

**Carbonic Anhydrase II: A Model System for Artificial  
Copper Center Design, Protein-guided Cycloadditions,  
Tethering Screenings and Fragment-based Lead Discovery**

**Dissertation**

zur

Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)



dem Fachbereich Pharmazie

der PHILIPPS-UNIVERSITÄT MARBURG

vorgelegt von

**Johannes Schulze Wischeler**

**aus Bergisch-Gladbach**

Marburg/Lahn 2010

---

Vom Fachbereich Pharmazie der Philipps-Universität Marburg  
als Dissertation angenommen am: 24.06.2010

Erstgutachter: Prof. Dr. Gerhard Klebe

Zweitgutachter: Prof. Dr. Ulrich Koert

Tag der mündlichen Prüfung: 25.06.2010

---

Die Untersuchungen zur vorliegenden Arbeit wurden auf Anregung von Herrn Prof. Dr. Gerhard Klebe am Institut für Pharmazeutische Chemie des Fachbereichs Pharmazie der Philipps-Universität Marburg in der Zeit von Januar 2007 bis Juni 2010 durchgeführt.



*Meinen Eltern*



**Abbreviations**

°C	Celsius
$\sigma$ (crystallography)	Sigma-level of an electron density map (refers to probability of finding an electron within the volume enclosed by the map. It provides a way of accessing the "noise" level of the respective electron density map.)
Å	Ångström ( $1\text{Å} = 10^{-10}\text{ m}$ )
A <sub>600</sub>	Absorption at 600 nm
Ala	Alanine
Arg	Arginine
Amp	Ampicillin
Asn	Asparagine
Asp	Aspartic acid
B-factor	Debye-Waller-factor
CA	Carbonic anhydrase
Cam	Chloramphenicol
CSD	Cambridge Structural Database
Cys	Cysteine
DMSO	Dimethylsulfoxide
dNTP	Desoxynucleosidtriphosphat
DTNB	5,5'-dithiobis-(2-nitrobenzoic acid)
DTT	Dithiothreitol
E. coli	Escherichia coli
F <sub>c</sub>	calculated structure amplitudes
F <sub>o</sub>	observed structure amplitudes
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
GST	Glutathione-S-Transferase

## Abbreviations

---

h	Hour
H-bond	Hydrogen bond
His	Histidine
HPLC-MS	High pressure liquid chromatography-mass spectrometry
Ile	Isoleucine
IPTG	Isopropylthio- $\beta$ -galactosid
I.U.	International unit
K	Kelvin
k	Kilo
Da	Dalton
$K_i$	Competitive inhibition constant
$K_m$	Michaelis Menten constant
Leu	Leucine
Lys	Lysine
M	Molarity (mol/l)
Met	Methionine
Min	Minute
NMR	Nuclear magnetic resonance
NTP	Nucleosidtriphosphat
PAGE	Polyacrylamid gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank
pH	Potentialis hydrogenii
Phe	Phenylalanine
Pro	Proline
R-factor	Reliability factor
RMSD	Root mean square deviation
rpm	Revolutions per minute



## Abbreviations

---

SDS	Sodiumdodecylsulfate
Ser	Serine
Thr	Threonine
Tris	Tris(hydroxymethyl)-aminomethane
Trp	Tryptophane
Tyr	Tyrosine
Val	Valine
w/v	Weight per volume
WT	wild type
X-ray	Electromagnetic radiation, also called Röntgen radiation (wavelength: 10 – 0.01 nanometers, typical wavelengths used for X-ray crystallography are around 1 Å (= 0.1 nanometer), which is in the range of covalent chemical bonds.)



**Table of contents**

<b>Abbreviations .....</b>	<b>7</b>
<b>Table of contents .....</b>	<b>11</b>
<b>1 Motivation .....</b>	<b>15</b>
1.1 Background .....	15
1.2 Aims of this Thesis.....	17
<b>2 Introduction Carbonic Anhydrase II.....</b>	<b>21</b>
2.1 Carbonic Anhydrases .....	21
2.2 Structure of CA II.....	22
2.3 Catalytic Mechanism of CA II.....	25
2.4 Carbonic Anhydrase Inhibitors .....	27
<b>3 Cloning, Expression, Purification and Crystallization of CA II and CA II Mutants.....</b>	<b>31</b>
3.1 Introductory Remarks.....	31
3.2 Cloning and Mutagenesis .....	31
3.3 Expression and Purification .....	33
3.4 Crystallization .....	35
<b>4 Introduction of an Artificial Cu Binding Site at the Surface of CA II: Pitfalls of Rational Design Finally Scooped by Serendipity.....</b>	<b>37</b>
4.1 Introductory Remarks.....	37
4.2 Introduction .....	37
4.3 Results .....	42
4.4 Conclusion.....	50
<b>5 Stereo- and Regioselective Azide/Alkyne Cycloadditions in Carbonic Anhydrase II via Tethering Followed by Protein Crystallography and Mass Spectrometry .....</b>	<b>53</b>
5.1 Introductory Remarks.....	53
5.2 Abstract .....	53

5.3	Introduction .....	54
5.4	Results and Discussion .....	57
5.5	Conclusion .....	70
5.6	Experimental Section.....	71
<b>6</b>	<b>Saccharin Inhibits Carbonic Anhydrases: Possible Explanation for its Unpleasant Metallic Aftertaste .....</b>	<b>73</b>
6.1	Introductory Remarks .....	73
6.2	Introduction .....	73
6.3	Results .....	74
<b>7</b>	<b>Tethering Approach on Carbonic Anhydrase II Supported by Pre-selecting Computational Docking Experiments .....</b>	<b>79</b>
7.1	Introductory Remarks .....	79
7.2	Introduction .....	79
7.3	Results .....	86
7.4	Experimental Section.....	110
<b>8</b>	<b>The Sulfamide Head Group: A Versatile Zinc Chelator for Carbonic Anhydrase Inhibition of Remarkable Structural Variability .....</b>	<b>117</b>
8.1	Introductory Remarks .....	117
8.2	Introduction .....	117
8.3	Results .....	120
8.4	Discussion.....	130
8.5	Experimental Section.....	135
<b>9</b>	<b>Bidentate Zinc Chelators for <math>\alpha</math>-Carbonic Anhydrases that Produce a Trigonal Bipyramidal Coordination Geometry .....</b>	<b>139</b>
9.1	Introductory Remarks .....	139
9.2	Abstract.....	139
9.3	Introduction .....	139
9.4	Results .....	141
9.5	Discussion.....	147
9.6	Experimental Section.....	149
<b>10</b>	<b>Conclusion and Outlook .....</b>	<b>151</b>

<b>11 Zusammenfassung und Ausblick.....</b>	<b>157</b>
<b>12 Materials and Methods.....</b>	<b>163</b>
12.1 Materials.....	163
12.2 Methods.....	167
<b>13 Appendix.....</b>	<b>173</b>
13.1 Crystal Data.....	173
13.2 Tethering Compound Library .....	177
13.3 Sequencing of CA II and CA II-Mutants .....	180
<b>14 Literature.....</b>	<b>183</b>
<b>Danksagung .....</b>	<b>199</b>
<b>Erklärung .....</b>	<b>201</b>
<b>Curriculum Vitae .....</b>	<b>203</b>



# 1 Motivation

## 1.1 Background

Protein crystallography is a powerful tool in drug discovery and development and is applied for, e.g. the determination of *de novo* protein structures as well as the investigation of protein-ligand complexes. The crystallographic analysis of unknown protein structures can help to understand in case of enzymes the involvement of the essential amino acids in the catalytic mechanism and reveal crucial substrate binding pockets that should be addressed by inhibitors. The investigation of protein-ligand complexes can provide exact spatial positioning of the compound within the active site helping to understand the key interactions of inhibition.

However, the application of protein crystallography was limited by several parameters in the past. Apart from the challenging identification of crystallization conditions, the application of protein crystallography required time consuming data collection and huge computational resources for structure refinement. To minimize this experimental effort, further analyses are performed in drug discovery. Prior to structure determination another faster screening method is preceded in order to achieve a pre-selection of promising compounds that are subsequently analyzed by crystallographic experiments. Besides these expenses, the number of experiments that are necessary for successful determination of a protein-ligand complex can vary from case to case. Parameters, such as ligand size, ligand solubility and crystal stability take inherent impact on the success of the crystallographic experiment.

Especially protein complexes with rather small molecules (molecular weight < 250 Da), so called fragments, are usually difficult to determine. Along the refinement process the localization of the compound within the active site is determined by the difference electron density. Apart from the resolution, the quality of the density is mainly dependent on the compound affinity, the applied ligand concentrations and thus on the occupation of the ligand at its position. Compounds of increased molecular weight substituted with several functional groups can perform more interactions to the protein compared to a small fragment and thus result in better defined difference electron

density. Accordingly, it is well comprehensible that fragments which can occupy diverse binding modes within the same active site are not easily placed into a protein structure. The success rate of determining protein fragment complexes suffers from these multiple binding modes and therefore, increased crystallographic effort is required to screen for optimal crystallization conditions as often several soaking and cocrystallization experiments have to be performed.

Systematical fragment screenings via protein crystallography were rather unusual in the past. Nevertheless, the entire field has been stimulated by a series of successful screenings using NMR. Stephen Fesik et al. first reported successful detection and utilization of fragments in drug design which became popular as the SAR-by-NMR approach (Structure-Activity-Relationships by Nuclear Magnetic Resonance) (Shuker, Hajduk et al. 1996). In their study, two-dimensional, isotope-edited NMR spectroscopy was used to detect two fragment leads that bind at two proximal sites on the protein surface. Using three-dimensional structural information about the bound ligands, the fragments were successfully joined together to produce a high affinity ligand.

Nevertheless, in the last 10 to 15 years protein crystallography has advanced and emerged to a kind of routine analytical tool also applicable for screening purposes. The availability of synchrotron beam time and improved in-house radiation sources has increased remarkably and in combination with improved computing power the collection of 50 datasets within 24 hours is possible. In addition, low temperature data collection, robotic-systems for protein crystallization, fast protein-crystal mounting and diffraction screening as well as automated structure refinement tools support the speed of this method. The advancement in these technologies can also be demonstrated by the number of protein structures that are deposited each year in the protein data bank (PDB) (Figure 1.1). The number of structures that are uploaded to the PDB has increased annually reaching nearly 7500 structures in 2009. Actually, the collected and publicly available data only represent a fraction of the determined protein structures since pharmaceutical companies usually do not publish their data. The increasing impact of protein crystallography being an analytical tool allows application of this method for library screening without cumbersome, tedious pre-selecting experiments. Formerly,



only in very legitimate case of promising compounds exhibiting high affinity in a functional bioassay the crystallographic analysis was attempted.

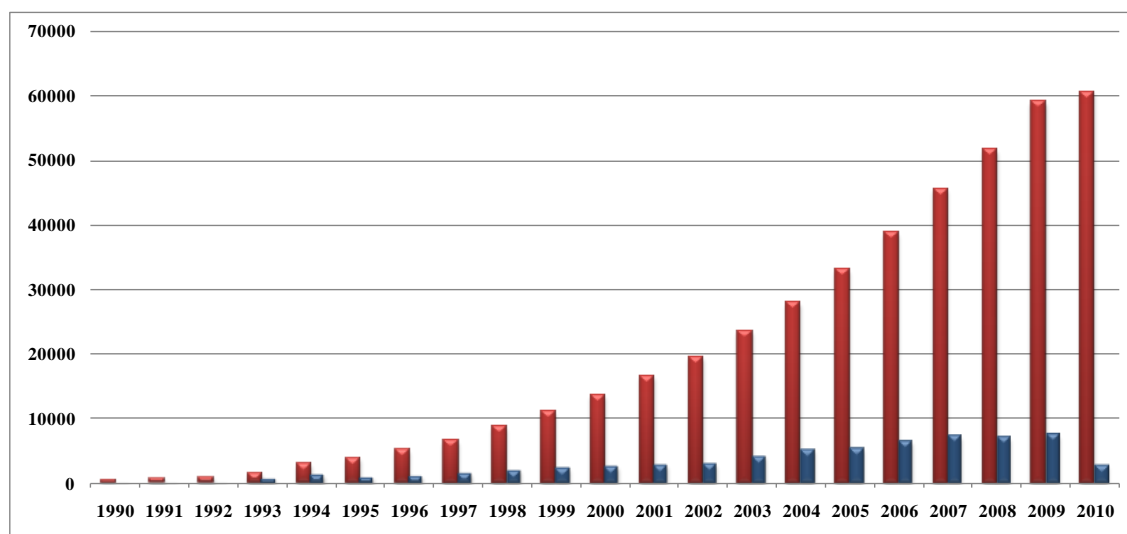


Figure 1.1 : Number of deposited protein structures in the Protein Data Bank (PDB; red: total; blue: year) ([www.pdb.org](http://www.pdb.org) – 03/2010).

## 1.2 Aims of this Thesis

The aim of this thesis was to take advantage of these enhancements in protein crystallography and predominantly apply this method to a variety of rather diverse fragment-based experiments. How can interactions of proteins with small molecules be exploited? Are we able to, e.g. monitor *in situ* reactions via protein crystallography? Will a tethering approach increase the success rate of fragment-based crystallography?

Carbonic anhydrase II has been selected as a model protein for these experiments as this enzyme has been studied exhaustively in our group. Several protein-ligand complexes, in most cases with high resolution, have been determined so far. Diverse experiments with small molecules should be performed and followed by protein crystallography. In principle, crystallographic investigation performed in each experiment should provide the structural basis to understand the processes executed in the active site of CA II and help to develop new approaches.

The following issues have been investigated:

1. **Cloning, expression and mutagenesis of CA II:** In order to perform the broad range of crystallographic studies, considering a scope of different mutant variants,

the gene coding for CA II has to be recombinantly available and modified by mutagenesis. The protein has to be accessible in high amounts to perform the planned crystallographic and in-solution experiments.

2. **Artificial copper center:** One initial idea of this work was the introduction of an artificial copper center at the surface of CA II. Copper-coordinating protein residues should be introduced into the target protein via site-directed mutagenesis. The investigation of each mutation and Cu affinity experiments should be followed by protein crystallography revealing structural information about the formed complexes.
3. ***In situ* click chemistry:** Addressing this metal center azide and alkyne building blocks should be brought to reaction to form a triazole product via Cu-catalyzed 2+3 cycloaddition. The cycloaddition reaction should be illustrated in the same way allowing comprehension of the reaction by determining protein-reactants, protein-intermediate and protein-product complexes.
4. **Tethering:** The tethering method allows investigation of fragments with weak binding affinity by covalently attaching the compound via a thiole containing moiety to a surface exposed cysteine residue (Erlanson, Braisted et al. 2000). The screening of fragment libraries can be investigated by HPLC-MS analysis and protein crystallography. Via a computational docking approach a large fragment library should be pre-selected to a number of fragments feasible to synthesize which should be tested subsequently by different HPLC-MS experiments. The most promising compounds should be investigated crystallographically for their key interactions with the protein. This allows identification of scaffolds for different binding regions of CA II. Subsequently, the fragments should be enlarged to improve their binding affinity to finally release the compound from the tether.
5. **Investigation of a sulfamide inhibitor series:** A series of sulfamides should be investigated with respect to their binding to the target CA II. Starting with compounds of fragment size the series is completed by inhibitors with a molecular weight of 350 Da. All compounds were synthesized by Haake et al. and Winum et

al. and kinetically characterized by Supuran et al. Protein-ligand key interactions should be revealed by protein crystallography.

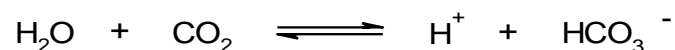
6. **Investigation of new zinc binding groups coordinating to the active site of CA II:** The sulfonamide is the most prominent CA II zinc coordinating moiety and is present in the majority of all known CA II inhibitors. In this study different fragment-like zinc coordinating groups should be investigated for the affinity to CA II and structurally characterized by protein crystallography.



## 2 Introduction Carbonic Anhydrase II

### 2.1 Carbonic Anhydrases

Carbonic anhydrases (CA) (EC 4.2.1.1) belong to the class of hydrolyases containing a catalytic active  $Zn^{2+}$  ion at the bottom of the binding pocket. CAs are ubiquitous present in all animals and plants as well as bacteria and divided into  $\alpha$ -,  $\beta$ -,  $\gamma$ - and  $\delta$ -families:  $\alpha$ -CAs are mainly present in vertebrates, also in plants, algae and some bacteria (Chegwidden and Carter 2000; Kerry and James 2000),  $\beta$ -CAs mainly in leaves of plants in eubacteria, archaea and some algae,  $\gamma$ -CAs mainly in archaea and some eubacteria and  $\delta$ -CAs in diatoms (Supuran and Scozzafava 2007). With a molecular weight between 28-30 kDa CAs catalyze the reversible hydration of  $CO_2$  to hydrogen carbonate:



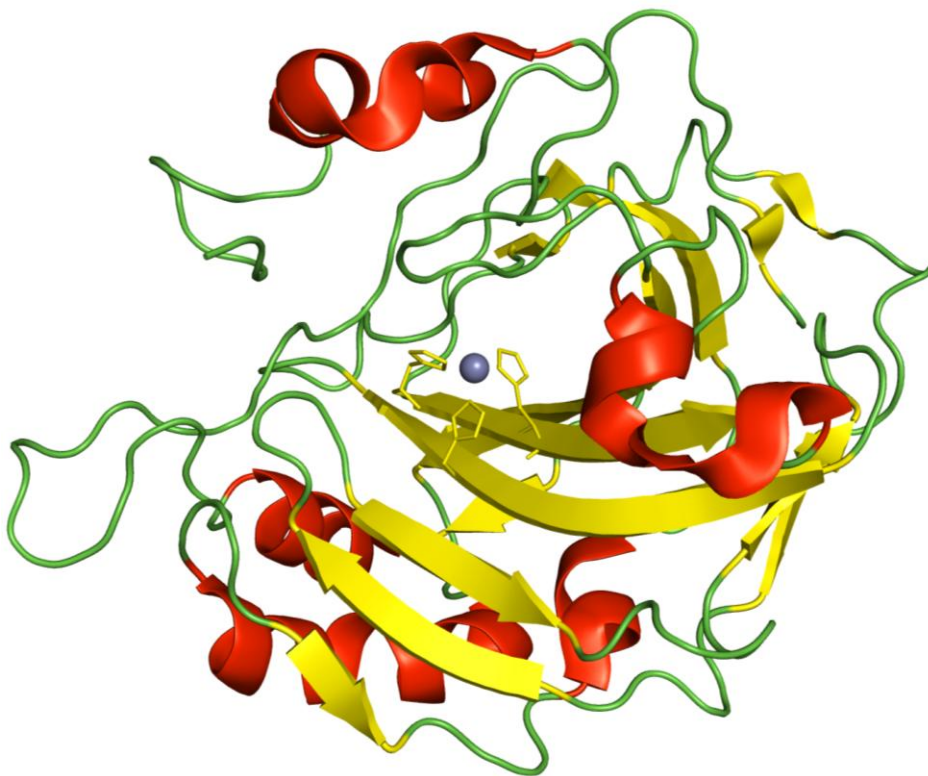
The 16 different  $\alpha$ -CA isoforms described so far in mammals, including *homo sapiens*, are involved in pH and  $CO_2$  homeostasis, respiration and transport of  $CO_2$ /bicarbonate between metabolizing tissues and lungs, electrolyte secretion in a variety of tissues/organs, biosynthetic reactions (e.g. gluconeogenesis, lipogenesis and ureagenesis), bone resorption, calcification, tumorigenicity, and many other processes (Maestrelli, Mura et al. 2002; Supuran, Scozzafava et al. 2003; Stiti, Cecchi et al. 2008; Supuran 2008; Winum, Rami et al. 2008).

The most prominent CA-isoenzyme, the carbonic anhydrase II (CA II) is located in many organs showing a very high catalytic efficacy. With a transformation rate of  $10^6/s$  CA II belongs to the fastest enzymes known so far (Khalifah 1971). The direction of the reaction is pH dependent. At a pH above 7,  $CO_2$  is hydrated to hydrogen carbonate; below 7 the backwards reaction occurs (Silverman, Tu et al. 1976). At a pH between 7 and 9 the enzyme shows the highest activity.

## 2.2 Structure of CA II

### 2.2.1 Folding of CA II

The ellipsoidal enzyme CA II exhibits the size of approximately 55 x 44 x 39 Å. 260 amino acids participate in the primary structure of CA II. Determining the tertiary structure, Liljas et al. showed that the protein exists of only one domain (Liljas, Kannan et al. 1972). The protein is divided by a 10-fold beta-sheet, which is the predominant secondary element (Figure 2.1). Besides two parallel strands the beta sheet is antiparallel.



**Figure 2.1:** Tertiary structure of human carbonic anhydrase II ( $\alpha$ -helices; sheets; loops). The active site  $Zn^{2+}$  ion (grey sphere) is coordinated by three histidine residues.

It has been shown via structure determination that there are no disulfide bridges present. Furthermore, all lysine residues are located at the surface of the enzyme which are in close contact to the side chains of the neighboring molecules. Within the folding pattern of CA II, there are two clusters formed by aromatic residues. The first cluster consists of four aromatic amino acids: Trp5, Trp7, Trp16 and Phe20. The other one is formed by eight residues Phe66, Phe70, Phe93, Phe95, Phe176, Phe179, Phe226 and Trp97. The

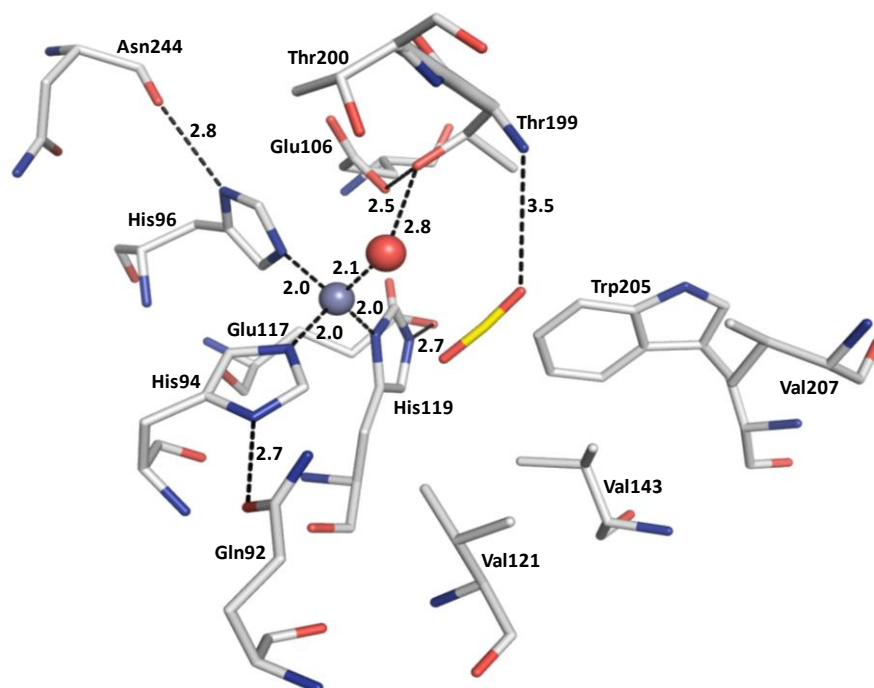
latter cluster is also present in carbonic anhydrase I and partially in carbonic anhydrase III (Eriksson, Jones et al. 1988).

### 2.2.2 Active Site of CA II

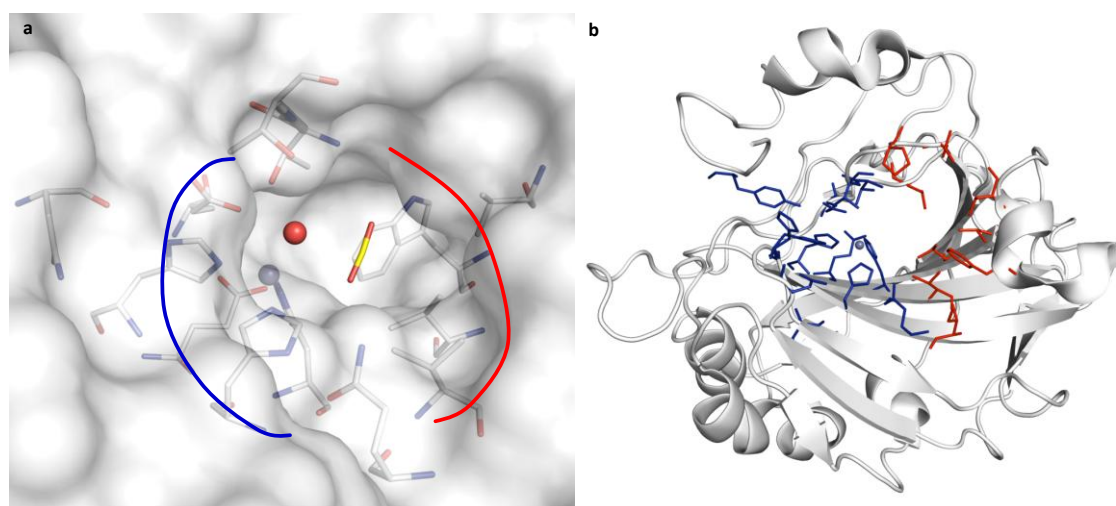
The active site of CA II is amphiphilic. Without a bound ligand the binding pocket is filled by a network of water molecules which provides direct contact to the solvent surrounding of CA II. At the bottom of the 15 Å deep binding pocket the catalytic active  $Zn^{2+}$  ion is coordinated by three histidine residues. His94 and His96 coordinate with the  $N\epsilon$ -Atom, whereas His119 uses its  $N\delta$ -Atom to interact with the metal ion. Via these primary histidine residues  $Zn^{2+}$  is connected to four additional amino acids. His94 ( $N\delta 1$ ) forms a hydrogen bond to the oxygen atom of Gln92 ( $O\epsilon 1$ ) with a distance of 2.7 Å, likewise His119 ( $N\epsilon 2$ ) with the carboxyl side chain of Glu117 ( $O\epsilon 2$ ) at a distance of 2.7 Å. Furthermore, His96 ( $N\delta 1$ ) is in close contact with the backbone oxygen of Asn244 via an H-bond (2.8 Å) (Figure 2.2) (Christianson and Fierke 1996). A water molecule, most probably present as  $OH^-$  ion, completes as fourth ligand the coordination geometry at the  $Zn^{2+}$  ion forming a distorted tetrahedron (Silverman and Lindskog 1988; Merz and Banci 1997). A further water molecule, the so called "deep water", is located in the conically shaped binding pocket. This water is displaced by  $CO_2$  during the catalytic reaction (Figure 2.2) (Eriksson, Jones et al. 1988).

Thr199 adopts a fundamental role in the catalytic mechanism stabilizing the  $Zn^{2+}$  bound hydroxide ion via H-bonding. Additionally, it forms an H-bond to the substrate  $CO_2$  (Figure 2.2 and Figure 2.3a). The substrate binding pocket is shaped by Val121, Val143 and Val207 as well as Trp209, Thr200, Glu106, Thr199 and His119. Considering the entire amphiphilic binding pocket the amino acids can be divided into two groups according to their hydrophilic and hydrophobic properties. The hydrophobic portion consists of Ile91, Val121, Phe131, Val135, Leu141, Val143, Leu198, Pro202 and Leu204. The hydrophilic amino acids in the binding pocket are: Tyr7, Asn62, His64, Asn67, Gln92, His94, His96, Glu106, Glu117, His119, Thr199 and Thr200 (Figure 2.3b). The active site  $Zn^{2+}$  ion can be replaced by a variety of metal ions. Only the  $Co^{II}$  containing protein shows a comparable activity as the  $Zn^{2+}$  enzyme (approx. 50%). Cu,

Fe, Ni and Mn develop only minor or no activity while showing similar properties in size, charge and  $pK_a$ -value compared to  $Zn^{2+}$ .



**Figure 2.2** Substrate and active site water in the binding pocket of CA II (PDB-code: 2VVA). The catalytic relevant residues are shown in stick representation (protein: C, N, O; ligand: C, O);  $Zn^{2+}$  ion as grey sphere; active site water molecule as red sphere.

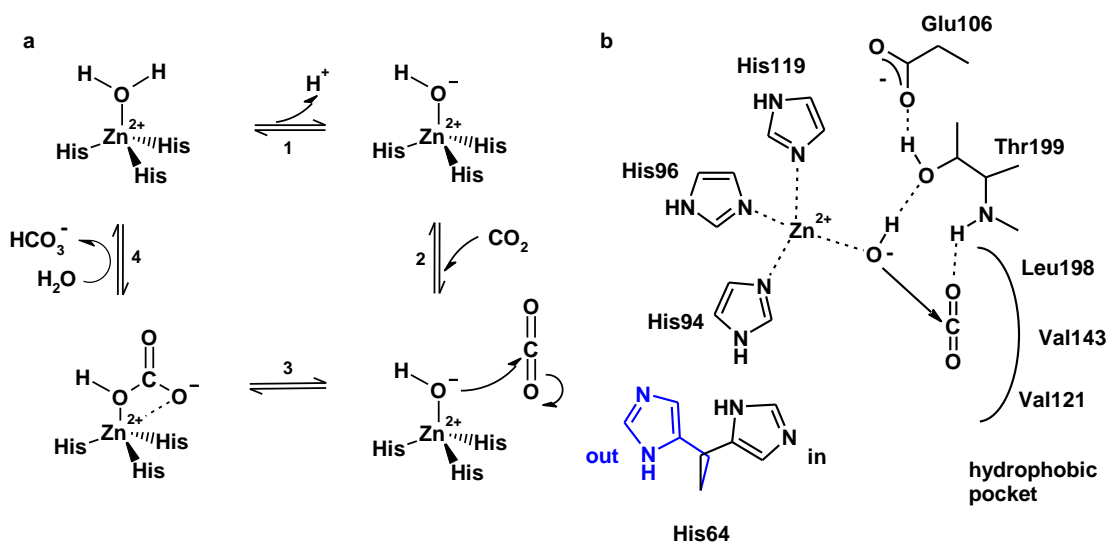


**Figure 2.3** Active site of CA II. (a) Substrate and active site water in the binding pocket of CA II (PDB-code: 2VVA). The solvent accessible surface is shown in white;  $Zn^{2+}$  is shown as grey sphere; water molecule as red sphere. Relevant residues and  $CO_2$  are shown in stick representation (protein: C, N, O; ligand: C, O). The binding pocket of CA II can be divided into a hydrophobic (indicated by red line) and a hydrophilic binding region (blue line). (b) Amphiphilic binding pocket of CA II; hydrophobic amino acids red sticks; hydrophilic residues blue sticks.



### 2.3 Catalytic Mechanism of CA II

During the catalytic mechanism the  $Zn^{2+}$  ion plays the central role. The impressive turnover rate of CA II is achieved by the local vicinity of all participating reactants. However, the actual mechanism is controversially discussed in literature (Silverman and Lindskog 1988; Hakansson, Carlsson et al. 1992; Kiefer and Fierke 1994; Jackman, Merz et al. 1996). The transient catalytic process cannot be investigated by X-ray diffraction techniques. The mechanism proposed by Lindskog was proven by extensive mutagenesis and QM/MM<sup>1</sup> calculations (Figure 2.4) (Behravan, Jonsson et al. 1990; Alexander, Nair et al. 1991; Nair and Christianson 1991b; Nair and Christianson 1991a; Kiefer, Krebs et al. 1993; Krebs and Fierke 1993; Nair and Christianson 1993; Lindskog 1997; Merz and Banci 1997; Toba, Colombo et al. 1999).



**Figure 2.4** (a) Catalytic mechanism and (b) orientation within the binding pocket of CA II. The  $Zn^{2+}$  ion facilitates the deprotonation of the active site water (1). The emerging hydroxide ion performs a nucleophilic attack towards the  $CO_2$  molecule that is fixed by an H-bond to Thr199 and lipophilic interactions to Leu198, Val143 and Val121 (2).  $HCO_3^-$  is formed next to the  $Zn^{2+}$  ion (3). The hydrogencarbonate ion is replaced by another water molecule and released into the solvent (4). The regeneration of the hydroxide ion at the  $Zn^{2+}$  ion is mediated by a proton transfer via two additional water molecules and the histidine 64 residue (1). The function of this "proton shuttle" is dependent on two conformations that His64 can adopt (in- and out-positions).

The residues that are important for the catalytic mechanism are on the hydrophilic side (Thr199, Glu106 and His64) and on the hydrophobic side (Val121, Val143, Leu198 and Trp209). The hydrophobic residues form a pocket which allows  $CO_2$  or the "deep

<sup>1</sup> QM/MM: Quantum Mechanical and Molecular Mechanical

water", respectively, to bind. The high catalytic rate is among others explained by the fact that after desolvation, CO<sub>2</sub> is only loosely bound while the interaction to the NH group of Thr199 is formed by the CO<sub>2</sub>/HCO<sub>3</sub><sup>-</sup> intermediate. Weak polar interactions to the Zn<sup>2+</sup> ion help orienting the CO<sub>2</sub> molecule within the binding pocket. Thereby, the Zn<sup>2+</sup> bound OH<sup>-</sup> ion is not replaced by CO<sub>2</sub>. HCO<sub>3</sub><sup>-</sup> is formed upon nucleophilic attack of OH<sup>-</sup> towards the immobilized CO<sub>2</sub>. Furthermore, it is proposed that Thr199 functions as a kind of "doorkeeper" allowing only those anions to bind to the active site Zn<sup>2+</sup> ion, which are able to form an H-bond to the side chain oxygen of Thr199 (Christianson and Fierke 1996). The distinct H-bond acceptor functionality of Thr199 is induced by a strong H-bond of its hydroxyl proton to Glu106 (Merz 1991). Different mutations of Thr199 have shown a stabilizing effect on the HCO<sub>3</sub><sup>-</sup>/Zn<sup>2+</sup>-complex thereby inhibiting the fast dissociation of HCO<sub>3</sub><sup>-</sup> to leave the binding pocket. This dissociation process is assumed to be the rate-determining step. An additional water molecule binding to the Zn<sup>2+</sup> ion pushes the hydrogencarbonate out of the active site and after deprotonation it takes part in the next catalytic cycle of hydrating CO<sub>2</sub>.

Mediated by two water molecules the residue His64 functions as a "proton shuttle" transporting the proton to the surrounding buffer solution. Crystal structure analysis at pH values of 5.7, 6.5, 8.5 and 9.5 revealed a pH dependent transition in conformational states of His64 (Nair and Christianson 1991b; Nair and Christianson 1991a). His64 clearly rotates away from the binding pocket at a pH of 5.7. Furthermore, mutational studies exchanging His64 by diverse amino acids demonstrated that histidine is the best catalyst for the transformation of CO<sub>2</sub> to HCO<sub>3</sub><sup>-</sup> since it regenerates the active Zn<sup>2+</sup>-OH<sup>-</sup> form and shows the lowest pH dependency (Lindskog 1997). The k<sub>cat</sub>-values are pH dependent as shown by modifying the buffer medium (Silverman and Lindskog 1988). At a pH value close to 8 the Lys64 mutant shows a five times decreased efficiency compared to the wild type while the Glu64 mutant is inactive. In contrast at pH 6 the Lys64 mutant is inefficient and the Glu64 mutant seems to be a 2-3 times more effective proton shuttle with respect to the His64 form (Engstrand, Forsman et al. 1992).

## 2.4 Carbonic Anhydrase Inhibitors

### 2.4.1 Anionic Inhibitors

Anions, such as  $\text{HS}^-$ ,  $\text{CN}^-$ ,  $\text{NCO}^-$ ,  $\text{N}_3^-$ ,  $\text{HSO}_3^-$ ,  $\text{I}^-$  and  $\text{HCOO}^-$  develop in general only weak inhibition of CA II by coordinating to the active  $\text{Zn}^{2+}$  ion. There are three possibilities for the coordination to the metal: Displacement of the catalytic water molecule forming a tetrahedral geometry, or occupation of the fifth coordination site thus forming a distorted trigonal pyramidal geometry or equilibrium between both states is observed.

Anions that provide a proton to coordinate to the  $\text{OH}^-$  group of Thr199 ( $\text{HS}^-$ ,  $\text{HSO}_3^-$ ) displace the metal coordinating water molecule without affecting coordination geometry. The  $\text{Zn}^{2+}$  coordination remains tetrahedral in a pH range of 5.7 to 8.0 (Lindskog 1997). Upon displacement of the "deep water",  $\text{HSO}_3^-$  forms an additional H-bond to the NH-group of Thr199.

Most proton free anionic inhibitors change only slightly the position of the  $\text{Zn}^{2+}$  bound water molecule and do not affect the H-bond to Thr199. Anions such as nitrate bind in the vicinity of the metal pushing the water molecule aside. They form an H-bond to the NH-group of Thr199. Furthermore, nitrate occupies the fifth coordination site at the  $\text{Zn}^{2+}$  ion forming a distorted trigonal pyramidal structure (Mangani and Kansson 1992).

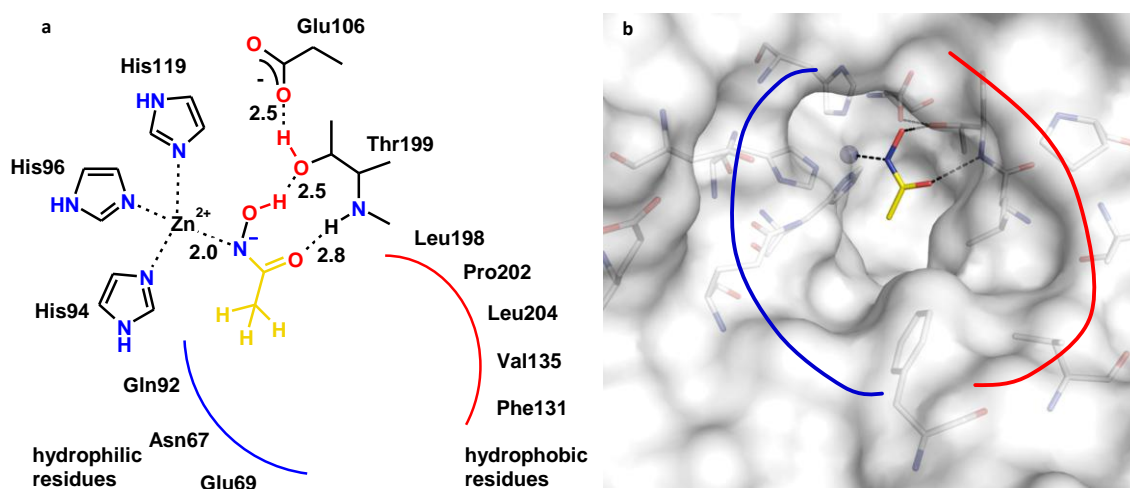
### 2.4.2 Inhibitors of the Proton Shuttle

$\text{Cu}^{2+}$  ions and  $\text{Hg}^{2+}$  ions develop an extraordinary inhibition mechanism. These coordinate to the imidazole nitrogen of His64 thereby prohibiting its function as a proton shuttle. With respect to CA II at pH 7.3 an  $\text{IC}_{50}$  of 0.5  $\mu\text{M}$  was determined for  $\text{Cu}^{2+}$  binding.

### 2.4.3 Hydroxamate Inhibitors

For the compounds acetohydroxamic acid and trifluorohydroxamic acid the  $\text{IC}_{50}$  values against CA II were measured ( $\text{IC}_{50}(\text{CH}_3\text{CONHOH}) = 47 \mu\text{M}$ ;  $\text{IC}_{50}(\text{CF}_3\text{CONHOH}) = 3.8 \mu\text{M}$ ) (Scolnick, Clements et al. 1997). Crystal structure analysis of a CA II –

acetohydroxamic acid complex reveals a binding mode analog to the binding geometry known from sulfonamide inhibitors. The carbonyl oxygen of the hydroxamic acid forms an H-bond to the NH-group of Thr199 (Figure 2.5). The hydroxyl oxygen of Thr199 operates as an H-bond acceptor towards the hydrogen of the hydroxamic hydroxyl group. Furthermore, the hydrogen atom of the Thr199 OH-group forms an H-bond to Glu106. The deprotonated nitrogen of the hydroxamic acid is tetrahedrally coordinated to the  $Zn^{2+}$  ion. By introducing suitable substituents this class of CA II inhibitors achieves inhibition in the nanomolar range (Scozzafava, Banciu et al. 2000).

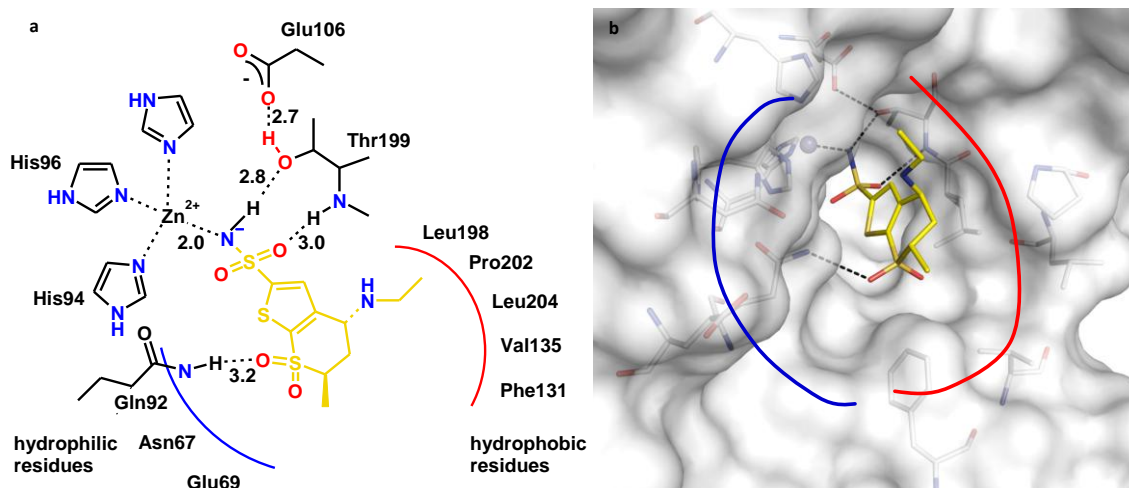


**Figure 2.5** Binding mode of acetohydroxamic acid within the binding pocket of CA II. (a) in schematic representation (b) Crystal structure (PDB-Code: 2AM6). The solvent accessible surface is shown in white. Relevant amino acids and the ligand are shown stick representation (protein: C, N, O; ligand: C, N, O).  $Zn^{2+}$  is shown as grey sphere. Hydrophobic residues are indicated by the red line and hydrophilic residues by the blue line.

#### 2.4.4 Sulfonamide Inhibitors

Sulfonamides show a highly conserved interaction pattern to CAs. The active site water molecule and the "deep water" are displaced and the sulfonamide group coordinates with its deprotonated nitrogen to the  $Zn^{2+}$  ion (Figure 2.6). The remaining hydrogen of the NH-group functions as a donor forming an H-bond to the side chain oxygen of Thr199. The first oxygen of the sulfonamide moiety acts as an acceptor forming an H-bond to the hydrogen of the main chain nitrogen of Thr199, while the second oxygen is in weak contact to the  $Zn^{2+}$  ion. The affinity of the sulfonamide is strongly modulated by the additional substituents connected to the sulfonamide.

Due to the binding mode of sulfonamides the substitution at the nitrogen of the primary, terminal sulfonamide group mostly leads to the loss of inhibitory potential (Krebs 1948). Hence it is astonishing that Saccharin develops inhibition against CAs. In contrast to common sulfonamide CA-inhibitors, in the latter compound the sulfonamide moiety is endocyclic and neighbored by a carbonyl group. Thereby a sulfimide is formed decreasing the  $pK_a$ -value for the NH-group to 1.5.

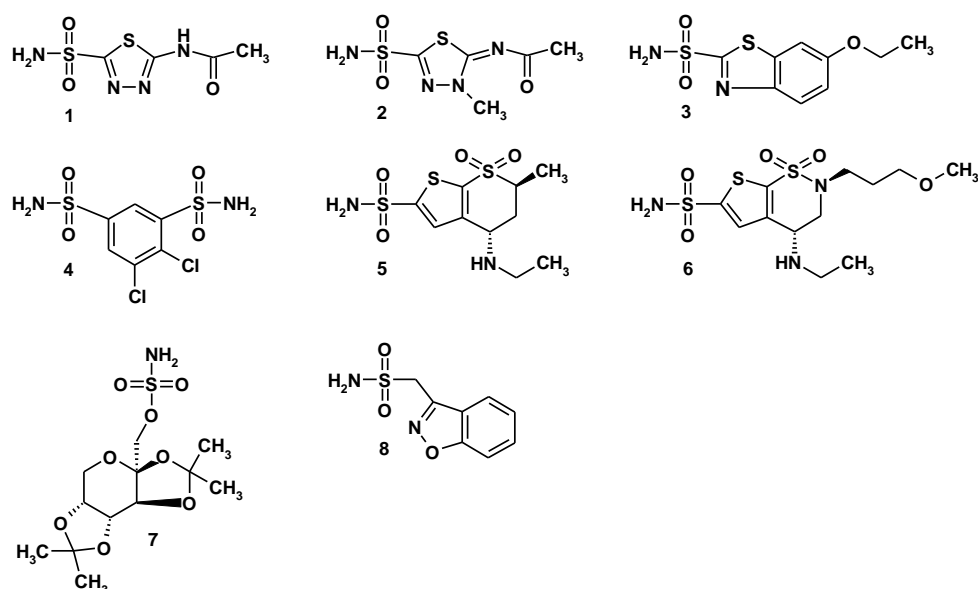


**Figure 2.6** Binding mode of dorzolamide within the binding pocket of CA II. (a) Schematic representation (b) Crystal structure (PDB-Code: 1CIL). The solvent accessible surface is shown in white. Relevant amino acids and the ligand are shown in stick representation (protein: C, N, O; ligand: C, N, O, S).  $Zn^{2+}$  is shown as grey sphere. Hydrophobic residues are indicated by the red line and hydrophilic residues by the blue line.

### 2.4.5 Therapeutically Applied Carbonic Anhydrase Inhibitors

For over 40 years acetazolamide (Figure 2.7) was applied as diuretic drug. However, since 1997 it is no longer approved for this indication. Inhibiting the carbonic anhydrase in the kidneys, acetazolamide decreases the resorption of  $HCO_3^-$  and consequently increases diuresis (Maren 1967). Classic CA-inhibitors **1-4** are also applied in the therapy of glaucoma. However, the systematic application of these unselective CA-inhibitors does not simply affect CAs in the eye. Also CAs in the kidney, the liver, the gastrointestinal tract, the pancreas and the brain are inhibited. This leads to undesired side effects such as anorexia, fatigue, depressions, renal calculi, gastrointestinal disorder and metabolic acidosis (Maren 1967; Supuran and Scozzafava 2000). To prevent these side effects, topically administered sulfonamides that endow only a local effect in the eyes were developed in the 90s (**5**, dorzolamide, Trusopt® and

**6**, brinzolamide, Azopt®). These novel CA inhibitors show sufficient water solubility and adequate lipophilic properties to still pass the cornea (Maren 1995; Supuran and Scozzafava 2000). The classic CA-inhibitors **1-4** are not appropriate for the topical use since the concentrations reached at the target tissues are not sufficient. These compounds show either rather lipophilic or hydrophilic properties. Next to the therapy of glaucoma, acetazolamide **1** is also applied in the therapy of different forms of epilepsy, of cerebral oedema and of anoxia (Larson, Roach et al. 1982; Reiss and Oles 1996; Carrion, Hertzog et al. 2001). Methazolamide **2**, topiramate **7** and zonisamide **8** are likewise used as anticonvulsant (Dodgson, Shank et al. 2000; Masereel, Rolin et al. 2001; Supuran and Scozzafava 2002).



**Figure 2.7** Therapeutically applied CA inhibitors. **1** acetazolamide, **2** methazolamide, **3** ethoxzolamide, **4** dichlorphenamide, **5** dorzolamide, **6** brinzolamide, **7** topiramate, **8** zonisamide.

### 3 Cloning, Expression, Purification and Crystallization of CA II and CA II Mutants

#### 3.1 Introductory Remarks

Cloning of the CA II gene was performed during a four week stay in the group of Prof. Pastorekova, Institute of Virology - Slovak Academy of Sciences (Bratislava). Subsequently, the expression system was optimized for high protein yields in our lab. Up to 30 mg pure protein can be achieved of 1 l expression culture. Site-directed mutagenesis allows creation of desired mutants.

#### 3.2 Cloning and Mutagenesis

**Cloning of CA II gene:** The gene coding for CA II was amplified by PCR from a construct, generously provided by the group of Pastorekova (12.2.1.1) applying the primers XhoICA2a and EcoRICA2s (Table 12.4). The PCR-fragment and the plasmid pGEX-4T1 (GE Healthcare) were each incubated with the restriction enzymes XhoI and EcoRI to create the sticky ends for ligation (12.2.1.2). After ligation (12.2.1.3) the plasmid was transformed into *E.coli* XL-2Blue (Table 12.5) competent cells and incubated for plasmid multiplication (12.2.1.5). A sequencing experiment showed that the sequence was identical to the original copy (Appendix 13.3). For high gene expression the plasmid was transformed into *E.coli* BL21 CodonPlus cells (Table 12.5;12.2.1.5).

**Creation of CA II mutants:** All CA II mutants were made by site-directed mutagenesis using the expression vector containing the wild-type CA II coding region with the primer listed in Table 3.1 and in Table 12.4(12.2.1.4). The point mutations were made using the QuikChange II Kit for site-directed mutagenesis. The mutations were introduced focusing on different objectives (Table 3.1). The mutants CA II-W5C, CA II-G63C, CA II-H64C, CA II-W5C-H64C and CA II-W5C-H64M were created in the context of *in situ* click chemistry to evaluate their affinity towards Cu ions monitored by protein crystallography. In order to perform tethering and tethering-assisted azide/alkyne cycloaddition experiments the mutant CA II-H64C was created. In

addition several other mutations were introduced into the active site replacing amino acids that are relevant for the binding of typical CA II inhibitors.

**Table 3.1 Objective of constructs and primers used for cloning of CA II-WT and creation of CAII mutants.<sup>2</sup>**

Mutant	Applied Plasmid	Primer	Objective
WT	pGEX-4T1	EcoRICA2s/XhoICA2a	Cloning of CA II
W5C	pGEX-4T1-CA II	CA2W5C s/a	2+3 C. <sup>3</sup>
G63C	pGEX-4T1-CA II	CA2G64C s/a	2+3 C.
H64C	pGEX-4T1-CA II	CA2H64C f/rev	2+3 C.
W5C-H64C	pGEX-4T1-CA II-W5C	CA2H64C f/rev	2+3C./Tethering
W5C-H64M	pGEX-4T1-CA II-W5C	CA2H64M f/rev	2+3 C.
W5C-H64C-H96E	pGEX-4T1-CA II-W5C-H64C	CA2H96E f/rev	2+3 C.
W5C-H64C-H96K	pGEX-4T1-CA II-W5C-H64C	CA2H96K f/rev	2+3 C.
W5C-H64C-E106L	pGEX-4T1-CA II-W5C-H64C	CA2E106L f/rev	2+3 C.
W5C-H64C-T198V	pGEX-4T1-CA II-W5C-H64C	CA2T198V f/rev	2+3 C.
H64C-H96K	pGEX-4T1-CA II-H64C	CA2H96K f/rev	2+3C./Tethering
H64C-H96E	pGEX-4T1-CA II-H64C	CA2H96E f/rev	2+3C./Tethering
H64C-E106L	pGEX-4T1-CA II-H64C	CA2E106L f/rev	2+3C./Tethering
H64C-T198V	pGEX-4T1-CA II-H64C	CA2T198V f/rev	2+3C./Tethering

<sup>2</sup> Plasmids have each been transformed into X12-blue and BL21Codon plus Cells. The strains are named: bacterial strain-vector-gene; double and triple mutations were introduced in the equivalent single or double mutant.

<sup>3</sup> 2+3 C. = 2+3 Cycloaddition



### 3.3 Expression and Purification

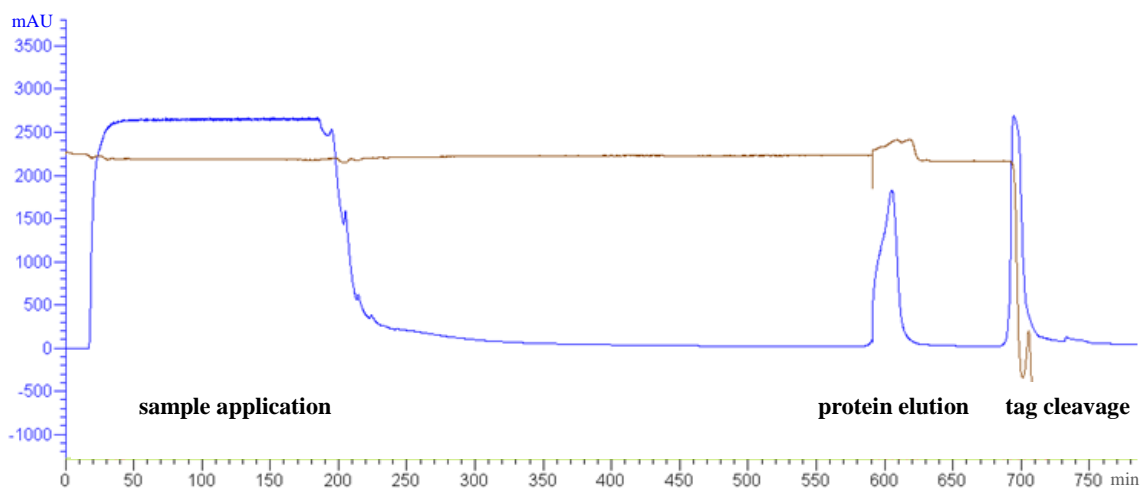
**Gene expression:** CA II and CA II mutants were expressed and purified according to the Glutathione-S-Transferase (GST) gene fusion system handbook (Amersham Biosciences) with the following adjusted protocol. A first 100 ml culture was incubated in the presence of 100 µg/ml amp and 34 µg/ml cam overnight at 37 °C. 10 ml of this solution was used to inoculate a 1000 ml expression culture with 100 µg/ml amp. Expression of the CA II glutathione-S-transferase fusion protein was induced at an optical density at 600 nm ( $OD_{600}$ ) of 0.6-0.8 with 2 mM IPTG, and incubation was continued for 4 h at 32 °C.

**Lysis:** Cells were harvested by centrifugation at 5,000 rpm for 15 min at 4 °C. The cell pellet was resuspended with 100 ml PBS buffer (Table 12.3), which contained a protease inhibitor cocktail tablet and approximately 100 mg of lysozyme. After chilling for 30 min on ice, the cells were lysed by ultrasonic sonication using a Branson Sonifier 250, applying 7 cycles of 2 min with duty cycle set to 70 and output control to 4. This solution was centrifuged at 20,000 rpm for 45 min at 4 °C and filtered (0.45 µM) to separate the soluble fractions from the insoluble material.

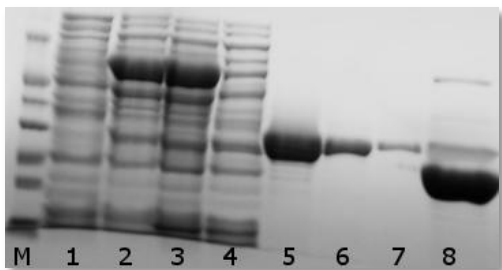
**Protein purification:** The flow-rate along all purification steps was adjusted to a maximum pressure of 0.25 MPa. The sample was applied to a HiPrep GSTPrep FF 16/10 column (GE Healthcare) at a flow-rate of 1 ml/min to ensure that all the fusion protein, which shows very low binding kinetics, binds to the sepharose (Figure 3.1). Afterwards the column was washed with 10 column-volumes of PBS buffer at a flow-rate of 3 ml/min. For cleaving the GST-tag 500 I.U. of thrombin (from Beriplast, CSL Behring) were added with 1 ml/min and the column was slightly shaken for 20 hours at room temperature. After incubation, the target protein was eluted from the column with approximately three column-volumes of PBS at a flow-rate of 1 ml/min. The elution volume was mixed with the same volume of 100 mM Tris-HCl pH 7.4, 1 M NaCl buffer. This mixture was purified over a HiTrap Benzamidine column (Amersham Bioscience) to separate the thrombin from CA II. As a final purification step, the probe was applied on a HiLoad 26/60 Superdex200 size exclusion column (GE Healthcare) to

purify the target enzyme from any last impurity and change the buffer to 50 mM Tris pH 7.8. To monitor the different purification steps, the probes were controlled by SDS-PAGE (Figure 3.2) (Laemmli 1970).

The final purification of 1 l expression culture yielded approximately 30 mg CA II. Finally the GSTPrep column was washed with reduced glutathione buffer to elute the GST-tag from the column and check for full thrombin cleavage. The purified enzyme was concentrated using Vivaspin20 (MW: 10 kDa) columns to a final concentration of 10 mg/ml, which was used for crystallization experiments. For further measurements the enzyme was diluted to the required concentration using a 50 mM Tris-HCl pH 7.8 buffer.



**Figure 3.1 Elution profile of CA II and CA II mutant purification.** The samples were applied on the column (first peak). After purification with 10 column volumes of buffer 500 I.U. of thrombin were added and the column incubated for 20 hours. Thrombin cleaves CA II from the GST-tag which is eluted after incubation (second peak). Finally adding a reduced glutathione buffer, the GST-tag can be eluted from the column (third peak).



**Figure 3.2 SDS-gel-** Fractions of CA II expression: M: Marker (68 kDa, 45 kDa, 36 kDa, 29 kDa, 24 kDa, 20 kDa 14 kDa, 6.5 kDa); 1: before induction; 2: after induction; 3: after lysis; 4: flow-through upon sample application; 5-7: fractions 1-3 upon elution of the target protein; 8: cleavage of the GST-Tag

### 3.4 Crystallization

The CA II-WT crystals were grown using the sitting drop vapor diffusion method at 18 °C by mixing 5 µl of the protein solution (~10 mg/ml) with 5 µl of a well solution (2.75 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.3 M NaCl, 0.15 mM *p*-chloromercuribenzoic acid 0.1 M Tris-HCl pH 7.8). After a few days the crystallization drops were seeded with a seeding solution from old CA II crystals. Crystals appeared within 1 – 4 weeks in space group *P*2<sub>1</sub> and took up to 6 months to reach their maximum size. Complex structures were obtained by cocrystallization of the enzyme with the compounds at concentrations of 1 mM or soaking in 3 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 50 mM Tris pH 7.8 + 0.5 - 1 mM inhibitor. For cryoprotection, crystals were briefly soaked in mother liquor containing 25% glycerol (Acros organics).

Since we have introduced certain cysteine residues at the surface of CA II, which could possibly be coordinated by the mercury, we applied for these mutants the following crystallization conditions: 2.75 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 mM Tris pH 7.8. As the mercury provides an exceeded crystal growth, the mercury-free crystallization of these mutants did not result in the typical crystal size, showing the shape of flat plates. However, the diffraction pattern showed still sufficient quality to collect in-house and synchrotron data sets.



## **4 Introduction of an Artificial Cu Binding Site at the Surface of CA II: Pitfalls of Rational Design Finally Scooped by Serendipity**

### **4.1 Introductory Remarks**

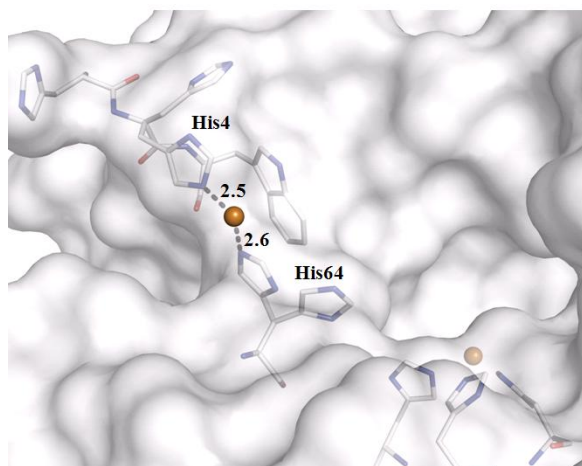
The following chapter has been prepared for submission to a scientific journal. The aim of this study was the introduction of an artificial copper center at the surface of CA II. This metal ion should show catalytic effect on enzyme bound azides and alkynes to perform 2+3 cycloadditions. Protein Crystallography was mainly applied to evaluate different mutants. Altogether approx. 100 datasets were collected in the context of this project. Extensive protein mutagenesis, gene expression and crystallization experiments have been performed.

### **4.2 Introduction**

Carbonic anhydrase II is a zinc metalloenzyme that catalyzes the reversible hydration of carbon dioxide to bicarbonate. It is involved in a variety of physiological processes (Alberty and Hammes 1958). At the bottom of the 15 Å deep active site of CA II the catalytically relevant  $Zn^{2+}$  ion is coordinated by His94, His96 and His119. During the catalytic reaction also His64 plays an important role by shuffling protons from the binding pocket to the surface and vice versa (Steiner, Jonsson et al. 1975). The preference of this residue to recognize and interact with positively charged species could be demonstrated by a recent crystallographic study in our group of a CA II inhibitor complex that revealed the density for a metal ion coordinating to this His64 at the rim of the binding pocket (s. also Chapter 6). According to the applied crystallization conditions it was identified as a  $Zn^{2+}$  ion.

Already previously, it has been demonstrated that this residue is capable to experience the coordination of metal ions. Silverman et al. observed that  $Cu^{2+}$  ions inhibit human CA II through binding to a site remote from the active site zinc ion. Even though lacking crystallographic information they concluded that the copper ions bind tightly to His64 since the ion inhibits the catalytic mechanism without affecting the equilibrium rate of  $CO_2$  hydration (Tu, Wynns et al. 1981). Evidence for this hypothesis was provided later by a crystal structure of CA II in complex with  $Cu^{2+}$  (Figure 4.1)

(Hakansson, Wehnert et al. 1994). Apart from replacing the active site zinc ion, a second  $\text{Cu}^{2+}$  ion is coordinating to His64 and His4 at a distance of approximately 2.5 Å each. The refinement revealed only partial occupation of the metal ion at this site. Obviously, occupancy is influenced by the applied pH conditions since the histidine residue needs to be deprotonated for copper coordination. Apart from  $\text{Cu}^{2+}$ , also  $\text{Hg}^{2+}$  is able to occupy this site while being coordinated by His64, Asp62 and a water molecule (Eriksson, Kylsten et al. 1988).



**Figure 4.1** Crystal structure of carbonic anhydrase II in complex with  $\text{Cu}^{2+}$  (PDB-Code: 1RZC). The solvent accessible surface of the protein is represented in white. Relevant amino acids are shown in stick representation (C, N, O). The copper ions are shown as brown spheres. The active  $\text{Zn}^{2+}$  ion is replaced by  $\text{Cu}^{2+}$  while a second copper ion is coordinated by His64 and His4.

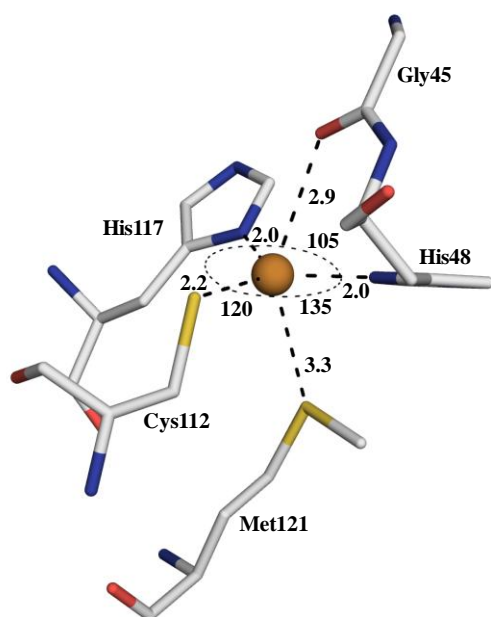
Stimulated by these observations we were interested to transform this apparently weak secondary metal ion recognition site into a solid copper binding site. Introduction of a Cu center at the surface of CA II could establish new catalytic opportunities since  $\text{Cu}^+$  is able to catalyze a variety of chemical reactions, among these the 2+3 cycloaddition of an azide with an alkyne. The reaction results in the formation of a triazole. In previous studies based on cocrystallization and soaking experiments with  $\text{Cu}^+$  or  $\text{Cu}^{2+}$  solutions we were not able to reliably achieve full metal ion population at this binding site. The accommodated metal ion showed only weak coordination and partial occupation. In order to generate a firm catalytic site we decided to design by appropriate introduction of novel residues an artificial Cu center in the proximity of His64 and Trp5.

Different approaches were successfully employed in literature for the design of artificial metal ion centers in the binding pocket of proteins. Schwarz et al. have shown that the

introduction of histidyl residues at the antagonist-binding site of the human NK-1 receptor gradually converts it into a high-affinity metal ion binding site (Elling, Nielsen et al. 1995). This procedure could be used as a general tool in structural and functional characterization of helix-helix interactions in G-protein-coupled receptors. Furthermore, by site-directed mutagenesis single amino acids or even entire turns were exchanged or introduced into the target protein (Regan and Clarke 1990; Klemba, Gardner et al. 1995; Pinto, Hellinga et al. 1997; Shields and Franklin 2004; Toyama, Sasaki et al. 2006). However, considering the small size of CA II it appears rather intractable to correctly predict the impact of such dramatic changes on protein folding simultaneously keeping the characteristics of this enzyme by replacing entire loops or inserting larger peptide chains. Therefore, we decided to apply a strategy based on rational design concepts. It appeared most promising to us to virtually copy the coordination geometry found in other copper coordinating proteins across to our case.

We selected the copper centers found in two well-known biological systems. As our first model reference served the so-called “type I” copper sites observed in blue copper proteins. These enzymes are involved in electron transfer, as they are capable to host the Cu ion in both oxidation states. In azurin, a prominent example, the copper is coordinated by two histidines and one cysteine residue with trigonal planar geometry exhibiting two additional weakly interacting groups in axial position (Figure 4.2) (Karlsson, Nordling et al. 1991). The additionally coordinating ligands can take important impact on the formation of the Cu center. They can stabilize certain oxidation states of the metal ion. In consequence, the redox potential of enzymes with such centers is dependent on the actual composition of the coordinating ligands (Li, Webb et al. 2004). Hellinga, Caradonna and Richards tried to introduce such a copper center into the binding site of thioredoxin from *Escherichia coli* by exchanging particular amino acids according to a computational prediction (Hellinga, Caradonna et al. 1991; Hellinga and Richards 1991). Following the computationally suggested site-directed mutagenesis into the fold of the target enzyme they were actually able to constitute a copper binding site at the anticipated place. However, the introduced artificial Cu<sup>2+</sup> center showed a coordination pattern deviating from the original design as they found two histidines and two additional carbonyl groups form the binding site to chelate the

metal ion. The especially introduced cysteine residues do not participate in the metal ion coordination. These results demonstrate that computational approaches are a good starting point and provide valuable suggestions; however, they hardly reflect all the impact amino acid mutations will develop on a highly functional biological system. The study suggests that a stepwise approach as planned in our case is perhaps a bit more conservative but provides higher conceptional flexibility since information gained upon a single mutation can be considered in subsequent steps of further design.

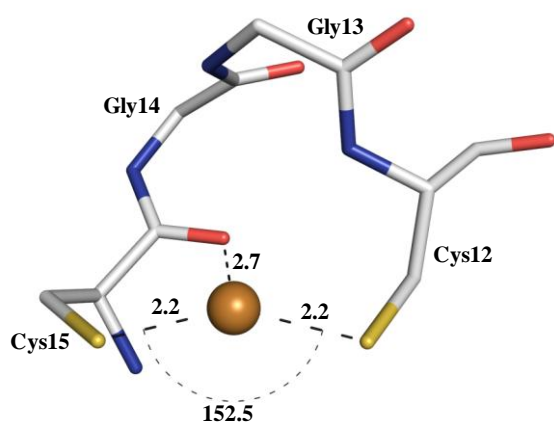


**Figure 4.2** Crystal structure of Azurin (PDB-Code: 2CCW). His48, His117 and Cys112 (C, N, O, S) form a trigonal-planar conformation coordinating the copper ion (brown sphere). Gly45 and a Met121 are located in axial positions influencing the oxidation state of the ion.

As the second reference system for our modeling considerations we selected the human metallochaperone HAH1, a small soluble protein, which delivers  $\text{Cu}^+$  ions to a target enzyme. HAH1 carries a  $\text{Cu}^+$ -ion at its surface (Klomp, Lin et al. 1997; Pufahl, Singer et al. 1997; Jordan, Natale et al. 2000; O'Halloran and Culotta 2000; Puig and Thiele 2002). This ion is coordinated by two cysteine residues at a distance of 2.2 Å each and a chelating angle of 152.5° (Figure 4.3). Both cysteines are located at the transition of a loop to an  $\alpha$ -helix and they are separated by two amino acids (Cys12-Gly13-Gly14-Cys15). In solution the copper ion appears to be coordinated by these two cysteine residues as the stoichiometry of the formed complex could be examined by EXAFS and isothermal titration calorimetry (ITC) while in the crystalline state two protein



monomers coordinate the ion via four cysteine residues (Ralle, Lutsenko et al. 2003; Wernimont, Yatsunyk et al. 2004). According to the latter structural data it appears reasonable that the introduction of two cysteine residues in similar fashion into CA II could allow creation of a copper center in the target protein. Considering the mobility of amino acids at protein surfaces we can assume a certain flexibility of the introduced residues that might assist adaptation to form the correct conformation required for metal coordination. In contrast, too pronounced residual flexibility could instead be detrimental for efficient coordination and possibly the residues will escape the planned metal chelation.



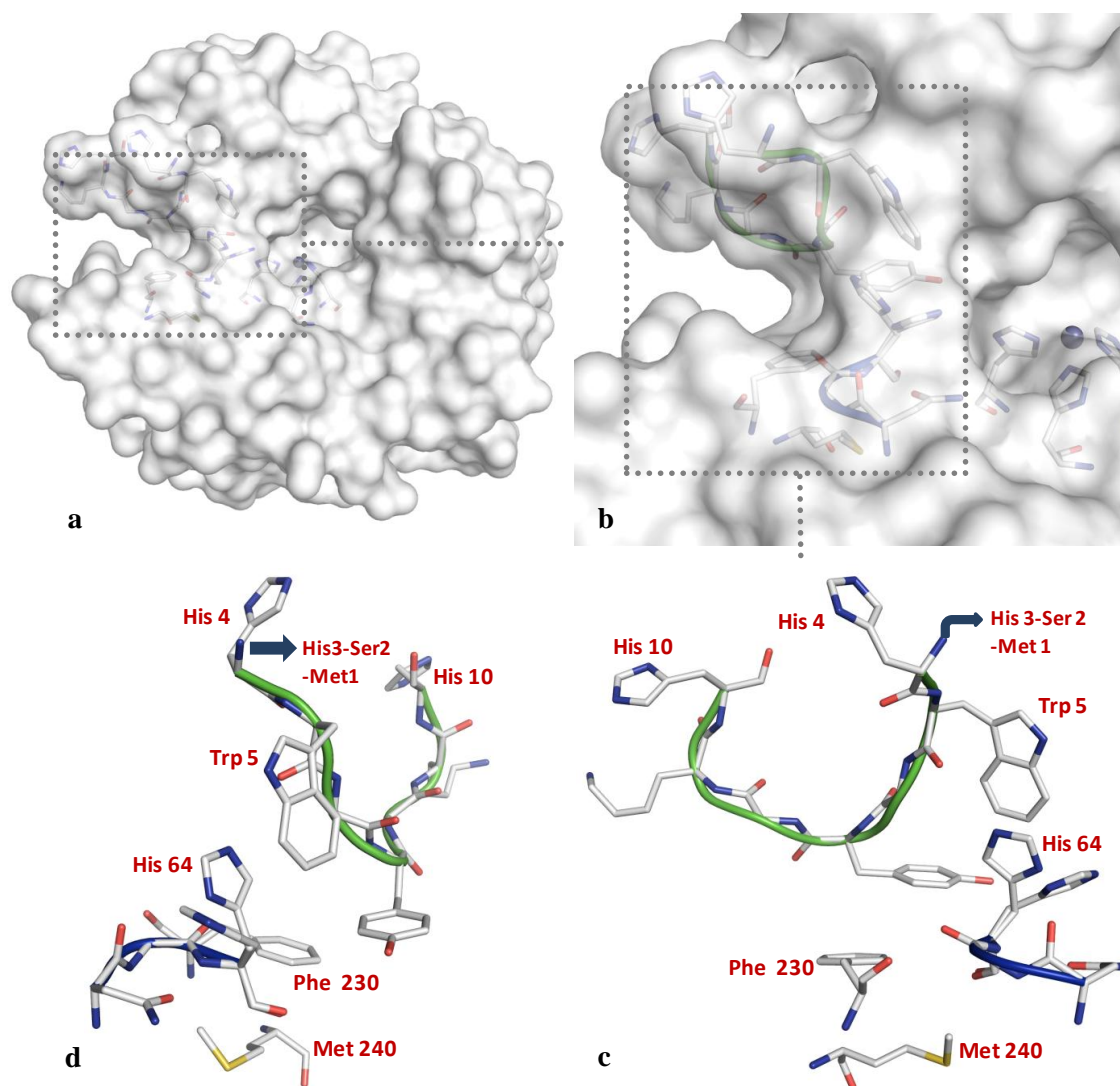
**Figure 4.3** Crystal structure of Human metallochaperone HAH1 (PDB-Code: 1TL4). The copper ion (brown sphere) is coordinated by Cys12 and Cys15 (C, N, O, S). Both cysteines are located at the transition of a loop to an  $\alpha$ -helix and are separated by two amino acids.

Here, we present the introduction of an iteratively designed artificial Cu ion binding site at the surface of CA II. We created different mutants designed according to the above-mentioned reference systems that actually show such centers. Following a stepwise approach, different mutants were subsequently created according to literature and the crystal structures obtained with our mutants. Our initial mutations were not yet competent to bind copper ions. Upon further mutations we finally obtained a copper center on the enzyme's surface. Crystal structure analysis shows also in our case that the actually obtained coordination pattern deviates from the anticipated one: a square planar geometry is formed by three nitrogen and one oxygen atom, whereas participation of the especially introduced cysteine residues is not observed. An azide group, which has been covalently tethered to Cys64 via a disulfide linker, obviously assists the copper center to form.

### 4.3 Results

In order to create an artificial metal ion binding site certain amino acids at the surface of CA II were exchanged. We started with single amino acid mutations, crystallized the mutated protein and used the obtained structural insights to plan our subsequent design. The enzyme was recombinantly produced, purified and either cocrystallized or after crystallization soaked with a  $\text{Cu}^+$  or  $\text{Cu}^{2+}$  metal ion solution, the applied concentrations ranged from 100  $\mu\text{M}$  to 5 mM. The obtained geometry was characterized by protein crystallography, which enables a detailed analysis of the coordination properties of a potentially created metal binding site.

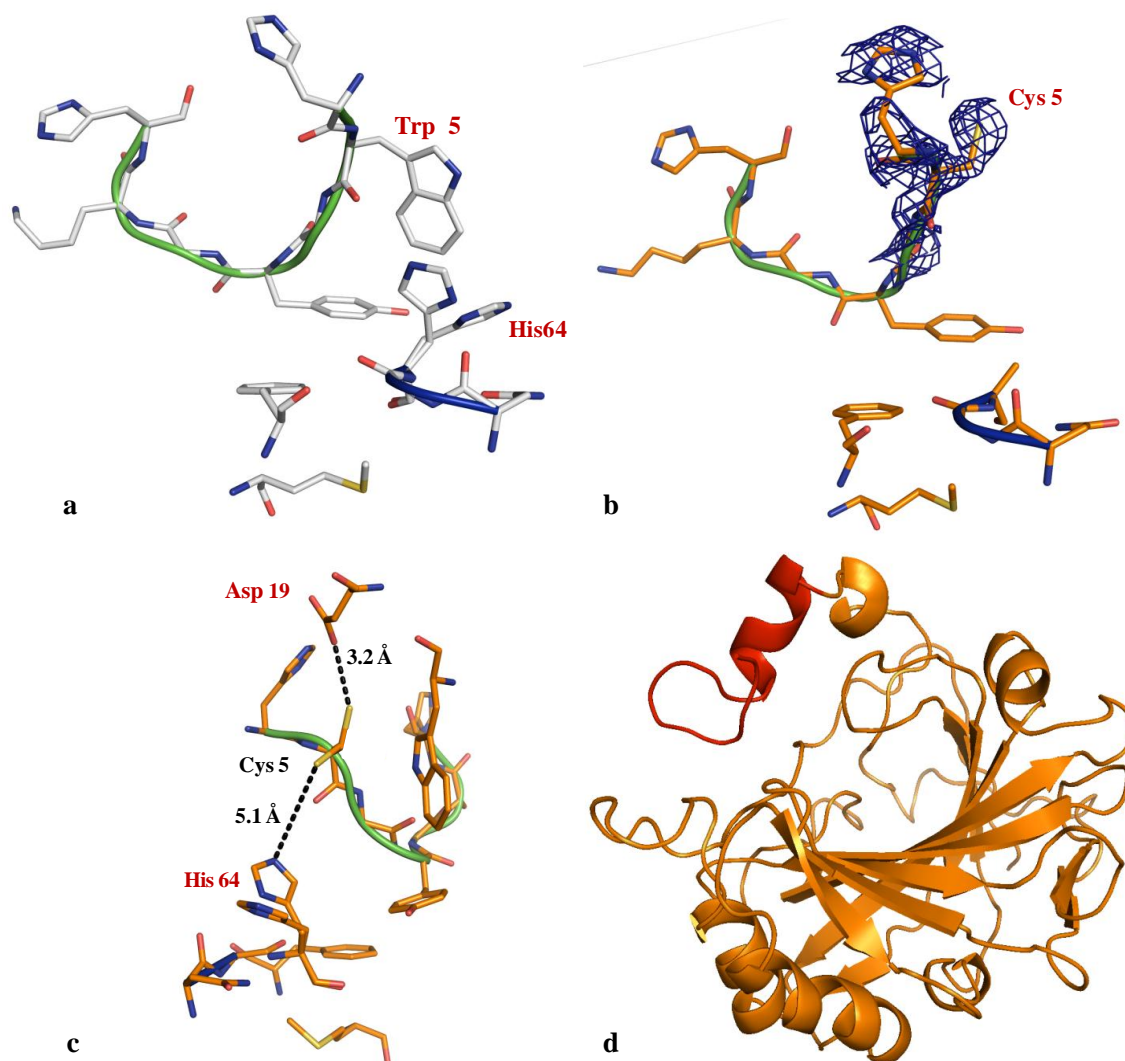
**Design of mutants:** As a first step, which amino acid mutations of CA II could be competent to create the required geometry? Figure 4.4 shows the crystal structure of native CA II (PDB-code: 3D92). Our selected target residue His64 is located at the rim of the binding pocket typically adopting two conformations important for proton shuffling. The N-terminus of CA II (green loop) adopts a conformation in the spatial vicinity of this residue. It ideally exhibits some mobility that could be favorable to assist formation of the required metal ion coordination geometry. In type I copper centers two histidine and one cysteine residue establish a trigonal planar Cu coordination site. As cysteine residues are not only present in this Cu-binding site but also involved in that of our second reference protein we attempted introduction of such a residue as initial step. Spatially close to His64, CA II exhibits two additional histidine residues, His3 and His4 (Figure 4.4c,d). Exchanging one of these residues by the favored cysteine should allow copper coordination. Furthermore, His10, also in the flexible part of the N-terminus, could be ideally located for coordination. Apart from the residues desired to form a trigonal planar coordination geometry, Met240 is found in a position to potentially interact as a coordination partner with a putative metal site. However, its accessibility will depend on the orientation of Phe230 which occupies the space between His64 and Met240. Finally, we decided to select Trp5 as a first candidate for cysteine introduction. Thereby His64, His3 and His4 could experience rearrangements of the N-terminus to adopt a geometry competent for metal coordination.



**Figure 4.4** Crystal structure of carbonic anhydrase II. (a) Overview of carbonic anhydrase II. The solvent accessible surface of the protein is represented in white in (a) and (b) (C, N, O, S). (b) The area for Cu-center introduction in surface representation from a close up view. (c) The area for Cu-center introduction in stick representation. The backbone of the flexible N-terminus is represented by a green cartoon, of the Cys64 region by a blue cartoon. (d) (c) rotated by 90°.

Similar considerations propose exchange of His64 by a cysteine residue. Such mutation would reduce the distance of the anticipated copper center to the binding pocket of CA II. This center was planned to host possible reagents for further chemical reactions involving the newly created copper center. Considering the human metallochaperone HAH1 the described double cysteine mutant, exchanging His64 as well as Trp5 was also planned as a designed mutant.

**Structural investigation of CA II-W5C:** The W5C mutation was performed and its crystal structure shows Cys5 to be rotated away from His64 (Figure 4.5).



**Figure 4.5** Crystal structure of CA II and CA II-W5C. (a) Crystal structure of CA II. The area for Cu-center introduction is shown in stick representation (C, N, O, S). The backbone of the flexible N-terminus is represented by a green cartoon, of the His64 region by a blue cartoon. (b) Crystal structure of CA II-W5C (C, N, O, S). The  $2F_o - F_c$  map for the mutated amino acid is displayed at a  $\sigma$  level of 1.0 as blue mesh. (c) (b) rotated by  $90^\circ$ . Cys5 is located at a distance of 3.2 Å to Asp19. A rotamer of Cys5 experiences a distance of 5.1 Å to His64. (d) Crystal structure of CA II-W5C after Cu soaking in cartoon representation. The N-terminus experiences high mobility which is indicated by weak or missing electron density. The residues that are missing after the copper soaking are shown in red.

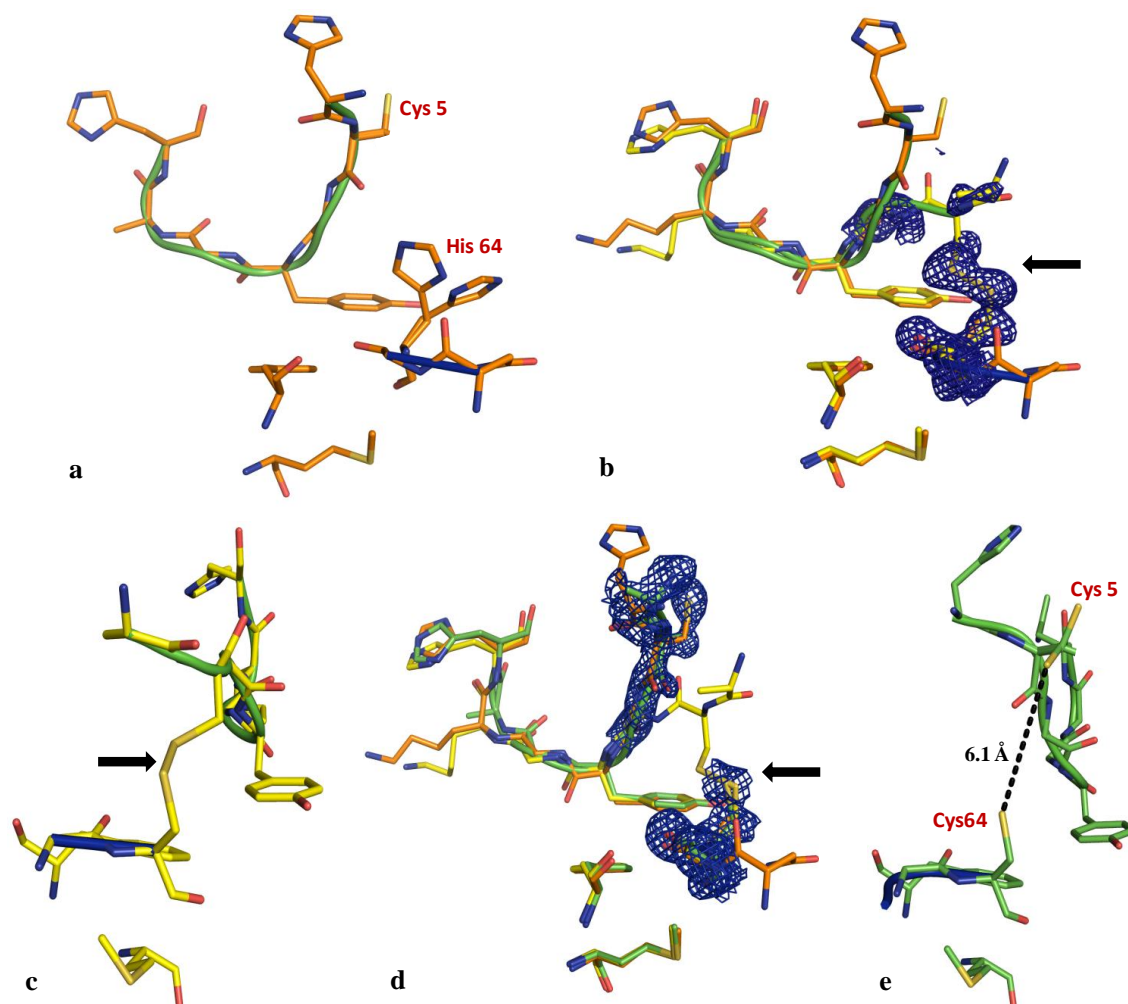
The sulfur atom is oriented towards Asp19 with a distance of 3.2 Å. The geometry of His3 cannot be determined as no difference electron density is observed for this residue. Similarly, the preceding two amino acids experience very high flexibility and can therefore also not be defined in crystal structures. The mutant CA II-W5C suggests a

spatial arrangement of the involved amino acid which seems capable to form the desired copper center. A possible rotamer of Cys5 would experience a distance of 5.1 Å to His64. The residual mobility of the N-terminus could assist formation of the coordination site. A trigonal-planar coordination with two histidine and one cysteine residue known from *Pseudomonas aeruginosa* azurin seems therefore achievable (Karlsson, Nordling et al. 1991). However, after soaking this mutant with Cu<sup>+</sup> and Cu<sup>2+</sup> solutions (10 mM) binding of Cu ions to CA II-W5C was not observed. Moreover, the crystal structure does not show any density for the N-terminal residues (Figure 4.5d). The amino acids 1 to 19 are disordered and seem to adopt multiple conformations. The weakly defined electron density prohibits any identification of single amino acid conformations.

**Structural investigation of CA II-W5C-H64C:** Taking HAH1 as a reference, the N-terminal flexibility suggests introduction of a second cysteine mutation within this area of the protein. Therefore, His64 was exchanged as the next step (Figure 4.6a). The crystal structure of CA II-W5C-H64C reveals extensive conformational changes at the protein surface (Figure 4.6b, PDB-code: 3M1W). The N-terminus is shifted towards Cys64 as a disulfide bond is formed between Cys5 and Cys64. This disulfide bond is apparently fully established as it is indicated by a very well resolved electron density. The amino acids His4 and His3 experience high mobility as they are not visible in the diffraction pattern. Residues Tyr7 and the following ones are not affected by this mobility and adopt a conformation commonly found in many structures.

From HAH1 it is known that the two cysteine residues coordinate the copper ion forming an angle of 152.5°. The observed terminal mobility of the CA II-W5C mutant provides the necessary prerequisite to accommodate Cu with reasonable affinity. Both cysteines should be able to arrange in the required manner. In order to verify this assumption the disulfide bond was reduced prior to Cu soaking. In previous thermostability measurements we were able to determine the TCEP and DTT concentration required for the disulfide reduction. In such an assay, the enzyme is gradually heated until denaturation of the protein is indicated. Shifts in the "melting" temperature correlate with enzyme stability. The experiment showed that in case of CA

II-W5C-H64C 1.5 mM TCEP and 5 mM DTT are sufficient to reduce the disulfide bond (data not shown). The protocol to grow protein crystals involves several steps starting with protein expression, purification and concentration to finally crystallization. In each step disulfide bond reduction and copper center constitution were tested, separately.



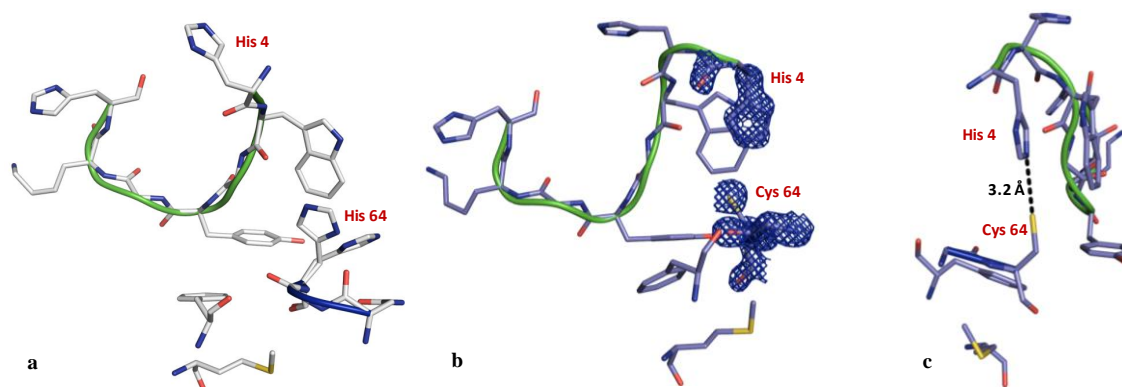
**Figure 4.6** Crystal structure of CA II-W5C and CA II-W5C-H64C. (a) Crystal structure of CA II-W5C. The area for Cu center introduction is shown in stick representation (C, N, O, S). The backbone of the flexible N-terminus is represented by a green cartoon, of the His64 region by a blue cartoon. (b) Crystal structure of CA II-W5C-H64C (C, N, O, S) superimposed with CA II-W5C (C, N, O, S). The  $2F_o - F_c$  map for the mutated amino acids is displayed at a  $\sigma$  level of 1.0 as blue mesh. The formed disulfide bond is indicated by an arrow. (c) b rotated by  $90^\circ$ . (d) Crystal structure of CA II-W5C-H64C (C, N, O, S) after disulfide bond reduction superimposed with the closed form (C, N, O, S) and CA II-W5C (C, N, O, S). The N-terminus readopts the original orientation. (e) (d) rotated by  $90^\circ$  (C, N, O, S). A rotamer of Cys5 is located at a distance of 6.1 Å to Cys64.

Each thus treated protein charge was crystallized and tested for Cu center formation by X-ray crystallography. None of the produced crystalline samples showed the desired formation of the copper center. As the crystallization process of these CA II mutants

takes between weeks to months, the DTT concentration supplemented to the crystallization drops was obviously not sufficient to keep the thiole groups in the reduced state until dataset collection. Most likely the initially ruptured disulfide bond was reformed possibly as a consequence of small traces of oxygen present during the crystallization experiments. As an alternative we tried to soak the reducing agent and copper ions, however also these soaking experiments failed to reduce the disulfide bond. In the crystallization protocol of the wild type *p*-chloromercuribenzoic acid is added to the crystallization drop to significantly increase crystal growth in all dimensions. Considering that both cysteine residues might be capable to coordinate the mercury ion and thus could possibly hamper Cu coordination we performed all crystallization trials in the absence of mercury. Unfortunately, under these conditions protein crystals grow rather slow, remain small and lack the required stability over a sufficient soaking time. We also performed our experiments under anaerobic conditions but no formation of a Cu center could be detected by crystallography. We also dialyzed the protein under anaerobic conditions, without success. Possibly, the applied Cu concentrations were too harsh and the enzyme was denaturated to some degree. Finally, after repeating soaking experiments and crystal growth for six months we revealed a crystal used for data collection that showed an opened disulfide bond (Figure 4.6d,e, PDB-code: 3M1Q). The N-terminus adopts the previously seen conformation. Cys5 and Cys64 are separated by a distance of 6.1 Å. Nevertheless, considering the inherent mobility of the residues of this mutant the formation of a copper center appeared feasible. Unfortunately, additional experiments with varying soaking conditions remained unsuccessful. We therefore considered further mutations.

**Structural investigation of CA II-H64C:** CA II-H64C shows the same backbone conformation as the wild type (Figure 4.7b, PDB-code: 3M5S). Interestingly enough, the newly introduced cysteine residue induces a rotation of the proximal His3. The imidazole moiety shows a very well defined electron density and adopts a conformation placing its side chain in close distance to Cys64 (3.2 Å). After crystals had grown to appropriate size, soaking experiments with Cu<sup>+</sup> and Cu<sup>2+</sup> solutions were performed. However, similar to the other mutants no constitution of a copper center could be detected.

In addition to the above described Cys and His mutants we performed additional site-directed mutagenesis to introduce the potentially Cu coordinating residues His, Cys and Met at various positions. Unfortunately, for the CA II-W5H, CA II-W5C-H64M, CA II-G63C and CA II-H64M mutants no crystallization conditions could be found and thus no insights into the putative formation of a geometry competent to coordinate a copper ion could be obtained.

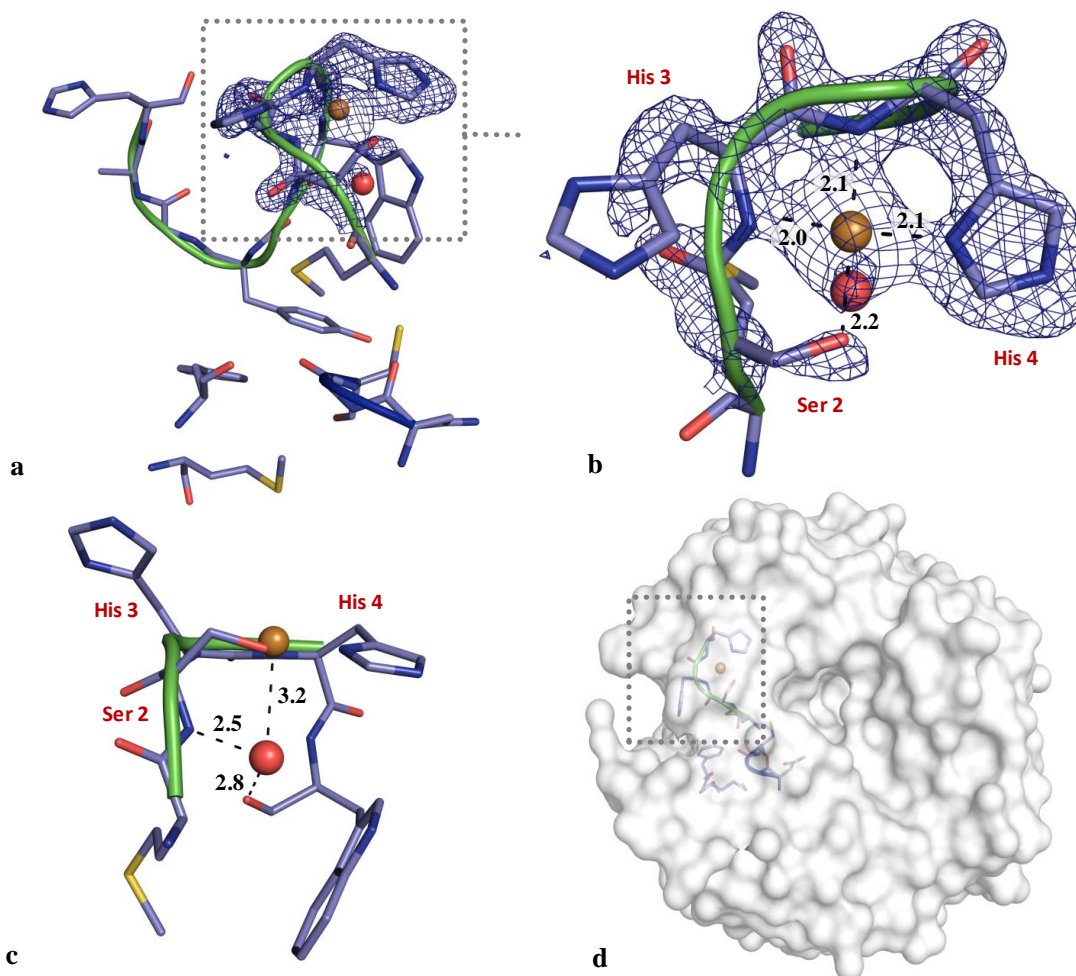


**Figure 4.7** Crystal structure of CA II and CA II-H64C. (a) Crystal structure of CA II. The area for Cu-center introduction is shown in stick representation (C, N, O, S). The backbone of the flexible N-terminus is represented by a green cartoon, of the His64 region by a blue cartoon. (b) Crystal structure of CA II-H64C. The  $2F_o - F_c$  map for the mutated amino acid and His3 is displayed at a  $\sigma$  level of 1.8 as blue mesh. His3 is rotated towards Cys64. (c) (b) rotated by  $90^\circ$ . His3 is located at a distance of 3.2 Å to Cys64.

We therefore reconsidered our initial design concept. The introduction of the copper center has been attempted to generate an additional catalytic site to bring reagents such as an azide and an alkyne to reaction. During the reaction the azide is assumed to show transient coordination with the copper center. By this coordination step a possible stabilization of the copper center can also be anticipated. In order to profit from this strategy of an additional stabilization of the designed copper coordination center we tethered different azide components via disulfide linkage to the introduced cysteine moieties in our mutants. Admittedly more by serendipity than by rational design concepts we succeeded to determine a crystal structure with an actually formed copper center. In this structure an azide was tethered to the CA II-H64C mutant followed by the exposure to  $\text{Cu}^+$  ions. This exposure was primarily attempted to initiate reaction with an alkyne component added to the specificity pocket of CA II. Crystal structure analysis of thus treated crystals surprisingly shows the desired copper center to be formed at the N-terminus of the mutant (Figure 4.8). Possibly supported by the additional presence of



the azide, the Cu center has been established. Interestingly, Ser2, His3 and His4 become well ordered in this structure and arrange with square planar coordination geometry around a copper ion exhibiting distances from 2.0 to 2.2 Å (Figure 4.8b).



**Figure 4.8** Crystal structure of CA II-H64C. (a) Artificially introduced copper binding site in CA II-H64C. The area of the Cu center is shown in stick representation (C, N, O, S). The  $2F_o - F_c$  map for the coordinating amino acids is displayed at a  $\sigma$  level of 1.0 as blue mesh in (a) and (b). The backbone of the flexible N-terminus is represented by a green cartoon, of the Cys64 region by a blue cartoon. (b) Cu binding site from a close-up. The metal ion is coordinated by His4, His3 and Ser2 at distances between 2.0 Å to 2.2 Å. (c) b from a different orientation. (d) Overview of CA II-H64C. The Cu binding site is located at the N-terminus on the left hand side of the binding pocket. The solvent accessible surface of the protein is schematically represented in white.

Remarkably this coordination pattern matches well with those known for type II copper centers, where the metal ion is coordinated by nitrogen and oxygen atoms in a square planar geometry. Residues Met1, Ser2 and His3 are usually not visible in the electron density of CA II crystal structures. Since the crystals were soaked with a  $\text{CuSO}_4$  solution containing five-fold excess of ascorbic acid we assume that a  $\text{Cu}^+$  ion is

coordinated by Ser2, His3 and His4. Additionally, a water molecule contributes to the coordination geometry and occupies a fifth site in axial position. The colorless crystals provide additional evidence for the presence of  $\text{Cu}^+$ , since enzyme crystals, such as of azurin show a blue color after  $\text{Cu}^{2+}$  soaking. Performing a CSD database search for known Cu coordination reveals that the observed arrangement matches well with a square planar geometry about a Cu ion and argues in favor of a  $\text{Cu}^+$  center.

Currently it is not obvious how the azide moiety might have contributed to the formation of the Cu binding site. We tested further azides with different linker length. For all azides the metal center was formed under the described conditions. We assume that the azide plays an important role in the pre-formation of the metal center. Possibly, as the azide finally engages in triazole formation it is only transiently available to contribute to the coordination sphere. For this it could adopt a minor conformation placing the terminal azide nitrogen in an axial position. Once the coordination site has been formed the involved amino acids seem to establish a stable arrangement. We have repeated this procedure with numerous protein crystals and in all cases we revealed the metal ion center to be formed after Cu soaking. This supports the assumption that a stable metal center has formed providing the properties to be available for further experiments. The formation of this center in solution has to be confirmed and its catalytic properties will be investigated in subsequent studies.

#### **4.4 Conclusion**

In summary, we have created more by serendipity than by rational design a Cu binding site at the surface of carbonic anhydrase II next to its catalytic zinc site. Focusing on known Cu binding sites observed in other structures we tried to imitate the required binding geometry for such a metal ion. Therefore, single amino acids were exchanged by site-directed mutagenesis. If possible, for each produced mutant a crystal structure analysis was performed to obtain insights into the revealed structural modifications. Subsequently, each mutant was soaked in crystalline state with a  $\text{Cu}^+$  and  $\text{Cu}^{2+}$  solution and the achieved modifications were investigated again by protein crystallography. Numerous mutants were created and tested. In case of CA II-W5C a conformational change of the N-terminus could be observed upon Cu soaking. This mobility seems to

be an important prerequisite for metal center constitution. However, optimization of the individual N-terminal residues remained impossible due to ill-defined electron density. We also attempted to introduce a copper center similar to the one observed in HAH1, by placing two mutually facing cysteine residues in CA II-W5C-H64C. The initial crystal structure of this mutant revealed a disulfide bond between Cys5 and Cys64. This disulfide bridge could be reduced by DTT soaking; however, a Cu center could not be introduced upon Cu supplementation.

Finally, the formation of a copper center could be achieved more by serendipity than design obviously temporarily supported by an azide substituent tethered to Cys64 in the CA II-H64C mutant. The geometry of this metal binding site is related to the type II copper sites composed exclusively by oxygens and nitrogens as coordinating ligand atoms.



## **5 Stereo- and Regioselective Azide/Alkyne Cycloadditions in Carbonic Anhydrase II via Tethering Followed by Protein Crystallography and Mass Spectrometry**

### **5.1 Introductory Remarks**

The following chapter has been submitted to the scientific journal for publication. In the following the original text is depicted. The introduction is elaborately described to provide a detailed view of click chemistry. In addition, extensive experimental work concerning HPLC-MS analysis is presented at the end of the results section.

The present study was accomplished in cooperation with Dong Sun and Nicola Sandner, group of Prof. Koert, and Uwe Linne, group of Prof. Marahiel from the Chemistry department, Philipps University Marburg. Dong Sun collaborated in the entire project, performed the synthesis of all compounds applied in the experiments and supported in the design of the building blocks. Nicola Sander joined the project in March 2009 and contributed to the refinement and deposition of crystal structures and to the development of the HPLC-MS experiments. Uwe Linne performed sample purification and mass spectrometric analysis of the HPLC-MS experiments.

### **5.2 Abstract**

A carbonic anhydrase II mutant has been created which can be employed for the selective synthesis of triazoles via the Huisgen reaction from azides and alkynes. The azide is covalently attached by tethering to a mutationally introduced cysteine residue on the surface of CA II. The alkyne has been fixed reversibly to the protein surface via a sulfonamide anchor which coordinates the zinc ion at the original catalytic center. The 2+3 cycloaddition is enormously accelerated by the thus achieved spatial preorganization of both building blocks. The triazole formation has been monitored by protein crystallography and HPLC-MS which allows the successive observation along the reaction time. The dependence of product formation on the applied reaction conditions was studied in detail. It can be shown that the CA II surface takes determining influence on product formation while  $\text{Cu}^+$  ions even accelerate the turnover. Apart from optimization of product yields we were able to develop building

blocks for regio- and stereoselective triazole synthesis resulting in the formation of a stereochemically pure *1,5-S*-triazole.

### 5.3 Introduction

Reactions, classified as click chemistry, are expected to be modular, stereospecific, wide in scope, excellent in yields and will generate only safe byproducts that are easily removable by non-chromatographic methods. Furthermore, the applied reaction conditions have to be simple, the starting materials and reagents readily available and at best water can be used as solvent (Kolb, Finn et al. 2001). The synthesis of *1,2,3*-triazoles via the Huisgen reaction has been established as one of the best-suited transformations out of the entire repertoire of organic reactions to meet the criteria of click chemistry (Huisgen, Szeimies et al. 1967). This cycloaddition based on azides and alkynes facilitates the generation of large triazole libraries, particularly as a large variety of alkyne and azide components can be introduced into a broad range of molecular skeletons. However, for more than 40 years the impact of this reaction was regarded as minor due to low selectivity yielding in mixtures of *1,4*- and *1,5*-regioisomers (Tron, Pirali et al. 2008). As further drawback these cycloadditions require high activation energy often resulting in rather slow turn-over. Therefore either elevated temperatures and/or pressures were applied (Zhang, Chen et al. 2005). To accelerate the reaction rate, different strategies have been developed for alkyne activation. The addition of electron-withdrawing groups to the alkyne, the reaction of azides with alkynes or the activation by metal ions such as  $\text{Cu}^+$  and  $\text{Ru}^{2+}$  are just a few features to mention (Zhang, Chen et al. 2005; Berkel, Dirks et al. 2008; Shi, Waldo et al. 2008). Particularly, the latter ions have productively changed click chemistry since they allow the selective synthesis of *1,4*- and *1,5*-substituted triazoles. The addition of  $\text{Ru}^{2+}$  ions to Huisgen reaction conditions leads to the regioselective synthesis of *1,5*-disubstituted triazoles in acceptable reaction time (Zhang, Chen et al. 2005).  $\text{Cu}^+$  shows an even higher impact on the Huisgen reaction. Sharpless and coworkers demonstrated that the rate of cycloaddition between azides and alkynes can be accelerated  $\sim 10^6$ -fold using catalytic amounts of  $\text{Cu}^+$ , leading to *1,4*-triazole products (Rostovtsev, Green et al. 2002; Tornøe, Christensen et al. 2002). Apparently, the mechanism of this click reaction is well understood. Most likely the reaction does not proceed through a concerted mechanism, as frequently discussed for

cycloadditions. Instead, a stepwise reaction pathway is favored, as confirmed by kinetic data and molecular modeling and supported by simulations of activation barriers (Hein, Liu et al. 2008). These studies suggest that up to two  $\text{Cu}^+$  ions are involved in the catalysis. In the first step, subsequent to the deprotonation of the terminal alkyne carbon, the two  $\text{Cu}^+$  ions and the acetylide form a complex. Thereafter the innermost nitrogen of the azide displaces one of the coordinating ligand atoms from the second  $\text{Cu}^+$  ion thus activating the terminal nitrogen for nucleophilic attack towards the alkyne. After cyclization and protonation, the catalyst releases the *1,2,3*-triazole product (Himo, Lovell et al. 2004; Rodionov, Fokin et al. 2005; Bock, Hiemstra et al. 2006; Rodionov, Presolski et al. 2007). The isolation of a crystalline  $\text{Cu}^+$  triazolide strongly supports the existence of such click-intermediates (Nolte, Mayer et al. 2007). Apart from  $\text{Cu}^+$  ions several other copper derivatives serve as catalysts with deviating reaction profiles in the Huisgen reaction; e.g. Cu colloids, Cu powder, Cu shavings or  $\text{CuSO}_4$ /ascorbate. For these systems a detailed mechanistic analysis is still of major interest. Remarkably, for all copper catalysts mentioned above an incubation, e.g. for  $\text{Cu}^{2+}$ /ascorbate of up to 5 hours, is required before any clickproduct can be detected, however no satisfactory explanation for this observation can yet be provided (Pachón, van Maarseveen et al. 2005).

With respect to pharmaceutical research click chemistry provides novel opportunities, e.g. for polymeric micelles, nanoscale delivery systems and DNA labeling (Hein, Liu et al. 2008). Furthermore, libraries produced by Click chemistry can be tested in High Throughput Screenings on pharmaceutically interesting targets. However, the need of a copper catalyst for the Huisgen reaction is detrimental to pharmaceutical applications as possibly remaining copper impurities might create toxic side effects, such as hepatitis, neurological disorder, kidney diseases, and Alzheimer's disease (Wang and Guo 2006). In light of such risks significant effort has been spent to develop metal-free click reactions with the required turn-over and stereochemistry.

It has recently been shown that the protein environment of an enzyme can function as such an activator for the Huisgen reaction catalyzing triazole formation without the need of supplementary metal ions (Krasinski, Radic et al. 2005; Mocharla, Colasson et al. 2005). This opportunity is particularly supported by the fact that alkynes and azides

are essentially inert and chemically orthogonal to most biological and organic conditions, including highly functionalized biological molecules, e.g. as given in the binding pockets of enzymes. Cycloaddition reactions can be accelerated tremendously once the reacting building blocks are brought into favorable spatial vicinity (Mock, Irra et al. 1983; Mocharla, Colasson et al. 2005). This strategy is exploited in the so called target-guided synthesis (TGS) using a protein environment as constraint. The enzyme is incubated with a mixture of diverse building blocks which exhibit graded affinities to the site of mutual interaction. This site will thus select mainly those fragments that experience highest affinity, in consequence, the product with the strongest binding affinity will overwhelmingly be found. The enzyme operates as a kind of blueprint, filtering out the most potent inhibitor from a large number of possible products. Applying this TGS approach, Sharpless et al. showed in several studies that enzymes such as acetylcholinesterase (AChE) or carbonic anhydrase II (CA II) can be exploited successfully for the assembly of triazoles (Krasinski, Radic et al. 2005; Mocharla, Colasson et al. 2005).

The goal of the present study was to examine the azide/alkyne cycloaddition in a protein environment by integrating a tethering approach into the target-guided synthesis concept (Erlanson, Braisted et al. 2000). The tethering method facilitates to introduce fragments with weak affinity into the binding pocket of proteins by covalently attaching the ligand to a cysteine residue close to or within the active site. This system should be appropriate to keep track of the reaction geometry by crystallographic diffraction techniques to obtain reliable insight in the produced stereochemistry. For this purpose we selected the binding pocket of carbonic anhydrase II. By covalently attaching the azide component to a mutationally introduced cysteine residue next to the binding pocket and linking the alkyne component via an appropriate sulfonamide anchor to coordinate efficiently the active site zinc ion we were able to monitor the reactions accomplished in the binding pocket by denaturing HPLC-MS analysis performed in parallel to protein crystallography. Qualitative and quantitative considerations are possible as the click reaction system is covalently attached and upon triazole product formation the protein mass increases by the mass of the added alkyne. In particular, we were interested in the influence of the enzyme binding pocket on the yield and the rate

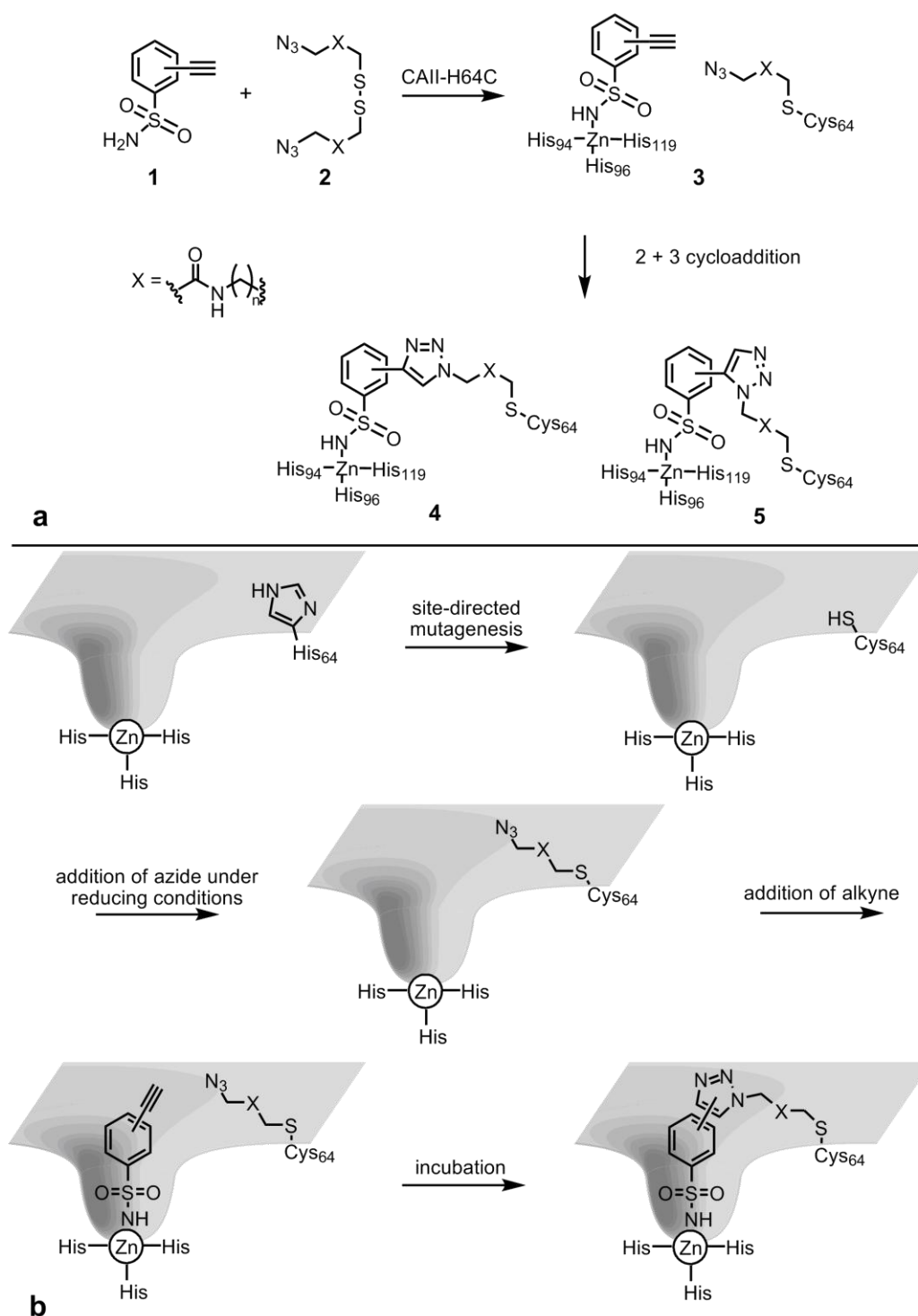


of triazole formation (1), the regioselectivity (2) and the stereodiscrimination of chiral starting materials (3).

Large scale test series with different reaction components showed that already the CA II-H64C mutant is able to induce triazole formation. However, the presence of  $\text{Cu}^+$  ions further accelerates the reaction whereas other ions such as  $\text{Ru}^{2+}$  or  $\text{Zn}^{2+}$  have only minor effects. We decided to monitor the influence  $\text{Zn}^{2+}$  takes on this system as a control as it is almost impossible to create Zn-free conditions in biological systems. Furthermore, reactant-dependent the degree of product formation along with the produced regio- and stereoselectivity could be recorded. Our study provides the perspective to design selective triazoles catalyzed by the protein environment since site-directed mutagenesis allows creating a specific protein surrounding that drives triazole formation towards a desired product.

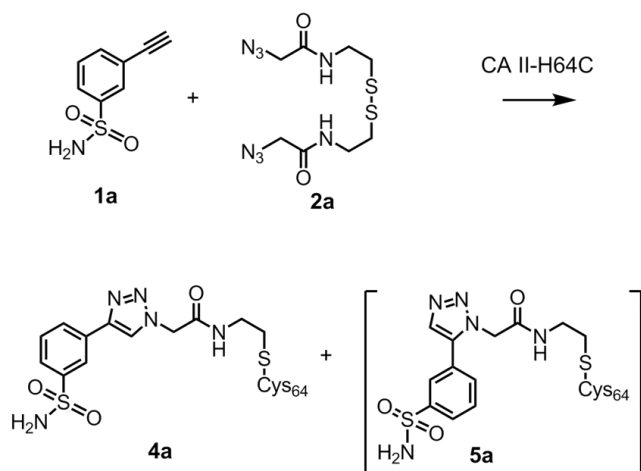
## 5.4 Results and Discussion

The general concept for a tethered approach to study azide/alkyne cycloaddition reactions in the binding pocket is shown in Scheme 5.1. An arylsulfonamide bearing a substituent with a terminal alkyne group **1** and an azide-functionalized disulfide **2** react with CA II-H64C to form the complex **3**. In **3** the alkyne component is bound to a lipophilic pocket and reversibly coordinated via its sulfonamide anchor to the zinc ion in the original catalytic center of CA II (His94, His96, His119). The azide component is connected covalently via a disulfide bond to Cys64. After 2+3 cycloaddition the formed triazole remains covalently attached to Cys64. Principally, the two regioisomers, the *1,4*-triazole **4** and the *1,5*-triazole **5** can be formed. This strategy facilitates formation of a complex with both reaction components in close vicinity and allows to record successive situations along chemical transition by crystallography. Furthermore, this setup allows tracing the reaction turn-over by HPLC-MS under denaturing conditions.



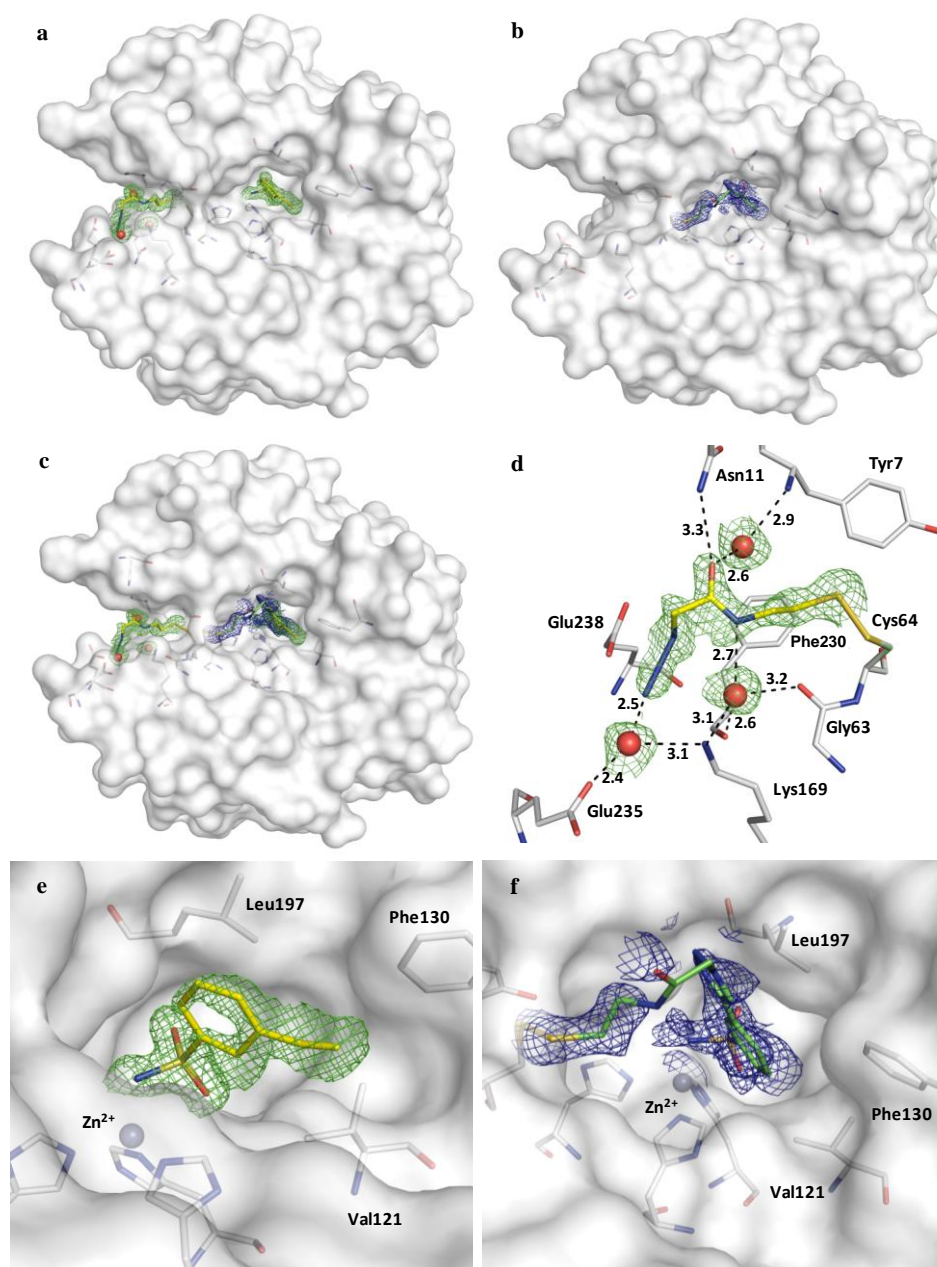
**Scheme 5.1 Tethered enzyme mediated alkyne-azide cycloadditions.** (a) Reaction of the alkyne component **1** and the azide component **2** with CA II-H64C yields the complex **3** which results in the formation of the tethered triazoles **4** and **5**. (b) By site-directed mutagenesis histidine 64 is replaced by a cysteine residue. The azide is covalently attached to this cysteine residue under mild reducing conditions. The alkyne coordinates with its sulfonamide anchor to the active site Zn<sup>2+</sup> ion. The triazole is formed upon incubation.

**Initial Experiments:** Molecular modeling considerations suggested **1a** and **2a** as first prospective candidates for azide and alkyne starting materials (Scheme 5.2).



**Scheme 5.2** Tethered cycloaddition (**1a** + **2a** + CA II-H64C → **4a**).

Both compounds were incubated with CA II-H64C. The formed ternary complex could be structurally characterized by crystallography (Figure 5.1a,d,e). The alkyne immobilized via the coordination of the sulfonamide to the zinc ion accommodates its ethynyl substituent into a small hydrophobic crevice flanked by Phe130, Val121 and Leu197 (Figure 5.1e). The azide component attached covalently via a propylene disulfide linker to Cys64 occupies a surface exposed hydrophilic canyon next to the exit of the catalytic site of CA II. Both components show well-defined electron density indicating only minor residual mobility. The terminal group of the azide component establishes an extensive hydrogen-bond network mediated by three water molecules with Tyr7, Asn11, Gly63, Lys169, Phe230 and one of the terminal oxygens of Glu235 (Figure 5.1d). Furthermore, the azide moiety lies in close vicinity to the terminal oxygens of Glu238 indicating a possible interaction. More detailed considerations about the established interaction pattern are impossible due to unresolved protonation states. With the cycloaddition components **1a** and **2a** no triazole formation was observed in the crystalline state at room temperature. We subsequently treated the protein crystals (CA II-H64C + **1a** + **2a**) with solutions containing Cu<sup>+</sup> ions at 18 °C to induce triazole formation. Crystal structure analysis performed with a crystal from the accordingly treated batch after 48 hours revealed clear changes in the difference electron density (Figure 5.1b,c,f).



**Figure 5.1** Crystal structure of (**1a** + **2a** + CA II-H64C → **4a**). Distances are shown in Å (**a**) Ligand geometries of **1a** and **2a** (protein: C, N, O, S; ligand: C, N, O, S) in the cocrystal structure with CA II-H64C. The  $F_o - F_c$  omit maps for the reactants are displayed in (a) and (c) - (e) at a  $\sigma$  level of 2.0 as green mesh (before cyclization). The solvent accessible surface of the protein is schematically represented in white in all images. Both building blocks bind with distinct orientation from each other to CA II-H64C, not facing each other. (**b**) Ligand geometry of the cyclization product (protein: C, N, O, S; ligand: C, N, O, S) formed out of **1a** and **2a** individually cocrystallized with CA II-H64C and incubated with  $\text{Cu}^+$ -solution for 48 hours. The  $2F_o - F_c$  map for the cyclization product is displayed in (b) and (c) at a  $\sigma$  level of 1.0 as blue mesh (after cyclization). (**c**) Ligand geometry of the click product superimposed with the reactants **1a** and **2a** before cyclization. (**d**) Ligand geometry of **2a** in the cocrystal structure with CA II-H64C. The azide is fixed to the surface by an H-bond network to Tyr7, Asn11, Gly63, Lys169, Phe230, Glu235, Glu238 mediated by three water molecules. (**e**) Ligand geometry of **1a** in the cocrystal structure with CA II-H64C. The sulfonamide adopts the typical conformation, coordinating to the  $\text{Zn}^{2+}$  ion, while the alkyne moiety lies within a lipophilic pocket formed by Phe130, Val121 and Leu197. (**f**) Ligand geometry of the click product **4a** seen from a close-up view.

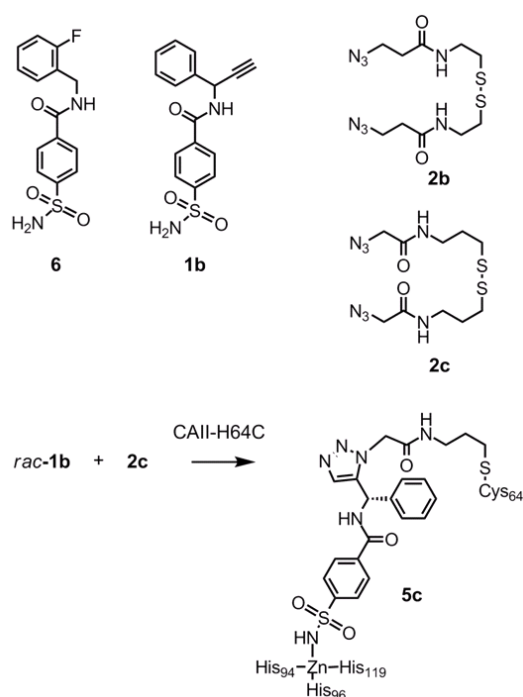
Even though the molecular arrangement detected in the first structure clearly argues against a geometry competent to execute the planned triazole formation, the electron density in the crystal of the second structure shows the alkyne component to be rotated by 180° now facing toward the azide component. The latter azide component experiences even larger rearrangements now orienting its side chain towards the rotated alkyne component. Even though the difference electron density is only weakly defined next to the region where the novel triazole ring should have assembled to the likely generated *1,4*-click product, the trace of the electron density agrees with the geometry of this product. We assume that enhanced residual mobility of the central part of the reaction product which is not in close surface contact with the protein and incomplete product formation across all unit cells of the protein crystal are responsible for the partially defined difference electron density.

To further characterize the anticipated product formation we applied HPLC-MS analysis under denaturing conditions. To increase the reaction rate we raised the temperature from 18 °C applied for our crystallographic studies to 37 °C in solution. After incubation for 48 hours in the presence of Cu<sup>+</sup> we subjected the protein to mass analysis. The enzyme mass was determined after purification over a size exclusion column to remove any protein-unbound compounds and therefore to prevent non-specific background reactions during denaturation. Furthermore, any unreacted and non-covalently attached sulfonamide component will be removed upon denaturation. The analysis of the turned-over protein variant shows a mass peak in agreement with the size of the formed triazole-sulfonamide reaction product covalently attached to the enzyme. In consequence, part of the covalently bound azide reacted with the non-covalently coordinated alkyne to the cyclization product resulting in the above-mentioned enzyme mass modification. In conclusion, the components **1a** and **2a** react in the presence of CA II-H64C to the triazole product only with Cu<sup>+</sup> assistance. Crystallographic evidence suggests formation of the *1,4*-regioisomer **4a**, which is in accordance with the general Cu<sup>+</sup> catalyzed Huisgen reaction. Within experimental accuracy no indication for the formation of the *1,5*-product **5a** were found.

However, as seen before in the crystallization experiments, the reaction rate (**1a** + **2a** + CA II-H64C) seemed to be rather slow, most likely resulting in an incomplete triazole

cyclization. For what reason is such a low reaction rate observed and why can the initial crystal structure with the separated reaction components prior to the click ring formation be trapped? Obviously, both components are rather tightly immobilized in local depressions of the protein surface additionally fixed through hydrogen bonds. Furthermore, the supplemented  $\text{Cu}^+$  ions appear to be essential to catalyze the triazole formation with these reactants. Finally, the building blocks might simply not feature the spatial requirement necessary for a fast cyclization. At least the rather strained arrangement of the formed product suggests that the geometry of the reaction components has not yet been ideal for the case (**1a** + **2a** + CA II-H64C).

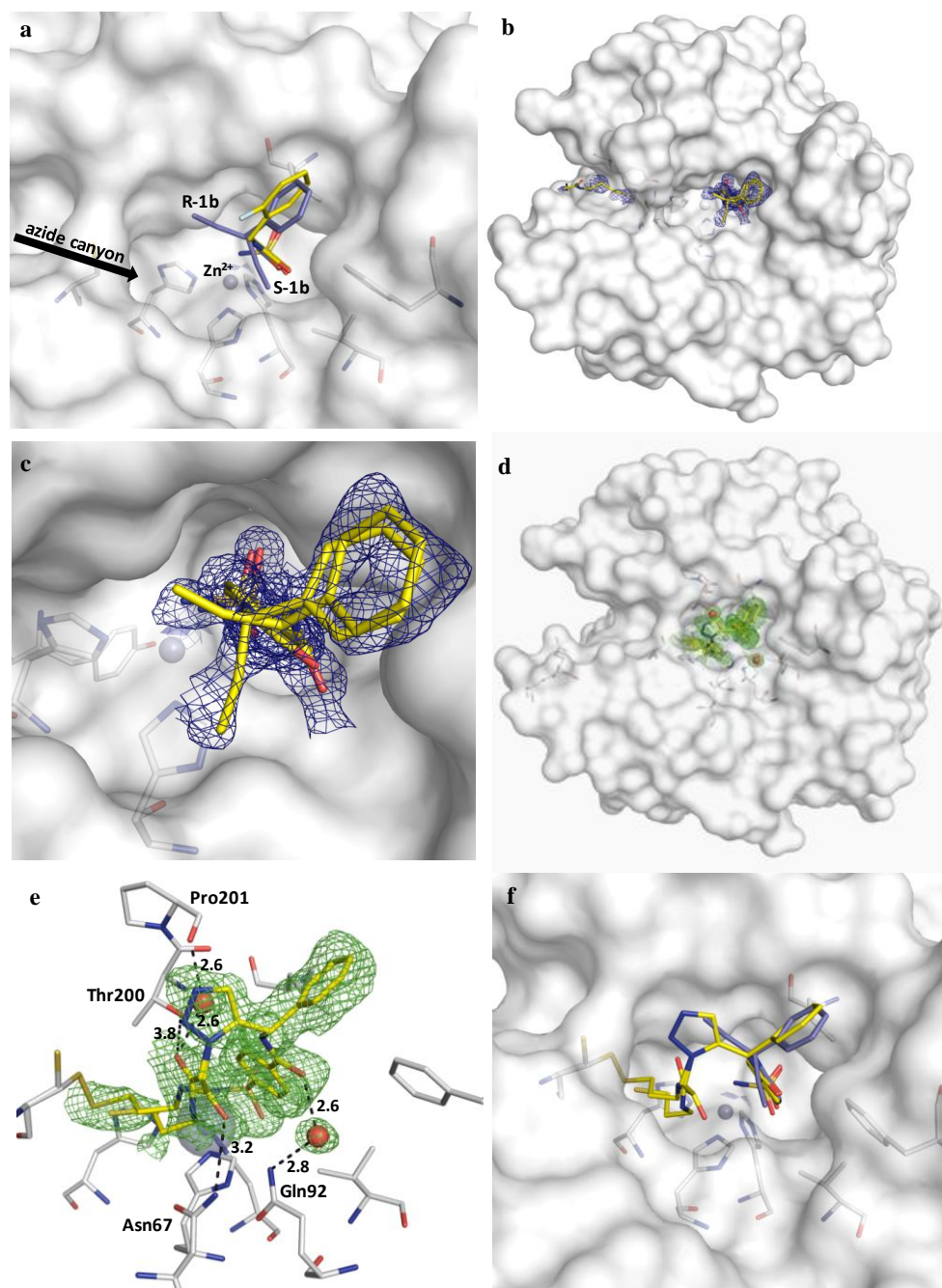
**Triazole formation with optimized building blocks:** Accordingly, we decided to replace **1a** by an extended reagent putatively orienting its alkyne component in a more favorable orientation. Therefore, we screened the known crystal structures in the PDB for binding geometries of CA II inhibitors to find a more promising scaffold.



**Scheme 5.3** Tethered cycloaddition (*rac*-**1b** + **2c** + CA II-H64C  $\rightarrow$  **5c**).

We selected the ligand **6** shown in Scheme 5.3 from an appropriate CA II complex crystal structure (PDB-code: 1G1D) as a starting point for the design of an improved alkyne component **1b** (Kim, Chang et al. 2000). Extended by an amide group the

carbon atom attached to the amide nitrogen and substituted by a second phenyl moiety suggests ideal spatial positioning to hook up the alkyne substituent (Figure 5.2a). However, attachment of the alkyne will create a stereocenter which in the following will give rise to two different orientations of this reacting component towards the azide. Parallel to the replacement of **1a** by **1b** we exchanged the likely too short **2a** by **2b** and **2c**. Both elongated azides with additional linker atoms should minimize the distance to the alkyne and avoid pronounced H-bond fixing of the azide on the protein surface. This reagent component should then experience the required mobility for reaction. *Rac-1b* (racemic mixture), **2b** and **2c** were synthesized and each cocrystallized with the enzyme. A first crystal structure of CA II-H64C cocrystallized with **2c** under mild reducing conditions and subsequently soaked for 2 hours with the racemic mixture **1b** could be determined. Crystals were shock-frozen for data collection shortly after soaking. The structure shows that the new reactants bind in a favorable orientation for the click reaction (Figure 5.2b,c). **2c** is no longer immobilized by an H-bond network on the surface and experiences residual mobility which is indicated by the weakly defined electron density and the high *B*-factors assigned to the terminal atoms of the azide. The soaked racemic *rac-1b* populates the binding pocket in terms of both enantiomeric forms and correct positioning for triazole formation can be anticipated. According to the fact that both enantiomers are present in the binding pockets, the density for the alkyne side chain is only weakly defined (Figure 5.2c). To evaluate whether also for this setup the click reaction is executed in crystalline state, we allowed the crystals of a CA II-H64C, *rac-1b* and **2c** cocrystallization setup to grow for 2 months. With these crystals we were able to observe a difference electron density of the formed click product **5c** (Figure 5.2d,e). The product **5c** was formed without any additional Cu<sup>+</sup> soaking. The electron density of the formed triazole is very well defined indicating high cyclization yields and regioselective formation of the *1,5*-regioisomer. The phenyl ring and aromatic sulfonamide anchor remain virtually in the binding pose attributed to the **1b** alkyne prior to cyclization. In contrast, the azide component has to rotate by 180° for the reaction, very similarly as observed for the cyclization of **1a** with **2a**. The linker atoms of the azide show some flexibility. They adopt two conformations in agreement with the electron density and the conformation of the formed triazole.



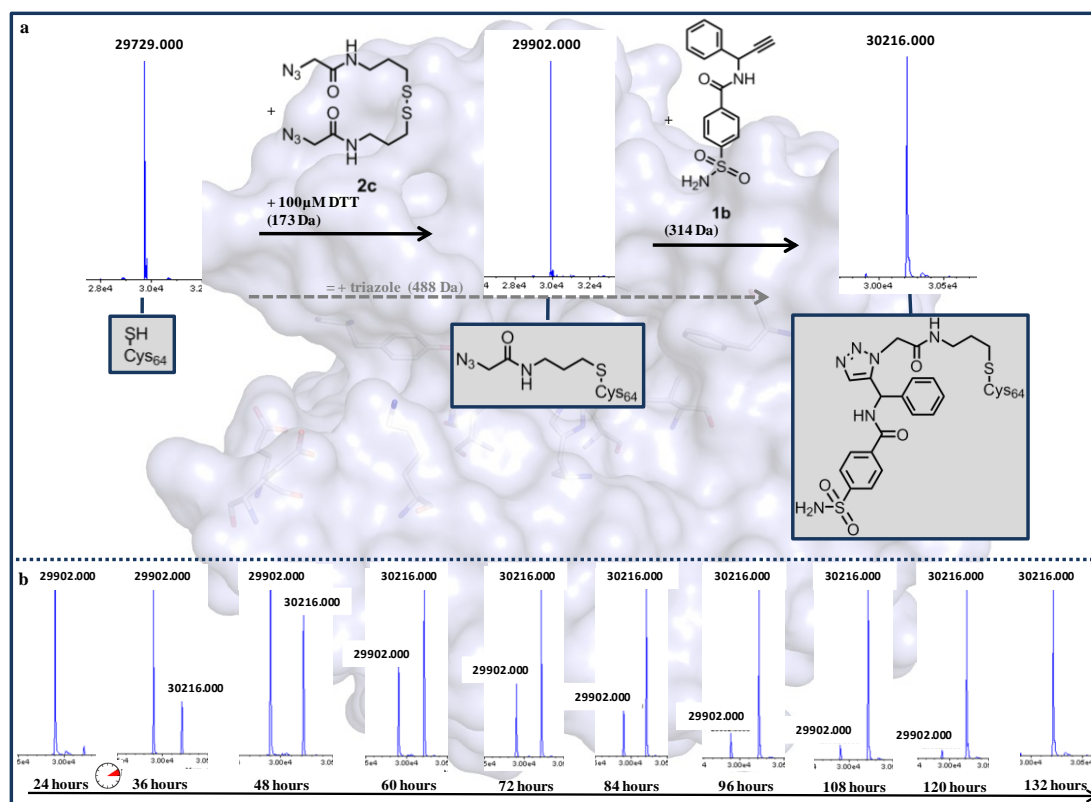
**Figure 5.2** Crystal structures of (*rac*-**1b** + **2c** + CA II-H64C → **5c**). Distances are shown in Å (a) Crystal structure of **6** (PDB-Code 1G1D; protein: C, N, O, S; ligand: C, N, O, S) used as template for the design of **1b** superimposed with the subsequently determined crystal structure of the *rac*-**1b** (C, N, O, S). The solvent accessible surface of the protein is schematically represented in white in all images. The two racemic alkynes place their triple bond into different directions which takes significant influence on product formation. (b) Crystal structure of *rac*-**1b** and **2c** in complex with CA II-H64C. The  $2F_o - F_c$  maps for the reactants are displayed in (b) – (c) at a  $\sigma$  level of 1.0 as blue mesh; (c) Crystal structure of *rac*-**1b** in complex with CA II-H64C seen from close-up. (d) Crystal structure of the *in situ* formed 1,5-*S*-triazole **5c** (C, N, O, S). The  $F_o - F_c$  omit map for the cyclization product is displayed in (d) and (e) at a  $\sigma$  level of 2.0 (green mesh). The product participates in hydrogen bonds to Thr200, Pro201, Asn67 and Gln92 partially mediated by two water molecules (e) Crystal structure of the *in situ* formed triazole **5c** seen from a close-up view. (f) Superimposition of the crystal structure of **1b** and the *in situ* formed click product. The alkyne **1b** adopts the appropriate orientation for triazole formation.



The carboxyl oxygen of the distal amide bond forms, due to the splitting in two conformations, either a hydrogen bond to Asn67 or to Thr200 (Figure 5.2e). In the latter orientation it is also involved in an H-bond to Pro 201 mediated through a water molecule. A second interstitial water molecule forms an H-bond to the proximal amide oxygen of the ligand and to Gln92. The crystals were grown at 18 °C, which indicates that the cyclization can already take place at this temperature. Even though, we started with **rac-1b** and detected accommodation of both enantiomers in the binding pocket prior to cycloaddition only one stereochemically unique product **5c** with *S*-configuration is observed. This stereochemistry results directly from **R-1b** as priority rules invert assignment of absolute configuration. Furthermore, the *1,5*-triazole is formed. Modeling considerations show that also the *1,4*-triazole product would geometrically fit into the binding pocket. Possibly only the *1,5*-product is observed as in this second experiment no Cu<sup>+</sup> ions were present.

**Reaction kinetics:** Our elaborate crystallographic analysis only defines the end point of the reaction. It hardly allows to record details about the reaction kinetics, particularly under varying conditions. Thus, we decided to monitor the triazole formation by mass spectrometric experiments in parallel under solvent conditions. The HPLC-MS analysis of CA II-H64C incubated with the thiole-azide component showed one definite mass peak for the covalent CA II-H64C azide complex (with **2c**: 29729 Da + 173 Da = 29902 Da; (Figure 5.3a). Subsequently, the racemic alkyne component (**rac-1b**) with the sulfonamide anchor was added to this mixture and the reaction batch was incubated for 12 to 144 hours at 37 °C. As long as the non-covalently attached alkyne has not reacted the enzyme mass is not altered since under denaturing conditions the non-covalent component will be removed. Once the reaction gradually forms the triazole product an additional mass peak is detected (29902 Da + 314 Da = 30216 Da). As time proceeded the mass spectra showed a shift in the ratio of enzyme-azide to enzyme-product peaks (Figure 5.3b). Relative quantitative rating could be revealed by integration over the peaks. Analyzing the click product after 30 hours at 37 °C we observed a yield of approximately 20% for either the reaction of **rac-1b** with **2a**, **2b**, and **2c** respectively.

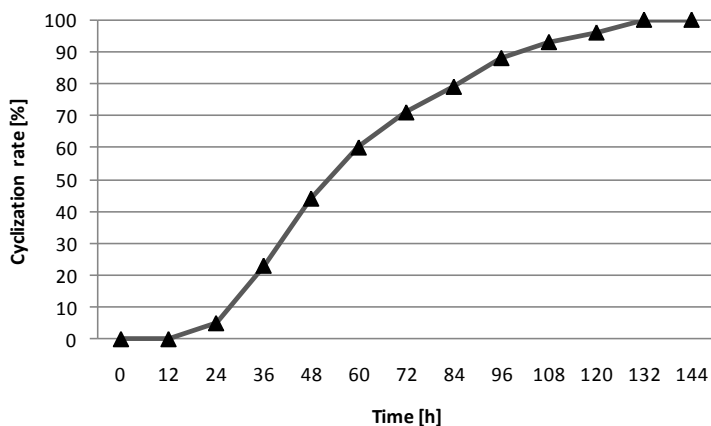
To further accelerate and influence this reaction, we studied the reaction of *rac*-**1b** with **2c**, as this combination showed the best yields. Figure 5.4 shows the cycloaddition rates along reaction time. Subsequent to an induction time, the triazole formation occurs with similar reaction rate until reaction components are fully transformed to the product. Recorded via approx. 1500 mass analyses, we systematically altered the reaction conditions by modifying temperature and reaction time along with the exposure of different metal ions.



**Figure 5.3 Monitoring reaction progress via HPLC-MS.** (a) Reaction pathway for (CA II-H64C + **2c** + *rac*-**1b** → **5c**) monitored by HPLC-MS experiments. CA II-H64C shows a mass of 29729 Da. After addition of **1b**, **2c** and 100 μM DTT to reduce the azide internal disulfide bridge, the HPLC-MS analysis shows one definite mass peak of 29902 Da representing the mass of the enzyme with the covalently attached azide. Upon incubation with the alkyne at 37 °C the triazole formation is induced by the enzyme and an additional mass peak of 30216 Da appears in the mass spectrum. (b) Incubating the setup mentioned in (a) for 12 to 144 hours taking samples every 12 hours we could monitor the rate of triazole formation by comparing the peaks corresponding of the enzyme-azide with the enzyme-triazole products. Within 132 hours full conversion of enzyme-bound azide and alkyne components can be observed.

**Temperature dependence:** Experiments were performed at 20 °C, 37 °C and 55 °C. The conversion was analyzed between 12 and 144 hours taking samples every 12 hours. A strong temperature dependence could be observed. At 20 °C a first product

peak could be detected after 96h indicating 10% click product. At 37 °C during the same time span the triazole yield was increased to 80%. After 48 hours at 55 °C already full product formation (100%) can be observed.



**Figure 5.4** Time dependent 2+3 cycloaddition (CA II-H64C + **2c** + *rac*-**1b** → **5c**). The triazole formation rate of the CA II-H64C-azide to CA II-H64C-triazole complex is shown in percent with respect to the reaction time in hours.

**Effect of metal ion supplementation:** The following metal ions have been tested as catalysts of the click-reaction in the binding pocket of CA II-H64C: Cu<sup>2+</sup>, Cu<sup>+</sup>, Zn<sup>2+</sup> and Ru<sup>2+</sup>. The solvated samples were incubated with metal ions for 12 to 144 hours at 37 °C. In case of Cu<sup>+</sup>, the samples showed pronounced acceleration also at room temperature transforming 20% of the azide within 96 hours compared to 10% for samples lacking Cu<sup>+</sup> presence. At 37 °C we accomplished 100% product formation after 84 hours, remarkably without copper 132 hours were not sufficient for complete turn-over. We performed the same experiments with exposure to Cu<sup>2+</sup> ions. A similar catalytic impact was observed, however with reduced efficiency compared to Cu<sup>+</sup>. It is difficult to elucidate whether actually the Cu<sup>2+</sup> ions are catalytically active, or whether they are reduced to Cu<sup>+</sup> ions by the always present low DTT concentrations applied to reveal the thiole tethering of the azide component. Addition of Ru<sup>2+</sup> and Zn<sup>2+</sup> did not reveal any significant influence on the reaction rate. Interestingly enough, in all test series the triazole formation starts with significant delay. For the first 10 to 15 hours we could not detect any product formation. Similar findings have been reported for the Huisgen reaction catalyzed by diverse copper systems (Pachón, van Maarseveen et al. 2005).

**Regio-/stereoselectivity:** Finally we addressed our attention to the produced regio- and stereoselectivity. In our crystallographic experiments we exposed the racemic **rac-1b** as alkyne component, interestingly enough resulting exclusively in the *1,5-S*-triazole product (Figure 5.2d,e). The previous study with **1a** and **2a** revealed *1,4*-triazole formation (Figure 5.1b). At least for the reaction of **rac-1b** and **2c** modeling considerations could not exclude *1,4*-triazole formation for steric reasons. However, in the crystal only the *1,5*-stereoisomer is observed notably in the absence of any copper ions. Our HPLC-MS experiments revealed product formation with and without the presence of copper ions, however, we have no information which regioisomers are produced possibly as a mixture and whether the presence or absence of copper takes any directional influence. Additional crystallographic studies will be necessary to get better insights into the role of Cu<sup>+</sup> in this system. Attempts to cocrystallize the system appear rather impracticable since crystal growth takes months and during this period the copper-free cyclization will already be fully executed. As an alternative, soaking experiments in presence of Cu<sup>+</sup> might help to reveal an answer to this important aspect of the copper-guided regio- and stereochemistry, but might be unfeasible due to lacking stability of the crystals over sufficient time.

Nevertheless, we further studied the influence of the stereochemistry of the chiral alkyne component on product formation. Even though, we observed crystallographically the binding of the racemic **rac-1b** to CA II-H64C we only detected the *1,5-S*-triazole. We therefore synthesized the alkyne enantiopurely and performed additional HPLC-MS measurements. **R-1b** and **S-1b** were separately reacted in the enzyme with **2c** for 144 hours. The HPLC spectra revealed that both enantiomers are capable to initiate triazole formation. In agreement with the above-described crystallographic evidence, the *R*-enantiomer showed higher reaction rates. The question remains why do we see in the crystal exclusively the *S*-product and not, at least to some extent, also the *R*-product? Does the selective synthesis depend on a raised affinity of one enantiomer compared to the other or on the spatial organization of both components?

The structure with the racemic starting material was obtained by short time soaking. Likely both enantiomers exhibit very similar binding affinity to the enzyme as

indicated by inhibition data measured in a bioassay. Crystallographic experience shows that the solid state is highly discriminative with respect to the binding of stereoisomers, particularly, if the crystals were grown over sufficiently long time. Then only one isomer is usually found even though a racemic (or diastereomeric) mixture of similarly potent ligands was exposed. Through equilibrium over the long time of crystal growth even small affinity differences are sufficient to finally populate only the more potent ligand in the binding pocket.

In our case the crystals had two months time at 18 °C to accomplish the cycloaddition reaction whereas for the complexes with the reagents where we could detect the racemic mixture of the starting material a short soaking over 2h has only been allowed. Most likely, as the racemic alkyne component is only bound to the enzyme reversibly, **S-1b** could be gradually exchanged by **R-1b** present in excess in the crystallization buffer. The latter *R*-enantiomer adopts a more favorable orientation for the cyclization reaction (Figure 5.2f) with better suited spatial orientation of the cyclization step, thus after two months exclusively the *1,5-S*-product is revealed.

**Control-experiments:** Up to now, it could only be evidenced indirectly that the triazole formation occurs within the active site of a protein (Lewis, Green et al. 2002). Triazoles formed in solution might imitate *in situ* generated click products. To verify in our case that the click reaction has actually been executed in the binding pocket we performed several control experiments. First, we added the potent CA II inhibitor ethoxzolamide to the reaction batch. After exposure at 37 °C for 72 hours the HPLC-MS spectra did not indicate any product formation. Only the mass-shift for the tethered **2c** bound to the enzyme (29729 Da → 29902 Da) could be recorded. Obviously, Ethoxzolamide replaces or prevents binding of the less potent sulfonamide **1b**. In case a background reaction would have occurred in solution, the generated clickproduct should also be tethered to Cys64, leading to the product mass shift. Similarly any reaction of the covalently tethered azide component with the protein unbound sulfonamide-type alkyne can be excluded. In consequence the alkyne is only competent for the cyclization once bound to the active site. To validate this conclusion we tried to react the azide **2c** with different alkynes lacking a sulfonamide anchor.

Also under these conditions no product formation could be observed. Finally, in another control experiment the enzyme was purified over a size exclusion column after short incubation at room temperature with the azide and the alkyne. Thereby the enzyme was purified from any unbound reactants. Only the tethered and zinc-coordinated building blocks remained in the samples. After exposure at 37 °C for 48 hours and HPLC-MS analysis of these samples, product formation could be observed with the same rate as observed for all the other test series. Furthermore, we analyzed the solutions from which the protein has been extracted. If any solvated reactants would have been present, a background reaction might have occurred. However no triazoles click product could be detected.

## 5.5 Conclusion

In summary, we have developed a CA II mutant that can be used for covalent disulfide attachment of azide click reactants on the enzyme surface. The immobilized components can be reacted with appropriate alkyne components fixed through an attached sulfonamide anchor that allows to reversibly coordinate the CA II active site zinc ion. This experimental setup permits unique structural and kinetic insights into the complex reaction conditions driving the *in situ* triazole formation, in parallel it gives access to structural information via crystal structure analysis and to kinetic data via the achieved reaction turn over along the reaction path by HPLC-MS. Systematically modifying the reaction conditions, we were able to elucidate the time-, temperature-, additive catalyst- and reactant-dependency of the *in situ* click chemistry performed on the protein surface. Elevated temperatures as well as the additional supplementation with Cu<sup>+</sup> ions accelerate the triazole cyclization. Furthermore, the stereochemical properties of both building blocks influence the reaction rates as interactions with the protein surface characterized by crystallography can diminish the mobility of the reactants and thus slow-down turn over. We have proven via appropriate control experiments that the studied click reaction occurs actually exclusively in the binding pocket of CA II. Furthermore, our *in situ* click reaction is performed under regio- and stereo-selective control as we could demonstrate that only the selective synthesis of the *1,5-S*-triazole is observed. Furthermore, the synthesis of *1,4*-triazoles can be achieved most likely induced by Cu<sup>+</sup> supplementation. The proposed system provides the

perspective to create protein surface exposed reaction sites. The next step would be to omit the tether in an efficient system evolved to go towards catalytic applications.

## 5.6 Experimental Section

**Cloning, expression, purification and mutagenesis of CA II:** This is described in Chapter 3.

**Crystallization of CA II complexes:** The CA II crystals were grown using the sitting drop vapor diffusion method at 18 °C by mixing the protein solution (5  $\mu$ l: ~10 mg/ml) with well solution (5  $\mu$ l: 2.75 M  $(\text{NH}_4)_2\text{SO}_4$ , 50 mM Tris-HCl pH 8.0). After a few days the crystallization drops were seeded with a seeding solution from old CA II crystals. Crystals appeared within 2 – 4 weeks in space group  $P2_1$  and took up to 2 months to reach their maximum size. Complex structures were obtained by cocrystallization or soaking of the enzyme with the building blocks (1 mM). For cryoprotection, crystals were briefly soaked in mother liquor containing 25% glycerol.

**Data collection, phasing, and refinement:** The data sets were collected at the synchrotron BESSY II in Berlin/Germany on PSF beamline 14.2 (PDB codes: 3KIG, 3KNE). Data were processed and scaled with Denzo and Scalepack as implemented in HKL2000 (Otwinowski, Minor et al. 1997). The structures were determined by the molecular replacement method with Phaser (Storoni, McCoy et al. 2004) with our first inhouse CA II-H64C structure as search model. Refinement was continued with CNS (Brünger, Adams et al. 1998) and SHELXL-97 (Sheldrick, Schneider et al. 1997). For each refinement step at least 10 cycles of conjugate minimization were performed, with restraints on bond distances, angles, and  $B$ -values. Intermittent cycles of model building were done with the program COOT (Emsley and Cowtan 2004). The coordinates have been deposited in the PDB (<http://www.rcsb.org/pdb/>) with the following access codes: 3KIG, 3KNE. The statistics are shown in the Appendix 13.1.

**HPLC-MS analysis:** For HPLC-MS measurements 10  $\mu$ l of the CA II-H64C mutant (75  $\mu$ M), 5  $\mu$ l DMSO, 10  $\mu$ l DTT-solution (500  $\mu$ M), 0.5  $\mu$ l azide (100 mM), and 0.5  $\mu$ l alkyne (100 mM) was diluted to a final volume of 50  $\mu$ l with buffer (50 mM

Tris-HCl pH 7.8). The samples were incubated at different temperatures (20 °C – 55 °C) between 12 and 144 hours. After incubation the sample solutions were pre-purified by size-exclusion chromatography (Illustra AutoSeq G-50 columns – GE Healthcare) to separate from excess substrates to avoid undesired side reactions. Complete desalting of the protein was achieved by HPLC using an Agilent 1100 system, samples were applied to a monolithic 50/1 ProSwift RP-4H column (Dionex). Desalted protein samples were eluted by the following gradient of buffer A (water/0.05% formic acid) and buffer B (acetonitrile/0.045% formic acid) at a column temperature of 40 °C and a flow rate of 0.2 mL/min: Isocratic elution with A (5%) for two minutes, followed by a linear gradient to B (95%) within 8 minutes and holding B (95%) for additional 4 minutes. Online mass spectrometric analysis was done with a Qstar Pulsar I mass spectrometer (Applied Biosystems) equipped with an ESI source. Measurement parameters were as follows: DP1 75, FP 265, DP2 15, CAD 2, GS1 35, CUR 25. The applied voltage was 5000 V. Positive ions within the mass range of 500-2000 m/z were detected. For better performance, the “enhance all” mode was activated.

**Metal ion incubation:** In some cases a solution of Cu<sup>+</sup>, Cu<sup>2+</sup> (CuSO<sub>4</sub>), Zn<sup>2+</sup> (ZnSO<sub>4</sub>) or Ru<sup>2+</sup> (RuCl<sub>2</sub>\*DMSO<sub>4</sub>) were added before incubation. Cu<sup>+</sup> was gained by adding five-fold excess of ascorbic acid to a Cu<sup>2+</sup> solution. We applied metal ion concentrations ranging from 10 μM to 1 mM. At concentrations above 40 μM the samples often showed spectra which were difficult to interpret most probable due to denaturation. To analyze the impact the metal ions develop on the mutant without any building block, we performed initial measurement by incubating CA II-H64C with each metal ion solution. In case of Cu<sup>+</sup> we observed in several samples an additional mass peak which shows the size of the mutant reduced by a mass of 35 Da. This by-product is dependent on the reaction time and the applied temperature and its occurrence increased with raised concentrations along solvent evaporation. We assign these changes to some redox instabilities of the protein upon exposure to Cu<sup>+</sup> ions.



## **6 Saccharin Inhibits Carbonic Anhydrases: Possible Explanation for its Unpleasant Metallic Aftertaste**

### **6.1 Introductory Remarks**

This chapter has been published in *Angewandte Chemie*. In the following the original text is depicted. The present study was accomplished in cooperation with Karen Köhler, Alexander Hillebrecht and Alessio Innocenti. The crystal structure 2Q38 has been contributed to this publication. The same structure has been used as a starting point for a tethering approach on carbonic anhydrase II (Chapter 7).

### **6.2 Introduction**

Saccharin is the oldest artificial sweetener, discovered by serendipity in 1879 by Fahlberg and Remsen (Fahlberg and Ira 1879). About 450 times as sweet as sucrose, it displayed a very important discovery, particularly for diabetics. Commercialization started shortly after its discovery and became increasingly popular in recent years as calorie-free sugar surrogate. Blended with cyclamate, it receives huge success, e.g. sold as ‘Sweet ‘n Low’. Saccharin is not only known for its extreme sweetness, but also for a bitter and metallic aftertaste. Even though there have been worries about its safety since its introduction, particularly in the 60ties and 70ties, speculations about its putative carcinogenicity have been put forward. Obviously at normal doses no clear evidence for a causal correlation of saccharin consumption and health risks for humans could be demonstrated (Weihrauch and Diehl 2004).

Saccharin is applied as its sodium salt. Chemically it represents a sulfimide with a lactame and cyclic sulfonamide moiety. The latter functionality is responsible for the acidic character of the molecule and suggests its potential to interact with the zinc ion at the floor of the binding pocket of carbonic anhydrases. The family of carbonic anhydrases is an important class of ubiquitously present proteins distributed over many compartments and organs of our body. At present 15 isozymes have been characterized (Supuran 2007) (presently 16 s. Chapter 1.1). They all catalyze the same reaction: the fixation of carbon dioxide by water to bicarbonate. As such, the different carbonic

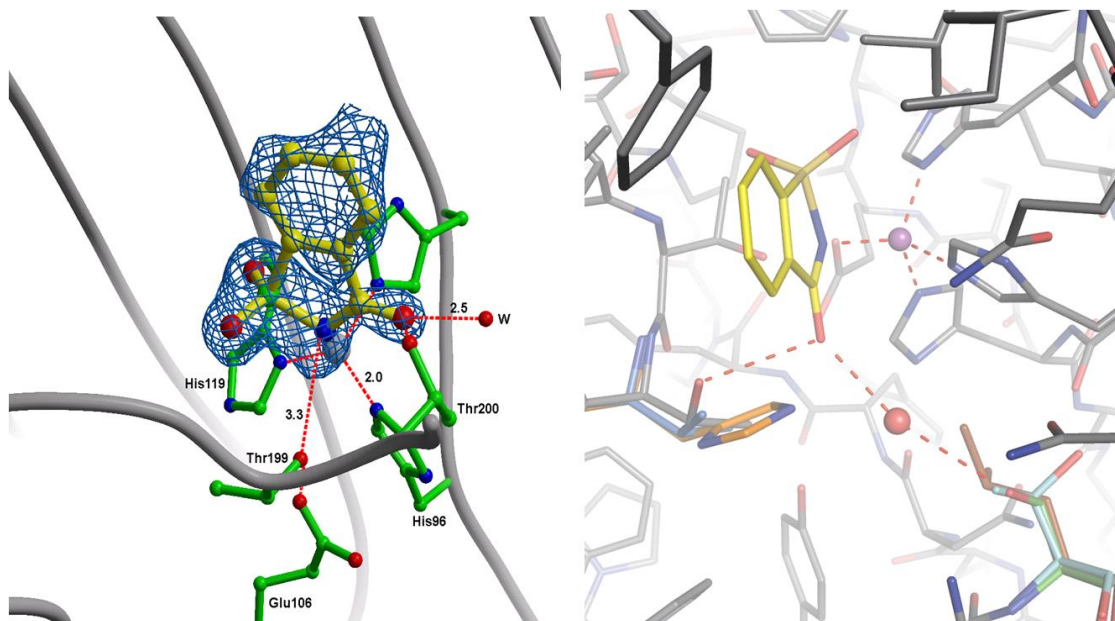
anhydrases are involved in various essential tasks, e.g. CO<sub>2</sub> transportation, pH regulation or delivery of C1 building blocks in biosynthesis (Supuran, Scozzafava et al. 2003).

### 6.3 Results

We determined the inhibition profile (Khalifah 1971) of saccharin against a panel of different carbonic anhydrase isoenzymes. Surprisingly enough, the compound shows up to nanomolar inhibition of some members of this protein family (Table 6.1). Compared to the classical and therapeutically well-established CA inhibitor acetazolamide (AAZ) and other sulfonamides, such as furosemide, hydrochlorothiazide, or topiramate (TPM), saccharin shows remarkable selectivity discrimination among the different isoforms. This fact is in particular impressive as all CAs possess largely conserved catalytic centers and medicinal chemists have to fight hard to equip their development compounds with a sufficient selectivity profile. With respect to the ubiquitously distributed CA II saccharin shows low micromolar inhibition. This made us confident to embark into the structure determination of saccharin in complex with carbonic anhydrase II (Figure 6.1, crystal data 13.1).

Saccharin coordinates, most likely in deprotonated state, with its nitrogen to the catalytic zinc producing a flattened distorted tetrahedral environment at the metal center. The nitrogen itself adopts a pyramidal local environment and prolonged along the assumed lone pair direction, the oxygen atom O<sub>γ</sub> of the Thr199 side chain is found. The mutual distance of 3.3 Å advocates for a weak H-bond formation, however, the latter oxygen atom is also in H-bond donor distance of one carboxylate oxygen of Glu106 (2.5Å). One oxygen of the neighbouring sulfoxy group expands the coordination at zinc towards distorted bipyramidal geometry (3.1 Å), whereas the second oxygen forms a weak hydrogen bond to the backbone NH group of Thr199. The lactam carbonyl oxygen is in hydrogen-bond distance to Thr200 O<sub>γ</sub>. Furthermore, it forms a hydrogen bond to an active site water molecule. Another peculiarity indicated by the present structure is the incorporation of a putative second zinc ion to the active site. His64 involved in shuttling protons in and out of the catalytic site is found distributed over two conformations. The first conformer is superimposed by a

strong density peak that could be refined as a second partially populated zinc ion (67%). It coordinates to His64 N $\delta$  in the second conformer, the backbone NH of the same residue and the backbone carbonyl group of Asn62. Furthermore, two water molecules complete the coordination sphere of this second zinc ion. Zinc ions are prevalently available, usually in micromolar concentration, thus they can be picked-up from the buffer conditions and incorporated into the structure.

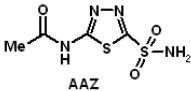
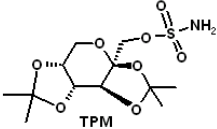
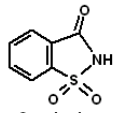
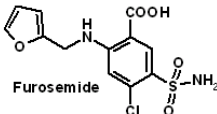
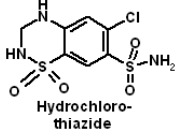


**Figure 6.1** Crystal structure of CAII Saccharin complex. (a) Binding site of carbonic anhydrase II (catalytic residues C, N, O, S) with bound saccharin (C, N, O, S, difference electron density (blue) contoured at 3  $\sigma$ ). The picked-up active-site water is shown to the right as red sphere. (b): Crystal structure from the left (protein residues: C, N, O, S; saccharin (C, N, O, S) and active site water (red sphere), zinc (violet sphere)) with superimposed residues mutated in the other isoforms: His200 (C, N, O, S, CA I), Val200 (C, N, O, S, mCA XIII), Ser65 (C, N, O, S, CA VII), Leu65 (C, N, O, S, CA Va), Thr65 (C, N, O, S, CA XIV).

On the basis of the present structure determination, it is tempting to speculate about the impressive selectivity discrimination of saccharin among the different isozymes. Nearly exclusively, the well-established and clinically approved CA inhibitors exhibit a terminal exocyclic sulfonamide, sulfamate or sulfimine functionality (Abbate, Supuran et al. 2002). In contrast, saccharin shows a cyclic structure and its lactam carbonyl group in the five-membered heterocycle manifests an exclusive feature not given in the other usually considered inhibitor skeletons. Saccharin shows moderate inhibitory potency towards CA II, but nanomolar binding is recorded for CA VII. As no crystal structure is available for this isoform, a homology model retrieved from the

Swiss-Model Repository was used.<sup>4</sup> Interestingly, at the far end of the binding pocket in the vicinity of the lactam carbonyl oxygen an alanine residue is replaced by a serine (Figure 6.1b). In the crystal structure of the saccharin CA II complex a water molecule is observed which is ideally placed to mediate a contact between the lactam carbonyl and the serine O $\gamma$ . This additional contact is likely to enhance ligand binding. In a recent study on aldose reductase we could show that an inhibitor which picks-up a contact-mediating water molecule gains a significant affinity advantage over a closely related ligand (Steuber, Heine et al. 2007). The latter one lacks formation of a similar contact to an incorporated interstitial water molecule. Interestingly, a sequence alignment shows that this Ala/Ser replacement is also given in CA I, CA IV, CA Vb, CA IX, CA XII and murine CA XIII.

**Table 6.1** Inhibition data of the clinically used compounds acetazolamide (AAZ), topiramate (TPM), furosemide and hydrochlorothiazide together with saccharin against isozymes I – XIV (stopped-flow, CO<sub>2</sub> hydrase method (Khalifah 1971)).

Isozyme*	K <sub>i</sub> ** (nM)					
	AAZ	TPM	Saccharin	Furosemide	Hydrochlorothiazide	
hCA I <sup>a</sup>	250	250	18540	62	328	
hCA II <sup>a</sup>	12	10	5950	65	290	
hCA IV <sup>a</sup>	74	4900	7920	564	427	
hCA Va <sup>a</sup>	63	63	10060	499	4225	
hCA Vb <sup>a</sup>	54	30	7210	322	603	
hCA VI <sup>a</sup>	11	45	935	245	3655	
hCA VII <sup>a</sup>	2.5	0.9	10	513	5010	
hCA IX <sup>b</sup>	25	58	103	420	367	
hCA XII <sup>b</sup>	5.7	3.8	633	261	355	
mCA XIII <sup>a</sup>	17	47	12100	550	3885	
hCA XIV <sup>a</sup>	41	1460	773	52	4105	

\* h = human; m = murine isozyme.

\*\* Errors in the range of 5-10 % of the reported value (from 3 different assays).

<sup>a</sup> Human (cloned) isozymes, by the CO<sub>2</sub> hydration method;

<sup>b</sup> Catalytic domain of human, cloned isozyme, by the CO<sub>2</sub> hydration method.<sup>[5]</sup>

Also CA IX and XII show enhanced saccharin binding, whereas for CA I and CA XIII a dramatic drop in affinity is observed. In the latter two isoforms Thr200 is replaced by either His200 or Val200. Both require significantly more space next to the lactam

<sup>4</sup> <http://swissmodel.expasy.org/repository>

carbonyl group. Thus, they are expected to produce steric conflicts with the bound saccharin. A dramatical drop in affinity can be expected. Also in CA Va weak inhibition is observed. Here, the crucial Ser to Ala replacement is exchanged by a spatially more demanding Leu. In addition, this residue would be inadequate to promote a water-mediated contact. As the Ala mutation, this residue will be detrimental for potent saccharin binding. Finally, also CA VI and CA XIV are remarkably inhibited by saccharin. Interestingly, they possess a Thr residue at the position of Ala65 in CA II. This residue should be equally well suited to pick-up a water-mediated contact from the ligand.

The question arises whether the observed potent *in-vitro* inhibition of some of the CA isoforms has any pharmacological consequences, e. g., in terms of desired or undesired cross-reactivities or side effects. Saccharin is expected to activate the sweetness-sensing receptors on the tongue. Apart of this function, an ideal sweetener should pass the intestinal system without an unwanted penetration into the blood stream or even passing the blood brain barrier. CA VII, the most strongly inhibited isoform, is located in the brain. Unlikely, considering the acidic character of saccharin, the compound will arrive in the brain at significant concentration by passive transportation. Only under acidic conditions of the stomach a considerable amount of saccharin will be present in the uncharged more hydrophobic state. In this state some passive penetration might be given, however absorption at pH 6-8 is rather unlikely. Nevertheless, it has been shown that saccharin is nearly completely absorbed from the gut and rapid elimination in the urine occurs as unchanged compound (Renwick 1985). In rat a distribution of saccharin across most organs has been described. A possible active transport has been suggested (Renwick 1985). Obviously, the compound is chemically rather inert and not easily metabolized *in vivo*. Nevertheless, the decrease in plasma levels after oral dosing has been reported as slow and limited by the rate of absorption from the gut (Renwick and Sweatman 1979). It remains, however, speculative, whether this latter observation can be attributed to putative inhibition of different CAs present in various compartments of the organism. Recently, CA VI, the only secreted form of CAs has been discussed for its putative involvement in olfaction, taste and pH regulation in the oral cavity (Charles, Antonio et al. 2006). As such, it contributes to the acidification of

the enamel pellicle and speculations about its inhibition with respect to caries protection have been put forward (Kimoto, Kishino et al. 2006; Nishimori, Minakuchi et al. 2006). Saccharin possesses an unpleasant bitter or metallic aftertaste, a property shared with clinically approved systemic carbonic anhydrase inhibitors (Supuran and Scozzafava 2000). This may in fact be due to the inhibition of the salivary CA VI for which also saccharin shows submicromolar inhibition. Nevertheless, as the various CAs are repeatedly discussed as putative targets for drug therapy with respect to diuresis, glaucoma, tumor suppression, or obesity, saccharin might tell the medicinal chemists a lesson how to equip their molecules with additional features to achieve impressive selectivity discrimination.

## **7 Tethering Approach on Carbonic Anhydrase II Supported by Pre-selecting Computational Docking Experiments**

### **7.1 Introductory Remarks**

The following chapter was prepared for submission to a scientific journal. This study was performed in cooperation with Dr. Sascha Brass and Dr. Uwe Linne. The compounds for the tethering experiments were synthesized by Dr. Sascha Brass and Thomas Kronenberger. All HPLC-MS samples were analyzed by Dr. Uwe Linne from the Chemistry Department.

### **7.2 Introduction**

The amount of well-characterized proteins is rapidly increasing due to multiple methodological improvements in modern molecular and structural biology. This enormous body of information is available and demands adequate techniques that can be used to screen and validate these targets for the development of drugs. In the last decades high-throughput screening (HTS) increasingly provided first hits as putative candidates for a subsequent lead optimization (Klebe 2009). At first up to a several million compounds are tested for binding to the target protein of interest. Upon discovery of an active compound a second in-depth screening is performed hopefully providing compounds of narrow chemical space to further profile and validate binding of the initial hit. Such data assists the hit validation step. Once the relevance of a detected hit can be approved it is taken to further optimization through iterative drug design cycles. Meanwhile, HTS has provided for various therapeutic applications valuable hits that could be translated into blockbusters. However, besides these benefits, HTS also suffers from certain disadvantages and is not always the gold standard.

The screening of large compound libraries under high throughput conditions requires an assay that reliably and effectively delivers affinity data for each tested candidate. This demands high financial resources to produce sufficient amount of the target protein, either isolated from the parent organism or recombinantly produced, a

sufficiently large compound library and appropriate automated test facilities. Besides these investment costs, there are also methodological limitations inherently related to HTS. The compounds detected by HTS usually reveal affinities in the micromolar range. To improve the interactions with the target protein structural changes are necessary, e.g. by the addition of further functional groups. This easily leads to molecular masses exceeding beyond 500 to 600 Da, which is usually defined as a border line for sufficient bioavailability. Therefore, a selected hit has often to be reduced in size prior to further optimization which parallels an initial loss in affinity. The FDA excludes certain functional groups, such as phenols, therefore current research tries to avoid such groups in detected hits to follow up. "Ligand efficiency" allows first rating of identified hits. It is calculated by dividing the binding affinity of a compound by its number of non-hydrogen atoms (Hopkins, Groom et al. 2004). The ligand efficiency of screening hits often shows that small compounds provide a better value and accordingly suggests them as a favorable starting point compared to larger compounds exhibiting the same affinity. The ligand efficiency therefore captures better the actual potency of discovered hits and ranks them with respect to the space left for further optimization (Hopkins, Groom et al. 2004; Abad-Zapatero and Metz 2005).

Keeping these requirements for screening hits in mind, consideration about compound libraries have been extended toward smaller and smaller size exceeding the range of so called fragments. All fragment-based experimental approaches to lead discovery attempt to identify low molecular weight (50 - 250 Da) molecules with high intrinsic binding affinity for the target. The affinity of these hits typically range from high micromolar to low millimolar binding. A crucial step in fragment-based lead discovery is an appropriate library design. There are numerous attempts for such design which vary in a number of parameters (Lyne 2005). Researchers first emphasize on molecular weight, cLogP distributions, numbers of hydrogen bond donors and acceptors and the number of rotatable bonds. In our group such a library was developed by Tobias Craan constrained by the following parameters:  $cLogP \leq 3$ ; heavy atoms  $< 20$ ; ring systems = 1; non-ring atoms  $\geq 1$ ; free rotatable bonds  $\leq 3$  and molecular weight 120 – 230 Da (in case of a bromide or sulfonamide group present also exceeding this threshold). Slight



deviations in these parameters can result in drastic changes of chemical properties of the library potentially excluding important parts of chemical space.

The screening of fragments against a given target can be performed by different methods. Nuclear magnetic resonance - (NMR) and X-ray crystallography are methods that reveal initial structural information about the binding mode of putative hits in the binding pocket of the target protein. Inhibition assays, ThermoFluor<sup>®</sup> assay technology and surface plasmon resonance (SPR – Biacore) technology provide additional information about the affinity of the fragments towards the enzyme (Ericsson, Hallberg et al. 2006). S. Fesik et al. first reported successful detection and utilization of fragments in drug design which became famous as the SAR by NMR approach (structure-Activity-Relationships by Nuclear Magnetic Resonance) (Shuker, Hajduk et al. 1996). In their study, two-dimensional, isotope-edited NMR spectroscopy was used to detect two fragment leads that bind to two proximal sites about a protein surface. Using three-dimensional structural information on the bound ligands, the fragments were successfully joined together to produce a high affinity ligand. This so called additivity in binding suggests assembly of drug candidates from pieces, which offers a tremendous combinatorial advantage compared to the screening of drug size molecules (Erlanson, Wells et al. 2004). Linking two different fragments taken from a 1000 entry-large library in binary combinatorial fashion can cover a virtual space of 1,000,000 compounds. Apart from this combinatorial advantage, the main benefits of fragment-based drug discovery methods are the following: better access to small pockets; synthetic accessibility; good solubility and simple handling of such a small library. Interestingly enough, in several cases, potent leads were produced using fragment-based drug design even though a conventional high-throughput screen failed to produce prospective leads (Hajduk 2006). This indicates that HTS libraries can possibly lack sufficient chemical diversity for certain targets while a fragment-based approach can more easily cover the required diversity in chemical space. Many pharmaceutical companies appreciate this fact and extended their efforts to both the depth and breadth of chemical diversity in their corporate compound repositories (Leach and Hann 2000; Jacoby, Schuffenhauer et al. 2005). Considering these aspects

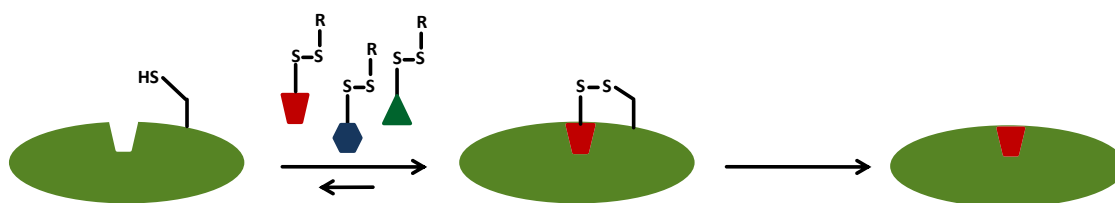
a combination of different techniques such as fragment-based approaches and HTS appears most advisable to perform a successful screening for a novel drug target.

However, there are also some major pitfalls in fragment-based lead discovery. As fragments experience rather low affinity to the target enzyme, they frequently remain undetected in an assay. To measure the affinity of a fragment properly, researchers have to apply high concentrations of the fragments in such an assay. This might lead on the one hand to solubility problems and on the other hand to false positives. Shoichet and colleagues have demonstrated that a number of molecules can act as "promiscuous binders" at a wide range of targets, probably by forming extensive aggregates (McGovern, Caselli et al. 2002). Furthermore, the evaluation of the binding mode of fragments within the binding pocket of the target protein can be quite difficult since crystallization conditions of a protein are highly dependent and extremely sensitive to externally modified properties for example added high ligand concentrations. Furthermore, protein crystals often do not survive harsh soaking conditions. There are numerous examples for different fragments that have successfully been soaked into protein crystals at concentrations about 50 - 100 mM. Jürgen Behnen has identified diverse probe molecules that can be used to map the binding pocket for favorable interaction sites. However, even in case of a successful soaking or cocrystallization experiment such cocrystal structure often lacks sufficient difference electron density for the bound fragment. This can result from various reasons, e.g. the concentration is not adequate or the fragment adopts multiple conformations resulting in low occupancies and therefore the obtained difference electron density can hardly be interpreted.

To circumvent these problems and still benefit from all the described advantages of a fragment-based approach the so called "tethering" method has been developed by J. Wells and coworkers of the Sunesis company in San Francisco (Erlanson, Wells et al. 2004). Only 30 to 50 mg of enzyme are needed for the screening of 10,000 compounds. This method allows the detection of fragments which develop only weak affinity towards the target enzyme and cannot be identified by common inhibition assays. Additionally, it facilitates the generation of 3D structures of the protein inhibitor complex by X-ray crystallography. Tethering relies on the reversible covalent

bond formation between a fragment and a protein of interest. To apply the method, a cysteine residue is required to be present within or in close vicinity to the active site (Figure 7.1). A distance of 5 – 10 Å to the site of interest appears most favorable. In case there is no cysteine residue present in this region of the protein it can be introduced by site-directed mutagenesis.

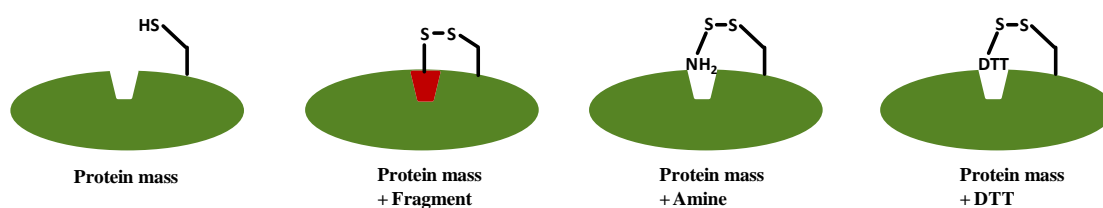
Thereafter, the enzyme is tested against a library of disulfide-containing fragments. In a screening campaign the target protein is incubated with a mixture of 5 to 10 entries from such a fragment library under mild reducing conditions. In principle, appropriately placed cysteine thiole groups should be able to form a disulfide with each of the exposed fragments, and the exchange should occur rapidly. If one of the fragments has inherent affinity for the protein, the thiole-disulfide equilibrium is shifted in favor of the disulfide-protein complex with this fragment.



**Figure 7.1 Schematic presentation of the tethering approach.** A cysteine-containing protein is equilibrated with a disulfide-containing mixture in presence of a reducing agent, such as 2-mercaptoethanol. Most of the library members will have little or no inherent affinity for the protein, and thus by HPLC-MS analysis of the equilibrium will lie toward the unmodified protein. However, if a library member experiences inherent affinity for the protein, the equilibrium shifts toward the modified protein. After affinity improvement through decoration with favorable functional groups the fragment can be released from the leash.

Since the substance is covalently attached to the enzyme the modification can easily be detected by HPLC-MS. With respect to the mass spectrometric experiments considering the reaction of the protein with one single disulfide, four deviating mass peaks can principally be expected: unreacted protein, protein + reducing agent, protein + fragment, protein and amine (Figure 7.2). To decide whether any of the fragments in the applied mixture has reacted, the mass of each fragment needs to be unique and favorably deviating by 5 to 10 mass units. Wells et al. have shown that identification of a single compound out of a pool of 100 different fragments is achievable (Erlanson, Braisted et al. 2000). However, mixtures of 10 compounds appear to be most convenient for screening purposes.

Considering the fact that the probability of the formation of disulfide bond between fragment and protein depends on the binding affinity of the non-covalent contact, the success rate of such screening can be adjusted to some degree by the concentration of the added reducing agent (e.g. 0.1 – 1%). Gradually increasing the concentration will successively reduce the formed disulfide bonds between enzyme and fragments and most likely the bond to the fragment with strongest affinity will remain intact up to the highest concentration. Such a sequential analysis to all successfully tethered fragments will help to rank them in terms of binding affinity. At Sunesis, Wells et al. have published successful tethering for more than 10 different drug targets (Erlanson, Wells et al. 2004). To mention a few examples, in case of interleukin-2 (IL-2) a protein-protein interface was targeted by the tethering approach. The chemical information that the experiment provided was applied to existing hits to generate higher-affinity compounds (Braisted, Oslob et al. 2003). Initial tethering hits found for thymidylate synthase could be optimized to result in fragments that also bind sufficiently strong to avoid the covalent attachment (Erlanson, Braisted et al. 2000). Finally, for proteins which only exhibit catalytically relevant cysteine residues close to the active site as attachment point; however, which have to remain untouched, Wells and coworkers suggested introduction of thiole groups beyond the binding pocket to hook up an extender for fragment coupling.



**Figure 7.2** Possible modifications of the protein that can occur along the tethering experiment. The protein mass can be increased by the mass of the fragment, the mass of an amine from the disulfide counterpart or the mass of a DTT molecule.

Tethering provides many opportunities in the field of drug design particularly facilitating the identification and the design of fragments. Compound libraries with more than 10,000 entities have been synthesized at Sunesis to cover chemical space rather exhaustively. Major commitment with respect to the synthesis resources are required which appears rather unpractical in academic research. Therefore, we decided to move part of the library selection into the computer and performed docking

experiments to select the most promising fragments for synthesis from a large virtual library. The disulfides composing the docking library are reduced to a number feasible for synthesis and subsequent HPLC-MS and crystallographic analysis. For our case study carbonic anhydrase II was selected as target protein.

The crystal structure of CA II in complex with Saccharin was recently determined in our group and we could show that two Saccharin molecules bind to CA II (Köhler, Hillebrecht et al. 2007). Apart from a molecule coordinating to the active site zinc ion a second molecule binds to the surface of CA II close to the catalytically relevant proton shuttle His64 (PDB-code: 2Q38). The sulfolactam moiety of Saccharin is fixed via an H-bond network, partially mediated by three water molecules, to six protein residues, among others His64. It is well established that His64 can adopt multiple conformations (in and out conformation) thus being able to shuffle a proton from the catalytic center to the protein solvent (Tu, Silverman et al. 1989; Nair and Christianson 1991b; Liang, E. et al. 1993; Smith, Alexander et al. 1994). Apart from His64, also His3, His4, His10, His15 and His17 are part of an extended histidine cluster which seems to perform the proton transfer. As this step is rate-determining for the catalytic turn-over (Steiner, Jonsson et al. 1975) the binding of compounds such as Saccharin in this area might take influence on the reaction rate. In consequence this additional binding site might emerge as interesting target site for drug design to modulate the properties of CA II. Interestingly enough, it has recently been shown that Timolol, which exhibits activating effects on CO<sub>2</sub> hydration in CA II, is proposed to bind to this site, similarly to Saccharin (Sugimoto, Ikeda et al.). Evidence for this assumption was provided by docking experiments, which were performed with the docking tool Autodock (Morris, Goodsell et al. 1998). The calculated binding mode, however, deviates to some degree from that of well known CA II activators, e.g. histamine or L-histidine. Consistent with the suggested Timolol docking pose, His64 also participates in the coordination of these compounds. Inconsistent with the proposed binding mode of Timolol, histamine and L-histidine are further coordinated by Asn67, Gln92 and via a water network to the zinc ion located in the active site of CA II (Briganti, Mangani et al. 1997; Scozzafava and Supuran 2002; Temperini, Scozzafava et al. 2005; Temperini, Scozzafava et al. 2006a; Temperini, Innocenti et al. 2007). This

described activator binding pocket is proposed to be an important binding site in CAs, in addition to the metal center to coordinate the nucleophile OH<sup>-</sup> and the hydrophobic pocket which recognizes the substrate (Briganti, Mangani et al. 1997).

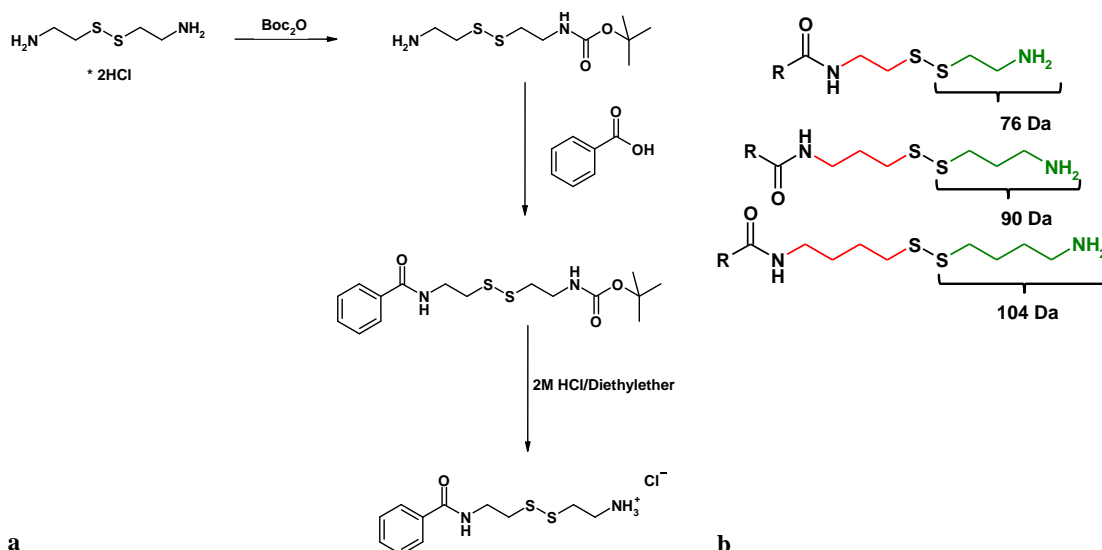
In this study, we performed a tethering approach on CA II to identify fragments with inherent affinity to the protein. Therefore, a cysteine residue was mutationally introduced at the entrance of the active site, replacing His64. This setup should allow exploring the different binding pockets of CA II: the active site zinc ion, the hydrophobic pocket for substrate recognition, the binding site to host CA II activators and the recently discovered Saccharin binding area at the CA II surface. A fragment library which was designed by a computational docking screen was incubated with the mutant and putative disulfide tethering was evaluated by HPLC-MS analysis. Fragments identified by this mass spectrometry study should be subsequently characterized with respect to their binding mode and analyzed in terms of activating or inhibiting properties.

### 7.3 Results

At the entrance of the CA II binding pocket His64 was exchanged by a cysteine residue using site-directed mutagenesis. Prior to the initial tethering procedures a control experiment was performed to verify that Cys64 is modified by such an approach. Apart from the mutationally introduced Cys64, CA II exhibits a second cysteine residue at the protein surface (Cys205). Therefore, the protein was incubated with Ellman's reagent (5,5'-Dithiobis(2-nitrobenzoic acid) - DTNB), a compound that binds virtually to any accessible cysteine residue at the protein surface (s. experimental section 7.4). This experiment revealed that primarily one cysteine residue is modified by DTNB while a second cysteine experiences compound tethering at increased DTNB concentrations. Tryptic digestions revealed that both residues are accessible for DTNB modification. We believe that most CA II-H64C mass modifications observed as a response to our tethering experiments only occur at the Cys64 residue as attached to this residue the fragments can experience interactions to the different binding pockets. A tryptic digestion which in principle can elucidate which cysteine has been modified

is not necessary in the present case. Nevertheless, for all promising fragments the exact binding mode has to be determined via protein crystallography.

In order to allow the covalent attachment of fragments to a cysteine residue, a thiole group has to be coupled to each candidate fragment. Therefore, an appropriate synthetic route had to be established which allows fast and reliable coupling of a broad range of diverse fragments. We applied the synthesis strategy similarly suggested by Wells et al. (Erlanson, Braisted et al. 2000). A cystamine scaffold can easily be linked to putative candidate fragments and additionally assures sufficient solubility in case increased compound concentrations are required, particularly if fragments are tested in terms of well assembled cocktails.



**Figure 7.3** Synthesis of tethering compounds. (a) Synthetic route to asymmetric disulfides. (b) Tethering Scaffolds. R: Fragment (Carboxylic acid); Red: Linker – diverse length with different scope; Green: Fragment counterpart for better solubility.

Insufficient compound solubility is improved by protecting one of the two amine functionalities of the cystamine skeleton along the synthesis (Figure 7.3). The  $(\text{CH}_2)_x$ -linker ( $x = 2-4$ ) between candidate fragment and thiole expands the leash of the tether in order to explore the different cavities in CA II with appropriate linker distance (s. experimental section 7.4). Almost any carboxylic acid can be attached to cystamine by this synthesis. This facilitates the generation of chemically diverse libraries as numerous carboxylic acids are readily available.

Nevertheless, to avoid costly and time consuming synthesis of large compound libraries a pre-selection via docking experiments has been accomplished. A dataset of commercially available carboxylic acids, generously provided by Dr. Christof Wegscheid-Gerlach, represented the starting point of our library design (s. experimental section 7.4) (Gerlach, Münzel et al. 2007). For the docking calculations the carboxylic acids were computationally attached to the thiole group with the desired linker, attached to Cys64 and docked into CA II using the program GOLD. Each generated docking pose was evaluated by a scoring function and subsequently ranked according to their docking score. With regards to this ranking, the most promising candidate fragments were selected by visual inspection, in the following synthesized and applied in the tethering experiments.

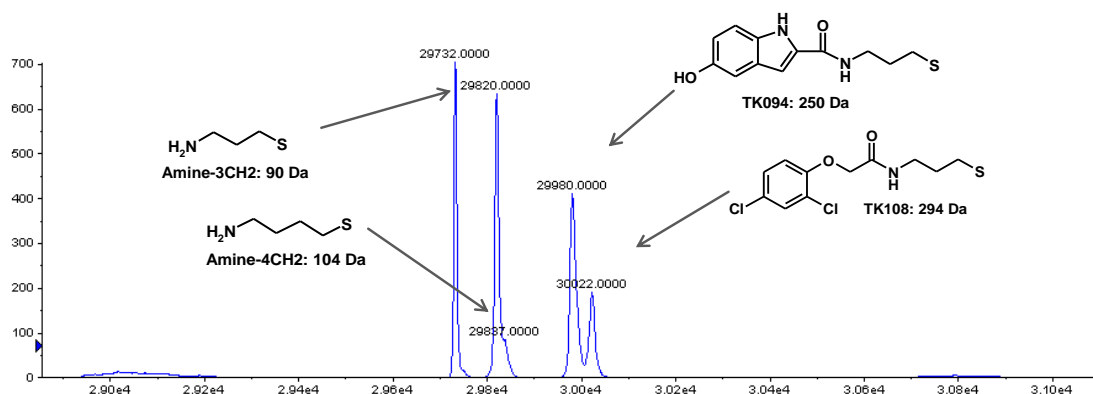
Altogether three fragment libraries were synthesized and evaluated for binding to CA II by means of HPLC-MS and in selected cases by protein crystallography (s. Appendix: tethering-series 1-3). The first series was composed of carboxylic acids available on stock in-house without any pre-selection considerations. This series was synthesized as symmetric disulfides as the more cumbersome asymmetric synthesis was not conceived for our initial experiments. However, candidate fragments from this library often showed poor solubility according to their overall hydrophobic properties. Therefore, only reduced compound concentration could be achieved in solution which can take significant impact on the scope of our experiments. The second series was guided by the results of our docking experiments with CA II. They suggested 14 promising and readily available fragments for synthesis with different linker lengths. In parallel a third series which was synthesized as originally planned for tethering experiments with another protein target (TGT) was also examined.

### **7.3.1 HPLC-MS Experiments**

In order to elucidate the affinity of the candidate fragments to CA II by HPLC-MS the enzyme was incubated with the disulfide under reducing conditions (1 mM DTT) for at least 1 hour at 20 °C. Subsequently, each sample was purified by size exclusion chromatography to remove any unbound fragment and upon denaturation the protein mass was determined.



In a typical mass spectrum of accordingly treated samples the following protein species can be observed: CA II-H64C-unmodified, CA II-H64C-fragment-complex and CA II-H64C-amine-complex (synthetic counterpart; Figure 7.3, green portion). To speed-up the screening attempts and to allow for a rough affinity scan under competitive conditions we did not only screen individual disulfides but also appropriately assembled mixtures or "screening cocktails". According to the applied fragment or fragment cocktails mass shifts of the amine-tethered complexes with a  $(\text{CH}_2)_2$ - (76 Da),  $(\text{CH}_2)_3$ - (90 Da) and  $(\text{CH}_2)_4$ -linker (104 Da) could be detected (Figure 7.4). As the amine fragments show most likely only weak affinity to the target protein, the occurrence of such mass peaks indicates very limited non-covalent interactions for the actual candidate fragment(s) and the protein in the studied cocktail samples.



**Figure 7.4** Example for HPLC-MS analysis of CA II-H64C tethered with a fragment cocktail. CA II-H64C: 29729 Da; CA II-H64C + Amine  $(\text{CH}_2)_3$ : 29729 Da + 90 Da = 29819 Da; CA II-H64C + Amine  $(\text{CH}_2)_4$ : 29729 Da + 104 Da = 29833 Da; CA II-H64C + TK094: 29729 Da + 250 Da = 29979 Da; CA II-H64C + TK108: 29729 Da + 294 Da = 29925 Da.

The binding affinities of all fragments (series 1 - 3) were evaluated in three different HPLC-MS experiments. In the first experiment four to six fragments (minimal mass difference of 10 Da) were mixed at concentrations of 200  $\mu\text{M}$  each and incubated with CA II-H64C. Thus, the strongest binder in each cocktail could be identified. Each candidate fragment was exposed to the protein in two different cocktails to elucidate its properties under varying conditions. In a second experiment all compounds were examined individually to avoid possibly imposed solubility problems in the cocktail or any undesired fragment-fragment interactions. Apart from the mentioned amine-tethered byproducts, the HPLC-MS spectra showed peaks for the unmodified protein and tethered CA II-H64C-candidate fragment complexes. The distribution of the latter

complexes depends on the relative binding affinity of the applied candidate fragments in the composed sample. This strategy allows ranking of all compounds according to the rate by which the protein has been modified. In a third experiment each fragment was incubated with CA II-H64C under the before-mentioned reaction conditions, however now in the presence of the reference compound **BO333** (formula p. 100) - a potent sulfonamide-disulfide-fragment - which should directly interact with the active-site zinc ion. Thereby, the relative ratio of complexes formed as CA II-H64C-fragment to CA II-H64C-**BO333** provides an additional option to estimate the relative compound affinity. The results of all these experiments served as basis for our considerations about the affinity of each fragment and were used for a final ranking of all tested compounds.

**HPLC-MS I:** In Table 7.1 the results of the first series of experiments are listed. 29 disulfide cocktails for tethering have been produced (column 1). Each compound has been present for testing the affinity towards CA II-H64C in at least two independent cocktails. Thereby, deviating rates of tethering can easily be detected. For each cocktail the percentages of fragment-tethered, amine-tethered and unmodified protein are given (column 4). In addition, the protein modification achieved in the second cocktail with the same compound is listed in column 5/6. If a compound name is indicated in bold it shows that this ligand originates from the first cocktail otherwise it indicates the second cocktail as the source (column 2). In red deviating modification rates detected in both cocktails are shown. From all cocktails the fragments with the highest affinity are highlighted in yellow.

Which impact does the individual composition of the cocktail take on the tethering results? In general, similar rates can be observed for each fragment in both differently composed cocktails. A compound which experiences increased affinity to the target protein in the first cocktail is in general also identified in a second cocktail. Similarly, weak or non-binding fragments consistently fail in all studied cocktails. Nevertheless, also deviating results can be observed in both cocktails (red font). This most likely depends on very highly potent compounds present in the sample such as, e.g. **SB114** in mix9 and mix24. In mix9, compound **SB114** can clearly be identified as highly potent

fragment as 65% of the protein is tethered. In addition, the amount of amine-tethered protein is low (15%). Nevertheless, in the second cocktail, mix24, the same fragment does not bind to CA II-H64C at all. This assignment as a "false negative" arises from the fact that **SB102** is present in mix24. **SB102** has been identified as one of the most potent fragments of all tethering series, therefore prohibiting binding of any other compounds to CA II-H64C. This example shows that the composition of a cocktail can take drastic impact on hit identification. Nevertheless, any information about the affinity of the other fragments present in the same sample helps to deconvolute such situations.

There are also examples of easily detectable compounds, e.g. in mix2. In this sample the protein is mainly modified by **SB103** as 71% of this protein-tethered species can be detected. This strong interaction is confirmed by a second cocktail, mix19. There, the same fragment is exclusively covalently attached to the protein (100%). In addition, only 0-15% of amine-tethered protein is detected in the sample (mix2).

In order to elucidate the compound with the highest affinity towards CA II-H64C, five strong binders (**SB100**, **SB114**, **SB103**, **SB102** and **SB227**) from cocktails 1-29 have been selected and mixed with each other (mix30) (Table 7.2). This sample clearly allows a ranking of these compounds. With protein modification of 50%, fragment **SB103** shows the highest affinity of all.

In Table 7.3 fragments with different linker length are compared. Compounds which have been synthesized with different linker length were mixed. The results show that in some cases the compounds seem to bind to the same binding pocket since the modification rates are similar. Nevertheless, in mix34 the tethering-rate deviates by 40% which could indicate that the expanded compound **TK108** is increasingly hindered from binding into the same area. Most likely, it is sterically impossible for this compound to adopt the same orientation as it is possible with the shorter analog. Further crystallographic investigations of these samples have to be performed to elucidate this question.

**HPLC-MS II & III:** In the second and third experiment each compound has been tested individually and in competition with the sulfonamide **BO333**. These experiments should allow a ranking of all compounds to estimate their relative affinity towards CA II-H64C. The results are listed in Table 7.4. For each compound, the percentage of protein-candidate fragment complex, protein-amine complex and unmodified protein is shown in columns 2 and 3. In columns 4 and 5 the results from competition-experiments with **BO333** are displayed. As the sulfonamide experiences high affinity to CA II, high yields of protein-fragment complex formation in such sample indicates strong binding of the corresponding candidate fragment. Candidate fragments with tethering rates of  $\geq 50\%$  in the individual experiment and/or rates of  $\geq 25\%$  in the competition experiment were highlighted in yellow. Except for two cases (**SB124** and **SB233**) these compounds have also been identified as promising candidates in the cocktail experiments. Remarkably, almost all fragments from series 1 show enhanced affinity even though their solubility is partially poor. This correlates with the  $(\text{CH}_2)_2$ -linker which will be explained in the crystallographic section.

## Tethering Approach on Carbonic Anhydrase II Supported by Pre-selecting Computational Docking Experiments

Mix	Compounds	Linker	Rate	2 <sup>nd</sup> Rate	2 <sup>nd</sup> Mix	Mix	Compounds	Linker	Rate	2 <sup>nd</sup> Rate	2 <sup>nd</sup> Mix
1	SB119	(CH <sub>2</sub> ) <sub>2</sub>	25%	15%	Mix28	12	TK110	(CH <sub>2</sub> ) <sub>2</sub>	-	-	Mix3
	SB228	(CH <sub>2</sub> ) <sub>2</sub>	5%	-	Mix14		SB226	(CH <sub>2</sub> ) <sub>2</sub>	25%	30%	Mix5
	SB104	(CH <sub>2</sub> ) <sub>4</sub>	20%	-	Mix12		TK104	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix23
	TK106	(CH <sub>2</sub> ) <sub>4</sub>	5%	-	Mix11		TK102	(CH <sub>2</sub> ) <sub>4</sub>	5%	-	Mix7
	amines unmodified	-	20%	-	-		SB229	(CH <sub>2</sub> ) <sub>2</sub>	30%	15%	Mix2
2	SB127	(CH <sub>2</sub> ) <sub>2</sub>	-	0%	Mix23	13	amines unmodified	-	40%	-	-
	SB103	(CH <sub>2</sub> ) <sub>2</sub>	71%	100%	Mix29		TK112	(CH <sub>2</sub> ) <sub>2</sub>	-	20%	Mix9
	SB233	(CH <sub>2</sub> ) <sub>2</sub>	-	30%	Mix15		SB232	(CH <sub>2</sub> ) <sub>2</sub>	10%	10%	Mix6
	TK107	(CH <sub>2</sub> ) <sub>3</sub>	-	-	Mix13		TK107	(CH <sub>2</sub> ) <sub>3</sub>	-	-	Mix2
	SB229	(CH <sub>2</sub> ) <sub>2</sub>	15%	30%	Mix12		TK101	(CH <sub>2</sub> ) <sub>4</sub>	-	6%	Mix8
3	amines unmodified	-	15%	-	-	SB227	(CH <sub>2</sub> ) <sub>2</sub>	55%	85%	Mix3	
	SB106	(CH <sub>2</sub> ) <sub>2</sub>	10%	20%	Mix25	amines unmodified	-	35%	-	-	
	TK110	(CH <sub>2</sub> ) <sub>2</sub>	-	-	Mix12	14	SB228	(CH <sub>2</sub> ) <sub>2</sub>	-	5%	Mix1
	TK097	(CH <sub>2</sub> ) <sub>3</sub>	-	-	Mix11		TK094	(CH <sub>2</sub> ) <sub>3</sub>	25%	30%	Mix8
	SB099	(CH <sub>2</sub> ) <sub>2</sub>	-	-	Mix27		TK109	(CH <sub>2</sub> ) <sub>3</sub>	-	-	Mix4
SB227	(CH <sub>2</sub> ) <sub>2</sub>	85%	55%	Mix13	TK098		(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix9	
amines unmodified	-	5%	-	-	TK108		(CH <sub>2</sub> ) <sub>3</sub>	20%	-	Mix4	
4	SB095	(CH <sub>2</sub> ) <sub>2</sub>	35%	25%	Mix26	15	amines unmodified	-	55%	-	-
	SB096	(CH <sub>2</sub> ) <sub>2</sub>	13%	0%	Mix22		SB233	(CH <sub>2</sub> ) <sub>2</sub>	30%	-	Mix2
	SB102	(CH <sub>2</sub> ) <sub>2</sub>	40%	100%	Mix24		TK096	(CH <sub>2</sub> ) <sub>3</sub>	-	-	Mix9
	TK109	(CH <sub>2</sub> ) <sub>3</sub>	-	0%	Mix14		TK100	(CH <sub>2</sub> ) <sub>4</sub>	10%	5%	Mix5
	TK108	(CH <sub>2</sub> ) <sub>3</sub>	-	20%	Mix14		TK111	(CH <sub>2</sub> ) <sub>2</sub>	-	-	Mix10
5	amines unmodified	-	12%	-	-	TK105	(CH <sub>2</sub> ) <sub>4</sub>	5%	-	Mix6	
	SB123	(CH <sub>2</sub> ) <sub>2</sub>	15%	-	Mix27	16	amines unmodified	-	55%	-	-
	SB097	(CH <sub>2</sub> ) <sub>2</sub>	15%	-	Mix24		BO333	(CH <sub>2</sub> ) <sub>2</sub>	45%	15%	Mix22
	SB226	(CH <sub>2</sub> ) <sub>2</sub>	30%	40%	Mix12		TK157	(CH <sub>2</sub> ) <sub>3</sub>	-	-	Mix25
	TK100	(CH <sub>2</sub> ) <sub>4</sub>	5%	10%	Mix15		TK148	(CH <sub>2</sub> ) <sub>3</sub>	20%	30%	Mix23
SB124	(CH <sub>2</sub> ) <sub>2</sub>	30%	-	Mix22	TK169		(CH <sub>2</sub> ) <sub>3</sub>	15%	10%	Mix21	
6	amines unmodified	-	5%	-	-	TK173	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix28	
	SB100	(CH <sub>2</sub> ) <sub>2</sub>	55%	30%	Mix28	17	amines unmodified	-	20%	-	-
	SB120	(CH <sub>2</sub> ) <sub>2</sub>	-	20%	Mix25		TK135	(CH <sub>2</sub> ) <sub>2</sub>	5%	-	Mix24
	SB232	(CH <sub>2</sub> ) <sub>2</sub>	20%	10%	Mix13		TK140	(CH <sub>2</sub> ) <sub>2</sub>	10%	15%	Mix23
	TK103	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix11		TK158	(CH <sub>2</sub> ) <sub>3</sub>	-	-	Mix25
TK105	(CH <sub>2</sub> ) <sub>4</sub>	-	5%	Mix15	TK170		(CH <sub>2</sub> ) <sub>4</sub>	70%	-	Mix22	
7	amines unmodified	-	25%	-	-	TK163	(CH <sub>2</sub> ) <sub>3</sub>	-	-	Mix29	
	SB109	(CH <sub>2</sub> ) <sub>2</sub>	10%	0%	Mix22	18	amines unmodified	-	15%	-	-
	SB115	(CH <sub>2</sub> ) <sub>2</sub>	25%	15%	Mix28		TK147	(CH <sub>2</sub> ) <sub>2</sub>	-	-	Mix29
	SB128	(CH <sub>2</sub> ) <sub>2</sub>	10%	0%	Mix29		TK141	(CH <sub>2</sub> ) <sub>2</sub>	25%	20%	Mix27
	TK102	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix12		TK174	(CH <sub>2</sub> ) <sub>4</sub>	-	10%	Mix26
TK099	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix11	TK175		(CH <sub>2</sub> ) <sub>4</sub>	10%	-	Mix24	
8	amines unmodified	-	10%	-	-	TK176	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix21	
	SB231	(CH <sub>2</sub> ) <sub>2</sub>	5%	5%	Mix23	19	amines unmodified	-	65%	-	-
	SB107	(CH <sub>2</sub> ) <sub>2</sub>	40%	0%	Mix29		TK164	(CH <sub>2</sub> ) <sub>4</sub>	10%	5%	Mix21
	TK094	(CH <sub>2</sub> ) <sub>3</sub>	30%	25%	Mix14		TK165	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix26
	TK101	(CH <sub>2</sub> ) <sub>4</sub>	5%	-	Mix13		TK162	(CH <sub>2</sub> ) <sub>3</sub>	50%	10%	Mix27
amines unmodified	-	20%	-	-	TK142		(CH <sub>2</sub> ) <sub>2</sub>	-	5%	Mix25	
9	SB114	(CH <sub>2</sub> ) <sub>2</sub>	65%	-	Mix24	20	amines unmodified	-	40%	-	-
	TK112	(CH <sub>2</sub> ) <sub>2</sub>	20%	-	Mix13		TK149	(CH <sub>2</sub> ) <sub>3</sub>	5%	-	Mix26
	TK096	(CH <sub>2</sub> ) <sub>3</sub>	-	-	Mix15		TK136	(CH <sub>2</sub> ) <sub>2</sub>	20%	25%	Mix21
	TK098	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix14		TK150	(CH <sub>2</sub> ) <sub>3</sub>	15%	5%	Mix28
	amines unmodified	-	15%	-	-		TK171	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix26
10	SB101	(CH <sub>2</sub> ) <sub>2</sub>	65%	25%	Mix27	21	amines unmodified	-	60%	-	-
	SB113	(CH <sub>2</sub> ) <sub>2</sub>	10%	10%	Mix22		SB104	(CH <sub>2</sub> ) <sub>2</sub>	-	-	Mix1
	TK095	(CH <sub>2</sub> ) <sub>3</sub>	-	15%	Mix11		TK164	(CH <sub>2</sub> ) <sub>4</sub>	5%	10%	Mix19
	TK111	(CH <sub>2</sub> ) <sub>2</sub>	-	-	Mix15		TK110	(CH <sub>2</sub> ) <sub>2</sub>	-	-	Mix12
	amines unmodified	-	25%	-	-		TK136	(CH <sub>2</sub> ) <sub>2</sub>	25%	20%	Mix20
11	TK097	(CH <sub>2</sub> ) <sub>3</sub>	-	-	Mix3	TK169	(CH <sub>2</sub> ) <sub>4</sub>	10%	15%	Mix16	
	TK095	(CH <sub>2</sub> ) <sub>3</sub>	15%	-	Mix10	TK176	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix18	
	TK103	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix6	amines unmodified	-	60%	-	-	
	TK106	(CH <sub>2</sub> ) <sub>4</sub>	-	5%	Mix1	amines unmodified	-	-	-	-	
	TK099	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix7						

Tethering Approach on Carbonic Anhydrase II Supported by Pre-selecting Computational Docking Experiments

Mix	Compounds	Linker	Rate	2 <sup>nd</sup> Rate	2 <sup>nd</sup> Mix	Mix	Compounds	Linker	Rate	2 <sup>nd</sup> Rate	2 <sup>nd</sup> Mix
<b>22</b>	BO333	(CH <sub>2</sub> ) <sub>2</sub>	15%	45%	Mix16	<b>26</b>	SB095	(CH <sub>2</sub> ) <sub>2</sub>	25%	25%	Mix4
	SB109	(CH <sub>2</sub> ) <sub>2</sub>	-	12%	Mix7		TK149	(CH <sub>2</sub> ) <sub>3</sub>	-	5%	Mix20
	SB096	(CH <sub>2</sub> ) <sub>2</sub>	-	13%	Mix4		TK165	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix19
	SB113	(CH <sub>2</sub> ) <sub>2</sub>	10%	10%	Mix10		TK174	(CH <sub>2</sub> ) <sub>4</sub>	10%	-	Mix18
	TK170	(CH <sub>2</sub> ) <sub>4</sub>	-	70%	Mix17		TK171	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix20
	SB124	(CH <sub>2</sub> ) <sub>2</sub>	-	30%	Mix5		amines	-	60%	-	-
	amines	-	-	-	-		unmodified	-	-	-	-
unmodified	-	75%	-	-	<b>27</b>	SB123	(CH <sub>2</sub> ) <sub>2</sub>	-	15%	Mix5	
<b>23</b>	SB127	(CH <sub>2</sub> ) <sub>2</sub>	-	-		Mix2	SB101	(CH <sub>2</sub> ) <sub>2</sub>	25%	65%	Mix10
	SB231	(CH <sub>2</sub> ) <sub>2</sub>	5%	5%		Mix8	TK141	(CH <sub>2</sub> ) <sub>2</sub>	20%	25%	Mix18
	TK140	(CH <sub>2</sub> ) <sub>2</sub>	15%	10%		Mix17	TK162	(CH <sub>2</sub> ) <sub>3</sub>	10%	50%	Mix19
	TK148	(CH <sub>2</sub> ) <sub>3</sub>	30%	20%		Mix16	SB099	(CH <sub>2</sub> ) <sub>2</sub>	-	-	Mix3
	TK104	(CH <sub>2</sub> ) <sub>4</sub>	-	-		Mix12	amines	-	45%	-	-
	amines	-	50%	-		-	unmodified	-	-	-	-
	unmodified	-	-	-	-	<b>28</b>	SB100	(CH <sub>2</sub> ) <sub>2</sub>	30%	55%	Mix6
<b>24</b>	TK135	(CH <sub>2</sub> ) <sub>2</sub>	-	5%	Mix17		SB119	(CH <sub>2</sub> ) <sub>2</sub>	15%	25%	Mix1
	SB114	(CH <sub>2</sub> ) <sub>2</sub>	-	65%	Mix9		SB115	(CH <sub>2</sub> ) <sub>2</sub>	15%	25%	Mix7
	SB097	(CH <sub>2</sub> ) <sub>2</sub>	-	15%	Mix5		TK150	(CH <sub>2</sub> ) <sub>3</sub>	5%	15%	Mix20
	SB102	(CH <sub>2</sub> ) <sub>2</sub>	100%	40%	Mix4		TK173	(CH <sub>2</sub> ) <sub>4</sub>	-	-	Mix16
	TK175	(CH <sub>2</sub> ) <sub>4</sub>	-	10%	Mix18		amines	-	25%	-	-
	amines	-	-	-	-		unmodified	-	10%	-	-
	unmodified	-	-	-	-	<b>29</b>	TK147	(CH <sub>2</sub> ) <sub>2</sub>	-	-	Mix18
<b>25</b>	SB106	(CH <sub>2</sub> ) <sub>2</sub>	20%	10%	Mix3		SB103	(CH <sub>2</sub> ) <sub>2</sub>	100%	70%	Mix2
	TK157	(CH <sub>2</sub> ) <sub>3</sub>	-	-	Mix16		SB107	(CH <sub>2</sub> ) <sub>2</sub>	-	40%	Mix8
	SB120	(CH <sub>2</sub> ) <sub>2</sub>	20%	-	Mix6		SB128	(CH <sub>2</sub> ) <sub>2</sub>	-	10%	Mix7
	TK158	(CH <sub>2</sub> ) <sub>3</sub>	-	-	Mix17		TK163	(CH <sub>2</sub> ) <sub>3</sub>	-	-	Mix17
	TK142	(CH <sub>2</sub> ) <sub>2</sub>	5%	5%	Mix19		amines	-	-	-	-
	amines	-	55%	-	-		unmodified	-	-	-	-
	unmodified	-	-	-	-						

**Table 7.1 Fragment cocktails investigated by HPLC-MS.** In each sample the modification in % is shown for all fragments and the unmodified- and amine-modified protein. In column five and six the results for each compound in its second cocktail is shown. A compound name written in bold font indicates its first cocktail and in regular font the second cocktail (2nd column), those with yellow background have been selected as the fragments with highest affinity. Red font fragments show diverse affinity in both measurements. The different results are dependent on other fragments present in the sample. Very high affine compounds are mainly bound and prohibit other compounds in the samples from binding to CA II-H64C.

**Table 7.2 Fragments with highest affinity have been tested in mix30.** SB103 shows the highest affinity to CA II-H64C

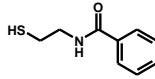
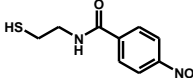
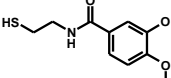
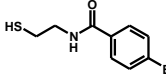
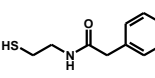
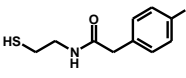
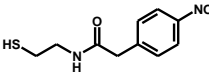
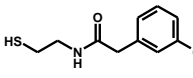
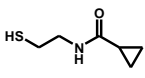
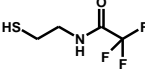
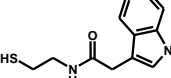
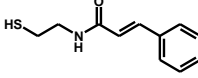
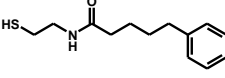
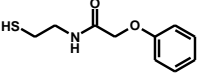
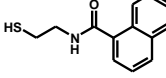
Mix	Compounds	Linker	Rate	1 <sup>st</sup> Rate	1 <sup>st</sup> Mix	2 <sup>nd</sup> Rate	2 <sup>nd</sup> Mix
<b>30</b>	SB100	(CH <sub>2</sub> ) <sub>2</sub>	-	55%	Mix6	30%	Mix28
	SB114	(CH <sub>2</sub> ) <sub>2</sub>	20%	65%	Mix9	-	Mix24
	SB103	(CH <sub>2</sub> ) <sub>2</sub>	50%	71%	Mix2	100%	Mix19
	SB102	(CH <sub>2</sub> ) <sub>2</sub>	25%	40%	Mix4	100%	Mix24
	SB227	(CH <sub>2</sub> ) <sub>2</sub>	5%	85%	Mix3	55%	Mix13
	amines	-	-	-	-	-	-
	unmodified	-	-	-	-	-	-

**Table 7.3** Same fragment with diverse linker lengths.

Mix	Compounds	Linker	Rate	1 <sup>st</sup> Rate	1 <sup>st</sup> Mix	2 <sup>nd</sup> Rate	2 <sup>nd</sup> Mix
31	TK148	(CH <sub>2</sub> ) <sub>3</sub>	35%	20%	Mix16	30%	Mix23
	TK170	(CH <sub>2</sub> ) <sub>4</sub>	40%	70%	Mix17	-	Mix22
	amines	-	25%	-	-	-	-
	unmodified	-	-	-	-	-	-
32	TK136	(CH <sub>2</sub> ) <sub>2</sub>	25%	20%	Mix20	25%	Mix21
	TK150	(CH <sub>2</sub> ) <sub>3</sub>	10%	15%	Mix20	5%	Mix28
	TK173	(CH <sub>2</sub> ) <sub>4</sub>	5%	-	Mix16	-	Mix28
	amines	-	-	-	-	-	-
unmodified	-	-	-	-	-	-	
33	SB233	(CH <sub>2</sub> ) <sub>2</sub>	15%	-	Mix2	30%	Mix15
	TK096	(CH <sub>2</sub> ) <sub>3</sub>	15%	-	Mix9	-	Mix15
	amines	-	70%	-	-	-	-
	unmodified	-	-	-	-	-	-
34	SB229	(CH <sub>2</sub> ) <sub>2</sub>	55%	15%	Mix2	30%	Mix12
	TK108	(CH <sub>2</sub> ) <sub>3</sub>	15%	-	Mix4	20%	Mix14
	amines	-	30%	-	-	-	-
	unmodified	-	-	-	-	-	-
35	TK135	(CH <sub>2</sub> ) <sub>2</sub>	25%	5%	Mix17	-	Mix24
	TK164	(CH <sub>2</sub> ) <sub>4</sub>	15%	10%	Mix19	5%	Mix21
	amines	-	60%	-	-	-	-
	unmodified	-	-	-	-	-	-
36	TK141	(CH <sub>2</sub> ) <sub>2</sub>	25%	25%	Mix18	20%	Mix27
	TK162	(CH <sub>2</sub> ) <sub>3</sub>	15%	50%	Mix19	10%	Mix27
	amines	-	60%	-	-	-	-
	unmodified	-	-	-	-	-	-
37	TK094	(CH <sub>2</sub> ) <sub>3</sub>	30%	30%	Mix8	25%	Mix14
	TK100	(CH <sub>2</sub> ) <sub>4</sub>	15%	5%	Mix5	10%	Mix15
	amines	-	55%	-	-	-	-
	unmodified	-	-	-	-	-	-

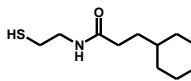
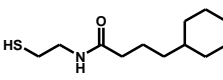
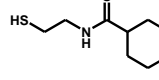
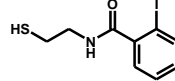
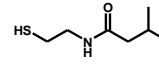
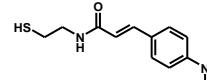
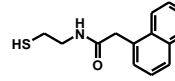
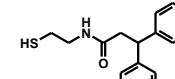
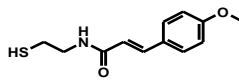
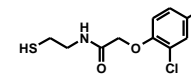
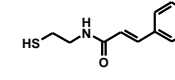
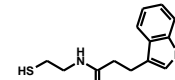
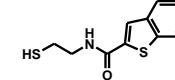
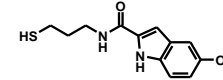
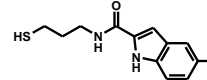
The HPLC-MS experiments cannot be applied to give an absolute statement on the affinity of each candidate fragment. Still, these in solution experiments can provide initial information about the protein-ligand interaction and in most cases reveals consistent results in all experiments. In addition, the sum of all experiments provides a satisfying overview of such large library in an acceptable amount of time. In order to understand the estimated affinity the binding mode of selected fragments should be investigated by protein crystallography. All identified candidate fragments should be cocrystallized with CA II-H64C and the data sets collected according to a priority list derived from the affinity estimated by HPLC-MS.

Tethering Approach on Carbonic Anhydrase II Supported by Pre-selecting Computational Docking Experiments

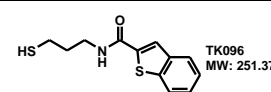
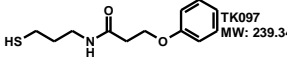
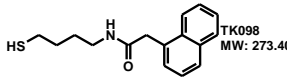
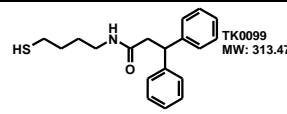
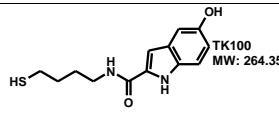
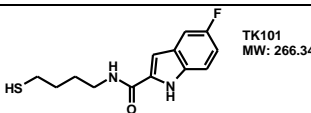
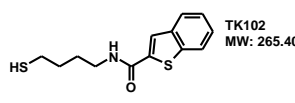
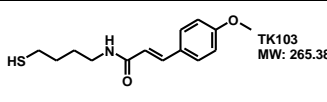
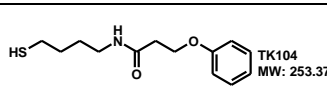
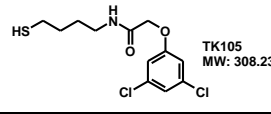
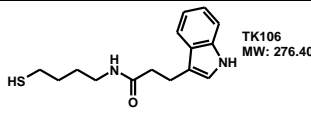
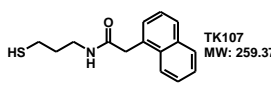
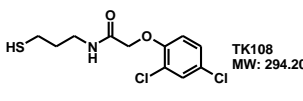
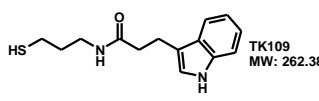
Compound	Individual	Rate	Competition with BO333	Rate	Structure
<b>SB095</b> (CH <sub>2</sub> ) <sub>2</sub>	SB095 no modification amine	100% - -	SB095 no modification amine	55% - -	 SB095 MW: 181.26
series 1	-	-	BO333	45%	
<b>SB096</b> (CH <sub>2</sub> ) <sub>2</sub>	SB096 no modification amine	100% - -	SB096 no modification amine	30% - -	 SB096 MW: 226.26
series 1	-	-	BO333	70%	
<b>SB097</b> (CH <sub>2</sub> ) <sub>2</sub>	SB097 no modification amine	85% 15% -	SB097 no modification amine	40% - -	 SB097 MW: 227.28
series	-	-	BO333	60%	
<b>SB099</b> (CH <sub>2</sub> ) <sub>2</sub>	SB099 no modification amine	- 100% -	SB099 no modification amine	- - -	 SB099 MW: 260.15
series 1	-	-	BO333	100%	
<b>SB100</b> (CH <sub>2</sub> ) <sub>2</sub>	SB100 no modification amine	85% 15% -	SB100 no modification amine	40% - -	 SB100 MW: 195.29
series 1	-	-	BO333	60%	
<b>SB101</b> (CH <sub>2</sub> ) <sub>2</sub>	SB101 no modification amine	100% - -	SB101 no modification amine	25% - -	 SB101 MW: 213.28
series 1	-	-	BO333	75%	
<b>SB102</b> (CH <sub>2</sub> ) <sub>2</sub>	SB102 no modification amine	100% - -	SB102 no modification amine	15% - -	 SB102 MW: 240.28
series 1	-	-	BO333	85%	
<b>SB103</b> (CH <sub>2</sub> ) <sub>2</sub>	SB103 no modification amine	100% - -	SB103 no modification amine	55% - -	 SB103 MW: 225.31
series 1	-	-	BO333	45%	
<b>SB104</b> (CH <sub>2</sub> ) <sub>2</sub>	SB104 no modification amine	70% 30% -	SB104 no modification amine	- - -	 SB104 MW: 145.22
series 1	-	-	BO333	100%	
<b>SB106</b> (CH <sub>2</sub> ) <sub>2</sub>	SB106 no modification amine	100% - -	SB106 no modification amine	35% - -	 SB106 MW: 173.16
series 1	-	-	BO333	65%	
<b>SB107</b> (CH <sub>2</sub> ) <sub>2</sub>	SB107 no modification amine	100% - -	SB107 no modification amine	50% - -	 SB107 MW: 234.32
series 1	-	-	BO333	50%	
<b>SB109</b> (CH <sub>2</sub> ) <sub>2</sub>	SB109 no modification amine	60% 40% -	SB109 no modification amine	30% - -	 SB109 MW: 207.30
series 1	-	-	BO333	70%	
<b>SB113</b> (CH <sub>2</sub> ) <sub>2</sub>	SB113 no modification amine	65% 35% -	SB113 no modification amine	40% - -	 SB113 MW: 237.37
series 1	-	-	BO333	60%	
<b>SB114</b> (CH <sub>2</sub> ) <sub>2</sub>	SB114 no modification amine	100% - -	SB114 no modification amine	65% - -	 SB114 MW: 211.28
series 1	-	-	BO333	35%	
<b>SB115</b> (CH <sub>2</sub> ) <sub>2</sub>	SB115 no modification amine	100% - -	SB115 no modification amine	55% - -	 SB115 MW: 231.32
series 1	-	-	BO333	45%	



Tethering Approach on Carbonic Anhydrase II Supported by Pre-selecting Computational Docking Experiments

Compound	Individual	Rate	Competition with BO333	Rate	Structure
<b>SB119</b> (CH <sub>2</sub> ) <sub>2</sub>	SB119 no modification amine	65% 35% -	SB119 no modification amine	15% -	
series 1	-	-	BO333	85%	
<b>SB120</b> (CH <sub>2</sub> ) <sub>2</sub>	SB120 no modification amine	- 100% -	SB120 no modification amine	- 85% -	
series 1	-	-	BO333	15%	
<b>SB123</b> (CH <sub>2</sub> ) <sub>2</sub>	SB123 no modification amine	10% 90% -	SB123 no modification amine	10% 70% -	
series 1	-	-	BO333	20%	
<b>SB124</b> (CH <sub>2</sub> ) <sub>2</sub>	SB124 no modification amine	- 100% -	SB124 no modification amine	5% 80% -	
series 1	-	-	BO333	15%	
<b>SB127</b> (CH <sub>2</sub> ) <sub>2</sub>	SB127 no modification amine	- 100% -	SB127 no modification amine	- 70% -	
series 1	-	-	BO333	30%	
<b>SB128</b> (CH <sub>2</sub> ) <sub>2</sub>	SB128 no modification amine	10% 90% -	SB128 no modification amine	15% 55% -	
series 3	-	-	BO333	30%	
<b>SB226</b> (CH <sub>2</sub> ) <sub>2</sub>	SB226 no modification amine	80% 20% -	SB226 no modification amine	30% -	
Series 3	-	-	BO333	50%	
<b>SB227</b> (CH <sub>2</sub> ) <sub>2</sub>	SB227 no modification amine	60% -	SB227 no modification amine	40% -	
series 3	-	40%	BO333	30%	
<b>SB228</b> (CH <sub>2</sub> ) <sub>2</sub>	SB228 no modification amine	10% -	SB228 no modification amine	10% -	
Series 2	-	90%	BO333	40%	
<b>SB229</b> (CH <sub>2</sub> ) <sub>2</sub>	SB229 no modification amine	80% -	SB229 no modification amine	30% -	
series 3	-	20%	BO333	30%	
<b>SB231</b> (CH <sub>2</sub> ) <sub>2</sub>	SB231 no modification amine	50% -	SB231 no modification amine	30% -	
series 3	-	50%	BO333	15%	
<b>SB232</b> (CH <sub>2</sub> ) <sub>2</sub>	SB232 no modification amine	40% -	SB232 no modification amine	15% -	
series 3	-	60%	BO333	30%	
<b>SB233</b> (CH <sub>2</sub> ) <sub>2</sub>	SB233 no modification amine	25% -	SB233 no modification amine	20% -	
series 3	-	75%	BO333	45%	
<b>TK094</b> (CH <sub>2</sub> ) <sub>3</sub>	TK094 no modification amine	45% -	T094 no modification amine	30% -	
series 3	-	55%	BO333	30%	
<b>TK095</b> (CH <sub>2</sub> ) <sub>4</sub>	TK095 no modification amine	50% -	TK095 no modification amine	30% -	
series	-	50%	BO333	25%	
		-		45%	

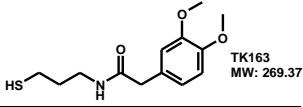
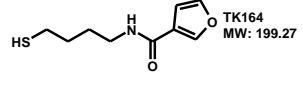
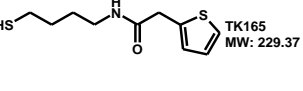
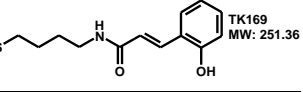
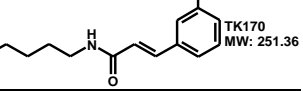
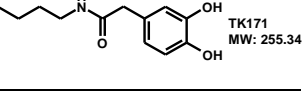
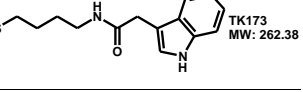
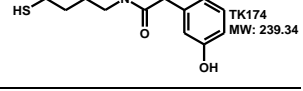
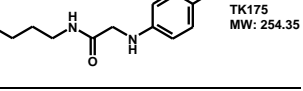
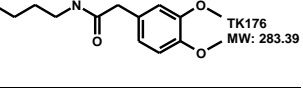
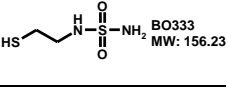
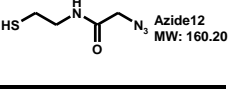
Tethering Approach on Carbonic Anhydrase II Supported by Pre-selecting Computational Docking Experiments

Compound	Individual	Rate	Competition with BO333	Rate	Structure
<b>TK096</b> (CH <sub>2</sub> ) <sub>3</sub>	TK096 no modification	45%	TK096	25%	 TK096 MW: 251.37
series 3	amine	-	no modification	-	
	-	55%	amine	35%	
<b>TK097</b> (CH <sub>2</sub> ) <sub>3</sub>	TK097 no modification	20%	TK097	10%	 TK097 MW: 239.34
Series 3	amine	10%	no modification	-	
	-	70%	amine	45%	
<b>TK098</b> (CH <sub>2</sub> ) <sub>4</sub>	TK098 no modification	5%	TK098	5%	 TK098 MW: 273.40
series 3	amine	80%	no modification	-	
	-	15%	amine	20%	
<b>TK099</b> (CH <sub>2</sub> ) <sub>4</sub>	TK099 no modification	10%	TK099	15%	 TK099 MW: 313.47
series 3	amine	80%	no modification	10%	
	-	10%	amine	15%	
<b>TK100</b> (CH <sub>2</sub> ) <sub>4</sub>	TK100 no modification	5%	TK100	15%	 TK100 MW: 264.35
Series 3	amine	85%	no modification	30%	
	-	10%	amine	10%	
<b>TK101</b> (CH <sub>2</sub> ) <sub>4</sub>	TK101 no modification	7%	TK101	7%	 TK101 MW: 266.34
series 3	amine	93%	no modification	73%	
	-	-	amine	-	
<b>TK102</b> (CH <sub>2</sub> ) <sub>4</sub>	TK102 no modification	-	TK102	-	 TK102 MW: 265.40
series 3	amine	100%	no modification	70%	
	-	-	amine	-	
<b>TK103</b> (CH <sub>2</sub> ) <sub>4</sub>	TK103 no modification	-	TK103	-	 TK103 MW: 265.38
series 3	amine	100%	no modification	75%	
	-	-	amine	-	
<b>TK104</b> (CH <sub>2</sub> ) <sub>4</sub>	TK104 no modification	-	TK104	5%	 TK104 MW: 253.37
series 3	amine	100%	no modification	60%	
	-	-	amine	5%	
<b>TK105</b> (CH <sub>2</sub> ) <sub>4</sub>	TK105 no modification	-	TK105	10%	 TK105 MW: 308.23
series 3	amine	0%	no modification	65%	
	-	-	amine	25%	
<b>TK106</b> (CH <sub>2</sub> ) <sub>4</sub>	TK106 no modification	10%	TK106	15%	 TK106 MW: 276.40
series 3	amine	90%	no modification	55%	
	-	-	amine	-	
<b>TK107</b> (CH <sub>2</sub> ) <sub>3</sub>	TK107 no modification	20%	TK107	25%	 TK107 MW: 259.37
series 3	amine	50%	no modification	10%	
	-	30%	amine	20%	
<b>TK108</b> (CH <sub>2</sub> ) <sub>3</sub>	TK108 no modification	40%	TK108	35%	 TK108 MW: 294.20
series 3	amine	40%	no modification	10%	
	-	20%	amine	15%	
<b>TK109</b> (CH <sub>2</sub> ) <sub>3</sub>	TK109 no modification	30%	TK109	20%	 TK109 MW: 262.38
series 3	amine	35%	no modification	10%	
	-	35%	amine	25%	
			BO333	45%	

Tethering Approach on Carbonic Anhydrase II Supported by Pre-selecting Computational Docking Experiments

Compound	Individual	Rate	Competition with BO333	Rate	Structure
<b>TK110</b> (CH <sub>2</sub> ) <sub>2</sub>	TK110 no modification amine	- 0% -	TK110 no modification amine	- 80% -	
series 3	-	-	BO333	20%	
<b>TK111</b> (CH <sub>2</sub> ) <sub>2</sub>	TK111 no modification amine	20% 30% 50%	TK111 no modification amine	10% 15% 20%	
series 3z	-	-	BO333	55%	
<b>TK112</b> (CH <sub>2</sub> ) <sub>2</sub>	TK112 no modification amine	35% 10% 55%	TK112 no modification amine	25% -	
series 3	-	-	BO333	50%	
<b>TK135</b> (CH <sub>2</sub> ) <sub>2</sub>	TK135 no modification amine	30% -	TK135 no modification amine	50% -	
series 2	-	70% -	BO333	20% 30%	
<b>TK136</b> (CH <sub>2</sub> ) <sub>2</sub>	TK136 no modification amine	50% -	TK136 no modification amine	30% -	
series 2	-	50% -	BO333	25% 45%	
<b>TK140</b> (CH <sub>2</sub> ) <sub>2</sub>	TK140 no modification amine	40% -	TK140 no modification amine	15% -	
series 2	-	60% -	BO333	35% 50%	
<b>TK141</b> (CH <sub>2</sub> ) <sub>2</sub>	TK141 no modification amine	50% -	TK141 no modification amine	25% -	
series 2	-	50% -	BO333	35% 40%	
<b>TK142</b> (CH <sub>2</sub> ) <sub>2</sub>	TK142 no modification amine	25% -	TK142 no modification amine	10% -	
series 2	-	75% -	BO333	40% 50%	
<b>TK147</b> (CH <sub>2</sub> ) <sub>2</sub>	TK147 no modification amine	30% -	TK147 no modification amine	10% -	
series 2	-	70% -	BO333	50% 40%	
<b>TK148</b> (CH <sub>2</sub> ) <sub>3</sub>	TK148 no modification amine	90% -	TK148 no modification amine	35% -	
series 2	-	10% -	BO333	10% 35%	
<b>TK149</b> (CH <sub>2</sub> ) <sub>3</sub>	TK149 no modification amine	- 100% -	TK149 no modification amine	10% 35% 20%	
series 2	-	-	BO333	35%	
<b>TK150</b> (CH <sub>2</sub> ) <sub>3</sub>	TK150 no modification amine	10% -	TK150 no modification amine	20% -	
series 2	-	90% -	BO333	45% 10% 25%	
<b>TK157</b> (CH <sub>2</sub> ) <sub>3</sub>	TK157 no modification amine	- 100% -	TK157 no modification amine	- 45% 15%	
series 2	-	-	BO333	40%	
<b>TK158</b> (CH <sub>2</sub> ) <sub>3</sub>	TK158 no modification amine	25% 70% 5%	TK158 no modification amine	40% 15% 15%	
series 2	-	-	BO333	30%	
<b>TK162</b> (CH <sub>2</sub> ) <sub>3</sub>	TK162 no modification amine	25% 65% 10%	TK162 no modification amine	40% 15% 15%	
series 2	-	-	BO333	30%	

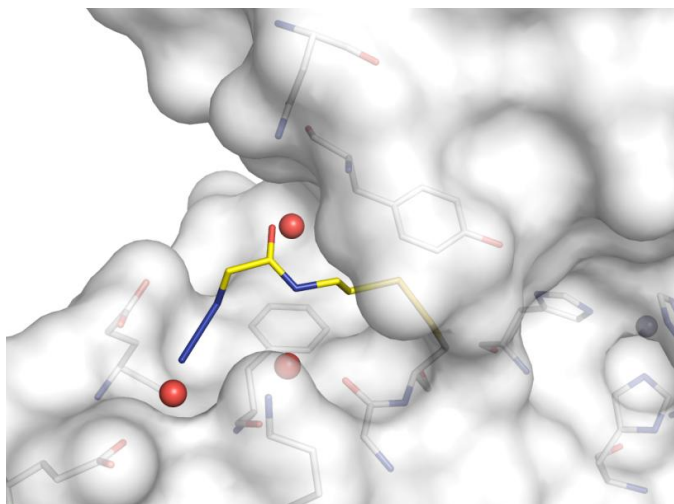
Tethering Approach on Carbonic Anhydrase II Supported by Pre-selecting Computational Docking Experiments

Compound	Individual	Rate	Competition with BO333	Rate	Structure
<b>TK163</b> (CH <sub>2</sub> ) <sub>3</sub>	TK163 no modification	10%	TK163	20%	
series 2	-	75%	no modification	20%	
	amine	15%	amine	20%	
		-	BO333	40%	
<b>TK164</b> (CH <sub>2</sub> ) <sub>4</sub>	TK164 no modification	10%	TK164	10%	
series 2	-	90%	no modification	45%	
	amine	-	amine	45%	
		-	BO333	-	
<b>TK165</b> (CH <sub>2</sub> ) <sub>4</sub>	TK165 no modification	5%	TK165	10%	
series 2	-	95%	no modification	40%	
	amine	-	amine	-	
		-	BO333	50%	
<b>TK169</b> (CH <sub>2</sub> ) <sub>4</sub>	TK169 no modification	10%	TK169	5%	
series 2	-	90%	no modification	35%	
	amine	-	amine	5%	
		-	BO333	55%	
<b>TK170</b> (CH <sub>2</sub> ) <sub>4</sub>	TK170 no modification	95%	TK170	20%	
series 2	-	5%	no modification	25%	
	amine	-	amine	5%	
		-	BO333	50%	
<b>TK171</b> (CH <sub>2</sub> ) <sub>4</sub>	TK171 no modification	60%	TK171	10%	
series	-	25%	no modification	-	
	amine	15%	amine	-	
		-	BO333	90%	
<b>TK173</b> (CH <sub>2</sub> ) <sub>4</sub>	TK173 no modification	5%	TK173	10%	
series 2	-	95%	no modification	35%	
	amine	-	amine	5%	
		-	BO333	50%	
<b>TK174</b> (CH <sub>2</sub> ) <sub>4</sub>	TK174 no modification	25%	TK174	5%	
series 2	-	75%	no modification	5%	
	amine	-	amine	-	
		-	BO333	90%	
<b>TK175</b> (CH <sub>2</sub> ) <sub>4</sub>	TK175 no modification	-	TK175	-	
Series 2	-	-	no modification	-	
	amine	100%	amine	-	
		-	BO333	-	
<b>TK176</b> (CH <sub>2</sub> ) <sub>4</sub>	TK176 no modification	-	TK176	10%	
Series 2	-	100%	no modification	25%	
	amine	-	amine	-	
		-	BO333	65%	
<b>BO333</b> (CH <sub>2</sub> ) <sub>2</sub>	BO333 no modification	100%	BO333	100%	
	amine	-	amine	-	
		-	BO333	-	
<b>Azide12</b> (CH <sub>2</sub> ) <sub>2</sub>	Azide 12 no modification	75%	Azide 12	40%	
	amine	25%	no modification	-	
		-	amine	-	
		-	-	60%	

**Table 7.4** All tethering compounds tested individually and in competition with reference BO333 for affinity to CA II-H64C. Candidate fragments with affinity to CA II are highlighted in yellow.

### 7.3.2 Protein Crystallographic Experiments

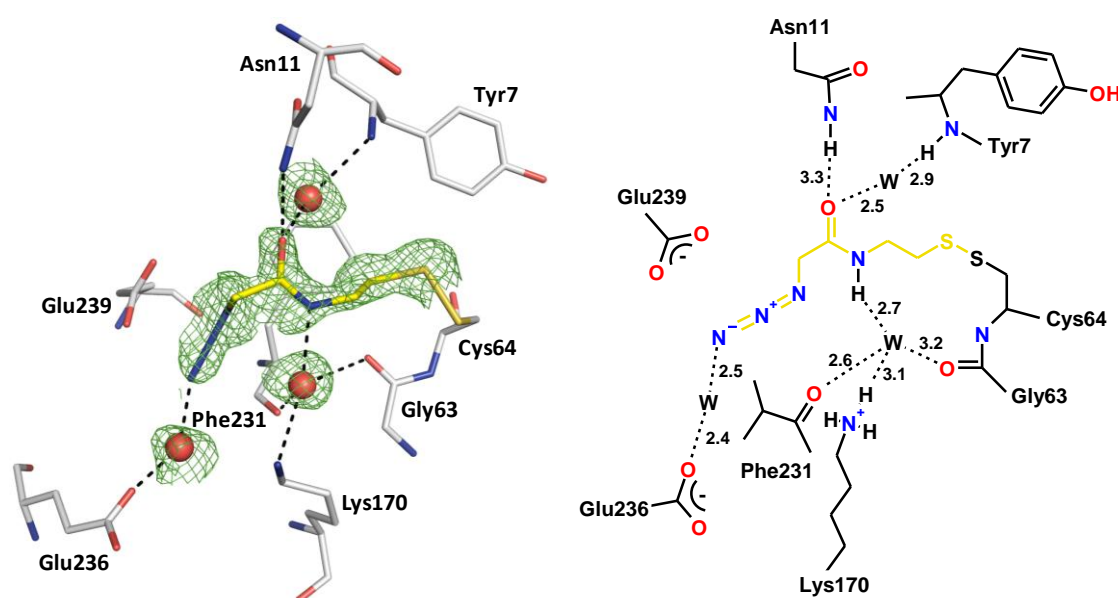
The evaluation of the HPLC-MS experiments revealed that several candidate fragments from the disulfide library with a  $(\text{CH}_2)_2$ -linker experience remarkable affinity to CA II-H64C. The explanation for this elevated hit rate is provided by the determination of the crystal structure of CA II-H64C in covalent complex with **Azide12** and **SB095**.



**Figure 7.5** **Azide12** covalently bound to CA II-H64C (PDB-code: 3KIG). The solvent accessible surface is presented in white. The interacting protein amino acids are shown in stick representation (protein: C, N, O, S; ligand: C, N, O, S). The protein is shown in an orientation with the proton shuttle His64 (here mutated to Cys; next to the protein gap) on the left hand side. Unless it is mentioned differently this mode of representation is applied in all figures with protein surfaces in this chapter. The ligand adopts a hydrophilic canyon at the surface of CA II-H64C. An H-bond network between the fragment and the protein is partially mediated by three water molecules.

**Crystal structure of CA II in complex with Azide12:** The crystal structure of **Azide12** tethered to CA II-H64C was determined at a resolution of 1.39 Å (PDB-Code 3KIG). The azide component covalently attached via a propylene disulfide linker to Cys64 occupies a surface exposed hydrophilic canyon next to the exit of the catalytic site of CA II (Figure 7.5). The carboxyl oxygen of **Azide12** participates in an H-bond network with Asn11 (3.3 Å) and Tyr7. The latter interaction is mediated by a water molecule with distances of 2.5 Å and 2.9 Å (Figure 7.6). The NH of **Azide12** is connected to a water molecule (2.7 Å) which is further bound to Gly63 (3.2 Å), Lys170 (3.1 Å) and Phe231 (2.6 Å). The azide moiety forms a hydrogen bond to Glu236 which is also mediated by a water molecule (2.5 Å; 2.4 Å). The azide group falls into close vicinity to the terminal oxygens of Glu239 indicating possible polar

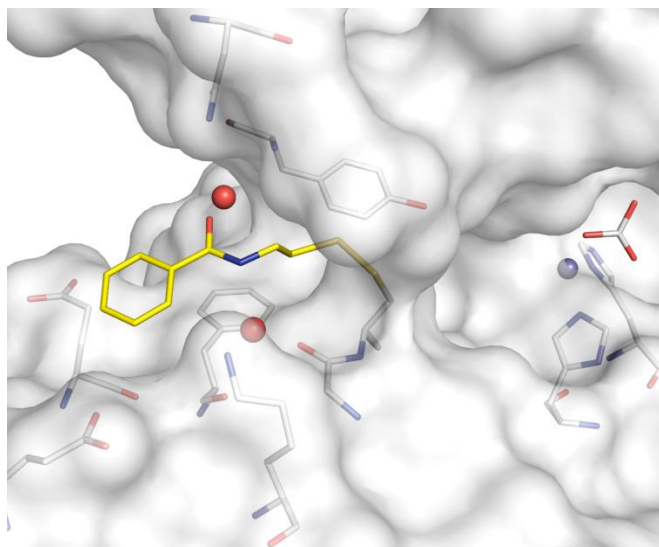
interactions. In addition to the described hydrogen bonds the inhibitor shows 29 van der Waals interactions to neighboring amino acids. The binding mode of **Azide12** at the surface of CA II is unambiguously defined according to a very well defined difference electron density ( $F_o-F_c$ -density: 2.0 sigma). However, the ligand exhibits some residual mobility which is indicated by the average overall ligand  $B$ -factor of  $31.5 \text{ \AA}^2$ . This mobility increases towards the terminal atoms of **Azide12** from  $21.8 \text{ \AA}^2$  for the sulfur to  $44.2 \text{ \AA}^2$  for the terminal nitrogen. A similar  $B$ -value distribution can be observed for the coordinating water molecules ( $26.5 \text{ \AA}^2$ ,  $27.2 \text{ \AA}^2$  for carboxylate-coordinating waters and  $42.2 \text{ \AA}^2$  for azide-coordinating water). Across the protein, the average  $B$ -values are for main-chain atoms  $13.2 \text{ \AA}^2$ , side-chain atoms  $17.8 \text{ \AA}^2$ , water molecules  $26.6 \text{ \AA}^2$  and zinc ion  $19.4 \text{ \AA}^2$ .



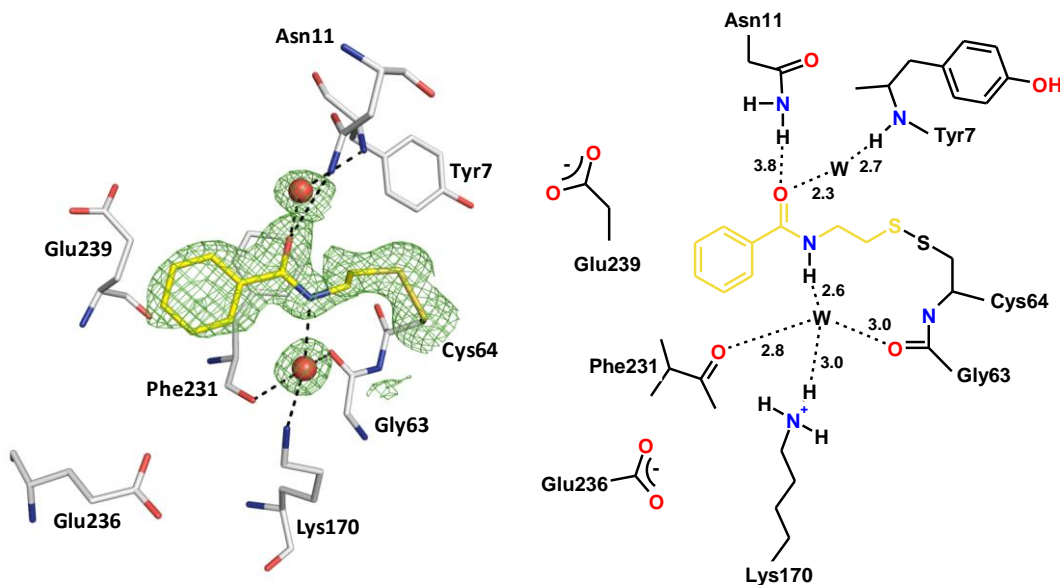
**Figure 7.6 Binding mode of Azide12 (PDB-code: 3KIG).** On the left hand side the crystal structure is shown with the important amino acids. The protein residues are shown in stick representation (protein: C, N, O, S; ligand: C, N, O, S). Water molecules are shown as red spheres. Hydrogen bond interactions are indicated by the dashed lines. The equivalent distances are shown in the schematic representation of the binding mode on the right hand side. The difference electron density ( $F_o-F_c$ ) for the ligand and directly interaction water molecules is shown at a sigma level of 2.0. Atomic distances are shown in  $\text{\AA}$ . Unless it is mentioned differently this mode of representation is applied in all binding mode figures in this chapter. The interaction to the residues Tyr7, Asn11, Gly63, Lys170, Phe231, Glu236 and Glu239 is partially mediated by three water molecules.

**Crystal structure of CA II in complex with SB095:** The crystal structure of **SB095** tethered to CA II-H64C was determined at a resolution of  $1.70 \text{ \AA}$  (PDB-Code 3M2Z).

The fragment is covalently attached to Cys64 and adopts a binding mode similar to **Azide12** (Figure 7.7).



**Figure 7.7** SB095 covalently tethered to CA II-H64C (protein: C, N, O, S; ligand: C, N, O, S) (PDB-Code: 3M2Z). The ligand accommodates a hydrophilic canyon at the surface of CA II-H64C. An H-bond network between fragment and protein is partially mediated by two water molecules.



**Figure 7.8** Binding mode of SB095 tethered to CA II-H64C (protein: C, N, O, S; ligand: C, N, O, S; PDB-code: 3M2Z). The interactions to the residues Tyr7, Asn11, Gly63, Lys170 and Phe231 is partially mediated by two water molecules.

The carboxyl oxygen is linked to Tyr7 by an hydrogen bond mediated via a water molecule (2.3 Å; 2.7 Å) (Figure 7.8). The terminal NH of Asn11 is positioned at 3.8 Å from this ligand oxygen which can be still seen as long hydrogen bond interaction. In

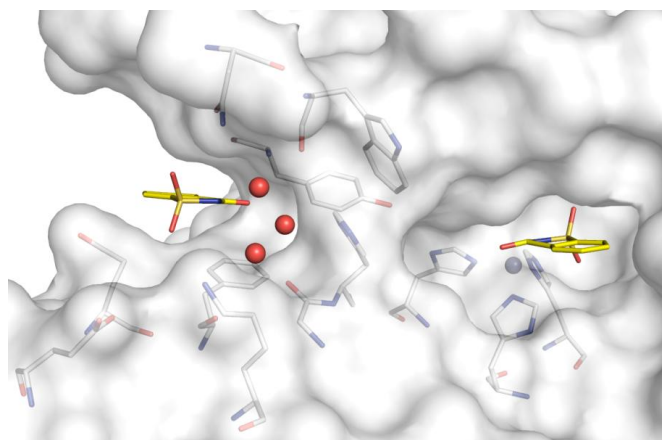
addition, the NH of **SB095** forms an H-bond to a second water which is coordinated to Gly63 (3.0 Å), Lys170 (3.0 Å) and Phe231 (2.8 Å). Besides these hydrogen bonds, the benzyl group performs hydrophobic interactions to Phe231. Calculations revealed that **SB095** experiences 29 van der Waals interactions. The binding mode is very well described by the difference electron density ( $F_o-F_c$ -density: 2.0 sigma). Average *B*-values of 25.9 Å<sup>2</sup> for the ligand and the coordinating water molecules (22.4 Å<sup>2</sup>; 27.7 Å<sup>2</sup>) indicate little mobility of the tethered fragment. The mean *B*-values across the protein are: main-chain atoms 13.7 Å<sup>2</sup>, side-chain atoms 17.5 Å<sup>2</sup>, water molecules 25.3 Å<sup>2</sup> and zinc ion 39.5 Å<sup>2</sup>.

**Crystal structure of CA II in complex with Saccharin:** This binding site for fragments **SB095** and **Azide12** has been observed as occupied previously in our group by determining a crystal structure of CA II in complex with **Saccharin** (Figure 7.9). **Saccharin** binds to CA II with a stoichiometry of 2:1. One molecule binds to the active site, coordinating with its sulfolactam nitrogen the Zn<sup>2+</sup> ion and in parallel a second molecule is fixed at the enzyme's surface in the neighborhood of His64 and Tyr7. **Saccharin** is bound to this second binding site by means of four hydrogen bonds (Figure 7.10). The carboxyl oxygen of the sulfolactam performs an H-bond to the backbone NH of Tyr7 with a distance of 2.7 Å and in addition one sulfonyl oxygen is connected to the terminal NH of Asn11 (3.2 Å). Furthermore, mediated via three water molecules the sulfolactam nitrogen participates in an H-bond network to Trp5, His64, Gly63 and Phe231. Furthermore, Phe231 performs hydrophobic  $\pi$ - $\pi$ -interactions to the aromatic system of **Saccharin**. 38 van der Waals interactions have been calculated for the surface-bound **Saccharin** molecule. The mean *B*-value indicates slight mobility of this ligand. The following averaged *B*-values have been observed: main-chain atoms 13.9 Å<sup>2</sup>, side-chain atoms 14.5 Å<sup>2</sup>, water molecules 23.9 Å<sup>2</sup>, **Saccharin** 30.5 Å<sup>2</sup> and zinc ion 29.4 Å<sup>2</sup>.

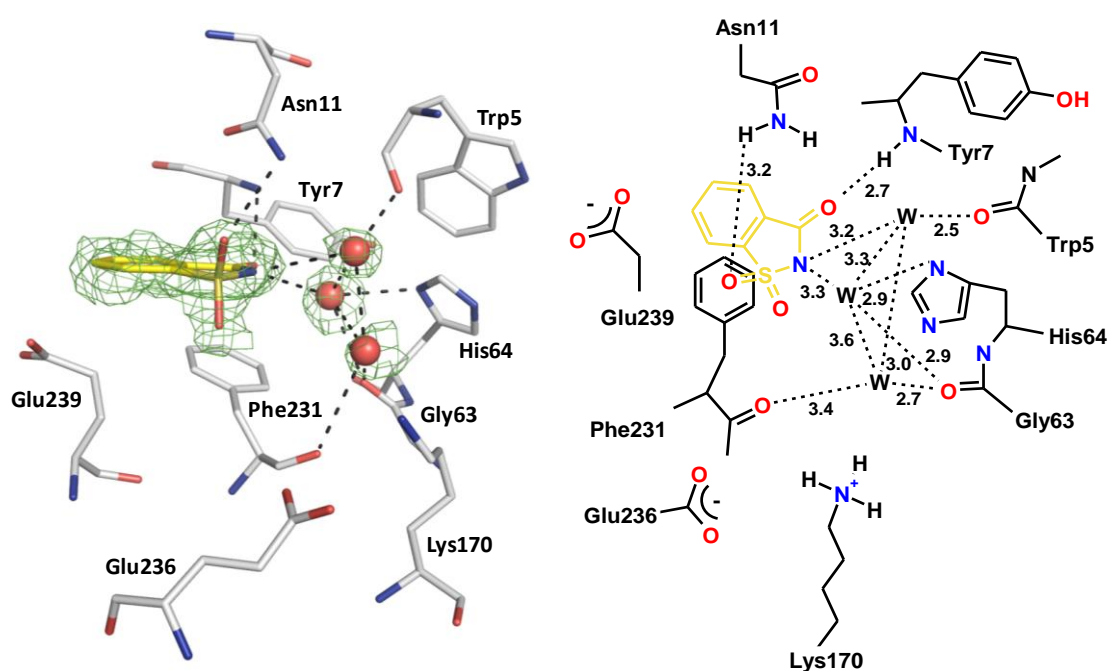
The superposition of **Saccharin** with the two tethered fragments **SB095** and **Azide12** reveals related binding modes. Asn11, Tyr7 and Gly63 are participating in the binding of all compounds (Figure 7.11). Furthermore, in all three complexes water molecules



are essential for fragment binding affinity as they mediate protein-to-ligand interactions.



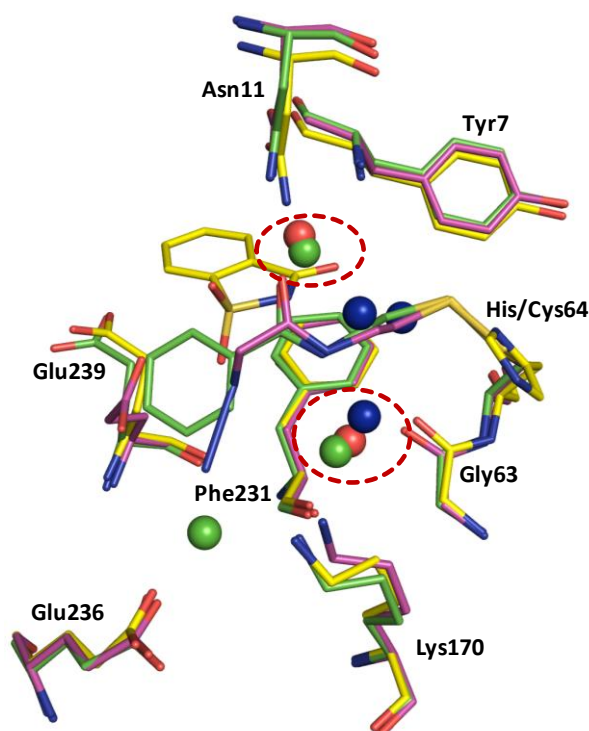
**Figure 7.9** Binding mode of Saccharin in complex with CA II (protein: C, N, O, S; ligand: C, N, O, S; PDB-Code: 2Q38). Saccharin binds to CA II with a stoichiometry of 2:1. The first molecule is coordinating to the active site zinc ion, and a second molecule is bound to the surface of CA II close to the His64 residue.



**Figure 7.10** Binding mode of Saccharin at the surface of CA II (protein: C, N, O, S; ligand: C, N, O, S; PDB-code: 2Q38). The interactions to the residues Tyr7, Asn11, Gly63, His64, and Phe231 are partially mediated via three water molecules.

The carboxyl oxygen of the **Saccharin** sulfolactam and one of the coordinating water molecules in the **SB095** and **Azide12** complexes adopt similar positions linking the fragments to the backbone NH of Tyr7. In addition, the interaction to the backbone

oxygen of Phe231, to the terminal nitrogen of Lys170 and to the backbone oxygen of Gly63 is in all three complexes mediated by a water molecule which is similarly located in all three structures. The binding affinity of the three compounds appears rather weak and full occupancy could not be achieved as indicated by the ligand-*B*-values assigned by refinement. Nevertheless, their binding modes appear quite similar and suggest interesting complex structures to be studied by linking fragments via a (CH<sub>2</sub>)<sub>2</sub>-tether.



**Figure 7.11** Superposition of CA II complex structures with Saccharin (C, N, O, S; waters as blue spheres), SB095 (C, N, O, S; waters as red spheres) and Azide12 (C, N, O, S; waters as green spheres). The analog interacting protein residues are colored in same fashion as the bound ligand. Similar interaction patterns are indicated by red ellipses.

The described crystal structures reveal that this tethering approach easily allows the identification of fragments binding into this area. Glu239 is available for additional binding interaction, that possibly increase fragment affinity. Decorating the phenyl ring of **SB095** with an H-bond donor functionality should increase the affinity thereby allowing to subsequently release the fragment from its covalent leash and to inhibit the proton shuttle.

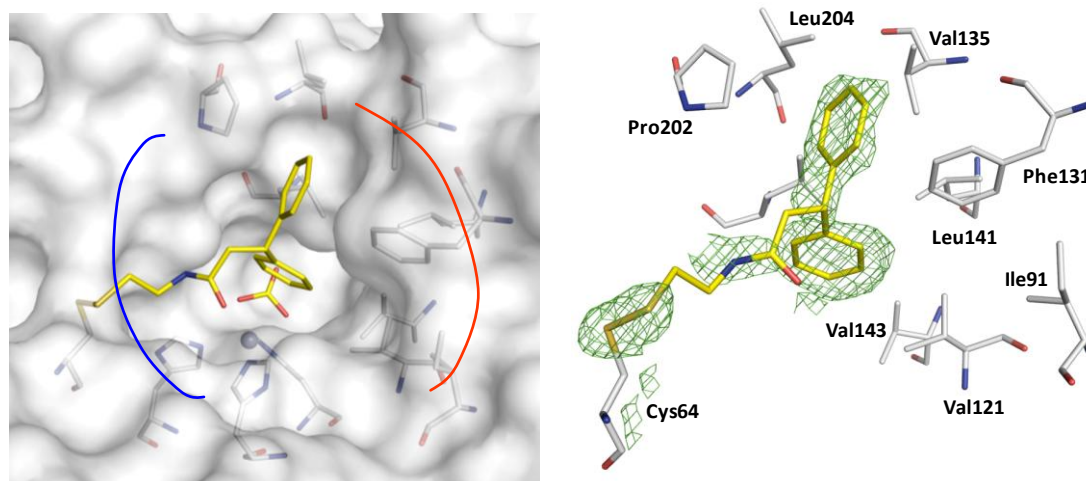
Timolol, a known CA II activator is proposed to bind into the same area, as suggested by docking experiments (Sugimoto, Ikeda et al.). Our crystal structures give a first hint for the existence of such a further binding site in CA II. More elaborate studies with enhanced fragment libraries could possibly identify new prospective fragments. CA II proton shuttle inhibitors could be developed by these studies.

**Crystal structure of CA II in complex with SB227:** Next to these novel "surface-binders", we were able to elucidate the binding mode of fragment **SB227** to CA II. The crystal structure of **SB227** tethered to CA II-H64C was determined at a resolution of 1.95 Å (PDB-Code 3M5T). The fragment is covalently attached to Cys64 which positions the attached fragment into the active site of CA II (Figure 7.12).

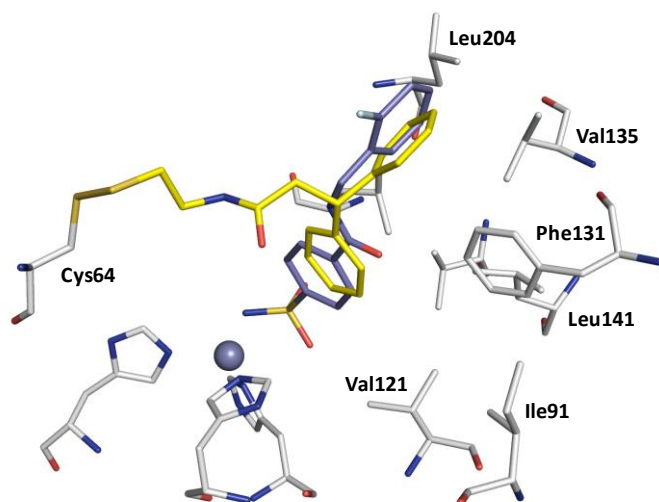
The biphenyl methane moiety of **SB227** is placed with its two phenyl groups next to the hydrophobic residues Val121, Ile91, Phe131, Leu141, Val206, Val135, Leu198 and Leu204. The upper phenyl ring performs edge-to-face interactions with Phe131 and simultaneously the second ring system penetrates deeper into the catalytic binding pocket. The *B*-values attributed to the atoms of the fragment indicate that the biphenyl methane moiety is fully populated in one orientation whereas the linker experiences weak interactions with protein residues. Therefore, it adopts multiple orientations. The mean *B*-values are: main-chain atoms 21.3 Å<sup>2</sup>, side-chain atoms 25.0 Å<sup>2</sup>, water molecules 30.2 Å<sup>2</sup>, ligand **SB227** 39.2 Å<sup>2</sup> and zinc ion 22.7 Å<sup>2</sup>. 31 van der Waals interactions have been calculated for **SB227**.

The lipophilic binding site is well known and addressed by many CA II inhibitors (e.g. Figure 7.13) and also neighbored to the recently described coumarin binding site (Maresca, Temperini et al. 2009a; Maresca, Temperini et al. 2009b; Temperini, Innocenti et al. 2009). However, since the crystallographic evaluation of molecules addressing this region requires in most cases (except the newly discovered coumarin type inhibitors) (Maresca, Temperini et al. 2009a; Maresca, Temperini et al. 2009b; Temperini, Innocenti et al. 2009) a sulfonamide or related zinc binding anchors to successfully accommodate protein crystals any screening with chemical diverse hydrophobic probes appears hardly feasible. Following the suggested tethering approach, new hydrophobic moieties can easily be examined by HPLC-MS and

crystallography to identify novel binders appropriate of CA II inhibitors. Apart from these lipophilic interactions the current complex crystal structure also provides new insights into the development of CA II activators. In Figure 7.14 the binding mode of the known activators histamine and L-histidine is shown.



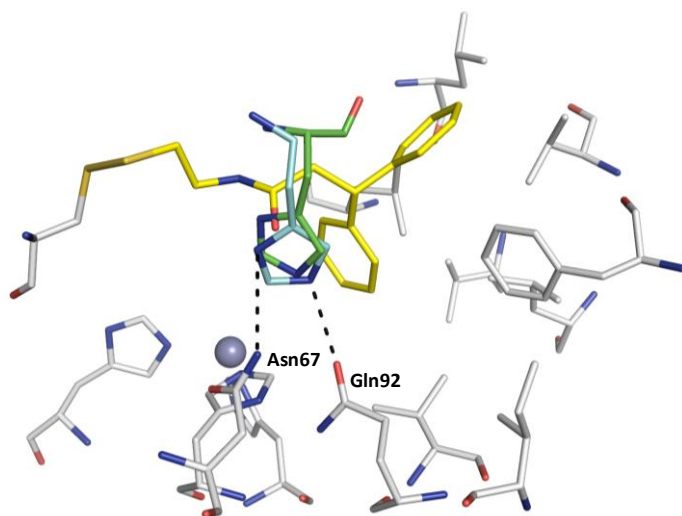
**Figure 7.12** Binding mode of SB227 in complex with CA II-H64C (protein: C, N, O, S; ligand: C, N, O, S; PDB-code: 3M5T). The binding pocket of CA II can be divided into a hydrophobic (indicated by red line) and a hydrophilic binding region (blue line). The fragment is covalently attached to Cys64 orienting with its biphenyl moiety towards the hydrophobic side chains.



**Figure 7.13** Superposition of SB227 in complex with CA II-H64C (protein: C, N, O, S; ligand: C, N, O, S) with sulfonamide inhibitor (C, N, O, S; PDB-Code 1G1D). The biphenyl-moiety of SB227 adopts similar positions as the two ring systems of the superposed ligand.

SB227 binds in the same region as these activators. The crystallographic determination of the binding mode for such small ligands is rather difficult due to their lower binding affinities. However, attaching a fragment covalently to Cys64 increases the local

occupancy of the low affinity binder thereby facilitating the crystallographic investigation.



**Figure 7.14** Superposition of CA II activators L-histidine (protein: C, N, O, S; ligand: C, N, O, S; PDB-Code 2ABE) and histamine (C, N, O, S; PDB-Code: 1AVN) with SB227 (C, N, O, S). The activators form hydrogen bonds to Asn67 and Gln92 and are further coordinated to the active site Zn<sup>2+</sup> ion via three water molecules. The carbonyl-group of SB227 is located in the area where the activator-heterocycle is bound.

The reported crystal structures evidence that the sole identification of binding fragments by HPLC-MS analysis is not sufficient to conclude on their most likely site of interaction. As apparent in the case of **Azide12** and **SB095**, the fragments can bind perfectly arranged to the surface of CA II. Moreover, this strong interaction pattern seems to prevent the tether to rotate towards the active site as long as no stronger interaction can be experienced within the active site. Fragment **SB227** is able to perform such interactions to the lipophilic portion of the cavity next to the entrance of the CA II active site. In order to screen fragments for putative binding to the active site zinc ion or the adjacent catalytically relevant amino acids the extended linkers ((CH<sub>2</sub>)<sub>3</sub>, (CH<sub>2</sub>)<sub>4</sub>) were used to attach to the terminal probe fragments, possibly they avoid stabilizing contributions arising from the placement of the linker. This concept has been successful in case of a CA II-H64C complex where a terminal azide group was linked by (CH<sub>2</sub>)<sub>2</sub> and subsequently by a (CH<sub>2</sub>)<sub>3</sub> group (s. Chapter 5.4). In the structure with the extended linker the mobility of the probe fragment was indicated by a poorly defined electron density which suggests scattering of the linker across multiple binding orientations. Compounds that were accordingly synthesized have been tested by

HPLC-MS. **TK094**, **TK095**, **TK096**, **TK107**, **TK108**, **TK148**, **TK162**, **TK170** and **TK171** showed increased affinity to CA II and will be tested by protein crystallography.

### 7.3.3 Comparison Crystal Structure and Docking Experiments

The crystallographic analysis of several (CH<sub>2</sub>)<sub>2</sub>-linked fragments in complex with CA II has shown in the current examples that mediated by water molecules such a linker distance allows placement of terminal probe skeletons to perfectly accommodate at the surface of CA II. Interestingly enough, two major orientations of fragments have been detected; the last example, the CA II-**SB227**-complex also probed the binding pocket. In the latter case the previously obtained docking poses agreed with the subsequently obtained crystal structure. Deviations in the exact binding mode can be observed; however, the hydrophobic interactions are equivalently suggested.

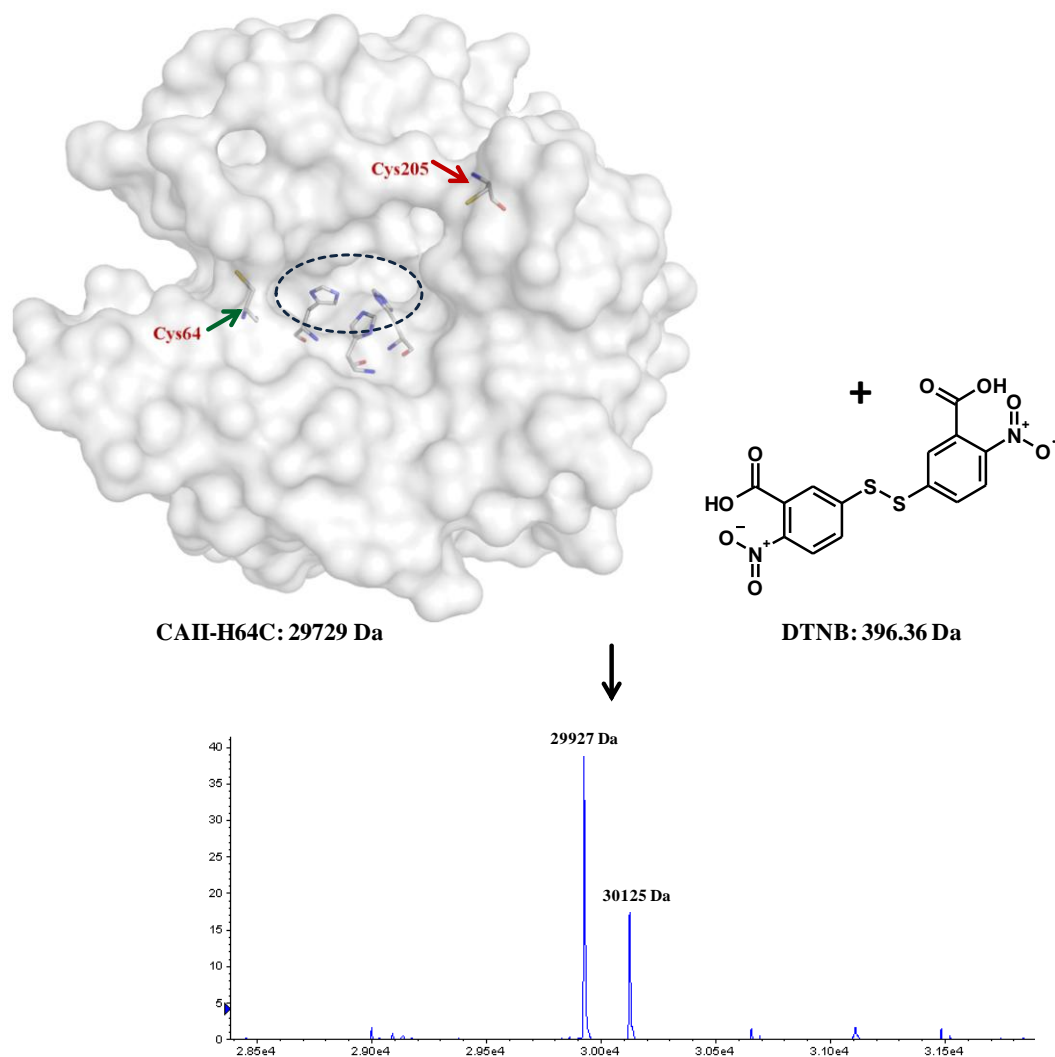
In case of the smaller fragments **SB095** and **Azide12** the docking results did not confirm the crystal structure determined subsequently. Nevertheless, among the proposed docking modes some approximate the crystal structure. A clear deficiency that might deteriorate the docking attempts is the complete neglect of the water network which is clearly essential and actually mediates the binding of the tethered fragment to the protein surface.

Nevertheless, the HPLC-MS and crystallographic experiments have shown so far that the selected fragments experience sufficient affinity to the surface of CA II-H64C to be successfully reacted with the accessible thiole group of Cys64. Most likely the size and introduced hydrogen-bond donor and acceptor functionalities of the tethered fragment take distinct impact on the docking modes and achieved scores. Repeating docking experiments considering the described water network could possibly improve the prediction of the fragment binding mode.

## 7.4 Experimental Section

**Accessible cysteine residues:** Prior to our initial tethering experiments, the overall accessibility of the cysteine residues of the target enzyme was determined. In best

case, only one cysteine residue shows the potential to be reacted with the added disulfide since only then a tryptic digestion can be avoided to clearly characterize the tether product. In case of CA II a cysteine residue has been introduced at the entrance of the binding pocket. This mutant CA II-H64C exhibits next to the mutationally introduced Cys64 a second cysteine residue close to the enzyme surface (Cys205) experiencing a distance of 7 Å to the binding pocket (Figure 7.15).



**Figure 7.15** Accessible Cysteine residues in CA II-H64C. The active site is indicated by the dashed ellipse. DTNB is used to elucidate cysteine residues that can be applied to covalently attach thiole groups. The green arrow shows the modifiable cysteine residue while the red arrow presents the rather buried cysteine that is less modified. CA II-H64C provides next to the mutationally introduced Cys64 a second cysteine at the protein surface (Cys205). At low DTNB concentrations Cys64 is mainly modified. Cys205 experiences modifications only at raised concentrations.

Even though it appeared likely that only Cys64 would be prominent to bind fragments as only attachment to this residue would allow a fragment to develop interactions with

the protein's active site. To check for the general accessibility of both cysteines for tethering the wild type and the Cys mutant were incubated with Ellman's reagent (5,5'-Dithiobis(2-nitrobenzoic acid) - DTNB) (Schwarzer, Mootz et al. 2002). Subsequently, the samples were purified over a size exclusion column to remove any unbound DTNB molecules and mass modification was determined by HPLC-MS. Unmodified CA II-H64C shows a mass of 29729 Da. Upon DTNB binding the enzyme mass must be increased by 198 Da (equivalent to half of DTNB). Different DTNB concentrations (10  $\mu$ M to 150  $\mu$ M) have been applied in these studies. Mass spectrometric analysis revealed that CA II-H64C is mainly modified by the anticipated mass difference which is indicated by the predominant mass peak of 29928 Da (10  $\mu$ M DTNB) ( $\Delta=198$  Da).

However, increasing the DTNB concentrations the protein mass is further increased indicating that a second DTNB attachment to CA II-H64C has been experienced (30124 Da;  $\Delta=395$  Da). Apparently, at increased concentrations apart from the first cysteine residue which is saturated the second available cysteine residue is now also modified. CA II-WT is also modified by DTNB at concentrations of 10  $\mu$ M. In this case Cys205 is modified, as there is only one cysteine among the protein residues.

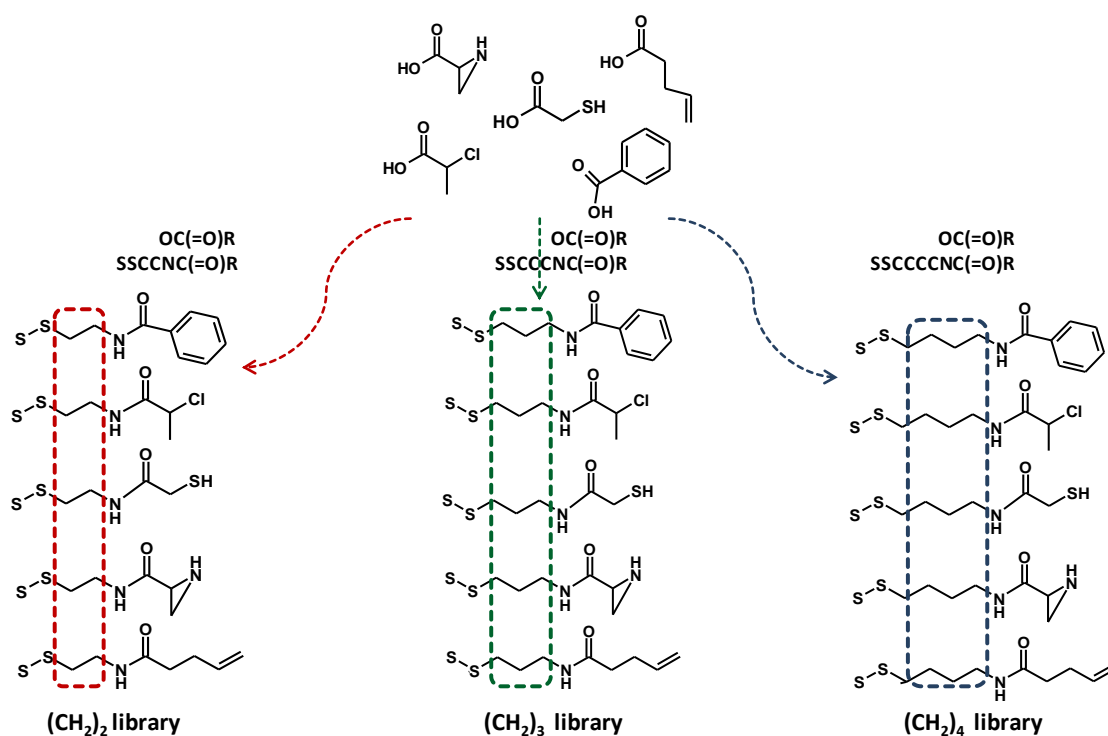
To confirm these results tryptic digestions of accordingly treated protein samples have been accomplished. We were able to identify the protein fragment with Cys64 modified by half of the mass equivalent of DTNB and the same for Cys205. Therefore, both cysteines are accessible for modification.

In conclusion, the DTNB-experiments confirmed success of the anticipated mutational design for tethering of CA II. CA II-H64C provides ideal preconditions for the tethering approach. Successful binding can be identified by HPLC-MS; however, the binding mode of the attached fragment has to be examined by protein crystallography. Only the evaluation of the binding mode allows precise interpretation of mass spectrometric data.

**Library Design:** Carboxylic acids are required for the synthesis of thiole containing fragments. Therefore, a virtual collection of carboxylic acids in mol2 format was used



for the design of an appropriate docking library. The carboxylic acids were converted into Smiles codes using the program Colibri. The corresponding linker together with the disulfide attachment, which are required to perform the covalent docking with GOLD, were added via this Smiles code to each ligand resulting in three distinct fragment sublibraries according to the linker lengths  $(\text{CH}_2)_2$  to  $(\text{CH}_2)_4$  (Figure 7.16). The Smiles codes have been translated into 3D structures using the program Corina and subsequently minimized in the force field implemented into the program Sybyl. After formal deprotonation of the terminal sulfur-hydrogen, the fragment libraries were completed.

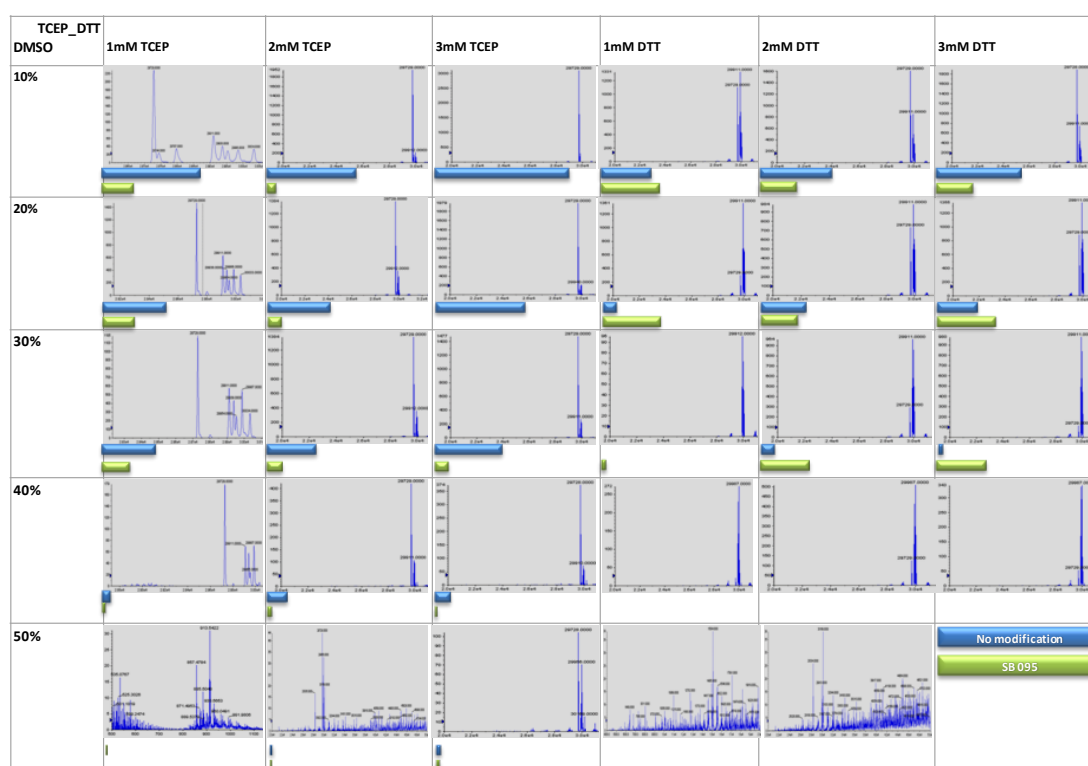


**Figure 7.16** Virtual library that has been applied in the docking experiments. Via smiles codes the library was modified by the linker-group and the disulfide group.

The protein geometry used for docking was extracted from a crystal structure of a CA II-H64C mutant determined in-house. To select an appropriate binding site area for docking a radius of 12 Å around the hydrogen of O $\gamma$  of Thr199 was used. All fragments were docked into CA II-H64C, covalently attached to Cys64 which itself was allowed to adopt all possible rotamers during the docking runs. The standard GOLD 4.0 parameters were applied and 30 solutions were generated. The docking solutions were ranked according to the achieved score. For the top 100 ranked

fragments the best five docking solutions were visually inspected and thus the most promising candidates selected for synthesis.

**Optimization of reaction conditions:** In general the reaction conditions described by Wells et al. were applied for sample preparations (Erlanson, Braisted et al. 2000). However, as we have preselected our fragment library by docking the affinity of the finally synthesized candidate fragments was anticipated to be sufficiently for tethering and subsequent HPLC-MS analysis. Apart from the concentration of the reducing agent also the solubility of the fragments and fragment cocktails influences the hit rate.



**Figure 7.17 DMSO vs. TCEP/DTT screening with Tethering fragment SB095 on CA II-H64C.** The bars below the spectra indicate ratio of modified and unmodified CA II-H64C: CA II-H64C-unmodified: blue bar; CA II-H64C-SB095: green bar.

Therefore, a reaction setup had to be established which assures a reasonable identification of binders. Multiple reaction conditions were tested with one fragment (SB095), systematically altering the DMSO as well as DTT concentrations. The mass spectra of accordingly treated samples were evaluated for optimal CA II-H64C modification. The analysis revealed that 10-30% DMSO and 1-2 mM DTT most likely produce the best results (Figure 7.17). The reducing agent 2-mercaptoethanol was

exchanged by DTT due to toxicity classification of the former. The reaction conditions were as following: 15  $\mu$ M CA II-H64C; 10% DMSO; 1 mM DTT; 200  $\mu$ M inhibitor.

**Inhibitor-mixtures:** All inhibitors were ranked according to their molecular weight. The ranking was applied to compose cocktails of four to six fragments with mass differences of 10 Da. Stock solutions of each mixture were prepared with 2 mM of each fragment. 5  $\mu$ l of this cocktail was added to a final sample volume of 50  $\mu$ l resulting in a final fragment concentration of 200  $\mu$ M each. In order to perform tethering experiments with a single fragment 1  $\mu$ l of a 10 mM compound stock solution were applied in 50  $\mu$ l sample volume.

**HPLC-MS analysis:** For HPLC-MS measurements 10  $\mu$ l of the CA II-H64C mutant (75  $\mu$ M), x  $\mu$ l DMSO ( $x = 5 \mu\text{l} - n \mu\text{l fragment(s)}$ ), 10  $\mu$ l DTT-solution (5 mM in Tris buffer) and 5  $\mu$ l/1  $\mu$ l of a fragment mixture/single fragment was diluted to a final volume of 50  $\mu$ l with buffer (50 mM Tris-HCl pH 8.0). The samples were incubated at room temperature for at least 1 hour.

After incubation the samples were pre-purified by size-exclusion chromatography (Illustra AutoSeq G-50 columns – GE Healthcare) to separate from excess substrates to avoid undesired side reactions. Complete desalting of the protein was achieved by HPLC using an Agilent 1100 system, samples were applied to a monolithic 50/1 ProSwift RP-4H column (Dionex). Desalted protein samples were eluted by the following gradient of buffer A (water/0.05% formic acid) and buffer B (acetonitrile/0.045% formic acid) at a column temperature of 40  $^{\circ}$ C and a flow rate of 0.2 mL/min: Isocratic elution with A (5%) for two minutes, followed by a linear gradient to B (95%) within 8 minutes and holding B (95%) for additional 4 minutes. Online mass spectrometric analysis was done with a Qstar Pulsar I mass spectrometer (Applied Biosystems) equipped with an ESI source. Measurement parameters were as follows: DP1 75, FP 265, DP2 15, CAD 2, GS1 35, CUR 25. The applied voltage was 5000 V. Positive ions within the mass range of 500-2000 m/z were detected. For better performance, the “enhance all” mode was activated.

**Crystallization of CA II complexes:** CA II-H64C crystals were grown using the sitting drop vapor diffusion method at 18 °C by mixing the protein solution (5 µl: ~10 mg/ml) with well solution (5 µl: 2.75 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 M Tris-HCl pH 8.0). Crystals appeared within 2 – 4 weeks in space group *P2<sub>1</sub>* and took up to 2 months to reach their maximum size. Complex structures were obtained by cocrystallization or soaking of the enzyme with DTT (500 µM) and the fragment (1 mM). For cryoprotection, crystals were briefly soaked in mother liquor containing 25% glycerol (Acros organics).

**Data Collection, Phasing, and Refinement:** The data sets were collected at the synchrotron BESSY II in Berlin/Germany on PSF beamline 14.2 (PDB codes: 3KIG, 3M2Z) and at our in-house source (2Q38, 3M5T). Data were processed and scaled with Denzo and Scalepack as implemented in HKL2000 (Otwinowski, Minor et al. 1997). The structures were determined by the molecular replacement method with Phaser (Storoni, McCoy et al. 2004) with our first inhouse CA II-H64C structure search model. Refinement was continued with CNS (Brünger, Adams et al. 1998) and SHELXL-97 (Sheldrick, Schneider et al. 1997). For each refinement step at least 10 cycles of conjugate minimization were performed, with restraints on bond distances, angles, and *B*-values. Intermittent cycles of model building were done with the program COOT (Emsley and Cowtan 2004). The coordinates have been deposited in the PDB (<http://www.rcsb.org/pdb/>) with the following accession codes: 3KIG, 3KNE. The data collection statistics are summarized in the Appendix 13.1.

## 8 The Sulfamide Head Group: A Versatile Zinc Chelator for Carbonic Anhydrase Inhibition of Remarkable Structural Variability

### 8.1 Introductory Remarks

In this chapter a series of CA II inhibitors (sulfamides) is presented which have been synthesized by Dr. Schümmelfeder in the former group of Prof. Haake (inhibitors **1 - 12**) and by Prof. Winum from Montpellier (inhibitor **13**) and have been enzyme kinetically tested in the group of Dr. Supuran from Florence, Italy. These compounds have been investigated by protein crystallography to determine the binding mode in CA II. The chapter has been prepared for submission to a scientific journal.

### 8.2 Introduction

$\alpha$ -Carbonic anhydrases are  $\text{Zn}^{2+}$ -dependent enzymes that catalyze the reversible hydration of carbon dioxide to hydrogen carbonate and a proton. They play an important role in physiological anion-exchange processes, pH homeostasy, respiration and fluid balance, electrolyte secretion, biosynthetic reactions, or tumorigenesis (Maren 1988). CAs are thus involved in many physiological and pathologic processes (Pastorekova, Parkkila et al. 2004; Maestrelli, Mura et al. 2002; Supuran, Scozzafava et al. 2003; Stiti, Cecchi et al. 2008; Supuran 2008; Winum, Rami et al. 2008). The 16 different  $\alpha$ -CA isoforms described so far in mammals, including *homo sapiens*, are involved in all these processes and are drug targets for various applications (diuretics, antiglaucoma, antiobesity, anticonvulsant or antitumor drugs among others) (Maren 1988). The dominant CA-isoenzyme, the carbonic anhydrase II (CA II) is located in many organs, showing very high catalytic efficacy. With a turnover rate of  $10^6 \text{ s}^{-1}$  CA II is one of the fastest enzymes known so far (Khalifah 1971).

At the bottom of the 15 Å deep funnel-like active site, the catalytic active  $\text{Zn}^{2+}$  ion is coordinated by three histidine residues, His94, His96 and His119. A water molecule/hydroxide ion completes the tetrahedral coordination of the  $\text{Zn}^{2+}$  ion and acts

as a nucleophile in the conversion of carbon dioxide (Silverman and Lindskog 1988; Christianson and Fierke 1996); (Merz and Banci 1997). Along the catalytic mechanism the metal-bound bicarbonate ion is displaced by a water molecule and liberated into solution, leading to the catalytically inactive acid form of the enzyme, with water coordinated to  $Zn^{2+}$ . Subsequently, a proton is transferred from the active site to the environment which regenerates the basic form of the enzyme. This proton shuffling may be assisted either by active site residues (such as His64 in several isoforms) or by buffers present in the medium. This step is rate limiting for the catalytic reaction (Nair and Christianson 1991b; Nair and Christianson 1991a).

Three main classes of CA inhibitors (CAIs) are known: the metal complexing anions, the unsubstituted sulfonamides (and their isosteres), which bind to the  $Zn^{2+}$  ion of the enzyme either by competitively substituting the substrate at  $Zn^{2+}$  or coordinating to the metal center by generating trigonal-bipyramidal species, and the coumarins (Supuran and Scozzafava 2000; Supuran and Scozzafava 2002; Supuran 2003; Supuran, Scozzafava et al. 2003; Pastorekova, Parkkila et al. 2004; Supuran, Scozzafava et al. 2004; Maresca, Temperini et al. 2009a; Maresca, Temperini et al. 2009b). The most investigated CAIs, the sulfonamides, coordinate with the terminal deprotonated nitrogen to the active site  $Zn^{2+}$  ion forming a tetrahedral geometry (Supuran and Scozzafava 2000; Supuran and Scozzafava 2002; Supuran 2003; Supuran, Scozzafava et al. 2003; Pastorekova, Parkkila et al. 2004; Supuran, Scozzafava et al. 2004). The proton bound to this nitrogen donates a hydrogen bond to Thr199 which itself is further linked as H-bond donor to Glu106. In addition, one of the sulfonamide oxygens interacts via an H-bond to the backbone NH of Thr199 (Abbate, Supuran et al. 2002; Casini, Antel et al. 2003; Fiore, Simone et al. 2005; Menchise, De Simone et al. 2005; Simone, Fiore et al. 2005). Carbonic anhydrase inhibitors belonging to the sulfonamide class led to important drugs, such as the antihypertensives of benzothiadiazine and high-ceiling diuretics type (Supuran and Scozzafava 2000), the sulfonamides with CA inhibitory properties mainly used as antiglaucoma agents, with at least 25 such agents in clinical use (Supuran and Scozzafava 2000; Supuran and Scozzafava 2002; Supuran, Scozzafava et al. 2003; Supuran, Scozzafava et al. 2004).

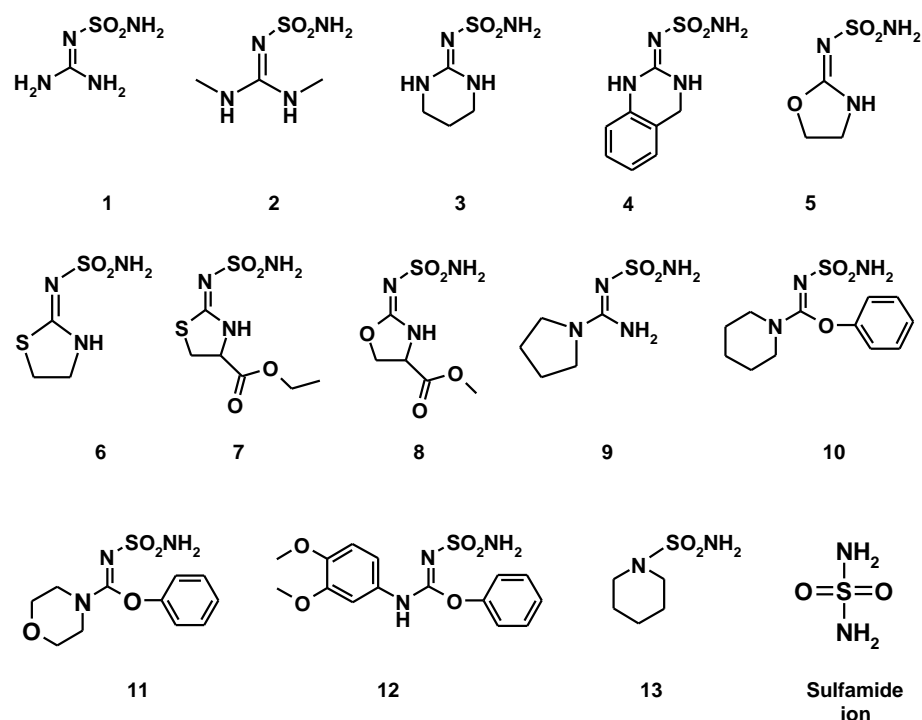
However, in the last years, the focus of CA II inhibitor design has changed toward other classes of compounds. Especially, new zinc binding groups have been tested (Winum, Scozzafava et al. 2007; Winum, Scozzafava et al. 2008). Compounds derived from the sulfamide skeleton were among the important ones (Winum, Scozzafava et al. 2006). Sulfamides are widely used by medicinal chemists for the design of a plethora of biologically active derivatives involved in different pharmacological applications. This scaffold can be altered to at least five types of derivatives, by successively substituting one of the four hydrogen atoms present at the two nitrogens. They have been tested on several protein targets such as carbonic anhydrases, proteases, serine proteases, metalloproteases and matrix metalloproteinases families (Winum, Scozzafava et al. 2006).

This study focuses on CA inhibitors of the sulfamide type. A first inhibition study with respect to some CA isoenzymes with the parent sulfamide skeleton ( $\text{H}_2\text{NSO}_2\text{NH}_2$ ) has been reported previously (Briganti, Pierattelli et al. 1996). This compound develops weak inhibition against the classical cytosolic isoforms CA I (0.31 mM) and CA II (1.13 mM). A crystal structure revealed the binding mode of such sulfamide to CA II (Abbate, Supuran et al. 2002). Similar to the sulfonamide type of inhibitors the sulfamide coordinates with one terminal deprotonated nitrogen to the active site  $\text{Zn}^{2+}$  ion (1.8 Å) and is further coordinated to Thr199 (2.8 Å) by means of a hydrogen bond. In addition, Thr199 is linked to Glu106 by an H-bond (2.6 Å). Another hydrogen bond involves one of the sulfamide oxygens and the backbone NH of Thr200 (2.8 Å). The latter amino acid forms a second H-bond to the non- $\text{Zn}^{2+}$  coordinating sulfamide nitrogen (3.3 Å). This well organized binding mode and the synthetic availability of the terminal nitrogen for additional substitution predicts a remarkable potential for this sulfamide scaffold. It has been shown that this group can be derivatized by means of reactions with sulfonyl halides, arylsulfonyl isocyanates, or aromatic/heterocyclic aldehydes, leading to derivatives with inhibition constants against isozymes CA I, CA II and CA IV in the low nanomolar range (Briganti, Pierattelli et al. 1996). Derivatives of the  $\text{RR}'\text{NH-SO}_2\text{-NH}_2$ -type such as *N,N*-disubstituted- and *N*-substituted-sulfamides were tested in another series. The disubstituted compounds showed only weak inhibition constants against CA I and CA II, whereas the mono-

substituted derivatives (incorporating aliphatic, cyclic, and aromatic moieties) as well as bis-sulfamide, exhibit a micro-to-nanomolar inhibition (Casini, Winum et al. 2003).

Here we present the crystallographic investigation of **13** (RR'NH-SO<sub>2</sub>-NH<sub>2</sub>-type) in complex with CA II, which has been kinetically evaluated earlier (Casini, Winum et al. 2003). In addition, we investigated a new series of 12 sulfamides of the RR'N=SO<sub>2</sub>-NH<sub>2</sub>-type against hCA I, hCA II, hCA IX and hCA XII with several inhibition constants in the one-digit nanomolar range. Remarkably, some compounds of this series show high selectivity towards CA II (nanomolar) compared to the other isozymes (micromolar). Three of these selective compounds were cocrystallized with CA II and investigated by means of the adopted binding poses.

### 8.3 Results



Scheme 8.1 Compounds that have been investigated for affinity against CAs.

#### 8.3.1 Inhibition Data

The inhibition data for compounds **1-13** and the sulfamide lead (Scheme 8.1) are shown in Table 8.1. The sulfamide lead is a very weak inhibitor against all isoforms



(310 – 1130  $\mu\text{M}$ ) except in case of hCA XI with 9.6  $\mu\text{M}$ . The inhibition constants of most inhibitors fall into the  $\mu\text{M}$  range. Some of the compounds experience increased affinity against certain isoforms, e.g. **2** against hCA XII; **11** and **13** against hCA II. In the contrary, some inhibitors are just very weak as, e.g. **10** against hCA I, hCA II and hCA XII or **3** against hCA I.

**Table 8.1** Inhibition data of compounds **1** to **13** against CA isoforms I, II, IX and XII.

Compound	hCA I <sup>a</sup>	K <sub>I</sub> * ( $\mu\text{M}$ )		
		hCA II <sup>a</sup>	hCA IX <sup>b</sup>	hCA XII <sup>b</sup>
<b>1</b>	624	6.0	8.0	7.7
<b>2</b>	750	6.3	8.9	115 nM
<b>3</b>	1248	4.5	7.8	66.9
<b>4</b>	14.5	12.0	16.3	12.8
<b>5</b>	15.2	6.4	7.7	7.4
<b>6</b>	13.8	5.9	7.5	7.6
<b>7</b>	10.2	10.9	476	395
<b>8</b>	5.9	10.6	10.0	8.8
<b>9</b>	589	11.6	9.1	14.6
<b>10</b>	2840	58.6	7.5	704
<b>11</b>	1547	15.1 nM	5.0	9.9
<b>12</b>	11.8	4.5	4.9	6.5
<b>13</b>	264 nM	179 nM	112 nM	63 nM
<b>H<sub>2</sub>NSO<sub>2</sub>NH<sub>2</sub></b>	310	1130	9.6	830

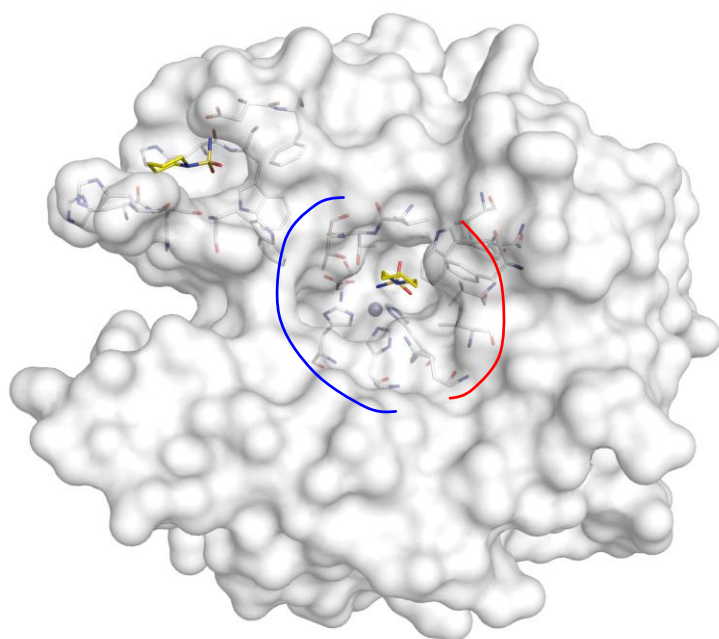
\*From 3 different assay (error +/- 10%)

<sup>a</sup> Recombinant, full length enzymes; <sup>b</sup> Recombinant enzyme, catalytic domain.

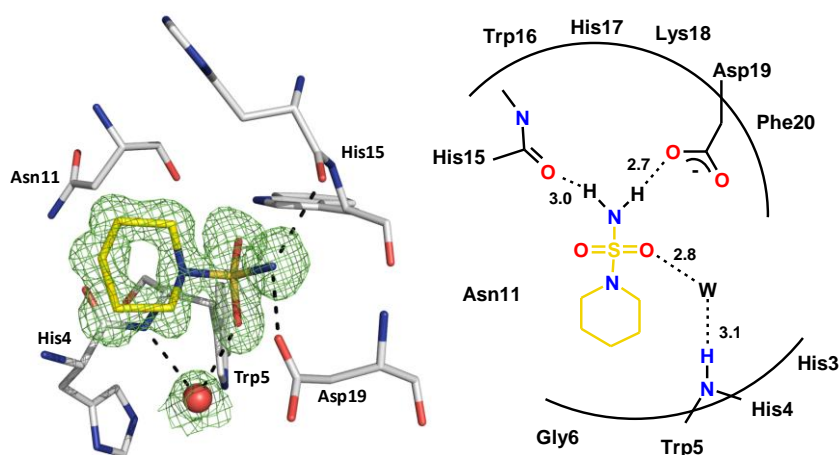
<sup>c</sup> From Casini et al., BMCL 2003, 13, 837-40

### 8.3.2 X-ray Crystallography

The binding modes of sulfamides **6**, **11**, **12** and **13** to hCA II were analyzed by determining the crystal structures of the protein-inhibitor complexes (Scheme 8.1). Crystals of the adducts were isomorphous with those of the uncomplexed protein (Eriksson, Jones et al. 1988), allowing for the determination of the crystal structures by difference Fourier techniques. The structures were refined at resolutions ranging from 1.17 to 1.87 Å. The statistics for data collection and refinement are shown in the Appendix 13.1.



**Figure 8.1** Binding mode of **13** in complex with CA II (PDB-Code: 2M2Y). The solvent accessible surface is presented in white. The interacting protein amino acids are shown in stick representation (protein: C, N, O, S; ligand: C, N, O, S). The binding pocket of CA II can be divided into a hydrophobic (red line) and a hydrophilic binding region (blue line). The protein is shown in an orientation with the proton shuttle His64 (next to the gap on the left hand side). This mode of representation is applied in all surface-figures in this chapter. **13** is bound to CA II in a stoichiometric ratio of 2:1. One inhibitor molecule is coordinated to the active site  $Zn^{2+}$  ion and a second molecule is located in a pocket formed by the N-terminal residues His4 to Phe20.



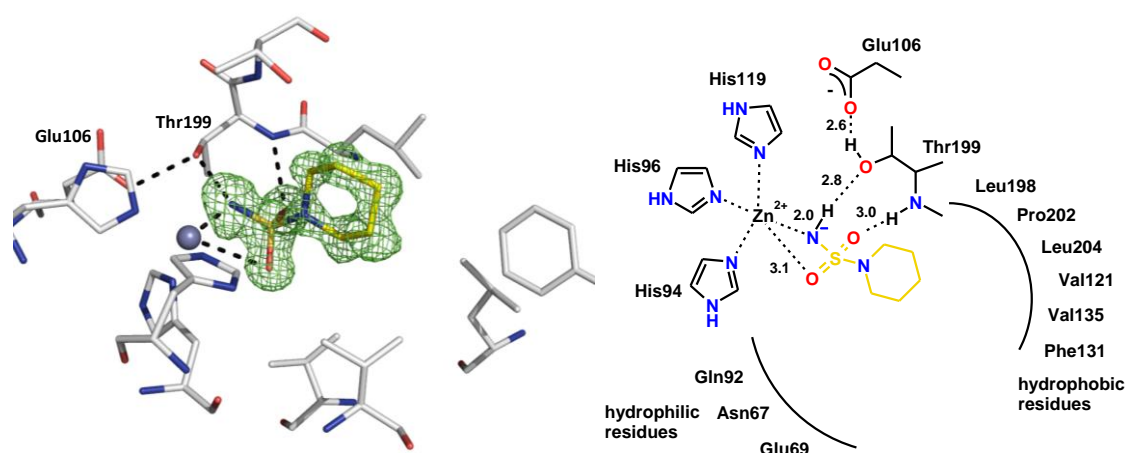
**Figure 8.2** Binding mode of **13** in complex with CA II (surface-bound molecule; PDB-Code: 2M2Y). On the left hand side the crystal structure is shown with the important amino acids. The protein residues are shown in stick representation (protein: C, N, O, S; ligand: C, N, O, S). Water molecules are shown as red spheres. Hydrogen bond interactions are indicated by the dashed lines. The equivalent distances are shown in the schematic representation of the binding mode on the right hand side. The difference electron density ( $F_o - F_c$ ) for the ligands and directly interacting water molecules are shown at a sigma level of 2.0. Atomic distances are shown in Å. This mode of representation is applied in all crystal structures in this chapter. **13** is bound to the protein surface by means of three hydrogen bonds to His15, Asp19 and mediated by a water molecule to the backbone NH of Trp5. The endocyclic nitrogen and one sulfamide oxygen do not interact with any protein residue.

**Complex structure CA II · 13:** The inhibitor binds to CA II in a stoichiometric ratio of 2:1. One inhibitor molecule is fixed to a cavity formed next to the N-terminus while a second molecule coordinates to the active site zinc ion (Figure 8.1). The remote pocket at the N-terminus is flanked by residues His4 to Phe20. Full occupancy of **13** in this pocket is clearly indicated by the excellent difference electron density ( $F_o - F_c$ -density: 2.0 sigma). The ligand is fixed by means of three H-bonds (Figure 8.2). The terminal sulfamide-nitrogen forms two hydrogen bonds to His15 and Asp19 of 3.0 Å (His 15) and 2.7 Å (Asp19). Furthermore, one sulfamide-oxygen interacts with the backbone nitrogen of Trp5 via an H-bond network that is mediated by a water molecule (2.8 Å and 3.1 Å). The second sulfamide-oxygen is not involved in any direct interactions. Asn11 (3.8 Å) is too distant to undergo sufficient H-bond interactions. Similarly, the cyclic sulfamide-nitrogen is not coordinated to further residues. 25 van der Waals contacts can be counted between ligand and protein. The average *B*-values provide an additional proof for the unambiguous fully populated binding mode of **13** at this surface-exposed site as residues 1 to 20 show an average *B*-value of 13.4 Å<sup>2</sup>, and the ligand atoms of 12.8 Å<sup>2</sup>.

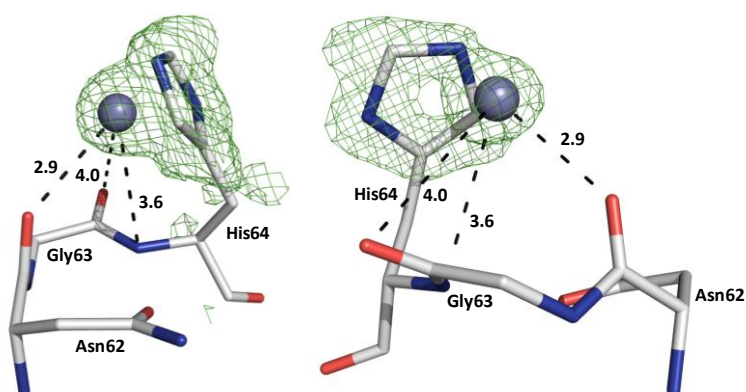
The second sulfamide ligand binds to the active site zinc ion adopting a binding mode that agrees well with those known for sulfonamides (Figure 8.3). The presumably deprotonated, negatively charged terminal nitrogen is coordinated to the zinc ion at a distance of 2.0 Å. The tetrahedral coordination geometry of the zinc ion is slightly distorted due to the sulfamide-oxygen which occupies the fifth coordination site (3.1 Å). Similar to sulfonamide inhibitors the sulfamide-nitrogen forms an H-bond to the terminal oxygen of Thr199 which is further coordinated by a hydrogen bond to Glu106. This H-bond network is established at common distances known from sulfonamide inhibitors (2.8 Å and 2.6 Å). The typical interaction pattern of CA II-sulfonamide complexes is completed by the hydrogen bond of the Thr199 backbone NH to the second sulfamide-oxygen at a distance of 3.0 Å.

Furthermore, the endocyclic nitrogen of **13** is not involved in any interactions with the protein. The ring system is oriented towards the lipophilic pocket which is formed by Leu198, Pro202, Leu204, Val121, Val135 and Phe131. In addition to the described hydrogen bonds, the inhibitor shows 26 van der Waals interactions to the neighboring

amino acids. The binding mode of **13** in the active site of CA II is well defined and clearly apparent from the difference electron density ( $F_o-F_c$ -density: 2.0 sigma). Additionally, the  $B$ -value of 12.3 Å<sup>2</sup> averaged across the ligand is similar to the average  $B$ -value for all protein atoms indicating full occupancy of the ligand in the binding site.



**Figure 8.3** Binding mode of **13** in complex with CA II (active site-bound molecule; PDB-Code: 2M2Y). The sulfamide moiety is coordinated to the active site  $Zn^{2+}$  ion and connected to Thr199 by means of H-bonds as is well established for many sulfonamide inhibitors. The ring system performs hydrophobic interactions to the side chains Leu198, Pro202, Leu204, Val121, Val135 and Phe131.



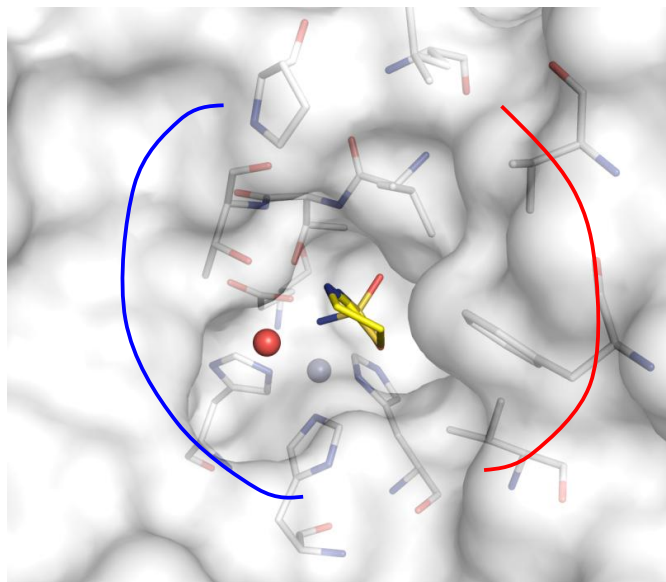
**Figure 8.4** Difference electron density at the proton shuttle His64. The sigma level of this additional density is approx. 28 and allows the assignment of a  $Zn^{2+}$  ion. On the right hand side the geometry is rotated by 90°. The density is shown at a sigma level of 2.0 left and 4.0 right, to illustrate the well defined binding mode. As both residues cannot be present at the same time the occupancies have been refined for His64 (62%) and  $Zn^{2+}$  (38%).

During the refinement of this structure we observed a difference density peak at very high sigma level (28 sigma) next to NE2 of His64. The imidazole moiety of His64 is very well defined and does not allow assignment of a second orientation. As we have

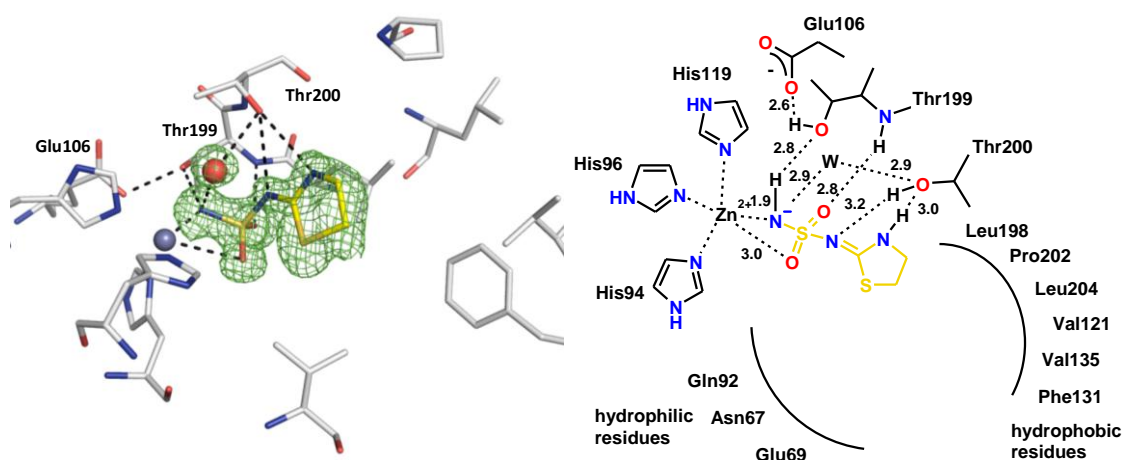
cocrystallized CA II solely with the inhibitor, and applied usual crystallization conditions, we propose that this density peak can be assigned to a second zinc ion. The spatial restraints prevent His64 and zinc to be localized at the same location in time and space (presumable distance NE2 – Zn<sup>2+</sup> 1.37 Å). Therefore, a partial occupation for both portions has been assigned. His64 is refined to a population to 62%, the zinc ion to 38% (*B*-value 11.0 Å<sup>2</sup>). The backbone oxygen of Asn62 coordinates the metal ion at a distance of 2.9 Å. Other neighboring potentially weakly coordinating residues are the backbone oxygen of Gly63 (4.0 Å) and the backbone NH of His64 (3.6 Å) (Figure 8.4).

**Complex structure CA II · 6:** In this complex, only one single inhibitor molecule is bound and located in the active site (Figure 8.5). In contrast to **13** described above, in this compound the non-terminal sulfamide-nitrogen is not part of a ring system. Furthermore, the attached thiazole ring contains a sulfur and a nitrogen atom. The terminal negatively charged nitrogen of the anchor group coordinates to the zinc ion with a distance of 1.9 Å (Figure 8.6). The tetrahedral coordination geometry at the zinc ion is slightly distorted due to one sulfamide oxygen which occupies the fifth coordination site (3.0 Å). Furthermore, the former nitrogen forms a hydrogen bond to the terminal oxygen of Thr199 (2.8 Å) which is additionally involved in an H-bond to Glu106 (2.6 Å). The binding mode is further fixed by the backbone nitrogen of Thr199 as it interacts via an H-bond to the second sulfamide oxygen at a distance of 2.8 Å. Thr200 forms with **6** three hydrogen bonds. Its terminal hydroxyl group functions as an H-bond donor to the inner-most sulfamide nitrogen (3.2 Å) and is addressed simultaneously as H-bond acceptor by the endocyclic nitrogen at a distance of 3.0 Å. In addition, a water molecule mediates an interaction between the hydroxyl group of Thr200 (2.9 Å) and the negatively charged terminal nitrogen atom (2.9 Å). The heterocycle of **6** is oriented towards the hydrophobic substrate binding pocket. The sulfur atom of the heterocycle is not involved in any interactions to the protein residues. In addition to the described hydrogen bonds, 28 van der Waals interactions can be observed. The ligand exhibits only minor mobility which is indicated by the very well resolved difference electron density (*F*<sub>o</sub>-*F*<sub>c</sub>-density: 2.0 sigma) and a

comparison of the average  $B$ -values: protein main-chain  $13.8 \text{ \AA}^2$ , side-chain  $17.7 \text{ \AA}^2$ , waters  $28.2 \text{ \AA}^2$ , inhibitor **6**  $16.5 \text{ \AA}^2$  and zinc ion  $6.6 \text{ \AA}^2$ .



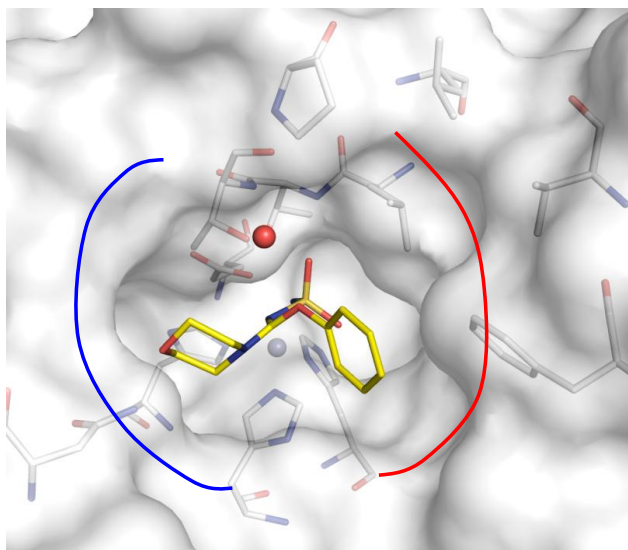
**Figure 8.5** CA II in complex with **6** (PDB-Code: 3M14). A water molecule (red sphere) participates in the interaction of **6** with CA II.



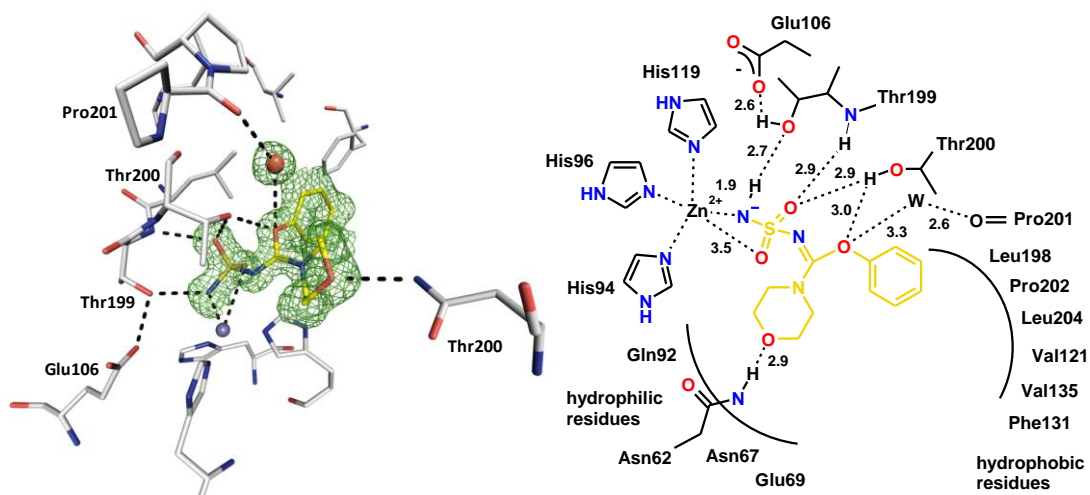
**Figure 8.6** Binding mode of **6** in complex with CA II (PDB-Code: 3M14). The sulfamide group is coordinated to the active site  $\text{Zn}^{2+}$  ion and connected to Thr199 and Thr200 by means of five hydrogen bonds with one water molecule participating in the binding mode. The ring system performs hydrophobic interactions to the lipophilic binding pocket.

**Complex structure CA II · 11:** The inhibitor **11** binds to the active site of CA II with its two ring moieties facing different parts of the large funnel-shaped sub-pockets (Figure 8.7). The phenoxy group is oriented towards the hydrophobic pocket wall of the active site whereas the morpholino moiety binds next to the hydrophilic side chains

Gln92, Asn67, Asn62 and Glu69 on the opposite side. An additional H-bond is formed between the terminal oxygen of the heterocycle and the carboxamide nitrogen of Asn62, a remarkable feature of the binding mode (Figure 8.8). With a distance of 2.9 Å the ring system is in good contact geometry and thereby pulled into the observed binding mode.



**Figure 8.7** CA II in complex with 11 (PDB-Code: 3M2X). The inhibitor 11 binds to the active site of CA II with both ring systems pointing towards different sub-pockets. The phenoxy moiety is bound to the lipophilic residues while the morpholino ring performs interactions with the hydrophilic side chains.



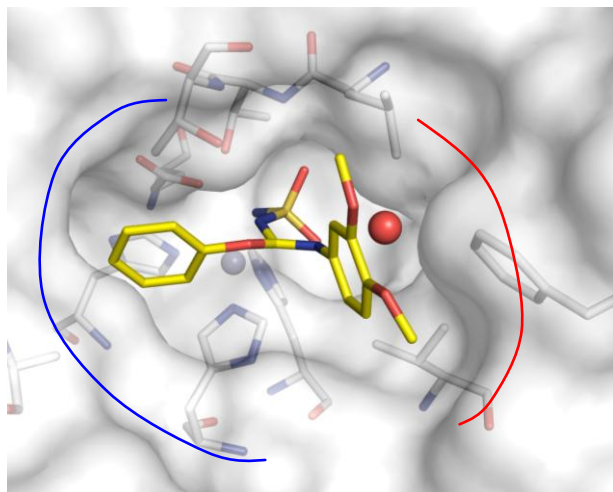
**Figure 8.8** Binding mode of 11 in complex with CA II (PDB-Code: 3M2X). The inhibitor is bound to four protein residues by means of six hydrogen bonds. One water molecule is involved in the coordination to Pro201. The sulfamide-anchor is coordinated to the active site zinc ion. The interaction to Asn62 pulls the inhibitor into the direction of the hydrophilic residues.

The second ring system, a phenoxy group, is pointing towards the lipophilic side chains of Leu198, Pro202, Leu204, Val121, Val135 and Phe131. In addition to these hydrophobic interactions, the observed binding mode is stabilized by a water-mediated H-bond (3.3 Å), formed by the iminoester oxygen towards the backbone carbonyl oxygen of Pro201 (2.6 Å) and a second hydrogen bond to Thr200 O $\gamma$ . This atom is also in favorable distance to hydrogen-bond to the second sulfamide oxygen (2.9 Å), most likely a furcated H-bond is formed. The binding geometry of this zinc anchor is characterized by a contiguous H-bond network starting at the terminal negatively charged nitrogen via Thr199 O $\gamma$  (2.7 Å) to Glu106 (2.6 Å). The first sulfamide-oxygen coordinates in bidentate fashion along with the negatively charged nitrogen (1.9 Å) to the active site zinc ion (3.5 Å) while the second oxygen forms a hydrogen bond to the backbone nitrogen of Thr199 at a distance of 2.9 Å. In comparison to ligand **6** the ring system is closer to the residues of the hydrophobic portion of the funnel-shaped pocket exhibiting 57 van der Waals interactions. Similar to **13** and **6** the difference electron density of **11** is very well defined ( $F_o-F_c$ -density: 2.0 sigma).

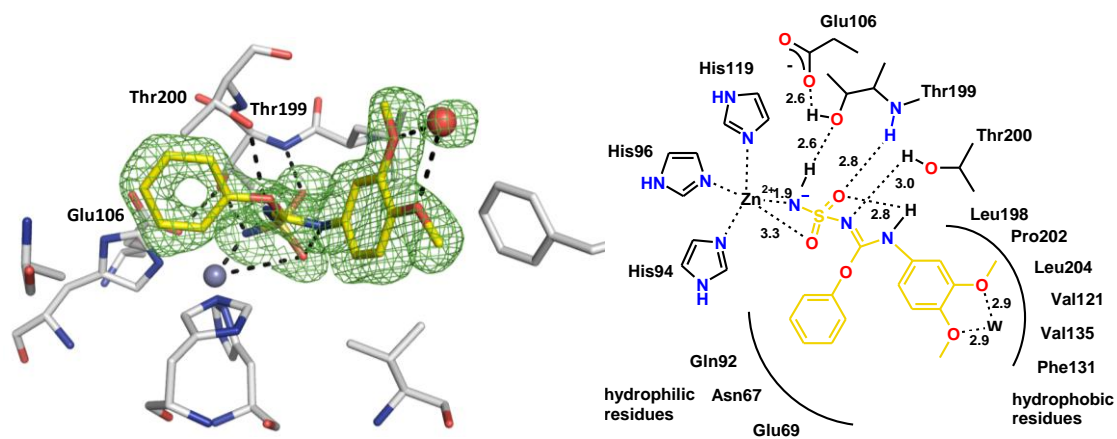
**Complex structure CA II · 12:** Similar to **11**, **12** bears two ring systems as terminal substituents. Apart from a phenoxy moiety, a second aniline-type substituent is present decorated by two methoxy-groups in 3- and 4- position. In contrast to **11**, the phenoxy group of **12** is now orienting towards the hydrophilic well of the binding pocket while the dimethoxy substituted ring system is placed next to the hydrophobic portion of the pocket (Figure 8.9). Similar to **11**, the difference electron density of **12** is very well defined ( $F_o-F_c$ -density: 2.0 sigma). The sulfamide anchor adopts the same orientation as in **11**. The terminal negatively charged nitrogen coordinates to the active site zinc ion (1.9 Å) and forms via its remaining hydrogen a bond to Thr199 (2.6 Å) which is further connected to Glu106 (2.6 Å) (Figure 8.10). One sulfamide oxygen is placed to the fifth coordination site at the zinc ion (3.3 Å) thereby distorting the tetrahedral coordination formed by His94, His96, His119 and the deprotonated sulfamide nitrogen. The second sulfamide oxygen forms, as in other complexes, an H-bond to the backbone NH of Thr199 (2.8 Å). Furthermore, the inner-most sulfamide nitrogen forms a hydrogen bond to the terminal oxygen of Thr200, similarly to **6** (3.0 Å). Interestingly enough, the nitrogen of the dimethoxyaniline substituent forms an



intramolecular H-bond with the neighboring sulfamide oxygen which is also engaged in an interaction with Thr199. The internal contact stabilizes the bound conformation. At the exit of the binding pocket the two neighboring methoxy groups capture an H-bonded water molecule via distances of 2.9 Å each. The interactions to this water molecule seem to be rather weak since its *B*-factor (38.7 Å<sup>2</sup>) is significantly augmented compared to those of the two methoxy groups (*B*-factors: 17.1 Å<sup>2</sup>; 21.8 Å<sup>2</sup>).



**Figure 8.9** CA II in complex with 12 (PDB-Code: 3M04). The ring systems of 12 are swapped in comparison to 11. The phenoxy-moiety is oriented towards the hydrophilic residues in the active site and the dimethoxy substituted second aromatic portion is interacting with the hydrophobic residues.



**Figure 8.10** Binding mode of 12 in complex with CA II (PDB-Code: 3M04). The sulfamide moiety is coordinating to the Zn<sup>2+</sup> ion, further stabilized by H-bonds to Thr199 and Thr200. In addition an intramolecular hydrogen bond and the interaction to a water molecule is observed.

Only minor ligand mobility is indicated as a very well defined difference electron density is given and the average *B*-values support this hypothesis: protein main-chain 18.2 Å<sup>2</sup>, side-chain 21.3 Å<sup>2</sup>, waters 28.2 Å<sup>2</sup>, ligand 12 18.8 Å<sup>2</sup> and zinc ion 12.7 Å<sup>2</sup>.

Similar to **11**, the aromatic portion of this substituent performs strong hydrophobic interactions with the lipophilic part of the binding pocket. The terminal methoxy groups enhance these contacts. In addition to the described hydrogen bonds 67 van der Waals interactions were calculated.

## 8.4 Discussion

### 8.4.1 Inhibition Data

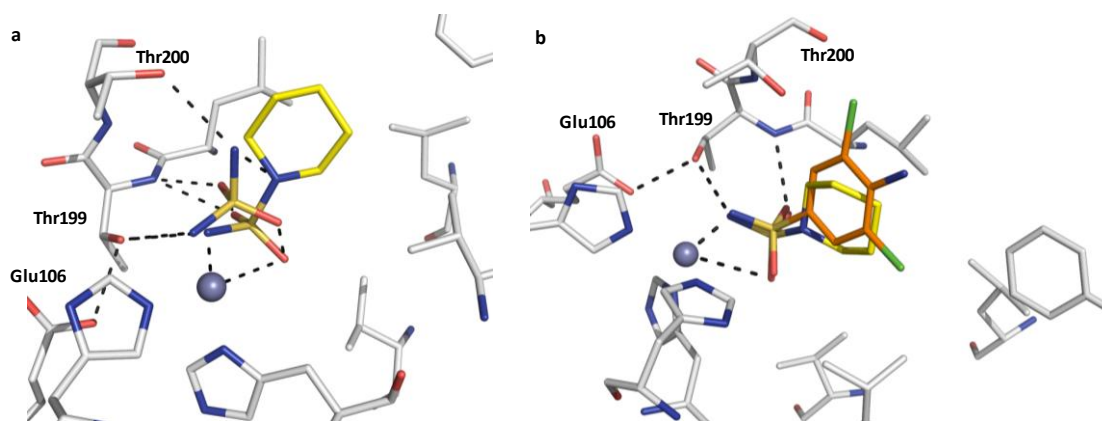
The inhibition data for all compounds against CA isoforms I, IX and XII show some inhibitors experiencing enhanced affinities and in parallel several with only very weak affinity to these isoforms. Since there is no data from crystal structure analysis available, detailed analysis of structure activity relationships are not possible at the moment.

The sulfamide ion is a very weak inhibitor for CA II ( $K_I$  of 1.13 mM), but all its derivatives show a highly enhanced affinity for this isoform, with inhibition constants in the range of 15.1 nM – 58.6  $\mu$ M (Table 8.1). Thus, compounds **1-9** show a rather compact behaviour of medium potency inhibitors, with a small variation of the  $K_I$ s (of 4.5 –12.0  $\mu$ M). They all incorporate a very similar motif, of  $\omega$ -sulfamoylated-guanidine/urea/thiourea. Inhibitor **11** shows the highest affinity with 15.1 nM. A similar inhibitor **10**, which is substituted by a piperidine ring instead of the morpholino ring in **11**, shows drastically decreased affinity. Also inhibitor **13** shows very strong affinity towards CA II with inhibition constant of 179 nM.

Depending on the substituted residues linked to the nitrogen and carbon of the imine functionality, the ligands can undergo fast E/Z-isomerization via rotation or inversion of the C=N bond. Considering, that compounds, such as **10 – 12**, which are substituted by a phenoxy group and a nitrogen show very fast isomerization (dissertation Dr. Schümmelfeder), both isomers are present in the experiments. In addition, the NH-tautomerism which is possible for some of these compounds also allows rotation around the imine nitrogen. Therefore, the protein can select the isomer that fits best into the active site.

### 8.4.2 Structural Comparison

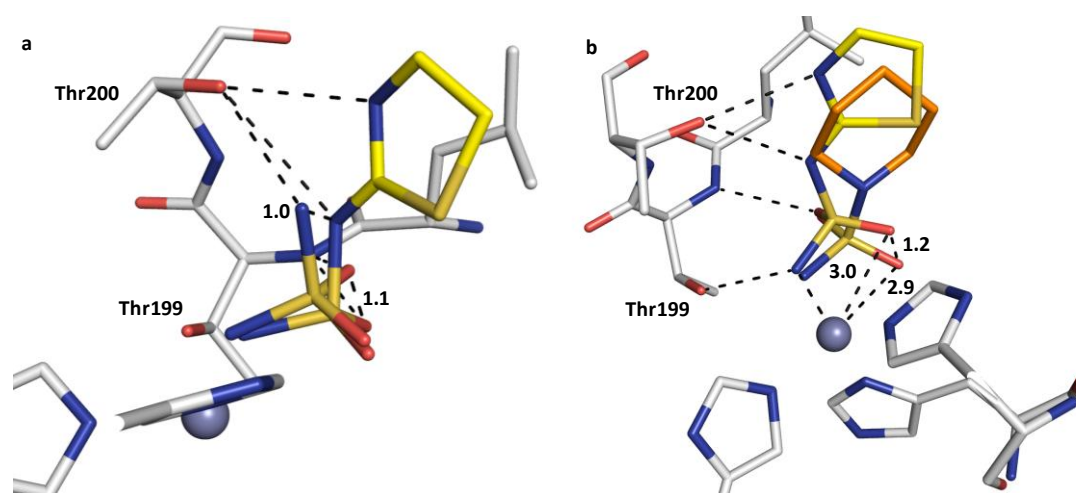
**Inhibitor 13:** The superposition of **13** with the parent sulfamide ion shows that **13** is tilted away from Thr200. It does not form a hydrogen bond with its cyclohexylamino nitrogen to this residue (Figure 8.11a). Obviously, the cycloalkyl substituent is too bulky to allow the same orientation as found for the parent sulfamide ion (Winum, Scozzafava et al. 2006). Supposedly, the gain of lipophilic contacts to the hydrophobic portion of the CA II binding pocket and the enhanced  $Zn^{2+}$  coordination compensates the unbalanced hydrogen bonding inventory of this ligand. The superposition of **13** with a substituted sulfonamide inhibitor (PDB-code 1ZGE) reveals that the binding modes for both chemically diverse anchor groups are virtually identical (Figure 8.11b) (Ilies, Vullo et al. 2003). The terminal nitrogen and both oxygen atoms occupy exactly the same positions. Since the piperidinyl nitrogen of **13** does not perform any interaction with the protein its ring system orients almost identical as the phenyl moiety of the superimposed ligand. The binding affinity of the superimposed inhibitor (50 nM) can explain to some degree the enhanced affinity of **13** (179 nM) as their binding modes are similar. Nevertheless, considering the observed interactions the measured affinity of **13** is surprisingly high.



**Figure 8.11** Superposition of **13** bound to CA II (ligand: C, N, O, S; protein: C, N, O, S). Waters are shown as red spheres. Dashed lines indicate interatomic distances in Å. This mode of representation is applied in all superpositions in this chapter. (a) Superposition of **13** with the structure of the sulfamide ion. **13** is shifted away from Thr200. (b) Superposition of **13** (C, N, O, S) with a typical sulfonamide inhibitor (C, N, O, S; PDB-code 1ZGE). The anchor groups adopt exactly the same orientation with the hydrogen bond network typical for sulfonamides with Thr199 O $\gamma$ . The ring systems fall in the same plane. The endocyclic nitrogen of **13** does not influence the binding mode, as it is not interacting with any protein residue.

**Inhibitor 6:** The superposition of **6** with the parent sulfamide ion reveals that **6** is shifted by about 1 Å in order to form hydrogen bonds to Thr200 (Figure 8.12a).

At first sight, as the anchor group distances of **6** are similar to those of **13** in complex with CA II, the binding mode of **6** seems to resemble that of typical sulfonamide inhibitors. However, detailed comparison reveals differences in the interaction pattern, giving **6** the advantage of two additional hydrogen bonds which might explain its improved binding affinity. Superposition of **6** with **13** illustrates the impact of the two nitrogen atoms of **6** on the binding mode (Figure 8.12b). The terminal negatively charged nitrogen and the oxygen atom which interact with Thr199 remain at the same positions. However, the second sulfamide oxygen experiences a shift of 1.2 Å due to an inclination of the inhibitor towards Thr200 to form the two additional H-bonds.

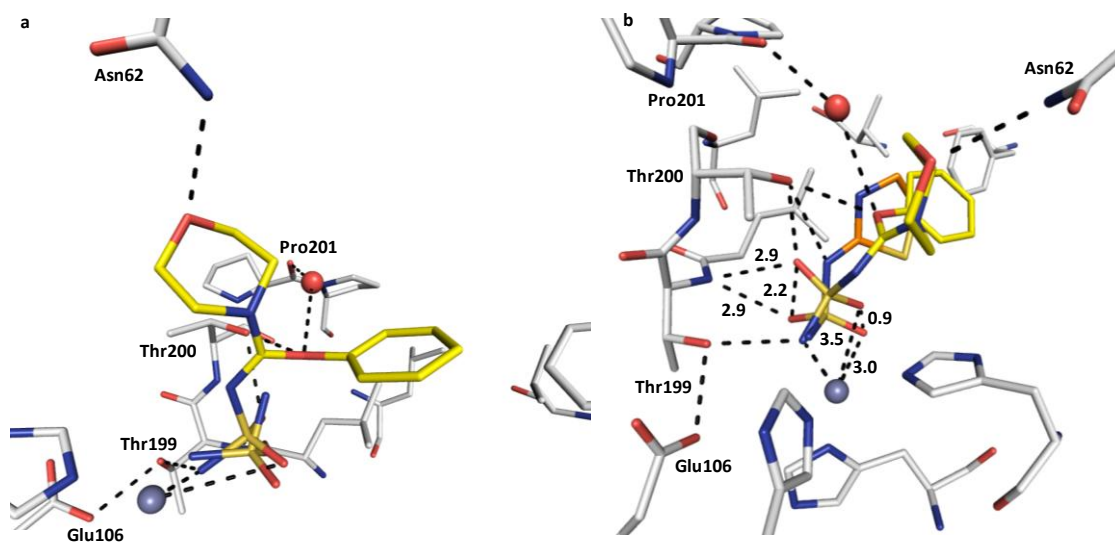


**Figure 8.12 Superposition of **6** bound to CA II.** (a) Superposition of **6** (C, N, O, S) with the sulfamide ion. **6** is shifted by 1 Å forming an additional hydrogen bond to Thr200. (b) Superposition of ligands **6** (C, N, O, S) and **13** (C, N, O, S) binding to CA II. The NH-Zn- and NH-Thr199- distances have almost the same values. However, one oxygen atom is shifted by 1.2 Å since thereby **6** can perform additional hydrogen bonds to Thr200.

The cyclic moieties of **13** and **6** adopt almost the same position filling the hydrophobic binding pocket. Furthermore, the interaction pattern to the active site zinc ion is similar as the respective distances remain the same. Also a similar amount of van der Waals interactions is counted. Considering all these facts it is difficult to understand the enhanced affinity of **13** compared to **6**. The formation of two additional, most likely strong, hydrogen bonds should increase the binding affinity of **6**. Nevertheless, in an inhibition assay compound **6** (5.9 μM) shows 30 fold decreased affinity compared to

**13** (179 nM). The present data do not provide an obvious explanation for this affinity difference. Additional experiments, e.g. ITC measurements, should be performed to elucidate the reasons.

**Inhibitor 11:** Superimposing **11** with the parent sulfamide ion shows that **11** is shifted toward the hydrophilic binding pocket, losing one hydrogen bond to Thr200 in order to gain an interaction with Asn62 (Figure 8.13a).



**Figure 8.13 Superposition of 11 bound to CA II.** (a) Superposition of **11** (C, N, O, S) with the sulfamide ion. (b) Superposition of **11** (C, N, O, S) and **6** (C, N, O, S) binding to CA II. Due to the hydrogen bond of the morpholino ring to Asn62 the sulfamide-group is shifted by 0.9 Å to 2.2 Å.

Comparing the hydrogen-bond network formed by **11** and **6**, similar distances can be observed (Figure 8.13b). However, similar to the previous case, the superposition of **6** and **11** reveals major rearrangements in the sulfamide coordination pattern. Due to the interaction of the morpholino ring with Asn62, the ligand is shifted towards the hydrophilic pocket. In contrast to **6**, the inner-most sulfamide nitrogen of **11** is rotated by 2 Å away from the terminal oxygen of Thr200 and cannot perform any interactions to this residue. Therefore, one hydrogen bond interaction is lost by this movement. However, the oxygen of the anchor group which typically interacts with the backbone NH of Thr199 is also shifted towards Thr200 by 2.2 Å compared to the situation with **6**. It regains the hydrogen bond by occupying a position similar to the nitrogen in **6**. The iminoester oxygen of **11** is located at a similar position as the endocyclic nitrogen

of **6**. The secondary amine of **6** donates an H-bond to Thr200. The spatial orientation of this nitrogen does not allow additional hydrogen-bond interactions. In contrast, the iminoester oxygen in **11** exhibits two acceptor functionalities and can, in addition to the interaction with Thr199, participate in a water-mediated H-bond with Pro201.

Due to the ligand's inclination towards the hydrophilic residues, **11** cannot form a short directed H-bond from its inner sulfamide-nitrogen group to Thr200 O $\gamma$ . This loss is compensated by hydrogen bonds between Thr200 O $\gamma$  the iminoester oxygen and the S=O group of the sulfamide. Compared to **6**, **11** loses the NH Thr200 O $\gamma$  contact but gains two additional hydrogen bonds via two oxygens. These supposedly more favorable interactions explain the increased binding affinity further enhanced by pronounced van der Waals interactions particularly formed by the phenoxy group with the hydrophobic portion of the binding pocket.

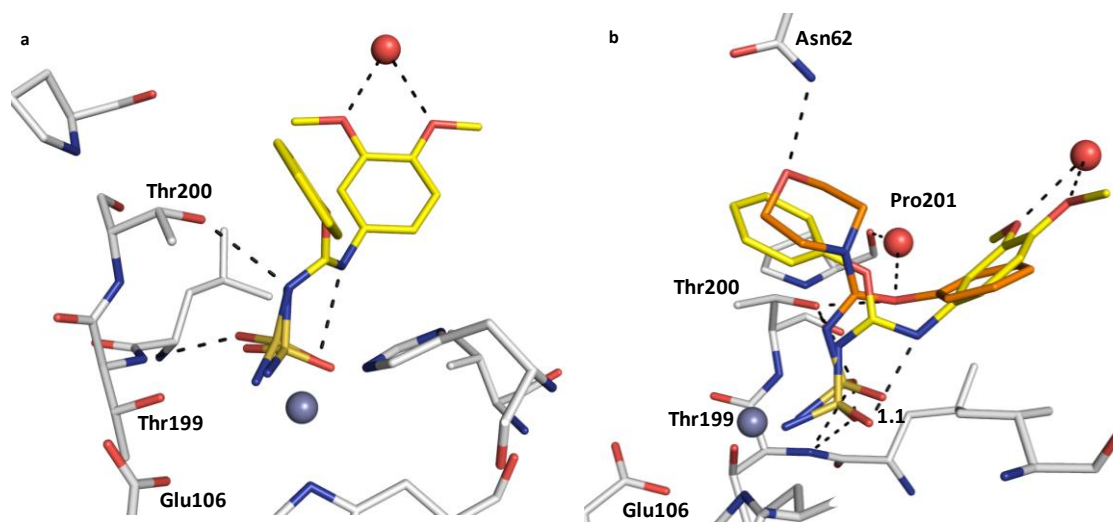
**Inhibitor 12:** Ligand **12** adopts virtually the same orientation with respect to its sulfamide anchor as the parent sulfamide ion (Figure 8.14a). Differences in the H-bonding pattern are observed between **11** and **12** (Figure 8.14b). The interactions to the active site zinc ion and Thr199 are similar in both structures (two hydrogen bonds and Zn<sup>2+</sup> coordination). However, the two oxygens and the inner nitrogen of the sulfamide group are shifted by approx. 1 Å towards Thr199 in **12**. Thereby, this nitrogen can interact with Thr200 O $\gamma$ , a contact not feasible to undergo for the corresponding nitrogen in **11**. In contrast, **11** forms additional H-bonds to Thr200, Pro201 and Asn62 via its oxygens. **12** shows, in comparison to **11**, different E/Z stereochemistry at the innermost imine nitrogen. The described E/Z equilibrium can easily adjust the most favorable conformation for binding. As a result, the iminoester oxygen moiety is now facing the hydrophilic wall of the funnel-shaped binding pocket and the dimethoxy aniline portion is oriented toward the hydrophobic side. It thereby places the NH group in a position to form an intramolecular H-bond to the sulfamide oxygen, thus stabilizing the geometry of the bound conformation.

Interestingly, the two adjacent methoxy groups capture a water molecule which is in no further contacts with the protein. In total, **11** forms six hydrogen bonds (two furcated ones) to the protein while **12** only participates in three H-bonds, but it

establishes two additional hydrogen bonds to a water molecule and experiences an intramolecular contact to fix the arrangement of the bound conformation.

Both ligands exhibit diverse affinities with inhibitor **11** binding approx. 265 fold stronger than **12**. It is remarkable that even so sharing parts of their scaffold in common they bind with reversed placement of their substituents. This is achieved by deviating E/Z-orientation at the C=N bond. Solely counting the interaction inventory **11** outnumbers **12** in terms of favorable H-bonds. Possibly the intramolecular H-bond which is formed in the complex with **12** is not accessible for inhibitor **11**. Thus the latter enables the coordination with its ring system to Asn62.

This observed E/Z-isomerization might also explain the dramatic loss of affinity of **10** compared to **11**. Possibly, as the piperidine ring system cannot form an H-bond to Asn62 (compared to **11**), inhibitor **10** might not be pushed into the same configuration and therefore, binds in a similar orientation as **12** forming the same intramolecular H-bond.



**Figure 8.14 Superposition of 12 bound to CA II.** (a) Superposition of **12** (C, N, O, S) with the sulfamide ion. (b) Superposition of ligand **12** (C, N, O, S) and **11** (C, N, O, S) in complex with CA II. The sulfamide oxygen atoms are shifted towards the  $Zn^{2+}$  ion by 0.9 Å and 1.1 Å, respectively. The innermost sulfamide nitrogen is rotated towards Thr200 thereby allowing a hydrogen bond to this residue.

## 8.5 Experimental Section

**CA inhibition:** An Applied Photophysics stopped-flow instrument has been used for assaying the CA catalyzed  $CO_2$  hydration activity (Khalifah 1971). Phenol red (at a concentration of 0.2 mM) has been used as indicator, working at the absorbance

maximum of 557 nm, with 20 mM Hepes (pH 7.5) as buffer, and 20 mM Na<sub>2</sub>SO<sub>4</sub> (for maintaining a constant ionic strength), following the initial rates of the CA-catalyzed CO<sub>2</sub> hydration reaction for a period of 10-100 s. The CO<sub>2</sub> concentrations ranged from 1.7 to 17 mM for the determination of the kinetic parameters and inhibition constants (5 different substrate concentrations have been used). For each inhibitor at least six traces of the initial 5-10% of the reaction have been applied for determining the initial velocity. The uncatalyzed rates were determined in the same fashion and subtracted from the totally observed rates. Stock solutions of inhibitor (0.1 mM) were prepared in distilled-deionized water and dilutions up to 0.01 nM were done thereafter with distilled-deionized water. Experiments were done using 6 different inhibitor concentrations, varying from 100 μM to 0.1 nM. Inhibitor and enzyme solutions were preincubated together for 15 min to 24 h at room temperature (15 min) or 4 °C (all other incubation times) prior to assay, in order to allow for the formation of the E-I complex or for the eventual active site mediated hydrolysis of the inhibitor. Data show the inhibition after 15 min incubation, as there were no differences of inhibitory power when the enzyme and inhibitors were kept for longer periods in incubation (Maren 1988). The inhibition constants were obtained by non-linear least-squares methods using PRISM 3, as reported earlier (Briganti, Mangani et al. 1997; Temperini, Scozzafava et al. 2005; Temperini, Scozzafava et al. 2006b; Temperini, Innocenti et al. 2007; Temperini, Innocenti et al. 2008) and represent the mean from at least three different determinations. CA isozymes were prepared as reported earlier (Bayram, Senturk et al. 2008; Innocenti, Hilvo et al. 2008; Innocenti, Vullo et al. 2008b; Innocenti, Vullo et al. 2008a; Maresca, Temperini et al. 2009a; Maresca, Temperini et al. 2009b; Temperini, Innocenti et al. 2009).

**Synthesis:** The synthesis of compound **13** was performed by the group of Prof. Winum, the synthesis of the other compounds by the group of Prof. Haake, in particular by Dr. Schümmelfeder.

**Crystallization, X-ray data collection and refinement:** Crystals of hCA II (from Sigma-Aldrich) in complex with each inhibitor were obtained by the sitting drop technique, using 10 mg/mL protein with 2 mM inhibitor (**13**, **6**, **11** and **12** respectively)



in 50 mM Tris-HCl buffer (pH 8.0). The drops consisted of 5  $\mu\text{L}$  of the enzyme-inhibitor solution and 5  $\mu\text{L}$  of the precipitant solution containing 3.0 M  $(\text{NH}_4)_2\text{SO}_4$  in 100 mM Tris-HCl (pH 8.0) and 0.15 mM *p*-chloromercuribenzoic acid to promote the growth of highly oriented crystals. The drops were equilibrated by vapor diffusion against the precipitant solution at 18 °C, and crystals appeared after 1-4 days.

The data sets of CA II-**13** and CA II-**11** complexes were collected at the synchrotron BESSY II in Berlin/Germany on PSF beamline 14.2. The experiment was performed at 100 K using synchrotron radiation of 0.91841 Å wavelength on a MAR CCD 165 mm detector. The data set of CA II-**6** complex was collected at the Swiss Light Source (SLS) in Villigen/Switzerland on Beamline SLS X06DA on a MAR CCD 225 mm detector at 100 K using radiation of 1.000 Å wavelength. The data of CA II-**12** complex were collected on an MAR 345 mm plate detector with Incoatec I $\mu$ S copper microfocus source at a wavelength of 1.54178 Å (100 K). The data collection parameters and refinement statistics are reported in the Appendix 13.1.

Data were processed and scaled with Denzo and Scalepack as implemented in HKL2000 (Otwinowski, Minor et al. 1997). The complex structures were analyzed by difference Fourier techniques, using the PDB file 2OQ5 as starting model for refinement. Refinement was continued with CNS (Brünger, Adams et al. 1998) and SHELXL-97 (Sheldrick, Schneider et al. 1997). For each refinement step at least 10 cycles of conjugate minimization were performed, with restraints on bond distances, angles, and *B*-values. For each crystal the  $2F_o-F_c$  and  $F_o-F_c$  maps were calculated and the inhibitor binding mode assigned from the difference electron density. Intermittent cycles of model building were done with the program COOT (Emsley and Cowtan 2004). The coordinates have been deposited in the PDB (<http://www.rcsb.org/pdb/>) under the accession codes 3M2Y, 3M14, 3M04 and 3M2X.



## 9 Bidentate Zinc Chelators for $\alpha$ -Carbonic Anhydrases that Produce a Trigonal Bipyramidal Coordination Geometry

### 9.1 Introductory Remarks

The following chapter has been submitted to a scientific journal for publication. In the following the original text is depicted. A series of new zinc binders is presented. The crystallographic investigation was performed during this thesis, the compounds have been provided by the group of Prof. Seth Cohen (San Diego, USA) and enzyme kinetically tested on 13 CA isoforms in the group of Dr. Supuran (Florence, Italy; enzyme preparation: Daniela Vullo; kinetic assay: Alessio Innocenti).

### 9.2 Abstract

A series of new zinc binding groups (ZBGs) has been evaluated kinetically on 13 carbonic anhydrase isoforms. The fragments show affinity for all isoforms with  $IC_{50}$  values in the range of 2 to 11  $\mu$ M. The crystal structure of hCA II in complex with one such fragment reveals a bidentate binding mode with a trigonal-bipyramidal coordination geometry at the  $Zn^{2+}$  center. The fragment also interacts via hydrogen bonding to Thr199 and Thr200 and participates in a water network. Further development of this ZBG should increase the binding affinity leading to a structurally distinct and promising class of CA inhibitors.

### 9.3 Introduction

Carbonic anhydrases constitute one of the most extensively studied group of metalloenzymes (Supuran, Scozzafava et al. 2003; Supuran, Scozzafava et al. 2004). They belong to a superfamily of ubiquitous metalloproteins present in prokaryotes and eukaryotes and are encoded by five distinct evolutionarily unrelated gene families:  $\alpha$ -class (present in vertebrates, bacteria, algae, and cytoplasm of green plants),  $\beta$ -class (predominantly in bacteria, algae, and chloroplasts),  $\gamma$ -class (mainly in *Archaea* and some bacteria), and  $\delta$ - and  $\xi$ -classes present in marine diatoms. They catalyze the reversible hydration of carbon dioxide into hydrogen carbonate and a proton. The catalytically relevant metal ion (which is  $Zn^{2+}$  in the  $\alpha$ -CAs but may be  $Fe^{2+}$  in  $\gamma$ - or

$\text{Cd}^{2+}$  in  $\xi$ -CAs, among others) of CAs is located (for the  $\alpha$ -class enzymes) at the bottom of the 15 Å deep active site, coordinated by three histidine residues - His94, His96 and His119 - and a water molecule forming a tetrahedral geometry at the active site (Supuran, Scozzafava et al. 2003; Supuran, Scozzafava et al. 2004).

Sixteen different isoforms have been described in mammals, each differing by their relative hydase activity, their sub-cellular localization, and their susceptibility to inhibition. Because these CAs are involved in several physiological processes it is not surprisingly that carbonic anhydrase inhibitors (CAIs) have been developed for the treatment of various pathologies such as glaucoma, neurological disorders, and osteoporosis (Scozzafava, Mastrolorenzo et al. 2006; Supuran 2008). The recent discovery of CA isoforms that are involved in cancer or obesity has demonstrated the need for potent, novel classes of CAIs. However, the large number of diverse  $\alpha$ -CAs leads to the lack of inhibitor selectivity, especially for first-generation CAIs, i.e., sulfonamides and their bioisosteres. In order to reduce unwanted side effects of CAI therapy, the design of novel inhibitors has to discriminate between the different isozymes (Supuran, Scozzafava et al. 2003; Pastorekova, Parkkila et al. 2004; Scozzafava, Mastrolorenzo et al. 2004; Supuran, Scozzafava et al. 2004).

So far, most rationally designed CAIs are based on a zinc-binding group (ZBG) that interacts with the  $\text{Zn}^{2+}$  ion of the enzyme and Thr199. To this ZBG a diversely substituted aromatic ring, heterocyclic, aliphatic group, or sugar moiety is linked that interacts with hydrophobic and hydrophilic pockets neighboring the active site. The sulfonamide group (Scheme 9.1) is the predominantly ZBG present in CAIs; this is in contrast to other  $\text{Zn}^{2+}$  metalloenzymes, such as the matrix metalloproteinases (MMPs), where hydroxamic acids serve as the dominant ZBG (Scheme 9.1) (Jacobsen, Major Jourden et al.). The sulfonamide moiety coordinates with its terminal deprotonated nitrogen atom to the  $\text{Zn}^{2+}$  ion (2.0 Å) and participates in a hydrogen-bonding network to Thr199 (2.8 Å) and Glu106 (2.7 Å). The latter amino acids are referred to as the "door-keepers" in the CA active site (Christianson and Fierke 1996). In addition, one sulfonamide oxygen forms a hydrogen bond to the backbone NH of Thr199 (3.0 Å). With this interaction pattern, many highly potent sulfonamide CAIs have been

developed (Supuran and Scozzafava 2000; Supuran and Scozzafava 2002; Supuran, Scozzafava et al. 2003; Supuran, Scozzafava et al. 2004).

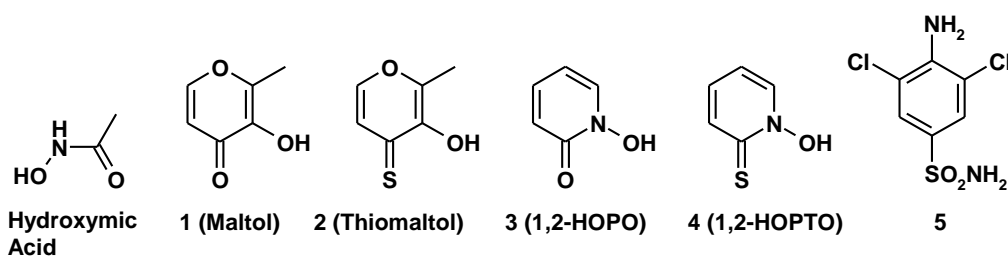
The side effects of sulfonamide CAIs have led to new strategies for selective inhibitor design, such as modifications of the inhibitor backbone (Supuran, Scozzafava et al. 2003; Supuran, Scozzafava et al. 2004; Scozzafava, Mastrolorenzo et al. 2006; Supuran 2008). One attractive approach to new CAIs is the use of non-sulfonamide ZBGs. A range of ZBGs have been tested so far, including sulfamates, sulfamides, substituted sulfonamides, Schiff bases, ureas and hydroxyureas, as well as hydroxamic acids. These studies revealed sulfamates and sulfamides, structurally related to sulfonamides, to be amongst the most promising alternative ZBGs (Winum, Scozzafava et al. 2005; Winum, Scozzafava et al. 2006; Winum, Scozzafava et al. 2007).

In order to find new ZBGs for use in CAIs, ligands from other  $Zn^{2+}$ -dependent metalloenzymes are of particular interest. Recently, Cohen et al. have developed a series of new ZBGs, which include hydroxypyr(thi)ones and hydroxypyridine(thi)ones, that have been examined as alternative ZBGs for MMP inhibitors (Scheme 9.1). These ZBGs showed potent affinity, were nontoxic, biocompatible, and showed promise for incorporation into a new family of MMP inhibitors (Lewis, Tran et al. 2005; Puerta, Griffin et al. 2006). The studies reported here examine the use of these alternative ZBGs as warheads for CAIs. Here we present the investigation of four ligands from this series. Compounds **1** to **4** have been tested in a kinetic enzyme assay on all catalytically active mammalian CAs (Scheme 9.1). The crystal structure of CA II in complex with **4** was determined and revealed an intriguing bidentate binding mode. Fragment **4** coordinates the active site  $Zn^{2+}$  ion in a chelating, bidentate mode and also interacts with Thr199 and Thr200. In addition, the compound participates in a hydrogen bonding network which is translated via water molecules within the entire active site of CAII.

## 9.4 Results

**Inhibition data:** The inhibition data for compounds **1** to **4** against 13 CA isoforms (hCA - mCA XV) are given in Table 9.1. All of these compounds show affinities

against the tested isoforms in the range of 2-11  $\mu\text{M}$ . The inhibition of each isoform with compounds **1** to **4** is generally very similar ( $<2.5 \mu\text{M}$  range). The best inhibition was obtained with maltol (**1**) against mCA XV with a  $K_i$  value of 2.74  $\mu\text{M}$ ; the weakest inhibition was observed for 1,2-HOPTO (**4**) against hCA XIII with a  $K_i$  value of 10.81  $\mu\text{M}$ . Isoforms hCA I, hCA III, hCA IX, hCA XIII, and hCA XIV show generally weaker inhibition, with  $K_i$  values ranging from 7.53 to 10.81  $\mu\text{M}$ , when compared to all other isoforms where  $K_i$  values of 2.74 to 4.89  $\mu\text{M}$  are found.



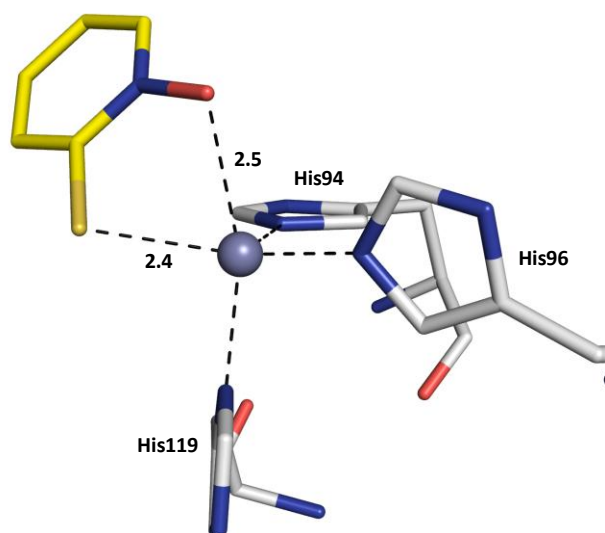
Scheme 9.1 Zinc-binding groups (ZBGs) used in this study as CA inhibitors.

Table 9.1 Inhibition data against **1** – **4** on CA isoforms.

Isoform	$K_i$ ( $\mu\text{M}$ )*				
	<b>1</b> (maltol)	<b>2</b> (thiomaltol)	<b>3</b> (HOPO)	<b>4</b> (HOPTO)	<b>5</b>
hCA I	9.07	8.31	9.66	9.37	7.00
hCA II	3.62	4.00	3.87	4.06	50nM
hCA III	9.50	8.43	9.12	7.53	-
hCA IV	4.25	4.61	3.90	4.37	-
hCA VA	4.60	4.15	4.35	4.02	-
hCA VB	4.31	4.02	4.43	3.81	-
hCA VI	3.47	3.24	3.89	3.62	-
hCA VII	3.44	3.57	3.81	3.55	-
hCA IX	9.54	9.95	9.27	9.23	27nM
hCA XII	3.96	4.45	4.74	4.76	-
hCA XIII	10.41	10.37	10.76	10.81	-
hCA XIV	7.15	6.71	6.80	7.14	-
mCA XV	2.74	3.14	4.89	4.71	-

\* Errors in the range of  $\pm 10\%$  of the reported value, from 3 different assays.

**X-ray Crystallography:** The crystal structure of CA II in complex with fragment **4** (1,2-HOPTO) was determined at a resolution of 1.35 Å. As expected, fragment **4** binds to CA II coordinating to the active site  $\text{Zn}^{2+}$  ion. The exocyclic sulfur- and oxygen atoms participate in the coordination of the metal ion at distances of 2.4 Å (S) and 2.5 Å (O) (Figure 9.1). Importantly, the coordination of the  $\text{Zn}^{2+}$  ion changes upon binding of **4** from a 4-coordinate, tetrahedral geometry (including the bound  $\text{OH}^-$  ion, PDB-Code: 3D92) (Domsic, Avvaru et al. 2008) to a distorted, 5-coordinate trigonal-bipyramidal geometry. Two nitrogen atoms (NE2-His94 and NE2-His96) and the sulfur atom of **4** form the equatorial plane, while the ND1 of His119 and the oxygen atom of **4** reside at the axial positions. A similar coordination change from 4- to 5-coordinate has been proposed for the binding of fragment **4** to MMPs (Puerta and Cohen 2003).



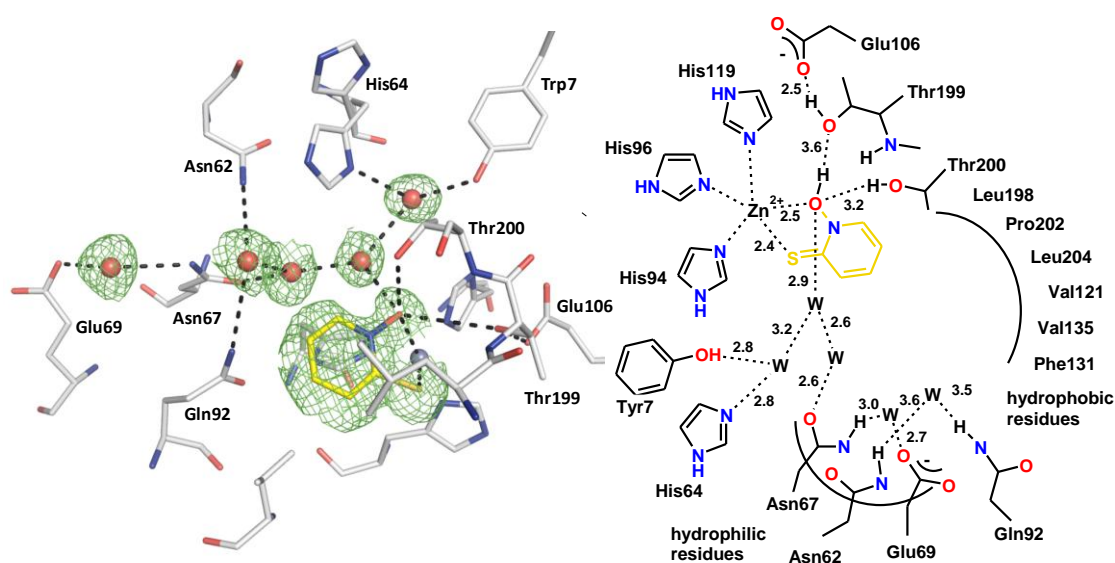
**Figure 9.1** Coordination geometry of the active site  $\text{Zn}^{2+}$  ion upon binding of **4**. The protein residues are shown in stick representation (protein: C, N, O, S; ligand: C, N, O, S). The tetrahedral geometry is changed to distorted trigonal-bipyramidal arrangement with NE2-His94 (2.0 Å), NE2-His96 (2.0 Å) and the sulfur of **4** (2.4 Å) in the equatorial plane, whereas ND1-His119 (2.1 Å) and the oxygen of **4** (2.5 Å) are in the axial positions. The angles in the plane are 143.5° (S-Zn-His96), 109.5° (S-Zn-His94) and 101.7° (His94-Zn-His96). The two axial ligands and the zinc ion form an angle of 161.2°.

The equatorial plane shows  $\text{Zn}^{2+}$  distances of 2.0 Å for NE2-His94, 2.0 Å for NE2-His96, and 2.4 Å for the sulfur atom with angles of 143.5° (S-Zn-His96), 109.5° (S-Zn-His94), and 101.7° (His94-Zn-His96). The  $\text{Zn}^{2+}$  ion is positioned 0.25 Å away from the plane that is formed by the equatorial coordinating ligands. The axial oxygen

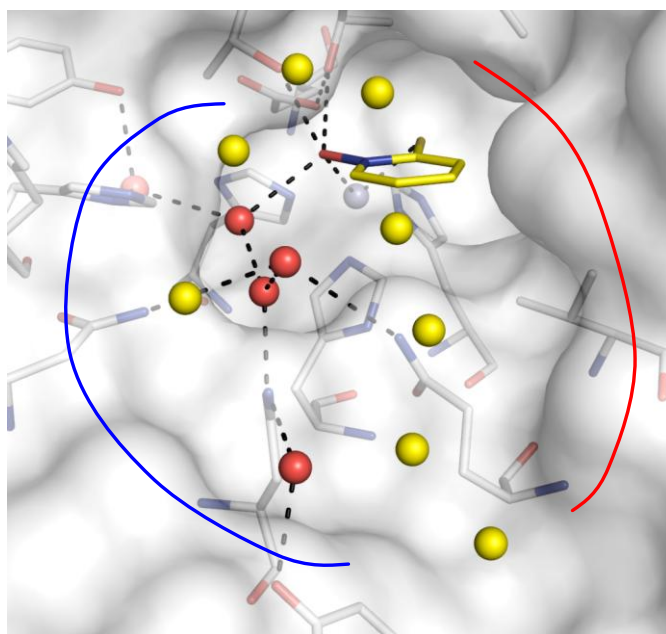
forms a distance of 2.5 Å and ND1-His199 is positioned 2.1 Å away from the metal ion. The two axial coordinating ligand atoms and the  $\text{Zn}^{2+}$  ion form an angle of 161.2°. In addition to metal coordination, **4** also participates in three hydrogen bonds (Figure 9.2). The terminal oxygen of Thr200 forms a hydrogen-bond to the oxygen in **4** at a distance of 3.2 Å. Furthermore, the same oxygen in **4** acts as the donor in a hydrogen-bond network via the terminal oxygen of Thr199 (3.6 Å) to the terminal group of Glu106 (2.5 Å). This network is well known for typical sulfonamide inhibitors; however, the distance to Thr199 is increased as the fragment most likely cannot move closer to Thr199. Interestingly, fragment **4** participates in a hydrogen-bonding network that is spread over the entire binding pocket, which provides additional stabilization in the active site of CA II. The interactions are mediated by five water molecules which address all residues in the hydrophilic pocket of CA II, i.e. Tyr7, Asn62, His64, Asn67, Glu69, and Gln92. Fragment **4** is linked via two water molecules to the terminal OH of Tyr7 and ND1 of the proton shuttle His64 in the "in" conformation. In addition, **4** is connected via two water molecules to Asn67 that forms an additional hydrogen-bond network to Glu69 at the exit of the binding pocket mediated by a water. The terminal NH of Gln92 and Asn62 are also participating in this network via a water molecule (Figure 9.3). In addition to these direct interactions with water molecules, several other water molecules close to the exit of the binding pocket are involved in this network (Figure 9.3). The distribution of *B*-values of the participating water molecules illustrates the well-coordinated water chain (20.5 Å<sup>2</sup>, 14.8 Å<sup>2</sup>, 26.3 Å<sup>2</sup>, 32.8 Å<sup>2</sup> and 17.1 Å<sup>2</sup>).

The conjugated ring system of **4** also experiences lipophilic interactions as it is pointing towards a hydrophobic binding pocket. In total, 26 van der Waals interactions can be identified. The difference electron density of **4** is very well defined and the orientation of the fragment is unambiguous because the sulfur atom and the ring system can be clearly assigned due to the difference electron density and assigned *B*-values.





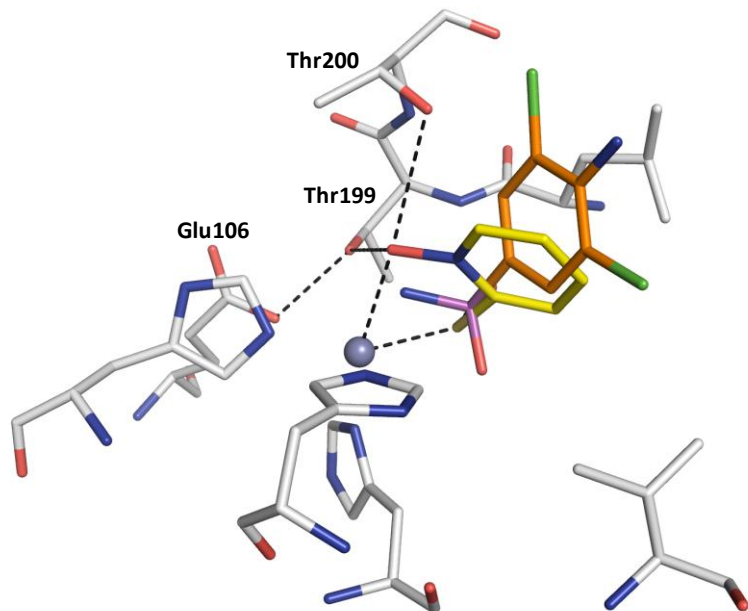
**Figure 9.2** Binding mode of **4** in complex with CA II. On the left hand side the crystal structure is shown with the important amino acids. The protein residues and the ligand are shown in stick representation (protein: C, N, O, S; ligand: C, N, O, S). Water molecules are shown as red spheres. Hydrogen bond interactions are indicated by the dashed lines. The equivalent distances are shown in the schematic representation of the binding mode on the right hand side. The difference electron density ( $F_o - F_c$ ) for the ligands and directly interaction water molecules is shown at a sigma level of 2.0. Atomic distances are shown in Å. The fragment coordinates with its sulfur and oxygen atom to the active site zinc ion. Hydrogen bonds to Thr200 and Thr199 as well as a water network which is spread over the entire binding pocket linking the fragment to the common hydrophilic residues can be observed.



**Figure 9.3** Binding mode of **4** in complex with CA II. The solvent accessible surface is represented in white. The interacting protein amino acids and the ligand are shown in stick representation (protein: C, N, O, S; ligand: C, N, O, S). The binding pocket of CA II can be divided into a hydrophobic (red line) and a hydrophilic binding region (blue line). The fragment is stabilized by an H-bond network which is mediated by five water molecules (red spheres) addressing the typical hydrophilic residues in the active site of CA II. The remaining waters in the binding pocket (yellow spheres) are connected to the H-bond network as well, however are not performing any further strong interactions to the protein.

The positioning of the oxygen was ambiguous at the beginning of the refinement. However, repeated refinement cycles using different fragment orientations (ring flip) and considerations about a consistent H-bond network render the assigned binding mode as the most likely. The *B*-values for **4** are somewhat larger compared to the protein residues which is probably related to weak fragment affinity or reduced occupancy. The *B*-values are: main chain 14.1 Å<sup>2</sup>, side-chain 18.2 Å<sup>2</sup>, waters 27.1 Å<sup>2</sup>, ligand **4** 24.1 Å<sup>2</sup> and Zn<sup>2+</sup> ion 8.4 Å<sup>2</sup>.

The superposition of **4** with a sulfonamide inhibitor **5** (PDB-Code 1ZGE) reveals that the sulfur atom of **4** adopts a position shifted by 0.9 Å to the sulfonamide sulfur (Figure 9.4) (Ilies, Vullo et al. 2003). The fact that in **4** no terminal nitrogen atom is available to coordinate the Zn<sup>2+</sup> ion influences the position of the sulfur. Furthermore, the ring system is shifted compared to the benzyl-group of the sulfonamide inhibitor. Because the oxygen atom is coordinated to the Zn<sup>2+</sup> ion and stabilized by Thr199 and Thr200, the ring system is slightly rotated and shifted towards Zn<sup>2+</sup> ion when compared to the sulfonamide fragment.



**Figure 9.4** Superposition of **4** and sulfonamide inhibitor **5** in complex with CA II. The protein residues and ligands are shown in stick representation (protein: C, N, O, S; **4**: C, N, O, S; sulfonamide **5**: C, N, O, S, Cl - PDB-Code 1ZGE). The sulfur atoms are located at the same position. Fragment **4** forms an additional coordination to the Zn<sup>2+</sup> ion. The ring system of **4** is shifted compared to the sulfonamide inhibitor which is dependent on the H-bonds of **4** to the oxygen of Thr199 and Thr200.

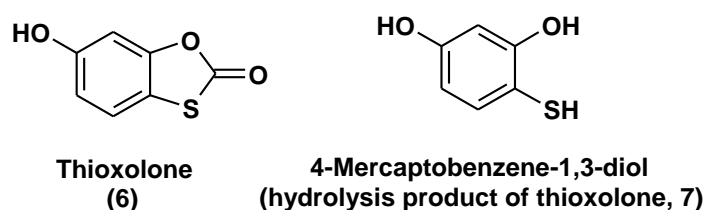
## 9.5 Discussion

The inhibition of **1** to **4** with respect to the considered CA isoforms is not exceptionally strong; however, they reveal a new and promising anchoring motif for CA inhibitors. The crystallographic analysis of the CA II-**4** complex reveals a bidentate binding mode to the  $\text{Zn}^{2+}$  ion with interactions to Thr199, which itself donates a hydrogen-bond to Glu106. Such a binding has been observed earlier only in the hCA II – N-hydroxyurea complex, in which the  $\text{Zn}^{2+}$  is coordinated by the N and O atoms of the inhibitor (which is presumably deprotonated) (Temperini, Innocenti et al. 2006). In addition, **4** accepts an H-bond from Thr200. The weaker affinity towards some of the isoforms can easily be understood since these show variable amino acid compositions within the active site, e.g. at position 200. For example, in hCA XIII a valine and in hCA I a histidine is present at this position (and many others, mainly towards the exit of the cavity). These residues cannot produce a similar interaction with **4**, which leads to a more than 2-fold loss in affinity against these isoforms.

From the inhibition, the crystallographic, and previously reported modeling data we can also infer the binding mode of **1** to **3** (Puerta and Cohen 2003). As the inhibition profile of all four compounds is the same against each isoform, and their metal complexes with a tris(histidine) active site model complex are all similar, the binding mode of **1** to **3** are most likely nearly identical to that of **4** (Puerta and Cohen 2003). Indeed, in a previously reported tris(pyrazolyl)borate  $\text{Zn}^{2+}$  model complex, fragment **4** was shown to bind in a bidentate fashion to the  $\text{Zn}^{2+}$  center, generating a 5-coordinate, distorted trigonal-bipyramidal coordination geometry at the  $\text{Zn}^{2+}$  ion. In this model complex, the Zn-S distance was 2.32 Å and the Zn-O distance was 2.08 Å (Puerta and Cohen 2003). In the structure reported here, the Zn-S distance is ~2.4 Å, very close to that observed in the model complex. The Zn-O distance in the complex of **4** with CA is longer than found in the model complex, at 2.5 Å, probably due to the asymmetric coordination of the ligand enforced by the CA active site. Nonetheless, it is likely that the related fragments **1**, **2** and **3** also coordinate in a chelating fashion with oxygen and/or sulfur atoms to the  $\text{Zn}^{2+}$  ion in bidentate manner.

A similar CA II complex has been determined crystallographically and recently published by Barrese et al (Barrese, Genis et al. 2008). Investigating the inhibition

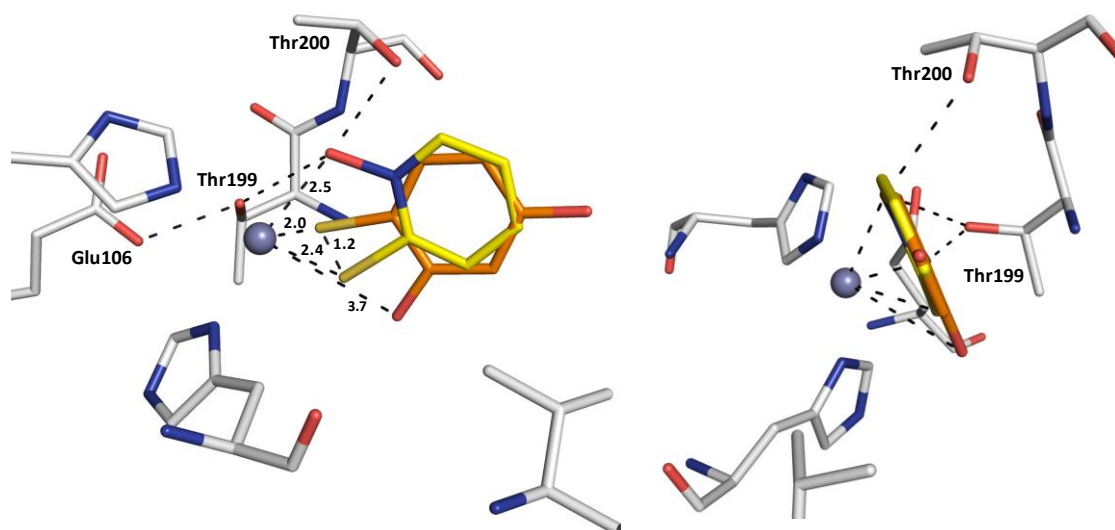
profile and binding mode of thioxolone **6** in CA II, they observed that the inhibitor, most likely a prodrug, is hydrolyzed upon binding forming the fragment 4-mercaptobenzene-1,3-diol **7** (Scheme 9.2, Figure 9.5).



**Scheme 9.2** Thioxolone (**6**) and its hydrolysis-fragment (**7**) from PDB-Codes 2OSE.

However, superposition of **7** with **4** shows significant differences in binding mode. Although, both inhibitor ring systems are similarly substituted with oxygen and sulfur donor atoms, the interaction pattern with CA II is completely different. The ring systems are rotated by 180° relative to each other and in **7** the oxygen donor atom is not interacting with Thr199 and Thr200. Also, the sulfur atom in **7** resides in a different position, in total shifted by 1.2 Å. It appears that the nitrogen atom in the heterocycle of **4**, which is not present in the compared fragment **7**, as well as the thione (C=S) nature of the donor functionality in **4** influences the binding mode and improves the Zn<sup>2+</sup> coordination in CA II. The trigonal-bipyramidal coordination of **4** and the interactions to Thr199 and Thr200 result in an excellent set of interactions of fragment **4** with the active site Zn<sup>2+</sup> ion. The thiol formed from thioxolone is a monodentate zinc ligand whereas **4** is a bidentate one.

The binding affinities of **1-4** to CA II are significantly better when compared to other alternative ZBGs such as the sulfamide (1130  $\mu$ M) and sulfamic acid (390  $\mu$ M) groups (Winum, Scozzafava et al. 2007). The observed bidentate binding mode of **4** to CA II is very promising and rarely observed in CAIs. The substitution of **4** with additional functional groups should notably increase the affinity toward CAs, as shown previously e.g. for sulfamides (Winum, Scozzafava et al. 2006). Lipophilic moieties can be attached to interact with the hydrophobic cavities found inside CAs. In addition, hydrophilic residues such as Asn62, Asn67, Glu69, and Gln92 should be targeted for hydrogen bonding. Further development of this ZBG could lead to a new, promising class of CAIs.



**Figure 9.5** Superposition of **4** with a structurally related fragment **7** from a crystal structure (PDB-code: 2OSF (C, N, O, S)). The sulfur atoms are shifted by 1.2Å. The ring system of the superimposed ligand is rotated by 180°. In contrast to this ligand, **4** can form hydrogen bonds to Thr199 and Thr200.

## 9.6 Experimental Section

**CA inhibition:** The inhibition assay is described in section 8.5.

**Synthesis:** Maltol (**1**), 1,2-HOPO (**3**), and 1,2-HOPTO (**4**) were obtained from commercial sources (Aldrich). Thiomaltol (**2**) was synthesized as previously described (Lewis, Puerta et al. 2003).

**Crystallization, X-ray data collection and refinement:** Crystals of hCA II (from Sigma-Aldrich) in complex with **4** were obtained by the sitting drop technique, using 10 mg/ml protein with 2 mM inhibitor in 50 mM Tris -HCl buffer (pH 8.0). The drops consisted of 5  $\mu$ L of the enzyme-inhibitor solution and 5  $\mu$ L of the precipitant solution containing 3.0 M  $(\text{NH}_4)_2\text{SO}_4$  in 100 mM Tris-HCl (pH 8.0) and 0.15 mM *p*-chloromerurybenzoic acid to promote the growth of highly oriented crystals. The drops were equilibrated by vapor diffusion against the precipitant solution at 18 °C and crystals appeared after 1-4 days.

The data were collected at the synchrotron BESSY II in Berlin/Germany on PSF beamline 14.2. The experiment was performed at 100 K using synchrotron radiation of 0.91841 Å wavelength on a MAR CCD 165 mm detector. The crystal was monoclinic

$P2_1$  with the following cell parameters:  $a = 42.3 \text{ \AA}$ ,  $b = 41.4 \text{ \AA}$ ,  $c = 72.4 \text{ \AA}$ ,  $\beta = 104.4^\circ$ . The data collection parameters and refinement statistics are reported in the Appendix 13.1.

Data were processed and scaled with Denzo and Scalepack as implemented in HKL2000 (Otwinowski, Minor et al. 1997). The complex structures were analyzed by difference Fourier techniques, using the PDB file 2OQ5 as starting model for refinement. Refinement was continued with CNS (Brünger, Adams et al. 1998) and SHELXL-97 (Sheldrick, Schneider et al. 1997). For each refinement step at least 10 cycles of conjugate minimization were performed, with restraints on bond distances, angles, and  $B$ -values. For each crystal the  $2F_o-F_c$  and  $F_o-F_c$  maps were calculated and the inhibitor binding mode assigned from the difference electron density. Intermittent cycles of model building were done with the program COOT (Emsley and Cowtan 2004). The coordinates have been deposited in the PDB (<http://www.rcsb.org/pdb/>) with the reference code 3M1K.

## 10 Conclusion and Outlook

In this thesis a variety of rather diverse fragment-based lead discovery approaches with the target protein carbonic anhydrase II were performed. The different projects were strongly supported and tailored towards protein crystallography; a method which is currently emerging as a routine analytical tool. This maturation mainly results from improved radiation sources and enhanced computing power. About 200-250 datasets were collected in due course of this thesis answering miscellaneous questions in the context of several ambitious projects focused on structural biology. How can interactions of proteins with small molecules be exploited? Are we able to, e.g. monitor *in situ* reactions using protein crystallography? Will a tethering approach fixing weak initial binders increase the success rate of fragment-based crystallography?

In order to cover this broad scope of diverse projects access to large amounts of CA II is a prerequisite. Therefore, the gene coding for CA II was cloned into a high-level expression system, the GST gene fusion system. The expression system was optimized yielding target protein of up to 30 mg from 1 L of original cell culture. Several CA II mutants were produced via site-directed mutagenesis and a mercury-free protein crystallization protocol had to be established for both the wild type and all mutants.

Initial goal of the first project of this thesis was the accomplishment of a prominent click chemistry reaction, the Huisgen reaction – Cu<sup>+</sup> catalyzed 2+3 dipolar cycloaddition – within or next to the active site of CA II. Therefore, an artificial copper center had to be introduced at the surface of CA II which should be used by azide and alkyne building blocks to catalyze triazole formation. Favorable Cu coordinating residues, such as cysteines or histidines which are well-known to participate in the metal center coordination in naturally occurring copper enzymes, were arranged at the surface of CA II according to a rational design concept partially introduced by site-directed mutagenesis. A variety of mutants was produced and investigated with respect to its Cu coordinating facilities. Unfortunately, this rational design strategy remained unsuccessful which shows that our limited understanding of

amino acid modifications still might provoke unpredictable effects while tampering with complex biological systems such as a highly functionalized protein. Finally, by serendipity the formation of an artificial Cu center at the surface of CA II could be achieved. A square planar coordination by residues Ser2, His3 and His4 occurred upon soaking crystals with a  $\text{Cu}^+$  solution of the CA II-H64C mutant which had previously been tethered with an azide fragment. In further experiments beyond the scope of this thesis a set of click reaction building blocks should be designed and synthesized to address this introduced copper ion. Thereby its catalytic potential can be screened again with strong support by protein crystallography. In addition, further mutational studies should be performed guided by the properties of human metallochaperone HAH1, a copper containing protein. A four-membered amino acid loop of the latter protein which coordinates the metal ion should be introduced at the N-terminus of CA II. Possibly the thus modified CA II mutants will show an accurate Cu coordination center.

In parallel to the formerly described metal center design, the click chemistry project was also promoted by a new tether-assisted approach. By covalently attaching one building block to the protein surface and linking the second reactant via a reversible sulfonamide anchor to the active site  $\text{Zn}^{2+}$  ion, both components were brought into vicinity favorable to undergo cycloaddition reaction. The reaction could be induced by either  $\text{Cu}^+$  ions or the environment of CA II, respectively. As the azide component was tethered to a mutationally introduced cysteine residue at position 64 and the alkyne reversibly bound to the active site, a crystal structure of two initially designed building blocks in complex with CA II-H64C could be determined. The crystal structure revealed that both reactants are fixed to the protein in a distinct binding mode. In order to initiate triazole formation based on this setup, crystals from the same crystallization well were soaked with a  $\text{Cu}^+$  solution. The crystallographic analysis following this experiment showed pronounced structural reorganization of the reagents now suggesting triazole formation at some level. Complementary mass spectrometric analysis of this reaction mixture showed very poor cyclization rates suggesting that the initial design of the reaction components had not been ideal. Similar experiments performed with improved building blocks revealed much higher yields of triazole



formation. This was studied by HPLC-MS even allowing recording a time-dependent transformation rate. Crystallographic analysis of CA II-H64C cocrystallized with the optimized azide and alkyne component supported the 'in solution' transformation. The protein environment of CA II-H64C induced triazole formation upon the exposure to crystallization without further supplementation of any catalyst such as  $\text{Cu}^+$ . Remarkably, the triazole formation was executed with regio- and stereoselectivity. The *1,5-S*-triazole was selectively formed from a racemic building block mixture. Most likely, the spatial organization of the reactants under exchange equilibrium conditions within the binding pockets facilitates the formation of one certain triazole product. In solution experiments analyzed by HPLC-MS revealed that this tethered cycloaddition can be further accelerated by copper supplementation and increased temperatures. In further studies, the stereo- and regioselectivity should be investigated in more detail by HPLC-MS. In crystallization experiments the cyclization product from both alkyne enantiomers, individually cocrystallized with differently tethered azides, should be investigated and compared with results under solution conditions. To improve the "catalytic" effect of CA II-H64C and to transform the protein into a full catalyst for click reactions, the cyclization product has to leave the binding pocket after the reaction. Possibly the substitution of the alkyne component with a rather weak zinc coordinating group and the application of a reducing agent to cleave the disulfide tether can accomplish this task.

Apart from this click chemistry approach, the tethering method was also applied for a fragment screening on CA II. Tethering allows the identification of fragments with rather weak affinity to the target protein which are extracted from a library of thiole containing compounds under reducing conditions. Stimulated by a CA II Saccharin complex crystal structure and the observation of a novel binding site at the surface of CA II, this approach was, among others, setup to explore this novel binding region. Thereby, possibly new CA II inhibitors can be developed which interact with this area adjacent to the proton shuttle His64, which is important for the catalytic rate of CA II. A computational docking study was applied to screen a virtual library and select a feasible number of compounds for synthesis. The 'in solution' experiments monitored by HPLC-MS revealed many candidate fragments with promising affinity to CA II.

The crystallographic analysis of several CA II-H64C tethered complexes showed that the Saccharin binding pocket at the surface could be successfully addressed by the tethering approach. Also a second surface exposed binding area, known from CA II activators could be screened. According to the ranking from the mass spectrometric experiments, further tethering complexes should be investigated by protein crystallography. Guided by the results from our crystal structure determinations an improvement of the potency of these first leads could be planned. Furthermore, fragments which have been identified to bind in separate areas of the CA II binding pocket could be linked together to one larger ligand to achieve increased binding affinity. In conclusion this tethering approach highly facilitates the crystallographic investigation of protein fragment complexes and provides the opportunity for further application, e.g. the identification of new building blocks for *in situ click* chemistry or the setup of an inhibition assay by variation of the inhibitor or reducing agent concentration.

In a different fragment based lead discovery approach, four novel head groups for zinc coordination, originally synthesized in the group of Prof. Seth Cohen (San Diego, USA), have been investigated with respect to CA isoform inhibition by the group of Dr. Claudiu Supuran (Florence, Italy). In our laboratory they have been analyzed by protein crystallography. A complex structure of CA II and 1,2-HOPTO revealed a novel bidentate  $Zn^{2+}$  binding mode with favorable interactions to Thr199. The interaction to Thr199 is usually found in typical sulfonamide inhibitors; in addition the fragment binds to Thr200. The fragment which is perfectly coordinated by a water network suggests a new class of putative CA inhibitors. The binding affinity can be improved by adding substituents to occupy to the hydrophilic and hydrophobic portion of the CA II binding pocket. In further crystallographic trials the binding mode of the remaining three inhibitors should be investigated and eventually optimized for improved inhibition.

Apart from this novel zinc binding functionality, also for a series of 13 sulfamides from a different thesis (group of Prof. Haake) binding affinities against several CA isoforms (group of Dr. Supuran) and for several examples the CA II complex crystal structures have been determined. Several of these inhibitors experience interactions

with Thr199 and Thr200. The different complex structures showed that the E/Z-isomerisation which all inhibitors can undergo allows the enzyme to select the isomer configuration that fits best into its binding pocket. Remarkably, two compounds with similar substituents showed strongly deviating binding modes. The interaction patterns were influenced on the one hand side by favorable hydrogen-bond interactions to the hydrophilic portion of the protein binding pocket, on the other hand side an intramolecular H-bond is established which stabilizes the bound conformation. The inhibition data of a series of further 41 sulfamides still leave some open questions to set up a conclusive structure-activity relationship. Further crystallographic analyses of CA II-sulfamide complexes will be necessary in order to understand the impact of *in situ* E/Z-isomerism with respect to the binding affinity.



## 11 Zusammenfassung und Ausblick

Im Rahmen dieser Dissertation wurde eine Auswahl recht verschiedener Ansätze der Fragment-basierten Leitstruktursuche (*fragment-based lead discovery*) mit dem Zielprotein Carboanhydrase II durchgeführt. Die verschiedenen Projekte wurden entscheidend durch die Protein Kristallographie unterstützt und methodisch gelenkt. Durch Verbesserung von Strahlungsquellen und der Rechenleistung von Computern hat sich die Protein-Kristallographie in den letzten Jahren zu einem analytischen Werkzeug entwickelt, welches routinemäßig zur Strukturaufklärung eingesetzt werden kann. Circa 200 bis 250 Datensätze wurden im Rahmen dieser Arbeit gesammelt um Fragestellungen der verschiedenen ambitionierten Projekte mit Focus auf die Strukturbiologie zu beantworten. Wie kann die Wechselwirkung zwischen Proteinen und kleinen Substanzen für eine Wirkstoffentwicklung ausgenutzt werden? Ist es möglich, z.B. Reaktionen, die *in situ* ablaufen, durch Protein Kristallographie zu verfolgen? Kann die Erfolgsquote von Fragment-basierter Kristallographie durch einen *Tethering*-Ansatz, welcher schwach bindende Fragmente am Protein fixiert, erhöht werden?

Die Verfügbarkeit großer Mengen an CA II ist die Grundvoraussetzung um solch eine Bandbreite sehr verschiedener Projekte durchführen zu können. Daher wurde das CA II Gen in das *GST gene fusion System* kloniert, ein System welches Gen-Expression auf sehr hohem Niveau ermöglicht. Das Expressionssystem wurde hinsichtlich der Proteinausbeute optimiert und stellt 30 mg Protein aus 1 L Zellkultur bereit. Etliche CA II Mutanten wurden durch positionsgerichtete Mutation hergestellt und eine quecksilberfreie Protein-Kristallisation für den Wildtyp und die Mutanten etabliert.

Die Durchführung einer prominenten Click-Chemie-Reaktion, der Huisgen Reaktion – eine  $\text{Cu}^+$  katalysierte 2+3 dipolare Cycloaddition – in der Bindetasche der CA II war das anfängliche Ziel des ersten Projektes dieser Dissertation. Dafür sollte ein künstliches Kupfer-Zentrum an der Oberfläche der CA II eingeführt, und dieses neue katalytische Zentrum für die Bildung von Triazolen aus Alkinen und Aziden

ausgenutzt werden. Gemäß eines rationalen Design-Konzeptes wurden Kupfer-koordinierende Aminosäuren, wie z.B. Cysteine oder Histidine, die bekannter Weise in Metall-Zentren natürlicher Kupfer-Proteine favorisiert vorkommen, teilweise per positionsgerichteter Mutation an der Oberfläche der CA II angeordnet. So wurde eine Vielfalt von CA II Mutanten hergestellt und auf ihre Kupfer-koordinierenden Eigenschaften hin untersucht. Leider blieben diese Versuche eines rationalen Designs erfolglos. Dies zeigt auf, dass unser begrenztes Verständnis von Aminosäure-Mutationen im Hinblick auf die Proteinarchitektur in komplexen biologischen Systemen, wie z.B. in hoch entwickelten Proteinen, schwer vorhersehbare Auswirkungen mit sich bringen kann. Schließlich, eher durch einen glücklichen Zufall als rationales Design, konnte ein Kupfer-Zentrum an der Oberfläche der CA II eingebracht werden. Nachdem ein Kristall der CA II-H64C Mutante, zuvor mit einem Azid-Fragment gekuppelt, mit einer  $\text{Cu}^+$  Lösung versetzt wurde, konnte die Bildung eines quadratisch-planaren Metall Zentrums durch die Reste Ser2, His3 und His4 beobachtet werden. In weiteren Experimenten, über diese Dissertation hinaus, sollten die Ausgangssubstanzen für die Click-Reaktion so entworfen und synthetisiert werden, dass diese das Kupfer-Zentrum adressieren. Somit kann das katalytische Potential mit Hilfe der Protein-Kristallographie untersucht werden. Zusätzlich sollten weitere Mutations-Studien durchgeführt werden, die sich an den Eigenschaften eines Kupfer-Proteins, des humanen Metallochaperons HAH1, orientieren. Ein Loop aus vier Aminosäuren, der das Metallion in diesem Protein koordiniert, sollte am N-Terminus der CA II eingefügt werden. Möglicherweise wird diese CA II Mutante die Eigenschaft haben ein Kupfer-Zentrum präzise zu formen.

Neben dem zuvor beschriebenen Design eines Metall-Zentrums, wurde das Click-Chemie-Projekt ebenso durch einen neuartigen *Tethering*-unterstützten Ansatz vorangetrieben. In diesem Ansatz wird eine Reaktionskomponente kovalent an die Proteinoberfläche gebunden, während die zweite Komponente durch einen Sulfonamid-Anker an das Zinkion im aktiven Zentrum bindet. Dadurch werden beide Edukte in eine für die nachfolgende Zyklisierung günstige räumliche Anordnung gebracht. Die Reaktion könnte entweder durch  $\text{Cu}^+$  Ionen oder aber durch die Proteinumgebung ausgelöst werden. Die Tatsache, dass die Azid-Komponente an ein

in Position 64 mutiertes Cystein und die Alkin-Komponente reversibel in der aktiven Tasche gebunden wird, ermöglichte die Aufklärung einer Kristallstruktur von CA II im Komplex mit den ursprünglich entworfenen Edukten. Diese Kristallstruktur zeigte, dass beide Edukte in einem eindeutigen Bindungsmodus getrennt voneinander an das Protein binden. Um die Triazolbildung in diesem System auszulösen, wurde ein Kristall aus dem gleichen Kristallisationstropfen mit einer  $\text{Cu}^+$  Lösung versetzt. Eindeutige Veränderungen im Bindungsmodus beider Komponenten konnten durch die Strukturanalyse beobachtet werden und verweisen auf eine partielle Triazolbildung. Ergänzende massenspektrometrische Untersuchungen dieser Reaktionsansätze zeigten eine sehr schwache Cycloadditionsgeschwindigkeit und verdeutlichten, dass das anfängliche Design der Reaktionskomponenten nicht ideal war. Die Edukte für die Click-Reaktion wurden optimiert und die gleichen Experimente erneut durchgeführt. Die Reaktionsgeschwindigkeiten und Ausbeuten konnten so deutlich gesteigert werden. Die HPLC-MS Untersuchungen ermöglichten zusätzlich die Aufzeichnung einer zeitabhängigen Transformationsrate. Kristallographische Untersuchungen von CA II-H64C im Komplex mit diesem optimierten Alkin und Azid unterstrichen die Versuche, die in Lösung durchgeführt wurden. Die CA II-H64C Proteinumgebung induzierte die Triazolbildung aus den Edukten während der Kristallisation ohne, dass die Zugabe eines Katalysators, wie z.B. einer  $\text{Cu}^+$  Lösung, von Nöten war. Bemerkenswerterweise erfolgte die Triazolbildung mit einer Regio- sowie Stereoselektivität. Aus dem racemischen Alkingemisch wurde selektiv das *1,5-S*-Triazol gebildet. Höchstwahrscheinlich wurde das langsamer reagierende *S*-Alkin in einer Gleichgewichtsreaktion gegen das *R*-Alkin ausgetauscht und so durch die räumliche Organisation beider Ausgangssubstanzen in der Bindetasche die Bildung eines stereochemisch eindeutigen Triazols ermöglicht. Durch Experimente, die in Lösung durchgeführt und per HPLC-MS untersucht wurden, konnte bewiesen werden, dass diese kovalent fixierte Cycloaddition zusätzlich durch Kupfer-Zugabe und erhöhte Temperaturen beschleunigt werden kann. In weiterführenden Experimenten sollte diese Stereo- und Regioselektivität mit Hilfe von HPLC-MS Messungen untersucht werden. Weiterhin sollten beide Alkin-Enantiomere individuell mit verschiedenen, kovalent an das Protein verknüpften Aziden kristallisiert und die entstehenden Cyclisierungsprodukte mit den Ergebnissen aus den

massenspektrometrischen Experimenten verglichen werden. Um den "katalytischen" Effekt von CA II-H64C zu erhöhen und das Protein in einen vollwertigen Katalysator für Click-Reaktionen zu überführen, sollte das Cyclisierungsprodukt nach der Reaktion die Bindetasche wieder verlassen. Dies könnte möglicherweise erreicht werden, indem das Alkin mit einer Gruppe verknüpft wird, die nur sehr schwach an das aktive Zink bindet und indem zusätzlich reduzierende Bedingungen angewandt werden um die Disulfidbrücke wieder zu lösen.

Die *Tethering*-Methode wurde abgesehen von diesem Click-Chemie-Ansatz ebenso für ein Fragment *Screening* mit der CA II angewandt. Der *Tethering*-Ansatz ermöglicht die Identifizierung von Fragmenten mit eher schwacher Affinität zum Zielprotein. Diese werden unter reduzierenden Bedingungen aus einer Bibliothek von Thiol-Fragmenten gefischt. Durch eine Kristallstruktur eines CA II Saccharin Komplexes konnten wir im Vorhinein eine neue Bindestelle für Inhibitoren an der Oberfläche von CA II identifizieren. Mit Hilfe der *Tethering*-Methode sollte diese Binderegion besser ausgeleuchtet werden. Auf diesem Wege könnte möglicherweise eine neue Gruppe von CA II Inhibitoren entwickelt werden, welche mit dem His64 wechselwirken. His64, das so genannte "*Proton Shuttle*", spielt eine entscheidende Rolle in der katalytischen Aktivität von CA II. Eine Inhibition dieses Mechanismus würde folglich die katalytische Umsatzrate drastisch verringern. Durch einen computerbasierten *Docking*-Ansatz wurde eine virtuelle Bibliothek von Fragmenten getestet um so eine plausible Anzahl an Substanzen für die Synthese auszuwählen. Verschiedene *Tethering* Experimente wurden in Lösung durchgeführt und mittels HPLC-MS analysiert. Dadurch konnten viele Fragmente mit einer vielversprechenden Affinität zu CA II identifiziert werden. Einige dieser Substanzen wurden im kovalenten Komplex mit CA II kristallographisch untersucht und zeigten, dass die Saccharin Bindetasche mit diesem *Tethering*-Ansatz erfolgreich untersucht werden kann. Zusätzlich konnte eine zweite Bindetasche an der Oberfläche, bekannt für die Bindung von CA II Aktivatoren, untersucht werden. Die massenspektrometrischen Experimente, die in Lösung durchgeführt wurden, und die dadurch identifizierten vielversprechenden Fragmente sollten weiterhin kristallographisch untersucht werden. Die Ergebnisse dieser Untersuchungen sollten schließlich angewandt werden um die Affinität dieser



ersten Komponenten zu erhöhen. Außerdem könnten Fragmente, die in verschiedenen Bereichen der aktiven Tasche der CA II binden, miteinander zu einem Liganden mit erhöhter Affinität verknüpft werden. Abschließend lässt sich sagen, dass dieser *Tethering*-Ansatz die kristallographische Untersuchung von Protein-Fragment-Komplexen in höchstem Maße vereinfacht und weitere Anwendungsmöglichkeiten mit sich bringt. So könnten z.B. neue Ausgangssubstanzen für *in situ* Click-Chemie oder aber auch ein Inhibitionsassay durch Variation der Inhibitor oder DTT Konzentration entwickelt werden.

In einem weiteren Ansatz der Fragment-basierten Leitstruktursuche wurden vier neue Zink koordinierende Kopfgruppen an CA Isoenzymen getestet. Die Substanzen, die ursprünglich in der Gruppe von Prof. Seth Cohen (San Diego, USA) synthetisiert bzw. zusammengestellt wurden, sind in der Gruppe von Dr. Claudiu Supuran (Florenz, Italien) auf ihre Inhibition von CA Isoenzymen hin getestet worden. Im Rahmen dieser Arbeit wurden diese neuen Kopfgruppen proteinkristallographisch untersucht. Eine Komplexstruktur von CA II mit 1,2-HOPTO zeigt einen neuartigen zweizahnigen  $Zn^{2+}$  koordinierenden Bindungsmodus mit Wechselwirkungen zu Thr199. Die Wechselwirkung zu Thr199 ist üblicherweise bekannt von Sulfonamid-Inhibitoren. Zusätzlich wechselwirkt die Substanz mit Thr200. Das Fragment, welches zusätzlich perfekt durch ein Wassernetzwerk koordiniert ist, verspricht eine neue Klasse von CA-Inhibitoren. Durch die Addition von Substituenten, die mit den hydrophilen und lipophilen Bereichen der CA II Bindetasche wechselwirken, kann die Bindungsaffinität verbessert werden. In weiteren kristallographischen Experimenten sollten die Bindungsmodi der verbleibenden drei Verbindungen untersucht und im Anschluss optimiert werden.

Abgesehen von dieser neuen Zink-koordinierenden Gruppe konnte ebenfalls für eine aus einer anderen Dissertation stammenden (Gruppe Prof. Haake) Serie von 13 Sulfamiden die Bindungsaffinitäten gegen einige CA-Isoformen (Group Dr. Supuran) und für einige ausgewählte Beispiele eine Komplexstruktur mit CA II aufgeklärt werden. Einige dieser Inhibitoren interagieren mit Thr199 und Thr200. Die verschiedenen Strukturen zeigten, dass die E/Z-Isomerisierung, welche einige der Inhibitoren erfahren, es dem Protein ermöglicht, die isomere Konfiguration

auszuwählen, die am besten in die Bindetasche passt. Bemerkenswerterweise zeigen zwei ähnlich substituierte Inhibitoren einen deutlich abweichenden Bindungsmodus. Die Wechselwirkungsprofile wurden auf der einen Seite durch vorteilhafte Wasserstoffbrückenbindungen zu hydrophilen Bereichen der Protein Bindetasche und auf der anderen Seite durch intramolekulare Wasserstoffbrücken beeinflusst. Die Inhibitionsdaten einer Serie von weiteren 41 Sulfamiden lässt noch einige Fragen offen, um zusammenhängende Struktur-Wirkungs-Beziehungen aufstellen zu können. Die kristallographische Untersuchung von weiteren CA II-Sulfamid-Komplexen ist erforderlich um den Einfluss der E/Z-Isomerie auf die Bindungsaffinität vollständig nachzuvollziehen.

## 12 Materials and Methods

### 12.1 Materials

#### 12.1.1 Chemicals

**Table 12.1** Chemicals in alphabetical order.

<b>Name</b>	<b>Company</b>
1,4-Dithiothreitol (DTT)	Hartmann Analytik or Sigma Aldrich
Acrylamide/Bisacrylamide (30%/0.8%)	ROTH
Agar-Agar	ROTH
Ammoniumpersulfate (APS)	ROTH
Ammoniumsulfate	Merck
Ampicilline	ROTH
Bromphenolbue	Merck
Chloramphenicol	ROTH
Complete mini protease inhibitor, EDTA free	Roche
Coomassie Brilliant Blue R250	ROTH
Dimethylsulfoxid (DMSO)	Sigma Aldrich
Acetic acid 100%	ROTH
Ethanol 96%	Riedel-deHaën
Ethidiumbromid	Merck
Glycerol 100%	Merck
Yeast extract	ROTH
Isopropyl- $\beta$ -D-thiogalactopyranosid (IPTG)	ROTH
Lysozyme	Merck
Methanol	Fischer Scientific
<i>p</i> -Chlormercuribenzoic acid	Sigma
Sodium chloride	ROTH
Peptone from Casein	ROTH
Rotiphorese gel 30	ROTH
Tris(hydroxamethyl)aminomethane (Tris)	ROTH

## 12.1.2 Laboratory Devices

**Table 12.2** Laboratory devices and the producing companies in alphabetical order.

<b>Name</b>	<b>Company</b>
ÄKTA FPLC	Amersham Pharmacia Biosciences
Autoklav	Syntec Tuttnauer 3850EI
Blockthermostat	Roth Blockthermostat TCR 100
Brutschrank	Heareus Instrumente
Entsalzungsanlage Purelab Plus	USF Elga
Feinwaage Type 404/13	Sauter
Heizrührer IKA-COMBIMAG REG	Janke&Kunkel
PCR Mini Cycler	MJ Research
pH-Meter 744	Metrohm
Pipetten	Eppendorf
Schüttelinkubator Innova 4200	New Brunswick Scientific
Schüttelinkubator Innova 4300	New Brunswick Scientific
SDS-Gelelektrophoresekammer Mini-V8*10	Life Technologie
DNA Mini Speedvaq	Heto Labequipment Wika
Sterilbank 8511	Köttermann
Thermomixer Comfort 2 ml	Eppendorf
Ultraschall Sonifier	Branson
UV-Leuchttisch Transilluminator	Biostep
UV-Vis-Spektrometer Smart Spec 3000	BioRad
Vakuumpumpe DNA Mini	Jouan Nordic
Vortexer VF2	Janke&Kunkel
Waage Typ 572/45	Kern
Zentrifuge Avanti J-25	Beckmann Coulter
Zentrifuge Biofuge pico	Heaeus
Zentrifuge Multifuge 3	Heraeus

### 12.1.3 Solutions and Buffers

**Table 12.3 Puffer und Lösungen.**

Name	Substances
APS-Lsg.	10% (w/v) Ammoniumpersulfat ad 10 ml
Ampicilin-Lsg.	1.0 g Ampicilin ad 10 ml, steril filtrieren
Chloramphenicol-Lsg	0.34 g Chloramphenicol ad 10 ml, steril filtrieren
Comassie-Blue Entfärbelsg.	10% (v/v) Essigsäure, 40% (v/v) Methanol ad 1 L
Coomasie-Blue Färbelsg.	10% Essigsäure, 40% Methanol, 1 g Coomasie Brilliant Blau R250 ad 1 L
IPTG-Lsg.	1 M IPTG, steril filtrieren
Laemmlli-Buffer	0.25 M Tris, 2 M Glycerin, 1% (w/v) SDS, pH 8.3
Natriumhydroxid-Lsg.	0.1 M NaOH
PBS-Buffer	140 mM NaCl, 10 mM Na <sub>2</sub> HPO <sub>4</sub> , 2.7 mM KCl, 1.8 mM KH <sub>2</sub> PO <sub>4</sub> , pH 7.3
Elutionspuffer	50 mM Tris-HCl, 10 mM reduziertes Glutathion, pH 8.0
Cleavage-Puffer	6 M Guanidine-HCl
SDS-Sammelgel-Buffer	1 M Tris, pH 6.8
10% SDS-Lsg.	10% (w/v) SDS ad 50 ml
SDS-Proben-Puffer	250 mM Tris, 8% (v/v) SDS, 40% (v/v) Glycerol, 0.04% (w/v) Bromphenolblau, 8% (v/v) β-Mercaptoethanol, pH 6.8
Trenngel-Puffer	1 M Tris, pH 8.8
Tris-Puffer pH 8.0	50 mM Tris-HCl pH 8.0
Kristallisationspuffer CA II	2.5 - 3 M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , 50 mM Tris pH 8.0 (evtl. 0.15 mM <i>p</i> -Chlormercuribenzoessäure)
Crypuffer CA II	3 Teile Kristallisationspuffer CA II, 1 Teile Glycerol
LB-Medium	1.0% (m/v) Pepton aus Casein, 0.5% (m/v) Hefe, 1% (m/v) NaCl, autoklaviert
Agar-LB-Medium	1.0% (m/v) Pepton aus Casein, 0.5% (m/v) Hefe, 1% (m/v) NaCl, 1.5% (m/v) Agar-Agar, autoklaviert
Lysepuffer	PBS-Puffer, Protease-Inhibitoretabelle (Complete™), Lysozym

### 12.1.4 Primer and Bacterial Strains

**Table 12.4** Primer

Name	Sequence (5' → 3')
EcoRICA2s	CGG AAT TCA TGT CCC ATC ACT GGG GGT A
XhoICA2a	ACC GCT CGA GTT ATT TGA AGG AAG CTT TGA TTT GC
CA2W5Cs	ATG TCC CAT CAC TGT GGG TAC GGC AAA CA
CA2W5Ca	TGT TTG CCG TAC CCA CAG TGA TGG GAC AT
CA2G63Cs	ATC CTC AAC AAT TGT CAT GCT TTC AAC
CA2G63Ca	GTT GAA AGC ATG ACA ATT GTT GAG GAT
CA2H64Cf	GAG GAT CCT CAA CAA TGG TTG CGC TTT CAA CGT GGA GTT TG
CA2H64Crev	CAA ACT CCA CGT TGA AAG CGC AAC CAT TGT TGA GGA TCC TC
CA2H64Mf	GAG GAT CCT CAA CAA TGG TAT GGC TTT CAA CGT GGA GTT TG
CA2H64Mrev	CAA ACT CCA CGT TGA AAG CCA TAC CAT TGT TGA GGA TCC TC
CA2H96Ef	GAT TGA TTC AGT TTC ACT TTG AGT GGG GTT CAC TTG ATG GAC
CA2H96Er	GTC CAT CAA GTG AAC CCC ACT CAA AGT GAA ACT GAA TCA ATC
CA2H96Kf	GAT TGA TTC AGT TTC ACT TTA AAT GGG GTT CAC TTG ATG GAC
CA2H96Kr	GTC CAT CAA GTG AAC CCC ATT TAA AGT GAA ACT GAA TCA ATC
CA2E106Lf	CTT GAT GGA CAA GGT TCA CTG CAT ACT GTG GAT AAA AAG
CA2E106Lr	CTT TTT ATC CAC AGT ATG CAG TGA ACC TTG TCC ATC AAG
CA2T198Vf	GAC CTA CCC AGG CTC ACT GGT GAC CCC TCC TCT TCT GGA ATG
CA2T108Vr	CAT TCC AGA AGA GGA GGG GTC ACC AGT GAG CCT GGG TAG GTC

**Table 12.5** Used *E. coli* strains

Strain	Characteristics	Company
XL2-blue	<i>(mcrA)183(mcrB-hsdSMR-mrr)173</i> <i>endA1 supE44 thi-1 recA1 gyrA96 relA1</i> <i>lac[F'proAB laqI<sup>fl</sup> ZΔM15 Tn5(Km<sup>r</sup>)]</i>	Stratagene
BL21Codon plus	F', <i>hsdS</i> (r <sub>B</sub> , m <sub>B</sub> ) <i>gal</i> , (DE3), <i>ompT</i> ., <i>endA</i> Hte[ <i>argU ileY leuW cam<sup>r</sup></i> ]	Stratagene

## 12.2 Methods

### 12.2.1 Biomolecular Methods

#### 12.2.1.1 Polymerase Chain Reaction (PCR)

The PCR-sample was prepared as follows:

10 µl	5x Polymerase buffer
1 µl	Template DNA (10 ng/µl)
1 µl	Primer 1 (200 µM)
1 µl	Primer 2 (200 µM)
1 µl	dNTP-Mix (25 mM each)
35 µl	dd H <sub>2</sub> O
1 µl	Polymerase

#### PCR-protocol:

1.	30 sec	98 °C
2.	10 sec	98 °C
3.	30 sec	65 °C x3
4.	30 sec	98 °C
5.	7 min	72 °C
6.		4 °C

The success of the PCR was controlled via standard agarose gel electrophoresis and the PCR-fragment extracted according to the standard protocol of the S.N.A.P. UV-Free Gel Purification Kit (Invitrogen).

### 12.2.1.2 Restriction

The PCR-fragment was incubated with the restriction endonuclease XhoI and EcoRI. The sample was incubated for 2 hours at 37 °C.

50 µl	PCR-fragment
0.8 µl	XhoI
0.8 µl	EcoRI
5.8 µl	10x buffer
0.6 µl	dd H <sub>2</sub> O

Subsequently the restricted PCR-fragment was purified with QIAquick spin columns (QIAGEN) according to the standard protocol.

The plasmid was incubated with EcoRI and XhoI for 2 h at 37 °C.

10 µl	pGEX-4T1
0.8 µl	EcoRI
0.8 µl	XhoI
3 µl	3x buffer
15.4 µl	dd H <sub>2</sub> O

The restricted plasmid was purified via agarose gel electrophoresis and subsequently extracted with the S.N.A.P. UV-Free Gel Purification Kit.

### 12.2.1.3 Ligation

The PCR-fragment and plasmid were incubated for ligation with T4-DNA-ligase for 2 hours at room temperature.

0.4 µl	T4-DNA-ligase
4 µl	5x buffer
2 µl	pGEX-4T1 (~ 40 – 50 ng)
2.6 µl	CA II (PCR-fragment)
11 µl	dd H <sub>2</sub> O

### 12.2.1.4 Mutagenesis of CA II-DNA

In order to introduce the desired mutation into the CA II DNA-sequence the QuickChange II site-directed mutagenesis kit (Quiagen) is applied in a polymerase chain reaction (PCR). The PCR-samples are prepared as follows:



10 µl	5x Phusion buffer
5 µl	pGEX-4T1-CA II plasmid (~ 6 ng/µl ≈ 30 ng)
1.25 µl	Primer 1 (100 ng/µl)
1.25 µl	Primer 2 (100 ng/µl)
1 µl	dNTP-Mix
30.5 µl	dd H <sub>2</sub> O
1 µl	Phusion polymerase

**PCR-protocol:**

1.	30 sec	95 °C	
2.	30 sec	95 °C	
3.	60 sec	55 °C	x16
4.	6 min	68 °C	
5.	10 min	68 °C	
6.	∞	4 °C	

Subsequently, 1 µl of *DpnI* is added to the sample which specifically cuts methylated DNA. Thereby, the parental DNA is digested and the DNA containing the desired mutation can be isolated. The plasmid is transformed into XL2-Blue cells (12.2.1.5). Upon plasmid isolation from these cells the plasmid is send for sequencing to the MWG-company. Sequencing results are listed in the Appendix 13.2. For gene expression the plasmid were transformed into BL21Codon plus cells.

**12.2.1.5 Plasmid Transformation**

Plasmids are transformed by heat shock into the desired competent cells. Therefore, the XL2-Blue or BL21Codon plus cells are thawed for 30 min on ice and supplemented with 100 to 200 ng plasmid DNA. Upon 10 min incubation on ice the cells are exposed to a heat shock of 45 sec at 42 °C. Subsequently, the cells were chilled for 2 min on ice and incubated at 37 °C overnight on agar plates which are treated with the accordingly antibiotic (XL2-Blue – amp; BL21codon plus – amp, cam).

**12.2.1.6 Determination of DNA- and Protein-Concentration**

DNA- and protein concentrations are determined by UV-spectroscopy. The nullification is performed with the according buffer.

**DNA-samples:**

$$\text{DNA concentration} = A_{260} * 50 \frac{\mu\text{g}}{\text{ml}} * \text{dilution}$$

**Protein-samples<sup>5</sup>:**

$$\text{CAII} - \text{WT}: \frac{A_{280}}{1.69} = \text{concentration} \left[ \frac{\text{mg}}{\text{ml}} \right]$$

$$\text{CAII} - \text{W5C}: \frac{A_{280}}{1.52} = \text{concentration} \left[ \frac{\text{mg}}{\text{ml}} \right]$$

$$\text{CAII} - \text{W5C} - \text{H64C}: \frac{A_{280}}{1.52} = \text{concentration} \left[ \frac{\text{mg}}{\text{ml}} \right]$$

$$\text{CAII} - \text{H64C}: \frac{A_{280}}{1.70} = \text{concentration} \left[ \frac{\text{mg}}{\text{ml}} \right]$$

**12.2.2 X-ray Analysis****12.2.2.1 Data Collection**

Apo crystals or soaked crystals are transferred for about 10 seconds into a solution containing 25% glycerol as cryo-protectant and afterwards directly flash-frozen in liquid nitrogen.

Data sets were collected at one of our two in-house sources, at the BESSY in Berlin or at the SLS in Villigen, Switzerland (s. also Appendix 13.1).

At the first in-house source data are collected at cryo conditions (100K) with CuK $\alpha$  radiation ( $\lambda = 1.54178 \text{ \AA}$ ) using a Rigaku RU-H300R rotating-anode generator at 50 kV and 90 mA equipped with focusing mirrors and an R-AXIS IV (Xenocs mirrors) or R-AXIS IV ++ (MCS, USA) image-plate system.

At the second in-house source data are collected at cryo conditions (100K) with I $\mu$ S copper microfocus source ( $\lambda = 1.54178 \text{ \AA}$ ) using a MAR345 DTB image plate detector system.

---

<sup>5</sup> calculated: <http://www.expasy.ch/tools/protparam.html>

Additionally, data sets are collected at the BESSY-PSF Beamline 14.1 and 14.2 in Berlin at cryo conditions (100K) (13.1). Synchrotron radiation is used. A MAR CCD 165 mm or MAR CCD 225 mm detector is utilized for data collection.

Furthermore, data sets are collected at the SLS PXIII X06DA beamline in Villigen, Switzerland at cryo conditions (100K). Synchrotron radiation is used. A MAR CCD 225-mm detector is utilized for data collection.

The protein usually crystallizes in the monoclinic space group  $P2_1$  containing one monomer per asymmetric unit with Matthews coefficient of 2.1 to 2.2. Unit cell dimensions for all crystals are listed in the Appendix 13.1. Data processing and scaling were performed using the HKL2000 package. Data collection and processing statistics are given in the Appendix 13.1.

#### **12.2.2.2 Structure Determination and Refinement**

For all crystals coordinates of the CA II apo-structure (1OQ5) or of an in-house CA II-H64C structure are directly used for an initial rigid-body refinement, followed by simulated annealing, positional and  $B$ -factor refinement using the CNS program package (Brünger, Adams et al. 1998). The generally high resolution of the structures allows further refinement with the program SHELXL (Sheldrick, Schneider et al. 1997). Here, at least 20 cycles of conjugate gradient minimization are performed with default restraints on bond geometry and  $B$ -values: 5% of all data are used for  $R_{\text{free}}$  calculation. Amino acid side-chains are fitted to  $\sigma_A$ -weighted  $|F_o| - |F_c|$  and  $2 |F_o| - |F_c|$  electron density maps using COOT (Emsley and Cowtan 2004). In the electron density water, glycerol molecules and the ligand are located and subsequently included in further refinement cycles. In a final refinement, riding H-atoms are placed for the protein (not for the ligand) without using additional parameters. A final validation of the model was performed with PROCHECK (Laskowski, MacArthur et al. 1993).



## 13 Appendix

### 13.1 Crystal Data

Crystal Data	WT-Saccharin	H64C-Click-reactants	H64C-Click	WT-HK1939
<b>PDB-Code</b>	<b>2Q38</b>	<b>3KIG</b>	<b>3KNE</b>	<b>3M04</b>
<b>A. Data collection and processing</b>				
Beamline	in-house	BESSY 14.2	BESSY 14.2	BESSY 14.2
No. Crystals used	1	1	1	1
Wavelength (Å)	1.54178	0.91841	0.91841	0.91841
Space group	<i>P2<sub>1</sub></i>	<i>P2<sub>1</sub></i>	<i>P2<sub>1</sub></i>	<i>P2<sub>1</sub></i>
<b>Unit cell parameters</b>				
<i>a</i> (Å)	42.2	42.2	42.4	42.2
<i>b</i> (Å)	41.4	41.4	41.5	41.9
<i>c</i> (Å)	72.2	72.3	72.0	72.3
$\beta$ (deg.)	104.6	104.1	104.3	104.2
<b>B. Diffraction data</b>				
Resolution range (Å)*	20-1.95 (2.00-1.95)	25-1.39 (1.41-1.39)	25-1.35 (1.37-1.35)	25-1.40 (1.42-1.40)
Unique reflections*	16649 (734)	47302 (2315)	52457 (2456)	46605 (1803)
<b>Final R-values</b>				
R(I) <sub>sym</sub> (%)*	6.4 (21.0)	3.8 (24.8)	7.4 (33.5)	5.1 (20.1)
Completeness (%)*	93.2 (61.7)	96.7 (95.7)	98.2 (93.1)	96.2 (75.4)
Redundancy*	2.4 (1.8)	2.7 (2.5)	3.0 (2.3)	1.9 (1.6)
I/ $\sigma$ (I)*	13.5 (3.4)	19.9 (3.3)	13.8 (2.8)	12.1 (3.6)
<b>C. Refinement</b>				
Program used for refinement	SHELXL	SHELXL	SHELXL	SHELXL
Resolution range (Å)	20-1.95	10-1.39	10-1.35	10-1.40
Reflections used in refinement	16249	45897	50551	45355
R <sub>free</sub> (F <sub>o</sub> ; F <sub>o</sub> > 4 $\sigma$ F <sub>o</sub> )	26.0; 23.6	19.0; 18.0	18.3; 17.3	17.0; 16.0
R <sub>work</sub> (F <sub>o</sub> ; F <sub>o</sub> > 4 $\sigma$ F <sub>o</sub> )	18.2; 17.0	13.4; 12.8	12.8; 12.0	12.8; 12.1
<b>No. of atoms (non hydrogen)</b>				
Protein atoms	2058	2065	2060	2021
Water molecules	295	220	219	234
Ligand atoms	24	32	51	29
RMSD, angle (deg.)	1.8	2.1	2.2	2.3
RMSD, bond (Å)	0.006	0.010	0.011	0.011
<b>Ramachandran Plot</b>				
Most favoured regions (%)	87.0	89.3	88.9	89.3
Additionally allowed regions (%)	12.5	10.2	10.6	10.2
Generously allowed regions (%)	0.5	0.5	0.5	0.5
Disallowed regions (%)	0	0	0	0
<b>Mean B-factor (Å<sup>2</sup>)</b>				
Main-chain	13.9	13.2	13.1	11.4
Side-chain	14.5	17.8	17.6	15.1
Water molecules	23.9	26.6	27.9	25.2
Ligand 1 atoms	10.5 (Saccharin 1)	31.5 (azide12)	23.8	12.4
Ligand 2 atoms	30.5 (Saccharin 2)	28.3 (alkyne)	-	-
Zinc atom	29.4	19.4	7.0	5.5
Matthews' coefficient (Å <sup>3</sup> /Da)	2.2	2.2	2.1	2.2
Solvent content (%)	43.0	43.3	42.3	43.9

Crystal Data	WT-HK5053	WT-1,2-HOPTO	W5C-H64C open	W5C-H64C closed
<b>PDB-Code</b>	<b>3M14</b>	<b>3M1K</b>	<b>3M1Q</b>	<b>3M1W</b>
<b>A. Data collection and processing</b>				
Beamline	SLS X06DA	BESSY 14.2	in-house	Bessy 14.2
No. Crystals used	1	1	1	1
Wavelength (Å)	1.000	0.91841	1.54178	1.37848
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
<b>Unit cell parameters</b>				
$a$ (Å)	42.1	42.3	42.2	42.3
$b$ (Å)	41.4	41.4	41.4	41.5
$c$ (Å)	72.0	72.4	72.0	72.2
$\beta$ (deg.)	104.3	104.4	104.3	104.4
<b>B. Diffraction data</b>				
Resolution range (Å)*	25-1.38 (1.40-1.38)	25-1.35 (1.37-1.35)	25-1.69 (1.72-1.69)	25-1.38 (1.40-1.38)
Unique reflections*	49221 (2281)	50925 (2347)	26920 (1150)	48728 (1878)
<b>Final R-values</b>				
$R(I)_{\text{sym}}$ (%)*	6.4 (16.2)	5.1 (25.4)	3.9 (13.2)	4.0 (15.6)
Completeness (%)*	99.1 (91.6)	95.2 (88.8)	98.0 (85.0)	97.2 (75.8)
Redundancy*	3.6 (2.6)	2.2 (1.8)	2.4 (2.0)	2.7 (1.9)
$I/\sigma(I)$ *	13.7 (5.5)	15.1 (2.9)	22.1 (5.6)	25.6 (5.2)
<b>C. Refinement</b>				
Program used for refinement	SHELXL	SHELXL	SHELXL	SHELXL
Resolution range (Å)	10-1.38	10-1.35	10-1.69	10-1.38
Reflections used in refinement	48501	49131	26305	47961
$R_{\text{free}}$ ( $F_o$ ; $F_o > 4 \sigma F_o$ )	17.6; 17.2	18.6; 17.6	20.9; 20.0	17.5; 17.0
$R_{\text{work}}$ ( $F_o$ ; $F_o > 4 \sigma F_o$ )	12.7; 12.3	12.8; 12.1	16.0; 15.5	13.1; 12.8
<b>No. of atoms (non hydrogen)</b>				
Protein atoms	2115	2080	2041	2040
Water molecules	255	239	198	212
Ligand atoms	20	18	-	5
RMSD, angle (deg.)	2.2	2.2	2.0	2.4
RMSD, bond (Å)	0.012	0.012	0.009	0.013
<b>Ramachandran Plot</b>				
Most favoured regions (%)	87.5	88.4	87.0	88.3
Additionally allowed regions (%)	12.0	11.1	13.0	11.2
Generously allowed regions (%)	0.5	0.5	0	0.5
Disallowed regions (%)	0	0	0	0
<b>Mean B-factor (Å<sup>2</sup>)</b>				
Main-chain	13.8	14.1	17.0	15.0
Side-chain	17.7	18.2	21.3	18.9
Water molecules	28.2	27.1	28.1	28.9
Ligand 1 atoms	16.5	24.1	-	-
Ligand 2 atoms	-	-	-	-
Zinc atom	6.6	8.4	12.4	9.1
Matthews' coefficient (Å <sup>3</sup> /Da)	2.2	2.2	2.2	2.2
Solvent content (%)	42.8	43.4	43.0	43.4

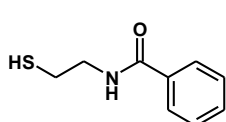
Crystal Data	WT-HK1375	WT-BO338	H64C-SB095	H64C
<b>PDB-Code</b>	<b>3M2X</b>	<b>3M2Y</b>	<b>3M2Z</b>	<b>3M5S</b>
<b>A. Data collection and processing</b>				
Beamline	in-house	Bessy 14.2	Bessy 14.2	Bessy 14.2
No. Crystals used	1	1	1	1
Wavelength (Å)	1.54178	0.91841	0.91841	0.91841
Space group	<i>P2<sub>1</sub></i>	<i>P2<sub>1</sub></i>	<i>P2<sub>1</sub></i>	<i>P2<sub>1</sub></i>
<b>Unit cell parameters</b>				
<i>a</i> (Å)	42.2	42.4	42.4	42.4
<i>b</i> (Å)	41.6	41.4	41.5	41.2
<i>c</i> (Å)	72.4	72.2	72.5	71.9
β (deg.)	104.3	104.6	104.0	104.3
<b>B. Diffraction data</b>				
Resolution range (Å)*	25-1.87 (1.90-1.87)	25-1.17 (1.19-1.17)	25-1.70 (1.73-1.70)	25-1.40 (1.42-1.40)
Unique reflections*	20128 (884)	79675 (3895)	25733 (1239)	46592 (2277)
<b>Final R-values</b>				
R(I) <sub>sym</sub> (%)*	5.7 (15.1)	4.8 (18.3)	9.2 (32.7)	5.6 (28.6)
Completeness (%)*	98.2 (84.5)	97.3 (95.4)	94.2 (92.7)	97.2 (94.1)
Redundancy*	2.8 (1.8)	3.0 (2.8)	2.3 (2.2)	2.4 (2.1)
I/σ(I)*	16.6 (4.0)	16.8 (5.7)	11.6 (3.1)	15.0 (2.9)
<b>C. Refinement</b>				
Program used for refinement	SHELXL	SHELXL	SHELXL	SHELXL
Resolution range (Å)	10-1.87	10-1.17	10-1.70	10-1.40
Reflections used in refinement	19475	77967	25630	44844
R <sub>free</sub> (F <sub>o</sub> ; F <sub>o</sub> > 4 σF <sub>o</sub> )	22.7; 21.3	16.1; 15.6	22.0; 21.0	17.9; 17.0
R <sub>work</sub> (F <sub>o</sub> ; F <sub>o</sub> > 4 σF <sub>o</sub> )	16.3; 15.3	12.7; 12.3	18.4; 16.7	13.7; 13.1
<b>No. of atoms (non hydrogen)</b>				
Protein atoms	2010	2073	2039	2045
Water molecules	180	266	187	177
Ligand atoms	34	29	16	4
RMSD, angle (deg.)	1.9	2.1	1.9	2.1
RMSD, bond (Å)	0.008	0.014	0.008	0.011
<b>Ramachandran Plot</b>				
Most favoured regions (%)	88.9	88.9	89.4	88.9
Additionally allowed regions (%)	10.2	10.6	10.1	10.6
Generously allowed regions (%)	0.9	0.5	0.5	0.5
Disallowed regions (%)	0	0	0	0
<b>Mean B-factor (Å<sup>2</sup>)</b>				
Main-chain	18.2	10.0	13.7	15.1
Side-chain	21.3	13.6	17.5	19.5
Water molecules	28.3	23.7	25.3	28.3
Ligand 1 atoms	18.8	12.3 (BO338 1)	25.9	-
Ligand 2 atoms		12.8 (BO338 2)		
Zinc atom	12.7	6.2	39.5	10.5
Matthews' coefficient (Å <sup>3</sup> /Da)	2.2	2.2	2.2	2.2
Solvent content (%)	43.6	43.3	43.8	42.9

Crystal Data	H64C-SB227
<b>PDB-Code</b>	<b>3M5T</b>
<b>A. Data collection and processing</b>	
Beamline	in-house
No. Crystals used	1
Wavelength (Å)	1.54178
Space group	$P2_1$
<b>Unit cell parameters</b>	
$a$ (Å)	42.8
$b$ (Å)	41.3
$c$ (Å)	72.1
$\beta$ (deg.)	104.3
<b>B. Diffraction data</b>	
Resolution range (Å)*	25-1.95 (1.98-1.95)
Unique reflections*	17724 (871)
<b>Final R-values</b>	
$R(I)_{\text{sym}}$ (%)*	8.8 (40.3)
Completeness (%)*	99.7 (99.8)
Redundancy*	3.4 (3.2)
$I/\sigma(I)$ *	14.3 (3.6)
<b>C. Refinement</b>	
Program used for refinement	SHELXL
Resolution range (Å)	10-1.95
Reflections used in refinement	16871
$R_{\text{free}}$ ( $F_o$ ; $F_o > 4 \sigma F_o$ )	25.0; 23.1
$R_{\text{work}}$ ( $F_o$ ; $F_o > 4 \sigma F_o$ )	17.6; 16.2
<b>No. of atoms (non hydrogen)</b>	
Protein atoms	2019
Water molecules	142
Ligand atoms	24
RMSD, angle (deg.)	1.8
RMSD, bond (Å)	0.007
<b>Ramachandran Plot</b>	
Most favoured regions (%)	88.0
Additionally allowed regions (%)	11.5
Generously allowed regions (%)	0.5
Disallowed regions (%)	0
<b>Mean B-factor (Å<sup>2</sup>)</b>	
Main-chain	21.3
Side-chain	25.0
Water molecules	30.2
Ligand 1 atoms	39.2
Ligand 2 atoms	-
Zinc atom	22.7
Matthews' coefficient (Å <sup>3</sup> /Da)	2.2
Solvent content (%)	43.7

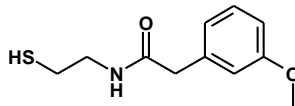


## 13.2 Tethering Compound Library

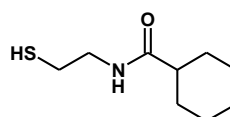
**Series 1** Initially synthesized compounds from in house carboxylic acids. The compounds were synthesized symmetrically (the molecular weight is shown in Da).



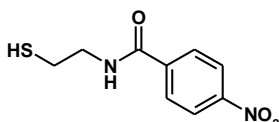
SB095 - MW: 181.26



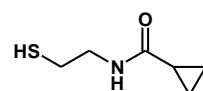
SB103 - MW: 225.31



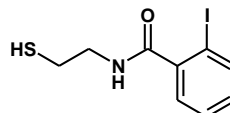
SB123 - MW: 187.31



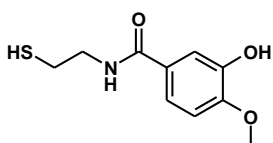
SB096 - MW: 226.26



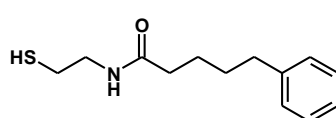
SB104 - MW: 145.22



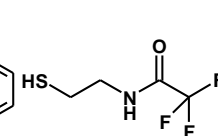
SB124 - MW: 307.15



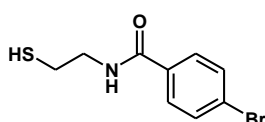
SB097 - MW: 227.28



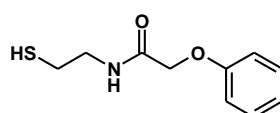
SB113 - MW: 237.37



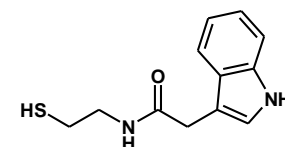
SB106 - MW: 173.16



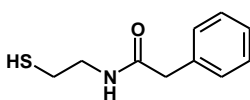
SB099 - MW: 260.15



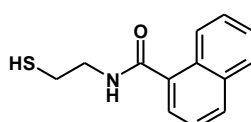
SB114 - MW: 211.28



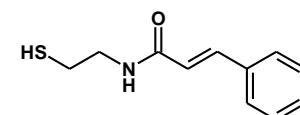
SB107 - MW: 234.32



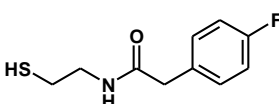
SB100 - MW: 195.29



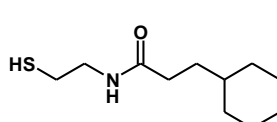
SB115 - MW: 231.32



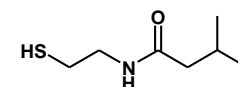
SB109 - MW: 207.30



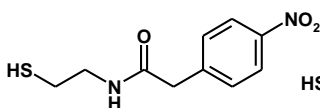
SB101 - MW: 213.28



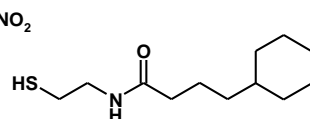
SB119 - MW: 215.36



SB127 - MW: 161.27

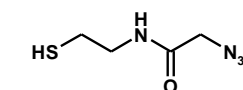
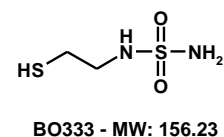
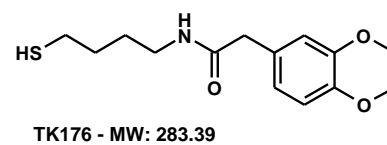
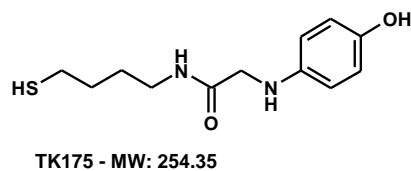
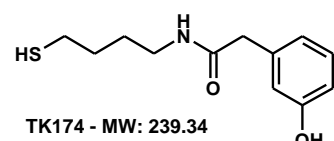
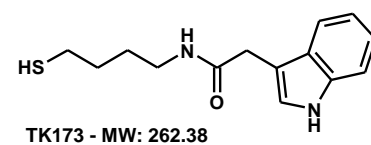
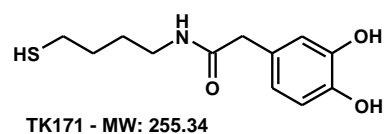
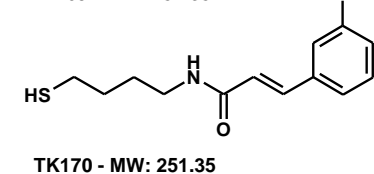
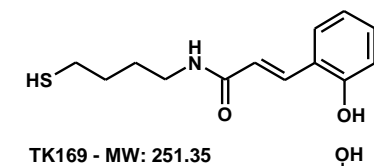
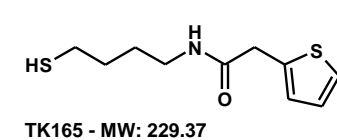
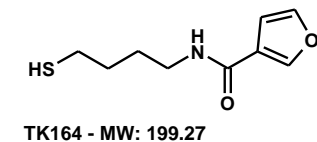
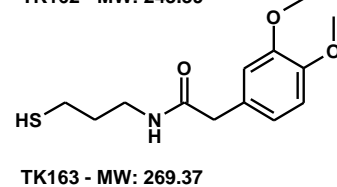
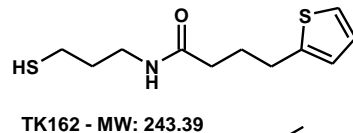
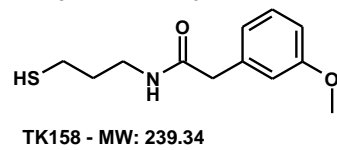
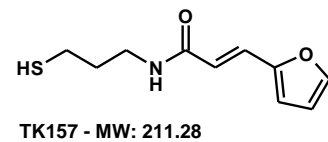
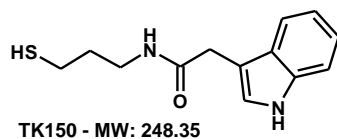
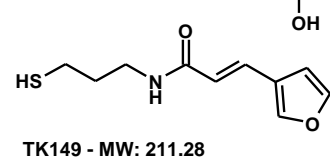
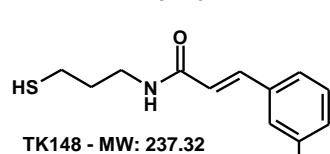
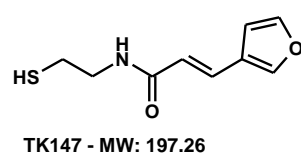
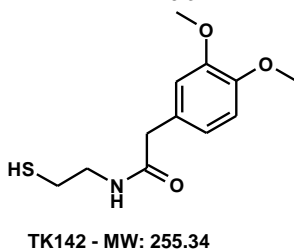
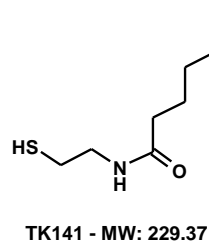
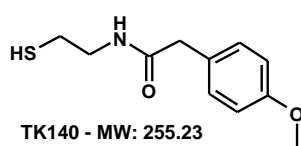
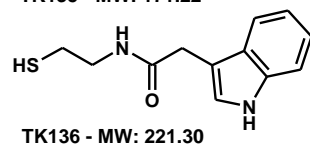
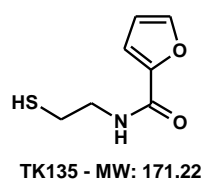


SB102 - MW: 240.28

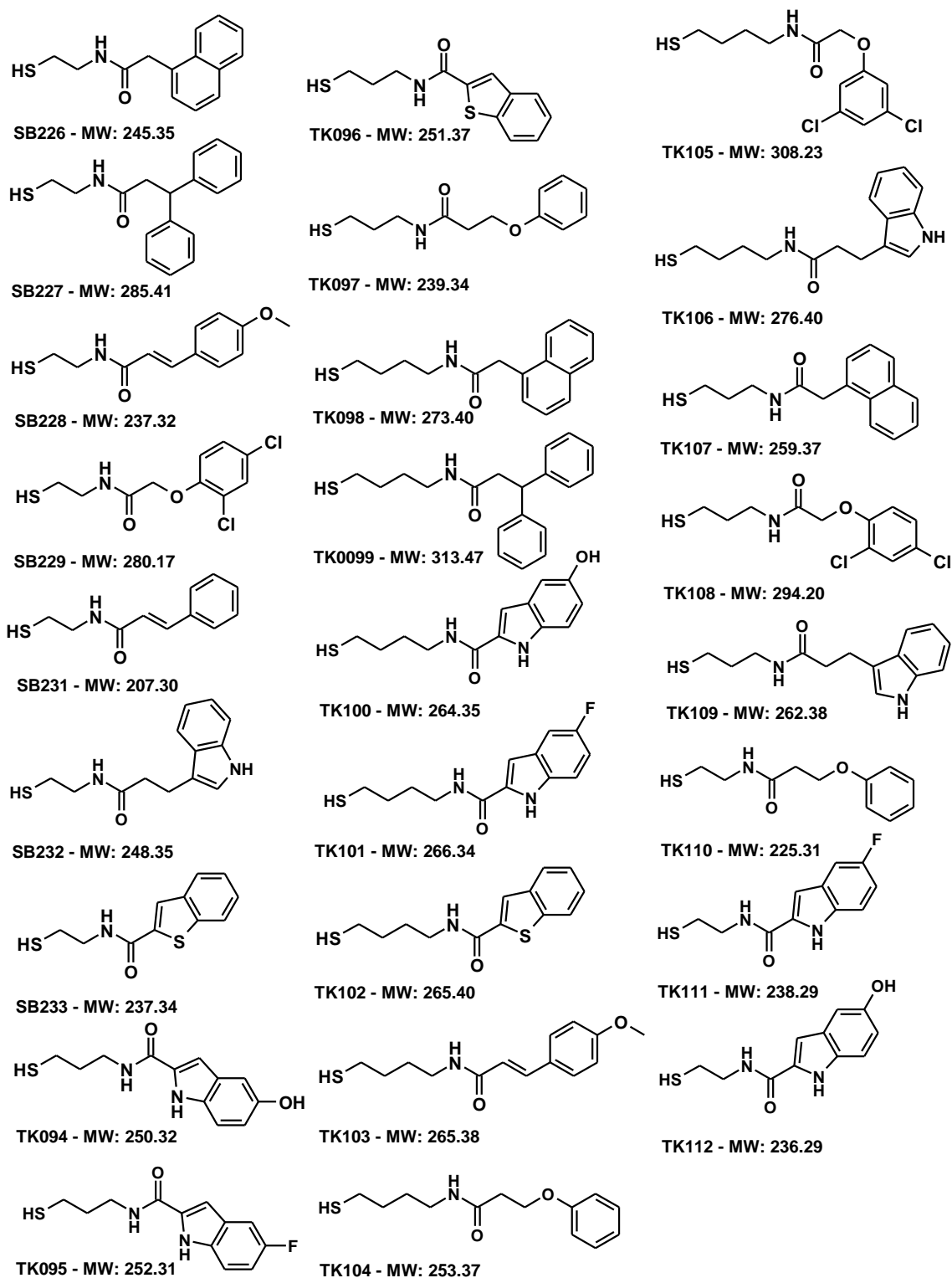


SB120 - MW: 229.39

**Series 2 Synthesized compounds from CA II-H64C docking.** These fragments were synthesized according to the skeletons shown in Figure 7.3 (the molecular weight is shown in Da). An amine is connected to a linker with the equivalent length of the corresponding fragment.



**Series 3 Synthesized compounds from TGT-V233G docking.** These fragments were synthesized according to the skeletons shown in Figure 7.3 (the molecular weight is shown in Da). An amine is connected to a linker with the equivalent length of the corresponding fragment.



### 13.3 Sequencing of CA II and CA II-Mutants

#### CA II

EMBOSS_001	1	MSHHWGYGKHNGPEHWHKDFPIAKGERQSPVDIDHTHTAKYDPSLKPLSVS 	50
EMBOSS_001	1	MSHHWGYGKHNGPEHWHKDFPIAKGERQSPVDIDHTHTAKYDPSLKPLSVS 	50
EMBOSS_001	51	YDQATSLRILNNGHAFNVEFDDSDKAVLKGKGLDGT YRLIQFHFHWGSL 	100
EMBOSS_001	51	YDQATSLRILNNGHAFNVEFDDSDKAVLKGKGLDGT YRLIQFHFHWGSL 	100
EMBOSS_001	101	DGQGSEHTVDKKKYAAELHLVHWNTKYDFGKAVQQPDGLAVLGI FLKVG 	150
EMBOSS_001	101	DGQGSEHTVDKKKYAAELHLVHWNTKYDFGKAVQQPDGLAVLGI FLKVG 	150
EMBOSS_001	151	SAKPLQKVVVDLDSIKTKGKSADFTNFDPRGLLPESLDYWTYPGSLTTP 	200
EMBOSS_001	151	SAKPLQKVVVDLDSIKTKGKSADFTNFDPRGLLPESLDYWTYPGSLTTP 	200
EMBOSS_001	201	PLLECVTWIVLKEPISVSSEQVLKFRKLNFNNGEPEELMVDNWRPAQPL 	250
EMBOSS_001	201	PLLECVTWIVLKEPISVSSEQVLKFRKLNFNNGEPEELMVDNWRPAQPL 	250
EMBOSS_001	251	KNRQIKASFK       260 	
EMBOSS_001	251	KNRQIKASFK       260 	

#### CA II-W5C

EMBOSS_001	1	MSHHWGYGKHNGPEHWHKDFPIAKGERQSPVDIDHTHTAKYDPSLKPLSVS 	50
EMBOSS_001	1	MSHHWGYGKHNGPEHWHKDFPIAKGERQSPVDIDHTHTAKYDPSLKPLSVS 	50
EMBOSS_001	51	YDQATSLRILNNGHAFNVEFDDSDKAVLKGKGLDGT YRLIQFHFHWGSL 	100
EMBOSS_001	51	YDQATSLRILNNGHAFNVEFDDSDKAVLKGKGLDGT YRLIQFHFHWGSL 	100
EMBOSS_001	101	DGQGSEHTVDKKKYAAELHLVHWNTKYDFGKAVQQPDGLAVLGI FLKVG 	150
EMBOSS_001	101	DGQGSEHTVDKKKYAAELHLVHWNTKYDFGKAVQQPDGLAVLGI FLKVG 	150
EMBOSS_001	151	SAKPLQKVVVDLDSIKTKGKSADFTNFDPRGLLPESLDYWTYPGSLTTP 	200
EMBOSS_001	151	SAKPLQKVVVDLDSIKTKGKSADFTNFDPRGLLPESLDYWTYPGSLTTP 	200
EMBOSS_001	201	PLLECVTWIVLKEPISVSSEQVLKFRKLNFNNGEPEELMVDNWRPAQPL 	250
EMBOSS_001	201	PLLECVTWIVLKEPISVSSEQVLKFRKLNFNNGEPEELMVDNWRPAQPL 	250
EMBOSS_001	251	KNRQIKASFK       260 	
EMBOSS_001	251	KNRQIKASFK       260 	

#### CA II-W5C-H64C

EMBOSS_001	1	MSHHWGYGKHNGPEHWHKDFPIAKGERQSPVDIDHTHTAKYDPSLKPLSVS 	50
EMBOSS_001	1	MSHHWGYGKHNGPEHWHKDFPIAKGERQSPVDIDHTHTAKYDPSLKPLSVS 	50
EMBOSS_001	51	YDQATSLRILNNGHAFNVEFDDSDKAVLKGKGLDGT YRLIQFHFHWGSL 	100
EMBOSS_001	51	YDQATSLRILNNGCAFNVEFDDSDKAVLKGKGLDGT YRLIQFHFHWGSL 	100
EMBOSS_001	101	DGQGSEHTVDKKKYAAELHLVHWNTKYDFGKAVQQPDGLAVLGI FLKVG 	150
EMBOSS_001	101	DGQGSEHTVDKKKYAAELHLVHWNTKYDFGKAVQQPDGLAVLGI FLKVG 	150
EMBOSS_001	151	SAKPLQKVVVDLDSIKTKGKSADFTNFDPRGLLPESLDYWTYPGSLTTP 	200
EMBOSS_001	151	SAKPLQKVVVDLDSIKTKGKSADFTNFDPRGLLPESLDYWTYPGSLTTP 	200
EMBOSS_001	201	PLLECVTWIVLKEPISVSSEQVLKFRKLNFNNGEPEELMVDNWRPAQPL 	250
EMBOSS_001	201	PLLECVTWIVLKEPISVSSEQVLKFRKLNFNNGEPEELMVDNWRPAQPL 	250
EMBOSS_001	251	KNRQIKASFK       260 	
EMBOSS_001	251	KNRQIKASFK       260 	

### CA II-H64C

```

EMBOSS_001      1 MSHHWGYGKHNNGPEHWHKDFPIAKGERQSPVDIDHTAKYDPSLKPLSVS   50
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
EMBOSS_001      1 MSHHWGYGKHNNGPEHWHKDFPIAKGERQSPVDIDHTAKYDPSLKPLSVS   50
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
EMBOSS_001      51 YDQATSLRILNNGHAFNVEFDDSQDKAVLKGGLDGTYRLIQFHFHWGSL   100
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
EMBOSS_001      51 YDQATSLRILNNGCAFNVEFDDSQDKAVLKGGLDGTYRLIQFHFHWGSL   100
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
EMBOSS_001     101 DGQGEHTVDKKKYAAELHLVHWNTKYDGFQKAVQPDGLAVLGIPLKVG   150
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
EMBOSS_001     101 DGQGEHTVDKKKYAAELHLVHWNTKYDGFQKAVQPDGLAVLGIPLKVG   150
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
EMBOSS_001     151 SAKPGLQKVVDVLDSEIKTKGKSADFTNFDPRGLLPESLDYWTYPGSLTTP   200
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
EMBOSS_001     151 SAKPGLQKVVDVLDSEIKTKGKSADFTNFDPRGLLPESLDYWTYPGSLTTP   200
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
EMBOSS_001     201 PLLECVTWIVLKEPISVSSEQVLKFRKLNFNGEGEPEELMVDNWRPAQPL   250
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
EMBOSS_001     201 PLLECVTWIVLKEPISVSSEQVLKFRKLNFNGEGEPEELMVDNWRPAQPL   250
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
EMBOSS_001     251 KNRQIKASFK      260
  | | | | | | | |
EMBOSS_001     251 KNRQIKASFK      260
  | | | | | | | |
  
```



## 14 Literature

- Abad-Zapatero, C. and J. T. Metz (2005). "Ligand efficiency indices as guideposts for drug discovery." Drug Discovery Today **10** (7): 464-469.
- Abbate, F., C. T. Supuran, A. Scozzafava, P. Orioli, M. T. Stubbs and G. Klebe (2002). "Nonaromatic Sulfonamide Group as an Ideal Anchor for Potent Human Carbonic Anhydrase Inhibitors: Role of Hydrogen-Bonding Networks in Ligand Binding and Drug Design." Journal of Medicinal Chemistry **45** (17): 3583-3587.
- Alberty, R. A. and G. G. Hammes (1958). "Application of the Theory of Diffusion-controlled Reactions to Enzyme Kinetics." The Journal of Physical Chemistry **62** (2): 154-159.
- Alexander, R. S., S. K. Nair and D. W. Christianson (1991). "Engineering the hydrophobic pocket of carbonic anhydrase II." Biochemistry **30** (46): 11064-11072.
- Barrese, A. A., C. Genis, S. Z. Fisher, J. N. Orwenyo, M. T. Kumara, S. K. Dutta, E. Phillips, J. J. Kiddle, C. Tu, D. N. Silverman, L. Govindasamy, M. Agbandje-McKenna, R. McKenna and B. C. Tripp (2008). "Inhibition of Carbonic Anhydrase II by Thioxolone: A Mechanistic and Structural Study" Biochemistry **47** (10): 3174-3184.
- Bayram, E., M. Senturk, O. Irfan Kufrevioglu and C. T. Supuran (2008). "In vitro inhibition of salicylic acid derivatives on human cytosolic carbonic anhydrase isozymes I and II." Bioorganic & Medicinal Chemistry **16** (20): 9101-9105.
- Behravan, G., B.-H. Jonsson and S. Lindskog (1990). "Fine tuning of the catalytic properties of carbonic anhydrase." European Journal of Biochemistry **190** (2): 351-357.
- Berkel, S. S. v., A. J. Dirks, S. A. Meeuwissen, D. L. L. Pingen, O. C. Boerman, P. Laverman, F. L. v. Delft, J. J. L. M. Cornelissen and F. P. J. T. Rutjes (2008). "Application of Metal-Free Triazole Formation in the Synthesis of Cyclic RGD-DTPA Conjugates." ChemBioChem **9** (11): 1805-1815.
- Bock, V. D., H. Hiemstra and J. H. v. Maarseveen (2006). "CuI-Catalyzed Alkyne-Azide "Click" Cycloadditions from a Mechanistic and Synthetic Perspective." European Journal of Organic Chemistry **2006** (1): 51-68.
- Braisted, A. C., J. D. Oslob, W. L. Delano, J. Hyde and R. S. McDowell (2003). "Discovery of a potent small molecule IL-2 inhibitor through fragment assembly." J. Am. Chem. Soc. **125**: 3714-3715.

- Briganti, F., S. Mangani, P. Orioli, A. Scozzafava, G. Vernaglione and C. T. Supuran (1997). "Carbonic Anhydrase Activators: X-ray Crystallographic and Spectroscopic Investigations for the Interaction of Isozymes I and II with Histamine." Biochemistry **36** (34): 10384-10392.
- Briganti, F., R. Pierattelli, A. Scozzafava and C. T. Supuran (1996). "Carbonic anhydrase inhibitors. Part 37. Novel classes of isozyme I and II inhibitors and their mechanism of action. Kinetic and spectroscopic investigations on native and cobalt-substituted enzymes." European Journal of Medicinal Chemistry **31** (12): 1001-1010.
- Brünger, A. T., P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson and G. L. Warren (1998). "Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination." Acta Crystallographica Section D **54** (5): 905-921.
- Carrion, E., J. H. Hertzog, M. D. Medlock, G. J. Hauser and H. J. Dalton (2001). "Use of acetazolamide to decrease cerebrospinal fluid production in chronically ventilated patients with ventriculopleural shunts." Archives of Disease in Childhood **84** (1): 68-71.
- Casini, A., J. Antel, F. Abbate, A. Scozzafava, S. David, H. Waldeck, S. Schäfer and C. T. Supuran (2003). "Carbonic anhydrase inhibitors: SAR and X-ray crystallographic study for the interaction of sugar sulfamates/sulfamides with isozymes I, II and IV." Bioorganic & Medicinal Chemistry Letters **13** (5): 841-845.
- Casini, A., J.-Y. Winum, J.-L. Montero, A. Scozzafava and C. T. Supuran (2003). "Carbonic anhydrase inhibitors: inhibition of cytosolic isozymes I and II with sulfamide derivatives." Bioorganic & Medicinal Chemistry Letters **13** (5): 837-840.
- Charles, E. S., N. Antonio and M. Pierre (2006). "Evidence by signal peptide trap technology for the expression of carbonic anhydrase 6 in rat incisor enamel organs." European Journal Of Oral Sciences **114** (1): 147-153.
- Chegwidden, W. R. and N. D. Carter (2000). "The Carbonic Anhydrases - New Horizons." Birkhäuser Verlag: 13-28.
- Christianson, D. W. and C. A. Fierke (1996). "Carbonic Anhydrase: Evolution of the Zinc Binding Site by Nature and by Design." Accounts of Chemical Research **29** (7): 331-339.
- Dodgson, S. J., R. P. Shank and B. E. Maryanoff (2000). "Topiramate as an Inhibitor of Carbonic Anhydrase Isoenzymes." Epilepsia **41** (5): 35-39.



- Domsic, J. F., B. S. Avvaru, C. U. Kim, S. M. Gruner, M. Agbandje-McKenna, D. N. Silverman and R. McKenna (2008). "Entrapment of Carbon Dioxide in the Active Site of Carbonic Anhydrase II." Journal of Biological Chemistry **283** (45): 30766-30771.
- Elling, C. E., S. M. Nielsen and T. W. Schwartz (1995). "Conversion of antagonist-binding site to metal-ion site in the tachykinin." Nature **374** (6517): 74-77.
- Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallographica Section D **60** (12 Part 1): 2126-2132.
- Engstrand, C., C. Forsman, Z. Liang and S. Lindskog (1992). "Proton transfer roles of lysine 64 and glutamic acid 64 replacing histidine 64 in the active site of human carbonic anhydrase II." Biochim-Biophys-Acta **1122**: 312-326.
- Ericsson, U. B., B. M. Hallberg, G. T. DeTitta, N. Dekker and P. Nordlund (2006). "Thermofluor-based high-throughput stability optimization of proteins for structural studies." Analytical Biochemistry **357** (2): 289-298.
- Eriksson, A. E., T. A. Jones and L. Anders (1988). "Refined structure of human carbonic anhydrase II at 2.0 Å resolution." Proteins: Structure, Function, and Genetics **4** (4): 274-282.
- Eriksson, A. E., P. M. Kylsten, T. A. Jones and L. Anders (1988). "Crystallographic studies of inhibitor binding sites in human carbonic anhydrase II: A pentacoordinated binding of the SCN<sup>-</sup> ion to the zinc at high pH." Proteins: Structure, Function, and Genetics **4** (4): 283-293.
- Erlanson, D. A., A. C. Braisted, D. R. Raphael, M. Randal, R. M. Stroud, E. M. Gordon and J. A. Wells (2000). "Site-directed ligand discovery." Proceedings of the National Academy of Sciences of the United States of America **97** (17): 9367-9372.
- Erlanson, D. A., J. A. Wells and A. C. Braisted (2004). "TETHERING: Fragment-Based Drug Discovery." Annual Review of Biophysics and Biomolecular Structure **33** (1): 199-223.
- Fahlberg, C. and R. Ira (1879). "Ueber die Oxydation des Orthotoluolsulfamids." Berichte der deutschen chemischen Gesellschaft **12** (1): 469-473.
- Fiore, A. D., G. D. Simone, V. Menchise, C. Pedone, A. Casini, A. Scozzafava and C. T. Supuran (2005). "Carbonic anhydrase inhibitors: X-ray crystal structure of a benzenesulfonamide strong CA II and CA IX inhibitor bearing a pentafluorophenylaminothioureido tail in complex with isozyme II." Bioorganic & Medicinal Chemistry Letters **15** (7): 1937-1942.

- Gerlach, C., M. Münzel, B. Baum, H.-D. Gerber, T. Craan, W. E. Diederich and G. Klebe (2007). "KNOBLE: Ein wissenschaftlicher Ansatz zu Design und Synthese einfach herstellbarer Liganden als Ausgangspunkt für eine strukturbasierte Leitstrukturoptimierung." Angewandte Chemie **119** (47): 9265-9269.
- Hajduk, P. J. (2006). "SAR by NMR: Putting the Pieces Together." Molecular Interventions **6** (5): 266-272.
- Hakansson, K., M. Carlsson, L. A. Svensson and A. Liljas (1992). "Structure of native and apo carbonic anhydrase II and structure of some of its anion-ligand complexes." Journal of Molecular Biology **227** (4): 1192-204.
- Hakansson, K., A. Wehnert and A. Liljas (1994). "X-ray analysis of metal-substituted human carbonic anhydrase II derivatives." Acta Crystallographica Section D **50** (1): 93-100.
- Hein, C. D., X.-M. Liu and D. Wang (2008). "Click Chemistry, A Powerful Tool for Pharmaceutical Sciences." Pharmaceutical Research **25** (10): 2216-2230.
- Hellinga, H. W., J. P. Caradonna and F. M. Richards (1991). "Construction of new ligand binding sites in proteins of known structure II. Grafting of a buried transition metal binding site into Escherichia coli thioredoxin." Journal of Molecular Biology **222** (3): 787-803.
- Hellinga, H. W. and F. M. Richards (1991). "Construction of new ligand binding sites in proteins of known structure : I. Computer-aided modeling of sites with pre-defined geometry." Journal of Molecular Biology **222** (3): 763-785.
- Himo, F., T. Lovell, R. Hilgraf, V. V. Rostovtsev, L. Noodleman, K. B. Sharpless and V. V. Fokin (2004). "Copper(I)-Catalyzed Synthesis of Azoles. DFT Study Predicts Unprecedented Reactivity and Intermediates." Journal of the American Chemical Society **127** (1): 210-216.
- Hopkins, A. L., C. R. Groom and A. Alex (2004). "Ligand efficiency: a useful metric for lead selection." Drug Discovery Today **9** (10): 430-431.
- Huisgen, R., G. Szeimies and L. Möbius (1967). "1,3-Dipolare Cycloadditionen, XXXII. Kinetik der Additionen organischer Azide an CC-Mehrfachbindungen." Chemische Berichte **100** (8): 2494-2507.
- Ilies, M. A., D. Vullo, J. Pastorek, A. Scozzafava, M. Ilies, M. T. Caproiu, S. Pastorekova and C. T. Supuran (2003). "Carbonic Anhydrase Inhibitors. Inhibition of Tumor-Associated Isozyme IX by Halogenosulfanilamide and Halogenophenylaminobenzolamide Derivatives." Journal of Medicinal Chemistry **46** (11): 2187-2196.

- Innocenti, A., M. Hilvo, A. Scozzafava, S. Parkkila and C. T. Supuran (2008). "Carbonic anhydrase inhibitors: Inhibition of the new membrane-associated isoform XV with phenols." Bioorganic & Medicinal Chemistry Letters **18** (12): 3593-3596.
- Innocenti, A., D. Vullo, A. Scozzafava and C. T. Supuran (2008a). "Carbonic anhydrase inhibitors: Inhibition of mammalian isoforms I-XIV with a series of substituted phenols including paracetamol and salicylic acid." Bioorganic & Medicinal Chemistry **16** (15): 7424-7428.
- Innocenti, A., D. Vullo, A. Scozzafava and C. T. Supuran (2008b). "Carbonic anhydrase inhibitors: Interactions of phenols with the 12 catalytically active mammalian isoforms (CA I-XIV)." Bioorganic & Medicinal Chemistry Letters **18** (5): 1583-1587.
- Jackman, J. E., K. M. Merz and C. A. Fierke (1996). "Disruption of the Active Site Solvent Network in Carbonic Anhydrase II Decreases the Efficiency of Proton Transfer " Biochemistry **35** (51): 16421-16428.
- Jacobsen, J. A., J. L. Major Jourden, M. T. Miller and S. M. Cohen "To bind zinc or not to bind zinc: An examination of innovative approaches to improved metalloproteinase inhibition." Biochimica et Biophysica Acta (BBA) - Molecular Cell Research **1803** (1): 72-94.
- Jacoby, E., A. Schuffenhauer, M. Popov, K. Azzaoui, B. Havill, U. Schopfer, C. Engeloch, J. Stanek, P. Acklin, P. Rigollier, F. Stoll, G. Koch, P. Meier, D. Orain, R. Giger, J. Hinrichs, K. Malagu, J. Zimmermann and H.-J. Roth (2005). "Key Aspects of the Novartis Compound Collection Enhancement Project for the Compilation of a Comprehensive Chemogenomics Drug Discovery Screening Collection." Current Topics in Medicinal Chemistry **5**: 397-411.
- Jordan, I. K., D. A. Natale and M. Y. Galperin (2000). "Copper chaperones in bacteria: association with copper-transporting ATPases." Trends in Biochemical Sciences **25** (10): 480-481.
- Karlsson, B. G., M. Nordling, T. Pascher, L.-C. Tsai, L. Sjolín and L. G. Lundberg (1991). "Cassette mutagenesis of Met121 in azurin from *Pseudomonas aeruginosa*." Protein Eng. **4** (3): 343-349.
- Kerry, S. S. and G. F. James (2000). "Prokaryotic carbonic anhydrases." FEMS Microbiology Reviews **24** (4): 335-366.
- Khalifah, R. G. (1971). "The Carbon Dioxide Hydration Activity of Carbonic Anhydrase." Journal of Biological Chemistry **246** (8): 2561-2573.

- Kiefer, L. L. and C. A. Fierke (1994). "Functional Characterization of Human Carbonic Anhydrase II Variants with Altered Zinc Binding Sites." Biochemistry **33** (51): 15233-15240.
- Kiefer, L. L., J. F. Krebs, S. A. Paterno and C. A. Fierke (1993). "Engineering a cysteine ligand into the zinc binding site of human carbonic anhydrase II." Biochemistry **32** (38): 9896-9900.
- Kim, C.-Y., J. S. Chang, J. B. Doyon, T. T. Baird, C. A. Fierke, A. Jain and D. W. Christianson (2000). "Contribution of Fluorine to Protein-Ligand Affinity in the Binding of Fluoroaromatic Inhibitors to Carbonic Anhydrase II." Journal of the American Chemical Society **122** (49): 12125-12134.
- Kimoto, M., M. Kishino, Y. Yura and Y. Ogawa (2006). "A role of salivary carbonic anhydrase VI in dental plaque." Archives of Oral Biology **51** (2): 117-122.
- Klebe, G. (2009). "Wirkstoffdesign: Entwurf und Wirkung von Arzneistoffen." Spektrum Akademischer Verlag **2. Auflage**: 97-98.
- Klemba, M., K. H. Gardner, S. Marino, N. D. Clarke and L. Regan (1995). "Novel metal-binding proteins by design." Nature Structural & Molecular Biology **2** (5): 368-373.
- Klomp, L. W. J., S.-J. Lin, D. S. Yuan, R. D. Klausner, V. C. Culotta and J. D. Gitlin (1997). "Identification and Functional Expression of HAH1, a Novel Human Gene Involved in Copper Homeostasis." Journal of Biological Chemistry **272** (14): 9221-9226.
- Köhler, K., A. Hillebrecht, J. Schulze Wischeler, A. Innocenti, A. Heine, C. T. Supuran and G. Klebe (2007). "Saccharin Inhibits Carbonic Anhydrases: Possible Explanation for its Unpleasant Metallic Aftertaste." Angewandte Chemie International Edition **46** (40): 7697-7699.
- Kolb, H. C., M. G. Finn and K. B. Sharpless (2001). "Click Chemistry: Diverse Chemical Function from a Few Good Reactions." Angewandte Chemie International Edition **40** (11): 2004-2021.
- Krasinski, A., Z. Radic, R. Manetsch, J. Raushel, P. Taylor, K. B. Sharpless and H. C. Kolb (2005). "In Situ Selection of Lead Compounds by Click Chemistry: Target-Guided Optimization of Acetylcholinesterase Inhibitors." Journal of the American Chemical Society **127** (18): 6686-6692.
- Krebs, H. A. (1948). "Inhibition of carbonic anhydrase by sulphonamides." Biochemical Journal **43** (4): 525-528.

- Krebs, J. F. and C. A. Fierke (1993). "Determinants of catalytic activity and stability of carbonic anhydrase II as revealed by random mutagenesis." Journal of Biological Chemistry **268** (2): 948-954.
- Laemmli, U. K. (1970). "Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4." Nature **227** (5259): 680-685.
- Larson, E. B., R. C. Roach, R. B. Schoene and T. F. Hornbein (1982). "Acute Mountain Sickness and Acetazolamide: Clinical Efficacy and Effect on Ventilation." The Journal of the American Medical Association **248** (3): 328-332.
- Laskowski, R. A., M. W. MacArthur, D. S. Moss and J. M. Thornton (1993). "PROCHECK: a program to check the stereochemical quality of protein structures." Journal of Applied Crystallography **26** (2): 283-291.
- Leach, A. R. and M. M. Hann (2000). "The in silico world of virtual libraries." Drug Discovery Today **5** (8): 326-336.
- Lewis, J. A., D. T. Puerta and S. M. Cohen (2003). "Metal Complexes of the trans-Influencing Ligand Thiomaltol." Inorganic Chemistry **42** (23): 7455-7459.
- Lewis, J. A., B. L. Tran, D. T. Puerta, E. M. Rumberger, D. N. Hendrickson and S. M. Cohen (2005). "Synthesis, structure and spectroscopy of new thiopyrone and hydroxypyridinethione transition-metal complexes." Dalton Transactions **2005**: 2588-2596.
- Lewis, W. G., L. G. Green, F. Grynszpan, Z. Radic, P. R. Carlier, P. Taylor, M. G. Finn and K. B. Sharpless (2002). "Click Chemistry In Situ: Acetylcholinesterase as a Reaction Vessel for the Selective Assembly of a Femtomolar Inhibitor from an Array of Building Blocks." Angewandte Chemie International Edition **41** (6): 1053-1057.
- Li, H., S. P. Webb, J. Ivancic and J. H. Jensen (2004). "Determinants of the Relative Reduction Potentials of Type-1 Copper Sites in Proteins." Journal of the American Chemical Society **126** (25): 8010-8019.
- Liang, Z., Y. Xue, G. Behravan, B.-H. Jonsson and S. Lindskog (1993). "Importance of the conserved active-site residues Try7, Glu106 and Thr199 for the catalytic function of human carbonic anhydrase II." European Journal of Biochemistry **211** (3): 821-827.
- Liljas, A., K. K. Kannan, P. C. Bergstén, I. Waara, K. Fridborg, B. Strandberg, U. Carlbom, L. Järup, S. Lövgren and M. P. . (1972). "Crystal structure of human carbonic anhydrase C." Nature New Biology. **235**: 131-137.
- Lindskog, S. (1997). "Structure and mechanism of carbonic anhydrase." Pharmacology & Therapeutics **74** (1): 1-20.

- Lyne, P. D. (2005). Rational and Structure-Based Approaches to Lead Discovery: American Association for cancer research 96<sup>th</sup> Annual Meeting: 130-133.
- Maestrelli, F., P. Mura, A. Casini, F. Mincione, A. Scozzafava and C. T. Supuran (2002). "Cyclodextrin complexes of sulfonamide carbonic anhydrase inhibitors as long-lasting topically acting antiglaucoma agents." Journal of Pharmaceutical Sciences **91** (10): 2211-2219.
- Mangani, S. and K. H. Kansson (1992). "Crystallographic studies of the binding of protonated and unprotonated inhibitors to carbonic anhydrase using hydrogen sulphide and nitrate anions." European Journal of Biochemistry **210** (3): 867-871.
- Maren, T. H. (1967). "Carbonic anhydrase: chemistry, physiology, and inhibition." Physiological Reviews **47** (4): 595-781.
- Maren, T. H. (1988). "The kinetics of  $\text{HCO}_3^-$  synthesis related to fluid secretion, pH control, and  $\text{CO}_2$  elimination." Annual Reviews of Physiology **50**: 695-717.
- Maren, T. H. (1995). "The Development of Topical Carbonic Anhydrase Inhibitors." Journal of Glaucoma **4** (1): 49-62.
- Maresca, A., C. Temperini, L. Pochet, B. Masereel, A. Scozzafava and C. T. Supuran (2009a). "Deciphering the Mechanism of Carbonic Anhydrase Inhibition with Coumarins and Thiocoumarins." Journal of Medicinal Chemistry **53** (1): 335-344.
- Maresca, A., C. Temperini, H. Vu, N. B. Pham, S.-A. Poulsen, A. Scozzafava, R. J. Quinn and C. T. Supuran (2009b). "Non-Zinc Mediated Inhibition of Carbonic Anhydrases: Coumarins Are a New Class of Suicide Inhibitors." Journal of the American Chemical Society **131** (8): 3057-3062.
- Masereel, B., S. Rolin, F. Abbate, A. Scozzafava and C. T. Supuran (2001). "Carbonic Anhydrase Inhibitors: Anticonvulsant Sulfonamides Incorporating Valproyl and Other Lipophilic Moieties." Journal of Medicinal Chemistry **45** (2): 312-320.
- McGovern, S. L., E. Caselli, N. Grigorieff and B. K. Shoichet (2002). "A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening." Journal of Medicinal Chemistry **45**: 1712-1722.
- Menchise, V., G. De Simone, V. Alterio, A. Di Fiore, C. Pedone, A. Scozzafava and C. T. Supuran (2005). "Carbonic Anhydrase Inhibitors: Stacking with Phe131 Determines Active Site Binding Region of Inhibitors As Exemplified by the X-ray Crystal Structure of a Membrane-Impermeant Antitumor Sulfonamide

- Complexed with Isozyme II." Journal of Medicinal Chemistry **48** (18): 5721-5727.
- Merz, K. M. (1991). "Determination of pKas of ionizable groups in proteins: the pKa of Glu 7 and 35 in hen eggs white lysozyme and Glu 106 in human carbonic anhydrase II." Journal of the American Chemical Society **113** (9): 3572-3575.
- Merz, K. M. and L. Banci (1997). "Binding of Bicarbonate to Human Carbonic Anhydrase II: A Continuum of Binding States." Journal of the American Chemical Society **119** (5): 863-871.
- Mocharla, V. P., B. Colasson, L. V. Lee, S. Röper, K. B. Sharpless, C.-H. Wong and H. C. Kolb (2005). "In Situ Click Chemistry: Enzyme-Generated Inhibitors of Carbonic Anhydrase II." Angewandte Chemie International Edition **44** (1): 116-120.
- Mock, W. L., T. A. Irra, J. P. Wepsiec and T. L. Manimaran (1983). "Cycloaddition induced by cucurbituril. A case of Pauling principle catalysis." The Journal of Organic Chemistry **48** (20): 3619-3620.
- Morris, G. M., D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olson (1998). "Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function." Journal of Computational Chemistry **19** (14): 1639-1662.
- Nair, S. K. and D. W. Christianson (1991a). "Structural properties of human carbonic anhydrase II at pH 9.5." Biochemical and Biophysical Research Communications **181** (2): 579-584.
- Nair, S. K. and D. W. Christianson (1991b). "Unexpected pH-dependent conformation of His-64, the proton shuttle of carbonic anhydrase II." Journal of the American Chemical Society **113** (25): 9455-9458.
- Nair, S. K. and D. W. Christianson (1993). "Structural consequences of hydrophilic amino acid substitutions in the hydrophobic pocket of human carbonic anhydrase II." Biochemistry **32** (17): 4506-4514.
- Nishimori, I., T. Minakuchi, S. Onishi, D. Vullo, A. Scozzafava and C. T. Supuran (2006). "Carbonic Anhydrase Inhibitors. DNA Cloning, Characterization, and Inhibition Studies of the Human Secretory Isoform VI, a New Target for Sulfonamide and Sulfamate Inhibitors." Journal of Medicinal Chemistry **50** (2): 381-388.
- Nolte, C., P. Mayer and B. F. Straub (2007). "Isolation of a Copper(I) Triazolide: A "Click" Intermediate." Angewandte Chemie International Edition **46** (12): 2101-2103.

- O'Halloran, T. V. and V. C. Culotta (2000). "Metallochaperones, an Intracellular Shuttle Service for Metal Ions." Journal of Biological Chemistry **275** (33): 25057-25060.
- Otwinowski, Z., W. Minor and Charles W. Carter, Jr. (1997). [20] Processing of X-ray diffraction data collected in oscillation mode. Methods in Enzymology, Academic Press. **Volume 276**: 307-326.
- Pachón, Laura D., J. H. van Maarseveen and G. Rothenberg (2005). "Click Chemistry: Copper Clusters Catalyze the Cycloaddition of Azides with Terminal Alkynes." Advanced Synthesis & Catalysis **347** (6): 811-815.
- Pastorekova, S., S. Parkkila, J. Pastorek and C. T. Supuran (2004). "Carbonic Anhydrases: Current State of the Art, Therapeutic Applications and Future Prospects." Journal of Enzyme Inhibition and Medicinal Chemistry **19** (3): 199-229.
- Pinto, A. L., H. W. Hellinga and J. P. Caradonna (1997). "Construction of a catalytically active iron superoxide dismutase by rational protein design." Proceedings of the National Academy of Sciences of the United States of America **94** (11): 5562-5567.
- Puerta, D. T. and S. M. Cohen (2003). "Examination of Novel Zinc-Binding Groups for Use in Matrix Metalloproteinase Inhibitors." Inorganic Chemistry **42** (11): 3423-3430.
- Puerta, D. T., M. O. Griffin, J. A. Lewis, D. Romero-Perez, R. Garcia, F. J. Villarreal and S. M. Cohen (2006). "Heterocyclic zinc-binding groups for use in next-generation matrix metalloproteinase inhibitors: potency, toxicity, and reactivity." The Journal of Biological Inorganic Chemistry **11**: 131-138.
- Pufahl, R. A., C. P. Singer, K. L. Peariso, S. J. Lin, P. J. Schmidt, C. J. Fahrni, V. C. Culotta, J. E. Penner-Hahn and T. V. O'Halloran (1997). "Metal Ion Chaperone Function of the Soluble Cu(I) Receptor Atx1." Science **278** (5339): 853-856.
- Puig, S. and D. J. Thiele (2002). "Molecular mechanisms of copper uptake and distribution." Current Opinion in Chemical Biology **6** (2): 171-180.
- Ralle, M., S. Lutsenko and N. J. Blackburn (2003). "X-ray Absorption Spectroscopy of the Copper Chaperone HAH1 Reveals a Linear Two-coordinate Cu(I) Center Capable of Adduct Formation with Exogenous Thiols and Phosphines." Journal of Biological Chemistry **278** (25): 23163-23170.
- Regan, L. and N. D. Clarke (1990). "A tetrahedral zinc(II)-binding site introduced into a designed protein." Biochemistry **29** (49): 10878-10883.



- Reiss, W. G. and K. S. Oles (1996). "Acetazolamide in the treatment of seizures." The Annals of Pharmacotherapy **30** (5): 514-519.
- Renwick, A. G. (1985). "The disposition of saccharin in animals and man--A review." Food and Chemical Toxicology **23** (4-5): 429-435.
- Renwick, A. G. and T. W. Sweatman (1979). "The absorption of saccharin from the rat urinary bladder." Journal of Pharmacy and Pharmacology **31**: 650.
- Rodionov, V. O., V. V. Fokin and M. G. Finn (2005). "Mechanism of the Ligand-Free CuI-Catalyzed Azide-Alkyne Cycloaddition Reaction." Angewandte Chemie International Edition **44** (15): 2210-2215.
- Rodionov, V. O., S. I. Presolski, D. Diaz Diaz, V. V. Fokin and M. G. Finn (2007). "Ligand-Accelerated Cu-Catalyzed Azide-Alkyne Cycloaddition: A Mechanistic Report." Journal of the American Chemical Society **129** (42): 12705-12712.
- Rostovtsev, V. V., L. G. Green, V. V. Fokin and K. B. Sharpless (2002). "A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes." Angewandte Chemie International Edition **41** (14): 2596-2599.
- Schwarzer, D., H. D. Mootz, U. Linne and M. A. Marahiel (2002). "Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases." Proceedings of the National Academy of Sciences of the United States of America **99** (22): 14083-14088.
- Scolnick, L. R., A. M. Clements, J. Liao, L. Crenshaw, M. Hellberg, J. May, T. R. Dean and D. W. Christianson (1997). "Novel Binding Mode of Hydroxamate Inhibitors to Human Carbonic Anhydrase II." Journal of the American Chemical Society **119** (4): 850-851.
- Scozzafava, A., M. D. Banciu, A. Popescu and C. T. Supuran (2000). "Carbonic Anhydrase Inhibitors: Inhibition of Isozymes I, II and IV by Sulfamide and Sulfamic Acid Derivatives." Journal of Enzyme Inhibition and Medicinal Chemistry **15** (5): 443-453.
- Scozzafava, A., A. Mastrolorenzo and C. T. Supuran (2004). "Modulation of carbonic anhydrase activity and its applications in therapy." Expert Opinion on Therapeutic Patents **14** (5): 667-702.
- Scozzafava, A., A. Mastrolorenzo and C. T. Supuran (2006). "Carbonic anhydrase inhibitors and activators and their use in therapy." Expert Opinion on Therapeutic Patents **16**: 1627-1664.
- Scozzafava, A. and C. T. Supuran (2002). "Carbonic anhydrase activators: Human isozyme II is strongly activated by oligopeptides incorporating the

- carboxyterminal sequence of the bicarbonate anion exchanger AE1." Bioorganic & Medicinal Chemistry Letters **12** (8): 1177-1180.
- Sheldrick, G. M., T. R. Schneider, W. C. Charles, Jr. and M. S. Robert (1997). [16] SHELXL: High-resolution refinement. Methods in Enzymology, Academic Press. **Volume 277**: 319-343.
- Shi, F., J. P. Waldo, Y. Chen and R. C. Larock (2008). "Benzyne Click Chemistry: Synthesis of Benzotriazoles from Benzyne and Azides." Organic Letters **10** (12): 2409-2412.
- Shields, S. B. and S. J. Franklin (2004). "De Novo Design of a Copper(II)-Binding Helix- Turn- Helix Chimera: The Prion Octarepeat Motif in a New Context." Biochemistry **43** (51): 16086-16091.
- Shuker, S. B., P. J. Hajduk, R. P. Meadows and S. W. Fesik (1996). "Discovering high-affinity ligands for proteins: SAR by NMR." Science **274**: 1531-1534.
- Silverman, D. N. and S. Lindskog (1988). "The catalytic mechanism of carbonic anhydrase: implications of a rate-limiting protolysis of water." Accounts of Chemical Research **21** (1): 30-36.
- Silverman, D. N., C. Tu and G. C. Wynns (1976). "Depletion of  $^{18}\text{O}$  from  $\text{C}^{18}\text{O}_2$  in erythrocyte suspensions. The permeability of the erythrocyte membrane to  $\text{CO}_2$ ." Journal of Biological Chemistry **251** (14): 4428-4435.
- Simone, G. D., A. D. Fiore, V. Menchise, C. Pedone, J. Antel, A. Casini, A. Scozzafava, M. Wurl and C. T. Supuran (2005). "Carbonic anhydrase inhibitors. Zonisamide is an effective inhibitor of the cytosolic isozyme II and mitochondrial isozyme V: solution and X-ray crystallographic studies." Bioorganic & Medicinal Chemistry Letters **15** (9): 2315-2320.
- Smith, G. M., R. S. Alexander, D. W. Christianson, B. M. McKeever, G. S. Ponticello, J. P. Springer, W. C. Randall, J. J. Baldwin and C. N. Habecker (1994). "Positions of His-64 and a bound water in human carbonic anhydrase II upon binding three structurally related inhibitors." Protein Science **3** (1): 118-125.
- Steiner, H., B.-H. Jonsson and S. Lindskog (1975). "The Catalytic Mechanism of Carbonic Anhydrase." European Journal of Biochemistry **59** (1): 253-259.
- Steuber, H., A. Heine and G. Klebe (2007). "Structural and Thermodynamic Study on Aldose Reductase: Nitro-substituted Inhibitors with Strong Enthalpic Binding Contribution." Journal of Molecular Biology **368** (3): 618-638.
- Stiti, M., A. Cecchi, M. Rami, M. Abdaoui, V. r. Barragan-Montero, A. Scozzafava, Y. Guari, J.-Y. Winum and C. T. Supuran (2008). "Carbonic Anhydrase Inhibitor Coated Gold Nanoparticles Selectively Inhibit the Tumor-Associated Isoform IX

- over the Cytosolic Isozymes I and II." Journal of the American Chemical Society **130** (48): 16130-16131.
- Storoni, L. C., A. J. McCoy and R. J. Read (2004). "Likelihood-enhanced fast rotation functions." Acta Crystallographica Section D **60** (3): 432-438.
- Sugimoto, A., H. Ikeda, H. Tsukamoto, K. Kihira, M. Ishioka, J. Hirose, T. Hata, H. Fujioka and Y. Ono "Timolol Activates the Enzyme Activities of Human Carbonic Anhydrase I and II." Biological & Pharmaceutical Bulletin **33** (2): 301-306.
- Supuran, C. T. (2003). "Carbonic anhydrase inhibitors in the treatment and prophylaxis of obesity." Expert Opinion on Therapeutic Patents **13** (10): 1545-1550.
- Supuran, C. T. (2007). "Carbonic Anhydrases as Drug Targets - An Overview." Current Topics in Medicinal Chemistry **7**: 825-833.
- Supuran, C. T. (2008). "Carbonic anhydrases: novel therapeutic applications for inhibitors and activators." Nat Rev Drug Discov **7** (2): 168-181.
- Supuran, C. T. and A. Scozzafava (2000). "Carbonic anhydrase inhibitors and their therapeutic potential." Expert Opinion on Therapeutic Patents **10** (5): 575-600.
- Supuran, C. T. and A. Scozzafava (2002). "Applications of carbonic anhydrase inhibitors and activators in therapy." Expert Opinion on Therapeutic Patents **12** (2): 217-242.
- Supuran, C. T. and A. Scozzafava (2007). "Carbonic anhydrases as targets for medicinal chemistry." Bioorganic & Medicinal Chemistry **15** (13): 4336-4350.
- Supuran, C. T., A. Scozzafava and A. Casini (2003). "Carbonic anhydrase inhibitors." Medicinal Research Reviews **23** (2): 146-189.
- Supuran, C. T., A. Scozzafava and J. Conway (2004). "Carbonic Anhydrase - Its inhibitors and activators." CRC Press: 1-363.
- Temperini, C., A. Innocenti, A. Scozzafava, A. Mastrolorenzo and C. T. Supuran (2007). "Carbonic anhydrase activators: l-Adrenaline plugs the active site entrance of isozyme II, activating better isoforms I, IV, VA, VII, and XIV." Bioorganic & Medicinal Chemistry Letters **17** (3): 628-635.
- Temperini, C., A. Innocenti, A. Scozzafava, S. Parkkila and C. T. Supuran (2009). "The Coumarin-Binding Site in Carbonic Anhydrase Accommodates Structurally Diverse Inhibitors: The Antiepileptic Lacosamide As an Example and Lead Molecule for Novel Classes of Carbonic Anhydrase Inhibitors." Journal of Medicinal Chemistry **53** (2): 850-854.

- Temperini, C., A. Innocenti, A. Scozzafava and C. T. Supuran (2006). "N-Hydroxyurea-A versatile zinc binding function in the design of metalloenzyme inhibitors." Bioorganic & Medicinal Chemistry Letters **16** (16): 4316-4320.
- Temperini, C., A. Innocenti, A. Scozzafava and C. T. Supuran (2008). "Carbonic anhydrase activators: Kinetic and X-ray crystallographic study for the interaction of d- and l-tryptophan with the mammalian isoforms I-XIV." Bioorganic & Medicinal Chemistry **16** (18): 8373-8378.
- Temperini, C., A. Scozzafava, L. Puccetti and C. T. Supuran (2005). "Carbonic anhydrase activators: X-ray crystal structure of the adduct of human isozyme II with l-histidine as a platform for the design of stronger activators." Bioorganic & Medicinal Chemistry Letters **15** (23): 5136-5141.
- Temperini, C., A. Scozzafava and C. T. Supuran (2006a). "Carbonic anhydrase activators: The first X-ray crystallographic study of an adduct of isoform I." Bioorganic & Medicinal Chemistry Letters **16** (19): 5152-5156.
- Temperini, C., A. Scozzafava, D. Vullo and C. T. Supuran (2006b). "Carbonic Anhydrase Activators. Activation of Isozymes I, II, IV, VA, VII, and XIV with L- and D-Histidine and Crystallographic Analysis of Their Adducts with Isoform II: Engineering Proton-Transfer Processes within the Active Site of an Enzyme." Chemistry - A European Journal **12** (27): 7057-7066.
- Toba, S., G. Colombo and K. M. Merz (1999). "Solvent Dynamics and Mechanism of Proton Transfer in Human Carbonic Anhydrase II." Journal of the American Chemical Society **121** (10): 2290-2302.
- Tornøe, C. W., C. Christensen and M. Meldal (2002). "Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides." The Journal of Organic Chemistry **67** (9): 3057-3064.
- Toyama, M., M. Sasaki, N. Hirayama, Y. Murooka and M. Yamashita (2006). "Construction of an additional metal-binding site in human metallothionein-2." Journal of Bioscience and Bioengineering **101** (4): 354-360.
- Tron, G. C., T. Pirali, R. A. Billington, P. L. Canonico, G. Sorba and A. A. Genazzani (2008). "Click chemistry reactions in medicinal chemistry: Applications of the 1,3-dipolar cycloaddition between azides and alkynes." Medicinal Research Reviews **28** (2): 278-308.
- Tu, C., D. N. Silverman, C. Forsman, B. H. Jonsson and S. Lindskog (1989). "Role of histidine 64 in the catalytic mechanism of human carbonic anhydrase II studied with a site-specific mutant." Biochemistry **28** (19): 7913-7918.

- Tu, C., G. C. Wynns and D. N. Silverman (1981). "Inhibition by cupric ions of 18O exchange catalyzed by human carbonic anhydrase II. Relation to the interaction between carbonic anhydrase and hemoglobin." Journal of Biological Chemistry **256** (18): 9466-9470.
- Wang, T. and Z. Guo (2006). "Copper in Medicine: Homeostasis, Chelation Therapy and Antitumor Drug Design." Current Medicinal Chemistry **13**: 525-537.
- Weihrauch, M. R. and V. Diehl (2004). "Artificial sweeteners do they bear a carcinogenic risk?" Annals of Oncology **15** (10): 1460-1465.
- Wernimont, A. K., L. A. Yatsunyk and A. C. Rosenzweig (2004). "Binding of Copper(I) by the Wilson Disease Protein and Its Copper Chaperone." Journal of Biological Chemistry **279** (13): 12269-12276.
- Winum, J.-Y., M. Rami, A. Scozzafava, J.-L. Montero and C. Supuran (2008). "Carbonic anhydrase IX: A new druggable target for the design of antitumor agents." Medicinal Research Reviews **28** (3): 445-463.
- Winum, J.-Y., A. Scozzafava, J.-L. Montero and C. T. Supuran (2005). "Sulfamates and their therapeutic potential." Medicinal Research Reviews **25** (2): 186-228.
- Winum, J.-Y., A. Scozzafava, J.-L. Montero and C. T. Supuran (2006). "Therapeutic potential of sulfamides as enzyme inhibitors." Medicinal Research Reviews **26** (6): 767-792.
- Winum, J. Y., A. Scozzafava, J. L. Montero and C. T. Supuran (2007). "Metal Binding Functions in the Design of Carbonic Anhydrase Inhibitors." Current Topics in Medicinal Chemistry **7**: 835-848.
- Winum, J. Y., A. Scozzafava, J. L. Montero and C. T. Supuran (2008). "Design of Zinc Binding Functions for Carbonic Anhydrase Inhibitors." Current Pharmaceutical Design **14**: 615-621.
- Zhang, L., X. Chen, P. Xue, H. H. Y. Sun, I. D. Williams, K. B. Sharpless, V. V. Fokin and G. Jia (2005). "Ruthenium-Catalyzed Cycloaddition of Alkynes and Organic Azides." Journal of the American Chemical Society **127** (46): 15998-15999.



## Danksagung

*Herrn Prof. Dr. Gerhard Klebe* danke ich herzlich für die interessante Themenstellung dieser Doktorarbeit und die sehr gute Betreuung während der gesamten Zeit. Ich danke ihm für die fachlichen und persönlichen Gespräche für die er stets Zeit gefunden hat, in denen viele Ideen, Konzepte und Lösungen entstanden sind. Auch danke ich ihm für die vielen ausführlichen Korrekturen von Manuskripten sowie das Vertrauen in mich die Projekte eigenständig lenken und weiterentwickeln zu können.

*Herrn Dr. Andreas Heine* danke ich für die ausführliche und fundierte Einführung in die Kristallographie. Seine Erfahrung, die er mit mir geteilt hat, hat maßgeblich zu der Qualität dieser Arbeit beigetragen.

Ich danke *Herrn Prof. Dr. Ulrich Koert* und *Herrn Dong Sun* für die Kooperation auf dem Gebiet der Click-Chemie, die Synthese der Verbindungen und die Diskussionen, die zum Erfolg des Projektes geführt haben. Weiterhin möchte ich mich sehr bei *Herrn Prof. Dr. Ulrich Koert* für Erstellung des Zweitgutachtens bedanken.

Herzlich danke ich *Herrn Prof. Dr. Manfred Haake* und seinen *ehemaligen Mitarbeitern* für die Synthese getesteter Substanzen, sowie *Herrn Prof. Dr. Manfred Haake* für die vielen konstruktiven Gespräche und seine Mitwirkung in der Prüfungskommission.

*Herrn Prof. Dr. Udo Bakowsky* danke ich sehr für die Bereitschaft in der Prüfungskommission mitzuwirken.

*Herrn Dr. Claudiu Supuran* (Florenz, Italien) und seinen *Mitarbeitern* danke ich für Vermessung vieler Inhibitoren und den freundlichen Aufenthalt in Florenz.

*Herrn Prof. Dr. Seth Cohen* (San Diego, USA) möchte ich für die Synthese und Bereitstellung neuer Fragmente danken.

*Der Gruppe von Prof. Pastorekova* (Bratislava, Slowakei) danke ich für die Unterstützung bei der Klonierung des CA II Gens.

Ich danke *Herrn Prof. Dr. Klaus Reuter* für die Leitung des S1 Labors und seine freundliche Hilfe bei allen Fragestellungen der letzten 4 Jahre.

*Frau Nicola Sandner* danke ich für die sehr gute und sehr angenehme Zusammenarbeit, die stetige Unterstützung und die Weiterführung der Projekte.

*Herrn Dr. Sascha Brass* und *Herrn Thomas Kronenberger* möchte ich für die Synthese der Tethering-Substanzen danken. *Dres. Nan-Si* und *Sascha Brass* danke ich zudem für ihre außerordentliche Hilfsbereitschaft und Unterstützung.

*Frau Prof. Dr. Wibke Diederich und Herrn Hans-Dieter Gerber* danke ich für den freundlichen Austausch bei Fragen aus dem Bereich der organischen Synthese.

Ich danke *Herrn Dr. Uwe Linne* und *Frau Natalia Fritzler* für die Präparation sowie die Vermessung der HPLC-MS Proben.

*Herrn Tobias Craan* danke ich für die sehr gute Einführung in das Docking Programm GOLD, für die Erstellung von Skripten, die bei der Auswertung von Docking-Berechnungen und Verfeinerung von Kristallstrukturen Anwendung gefunden haben, sowie jegliche Unterstützung bei administrativen Problemen.

*Frau Lydia Hartleben* und *Frau Angela Scholz* möchte ich für Ihre große Hilfsbereitschaft und Unterstützung danken.

Ein besonderer Dank gilt *Herrn Christian Sohn* für die Betreuung und Wartung der Röntengeräte sowie die stete Bereitschaft zur Hilfe.

Mein Dank gilt *allen Angestellten* des *BESSY II, Berlin* sowie des *SLS, Villigen*.

*Allen Administratoren* unserer Arbeitsgruppe danke ich für die Wartung der Rechner und des Netzwerkes.

*Herrn Christian Hasewinkel, Herrn Florian Immekus* und *Frau Daniela Meyer* danke ich für die angenehme und produktive Atmosphäre im S1 Labor sowie die ausgezeichnete Beratung bei Problemen.

*Allen Vertiefungsstudenten* danke ich für die geleistete Arbeit bei der Mutagenese und Expression von CA II Mutanten.

Ich danke den *Mitgliedern der AG Klebe* für die gute Zusammenarbeit. Vor allem danke ich *Jürgen, Tobias, Florian, Kan, Christian S, Andreas S., Michael, Patrick, Tina, Björn, Jark, Andreas B., Nan-Si, Sascha, Nicola, Daniela und Frank (AG Steinmetzer), und Markus (AG Kissel)* für die schöne Zeit am Institut.

Danken möchte ich *Herrn Dr. Andreas Heine, Herrn Daniel Schmidt, Herrn Florian Immekus, Herrn Tobias Craan* und *Herrn Thomas Schulze Wischeler* für das Korrekturlesen des Manuskripts meiner Dissertation.

Ich danke *Herrn Jürgen Behnen* für die ausgezeichnete Büroatmosphäre, die vielen interessanten und lustigen Gespräche und seine großen Weisheiten.

*Meiner ganzen Familie* und *meinen Freunden* danke ich für die große Unterstützung im Studium und während der Promotion. Ohne sie wäre diese Arbeit nicht möglich gewesen. Von ganzem Herzen danke ich *meinen Eltern* und *Vanessa* für die allgegenwärtige, unermüdliche Unterstützung, ihren Rückhalt, ihre Förderung und ihr Vertrauen.



## **Erklärung**

Ich versichere, dass ich meine Dissertation

**Carbonic Anhydrase II: A Model System for Artificial Copper Center Design,  
Protein-guided Cycloadditions, Tethering Screenings and Fragment-based Lead  
Discovery**

selbständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den 21.05.2010

.....  
(Johannes Schulze Wischeler)



## Curriculum Vitae

### PERSÖNLICHE DATEN

---

Familienstand: ledig  
Geburtsdatum/-Ort: 30. März 1981 in Bergisch-Gladbach  
Staatsangehörigkeit: deutsch

### HOCHSCHULBILDUNG

---

Seit 01/2007 Promotion am Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Arbeitsgruppe Prof. Dr. Klebe  
03/2008 – 08/2009 Aufbaustudium zum "Betriebswirt (IWW)" am Institut für Wirtschaftswissenschaftliche Forschung und Weiterbildung der Fernuniversität Hagen  
10/2007 – 03/2009 „Zusatzqualifikation Pharmarecht“ der Forschungsstelle für Pharmarecht, Philipps-Universität Marburg  
12/2006 Erteilung der Approbation als Apotheker  
11/2005 – 04/2006 Diplomarbeit zum Thema: „Erstellung und Optimierung verschiedener Expressionssysteme für die Shigellen-Invasine IpaA, IpgB2 und OspD1 mit dem Chaperon Spa15“, Philipps-Universität Marburg, Arbeitsgruppe Prof. Dr. Klebe  
10/2001 – 09/2005 Studium der Pharmazie an der Philipps Universität Marburg

### SCHULBILDUNG

---

08/1991 – 06/2001 Norbert-Gymnasium Knechtsteden in Dormagen: Abitur  
08/1997 – 06/1998 East Union High School in Afton, Iowa, USA: High School Diploma  
08/1987 – 06/1991 Tannenbuschgrundschule in Dormagen-Delhoven