of  $AgNO_3$ , the binding of newly introduced metal appears to be similarly strong as the copper binding. Calculations have been performed by others to determine the binding energies of copper and silver in complex with free hydrogen sulfide.<sup>[107]</sup> These studies show a decreasing binding energy for the series  $Au > Cu > Ag$ . However, other ligands from the the active site surroundings might additionally influence the metal binding energies of copper and silver. Upon addition of one equivalent of  $Ag^+$  to the reaction mixture the rate of product formation is decreased to 30 % of the initial rate in the absence of silver (only copper), indicating a stronger binding of silver. Following our previous experiments to determine copper off rates from FGE<sub>4C</sub> (see Fig. 5-2) similar efforts were made to confirm the earlier data (see Fig. 5-5). Upon the addition of 10 equivalents of  $AgNO_3$  to an ongoing reaction with FGE<sub>4C</sub> in the presence of Cu(I) the rate of product formation drops within 30 – 60 mins to a value even below the rate of FGE<sub>4C</sub> in the absence of copper ( $k_{Cu} = 0.35$  min<sup>-1</sup>,  $k_{+Ag, t=45\text{ min}} = 0.07$  min<sup>-1</sup>). Along with our previous hypothesis, that the inherent activity of FGE is attributed to residual, copurified copper, we believe, that silver also competes with the residual copper in the enzyme sample and is thereby capable of lowering the reaction rate even below that of native FGE. Again, the rate at which the product formation is slowed down describes the rate at which copper leaves the active site of FGE. In good agreement with the previous results the copper off rate in the presence of silver was determined to be  $k_{Cu, off} = 0.14$  min<sup>-1</sup> (see also in the presence of inactive mutants:  $k_{Cu, off} = 0.06 - 0.09$  min<sup>-1</sup>).

From the displayed data it is not clear, whether the reduced reaction rate is indeed derived from a replacement of copper from the active site. Inactivation of FGE by other pathways is likewise possible. As elevated concentrations of metal ions in protein solutions can cause agregation and inactivation of enzymes, we analyzed the resistance of the FGE : Ag complex to thermal stress. The point of thermal transition of FGE<sub>WT</sub> in the presence of one equivalent of silver nitrate is similarly increased as described previously for copper(I) (see Fig. 5-6 and chapter 4.1). Likewise, a three state transition is observed with part of the structure denaturing at 60 °C, whereas global unfolding is only achieved at temperatures above 80 °C. The extend of the low-temperature



**Fig. 5-5: Top:** Silver replaces the active site bound copper ion. **Bottom:** Rate of product formation of FGE<sub>4C</sub> in the presence of copper and silver. **Left:** Reaction mixtures containing a FGE:Cu mixture at a ratio of 1:1 were supplied with increasing amounts of AgNO<sub>3</sub> up to a Cu:Ag ratio of 1:10. Increased concentrations of silver lead to decreased reaction rates. **Right:** Time dependency of silver inhibition. Silver was added one minute after initiating the reaction. The off rate of copper from the active complex was determined as described above (see also Equation 3) to be  $k_{\text{off,Cu}} = 0.14 \text{ min}^{-1}$ .

unfolding, however, appears to be greater than previously observed for copper containing FGE. The comparable shift of the overall protein stabilization, however, suggests an analogous binding mode. It favours the hypothesis that silver moves into the copper binding site, where it replaces copper. Silver does therefore not inactivate the protein by unspecific interactions, such as oxidation or aggregation. Additionally, this observation suggests that the copper induced rate enhancement is decoupled from reorganization and stabilization of the tertiary structure. Stabilization of the overall structure is achieved even for inactive metal:FGE complexes. The increased reaction rate of the Cu:FGE complex is likewise not a result of the structural change.

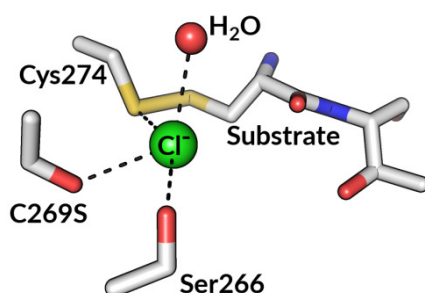


**Fig. 5-6:** Change in CD signal of FGE<sub>WT</sub> at  $\lambda_{\text{CD}} = 222 \text{ nm}$  at increasing temperature in the presence of silver nitrate. Silver ions induce a similar protein stability as observed for copper(I).

## 5.2 Conclusion

The presented data identifies FGE as a copper-dependent oxidase. While product formation occurs at rates of  $k_{\text{react}} = 0.7 - 1 \text{ min}^{-1}$  the rate of copper release has been determined to be  $k_{\text{Cu,off}} = 0.06 - 0.1 \text{ min}^{-1}$ . Copper binds tightly to the two cysteine residues previously identified in the active site, with a  $K_{\text{D,WT}} = 10^{-17.2} \text{ M}$ , which is comparable to other copper(I) containing proteins such as copper chaperon Atx1 or the copper transporting ATPase WLN5-6 ( $K_{\text{D,Atx1}} = 10^{-17.4} \text{ M}$ ,  $K_{\text{D,WLN5-6}} = 10^{-17.6} \text{ M}$  both at pH 7).<sup>[100]</sup> Mutants lacking either of the two highly conserved cysteine residue show no detectable affinity for Cu(I). Previously, these residues were thought to be involved in substrate binding through an intermolecular disulfide bond of the enzyme and the substrate.<sup>[56]</sup> In light of the presented affinities, together with the importance of copper for catalysis, we postulate that mutants lacking either of the two cysteine residues in the active site show inability to perform turnover due to insufficient copper binding. The mutant FGE<sub>S290K</sub> shows comparable copper affinities, but fails to compete for copper if other FGE variants such as FGE<sub>4C</sub> are present, indicating the participation of the substrate cysteine in the complex. A published crystal structure of FGE<sub>human</sub> (see Fig. 5-7) shows a chloride ion coordinated by Cys274 and the mutated Ser269 (cysteine in the original structure). The introduced serine residue contains a protonated oxygen ( $\text{p}K_{\text{a}} \sim 15$ ), while the original cysteine would rather be deprotonated ( $\text{p}K_{\text{a}} \sim 7-8$ ) and favour metal binding. The substrate cysteine is neatly positioned to further stabilize copper binding with a free coordination position on top to accommodate a potential oxygen molecule (water in the structure). While Ser266 appears to be less crucial for copper binding, we hypothesized, that copper will be bound at this position. Current crystal structures however fail to illustrate this dependency. The enzyme-substrate complex was only crystalized after introduction of a Cys  $\rightarrow$  Ser mutation in the active site.<sup>[56]</sup> As the reduced oxidation state of the active site cysteines is crucial for copper binding the presence of a disulfide between enzyme and substrate most likely renders the complex unable to bind copper.

Furthermore, the low affinity of oxidized FGE for copper ions might provide rationale for the absence of bound metal in any of the crystal structures published to date. Keeping FGE in a reducing environment during the crystallization experiments might be one route to circumvent this issue when co-crystallizing the FGE with copper. Replacement of copper by silver ions, could be a second route for successful crystallization of a metal bound FGE. The introduction of a different metal with the resulting loss in activity but remaining structural features (i.e. increased



**Fig. 5-7:** Possible copper-binding site in FGE. Active site of FGE<sub>human</sub> (PDB 2AIK, numbering based on FGE from *T. curvata*) with a co-crystallized chloride ion (green sphere) coordinated by three catalytically important residues as well as water.

---

thermal stability) further supports the participation of copper in redox chemistry. EPR spectroscopy showed an EPR-silent copper species, which suggests that copper(I) is the active metal in the active site. However, if copper is involved in redox chemistry, one would expect a +2 species to be present as well. Short living species that do not reach considerable concentrations during the ongoing reactions and therefore remain below detection limits as well as the generation of EPR silent cupric oxygen species have to be considered though.<sup>[108][109]</sup> Head-on oxygen binding was observed in synthetic Cu(II) complexes that do lack any EPR signal, which was attributed to the presence of a triplet state.

With copper only being released at a rate of  $k_{\text{off,Cu}} = 0.06 - 0.1 \text{ min}^{-1}$  and reaction rates of  $k_{\text{react}} = 0.5 - 1.0 \text{ min}^{-1}$  FGE should be categorized as a copper-dependent oxidase.

### 5.3 Experimental

**Plasmid Construction.** Mutations were introduced by site directed mutagenesis using the vector pET28-FGE<sub>WT</sub> as a template. Synthetic primers were ordered from Mycosynth and are shown in Tab. 5-2. After PCR, the fragments, as well as the pET28 target vector were treated with NdeI and XhoI restriction enzymes, followed by ligation of the gene into the target vector by T4 DNA ligase. The identity of final plasmids was confirmed by sequencing analysis (Microsynth).

**Tab. 5-2:** Primer sequences used for the mutation of the FGE variants described in this chapter.

FGE variant	Primer 3' → 5'	
	Sense	Antisense
FGE <sub>S266A</sub>	AAGGGCGGCGCTTTCCTGTGCCACGA GT	GGCACAGGAAAGCGCCGCCCTTGGT GAC
FGE <sub>S290K</sub>	CCGACTCCAAAGCCGCCCA	TGGGCGGCTTTGGAGTCGG
FGE <sub>W108A,W109A</sub>	GCGCCCGCGGCGGTGGCG	CGCCACCGCCGCGGGCGC
FGE <sub>C269S</sub>	GTTCTGTCCCACGAGTC	GACTCGTGGGACAGGAAC
FGE <sub>C274S</sub>	AGTCGTACTCCAACCGCTA	TAGCGTTGGAGTACGACT

**Protein Expression and Purification.** FGE mutants were produced as previously described (see chapter 4.3). Genes were transferred into BL-21 pLys *E. coli* cells. Precultures were inoculated with a single colony and incubated at  $T_{inc} = 37\text{ °C}$  overnight. Expression cultures of 1 L LB medium supplemented with kanamycin and chloramphenicol (50 mg/L and 34 mg/L, respectively) were inoculated with 1 mL of preculture. At OD = 0.6 – 0.8 cultures were induced by the addition of IPTG. After an additional 3 hrs of incubation at  $T_{express} = 37\text{ °C}$  the cells were harvested by centrifugation. The protein was then isolated by Ni<sup>2+</sup>- affinity chromatography and dialysed into NaCl supplemented Tris buffer pH 8 before storage at  $T_{store} = -80\text{ °C}$ .

**Tab. 5-3:** Calculated and observed molecular weights of proteins described in chapter 5 with comments on the mass difference.

Protein	MW <sub>calc</sub> [Da]	MW <sub>obs</sub> [Da]	Comment
FGE <sub>S266A</sub>	35292.1	35290.4	Disulfide bond
FGE <sub>S290K</sub>	35363.2	35360.4	Disulfide bond
FGE <sub>W108A,W109A</sub>	35077.8	35075.9	Disulfide bond

**Analysis of reaction rates.** FGE reaction rates were determined as described previously (see chapter 4.3).

**Peptide synthesis.** Peptides were synthesized by Fmoc solid phase peptide synthesis as described above (see chapter 4.3). The identity of the peptide was confirmed by ESI-MS (sequence: Abz-SALSSPTRA-NH<sub>2</sub>; MW<sub>calc</sub> = 1006.5 Da, [M+H]<sup>+</sup><sub>obs</sub> = 1007.5 Da, [M+2H]<sup>2+</sup><sub>obs</sub> = 504.2 Da).

**Copper(I) affinity determination.** Protein samples were diluted in a 50 mM Tris buffer pH 8 containing 20 μM CuSO<sub>4</sub>, 200 μM bathocuproine disulfonic acid (BCS), 2 mM cysteamine and 4 mM Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub>. The reaction mixture was vortexed and the absorbance at λ<sub>BCS</sub> = 483 nm was determined on a Cary 300 Bio UV/vis spectrophotometer. The concentration of free Cu : BCS<sub>2</sub> was calculated using the molar extinction coefficient of BCS ε<sub>BCS</sub> = 13000 M<sup>-1</sup>cm<sup>-1</sup>. The equilibrium is described by Eq. 4<sup>[100]</sup> which can be linearized to Eq. 5. The slope of the plot yields the K<sub>D</sub> value by using β<sub>2</sub> = 10.<sup>19.8</sup> M<sup>-2</sup>.<sup>[100]</sup> For measurements using reduced FGE, the protein was incubated with 5 mM DTT on ice for T<sub>inc</sub> = 20 min prior to removal of excess DTT by size exclusion chromatography. Oxidized FGE was prepared by treatment with H<sub>2</sub>O<sub>2</sub> prior to removal of excess oxidant by size exclusion chromatography. The oxidation state was confirmed by labeling with 2-iodoacetic acid and analysis by HRMS.

$$K_D \beta_2 = \frac{\frac{[P]_{total} - 1}{[MP]}}{\left(\frac{[L]_{total} - 2}{[ML_2]}\right)^2 [ML_2]} \quad (4)$$

K<sub>D</sub>: dissociation constant, β<sub>2</sub>: formation constant of CuBCS<sub>2</sub>, [P]<sub>total</sub>: total concentration of protein, [L]<sub>total</sub>: total concentration of BCS, [MP]: concentration of protein bound metal, [ML<sub>2</sub>]: concentration of ligand bound metal

$$\left(\frac{[L]_{total}}{[ML_2]} - 2\right)^2 [ML_2]([M]_0 - [ML_2]) = \frac{1}{K_D \beta_2} ([P]_{total} - ([M]_0 - [ML_2])) \quad (5)$$

with 
$$y = \left(\frac{[L]_{total}}{[ML_2]} - 2\right)^2 [ML_2]([M]_0 - [ML_2])$$

$$x = [P]_{total} - ([M]_0 - [ML_2])$$

and 
$$m = \frac{1}{K_D \beta_2}$$

$$y = \frac{1}{K_D \beta_2} \cdot x$$



---

**Labeling with 2-iodo acetamide.** FGE was diluted to a final concentration of  $c_{\text{FGE}} = 45 \mu\text{M}$  in a reaction mixture containing 5 mM cysteamine, 50 mM NaCl and 50 mM Tris-HCl pH 8. The reaction was incubated at RT for 30 min. Subsequently, 10 mM IAA was added and incubated for 60 min on ice. Protein modification was subsequently analyzed by LC/HRMS.

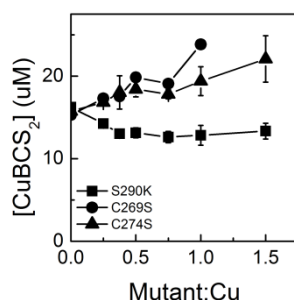
**Determination of Cu(I) off rate.** Analysis of metal release from the active site of active FGE was performed by diluting FGE<sub>4C</sub> to a final concentration of  $c_{\text{FGE}} = 0.5 \mu\text{M}$  in a 50 mM Tris buffer pH 8 containing 0.5  $\mu\text{M}$  CuSO<sub>4</sub>, 200  $\mu\text{M}$  substrate, 2 mM DTT, 50 mM EDTA and 50 mM NaCl. Reactions were initiated by addition of enzyme as the last step. After one minute of incubation the competing ligand was added to a final concentration of  $c_{\text{competitor}} = 5 \mu\text{M}$ . Product concentration was determined as described previously (see chapter 4.3). The resulting plot was fitted to Eq. 3, from which the rate of equilibration has been extracted as  $k_{\text{off,Cu}}$ .

$$[P] = A \cdot (1 - \exp(k_{\text{off,Cu(I)}} \cdot t)) + m \cdot t \quad (3)$$

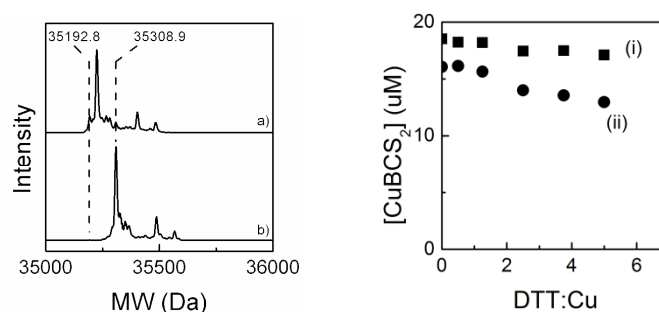
**EPR measurements.** For analysis of the copper oxidation state, samples of FGE<sub>4C</sub> were diluted to a final concentration of  $c_{\text{FGE}} = 50 \mu\text{M}$  in a 50 mM Tris buffer pH 8 containing 500  $\mu\text{M}$  substrate, 0 - 50  $\mu\text{M}$  CuSO<sub>4</sub>, 0 – 5 mM DTT, 50 mM EDTA and 50 mM NaCl. Reactions in the presence of DTT were freeze quenched after  $t_{\text{inc}} = 1$  min, whereas samples without reducing agent were incubated for  $t_{\text{inc}} = 15$  min prio to freeze quenching in liquid N<sub>2</sub>. Temperature: 150 K, modulation amplitude: 1 mT, microwave attenuation: 15 dB (6.346 mW power), 160 scans each. Spectra were acquired on a Bruker Elexsys 500 spectrometer equipped with a super-high Q resonator.

**Analysis of thermal unfolding by circular dichroism.** FGE was diluted to a final concentration of  $c_{\text{FGE}} = 10 \mu\text{M}$  in a buffer containing 0 – 10  $\mu\text{M}$  AgNO<sub>3</sub>, 50 mM Tris, 50 mM NaCl and 2 mM DTT. Thermal unfolding was monitored at  $\lambda_{\text{CD}} = 222$  nm in a quartz cuvette of 1 mm pathlength while heating at 1 °C/min in steps of 1 °C.

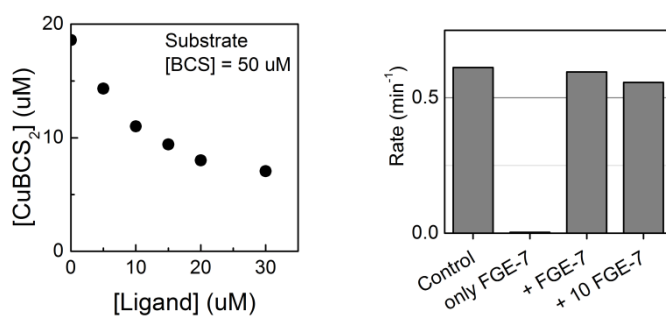
## 5.4 Appendix



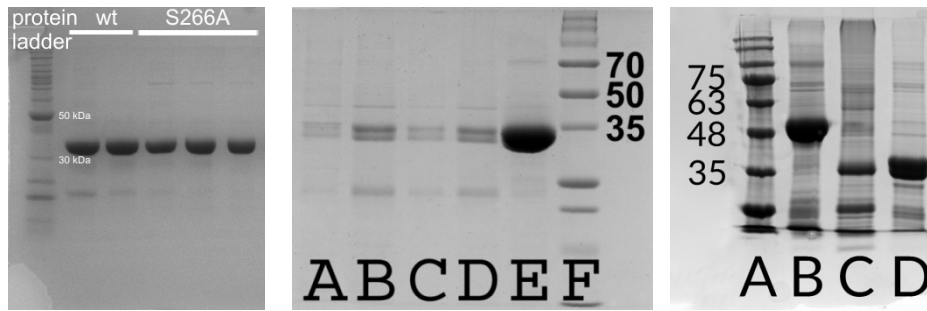
**Fig. 5-8:** Concentration of CuBCS<sub>2</sub> in the presence of increasing amounts of different FGE variants. The absorbance of the complex was assayed at  $\lambda = 483$  nm under the following condition: [CuSO<sub>4</sub>] = 20  $\mu$ M, [BCS] = 200  $\mu$ M, [cysteamine] = 2 mM, [FGE] = 0 – 30  $\mu$ M.



**Fig. 5-9: Left:** HRMS spectrum of FGE<sub>4C</sub> after labeling with 2-iodoacetamide. Native FGE does not react with IAA (a), whereas FGE<sub>4C</sub> treated with 5 mM cysteamine is readily alkylated (b, MW<sub>calc,4C,IAA</sub> = 35309.5 Da; MW<sub>obs</sub> = 35308.9 Da). **Right:** Determination of DTT binding affinity in the absence of cysteamine (i) and in the presence of 2 mM cysteamine (ii). Cysteamine does not significantly interfere with the assay ( $pK_{D, DTT} = 15.2$ ;  $pK_{D, DTT+cysteamine} = 15.9$ )



**Fig. 5-10: Left:** Concentration of CuBCS<sub>2</sub> in the presence of increasing amounts of cysteine containing substrate peptide. The absorbance of the complex was assayed at  $\lambda = 483$  nm under the following condition: [CuSO<sub>4</sub>] = 20  $\mu$ M, [BCS] = 50  $\mu$ M, [peptide] = 0 – 30  $\mu$ M. **Right:** Reaction rate of FGE<sub>4C</sub> in the presence of increasing amounts of serine containing peptide (FGE-7). FGE<sub>4C</sub> was diluted to a final concentration of  $c_{FGE} = 2$   $\mu$ M in a reaction mixture containing 200  $\mu$ M substrate (Abz-SALSCPTRA-NH<sub>2</sub>), 0 – 2000  $\mu$ M serine peptide (FGE-7; Abz-SALSSPTRA-NH<sub>2</sub>), 2  $\mu$ M CuSO<sub>4</sub>, 50 mM EDTA, 50 mM NaCl and 50 mM Tris pH 8.0. After various time points the reaction has been quenched for product analysis by addition of 1 volume equivalent of 1% TFA in 4 M urea. Product formation has been analyzed by RP-HPLC.



**Fig.5-11: Left:** SDS-PAGE picture of collected fractions of FGE<sub>S266A</sub>. **Middle:** SDS-PAGE picture of collected fractions of FGE<sub>W108A,W109A</sub>. A: Pellet; B: Flow through; C: wash with 10 mM imidazole; D: wash with 20 mM imidazole; E: Eluted Protein sample; F: MWM. **Right:** SDS-PAGE pictures of different FGE variants. A: Molecular weight marker; B: FGE<sub>smegmatis</sub>; C: FGE<sub>C269S</sub>; D: FGE<sub>S290K</sub>.

---

---

---

## 6. Identification of Reducing Agents for FGE

The FGE catalyzed reaction from a thiol to fGly is a 2-electron reaction with oxygen being the terminal electron acceptor. It has been suggested in the literature that oxygen is fully reduced to two molecules of water.<sup>[65]</sup> To reduce O<sub>2</sub> to H<sub>2</sub>O, however it requires four electrons. While during catalysis in FGE the substrate is oxidized by two electrons it requires an additional electron donor to fully reduce oxygen..

We and others have established that the FGE activity depends on thiol-containing reducing agents. *In vitro* reactions are therefore typically supplemented with DTT, mainly due to its low redox potential ( $E'_0 = -0.32 \text{ V}^{[117]}$ ). DTT however is an artificial compound and is not found in nature. The question remains, what is the reducing agent of FGE *in vivo*? While others have reported FGE catalysis in the presence of GSH, we did observe low reaction rates in the presence of glutathione (see also chapter 4.1) and therefore suggest that FGE requires a different electron donor *in vivo*.

In the following chapter we show that FGE performs a single turnover in the absence of an efficient reducing agent. The active site cysteines are then reversibly oxidized via a two electron oxidation. The resulting modification is equivalent to a disulfide bond in terms of oxidation equivalents but cannot be identical since copper binding is still possible. Furthermore, we provide evidence that an N-terminal fragment containing a CxC-motif is a preferred redox partner and efficiently reduces the enzyme during turnover.

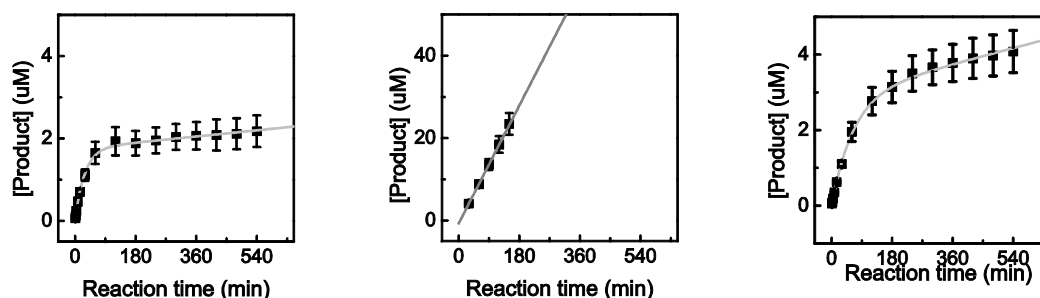
### 6.1 FGE reduction *in vitro*

#### 6.1.1 FGE activity in the absence of reducing agent

FGE catalysis requires a reducing agent for multiple turnovers to occur. To gain knowledge on the redox state of FGE before reduction we performed the reaction in the absence of reducing agents. To simplify the setup we chose to initially perform the reaction in absence of copper. Without reducing agent the product formation is slowed down within a few hours and reaches a plateau (see Fig. 6-1 and Ref. [65]). Addition of DTT after equilibration recovers the initial activity as describe above (see chapter 4.1). The plot of the product concentration over time shows a burst phase with a steady, linear phase after three hours of incubation. The data is therefore best fitted with an exponential fit adjusted by a linear function (see Eq. 7). When using FGE<sub>4C</sub> the amount of product formed in the absence of reducing agent before reaching the linear phase matches the enzyme concentration in the reaction mixture ( $C_{\text{FGE}} = 2 \mu\text{M}$ ,  $C_{\text{Prod,exp}} = 1.8 \mu\text{M}$ ), which is a strong indication for inactivation after a single turnover. Since the substrate peptide contains a thiolate, the second linear phase most likely represents the steady state reaction using the substrate as a reducing agent at a low reaction rate of  $k_{\text{substrate}} = 0.4 \cdot 10^{-3} \text{ min}^{-1}$  (vs.  $k_{\text{DTT}} = 0.16 \text{ min}^{-1}$ ). When repeating the reducing agent-free experiment with FGE<sub>WT</sub> product formation is similarly slowed down. The qualitative observation remains the same. Both reactions, with FGE<sub>4C</sub> and FGE<sub>WT</sub> are slowed down within few hours. FGE<sub>4C</sub> shows enhanced reactivity during the

$$[P] = -[P]_{exp} \cdot \exp(-t_{react} \cdot k_{inact}) + k_{linear} \cdot t_{react} + b \quad (7)$$

$[P]_{exp}$ : product concentration after the exponential phase  
 $k_{inact}$ : rate of FGE inactivation  
 $k_{linear}$ : reaction rate of FGE reduced by the substrate  
 $b$ : offset of linear phase

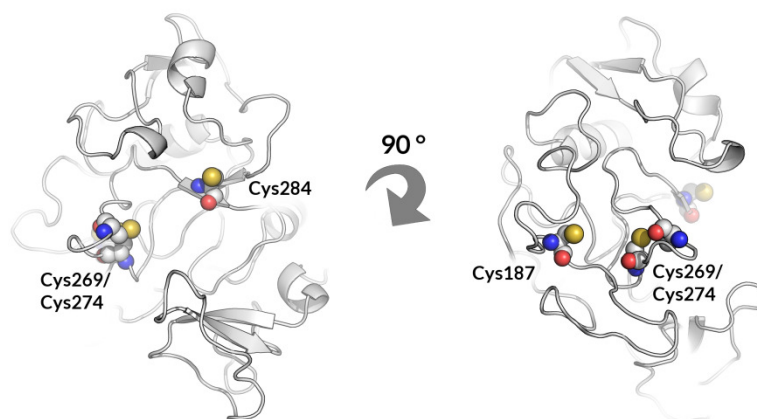


**Fig. 6-1:** Product formation of FGE in the absence and in the presence of reducing agent. **Left:** If no reducing agent is added, FGE<sub>4C</sub> performs only one turnover ( $C_{FGE} = 2 \mu\text{M}$ ) after which it reaches its oxidized state, waiting for reduction to occur. **Middle:** In the presence of DTT multiple turnover ( $C_{FGE} = 2 \mu\text{M}$ ) are performed without any sign of inactivation. **Right:** Reducing agent free reaction of FGE<sub>WT</sub>. The presence of additional cysteine residues appears to favor the reduction reaction by internal disulfide shuffling.

*in vitro* assays described earlier (see chapter 4.1). This is also observed in the absence of reducing agent. The rate at which FGE<sub>4C</sub> is inactivated is 2.3-fold higher than that of FGE<sub>WT</sub> (see also Tab. 6-1). While the maximal product formation when using FGE<sub>4C</sub> is equivalent to the enzyme concentration, the product concentrations with FGE<sub>WT</sub> exceed  $2 \mu\text{M}$  ( $C_{FGE} = 2 \mu\text{M}$ ). The reason for the higher turnover number is most likely the presence of additional cysteine residues on FGE<sub>WT</sub>. These might aid to use the substrate as a reducing agent by providing additional sites of disulfide exchange. This might be a slow process which only occurs if no suitable reducing agent is present to perform reactivation of the enzyme. The closest cysteine residues in FGE<sub>WT</sub> can be found within a distance of 9 and 13 Å for Cys187 and Cys284, respectively (see Fig. 6-2, estimation based on a theoretical structure modeled on the basis of a published structure from FGE<sub>human</sub> using swiss model). Since these cysteine residues reside on the surface of FGE, disulfide exchanges between FGEs are also to be taken into account. To test for oligomers of FGE we assayed the inactivated FGE by SEC. Interestingly, the chromatogram of FGE<sub>WT</sub> shows multiple peaks corresponding to dimers and multimers, after incubation with excess of substrate (see Fig. 6-3, left). In contrast the retention volume of FGE<sub>4C</sub> after turnover is indistinguishable from that of unreacted FGE<sub>4C</sub> (see Fig. 6-3, middle), suggesting that scaffold cysteines are indeed involved in disulfide exchange within multiple FGEs. The active site cysteines however are not involved in dimerization. To further characterize these isolated FGE intermediates we analyzed the content of thiolates as well as the protein mass. After incubation with the substrate in absence of DTT, the wild type protein shows less free thiols than the unreacted FGE<sub>WT</sub> ( $4.3 \pm 0.2 \text{ SH/FGE}$  before and  $2.8 \pm 0.1 \text{ SH/FGE}$  after turnover). Along with the previous observation that FGE<sub>WT</sub> oligomerizes in absence of DTT we propose that these additional disulfides are formed within scaffold cysteines outside of the active site. In agreement with this hypothesis FGE<sub>4C</sub>

**Tab. 6-1:** Fitting parameters obtained from fitting the reducing agent independent data to Eq. (7).

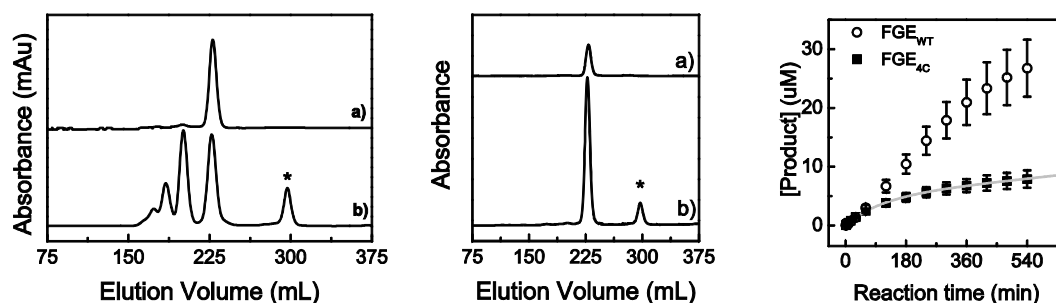
Protein	$k_{\text{inact}}$ [ $\text{min}^{-1}$ ]	$[P]_{\text{exp}}$ [ $\mu\text{M}$ ]	$k_{\text{linear}}$ [ $\mu\text{M}\cdot\text{min}^{-1}$ ]
FGE <sub>WT</sub>	0.014	3.2	$1.7\cdot 10^{-3}$
FGE <sub>4C</sub>	0.032	1.8	$0.58\cdot 10^{-3}$
FGE <sub>WT</sub> + Cu	n.a.	n.a.	n.a.
FGE <sub>4C</sub> + Cu	0.009	4.7	$6.1\cdot 10^{-3}$



**Fig. 6-2:** Modelling of structure of FGE<sub>WT</sub> based on the published structure of FGE<sub>human</sub> (1Y1E). Shown are the two active site cysteines Cys269 & Cys274 together with other present cysteines in the vicinity. The closest distance between the sulfur atoms of Cys284 and Cys274 is 13 Å, whereas closest distance between the sulfur atoms of Cys187 and Cys269 is 9 Å. **Left:** Front view. **Right:** Side view.

shows an oxidized redox state before and after turnover ( $0.3 \pm 0.1$  SH/FGE after turnover). Along with this observation, the mass spectrum of FGE<sub>4C</sub> is indistinguishable from the unreacted enzyme with only a mass shift of  $\Delta MW_{\text{obs}} = 2$  Da which is reversible by addition of DTT ( $MW_{\text{calc},4C} = 35195.5$  Da,  $MW_{\text{obs}} = 35193.5$  Da, and  $MW_{\text{obs}+\text{DTT}} = 35195.5$  Da). Importantly, the isolated species shows does catalyze additional product formation if incubated with new substrate (see appendix Fig. 6-9). When DTT is added, though, the reaction proceeds and product is generated over many turnover. When repeating the DTT free experiments in presence of copper, the product formation is similarly slowed down. Both, FGE<sub>WT</sub> and FGE<sub>4C</sub>, show an increased product formation, though. While the general scheme is reproducible even in the presence of the redox active cofactor, other pathways occur that complicate the system. One possible explanation is that copper coordinates free thiols and thereby serves as an electron transporter to assist reduction by the substrate. Data evaluation, however, is more difficult (see Fig. 6-3, right). Again, the presence of multiple cysteine residues in FGE<sub>WT</sub> causes the reaction to proceed at a higher reaction rate without strong evidence for inactivation within 9 hrs (for comparison: copper free FGE<sub>WT</sub> reaching the linear phase after 3 – 4 hrs) making fitting to Eq. 7 impossible.

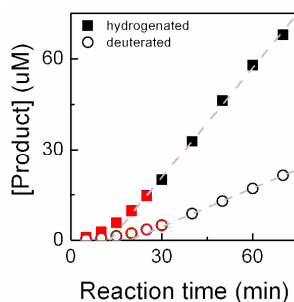
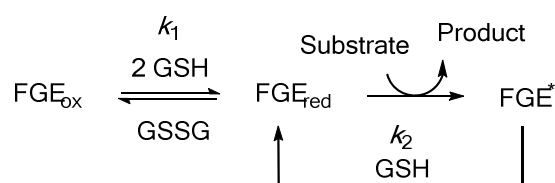
The described observations are consistent with the presence of a disulfide bond between the two active site cysteines as well as within multiple FGE molecules if additional cysteines are present. Importantly, even in excess of substrate, mixed disulfides of FGE and the peptides are not observed during LC-MS measurements. Previously, disulfide bond formation has been observed



**Fig. 6-3:** **Left:** Size exclusion chromatogram showing FGE<sub>WT</sub> after DTT independent turnover (b) in comparison to the unreacted enzyme (a). Asterisk indicates the elution peak of the reaction product. **Middle:** Size exclusion chromatogram showing FGE<sub>4C</sub> after DTT independent turnover (b) in comparison to the unreacted enzyme (a). Asterisk indicates the elution peak of the reaction product. **Right:** Product formation of FGE in the absence of reducing agent but in presence of copper. Copper addition facilitates the reduction by the substrate and thereby leads to enhanced enhanced product formation compared to the copper free reaction.

in crystal structures with co-crystallized substrate. However, during our experiments we do not have evidence for such covalent interactions. Heterodimer formation with the active site cysteines, as well as to other solvent exposed cysteines throughout the enzyme scaffold, is not observed.

Analysis of the structure of FGE after a single turnover is difficult. Potential instability of the oxidized species or presumably short living intermediates and the presence of other thiols such as the substrate might conflict with the isolation and identification of the precise modification in the active site. We have previously established the presence of a disulfide bond in the active site of FGE<sub>4C</sub> and FGE<sub>WT</sub> (see section 3.1.2). Enzymes in the absence of reducing agents will hence carry said sulfur-sulfur bond. When incubating FGE (FGE<sub>4C-Y273F</sub>, for details see chapter 7.2) with glutathione, the product formation shows a lag phase with the initial few turnover being slower than the later turnover in the second linear phase. This lag phase is not observed with DTT.



**Fig. 6-4: Top:** Reaction scheme describing the reduction of the disulfide containing FGE<sub>ox</sub> by GSH and the reactivation of FGE<sup>\*</sup>. **Bottom:** Product formation of FGE<sub>4C-Y273F</sub> with deuterated (empty circles) and non-deuterated (solid squares) substrate using GSH (c<sub>GSH</sub> = 2 mM) as a reducing agent shows lag phase. The linear phase (black) is fitted with a linear curve (dashed lines). The rates of reduction of FGE<sub>ox</sub> have been determined by fitting the data points of the initial 30 min with  $[P]=A \cdot \exp(-k \cdot t_{\text{react}})+Y_0$ .



---

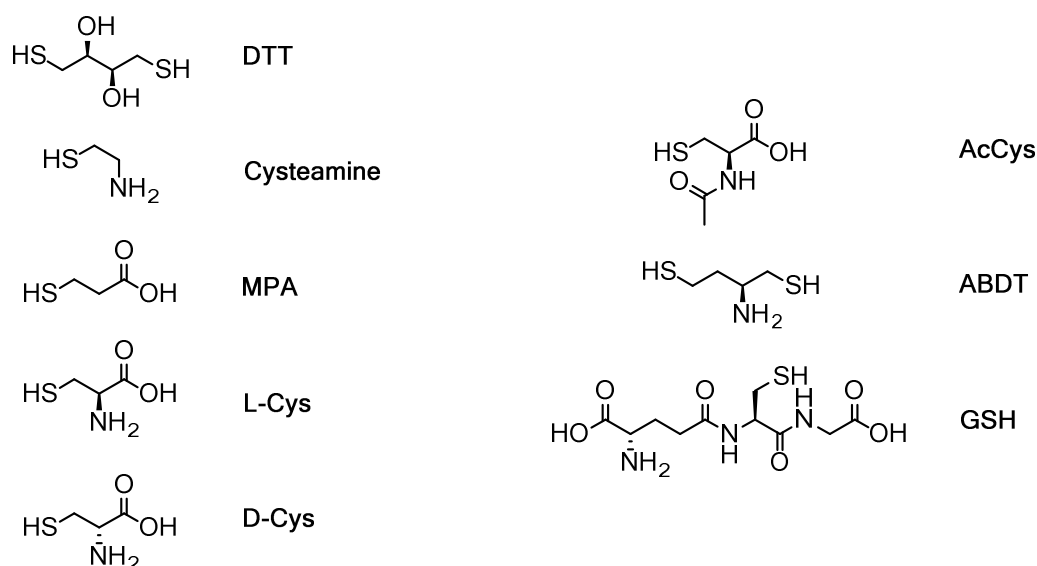
The increasing reaction rate during catalysis in presence of GSH indicates that reactivation of the inactivated FGE (FGE\* in Fig. 6-4) is faster than the initial reduction of the disulfide bond in FGE<sub>ox</sub> ( $k_2 > k_1$ ). The extent of the lag phase though, is independent of the reaction rate, which is observed when using deuterated substrate ( $\beta$ -proton substituted by deuterium). While the secondary linear phase shows the expected isotope effect of KIE = 2.9, the rate of reduction of FGE<sub>ox</sub> in the lag phase is constant at  $k_{\text{red,H}} = 0.06 \text{ min}^{-1}$  and  $k_{\text{red,D}} = 0.07 \text{ min}^{-1}$  (at  $c_{\text{GSH}} = 2 \text{ mM}$ ) with the non-deuterated and the deuterated substrate respectively, indicating that the initial reduction is decoupled from the C-H bond breaking reaction. Since the linear phase shows the isotope effect, the reactivation reaction ( $k_2$ ) must be faster than the rate limiting step. The structure of FGE\* and FGE<sub>ox</sub> must thereby be different.

In summary, FGE cycles through an oxidized intermediate which is not identical to the resting state of the enzyme containing a disulfide bond. Additionally the formation of a disulfide bond during turnover, would cause release of the metal ion due to the low binding affinity of the disulfide-containing FGE<sub>ox</sub>. Unbinding of copper, however, is not observed since the rate of copper release from the enzyme was shown to be lower than the rate of product formation (see chapter 5). Possible oxidized enzyme intermediates that are capable of metal binding are sulfinic and sulfenic acid derivatives of the active site cysteines as found in nitrile hydratases.<sup>[111]</sup> Sulfinic acids however are not readily reduced by thiols at ambient conditions and require complex enzymatic systems to be reduced *in vivo*.<sup>[112]</sup> Along with previous work of others, we therefore suggest the formation of a sulfenic acid in the active site of FGE, even though it has not been detected to date.<sup>[65]</sup>

### 6.1.2 Scope of reducing agents compatible with FGE

Previous reports suggest a stimulation of FGE activity in the presence of low millimolar concentrations of glutathione and DTT, with high concentrations of GSH inhibiting fGly generation.<sup>[3]</sup> In order to question the dependence on reducing agents of FGE we performed further experiments with a broader range of reducing agents (see Scheme 6-1 and Fig. 6-5, left). Others have previously suggested that FGE can catalyze thiol oxidation in the presence of GSH. In contrast this, during our *in vitro* assays, the presence of GSH does not support FGE catalysis (insufficient reduction of FGE<sub>ox</sub>, see above). Reactions typically contain EDTA as a chelator for excess and undesired metal ions. To test whether trace metals in the reaction buffer could facilitate the reaction with GSH we reacted FGE with GSH in absence of the metal chelator. Interestingly the GSH-mediated reaction in presence of EDTA is significantly slower than in absence of EDTA (see Fig. 6-5, middle). We therefore hypothesize that trace metals in the reaction buffer assist the transfer of reduction equivalent from GSH towards FGE. However, there is no extensive study investigating this hypothesis to date.

Only few small molecular weight thiols do support efficient catalysis with FGE. Dithiol containing reducing agents like DTT and 2-amino butanedithiole (ABDT) perform just as well as very small molecules like cysteamine. In contrast, none of the other tested reductants did enhance the reaction rate to that extent.

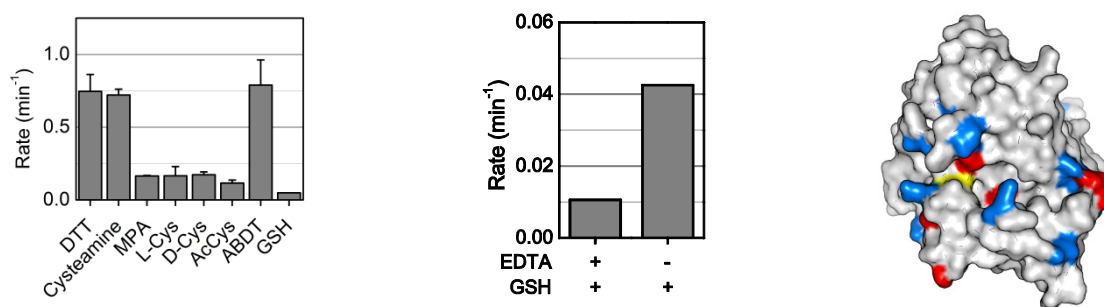


**Scheme 6-1:** Scope of reducing agents tested for their potential to efficiently support the FGE catalyzed aldehyde formation.

One possible explanation for the different reactivities is the different  $pK_a$  of the thiol group in each of the reducing agents. Comparison of the  $pK_a$  and the activity of FGE in the presence of the respective reductant suggests that there is no correlation between thiol  $pK_a$  and reactivity. Typical reaction mixtures for analysis of formylglycine generation are buffered at pH 8, which results in the majority of reductants being protonated (see Tab. 6-2). The fastest product formation is obtained for thiols with a  $pK_a$  range of 8.2 – 9.2, while other thiols like cysteine also share similar  $pK_a$  values but do not accelerate the reaction. One common characteristic of well performing thiols is the absence of carboxylic acid groups, suggesting a hindered binding of negatively charged reducing agents. Under the assay conditions the acid groups will be negatively charged and potentially conflict with negative charge on the protein surface. The active site of FGE<sub>WT</sub> is flanked by several negatively charged residues (namely Glu40, Glu105 & Glu271, see also Fig. 6-5, right) which might create a degree of repulsion and hence hinder the movement of the reducing agent to the active site. In the human variant these residues are found to be Gln110, Ala176 and Arg338, hence either positive or neutral residues. It is therefore not clear

**Tab. 6-2:**  $pK_a$  values and redox potentials of thiol groups of different reducing agents together with their relative activity using FGE<sub>4c</sub>.

Red. Agent	$pK_a$ of RS-H	$E'_0$ [V]	Rel. Activity [%]
DTT	9.2	-0.32 <sup>[117]</sup>	100
Cysteamine	8.2 <sup>[113]</sup>	-0.24 <sup>[116]</sup>	96.6
MPA	10.3 <sup>[114]</sup>	-	22.2
L-Cys	8.3 <sup>[113]</sup>	-0.24 <sup>[116]</sup>	22.2
D-Cys	8.3 <sup>[113]</sup>	-0.24 <sup>[116]</sup>	23.3
AcCys	9.5 <sup>[113]</sup>	-	15.5
ABDT	8.2 <sup>[115]</sup>	-0.32 <sup>[117]</sup>	105.7
GSH	9.6	-0.25 <sup>[116]</sup>	6.5



**Fig. 6-5:** Rate of product formation of FGE<sub>4C</sub> in the presence of different reducing agents. **Left:** Rate of product formation of FGE<sub>4C</sub> in the presence of different reducing agents (see Scheme 6-1). Either small molecular weight thiols or dithiols efficiently perform reducing reactions. **Middle:** Rate of product formation in the presence of DTT and glutathione is dependent on the presence of metal chelators in the reaction mixture. **Right:** Homology model of the structure of FGE from *T. curvata* based on a published structure of FGE from *H. sapiens* (PDB 1Y1E). Aspartates and glutamates are highlighted in blue, while histidines, arginines and lysines are highlighted in red. The active site cysteines are colored yellow.

whether a similar dependency will be observed for FGE<sub>human</sub>. An extensive analysis of the efficiency of a wide array of reducing agents was not performed to date.

In conclusion, FGE from *T. curvata* shows a preference for reducing agents with two mercapto groups or very small reducing agents. Negative charges other than the thiolate decrease the reactivity.

### 6.1.3 Reducing agent *in vivo*

The best reducing agents for the *in vitro* reaction contain two thiol groups or are very small. DTT and ABDT both display a very low redox potential making them extraordinary good reducing agents (see Tab. 6-2). These, however, are not available in the cellular context. The question of what is the electron source *in vivo* remains. Typical reducing agents *in vivo* are GSH, mycothiol, thioredoxins or disulfide isomerases. The latter require NADH as a terminal electron donor. While the reduction potential of these thioredoxins is lower than that of DTT or ABDT ( $E'_{0,Thx} = -0.23$  to  $-0.27$  V<sup>[118]</sup>) they do share the common feature of the bis-thiol functionality.

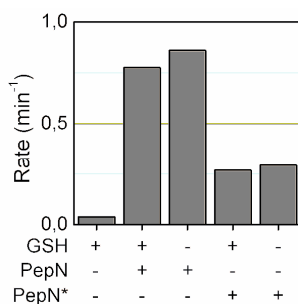
In humans FGE interacts with ERp44 (see also section 2.3.2, p. 13) through the N-terminal domain of FGE (residue 34-68 of the human variant) through disulfide bond formation as well as non-covalent interactions (see also chapter 2.3.2).<sup>[51]</sup> This N-terminal domain has been identified as being crucial for retention in the ER and hence for *in vivo* activity, but appears to be dispensable for *in vitro* reactivity.<sup>[51]</sup> It contains two cysteine residues (Cys50 and Cys52) that are involved in binding to ERp44 but also lead to homodimer or intramolecular disulfide bond formation.<sup>[45]</sup> The homodimer, however, appears to be labile and is readily reduced to the monomers in the presence of a reducing agent. The *in vivo* activity only depends on the presence of Cys52 but not Cys50.<sup>[51]</sup> Such a CxC motif has been described as a mimic of disulfide isomerases, which utilize a CxxC motif to catalyze disulfide shuffling.<sup>[110]</sup> With the N-terminal domain in eukaryotic FGEs containing the CxC motif, we envision that FGE activity might be restored by the presence of the N-terminal domain even if additional reducing agents are absent.

To test whether this domain can support the *in vitro* reaction, we synthesized a peptide encoding the core motif of the N-terminal domain from human FGE (see Fig. 6-6). In addition to the CxC motif the sequence of the N-terminal domain shows a second interesting characteristic. The

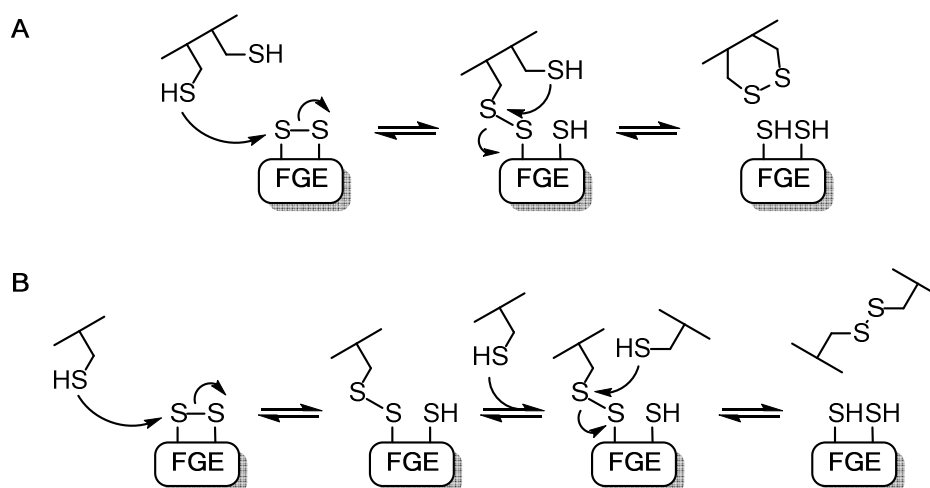
N-terminal Domain	SQEAGTGAGAGSLAGSGGGTTPQRPGAHGSSAAAHRYSR
PepN	-----BAGSGGGTTPQRPGAHGSS-----
PepN*	-----BAGSGGGTRQPHGAPGSS-----
Substrate 1	-----LC-TPSRA-----
Substrate 2	-----BSALC-SPTRA-----

**Fig. 6-6:** Sequence comparison of the N-terminal domain from FGE<sub>human</sub> (residue 34 – 72 of the human FGE gene), the two synthetic peptides used for the reducing agent analysis (PepN and PepN\*, single letter code ‘B’ for 2-aminobenzoic acid) as well as two substrates (substrate 1 used for cocrystalization in Ref [56]; substrate 2 used in this study). Cysteine residues are highlighted in green (Cys50 and Cys52 of FGE<sub>human</sub>). Conserved residues, that are potentially involved in binding are highlighted in yellow.

second two cysteines are very close in proximity to an amino acid sequence that is similar to the substrate recognition sequence (CxxPxR vs. CxPxR in the substrate). We therefore hypothesize, that the N-terminal domain binds to the enzyme via similar interactions as the substrate. To test this we additionally synthesized a second peptide with potentially reducing properties (PepN\*), but with key residues in the recognition motif (Pro, Arg and His) being displaced. With the two peptides in hand we assayed FGE<sub>4c</sub> activity in the presence of either of the two peptides. The variant with only two cysteine residues was chosen in order to reduce possible side effects and minimize disulfide shuffling of scaffold cysteines. As a terminal electron donor we selected GSH and supplemented each reaction additionally with 2 mM GSH to ensure that PepN and PepN\* would be present in the reduced oxidation state. The reaction rates in the presence of PepN (see Fig. 6-7) are comparable to previously observed reaction rates in the presence of DTT or ABDT ( $k_{\text{DTT}} = 0.75 \text{ min}^{-1}$ ;  $k_{\text{PepN}} = 0.78 \text{ min}^{-1}$ ). In the control reaction with PepN\* product formation was observed at a 3-fold lower rate, which is in agreement with a less efficient binding of the scrambled peptide sequence. The presence of GSH however does not effect the reaction rate. Despite the long sequence of PepN and PepN\*, catalysis under participation of either of the two, occurs at significantly higher rates than in the presence of the tripeptide GSH, indicating that size is infact not the limiting factor. The presence of a two mercapto-groups in close vicinity to each other is most likely more important. The rational design of the two reducing agents suggests a disadvantageous binding of the randomized peptide structure. In good agreement with this hypothesis the  $K_M$  value with FGE<sub>4c</sub> is  $K_{M,\text{PepN}} = 40 \mu\text{M}$  and  $K_{M,\text{PepN}^*} = 70 \mu\text{M}$  while in the case of DTT this was determined to be  $K_{M,\text{DTT}} = 1000 \mu\text{M}$  (Fig. 6-8, left and middle). The catalytic efficiency of FGE<sub>4c</sub> in the presence of PepN is therefore 25-fold higher than in presence of DTT. With DTT being a non-natural reducing agent, in the light of this data, it is more likely to be the case, that

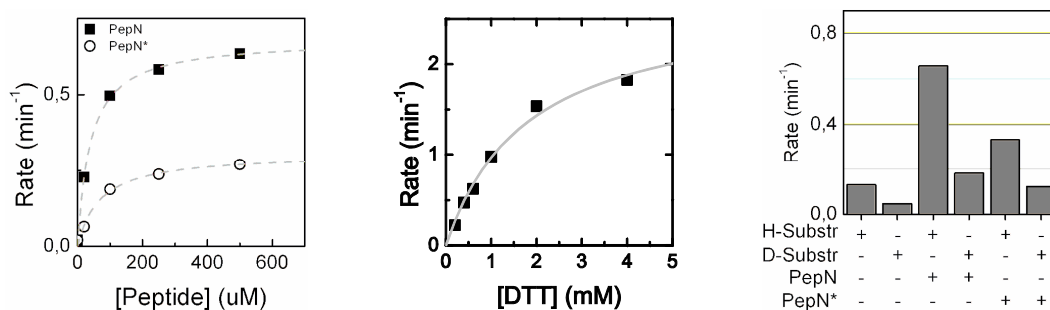


**Fig. 6-7:** Rates of product formation of FGE<sub>4c</sub> in the presence of PepN and PepN\* as a reducing agent. While both peptides are capable of reducing FGE only PepN performs to a similar extend as DTT or ABDT. Supplementation with GSH does not significantly affect product formation.



**Scheme 6-2.** Reaction scheme of dithiol reduction of oxidized FGE (A) versus reduction by two individual thiols (B). The close vicinity of the second thiol in A favours the direct attack and reduces the enzyme. If dependent on two individual thiol molecules, the reaction is slowed down by the back reaction of the mixed disulfide to the oxidized FGE and by the limitations in diffusion of the second thiol towards the active site.

FGE requires a peptide type reducing agent. Due to the lack of an extensive investigation of reducing peptides of different length and sequence one cannot further comment on the ideal structure and sequence, though. Importantly, the KIE of FGE is maintained with the peptidyl reducing agent ( $KIE_{\text{PepN}} = 3.6$ ,  $KIE_{\text{PepN}^*} = 2.7$ , see Fig. 6-8, right), indicating that the mechanism of action is not perturbed. DTT and ABDT exhibit very low reduction potentials, the reduction potentials of PepN and PepN\* are yet to be determined. In a cellular context, thioredoxin often perform reduction reactions. The redox potential of such thioredoxins are typically carefully tuned to a value that allows for oxidation by other oxidized substrates, but likewise allows for subsequent reduction by a secondary (most often NADH driven) system to regenerate the active thioredoxin.<sup>[119]</sup> A too low reduction potential would thereby trap the Trx in the oxidized state and hence render it impossible to be reduced (compare  $E'_{0,\text{Trx}} = -0.23$  to  $-0.27$  V;  $E'_{0,\text{GS}} = -0.25$  V). We therefore hypothesize that the redox potential of the N-terminal domain should be in a similar range. With this we propose, that a reducing agent for FGE *in vivo* should display two cysteines in close vicinity to each other.



**Fig. 6-8:** **Left:** Rate of product formation in dependency of increasing amounts of PepN (solid squares) and PepN\* (empty circles). Both reducing peptides efficiently reduce FGE in the low micromolar range. **Middle:** Rate of product formation in dependency of the concentration of DTT. **Right:** Kinetic isotope effect of FGE<sub>4C</sub> in the presence of PepN and PepN\*. None of the reducing peptides changes the rate limiting step of the FGE catalyzed reaction.

---

In case of two individual thiolates (see Scheme 6-2, B), after the initial attack of the first thiol one requires the presence of a second molecule. If this binding to the active site is slow (due to steric hindrance in the packed enzyme or colombic repulsion as described above) the back reaction can occur, which then releases the oxidized enzyme and the reduced thiol again. Additionally, the combination of two individual molecules to one oxidized homodimer disfavors the reaction entropically.

We then went on to test the observed reducing capability with FGE<sub>human</sub>. While its N-terminal tag was shown to be without function *in vitro* earlier, these studies included additional DTT in the reaction mixture.<sup>[45][49]</sup> To test for possible effects, we used a full length version of human FGE with the N-terminal domain (residues 36-72) still in place (FGE<sub>human-FL</sub>). Although the overall reaction rates of (truncated) FGE<sub>human</sub> and FGE<sub>human-FL</sub> appear to differ significantly (inactivation of the protein samples, upon long time storage at T<sub>store</sub> = 4 °C, see Appendix Fig. 6-10), one can observe an overall trend. Along with FGE<sub>WT</sub>, the truncated FGE<sub>human</sub> does not react with GSH, whereas the reaction rate of FGE<sub>human-FL</sub> with GSH is less reduced (see Tab. 6-3). While this assessment is far from being complete, it should point out that the presence of the N-terminal domain of FGE as a covalently linked reducing agent can increase product formation and hint at a potential role of this domain also *in vivo*. The reader is thereby invited to follow the discussion in future publications.

While working with human proteins tend to be difficult not only in terms of production and purification but also with regards to stability, bacterial homologues are often used to circumvent these issues. Synthesis of a fusion of a bacterial FGE (possibly FGE<sub>WT</sub> or FGE<sub>4C</sub>) with the human N-terminal domain could lead to data further supporting the peptide supplementation assay.

**Tab. 6-3:** Relative activity of FGEs in the presence of GSH with respect to the rate of product formation in the presence of DTT.

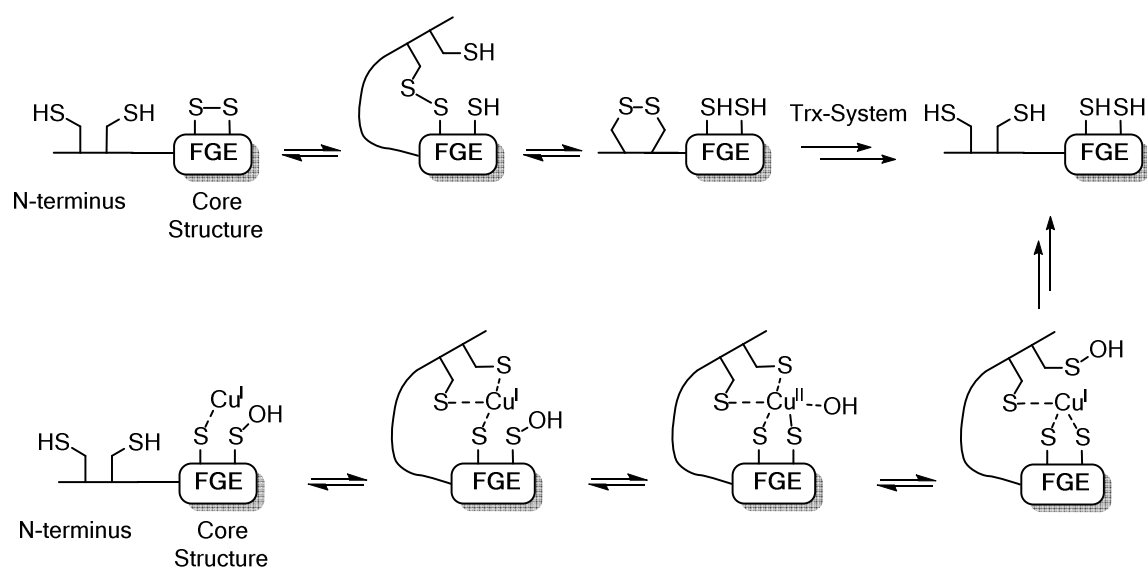
Variant	$k_{\text{GSH}}/k_{\text{DTT}}$ [%]
FGE <sub>WT</sub>	12
FGE <sub>human</sub>	16
FGE <sub>human-FL</sub>	43

## 6.2 Conclusion

This chapter shows the dependency of FGE on reducing agents. In the absence of a suitable reducing agent the enzyme is inactivated to a (most likely) sulfenic acid derivative of the active site cysteines. Which under participation of a reducing agent is reduced to the thiol. The scope of reducing agents has been expanded towards a more general view. While cysteine and cysteine derivatives, as well as glutathione do not efficiently reduce the enzyme, dithiol containing molecules readily reactivate FGE. The very small cysteamine reacts similarly fast, most likely due to the possibility to quickly diffuse in the packed active site. The requirement of two thiols being present in close vicinity to each other rather than having a particular size was additionally demonstrated by usage of large peptides containing the thioredoxin like CxC motif. The peptide synthesis was inspired by the N-terminal domain of human FGE which is required for its *in vivo*-reactivity. This domain has previously been reported to aid in localization in the ER where FGE activity is required. The presented data however suggests an additional role of this domain during catalysis (see Scheme 6-3). An *in vitro* reaction complemented with a peptide derivative of the N-terminal domain, showed fast catalysis with good affinities for the large reducing agent. While the active site could be too packed to be reduced by a thioredoxin/thioredoxin reductase system directly, intermediate reduction by the N-terminal domain might facilitate this process.

We and others have proposed the formation of a sulfenic acid during turnover. Along with copper being present throughout turnover, metal assisted reduction of the sulfenic acid is likewise possible. The oxidation equivalent will thereby be transferred to the metal centre from which the reducing N-terminal is oxidized to the sulfenic acid. Disulfide bond formation and release of water is then followed by reduction by a Trx-system to regenerate the active FGE catalyst.

We do envision that these findings will drive the mechanistic investigations on FGE and that future investigations will identify similar reducing agents in FGEs from other organisms.



**Scheme 6-3:** Hypothetical mechanism of the reduction of FGE by the N-terminal domain encoding the CxC motif. The ideal positioning of the two cysteines favors reduction of the active site cysteines. The exposed location of the oxidized N-terminal cysteines then allows for reduction by a suitable thioredoxin system.

---

### 6.3 Experimental

**Protein Expression and Purification.** FGE mutants were produced as described previously (see chapter 4.3). Genes were transferred into BL-21 pLys *E. coli* cells. Precultures were inoculated with a single colony and incubated at  $T_{inc} = 37\text{ }^{\circ}\text{C}$  overnight. Expression cultures of 1 L LB medium supplemented with kanamycine and chloramphenicol (50 mg/L and 34 mg/L, respectively) were inoculated with 1 mL of preculture. At OD = 0.6 – 0.8 cultures were induced by addition of IPTG. After additional 3 hrs of incubation at  $T_{express} = 37\text{ }^{\circ}\text{C}$  the cells were harvested by centrifugation. The protein was then isolated by  $\text{Ni}^{2+}$ - affinity chromatography and dialysed into Tris buffer pH 8 before storage at  $T_{store} = -80\text{ }^{\circ}\text{C}$ .

Samples of FGE<sub>human</sub> and FGE<sub>human-FL</sub> were kindly supplied by Prof. Dr. Thomas Dierks, University of Bielefeld.

**Analysis of product formation in absence of reducing agent.** If not stated otherwise, FGE was diluted to a final concentration of  $c_{FGE} = 2\text{ }\mu\text{M}$  in a reaction buffer containing 200  $\mu\text{M}$  substrate peptide, 50 mM EDTA, 50 mM NaCl and 50 mM Tris-HCl buffer pH 8. Samples containing additional copper salts were supplemented with  $\text{CuSO}_4$  at a final concentration of  $c_{metal} = 2\text{ }\mu\text{M}$ . Control reactions with DTT were additionally supplemented with the reducing agent at a final concentration of  $c_{DTT} = 2\text{ mM}$ . The reaction was incubated at  $T_{react} = 25\text{ }^{\circ}\text{C}$ . Throughout the reaction time, samples were quenched by addition of one volume equivalent of 4 M urea in 2 % TFA, which were subsequently analyzed by RP-HPLC (Gemini-NX, 5 $\mu\text{m}$  C18 250 x 4.6 mm column, gradient see Tab. 6-4,  $RT_{substrate} = 10.8\text{ min}$ ,  $RT_{product} = 13.1\text{ min}$ ). Substrate concentration was determined by comparison to a calibration curve established with free 2-amino benzoic acid. Product concentration was determined by comparison to a calibration curve established with samples of fixed product concentration (known substrate concentration after complete turnover).

**Tab. 6-4:** HPLC gradient used for the separation of the substrate peptide and the product peptide of the FGE catalyzed reaction. Buffer A:  $\text{H}_2\text{O}$ , 1 % MeCN, 0.1 % TFA. Buffer B: MeCN, 0.085 % TFA. Flow rate 1 mL/min.

Time [min]	Buffer B [%]
0	15
0.5	15
12.0	20
17.0	95
20.7	95
22.0	15
25.0	Stop

**Analysis of structure after single turnover.** To quantitatively produce FGE in the inactive state FGE was diluted to a final concentration of  $c_{FGE} = 90\text{ }\mu\text{M}$  in a reaction buffer containing 140  $\mu\text{M}$  substrate peptide, 50 mM EDTA, 50 mM NaCl and 50 mM Tris-HCl buffer pH 8. After incubation overnight the samples were analyzed by SEC and HRMS.



---

**Size exclusion chromatography.** Size exclusion chromatography was performed on an Äkta Prime with a HiLoad Superdex 26/600 column at a flow rate of 1.5 – 2 mL/min in a 50 mM Tris buffer pH 8 supplemented with 200 mM NaCl.

**Analysis of reaction rates in the presence of different reducing agents.** Reaction rates were determined using a discontinuous HPLC assay as described previously (see chapter 4.3). Shortly, FGE was diluted to a final concentration of  $c_{\text{FGE}} = 2 \mu\text{M}$  in a reaction buffer containing 200  $\mu\text{M}$  substrate peptide, 2 mM reducing agent, 2  $\mu\text{M}$   $\text{CuSO}_4$ , 50 mM EDTA, 50 mM NaCl and 50 mM Tris-HCl buffer pH 8. Throughout the reaction time, samples were quenched by addition of one volume equivalent of 4 M urea in 2 % TFA, which were subsequently analyzed by RP-HPLC.

**Peptide synthesis.** Peptide substrate were synthesized by Fmoc solid phase peptide synthesis following standard protocols using a peptide synthesizer. Shortly, the most C-terminal amino acid was coupled to Rink amide resin. Deprotection of the primary amine is then accomplished by incubation with 40 % piperidine in DMF. The subsequent amino acid is added and activated by HCTU in DMF upon which coupling occurs. Unreacted peptide is quenched by addition of acetic anhydride in DMF. The peptide is sequentially synthesized with the last step being coupling of 2-amino benzoic acid. Cleavage from the resin and deprotection of the amino acid side chains is simultaneously achieved by incubation with the cleavage cocktail (TFA:TIPS:DCM 96.5 : 2.5 : 1) for 3 hrs. Impurities have been removed by RP-HPLC after which lyophilization yields the final peptide substrate. The identity of the peptide was confirmed by ESI-MS.

**Tab. 6-5:** Sequences and molecular weights of reducing peptides described in this chapter.

ID	Sequence	MW <sub>calc</sub> [Da]	[M+H] <sup>+</sup> <sub>obs</sub> [Da]
PepN	Abz-AGSCGCGTPQRPGAHGSS-NH <sub>2</sub>	1746.7	1747.0
PepN*	Abz-AGSCGCGTRQPHGAPGSS-NH <sub>2</sub>	1746.7	1747.0

**Analysis of reaction rates in the presence of PepN and PepN\*.** Reaction rates were determined using a discontinuous HPLC assay as described previously (see chapter 4.3). Shortly, FGE was diluted to a final concentration of  $c_{\text{FGE}} = 2 \mu\text{M}$  in a reaction buffer containing 200  $\mu\text{M}$  substrate peptide, 0 - 500  $\mu\text{M}$  of either PepN or PepN\*, 2 mM GSH, 2  $\mu\text{M}$   $\text{CuSO}_4$ , 50 mM EDTA, 50 mM NaCl and 50 mM Tris-HCl buffer pH 8. Throughout the reaction time, samples were quenched by addition of one volume equivalent of 4 M urea in 2 % TFA, which were subsequently analyzed by RP-HPLC ( $\text{RT}_{\text{substrate}} = 18.4 \text{ min}$ ,  $\text{RT}_{\text{product}} = 19.8 \text{ min}$ ,  $\text{RT}_{\text{PepN}} = 13.1 \text{ min}$ ,  $\text{RT}_{\text{PepN}^*} = 13.3 \text{ min}$ ).

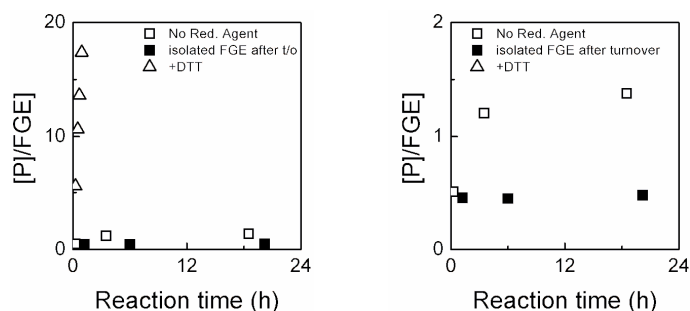
---

**Tab. 6-6:** HPLC gradient used for the separation of the substrate peptide and the product peptide of the FGE catalyzed reaction in the presence of PepN or PepN\*. Buffer A: H<sub>2</sub>O, 1 % MeCN, 0.1 % TFA. Buffer B: MeCN, 0.085 % TFA. Flow rate 1 mL/min.

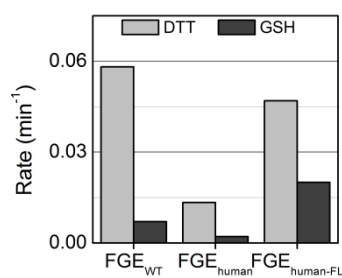
Time [min]	Buffer B [%]
0	5
0.5	5
20.0	25
22.0	95
27.0	95
28.0	5
31.0	Stop

**Kinetic isotope effects.** Reaction rates were determined using a discontinuous HPLC assay as described previously (see chapter 4.3). Shortly, FGE was diluted to a final concentration of  $c_{\text{FGE}} = 2 \mu\text{M}$  in a reaction buffer containing 200  $\mu\text{M}$  substrate peptide, 0 - 400  $\mu\text{M}$  of either PepN or PepN\*, 2 mM GSH, 2  $\mu\text{M}$  CuSO<sub>4</sub>, 50 mM EDTA, 50 mM NaCl and 50 mM Tris-HCl buffer pH 8. Throughout the reaction time, samples were quenched by addition of one volume equivalent of 4 M urea in 2 % TFA, which were subsequently analyzed by RP-HPLC. The deuterated substrate was kindly provided by Pacal Engi and synthesized as described before (see chapter 4.3).

## 6.4 Appendix



**Fig. 6-9:** Product formation in absence of reducing agent (open square) and after isolation of the inactivated FGE without additional reducing agent (solid square) and with added DTT (open triangle). The product formation is given as equivalents compared to the enzyme concentration. Left: Full scale and Right: zoom.



**Fig. 6-10:** Rates of product formation of full length FGE from *H. sapiens* (FGE<sub>human-FL</sub>) in comparison to the truncated form (FGE<sub>human</sub>) and FGE from *T. curvata* (FGE<sub>WT</sub>). Rates were determined using DTT or GSH as a reducing agent. Low activity of FGE<sub>human</sub> is most likely an expression of the prolonged storage time of this very sample at T<sub>store</sub> = 4 °C which led to decomposition and thereby lost activity.

---

---

## 7. Optimization of FGE driven aldehyde formation *in vitro*

A large number of investigations were carried out aiming to utilize the unique aldehyde residue for biotechnological applications.<sup>[10][81][120]</sup> The process mainly relies on the formation of an aldehyde functional group on the protein of interest and the following labeling reaction by chemical means. The CxPxR recognition motif can be introduced at either the C- or N-terminus or even within solvent accessible, flexible loop regions of the target protein.<sup>[83][80]</sup> Current protocols however display one major disadvantage, which is the low reactivity of the applied FGEs. Some setups require elevated temperatures that might conflict with the stability of the target protein, i.e. 30 °C.<sup>[84]</sup> Other reactions were performed at lower temperatures (i.e. 4 °C), but consequently demand longer incubation times of up to 48 hrs.<sup>[85]</sup> When copper was identified as an electron donor in the FGE catalyzed reaction, culture media were supplemented with the metal salts. The expressed FGE was thereby efficiently supplied with copper ions which was shown to successfully enhance the aldehyde formation on antibodies *in vivo*.<sup>[97]</sup> *In vitro* reactions were similarly enhanced and reactions were then performed at 18 °C for 16 hrs with a 10 mol% catalyst loading.<sup>[94]</sup> We likewise developed protocols that enabled us to reach full conversion on a protein target within less than one hour with only 1 % of catalyst loading.<sup>[64]</sup>

The utility of FGE for biochemical and biotechnological research has expanded to a wide array of applications, such as labeling reactions, the generation of protein-DNA conjugates and protein (or even virus) surface immobilization.<sup>[121][122][123]</sup> For a successful implementation as a biotechnological tool, however, one requires high turnover rates and high catalytic efficiencies. The threshold for such applications appears to be  $k_{cat}/K_M > 10^3 \text{ M}^{-1}\text{s}^{-1}$  (see also Fig. 7-1). While others have reported efficiencies of  $k_{cat}/K_{M, \text{human}} \sim 10^4 \text{ M}^{-1}\text{s}^{-1}$ <sup>[94]</sup> for the human FGE variant, FGE<sub>WT</sub> from *T. curvata* fails to meet these requirements. We envision that the use of a bacterial protein that can simply be expressed in *E. coli* cells and shows resistance to high temperatures could be advantageous with regards to applicability in the laboratory. The incorporation of copper into the reaction mixture increased the activity of FGE<sub>WT</sub> to some extent (see also section 3.1). In pursuit of the mechanistic details of the FGE catalyzed reaction we generated an FGE variant with only two cysteine residues being present in the active site of FGE from *T. curvata* (FGE<sub>4C</sub>, see also section 3.1). The resulting catalyst happens to be almost 7-fold more active than the parent protein ( $k_{cat}/K_{M, \text{WT}} = 50 \text{ M}^{-1}\text{s}^{-1}$ ;  $k_{cat}/K_{M, 4C} = 330 \text{ M}^{-1}\text{s}^{-1}$ ). We now show the identification of the residue responsible for the increased activity in FGE<sub>4C</sub>. Furthermore, we insert a Tyr->Phe mutation in close vicinity to the active site cysteines that additionally facilitates the reaction by

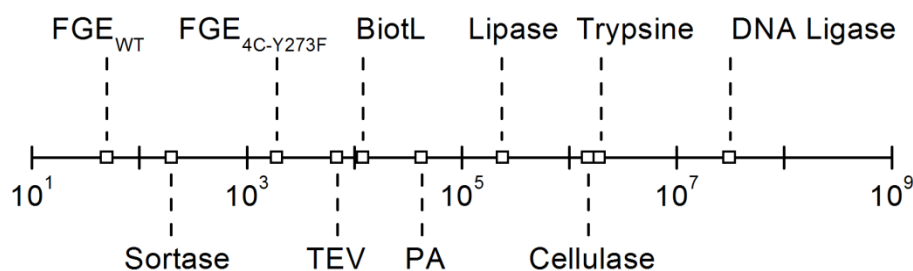


Fig. 7-1: Catalytic efficiencies in  $\text{M}^{-1}\text{s}^{-1}$  of different enzymes used in biotechnological applications (see also Tab. 7-1).

6 fold. The rate enhancing effects found in FGE<sub>4C</sub> and FGE<sub>Y273F</sub> are then combined in a highly active variant with a catalytic efficiency of  $k_{cat}/K_M,4C-Y273F = 1.9 \cdot 10^3 \text{ M}^{-1}\text{s}^{-1}$ . With this we envision that this easy-to-obtain, thermally stable and fast FGE variant will be of great advantage when aiming to introduce aldehyde functionalities on protein targets.

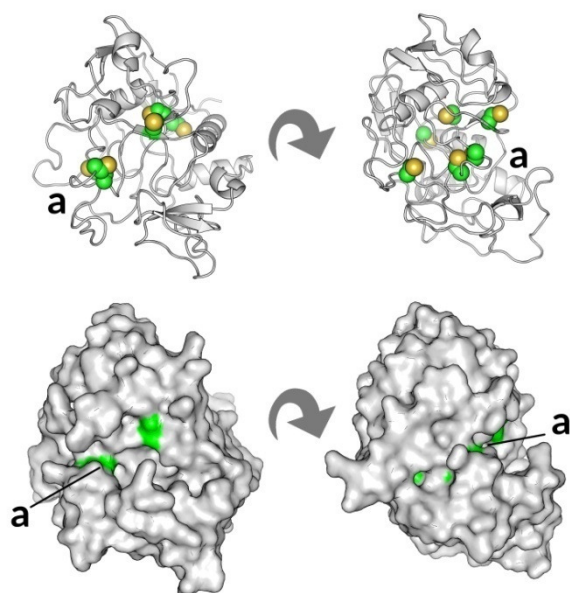
**Tab. 7-1:** Catalytic efficiencies of selected enzymes used in biotechnological applications in comparison to FGE<sub>WT</sub>.

Protein	Use	$k_{cat}/K_M [\text{M}^{-1}\text{s}^{-1}]$
TEV protease	Protein hydrolysis	$6.9 \cdot 10^3$ [124]
T4 DNA ligase	DNA ligation	$(3.1 - 4.7) \cdot 10^7$ [125]
Trypsine	Protein hydrolysis	$1.9 \cdot 10^6$ [126]
Penicillin amidase	Antibiotic semi-synthesis	$4.2 \cdot 10^4$ [127]
Lipase	Industrial Synthesis	$2.4 \cdot 10^5$ [128]
Cellulase	Paper industry	$1.5 \cdot 10^6$ [129]
Sortase	Bioconjugation	200 [130]
Biotin ligase	Biotinylation	$1.2 \cdot 10^4$ [131]
FGE <sub>WT</sub>	Bioconjugation	$50$ [64] - $3.0 \cdot 10^4$ [94]
FGE <sub>4C-Y273F</sub>		$1.9 \cdot 10^3$

### 7.1 Identification of *in vitro* gain of function cysteine mutation

Gain of function mutations are found rarely during *in vitro* studies. However, the observation that the activity of FGE<sub>4C</sub> is enhanced demonstrates a key difference between *in vivo* and *in vitro* applications. While nature has evolved the amino acid sequence of each enzyme over millions of years, outside of the cellular context the function of a residue might be significantly different. After our initial observation of the rate acceleration in FGE<sub>4C</sub> we then sought out to define the cysteine residue that is responsible for the increased activity.

Wild type FGE contains a total of six cysteine residues, two of which are located in the active site and are required for turnover to occur.<sup>[48]</sup> The remaining four cysteines are found in the protein scaffold (see Fig. 7-2). It is noteworthy that, these cysteines scatter throughout the protein structure and are positioned in distances that do not allow for disulfide bond formation ( $d_{S-S} > 8 \text{ \AA}$ , estimation based on a homology model of FGE<sub>WT</sub> based on a published crystal structure from human FGE, PDB 1Y1E). The residue with the highest degree of conservation is Cys298, which can be found in 83 % of the 500 closest homologs of FGE<sub>WT</sub> (after BLAST search). Two additional cysteine residues show high degrees of conservation with 68 % and 67 % for Cys187 and Cys231, respectively. The residue with the lowest degree of conservation is Cys284, which can be found in only 0.6 % of the analyzed sequences. To address the question of which

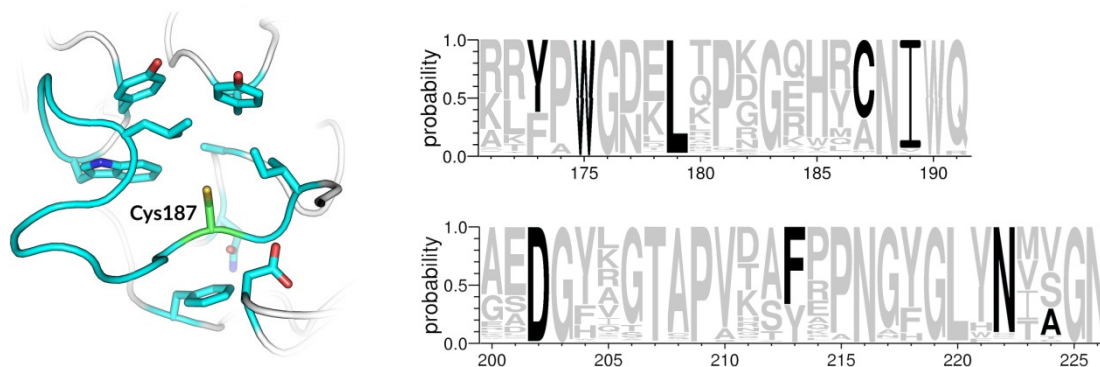


**Fig. 7-2:** Homology model of FGE<sub>WT</sub> based on a published crystal structure of human FGE (PDB 1Y1E). Cysteine residues are highlighted in green, with the active site cysteines assigned with **a**. Shown is the cartoon representation (top) and the surface representation (bottom) to indicate solvent exposure.

mutation causes the increased activity, we designed four variants of the activated FGE<sub>4C</sub> variant. In each of these mutants one cysteine residue was mutated back to its original position (see Tab. 7-2), yielding four proteins with three cysteine residues each. Most of the new variants showed reaction rates comparable to that of FGE<sub>4C</sub>. Only FGE<sub>CASA</sub>, with Cys187 back in place, displayed a reduced activity to the level of FGE<sub>WT</sub> indicating that this residue causes the change in activity. The alanine residue replacing Cys187 is found at this position in 27 % of the close homologues of FGE<sub>WT</sub>, indicating that replacement of cysteine by alanine has similarly occurred in many organisms during evolution and should therefore be a tolerated mutation. We consequently addressed this question with an additional control by generating a mutant with only cysteine at position 187 being replaced to alanine. In agreement with the previous observation this mutant shows elevated rates of product formation similar to FGE<sub>4C</sub>. Cysteine 187 is located in a hydrophobic pocket in the close vicinity to the active site (see Fig. 7-3). To assay the solvent

**Tab. 7-2:** FGE variants with different cysteine patterns to identify the cysteine residue that causes the increased activity of FGE from *T. curvata*. Shown are the residues at the various positions as well as the corresponding reactivity and midpoints of thermal transition.

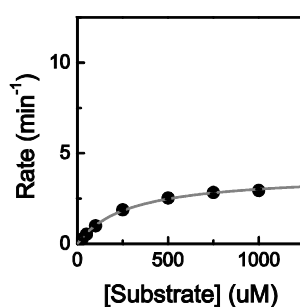
Protein	Position				$k_{\text{obs}}$ [min <sup>-1</sup> ]	$T_m$ [°C]
	187	231	284	298		
FGE <sub>WT</sub>	C	C	C	C	0.3	68
FGE <sub>4C</sub>	A	A	S	A	0.7	63
FGE <sub>CASA</sub>	C	A	S	A	0.3	65
FGE <sub>ACSA</sub>	A	C	S	A	1.0	62
FGE <sub>AACA</sub>	A	A	C	A	1.6	57
FGE <sub>AASC</sub>	A	A	S	C	1.0	61
FGE <sub>ACCC</sub>	A	C	C	C	0.9	63



**Fig. 7-3:** **Left:** Cysteine 187 (green) with the surrounding loop region (cyan) and conserved hydrophobic residues shown as sticks (residues within 5 Å distance of the cysteine sulfur atom). **Right:** Sequence logo showing the conservation of the hydrophobic residues (black). The degree of conservation for each residue was determined to be: Y173: 63 %; W175: 100 %; L179: 98 %; I189: 94 %; D202: 87 %; F213: 67 %; N222: 91 %; A224: 24 %.

accessibility of this residue we incubated FGE with 2-iodoacetamide (IAA). While other cysteines can efficiently be alkylated using IAA, Cys187 does not form covalent adducts (see Appendix Fig. 7-9), indicating that the surrounding loop structure is stable and forms a tight pocket preventing nucleophilic attack of IAA. Many residues in this hydrophobic area surrounding Cys187 are highly conserved and point towards the cysteine residue (60 - 100 % in in the 500 closest homologues after BLAST search, except for Ala224) indicating the importance of the hydrophobicity of this pocket (see Fig. 7-3).

Introduction of mutations into a protein sequence can perturb the structure or fold of a protein. We therefore performed circular dichroism spectroscopy to analyze changes in secondary structure and compare the mutants resistance to increasing temperatures (see Tab. 7-2 and Appendix Fig. 7-10). The effect of the presence of copper on the thermal resistance has been discussed before (see section 4.1.3). Due to the increased denaturing temperature we were not able to precisely determine the midpoint of transition since the second state is not reached until the maximal temperature of  $T_{\max} = 90\text{ }^{\circ}\text{C}$ . We therefore only evaluated the temperature dependend CD signal in absence of copper(I). The secondary structure of the different variants does not differ significantly. With regards to thermal stress, however, Cys187 appears to have a slight effect. While mutants containing a cysteine residue at position 187 unfold at higher temperature ( $T_{\text{melt}} \geq 65\text{ }^{\circ}\text{C}$ ), enzymes with an alanine residue at this position show lower tolerance for thermal stress ( $T_{\text{melt}} \leq 63\text{ }^{\circ}\text{C}$ ). The difference might be minor, but the observation shows, that perturbation of the carefully crafted loop environment decreases the protein stability.



**Fig. 7-4:** Reaction rate of  $\text{FGE}_{\text{acc}}$  at increasing substrate concentration.  $v_{\max} = 3.9\text{ min}^{-1}$ ;  $K_M = 280\text{ }\mu\text{M}$ .

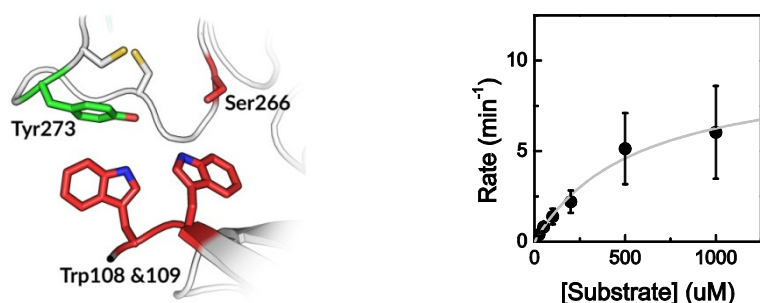


The effect of the cysteine to alanine mutation in FGE<sub>ACC</sub> is mostly based on an increased  $k_{cat}$  value of  $k_{cat,ACC} = 3.9 \text{ min}^{-1}$  in comparison to FGE<sub>WT</sub> ( $k_{cat,WT} = 1.6 \text{ min}^{-1}$ ; see Fig. 7-4 and chapter 4). The  $K_M$ -value is similarly decreased as for FGE<sub>4C</sub> with half the  $K_M$  value of FGE<sub>WT</sub> ( $K_{M,WT} = 580 \text{ }\mu\text{M}$ ;  $K_{M,4C} = 200 \text{ }\mu\text{M}$ ;  $K_{M,ACC} = 280 \text{ }\mu\text{M}$ ). In summary these effects add up to a catalytic efficiency comparable to that of FGE<sub>4C</sub>, which is six to seven fold higher than that of the parent protein ( $k_{cat}/K_{M,WT} = 50 \text{ M}^{-1}\text{s}^{-1}$ ;  $k_{cat}/K_{M,4C} = 330 \text{ M}^{-1}\text{s}^{-1}$ ;  $k_{cat}/K_{M,ACC} = 230 \text{ M}^{-1}\text{s}^{-1}$ ). In summary the variant with only Cys187 mutated to alanine (FGE<sub>ACC</sub>) shares the kinetic properties of FGE<sub>4C</sub> as well as the reduced thermal stability, suggesting that this mutation is indeed responsible for the enhanced *in vitro* reactivity of FGE<sub>4C</sub>.

Conclusive rationale, however, still lacks. While Cys187 is found in 92 % of the 500 closest prokaryotic homologues of FGE, it is not found in a single one of the 500 closest eukaryotic FGE-like proteins. This shows that Cys187 is exclusive to bacteria and archaea, suggesting that it was introduced after these classes of organisms have evolutionary diverged.

## 7.2 Mutation of active site Tyr273

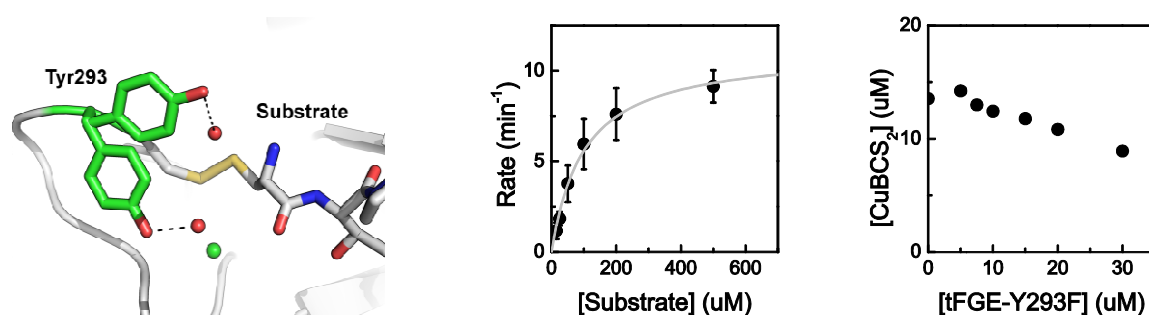
In pursuit of a more active catalyst for possible *in vitro* applications, we additionally mutated other residues in close vicinity to the active site cysteine residues (see Fig. 7-5). Serine at position 266, as well as the two more distant tryptophane side chains at position 108 and 109 were shown to be essential for turnover. Mutation of these residues led to lower catalytic efficiencies ( $k_{cat}/K_{M,W108A,W109A} = 0.03 \text{ M}^{-1}\text{s}^{-1}$ ;  $k_{cat}/K_{M,S266A} = 0.82 \text{ M}^{-1}\text{s}^{-1}$  vs.  $k_{cat}/K_{M,WT} = 50 \text{ M}^{-1}\text{s}^{-1}$ , see also section 5.1). The two tryptophane residues are located in a distance of  $d_{\text{Cys-SH}:\text{Trp-NH}} \geq 10 \text{ \AA}$  to the active site cysteines, whereas Ser266 is more closely positioned ( $d_{\text{Cys-SH}:\text{Ser-OH}} = 6.2 \text{ \AA}$ ). The entry towards the active site cysteines is mainly dominated by the presence of the two tryptophane residues in Fig. 7-5 as well as a tyrosine functional group even closer to the cysteines (Tyr273,  $d_{\text{Cys-SH}:\text{Tyr-OH}} = 4.5 \text{ \AA}$ ). This specific tyrosine is present in 99 % of the 500 closest homologues of FGE, additionally suggesting a role in catalysis. A side chain in such close vicinity most likely has an effect on the reactivity of the nearby cysteine. The role of this tyrosine residue (Tyr340 in human FGE) has been investigated before. The findings, however, are contrasting. Earlier publications describe the mutation Tyr340Phe to not affect the activity significantly, whereas recent investigations characterize this tyrosine residue as having an effect on FGEs activity.<sup>[48][94]</sup> However, further details are not given. We therefore investigated the role of this residue.



**Fig. 7-5: Left:** Active site of FGE<sub>WT</sub> (homology model based on a published crystal structure of FGE<sub>human</sub>) with Ser266, Trp108 and Trp109 (red) as well as Tyr273 (green). Mutation of the serine or tryptophane residues yields a less active variant, whereas mutation of Tyr273 to Phe provides a more active catalyst during *in vitro* experiments. **Right:** Michaelis-Menten plot of FGE<sub>Y273F</sub>, showing the rates of product formation at increasing substrate concentration.

Surprisingly, during our *in vitro* studies, the mutant FGE<sub>Y273F</sub> shows reactivity above that of FGE<sub>WT</sub> ( $k_{\text{cat}}/K_{\text{M,Y273F}} = 340 \text{ M}^{-1}\text{s}^{-1}$ ). Substrate binding appears not to be influenced by the presence of Tyr273, since the  $K_{\text{M}}$  does not significantly differ from that of FGE<sub>WT</sub> ( $K_{\text{M,WT}} = 580 \pm 40 \mu\text{M}$  vs.  $K_{\text{M,Y273F}} = 560 \mu\text{M}$ ). In an X-ray structure of FGE<sub>human</sub> there is electron density for Tyr340 in two distinct orientations upon substrate binding (see Fig. 7-6, left), whereas the structure without substrate shows a single orientation only.<sup>[56]</sup> While this might be an artefact of the crystalization prozess, during our *in vitro* studies the phenol group seems to be insignificant for substrate binding. To test whether the displacement of Tyr273 changes the rate limiting step, we investigated the rate of product formation in the presence of the previously described deuterated substrate (see also section 4.1). In good agreement with the observations with FGE<sub>WT</sub> and FGE<sub>4C</sub> the removal of the  $\beta$ -proton of the substrate cysteine by FGE<sub>Y273F</sub> is rate limiting ( $\text{KIE}_{\text{Y273F}} = 2.8 \pm 0.1$ , Appendix Fig. 7-12). We therefore conclude that the overall mechanism remains the same. Similarly the temperature of thermal unfolding is comparable to that of the wild type enzyme ( $T_{\text{unfold,Y273F}} = 68 \text{ }^\circ\text{C}$ , see Appendix Fig. 7-10).

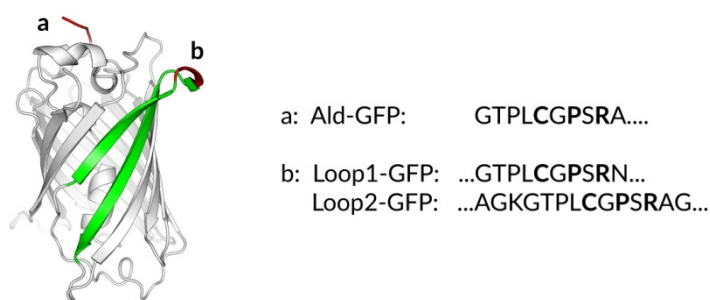
We then went on to test, whether this mutation and the previously described cysteine mutations of FGE<sub>4C</sub> would interfere with each other. Of particular interest is whether the rate enhancements of each mutant would be additive in an FGE<sub>4C-Y273F</sub> variant containing both, the Tyr273Phe mutation and the four cysteine mutations. We therefore created this very FGE variant. The newly produced enzyme shows a catalytic efficiency which lies more than 5-fold above that of FGE<sub>Y273F</sub> and more than 5-fold over that of FGE<sub>4C</sub> ( $k_{\text{cat}}/K_{\text{M,4C-Y273F}} = 1900 \text{ M}^{-1}\text{s}^{-1}$ ). We therefore conclude that the rate enhancing effects of the mutations in FGE<sub>4C</sub> and FGE<sub>Y273F</sub> are added up in FGE<sub>4C-Y273F</sub>. Overall, the resulting variant shows a 35-fold higher efficiency than the wild type enzyme. The  $K_{\text{M}}$  value, in particular, is similarly decreased as observed for FGE<sub>4C</sub> ( $K_{\text{M,4C}} = 230 \pm 40 \mu\text{M}$ ,  $K_{\text{M,4C-Y273F}} = 100 \pm 30 \mu\text{M}$ ), which again is indication that Tyr273 is not involved in substrate binding. Tyrosine residues perform key interactions in some copper(II) containing proteins such as galactose oxidase or amine oxidases (see also section 2.2). To ensure that copper binding is not perturbed in FGE mutants lacking Tyr273, we determined the binding affinity of FGE<sub>Y273F</sub> using the previously described method for copper affinity determinations (see also chapter 5). The  $K_{\text{D}}$  of FGE<sub>Y273F</sub> was determined to be  $K_{\text{D,Y273F}} = 10^{-17.0} \text{ M}$  (see Fig. 7-6, right) which is in good agreement with the binding affinities of FGE<sub>WT</sub> and other FGE variants ( $K_{\text{D,WT}} = 10^{-17.1}$ ), hence we conclude that copper binding is not perturbed in the active site of FGE<sub>Y273F</sub>.



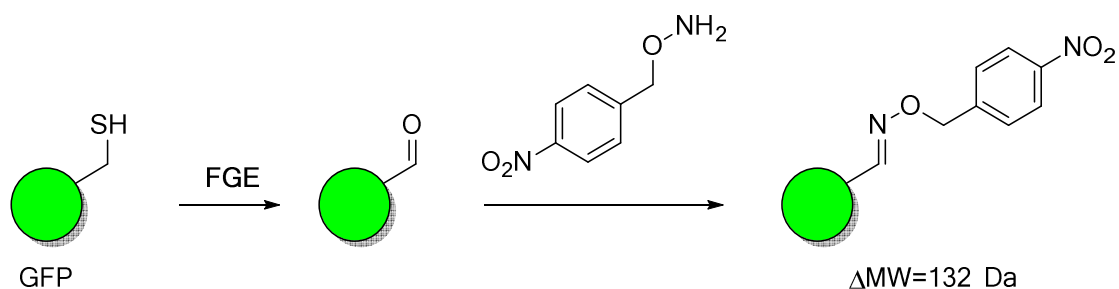
**Fig. 7-6: Left:** Crystal structure of human FGE with bound substrate peptide showing different orientations of the Tyr340 (PDB 2AIJ).<sup>[56]</sup> **Middle:** Michaelis-Menten plot of FGE<sub>4C-Y273F</sub>. The rate of product formation is assayed at increasing substrate concentrations. The more active variant shows an increased  $k_{\text{cat}}$  value as well as a decreased  $K_{\text{M}}$  value, in comparison to FGE<sub>WT</sub>. **Right:** Copper binding of FGE<sub>Y273F</sub> in competition with bathocuproinedisulfonic acid (BCS).

Copper(II) containing LPMOs such as glucoside hydrolases (GHs) likewise display a tyrosine residue in the coordination sphere of the metal ion. A very similar enzyme class, the “33 family carbohydrate binding models” (CBM33), catalyze copper dependent cleavage of chitin chains. While crystal structures of these enzymes show similar metal binding sites, the conserved and metal coordinating tyrosine residue in GHs is displaced by a phenylalanine in CBMs. The metal binding capability and hence reactivity however remains.<sup>[133]</sup> Although, the metal binding affinity in GHs are higher than in CBMs, in FGE, the mutation of tyrosine to phenylalanine does not affect the metal binding affinity.

Small peptide substrates provides an ideal tool for mechanistic investigations and first explorational experiments for method development. However, different properties might influence the interactions with the target protein *in vivo*. Such factors are the concentration of available oxygen *in vivo* or the participation of a different reducing agent (see also chapter 6). Additionally, since FGE catalysis is believed to occur co- or post-translational with sulfatases, the substrates *in vivo* are much bigger proteins. A prolonged peptide sequence might interact differently with FGE than a short peptide. Larger substrates might interfere with FGE by steric hindrance or bending of the CxPxR recognition site, which could reduce FGEs capability to bind the substrate. We addressed these concerns by investigating the activity of FGE<sub>4C-Y273F</sub> with different protein substrates. The selected GFP substrates have the recognition motif installed at different locations (see Fig. 7-7). The FGE recognition site (ald-tag) was positioned either at the N-terminus or in a loop region that has been described to tolerate modifications without loss of the characteristic fluorescence.<sup>[80][132]</sup> GFP with ald-tags in the loop was used with different lengths of sequences, since we anticipated that the loop formation adds constraints to the substrate. A longer sequence might lower this substrate bending and potentially allow for better recognition by FGE. We therefore incubated these different GFP samples with either FGE<sub>WT</sub> or FGE<sub>4C-Y273F</sub> and analyzed the substrate to product distribution by means of HRMS (see Fig. 7-8). To increase the mass difference between the thiol substrate and the aldehyde containing product we incubated the fGly containing product with O-(4-nitrobenzyl)hydroxylamine. While the FGE reaction is performed at pH 8, labeling is achieved at pH 4.5. This lower pH will simultaneously quench the FGE reaction (see chapter 4). Interestingly, loop1-GFP with the short insert is not accepted as a substrate and no product is observed after 90 mins of incubation with either FGE<sub>WT</sub> or FGE<sub>4C-Y273F</sub>. In contrast, the N-terminal tag as well as the longer sequence in the loop region (loop2-GFP) are readily converted. In agreement with the previous observations with the peptide

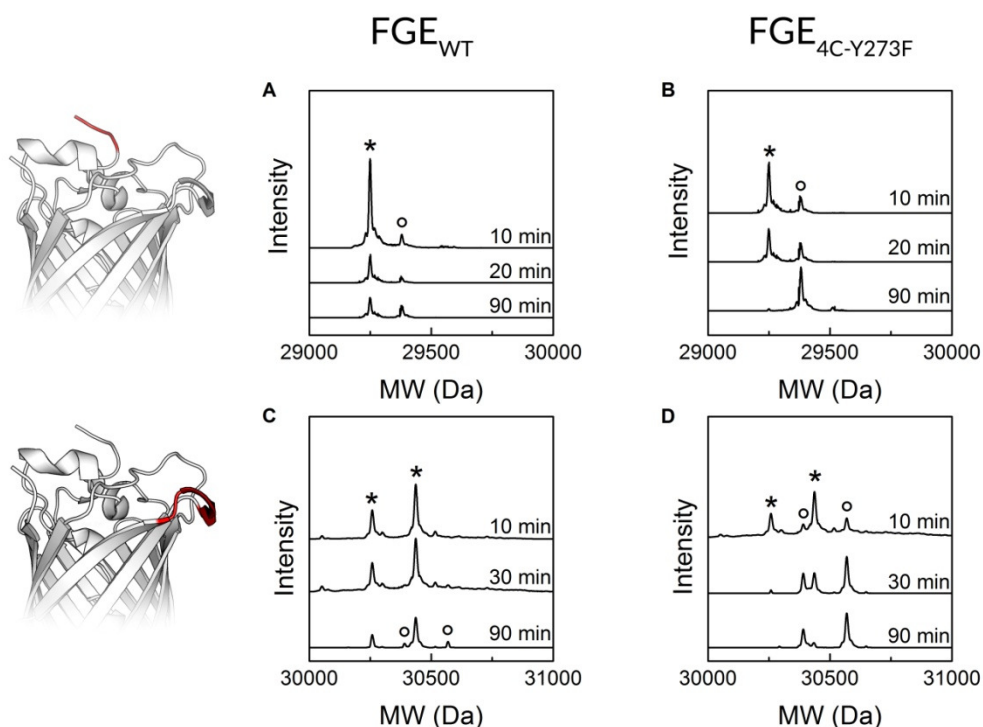


**Fig. 7-7:** Structure of GFP (PDB 1GFL). Sites where FGE recognition sequences have been introduced are highlighted in red. Beta sheets surrounding the site of the internal loop are highlighted in green. Sequences that were introduced are given with their respective location (a or b).



**Scheme 7-1:** Labeling of GFP with hydroxylamine after FGE catalyzed oxidation.

substrate, FGE<sub>4C-Y273F</sub> performs faster catalysis on both substrate peptides. Catalysis of FGE<sub>WT</sub> on the N-terminal tag is incomplete after 90 min. The percentage of product is 38 %. The internal cysteine in the loop-tag seems to be less accessible since only 24 % of product can be identified after 90 min of incubation. The more constrained structure of the recognition motif in the loop region seems to conflict with efficient catalysis. FGE<sub>4C-Y273F</sub> however, performs efficient catalysis on both substrates. In crystal structures of FGE<sub>human</sub> with co-crystallized substrate the corresponding Tyr340 adopts two distinct orientations. A second structure, co-crystallized with a longer substrate, shows the tyrosine only in the orientation facing towards the substrate.<sup>[56]</sup> Removal of the hydroxyl group from this tyrosine residue might therefore favour the orientation with the phenylalanine facing away from the substrate. The phenylalanine would then open up



**Fig. 7-8:** Conversion of ald-tag containing GFP with the recognition motif installed at either the N-terminus (top) or in a loop (bottom) with FGE<sub>WT</sub> (A & C) and FGE<sub>4C-Y273F</sub> (B & D). The substrate is indicated by the asterisk whereas the product peak is highlighted by the open circle.  $MW_{calc,ald-GFP} = 29253.1$  Da,  $MW_{obs} = 29249.6$  Da.  $MW_{calc,loop2-GFP} = 30280$  Da,  $MW_{obs} = 30258.5$  Da.  $\Delta MW_{modif,calc} = 132.1$  Da,  $\Delta MW_{modif,obs} = 130.5 - 132.1$  Da. The additional product peak in loop2-GFP samples represents glycosylation on the His<sub>6</sub>-tag ( $\Delta MW_{glyco} = 178$  Da).

---

the substrate channel allowing for the substrate to enter the active site even in a constrained sequence. Not only is the rate of N-terminal oxidation with FGE<sub>4C-Y273F</sub> significantly faster but also is the bend substrate oxidized similarly fast. In both cases the reaction mixture is absent of substrate after 90 min and contains only the labeled product.

We therefore conclude, that the most active FGE variant described here shows an elevated reaction rate as well as a higher tolerance to substrate constraints. While the wild type protein does catalyze oxidation of the protein substrates, catalysis is significantly more efficient when using FGE<sub>4C-Y273F</sub>. We do envision that this variant will be valuable especially when modifications are to be introduced in sterically demanding protein regions.

---

### 7.3 Conclusion

This chapter shows the development of a set of FGE variants with increased activity. We identified the mutation causing the increased activity in FGE<sub>4C</sub> to be Cys187Ala. A variant with only Cys187 replaced by alanine shows comparable reaction rates as FGE<sub>4C</sub> and shows similar resistance to increased temperatures.

An additional mutation in very close proximity to the active site ( $d = 4 - 5 \text{ \AA}$ ) has beneficial effects on the reaction rate. The mutant FGE<sub>Y273F</sub> is seven fold more active than the parent protein but shows similar copper binding affinities and is similarly stable as the wild type. Importantly the effects of the individual mutations are additive in a mutant containing the activating cysteine mutations as well as the Tyr273Phe mutation. The resulting mutant shows a catalytic efficiency 38-fold higher than the wild type protein. With a catalytic efficiency of  $k_{\text{cat}}/K_M = 1900 \text{ M}^{-1}\text{s}^{-1}$  FGE<sub>4C-Y273F</sub> suffices to serve as a tool for biotechnological application. The applicability of this highly active variant was shown by reacting FGE<sub>4C-Y273F</sub> with different authentic target proteins. GFP with the FGE recognition motif installed at either the N-terminus or in a solvent accessible loop was effectively converted within 90 min at a target to catalyst ratio of 380 : 1.

With the high degree of resistance to heat, we therefore believe, that this mutant will prove to be very valuable for the incorporation of aldehyde functionalities into recombinant proteins.

## 7.4 Experimental

**Plasmid Construction.** Mutations were introduced by site directed mutagenesis using the vector pET28-FGE<sub>WT</sub> or pET28-FGE<sub>4C</sub> as a template. Synthetic primers were ordered from Microsynth and are shown in Tab. 7-3. After PCR, the fragments as well as the pET28 target vector were treated with NdeI and XhoI restriction enzymes, followed by ligation of the gene into the target vector by T4 DNA ligase. The identity of final plasmids was confirmed by sequencing analysis (Microsynth).

Plasmids and protein samples of ald-GFP, Loop1-GFP and Loop2-GFP were kindly supplied by Dr. Roxana Lemnar, University of Basel.

**Tab. 7-3:** Primer sequences used for the mutation of the FGE variants described in this chapter.

FGE variant	Primer 3' → 5'	
	Sense	Antisense
FGE <sub>WT</sub>	ATATCATATGCCCTCGTTGCACTTCG	ATATCTCGAGCTACAGCGGGTC
FGE <sub>4C</sub>	ATATCATATGCCGTCGTTTGACTTTG	ATATCTCGAGTTACAGCGGATC
FGE <sub>CASA</sub>	GCCATCGTTGCAATATTTGGCAAGGTA	GCCAAATATTGCAACGATGGCGGCCAC
FGE <sub>ACSA</sub>	GGGAATGGTGCGCCGATTGGTGGTCAG	AATCGGCGCACCATTCCCACACATTGC
FGE <sub>AACA</sub>	TCGTACGTGCAATACGCCGACTCAAG	GCGTATTGCACGTACGAGCTGCAACAC
FGE <sub>AASC</sub>	ATATCTCGAGTTACAGCGGATCGGCGGCGCAGC GAAAACCCGT	-
FGE <sub>ACCC</sub>	ATATGGCACCGCGCAAACATCTGGCAG	ATATAGATGTTTGCGCGGTGCCGGCC
FGE <sub>Y273F</sub>	ACGAGTCGTTCTGCAACCG	CGGTTGCAGAACGACTCGT
FGE <sub>4C-Y273F</sub>	CACGAATCCTTTTGTAAATCGCT	AGCGATTACAAAAGGATTCTGTG

**Protein Expression and Purification.** FGE mutants were produced as described previously (see chapter 4.3). Genes were transferred into BL-21 pLys *E. coli* cells. Precultures were inoculated with a single colony and incubated at  $T_{inc} = 37\text{ °C}$  overnight. Expression cultures of 1 L LB medium supplemented with kanamycine and chloramphenicol (50 mg/L and 34 mg/L, respectively) were inoculated with 1 mL of preculture. At OD = 0.6 – 0.8 cultures were induced by addition of IPTG. After additional 3 hrs of incubation at  $T_{express} = 20 - 37\text{ °C}$  the cells were harvested by centrifugation. The protein was then isolated by Ni<sup>2+</sup>- affinity chromatography and dialysed into Tris buffer pH 8 before storage at  $T_{store} = -80\text{ °C}$ .

**Tab. 7-4:** Calculated and observed molecular weights of proteins described in chapter 6 with comments on the mass difference.

Protein	MW <sub>calc</sub> [Da]	MW <sub>obs</sub> [Da]	Comment
FGE <sub>CASA</sub>	35227.9	35224.2	Disulfide bond
FGE <sub>ACSA</sub>	35227.9	35224.8	Disulfide bond
FGE <sub>AACA</sub>	35211.9	35182.6	Undesired mutation Y277H
FGE <sub>AASC</sub>	35227.9	35225.0	Disulfide bond
FGE <sub>ACCC</sub>	35276.0	35272.5	Disulfide bond
FGE <sub>Y273F</sub>	35292.1	35291.0	-
FGE <sub>4C-Y273F</sub>	35179.9	35179.1	-

**Labeling with 2-iodo acetamide.** FGE was diluted to a final concentration of  $c_{\text{FGE}} = 45 \mu\text{M}$  in a reaction mixture containing 2 mM DTT, 50 mM EDTA, 50 mM NaCl and 50 mM Tris-HCl pH 8. The reaction was incubated at RT for 30 min, after which 2-iodo acetamide was added to a final concentration of  $c_{\text{IAA}} = 5 \text{ mM}$ . The reaction was then incubated at RT for 120 min. Protein modification was subsequently analyzed by LC/HRMS.

**Analysis of reaction rates.** Reaction rates were determined using a discontinuous HPLC assay as described previously (see chapter 4.3). Shortly, FGE was diluted to a final concentration of  $c_{\text{FGE}} = 2 \mu\text{M}$  in a reaction buffer containing 200  $\mu\text{M}$  substrate peptide, 2 mM DTT, 2  $\mu\text{M}$   $\text{CuSO}_4$ , 50 mM EDTA, 50 mM NaCl and 50 mM Tris-HCl buffer pH 8. Throughout the reaction time, samples were quenched by addition of one volume equivalent of 4 M urea in 2 % TFA, which were subsequently analyzed by RP-HPLC.

The deuterated substrate was kindly provided by Pascal Engi and synthesized as described before (see section 4.3).

**Michaelis Menten Kinetics.** Reaction rates at increasing amounts of substrate were determined using a discontinuous HPLC assay. FGE was diluted to a final concentration of  $c_{\text{FGE}} = 2 \mu\text{M}$  in a reaction buffer containing 0 - 1000  $\mu\text{M}$  substrate peptide, 2  $\mu\text{M}$   $\text{CuSO}_4$ , 5 mM DTT, 50 mM EDTA, 50 mM NaCl and 50 mM Tris-HCl buffer pH 8. The reaction was incubated at  $T_{\text{react}} = 25 \text{ }^\circ\text{C}$ . Throughout the reaction time, samples were quenched by addition of one volume equivalent of 4 M urea in 2 % TFA, which were subsequently analyzed by RP-HPLC (see above).

**CD measurements.** Circular dichroism measurements have been performed on a Applied Photophysics ChiraScan at  $T_{\text{CD}} = 20 \text{ }^\circ\text{C}$ . Mean residue ellipticity was determined in a 50 mM Tris Buffer pH 8 containing 50 mM NaCl and 2 mM DTT at a final protein concentration of  $c_{\text{FGE}} = 10 \mu\text{M}$ .

**Determination of thermal unfolding by circular dichroism.** FGE was diluted to a final concentration of  $c_{\text{FGE}} = 10 \mu\text{M}$  in a buffer containing 50 mM Tris, 50 mM NaCl and 2 mM DTT. Thermal unfolding was monitored at  $\lambda_{\text{CD}} = 222 \text{ nm}$  in a quartz cuvette of 1 mm pathlength while



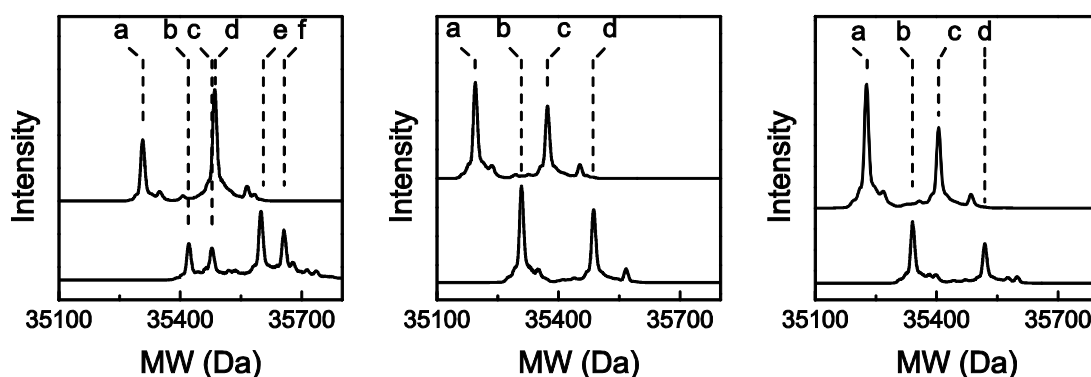
---

heating at 1 °C/min in steps of 1 °C. Data points were fitted with a boltzman distribution curve (see chapter 4.3).

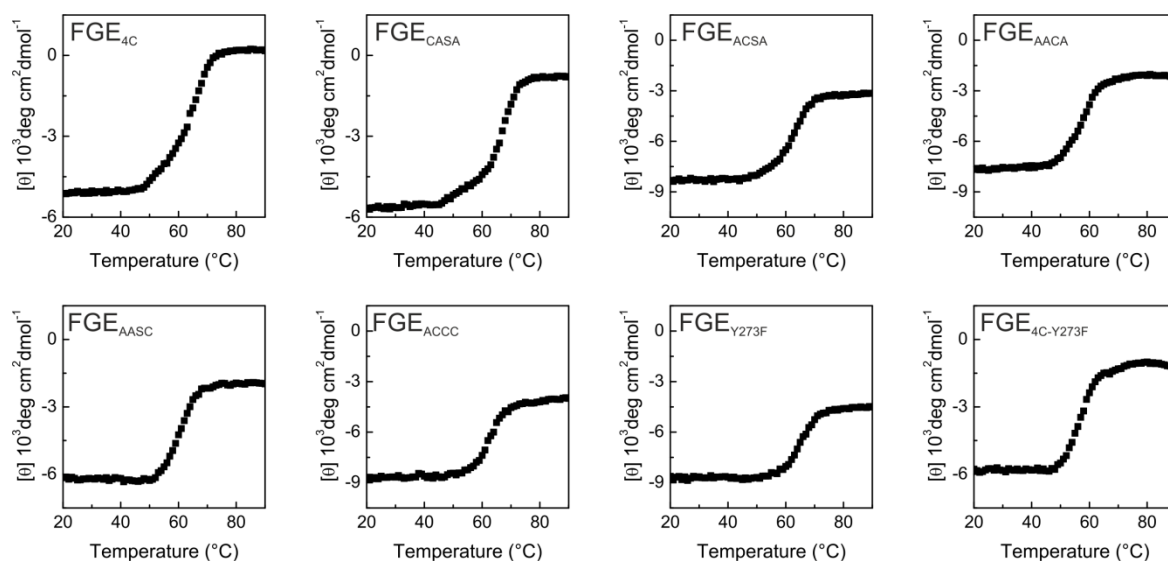
**Copper(I) affinity determination.** Copper affinities were assayed as described above (see also chapter 5.3) Shortly, protein samples were diluted in a 50 mM Tris buffer pH 8 containing 20 μM CuSO<sub>4</sub>, 200 μM bathocuproine disulfonic acid (BCS), 2 mM cysteamine and 4 mM Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub>. The reaction mixture was vortexed and the absorbance at λ<sub>BCS</sub> = 483 nm was determined on a Cary 300 Bio UV/vis spectrophotometer. The concentration of free Cu : BCS<sub>2</sub> was calculated using the molar extinction coefficient of BCS ε<sub>BCS</sub> = 13000 M<sup>-1</sup>cm<sup>-1</sup>. The equilibrium is described by Eq. 4 (see chapter 5.3 and Ref. [100]). Upon linearization the metal affinity is determined.

**Conversion of protein substrate.** To assay the reactivity of FGE in presence of an authentic protein substrate we incubated 0.2 μM of FGE with 75 μM of substrate in a Tris buffer pH 8 containing 2 mM DTT, 50 mM EDTA, 50 mM NaCl at T<sub>inc</sub> = 25 °C. After various time points the reaction was quenched by addition of one volume equivalent of 10 mM 4-nitro-benzylhydroxylamine in a 100 mM acetic acid buffer pH 4.5 supplemented with 6 M urea to ensure complete accessibility of the fGly side chain and inactivation of FGE. The labeling reaction was incubated for 1 – 2 hrs at RT prior to acidification by addition of formic acid and analysis of the protein mass by LC/HRMS.

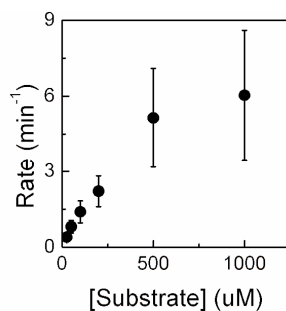
## 7.5 Appendix



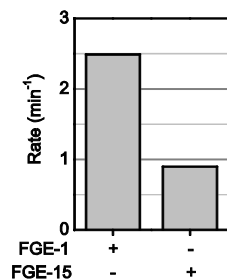
**Fig. 7-9:** Different FGE variants after covalent labeling with 2-iodo acetamide (IAA,  $\Delta MW = 57$  Da). Spectra without addition of alkylating agent are shown above, whereas IAA-labeled samples are shown below. **Left:** FGE<sub>WT</sub> is labeled with 2–3 IAA units.  $MW_{calc,WT} = 35309.1$  Da; a:  $MW_{obs} = 35307.1$  Da, b:  $MW_{obs} = 35421.4$  Da, c:  $MW_{obs} = 35477.6$  Da, d:  $MW_{obs} = 35484.8$  Da, e:  $MW_{obs} = 35599.6$  Da, f:  $MW_{obs} = 35655.8$  Da. **Middle:** FGE<sub>4C</sub> is labeled with 2 IAA units.  $MW_{calc,4C} = 35195.9$  Da; a:  $MW_{obs} = 35194.7$  Da, b:  $MW_{obs} = 35308.7$  Da, c:  $MW_{obs} = 35372.9$  Da, d:  $MW_{obs} = 35486.9$  Da. **Right:** FGE<sub>casa</sub> is labeled with 2 IAA units.  $MW_{calc,casa} = 35227.9$  Da; a:  $MW_{obs} = 35227.0$  Da, b:  $MW_{obs} = 35340.2$  Da, c:  $MW_{obs} = 35405.0$  Da, d:  $MW_{obs} = 35518.8$  Da. Mass increments of  $\Delta MW = 178$  Da result from glycosylation of the N-terminal His<sub>6</sub>-tag.<sup>[88]</sup>



**Fig. 7-10:** Determination of mid-point of thermal transition of different FGE variants in the absence of copper salts.



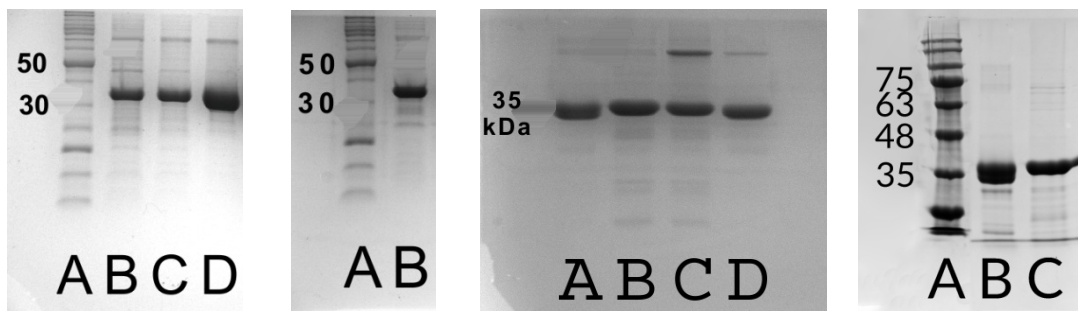
**Fig. 7-11:** Michaelis-Menten plot of FGE<sub>Y273F</sub> showing the reaction rate at increasing substrate concentration. The data was fitted with a hyperbolic curve ( $k_{obs} = ([S] \cdot k_{cat}) / ([S] + K_M)$ ) to obtain  $k_{cat}/K_{M,Y273F} = 21000 \pm 7000 \text{ M}^{-1}\text{min}^{-1}$ .



**Fig. 7-12:** Reaction rate of FGE<sub>Y273F</sub> in the presence of deuterated substrate. The kinetic isotope effect when using a substrate with the  $\beta$ -proton being replaced by deuterium is  $KIE_{Y273F} = 2.8 \pm 0.1$  and hence comparable to that of FGE<sub>WT</sub>.

**Tab. 7-5:** Amino acid sequences of GFPs with different tags as protein substrates for FGE conversion. The FGE recognition motif CxPxR is highlighted as bold letters, the inserted sequence is underlined.

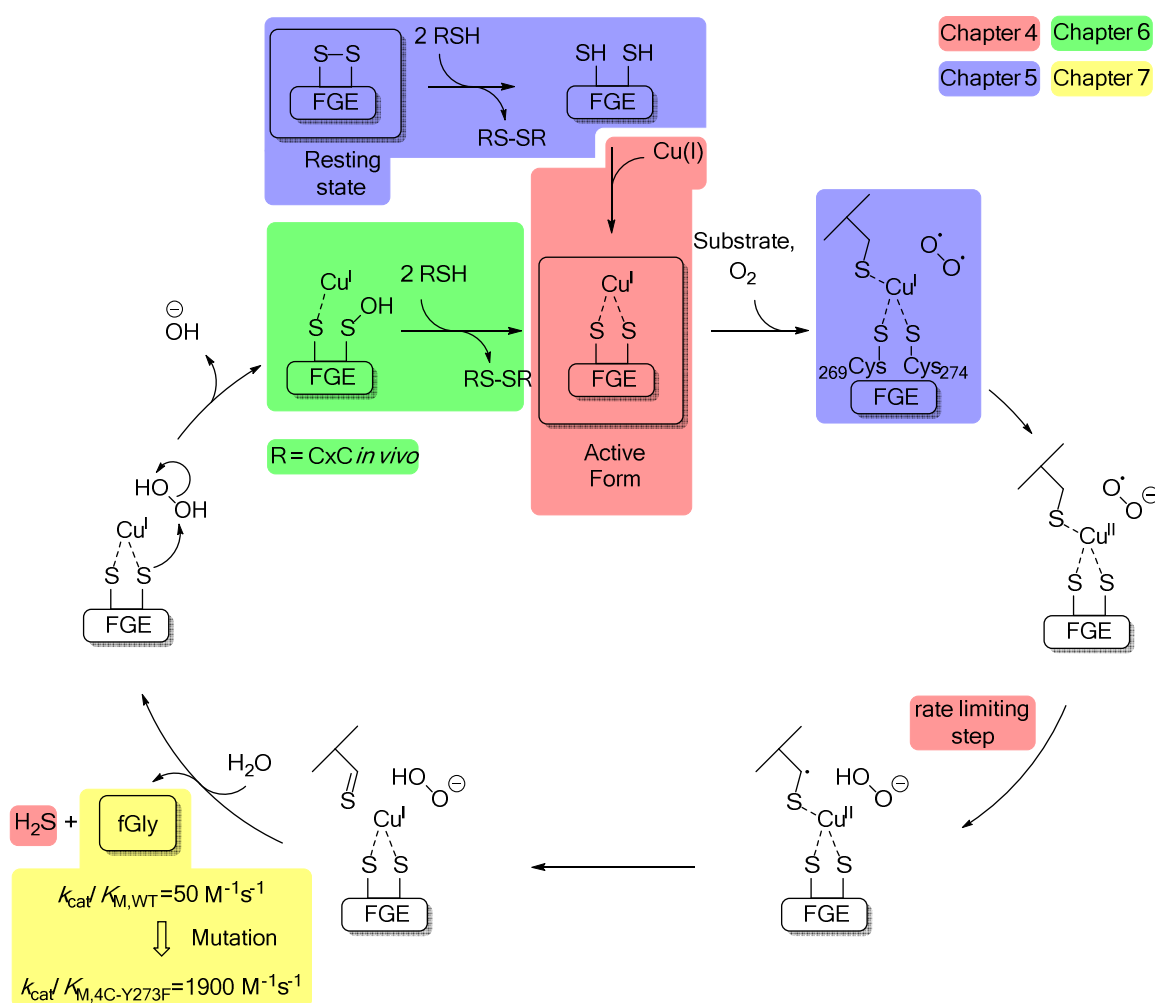
Protein ID	Sequence
<i>Aid-GFP</i>	GTPL <b>CGPSR</b> AHMGSGASKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATY GKLT <del>LF</del> ICTTGKLPVPWPTLVTTLCYGVQCFSRYPDHMKRHDFFKSAMPEGYVQ ERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNNSHNVY IMADK <u>Q</u> NGIKVNFKTRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQS ALSKDPNEKRDHMALLEFVTAAGITHGMDELYKLEHHHHHH
<i>Loop1-GFP</i>	GSSHHHHHSSGLVPRGSHIHMGSGASKGEELFTGVVPILVELDGDVNGHKFSVS GEGEGDATYGKLT <del>LF</del> ICTTGKLPVPWPTLVTTLCYGVQCFSRYPDHMKRHDFFK SAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKL EYNSHNVYIMADK <u>GTPLCGPSRN</u> GIKVNFKTRHNIEDGSVQLADHYQQNTPI GDGPVLLPDNHYLSTQSALSKDPNEKRDHMALLEFVTAAGITHGMDELY
<i>Loop2-GFP</i>	GSSHHHHHSSGLVPRGSHIHMGSGASKGEELFTGVVPILVELDGDVNGHKFSVS GEGEGDATYGKLT <del>LF</del> ICTTGKLPVPWPTLVTTLCYGVQCFSRYPDHMKRHDFFK SAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKL EYNSHNVYIMADK <u>AGKGTPLCGPSR</u> AGNGIKVNFKTRHNIEDGSVQLADHYQ QNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMALLEFVTAAGITHGMDELY



**Fig. 7-13:** SDS-PAGE picture of Different FGE variants. **Left:** A: MWM; B: FGE<sub>CASA</sub>; C: FGE<sub>AACA</sub>; D: FGE<sub>AASC</sub>. **Middle left:** A: MWM; B: FGE<sub>ACSA</sub>. **Middle right:** A: FGE<sub>WT</sub>; B: FGE<sub>4C</sub>; C: FGE<sub>CASA</sub>; D: FGE<sub>ACC</sub>. **Right:** A: MWM; B: FGE<sub>Y273F</sub>; C: FGE<sub>4C-Y273F</sub>.

## 8. Final Conclusion and Outlook

This thesis describes the identification of FGE as a copper-dependent oxidase. It was shown that FGE requires equimolar amounts of copper to reach its highest activity, whereas other transition metals were not accepted (see chapter 4). Copper binds to the reduced active site cysteines that were previously described as being crucial for turnover. The copper(I) ion remains in the active site for multiple turnovers, but can be replaced by silver which renders FGE inactive (see chapter 5). This observation further supports the role of copper as a redox cofactor. The binding of copper to FGE from *Thermomonospora curvata* is accompanied by a conformational change in the protein structure and a shift in resistance to high temperatures (up to  $T \sim 80^\circ\text{C}$ ). This conformational change also occurs with  $\text{Ag}^+$  indicating that the structural adjustment and the increased activity in presence of copper are decoupled. It was moreover shown that the FGE : copper complex is further stabilized in the presence of the substrate, suggesting that the sulfur atom of the substrate cysteine is involved in the active complex (see also chapter 5). Furthermore, we propose that the reducing agent, involved in FGE reactivation *in vivo*, depends



**Scheme 8-1:** Mechanistic proposal of fGly formation under participation of copper as a cofactor based on biochemical investigations performed in this thesis.

---

on a CxC motif similar to thioredoxin-like reducing proteins (see chapter 6). This motif can be found in the previously described N-terminal domain of human FGE, suggesting a participation of this peptide chain in the reduction of FGE.

In addition to these mechanistic findings, the utility of FGE from *T. curvata* for protein labeling applications was demonstrated (see chapter 7). We introduced several cysteine mutations outside the active site and a tyrosine mutation inside the active site. This new FGE<sub>4C-Y273F</sub> variant was shown to be highly active and was able to oxidize more than a 300-fold excess of target protein within 90 min at RT. It was shown that this variant is easy to produce, resists high temperatures and has 30-fold higher catalytic efficiency than the parent protein. We believe that this variant will be of great advantage in the development of new FGE-based labeling applications.

Further experiments are needed to assay the role of the N-terminal domain during the reduction of FGE. Testing a fusion protein of FGE<sub>WT</sub> with PepN attached to the N-terminus should show whether FGE can potentially accept other reducing agents *in vivo*. Moreover, the presence of a sulfenic acid intermediate of FGE has not directly been shown. The development of new sulfenic acid labeling methods might allow for the detection of this proposed intermediate. If the release of the product and the subsequent reduction of FGE follow a ping-pong mechanism, a dimedone derivative could be localized in a substrate-like peptide. At reducing agent limiting conditions, the dimedone-containing peptide might bind to the active site after the product has been released and possibly be covalently linked to the sulfenic acid.

---

## 9. References

- [1] Payen, A; Persoz, J.-F. *Annales de chimie et de physique* **1833**, 2<sup>nd</sup> series, 52, 73-92.
- [2] The Official Web Site of the Nobel Prize.  
[http://www.nobelprize.org/nobel\\_prizes/chemistry/laureates/1946.html](http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1946.html) (accessed Sept 23, 2016).
- [3] Fey, J.; Balleininger, M.; Borissenko, L. V; Schmidt, B.; von Figura, K.; Dierks, T. Characterization of Posttranslational Formylglycine Formation by Luminal Components of the Endoplasmic Reticulum. *J. Biol. Chem.* **2001**, 276 (50), 47021–47028.
- [4] Schmidt, B.; Selmer, T.; Ingendoh, a; von Figura, K. A Novel Amino Acid Modification in Sulfatases That Is Defective in Multiple Sulfatase Deficiency. *Cell* **1995**, 82 (2), 271–278.
- [5] Frese, M. A.; Dierks, T. Formylglycine Aldehyde Tag - Protein Engineering through a Novel Post-Translational Modification. *ChemBioChem* **2009**, 10 (3), 425–427.
- [6] Carrico, I. S.; Carlson, B. L.; Bertozzi, C. R. Introducing Genetically Encoded Aldehydes into Proteins. *Nat. Chem. Biol.* **2007**, 3 (6), 321–322.
- [7] Elstner, E. F. Oxygen Activation and Oxygen Toxicity. *Annu. Rev. Plant Physiol.* **1982**, 33 (1), 73–96.
- [8] Fetzner, S.; Steiner, R. a. Cofactor-Independent Oxidases and Oxygenases. *Appl. Microbiol. Biotechnol.* **2010**, 86 (3), 791–804.
- [9] Fetzner, S. Oxygenases without Requirement for Cofactors or Metal Ions. *Appl. Microbiol. Biotechnol.* **2003**, 60 (3), 243–257.
- [10] Appel, M. J.; Bertozzi, C. R. Formylglycine, a Posttranslationally Generated Residue with Unique Catalytic Capabilities and Biotechnology Applications. *ACS Chem. Biol.* **2014**, No. 10, 72–84.
- [11] Bekker, A.; Holland, H. D.; Wang, P.-L. L.; Rumble, D.; Stein, H. J.; Hannah, J. L.; Coetzee, L. L.; Beukes, N. J. Dating the Rise of Atmospheric Oxygen. *Nature* **2004**, 427 (6970), 117–120.
- [12] Schirrmeister, B. E.; de Vos, J. M.; Antonelli, A.; Bagheri, H. C. Evolution of Multicellularity Coincided with Increased Diversification of Cyanobacteria and the Great Oxidation Event. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, 110 (5), 1791–1796.
- [13] Baron S. *Medical Microbiology*, 4<sup>th</sup> ed.; Galveston, **1996**.
- [14] Sosa Torres, M. E.; Kroneck, P. M. H. Sustaining Life on Planet Earth: Metalloenzymes Mastering Dioxygen and Other Chewy Gases; **2015**; Vol. 15.
- [15] Bugg, T. D. H. How to Break the Rules of Dioxygen Activation. *Chem. Biol.* **2014**, 21 (2), 168–169.
- [16] Thierbach, S.; Bui, N.; Zapp, J.; Chhabra, S. R.; Kappl, R.; Fetzner, S. Substrate-Assisted O<sub>2</sub> Activation in a Cofactor-Independent Dioxygenase. *Chem. Biol.* **2014**, 21 (2), 217–225.
- [17] Massey, V. Activation of Molecular Oxygen by Flavins and Flavoproteins. *J. Biol. Chem.* **1994**, 269 (36), 22459–22462.
- [18] Que, L.; Ho, R. Y. N. Dioxygen Activation by Enzymes with Mononuclear Non-Heme Iron Active Sites. *Chem. Rev.* **1996**, 96 (lii), 2607–2624.
- [19] Hille, R. Structure and Function of Xanthine Oxidoreductase. *Eur. J. Inorg. Chem.* **2006**, No. 10, 1913–1926.
- [20] Goto, T. Chemistry of Bioluminescence. *Pure Appl. Chem.* **1968**, 17, 421–442.
- [21] Gabison, L.; Chopard, C.; Colloc'h, N.; Peyrot, F.; Castro, B.; Hajji, M. El; Altarsha, M.; Monard, G.; Chiadmi, M.; Prangé, T. X-Ray, ESR, and Quantum Mechanics Studies

- 
- Unravel a Spin Well in the Cofactor-Less Urate Oxidase. *Proteins Struct. Funct. Bioinforma.* **2011**, 79 (6), 1964–1976.
- [22] Oja, T.; Klika, K. D.; Appassamy, L.; Sinkkonen, J.; Mantsala, P.; Niemi, J.; Metsa-Ketela, M. Biosynthetic Pathway toward Carbohydrate-like Moieties of Alnumycins Contains Unusual Steps for C-C Bond Formation and Cleavage. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, 109 (16), 6024–6029.
- [23] Bonnot, F.; Iavarone, A. T.; Klinman, J. P. Multistep, Eight-Electron Oxidation Catalyzed by the Cofactorless Oxidase, PqqC: Identification of Chemical Intermediates and Their Dependence on Molecular Oxygen. *Biochemistry* **2013**, 52 (27), 4667–4675.
- [24] Loening, A. M.; Fenn, T. D.; Wu, A. M.; Gambhir, S. S. Consensus Guided Mutagenesis of Renilla Luciferase Yields Enhanced Stability and Light Output. *Protein Eng. Des. Sel.* **2006**, 19 (9), 391–400.
- [25] Grocholski, T.; Koskiniemi, H.; Lindqvist, Y.; Mäntsälä, P.; Niemi, J.; Schneider, G. Crystal Structure of the Cofactor-Independent Monooxygenase SnoaB from *Streptomyces Nogalater*: Implications for the Reaction Mechanism. *Biochemistry* **2010**, 49 (5), 934–944.
- [26] Dudev, T.; Lim, C. Competition among Metal Ions for Protein Binding Sites: Determinants of Metal Ion Selectivity in Proteins. *Chem. Rev.* **2014**, 114 (1), 538–556.
- [27] Frausto da Silva, J. J. R.; Williams, R. J. P. *The Biological Chemistry of the Elements: The Inorganic Chemistry of Life*; Oxford Press: New York, **2004**.
- [28] 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. Chem. Commun* **2016**, 52 (52), 1945–1948.
- [29] Breuer, W.; Epsztejn, S.; Cabantchik, Z. I. Iron Acquired from Transferrin by K562 Cells Is Delivered into a Cytoplasmic Pool of Chelatable iron(II). *J. Biol. Chem.* **1995**, 270 (41), 24209–24215.
- [30] Lippard, S. J.; Berg, J. M.; *Principles of bioinorganic chemistry*. University Science Books: Mill Valley, CA, **1994**.
- [31] Solomon, E. I.; Heppner, D. E.; Johnston, E. M.; Ginsbach, J. W.; Cirera, J.; Qayyum, M.; Kieber-Emmons, M. T.; Kjaergaard, C. H.; Hadt, R. G.; Tian, L. Copper active sites in biology. *Chem. Rev.* **2014**, 114 (7), 3659–3853.
- [32] Firbank, S.; Rogers, M.; Guerrero, R. H.; Dooley, D. M.; Halcrow, M. a; Phillips, S. E. V.; Knowles, P. F.; McPherson, M. J. Cofactor Processing in Galactose Oxidase. *Biochem. Soc. Symp.* **2004**, 31 (71), 15–25.
- [33] Whittaker, J. W. The Radical Chemistry of Galactose Oxidase. *Arch. Biochem. Biophys.* **2005**, 433 (1), 227–239.
- [34] Ito, N.; Phillips, S. E. V.; Stevens, C.; Ogel, Z. B.; McPherson, M. J.; Keen, J. N.; Yadav, K. D. S.; Knowles, P. F. Novel Thioether Bond Revealed by a 1.7 Å Crystal Structure of Galactose Oxidase. *Nature* **1991**, 350 (6313), 87–90.
- [35] Klinman, J. P. Mechanisms Whereby Mononuclear Copper Proteins Functionalize Organic Substrates. *Chem. Rev.* **1996**, 96 (7), 2541–2562.
- [36] Finney, J.; Moon, H. J.; Ronnebaum, T.; Lantz, M.; Mure, M. Human Copper-Dependent Amine Oxidases. *Arch. Biochem. Biophys.* **2014**, 546, 19–32.
- [37] Rogers, M. S.; Tyler, E. M.; Akyumani, N.; Kurtis, C. R.; Spooner, R. K.; Deacon, S. E.; Tamber, S.; Firbank, S. J.; Mahmoud, K.; Knowles, P. F.; et al. The Stacking Tryptophan of Galactose Oxidase : A Second-Coordination Sphere Residue That Has Profound Effects on Tyrosyl Radical Behavior and Enzyme Catalysis. **2007**, 46, 4606–4618.
- [38] Kishishita, S.; Okajima, T.; Kim, M.; Yamaguchi, H.; Hirota, S.; Suzuki, S.; Kuroda, S.; Tanizawa, K. Role of Copper Ion in Bacterial Copper Amine Oxidase : Spectroscopic and Crystallographic Studies of Metal-Substituted Enzymes. **2003**, No. 3, 1041–1055.
- [39] Decker, H.; Hellmann, N.; Jaenicke, E.; Lieb, B.; Meissner, U.; Markl, J. Minireview: Recent Progress in Hemocyanin Research. *Integr. Comp. Biol.* **2007**, 47 (4), 631–644.
-



- 
- [40] Markl, J. Evolution of Molluscan Hemocyanin Structures. *Biochim. Biophys. Acta - Proteins Proteomics* **2013**, *1834* (9), 1840–1852.
- [41] Singleton, C.; Le Brun, N. E. Atx1-like Chaperones and Their Cognate P-Type ATPases: Copper-Binding and Transfer. *BioMetals* **2007**, *20* (3–4), 275–289.
- [42] Polishchuk, R.; Lutsenko, S. Golgi in Copper Homeostasis: A View from the Membrane Trafficking Field. *Histochem. Cell Biol.* **2013**, *140* (3), 285–295.
- [43] Lamb, A. L.; Wernimont, A. K.; Robert, A.; Culotta, V. C.; Halloran, T. V. O.; Rosenzweig, A. C. Letters Crystal Structure of the Copper Chaperone for Superoxide Dismutase. *Nat. Struct. Biol.* **1999**, *6* (8), 724–729.
- [44] Banci, L.; Bertini, I.; Cefaro, C.; Ciofi-Baffoni, S.; Gallo, A. Functional Role of Two Interhelical Disulfide Bonds in Human Cox17 Protein from a Structural Perspective. *J. Biol. Chem.* **2011**, *286* (39), 34382–34390.
- [45] Preusser-Kunze, A.; Mariappan, M.; Schmidt, B.; Gande, S. L.; Mutenda, K.; Wenzel, D.; von Figura, K.; Dierks, T. Molecular Characterization of the Human Calpha-Formylglycine-Generating Enzyme. *J. Biol. Chem.* **2005**, *280* (15), 14900–14910.
- [46] Bojarová, P.; Williams, S. J. Sulfotransferases, Sulfatases and Formylglycine-Generating Enzymes: A Sulfation Fascination. *Curr. Opin. Chem. Biol.* **2008**, *12* (5), 573–581.
- [47] Recksiek, M.; Selmer, T.; Schmidt, B.; Von, K.; Chem, J. B.; Dierks, T.; Figura, K. Von. Sulfatases , Trapping of the Sulfated Enzyme Intermediate by Substituting the Active Site Formylglycine. *J. Biol. Chem.* **1998**, *273* (11), 6096–6103.
- [48] Dierks, T.; Dickmanns, A.; Preusser-Kunze, A.; Schmidt, B.; Mariappan, M.; von Figura, K.; Ficner, R.; Rudolph, M. G. Molecular Basis for Multiple Sulfatase Deficiency and Mechanism for Formylglycine Generation of the Human Formylglycine-Generating Enzyme. *Cell* **2005**, *121* (4), 541–552.
- [49] Dierks, T.; Schmidt, B.; Borissenko, L. V; Peng, J.; Preusser, A.; Mariappan, M.; von Figura, K. Multiple sulfatase deficiency is caused by mutations in the gene encoding the human C(alpha)-formylglycine generating enzyme. *Cell* **2003**, *113*, 435–444.
- [50] Fraldi, A.; Zito, E.; Annunziata, F.; Lombardi, A.; Cozzolino, M.; Monti, M.; Spampinato, C.; Ballabio, A.; Pucci, P.; Sitia, R.; Cosma, M. P. Multistep, sequential control of the trafficking and function of the multiple sulfatase deficiency gene product, SUMF1 by PDI, ERGIC-53 and ERp44. *Hum. Mol. Genet.* **2008**, *17*, 2610–2621.
- [51] Mariappan, M.; Radhakrishnan, K.; Dierks, T.; Schmidt, B.; Von Figura, K. ERp44 mediates a thiol-independent retention of formylglycine-generating enzyme in the endoplasmic reticulum. *J. Biol. Chem.* **2008**, *283*, 6375–6383.
- [52] Zito, E.; Buono, M.; Pepe, S.; Settembre, C.; Annunziata, I.; Surace, E. M.; Dierks, T.; Monti, M.; Cozzolino, M.; Pucci, P.; et al. Sulfatase Modifying Factor 1 Trafficking through the Cells: From Endoplasmic Reticulum to the Endoplasmic Reticulum. *EMBO J.* **2007**, *26* (10), 2443–2453.
- [53] Ennemann, E. C.; Radhakrishnan, K.; Mariappan, M.; Wachs, M.; Pringle, T. H.; Schmidt, B.; Dierks, T. Proprotein Convertases Process and Thereby Inactivate Formylglycine- Generating Enzyme. *J. Biol. Chem.* **2013**, *288* (8), 5828–5839.
- [54] Carlson, B. L.; Ballister, E. R.; Skordalakes, E.; King, D. S.; Breidenbach, M. a; Gilmore, S. a; Berger, J. M.; Bertozzi, C. R. Function and Structure of a Prokaryotic Formylglycine-Generating Enzyme. *J. Biol. Chem.* **2008**, *283* (29), 20117–20125.
- [55] Le Coq, J.; Ghosh, P. Conservation of the C-Type Lectin Fold for Massive Sequence Variation in a Treponema Diversity-Generating Retroelement. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (35), 14649–14653.
- [56] Roeser, D.; Preusser-Kunze, A.; Schmidt, B.; Gasow, K.; Wittmann, J. G.; Dierks, T.; von Figura, K.; Rudolph, M. G. A General Binding Mechanism for All Human Sulfatases by
-

- 
- the Formylglycine-Generating Enzyme. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (1), 81–86.
- [57] Luther, G. W.; Findlay, A. J.; MacDonald, D. J.; Owings, S. M.; Hanson, T. E.; Beinart, R. A.; Girguis, P. R. Thermodynamics and Kinetics of Sulfide Oxidation by Oxygen: A Look at Inorganically Controlled Reactions and Biologically Mediated Processes in the Environment. *Front. Microbiol.* **2011**, *2* (APR), 1–9.
- [58] Wu, P.; Shui, W.; Carlson, B. L.; Hu, N.; Rabuka, D.; Lee, J.; Bertozzi, C. R. Site-Specific Chemical Modification of Recombinant Proteins Produced in Mammalian Cells by Using the Genetically Encoded Aldehyde Tag. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (9), 3000–3005.
- [59] Lo Conte, M.; Carroll, K. S. The Redox Biochemistry of Protein Sulfenylation and Sulfinylation. *J. Biol. Chem.* **2013**, *288* (37), 26480–26488.
- [60] Couvertier, S. M.; Zhou, Y.; Weerapana, E. Chemical-Proteomic Strategies to Investigate Cysteine Posttranslational Modifications. *Biochim. Biophys. Acta - Proteins Proteomics* **2014**, *1844* (12), 2315–2330.
- [61] Marquardt, C.; Fang, Q.; Will, E.; Peng, J.; Von Figura, K.; Dierks, T. Posttranslational Modification of Serine to Formylglycine in Bacterial Sulfatases: Recognition of the Modification Motif by the Iron-Sulfur Protein AtsB. *J. Biol. Chem.* **2003**, *278* (4), 2212–2218.
- [62] Benjdia, A.; Leprince, J.; Sandström, C.; Vaudry, H.; Berteau, O. Mechanistic Investigations of Anaerobic Sulfatase-Maturing Enzyme: Direct C-H Atom Abstraction Catalyzed by a Radical AdoMet Enzyme. *J. Am. Chem. Soc.* **2009**, *131* (24), 8348–8349.
- [63] Grove, T. L.; Ahlum, J. H.; Qin, R. M.; Lanz, N. D.; Radle, M. I.; Krebs, C.; Booker, S. J. Further Characterization of Cys-Type and Ser-Type Anaerobic Sulfatase Maturing Enzymes Suggests a Commonality in the Mechanism of Catalysis. *Biochemistry* **2013**, *52* (17), 2874–2887.
- [64] Knop, M.; Engi, P.; Lemnar, R.; Seebeck, F. P. In Vitro Reconstitution of Formylglycine-Generating Enzymes Requires Copper(I). *ChemBioChem* **2015**, *16*, 2147–2150.
- [65] Peng, J.; Alam, S.; Radhakrishnan, K.; Mariappan, M.; Rudolph, M. G.; May, C.; Dierks, T.; von Figura, K.; Schmidt, B. Eukaryotic Formylglycine-Generating Enzyme Catalyses a Monooxygenase Type of Reaction. *FEBS J.* **2015**, n/a-n/a.
- [66] Frandsen, K. E. H.; Simmons, T. J.; Dupree, P.; Poulsen, J. N.; Hemsworth, G. R.; Ciano, L.; Johnston, E. M.; Tovborg, M.; Johansen, K. S.; von Freiesleben, P.; et al. The Molecular Basis of Polysaccharide Cleavage by Lytic Polysaccharide Monooxygenases. *Nat. Chem. Biol.* **2016**, *12* (4), 298–303.
- [67] Johansen, K. S. Lytic Polysaccharide Monooxygenases: The Microbial Power Tool for Lignocellulose Degradation. *Trends Plant Sci.* **2016**, [epub ahead of print].
- [68] Beeson, W. T.; Vu, V. V.; Span, E. A.; Phillips, C. M.; Marletta, M. A. Cellulose Degradation by Polysaccharide Monooxygenases. *Annu. Rev. Biochem.* **2015**, *84* (1), 923–946.
- [69] Rittle, J.; Green, M. T. Cytochrome P450 Compound I. *Science* **2010**, *330*, 933–938.
- [70] Donoghue, P. J.; Tehranchi, J.; Cramer, C. J.; Sarangi, R.; Solomon, E. I.; Tolman, W. B. Rapid C-H Bond Activation by a monocopper(III)-Hydroxide Complex. *J. Am. Chem. Soc.* **2011**, *133* (44), 17602–17605.
- [71] Kracher, D.; Scheiblbrandner, S.; Felice, A. K. G.; Breslmayr, E.; Preims, M.; Ludwicka, K.; Haltrich, D.; Eijssink, V. G. H.; Ludwig, R. Extracellular Electron Transfer Systems Fuel Cellulose Oxidative Degradation. *Science* **2016**, *352* (6289), 1098–1101.
- [72] Trommsdorff, J. B.; Brandes, R.; Geiger, P. L.; Liebig, J. *Annalen der Pharmacie, Universitäts-Buchhandlung von C. F. Winter: Heidelberg*, **1834**.
-

- 
- [73] Sumner, J. B.; Somers, G. F. *Chemistry and Methods of Enzymes 3<sup>rd</sup> Ed*; Academic Press Inc.: New York, **1953**.
- [74] Drauz, K.; Gröger, H.; May, O. *Enzymes Catalysis in Organic Synthesis, 3<sup>rd</sup> Ed*; Wiley-VCH Verlag: Weinheim, **2012**.
- [75] Aehle, W. *Enzymes in Industry: Production and Applications 2<sup>nd</sup> Ed*; Wiley-VCH Verlag: Weinheim, **2003**.
- [76] Hermanson, G. T. *Bioconjugate Techniques 3<sup>rd</sup> Ed*; Academic Press Inc.: New York, **2013**.
- [77] Williams, E. H.; Davydov, A. V.; Motayed, A.; Sundaresan, S. G.; Bocchini, P.; Richter, L. J.; Stan, G.; Steffens, K.; Zangmeister, R.; Schreifels, J. A.; et al. Immobilization of Streptavidin on 4H-SiC for Biosensor Development. *Appl. Surf. Sci.* **2012**, *258* (16), 6056–6063.
- [78] Carrico, I. S.; Carlson, B. L.; Bertozzi, C. R. Introducing Genetically Encoded Aldehydes into Proteins. *Nat. Chem. Biol.* **2007**, *3* (6), 321–322.
- [79] Rush, J. S.; Bertozzi, C. R. New Aldehyde Tag Sequences Identified by Screening Formylglycine Generating Enzymes in Vitro and in Vivo. *J. Am. Chem. Soc.* **2008**, *130* (37), 12240–12241.
- [80] Lemnar, R. A Chemoenzymatic Strategy for Protein-Nanocellulose Conjugates University of Basel, **2016**.
- [81] Smith, E. L.; Giddens, J. P.; Iavarone, A. T.; Godula, K.; Wang, L. X.; Bertozzi, C. R. Chemoenzymatic Fc Glycosylation via Engineered Aldehyde Tags. *Bioconjug. Chem.* **2014**, *25* (4), 788–795.
- [82] Zang, B.; Ren, J.; Xu, L.; Jia, L. Direct Site-Specific Immobilization of Protein A via Aldehyde-Hydrazide Conjugation. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2016**, *1008*, 132–138.
- [83] Rabuka, D.; Rush, J. S.; deHart, G. W.; Wu, P.; Bertozzi, C. R. Site-Specific Chemical Protein Conjugation Using Genetically Encoded Aldehyde Tags. *Nat. Protoc.* **2012**, *7* (6), 1052–1067.
- [84] Liu, J.; Hanne, J.; Britton, B. M.; Shoffner, M.; Albers, A. E.; Bennett, J.; Zatezalo, R.; Barfield, R.; Rabuka, D.; Lee, J.-B.; et al. An Efficient Site-Specific Method for Irreversible Covalent Labeling of Proteins with a Fluorophore. *Sci. Rep.* **2015**, *5*, 16883.
- [85] Mariappan, M.; Gande, S. L.; Radhakrishnan, K.; Schmidt, B.; Dierks, T.; von Figura, K. The Non-Catalytic N-Terminal Extension of Formylglycine-Generating Enzyme Is Required for Its Biological Activity and Retention in the Endoplasmic Reticulum. *J. Biol. Chem.* **2008**, *283* (17), 11556–11564.
- [86] Del Rio, T. G.; Tice, H.; Cheng, J.; Goodwin, L.; Pitluck, S. Complete Genome Sequence of *Thermomonospora Curvata* Type. *Stand. Genomic Sci.* **2011**, *4*, 13–22.
- [87] Engi, P. D. Personal communication, unpublished data.
- [88] Geoghegan, K. F.; Dixon, H. B.; Rosner, P. J.; Hoth, L. R.; Lanzetti, A. J.; Borzilleri, K. A.; Marr, E. S.; Pezzullo, L. H.; Martin, L. B.; LeMotte, P. K.; et al. Spontaneous Alpha-N-6-Phosphogluconoylation of a “His Tag” in *Escherichia Coli*: The Cause of Extra Mass of 258 or 178 Da in Fusion Proteins. *Anal. Biochem.* **1999**, *267*, 169–184.
- [89] Derrick, T. S.; Kashi, R. S.; Durrani, M.; Jhingan, A.; Middaugh, C. R. Effect of Metal Cations on the Conformation and Inactivation of Recombinant Human Factor VIII. *J. Pharm. Sci.* **2004**, *93* (10), 2549–2557.
- [90] Tamás, M. J.; Sharma, S. K.; Ibstedt, S.; Jacobson, T.; Christen, P. Heavy Metals and Metalloids as a Cause for Protein Misfolding and Aggregation. *Biomolecules* **2014**, *4* (1), 252–267.
- [91] Moffet, J. W.; Zika, R. G.; Petasne, R. G. Evaluation of bathocuproine for the spectrophotometric determination of copper(I) in copper redox studies with applications in studies of natural waters. *Anal. Chim. Acta* **1985**, *175*, 171–179.
-

- 
- [92] Riener, C. K.; Kada, G.; Gruber, H. J. Quick Measurement of Protein Sulfhydryls with Ellman's Reagent and with 4,4'-dithiodipyridine. *Anal. Bioanal. Chem.* **2002**, *373* (4–5), 266–276.
- [93] Thannhauser, T. W.; Konishi, Y.; Scheraga, H. A. Sensitive Quantitative Analysis of Disulfide Bonds in Polypeptides and Proteins. *Anal. Biochem.* **1984**, *138* (1), 181–188.
- [94] Holder, P. G.; Jones, L. C.; Drake, P. M.; Barfield, R. M.; Bañas, S.; de Hart, G. W.; Baker, J.; Rabuka, D. Reconstitution of Formylglycine-Generating Enzyme with Copper(II) for Aldehyde Tag Conversion. *J. Biol. Chem.* **2015**, *290* (25), 15730–15745.
- [95] Hussain, F.; Wittung-Stafshede, P. Impact of Cofactor on Stability of Bacterial (CopZ) and Human (Atox1) Copper Chaperones. *Biochim. Biophys. Acta - Proteins Proteomics* **2007**, *1774* (10), 1316–1322.
- [96] Solioz, M. Proteolytic Degradation of CopZ c CopA 4 Y Z Copy. *Biochem. Soc. Trans.* **2002**, *30*, 688–691.
- [97] York, D.; Baker, J.; Holder, P. G.; Jones, L. C.; Drake, P. M.; Barfield, R. M.; Bleck, G. T.; Rabuka, D. Generating Aldehyde-Tagged Antibodies with High Titers and High Formylglycine Yields by Supplementing Culture Media with copper(II). *BMC Biotechnol.* **2016**, *16* (1), 23.
- [98] Schlotawa, L.; Ennemann, E. C.; Radhakrishnan, K.; Schmidt, B.; Chakrapani, A.; Christen, H.-J.; Moser, H.; Steinmann, B.; Dierks, T.; Gärtner, J. SUMF1 Mutations Affecting Stability and Activity of Formylglycine Generating Enzyme Predict Clinical Outcome in Multiple Sulfatase Deficiency. *Eur. J. Hum. Genet.* **2011**, *19* (3), 253–261.
- [99] Cline, J. D.; Broenkow, W. W. Spectrophotometric Determination of Hydrogen Sulfide in Natural Waters. *Limnol. Oceanogr.* **1969**, *14* (3), 454–458.
- [100] Xiao, Z.; Brose, J.; Schimo, S.; Ackland, S. M.; La Fontaine, S.; Wedd, A. G. Unification of the copper(I) binding affinities of the metallo-chaperones Atx1, Atox1, and related proteins: detection probes and affinity standards. *J. Biol. Chem.* **2011**, *286* (13), 11047–11055.
- [101] Rosenzweig, A. C.; Huffman, D. L.; Hou, M. Y.; Wernimont, A. K.; Pufahl, R. A.; O'Halloran, T. V. Crystal structure of the Atx1 metallochaperone protein at 1.02 Å resolution. *Structure* **1999**, *7* (6), 605–617.
- [102] Allen, S.; Badarau, A.; Dennison, C. The influence of protein folding on the copper affinities of trafficking and target sites. *Dalt. Trans.* **2013**, *42* (9), 3233–3239.
- [103] Stewart, E. J.; Aslund, F.; Beckwith, J. Disulfide bond formation in the Escherichia coli cytoplasm: an in vivo role reversal for the thioredoxins. *Embo J.* **1998**, *17* (19), 5543–5550.
- [104] Katz, A. K.; Shimoni-Livny, L.; Navon, O.; Navon, N.; Bock, C. W.; Glusker, J. P. Copper-binding motifs: Structural and theoretical aspects. *Helv. Chim. Acta* **2003**, *86* (5), 1320–1338.
- [105] Drescher, M.; Jeschke, G. *EPR Spectroscopy Applications in Chemistry and Biology*; Springer-Verlag: New York, **2012**.
- [106] Stillman, M. J. Spectroscopic studies of copper and silver binding to metallothioneins. *Met. Based. Drugs* **1999**, *6* (4–5), 277–290.
- [107] Granatier, J.; Urban, M.; Sadlej, A. J. Van der Waals complexes of Cu, Ag, and Au with hydrogen sulfide. The bonding character.. *J. Phys. Chem. A* **2007**, *111*, 13238–13244.
- [108] Donoghue, P.; Gupta, A. An Anionic, Tetragonal Copper (II) Superoxide Complex. *J. Am. Chem. Soc.* **2010**, *132*, 15869–15871.
- [109] Lanci, M. P.; Smirnov, V. V.; Cramer, C. J.; Gauchenova, E. V.; Sundermeyer, J.; Roth, J. P. Isotopic probing of molecular oxygen activation at copper(I) sites. *J. Am. Chem. Soc.* **2007**, *129* (47), 14697–14709.
- [110] Woycechowsky, Kenneth, J.; Raines, R. T. CXC Motif: A Functional Mimic of Protein Disulfide Isomerase. *Biochemistry* **2003**, *42*, 5387.
-

- 
- [111] Martinez, S.; Wu, R.; Sanishvili, R.; Liu, D.; Holz, R. The Active Site Sulfenic Acid Ligand in Nitrile Hydratases Can Function as a Nucleophile. *J. Am. Chem. Soc.* **2014**, *136* (4), 1186–1189.
- [112] Jeong, W.; Sung, J. P.; Chang, T. S.; Lee, D. Y.; Sue, G. R. Molecular mechanism of the reduction of cysteine sulfinic acid of peroxiredoxin to cysteine by mammalian sulfiredoxin. *J. Biol. Chem.* **2006**, *281* (20), 14400–14407.
- [113] PubChem OPEN CHEMISTRY DATABASE. **2015**. PubChem OPEN CHEMISTRY DATABASE. [ONLINE] Available at: <http://pubchem.ncbi.nlm.nih.gov/>. [Accessed 11 May 15].
- [114] Srinivasan, U; Mieyal, P. A.; Mieyal, J. pH Profiles indicative of Rate Limiting Nucleophilic Displacement in Thioltransferase (Glutaredoxin) Catalysis. *Biochemistry* **1997**, *36*, 3199–3206.
- [115] Domkin, V.; Chabes, A. Phosphines are ribonucleotide reductase reductants that act via C-terminal cysteines similar to thioredoxins and glutaredoxins. *Scientific Reports* **2014**, *4*, 5539.
- [116] Lenaz, G.; Milazzo, G. *Bioelectrochemistry of Biomacromolecules*; Birkhäuser-Verlag: Basel **1997**.
- [117] Lukesh, J. C.; Palte, M. J.; Raines, R. T. A potent, versatile disulfide-reducing agent from aspartic acid. *J. Am. Chem. Soc.* **2012**, *134* (9), 4057–4059.
- [118] Chambers, J. E.; Tavender, T. J.; Oka, O. B. V; Warwood, S.; Knight, D.; Bulleid, N. J. The reduction potential of the active site disulfides of human protein disulfide isomerase limits oxidation of the enzyme by Ero1 $\alpha$ . *J. Biol. Chem.* **2010**, *285* (38), 29200–29207.
- [119] Mustacich, D.; Powis, G. Thioredoxin reductase. *Biochem. J.* **2000**, *346 Pt 1*, 1–8.
- [120] Hudak, J. E.; Yu, H. H.; Bertozzi, C. R. Protein Glycoengineering Enabled by the Versatile Synthesis of Aminooxy Glycans and the Genetically Encoded Aldehyde Tag. *J. Am. Chem. Soc.* **2011**, *133* (40), 16127–16135.
- [121] Liang, S. I.; McFarland, J. M.; Rabuka, D.; Gartner, Z. J. A Modular Approach for Assembling Aldehyde-Tagged Proteins on DNA Scaffolds. *J. Am. Chem. Soc.* **2014**, *136* (31), 10850–10853.
- [122] Cho, H.; Jaworski, J. Enzyme Directed Formation of Un-Natural Side-Chains for Covalent Surface Attachment of Proteins. *Colloids Surfaces B Biointerfaces* **2014**, *122*, 846–850.
- [123] Kwak, E.-A.; Jaworski, J. Controlled Surface Immobilization of Viruses via Site-Specific Enzymatic Modification. *J. Mater. Chem. B* **2013**, *1*, 3486.
- [124] Ranan-Kurussi, S.; Tözser, J.; Cherry, S.; Tropea, J. E.; Waugh, D. S. Differential Temperature Dependence of Tobacco Etch Virus and Rhinovirus 3C Proteases. *Anal. Biochem.* **2013**, *436* (2), 142–144.
- [125] Lamarche, B. J.; Showalter, A. K.; Tsai, M.-D. D. An Error-Prone Viral DNA Ligase. *Biochemistry* **2005**, *44* (23), 8408–8417.
- [126] Asgeirsson, B.; Fox, J. W.; Bjarnason, J. B. Purification and Characterization of Trypsin from the Poikilotherm *Gadus Morhua*. *Eur. J. Biochem.* **1989**, *180* (1), 85–94.
- [127] Forney, L. J.; Wong, D. C. L. Alteration of the Catalytic Efficiency of Penicillin Amidase from *Escherichia Coli*. *Appl. Environ. Microbiol.* **1989**, *55* (10), 2556–2560.
- [128] Wang, F.; Hou, S.; Wang, Q.; Wang, P.; Liu, J.; Yang, B.; Wang, Y. Impact of Leucine 278 Residue on Fatty Acid Length Specificity of *Candida Antarctica* Lipase B. **2015**, No. July, 493–499.
- [129] Lee, H. L.; Chang, C. K.; Jeng, W. Y.; Wang, A. H. J.; Liang, P. H. Mutations in the Substrate Entrance Region of  $\beta$ -Glucosidase from *Trichoderma Reesei* Improve Enzyme Activity and Thermostability. *Protein Eng. Des. Sel.* **2012**, *25* (11), 733–740.
- [130] Dorr, B. M.; Ham, H. O.; An, C.; Chaikof, E. L.; Liu, D. R. Reprogramming the Specificity of Sortase Enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (37), 13343–13348.
-

- 
- [131] Beckett, D.; Kovaleva, E.; Schatz, P. J. A Minimal Peptide Substrate in Biotin Holoenzyme Synthetase-Catalyzed Biotinylation. *Protein Sci.* **2008**, *8* (4), 921–929.
- [132] Abedi, M. R.; Caponigro, G.; Kamb, A. Green Fluorescent Protein as a Scaffold for Intracellular Presentation of Peptides. *Nucleic Acids Res.* **1998**, *26* (2), 623–630.
- [133] Wu, M.; Beckham, G. T.; Larsson, A. M.; Ishida, T.; Kim, S.; Payne, C. M.; Himmel, M. E.; Crowley, M. F.; Horn, S. J.; Westereng, B.; et al. Crystal Structure and Computational Characterization of the Lytic Polysaccharide Monooxygenase GH61D from the Basidiomycota Fungus *Phanerochaete Chrysosporium*. *J. Biol. Chem.* **2013**, *288* (18), 12828–12839.

---

## Acknowledgement

I would like to express my gratitude to Prof. Florian P. Seebeck for believing in my knowledge and skills and for giving me the opportunity to work in his research group. It has been a joyful journey and your impact on my scientific but also personal development in the last years is without peer.

I would like to thank Prof. Thomas R. Ward for accepting the co-examination and for taking the time to read my doctoral thesis. I would like to thank Prof. Dennis Gillingham for chairing the PhD defense session.

I would further like to thank Prof. Gunnar Jeschke for performing EPR measurements and having the patience to answer every question I might have had. Prof. Thomas Dierks and his team is highly acknowledged for the supply of human FGE samples.

Furthermore I would like to highlight the collaboration and participation of Roxana Lemnaru, Pascal Engi and Thanh Q. Dang who greatly supported this project and contributed to the successful outcome.

My special gratitude goes to the current and previous members of the Seebeck research group for the cheerful but also stimulating atmosphere. It has always been a pleasure to come to the laboratory and work in such a great team. Special thanks goes to the members of lab 304. Even in the most difficult times I always enjoyed working and laughing with you guys.

Special thanks goes to Anja Stampfli, Pascal Engi and Sabine Malzkuhn for critical comment on the manuscript.

I would like to thank my friends from the world outside of chemistry for helping me to get my mind free after work.

Writing a PhD thesis is sometimes distressful, troublesome and annoying. Hence, the importance of a supporting backup outside the laboratory in such a time is of tremendous importance. With this, my deepest gratitude goes to my lovely girlfriend Sabine for supporting me every day, every hour and every minute of the last year.

---



---

## CURRICULUM VITAE

Matthias Knop  
Im Winkel 15/1  
D-79576 Weil am Rhein

[E-mail: mtths.knp@gmail.com]

---

### PERSONAL DATA

Nationality	Born 28 Oct 1986 in Ribnitz-Damgarten, Germany german
-------------	--

---

### STUDIES

10/2012 – 03/2017	PhD student in Biochemistry in the group of Prof. Dr. F. P. Seebeck at the department of organic chemistry, University of Basel, Switzerland
04/2012 – 09/2012	Master thesis in the group of Prof. Dr. F. P. Seebeck at the department of organic chemistry, University of Basel, Switzerland
08/2010 – 03/2012	Graduate studies at the school of chemistry at Technical University "Carolo-Wilhelmina zu Braunschweig"
08/2010 – 05/2011	Scholarship of DAAD for two semester studies abroad at the University of Utah, Salt Lake City, United States of America
10/2007 – 07/2010	Undergraduate studies at the school of chemistry at Technical University "Carolo-Wilhelmina zu Braunschweig"

---

### PROFESSIONAL EXPERIENCE

01/2012 – 03/2012	Departement of Organic Chemistry, Research group of Dr. Jeroen Dickschadt, Technical University "Carolo-Wilhelmina zu Braunschweig", research assistant
-------------------	---

---

Synthesis of  $^{13}\text{C}$  labeled lactones for evaluation of the biosynthesis of complex natural products

05/2011 – 12/2011

Departement of High Frequency Technology, Laboratory of electro optics, Technical university "Carolo-Wilhelmina zu Braunschweig", research assistant

Organische und Anorganische Synthesen zur Entwicklung und Verbesserung von neuen Materialien für OLED-artige Emitter

---

#### **PUBLICATIONS AND CONFERENCES**

---

Publications

Knop, M.; Engi, P.; Lemnar, R.; Seebeck, F. P. *ChemBioChem*, **2015**, *16*, 2147-50.

Knop, M.; Dang, T. Q.; Jeschke, G.; Seebeck, F. P. *ChemBioChem*, **2016**, *17*, *accepted author manuscript*.

Conferences

67. Mosbacher Kolloquium 31.03. – 02.04. 2016

Poster presentation: "Enhancement of in vitro activity of the formylglycine generating enzyme by copper(I)"

35. REGIO-Symposium 14.-16.09. 2015

Oral presentation: "Understanding Activity of the Formylglycine Generating Enzyme"

World Expo Milano 22.05.2015

Presentation of research projects of the university of Basel as part of the swiss pavillon