Charles University in Prague
First Faculty of Medicine

Postgraduate Doctoral Studies in Biomedicine
Molecular and Cell Biology, Genetics and Virology

Mgr. Pavel Mader

Structure, Function and Inhibition of Human Carbonic Anhydrases

PhD Thesis

Supervisor: Doc. RNDr. Jiří Brynda, CSc.
Institute of Molecular Genetics, v.v.i.
Academy of Sciences of the Czech Republic

Prague 2010

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje. Současně dávám svolení k tomu, aby tato závěrečná práce byla archivována v Ústavu vědeckých informací 1. lékařské fakulty Univerzity Karlovy v Praze a zde užívána ke studijním účelům. Za předpokladu, že každý, kdo tuto práci použije pro svou přednáškovou nebo publikační aktivitu, se zavazuje, že bude tento zdroj informací řádně citovat. Souhlasím se zpřístupněním elektronické verze mé práce v Digitálním repozitáři Univerzity Karlovy v Praze (http://repozitar.cuni.cz). Práce je zpřístupněna pouze v rámci Univerzity Karlovy v Praze

V Praze, 12. 6. 2010

Pavel Mader

Identification record:

MADER, Pavel. *Structure, Function and Inhibition of Human Carbonic Anhydrases*. Praha, 2010. 122 s. Disertační práce. Universita Karlova v Praze, 1. lékařská fakulta. Vedoucí práce Doc. RNDr. Jiří Brynda, CSc.

Abstrakt:

Lidské karbonické anhydrasy jsou zinkové metaloenzymy hrající klíčovou roli v řadě fyziologických a patofyziologických procesů. CA IX je transmembránový isoenzym asociován s řadou lidských nádorů. Výsledky exprese, purifikace a krystalizace CA IX spolu se studiem struktury isoenzymu CA II v komplexu s novou třídou nízkomolekulárních inhibitorů jsou prezentovány. Dále strukturní studie Fab fragmentu monoklonální protilátky M75 a jeho komplexu s epitopovým peptidem nacházejícím se v unikátní "proteoglycan-like" doméně CA IX jsou předmětem této práce.

Klíčová slova: nádory, karbonická anhydrasa, inhibitory karbonických anhydras, krystalová struktura Fab, rozpoznávání antigenu protilátkou

Abstract:

Human carbonic anhydrases are zinc metalloenzymes playing a key role in several physiological and pathophysiological processes. CA IX is a tumor associated transmembrane isozyme representing a valuable therapeutic target. Results concerning expression, purification, and crystallization of CA IX as well as structural studies of CA II in complex with novel class of small molecular inhibitors are presented. Furthermore, structural studies of Fab fragment of monoclonal antibody M75 and its complex with epitope peptide derived from unique proteoglycan-like domain of CA IX are part of this work.

Keywords: cancer, carbonic anhydrase, carbonic anhydrase inhibitors, Fab crystal structure, antigen-antibody recognition

ACKNOWLEDGEMENTS

The work presented in this thesis was carried out in the Laboratory of Structural Biology at the Institute of Molecular Genetics, v.v.i., Academy of Sciences of the Czech Republic.

I would like to express my deep gratitude to the following people:

Jiří Brynda, my supervisor, for assigning me the project of my PhD thesis, for professional guidance, and for leading my first steps in the field of protein crystallography.

Pavlína Řezáčová, the head of our department, for creating excellent working conditions, and for sharing her vast experience in protein crystallization and protein structure analysis.

Juraj Sedláček for support of my work from the first day in the laboratory, for guidance in manuscript writing, and also for many valuable suggestions.

Milan Fábry for many inspiring discussions, for focusing the course of my work, and for introducing me to molecular biology techniques, which I have used in my work.

Vlastimil Král, Irena Sieglová, Petr Pachl and Kateřina Procházková, my fellow colleagues, for productive collaboration and personal support.

And finally, I thank all the remaining members of the Laboratory, for their help during the course of my work, and also for creating a stimulating and friendly atmosphere.

Many thanks also belong to collaborators from Jan Konvalinka's lab from the Institute of Organic Chemistry and Biochemistry, v.v.i., AS CR, namely to Petra Mlčochová, Pavel Šácha, and Jana Starková for their help with the S2 cell eukaryotic expression system.

I would also like to thank the staff of Hamburg synchrotron beamline X12 for providing measurement time.

This work is dedicated to my family and to the memory of my parents for their love, understanding and support.

This study was financially supported by

- Academy of Sciences of the Czech Republic, Project No. AV0Z50520514

- Ministry of Education of the Czech Republic, Grant No. 1M0505

- Czech Grant Foundation, Project No. GA203/09/0820

TABLE OF CONTENTS

1 INTRODUCTION 8

2 CURRENT STATE OF THE RESEARCH FIELD 10

2.1 CARBONIC ANHYDRASES IN NATURE 10

2.1.1 The α-class 11

2.1.2 The β-class 14

2.1.3 The ancient γ-class 15

2.1.4 Other classes (δ and ζ) 16

2.2 HUMAN CARBONIC ANHYDRASES 18

2.3 PHYSIOLOGICAL ROLE OF HUMAN CARBONIC ANHYDRASES 20

2.4 CAS AS THERAPEUTIC TARGETS 20

2.5 CARBONIC ANHYDRASE IX AS A NOVEL CANCER THERAPY TARGET 21

2.5.1 Molecular properties of CA IX 23

3 AIMS OF THE THESIS 24

4 MATERIAL AND METHODS 25

4.1 MOLECULAR CLONING – CONSTRUCTION OF CA IX EXPRESSION PLASMIDS 25

4.1.1 Plasmids for bacterial expression 26

4.1.2 Plasmids for expression in Drosophila S2 cells 27

4.1.2.1 Mutagenesis to remove glycosylation signal 27

4.1.2.2 N-terminal His-tag with TEV cleavage site and C-terminal extension 28

4.1.2.3 N-terminal His-tag with TEV cleavage site without C-terminal extension 29

4.1.2.4 N-terminal Avi-tag with TEV cleavage site 29

4.1.2.5 C-terminal Avi-tag 29

4.1.2.6 Mutagenesis Cys(41)Ser 31

4.1.2.7 CA domain (Ser41) extensions 31

4.2 EXPRESSION OF CA IX CONSTRUCTS IN *E. COLI* 32

4.2.1 Bacterial media 32

4.2.2 Host strains 32

4.2.3 Cultivation conditions 33

4.3 EXPRESSION OF CA IX CONSTRUCTS IN S2 CELLS 34

4.3.1 Insect cell cultivation 34

4.3.2 Transfection of insect cells and stable cell line generation 34

4.3.3 Recombinant CA IX large scale expression 35

4.4 PURIFICATION OF RECOMBINANT CA IX 35

4.4.1 Sulfonamide agarose affinity chromatography 36

4.4.2 Ion exchange chromatography 36

4.4.3 Purification of His-tagged proteins on Ni-affinity resin *37*
4.4.4 Expression and purification of Avi-tagged proteins on Streptavidin mutein matrix *37*
4.5 CARBONIC ANHYDRASE ACTIVITY ASSAY 38
4.6 PREPARATION OF FAB M75 FOR CRYSTALLIZATION 40
4.7 PROTEIN CRYSTALLOGRAPHY 40
4.7.1 Protein crystallization *40*
4.7.2 X-ray data collection *41*
4.7.3 Structure determination, refinement, and analysis *41*

5 RESULTS AND DISCUSSION 42

5.1 TARGETING ACTIVE SITE OF HUMAN CARBONIC ANHYDRASES 42
5.1.1 Structural studies of CA II in complex with inhibitors based on 3,4-dihydroisoquinoline-2(1H)-sulfonamide *42*
5.1.2 Crystallization of CA II complexes *44*
5.1.3 Data collection and structure determination *45*
5.1.4 Refined models and overall structures *48*
5.1.5 Details of inhibitor DT1 and DT2 binding *50*
5.1.6 Comparison of inhibitor DT1 and DT2 binding mode *54*
5.1.7 Conclusions from structural analysis of CA II + inhibitor complexes *56*
5.2 STRUCTURAL STUDIES OF CA IX 58
5.2.1 Recombinant CA IX constructs for E. coli expression *58*
5.2.2 Recombinant CA IX constructs for eukaryotic expression *64*
5.2.3 Conclusion from structural studies of CA IX *72*
5.3 STRUCTURE OF MOUSE MAB M75 IN COMPLEX WITH EPITOPE PEPTIDE 73
5.3.1 Crystallization, data collection and structure determination *73*
5.3.2 Refined models and overall structures *78*
5.3.3 Comparison of Fab M75 free and complexed structure *82*
5.3.4 Peptide-antibody interactions *85*
5.3.4.1 Polar interactions 85
5.3.4.2 Van der Waals interactions 89
5.3.5 Conclusion from Fab M75 structural analyses *89*

6 SUMMARY 91

7 ABBREVIATIONS 92

8 PUBLICATIONS 94

8.1 ORIGINAL PAPERS 94
8.2 ABSTRACTS AND SHORT COMMUNICATIONS 95
8.3 REPRINTS OF PUBLISHED ORIGINAL PAPERS 96

9 REFERENCES 117

1 INTRODUCTION

Carbonic anhydrases (CA, EC 4.2.1.1) are ubiquitous metalloenzymes catalyzing reversible hydration of carbon dioxide. Human CAs belong to the α-class of carbonic anhydrases, containing zinc ion in their active site. To date, fifteen human CA isozymes, displaying differences in activity, subcellular localization, and tissue expression profiles have been identified. They play key role in intracellular and extracellular pH homeostasis, in transport of CO_2 and bicarbonate in respiration, and in several biochemical pathways where either CO_2 or bicarbonate is required (Krishnamurthy *et al.*, 2008): bone resorption, production of gastric acid, renal acidification, and lipogenesis representing just a few examples.

Some of the human isozymes are established diagnostic and therapeutic targets. Several carbonic anhydrase inhibitors (CAIs) are clinically used as anti-glaucoma drugs, anti-convulsants, and anti-obesity agents (Mincione *et al.*, 2007, Supuran, 2007).

Human CA IX isozyme, discovered by team led by Dr. Jan Závada (Pastorekova *et al.*, 1992), is a special member of the family. Under normal conditions, its physiological expression is limited to a very narrow range of tissues. However, this membrane-bound isozyme is highly overexpressed on the cell surface of a variety of solid malignant tumors derived from kidney, cervix, uteri, colon, lung, oesophagus and breast (Zavada *et al.*, 1993). Interesting feature of this isozyme is the presence of N-terminal proteoglycan-like (PG) domain preceding the catalytic CA domain (Opavsky *et al.*, 1996). The overexpression of CA IX is induced by hypoxia, thus CA IX is used as a marker of tumor hypoxia and also as a prognostic factor for several human cancers (Tunuguntla & Jorda, 2008). For these reasons, CA IX serves as a valuable target for diagnostics and became a target for antitumor therapy (Winum *et al.*, 2008).

Due to the presence of the unique PG domain, human CA IX lends itself to a two-pronged “attack”, one arm targeting the PG moiety and the other the active site at the catalytic domain CA. First, the inhibition of the enzymatic activity, crucial for hypoxic tumor growth and progression (Poulsen, 2010) was studied. Structure based drug design requires detailed structural analysis of the active site, so that small molecular inhibitors, that would specifically inhibit the CA IX isozyme, can be designed, with minimal effect on other members of the CA family. The prerequisite for

these studies was a sufficient supply of recombinant CA IX protein of crystallization grade. Results concerning expression, purification and crystallization trials of various CA IX constructs are presented. Complementary to these efforts were structural studies of CA II, abundant, physiologically important and easily available isozyme, which was chosen as a representative model of other family members. The obtained structures of hCA II, free and complexed with a novel class of inhibitors based on isoquinoline scaffold, allowed discerning the fine details of the inhibitor binding mode to the active site, thus providing clues for design of inhibitors selective for CA IX.

As for the PG domain, an excellent tool for diagnostics and possibly targeted therapy is a specific mouse monoclonal antibody M75 (Zavada *et al.*, 2000). This antibody recognizes a linear epitope GDEELP localized in the PG domain of CA IX. Crystal structure of M75 Fab fragment complexed with this epitope peptide was determined in the present work with the aim to better understand the antibody-antigen interactions and to obtain structural information that would help in future attempts to humanize the antibody and/or its recombinant fragments.

Overall, the results of this work significantly improve our understanding of antigen-antibody recognition as well as structure of human CA isozymes, and should lead to improved design of both macromolecular (single chain Fv antibody fragments) as well as small molecular compounds selective for clinically important CAs, namely the tumor associated isozyme CA IX.

2 CURRENT STATE OF THE RESEARCH FIELD

The first carbonic anhydrase (CA, EC 4.2.1.1) was described nearly 80 years ago (Meldrum & Roughton, 1933a) as an enzyme present in erythrocytes, which catalyzes the reversible hydration of carbon dioxide:

$$CO_2 + H_2O \rightleftarrows HCO_3^- + H^+$$

Later, it was shown that carbonic anhydrases are nearly ubiquitous enzymes, with different classes identified in nearly all types of organisms from prokaryotes to humans.

2.1 Carbonic anhydrases in nature

Since the discovery of carbonic anhydrase in erythrocytes in 1933, several mammalian isozymes have been characterized and, for a long time, they have dominated carbonic anhydrase research. As concerns plants, since about 1940s it is known that carbonic anhydrase plays a crucial role in CO_2 fixation (Bradfield, 1947). As demonstrated by comparison of amino acid sequences and crystal structures, the mammalian and plant enzymes have evolved independently and have been designated the α- and β-class, respectively. In 1994, an independently evolved γ-class was reported (Alber & Ferry, 1994). This is an archean class of ancient origin; having evolved between 3.0 to 4.5 billion years ago (preceding evolution of α-class by more than 200 – 300 million years) (Smith *et al.*, 1999, Hewett-Emmett & Tashian, 1996, Jiang & Gupta, 1999). Finally, two other CAs belonging to new δ and ζ classes were isolated from marine diatoms and characterized recently (Park *et al.*, 2007).

Despite significant sequence and structural differences, all known CAs are metalloenzymes performing the same function, and vast majority contain a catalytic zinc ion in their active site. Thus carbonic anhydrase was repeatedly invented in nature, and the various CAs serve as an excellent example of convergent evolution (Smith & Ferry, 2000).

2.1.1 The α-class

The α-class is the best characterized group with 15 isozymes identified in mammals. In comparison to the other two major classes (β and γ), relatively few members of this class have been identified in prokaryotes. Several of the human isozymes from the α-class family are implicated in various disease states, as will be discussed further; and the treatment frequently involves application of sulfonamide based carbonic anhydrase inhibitors.

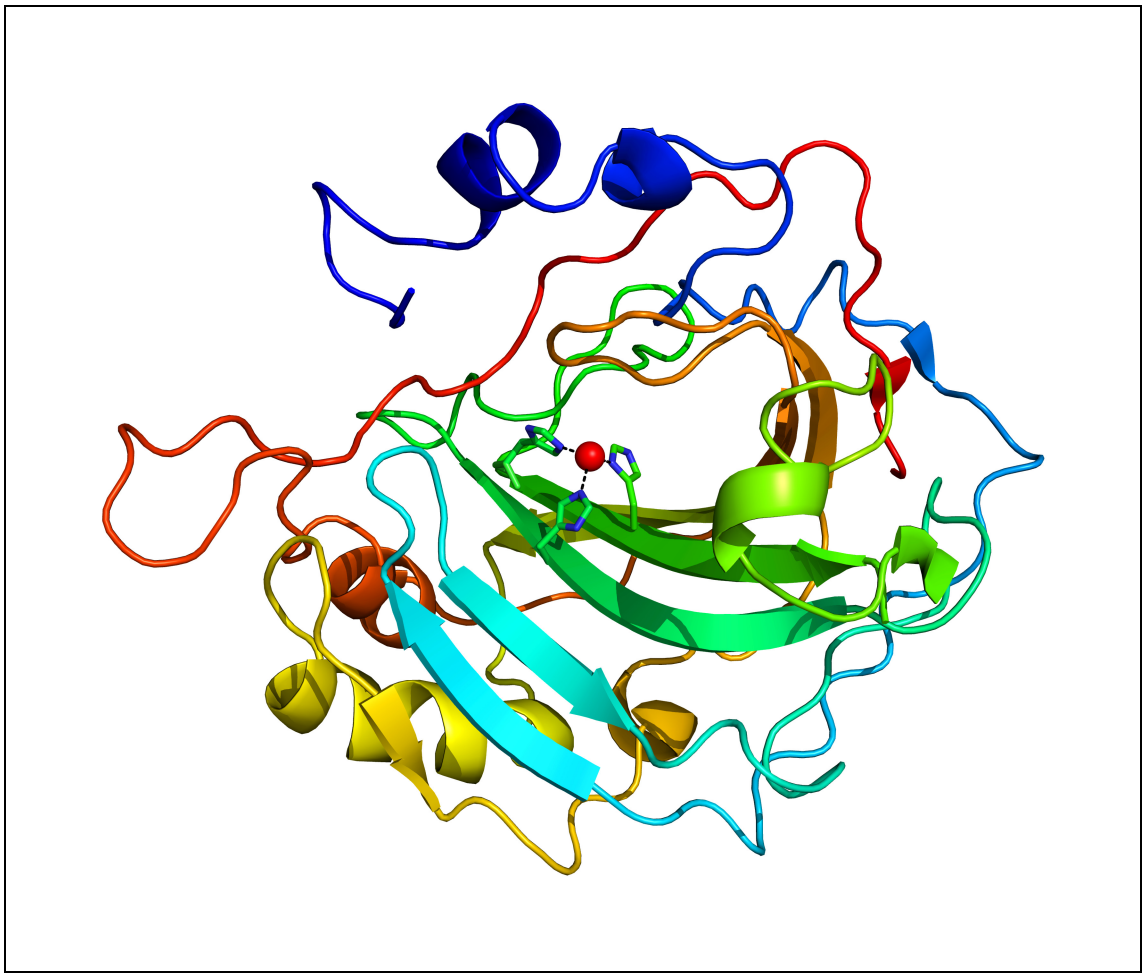


Figure 1 Three-dimensional structure of α-class of carbonic anhydrase.

Ribbon representation of human isozyme CA II. The active site Zn^{2+} ion (shown as a red sphere) is coordinated by three histidine residues. The figure was prepared using the structure deposited in PDB under the accession code 4CA2 (Alexander et al., 1991) and program PyMOL (DeLano, 2002).

The active site of α-class of CAs contains a catalytic Zn^{2+} ion coordinated by three histidines. Three-dimensional structure of human isozyme CA II, a representative member of the α-class of CAs, is depicted in Figure 1. Water molecule in form of hydroxide ion serves as the fourth ligand in a distorted tetrahedral geometry. The active site is located at the base of a cone-shaped depression with one wall dominated by hydrophobic and the other by hydrophilic residues (Krishnamurthy *et al.*, 2008).

Biochemical and biophysical properties of the most studied human members CA I and CA II were summarized recently in a review by Prof. Whitesides group (Krishnamurthy *et al.*, 2008). The reversible hydration of CO_2 to bicarbonate is catalyzed in a two-step "ping-pong" mechanism (Lindskog & Silverman, 2000). Schematic representation of this mechanism is shown in Figure 2.

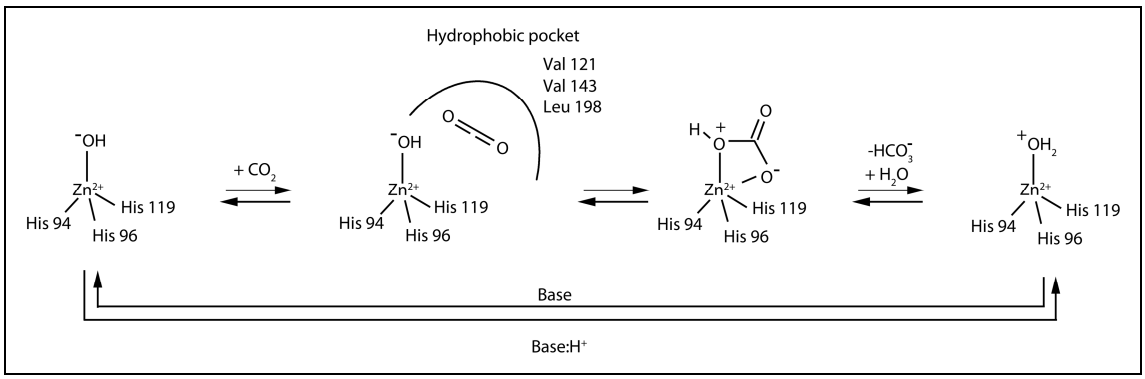


Figure 2 Schematic representation of CO_2 hydration by α-CAs.

The active site Zn^{2+} is coordinated to three histidine residues and H_2O molecule. This water molecule is deprotonated to generate strongly basic hydroxide anion, which is the reactive species in the hydration of CO_2 converting it to HCO_3^-. The hypothesized hydrophobic pocket in human isozyme CA II for the binding of substrates is shown schematically with CO_2 molecule. Figure adapted partly from Supuran et al., 2004 and Poulsen, 2010.

In the first step, the Zn^{2+} bound OH^- group performs nucleophilic attack on CO_2 resulting in formation of bicarbonate. Its release from the metal is mediated by displacement with a water molecule. In the second step, a proton is transferred from the Zn^{2+} bound water to the buffers in the solvent and the Zn^{2+} bound hydroxide is regenerated. The shuttling of proton via a series of intramolecular and intermolecular

proton transfer steps appears to be the rate limiting step in the catalysis (Silverman & Lindskog, 1988). Residue His64 seems to play a key role in this process, mutations of this residue to Ala cause the catalytic activity drop to 6 – 12 % of the native enzyme values (Tu *et al.*, 1989).

The catalytic efficiency of carbonic anhydrases is generally very high, in fact, their catalytic turnovers are among the highest known (k_{cat} in the range of 10^6 s^{-1}) and the second order rate constants are reaching diffusion controlled limit (k_{cat}/K_M in the range of 10^8 $M^{-1}s^{-1}$). Kinetic constants for various CA isozymes as well as some other highly efficient enzymes are listed in Table 1. Data were adapted from Krishnamurthy et al., 2008.

Table 1 Steady-state kinetic constants for selected α-class CAs and other highly-efficient enzymes
Data adapted from Krishnamurthy et al., 2008.

Enzyme	**Source**	**k_{cat} [s^{-1}]**	**k_{cat}/K_M [$M^{-1}s^{-1}$]**
CA I	human	2×10^5	5×10^7
CA II	human	14×10^5	15×10^7
CA III	human	0.1×10^5	0.03×10^7
CA III	bovine	0.064×10^5	0.04×10^7
CA IV	human	11×10^5	5×10^7
CA IV	murine	11×10^5	3.2×10^7
CA VA	murine	3×10^5	3×10^7
CA VB	human	9.5×10^5	9.8×10^7
CA VI	human	3.4×10^5	4.9×10^7
CA VI	rat	0.7×10^5	1.6×10^7
CA VII	murine	9.4×10^5	7.6×10^7
CA IX	human	3.8×10^5	5.5×10^7
CA XII	human	4×10^5	7.4×10^7
CA XIII	murine	0.83×10^5	4.3×10^7
CA XIV	human	3.12×10^5	3.9×10^7
acetylcholinesterase	eel electric-organ	0.14×10^5	16×10^7
catalase	human	5.5×10^5	0.7×10^7
catalase	horse liver	380×10^5	3.5×10^7
β-lactamase	human	0.02×10^5	10×10^7
superoxide dismutase	human	0.04×10^5	800×10^7
triosephosphate isomerase	human	0.043×10^5	24×10^7

In 2008, Hilvo *et al.* have shown that under appropriate conditions (namely $ZnCl_2$ concentration) the catalytic activity of some constructs of the human CA IX can increase to the highest level ($k_{cat}/K_M = 3.4 \times 10^9$ $M^{-1}s^{-1}$) ever measured for a CA isozyme (Hilvo *et al.*, 2008).

2.1.2 The β-class

As mentioned above, β-carbonic anhydrases (β-CAs) were originally discovered in plants (Neish, 1939) as constituents of plant leaf chloroplasts. It was nearly five decades later, when they were recognized as evolutionary and structurally distinct form of CAs, thus assigned the β-designation (Rowlett, 2010). Since 1990s, β-class CAs were isolated from a variety of photosynthetic organisms, but also from eubacteria, yeast and Archea species (Smith *et al.*, 1999). For many organisms, β-CAs are essential for growth at atmospheric concentration of CO_2. Distribution of this enzyme among phylogenetically and physiologically diverse prokaryotes indicates an important role of this enzyme in nature (Tripp *et al.*, 2001).

The first structure of β-CA enzyme has been published rather recently (Mitsuhashi *et al.*, 2000) in comparison to the α-class, e.g. structure of hCA II from 1972 (Liljas *et al.*, 1972). From the structural point of view, all β-CAs share a unique α/β fold which is not found in any other proteins (Rowlett, 2010). In contrast to other two classes, members of β-CAs form dimers which can further associate into tetramers, hexamers, and octamers. The members of the α-class enzymes are almost strictly monomers (with some exceptions) and γ-class enzymes form trimers. Figure 3 shows ribbon diagram of the fundamental dimer of β-CA from red algea *Porfiridium purpureum* (PDB ID 1DDZ). In addition to the differences in secondary structures, β-class crystal structures show that zinc is coordinated by two conserved cysteines and one conserved histidine (Mitsuhashi *et al.*, 2000, Kimber & Pai, 2000, Strop *et al.*, 2001).

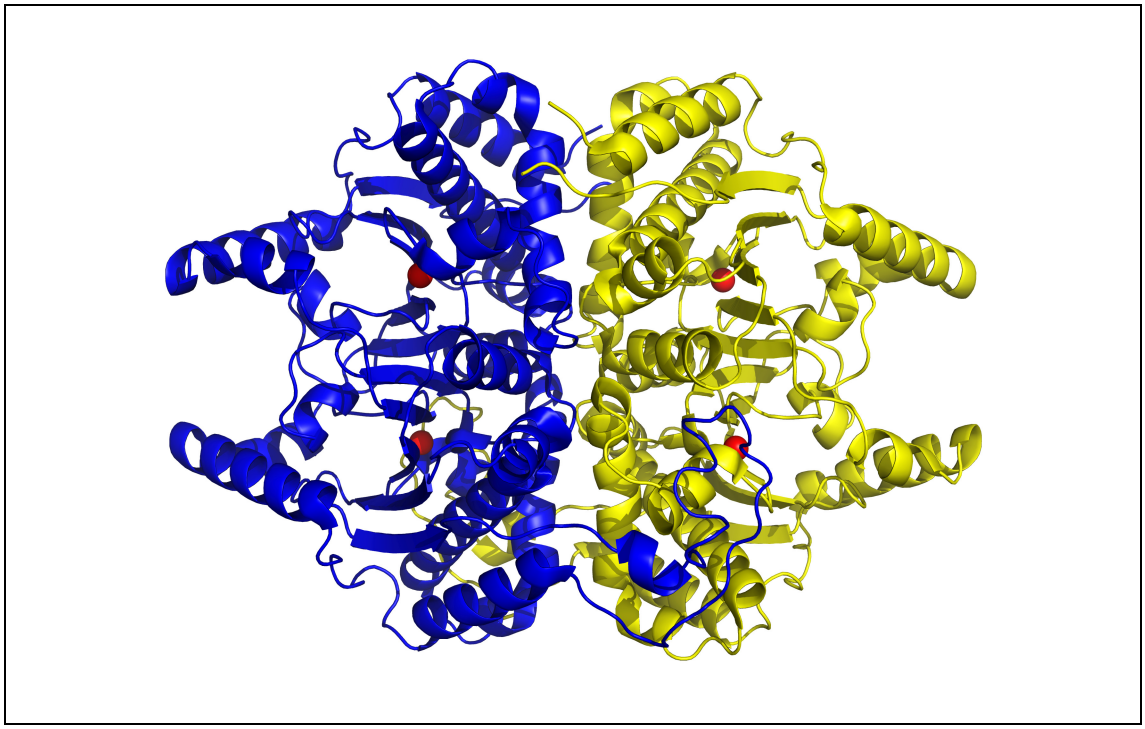


Figure 3 Three-dimensional structure of a β-class carbonic anhydrase.
Ribbon representation of carbonic anhydrase from algea Porfiridium purpureum with active site Zn^{2+} ions colored in red and individual monomeric subunits of the dimer colored in blue and yellow. The figure was prepared using the structure deposited in PDB under accession code 1DDZ (Mitsuhashi et al., 2000), and program PyMOL (DeLano, 2002).

2.1.3 The ancient γ-class

Another independently evolved CA class is the γ-class, first reported in 1994 (Alber & Ferry, 1994). According to phylogenetic analysis, this class has probably evolved 3.0 to 4.5 billion years ago (Smith *et al.*, 1999), preceding the evolution of α-class at 200-300 million years ago (Hewett-Emmett & Tashian, 1996, Jiang & Gupta, 1999).

The first γ-class enzyme to be characterized was "Cam" from the archaeon *Methanosarcina thermophila* (Alber & Ferry, 1994). Later, homologues of γ-class of CAs were found in various species from all three domains of life as reviewed in (Ferry, 2010). The structure of Cam from the archea domain is a homotrimer, the monomers of which adopt a left-handed parallel β-helical fold (Kisker *et al.*, 1996), see Figure 4. The

three active site metals are coordinated by histidines from adjacent monomers. In case of Cam, the enzyme is the most effective with a physiologically relevant Fe^{2+}, however it can function with zinc as well as some other transition metals, like cobalt (Alber *et al.*, 1999). The protein used for structure determination by (Kisker *et al.*, 1996) was a product of heterologous expression in *E. coli* and contained zinc in its active site.

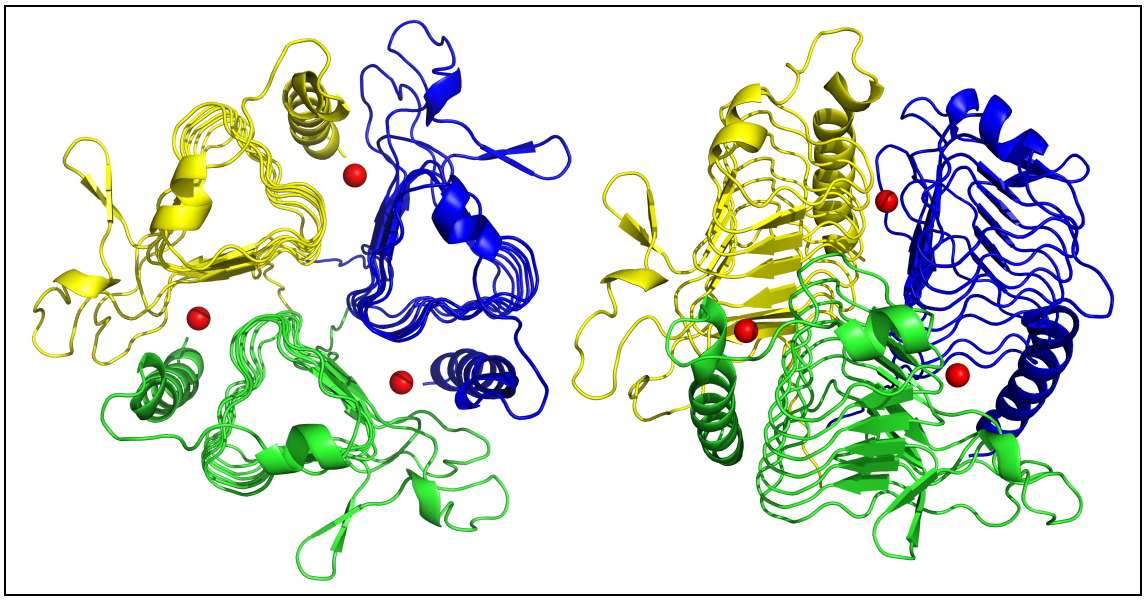


Figure 4 Three-dimensional structure of γ-class CA.

Ribbon representation of CA from Methanosarcina thermophila with active site zinc ions shown as red spheres and individual monomeric subunits colored in yellow, green and blue (on the left: top view, right: tilted/side view). The figure was prepared using the structure deposited in PDB under accession code 1THJ (Kisker et al., 1996) and program PyMOL (DeLano, 2002).

2.1.4 Other classes (δ and ζ)

Recently, two new significantly different classes of CAs were discovered. Only one member from each of these δ and ζ classes was characterized so far, both were isolated from marine diatoms (Park *et al.*, 2007). Carbonic anhydrase from *Thalassiosira weissflogii*, the sequence and preliminary characterization of which was published in (Lane *et al.*, 2005), is a member of the ζ class, and contains cadmium in its active site, hence it was named CDCA1 (cadmium containing carbonic anhydrase 1).

The concentration of zinc in ocean is so low that it limits the growth of phytoplankton like diatoms, which is the likely reason for CDCA1 utilizing cadmium in absence of zinc. Before this discovery, cadmium was considered only contaminating heavy metal in environment.

To summarize, carbonic anhydrases are expressed in nearly all types of organisms, and they form a group of repeatedly invented enzymes, which further demonstrates their physiological importance in nature.

2.2 Human carbonic anhydrases

Recent carbonic anhydrase research is undergoing rapid progress mainly in the biomedical area. The majority of over two thousand papers published in the CA field in the last five years discuss CA role as diagnostic and prognostic markers or CAs as therapy targets (reviewed in Pastorekova, 2009). Up to date, fifteen isozymes of human CA have been described. These isozymes can be divided according to various criteria, such as tissue distribution, expression level, catalytic activity and subcellular localization. List of catalytically active isozymes is given in Table 2.

Table 2 List of human catalytically active CA isozymes

Values of kinetic and inhibitory constants are from (Hilvo et al., 2008).

Isozyme	Molecular mass [kDa]	Subcellular localization	Typical tissue localization	k_{cat} [s^{-1}]	k_{cat}/K_M [M^{-1} s^{-1}]	K_I (acetazolamide) [nM]
CA I	29	cytosol	Erythrocytes, GI tract	2.0×10^5	5.0×10^7	250
CA II	29	cytosol	widely distributed	1.4×10^6	1.5×10^8	12
CA III	29	cytosol	Type I muscle	1.3×10^4	2.5×10^5	240 000
CA IV	35	extracellular GPI-anchored	GI tract, kidney, endothelium	1.1×10^6	5.1×10^7	74
CA VA	34.7	mitochondria	liver	2.9×10^5	2.9×10^7	63
CA VB	36.4	mitochondria	widely distributed	9.5×10^5	9.8×10^7	54
CA VI	39-42	secreted	saliva and milk	3.4×10^5	4.9×10^7	11
CA VII	29	cytosol	widely distributed	9.5×10^5	8.3×10^7	2.5
CA IX	54, 58	transmembrane	normal GI tract, several cancers	3.8×10^5	5.5×10^7	16
CA XII	44	transmembrane	kidney, certain cancers	4.2×10^5	3.5×10^7	5.7
CA XIII	29	cytosol	widely distributed	1.5×10^5	1.1×10^7	16
CA XIV	54	transmembrane	kidney, heart, skeletal muscle, brain	3.1×10^5	3.9×10^7	41

As documented in Table 2 and Figure 5, the twelve catalytically active human isozymes can be divided according to their subcellular localization to: cytoplasmic

(CA I, II, III, VII, XIII), mitochondrial (CA VA and CA VB), secreted (CA VI), and finally membrane associated isozymes with their catalytically active domain localized in the extracellular region (CA IV, CA IX, CA XII, and CA XIV). The remaining three (CA VIII, CA X, and CA XI) are catalytically inactive cytoplasmic isoforms with unknown physiological function (Chegwidden *et al.*, 2000, Supuran *et al.*, 2003), they are sometimes referred to CARPs (carbonic anhydrase related proteins).

Some of these isozymes are wide-spread among many different types of cells (CA II serving as a good example), while others are restricted to few tissues (CA I, III, VA, VI, VII, IX) (Pastorekova & Zavada, 2004).

Combination of these different properties creates a diversity allowing each isoform to fulfill a unique role in a specific physiological context (Pastorekova & Zavada, 2004).

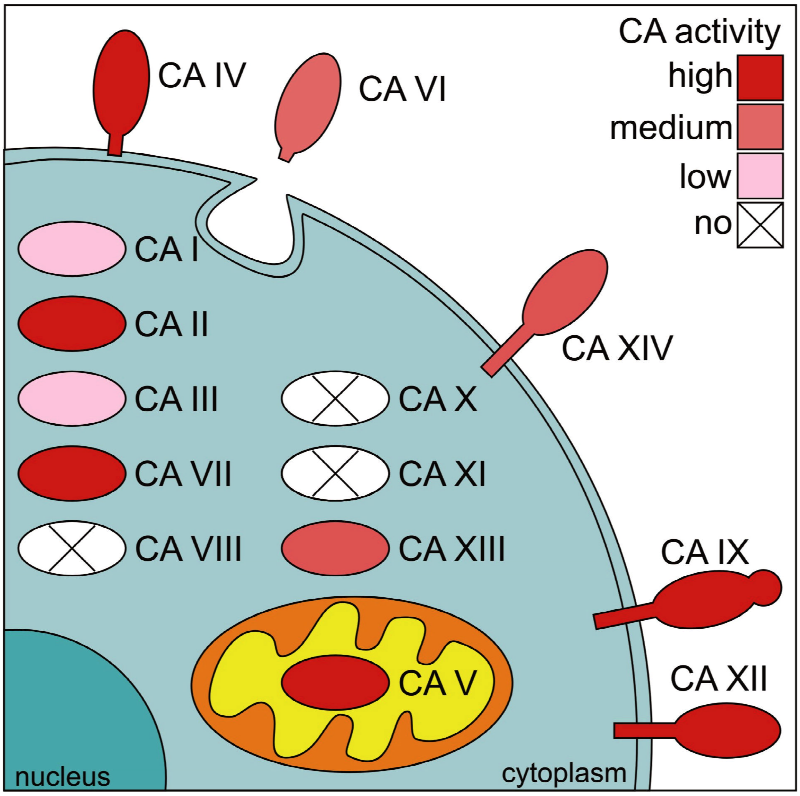


Figure 5 Enzymatic activity, subcellular localization and domain composition of human CAs.
CA IX is the only carbonic anhydrase with an N-terminal proteoglycan-like sequence, engaged in cell-to-cell adhesion. Figure was adapted from Pastorekova & Zavada, 2004.

2.3 Physiological role of human carbonic anhydrases

In addition to their role in CO_2 and bicarbonate transport in respiration, reported by Meldrum and Roughton in 1933 (Meldrum & Roughton, 1933b), human CA isozymes are involved in a broad range of physiological processes including pH homeostasis, ion transport, bone resorption, production of gastric acid, renal acidification, and lipogenesis (Sly & Hu, 1995). They are involved in various biochemical pathways where CO_2 or bicarbonate is required, such as gluconeogenesis, synthesis of certain amino acids (via pyruvate carboxylase), lipogenesis (via acetyl-CoA carboxylase), ureagenesis (via carbamoyl phosphate synthetase I), and others (Chegwidden *et al.*, 2000, Sly & Hu, 1995).

Several diseases are implicated to result from loss or deregulation of activity of certain isozymes, including glaucoma, osteopetrosis, oedema from heart and renal failure, neurological and neuromuscular disorders (Pastorekova *et al.*, 2004). However, despite many efforts, no consistent link to cancer had been found before the discovery of CA IX (Pastorekova *et al.*, 1992).

2.4 CAs as therapeutic targets

The clinical modulation of CA activity with small molecule inhibitors has been utilized in antiglaucoma, anticonvulsant and antiepileptic, antihypertensive, antithyroid, and hypoglycemic treatment regimes for several decades (Thiry *et al.*, 2007, Supuran, 2008, Poulsen, 2010). At least 25 clinically used drugs possess significant CA inhibitory properties (Supuran, 2008). Most of these drugs (e.g. acetazolamide, dorzolamide, brinzolamide) are small organic molecules with sulfonamide moiety, serving as carbonic anhydrase inhibitors (CAIs).

Of the fifteen isozymes, at least two-thirds (e.g. CA II, IV, V, VA, VB, VII, IX, XII, XIII, and XIV) constitute valid targets of development of novel anticancer (Pastorekova *et al.*, 2007, Thiry *et al.*, 2008, Poulsen, 2010), anti-obesity (Supuran *et al.*, 2008), antiglaucoma, and anticonvulsant drugs (Supuran *et al.*, 2003, Thiry *et al.*, 2007). However, due to the diffuse localization of CAs in many tissues and organs,

there is a need for development of highly selective and potent inhibitory compounds against specific isoforms (Gitto *et al.*, 2010). This strategy should lead to compounds without side-effects and improved therapeutic safety.

2.5 Carbonic anhydrase IX as a novel cancer therapy target

Human CA IX isozyme (initially named MN) was discovered in 1992 (Pastorekova *et al.*, 1992) using monoclonal antibody M75. This protein is a peculiar member of the human CA family. Under normal conditions, this membrane associated CA isoform is expressed on the surface of very limited number of tissues, namely in the gastric mucosa (Liao *et al.*, 1994) and in rapidly proliferating cells of small intestine (Saarnio *et al.*, 1998). On the other hand, this enzyme is ectopically overexpressed on the cell surface of a large number of solid tumors arising from originally CA IX negative tissues.

Solid tumors generally suffer from poor vasculature and hypoxia, which in turn leads to the activation of hypoxia-inducible factor (HIF) (Brahimi-Horn & Pouyssegur, 2009). Human CA IX tumor-related expression is determined by strong activation of CA9 gene transcription induced by HIF-1, which binds to hypoxia responsive element (HRE) localized in the CA9 promoter (Wykoff *et al.*, 2000). The *in vivo* expression pattern of CA IX clearly mirrors the distribution of hypoxic areas (Pastorekova & Zavada, 2004). Furthermore, CA IX overexpression is often associated with poor responsiveness to classical chemotherapy and radiotherapy.

Recent evidence strongly implicates that extracellular membrane bound CAs, namely CA IX and to some extent also CA XII are key regulatory molecules for countering acidosis during hypoxia (Chiche *et al.*, 2009, Swietach *et al.*, 2009, Poulsen, 2010). Elevated metabolism in tumor cells leads to increased acid production in comparison to healthy cells (Brahimi-Horn & Pouyssegur, 2009, Chiche *et al.*, 2009, Swietach *et al.*, 2007). It is known from a number of studies that tumor cells maintain normal intracellular pH_i and that they have evolved several mechanisms for expelling acids into the extracellular environment, thus contributing to lowering extracellular pH_e (Swietach *et al.*, 2007), which in turn favors tumor invasive behavior and development

(Martinez-Zaguilan *et al.*, 1996). Acidic extracellular milieu induces production of growth factors, induces genomic instability, perturbs cell-cell adhesion and facilitates tumor spread and metastasis (Stubbs *et al.*, 2000, Pastorekova & Zavada, 2004). In addition to lactic acid, the cell-generated CO_2 is also responsible for rapid removal of acid equivalents from tumor cells (Swietach *et al.*, 2008, Swietach *et al.*, 2009, Chiche *et al.*, 2009). Provided that there is a sufficient outward gradient, the cell derived CO_2 can freely diffuse across the cellular membrane to the extracellular space, where in poorly vascularized hypoxic tumor tissue it is hydrated to HCO_3^- and H^+ by CA IX (and/or CA XII) (Swietach *et al.*, 2009, Poulsen, 2010). A schematic representation demonstrating the prosurvival role of carbonic anhydrases in hypoxic tumors through maintaining normal intracellular pH (pH_i) is shown in Figure 6. As a net effect, the CA IX/XII trap the acid extracellularly, thus lower the pH_e and maintain normal pH_i with HCO_3^- recycled back to the cell.

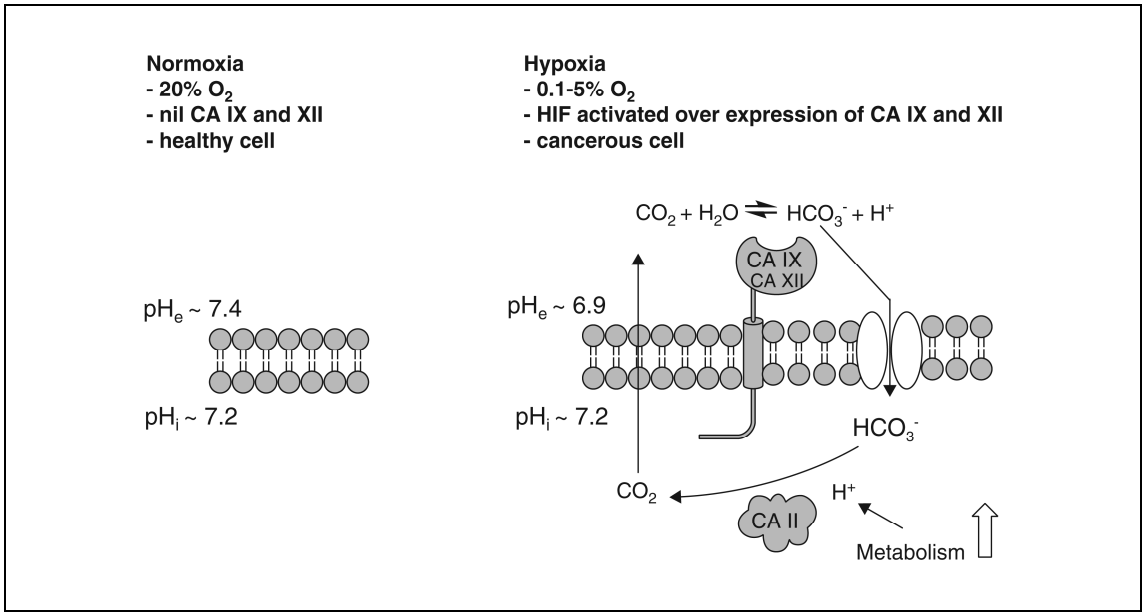


Figure 6 Prosurvival role of carbonic anhydrases IX and XII in hypoxic tumors.

Membrane bound extracellular carbonic anhydrases CA IX and CA XII help tumor cells in maintaining normal intracellular pH. Figure was adapted from Poulsen, 2010.

2.5.1 Molecular properties of CA IX

The analysis of genomic structure of CA IX (Opavsky *et al.*, 1996) revealed another interesting feature of this tumor-associated CA isoform, which is the presence of an extra proteoglycan-like sequence linked to the catalytic domain. The CA9 gene codes for 459 amino acid (aa) multidomain protein with 37 aa long signal peptide (SP; cleaved off during posttranslational processing) 75 aa long extracellular N-terminal proteoglycan-like (PG) sequence (with similarity to keratan sulfate-binding domain of a large proteoglycan aggrecan) followed by 302 aa long catalytic domain (CA) linked to 21 aa long helical transmembrane segment (TM) and 24 aa C-terminal intracellular tail (IC). The schematic representation of CA IX domain arrangement is depicted in Figure 7.

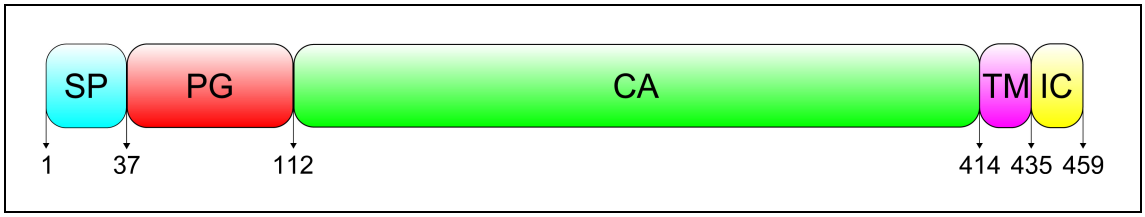


Figure 7 Schematic representation of CA IX domain composition.
The N-terminal proteoglycan-like (PG) sequence is unique for CA IX isozyme, it is followed by carbonic anhydrase catalytic domain (CA), short transmembrane segment (TM) and intracellular tail (IC).

The catalytic domain of CA IX displays significant sequence similarity with other human extracellular CAs, namely with recently discovered membrane associated isozyme CA XIV (Hilvo *et al.*, 2007) (sequence identity 42 %), secreted isozyme CA VI (sequence identity 39 %), and another membrane associated isozyme CA XII (also 39 % sequence identity).

To summarize, human CAs are studied in relationship to several human disorders including glaucoma, epilepsy, obesity, and finally most attention is paid to the relationship of CAs to cancer (Pastorekova & Zavada, 2004). A significant shift from early descriptive studies on distribution of CAs in various tissues towards correlation and functional studies (Pastorekova & Zavada, 2004) elucidating the roles and regulation of CAs in different physiological and pathophysiological situations is apparent (Pastorekova, 2009).

3 AIMS OF THE THESIS

The presented work is a part of an ongoing project focusing on study of human carbonic anhydrase isozymes and their role in human physiology and pathophysiology. The study of CA IX as a cancer related isozyme of the CA family started shortly after its discovery by Dr. Jan Závada and his co-workers and is still in progress. Development of novel compounds with high inhibitory potency and selectivity to human CA isozymes, namely to CA IX represents an attractive strategy to obtain pharmacological tools, while avoiding possible side effects and improving therapeutic safety.

The aims of this thesis follow two lines of studies focused on (1) obtaining structural information on CA isozymes which would be utilizable in structure-based drug design and (2) understanding the recognition of CA IX by specific antibodies.

To obtain structural information for structure-based drug design, following specific aims were proposed:

- Production of recombinant CA IX in *E. coli* or eukaryotic expression system. Development of purification protocol to obtain active CA IX in amount and quality suitable for crystallization trials and enzymatic activity assay.
- Preparation and crystallization of complexes of selected small molecule inhibitors with human carbonic anhydrase isozymes CA II and CA IX. Diffraction measurements with crystals from successful crystallization trials. Structure determination and analysis.

To understand the recognition of CA IX by a specific antibodies the following specific aim was proposed:

- Determination of crystal structure of fragment of specific antibody recognizing proteoglycan –like domain of CA IX by X-ray crystallography.

4 MATERIAL AND METHODS

4.1 Molecular cloning – construction of CA IX expression plasmids

Standard procedures of molecular cloning (plasmid DNA isolation, vector and insert preparation, ligation of DNA fragments and bacterial transformation) as given in Sambrook manual (Sambrook & Russell, 2001) were used. Reaction conditions for digests with restriction endonucleases, ligation reaction, site-directed mutagenesis, PCR etc. were according to instructions of enzyme and/or kit suppliers.

DNA coding for CA IX extracellular domains, PG+CA or CA, was obtained by PCR using full-length cDNA of CA IX as a template (kindly provided by Dr. Jan Závada) and suitable pairs of primers that introduced restriction sites into DNA fragments, allowing insertion into expression vectors.

Typically, the PCR reaction mixture contained in total volume of 50 µl about 20 ng template DNA, suitable pair of primers (final concentration 0.4 µM each), dNTPs (final concentration 0.2 mM), 5 µl 10 × concentrated PCR buffer (supplied by the polymerase manufacturer), and 0.5 µl *Pfu* DNA polymerase (Fermentas, 2.5 U/µl). The reaction began with initial denaturation step (94 °C for 3 min), followed by 30 – 35 amplification cycles (40 s at 94 °C, 1 min at 58 °C, 2 min at 72 °C), and ended with a final extension step (7 min at 72 °C). The resulting PCR fragment was purified by agarose gel electrophoresis (1 % w/v) and, after cloning into pUC18 vector previously opened with Sma I, sequence verified (at the core facility of the Institute of Microbiology, v.v.i., AS CR, Dr. J. Felsberg). The exact primer annealing temperature was adjusted in each case according to the primer sequences and the length of extension time according to the length of the amplified sequence.

In the course of this work, several constructs were prepared, first for bacterial expression, cytoplasmic as well as periplasmic, then for expression in S2 insect cells. The outline of their construction will be given below.

4.1.1 Plasmids for bacterial expression

In the PCR reactions set up for obtaining the coding sequences PG+CA and CA, the full-length CA IX cDNA in pBluescript as a template, and the following primers were used:

Table 3 Oligonucleotide primers used in PCR amplification of PG+CA and CA sequences.
Incorporated restriction sites are underlined.

N-primer PG+CA	5'-AGGCCATATGGGGGAAGATGACCC *NdeI*
N-primer CA	5'-GGATCATATGAGTCATTGGCGCTATGGA *NdeI*
C-primer (same for both fragments)	5'-CTCGAATTCAGAAGGAGGCCTCAATCAC *EcoRI*

N-primers introduced NdeI restriction site (CATATG, underlined), whereas C- primer brought in the stop codon and EcoRI site (GAATTC, underlined).

The PCR products were cloned between the NdeI and EcoRI sites of the pET-like vector pT7Q10, which contains inducible T7 promoter for heterologous expression, and genes for $lacI^q$ repressor and ampicillin resistance.

For periplasmic expression, both CA IX coding sequences, PG+CA and CA, were mutagenized to change their N-terminal NdeI sites to NcoI, and inserted as NcoI-EcoRI fragments into pET-22b vector (Novagen). The QuickChange site-directed mutagenesis kit (Stratagene, La Jolla, CA) was used for PG+CA insert with mutagenic primers:

(5'-GATATAACCATGGGGGAAGAT) and
(5'- ATCTTCCCCCATGGTTATATC).

For CA domain alone, PCR was performed using primers:

(5'-ATACCCATGGGTCATTGGCGCTATGGAG) and
(5'-CTCGAATTCAGAAGGAGGCCTCAATCAC).

Restriction sites for NcoI (CCATGG) and EcoRI (GAATTC) are underlined in primer sequences.

PG-CA insert was also cloned as NdeI-EcoRI fragment into the vector pT7Q10 coding for N-terminal $(His)_6$ tag (this vector was available in the laboratory).

4.1.2 Plasmids for expression in *Drosophila* S2 cells

The expression vector pMT/BiP/V5-HisA (Invitrogen, kindly provided by Dr. Šácha from Institute of Organic Chemistry and Biochemistry, v.v.i., AS CR) was first modified (cut with BglII, treated with Klenow polymerase to fill-in the ends, and religated) to achieve a "+1" shift of the reading frame at the adjacent NcoI site required for in-frame cloning of CA IX inserts. The NcoI-EcoRI fragments corresponding to PG+CA and CA domains (see above, periplasmic expression) were then inserted into this modified expression vector pMT/BiP/V5-HisA. The final constructs, designated pPG-CA and pCA were used in the first series of expression experiments. These, as well as the subfragments PG+CA and CA, also served as starting material for additional constructs comprising various N- and C-terminal tags, mutagenesis of a glycosylation signal within the CA domain (NQT to DQT), and others as detailed below.

4.1.2.1 Mutagenesis to remove glycosylation signal

The Asn residue in a triplet NQT was mutagenized to Asp using QuickChange site-directed mutagenesis kit (Stratagene, La Jolla, CA) with mutagenic primers:

NDU (5'- GGACTGTGTTTGACCAGACAGTGATG)

and NDL (5'- CATCACTGTCTGGTCAAACACAGTCC).

The Asp codon GAC in the sense oligonucleotide is underlined.

4.1.2.2 N-terminal His-tag with TEV cleavage site and C-terminal extension

The His+TEV fragment was assembled by PCR from four partly overlapping primers, designated [1], [2], [3], [4] (see below) and the product (77bp, containing His_6 - $(GlySer)_3$ - TEV site - BspEI-EcoRV) was, following gel-purification and cutting with BglII and XhoI, cloned into the appropriately cut vector pMT/BiP/V5-HisA.

```
      BglII         1 ->                              2 ->          BspEI            XhoI
5'  ttaAGATCTCACCACCATCATCACCATGGCAGCGGCAGCGGCAGCGAGAACCTGTACTTCCAGTCCGGAgcgGATATCaaCTCGAGggc  3'
3'                              CCGTCGCCGTCGCCGTCGCTCTTGGACATGAAGGTC        cCTATAGttGAGCTCccg  5'
                                                  <- 3                          <- 4
```

The sequence was verified using the vector specific primer (MT forward). Into this modified vector the PG+CA and CA inserts were cloned using BspEI and XhoI restriction sites. The inserts were obtained by PCR using full-length CA IX cDNA in pBluescript as template and the following primers, which introduced the needed restriction sites (underlined) BspEI and XhoI, and a stop codon upstream of XhoI site:

5'- GGCTCTTCCGGAGAAGATGACCCACTG for PG+CA N-terminus

5'- GAAGGGGATTCCGGAAGTCATTGGCGC for CA N-terminus

5'- AACCAGGGCCTCGAGTCAACCAGCAGCCAG for both inserts extended C-termini

The final constructs, designated pHisTEV-PG-CA-ext and pHisTEV-CA-ext thus coded for PG+CA and CA domains with His-tag and TEV cleavage site at the N-terminus, and additional amino residues PAGVDSSPRAAEPVQLNS**C**LAAG at the C-terminus.

4.1.2.3 N-terminal His-tag with TEV cleavage site without C-terminal extension

To obtain expression plasmids for "original" CA domain (i.e. without C-extensions) but with His-tag and TEV site, the above plasmid pHisTEV-CA-ext was used as a vector for two EcoRV-XbaI inserts from plasmids coding for CA domain either with glycosylation signal NQT or CA domain in which NQT was mutated to DQT. The resultant plasmids coded for (non)glycosylated versions of "short" CA domain with His-tag + TEV site at the N-terminus (pHisTEV-CA/D and pHisTEV-CA/N).

4.1.2.4 N-terminal Avi-tag with TEV cleavage site

Constructs in which the N-terminal His-tag was replaced by Avi-tag were prepared in two steps. First, the sequence of Avi-tag + TEV was obtained by PCR using plasmid pMTBiPAviTEV (kindly provided by Dr. Šácha from Institute of Organic Chemistry and Biochemistry, v.v.i., AS CR) as a template, and primers amplifying the region upstream of Avi-tag (from EagI site in the vector backbone), the tag+TEV and introducing BspEI site downstream of TEV site to facilitate insertion of the PCR product as EagI-BspEI fragment into the appropriately cut vector pHisTEV-CA/D in the second step, yielding the plasmid pAviTEV-CA/D. The primers were as follows:

5'- CACTCGAATTTGGAGCCGGCCGGCGTGTGC (EagI underlined)

5'- GGTCCGGAAGCTGCATTGGACTGGAAGTACAGGTTCTC (BspEI underlined).

4.1.2.5 C-terminal Avi-tag

The C-terminal Avi-tag was introduced at the C-terminus of the CA domain using a PCR technique SOE (Splicing by Overlap Extension):

a -----> b ----->

<----- c <----- **d**

where PCR products [ac] and [bd], obtained in separate reactions, (and where primers b and c are complementary) were isolated and mixed for final PCR reaction with primers a and d, yielding a fragment [ad].

The primers were as follows:

a (5‘- CCGGTGGATATCCGCCCCCAGCTC**);**

b (5‘- GAGGCCTCCTTCTCCGGACTGAACGACATCTTCGAG);

c (5‘- GAAGATGTCGTTCAGTCCGGAGAAGGAGGCCTCAAT);

d (5‘- CCCCGAATTCACTCGTGCCACTCGAT)

In this scheme, PCR [ac] amplified from pAviTEV-CA/D template the CA domain

starting at the EcoRV site (aa residue 166) to carboxy terminus AlaSerPhe and added BspEI site (SerGly) and several amino acid residues of the avi-tag amino terminus. In the PCR [bd] from the same template, pAviTEV-CA/D, primer d introduced a stop codon and EcoRI site at the 3’ end of Avi-tag, whereas the b primer, complementary to c, placed the last four C-terminal amino acid residues of the CA domain upstream of the Avi-tag. The final PCR reaction [ad] then yielded a fragment where CA domain (starting at the EcoRV site) is spliced with Avi-Tag at its C-terminus, allowing a replacement of EcoRV-XbaI fragment in pPG-CA or pCA, resulting in pPG-CA/D-Avi and p-CA/D-Avi.

By replacing NcoI-BlpI part in p-CA/D-Avi with inserts NcoI-BlpI from p-PG-CA or p-CA, constructs coding for glycosylated versions, p-PG-CA/N-Avi (p947) and p-CA/N-Avi were obtained.

4.1.2.6 Mutagenesis Cys(41)Ser

Mutagenesis was performed on a pUC18-based plasmid carrying isolated CA/N domain using a QuickChange site-directed mutagenesis kit (Stratagene, La Jolla, CA) with mutagenic primers:

CSU (5'-GCCGCCTTCTCGCCGGCCCTG)

and CSL (5'-CAGGGCCGGCGAGAAGGCGGC).

The Ser codon TCG in the sense oligonucleotide CSU is underlined. The fragment EcoRV-EcoRI containing the Cys(41)Ser mutation was then used to replace the wt region of the construct pAviTEV-CA/D, yielding pAviTEV-CA(S41)/N.

4.1.2.7 CA domain (Ser41) extensions

The CA domain, Cys>Ser mutation notwithstanding, still differs from the one leading to successful crystallization as described (Alterio *et al.*, 2009). Therefore, N-terminal DQ and C-terminal P residues were be added. First, a 406 bp EagI-EcoRV fragment from pAviTEV-CA(S41)/N was subcloned into pBluescript for site-directed mutagenesis, in which SerGly preceding N-terminus of the CA domain was replaced by AspGln (in the process, BspEI site was lost); after mutation, the fragment was returned to pAviTEV-CA(S41)/N, yielding an intermediate pAviTEV-CA(DQ/S41)/N.

The mutagenic primers were:

DQU (5'-TCCAATGCAGCTGACCAAAGTCATTGGCGC) and

DQL (5'-GCGCCAATGACTTTGGTCAGCTGCATTGGA).

The AspGln codons GACCAA in the sense oligonucleotide DQU are underlined.

To add C-terminal Pro residue, the pUC18-based plasmid carrying CA(S41)/N domain was used as a template for PCR, in which the region (~ 700bp) between EcoRV site and C-terminus was amplified using primers:

NPRO (5'-CCGGTGGATATCCGCCCCCAGCTC) and

CPRO (5'-GGTCTAGACTTAAGGGAAGGAGGCCTCAATC);

the CPRO primer introduced into the fragment C-terminal Pro residue, stop codon and an XbaI site. Finally, the PCR product EcoRV-XbaI was used to replace EcoRV-XbaI part of pAviTEV-CA(DQ/S41)/N, yielding a complete construct pAviTEV-CA(DQ/S41/EASFP)/N.

4.2 Expression of CA IX constructs in *E. coli*

4.2.1 Bacterial media

Luria-Bertani (LB) medium (10 g/l Bacto Trypton (Difco, USA), 5 g/l yeast extract (Difco, USA), 5 g/l sodium chloride, pH 7.4 (adjusted with Tris HCl)) (Sambrook & Russell, 2001) supplemented with appropriate antibiotics was used for small-scale bacterial cultures.

Large-scale cultivations were grown in the same LB medium supplemented with 3 % glycerol.

2YT medium (16 g/l Bacto Trypton, 10 g/l yeast extract, 5 g/l sodium chloride) was used in some pilot expression experiments.

LB agar plates contained LB medium, agar (Difco, USA) (15 g/l), appropriate antibiotics and 1 % glucose.

Depending on the combination of bacterial host strain and used plasmid, following concentrations of antibiotics were used: ampicillin (100 mg/l), kanamycin (25 mg/l), tetracycline (15 mg/l), and chloramphenicol (50 mg/l)

4.2.2 Host strains

***Escherichia coli* DH5 alpha**: *[F-, Φ80dlacZΔM15, endA1, recA1, hsdR17 (rk-,mk+), supE44, thi-1, gyrA96, relA1, Δ (lacZYA-argF)U169, λ-]* was used as the host during the cloning procedure.

***Escherichia coli* BL21 (DE3)**: F^- *ompT [lon] hsd*S_B (r_B^- m_B^-) an *E. coli* B strain with DE3, a λ prophage carrying the T7 RNA polymerase gene under inducible *lacUV* promoter (Studier & Moffatt, 1986), was tested as the first host for expression of the extracellular domains of CA IX.

***Escherichia coli* Origami B (DE3)**: F^- *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm lacY1 ahpC* (DE3) *gor522*:: Tn*10 trxB* (Kan^R, Tet^R) host strain is derived from a *lacZY* mutant of BL21, that has mutations in both the thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes. These mutations enhance disulfide bond formation in the cytoplasm. The *trxB* and *gor* mutations are selectable on kanamycin and tetracycline, respectively. This strain is suitable for use with plasmids carrying ampicillin resistance marker *bla*.

***Escherichia coli* Rosetta-gami B (DE3)**: F^- *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm lacY1 ahpC* (DE3) *gor522*::Tn*10 trxB* pRARE (Cam^R, Kan^R, Tet^R). This host strain is an Origami B derivative that combines enhanced disulfide bond formation resulting from *trxB*/*gor* mutations with enhanced expression of eukaryotic proteins that contain codons rarely used in *E. coli*. The tRNA genes for AGG, AGA, AUA, CUA, CCC, GGA are driven by their native promoters and are located on a chloramphenicol resistant plasmid. The *trxB* and *gor* mutations are selectable on kanamycin and tetracycline, respectively. Therefore, this strain is also suitable for use with plasmids carrying ampicillin resistance marker *bla*.

4.2.3 Cultivation conditions

To achieve expression of CA IX in *E. coli* various cultivation conditions in combination with above listed host strains were tested. Cultivation temperature, time, and concentration of inducer were the main variables.

A culture of *E. coli* cells was transformed with expression plasmid and plated on LB agar plates with appropriate antibiotics. Next day, the bacterial colonies were washed from the plates using fresh LB medium containing appropriate antibiotics and 3 % glycerol (typically 2 × 5 ml of medium were used for one plate). The optical density of thus prepared inoculum was measured at 550 nm and the inoculum was used

for setting up bacterial cultivation cultures in 500 ml media (for composition see section 4.2.1) at OD_{550nm} starting approximately at 0.1 value. The cultivations were performed in 2 l Erlenmayer flasks on a rotary shaker Gallenkamp (250 rpm) at 37 °C. Growth of the bacterial culture was monitored every hour by measuring optical density (OD) at 550 nm. When the OD_{550nm} reached value between 1.0 and 1.3, the expression was induced by adding ETG (ethylthiogalactoside) to a final concentration of 1 mM to the culture.

After the induction, the bacterial culture was grown under the same conditions for additional 3 hours. In some experiments the temperature of cultivation was lowered to 18 or 20 °C about 30 min before induction and kept at this level for the rest of cultivation. Finally, the cells were harvested by centrifugation (5000 rpm, 15 min, 4°C, Sorvall Instruments RC-3B, rotor H 400) and the cell paste was stored frozen at -25 °C.

4.3 Expression of CA IX constructs in S2 cells

Schneider S2 cells were used for eukaryotic heterologous expression of rCA IX constructs. This cell line, derived from a primary culture of late stage (20-24 hours old) *Drosophila melanogaster* embryos (Schneider, 1972), was supplied as a part of the *Drosophila* Expression System (Invitrogen).

4.3.1 Insect cell cultivation

S2 cells were cultured in SF900II medium (Gibco, Rockville, MD, USA) supplemented with 10% fetal bovine serum (FBS; Sigma) at 22-24°C without CO_2. The cells grew as a loose, semi-adherent monolayer in tissue culture flasks and in suspension in spinner flasks.

4.3.2 Transfection of insect cells and stable cell line generation

Schneider S2 cells were seeded in a 35-mm dish and grown until they reached a density of 2-4 × 10^6 cells/ml. The cells were co-transfected with 19 μg of expression

vector and 1 μg of a pCoBlast selection vector (Invitrogen) using a kit for calcium phosphate transfection (Invitrogen). The calcium phosphate solution was removed 16 hours after the transfection and fresh SF900II medium supplemented with 10% FBS was added (a complete medium). The cells were grown for additional 2 days and then the medium was replaced with the complete medium containing 300 μg/ml Blasticidine (Invitrogen). This selection medium was changed every 4 days. Extensive cell death of non-transfected cells was evident after about 1 week and cells resistant to Hygromycin B started grow out 3 - 4 weeks posttransfection.

4.3.3 Recombinant CA IX large scale expression

The hygromycin-resistant cells were seeded in 100 ml of serum-free expression medium [SF900II medium supplemented with 1 ml chemically defined Lipid Concentrate (Gibco) and 2 ml Yeastolate Ultrafiltrate (Gibco)] in a spinner flask (250 cm^3; Bellco Glass Inc., Vineland, NJ, USA) at density 2×10^6 cells/ml. The spinner was incubated at 24 °C with a constant stirring rate 100 rpm until the cells reached a density of 10×10^6 cells/ml. At this point, the cell suspension was transferred into a larger spinner flask (3000 cm^3; Bellco), 400 ml expression medium were added, and the cells were grown to a density of 8×10^6 cells/ml under the same conditions. Induction of protein expression was achieved by adding 500 μM $CuSO_4$ (Sigma). The stirring rate was increased to 130 rpm and incubation continued for additional 5-7 days. Every two days 5 ml of 20% D-(+)-glucose (Sigma) and 2 ml of 200 mM L-glutamine (Sigma) were added to the cell suspension. The cells were harvested by centrifugation at 1,600 × g for 10 min followed by the second centrifugation step at 3400 × g for 15 min. The conditioned medium was frozen and stored at – 70 °C until further use.

4.4 Purification of recombinant CA IX

The purification protocol was adapted according to the nature of the expressed recombinant construct and to the presence of affinity tags.

4.4.1 Sulfonamide agarose affinity chromatography

Affinity chromatography purification of CA IX constructs was performed using chromatographic resin p-aminomethyl-benzenesulfonamide agarose (Sigma) and the purification protocol was based on previously published procedure (Falkbring *et al.*, 1972). Unless otherwise stated, all procedures were done 4 °C. The sample was dialyzed against wash buffer (50 mM Tris-SO_4 pH 7.4; 150 mM Na_2SO_4). The affinity resin was pre-equilibrated in the same buffer, added to the dialyzed sample (typically 2 ml of resin per 100 ml dialyzed media were used), the mixture was gently stirred for 16 h, and finally the mixture was passed through a sintered glass filter. Retained affinity gel was subsequently extensively washed with about 30 – 40 column volumes of the wash buffer, to remove any nonspecifically bound protein. The elution of target protein was performed with the same wash buffer containing 0.1 mM acetazolamide (Sigma) or with a gradient of 0 – 1 M NaCl. Binding and elution profiles were monitored by absorbance at 280 nm.

4.4.2 Ion exchange chromatography

Purification of recombinant CA IX constructs, which did not contain any affinity purification tags, was in several cases performed using ion exchange chromatography on DEAE sepharose (DEAE Sephacel, Pharmacia, weak anionic exchanger) as the first purification step. The column (diameter 55 mm, height 50 mm) was equilibrated in 15 mM Tris-Cl pH 7.4 and the samples were dialyzed and applied on column using the same buffer. Unbound proteins were washed away with the 10 – 15 column volumes of the above buffer. Elution was performed in the same pH by 30 min gradient of increasing NaCl concentration from 0 to 1 M at flow rate 1 ml/min.

Ion exchange chromatography on MonoQ GL 5/50 column (GE Healthcare) was used as the second purification step of CA IX constructs previously purified by affinity chromatography. All recombinant CA IX constructs had pI $\leq$ 6, so 15 mM Tris-Cl pH 7.4 or 15 mM HEPES pH 7.0 buffers were used for binding and equilibration. The elution of bound proteins was performed using gradient of increasing NaCl

concentration, according to the following scheme: linear gradient 0 – 400 mM NaCl (0 – 20 ml); 400 – 1000 mM NaCl (20 – 25 ml). The flow rate (1 ml/min) and max pressure (5 MPa) were adjusted according to the manufacturer's recommendation.

4.4.3 Purification of His-tagged proteins on Ni-affinity resin

The affinity chromatography on Ni-CAM resin (Sigma) was used as the first purification step of CA IX constructs containing N-terminal $(His)_6$-tag. This method as well as other affinity and ion exchange chromatography techniques also had the benefit of concentrating the protein at least by a factor of 10 in comparison to original concentration in the harvested medium or cell supernatant.

Unless otherwise stated, all procedures were done 4 °C. The same buffer (50m M NaH_2PO_4 pH 8,0; 300 mM NaCl) was used for initial dialysis of samples, column equilibration, and washing of unbound proteins. The dialyzed samples were applied on pre-equilibrated column at flow rate 0.5 ml/min. Unbound proteins were washed away with the 10 – 15 column volumes of the above buffer. Elution was performed with gradient of 0 – 250 mM imidazole. Due to imidazole absorbance at 280 nm, eluted fractions were analyzed on SDS-PAGE. The volume of eluted fractions was equivalent to the bed volume of the affinity resin.

4.4.4 Expression and purification of Avi-tagged proteins on Streptavidin mutein matrix

Avi-tag is a 15 aa long sequence (GLNDIFEAQ**K**IEWHE), containing one lysine residue (marked in bold) which is specifically biotinylated on its ε-amino group by BirA ligase (*E. coli* biotin-(acetyl-CoA-carboxytransferase), EC 6.3.4.15) (Beckett *et al.*, 1999). *In vivo* biotinylated protein fusions with the above avi-tag sequence were previously expressed in insect cells (Duffy *et al.*, 1998, Yang *et al.*, 2004) and mammalian cells (de Boer *et al.*, 2003). Procedures used in this work were mainly based on protocols developed and optimized by Jan Tykvart in his diploma thesis (Tykvart, 2009).

A strain of S2 cells previously transfected with a vector coding for BirA ligase fused to *Drosophila* N-terminal signal peptide *BiP*, ensuring transport to endoplasmic reticulum (ER) and C-terminal retention signal to ER (highly conserved KDEL sequence) was kindly provided by Dr. Pavel Šácha and MSc. Jan Tykvart from Institute of Organic Chemistry and Biochemistry, v.v.i. AS CR.

Additional transfection with recombinant CA IX vector was performed as described above, and specifically *in vivo* biotinylated protein was expressed and secreted to serum free medium.

The purification of avi-tagged biotinylated proteins was performed on Streptavidin Mutein Matrix (Roche) according to manufacturer's protocol. Briefly, the harvested medium from S2 cells was mixed with equilibration buffer (450 mM NaCl, 300 mM Tris-HCl (pH 7.2)) in ratio 2:1, and applied on pre-equilibrated Streptavidin Mutein Matrix (Roche) resin in wash buffer (150 mM NaCl, 100 mM Tris-HCl (pH 7.2)). Typically, the binding was performed by adding 2 ml of resin to 500-600 ml of media in equilibration buffer and leaving overnight on a rocker (20 rpm). All procedures were done 4 °C. Next day, the resin was transferred to a plastic disposable column (Pierce) and washed extensively with wash buffer (typically 10 column volumes) and the elution was performed with elution buffer (wash buffer containing 2 mM biotin, Sigma)

Biotinylated proteins were detected on Western blots via NeutAvidin-HRP conjugate (Thermo scientific, USA).

4.5 Carbonic anhydrase activity assay

Several methods for measuring carbonic anhydrase activity are available, unfortunately the most sensitive and accurate ones require specialized equipment (stopped-flow instrument or dedicated mass spectrometer) and/or use of compounds prepared from non-common and sometimes radioactive isotopes e.g. (Stemler, 1993).

In this work, one of the oldest, simple and still commonly used method was chosen. The procedure is based on Maren's endpoint colorimetric technique (Maren &

Ellison, 1967) with some modifications detailed in (Brion *et al.*, 1988). The principle of this method is based on rate of acidification of solution monitored by change in acid-base indicator (phenol red) color.

The reaction was carried out in a glass tube kept in ice-water bath (0 °C). An 18G needle connected to a tubing with CO_2 flowing at 50 ml/min was inserted into the glass tube. Reaction buffer was prepared fresh by adding 30 ml of stock 1 M 1 M Na_2CO_3, 20.6 ml 1 M $NaHCO_3$, and water to 100 ml (final pH 9 – 10).

As a first step, 800 µl of phenol red solution (12.5 mg/l phenolphtalein in 2.6 mM 1 M $NaHCO_3$,) and 100 µl of sample and/or water were mixed in the reaction vessel and continuously bubbled with CO_2 until the color turned yellow. Then, 100 µl of the reaction buffer were rapidly added to this solution and timing was started. After addition of this basic buffer the solution immediately turned purple, and timing was stopped when the indicator color turned back to yellow. Another tube, previously acidified, was used for visual determination of the endpoint.

The endpoint in blank samples was achieved in 58-63 s. When active enzyme was present the times were proportionally shorter. One enzyme unit (EU) of CA activity was defined as the amount of enzyme needed to halve the time measured for the control sample (Brion *et al.*, 1988). To obtain reasonable time readings (measurement times above 10 s) the original samples were appropriately diluted, and the dilution factor was used in calculating the activity in the original sample.

Specific CA activity was calculated using the formula:

$$\text{CA (EU/mg protein)} = \frac{\log (t_B/t_S)}{m_{prot} \log 2}$$

where t_B and t_S are times measured for blank and active samples, respectively, m_{prot} is the amount of protein in milligrams present in the volume of sample used for the particular measurement and $\log 2 \approx 0.301$.

4.6 Preparation of Fab M75 for crystallization

IgG 2b M75 was expressed in tissue cultures of hybridoma cell line VU-M75 (deposited in ATCC under code HB 11128). Tissue culture media was applied to Protein A Sepharose (BioRad) column, and IgG fraction was eluted with 0.1 M sodium citrate, pH 3.0. The eluted fractions were pooled, dialyzed against 10 mM sodium phosphate, pH 8.0, loaded onto a DEAE Sephadex column equilibrated with the same buffer, and eluted with a 10–200 mM sodium phosphate, pH 8.0, gradient. Eluted IgG was concentrated to 5 mg/ml and dialyzed against 100 mM potassium phosphate, pH 7.2. The samples were then supplemented with 0.13 % 2-mercaptoethanol, and 0.9 mM EDTA, and papain (Worthington) was added to a final concentration of 2 μg/ml. The course of reaction was monitored by gel filtration on Superdex 200 HR (GE Healthcare), and the reaction was stopped after 3 h by addition of iodoacetamide (final concentration 5.4 mM). Undigested IgG as well as Fc fragments were removed using Protein A Sepharose (BioRad).

4.7 Protein crystallography

4.7.1 Protein crystallization

Crystallization trials were performed with the sparse matrix method (Jancarik & Kim, 1991) as well as with other crystallization screens designed on the basis of results of structural genomics consortia (e.g. JCSG+ screen, Qiagen). Vapor diffusion method of protein crystallization was used for screening and optimization of crystallization conditions at temperature of 18 ºC.

The vapor diffusion method was typically performed in 24-well Linbro plates (Hampton Reserach) or EasyXtal tools (Quiagen). For the sitting drop technique the 4 μl droplet containing mixture of protein solution with precipitant (mixed typically in 1:1 ratio) was placed onto the holder Micro-Bridge (Hampton Research) inserted into the well filled with 1 ml of precipitant. Wells were then sealed with clear sealing tape. When the hanging drop technique of vapor diffusion method (Ducruix & Giegé, 2000) was used, the droplet of the protein sample mixed with precipitant solution (typically in

the ratio of 2:1) was placed on a siliconized glass cover slide inverted over a reservoir of 1 ml of the precipitating solution.

For improving crystal quality, streak seeding technique using cat whisker as the probe (Ducruix & Giegé, 2000) was applied, as well as other macro- and micro-seeding techniques (Thaller *et al.*, 1981).

4.7.2 X-ray data collection

The testing of crystals for diffraction and data collection on sufficiently large crystals were performed using an in-house diffractometer (Nonius FR 591) connected to 345 mm MarResearch image plate detector at 120 K. Diffraction data collection using synchrotron radiation was performed at 100 K using the X12 EMBL beamline at DESY, Hamburg, Germany. For cryoprotection, the crystals were transferred into reservoir solution supplemented with 10 - 30 % of various tested cryoprotecting compounds (e.g. glycerol and PEG400), and flash-cooled by plunging into liquid nitrogen.

Diffraction data were integrated and reduced using MOSFLM (Leslie, 1999) and scaled using SCALA (Evans, 1993) from the CCP4 suite of programs (*The CCP4 suite: programs for protein crystallography*, 1994).

4.7.3 Structure determination, refinement, and analysis

Crystal structures were determined by the difference-Fourier method or, in cases where crystals were not isomorphous, by molecular replacement using the program Molrep (Vagin & Teplyakov, 2000). Model refinement was carried out using the program REFMAC 5.2 (Murshudov *et al.*, 1997) from the CCP4 package (The CCP4 suite: programs for protein crystallography, 1994), interspersed with manual adjustments using Coot (Emsley & Cowtan, 2004). The final steps included TLS refinement (Winn *et al.*, 2001). The quality of the final models was validated with Molprobity (Lovell *et al.*, 2003). All figures showing structural representations were prepared with the program PyMOL (DeLano, 2002).

Atomic coordinates and experimental structure factors have been deposited with the Protein Data Bank with the codes 2HKH and 2HKF for anti-CA IX antibody fragments and 3IGP for CA II + inhibitor complex.

5 RESULTS AND DISCUSSION

5.1 Targeting active site of human carbonic anhydrases

As already mentioned, we have decided to target the tumor-associated CA IX isozyme using a two-pronged approach: (1) by inhibition of the active site with small molecular inhibitors and (2) by targeting with specific monoclonal antibody recognizing specific epitope in the unique proteoglycan-like (PG) domain.

We initiated structural studies of CA IX, therapeutically very interesting isozyme for which the crystal structure was not available at the time of commencement of this thesis and was subject of research of many academic as well as pharmaceutical laboratories. As far as the active site is concerned, all human isozymes share a high degree of similarity due to significant sequence conservation of mammalian α-CAs. Some variation in amino acid composition in residues which are not directly involved in catalysis results in subtle changes in the shape of substrate binding cavity. To gain structural information needed in development of specific inhibitor of CA IX we also initiated structural studies on other CA isozymes, namely CA II.

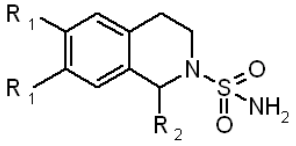
5.1.1 Structural studies of CA II in complex with inhibitors based on 3,4-dihydroisoquinoline-2(1H)-sulfonamide

For initial structural studies with a series of inhibitors we have selected isozyme CA II. This isoform has been extensively studied on structural level (for review see Krishnamurthy et al, 2008) and is typically used as a comparison partner for CA IX. The CA II is abundant and physiologically important and large quantities of protein for crystallization is readily available.

A series of novel 6,7-dimethoxy-3,4-dihydroisoquinoline-2(1H)-sulfonamides was prepared in collaboration with Prof. Chimirri group of medicinal chemists from University of Messina (Italy) (Gitto *et al.*, 2010). The enzymatic inhibitory activity of the series of compounds was tested on four important carbonic anhydrase isozymes involved in several physiological and pathophysiological processes: cytosolic CA I and

CA II isozymes, the tumor-associated transmembrane isoform CA IX, and the neuronal CA XIV isoform. The parent structural formula for this compounds series, inhibition constants and selectivity ratios for tested isozymes are summarized in Table 4.

Table 4 Inhibition of hCA I, hCA II, hCA IX, and hCA XIV.
Inhibitory constants of 6,7-Dimethoxy-3,4-dihydroisoquinoline-2(1H)-sulfonamides (1-10), 6,7-Dimethoxy-1-phenyl-3,4-dihydroisoquinoline-2(1H)-sulfonamide (I), zonisamide, acetazolamide, and topiramate and selectivity ratios K_I(hCA II)/K_I(hCA IX) and K_I(hCA II)/K_I(hCA XIV).



			K_I (nM)[a]				*Selectivity ratios*[b]	
R_1	R_2		hCA I	hCA II	hCA IX	hCA XIV	K_I(hCAII)/ K_I(hCAIX)	K_I(hCAII)/ K_I(hCAXIV)
H	**H**	**1**	1510	32.8	3.7	6.0	8.86	5.47
MeO	**H**	**2**	6410	94.5	9.5	9.8	9.95	9.64
MeO	**methyl**	**3**	2800	87.3	9.4	9.6	9.29	9.09
MeO	**ethyl**	**4**	1820	1975	76.3	2754	25.9	0.72
MeO	**n-propyl**	**5**	3150	235000	330	302	712.12	778.15
MeO	**i-propyl**	**6**	3780	1050	6.1	4.7	172.13	223.40
MeO	**c-propyl**	**7**	3950	820	8.5	6.4	96.47	128.13
MeO	**n-butyl**	**8**	4100	350000	400	1650	875.00	212.12
MeO	**c-pentyl**	**9**	100	650	55	178	11.82	3.65
MeO	**c-hexyl**	**10**	1180	18890	706	7950	26.76	2.38
MeO	**phenyl**	**I**[c]	8980	15700	8440	3860	1.86	4.07
acetazolamide[d]			250	12	25	41	0.48	0.29
zonisamide[d]			56	35	5.1	5250	6.86	<0.01
topiramate[d]			250	10	58	1460	0.17	<0.01

We have performed successful crystallization experiments with compounds 2 and 3 from the table, which were named by our collaborators DT1 and DT2, respectively. Figure 8 shows structures of DT1 and DT2 with atom names as used in the crystal structured described further.

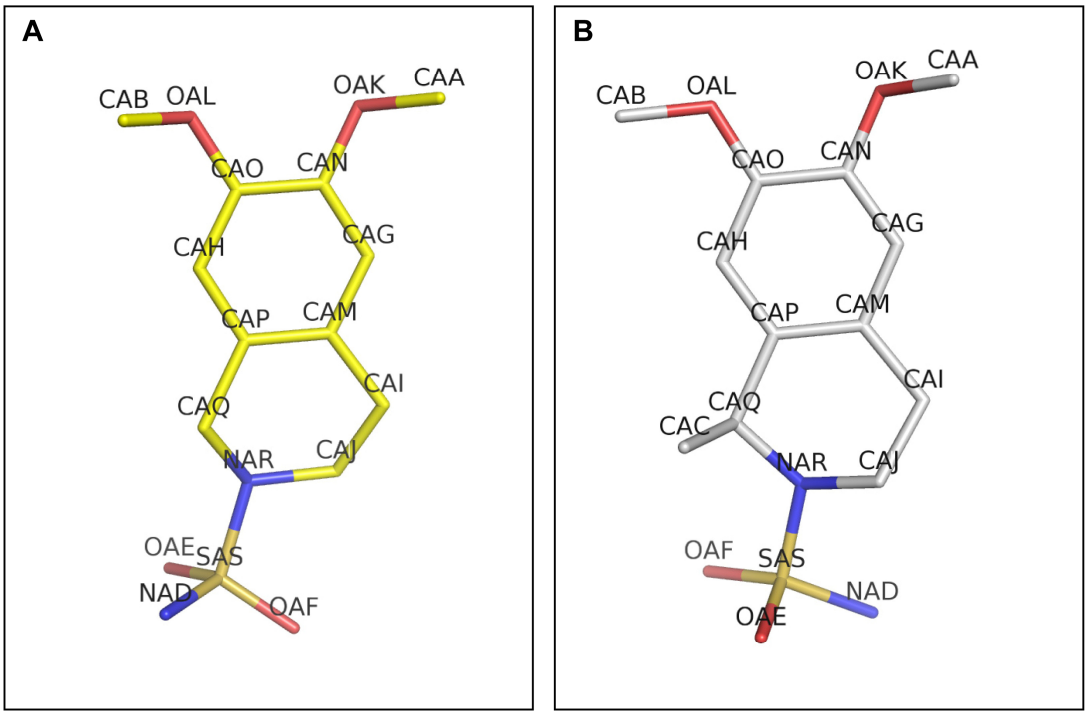


Figure 8 Structure of inhibitors used in structural studies with atom names indicated. *(A) Structure of DT1, 6,7-dimethoxy-3,4-dihydroisoquinoline- 2(1H)-sulfonamide and (B) DT2, (1R)-6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline- 2(1H)-sulfonamide.*

5.1.2 Crystallization of CA II complexes

The complexes of hCA II with sulfonamide inhibitors were prepared by adding 5-fold molar excess of inhibitor (in dimethyl sulfoxide) to 10 mg/ml protein solution of hCA II in 100mM Tris-Cl pH 8.5. The initial crystallization conditions were selected based on successful conditions used by others (Weber *et al.*, 2004). The hanging drop vapor diffusion method was used for screening and optimization as described in Methods (sub 4.7.1): drops containing 1 µl of protein solution were mixed with 1 µl of crystallization cocktail and placed over a 1 ml reservoir.

The best crystals of the complex were obtained by the hanging-drop vapor diffusion method, under the following conditions: an amount of 2 µl of complex solution was mixed with 2 µl of precipitant solution [2.5 M $(NH_4)_2SO_4$, 0.3 M NaCl, 100 mM Tris-Cl, pH 8.2] and equilibrated over a reservoir containing 1 ml of

precipitant solution at 18 °C. Crystals with dimensions 0.4 mm × 0.2 mm × 0.1 mm grew within 10 days (Figure 9).

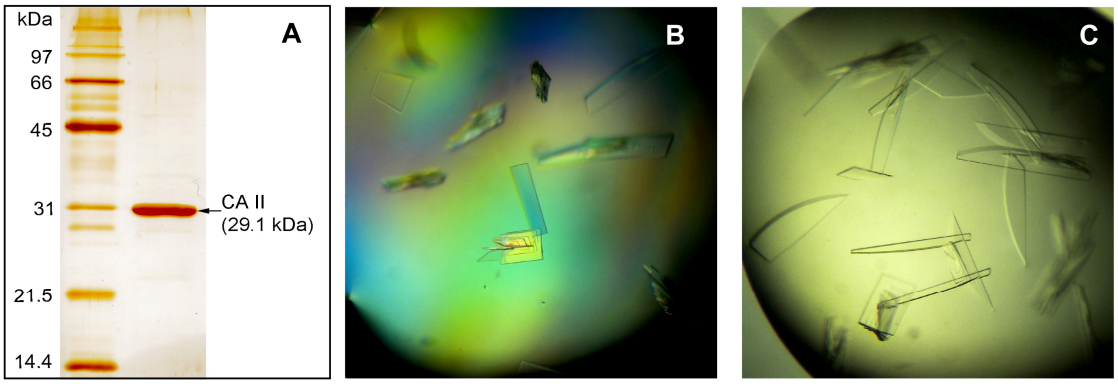


Figure 9 Crystallization of CA II in complex with inhibitors.

(A) SDS-PAGE analysis of CA II used for crystallization experiments. Silver stained 14 % SDS-PAGE gel with 1 µg of protein applied to the gel. (B) Crystals of CA II in complex with compound 2 (DT1), is ~ 0.4 × 0.2 × 0.1 mm. (C) Crystals of CA II in complex with compound 3 (DT2), is ~ 0.4 × 0.2 × 0.1 mm.

5.1.3 Data collection and structure determination

For data collection, the crystals were soaked in the reservoir solution supplemented with 20 % glycerol and transferred to liquid nitrogen.

Diffraction data for CA II + DT1 complex were collected at 120 K using an in-house diffractometer (Nonius FR 591) connected to 345 mm MarResearch image plate detector. The best crystal diffracted up to 1.65 Å resolution, and diffraction data were integrated and reduced using MOSFLM (Leslie, 1999) and scaled using SCALA (Evans, 1993) from the CCP4 suite of programs (*The CCP4 suite: programs for protein crystallography*, 1994).

Diffraction data for CA II + DT2 complex were collected at 100 K at Hamburg DESY, beamline X12, wavelength 0.9537 Å, using MAR Mosaic 225 CCD detector. The best crystal diffracted up to 1.47 Å resolution, and diffraction data were integrated and reduced using the same procedure described above for CA II + DT1 complex.

Both CA II complexes crystallized in monoclinic spacegroup $P2_1$, crystal parameters and data collection statistics are summarized in Table 5.

The structures of both CA II complexes were solved using the difference Fourier method, using CA II structure (Protein Data Bank entry 1H9N (Lesburg *et al.*, 1997)) as the initial model. Initial rigid-body refinement and subsequent restrained refinement were performed using the program REFMAC 5.48.

Both structures were refined with two inhibitor molecules, one in the enzyme active site and the other located in the surface pocket in the vicinity of the enzyme N-terminus. Atomic coordinates and geometry library for the inhibitor were generated using the PRODRG server (Schuttelkopf & van Aalten, 2004). The Coot program (Emsley & Cowtan, 2004) was used for inhibitor fitting, model rebuilding, and addition of water molecules. In final refinement stages, TLS refinement cycles in the program REFMAC 5 were introduced (Winn *et al.*, 2001). Atomic coordinates and experimental structure factors for CA II + DT1 complex have been deposited with the Protein Data Bank with the codes 3IGP. The final refinement statistics are summarized in Table 5.

Table 5 Crystal parameters, data collection and refinement statistics for complex of CA II with DT1 and DT2 compounds.

Crystal parameters, data collection and refinement statistics for CA II in complex with DT1 and DT2

Data collection statistics	**Complex of CA II with DT1**	**Complex of CA II with DT2**
Space group	$P2_1$	$P2_1$
Unit cell parameters (Å)	a = 42.31 Å b = 41.27 Å c = 72.04 Å α = 90° β = 104.22° γ = 90°	a = 42.18 Å b = 41.21 Å c = 72.00 Å α = 90° β = 104.31° γ = 90°
Number of molecules in AU	1	1
Wavelength (Å)	1.5418	0.95370
Resolution range (Å)	19.95 - 1.65 (1.69 - 1.65)	26.62 - 1.47 (1.51 – 1.47)
Number of unique reflections	26,928	40,809
Redundancy	2.2 (2.2)	3.6 (3.4)
Completeness (%)	92.5 (88.6)	99.8 (99.8)
R_{merge} [a]	0.079 (0.362)	0.053 (0.238)
Average I/σ(I)	5.9 (2.0)	7.8 (3.0)
Wilson B ($Å^2$)	13.3	13.7
Refinement statistics		
Resolution range (Å)	19.95 – 1.65 (1.69 - 1.65)	23.8 – 1.47 (1.51 – 1.47)
No. of reflections in working set	25,556 (1,811)	38,748 (2,849)
No. of reflections in test set	1,353 (84)	2,045 (132)
R [b] (%)	16.08 (22.40)	14.3 (18.2)
R_{free} [c] (%)	19.55 (28.90)	17.2 (21.2)
RMSD bond length (Å)	0.011	0.012
RMSD angle (°)	1.43	1.57
Number of atoms in AU	4,949	5,314
Number of protein atoms in AU	4,138	4,301
Number of solvent molecules in AU	237	1,013
Mean B value ($Å^2$)	14.0	13.8
PDB code	3IGP	

The data in parentheses refer to the highest-resolution shell.

[a] $R_{merge} = \Sigma_{hkl}\Sigma_i I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where the $I_i(hkl)$ is an individual intensity of the *i*th observation of reflection hkl and $\langle I(hkl)\rangle$ is the average intensity of reflection *hkl* with summation over all data.

[b] R-value = $||F_o| - |F_c||/|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively.

[c] R_{free} is equivalent to R value but is calculated for 5 % of the reflections chosen at random and omitted from the refinement process (Brunger, 1992).

5.1.4 Refined models and overall structures

The crystal structures of hCA II in complex with 6,7-dimethoxy- 3,4-dihydroisoquinoline-2(1H)-sulfonamide (DT1) and 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline-2(1H)-sulfonamide (DT2) were determined by difference Fourier techniques and refined using diffraction data to 1.65 Å and 1.47 Å resolution, respectively. Both complexes crystallized in the monoclinic $P2_1$ space group with one hCA II molecule in the asymmetric unit and solvent content of 41.3% and 42.2% for CA II+DT1 and CA II+DT2 complex, respectively.

In case of CA II+DT1 complex, all protein residues could be traced into a well defined electron density map with the exception of side chains of several terminal amino acid residues (Ser2, His3, His4, and Lys261) and side chains of two surface residues Lys9 and Lys133.

All residues of complex CA II + DT2 were well traceable into the electron density map, with the exception of N-terminal AA Ser2.

In both structures, two continuous non-protein electron densities were noticeable during the course of the crystallographic refinement: first in the active site and second on the surface of the protein molecule. Both could be unambiguously modeled as the inhibitor molecules. Binding of inhibitor into the active site will be discussed in the following chapter.

The inhibitor second binding site on the protein surface probably has no biological relevance and represents a crystallization artifact caused by high concentrations of inhibitor employed in the co-crystallization experiments. This inhibitor second binding site is also occupied by sulfonamide inhibitors in other hCA II-inhibitor complexes (e.g., PDB codes 2FOS, 2FOV, 2FOQ, 2FOU (Jude et al., 2006); 2NNO, 2NNS, 2NNV (Srivastava et al., 2007); and 1ZFQ).

The crystal structure of CA II + DT2 complex revealed an interesting result regarding the binding activity of two possible enantiomers of DT2 compound. The crystals were grown from solution of CA II with a racemic mixture of R and S enantiomers of 6,7-dimethoxy-1-methyl-3,4-dihydroisoquinoline- 2(1H)-sulfonamide.

The R enatiomer was found in active site of enzyme, whereas the S enantiomer was found bound to the surface of the protein molecule.

The inhibitor binding does not cause any major structural changes to the protein; the root-mean-square deviation (rmsd) for superposition of 256 Cα atoms (residues 4-259) of our complex structures with free hCA II (PDB code 1CA2 (Eriksson et al., 1988)) is 0.3-0.4 Å, which is below the value observed for identical structures (Betts & Sternberg, 1999).

Also the structures of CA II in complex with DT1 and DT2, respectively, are almost identical with the exception of N-terminal part. When first 6 residues are omitted, the rmsd for superimposition Cα atoms is 0.153 Å and for all atoms 0.132 Å, respectively.

Overall structure of CA II+DT1 with the two inhibitor molecules bound is shown in Figure 10.

All figures showing structural representations were prepared using PyMOL (DeLano, 2002) and the APBS tools (Baker *et al.*, 2001) plugin was used for generating solvent accessible surface colored by electrostatic potential.

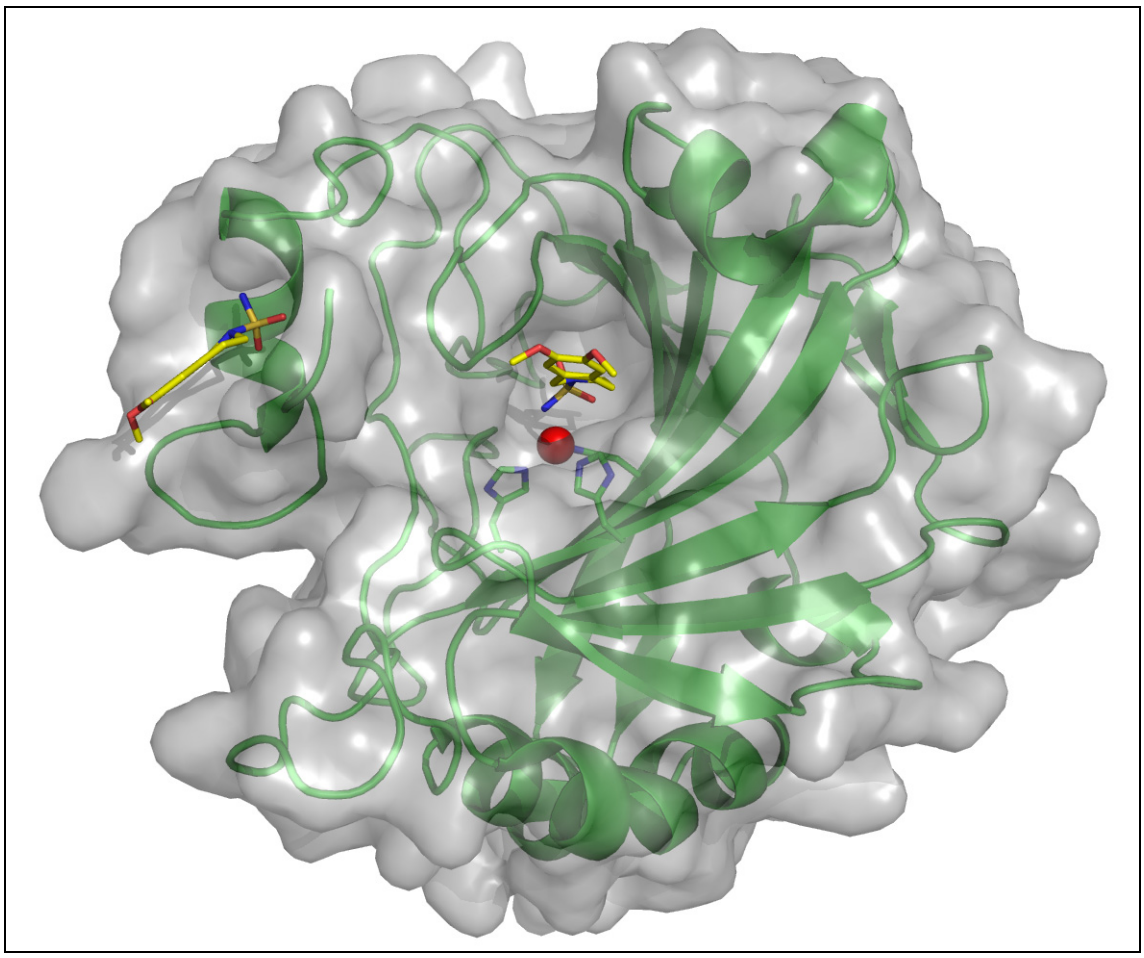


Figure 10 Overall structure of CA II + DT1 inhibitor. *CA II protein is represented by ribbon with transparent solvent accessible surface. Zinc ion is shown as a red sphere with three coordinating histidine residues in sticks. Two molecules of inhibitor DT1 bound to CA II are shown in stick model with carbon atoms colored yellow and oxygen and nitrogen atoms colored red and blue, respectively.*

5.1.5 Details of inhibitor DT1 and DT2 binding

As shown in Figure 11, the compounds DT1 and DT2 bind into the cavity of the CA II active site with the deeply buried sulfonamide group. The ionized nitrogen atom of the sulfonamide moiety is coordinated to the zinc ion at a distance ~ 2.0 Å. The sulfonamide nitrogen also accepts a hydrogen bond from hydroxyl group of Thr199 sidechain, and one oxygen from the sulfonamide moiety forms a hydrogen bond with backbone amine group of Thr199. These key hydrogen bonds between the sulfonamide moiety of the inhibitor and enzyme active site are also found in other structurally

characterized CA II-sulfonamide complexes (Eriksson *et al.*, 1988, Supuran *et al.*, 2003, Winum *et al.*, 2005, Srivastava *et al.*, 2007).

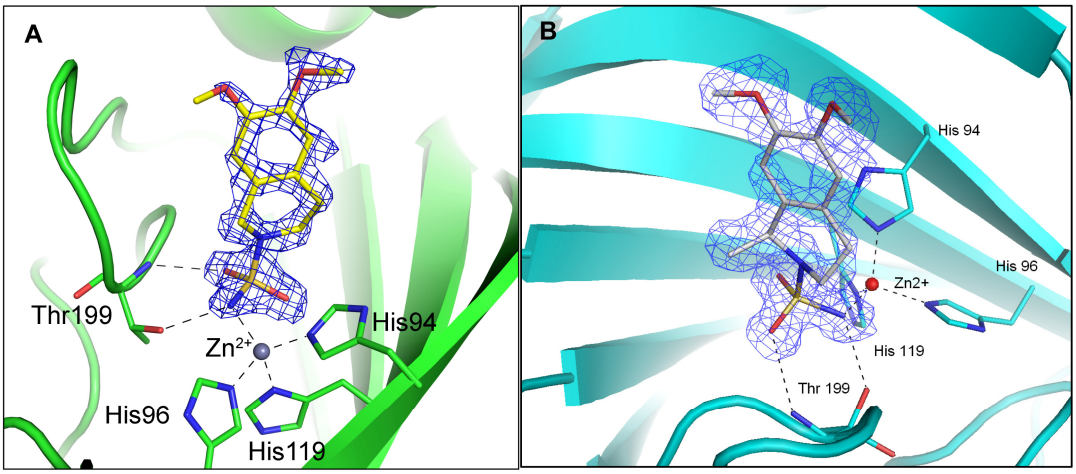


Figure 11 Binding of DT1 and DT2 inhibitor on CA II. *Detail of the CA II active site with the inhibitor DT1 (A) and DT2 (B) Protein is represented in green (panel A) and cyan (panel B) with residues forming polar contacts (gray dashed lines) with inhibitor highlighted in sticks. Also three histidine residues coordinating zinc ion are shown. In panel (A) DT1 inhibitor is represented as a stick model (with carbon and oxygen atoms colored yellow and red, respectively). The 2Fo - Fc electron density maps are contoured at 1.5σ. In panel (B) the DT2 inhibitor represented as a stick model (with carbon and oxygen atoms colored gray and red, respectively). The 2Fo - Fc electron density maps are contoured at 0.8σ.*

The sulfonamide moiety seems to serve as an anchor for specific binding of DT1 and DT2 compounds into the enzyme active site. In addition to the polar interactions mediated by the sulfonamide group, hydrophobic interactions of the substituted isoquinoline moiety strongly stabilize the inhibitor within the active site cavity (Table 6).

Table 6 List of contacts between CA II and inhibitors. *All contacts with a distance below 4.2 Å between inhibitor atoms on one side, and protein residues and coordinated Zn ion on the other, are listed. The inhibitor atom names are depicted in Figure 8.*

CAII			DT1		CAII			DT2	
Residue #	Residue	Atom name	Atom name	distance [Å]	Residue #	Residue	Atom name	Atom name	distance [Å]
92	GLN	NE2	CAI	3.82	62	ASN	ND2	CAA	3.82
94	HIS	CE1	NAD	3.93	67	ASN	ND2	CAO	4.15
94	HIS	NE2	NAD	3.42	67	ASN	ND2	CAN	3.40
94	HIS	CE1	OAF	3.40	67	ASN	CG	OAK	3.97
94	HIS	NE2	OAF	3.54	67	ASN	ND2	OAK	2.75
94	HIS	CE1	CAI	3.96	67	ASN	ND2	CAA	3.67
94	HIS	CE1	CAJ	3.80	67	ASN	ND2	CAG	3.98
94	HIS	CE1	NAR	4.07	91	ILE	CG2	CAB	4.04
94	HIS	CE1	SAS	3.99	91	ILE	CD1	CAB	4.11
94	HIS	NE2	SAS	3.96	92	GLN	NE2	CAP	3.52
96	HIS	CE1	NAD	3.84	92	GLN	CD	CAH	3.63
96	HIS	NE2	NAD	3.44	92	GLN	NE2	CAH	2.94
106	GLU	OE1	NAD	4.18	92	GLN	CG	CAH	4.10
119	HIS	ND1	NAD	3.51	92	GLN	CD	CAO	3.50
119	HIS	CE1	NAD	4.06	92	GLN	NE2	CAO	2.52
119	HIS	CB	OAF	3.72	92	GLN	CG	CAO	3.88
119	HIS	CG	OAF	3.80	92	GLN	CD	OAL	3.62
119	HIS	ND1	OAF	3.44	92	GLN	NE2	OAL	2.89
119	HIS	ND1	SAS	4.07	92	GLN	CG	OAL	3.48
121	VAL	CG2	OAF	4.02	92	GLN	CD	CAB	4.08
121	VAL	CG2	CAI	3.82	92	GLN	NE2	CAB	3.75
121	VAL	CG1	CAI	4.02	92	GLN	CG	CAB	3.52
121	VAL	CG2	CAJ	3.67	92	GLN	CD	CAN	4.02
121	VAL	CG1	CAJ	4.11	92	GLN	NE2	CAN	2.76
131	PHE	CZ	CAG	3.83	92	GLN	NE2	OAK	3.32
131	PHE	CE2	OAK	3.96	92	GLN	NE2	CAG	3.36
143	VAL	CG1	OAF	3.77	92	GLN	NE2	CAM	3.71
143	VAL	CG2	OAF	4.06	94	HIS	CE1	CAQ	3.69
197	SER	O	OAE	4.06	94	HIS	CE1	CAP	3.90
198	LEU	CA	OAE	3.27	94	HIS	CE1	CAM	3.97
198	LEU	C	OAE	3.54	94	HIS	CE1	CAI	3.93
198	LEU	CB	OAE	3.99	94	HIS	NE2	CAI	4.10
198	LEU	CD2	OAE	4.17	94	HIS	CE1	CAJ	4.16
198	LEU	CD2	NAR	4.02	94	HIS	CE1	NAR	3.22
199	THR	CB	NAD	3.92	94	HIS	NE2	NAR	3.60
199	THR	OG1	NAD	2.70	94	HIS	CE1	SAS	3.76
199	THR	N	NAD	3.98	94	HIS	NE2	SAS	3.81
199	THR	CA	OAE	4.02	94	HIS	CE1	OAF	3.52
199	THR	CB	OAE	4.08	94	HIS	NE2	OAF	3.83
199	THR	OG1	OAE	3.64	94	HIS	CE1	NAD	3.87
199	THR	CG2	OAE	4.03	94	HIS	NE2	NAD	3.32
199	THR	N	OAE	2.91	96	HIS	NE2	NAD	3.31
199	THR	OG1	SAS	3.80	96	HIS	CE1	NAD	3.70

Table 6 (continued) List of contacts between CA II and inhibitors. *All contacts with a distance below 4.2 Å between inhibitor atoms on one side, and protein residues and coordinated Zn ion on the other, are listed. The inhibitor atom names are depicted in Figure 8.*

CAII			DT1		CAII			DT2	
Residue #	Residue	Atom name	Atom name	distance [Å]	Residue #	Residue	Atom name	Atom name	distance [Å]
200	THR	OG1	CAB	3.46	119	HIS	ND1	SAS	4.03
200	THR	OG1	CAH	3.09	119	HIS	CB	OAF	3.91
200	THR	OG1	CAO	4.12	119	HIS	CG	OAF	4.00
200	THR	OG1	CAP	3.77	119	HIS	ND1	OAF	3.67
200	THR	OG1	CAQ	3.57	119	HIS	ND1	NAD	3.33
201	PRO	C	CAB	4.12	119	HIS	CE1	NAD	3.86
201	PRO	O	CAB	2.98	121	VAL	CG2	CAQ	3.77
201	PRO	O	OAL	3.79	121	VAL	CG2	OAF	3.80
202	PRO	CD	CAB	4.07	130	PHE	CD1	CAB	4.17
202	PRO	CD	OAL	3.92	130	PHE	CE1	CAB	3.61
209	TRP	CZ2	OAE	3.72	130	PHE	CZ	CAB	3.45
209	TRP	CH2	OAF	3.96	130	PHE	CE2	CAB	3.89
					142	VAL	CG1	OAF	3.73
					142	VAL	CG2	OAF	3.81
					197	LEU	CD2	CAC	3.34
					197	LEU	CA	OAE	3.19
					197	LEU	CB	OAE	3.81
					197	LEU	C	OAE	3.53
					198	THR	OG1	SAS	3.94
					198	THR	N	SAS	4.05
					198	THR	OG1	OAE	3.86
					198	THR	N	OAE	2.96
					198	THR	CA	OAE	4.12
					198	THR	CB	NAD	3.84
					198	THR	OG1	NAD	2.63
					198	THR	N	NAD	3.94
					199	THR	OG1	CAI	3.31
					199	THR	CG2	CAI	3.99
					199	THR	CB	CAJ	4.16
					199	THR	OG1	CAJ	3.19
					199	THR	CG2	CAJ	4.00
					208	TRP	CZ2	OAE	4.07
					208	TRP	CH2	OAF	3.99
CAII coordinated Zn ion			DT1 inhibitor		CAII coordinated Zn ion			DT2 ihnibitor	
		Atom name	Atom name	distance [Å]			Atom name	Atom name	distance [Å]
	ZN	ZN	NAD	2.06		ZN	ZN	NAR	3.64
	ZN	ZN	OAF	3.00		ZN	ZN	SAS	3.05
	ZN	ZN	SAS	3.09		ZN	ZN	OAF	3.33
						ZN	ZN	CAJ	4.17
						ZN	ZN	NAD	1.94

In the vicinity of the inhibitor, a glycerol molecule could be modeled into the CA II active site, making hydrogen bonds with Asn62, Asn67, and Gln92 (Figure 12). The binding of glycerol to an identical site was also observed in several other hCA II-inhibitor complexes (e.g., 2NNG, 2NNS, 2NNO, 2NNV (Srivastava *et al.*, 2007)).

Glycerol was used at high concentration (20% v/v) for crystal cryoprotection, and its binding probably represents a crystallization artifact. In the complex of CA II with DT2 inhibitor, the glycerol molecule in the active site is not present as its position is occupied by the isoquinoline moiety of DT2.

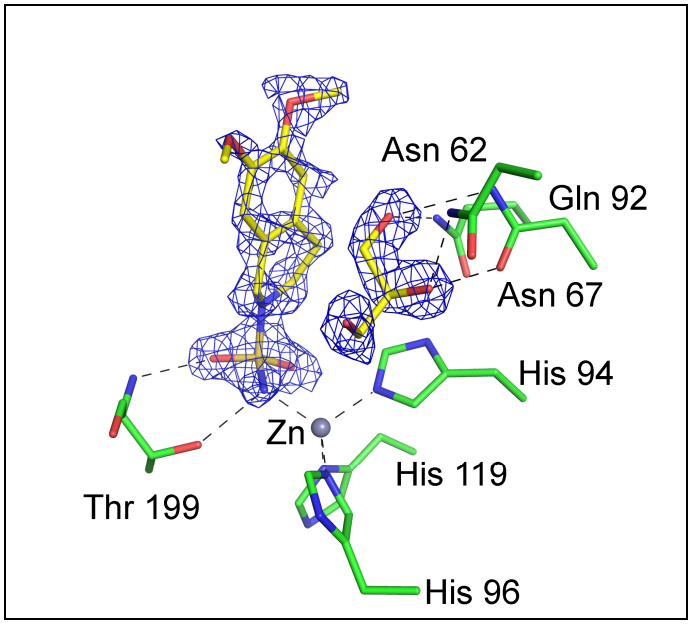


Figure 12 Binding of glycerol molecule into the active cavity of CA II +DT1 complex. *Detail of the CA II active site with the DT1 compound and glycerol represented as a stick models (with carbon and oxygen atoms colored yellow and red, respectively). The 2Fo-Fc electron density maps are contoured at 1.5 σ. Protein is represented in green sticks. Also three histidines coordinating zinc ion are shown.*

5.1.6 Comparison of inhibitor DT1 and DT2 binding mode

Comparison of the two CA II complex structures revealed interesting differences in binding of inhibitors DT1 and DT2 into the CA II active site. Although these two inhibitors differ only by the presence of an extra methyl group in DT2 compared to DT1 (see Figure 8), their binding mode is quite different (Figure 13).

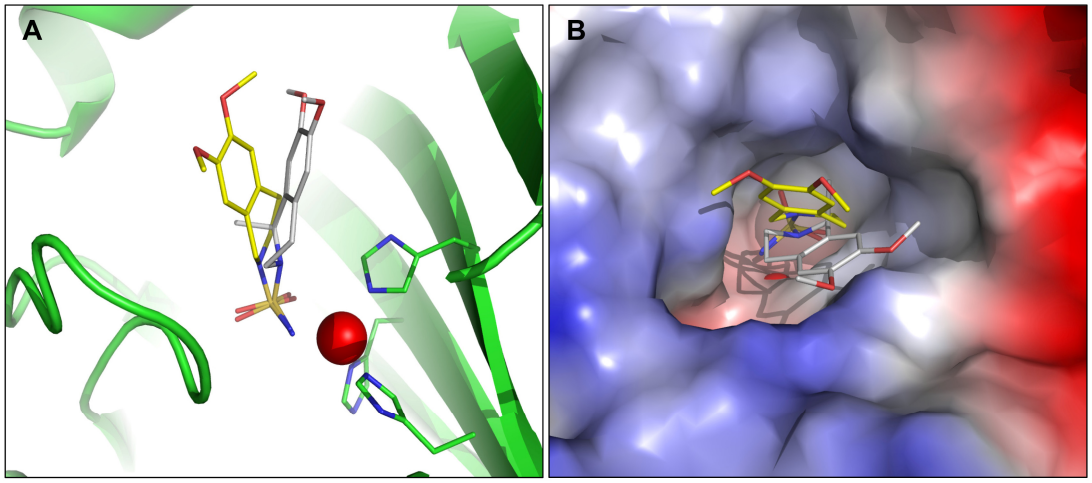


Figure 13 Comparison of DT1 and DT2 binding mode in CA II active site. *Superposition of the complex structures is based on the best fit for Cα atoms of CA II residues from Gly6 until C-terminus. DT1 inhibitor is represented with yellow carbon atoms, while DT2 carbon atoms are colored gray. In panel (A) protein is represented by ribbon and histidine residues coordinating Zn ion (red sphere) are shown as sticks. Panel (B)shows a top view into the active site where protein is represented by its van der Waals radii colored by the electrostatic potential from negative (red) to positive (blue).*

While the position of the sulfonamide group coordinating the Zn ion in the CA II active site is in DT1 and DT2 inhibitors is similar, the isoquinoline moiety positioning is different. In DT2 the ring system is rotated around the covalent bond between sulfur of sulfonamide group and nitrogen of six member ring by almost 180º when compared to DT1 conformation (Figure 13B).

In DT1 complex, the ring system forms numerous van der Waals interactions (distance of < 4.2 Å) with the bottom of the active site (residues Gln92, Val121, Val143, Trp209) and the loop formed by residues Ser197-Pro202. Also, the side chain of Phe131 from helix 4 contributes to interaction with the substituting methoxy-groups. In DT2 complex, the isoquinoline moiety is in contact with residues of seven stranded β-sheet located on the opposite side of the active site, mainly with Gln 92 and Val 121 (see Table 6).

When the conformations of isoquinoline moieties of DT1 and DT2 inhibitor are overlaid, it is obvious that that not only a torsion angle between sulfur of sulfonamide group and nitrogen of six member ring is different (Figure 14). The substituted

heterocyclic ring is deformed differentially from a non-substituted ring of DT1 molecule. This is probably caused by adaptation of the molecular shape necessary for proper interaction of sulfonamide group in the bottom of active site.

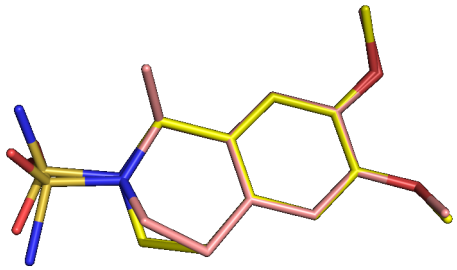


Figure 14 Comparison of DT1 and DT2 inhibitor conformation when bound to CA II active site. *The superposition was based on the conjugated rings of heterocycle atoms of isoquinoline moiety. DT1 inhibitor is represented with yellow carbon atoms, while DT2 carbon atoms are colored pink. Oxygen and nitrogen atoms are colored red and blue, respectively.*

5.1.7 Conclusions from structural analysis of CA II + inhibitor complexes

The results of the structural analyses presented in this thesis have become part of a publication Gitto *at al.*, 2009. Solved and analyzed crystal structures of CA II + DT1 and CA II + DT2 complexes revealed valuable structural information about interactions of 3,4-dihydroisoquinoline-2(1H)-sulfonamides with components of active site of CA II protein. Although it might appear that these two structures are just two more additions into a long list of structures of CA II + inhibitor complexes already deposited in PDB, they uncovered some unique features which can be used in future structure-based drug design effort.

The position of the sulfonamide group in the active site is highly similar to all other structures deposited in PDB and proves the importance of tetrahedral coordination of Zn^{2+} cation by four nitrogen atoms with their lone pair. Positions of the nitrogen atoms, Nε His 94, Nε His 96 and Nδ His 119 form a ligand field which dictates the

position of coordinated Zn^{2+} cation and consecutively determine the placing of acidic group of sulfonamide, to complete the tetrahedral coordination of central zinc ion.

Isoquinoline moiety of inhibitor molecule in the crystal structure of CA II + DT1 complex is stabilized by numerous van der Waals interactions in the active site cavity formed mainly by N-terminal part of CA II, whereas the substituted isoquinoline moiety of CA II + DT2 complex is rotated around the sulfur – nitrogen covalent bond, and ring system is stabilized by van der Waals interactions interaction with the amino acid residues coming from the central part seven stranded β-sheet on the opposite side of the active site cavity. Large distance difference, more than 4 Å, between methoxy oxygen atoms of the conjugated ring for the two placement of isoquinoline moiety in the solved structures (see Figure 13) indicated, that more voluminous substituent might give a better fit for the active site cavity. These findings allowed us to design a novel compound with optimized shape of the substituent as part of a project, which is out of scope of this thesis. The binding of such a lead compound has already been experimentally proved and the structural studies are in progress.

An important result from the structural studies which will have consequences for design of future experiments is the finding, that absolute configuration of an enantiomer bound in the active site is R. This suggests that only one enantiomer represents an active compound and this discovery will have large impact to selectivity studies of this series of molecules. All activity measurements until now were carried out with mixtures of both enantiomers and the values of activity and selectivity were thus affected by the ratio of the active and non-active enantiomers. This subject should thus be addressed in future research.

In conclusion, the structural study of series of isoquinoline sulfonamides complexes with CA II brought new insight to design of sulfonamides derivatives, especially usage of more voluminous substituent. From the structural point of view, molecules DT1 a DT2 (better expressed as the mixture of R and S enantiomers of DT2) served as excellent molecular tools for detailed mapping of active site cavity. These results will be extensively exploited for rational design of molecules which could be highly selective inhibitors for various CA isozymes.

5.2 Structural studies of CA IX

CA IX in contrast to CA II is a transmembrane isozyme composed of several domains. While CA II can be either prepared in recombinant form or isolated from natural tissue (erythrocytes), CA IX needs to be prepared in recombinant form in order to get sufficient supply of protein for structural studies. Furthermore, due to known problems regarding work and crystallization of transmembrane proteins, which is still an extremely challenging task, we had cloned and expressed truncated constructs composed of extracellular domains. As will be detailed below, some of these constructs contained both the proteoglycan-like (PG) domain and the catalytic carbonic anhydrase (CA) domain, while other constructs contained only the CA domain. Also, N-terminal and C-terminal tags facilitating purification of the recombinant product were included. First, constructs for heterologous expression in *E. coli*, and later also constructs for eukaryotic expression system (*Drosophila* S2 cells) were prepared.

In case of the CA IX constructs for *E. coli* expression, the yield of expressed protein was far too low for structural studies; moreover no enzymatic activity of CA was observed. Switch to eukaryotic expression system proved beneficial both for expressed protein yield as well as activity. Similar results (good yield and higher activity in eukaryotic expression system) were recently reported by another group expressing rCA IX constructs (Hilvo *et al.*, 2008).

5.2.1 Recombinant CA IX constructs for *E. coli* expression

Several constructs for prokaryotic heterologous expression of recombinant CA IX (rCA IX) were prepared and their expression tested. Originally, we were inspired by a published work on *E. coli* expression and structure determination of CA XII (Whittington *et al.*, 2001), a homologous transmembrane CA isozyme. Accordingly, first constructs of PG+CA and CA domains were designed for cytoplasmic expression with the aim to obtain a soluble product. Since the CA domain contains an intramolecular disulfide bridge, mainly Origami B (DE3) or Rosetta-gami B (DE3) strains were used, which have mutations in thioredoxin reductase (*trxB*) and glutathione

reductase (*gor*) genes, enhancing thus disulfide bond formation in the cytoplasm. Additionally, Rosetta-gami B (DE3) strain supplies tRNAs for rare codons. Some experiments were performed with standard BL21 (DE3) strain. Some of the constructs prepared for expression in *E. coli* are listed in Table 7, describing the domain arrangement, type of the expression vector, and expected size of the product.

Table 7 List of plasmids prepared for rCA IX expression in *E. coli*

No.	CA IX construct[a]	Construct description, additional N/C-terminal tags, target protein localization	MW [kDa]
p597	residues 43-390	PG+CA in *E. coli* expression vector T7Q10 (cytoplasmic)	38.0
p598	residues 55-390	PG+CA in *E. coli* expression vector T7Q10 (cytoplasmic)	36.9
p599	residues 139-390	CA domain only in *E. coli* expression vector T7Q10 (cytoplasmic)	26.0
p631	residues 55-390	PG+CA in *E. coli* expression vector pET22b+ (periplasmic expression; N-terminal pelB signal sequence)	36.9
p640	residues 43-390	PG+CA in *E. coli* expression vector T7Q10 (cytoplasmic, N-terminal $(His)_6$-tag)	39.0

[a] Residue numbers are derived from Swiss-prot entry Q16790

Various attempts with most of these constructs to increase the yield of expression by modifying cultivation parameters, such as temperature prior and after induction, concentration of the inductor, strains used, media composition, etc. are detailed in Table 8. However, in all cases the expression level was too low for structural studies. Moreover, no enzymatic activity of CA was observed in these preparations.

Table 8 Summary of prokaryotic expression experiments.

Construct	Expression strain	Temperature [°C] (before / after induction)	Medium used	ETG concentration [mM]	Scale of experiment	Recombinant protein expression			Inclusion bodies presence (+/-)
						SDS-PAGE (difference before/after induction)	Detected on WB (mAb M75-HRP)	Yield (estimated from SDS-PAGE, WB, or A_{280})	
p597 (PG+CA)	Origami B (DE3)	37/37	LB	1	pilot	no	+	< 0.1 mg.l^{-1}	n/a
					medium	no	+	~ 0.1 mg.l^{-1}	+
		37/20	LB	1	pilot	no	+	~ 0.1 mg.l^{-1}	n/a
			LB, transformation plate with 1 % Glc		medium	no	++	~ 0.2 mg.l^{-1}	-
	BL21 (DE3)	37/37	LB	1	pilot	no	+	< 0.1 mg.l^{-1}	n/a
	BL21 (DE3) pLysS	37/37	LB	1	pilot	no	-	none detected	n/a
				3	pilot	no	+	< 0.1 mg.l^{-1}	n/a
	Rosetta-gami B	37/37	LB	3	pilot	no	+	~ 0.1 mg.l^{-1}	n/a
			2YT	3	pilot	no	++	~ 0.2 mg.l^{-1}	n/a
p598 (PG+CA)	BL21 (DE3)	37/37	LB	1	pilot	no	+	< 0.1 mg.l^{-1}	n/a
p599 (CA)	Origami B (DE3)	37/37	LB	1	pilot	no	n/a	none detected	n/a
		37/30	LB	1		no			
	BL21 (DE3)	37/37	LB	1	pilot	no			
p631 (periplasmic pelB/PG+CA)	BL21 (DE3)	37/25	LB	0.5	pilot	no	+++ (degradation products visible)	~ 0.4 mg.l^{-1}	-
p640 (His_6+PG+CA)	BL21 (DE3)	37/37	LB	1	medium	band observed after purification	+++	~ 1 mg.l^{-1}	+; but all product in soluble fraction

An example of a typical (i.e. disappointing) result of *E. coli* expression of PG+CA construct is shown below in Figure 15. The comparison of protein content in cellular fraction before and after induction separated on SDS-PAGE does not show any increase in production of protein with expected molecular weight ~ 38 kDa (Figure 15, left). Only by using a highly sensitive and specific detection by Western blot analysis allowed identification of a minute amount of product in the soluble cellular fraction (Figure 15, right, lanes 2a and 2b).

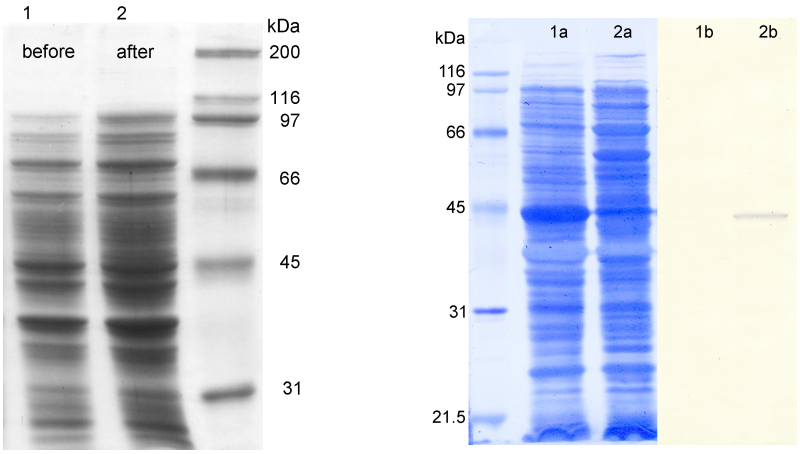


Figure 15 Expression of recombinant PG+CA construct in *E. coli* Origami B cells.

The 10 % SDS-PAGE gel analyzing protein content of bacterial cells before (lane 1) and 4 hours after induction with 1 mM ETG (lane 2,). The 12.5 % SDS-PAGE gel (lanes 1a and 2a) and Western blot analysis of the same gel (lanes 1b and 2b) stained with mAb M75-HRP conjugate on the right shows insoluble fraction (possibly containing inclusion bodies; lanes 1a and 1b) and soluble fraction (cell supernatant; lanes 2a and 2b).

To increase the amount of recombinant product, the effect of growth conditions were investigated including the effect of media composition. An example of an effect of using regular or rich media, i.e. LB or 2YT (for composition see Materials, chapter 4.2.1) on culture growth is shown in Figure 16. Although the yield of bacterial cell paste was only marginally higher for the rich 2YT media, the yield of recombinant product (as detected by Western blot) was approximately twice as high (Table 8). However, the rich media did not lead to sufficiently high protein yield.

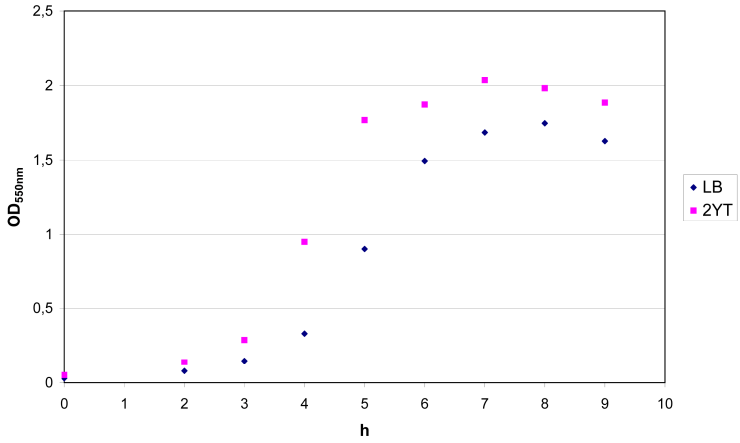


Figure 16 Example of the effect of media composition on the growth curves of *E. coli* Origami B cells. *Growth curves of E. coli Origami B cells grown on LB and 2YT media are shown as optical density of bacterial culture measured as a function of time.*

The estimated yield of expressed recombinant protein was in the optimized experiments less than or equal to 0.2 mg per 1 liter of bacterial culture. Attempts to purify the recombinant product using the affinity chromatography on sulfonamide agarose were not successful. An example of such unsuccessful purification is documented in Figure 17.

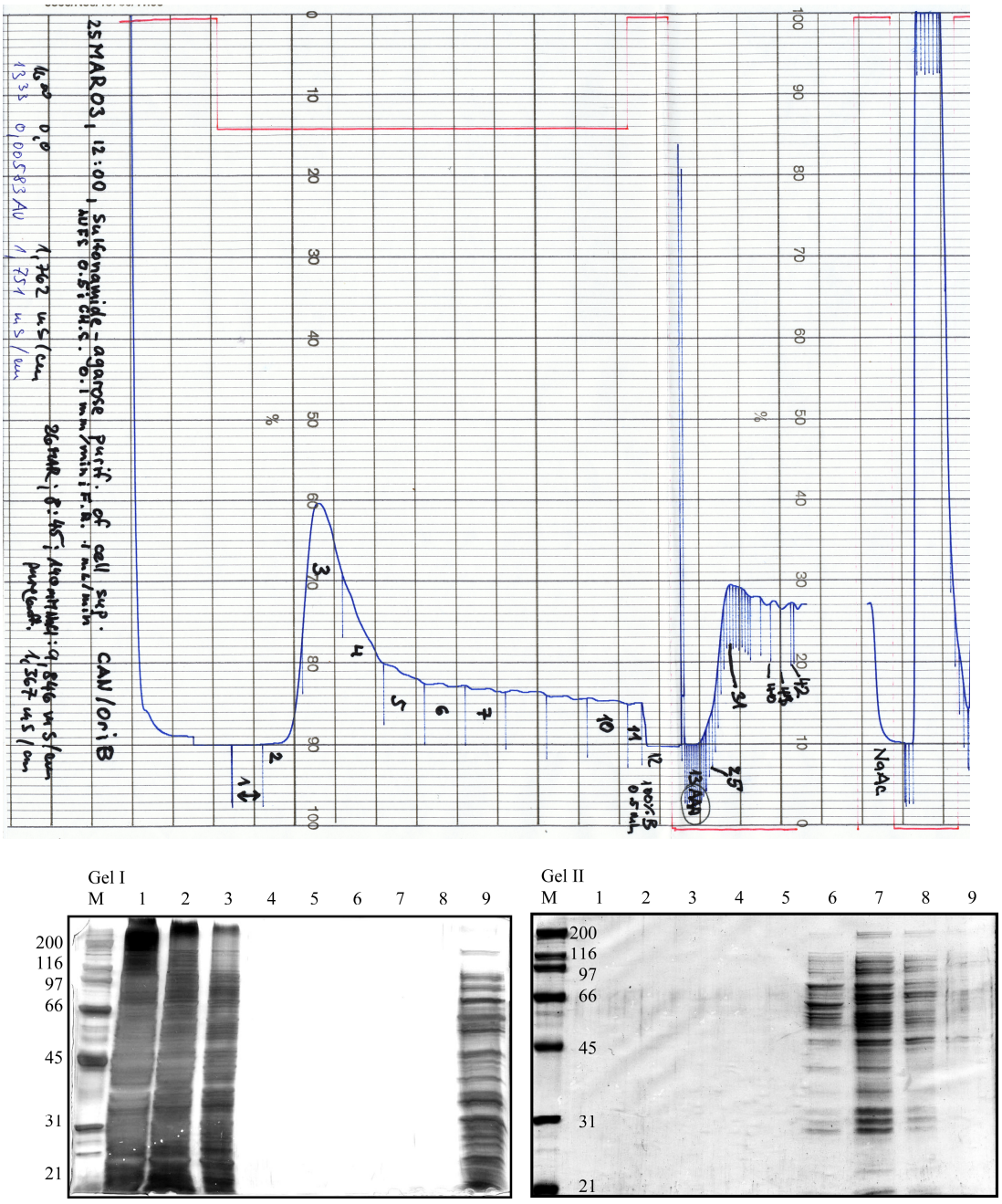


Figure 17 Purification of CA (p599) on sulfonamide agarose. *Above: chromatogram showing the course of the purification; fractions 1-12 correspond to application and washing of the column; fractions 13-42 to elution with 0.1 mM acetazolamide (nanomolar sulfonamide inhibitor of CAs; this compound shows strong absorbance at 280 nm) Below: the samples from chromatography fractions were analyzed on 12.5 % SDS-PAGE. Gel I, lane 1: Origami B cell lysate, lane 2: soluble fraction (cell supernatant after French press disintegration and centrifugation) applied on sulfonamide agarose column.; lane 3: flow through fraction; lanes 4-8: fractions from elution with 0.1 mM acetazolamide; lane 9: regeneration of the column with 0.1 M Na-acetate pH 5.0; 1 M NaCl; Gel II lanes 1-5 show elution fractions after TCA precipitation; lanes 6-9: fractions from column regeneration.*

As the heterologous expression of recombinant product targeted to cytosol, despite of enormous optimization effort, yielded only minute amounts of protein product, targeting recombinant product to periplasmic compartment was thus selected as an alternative strategy. The pilot experiment with periplasmic expression of PG+CA gave encouragingly higher yield; however, significant (probably proteolytic) degradation of the periplasmic product was observed (Figure 18), and the yield was not sufficient for structural studies.

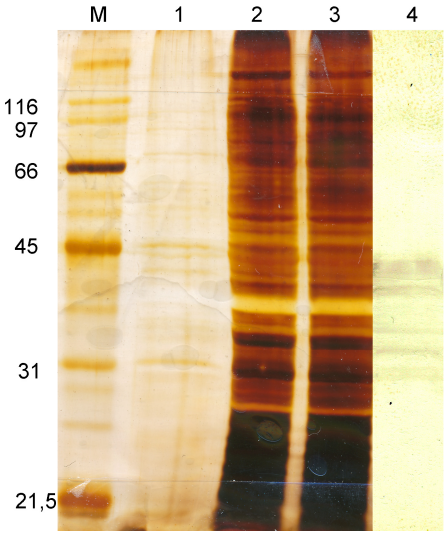


Figure 18 Periplasmic PG+CA pilot expression experiment (p631). *12.5 % SDS-PAGE was used for the analysis of expression; lane 1: LB medium after periplasmic expression (control sample); lane 2: cells before induction; lane 3: cells after induction; lane 4: WB of cells after induction. Again, the target protein could be detected only on WB, expression level was above average, though degradation could be observed.*

Including a His tag on the N-terminus of PG+CA was the final attempt to optimize recombinant production in prokaryotic system. The His tag apparently increased protein expression probably by affecting product solubility and provided an extra advantage in simple protein purification and detection by antibodies.

Experiments with His-tagged PG+CA expression in *E. coli* BL21 (DE3) cells produced soluble protein product which could be purified on Ni affinity resin (Figure 19). Nevertheless, no enzymatic activity was detected in the purified fractions

(Table 8). At this point, the prokaryotic expression experiments were not pursued further and we concentrated our effort on eukaryotic expression (as described in section 5.2.2.)

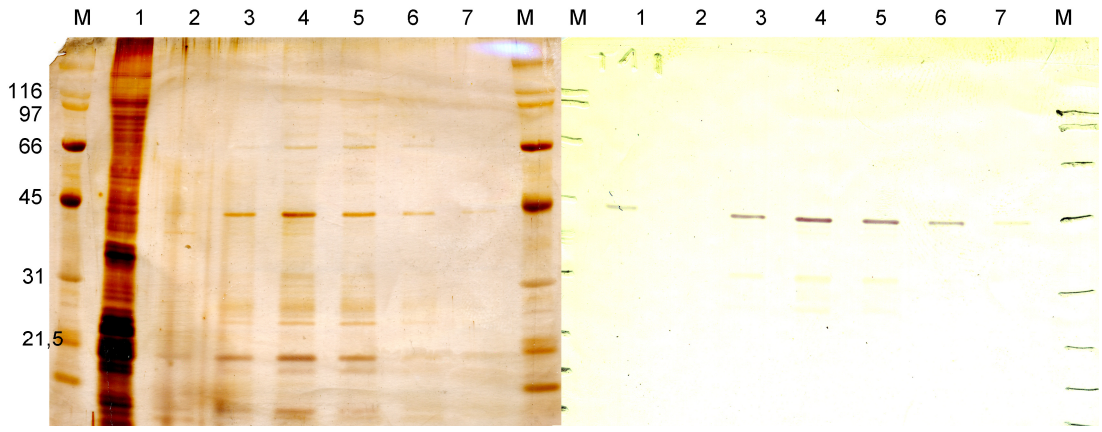


Figure 19 Ni column purification of His-tagged PG+CA (p640). *12.5 % SDS-PAGE and WB (stained with M75-HRP conjugate) lane 1: sample applied on column; lane 2: wash with 5 mM imidazole; lanes 3-7: elution fractions of target protein in 250 mM imidazole.*

5.2.2 Recombinant CA IX constructs for eukaryotic expression

Unsuccessful attempts to obtain sufficient supply of crystallizable recombinant CA IX by prokaryotic expression led us to consider a eukaryotic expression system, which proved beneficial both for the expression yield as well as the enzymatic activity of recombinant product. Thus, with the benefit of hindsight, it is clear that too much effort and trust have been spent on prokaryotic system, based solely on positive results published for CA XII expression (Whittington *et al.*, 2001).

Several recombinant CA IX constructs containing PG+CA domains as well as constructs containing solely the catalytic CA domain were prepared. Due to difficulties with purification of CA IX on commercially available sulfonamide agarose resin, some constructs were engineered with N-terminal or C-terminal affinity tags for purification on Ni-CAM (in case of His-tag) or Streptavidin Mutein Matrix (in case of biotinylated

Avi-tag). The affinity tags were in some constructs separated from the coding sequence by specific cleavage site for tobacco etch virus (TEV) protease to allow tag removal. Constructs with C-terminal extension of CA domain to the transmembrane segment (TM) were also prepared with the aim to investigate potential dimerization via a cystein residue located in sequence directly preceding the TM segment.

The *Drosophila* Expression System (Invitrogen) utilizing Schneider 2 (S2) cells with transfection reagents, expression and selection plasmids, media and additives were used according to the manufacturer's recommendations as described in Methods (chapter 4.3.1.). The expression of CA IX was targeted to growth media and this system produced the recombinant product in several orders of magnitude higher yield than the prokaryotic system. An additional benefit was the expression of correctly folded and enzymatically active protein.

Figure 20 shows sequence of PG+CA in the context of pMT/BiP expression vector (p661 in Table 9) with the relevant restriction sites that were used throughout the cloning of all other constructs. For details of plasmid construction see Methods.

List of prepared constructs for expression in *Drosophila* S2 cells is summarized in Table 9.

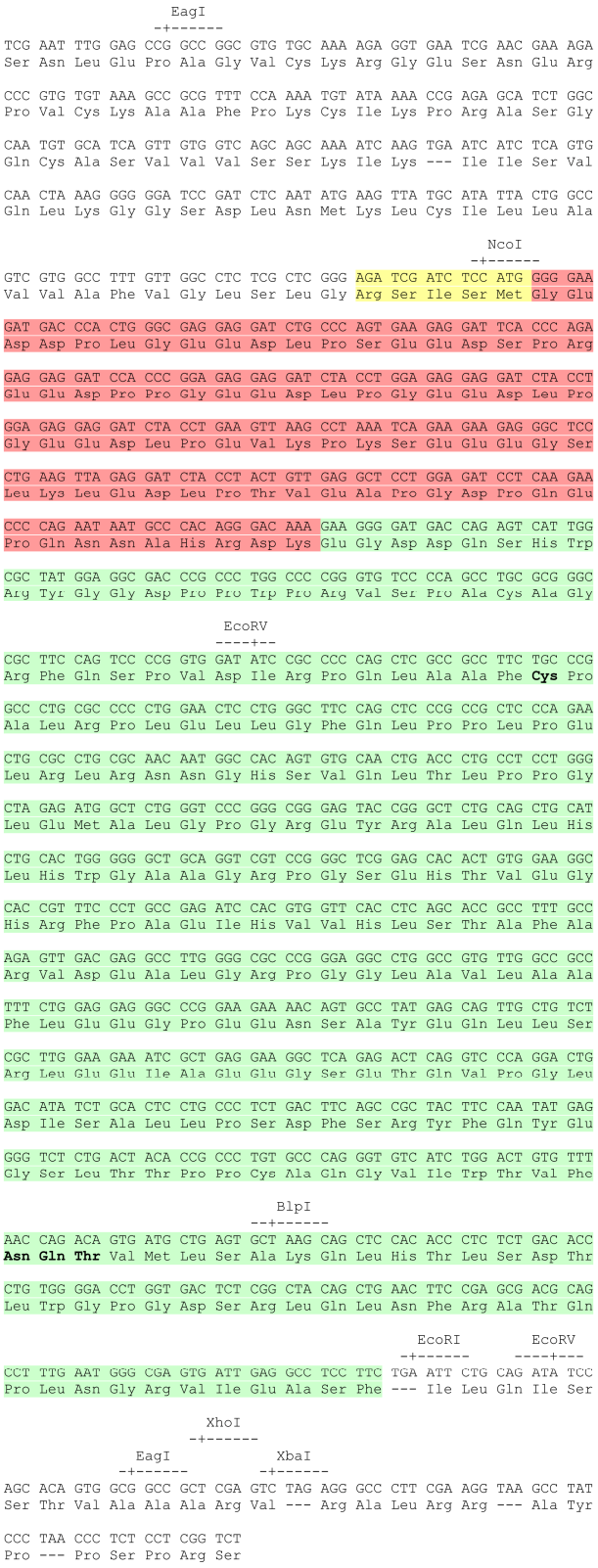


Figure 20 Sequence of PG+CA (p661) construct in pMT/BiP expression vector for S2 cell transfection. *Coding sequence for the PG and CA domains are highlighted in red and green colors, respectively. Restriction sites used throughout the cloning are indicated above the nucleotide sequence.*

Table 9 List of constructs prepared for expression in S2 cells

No.	CA IX Construct[a] in pMT/BiP vector	Construct description, additional N/C-terminal tags, cleavage sites and mutations	MW [kDa][b]	Expression yield[c] [mg.l^{-1}]	Enzymatic activity	Crystallization
p661	residues **55-390**	**PG+CA** domains of CA IX; CA domain with native N-glycosylation signal on Asn346 (NQT); no additional tag	37.4	50	++	Low quality crystals, no diffraction
p758	residues **139-390**	**CA** domain only; Asn346 (native N-glycosylation site **NQT**) no additional tag	28.2	2.2	++	-
p759	residues **139-390**	**CA** domain; **mutation in N-glycosylation signal Asn346Asp** (NQT→**DQT**), no additional tag	28.2	2.5	++	Low quality crystals, no diffraction
p847	residues **139-413**	**CA** domain (NQT) **extended** to TM domain N-terminal **(His)$_6$-tag**+TEV cleavage site	41.4	1.5	+	Construct not for crystallization
p848	residues **139-413**	**CA** domain (**DQT**) **extended** to TM domain N-terminal **(His)$_6$-tag**+TEV cleavage site;	32.3	1.4	+	Construct not for crystallization
p851	residues **139-390**	**CA** domain (NQT) N-terminal **(His)$_6$-tag**+TEV cleavage site	30.1	n/a	n/a	-
p852	residues **139-390**	**CA** domain (**DQT**) N-terminal **(His)$_6$-tag**+TEV cleavage site;	30.1	2.5	+	Low quality crystals, no diffraction
p912	residues **139-390**	**CA** domain N-terminal **Avi-tag**+TEV cleavage site; (**DQT**)	31.4	1.0	+	-
p929	residues **55-390**	**PG+CA** domains; **C-terminal Avi-tag**; (**DQT**)	39.3	0.7	+	Construct not for crystallization
p930	residues **140-390**	**CA** domain **C-terminal Avi-tag**; (**DQT**)	30.1	0.3	+	Construct not for crystallization
p947	residues **139-390**	**CA** domain with **Cys174Ser mutation** N-terminal **Avi-tag**+TEV cleavage site	31.4	1.3	+	-
p952	residues **137-390**	**CA** domain starting with 2 residues upstream (**DQ**): **Cys174Ser mutation;** N-terminal **Avi-tag**+TEV cleavage site	31.5	0.8	+	-
p955	residues **137-391**	**CA** domain starting with 2 residues upstream (**DQ**) and ending with one additional residue (**P**) at C-terminus; **Cys174Ser mutation;** N-terminal **Avi-tag**+TEV cleavage site	31.6	1.0	+	Experiments in progress
p957	residues **55-390**	**PG+CA** domains; **C-terminal Avi-tag**; (**NQT**)	39.3	n/a	n/a	Construct not for crystallization
p958	residues **140-390**	**CA** domain **C-terminal Avi-tag**; (**NQT**)	30.1	n/a	n/a	Construct not for crystallization

[a] Residue numbers are derived from Swiss-prot entry Q16790
[b] Theoretical molecular weight of protein construct without posttranslational modifications (glycosylation).
[c] Expression yield expressed in mg of target protein per 1 l of expression medium.

In a typical expression experiment, the pMT/BiP vector containing the cloned sequence of CA IX construct and selection vector (pCoBlast) were co-transfected using the calcium phosphate procedure as described in section 4.3.2. The cell culture was first cultivated on a selection medium containing 300 µg/ml blasticidine and 10 % fetal calf serum. After approximately 3 weeks of selection of stably transfected clones, the cells were transferred to serum free medium and large scale cultivations were performed in glass spinner flasks (1000 ml, Corning) at 24 °C, 120 rpm.

The cells were usually seeded at density 2 – 4 × 10^6 cell/ml, heterologous expression of genes under metallothionein promoter was induced by adding 1 mM $CuSO_4$ at cell density around 8 × 10^6 cell/ml, and the cells were further cultivated for about 6 days. Typical growth curves of cultivation carried in two spinner flasks in parallel are shown below (Figure 21).

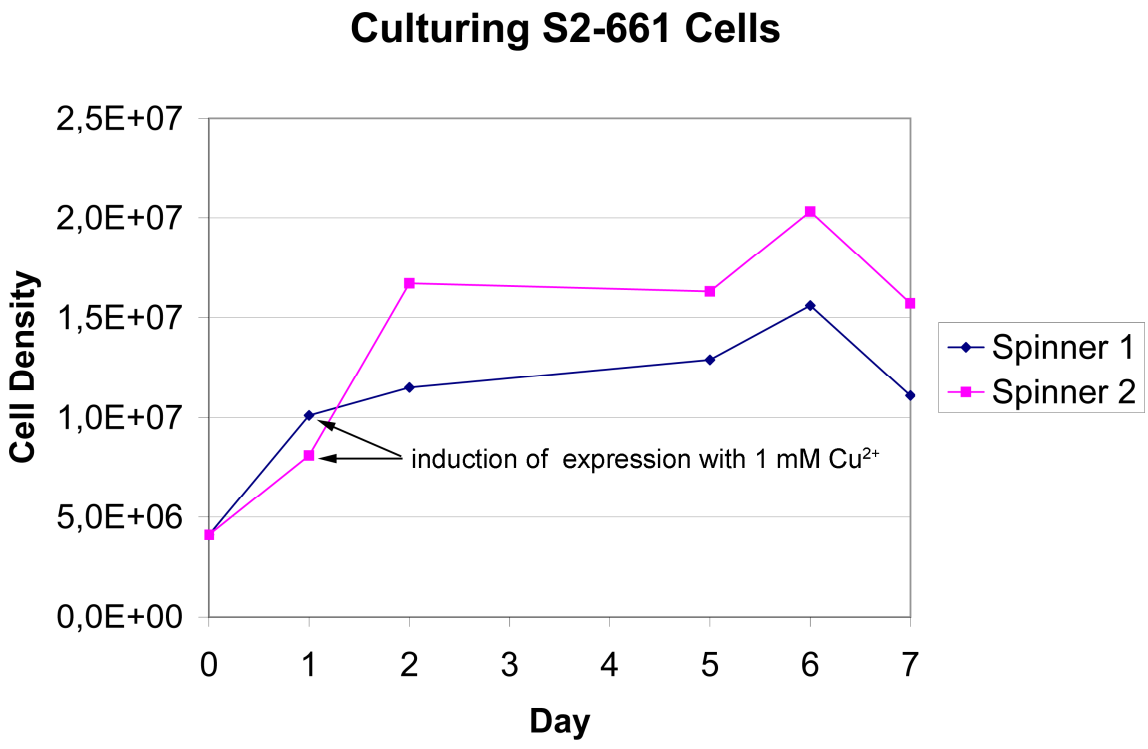


Figure 21 Growth curves of S2 cells expressing recombinant PG+CA (p661) construct. *The induction of recombinant expression with 1 mM CuSO4 is indicated at cell density approx. 8 × 10^6 cell.ml^{-1}.*

The secreted recombinant product accumulated in the serum-free medium and was harvested by centrifugation at the end of cultivation. An example of expressed recombinant CA IX protein in serum-free medium is shown in Figure 22. The conditioned medium was either immediately used for purification of target protein or frozen and stored at -70 °C until further use.

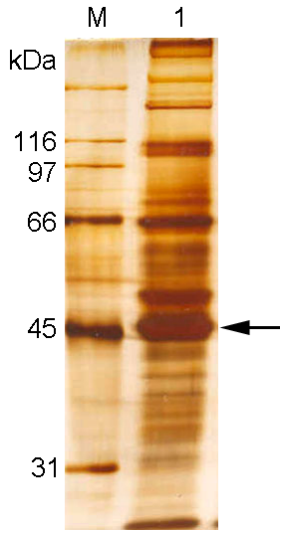


Figure 22 Expression of PG+CA construct (p661) in *Drosophila* S2 cells. *SDS-PAGE (12.5 % gel) analysis of serum-free medium containing protein of expected size (37.4 kDa indicated by arrow). N-glycosylation of NQT motif present in the CA domain of CA IX is the cause of slight increase in molecular weight resulting in decreased electrophoretic mobility.*

Purification on sulfonamide agarose was done according to the protocol described in Methods section 4.4.1, one such typical chromatography is depicted in Figure 23. In general, this method gave rather unsatisfactory results as far as the selectivity and specificity is concerned. Several contaminants from the medium bound non-specifically to this resin, and as shown on the SDS-PAGE below, the eluted protein was not much purified even after elution with relatively high salt concentration.

As concerns specific elution of CAs with the use of free sulfonamide (acetazolamide), this gave generally even worse results. In the end we have found that standard purification with ion exchange chromatography gave much better and reproducible results, and also the yield of eluted protein was higher.

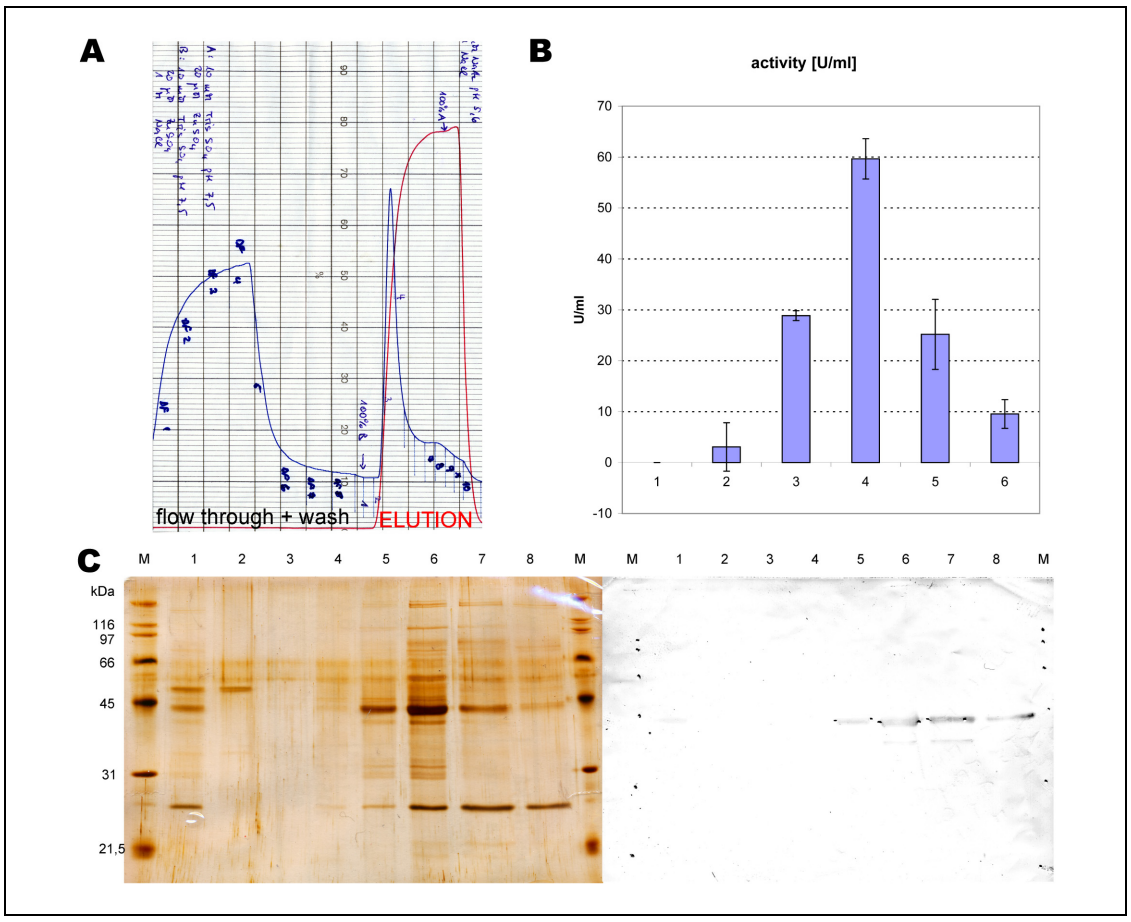


Figure 23 Purification of PG+CA (p661) on sulfonamide agarose. *(A) Chromatogram of purification on sulfonamide-agarose (elution with gradient of 0-1 M NaCl) (B) CA activity in individual elution fractions (C) On the left 12.5 % SDS-PAGE Lane 1: sample applied on the column; 2: flow-through fraction; 3 – 8: elution fractions 1-6 from above purification. On the right: the same samples on WB stained with mAb M75-HRP conjugate.*

Further purification was carried out using either ion exchange chromatography on MonoQ GL 5/50 column (GE Healthcare) and/or by gel filtration on Superdex 200 GL (GE Healthcare). Before crystallization trials, the samples were concentrated by ultrafiltration to 10 mg/ml in a buffer composed of 10 mM TrisCl pH 7.5 and 100 mM NaCl. Crystallization trials were performed using vapor diffusion hanging drop method with several commercially available crystal screening kits as described in the methods.

Some examples of partly successful hits are shown in Figure 24. Unfortunately, extensive optimization of initial crystallization conditions did not lead to any crystals usable for diffraction analysis and structure determination. Either the size of the crystal

or the crystalline order was never sufficient to obtain X-ray diffraction (neither using the home X-ray source nor the highly intense synchrotron source).

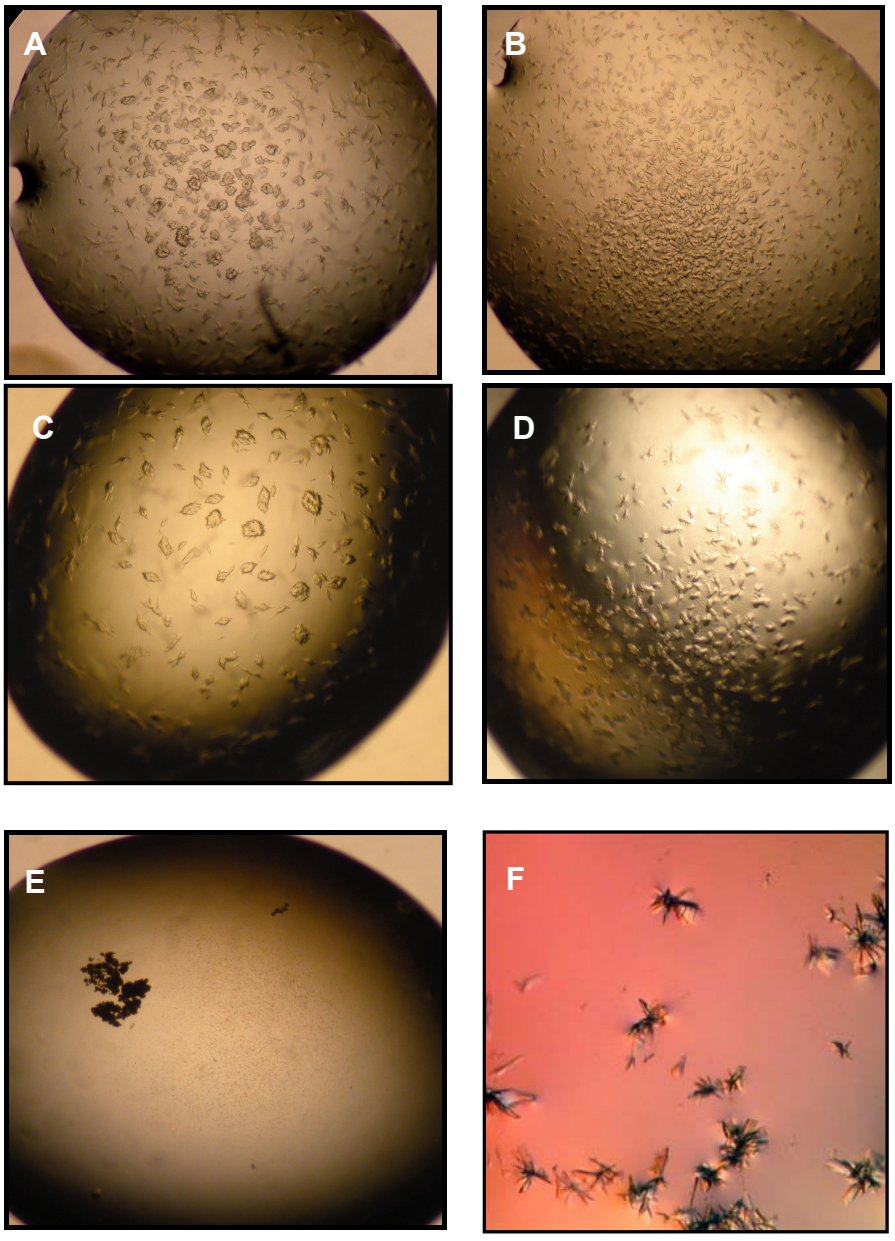


Figure 24 Some crystallization hits obtained with various rCA IX constructs. *Panels (A), (B), (C),and (D) 10 mg/ml PG+CA (p661) in 0.2 M Magnesium acetate; 0.1 M Sodium Cacodylate pH 6.5; 20 % PEG 8000 (micro-variation in pH and seeding experiments); (E) same construct in 0.2 M Calcium Chloride dihydrate, 0.1 M HEPES pH 7.5, 28 % PEG 400; (F) 8 mg/ml CA (p759) (1 M ammonium sulfate; 100 mM HEPES pH 7.5).*

In the course of optimizing our crystallization conditions a paper reporting crystal structure of CA IX was published (Alterio *et al.*, 2009). This led us to conclude our crystallization attempts at this stage and admit that we did not succeed in this highly competitive field and it was not for us to announce the "long-sought-for" CA IX structure.

5.2.3 Conclusion from structural studies of CA IX

Efforts to produce recombinant human CA IX of crystallization quality in prokaryotic or eukaryotic expression systems, which comprised in total 19 different gene constructs, resulted either in too low yields or in proteins unable to grow as crystals needed for X-ray analysis.

From discussions with colleagues working in the CA field we have learned that our experience was not unique and that others encountered similar pitfalls, e.g. difficulties with affinity chromatography on sulfonamide agarose. Other groups have developed protocols for carbonic anhydrase purification by using metal affinity resins of their own, giving better results after one step purification (Banerjee *et al.*, 2004).

After more than a decade of unsuccessful efforts by many teams, including big pharmaceutical companies, by the end of year 2009 a paper announcing the first 3D structure of CA IX was published (Alterio *et al.*, 2009). Based on the exact sequence of the CA domain (N- and C- termini) in their construct, including the Cys(41)Ser mutation, which finally led to well crystallizable CA IX in the Alterio *et al* paper, we have also designed such a CA construct (designated with internal number p955, Table 8) and we plan to use it in structure determination experiments.

We also are still interested in the role and structure of the PG domain, so we have not given up yet on crystallization of PG+CA construct either.

5.3 Structure of mouse mAb M75 in complex with epitope peptide

Monoclonal antibody M75 recognizes its specific linear epitope PGEEDLP in the proteoglycan-like segment (residues 38-112 in Swiss-prot entry Q16790) of CA IX (Zavada *et al.*, 2000). In addition to its enzymatic function, CA IX also acts as a cell adhesion molecule (CAM) that can mediate the attachment of cells to a nonadhesive solid support (Zavada *et al.*, 2000). The adhesion site of CA IX appears to overlap with the PGEEDLP epitope of the monoclonal antibody M75 since the antibody blocks attachment of cells to immobilized CA IX protein (Zavada *et al.*, 2000). The counterpart receptor component of the adhesion cell remains unknown. Nevertheless, it might be assumed that similar structural elements are involved in CA IX binding with both M75 antibody and with the CAM receptor, since formation of both complexes is inhibited in an acidic environment in the same manner (Zavadova & Zavada, 2005).

As this PG-like segment represents a unique feature of CA IX compared to other CA isozymes, this segment could be important in oncogenesis. Therefore the nature of its interactions with M75 antibody was investigated as one of the aims of this thesis. Other motivation for this study was to obtain structural information necessary for further protein engineering of M75 antibody with diagnostic and therapeutic potential.

We set to determine and analyze crystal structure of M75 Fab fragment alone and in complex with an epitope peptide representing part of the PG-segment.

5.3.1 Crystallization, data collection and structure determination

Fab fragment of antibody M75 for crystallization was prepared by the cleavage of isolated immunoglobulin M75 by papain as described in Methods (sub 4.6). The Fab fragment was concentrated by ultrafiltration to concentration 10 mg/ml in buffer composed of 100 mM potassium phosphate, pH 7.2, 0.13% 2-mercaptoethanol, and 0.9 mM EDTA and subjected to crystallization screening. Crystallization cocktails containing polyethylene glycol (PEG) of various molecular weight (MW 1000 – 6000)

were screened to find initial crystallization condition for free Fab M75 as well as the Fab M75 in complex with the epitope peptide PGEEDLPGEEDL. The hanging drop vapor diffusion method was used for screening and optimization as described in Methods (sub 4.7.1): drops containing 1 µl of protein solution were mixed with 1 µl of crystallization cocktail and placed over a 1 ml reservoir.

Optimal crystals for free Fab M75 were obtained under the following crystallization conditions: Fab M75 at 10 mg/ml; 100 mM Tris-HCl, pH 7.5; 15% PEG 4000; 18 °C. Crystals reached their final size within 1 week (Figure 25A).

Crystallization conditions for the Fab M75 + epitope peptide complex were: Fab M75 at 7.8 mg/ml; peptide at 10 mg/ml; 100 mM Tris-HCl, pH 7.4; 20% PEG 2000; 18 °C (Figure 25B). Streak seeding technique was utilized for obtaining the highest quality crystals of the complex, (Figure 25 C, D).

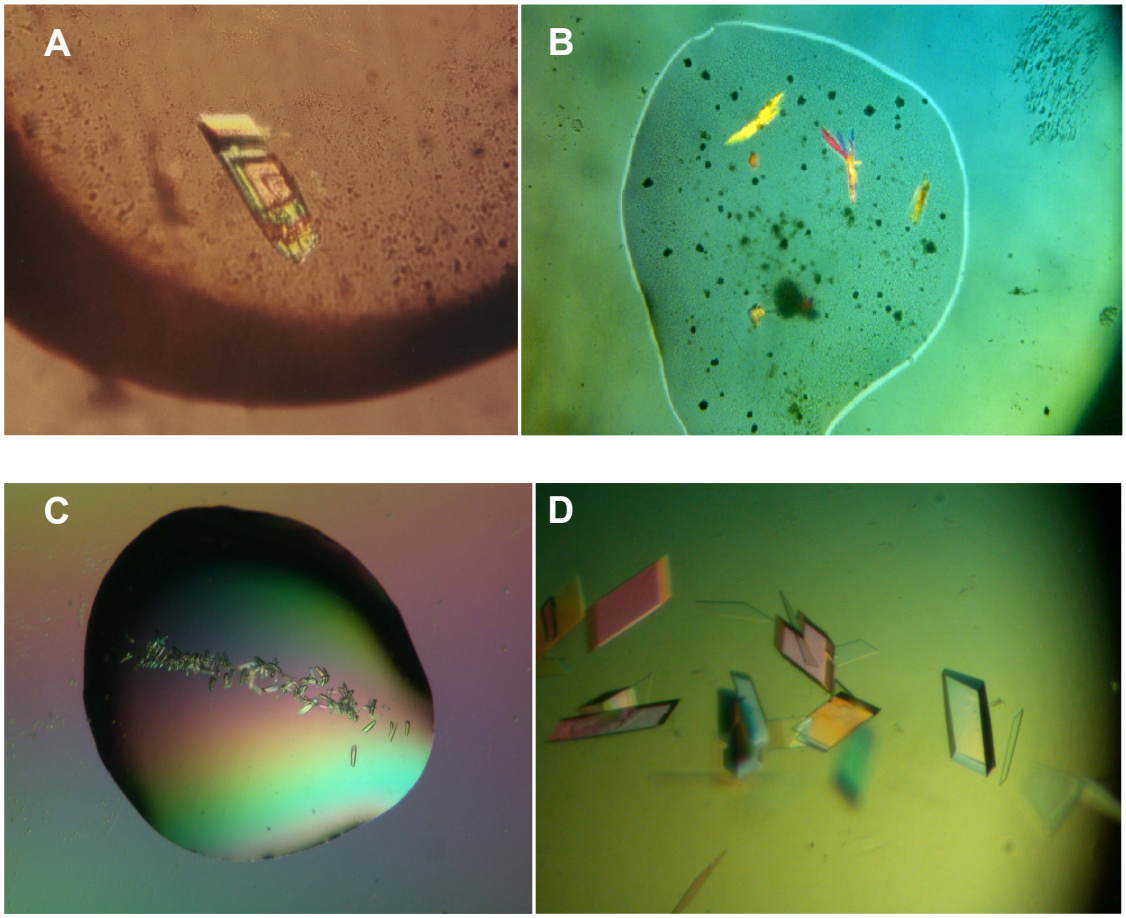


Figure 25 Crystals of Fab M75.
(A) Optimized crystals of free Fab M75; the crystal size is ~ 0.3 × 0.1 × 0.05 mm. (B) Initial crystals of M75 + epitope peptide complex whose quality was improved using strike seeding (C). (D) Optimized crystals of M75 + epitope peptide complex; the crystal size is ~ 0.25 × 0.1 × 0.05 mm. Crystallization conditions are detailed in the text.

For cryoprotection, the crystals were soaked in corresponding reservoir solution supplemented with 25% glycerol. Diffraction data for both crystals were collected at 150 K using an in-house diffractometer (Nonius FR 591 generator, 345-mm MarResearch Image Plate detector). Crystals belonged to triclinic spacegroup P1. Diffraction data were processed using procedures described in the Methods (sub 4.7.2). Crystal parameters and data collection statistics are summarized in Table 10.

The structure of free Fab M75 was solved by molecular replacement by EPMR program (Kissinger *et al.*, 1999) using the structure of Fab Bv04-01 (PDB code: 1NBV

(Herron *et al.*, 1991) as a search model. Structure factors in resolution range 15–4 Å were used for both, rotational and translational searches. In the rigid body refinement (Brunger *et al.*, 1998) individual subdomains light chain variable domain (V_L), heavy chain variable domain (V_H), and heavy and light chain constant domains C_H and C_L were allowed to move independently. In the next step of refinement, all the hypervariable complementarity determining regions (CDRs) were omitted from the initial model, which was then subjected to alternate cycles of molecular dynamics refinement with manual model building using the program Coot (Emsley & Cowtan, 2004). For final refinement cycles, TLS (translation-libration-screw) refinement cycles were used with subdomains V_L, C_L, V_H, and C_H as independent TLS groups. Final refinement statistics is summarized in Table 10, coordinates and structure factors were deposited to PDB under the accession code 2HKH.

The crystal of the Fab M75 + epitope peptide complex was isomorphous with the structure of free Fab M75 (see above), therefore, the structure of free Fab M75 could be used as a starting model for rigid body refinement with free movements of V_L, C_L, V_H, and C_H subdomains. Peptide tracing was carried out using electron density maps calculated with weighted Fourier coefficients. (Read, 1986). For final stage refinement of V_L, C_L, V_H, and C_H subdomains, program REFMAC 5.2.0019 with TLS cycles was used. Final refinement statistics is summarized in Table 10, coordinates and structure factors were deposited to PDB under the accession code 2HKF.

Table 10 Crystal parameters, data collection and refinement statistics for free Fab M75 and for its complex with epitope peptide.

Data collection statistics	**Free Fab M75**	**Complex Fab + epitope peptide**
Space group	P1	P1
Unit cell parameters (Å)	a = 40.98 Å b = 43.15 Å c = 57.37 Å α = 85.30° β = 88.45° γ = 84.29°	a = 40.39 Å b = 43.15 Å c = 58.02 Å α = 85.03° β = 88.47° γ = 85.67°
Number of molecules in AU	1	1
Wavelength (Å)	1.5418	1.5418
Resolution range (Å)	30.00 – 2.08 (2.18 – 2.08)	28.87 – 2.00 (2.11 – 2.00)
Number of unique reflections	22,667	24,418
Redundancy	2.52	2.40 (2.30)
Completeness (%)	90.7 (68.6)	93.5 (93.5)
R_{merge} [a]	0.069 (0.304)	0.049 (0.191)
Average I/σ(I)	6.5 (2.0)	12.0 (3.2)
Wilson B ($Å^2$)	39.4	19.4
Refinement statistics		
Resolution range (Å)	30.00 – 2.10 (2.18 – 2.10)	28.87 – 2.00 (2.11 – 2.00)
No. of reflections in working set	15,795	23,175
No. of reflections in test set	1,417	1,243
R [b] (%)	21.9	17.5
R_{free} [c] (%)	27.2	24.2
RMSD bond length (Å)	0.011	0.012
RMSD angle (°)	1,547	1.38
Number of atoms in AU	3,381	3,663
Number of protein atoms in AU	3,279	3,360
Number of solvent molecules in AU	102	303
Mean B value ($Å^2$)	31.07	17.78
Ramachandran plot statistics		
Residues in favored regions (%)	95.48	97.43
Residues in allowed regions (%)	3.81	2.34
PDB code	2HKH	2HKF

The data in parentheses refer to the highest-resolution shell.

[a] $R_{merge} = \Sigma_{hkl}\Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl)$, where the $I_i(hkl)$ is an individual intensity of the *i*th observation of reflection hkl and $\langle I(hkl)\rangle$ is the average intensity of reflection *hkl* with summation over all data.

[b] R-value = $||F_o| - |F_c|| / |F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively.

[c] R_{free} is equivalent to R value but is calculated for 5 % of the reflections chosen at random and omitted from the refinement process (Brunger, 1992).

5.3.2 Refined models and overall structures

The final Fab M75 model, refined at 2.1 Å resolution comprised a single protein molecule in the asymmetric unit with 219 residues for the light chain and 208 residues for the heavy chain, 98 water molecules and one molecule of glycerol. Residues belonging to the hypervariable CDR loops were well-defined in the electron density map, except for five residues of CDR-H3, namely H100–H104. Other missing residues in the model included N-terminal GluH1, and residues belonging to the constant region C_H1 (H134–H137 and H162–H166). Residues falling out of Ramachandran (Ramachandran & Sasisekharan, 1968) energy space are ValL56 (which is characteristic for the canonical conformation of CDR-L2, (Milner-White *et al.*, 1988), SerH27, and ThrH28 from CDR-H1 solvent exposed loop; generously allowed are SerH155 and SerH178 surface residues of C_H1 loops.

The final model of Fab M75 + epitope peptide, refined at 2.0 Å resolution, comprises one complex molecule in the asymmetric unit: 219 residues for the light chain, 209 residues for the heavy chain, 9 peptide residues, and 303 water molecules. Overall view of the Fab M75 + epitope peptide is depicted in Figure 26.

Figure 26 Structure of Fab M75 antibody fragment in complex with epitope peptide.
Heavy chain is colored blue, light chain red, peptide yellow, the respective CDR loops are colored pale blue for H1, green H2, gray H3, light magenta L1, mauve L2, dark magenta L3

Similarly to the free Fab M75 structure, ValL56 falls into the disallowed region of the Ramachandran energy space, and SerH155 into the generously allowed region. All residues belonging to the hypervariable loops were well-defined in the electron density map. The residues missing from the protein model belong to the C_H1 region. Eight out of 12 peptide residues (PGEEDLPG) could be conclusively traced into well defined electron density map in the antigen binding site Figure 27.

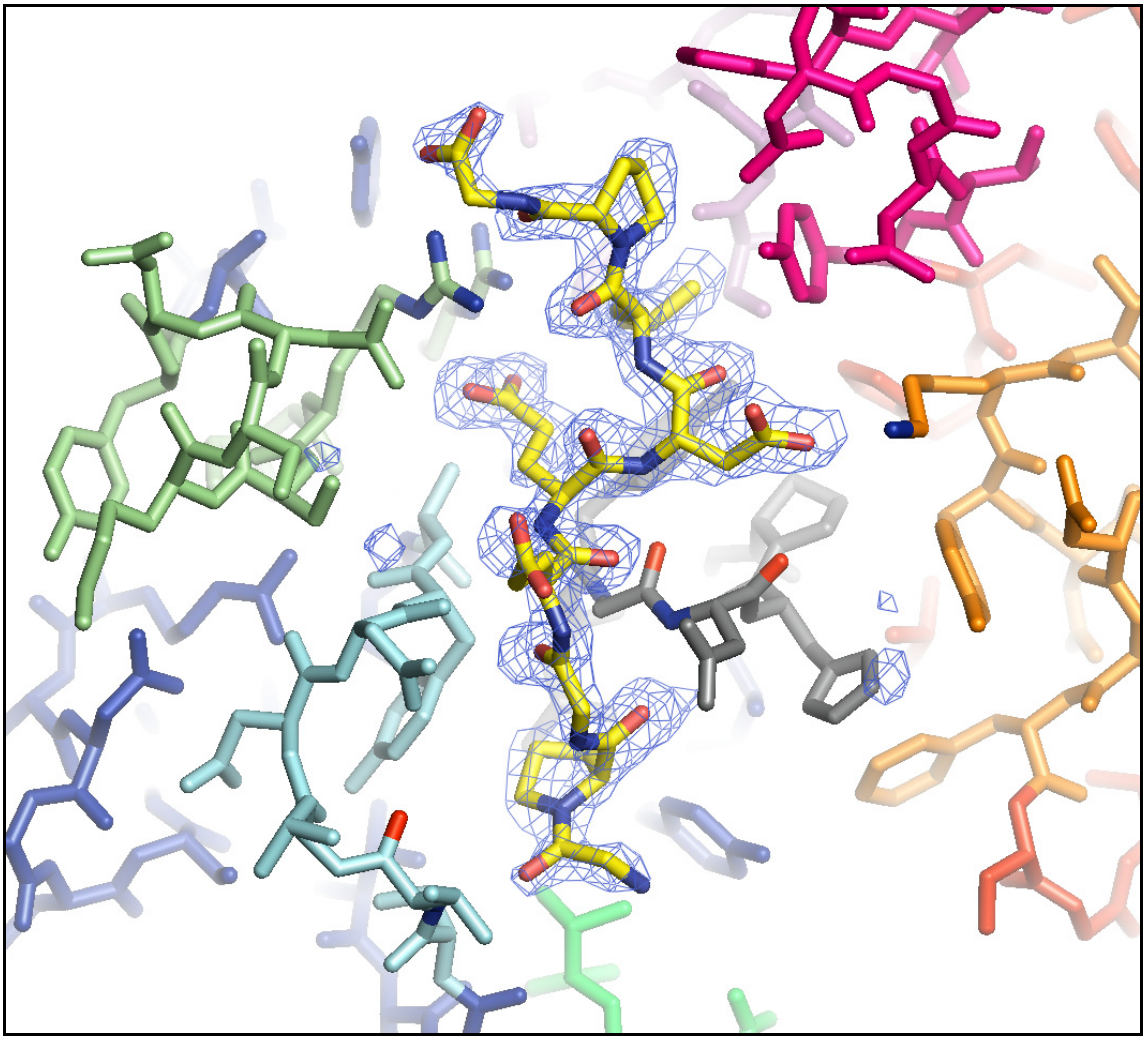


Figure 27 Structure of the complex Fab + epitope peptide: detail view of binding site.
Omit map of the bound peptide (yellow) contoured at 3σ level. The respective CDR loops are colored pale blue for H1, green H2, gray H3, light magenta L1, mauve L2, dark magenta L3. Peptide is represented by stick model with carbon atoms colored yellow and oxygen and nitrogen atoms shown in red or blue, respectively.

During the map interpretation, an extra electron density was observed preceding the amino end of the traced peptide that could be indisputably traced as a main chain of additional amino acid residue and identified as a leucine. We believe that the final crystallographic model represents an average of alternate epitope peptide binding by its amino-terminal segment (**PGEEDLP**GEEDL) and by the carboxy-terminal segment (PGEEDL**PGEEDL**), respectively (relevant parts of the epitope peptide are shown in bold). In the deposited complex structure (PDB entry 2HKF), the bound ‘‘averaged’’ peptide has nine residues and its first residue is approximated with glycine. The values of atomic displacement parameters (ADP) for the peptide bound to Fab indicate higher stability of the central peptide part (ProP2 to LeuP7) compared to the peptide terminal residues (Figure 28). The ambiguity of binding mode of the peptide PGEEDLPGEEDL clearly shows ability of binding site to bind repetitive sequence GEEDLP, which is repeated 4 times in the PG domain of CA IX.

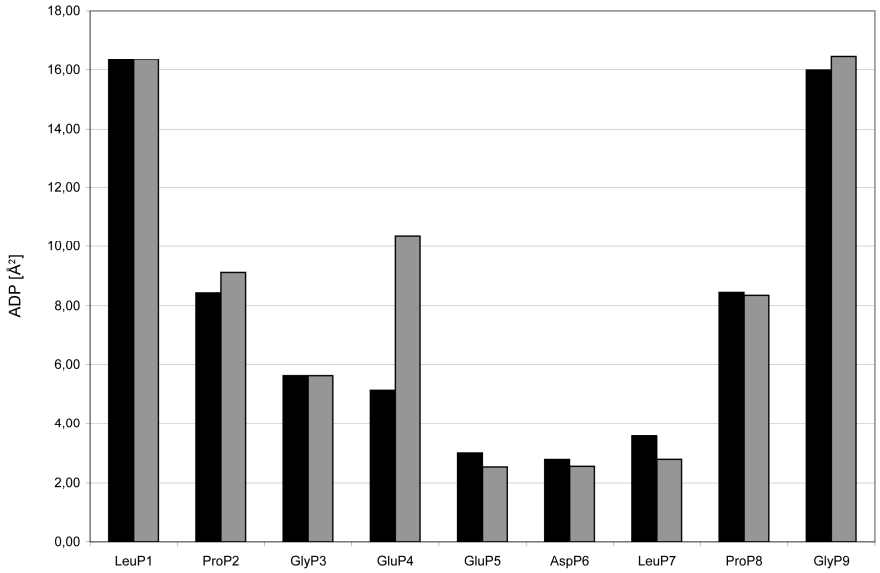


Figure 28 Average values of residual atomic displacement parameters (ADP) of the peptide residues in complex.
Black bars correspond to average ADP values of main chain atoms of the peptide residue and grey bars to all atoms of the residue. Note the difference in the ADP of GluP4 main chain in respect to whole residue ADP. The side-chain of GluP4 is not involved in any contacts with the antibody molecule.

5.3.3 Comparison of Fab M75 free and complexed structure

Comparison of free and complexed structures revealed that major transitions of the antibody main chains upon epitope peptide binding occur in the regions of CDR-H1, CDR-H3, and CDR-L1 (Figure 29).

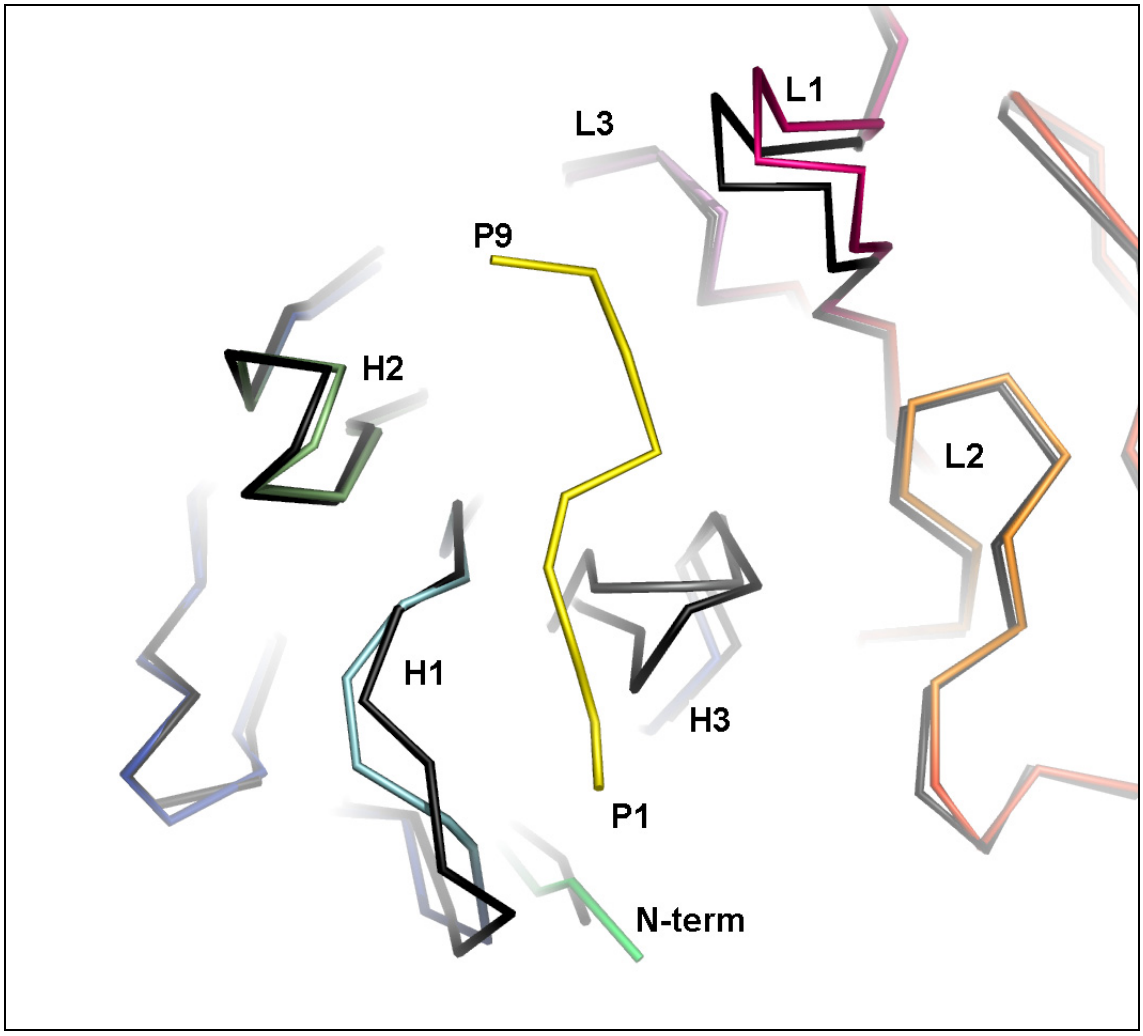


Figure 29 Superimposed Cα traces of the free and complexed Fab M75.
A ''top view'' into the binding cleft. The peptide and CDRs of the complex are color coded as in Figure 27; the ''free'' structure is colored black.

While the hypervariable loops of CDR-H1 and CDR-H3 acquire altered conformations, the entire CDR-L1 is shifted to a new position as almost rigid structure without changing the overall loop conformation. Upon the binding of the epitope

peptide, the conformation of the CDR-H1 loop is altered between residues GlyH26 and LeuH29. The loop is moving in the direction toward the bound epitope peptide with the maximum shift observed for the SerH27 Cα atom (3.06 Å). The conformational change in the CDR-H3 loop involves rearrangement of TyrH101-AsnH103 residues. The H3 loop moves toward the C-terminal part of CDR-L2, and provides the contacts to the epitope peptide from the L2 direction. The most dramatic movement is observed for GlyH102 Cα atom (3.68 Å). The new conformation of the CDR-H3 loop opens the groove in the antigen binding cleft for the binding of the N-terminal part of the epitope. The displacement of the hypervariable loop of CDR-L1 is smaller, with amplitude of 1.98 Å observed for the residue SerL32 Cα atom.

From comparison of the two crystal structures (free and complexed), it can be also noted that the structure of Fab M75 + epitope peptide complex has smaller number of disordered residues, slightly better resolution (2.0 Å vs. 2.1 Å), and a lower average atomic displacement parameter (ADP). The decrease in average ADP value from 31.1 $Å^2$ for free Fab to 17.8 $Å^2$ for the complex indicates a global stabilization of the complexed structure. Comparison of local (per residue) ADP values along the protein chains indicates one short segment with pronounced differences between the free and the complexed structure. The plot of ADP values for Cα atoms along the heavy chain (Figure 30) shows a distinctive peak for the CDR-H3 segment of the free antibody. This peak is absent in the corresponding plot of the complex structure ADPs, while the two plots are essentially identical for other heavy chain segments. In summary, two transitions in the heavy chain of two CDRs are observed, both generating shape complementarity, whereas only the readjustment of CDR-H3 contributes to local stability.

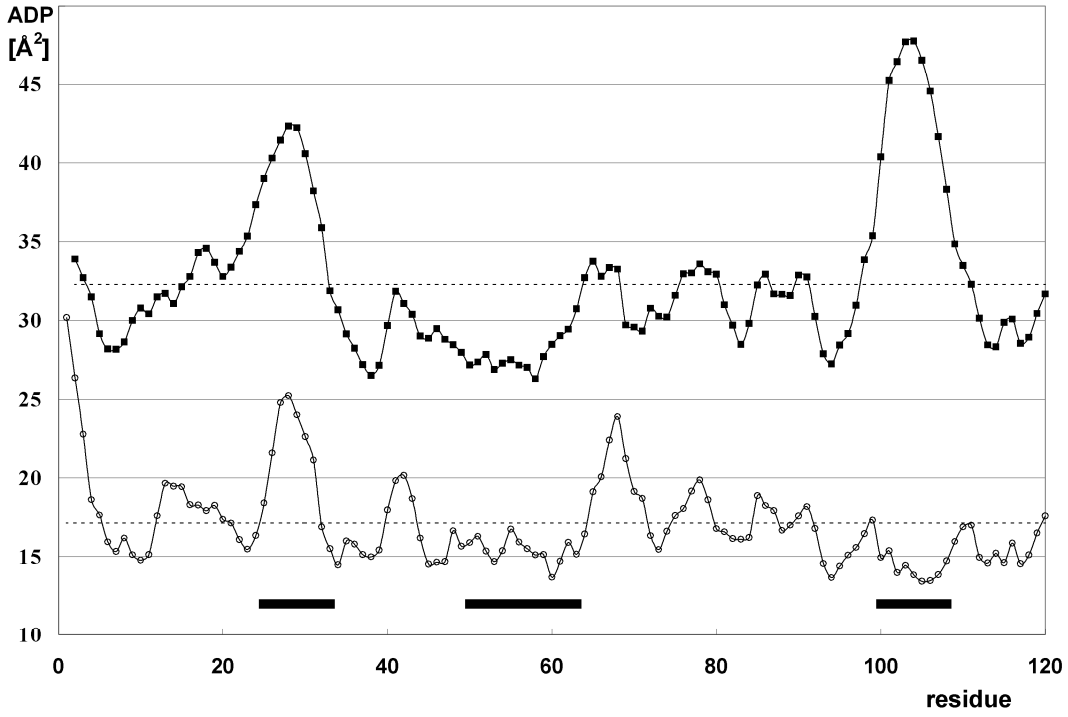


Figure 30 Residual atomic displacement parameters plotted for Cα atoms of the heavy chain residues. *(Free Fab, full squares; complex, open circles) Horizontal bars indicate CDRs H1, H2, and H3. Two dashed abscissa lines show the value of the average atomic displacement parameter for the free and complexed Fab, respectively.*

From the structure comparison it is also evident that some tightly bound water molecules are displaced upon peptide binding, such as the water molecule originally positioned between the CDR-H2 and CDR-L1 loops, see Figure 31.

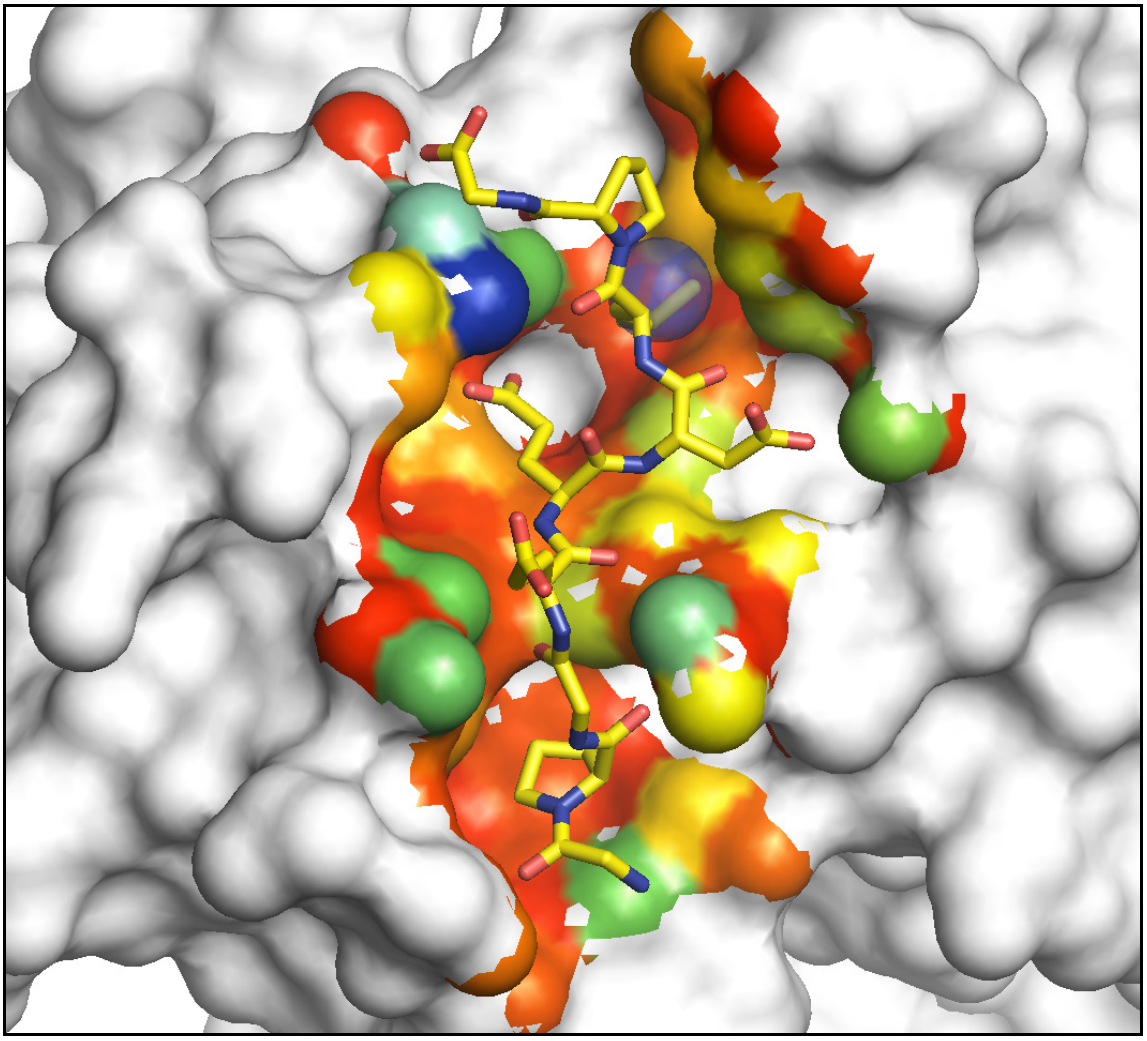


Figure 31 Protein surface at van der Waals radii colored by the protein surface areas buried upon epitope peptide binding. *Rainbow coloring from yellow (< 0 Å2) to blue (40 Å2); white (0 Å2); transparent blue sphere indicates a water molecule displaced upon the peptide binding. Peptide is represented by stick model with carbon atoms colored yellow and oxygen and nitrogen atoms shown in red or blue, respectively.*

5.3.4 Peptide-antibody interactions

5.3.4.1 Polar interactions

In general terms, antigen-antibody interactions reflect shape and charge complementarities. Charge (electrostatic) complementarity is assumed to be a prominent factor in specific antigen recognition (Rini *et al.*, 1992). In the structure of

the complex described here, the antigen binding cleft displays a continuous positively charged surface Figure 32 and the charge complementarities between antibody and peptide include the side chains of ArgH50 and ArgH52 of Fab M75 as counterparts of the GluP5 residue of the epitope peptide and the side chain of LysL55 of Fab M75 as the counterpart to the AspP6 residue of the epitope peptide.

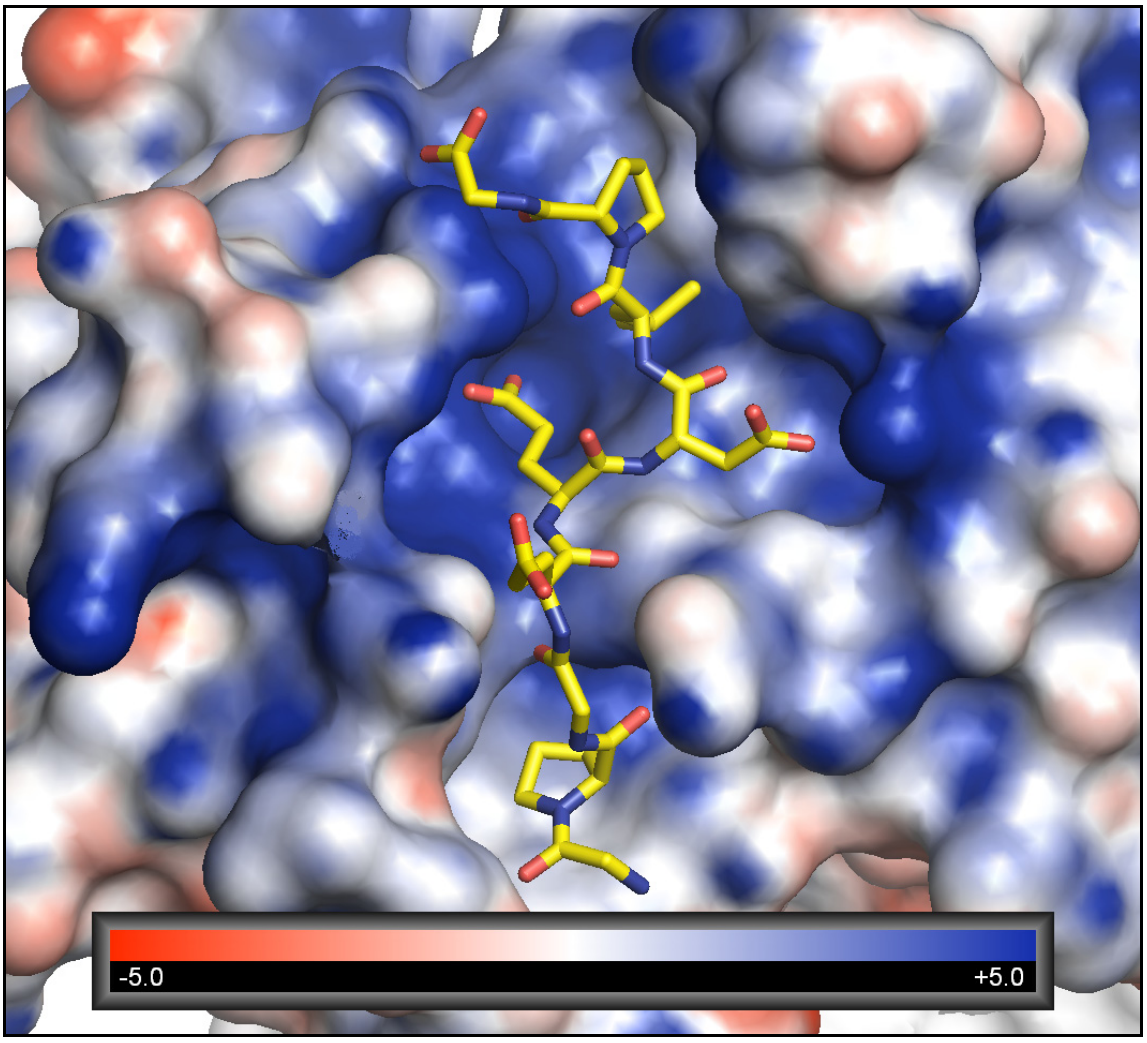


Figure 32 Electrostatic potential of the FabM75 the antigen binding site. *Protein surface at van der Waals radii are colored by the electrostatic potential from negative (red) to positive (blue); the scale bar is from -5 to +5 kT/e. Peptide is represented by stick model with carbon atoms colored yellow and oxygen and nitrogen atoms shown in red or blue, respectively.*

The two successive carboxylic amino acid residues of the epitope (GluP5, AspP6) form salt bridges with the positively charged residues of CDR-H2 and CDR-L2 (see Figure 33).

The side chain of the remaining carboxylic amino acid residue of the epitope (GluP4) is oriented toward the solvent.

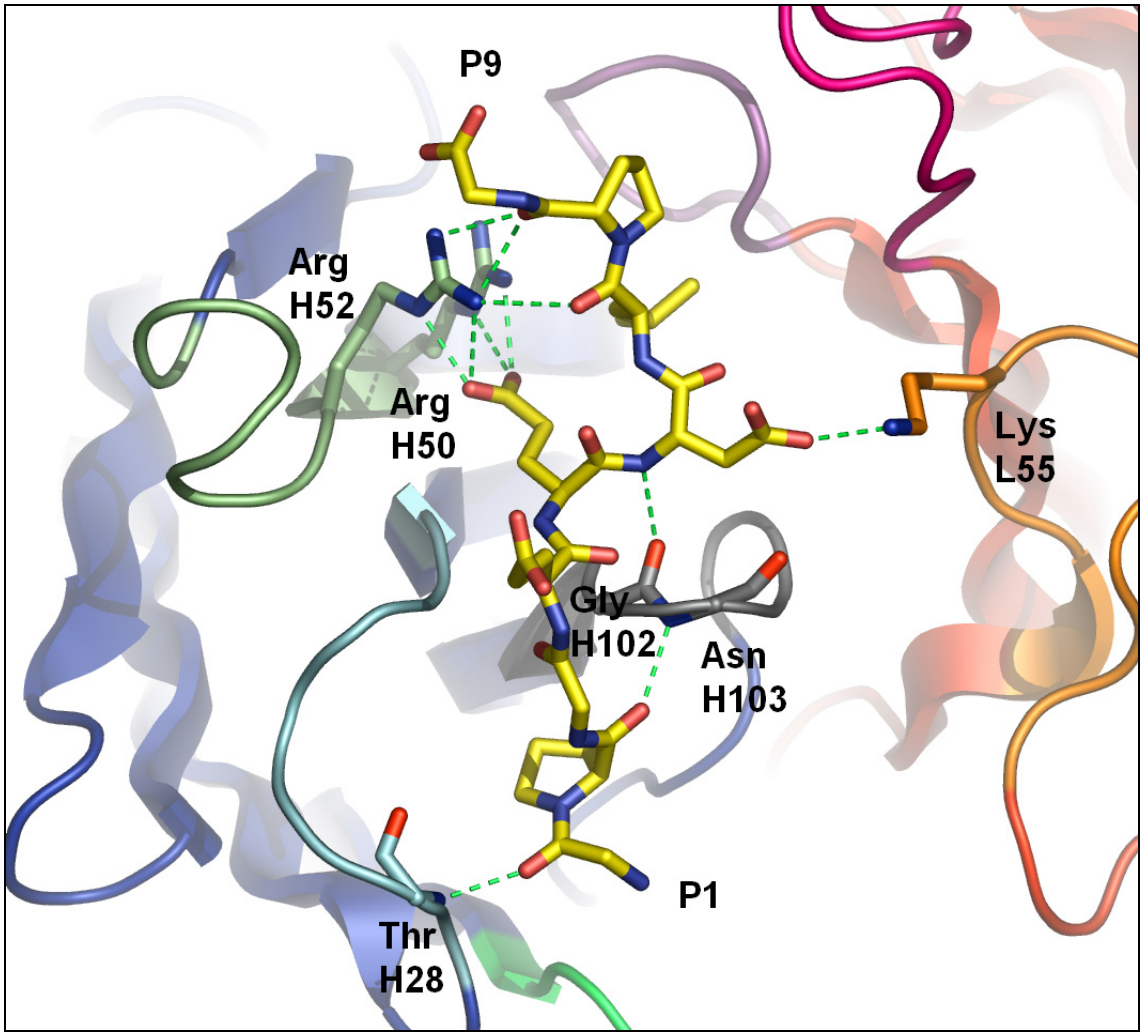


Figure 33 Polar interactions of the FabM75 antigen binding site with epitope peptide. *The interacting polar interactions are shown as green dotted lines. CDR loops are colored pale blue for H1, green H2, gray H3, light magenta L1, mauve L2, dark magenta L3. Peptide is represented by stick model with carbon atoms colored yellow and oxygen and nitrogen atoms shown in red or blue, respectively.*

Comparison of mutual positions of positively charged antibody residues revealed that distances [ArgH50:NH2 - LysL55:NZ] and [ArgH52:NH2 - LysL55:NZ] are very similar in free Fab and the complexed structure (differing by 0.81 and 0.07 Å, respectively). This indicates that the epitope-binding amino acid residues of Fab CDR-H2 and CDR-L2 are not involved in an antibody structural rearrangement associated with the antigen binding. From the point of view of charge complementarities, the antibody thus forms a properly established binding template without any prerequisite of adaptation.

Contacts of the Fab ArgH50, ArgH52, and LysL55 residues and the carboxylic amino acid residues of the epitope peptide can be also analyzed with respect to hydrogen bonding parameters. All polar contacts are listed in Table 11.

Table 11 List of polar contacts between FabM75 and epitope peptide.

Atom of peptide	Atom of Fab M75 (CDR loop)	Distance (Å)
GlyP1 O	ThrH28 N (H1)	3.03
ProP2 O	AsnH103 N (H3)	2.85
GluP4 O	AsnH103 N (H3)	3.07
GluP4 Oε1	ArgH52 Nε (H2)	2.79
GluP4 Oε1	ArgH52 Nη2 (H2)	3.11
GluP4 Oε2	ArgH50 Nε (H2)	3.09
GluP4 Oε2	ArgH50 Nη2 (H2)	2.89
AspP6 N	GlyH102 O (H3)	2.91
AspP6 Oδ1	LysL55 Nη (L2)	2.65
ProP8 O	ArgH52 Nη1 (H2)	3.14
ProP8 O	ArgH52 Nη1 (H2)	3.00

Figure 33 and Table 11 show that there are five contacts of the Fab ArgH50, ArgH52, and LysL55 residues with the carboxylic amino acid residues of the epitope, matching the criteria of hydrogen bonding distances and angles. Other hydrogen bonds involve ThrH28 of CDR-H1 and the amino terminal residue of the peptide (denoted as GlyP1, see above), three hydrogen bonds between ArgH52 of CDR-H2 and ProP8/GlyP9, and four hydrogen bonds between residues GlyH102 and AsnH103 of CDR-H3 and ProP2, GluP4, AspP6 (three of them are of ‘‘main chain-to-main chain’’ type). In total, heavy chain residues form 10 (out of 11) hydrogen bonds to the epitope peptide.

5.3.4.2 Van der Waals interactions

The complexed Fab M75 structure displays areas of extensive van der Waals contacts between Fab and the epitope peptide. The heavy chain residues are responsible for majority of these contacts (80%). As shown in Figure 31, the protein surface areas buried upon binding of the epitope peptide are approximately four times larger for the heavy chain than for the light chain. The contributions of individual heavy chain CDRs are not substantially different and decrease in the order CDR-H2 > CDR-H3 > CDR-H1. Interestingly, a van der Waals contact is also found between the amino terminal residue of the constant domain of the heavy chain, ValH2, and ProP2 of the epitope peptide.

5.3.5 Conclusion from Fab M75 structural analyses

Crystal structure analyses presented here were one part of a wider study which combined various techniques to characterize recognition binding of antibody M75 to its epitope in CA IX. The results of structural studies presented in this thesis have become part of a publication Kral, Mader *et al.*, 2008. The combination of X-ray crystallography, ITC experiments, and MD simulations allowed us to assess the key structural parameters responsible for Fab M75 antibody epitope recognition and binding.

We observed good correlation between structural elements and thermodynamic parameters of the association of antibody fragment to the epitope. Comparisons of the crystal structures of free Fab M75 and its complex with the epitope peptide reveal major readjustments of CDR-H1 and CDR-H3. In contrast, the overall conformations and positions of CDR-H2 and CDR-L2 remain unaltered, and their positively charged residues may thus present a fixed frame for epitope recognition. Adoption of the altered CDR-H3 conformation in the structure of the complex is accompanied by an obvious local stabilization.

Isothermal titration microcalorimetry experiments revealed a highly unfavorable entropy term, which can be attributed mainly to the decrease in the degrees of freedom of the system, the loss of conformational freedom of peptide and partially to a local stabilization of CDR-H3. Moreover, it was observed that one proton is transferred from the environment to the protein-ligand complex upon binding. Molecular dynamics simulations and calculations of the ligand (epitope peptide) binding energy yielded energy values that were in agreement with the ITC measurements and indicated that the charged residues play crucial role in the epitope binding. Theoretical arguments indicate that two adjacent arginine residues (ArgH50 and ArgH52) are responsible for the observed proton transfer.

By analysis of the two crystal structures, we also gained critical information necessary for further protein engineering of this important antibody. The analysis of interaction of epitope peptide with individual CDRs of M75 suggest that the antigen binding, especially its interaction with the light chain of the antibody, is not optimal and can be substantially improved.

In conclusion, the structural study performed as part of this PhD thesis helped understanding the details of antigen-antibody recognition and the design of improved antibody fragments of potential therapeutic and diagnostic interest.

6 SUMMARY

This work presents a part of an ongoing project focused on isozymes of human carbonic anhydrase (CA) and their roles in human physiology and pathophysiology. Carbonic anhydrase IX (CA IX), a cancer related isozyme of CA, has for long time been a center of attempts of many laboratories, both in academia and in pharmaceutical industry, to develop potent, yet selective CA inhibitors: these would provide promising pharmaceutical tools of improved therapeutic safety and reduced incidence of side effects.

One of major aims of this work was to get detailed structural information on isozymes of human CA as the essential precondition for structure-based drug design. Crystal structures of CA II in complexes with two inhibitors of a novel isoquinoline class have shown their binding modes in the isozyme active site to considerable details: such structural information permits fine-tuning of structures of the CA II inhibitors, and gives at the same time clues for the design of this class inhibitors selective toward CA IX. Inhibitory potency of such designed compounds will be corroborated with crystal structures of their complexes with recombinant CA IX isozyme expressed from a recently developed construct, which apparently, and so far uniquely, provides recombinant CA IX protein amenable to structural studies. All our previous efforts to produce a recombinant human CA IX of protein-crystallization quality (using in total 19 different constructs in both prokaryotic and eukaryotic expression systems) failed because of insufficient yields or because of inadequate crystal growths.

The other proposed main aim was to determine crystal structure of Fab fragment of monoclonal antibody M75 which recognizes the proteoglycan-like (PG) domain of CA IX. Two structures were solved, of free Fab fragment and of its complex with the epitope peptide. The fact that PG domain is a unique feature of CA IX, absent in other members of the CA family, makes this domain an attractive diagnostic and therapeutic target. The information obtained from the two crystal structures is valuable not only for its contribution to understanding of the antibody-antigen recognition in its own right, but also for the detailed knowledge on the M75 antibody variable domain structure: such insight will be useful in the planned engineering of the antibody to higher affinity, stability and humanization.

7 ABBREVIATIONS

aa	amino acid
Ab	antibody
ADP	atomic displacement parameter
AS CR	Academy of Sciences of the Czech Republic
CA	carbonic anhydrase
CAI	carbonic anhydrase inhibitor
CARP	carbonic anhydrase related protein
CCD	charge-coupled device
CDR	complementarity determining region
DMSO	dimethyl sulfoxide
dNTP	deoxynucleotide triphosphate
ECL	enhanced chemiluminescence
Fab	antibody fragment - antigen binding
Fc	antibody fragment – crystallizable
FCS	fetal calf serum
GI	gastro-intestinal
hCA	human carbonic anhydrase
HIF	hypoxia-inducible factor
HRP	horse radish peroxidase
mAb	monoclonal antibody
NC	nitrocellulose (membrane)
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data bank

PG	proteoglycan-like domain of CA IX
rmsd	root mean square deviation
scFv	single-chain variable antibody fragment
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate poly-acrylamide gel electrophoresis
TEV	tobacco etch virus
TLS	translation, libration, screw
WB	western blot

8 PUBLICATIONS

The results presented in this PhD thesis were published in two original papers and as five short communications and abstracts. Copies of the two published original papers are attached.

8.1 Original papers

Král, V.*, **Mader, P.***, Collard, R., Fábry, M., Hořejší, M., Řezáčová, P., Kožíšek, M., Závada, J., Sedláček, J., Rulíšek, L., Brynda, J. Stabilization of antibody structure upon association to a human carbonic anhydrase IX epitope studied by X-ray crystallography, microcalorimetry, and molecular dynamics simulations. (2008) ***Proteins***;71(3):1275-1287.
(IF 3.419)

*Vlastimil Král and Pavel Mader contributed equally to this work.

Gitto, R., Agnello, S., Ferro, S., De Luca, L., Vullo, D., Brynda, J., **Mader, P.**, Supuran, C.T., Chimirri, A. Identification of 3,4-Dihydroisoquinoline-2(1H)-sulfonamides as Potent Carbonic Anhydrase Inhibitors: Synthesis, Biological Evaluation, and Enzyme-Ligand X-ray Studies. (2010) ***J Med Chem***; 53 (6): 2401-2408
(IF 4.898)

8.2 Abstracts and short communications

Mader, P., Collard, R. Brynda, J. Závada J.. Structural Studies of anti-CA IX Monoclonal Antibody M75 Fab Fragment in Complex with its Epitope Peptide. (2004) *Materials Structure in Chemistry, Biology, Physics and Technology* S. 62

Brynda, J., Závada, J., **Mader**, **P.**, Fábry, M., Král, V., Kožíšek, M., Sedláček, J. Stabilization of antibody structure upon binding of an epitope rich in dicarboxylic amino acids: an X-ray and microcalorimetry study (2005) *Murnau Conference on Structural Biology of Molecular Recognition, Murnau, Germany*

Mader, **P.**, Král, V., Řezáčová, P.,. Štouračová, R., Fábry, M., Závada, J., Rulíšek, L., Kožíšek, M., Sedláček, J. Stabilization of antibody structure upon binding epitope peptide: X-ray, microcalorimetry and molecular dynamics study. (2006) *International School on Biological Crystallization, Granada, Spain*

Mader, **P.**, Král, V., Fábry, M., Kožíšek, M., Rulíšek, L., Sedláček, J., Řezáčová, P., Brynda, J. Structural Studies of anti CA IX antibody fragments (2009) *Proceedings of the 8th International Conference on Carbonic Anhydrases (Florence, Italy)*, 46

Gitto, R., Agnello, S., Ferro, S., Supuran, C.T., Brynda, J., **Mader, P.**, Chimirri, A. Identification of potent carbonic anhydrase inhibitors: synthesis, biological evaluation, and enzyme-ligand X-ray studies (2009) *Drugs of the Future* **34**, 212 Suppl. A

PROTEINS
STRUCTURE • FUNCTION • BIOINFORMATICS

Stabilization of antibody structure upon association to a human carbonic anhydrase IX epitope studied by X-ray crystallography, microcalorimetry, and molecular dynamics simulations

Vlastimil Král,[1,2] Pavel Mader,[1,3] Renata Collard,[1] Milan Fábry,[1] Magdaléna Hořejší,[1] Pavlína Řezáčová,[1] Milan Kožíšek,[2,4] Jan Závada,[4] Juraj Sedláček,[1] Lubomír Rulíšek,[4]* and Jiří Brynda[1,4]*

[1] Department of Recombinant Expression and Structural Biology, Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, 166 37 Prague, Czech Republic

[2] Department of Biochemistry, Charles University, Faculty of Natural Science, 128 43 Prague, Czech Republic

[3] 1st Faculty of Medicine, Charles University, 121 08 Prague, Czech Republic

[4] Gilead and IOCB Research Center, Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, 166 10 Prague, Czech Republic

ABSTRACT

Specific antibodies interfere with the function of human tumor-associated carbonic anhydrase IX (CA IX), and show potential as tools for anticancer interventions. In this work, a correlation between structural elements and thermodynamic parameters of the association of antibody fragment Fab M75 to a peptide corresponding to its epitope in the proteoglycan-like domain of CA IX, is presented. Comparisons of the crystal structures of free Fab M75 and its complex with the epitope peptide reveal major readjustments of CDR-H1 and CDR-H3. In contrast, the overall conformations and positions of CDR-H2 and CDR-L2 remain unaltered, and their positively charged residues may thus present a fixed frame for epitope recognition. Adoption of the altered CDR-H3 conformation in the structure of the complex is accompanied by an apparent local stabilization. Analysis of domain mobility with translation-libration-screw (TLS) method shows that librations of the entire heavy chain variable domain (V_H) decrease and reorient in the complex, which correlates well with participation of the heavy chain in ligand binding. Isothermal titration microcalorimetry (ITC) experiments revealed a highly unfavorable entropy term, which can be attributed mainly to the decrease in the degrees of freedom of the system, the loss of conformational freedom of peptide and partially to a local stabilization of CDR-H3. Moreover, it was observed that one proton is transferred from the environment to the protein-ligand complex upon binding. Molecular dynamics simulations followed by molecular mechanics/generalized Born surface area (MM-GBSA) calculations of the ligand (epitope peptide) binding energy yielded energy values that were in agreement with the ITC measurements and indicated that the charged residues play crucial role in the epitope binding. Theoretical arguments presented in this work indicate that two adjacent arginine residues (ArgH50 and ArgH52) are responsible for the observed proton transfer.

Proteins 2008; 71:1275–1287.

***Key words:** antigen-antibody recognition; Fab crystal structure; CDR loop rearrangement; TLS method; microcalorimetry, MM-GBSA calculations.*

The Supplementary Material referred to in this article can be found online at http://www.interscience.wiley.com/jpages/0887-3585/suppmat.

Abbreviations: ADP, atomic displacement parameter (formerly B-factor); CDR-H1, CDR-H2, CDR-H3, first, second, and third complementarity determining regions (CDR) of the heavy chain; CDR-L1, CDR-L2, CDR-L3, first, second, and third CDR of the light chain; C_H1, first constant domain of heavy chain; Fab, fragment antigen binding; ITC, isothermal titration microcalorimetry; mAb, monoclonal antibody; MM-GBSA, molecular mechanics–generalized Born surface area; PDB, protein data bank; scFv, single chain variable fragment; TLS, translation-libration-screw; V_H and V_L, heavy and light chain variable domain, respectively.

Grant sponsor: Academy of Sciences of the Czech Republic; Grant numbers: AV0Z50520514, Z0550506; Grant sponsor: Czech Ministry of Industry and Commerce; Grant number: 2A-2TP1/076; Grant sponsor: Czech Ministry of Education; Grant numbers: 1M0508, 1M0505, LC512, E13537; Grant sponsor: Grant Agency of the Czech Republic; Grant number: 203/02/0405.

Vlastimil Král and Pavel Mader contributed equally to this work.

Renata Collard's current address is Department of Pediatrics, University of Colorado at Denver and Health Sciences Center, Denver, CO 80217-3364.

*Correspondence to: Lubomír Rulíšek, Gilead and IOCB Research Center, Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, 166 10 Prague, Czech Republic. E-mail: lubos@uochb.cas.cz or Jiří Brynda, Department of Recombinant Expression and Structural Biology, Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, 166 37 Prague, Czech Republic. E-mail: brynda@img.cas.cz

Received 22 May 2007; Revised 14 August 2007; Accepted 11 September 2007
Published online 27 November 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21821

INTRODUCTION

Abnormal expression of human carbonic anhydrase IX (CA IX) in various commonly occurring carcinomas suggests that this molecule is involved in oncogenic pathways.[1–3] The CA IX is a promising diagnostic and therapeutic target for a variety of cancers due to its almost exclusive ectopic expression in tumors, very limited expression in normal tissues, and the availability of specific monoclonal antibodies including mAb M75,[4] mAb G250,[5] and several others.[6] In the search for selective anticancer agents that differentiate between malignant and nonmalignant cells, several strategies for targeting CA IX were designed and tested.[7] Detailed description of the interactions involved in anti-CA IX antibody recognition can help in development of further advanced strategies, such as generation of recombinant forms of the antibodies.

Of the 15 human carbonic anhydrases characterized to date,[8,9] only the CA IX isoform is strongly associated with certain types of cancer.[1,2] A unique structural feature of CA IX is its proteoglycan (PG)-like segment, which is located at the amino terminus of the molecule and comprise epitope recognized by mAb M75. Conceivably, this segment could be important in oncogenesis and therefore, the nature of its interactions with M75 antibody deserves more detailed investigation. CA IX also acts as a cell adhesion molecule (CAM) that can mediate the attachment of cells to a nonadhesive solid support.[4] The adhesion site of CA IX appears to overlap with the PGEEDLP epitope of the monoclonal antibody M75 since the antibody blocks attachment of cells to immobilized CA IX protein.[4] The counterpart receptor component of the adhesion cell remains unknown. Nevertheless, it is assumed that similar structural elements are involved in CA IX binding with M75 antibody and with the CAM receptor since formation of both complexes is inhibited in an acidic environment in the same manner.[7]

The PG amino acid sequence[10] comprises 58 residues and includes the epitope recognized by the anti-CA IX monoclonal antibody M75 (Ref. 4):

SSGEDDPL**GEEDLP**SEEDSPREEDPP
GEEDLPGEEDLPGEEDLPEVKPKSEEEGSLKE

A remarkable feature of the PG segment of CA IX is high content of carboxylic amino acids (8 Asp and 18 Glu in a total of 58 residues) and the low content of basic residues (1 Arg and 3 Lys residues). Most of the carboxylic amino acids are grouped in four identical GEEDLP repeating motifs (shown in bold) and in three other motifs (denoted with lines on the top).

Currently, there are several antibody structures in both the free and the antigen-bound form that have been determined by X-ray crystallography,[11–15] which provide invaluable information about mechanism of antigen recognition and the structural changes that occur during antibody-antigen association.

In the present work, the structural elements and thermodynamics of association of the Fab fragment of the anti-CA IX antibody M75 to an epitope peptide derived from its protein antigen, the proteoglycan-like (PG) segment of CA IX are reported. The epitope peptide used in the crystallographic part of this study (M peptide, PGEEDLPGEEDL),[4] underlined in the sequence above) seems to be an adequate representation of the protein antigen since structural predictions (e.g., DisEMBL program)[16] suggested a lack of secondary structure in the PG domain. Flexible epitope peptide that possesses no secondary structure is expected to adopt a complementary shape and undergo stabilization in complex with its cognate antibody. This is accompanied by a loss of conformational freedom, resulting in an unfavorable entropy contribution, which is measurable with microcalorimetry methods.[17] The combination of three methods presented in the study (crystallography, microcalorimetry, and molecular dynamics simulations) allowed us to discuss the process of antibody-epitope recognition in terms of simple mechanisms, such as "conformational selection" or "induced fit." An understanding of the structural and thermodynamic parameters associated with M75 antibody binding provides certain guidelines for rational design of antibody-based agents interfering with the presumed function of the PG domain of CAIX in oncogenesis.

RESULTS

Quality of models and overall structures

Refinement statistics for the Fab M75 and the [Fab M75 + M peptide] complex crystal structures (PDB entry **2HKH** and **2HKF,** respectively) are given in Table I. The final Fab M75 model, refined at 2.1 Å resolution comprised of a single protein molecule in the asymmetric unit (219 residues for the light chain and 208 residues for the heavy chain), 98 water molecules and one molecule of glycerol. Residues belonging to the hypervariable loops were well-defined in the electron density map, with the exception of five CDR-H3 residues (H100–H104; Antibody residues are numbered consecutively, preceded by L or H for light and heavy chain, respectively, whereas peptide residue numbers are preceded by P). Other missing residues in our model included N-terminal GluH1, and residues belonging to the constant region C_H1 (H134–H137 and H162–H166). Residues disallowed in the Ramachandran energy space are ValL56 (which is characteristic of the canonical conformation of complementarity defining region CDR-L2),[18] SerH27, and ThrH28 from CDR-H1 solvent exposed loop; generously allowed are SerH155 and SerH178 surface residues of C_H1 loops.

Table I
Data Collection and Refinement Statistics

Crystals	Free Fab M75	Complex Fab M75 + peptide
Space group	P1	P1
Unit cell parameters	a = 40.98 Å b = 43.15 Å c = 57.37 Å α = 85.30° β = 88.45° γ = 84.29°	a = 40.388 Å b = 43.145 Å c = 58.023 Å α = 85.03° β = 88.47° γ = 85.67°
Resolution[a]	30.00–2.10 (2.18–2.10) Å	28.87–2.00 (2.11–2.00) Å
Total number of reflections	57110	57489
Number of unique reflections	22667	24418
Completeness[a]	90.7 (68.6)%	93.5 (93.5)%
R_{merge}[a]	6.9 (30.4)%	4.9 (19.1)%
$I/\sigma(I)$[a]	6.5 (2.0)	12.0 (3.2)
Wilson plot B factor	39.4 Å^2	19.4 Å^2
Models		
PDB ID	**2HKH**	**2HKF**
NonH Atoms + Water molecules	3288 + 98	3360 + 303
R_{work}	0.219	0.175
R_{free}	0.272	0.242

[a]Values in parentheses correspond to the last resolution shell.

The final model of [Fab M75 + M peptide], refined at 2.0 Å resolution, comprises one complex molecule in the asymmetric unit: 219 residues for the light chain, 209 residues for the heavy chain, 9 peptide residues, and 303 water molecules. Similarly to the free Fab M75 structure, ValL56 falls into the disallowed region of the Ramachandran energy space, and SerH155 into the generously allowed region. All residues belonging to the hypervariable loops were well-defined in the electron density map. The residues missing from the protein model belong to the C_H1 region. Eight out of 12 peptide residues (PGEEDLPG) could be conclusively traced into well defined electron density map in the antigen binding site [Fig. 1(A)]. During the peptide building, an extra electron density was observed preceding the amino end of the traced peptide that could be indisputably traced as a main chain of additional amino acid residue and identified as a leucine. We believe that the final crystallographic model represents an average of alternate M peptide binding by its amino-terminal segment (**PGEEDLPG**EEDL) and by the carboxy-terminal segment (PGEED**LPGEEDL**), respectively (relevant parts of M peptide are shown in bold). In the deposited complex structure (PDB entry **2HKF**), the bound "averaged" peptide has nine residues and its first residue is approximated with glycine. The values of atomic displacement parameters (ADP) for the peptide bound to Fab (Fig. 2) demonstrated higher stability of the central peptide part (ProP2 to LeuP7) compared to the peptide terminal residues.

From comparison of the two crystal structures (free and complexed), it can be noted that the structure of [Fab M75 + M peptide] complex has smaller number of disordered residues, slightly better resolution (2.0 Å vs. 2.1 Å), and a lower average ADP (average B value). The decrease in average ADP value from 31.1 Å^2 for free Fab to 17.8 Å^2 for the complex indicates a global stabilization of the complexed structure. Comparison of local (per residue) ADP values along the protein chains indicates one short segment with pronounced differences between the free and the complexed structure. This is discussed below, in connection with the rearrangement of the CDR-H1 and CDR-H3 loops upon epitope binding. Moreover, it is evident that certain tightly bound water molecules are displaced upon peptide binding, such as the water molecule originally positioned between the CDR-H2 and CDR-L1 loops [Fig. 1(D)].

Peptide-antibody polar interactions

In general terms, antigen-antibody interactions reflect shape and charge complementarities. Charge (electrostatic) complementarity is assumed to be a prominent factor in specific antigen recognition.[14] In the structure of the complex described here, the antigen binding cleft displays a continuous positively charged surface [Fig. 1(B)] and the charge complementarities between antibody and peptide include the side chains of ArgH50 and ArgH52 of Fab M75 as counterparts of the GluP5 residue of the epitope (PGE**E**DLP) and the side chain of LysL55 of Fab M75 as the counterpart to the AspP6 residue of the epitope (PGEE**D**LP). The two successive carboxylic amino acid residues of the epitope (GluP5, AspP6) form salt bridges with the positively charged residues of CDR-H2 and CDR-L2 [Fig. 1(C)]. The side chain of the remaining carboxylic amino acid residue of the epitope (GluP4, PG**E**EDLP) is oriented toward the solvent. Comparison of mutual positions of positively charged antibody residues revealed that distances [ArgH50:NH2...LysL55:NZ] and [ArgH52:NH2...LysL55:NZ] are very similar in free Fab and the complexed structure (differing by 0.81 and 0.07 Å, respectively). This indicates that the epitope-binding amino acid residues of Fab CDR-H2 and CDR-L2 are not involved in an antibody structural rearrangement associated with the antigen binding. From the point of view of charge complementarities, the antibody

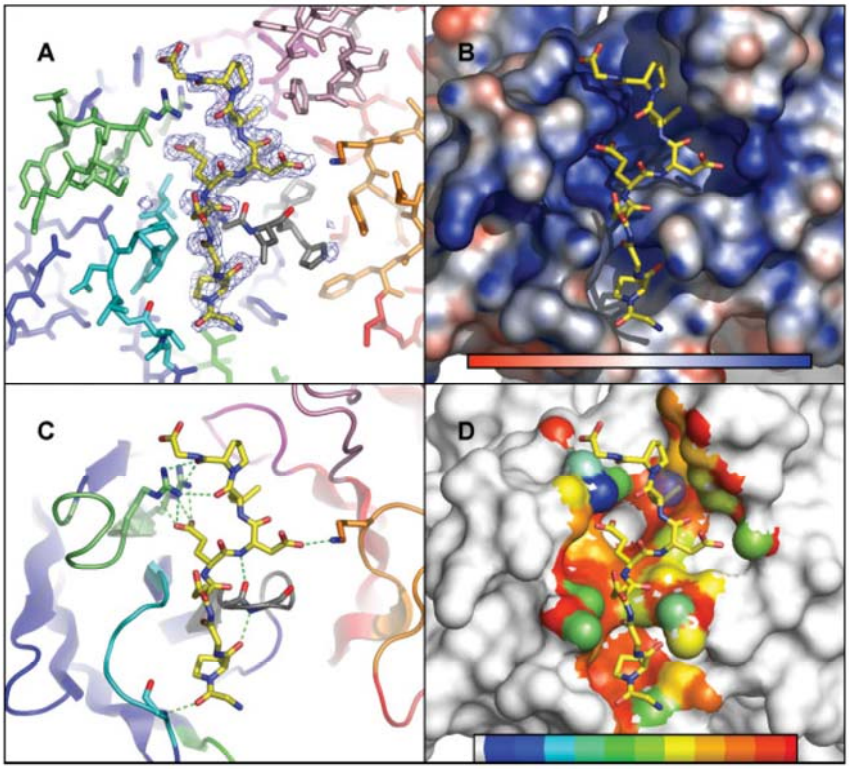


Figure 1

*Structure of the antibody combining site. Panel (**A**): Omit map of the bound peptide (yellow) contoured at 3σ level. The respective CDR loops are colored pale blue for H1, green for H2, gray for H3, light magenta for L1, mauve for L2, and dark magenta for L3. Panel (**B**): Protein surface at van der Waals radii colored by the electrostatic potential from negative (red) to positive (blue); the scale bar is from −7 to +7 kT/e. Panel (**C**): Hydrogen bonds formed between the antibody and the epitope peptide. The respective CDR loops are colored as in panel (A), the amino terminal segment of the heavy chain is colored green. Panel (**D**): Areas of protein surface at van der Waals radii buried upon peptide binding. The rainbow coloring scheme ranges from white (0 Å^2) through blue (5 Å^2) to red (50 Å^2); transparent blue sphere indicates a water molecule displaced upon the peptide binding. All panels were drawn using program PyMOL[13] (http://www.pymol.org).*

thus forms an adequate binding template without any prerequisite of induced fit. Contacts of the Fab ArgH50, ArgH52, and LysL55 residues and the carboxylic amino acid residues of the epitope (M peptide) can be also analyzed with respect to hydrogen bonding parameters. Figure 1(C) shows that there are five contacts of the Fab ArgH50, ArgH52, and LysL55 residues with the carboxylic amino acid residues of the epitope, matching the criteria of hydrogen bonding distances and angles.

Other hydrogen bonds involve ThrH28 of CDR-H1 and the amino terminal residue of the peptide (denoted as GlyP1, see above), three hydrogen bonds between ArgH52 of CDR-H2 and ProP8/GlyP9, and four hydrogen bonds between residues GlyH102 and AsnH103 of CDR-H3 and ProP2, GluP4, AspP6 (three of them are of "main chain-to-main chain" type). In total, heavy chain residues form 10 (out of 11) hydrogen bonds to the epitope peptide. All polar contacts are listed in Table II.

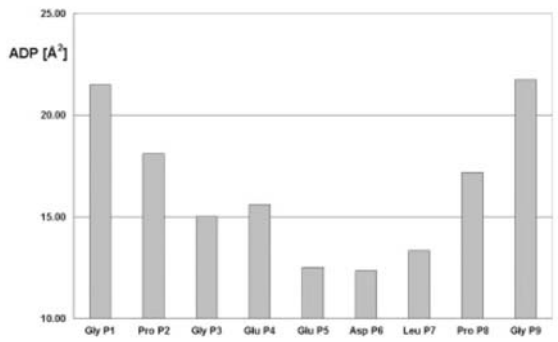


Figure 2

Average values of residual atomic displacement parameters of the peptide residues in complex.

Table II
The Polar Contacts Between the Peptide Residues and Fab M75 as Observed in the Crystal Structure of the Complexed Form as Calculated by ACT Program of CCP4 Suite

Peptide atom	Fab M75 atom (CDR loop)	Distance (Å)
GlyP1 O	ThrH28 N (H1)	3.03
ProP2 O	AsnH103 N (H3)	2.85
GluP4 O	AsnH103 N (H3)	3.07
GluP4 Oε1	ArgH52 Nε (H2)	2.79
GluP4 Oε1	ArgH52 Nη2 (H2)	311
GluP4 Oε2	ArgH50 Nε (H2)	3.09
GluP4 Oε2	ArgH50 Nη2 (H2)	2.89
AspP6 N	GlyH102 O (H3)	2.91
AspP6 Oδ1	LysL55 Nη (L2)	2.65
ProP8 O	ArgH52 Nη1 (H2)	3.14
ProP8 O	ArgH52 Nη1 (H2)	3.00

Van der Waals peptide-antibody contacts

The complexed structure displays areas of extensive van der Waals contacts between Fab and the epitope peptide. The heavy chain residues are responsible for majority of these contacts (~80%). As shown in Figure 1(D), the protein surface areas buried upon binding of the epitope peptide are approximately four times larger for the heavy chain than for the light chain. The contributions of individual heavy chain CDRs are not substantially different and decrease in the order CDR-H2 > CDR-H3 > CDR-H1. A van der Waals contact is also found between the amino terminal residue of the constant domain of the heavy chain, ValH2, and ProP2 of the epitope peptide.

Conformational changes of the antibody main chains

Comparison of free and complexed structures revealed that major transitions of the antibody main chains upon epitope peptide binding occur in the regions of CDR-H1, CDR-H3, and CDR-L1. While the hypervariable loops of CDR-H1 and CDR-H3 acquire altered conformations, the entire CDR-L1 is shifted to a new position as almost rigid structure without changing the overall loop conformation.

Superposition of the free and complexed structures is depicted in Figure 3 (using Cα-traces), and differences in the positions of Cα atoms of corresponding residues in the two models are plotted in the graph of Figure 4. Upon the binding of the epitope peptide, the conformation of the CDR-H1 loop is altered between residues

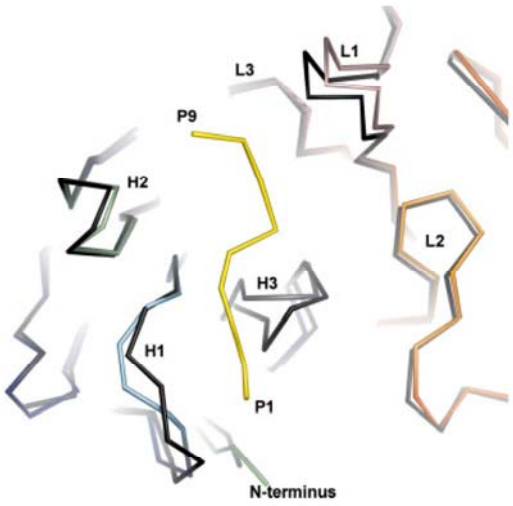


Figure 3
Superimposed Cα traces of the free and complexed Fab M75. A "top view" into the binding cleft. The peptide and CDRs of the complex are color coded as in Figure 1; the "free" structure is colored black. Figure was drawn using program PyMOL[13] (http://www.pymol.org).

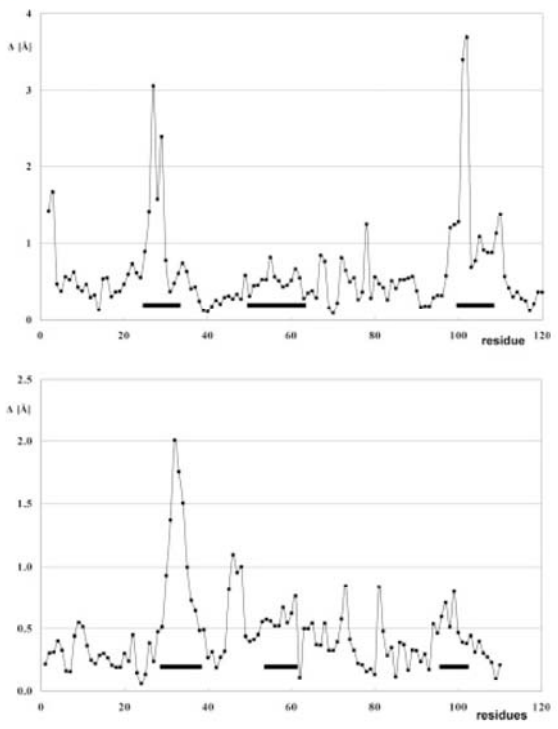


Figure 4
Differences in positions of Cα atoms of corresponding residues in free and complexed FabM75 plotted for the heavy chain (upper graph), and for the light chain (lower graph). Horizontal bars indicate CDRs H1, H2, and H3 and L1, L2, and L3, respectively.

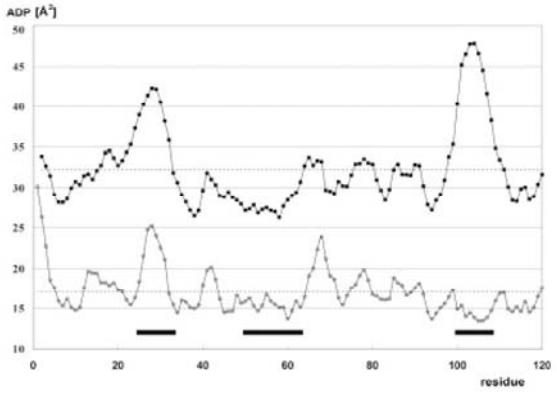


Figure 5

Residual atomic displacement parameters of Cα atoms of the heavy chain residues. Free Fab, full squares; complex, open circles. Horizontal bars indicate CDRs H1, H2, and H3. Two dashed lines show the average of atomic displacement parameters for the heavy chain of free and complexed Fab.

GlyH26 and LeuH29. The loop is moving in the direction of the bound epitope peptide with the maximum shift observed for the SerH27 Cα atom (3.06 Å). The conformational change in the CDR-H3 loop involves rearrangement of TyrH101-AsnH103 residues. The H3 loop moves toward the C-terminal part of CDR-L2, and provides the contacts to the epitope peptide from the L2 direction (cf., Fig. 3). The most dramatic movement is observed for GlyH102 Cα atom (3.68 Å). The new conformation of the CDR-H3 loop opens the groove in the antigen binding cleft for the binding of the N-terminal part of the epitope. The displacement of the hypervariable loop of CDR-L1 is smaller, with an amplitude of 1.98 Å observed for the residue SerL32 Cα atom. The loop moves away from the heavy chain and its rearrangement is accompanied by the displacement of framework loop F2-L (in the opposite direction and with similar amplitude). If the displacement is viewed as a rotation of a rigid structure, the hinge can be identified at the residue TrpL40.

The adoption of the altered CDR loop conformations in the complex is accompanied by local stabilization, that is decreased mobility, as measured with residual ADPs. The plot of ADP values for Cα atoms along the heavy chain (Fig. 5) shows a distinctive peak for the CDR-H3 segment of the free antibody. This peak is absent in the corresponding plot of the complex structure ADPs, whereas the two plots are essentially identical for other heavy chain segments. In summary, two transitions in the heavy chain of two CDRs are observed, both generating shape complementarity (cf., Figs. 3 and 4), whereas only the readjustment of CDR-H3 contributes to local stability.

Domain mobility

Investigation of the global stability of protein domains can be performed using TLS analysis.[19] In this method displacements of protein domains are separated into contributions of translation, libration, and screw-rotation displacements. For larger protein domains, libration is usually predominant and this was also observed for our structures. In Figure 6, the principal libration axes and magnitudes for free and complexed Fab M75 structures are depicted. The orientation of the libration axes and the values of libration magnitudes are similar for the free and complexed structures in the Fab constant domain (C_H1 + C_L). While for the variable light chain domain (V_L), the orientations of the libration axes remain essentially unaltered but the magnitudes are decreased in the complex, the variable heavy chain domain (V_H) in the complex librates along reoriented axes and with reduced magnitudes. These changes indicate an altered mode of librations for V_H and global stabilization for both the V_H and V_L domains, as a result of restrictions introduced by

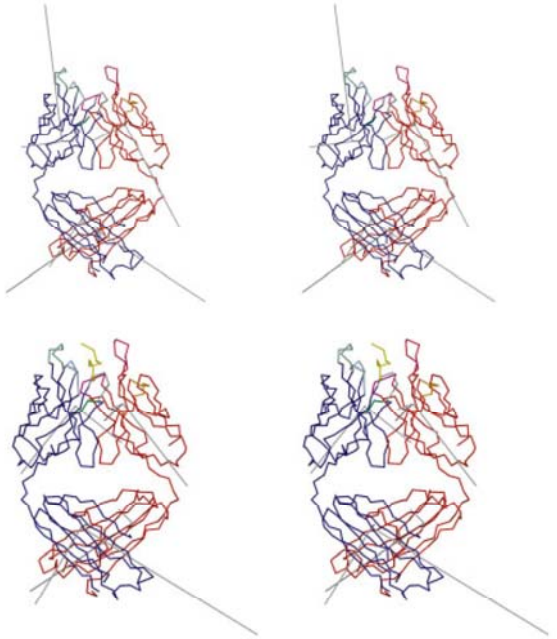


Figure 6

Principal axes of the libration tensors for Fab M75 domains (V_H, V_L, C_H, and C_L) superimposed on stereo views of crystal structures. The Cα traces of the free Fab (upper block) and of the complex (lower block) are shown. The heavy chain is colored mainly blue, the light chain mainly red, and the CDR loops and the epitope peptide using the same color scheme as in Figure 1. The libration axes represented by black lines are the output of the AXES option in TLSANL program.[14] The length of each axis is proportional to the mean square libration about that axis. Figure was drawn using programs MOLSCRIPT[15] and Raster3D.[16]

the peptide binding. The large effect of peptide binding on librations of the V_H domain qualitatively agrees with the extensive involvement of the heavy chain CDR loops in the antibody-peptide interactions, as described above.

Thermodynamic parameters

The thermodynamic parameters of the epitope binding were determined using isothermal titration calorimetry (ITC) for a minimized model system consisting of scFv M75 and the peptide DLPGEEDLPGEE. This peptide represents the interacting part of the peptide in the complexed structure extended by one amino acid residue on the N-terminus and two amino acid residues on the C-terminus (DLPGEEDLPGEE; underlined is a part of peptide which was conclusively traced in the crystallographic model). This particular sequence of peptide was selected to avoid the ambiguity of binding mode which was observed in crystal structure. The results of calorimetric titration performed in Tricine buffer are shown in Figure 7. The enthalpy is plotted as a function of the molar ratio of peptide to scFv M75 and the fit for a single binding site model yielded a binding stoichiometry of 0.95(4). The affinity strength calculated from the plot corresponds to association constant $K_a = 2.4(7) \times 10^6 M^{-1}$ (or the dissociation constant $K_d = 0.42\ \mu M$). The buffer-independent binding enthalpy was determined by performing identical titration experiments in three different buffers with known ionization enthalpies (results in Supplementary Material). The result reveals that one proton is absorbed during the binding of the DLPGEEDLPGEE peptide to scFv M75 [$\Delta n_H = 1.1(2)$, cf. Eq. (2)]. The total free (Gibbs) energy change associated with ligand binding was determined to be $\Delta G = -36.4(7)$ kJ mol^{-1}. It can be decomposed into favorable enthalpic ($\Delta H_b = -133(7)$ kJ mol^{-1}) and unfavorable entropic contributions ($-T\Delta S = 97(7)$ kJ mol^{-1}). By performing experiments at 15 and 25°C (288.15 and 298.15 K), it was possible to determine the heat capacity change of $\Delta c_p = -795$ J mol^{-1} K^{-1}, which can be related to changes in solvation/desolvation associated with ligand binding. This observation is typical for peptide-protein interactions and indicates the burial of the peptide into scFv M75 antigen binding site.[20]

Molecular dynamics simulations

A series of molecular dynamics (MD) simulations was carried out with the following aims: (i) to reproduce the experimental values obtained by ITC and estimate entropy versus enthalpy contributions; (ii) to decompose the interaction energy of ligand with protein into individual contributions originating from van der Waals, electrostatic, and solvation terms on a per-residue basis to evaluate the importance of individual amino acid residues in ligand binding. Moreover, the outcome of the molecular mechanics/dynamics calculations is the structure prediction of the scFv M75 protein in free form and in complex with peptide DLPGEEDLPGEE (used for ITC). The two model structures, scFv M75 and [scFv M75 + DLPGEEDLPGEE peptide] complex can be found in the Supplementary Material. We also expected that the MD simulations might also elucidate the origin of the proton absorption during ITC measurements and thus corroborate the experimental data. It should be noted that accurate calculation of the ligand-binding energy of an extended ligand (such as the studied epitope peptide) with protein is a difficult task. In this work, it is further complicated by the limited structural information concerning the $(Gly_4Ser)_4$ linker in scFv M75 structure. Therefore, the decomposition of the total free energy change associated with ligand binding is of semiquantitative character, allowing the prediction of trends rather than accurate values. Also, the current methods, such as

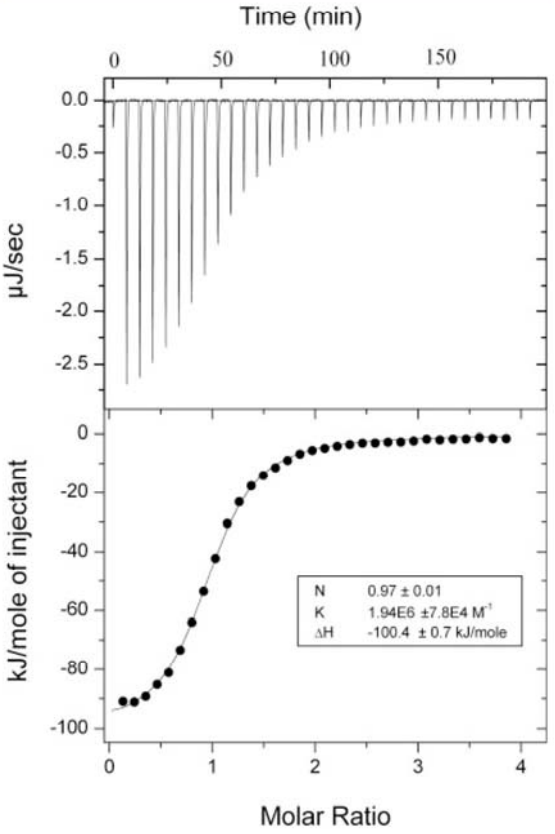


Figure 7

Isothermal titration of scFv M75 with peptide DLPGEEDLPGEE in 20 mM Tricine, pH 8.4, 100 mM NaCl at 25°C Upper graph: experimental data. Lower graph: Fit (line) to the integrated heats (full circles) from each injection, corrected for the heat of dilution of the peptide.

Table III
The Interaction Free (Gibbs) Energies Computed Using MM-GBSA Method on the Collection of 100 Snapshots From MD Trajectory

	ΔE_b[a]	ΔH_b[b]	$T\Delta S_{conf}$[c]	ΔG_b
[$R_{H50}^{(+)}$, $R_{H52}^{(+)}$]	−3041.8	−202.3	211.9	9.6
[$R_{H50}^{(0)}$, $R_{H52}^{(+)}$]	−2439.9	−151.7	142.5	−9.2

All values are in kJ mol^{-1}.
[a]Calculated in vacuo binding energy.
[b]Estimated binding enthalpy (containing solvation entropy change).
[c]Conformational entropy change (i.e., total entropy change without solvation entropy component).

MM-GB(PB)SA or polarized continuum models do not allow for the decomposition of total binding free energy into the total binding enthalpy and total binding entropy terms, thus preventing a direct comparison with ITC.

Generalized Born/surface area (MM-GBSA) calculations were performed on a collection of snapshots acquired during the MD simulations of the complex, as described in detail under Materials and Methods. We postulate that the net proton transfer observed during the microcalorimetric experiments might be attributed to the protonation of either ArgH50 or ArgH52. In the crystal structure of the free protein at pH = 8.4, the two charged residues are in close proximity and their positive charge is not compensated by any negatively charged residues in vicinity. Thus it might be expected, that one of the two arginine residues has to be in a neutral state to avoid the presence of two charged residues in the vicinity. However, upon complex formation, strong salt bridge is formed between two arginine residues and side chain of GluP5 of the ligand. The uptake of the proton by the neutral Arg residue would significantly stabilize the complex formation. Therefore, we have performed MD simulations and subsequent MM-GBSA analysis on both [ArgH50$^{(+)}$, ArgH52$^{(+)}$] and [ArgH50$^{(0)}$, ArgH52$^{(+)}$] forms of the scFv M75 protein. The results are summarized in Table III.

In agreement with ITC experiments, we have obtained high values of conformational entropy changes ($-T\Delta S$ = 142–212 kJ mol^{-1}) and high values of estimated binding enthalpies (−202 and −152 kJ mol^{-1}). The calculated value of ΔG_b = 9.6 kJ mol^{-1} must be corrected for the observed protonation event. The following process has been used for the calculation of $\Delta G_{b(H+)}$ which can be directly compared to the experimental value (i.e., $\Delta G_{b(H+)} \approx \Delta G_{b,exp}$):

$$[\mathrm{ArgH50}^{(0)}, \mathrm{ArgH52}^{(+)}] + \mathrm{L} \overset{\Delta G_1}{\rightarrow} [\mathrm{ArgH50}^{(+)}, \mathrm{ArgH52}^{(+)}] + \mathrm{L} \overset{\Delta G_b}{\rightarrow} [\mathrm{ArgH50}^{(+)}, \mathrm{ArgH52}^{(+)} \ldots \mathrm{L}], \quad (1)$$

where L is ligand–epitope peptide. It can be seen that $\Delta G_{b(H+)} = \Delta G_1 + \Delta G_b$. The latter has been calculated as 9.6 kJ mol^{-1}, whereas the value of ΔG_1 (by definition, $\Delta G_1 = RT \ln K_a$ for protonation) can be estimated from the pK_a value of the arginine residue that is assumed to be responsible for the observed protonation event, that is, $\Delta G_{b(H+)} = -2.303RT\,\mathrm{p}K_a(\mathrm{est}) + \Delta G_b$. Assuming p$K_a$(ArgH50,est) ≈ 5–8 (note that the experiments were carried out at pH = 8.4), we obtain $\Delta G_{b(H+),calc} \approx -19$ to −36 kJ mol^{-1}. This value is in very good agreement with the experimental value (considering errors of at least 30–40 kJ mol^{-1} inherent in the MM-GBSA procedure.[21]

A reasonable agreement between the calculated and experimental values of ligand binding energies allows the per-residue decomposition of the total ligand binding enthalpy. This decomposition can be only done from the trajectory of the [scFv M75 + DLPGEEDLPGEE peptide] complex. The total interaction enthalpy $\Delta H'_{b(cplx)}$ calculated for the complex was −252 kJ mol^{-1}. This value is larger (more negative) than the value calculated for the free scFv M75 reported in Table III and the difference $\Delta H'_{relax} = \Delta H'_b - \Delta H'_{b(cplx)}$ = 50 kJ mol^{-1}can be attributed to the relaxation (strain) energy of protein and ligand. The plot of the interaction enthalpies for the peptide residues is given in Figure 8. The van der Waals con-

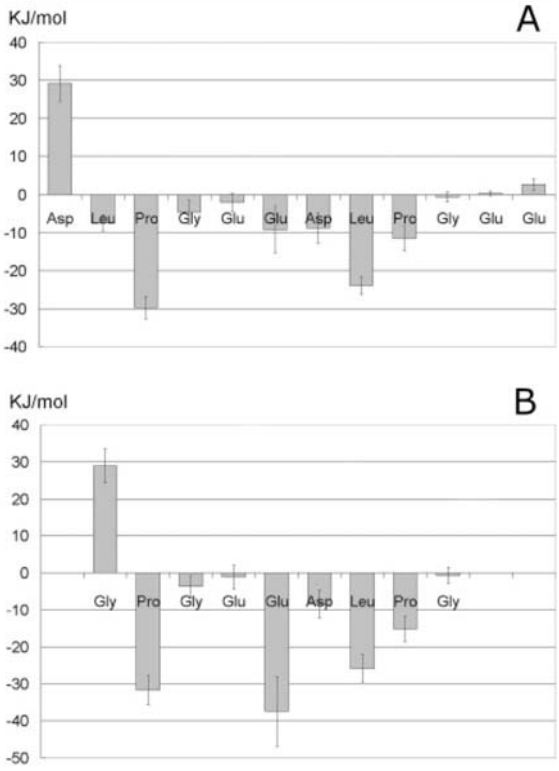


Figure 8
*Interaction energies decomposed for the epitope peptide residues. (**A**) peptide DLPGEEDLPGEE, (**B**) peptide GPGEEDLPG ("averaged peptide" determined by crystallography). The error bars show standard deviations of MD snapshots.*

tacts between the protein and the peptide residues ProP2 and LeuP7 contributed significantly to the overall interaction (−20 to −30 kJ mol^{-1}). In contrast, the negatively charged carboxylate residues (due to the very high solvation/desolvation enthalpies) contributed only by ∼10 kJ mol^{-1}. This analysis highlighted the relative importance of hydrophobic residues for binding once the peptide is recognized by long-range electrostatic interactions mediated by GluP5, AspP6. In qualitative agreement with the experimental data, all the residues that are not tightly bound to the protein (AspP-1, GlyP9, GluP10, GluP11) interact negligibly or even destabilize the peptide binding. On the protein side, the highest interaction enthalpies were calculated for the arginine pair (ArgH50, ArgH52), the counterparts of GluP6, with $\Delta H'_{b(ArgH50)}(GB) = \Delta H'_{b(ArgH52)}(GB) \approx -30$ kJ mol^{-1}. This supports the view that the formation of these salt bridges is essential for antibody recognition.

For comparison, the trajectories of the Fab M75 protein represented by the V_L and V_H domains in the free form and in complex with the peptide GPGEEDLPG were analyzed. The aim was to provide more supporting evidence for the above decomposition and compare the binding characteristics of the two different peptides used for X-ray crystallography and ITC experiments, respectively. The results, depicted in Figure 8(B), showed that the GPGEEDLPG ("averaged peptide" binding determined by crystallography) interaction is, indeed, very similar to the one of the DLPGEEDLPGEE peptide used for ITC. The only difference is the interaction of the GluP5 residue which increased by 28 kJ mol^{-1}; its binding is comparable to the sum of three hydrophobic residues (ProP2, LeuP7, and ProP8). In the protein, it has been calculated that $\Delta H'_{b(ArgH50)}(GB) = -21$ kJ mol^{-1} and $\Delta H'_{b(ArgH52)}(GB) \approx -45$ kJ mol^{-1} (i.e., the sum of the interaction of the arginine pair is almost conserved in both our model systems). These findings further demonstrate the crucial role of both salt bridge formation and short-range nonpolar interactions.

DISCUSSION

The combination of X-ray crystallography, ITC experiments, and MD simulations allowed us to assess the key structural parameters responsible for Fab M75 antibody epitope recognition and binding. The thermodynamic parameters obtained with ITC are consistent with the structural elements of the antibody-antigen association obtained from the experimentally determined crystal structures and are further supported by the analysis of MD trajectories.

The crystal structure of the [Fab M75 + peptide M] complex revealed, that charge interactions involve CDR-H2 and CDR-L2 of the antibody, whereas the hydrogen bonding involves, except for one H-bond, only the heavy chain CDRs. The same holds true for the van der Waals contacts, which predominantly involve the CDRs of the heavy chain. This interaction pattern provides us with a qualitative explanation for the relatively high value of the $-T\Delta S$ entropy term obtained from ITC measurements. The local stabilization (loss of flexibility) of the CDR-H3 loop represents unfavorable entropy contribution that adds on to the standard decrease in entropy associated with the ligand-protein binding in general. These contributions do not seem to be compensated for by a favorable entropy gain from displacement of bound water molecules.

While the ITC measurements permitted a decomposition of the binding free energy into enthalpic and entropic terms, MD simulations enabled a further separation of the total interaction energy into two contributions: (i) electrostatic, including hydrogen bonding, and (ii) the van der Waals contacts (hydrophobic interactions). The results of this study indicate that both types of interactions are important for antigen binding. The long-range electrostatic interactions, which are predicted to be strongest between ArgH50, ArgH52, and GluP5 residues, are responsible for antigen recognition. This could explain the sensitivity of the complex formation to an acidic environment. Once the complex is formed, short range van der Waals interactions between the antibody and ProP2/LeuP7/(ProP8) residues of the peptide represent an important stabilizing factor. The protonation of the carboxylic groups of the epitope (i.e., acidic environment) results in the loss of key interactions between the positively charged binding cleft of the antibody and the epitope; the same holds true for the antibody light chain due to the LysL55 and AspP6 interaction.[7] Though the results of MD simulations must be considered as semi-quantitative or qualitative due to the errors of 30–40 kJ mol^{-1} presented in the calculated values, they can give an indication for the possible mutations in the antibody that would change its binding.

The repetition of the GEEDLP sequence in the M peptide results in the ambiguous binding mode observed in the crystal structure. The implications of having a repetitive sequence to enhance binding affinity (more precisely, avidity) were previously reported by Závada in more detail,[4] whereas general mechanistic implications of multiple binding modes are discussed by Gestwicki *et al.*[22] and Hlavacek *et al.*[23]

Two mechanisms can be considered relevant for the antibody-antigen interaction: (i) "induced fit," that is, the complementary shape of the antibody binding site is formed upon ligand binding; (ii) "conformational selection," that is, the conformation favorable for antigen binding is already present in a fraction of the population of free antibody molecules.

The argument in favor of the induced fit mechanism in case of Fab M75 is the long-range nature of electrostatic interactions, and the presence of positive charges in

the ligand-binding cleft of the antibody. The almost identical distances between the Arg residues of CDR-H2 and the Lys residue of CDR-L2 in the free and complexed forms of Fab M75 suggest that this structural element make a preformed binding site, important for antigen recognition by the long-range electrostatic interactions. Subsequent rearrangements of CDR loops H1, H3, and L1 leads to formation of the final binding site that accommodates the antigen.

The concept of conformational selection[24] provides an alternate view. In this mechanism, the antigen binds selectivcly to the fraction of antibody molecules that temporarily adopt the complementary conformation to the epitope. If this mechanism is operational in the [M peptide...Fab M75] association, the charged ArgH50, ArgH52, and LysL55 residues are already in the prerequisite positions in all free Fab molecules and the peptide can be selected on the basis of an interaction with its central part (GluP5 and AspP6). The importance of this interaction is clearly demonstrated by the stabilization of bound peptide, judged by average ADP of peptide residues (Fig. 2); GluP5 and AspP6 have the lowest ADP values among peptide residues. At this point, it can be also noticed that average ADP of peptide is lower than the average ADP of the whole complex.

However, the presented data do not allow us to make a conclusive statement concerning the preference of one of the mechanisms in the complex process of antibody binding. The likely scenario probably include combination of both mechanisms: the proper peptide conformation (in the central part: GluP5 and AspP6) is first selected by an empty binding site of antibody, whose conformation (ArgH50, ArgH52, and LysL55) is relatively close to the complexed state and the induced fit is the final step in accommodation of the flexible parts of antibody and peptide to maximize the favorable interactions and shape complementarities of the interacting surfaces.

Examination of the domain movements using TLS analysis showed that the libration movements of the V_H and V_L domains have a closer mutual resemblance in the complex than in the free Fab molecule (cf., Fig. 6). This suggests that the librations of the V_H and V_L domains may become coupled upon the ligand binding or that these two variable domains tend to librate as a single body when connected by the ligand. This view is supported by the fact that the origins of the libration axes of the V_H and V_L domains are closer in the complex than in the free Fab molecule.

The main motivation for this study was to obtain structural, energetic, and physicochemical information necessary for further protein engineering of this important antibody. Our findings suggest that the antigen binding, especially its interaction with the light chain of the antibody, is not optimal and can be substantially improved. In fact, the construction of scFv derivatives of the V_H-linker-V_L type preserving the favorable interaction of the free amino terminus of the heavy chain (which is inevitably lost in scFv of the reverse V_L-linker-V_H type) can be viewed as the initial step in the rational design of antibody-based agents potentially interfering with the presumed function of the PG domain in oncogenesis. The next step can be initiated by directed mutagenesis, based on the results of the MD simulations, that is, by the replacement of the residues of the antibody predicted to contribute by a small per-residue binding enthalpy to the total binding enthalpy.

MATERIALS AND METHODS

Preparation of Fab M75

IgG 2b M75 was expressed in tissue cultures of hybridoma cell line VU-M75 (deposited in ATCC under code **HB 11128**), adsorbed on a Protein A Sepharose (Sigma) column, and eluted with 0.1*M* sodium citrate, pH 3.0. The eluted fractions were pooled, dialyzed against 10 m*M* sodium phosphate (pH 8.0), loaded onto a DEAE Sephadex column equilibrated with the same buffer, and eluted with a 10–200 m*M* sodium phosphate, pH 8.0, gradient. Eluted IgG was concentrated to 5 mg mL^{-1} and dialyzed against 100 m*M* potassium phosphate, pH 7.2. The samples were then supplemented with 0.13% 2-mercaptoethanol, and 0.9 m*M* EDTA, and papain (Worthington) was added to a final concentration of 2 μg/mL. The course of reaction was monitored by gel filtration on Superdex 200 HR, and the reaction was stopped after 3 h by addition of iodoacetamide (final concentration 5.4 m*M*). Undigested IgG as well as Fc fragments were removed using Protein A Sepharose (BioRad).

Construction and preparation of scFv M75

The coding sequences for V_L and V_H variable domains of mAb M75 were obtained by RT-PCR using total RNA from hybridoma M75 cells. The N-terminal primers reflected the experimentally determined amino acid sequences, and the C-terminal primers corresponded to constant regions adjacent to the variable domains.[25] PCR products were sequenced (providing thus information for Fab M75 crystallographic model building) and used as templates for reamplification with modified primers allowing assembly of the V_L and V_H coding DNA fragments into scFv molecule in the format (V_H)-$(Gly_4Ser)_4$-(V_L). The scFv molecule contains 118 N-terminal residues of the heavy chain linked to113 N-terminal residues of the light chain, followed by a C-myc tag sequence EQKLISEEDL to allow for the detection of scFv M75 by anti-myc antibody. For expression in *E. coli* BL21(DE3) cells, a modified pET-22(b) vector was used, in which the scFv coding sequence is preceded by the PelB signal sequence, allowing translocation of the prod-

uct into the periplasmic space, followed by His_5 tag, allowing product isolation and purification by IMAC chromatography on Ni-CAM (Sigma).

Crystallization and data collection

Crystals of free Fab M75 and crystals of Fab M75 in complex with the epitope peptide PGEEDLPGEEDL (M peptide) were obtained by the hanging drop vapor diffusion technique. Crystallization conditions for free Fab M75 were: Fab M75 10 mg/mL; 100 m*M* Tris-HCl, pH 7.5; 15% PEG 4000; 18°C. Crystallization conditions for the [Fab M75 + M peptide] complex were: Fab M75 7.8 mg/mL; M peptide 10 mg/mL; 100 m*M* Tris-HCl, pH 7.4; 20% PEG 2000; 18°C. The best crystals of the complex were prepared by streak seeding. For cryoprotection, the crystals were soaked in the reservoir solution supplemented with 25% glycerol. Diffraction data for both crystals were collected at 150 K using an in-house diffractometer (Nonius FR 591 generator, 345-mm MarResearch Image Plate detector). All data were integrated and reduced using MOSFLM[26] and scaled using SCALA[27] from the CCP4 suite.[28]

Structure determination and refinement

The free Fab M75 structure was solved by molecular replacement using EPMR program.[29] The structure of Fab Bv04-01 (**PDB code: 1NBV**) was used as a search model. Structure factors in the 15–4 Å resolution range were used for both, rotational and translational searches. In the rigid body refinement (part of the CNS[30] program package) V_L, C_L, V_H, and C_H subdomains were allowed to move independently. In the next step of refinement, all the hypervariable complementarity determining regions (CDRs) were omitted from the initial model, which was then subjected to alternate cycles of molecular dynamics refinement with the programs CNS, using the parameter set of Engh and Huber,[31] and manual model building using the program XtalView.[32] For final refinement stage of V_L, C_L, V_H, and C_H subdomains, TLS refinement cycles in program REFMAC 5.2.0019[33] were used.

The structure of the [Fab M75 + peptide] complex was isomorphous with the structure of free Fab M75 (see above), therefore, the structure of free Fab M75 could be used as a starting model for rigid body refinement with free movements of V_L, C_L, V_H, and C_H subdomains. Peptide tracing was carried out using electron density maps calculated with weighted Fourier coefficients.[34] For final stage refinement of V_L, C_L, V_H, and C_H subdomains, program REFMAC 5.2.0019 with TLS cycles was used.

TLS analyses, including determination of residual atomic displacement parameters, were done by TLSANL program[35] from the CCP4 suite.[28]

Isothermal calorimetry

Binding of epitope peptide DLPGEEDLPGEE to scFv M75 was monitored using the VP-ITC titration microcalorimeter (MicroCal, Northampton) at 25°C. The solutions of reactants were prepared in 20 m*M* Tris (or triethanolamine, tricine) pH 8.4, containing 100 m*M* NaCl, and their exact concentrations were determined by amino acid analysis. Nine-microliter aliquots of 130 μ*M* peptide were injected stepwise into a sample cell containing 1.43 mL of 8 μ*M* scFv M75 until complete saturation was attained. Each experiment was accompanied by a corresponding control experiment in which the peptide was injected into buffer alone. The heat of the binding reaction between the peptide and scFv M75 was obtained as the difference between the measured heat of reaction and the corresponding heat of dilution determined in a control experiment.[36] The thermodynamic parameters, stoichiometry and association constant K_a were calculated by Origin 7.0 software (MicroCal) using a single binding site model. A proton transfer between the binding complex and the buffer molecules was evaluated by performing the titration experiments in three buffers with different ionization enthalpies (47.2 kJ mol^{-1} for Tris buffer, 33.6 kJ mol^{-1} for triethanolamine, and 32.0 kJ mol^{-1} for tricine). The calculation of parameters n_H and H_b was performed as described previously[37] according to

$$\Delta H_{ITC} = \Delta H_b + n_H \cdot \Delta H_{ion} \quad (2)$$

where n_H is the number of absorbed protons upon binding (or released if n_H is negative), ΔH_b is the buffer-independent binding enthalpy, ΔH_{ion} is the ionization enthalpy of the buffer, and ΔH_{ITC} is an apparent enthalpy. The analysis of data yields ΔH_b and $\Delta G = -RT \ln K_a$. The entropy change was obtained by using the standard thermodynamic equation $\Delta G = \Delta H - T\Delta S$. By repeating a titration at different temperatures we can determine the change in heat capacity ΔC_p according to

$$\Delta H_{ITC} = \Delta H_0 + \Delta C_p(T - T_0) \quad (3)$$

where ΔH_0 is the binding enthalpy at a reference temperature T_0.[17,38]

Molecular dynamics simulations

General

Molecular dynamics (MD) calculations were carried out using the AMBER 8 program.[39] Throughout the simulations, Duan *et al.* force field (*all_amino03.in*, *parm99.dat* parameter sets in standard AMBER libraries) was used.[40] In total, six trajectories were obtained, viz. Fab M75 protein represented by V_L and V_H domains in free and complexed form with GPGEEDLPG peptide

(trajectories 1, 2), scFv M75 protein in its free (both [ArgH50$^{(+)}$,ArgH52$^{(+)}$] and [ArgH50$^{(0)}$,ArgH52$^{(+)}$] states) and complexed (with DLPGEEDLPGEE peptide) form (3–5), and free DLPGEEDLPGEE peptide (in order to faithfully reproduce the ITC measurements) (6).

The computational protocol consisted of the following steps (i) 20 ps MD equilibration at constant volume (NVT) and T = 298.15 K with all nonhydrogen atoms from protein restrained to their original positions in the crystal structure in a sphere of water molecules with radius of 36 Å from the protein center of mass, using TIP3P water model,[41] and the weak-coupling algorithm[42] with a time constant of 1.0 ps for the heat bath coupling; (ii) 200 ps MD equilibration + 500 ps data collection at constant pressure (1 atm), T =298.15 K, no restraints, and bonds between heavy atoms and hydrogens restrained by SHAKE algorithm[43]; (iii) MM-GBSA calculations (described below) with all the waters stripped off, using the set of 100 snapshots (every 5 ps) from MD trajectories, including a decomposition of the total (free) energy on a per-residue basis for the trajectory of the complex; (iv) energy minimization (MM) of the final MD structures.

System setup and structure preparation

The scFv M75 model complex was built from the truncated crystal structure of the complex (PDB 2HKF; AspL1 to ArgL113; GluH1 to AlaH120; peptide GlyP1 to GlyP9). The hydrogen atoms were added as to represent the standard protonation states at pH 7 and subject to MM minimization. All histidines were considered in neutral form with $N_{\varepsilon 2}$ nitrogen protonated. The S—S bonds between CysL23-CysL93 and CysH22-CysH98 and all crystallographic water molecules were preserved in the model. The scFv M75 protein was prepared by adding $(Gly_4Ser)_4$ linker to the Fab M75 structure, mutating the appropriate residues, and extending the peptide N- and C-termini by one and two residues (respectively) using program XtalView. The scFv M75 model was subject to MM minimization of the added $(Gly_4Ser)_4$ loop with all nonhydrogen atoms and crystal water oxygens constrained to their crystallographic positions. This was followed by 300 ps simulated annealing procedure with temperature raised to 450 K for the first 50 ps, kept on this value for 100 ps, and cooling the system down to 0 K for the remaining 150 ps. Finally, the structure was minimized without any constraints.

MM-GBSA procedure

The *mmpbsa* module of Amber was used to calculate the free energy change upon binding. In MM-GBSA procedure,[44] this energy is defined as

$$\Delta G_{\text{bind}} = G(\text{LP}) - G(\text{L}) - G(\text{P}), \qquad (4)$$

where $G(X)$ is free energy of system X (X = LP, [ligand. . .protein] complex; L, ligand; and P, free protein) in water solution.

It is calculated according to the following formula:

$$G = \langle E_{\text{MM}} \rangle + \langle G_{\text{solv}} \rangle + \langle G_{\text{np}} \rangle - T \langle S_{\text{MM}} \rangle . \qquad (5)$$

In Eq. 5, $\langle X \rangle$ is an average value of X taken from the snapshots generated during the MD simulations, E_{MM} is the internal molecular mechanics energy (comprises internal bonding energy terms, nonbonding electrostatic and van der Waals interations; $E_{\text{MM}} = E_{\text{int}} + E_{\text{elstat}} + E_{\text{vdW}}$), G_{solv} is the solvation free energy calculated by solving generalized Born[44] equations; G_{np} is the nonpolar part of the solvation free energy calculated from solvent-accessible surface area (SASA), and S_{MM} is the entropy calculated from the normal-mode analysis in the ideal gas approximation using the standard formulas of statistical mechanics. As can be seen from Eq. (5), we cannot directly compare the total entropy and enthalpy contributions with the experimental values, since the solvation entropies cannot be simply obtained from G_{solv} and G_{np}. If we denote S_{MM} as the conformational entropy, this, however, represents the major entropy contribution to the binding free energies.

ACKNOWLEDGMENTS

We thank Hillary Hoffman for critical reading of the manuscript and Dr. Mojmír Kývala for his help with statistical analysis of ITC data.

REFERENCES

1. Ivanov S, Liao SY, Ivanova A, Danilkovitch-Miagkova A, Tarasova N, Weirich G, Merrill MJ, Proescholdt MA, Oldfield EH, Lee J, Zavada J, Waheed A, Sly W, Lerman MI, Stanbridge EJ. Expression of hypoxia-inducible cell-surface transmembrane carbonic anhydrases in human cancer. Am J Pathol 2001;158:905–919.
2. Chrastina A, Zavada J, Parkkila S, Kaluz S, Kaluzova M, Rajcani J, Pastorek J, Pastorekova S. Biodistribution and pharmacokinetics of 125I-labeled monoclonal antibody M75 specific for carbonic anhydrase IX, an intrinsic marker of hypoxia, in nude mice xenografted with human colorectal carcinoma. Int J Cancer 2003;105:873–881.
3. Pastorekova S, Zavada J. Carbonic anhydrase IX (CA IX) as a potential target for cancer therapy. Cancer Ther 2004;2:245–262.
4. Zavada J, Zavadova Z, Pastorek J, Biesova Z, Jezek J, Velek J. Human tumour-associated cell adhesion protein MN/CA IX: identification of M75 epitope and of the region mediating cell adhesion. Br J Cancer 2000;82:1808–1813.
5. Oosterwijk E, Ruiter DJ, Hoedemaeker PJ, Pauwels EK, Jonas U, Zwartendijk J, Warnaar SO. Monoclonal antibody G 250 recognizes a determinant present in renal-cell carcinoma and absent from normal kidney. Int J Cancer 1986;38:489–494.
6. Zat'ovicova M, Tarabkova K, Svastova E, Gibadulinova A, Mucha V, Jakubickova L, Biesova Z, Rafajova M, Ortova Gut M, Parkkila S, Parkkila AK, Waheed A, Sly WS, Horak I, Pastorek J, Pastorekova S. Monoclonal antibodies generated in carbonic anhydrase IX-deficient mice recognize different domains of tumour-associated hypoxia-induced carbonic anhydrase IX. J Immunol Methods 2003;282:117–134.

7. Zavadova Z, Zavada J. Carbonic anhydrase IX (CA IX) mediates tumor cell interactions with microenvironment. Oncol Rep 2005;13: 977–982.
8. Alterio V, Vitale RM, Monti SM, Pedone C, Scozzafava A, Cecchi A, De Simone G, Supuran CT. Carbonic anhydrase inhibitors: X-ray and molecular modeling study for the interaction of a fluorescent antitumor sulfonamide with isozyme II and IX. J Am Chem Soc 2006;128:8329–8335.
9. Pastorekova S, Parkkila S, Zavada J. Tumor-associated carbonic anhydrases and their clinical significance. Adv Clin Chem 2006;42: 167–216.
10. Opavsky R, Pastorekova S, Zelnik V, Gibadulinova A, Stanbridge EJ, Zavada J, Kettmann R, Pastorek J. Human MN/CA9 gene, a novel member of the carbonic anhydrase family: structure and exon to protein domain relationships. Genomics 1996;33:480–487.
11. Bhat TN, Bentley GA, Boulot G, Greene MI, Tello D, Dall'Acqua W, Souchon H, Schwarz FP, Mariuzza RA, Poljak RJ. Bound water molecules and conformational stabilization help mediate an antigen-antibody association. Proc Natl Acad Sci USA 1994;91:1089–1093.
12. Rini JM, Schulze-Gahmen U, Wilson IA. Structural evidence for induced fit as a mechanism for antibody-antigen recognition. Science 1992;255:959–965.
13. James LC, Roversi P, Tawfik DS. Antibody multispecificity mediated by conformational diversity. Science 2003;299:1362–1367.
14. Rezacova P, Lescar J, Brynda J, Fabry M, Horejsi M, Sedlacek J, Bentley GA. Structural basis of HIV-1 and HIV-2 protease inhibition by a monoclonal antibody. Structure 2001;9:887–895.
15. Keitel T, Kramer A, Wessner H, Scholz C, Schneider-Mergener J, Hohne W. Crystallographic analysis of anti-p24 (HIV-1) monoclonal antibody cross-reactivity and polyspecificity. Cell 1997;91:811–820.
16. Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russell RB. Protein disorder prediction: implications for structural proteomics. Structure 2003;11:1453–1459.
17. Jelesarov I, Bosshard HR. Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. J Mol Recognit 1999;12: 3–18.
18. Milner-White E, Ross BM, Ismail R, Belhadj-Mostefa K, Poet R. One type of gamma-turn, rather than the other gives rise to chain-reversal in proteins. J Mol Biol 1988;204:777–782.
19. Winn MD, Murshudov GN, Papiz MZ. Macromolecular TLS refinement in REFMAC at moderate resolutions. Methods Enzymol 2003; 374:300–321.
20. Mandiyan V, O'Brien R, Zhou M, Margolis B, Lemmon MA, Sturtevant JM, Schlessinger J. Thermodynamic studies of SHC phosphotyrosine interaction domain recognition of the NPXpY motif. J Biol Chem 1996;271:4770–4775.
21. Weis A, Katebzadeh K, Soderhjelm P, Nilsson I, Ryde U. Ligand affinities predicted with the MM/PBSA method: dependence on the simulation method and the force field. J Med Chem 2006;49:6596–6606.
22. Gestwicki JE, Cairo CW, Strong LE, Oetjen KA, Kiessling LL. Influencing receptor-ligand binding mechanisms with multivalent ligand architecture. J Am Chem Soc 2002;124:14922–14933.
23. Hlavacek WS, Posner RG, Perelson AS. Steric effects on multivalent ligand-receptor binding: exclusion of ligand sites by bound cell surface receptors. Biophys J 1999;76:3031–3043.
24. Berger C, Weber-Bornhauser S, Eggenberger J, Hanes J, Pluckthun A, Bosshard HR. Antigen recognition by conformational selection. FEBS Lett 1999;450:149–153.
25. Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C. Sequences of proteins of immunological interest. Bethesda, MD: US Department of Health and Human Services, Public Health Service, National Institutes of Health; 1991.
26. Leslie AG. Integration of macromolecular diffraction data. Acta Crystallogr D Biol Crystallogr 1999;55(Part 10):1696–1702.
27. Evans PR. Data reduction. In: Sawyer L, Isaacs N, Bailey S, editors. Data collection and processing. Warrington: Daresbury Laboratory; 1993. pp 114–122.
28. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 1994;50:760–763.
29. Kissinger CR, Gehlhaar DK, Fogel DB. Rapid automated molecular replacement by evolutionary search. Acta Crystallogr D Biol Crystallogr 1999;55:484–491.
30. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography and NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998;54:905–921.
31. Engh RA, Huber R. Accurate bond and angle parameters for X-ray protein structure refinement. Acta Crystallogr A 1991;47:392–400.
32. McRee DE. Practical protein crystallography. San Diego: Academic Press; 1993.
33. Winn M, Isupov M, Murshudov GN. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr D Biol Crystallogr 2001;57:122–133.
34. Read RJ. Improved Fourier coefficients for maps using phases from partial structures with errors. Acta Crystallogr A 1986;42:140–149.
35. Howlin B, Butler SA, Moss DS, Harris GW, Driessen HPC. TLSANL: TLS parameter analysis program for segmented anisotropic refinement of macromolecular structures. J Appl Cryst 1993; 26:622–624.
36. Pierce MM, Raman CS, Nall BT. Isothermal titration calorimetry of protein-protein interactions. Methods 1999;19:213–221.
37. Fukada H, Takahashi K. Enthalpy and heat capacity changes for the proton dissociation of various buffer components in 0.1 M potassium chloride. Proteins 1998;33:159–166.
38. Todd MJ, Luque I, Velazquez-Campoy A, Freire E. Thermodynamic basis of resistance to HIV-1 protease inhibition: calorimetric analysis of the V82F/I84V active site resistant mutant. Biochemistry 2000;39:11876–11883.
39. Case DA, Darden TA, Cheatham TE, III, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Wang B, Pearlman DA, Crowley M, Brozell S, Tsui V, Gohlke H, Mongan J, Hornak V, Cui G, Beroza P, Schafmeister C, Caldwell JW, Ross WS, Kollman PA. AMBER 8. San Francisco: University of California; 2004.
40. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 2003;24:1999–2012.
41. Jorgensen WL, Chandrasekhar J, Madura J, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
42. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
43. Ryckaert J-P, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 1977;23:327–341.
44. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE, III. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 2000;33:889–897.

Journal of Medicinal Chemistry Article

J. Med. Chem. **2010,** *53,* 2401–2408 **2401**
DOI: 10.1021/jm9014026

Identification of 3,4-Dihydroisoquinoline-2(1*H*)-sulfonamides as Potent Carbonic Anhydrase Inhibitors: Synthesis, Biological Evaluation, and Enzyme−Ligand X-ray Studies†

Rosaria Gitto,*,‡ Stefano Agnello,‡ Stefania Ferro,‡ Laura De Luca,‡ Daniela Vullo,§ Jiri Brynda,∥,⊥ Pavel Mader,∥,⊥ Claudiu T. Supuran,§ and Alba Chimirri‡

‡*Dipartimento Farmaco-Chimico, Università di Messina, Viale Annunziata, I-98168 Messina, Italy,* §*Laboratorio di Chimica Bioinorganica, Università degli Studi di Firenze, Firenze, Italy,* ∥*Department of Structural Biology, Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, CZ-16637 Prague, Czech Republic, and* ⊥*Department of Structural Biology, Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, CZ-16637 Prague, Czech Republic*

Received September 21, 2009

Following previous studies we herein report the exploration of the carbonic anhydrase (CA, EC 4.2.1.1) inhibitory effects and enzyme selectivity of a small class of 1-(cyclo)alkylisoquinolines containing a sulfonamide function considered a key feature for inhibiting CA. The results of enzymatic assays against human (h) CA isoforms, hCA I and hCA II (cytosolic, ubiquitous enzymes), hCA IX (transmembrane, tumor-associated), and hCA XIV (transmembrane), suggested that the presence of C-1 small substituents on isoquinoline scaffold controls both inhibitory potency and selectivity. Some derivatives showed potent hCA IX and hCA XIV inhibitory effects at nanomolar concentrations as well as low affinity for the ubiquitous hCA II. Moreover, we report the X-ray crystal structure of one of these derivatives in complex with dominant human isoform II, thus confirming the sulfonamide−zinc interactions. Finally, the results of docking experiments suggested the hypothetic interactions in the catalytic binding site for the most active and selective hCA IX and hCA XIV inhibitor.

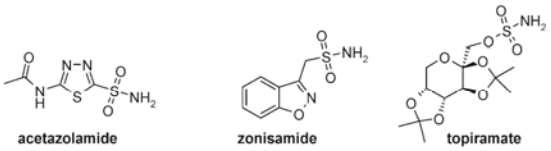
Introduction

The carbonic anhydrases (CAs, EC 4.2.1.1) are a family of monomeric zinc metalloenzymes that catalyze the reversible hydration of CO_2. Since this reaction regulates a broad range of physiological functions, the pharmacological modulation of CA activity could be useful for the treatment of several human diseases. There are 15 human known CA (hCA[a]) isoforms with different tissue distribution, expression levels, and subcellular locations. Some of these isozymes (e.g., hCA II, IV, VA, VB, VII, IX, XII, XIII, and XIV) constitute valid targets for the development of anticancer, antiglaucoma, antiobesity, or anticonvulsant drugs.[1−8] However, the CA diffuse localization in many tissues and organs limits potential clinical applications. So the development of CA inhibitors (CAIs) possessing high potency and selectivity against some specific isoforms represents an attractive strategy to obtain pharmacological tools, thus avoiding side effects and improving therapeutic safety.

In particular, there is significant interest in the development of selective inhibitors targeting isozymes CAIX and CAXIV. In particular, CA IX is a peculiar member of the CA family, since it is expressed in a limited number of normal tissues (mainly the gastrointestinal tract), whereas its overexpression is seen on the cell surface of a large number of solid tumors showing inadequate supply of oxygen as key feature.[5] This tumor hypoxia regulates the expression of several genes, including CA IX, through the hypoxia inducible factor 1 (HIF-1) cascade. Furthermore, CA IX overexpression is often associated with a poor responsiveness to the classical radiotherapy and chemotherapy. So the development of selective CAIX isoform inhibitors represents a new strategy to design anticancer drugs with a novel mechanism of action.[9] Similar to CA IX, CA XIV is a transmembrane isozyme with the active site oriented extracellularly; it is highly abundant in neurons and axons in the murine and human brain, where it seems to play an important role in modulating excitatory synaptic transmission.[10]

Most of known CAIs contain a sulfonamide/sulfamate moiety able to coordinate the zinc ion of catalytic binding site (e.g., acetazolamide, zonisamide, and topiramate, Chart 1), inhibiting in this way the enzymatic activity.[11−18] These inhibitors bear specific functional groups that interact with important amino acid residues, thus driving the selectivity against the different isoforms as confirmed by X-ray

Chart 1. Chemical Structures of the Carbonic Anhydrase Inhibitors: Acetazolamide, Zonisamide, and Topiramate



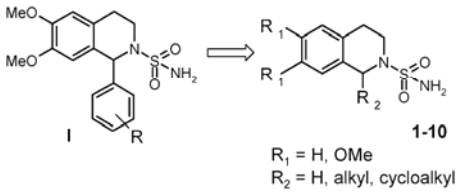
†The atomic coordinates and crystal structure have been deposited in the Protein Data Bank as entry 3IGP.

*To whom correspondence should be addressed. Phone: 00390906766413. Fax: 00390906766402. E-mail: rgitto@pharma.unime.it.

[a]Abbreviations: hCA, human carbonic anhydrase; HIF-1, hypoxia inducible factor 1; MW, microwave; PDB, Protein Data Bank; rmsd, root-mean-square deviation; TFA, trifluoroacetic acid.

Published on Web 02/19/2010

Chart 2. New Designed Isoquinoline Derivatives **1**−**10** as Carbonic Anhydrase Inhibitors



crystallographic data of complexes between CAIs and isozymes available in the literature.[19−26]

In a previous paper[27] we identified some 1-aryl-6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamides (**I**, Chart 2) that proved to inhibit some CA isoforms at micromolar concentration but show poor selectivity; our hypothesis was that the presence of bulky C-1 aryl substituent could produce steric hindrance for interaction between the sulfonamide group and the metal ion in the catalytic site.

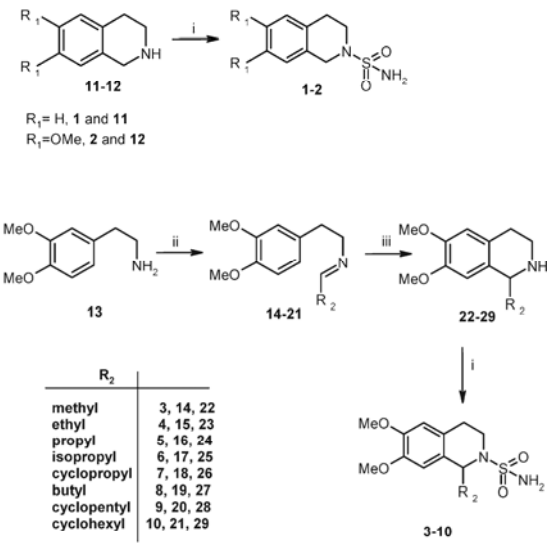
So in this paper we extended our study to a new series of 3,4-dihydroisoquinoline-2(1*H*)-sulfonamides (**1**−**10**, Chart 2) containing some structural modifications in search of the enhancement of both activity and selectivity. In particular, (a) to evaluate the steric and electronic effects, we planned the replacement of the C-1 aryl substituent with alkyl and cycloalkyl groups; (b) to test the role of the substituents on the isoquinoline scaffold, we synthesized compounds lacking of the C-1 substituent as well as the two methoxy groups. The synthesis of target compounds was accomplished in microwave conditions and the screening against some relevant CA isoforms was performed. In particular, the obtained compounds were tested against the two physiologically human cytosolic hCA I and hCA II isozymes, the tumor-associated transmembrane isoform hCA IX, and the neuronal hCA XIV isoform.

Moreover, we report the X-ray crystal structure of the dominant hCA II isoform complexed with a new derivative, allowing the determination of the sulfonamide−zinc interactions in the catalytic binding site. Docking experiments have been performed with the aim to identify the interactions promoting the selectivity toward hCA IX and hCA XIV isoforms.

Results and Discussion

Chemistry. The synthetic pathways for the 3,4-dihydroisoquinoline-2(1*H*)-sulfonamides (**1**−**10**) can be found in Scheme 1. Derivatives **1** and **2** were readily synthesized in microwave conditions starting from commercially available isoquinolines **11** and **12** and a large excess of sulfamide as previous reported by us.[27] The precursors 1-(cyclo)alkyl-6,7-dimethoxy-1,2,3,4-tetrahydroisoquinoline **22**−**29** were prepared in accordance with an optimized microwave-assisted approach through the Pictet−Spengler condensation.[27,28] Starting from the commercially available 2-(3′,4′-dimethoxyphenyl)ethylamine (**13**), the imine intermediates **14**−**21** were obtained by reaction with the appropriate aldehyde in solvent-free conditions and successively cyclized in acidic medium to give the desired compounds **22**−**29**. Finally, the intermediates **22**−**29** were coupled with a large excess of sulfamide, affording the corresponding 6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamides (**3**−**10**). The structures of all obtained compounds were supported by elemental analyses and spectroscopic measurements.

Scheme 1. Synthesis of 3,4-Dihydroisoquinoline-2(1*H*)-sulfonamides (**1**−**10**)[a]



[a] Reagents and conditions. (i) $CH_3CH(OCH_3)_2$, $NH_2SO_2NH_2$, two steps in the same conditions: 20 min, 100 °C, 200 psi, 150 W. (ii) R_2CHO, MW: 5 min, 90 °C, 200 psi, 150 W. (iii) TFA, MW: 5 min, 90 °C, 200 psi, 150 W.

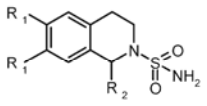
Carbonic Anhydrase Inhibition. To determine the enzymatic inhibitory activity, the new series of 6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamides (**1**−**10**) was assayed on four important carbonic anhydrase isoforms involved in several physiological and pathological processes, hCA I, hCA II, hCA IX, and hCA XIV (Table 1).

The enzymatic screening showed that the 3,4-dihydroisoquinoline-2(1*H*)-sulfonamides **1**−**10** have a moderate inhibitory activity against human carbonic anhydrase I (hCA I), with inhibition constants (K_I) in the range 0.10−6.41 μM (see Table 1).

Our biological data reported in Table 1 highlighted that the isozyme hCA II was inhibited by the 6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamides (**1**−**10**) with K_I values ranking from 32.8 nM to 350 μM, suggesting that there is significant impact of the size of the C-1 substituent on inhibitory efficacy; in fact the most active sulfonamide derivatives of this series were unsubstituted and methyl-substituted derivatives **1**, **2**, and **3**. For the other (cyclo)alkyl derivatives **4**−**10** the inhibitory effects were much lower than those of compounds **1**−**3**. In particular, the 6,7-dimethoxy-1-propyl-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (**5**) and the corresponding superior homologous 6,7-dimethoxy-1-butyl-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (**8**) were about 2500- to 3700-fold less active than the parent unsubstituted 6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (**2**) on hCA II; it is probably due to the key role of C-1 substituent in the recognition process into the active catalytic site.

All synthesized compounds **1**−**10** demonstrated generally very high and similar potency against hCA IX and hCA XIV isoforms with a significant variation of the inhibitory effects that appear related to the nature of C-1 substituent. We found that the unsubstituted (R_2 = H) derivatives **1** and **2** as well as derivatives containing small (cyclo)alkyl substituents

Table 1. Inhibition of hCA I, hCA II, hCA IX, and hCA XIV Isoforms by 6,7-Dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamides (**1**−**10**), 6,7-Dimethoxy-1-phenyl-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (**I**), Zonisamide, Acetazolamide, and Topiramate and Selectivity Ratios K_I(hCAII)/K_I(hCAIX) and K_I(hCAII)/K_I(hCAXIV)



			K_I (nM)[a]				selectivity ratios[b]	
R_1	R_2	compd	hCA I	hCA II	hCA IX	hCA XIV	K_I(hCAII)/K_I(hCAIX)	K_I(hCAII)/K_I(hCAXIV)
H	H	**1**	1510	32.8	3.7	6.0	8.86	5.47
MeO	H	**2**	6410	94.5	9.5	9.8	9.95	9.64
MeO	methyl	**3**	2800	87.3	9.4	9.6	9.29	9.09
MeO	ethyl	**4**	1820	1975	76.3	2754	25.9	0.72
MeO	*n*-propyl	**5**	3150	235000	330	302	712.12	778.15
MeO	isopropyl	**6**	3780	1050	6.1	4.7	172.13	223.40
MeO	cyclopropyl	**7**	3950	820	8.5	6.4	96.47	128.13
MeO	*n*-butyl	**8**	4100	350000	400	1650	875.00	212.12
MeO	cyclopentyl	**9**	100	650	55	178	11.82	3.65
MeO	cyclohexyl	**10**	1180	18890	706	7950	26.76	2.38
MeO	phenyl	**I**[c]	8980	15700	8440	3860	1.86	4.07
acetazolamide[d]			250	12	25	41	0.48	0.29
zonisamide[d]			56	35	5.1	5250	6.86	< 0.01
topiramate[d]			250	10	58	1460	0.17	< 0.01

[a] Errors in the range of ±10% of the reported value, from three different assays. Recombinant full length hCA I, II, and XIV and catalytic domain of hCA IX were used. [b] The K_I ratios indicate the inhibition selectivity. [c] From ref 27. [d] From ref 5.

(e.g., **3**, **6**, and **7**) showed the strongest inhibitory activity at nanomolar concentration against hCA IX and hCA XIV isoforms with inhibition constants in the range 3.7−9.5 nM for hCA IX and 4.7−9.8 nM for hCA XIV. The best inhibitor was the 6,7-dimethoxy-1-isopropyl-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (**6**) showing K_I values of 6.1 and 4.7 nM against hCA IX and hCA XIV, respectively. Furthermore compound **4** (R_2 = ethyl) showed significant selectivity for inhibiting hCA IX (K_I of 76.3 nM) over hCA XIV (K_I of 2754 nM). We observed that the presence of bulky (cyclo)alkyl groups (e.g., **5**, **8**, **9**, and **10**) led to a decrease of the activity, with K_I in the range 55.0−706 nM for hCA IX and 178−7950 nM for hCA XIV. Nevertheless, the inhibitor 6,7-dimethoxy-1-cyclopentyl-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (**9**) was ∼3-fold more active on hCA IX than hCA XIV (K_I of 55.0 nM versus K_I of 178.0 nM).

Interestingly some of these compounds showed higher selectivity for hCA IX and hCA XIV isoforms over the ubiquitous isoform hCA II, which can be considered an off-target. In fact, we found that for our newly synthesized CAIs **4**−**10** the selectivity ratios (see Table 1) for inhibiting hCA II over hCA IX were in the range 11.8−875, whereas their selectivity ratios for inhibiting hCA II over hCA XIV were in the range 128−778.

The most remarkable result is that the potent hCA IX and hCA XIV inhibitor 6,7-dimethoxy-1-isopropyl-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (**6**) is 172-fold and 223-fold more potent on hCA IX and hCA XIV than on hCA II. Moreover, the analysis of the selectivity ratio results highlighted that the presence of bulky C-1 substituents on isoquinoline scaffold led to the most selective inhibitors (e.g., **5** and **8**, with R_2 = propyl or butyl group) even if they are less potent than the compounds that bear small groups on tetrahydroisoquinoline skeleton (e.g., compounds **1**−**3**). Apparently for this class of compounds the introduction of cyclopropyl and isopropyl groups produces the optimization of both potency and selectivity, suggesting that in this region there is a restrictive steric requirement for the catalytic binding site interaction. The enhancement of potency and selectivity was very significant when we compared the obtained results for compounds **1**−**10** with those of other known inhibitors such as acetazolamide, zonisamide, and topiramate (K_I values and selectivity ratios reported in Table 1).

Finally, by considering the results of the current assays with the biological data obtained for the earlier reported prototype (**I**) (see Table 1), it may be observed that the presence of a phenyl group at the C-1 position drives to a flat selectivity and poor inhibitory efficacy for all tested isoforms.[27] So the most important consideration is that the introduction of a suitable small (cyclo)alkyl group in proximity of sulfonamide function could influence its geometrical disposition in the catalytic binding site, thus controlling both the inhibitory potency and selectivity.

Crystallographic Studies. To determine the binding mode and decipher the key interactions contributing to the inhibitory properties for this class of 3,4-dihydroisoquinoline-2(1*H*)-sulfonamides, one of the most active compounds has been cocrystallized with hCA II. The statistics for data collection and refinement are summarized in Table 2. The crystal structure of hCA II in complex with 6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (**2**) was determined by difference Fourier techniques and refined using diffraction data to 1.65 Å resolution. The complex crystallized in the monoclinic $P2_1$ space group with one hCA II molecule in the asymmetric unit and solvent content of 41.3%. All hCA II residues could be traced into a well-defined electron density map with the exception of side chains of several terminal amino acid residues (Ser2, His3, His4, and Lys261) and side chains of two surface residues Lys9 and Lys133. During the course of the crystallographic refinement two continuous non-protein electron densities

Table 2. Crystal Data and Diffraction Data Collection and Refinement Statistics[a]

parameter	
data collection statistics	
space group	$P2_1$
unit cell length (Å)	a = 42.31, b = 41.27, c = 72.04
angle (deg)	α = 90, β = 104.22, γ = 90
no. of molecules in AU	1
wavelength (Å)	1.5418
resolution range (Å)	19.95−1.65 (1.69−1.65)
no. of unique reflections	26,928
redundancy	2.2 (2.2)
completeness (%)	92.5 (88.6)
R_{merge} [b]	0.079 (0.362)
average $I/\sigma(I)$	5.9 (2.0)
Wilson B (Å^2)	13.3
refinement statistics	
resolution range (Å)	19.95−1.65 (1.69−1.65)
no. of reflections in working set	25556 (1811)
no. of reflections in test set	1353 (84)
R [c] (%)	16.08 (22.40)
R_{free} [d] (%)	19.55 (28.90)
rmsd bond length (Å)	0.011
rmsd angle (deg)	1.43
no. of atoms in AU	4949
no. of protein atoms in AU	4138
no. of water molecules in AU	237
mean B (Å^2)	14.0
Ramachandran plot statistics	
residues in favored regions (%)	97.3
residues in allowed regions (%)	2.7
PDB code	3IGP

[a] The data in parentheses refer to the highest-resolution shell. [b] $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where the $I_i(hkl)$ is an individual intensity of the ith observation of reflection hkl and $\langle I(hkl)\rangle$ is the average intensity of reflection hkl with summation over all data. [c] $R = \sum ||F_o| - |F_c|| / \sum |F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively. [d] R_{free} is equivalent to R but is calculated for 5% of the reflections chosen at random and omitted from the refinement process.[58]

were noticeable in the active site and on the surface of the molecule, respectively; both could be unambiguously modeled as the 6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (**2**).

As shown in Figure 1, the compound **2** binds into the cavity of the hCA II active site with the deeply buried sulfonamide group. The ionized nitrogen atom of the sulfonamide moiety is coordinated to the zinc ion at a distance of 2.07 Å. The sulfonamide nitrogen also donates a hydrogen bond to Oγ of Thr199, and one oxygen from the sulfonamide moiety forms a hydrogen bond with backbone amine group of Thr199 (Figure 1). These key hydrogen bonds between the sulfonamide moiety of the inhibitor and enzyme active site are also found in other structurally characterized hCA II−sulfonamide complexes.[4,11,29,30]

The sulfonamide moiety seems to serve as an anchor for specific binding of compound **2** into the enzyme active site. In addition to the polar interactions mediated by the sulfonamide group, hydrophobic interactions of the (substituted) isoquinoline moiety strongly stabilize the inhibitor within the active site cavity. The isoquinoline moiety forms numerous van der Waals interactions (distance of < 4.2 Å) with the bottom of the active site (residues Gln92, Val121, Val143, Trp209) and the loop formed by residues Ser197−Pro202.

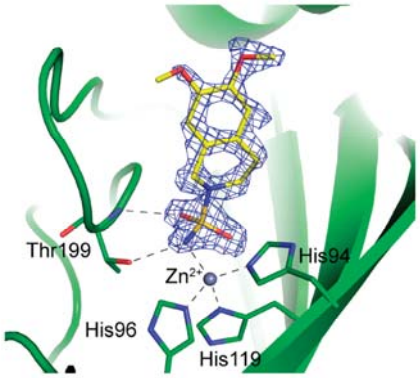


Figure 1. Binding of 6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (**2**) to the hCA II. Detail of the hCA II active site with the inhibitor represented as a stick model (with carbon and oxygen atoms colored yellow and red, respectively). The $2F_o - F_c$ electron density maps are contoured at 1.5σ. Protein is represented in green as a cartoon model with residues forming polar contacts with inhibitor highlighted in sticks. Also three histidine residues coordinating zinc ion are shown.

Also, the side chain of Phe131 from helix 4 contributes to interaction with the substituting methoxy-groups (Table S1 in Supporting Information). In the vicinity of the inhibitor, a glycerol molecule could be modeled into the hCA II active site, making hydrogen bonds with Asn62, Asn67, and Gln92 (Figure S1 in Supporting Information). The binding of glycerol to an identical site was also observed in several other hCA II−inhibitor complexes (e.g., 2NNG, 2NNS, 2NNO, 2NNV[29]). Glycerol was used at high concentration (20% v/v) for crystal cryoprotection, and its binding probably represents a crystallization artifact.

Also, the inhibitor second binding site identified on the protein surface probably has no biological relevance and represents a crystallization artifact caused by high concentrations of inhibitor employed in the cocrystallization experiments. The inhibitor **2** interacts through direct hydrogen bonds with surface residues His15 and Asp19 and through a water mediated hydrogen bond with Ser2 (Figure S2 in Supporting Information). This inhibitor second binding site is also occupied by sulfonamide inhibitors in other hCA II−inhibitor complexes (e.g., PDB codes 2FOS, 2FOV, 2FOQ, 2FOU;[31] 2NNO, 2NNS, 2NNV;[29] and 1ZFQ).

The inhibitor binding does not cause any major structural changes to the protein; the root-mean-square deviation (rmsd) for superposition of 256 Cα atoms (residues 4−259) of our complex structure with free hCA II (PDB code 1CA2[30]) is 0.34 Å, which is below the value observed for identical structures.[32]

Docking Studies. With the aim to clarify the differences observed in the affinity and selectivity of our compounds toward hCA IX or hCA XIV over hCA II, preliminary docking analysis[33] has been also performed. These computational studies were carried out using crystal structures of hCA II, hCA IX, and hCA XIV complexed with acetazolamide from the RCSB Protein Data Bank (1YDB,[34] 3IAI,[35] and 1RJ6[36]). Then we docked the most active compound 6, 7-dimethoxy-1-isopropyl-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (**6**) in the active site (see details in Experimental Section). The active sites of hCA II, hCA IX, and hCA XIV isoforms appear as a compact globular domain, located in a

large conical cavity that spans from the surface to the center of the protein, and present only minor differences in the region 125−137 and in the loop incorporating residues 198−204.[35] In particular, the hCA II residues Phe131, Val135, and Leu204 are replaced by Val131, Leu135, and Ala204 (hCA IX) and by Leu131, Ala135, and Tyr204 (hCA XIV).

Figure 2 displays the superposition of docking pose of compound **6** with X-ray position of acetazolamide in hCA II. Moreover, it also shows the X-ray position of **2** (see X-ray section) obtained through protein backbone-based structural superposition. It is possible to note that in all three inhibitors the sulfonamide groups assume similar orientations, sharing the key hydrogen bond interactions with the zinc ion and Thr199. The acetazolamide and compound **2** show a very similar binding mode stabilized by both interactions with the loop and a hydrophobic contact with residue Phe131, whereas the compound **6** appears locked in a different orientation, probably due to hydrophobic interactions with Val121, Ala142, and Ala143, through the isopropyl moiety. In this disposition compound **6** lacks the interaction with the loop, thus explaining its lower hCA II inhibitory activity (Table 1). Figure 3 shows the docking results of compound **6** compared with the X-ray position of acetazolamide in complex with hCA IX (Figure 3A) and hCAXIV (Figure 3B). It is possible to observe (Figures 3) that the acetazolamide presents the same sort of interactions as well as similar K_I values (Table 1), whereas compound **6** seems able to establish new additional hydrophobic interactions with the residues Leu135/Ala204 (hCA IX) and Ala135/Tyr204 (hCA XIV), located in the upper part of the binding site, thus justifying the selectivity ratios reported in Table 1.

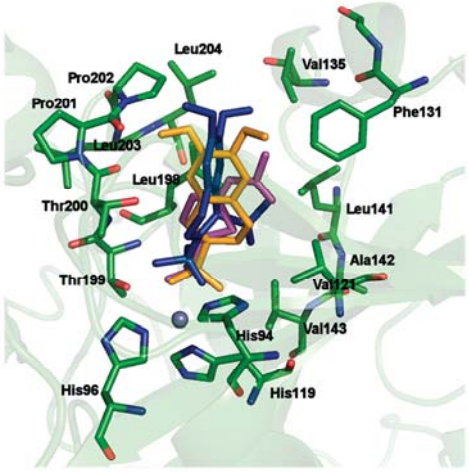


Figure 2. Superposition of inhibitors (acetazolamide, **2** and **6**) into the hCA II catalytic site. The residues participating in recognition of the inhibitors and the catalytic triad (His94, His96, and His119) coordinating zinc ion are also shown. The zinc ion is shown as a gray sphere. Acetazolamide is reported in magenta, **2** in orange, and **6** in blue.

Conclusions

We have identified new isoquinoline derivatives containing sulfonamide moiety that showed very high affinity toward tumor-associated isoform hCA IX as well as neuronal hCA XIV displaying inhibitory effects at nanomolar concentration. In addition, these new compounds proved to be selective CAIs with significant selectivity ratios for inhibiting hCA IX over hCA II in the range 8.86−875 and for inhibiting hCA XIV over hCA II in the range 9.09−778. This study furnished some SAR consideration for this new class of CAIs, suggesting that the nature of the C-1 substituent on the isoquinoline scaffold could play a key role in the inhibitory effects. Considering the importance and the difficulty of obtaining selective CAIs, these results give relevant insights useful for designing new inhibitors having low affinity to the physiological ubiquitous hCA I and hCA II but maintaining inhibitory activity on other druggable isoforms. By X-ray crystallography, we confirmed that these inhibitors could bind the catalytic site of CAs through the sulfonamide moiety. In particular, we studied one of the most active 3,4-dihydroisoquinoline-2(1*H*)-sulfonamide derivative (**2**) in complex with hCA II. Moreover, the key interactions promoting the hCA IX and hCA XIV activity

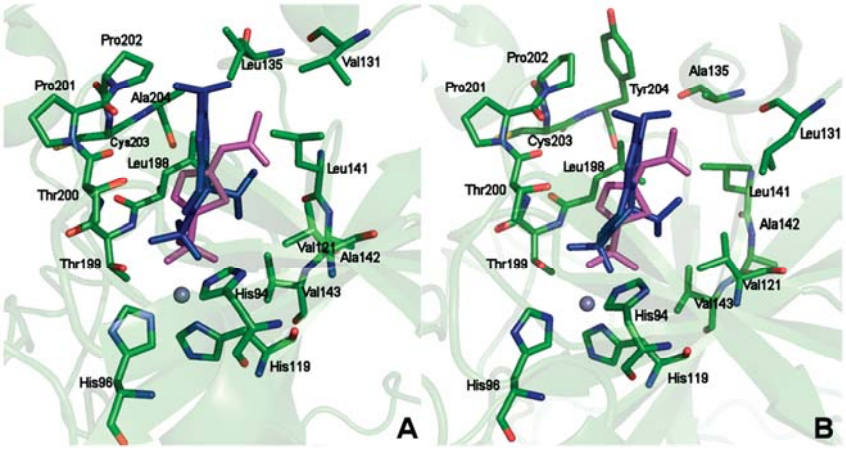


Figure 3. Structures of hCA IX (A) and hCA XIV(B) in complex with acetazolamide (magenta) overlaid with docked structure of compound **6** (blue). The residues participating in recognition of the inhibitors and the three histidine residues coordinating zinc ion are also shown. The zinc ion is shown as gray sphere.

and selectivity of compound **6** were analyzed by docking experiments.

Experimental Section

Chemistry. All reagents were purchased from Sigma Aldrich and were used without further purification. Microwave-assisted reactions were carried out in a CEM focused microwave synthesis system. Melting points were determined on a Buchi melting point B-545 apparatus and are uncorrected. Elemental analyses (C, H, N) were carried out on a Carlo Erba model 1106 elemental analyzer, and the results are within ±0.4% of the theoretical values. Merck silica gel 60 F254 plates were used for analytical TLC; R_f values were determined employing TLC plates and using $CHCl_3$/MeOH (95:5) as eluent. ^{1}H NMR and ^{13}C NMR spectra were measured in $CDCl_3$ (TMS as internal standard) or DMSO-d_6 with a Varian Gemini 300 spectrometer; chemical shifts are expressed in δ (ppm) and coupling constants (J) in Hz. All exchangeable protons were confirmed by addition of deuterium oxide (D_2O). GC−MS spectra for selected compounds were recorded on a Shimadzu QP500 EI 151 mass spectrometer (see Supporting Information).

General Procedure for the Synthesis of 1-(Cyclo)alkyl-6,7-dimethoxy-1,2,3,4-tetrahydroisoquinolines (22−29). The new one-pot procedure for the synthesis of free amines **22−29** was carried out in microwave assisted conditions starting from commercially available 2-(3′,4′-dimethoxyphenyl)ethylamine (**13**) and the suitable aldehydes via imine intermediates **14−21** following a previously reported synthetic approach to obtain 1-aryl-6,7-dimethoxy-1,2,3,4-tetrahydroisoquinolines.[27,28] Analytical and spectroscopic data for the isoquinoline derivatives **22−29** are in accordance with literature results.[37−40]

General Procedure for the Synthesis of 3,4-Dihydroisoquinoline-2(1*H*)-sulfonamides (1−10). A mixture of the appropriate 1,2,3,4-tetrahydroisoquinoline (**11**) or 6,7-dimethoxy-1,2,3,4-tetrahydroisoquinoline (**12**, **22−29**) (1.0 mmol) and sulfamide (6 mmol, 576 mg) in dimethoxyethane (2 mL) was placed in a cylindrical quartz tube (Ø2 cm), then stirred and irradiated in a microwave oven at 150 W for two steps of 20 min at 90 °C. The reaction was quenched by adding water (5 mL) and extracted with ethyl acetate (3 × 5 mL). The organic layer was washed with an aqueous saturated solution of $NaHCO_3$ (2 × 5 mL), dried over Na_2SO_4, and concentrated until dryness under reduced pressure. The residue crystallized from diethyl ether to give the desired compounds **1−10**.

3,4-Dihydroisoquinoline-2(1*H*)-sulfonamide (1). Yield 43%; mp 161−163 °C. R_f = 0.415. ^{1}H NMR (DMSO-d_6) δ 2.89 (t, J = 5.77, 2H, CH_2), 3.24 (t, J = 5.77, 2H, CH_2), 4.18 (s, 2H, CH_2), 6.91 (bs, 2H, NH_2), 7.15 (s, 4H, ArH). Anal. ($C_9H_{12}N_2O_2S$) C, H, N.

6,7-Dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (2). Yield 51%; mp 164−166 °C. R_f = 0.405. ^{1}H NMR (DMSO-d_6) δ 2.79−2.81 (m, 2H, CH_2), 3.19−3.23 (m, 2H, CH_2), 3.70 (s, 6H, OCH_3), 4.09 (s, 2H, CH_2), 6.71 (s, 1H, ArH), 6.72 (s, 1H, ArH), 6.85 (bs, 2H, NH_2). Anal. ($C_{11}H_{16}N_2O_4S$) C, H, N.

6,7-Dimethoxy-1-methyl-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (3). Yield 35%; mp 172−174 °C. R_f = 0.465. ^{1}H NMR ($CDCl_3$) δ 1.54 (d, J = 6.87, 3H, CH_3), 2.64−2.70 (m, 1H, CH), 2.96−3.12 (m, 1H, CH), 3.35.3.45 (m, 1H, CH), 3.82−3.95 (m, 1H, CH), 3.85 (s, 3H, OCH_3), 3.86 (s, 3H, OCH_3), 4.35 (bs, 2H, NH_2), 4.95 (q, J = 6.87, 1H, CH), 6.56 (s, 1H, ArH), 6.59 (s, 1H, ArH). Anal. ($C_{12}H_{18}N_2O_4S$) C, H, N.

6,7-Dimethoxy-1-ethyl-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (4). Yield 33%; mp 201−203 °C. R_f = 0.470. ^{1}H NMR ($CDCl_3$) δ 1.07 (t, J = 7.42, 3H, CH_3), 1.75−1.85 (m, 2H, CH_2), 2.60−2.66 (m, 1H, CH), 3.04−3.16 (m, 1H, CH), 3.36−3.47 (m, 1H, CH), 3.78−3.99 (m, 1H, CH), 3.85 (s, 3H, OCH_3), 3.86 (s, 3H, OCH_3), 4.27 (bs, 2H, NH_2), 4.61−4.66 (m, 1H, CH), 6.57 (s, 1H, ArH), 6.59 (s, 1H, ArH). Anal. ($C_{13}H_{20}N_2O_4S$) C, H, N.

6,7-Dimethoxy-1-propyl-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (5). Yield 46%. Mp 185−187 °C. R_f = 0.520. ^{1}H NMR ($CDCl_3$) δ 0.98 (t, J = 7.42, 3H, CH_3), 1.48−1.81 (m, 4H, CH_2−CH_2), 2.60−2.66 (m, 1H, CH), 3.04−3.16 (m, 1H, CH), 3.38−3.48 (m, 1H, CH), 3.85 (s, 3H, OCH_3), 3.87 (s, 3H, OCH_3), 3.93−4.00 (m, 1H, CH), 4.27 (bs, 2H, NH_2), 4.70−4.74 (m, 1H, CH), 6.56 (s, 1H, ArH), 6.58 (s, 1H, ArH). Anal. ($C_{14}H_{22}N_2O_4S$) C, H, N.

6,7-Dimethoxy-1-isopropyl-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (6). Yield 44%; mp 187−189 °C. R_f = 0.505. ^{1}H NMR (DMSO-d_6) δ 0.86−0.90 (m, 6H, CH_3), 1.96−2.02 (m, 1H, CH), 2.57−2.62 (m, 1H, CH), 2.84−2.92 (m, 1H, CH), 3.33−3.39 (m, 1H, CH), 3.48−3.56 (m, 1H, CH), 3.70 (s, 6H, OCH_3), 4.27 (d, J = 7.14, 1H, CH), 6.57 (bs, 2H, NH_2), 6.68 (s, 1H, ArH), 6.70 (s, 1H, ArH). Anal. ($C_{14}H_{22}N_2O_4S$) C, H, N.

1-Cyclopropyl-6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (7). Yield 35%; mp 217−219 °C. R_f = 0.468. ^{1}H NMR (DMSO-d_6) δ 0.42−0.61 (m, 4H, CH_2), 1.13−1.17 (m, 1H, CH), 2.55−2.60 (m, 1H, CH), 2.88−2.99 (m, 1H, CH), 3.34−3.65 (m, 2H, CH), 3.99 (d, J = 7.14, 1H), 6.62 (bs, 2H, NH_2), 6.68 (s, 1H, ArH), 6.76 (s, 1H, ArH). Anal. ($C_{14}H_{20}N_2O_4S$) C, H, N.

1-*n*-Butyl-6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (8). Yield 60%; mp 171−173 °C. R_f = 0.533. ^{1}H NMR ($CDCl_3$) δ 0.93 (t, J = 7.42, 3H, CH_3), 1.25−1.81 (m, 6H, CH_2−CH_2−CH_2), 2.60−2.66 (m, 1H, CH), 3.07−3.16 (m, 1H, CH), 3.38−3.48 (m, 1H, CH), 3.85 (s, 3H, OCH_3), 3.86 (s, 3H, OCH_3), 3.93−4.00 (m, 1H, CH), 4.28 (bs, 2H, NH_2), 4.68−4.73 (m, 1H, CH), 6.56 (s, 1H, ArH), 6.58 (s, 1H, ArH). Anal. ($C_{15}H_{24}N_2O_4S$) C, H, N.

1-Cyclopentyl-6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (9). Yield 67%; mp 193−195 °C. R_f = 0.575. ^{1}H NMR ($CDCl_3$) δ 1.39−1.92 (m, 9H, CH), 2.13−2.16 (m, 1H, CH), 2.71−2.76 (m, 1H, CH), 3.07−3.09 (m, 1H, CH), 3.53−3.61 (m, 1H, CH), 3.86 (s, 6H, OCH_3), 3.91−3.98 (m, 1H, CH), 4.17 (bs, 2H, NH_2), 4.45 (d, J = 8.89, 1H), 6.60 (s, 1H, ArH), 6.63 (s, 1H, ArH). Anal. ($C_{16}H_{24}N_2O_4S$) C, H, N.

1-Cyclohexyl-6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (10). Yield 34%; mp 177−179 °C. R_f = 0.560. ^{1}H NMR ($CDCl_3$) δ 1.03−1.99 (m, 11H, CH), 2.73−2.80 (m, 1H, CH), 2.98−3.11 (m, 1H, CH), 3.48−3.61 (m, 1H, CH), 3.83−3.90 (m, 1H, CH), 3.85 (s, 3H, OCH_3), 3.86 (s, 3H, OCH_3), 4.09 (bs, 2H, NH_2), 4.37 (d, J = 8.79, 1H, CH), 6.57 (s, 1H, ArH), 6.62 (s, 1H, ArH). Anal. ($C_{17}H_{26}N_2O_4S$) C, H, N.

CA Inhibition Assay. An Applied Photophysics stopped-flow instrument has been used for assaying the CA catalyzed CO_2 hydration activity.[41] Phenol red (at a concentration of 0.2 mM) has been used as indicator, working at the absorbance maximum of 557 nm, with 10−20 mM Hepes (pH 7.5) or Tris (pH 8.3) as buffers and 20 mM Na_2SO_4 or 20 mM $NaClO_4$ (for maintaining constant the ionic strength), following the initial rates of the CA-catalyzed CO_2 hydration reaction for a period of 10−100 s. The CO_2 concentrations ranged from 1.7 to 17 mM for the determination of the kinetic parameters and inhibition constants. For each inhibitor at least six traces of the initial 5−10% of the reaction have been used for determining the initial velocity. The uncatalyzed rates were determined in the same manner and subtracted from the total observed rates. Stock solutions of inhibitor (10 mM) were prepared in distilled−deionized water, and dilutions up to 0.01 nM were done thereafter with distilled−deionized water. Inhibitor and enzyme solutions were preincubated together for 15 min at room temperature prior to assay in order to allow for the formation of the E−I complex. The inhibition constants were obtained by nonlinear least-squares methods using PRISM 3, as reported earlier, and represent the mean from at least three different determinations. CA isoforms were recombinant ones obtained as reported earlier by this group.[42−45]

Protein Crystallography. Protein Crystallization and X-ray Data Collection. The complex was prepared by adding 5-fold

molar excess of 6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (**2**) (in dimethyl sulfoxide) to 10 mg·mL^{-1} protein solution of hCA II (Sigma) in 100 mM Tris-Cl, pH 8.5. The best crystals of the complex were obtained by the hanging-drop vapor diffusion method, under the following conditions: an amount of 2 μL of complex solution was mixed with 2 μL of precipitant solution [2.5 M $(NH_4)_2SO_4$, 0.3 M NaCl, 100 mM Tris-Cl, pH 8.2] and equilibrated over a reservoir containing 1 mL of precipitant solution at 18 °C. Crystals with dimensions 0.4 mm × 0.2 mm × 0.1 mm grew within 10 days.

For data collection, the crystals were soaked in the reservoir solution supplemented with 20% (v/v) glycerol and transferred to liquid nitrogen. Diffraction data were collected at 120 K using an in-house diffractometer (Nonius FR 591) connected to 345 mm MarResearch image plate detector. The best crystal diffracted up to 1.65 Å resolution, and diffraction data were integrated and reduced using MOSFLM[46] and scaled using SCALA[47] from the CCP4 suite of programs.[48] Crystal parameters and data collection statistics are summarized in Table 2.

Structure Determination, Refinement, and Analysis. The structure of hCA II in complex with 6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamide (**2**) was solved using the difference Fourier method, using hCA II structure (Protein Data Bank entry 1H9N[49]) as the initial model. Initial rigid-body refinement and subsequent restrained refinement were performed using the program REFMAC 5.[48] The structure was refined with two inhibitor molecules, one in the enzyme active site and the other located in the surface pocket in the vicinity of the enzyme N-terminus. Atomic coordinates and geometry library for the inhibitor were generated using the PRODRG server.[50] The Coot program[51] was used for inhibitor fitting, model rebuilding, and addition of water molecules. In final refinement stages, TLS refinement cycles in the program REFMAC 5 were introduced.[52] The quality of the crystallographic model was assessed with MolProbity.[53] Program CONTACT/ACT from the CCP4 suite[48] was used for finding contacts between inhibitor and protein molecules. The final refinement statistics are summarized in Table 2. All figures showing structural representations were prepared using PyMOL,[54] and the APBS[55] tools plugin was used for generating solvent accessible surface colored by electrostatic potential.

Docking Studies. The crystal structures of hCA II, hCA IX, and hCA XIV in complex with the inhibitor acetazolamide were retrieved from the RCSB Protein Data Bank (entry code 1YDB,[34] 3IAI,[35] and 1RJ6[36]). The complex 1YDB presents the mutant Phe198; thus, for our docking studies, it was converted in Leu198, placing its side chain in the same position observed in our X-ray crystal structure of hCAII. Hydrogen atoms were added to proteins by the Biopolymer module in SYBYL 8.0.1.[56] The acetazolamide structure was extracted from an X-ray complex, and the other structures of the ligands were constructed using standard bond lengths and angles from the SYBYL 8.0 fragment library. All inhibitors were fully optimized by the semiempirical quantum mechanical method AM1. The ligands minimized in this way were docked in their corresponding proteins by means of Gold 3.1.1.[33] The region of interest used by Gold was defined in order to contain the residues within 10 Å from the original position of the ligand in the X-ray structures; the zinc ion was set as possessing a trigonal-bipyramidal coordination. The "allow early termination" command was deactivated, while the possibility for the ligand to flip ring corners was activated. ChemScore[33] was chosen as the fitness function, and the formation of a H bond between the hydroxyl group of Thr199 and the ligands was also imposed.[57] As regards to all the other parameters, the Gold default ones were used, and the ligands were submitted to 100 genetic algorithm runs.

Acknowledgment. Financial support for this research by Fondo di Ateneo per la Ricerca (PRA 2006, Università di Messina) is gratefully acknowledged. Research from CTS lab was financed by a grant of the sixth Framework Programme (FP) of the European Union (Project DeZnIT) and by a seventh FP grant (Project METOXIA). In part, this work was supported by Projects AV0Z50520514 and AV0Z40550506 awarded by the Academy of Sciences of the Czech Republic, grant No. 1M0505 awarded by the Ministry of Education of the Czech Republic, and grant No. GA203/09/0820 awarded by Czech Grant Foundation. P. Mader is a PhD. student registered at the First Faculty of Medicine, Charles University in Prague.

Supporting Information Available: Other spectroscopic and crystal data; table of contacts between inhibitor and protein. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Thiry, A.; Dogne, J. M.; Supuran, C. T.; Masereel, B. Carbonic anhydrase inhibitors as anticonvulsant agents. *Curr. Top. Med. Chem.* **2007**, *7*, 855–864.
(2) Supuran, C. T.; Di Fiore, A.; De Simone, G. Carbonic anhydrase inhibitors as emerging drugs for the treatment of obesity. *Expert Opin. Emerging Drugs* **2008**, *13*, 383–392.
(3) Pastorekova, S.; Kopacek, J.; Pastorek, J. Carbonic anhydrase inhibitors and the management of cancer. *Curr. Top. Med. Chem.* **2007**, *7*, 865–878.
(4) Supuran, C. T.; Scozzafava, A.; Casini, A. Carbonic anhydrase inhibitors. *Med. Res. Rev.* **2003**, *23*, 146–189.
(5) Supuran, C. T. Carbonic anhydrases: novel therapeutic applications for inhibitors and activators. *Nat. Rev. Drug Discovery* **2008**, *7*, 168–181.
(6) Winum, J. Y.; Poulsen, S. A.; Supuran, C. T. Therapeutic applications of glycosidic carbonic anhydrase inhibitors. *Med. Res. Rev.* **2009**, *29*, 419–435.
(7) Guzel, O.; Temperini, C.; Innocenti, A.; Scozzafava, A.; Salman, A.; Supuran, C. T. Carbonic anhydrase inhibitors. Interaction of 2-(hydrazinocarbonyl)-3-phenyl-1*H*-indole-5-sulfonamide with 12 mammalian isoforms: kinetic and X-ray crystallographic studies. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 152–158.
(8) Thiry, A.; Supuran, C. T.; Masereel, B.; Dogne, J. M. Recent developments of carbonic anhydrase inhibitors as potential anticancer drugs. *J. Med. Chem.* **2008**, *51*, 3051–3056.
(9) Chiche, J.; Ilc, K.; Laferriere, J.; Trottier, E.; Dayan, F.; Mazure, N. M.; Brahimi-Horn, M. C.; Pouyssegur, J. Hypoxia-inducible carbonic anhydrase IX and XII promote tumor cell growth by counteracting acidosis through the regulation of the intracellular pH. *Cancer Res.* **2009**, *69*, 358–368.
(10) Parkkila, S.; Parkkila, A. K.; Rajaniemi, H.; Shah, G. N.; Grubb, J. H.; Waheed, A.; Sly, W. S. Expression of membrane-associated carbonic anhydrase XIV on neurons and axons in mouse and human brain. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 1918–1923.
(11) Winum, J. Y.; Scozzafava, A.; Montero, J. L.; Supuran, C. T. Sulfamates and their therapeutic potential. *Med. Res. Rev.* **2005**, *25*, 186–228.
(12) Winum, J. Y.; Scozzafava, A.; Montero, J. L.; Supuran, C. T. Therapeutic potential of sulfamides as enzyme inhibitors. *Med. Res. Rev.* **2006**, *26*, 767–792.
(13) Supuran, C. T.; Casini, A.; Scozzafava, A. Protease inhibitors of the sulfonamide type: anticancer, antiinflammatory, and antiviral agents. *Med. Res. Rev.* **2003**, *23*, 535–558.
(14) Winum, J. Y.; Scozzafava, A.; Montero, J. L.; Supuran, C. T. New zinc binding motifs in the design of selective carbonic anhydrase inhibitors. *Mini-Rev. Med. Chem.* **2006**, *6*, 921–936.
(15) Supuran, C. T. Diuretics: from classical carbonic anhydrase inhibitors to novel applications of the sulfonamides. *Curr. Pharm. Des.* **2008**, *14*, 641–648.
(16) Nasr, G.; Petit, E.; Vullo, D.; Winum, J. Y.; Supuran, C. T.; Barboiu, M. Carbonic anhydrase-encoded dynamic constitutional libraries: toward the discovery of isozyme-specific inhibitors. *J. Med. Chem.* **2009**, *52*, 853–859.
(17) D'Ambrosio, K.; Masereel, B.; Thiry, A.; Scozzafava, A.; Supuran, C. T.; De Simone, G. Carbonic anhydrase inhibitors: binding of indanesulfonamides to the human isoform II. *ChemMedChem* **2008**, *3*, 473–477.
(18) Winum, J. Y.; Scozzafava, A.; Montero, J. L.; Supuran, C. T. Metal binding functions in the design of carbonic anhydrase inhibitors. *Curr. Top. Med. Chem.* **2007**, *7*, 835–848.

(19) Temperini, C.; Cecchi, A.; Scozzafava, A.; Supuran, C. T. Carbonic anhydrase inhibitors. Comparison of chlorthalidone and indapamide X-ray crystal structures in adducts with isozyme II: when three water molecules and the keto-enol tautomerism make the difference. *J. Med. Chem.* **2009**, *52*, 322–328.

(20) Abbate, F.; Casini, A.; Scozzafava, A.; Supuran, C. T. Carbonic anhydrase inhibitors: X-ray crystallographic structure of the adduct of human isozyme II with a topically acting antiglaucoma sulfonamide. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2357–2361.

(21) Casini, A.; Antel, J.; Abbate, F.; Scozzafava, A.; David, S.; Waldeck, H.; Schafer, S.; Supuran, C. T. Carbonic anhydrase inhibitors: SAR and X-ray crystallographic study for the interaction of sugar sulfamates/sulfamides with isozymes I, II and IV. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 841–845.

(22) Di Fiore, A.; Monti, S. M.; Hilvo, M.; Parkkila, S.; Romano, V.; Scaloni, A.; Pedone, C.; Scozzafava, A.; Supuran, C. T.; De Simone, G. Crystal structure of human carbonic anhydrase XIII and its complex with the inhibitor acetazolamide. *Proteins* **2009**, *74*, 164–175.

(23) Temperini, C.; Cecchi, A.; Scozzafava, A.; Supuran, C. T. Carbonic anhydrase inhibitors. Interaction of indapamide and related diuretics with 12 mammalian isozymes and X-ray crystallographic studies for the indapamide−isozyme II adduct. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2567–2573.

(24) Winum, J. Y.; Temperini, C.; El Cheikh, K.; Innocenti, A.; Vullo, D.; Ciattini, S.; Montero, J. L.; Scozzafava, A.; Supuran, C. T. Carbonic anhydrase inhibitors: clash with Ala65 as a means for designing inhibitors with low affinity for the ubiquitous isozyme II, exemplified by the crystal structure of the topiramate sulfamide analogue. *J. Med. Chem.* **2006**, *49*, 7024–7031.

(25) Alterio, V.; Vitale, R. M.; Monti, S. M.; Pedone, C.; Scozzafava, A.; Cecchi, A.; De Simone, G.; Supuran, C. T. Carbonic anhydrase inhibitors: X-ray and molecular modeling study for the interaction of a fluorescent antitumor sulfonamide with isozyme II and IX. *J. Am. Chem. Soc.* **2006**, *128*, 8329–8335.

(26) Menchise, V.; De Simone, G.; Alterio, V.; Di Fiore, A.; Pedone, C.; Scozzafava, A.; Supuran, C. T. 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. *J. Med. Chem.* **2005**, *48*, 5721–5727.

(27) Gitto, R.; Ferro, S.; Agnello, S.; De Luca, L.; De Sarro, G.; Russo, E.; Vullo, D.; Supuran, C. T.; Chimirri, A. Synthesis and evaluation of pharmacological profile of 1-aryl-6,7-dimethoxy-3,4-dihydroisoquinoline-2(1*H*)-sulfonamides. *Bioorg. Med. Chem.* **2009**, *17*, 3659–3664.

(28) Gitto, R.; Ficarra, R.; Stancanelli, R.; Guardo, M.; De Luca, L.; Barreca, M. L.; Pagano, B.; Rotondo, A.; Bruno, G.; Russo, E.; De Sarro, G.; Chimirri, A. Synthesis, resolution, stereochemistry, and molecular modeling of (*R*)- and (*S*)-2-acetyl-1-(4′-chlorophenyl)-6,7-dimethoxy-1,2,3,4-tetrahydroisoquinol ine AMPAR antagonists. *Bioorg. Med. Chem.* **2007**, *15*, 5417–5423.

(29) Srivastava, D. K.; Jude, K. M.; Banerjee, A. L.; Haldar, M.; Manokaran, S.; Kooren, J.; Mallik, S.; Christianson, D. W. Structural analysis of charge discrimination in the binding of inhibitors to human carbonic anhydrases I and II. *J. Am. Chem. Soc.* **2007**, *129*, 5528–5537.

(30) Eriksson, A. E.; Jones, T. A.; Liljas, A. Refined structure of human carbonic anhydrase II at 2.0 Å resolution. *Proteins* **1988**, *4*, 274–282.

(31) Jude, K. M.; Banerjee, A. L.; Haldar, M. K.; Manokaran, S.; Roy, B.; Mallik, S.; Srivastava, D. K.; Christianson, D. W. Ultrahigh resolution crystal structures of human carbonic anhydrases I and II complexed with "two-prong" inhibitors reveal the molecular basis of high affinity. *J. Am. Chem. Soc.* **2006**, *128*, 3011–3018.

(32) Betts, M. J.; Sternberg, M. J. An analysis of conformational changes on protein−protein association: implications for predictive docking. *Protein Eng.* **1999**, *12*, 271–283.

(33) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.

(34) Nair, S. K.; Krebs, J. F.; Christianson, D. W.; Fierke, C. A. Structural basis of inhibitor affinity to variants of human carbonic anhydrase II. *Biochemistry* **1995**, *34*, 3981–3989.

(35) Alterio, V.; Hilvo, M.; Di Fiore, A.; Supuran, C. T.; Pan, P.; Parkkila, S.; Scaloni, A.; Pastorek, J.; Pastorekova, S.; Pedone, C.; Scozzafava, A.; Monti, S. M.; De Simone, G. Crystal structure of the catalytic domain of the tumor-associated human carbonic anhydrase IX. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 16233–16238.

(36) Whittington, D. A.; Grubb, J. H.; Waheed, A.; Shah, G. N.; Sly, W. S.; Christianson, D. W. Expression, assay, and structure of the extracellular domain of murine carbonic anhydrase XIV: implications for selective inhibition of membrane-associated isozymes. *J. Biol. Chem.* **2004**, *279*, 7223–7228.

(37) Craig, P. N.; Nabenhauer, F. P.; Williams, P. M.; Macko, E.; Toner, J. Tetrahydroisoquinolines. I. 1-Alkyl-6,7-dihydroxy-1,2,3,4-tetrahydroisoquinolines. *J. Am. Chem. Soc.* **1952**, *74*, 1316–1317.

(38) Barbier, A. M.; Rumpf, P. Compounds related to emetine. I. Synthesis of 1-(1,2,3,4-tetrahydro-6,7-dimethoxy-1-propyl-2-isoquinolyl)-4-(1,2,3,4-tetrahydro-6,7-dimethoxy-1-isoquinolyl)butane acid oxalate. *Bull. Soc. Chim. Fr.* **1953**, 293–296.

(39) Leseche, B.; Gilbert, J.; Viel, C. Study of the oxidation of 6,7- and 7,8-dimethoxytetrahydroisoquinolin-4-ols and the reactivity of the carbonyl group of 6,7-dimethoxytetrahydroisoquinol-4-one. *J. Heterocycl. Chem.* **1981**, *18*, 143–153.

(40) Hegedues, A.; Hell, Z. One-step preparation of 1-substituted tetrahydroisoquinolines via the Pictet−Spengler reaction using zeolite catalysts. *Tetrahedron Lett.* **2004**, *45*, 8553–8555.

(41) Khalifah, R. G. The carbon dioxide hydration activity of carbonic anhydrase. I. Stop-flow kinetic studies on the native human isoenzymes B and C. *J. Biol. Chem.* **1971**, *246*, 2561–2573.

(42) Nishimori, I.; Vullo, D.; Innocenti, A.; Scozzafava, A.; Mastrolorenzo, A.; Supuran, C. T. Carbonic anhydrase inhibitors. The mitochondrial isozyme VB as a new target for sulfonamide and sulfamate inhibitors. *J. Med. Chem.* **2005**, *48*, 7860–7866.

(43) Nishimori, I.; Vullo, D.; Innocenti, A.; Scozzafava, A.; Mastrolorenzo, A.; Supuran, C. T. Carbonic anhydrase inhibitors: inhibition of the transmembrane isozyme XIV with sulfonamides. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3828–3833.

(44) Nishimori, I.; Minakuchi, T.; Onishi, S.; Vullo, D.; Scozzafava, A.; Supuran, C. T. Carbonic anhydrase inhibitors. DNA cloning, characterization, and inhibition studies of the human secretory isoform VI, a new target for sulfonamide and sulfamate inhibitors. *J. Med. Chem.* **2007**, *50*, 381–388.

(45) Nishimori, I.; Minakuchi, T.; Morimoto, K.; Sano, S.; Onishi, S.; Takeuchi, H.; Vullo, D.; Scozzafava, A.; Supuran, C. T. Carbonic anhydrase inhibitors: DNA cloning and inhibition studies of the alpha-carbonic anhydrase from *Helicobacter pylori*, a new target for developing sulfonamide and sulfamate gastric drugs. *J. Med. Chem.* **2006**, *49*, 2117–2126.

(46) Leslie, A. G. W. Integration of macromolecular diffraction data. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*, 1696–1702.

(47) Evans, P. R. In *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*; Daresbury Laboratory: Warrington, U.K., 1993.

(48) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1994**, *50*, 760−763.

(49) Lesburg, C. A.; Huang, C.; Christianson, D. W.; Fierke, C. A. Histidine → carboxamide ligand substitutions in the zinc binding site of carbonic anhydrase II alter metal coordination geometry but retain catalytic activity. *Biochemistry* **1997**, *36*, 15780–15791.

(50) Schuttelkopf, A. W.; van Aalten, D. M. PRODRG: a tool for high-throughput crystallography of protein−ligand complexes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 1355–1363.

(51) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126–2132.

(52) Winn, M. D.; Isupov, M. N.; Murshudov, G. N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2001**, *57*, 122–133.

(53) Lovell, S. C.; Davis, I. W.; Arendall, W. B., 3rd; de Bakker, P. I.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **2003**, *50*, 437–450.

(54) De Lano, W. L. *The PyMOL Molecular Graphics System*; DeLano Scientific: Palo Alto, CA, 2002.

(55) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad Sci. U.S.A.* **2001**, *98*, 10037–10041.

(56) *SYBYL*, version 8.0; Tripos Inc. (1699 South Hanley Road, St. Louis, MO 63144), 2005.

(57) Tuccinardi, T.; Nuti, E.; Ortore, G.; Supuran, C. T.; Rossello, A.; Martinelli, A. Analysis of human carbonic anhydrase II: docking reliability and receptor-based 3D-QSAR study. *J. Chem. Inf. Model.* **2007**, *47*, 515–525.

(58) Brunger, A. T. Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **1992**, *355*, 472–475.

9 REFERENCES

Alber, B. E., Colangelo, C. M., Dong, J., Stalhandske, C. M., Baird, T. T., Tu, C., Fierke, C. A., Silverman, D. N., Scott, R. A. & Ferry, J. G. (1999). *Biochemistry* **38**, 13119-13128.

Alber, B. E. & Ferry, J. G. (1994). *Proc Natl Acad Sci U S A* **91**, 6909-6913.

Alexander, R. S., Nair, S. K. & Christianson, D. W. (1991). *Biochemistry* **30**, 11064-11072.

Alterio, V., Hilvo, M., Di Fiore, A., Supuran, C. T., Pan, P., Parkkila, S., Scaloni, A., Pastorek, J., Pastorekova, S., Pedone, C., Scozzafava, A., Monti, S. M. & De Simone, G. (2009). *Proc Natl Acad Sci U S A* **106**, 16233-16238.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc Natl Acad Sci U S A* **98**, 10037-10041.

Banerjee, A. L., Swanson, M., Mallik, S. & Srivastava, D. K. (2004). *Protein Expr Purif* **37**, 450-454.

Beckett, D., Kovaleva, E. & Schatz, P. J. (1999). *Protein Sci* **8**, 921-929.

Betts, M. J. & Sternberg, M. J. (1999). *Protein Eng* **12**, 271-283.

Bradfield, J. R. (1947). *Nature* **159**, 467.

Brahimi-Horn, M. C. & Pouyssegur, J. (2009). *J Cell Sci* **122**, 1055-1057.

Brion, L. P., Schwartz, J. H., Zavilowitz, B. J. & Schwartz, G. J. (1988). *Anal Biochem* **175**, 289-297.

Brunger, A. T. (1992). *Nature* **355**, 472-475.

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Crystallogr D Biol Crystallogr* **54**, 905-921.

The CCP4 suite: programs for protein crystallography (1994). **50**, 760-763.

de Boer, E., Rodriguez, P., Bonte, E., Krijgsveld, J., Katsantoni, E., Heck, A., Grosveld, F. & Strouboulis, J. (2003). *Proc Natl Acad Sci U S A* **100**, 7480-7485.

DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.*

Ducruix, A. & Giegé, R. (2000). *Crystallization of Nucleic Acids and Proteins: A Practical Approach*. Oxford University Press, USA.

Duffy, S., Tsao, K. L. & Waugh, D. S. (1998). *Anal Biochem* **262**, 122-128.

Emsley, P. & Cowtan, K. (2004). *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132.

Eriksson, A. E., Jones, T. A. & Liljas, A. (1988). *Proteins* **4**, 274-282.

Evans, P. R. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing.*

Falkbring, S. O., Gothe, P. O., Nyman, P. O., Sundberg, L. & Porath, J. (1972). *FEBS Lett* **24**, 229-235.

Ferry, J. G. (2010). *Biochim Biophys Acta* **1804**, 374-381.

Gitto, R., Agnello, S., Ferro, S., De Luca, L., Vullo, D., Brynda, J., Mader, P., Supuran, C. T. & Chimirri, A. (2010). *J Med Chem* **53**, 2401-2408.

Herron, J. N., He, X. M., Ballard, D. W., Blier, P. R., Pace, P. E., Bothwell, A. L., Voss, E. W., Jr. & Edmundson, A. B. (1991). *Proteins* **11**, 159-175.

Hewett-Emmett, D. & Tashian, R. E. (1996). *Mol Phylogenet Evol* **5**, 50-77.

Hilvo, M., Baranauskiene, L., Salzano, A. M., Scaloni, A., Matulis, D., Innocenti, A., Scozzafava, A., Monti, S. M., Di Fiore, A., De Simone, G., Lindfors, M., Janis, J., Valjakka, J., Pastorekova, S., Pastorek, J., Kulomaa, M. S., Nordlund, H. R., Supuran, C. T. & Parkkila, S. (2008). *J Biol Chem* **283**, 27799-27809.

Hilvo, M., Supuran, C. T. & Parkkila, S. (2007). *Current Topics in Medicinal Chemistry* **7**, 893-899.

Chegwidden, W. R., Dogson, S. J. & Spencer, I. M. (2000). *The Carbonic Anhydrases: New Horizons*, edited by W. R. Chegwidden, N. D. Carter & Y. Edwards, pp. 343-363: Birkhauser Verlag, Basel.

Chiche, J., Ilc, K., Laferriere, J., Trottier, E., Dayan, F., Mazure, N. M., Brahimi-Horn, M. C. & Pouyssegur, J. (2009). *Cancer Res* **69**, 358-368.

Jancarik, J. & Kim, S. H. (1991). *Journal of Applied Crystallography* **24**, 409-411.

Jiang, W. & Gupta, D. (1999). *Biochem J* **344 Pt 2**, 385-390.

Jude, K. M., Banerjee, A. L., Haldar, M. K., Manokaran, S., Roy, B., Mallik, S., Srivastava, D. K. & Christianson, D. W. (2006). *J Am Chem Soc* **128**, 3011-3018.

Kimber, M. S. & Pai, E. F. (2000). *EMBO J* **19**, 1407-1418.

Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G. & Rees, D. C. (1996). *EMBO J* **15**, 2323-2330.

Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Crystallogr D Biol Crystallogr* **55**, 484-491.

Krishnamurthy, V. M., Kaufman, G. K., Urbach, A. R., Gitlin, I., Gudiksen, K. L., Weibel, D. B. & Whitesides, G. M. (2008). *Chem Rev* **108**, 946-1051.

Lane, T. W., Saito, M. A., George, G. N., Pickering, I. J., Prince, R. C. & Morel, F. M. (2005). *Nature* **435**, 42.

Lesburg, C. A., Huang, C., Christianson, D. W. & Fierke, C. A. (1997). *Biochemistry* **36**, 15780-15791.

Leslie, A. G. (1999). *Acta Crystallogr D Biol Crystallogr* **55**, 1696-1702.

Liao, S. Y., Brewer, C., Závada, J., Pastorek, J., Pastorekova, S., Manetta, A., Berman, M. L., DiSaia, P. J. & Stanbridge, E. J. (1994). *The American Journal of Pathology* **145**, 598-609.

Liljas, A., Kannan, K. K., Bergsten, P. C., Waara, I., Fridborg, K., Strandberg, B., Carlbom, U., Jarup, L., Lovgren, S. & Petef, M. (1972). *Nat New Biol* **235**, 131-137.

Lindskog, S. & Silverman, D. N. (2000). *Carbonic Anhydrases: New Horizons*, edited by W. R. Chegwidden, N. D. Carter & Y. H. Edwards. Basel, Switzerland: Birkhäuser Verlag.

Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). *Proteins* **50**, 437-450.

Maren, T. H. & Ellison, A. C. (1967). *Mol Pharmacol* **3**, 503-508.

Martinez-Zaguilan, R., Seftor, E. A., Seftor, R. E., Chu, Y. W., Gillies, R. J. & Hendrix, M. J. (1996). *Clin Exp Metastasis* **14**, 176-186.

Meldrum, N. U. & Roughton, F. J. (1933a). *J Physiol* **80**, 113-142.

Meldrum, N. U. & Roughton, F. J. (1933b). *J Physiol* **80**, 143-170.

Milner-White, E., Ross, B. M., Ismail, R., Belhadj-Mostefa, K. & Poet, R. (1988). *J Mol Biol* **204**, 777-782.

Mincione, F., Scozzafava, A. & Supuran, C. T. (2007). *Curr Top Med Chem* **7**, 849-854.

Mitsuhashi, S., Mizushima, T., Yamashita, E., Yamamoto, M., Kumasaka, T., Moriyama, H., Ueki, T., Miyachi, S. & Tsukihara, T. (2000). *J Biol Chem* **275**, 5521-5526.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Crystallographica Section D Biological Crystallography* **53**, 240-255.

Neish, A. C. (1939). *Biochem J* **33**, 300-308.

Opavsky, R., Pastorekova, S., Zelnik, V., Gibadulinova, A., Stanbridge, E. J., Zavada, J., Kettmann, R. & Pastorek, J. (1996). *Genomics* **33**, 480-487.

Park, H., Song, B. & Morel, F. M. (2007). *Environ Microbiol* **9**, 403-413.

Pastorekova, S. (2009). *Proceedings of the 8th International Conference on the Carbonic Anhydrases*, p. 18. Firenze, Italy

Pastorekova, S., Kopacek, J. & Pastorek, J. (2007). *Curr Top Med Chem* **7**, 865-878.

Pastorekova, S., Parkkila, S., Pastorek, J. & Supuran, C. T. (2004). *J Enzyme Inhib Med Chem* **19**, 199-229.

Pastorekova, S. & Zavada, J. (2004). *Cancer Therapy* **2**, 245-262.

Pastorekova, S., Zavadova, Z., Kostal, M., Babusikova, O. & Zavada, J. (1992). *Virology* **187**, 620-626.

Poulsen, S. A. (2010). *Expert Opin Ther Pat* **20**, 795-806.

Ramachandran, G. N. & Sasisekharan, V. (1968). *Adv Protein Chem* **23**, 283-438.

Read, R. J. (1986). *Acta Crystallogr. Sect. A* **42**, 140-149.

Rini, J. M., Schulze-Gahmen, U. & Wilson, I. A. (1992). *Science* **255**, 959-965.

Rowlett, R. S. (2010). *Biochim Biophys Acta* **1804**, 362-373.

Saarnio, J., Parkkila, S., Parkkila, A. K., Waheed, A., Casey, M. C., Zhou, X. Y., Pastorekova, S., Pastorek, J., Karttunen, T., Haukipuro, K., Kairaluoma, M. I. & Sly, W. S. (1998). *J Histochem Cytochem* **46**, 497-504.

Sambrook, J. & Russell, D. W. (2001). *Molecular cloning: a laboratory manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Schneider, I. (1972). *J Embryol Exp Morphol* **27**, 353-365.

Schuttelkopf, A. W. & van Aalten, D. M. (2004). *Acta Crystallogr D Biol Crystallogr* **60**, 1355-1363.

Silverman, D. N. & Lindskog, S. (1988). *Accounts of Chemical Research* **21**, 30-36.

Sly, W. S. & Hu, P. Y. (1995). *Annu Rev Biochem* **64**, 375-401.

Smith, K. S. & Ferry, J. G. (2000). *FEMS Microbiol Rev* **24**, 335-366.

Smith, K. S., Jakubzick, C., Whittam, T. S. & Ferry, J. G. (1999). *Proc Natl Acad Sci U S A* **96**, 15184-15189.

Srivastava, D. K., Jude, K. M., Banerjee, A. L., Haldar, M., Manokaran, S., Kooren, J., Mallik, S. & Christianson, D. W. (2007). *J Am Chem Soc* **129**, 5528-5537.

Stemler, A. (1993). *Anal Biochem* **210**, 328-331.

Strop, P., Smith, K. S., Iverson, T. M., Ferry, J. G. & Rees, D. C. (2001). *J Biol Chem* **276**, 10299-10305.

Stubbs, M., McSheehy, P. M., Griffiths, J. R. & Bashford, C. L. (2000). *Mol Med Today* **6**, 15-19.

Studier, F. W. & Moffatt, B. A. (1986). *J Mol Biol* **189**, 113-130.

Supuran, C. T. (2007). *Curr Top Med Chem* **7**, 825-833.

Supuran, C. T. (2008). *Nat Rev Drug Discov* **7**, 168-181.

Supuran, C. T., Di Fiore, A. & De Simone, G. (2008). *Expert Opin Emerg Drugs* **13**, 383-392.

Supuran, C. T., Scozzafava, A. & Casini, A. (2003). *Med Res Rev* **23**, 146-189.

Swietach, P., Patiar, S., Supuran, C. T., Harris, A. L. & Vaughan-Jones, R. D. (2009). *J Biol Chem* **284**, 20299-20310.

Swietach, P., Vaughan-Jones, R. D. & Harris, A. L. (2007). *Cancer Metastasis Rev* **26**, 299-310.

Swietach, P., Wigfield, S., Cobden, P., Supuran, C. T., Harris, A. L. & Vaughan-Jones, R. D. (2008). *J Biol Chem* **283**, 20473-20483.

Thaller, C., Weaver, L. H., Eichele, G., Wilson, E., Karlsson, R. & Jansonius, J. N. (1981). *Journal of Molecular Biology* **147**, 465-469.

Thiry, A., Dogne, J. M., Supuran, C. T. & Masereel, B. (2007). *Curr Top Med Chem* **7**, 855-864.

Thiry, A., Supuran, C. T., Masereel, B. & Dogne, J. M. (2008). *J Med Chem* **51**, 3051-3056.

Tripp, B. C., Smith, K. & Ferry, J. G. (2001). *J Biol Chem* **276**, 48615-48618.

Tu, C. K., Silverman, D. N., Forsman, C., Jonsson, B. H. & Lindskog, S. (1989). *Biochemistry* **28**, 7913-7918.

Tunuguntla, H. S. & Jorda, M. (2008). *J Urol* **179**, 2096-2102.

Tykvart, J. (2009). Diploma Thesis thesis, Charles University in Prague, Prague.

Vagin, A. & Teplyakov, A. (2000). *Acta Crystallographica Section D Biological Crystallography* **56**, 1622-1624.

Weber, A., Casini, A., Heine, A., Kuhn, D., Supuran, C. T., Scozzafava, A. & Klebe, G. (2004). *J Med Chem* **47**, 550-557.

Whittington, D. A., Waheed, A., Ulmasov, B., Shah, G. N., Grubb, J. H., Sly, W. S. & Christianson, D. W. (2001). *Proc Natl Acad Sci U S A* **98**, 9545-9550.

Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Crystallogr D Biol Crystallogr* **57**, 122-133.

Winum, J. Y., Rami, M., Scozzafava, A., Montero, J. L. & Supuran, C. (2008). *Med Res Rev* **28**, 445-463.

Winum, J. Y., Scozzafava, A., Montero, J. L. & Supuran, C. T. (2005). *Med Res Rev* **25**, 186-228.

Wykoff, C. C., Beasley, N. J., Watson, P. H., Turner, K. J., Pastorek, J., Sibtain, A., Wilson, G. D., Turley, H., Talks, K. L., Maxwell, P. H., Pugh, C. W., Ratcliffe, P. J. & Harris, A. L. (2000). *Cancer Res* **60**, 7075-7083.

Yang, J., Jaramillo, A., Shi, R., Kwok, W. W. & Mohanakumar, T. (2004). *Hum Immunol* **65**, 692-699.

Zavada, J., Zavadova, Z., Pastorek, J., Biesova, Z., Jezek, J. & Velek, J. (2000). *Br J Cancer* **82**, 1808-1813.

Zavada, J., Zavadova, Z., Pastorekova, S., Ciampor, F., Pastorek, J. & Zelnik, V. (1993). *Int J Cancer* **54**, 268-274.

Zavadova, Z. & Zavada, J. (2005). *Oncol Rep* **13**, 977-982.

Univerzita Karlova v Praze, 1. lékařská fakulta
Kateřinská 32, Praha 2

Prohlášení zájemce o nahlédnutí
do závěrečné práce absolventa studijního programu
uskutečňovaného na 1. lékařské fakultě Univerzity Karlovy v Praze

Jsem si vědom/a, že závěrečná práce je autorským dílem a že informace získané nahlédnutím do zpřístupněné závěrečné práce nemohou být použity k výdělečným účelům, ani nemohou být vydávány za studijní, vědeckou nebo jinou tvůrčí činnost jiné osoby než autora.

Byl/a jsem seznámen/a se skutečností, že si mohu pořizovat výpisy, opisy nebo kopie závěrečné práce, jsem však povinen/a s nimi nakládat jako s autorským dílem a zachovávat pravidla uvedená v předchozím odstavci.

Příjmení, jméno (hůlkovým písmem)	Číslo dokladu totožnosti vypůjčitele (např. OP, cestovní pas)	Signatura závěrečné práce	Datum	Podpis