Vilnius University

Alexey Smirnov

CRYSTALLOGRAPHIC STUDIES OF CARBONIC ANHYDRASE ISOFORMS AND THEIR COMPLEXES WITH INHIBITORS

Summary of doctoral dissertation
Technological Sciences, Chemical Engineering (05 T)

Vilnius, 2018

The work described in this dissertation was performed from October 2013 to September 2017 at the Institute of Biotechnology, Vilnius University, Lithuania.

Supervisor:
Prof. dr. Daumantas Matulis (Vilnius university, technological sciences, chemical engineering – 05 T).

Consultant:
Dr. Elena Manakova (Vilnius university, physical sciences, chemistry – 03 P).

Dissertation will be defended at the public hearing before the defense committee.

Chairman:
Prof. habil. dr. Eugenijus Butkus (Vilnius University, technological sciences, chemical engineering – 05 T).

Members:
Prof. dr. Saulius Bagdonas (Vilnius University, physical sciences, physics – 02 P),

Dr. Rima Budvytytė (Vilnius University, technological sciences, chemical engineering – 05 T),

Prof. dr. Robert McKenna (University of Florida, physical sciences, physics – 02 P),

Prof. dr. Edita Sužiedėlienė (Vilnius University, physical sciences, biochemistry – 04 P).

The thesis defence will take place at the public defense hearing at Life Science Center, Vilnius University (auditorium R108, Saulėtekio 7, LT-10257, Vilnius, Lithuania) at 1 p.m. on the 27 of September, 2018. The thesis is available at the Library of Vilnius University.

The summary of doctoral thesis was sent on the 15 of August, 2018.
The thesis is available at the Library of Vilnius University and on the website:
https://www.vu.lt/naujienos/ivykiu-kalendorius.

Vilniaus universitetas

Alexey Smirnov

ŽMOGAUS KARBOANHIDRAZIŲ IZOFORMŲ IR JŲ SĄVEIKŲ SU SLOPIKLIAIS KRISTALOGRAFINIAI TYRIMAI

Daktaro disertacijos santrauka

Technologijos mokslai, chemijos inžinerija (05 T)

Vilnius, 2018

Disertacija rengta Vilniaus universiteto Biotechnologijos institute 2013-2017 metais.

Mokslinis vadovas:
Prof. dr. Daumantas Matulis (Vilniaus universitetas, technologijos mokslai, chemijos inžinerija – 05 T).

Mokslinis konsultantas:
Dr. Elena Manakova (Vilniaus universitetas, fiziniai mokslai, chemija – 03 P).

Disertacija ginama viešame disertacijos gynimo tarybos posėdyje.

Pirmininkas:
Prof. habil. dr. Eugenijus Butkus (Vilniaus universitetas, technologijos mokslai, chemijos inžinerija – 05 T).

Nariai:
Prof. dr. Saulius Bagdonas (Vilniaus universitetas, fiziniai mokslai, fizika – 02 P),

Dr. Rima Budvytytė (Vilniaus universitetas, technologijos mokslai, chemijos inžinerija – 05 T),

Prof. dr. Robert McKenna (Floridos universitetas, fiziniai mokslai, fizika – 02 P),

Prof. dr. Edita Sužiedėlienė (Vilniaus universitetas, fiziniai mokslai, biochemija – 04 P).

Disertacija bus ginama viešame disertacijos gynimo tarybos posėdyje 2018 m. rugsėjo 27 d. 13 val. Vilniaus universiteto Gyvybės mokslų centro R108 auditorijoje. Adresas: Saulėtekio al. 7, LT-10257 Vilnius, Lietuva.

Disertacijos santrauka išsiuntinėta 2018 m. rugpjūčio 15 d. Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje ir VU interneto svetainėje adresu: https://www.vu.lt/naujienos/ivykiu-kalendorius.

Acknowledgements

I would like to express my gratitude to Prof. Daumantas Matulis for the supervision, discussions and suggestions during the doctoral studies in his laboratory. Special thanks also goes to Prof. Virginijus Šikšnys for the possibility to work in the protein crystallography group of his laboratory.

I thank my thesis consultant Dr. Elena Manakova and Dr. Saulius Gražulis for the knowledge learned in the crystallography field. I would like to thank Dr. Virginija Dudutienė and Dr. Edita Čapkauskaitė for the new synthesized chemical compounds designed as CA inhibitors. Special thanks to Vilma Michailovienė, Aurelija Mickevičiūtė, Dr. Jurgita Matulienė, PhD student Justina Kazokaitė, Aistė Kasiliauskaitė, Jelena Jachno and Dr. Vaida Juozapaitienė for the cloning and purification of CA proteins that I used for crystallization. I also thank Dr. Asta Zubrienė, Dr. Vaida Linkuvienė, and Miglė Kišonaitė for the thermodynamic and kinetic data of CA-ligand binding.

I am very grateful to the collegues and students of the Department of Protein-DNA Interactions and the Department of Biothermodynamics and Drug Design for the friendly atmosphere.

I am deeply grateful to my familly for their support. Especially big thanks goes to my lovely wife Evelina for believing in me and critical reading my thesis and notes.

Alexey

Contents

Abbreviations

a.a. amino acids

CA human carbonic anhydrase

DBDD Department of Biothermodynamics and Drug Design

FTSA Fluorescence-based Thermal Shift Assay

ITC Isothermal titration calorimetry

PDB Protein Data Bank

PDB ID Protein Data Bank identifier

PEG polyethylene glycol

SPR Surface plasmon resonanse

Introduction

There are numerous diseases that may be treated by the inhibition of a certain protein target by small molecular weight compounds - drugs that should recognize only the target protein in the human body and selectively inhibit without affecting other proteins. Unfortunately this task is not easy to achieve because the mechanism of specific recognition of the target protein by small molecule is poorly understood. It is not yet possible to rationally design drugs (ligands) *in silico* that would selectively and with high affinity bind to target proteins.

Currently, modern biophysical methods are actively used in drug design. One of the most informative methods is X-ray crystallography which provides the 3D atomic resolution models of protein–ligand complexes and reveals important interactions that determine molecular recognition between the protein and the ligand. In this thesis, we have determined the crystal structures of small molecular weight ligands bound to human carbonic anhydrases (CA) and have analyzed the details of inhibitor binding to several human CA isoforms.

The CAs (EC 4.2.1.1) are ubiquitous metallo-enzymes which catalyze the hydration reaction of carbon dioxide (CO_2) into bicarbonate (HCO_3^-) and protons (H^+). All mammalian CAs belong to the α-class. There are fifteen CA isoforms in human. Most of them (12 out of 15) are catalytically active, while the remaining three are inactive because the catalytic Zn(II) is not present in them. Human CAs participate in many physiological and pathological processes. For example, CA IX and CA XII are associated with certain types of cancer [19, 26] and their inhibition is thought to help stop the proliferation of cancer cells.

Primary sulfonamide compounds have been known for many years to be highly selective inhibitors of CAs. However, most of them bind to all 12 CA isoforms and exhibit little selectivity. Sulfonamides that inhibit the ubiquitous CA I and CA II isoforms are in clinical use as diuretics, antimicrobial agents, etc. The clinically used CA inhibitors are also used against glaucoma, obesity and other diseases. The CA XII isoform is also known as a target to treat glaucoma [20].

Human CAs share high degree of similarity [29] and exhibit a similar fold formed by a central 10-stranded beta-sheet. Small structural differences are found between isoforms that differ in cell and tissue location, some membrane-bound CAs form dimers and carry the trans-membrane helical domain.

The amino acids which form the surface of the conical active site are nearly identical between isoforms, especially in its deeper parts. The active site cavity is approximately 15 Å wide and deep. The catalytically important zinc ion is coordinated by three conservative histidine side chains at the bottom of the cavity. The minor structural differences at the surface of the active site represent a challenge for the design of isoform-selective inhibitors.

Detailed structural information is very important for the understanding of the binding between a ligand and its target protein. Our crystallographic studies seek to provide a structural information for a search of the benzenesulfonamide-based inhibitors selective between human isoforms. We have solved crystal structures of five CA isoforms (CA I, CA II, CA IV, CA XII, and CA XIII) and CA II mutants that mimic the active site of other CA isoforms with series of benzenesulfonamide inhibitors synthesized by Dr. Edita Čapkauskaitė and Dr. Virginija Dudutienė. Chimeric versions of CA IX (chCA IX) and CA XII (chCA

XII) were produced by introducing point mutations into easily crystallizable CA II in order to make the active site cavity similar to that in corresponding isoforms. Mutant proteins (chimeras of CA IX and CA XII) could be easily crystallized like CA II and contained active site residues characteristic for CA IX and CA XII [5].

This structural information was correlated with the binding thermodynamics in search of structural features that determine high affinity and selective recognition.

The goal of this study

The goal of this study was to solve the crystal structures of the complexes of several CA isoforms with a series of sulfonamide inhibitors using X-ray crystallography technique, to analyze the interactions between the ligands and the CA proteins, determine the mechanisms of binding selectivity, affinity, and search for the correlations between the structure and the binding thermodynamic parameters.

The tasks of the study:

(1) to determine the crystallization conditions of CA isoforms (CA IV, CA VA, CA VB, CA VI, CA VII, and CA IX),

(2) to solve the X-ray crystal structures of newly synthesized aromatic benzensulfonamides bound to several CA isoforms and mutant CA II proteins,

(3) to explain the structural reasons of high affinity and high selectivity of the lead compound **VD11-4-2** binding to CA IX over CA I, CA II, CA XII, and CA XIII,

(4) to analyze the interactions of ligands with the amino acids in the CA active sites, differences in the binding modes of sulfonamide inhibitors, and analyze the correlations between the crystallographic structures and the thermodynamic parameters of binding.

Scientific novelty and practical value:

(1) In this work, we have published 61 crystal structures of newly synthesized sulfonamide-based CA inhibitors with 5 CA isoforms (CA I, CA II, CA IV, CA XII, and CA XIII) and also 2 mutant CA II proteins. Out of them, 25 crystal structures were solved with CA II (there are ~500 crystal structures of CA II in the PDB), 14 – with CA XII (there are 17 in the PDB), 12 – with CA XIII (there are 14 in the PDB), 4 – with CA I (there are 25 in the PDB), and 1 – with CA IV (there are 10 in the PDB). Crystal structures of several compounds were determined in complexes with several CA isoforms. This is especially important for the studies of binding selectivity.

(2) The binding of a series of newly synthesized sulfonamides in the active sites of 5 CA isoforms was studied in detail. The mechanism of molecular recognition that increases the binding affinity to the target CA was proposed. The recognition mechanism of CA IX selective binder was studied in detail explaining the crystallographic data.

(3) We have described the tendencies of the changes in the binding thermodynamics between structurally similar compounds. This helps in the design of new compounds with the desired properties.

(4) The set of crystallographic data for the studied CA-ligand complexes is a significant contribution to the existing experimental data that could be used for the future bioinformatic studies of protein-ligand binding.

Defending statements:

(1) The lead compound **VD11-4-2** binds to CA IX differently than to CA I, CA II, and CA XIII. The *meta*-cyclooctyl ring of this compound interacts with the hydrophobic pocket available only in the CA IX isoform and not in CA I, CA II, and CA XIII.

(2) The active cite of chCA IX (multiple mutant of CA II mimicking the active center of CA IX) does not fully resemble the molecular surface of the active center of CA IX at the hydrophobic pocket.

(3) For the pair of compounds differing in the substituents but where the benzene ring is positioned in the protein in a similar position, the addition of the substituent usually does not change the affinity but significantly changes the enthalpy and entropy of binding and the enthalpy-entropy compensation phenomenon is usually observed where an unfavorable enthalpy change is compensated by the entropy change. However, if the positions of bound compounds differ (positions of the benzene rings of the compounds are dissimilar), then the thermodynamic parameters may vary.

1 Materials and Methods

1.1 Materials and Equipment

1.1.1 Equipment

Spectrophotometer 'NanoDrop 1000'; cooling microcentrifuge 'Eppendorf Centrifuge 5415 R'; cooling centrifuge 'Eppendorf Centrifuge 5810-R'; protein crystallization robot 'Oryx8'; X-ray diffractometer 'RIGAKU' MM007-HF; thermostated room for crystal growth; macromolecular X-ray crystallography EMBL beamlines, P13 and P14, at the PETRA III storage ring at DESY (Hamburg, Germany); beamline ID11-3 for macromolecular X-ray crystallography at the MaxLAB IV storage ring in Lund (Sweden); light microscopes 'Carl Zeiss Stemi SN11' and 'ОГМЭ-П2'; protein concentrators 'Amicon-Ultra-0,5ml' 10kDa, protein concentrators Pierce Protein Concentrators (9K); crystallization plates: 'MRC Maxi 48-Well Crystallization; Plate (Swissci)', 'Cryschem Plate', 'CrystalEX (Corning)'; PC.

1.1.2 Materials

The crystallization buffers were home made using chemicals of highest available purity $\geq$ (99-97) %. Manufacturers of chemical reagents: **Sigma, Roth, Fluka, Serva**.

The main buffers: **MES** (2-(N-morpholino)ethanesulfonic acid), **HEPES** (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), **BIS-TRIS** (2,2-Bis(hydroxymethyl)-2,2',2"-nitrilotriethanol), **citric acid/citrate**, **TRIS/HCl** (2-Amino-2-(hydroxymethyl)-1,3-propanediol hydrochloride), **bicine** (2-(Bis(2-hydroxyethyl)amino)acetic acid), **acetic acid**, **acetate**, **tricine** (N-(2-Hydroxy-1,1-bis(hydroxymethyl)ethyl)glycine), **succinate**, **malonate**.

The most used salts: potassium chloride, sodium chloride, ammonium sulphate.

Precipitants: PEG8000, PEG4000, PEG3350, PEG400, iPrOH, PVP (polyvinylpyrrolidone), sodium malonate.

"Hampton" crystallization screens: grid screens (PEG6k, PEG/LiCl, Sodium Malonate), Index, PEG/Ion, crystal Screen Cryo, crystal screen 2 Cryo, and others.

1.1.3 Proteins

CA proteins used in this study were cloned, expressed and purified in the Department of Biothermodynamics and Drug Design by Dr. J. Matulienė, Dr. V. Jogaitė (Juozapaitienė), V. Michailovienė, J. Kazokaitė, J. Jachno, D. Timm, M. Gedgaudas, A. Kasiliauskaitė, and G. Milinavičiūtė.

The mutations of CA II that mimic the active sites of CA IX and CA XII were proposed by Dr. V. Kairys (Department of Bioinformatics, Institute of Biotechnology, Vilnius University) and the corresponding constructs were prepared by the PhD student David Daniel Timm.

The recombinant proteins used for the crystallization experiments in this study: **CA I** (3-261 a.a.), **CA II** pL0059 pET15b-CAII (1-260 a.a.), **CA IV** two proteins: (19-284 a.a.) and (19-284 a.a.), **CA VI** two proteins: (21-280 a.a.) and (31-280 a.a.), **CA VA** (34-305 a.a.), **CA VB** two proteins: (40-317 a.a.) and (40-317 a.a.), **CA VII** (1-264 a.a.), **CA IX** two proteins: (138-390 a.a.) and (138-392, Cys41Ser a.a.), **CA XII** (30-291 a.a.), **CA XIII** (1-262 a.a.).

1.2 Methods

1.2.1 Crystallization of CA I, II, IV, VI, CA IX, CA XII, and CA XIII isoforms

CA protein crystals were obtained at 19 °C using vapor diffusion method in a sitting drop format. The mixed volumes of starting drop were 0.1-5 μL, the volume of crystallization buffer was 0.1-0.5 mL in the wells of crystallization plate. The results of crystallization experiments are listed in Table 1.1.

Table 1.1: CA isoforms and their crystallization.

Isoform	Resolution, (Å)	Comment
CA I	1.3-2.0	The crystallization conditions of CA I were preliminary optimized by Dr. E. Manakova. The crystals were grown for about one week.
CA II	1.0-2.0	The crystallization conditions of CA II were preliminary optimized by Dr. E. Manakova. The crystals were grown for one week.
CA IV	$>$ 2.5	This isoform crystallizes very easily in many different buffers. The attempts to improve resolution were unsuccessful. One crystal structure was obtained. The crystals were grown for about one week.
CA VA	-	The crystallization conditions were not found (800 crystallization buffers were tested).
CA VB	-	The crystallization conditions were not found (800 crystallization buffers were tested). Two protein constructs were used in the crystallization studies (constructs were built by A. Kasiliauskaitė). 1400 crystallization buffers were tested.
CA VI	$>$ 1.6	The crystallization conditions were optimized. This isoform forms homodimers in crystal structures. The active sites are inaccessible for inhibitor molecules. We have not obtained any crystal structure of the protein-ligand complex. The crystals were grown for about 3-4 weeks.
CA VII	-	The crystallization conditions were not optimized. 1000 crystallization buffers were tested.
CA IX	$>$ 3	The crystallization conditions were optimized. Small and geometrically regular crystals were obtained, but diffraction was poor. The crystals were grown for about 4 weeks.
CA XII	1.2-2.0	The crystallization conditions were optimized. The crystals were grown for about 1 week.
CA XIII	1.6-2.0	The crystallization conditions of CA XIII were preliminary optimized by Dr. E. Manakova and M. Kisonaite. The crystals were grown for about 2-3 weeks.

The lowest diffraction resolution at 2.5 Å is still suitable for protein-small ligand complexes, when the position of ligand could be correctly determined in the CA active site. The peaks of water molecules are nearly absent in the electron-density maps at 2.5 Å. A better resolution ($<$ 2 Å) allows the obser-

vation of water molecules and ligand. All CA proteins were concentrated by ultrafiltration to 5-50 mg mL^{-1}. The protein concentration was directly determined by measuring the absorbance at 280 nm (by NanoDrop 1000 instrument). The extinction coefficients of proteins were calculated using ProtParam tool (https://web.expasy.org/protparam/). The crystallization buffers and protein concentrations for every crystallized CA isoform are listed below.

Crystallization buffers and protein concentrations:

The crystallization conditions of the published crystal structures are accessible by PDB ID.

(1) **CA I**: protein concentration 15-25 mg mL^{-1}

1) 0.2 M sodium chloride, 0.1 M Tris-HCl (pH 8.5), 28 % PEG3350
2) 0.2 M sodium chloride, 0.1 M Tris-HCl (pH 8.5), 20 % PEG8000
3) 0.2 M ammonium sulfate, 0.1 M Tris-HCl (pH 8.5), 24-28 % PEG3350
4) 0.2 M ammonium sulfate, 0.1 M Tris-HCl (pH 8.5), 24-28 % PEG4000
5) 0.2 M ammonium sulfate, 0.1 M Tris-HCl (pH 8.5), 20 % PEG8000
6) 0.2 M ammonium acetate, 0.1 M Tris-HCl (pH 8.5), 24-28 % PEG3350
7) 0.2 M ammonium acetate, 0.1 M Tris-HCl (pH 8.5), 24-28 % PEG4000
8) 0.2 M ammonium acetate, 0.1 M Tris-HCl (pH 8.5), 20 % PEG8000
9) 0.2 M sodium acetate (pH 8.3), 0.1 M Tris-HCl (pH 8.5), 24-28 % PEG3350
10) 0.2 M sodium acetate (pH 8.3), 0.1 M Tris-HCl (pH 8.5), 24-28 % PEG4000
11) 0.2 M sodium acetate (pH 8.3), 0.1 M Tris-HCl (pH 8.5), 20 % PEG8000

(2) **CA II**: protein concentration 15-50 mg mL^{-1}

1) 0.1 M sodium bicine (pH 9.0), 0.2 M ammonium sulfate, 2 M sodium malonate (pH 7);
2) 0.1 M sodium bicine (pH 9.0), 2 M sodium malonate (pH 7);

(3) **CA IV**: protein concentration 15-25 mg mL^{-1}

1) 0.2 M ammonium sulfate, 0.1 M sodium acetate (pH 4.5), 23-25 % PEG3350
2) 0.2 M ammonium chloride, 0.1 M Mes sodium (pH 6.5), 15 % PEG4000
3) 0.2 M ammonium sulfate, 0.1 M Mes sodium (pH 6.5), 20 % PEG2000mme
4) 0.2 M ammonium sulfate, 0.1 M Mes sodium (pH 6.5), 20 % PEG3350

(4) **CA VI**: protein concentration 10-20 mg mL^{-1}

1) 0.1 M ammonium citrate (pH 5.0) 20 % PEG2000mme
2) 0.1 M ammonium citrate (pH 5.0) 15 % PEG3350

(5) **CA IX**: protein concentration 3-6 mg mL^{-1}

1) 0.25-0.55 M $NH_4H_2PO_4$

(6) **CA XII**: protein concentration 20-30 mg mL^{-1}

1) 0.1 M ammonium citrate (pH 7.23), 0.2 M ammonium sulfate, 26-30 % PEG4000

(7) **CA XIII**: protein concentration 15-50 mg mL^{-1}

1) 0.1 M sodium citrate (pH 5.5), 0.2 M sodium acetate (pH 4.5), 30 % PEG4000;
2) 0.1 M sodium citrate (pH 5.5), 0.1 M sodium acetate (pH 4.5), 26 % PEG4000;

3) 0.1 M ammonium citrate (pH 7.0), 0.1 M sodium acetate (pH 4.5), 26 % PEG3350;

4) 0.1 M sodium citrate (pH 5.5), 0.2 M ammonium sulfate, 26 % PEG3350;

5) 0.1 M ammonium citrate (pH 5.0), 18-20 % PEG4000;

1.2.2 Preparation of CA complexes with inhibitors

The crystals of CA II, CA IV, CA XII, and CA XIII isoforms were soaked before diffraction experiment for about one week in the crystallization buffer containing 0.5-2 mM of inhibitor. Crystallization buffers for CA crystal soaking were made by the mixing of 0.5-2 μl inhibitor solution (50 mM) and 50 μl of crystallization buffer taken out the reservoir of crystallization plate. The CA I-ligand complexes were obtained by crystallization of pre-mixed complexes CA I with ligand of interest.

1.2.3 Collection of diffraction dataset

X-ray diffraction datasets of CA crystals were collected by the author and colleagues (Dr. S. Gražulis, Dr. E. Manakova) using: **(1)** Rotating Anode X-ray source RIGAKU MM007-HF for protein crystallography; **(2)** the beamlines P13 and P14 at the PETRA III storage ring at DESY (Hamburg, Germany) and **(3)** the Lund synchrotron radiation facility (Lund, Sweden) with the help of Dr. G. Tamulaitiene, Dr. G. Sasnauskas, PhD students A. Merkys and E. Golovinas.

1.2.4 X-ray data processing and solution of the phase problem

Diffraction datasets were processed on the Linux system using the CCP4 software for macromolecular X-Ray crystallography [37]. The programs used in the processing of crystallographic data and structure solution and refinement are listed in the Table 1.2.

Table 1.2: Software used for X-ray diffraction data processing.

Program	Description
MOSFLM [2]	Mosflm is a program for integrating single crystal diffraction data (2D images) from area detectors. The output is an MTZ file which contains the integrated intensities.
XDS [13]	**X**-ray **D**etector **S**oftware for processing single-crystal monochromatic diffraction data recorded by the rotation method. Diffraction datasets from beamlines P13 and P14 (Hamburg, EMBL) were processed by XDS. The output of XDS data processing package is the Text file XDS_ASCII.HKL. (http://xds.mpimf-heidelberg.mpg.de)
COMBAT	COMBAT produces an MTZ file in multirecord form suitable for input to SCALA (from file XDS_ASCII.HKL). (http://www.ccp4.ac.uk/html/combat.html)
SCALA	Scales together multiple observations of reflections, and merges multiple observations into an average intensity.

UNIQUE	Creates a unique list of reflections for a given unit cell with a given symmetry up to a specified high resolution limit. (http://www.ccp4.ac.uk/html/unique.html)
FREERFLAG	Tags each reflection in the MTZ file with a flag for cross-validation. The tagged reflections used for the calculation of 'Free R factors'. (http://www.ccp4.ac.uk/html/freerflag.html)
MTZUTILS	Reflection data files utility program. The MTZ utility program MTZUTILS is provided for the purpose of creating a new re-arranged or edited MTZ reflection data file from one or two existing files. (http://www.ccp4.ac.uk/html/mtzutils.html)
TRUNCATE	Obtains structure factor amplitudes using Truncate procedure and/or generates useful intensity statistics. (http://www.ccp4.ac.uk/html/truncate.html)

We examined several parameters for the quality estimation of processed diffraction data (Table 1.3).

Table 1.3: Diffraction data quality indicators.

Parameter	Description
Diffraction resolution	Resolution is the smallest distance between crystal lattice planes. High numeric value 1.5 Å means good resolution, whereas 4 Å – poor resolution. If the distance between two atoms is larger than high resolution value, these two atoms can be separated.
R_{merge}	A measure of agreement between multiple measurements of the same (not symmetry-related – see Rs_{symm}) reflections, with the different measurements being in different frames of data or different data sets.
I/σ	This parameter correlates with R_{merge}. We limit the resolution (in Å) at $I/\sigma = 2$ that means the choice of reflections whose intensities are more than 2 times stronger than the level of experimental noise.
Completeness	The number of crystallographic reflections measured in the data set, expressed as a percentage of the total number of reflections present at the specified resolution.

The dataset of collected reflections contains only intensities of scattered X-rays measured by detector. The phase information is lost during data collection. There are numerous methods developed in the modern protein crystallography for the solution of "phase problem". Experimental phasing methods are based on the introduction of the heavy or anomalously scattering atoms into the protein crystals that give a strong signal and could be used for phasing.

The other possible approach is the molecular replacement (MR) when the initial phases are calculated from the previously solved protein structure if the model protein is structurally similar to the target protein. We use MOLREP and AMORE programs for the molecular replacement (Table 1.4). MOLREP program

performs the MR in two steps: 1) Rotation function (RF) performs the search orientation of model; 2) Cross Translation function (TF) and Packing function (PF) perform the search position of oriented model.

We use SFALL and SIGMAA programs for the generation of electron-density maps. These programs create the MTZ type file which contains the phases and coefficients. We analyze the electron-density maps and build the atomic model using COOT software. Two types of electron-density maps are important: **(1)** Direct density map (2F(o)-F(c)); **(2)** Difference density map (F(o)-F(c)) shows the difference between the measured electron densities and the electron densities from the current model.

Table 1.4: Software used for crystal model building.

Program	Description
MOLREP	Automated program for molecular replacement. MR is a method for solution of the phase problem in X-ray crystallography. For molecular replacement it is necessary to find solved protein structure which is structurally similar to unknown structure which diffraction data is processed.
AMORE	Amore is the similar program to MOLREP.
COOT [8]	Coot is an interactive software used for macromolecular model building, model completion and validation, particularly suitable for protein modelling using X-ray data.
REFMAC [24, 36]	Refmac is the macromolecular refinement program which performs the refinement of the atomic model against the diffraction data. http://www.ccp4.ac.uk/html/refmac5.html
SFALL	Structure factor calculation and X-ray refinement using forward and reverse FFT. SFALL program calculates the structure factors and phases from the model. http://www.ccp4.ac.uk/dist/html/sfall.html
SIGMAA	This program improves Fourier coefficients using calculated phases. It calculates weighted Fourier coefficients either from the calculated phase from a (partial) model structure. http://www.ccp4.ac.uk/html/sigmaa.html
AVOGADRO [11]	Avogadro is a molecule editor and visualizer. We use this program for generation of small-molecule inhibitors which are the models for crystal structure building.
LIBCHECK	Generates and manages the library files which provide complete chemical and geometric descriptions of residues and ligands used in refinement. The chemical and geometric descriptions were used by REFMAC refinement program.

REFMAC software was used for automatic refinement of added molecules positions and geometry. The changes of electron-density maps were analyzed every time after the model refinement. The final model should not contain the strong positive peaks of difference density map. The models of the inhibitor were added in the active site when the electron densities were detected which can be addressed by compound chemical formula.

2 Results and Discussion

2.0.1 Solved crystal structures

During this work, we have collected 181 X-ray diffraction datasets from CA crystals 2.2. These datasets represent mostly CA-ligand complexes (Table 2.2). The 61 crystal structures of good quality were deposited in PDB database (https://www.rcsb.org/). Primary citation for deposited structures are listed in the Table 2.1.

One third of the crystal structures (62) were empty, meaning that the electron densities of inhibitors were not detected in the active sites of CA, as well as crystal structures of inhibitor-free CA. Most of deposited in PDB crystal structures were used in 11 publications (see publication list 2.1, [32] – as the first author).

The CA inhibitors analyzed in this thesis were synthesized in the Department of Biothermodynamics and Drug Design and can be grouped into two classes:

(1) *mono/di*-substituted fluorinated benzensulfonamides (dr. V. Dudutienė, V class).

(2) *mono/di*-substituted monochlorbenzensulfonamides (dr. E.Čapkauskaitė, E class).

Table 2.1: PDB structures and related primary citations.

PDB ID	Reference	Notes
5LLH, 5LLC, 5LLE, 5LLG, 5LLN, 5LLP, 5LLO, 5MSA, 5MSB	[32]	The authors tried to establish the correlation between the binding of inhibitors in CA-ligand crystal structures and *intrinsic* parameters of binding thermodynamics.
5E2N, 5DOG, 5DOH, 5DRS, 5E2M, 5EHE	[38]	The *intrinsic* binding parameters of *meta*-, *ortho*-substituted fluorinated benzensulfonamides were analyzed.
5IPZ CA IV	[21]	Our crystal structure of the CA IV complex with inhibitor is published. The article also contains the *intrinsic* binding data.
5LL4, 5LL5, 5LL9, 5LLA	[3]	Analysis of *intrinsic* parameters of binding the series of N-alkylated benzimidazoles revealed compounds selective towards CA VA isoform. The analysis of thermodynamics of binding and crystal structures are presented.
4QIY, 4QIZ, 4QJ0, 4QJM, 4QJO, 4QJP, 4QJW, 4QJX, 4QTL	[7]	First publication of *para*-, *ortho*-, *meta*-substituted fluorinated benzensulfonamides. The binding data contains only the *observed* binding parameters.
4WR7, 4WUP, 4WUQ, 4WW6, 4WW8	[39]	The *intrinsic* binding parameters of *para*-substituted fluorinated benzensulfonamides.
4QSA, 4QSB, 4QSI, 4QSJ	[14]	The *intrinsic* binding parameters of compounds synthesized by dr. E. Čapkauskaitė.

4PYX, 4PYY, 4Q0L, 4PZH, 4Q06, 4Q08, 4Q09, 4Q07	[5]	The first publication describing inhibitors selective towards isoform CA IX (synthesized by dr. V. Dudutienė). 8 crystal structures, including mutant proteins of CA II isoform.
4Q6D, 4Q6E	[30]	*Para*-substituted benzensulfonamides synthesized by dr. K. Rutkauskas from Kaunas University of Technology.
4KNJ, 4KNN, 4KP5, 4KNI, 4KNM, 4KP8	[4]	The first publication of several compounds of E class (dr. E. Čapkauskaitė). The *observed* binding data shown for 6 CA isoforms.
4HT2, 4HU1, 4HT0	[6]	First publication of *para*-substituted fluorinated benzensulfonamides synthesized by dr. V. Dudutienė. The *observed* binding parameters are listed.
5OGJ, 5OGO, 5OHH	-	Unreleased
4LHI	-	Unpublished yet

Table 2.2: The summary of X-ray diffraction datasets collected during PhD study.

Diffraction experiments	Notes
169 datasets processed	The number of diffraction datasets which were successfully processed and electron-density maps were drawn.
61 crystal structures deposited into PDB database	These crystal structures have PDB identification numbers.
62 crystal structures contained no inhibitor	The electron-densities of compounds were not present in the active sites of analyzed structures.
28 crystal structures had an unclear electron density in the active site	The electron-density maps could not be completely correlated with certain compounds.
18 structures of crystals grown in the absence of CA inhibitors.	Crystal structures of CA isoforms and mutant CA II proteins crystallized without CA inhibitors.

2.0.2 Crystallographic water molecules in the active site

The crystal structures with the resolution better than 1.8 Å could be used for the analysis of the distribution of crystallographic water molecules in the active site. The rearrangement of solvent molecules after binding event is an open question. The dynamics of the crystallographic water molecules in the active site can be the reason of the entropy-enthalpy compensation effect ([15, 33], software Watermap). It can be a "cornerstone" for the understanding of the protein-ligand binding thermodynamics. CA I, II, XII and XIII usually produce crystals diffracting to the good resolution and could be used for the analysis

of crystallographic solvent molecules. The number of ordered water molecules is influenced by several factors:

(a) **Resolution of the crystal structure**. Protein surface contains mostly charged and polar side chains, such as lysine, aspartate, arginine, glutamine etc. Water molecules bound by hydrogen bonds to the protein surface are often ordered in the crystal structure. The amount of water molecules visible in the structure depends on the resolution. The spacious active site of CA is also filled with solvent molecules and some of them occupy well defined positions (deep water, zinc-bound water etc.) [9, 16, 17]. The number of water molecules in the active site of crystal structure strongly depends on the resolution. Weak peaks of electron-density map can be interpreted as the peaks of more mobile water molecules. Different crystallographic software can generate slightly different electron-density maps.

(b) **Distribution of water molecules could vary between different protein subunits in the same structure**. The crystal structure of CA XII often contains four CA XII molecules. Interestingly, the number of water molecules and quality of water-like peaks is different between active sites of CA XII molecules, as well the conformation of ligands may vary between subunits.

(c) **Crystallization buffer composition**. The number of water molecules observed in the active site depends on the buffer composition resulting in buffer molecules observed at the protein surface instead of water molecules. Thus bicine molecules were often found in CA II crystal structures at fixed position. The hydrophobic DMSO molecules carried over from the stock solution of the inhibitor also occupy the fixed positions in CA II active site.

These reasons complicate the study of the correlation between the dynamics of water molecules in the active sites and thermodynamic parameters of binding.

2.0.3 The *intrinsic* binding parameters of sulfonamide-based inhibitors

The binding thermodynamics in water solution is a combination of thermodynamic processes. The binding of sulfonamide-based inhibitors depends on pH and buffer. The deprotonated sulfonamide ($R\text{-}SO_2NH^-$) is the active form of the inhibitor which can bind to Zn(II) ion in the CA active site. The proton transfer from the sulfonamide group to the buffer is a thermodynamic process, which has some contribution to observed energies (ΔH, ΔG) [27, 35]. The influence of pH on the binding process can be determined and dissected from observed energies [1, 12, 18, 22]. The *intrinsic* thermodynamic parameters describe the reaction of binding of ligand with CA active site, that is free from the interaction of binding buffer with ligand.

The *intrinsic* binding data of synthesized compounds (E and V classes) to human CA isoforms are published in the articles [14, 38, 39]. The increasing number of solved crystal structures and *intrinsic* binding data allowed to perform a comprehensive analysis of binding structure-thermodynamics relationships [32].

2.1 Crystal structure correlations with the *intrinsic* thermodynamics of human CA inhibitor binding.

We apply the strategy of matched molecular pairs for minimization of the possible factors which influence parameters of binding thermodynamics. The matched molecular pairs are composed from two structurally similar compounds, which differ in the substituents in *meta*- or *ortho*- or *para*-position. Moreover, we selected such molecular pairs where the compounds bind to the same CA isoform in the pair.

A series of 24 crystal structures of CAs co-crystallized with 16 benzenesulfonamides bearing different substituents was analyzed by grouping 15 inhibitors into 13 matched molecular pairs where compounds differed by the hydrophobicity of substituents whose contribution is being studied (Fig. 2.1). The 13 matched molecular pairs included six pairs with CA II, two – with CA XIII, one – with CA I, and four – with CA XII. Additionally, two crystal structures of compound **1** were compared when bound in the active sites of CA II and CA XII.

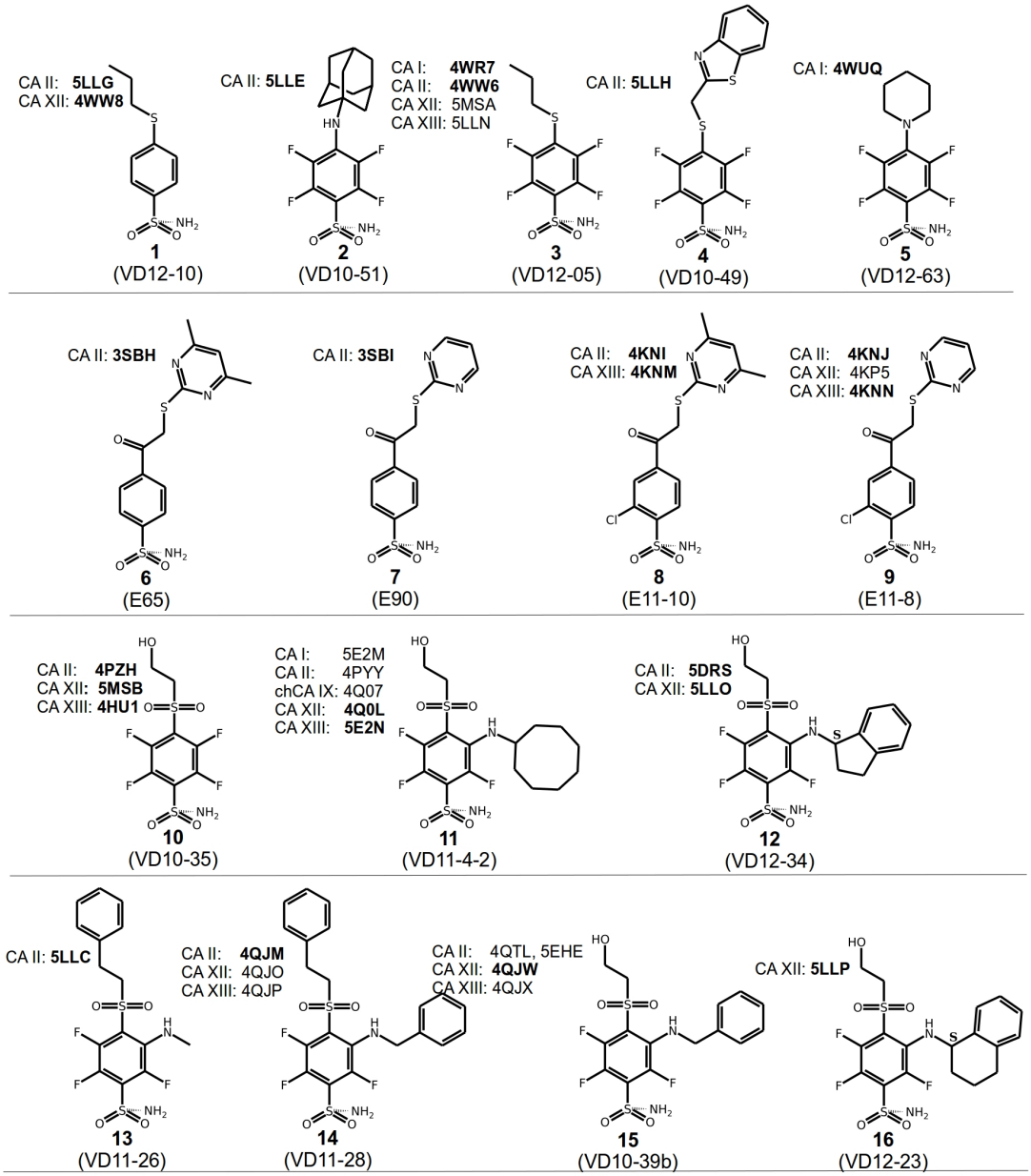

Figure 2.1: The inhibitors and PDB accession codes of related crystal structures.

2.1.1 Similar binders

Nine matched molecular pairs exhibited similar binding mode (group **S1-9**), i.e. the benzene ring was found in the same orientation while only the positions of substituents may have differed in each pair (Fig. 2.2). The remaining four matched molecular pairs exhibited a dissimilar binding mode of the benzenesulfonamide ring (**D1-4**) (Fig. 2.5).

The binding energies are plotted against the sum of accessible and buried surface areas of the hydrophobic substituents in each matched molecular pair (Fig. 2.3). The accessible and buried surface areas were (ASA and BSA, respectively) calculated by VORONOTA algorithm [25] for areas of the marked portion of the compound (green or magenta in Fig. 2.2 and Fig. 2.5). The crystallographic data, electron density maps

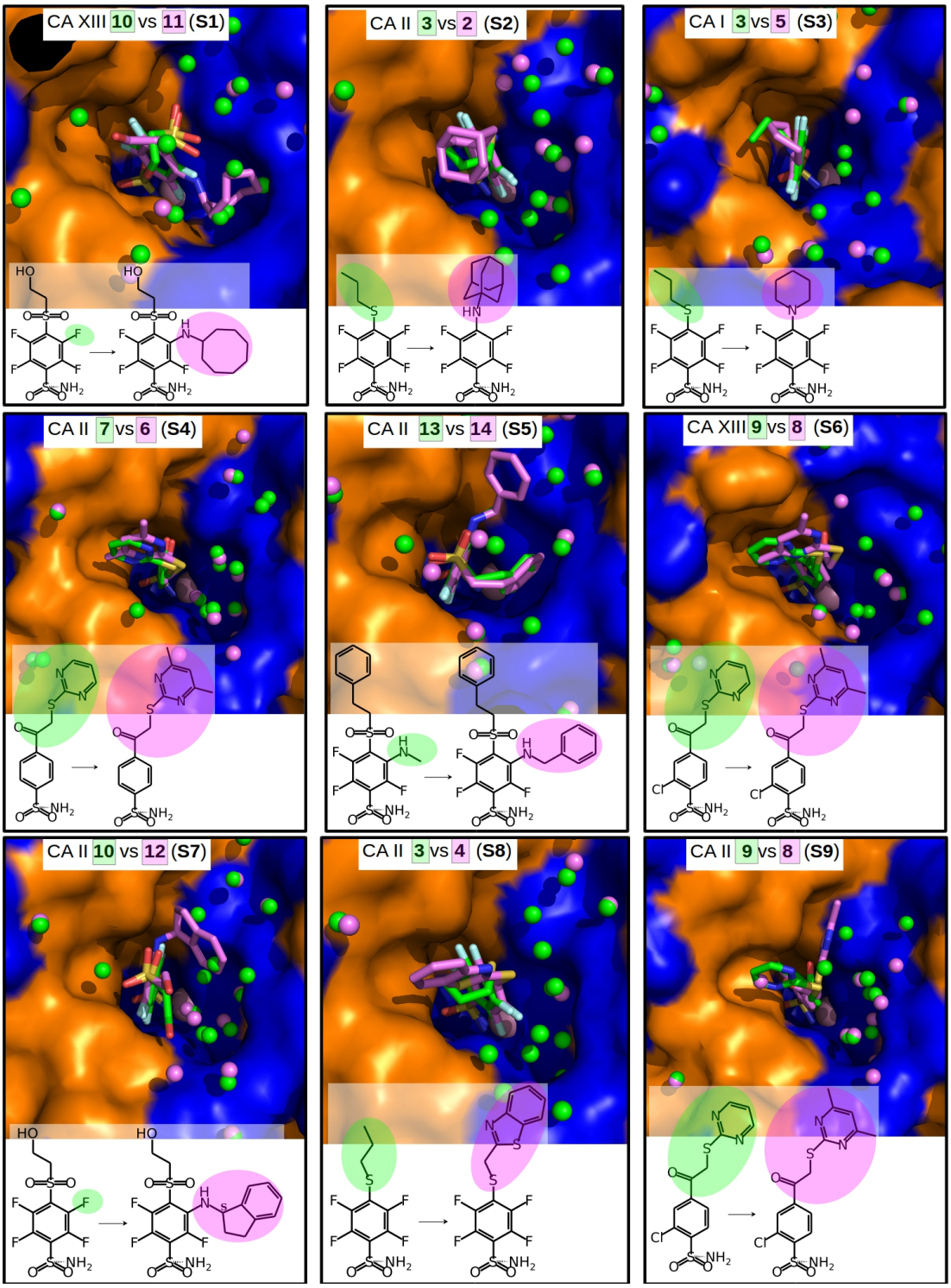

Figure 2.2: "Similar binders": matched pairs of inhibitors bound in the crystal structures of CA isoforms. Inhibitors bound in CA active sites are shown in the same orientation. The structural differences of the inhibitors in pairs are shaded by green or pink. The ligands in crystal structures are colored accordingly. Water molecules found in crystal structures are shown as green and pink spheres. Zn(II) ion is shown as a magenta sphere. The protein surface of CA active site is colored orange for hydrophobic residues (Val, Ile, Leu, Phe, Met, Ala, Gly, and Pro) and blue for the residues with charged and polar side chains (Arg, Asp, Asn, Glu, Gln, His, Lys, Ser, Thr, Tyr, Trp, and Cys). (pair **S1**) Compounds **10** (green, PDB ID 4HU1) and **11** (pink, PDB ID 5E2N) bound in CA XIII.

Figure 2.2: *continued*

(pair **S2**) Compounds **3** (green, PDB ID 4WW6) and **2** (pink, PDB ID 5LLE) bound to CA II. (pair **S3**) Compounds **3** (green, PDB ID 4WR7) and **5** (pink, PDB ID 4WUQ) bound to CA I. (pair **S4**) Compounds **7** (green, PDB ID 3SBI) and **6** (pink, PDB ID 3SBH) bound to CA II. (pair **S5**) Compounds **13** (green, PDB ID 5LLC) and **14** (pink, PDB ID 4QJM) bound to CA II. (pair **S6**) Compounds **9** (green, PDB ID 4KNN) and **8** (pink, PDB ID 4KNM) bound to CA XIII. (pair **S7**) Compounds **10** (green, PDB ID 4PZH) and **12** (pink, PDB ID 5DRS) bound to CA II. (pair **S8**) Compounds **3** (green, PDB ID 4WW6) and **4** (pink, PDB ID 5LLH) bound to CA II. (pair **S9**) Compounds **9** (green, PDB ID 4KNJ) and **8** (pink, PDB ID 4KNI) bound to CA II.

of the compounds and other data are available in ([32], open access).

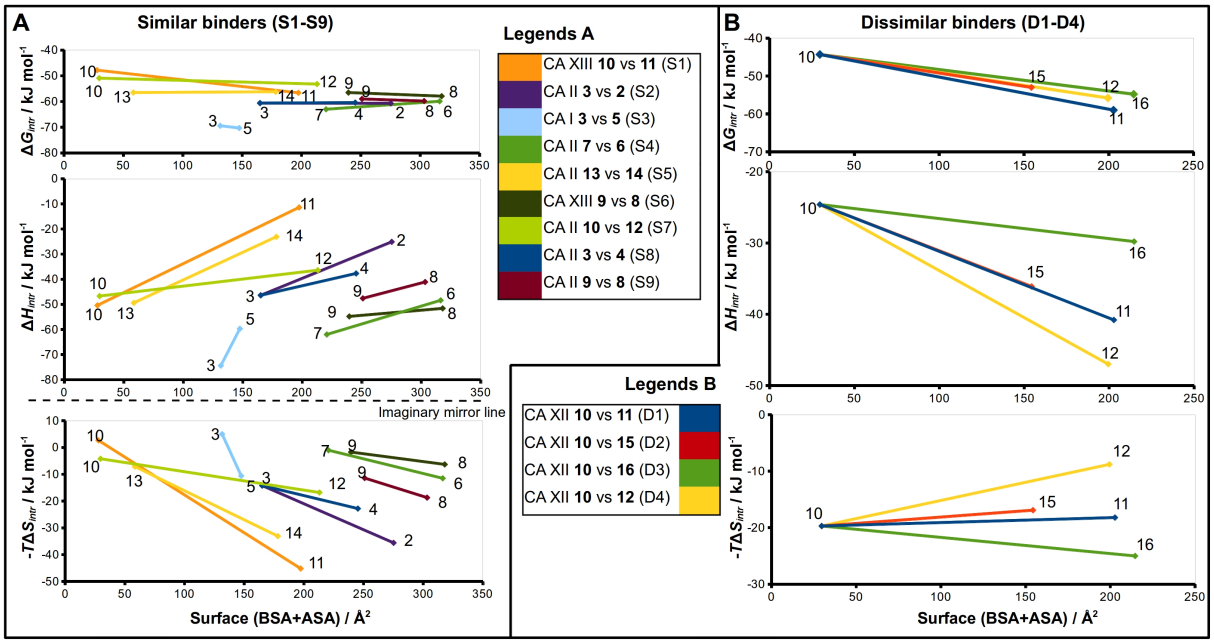

Figure 2.3: Correlations between structural changes and the intrinsic thermodynamics of binding to CA isoforms in the matched pairs for similar **(A)** and dissimilar binders **(B)**. The surface areas, (BSA + ASA), in Å^2) of the functional groups marked by pink and green in (Fig. 2.2 and 2.5) are shown on x-axis. The values of the surface area is the sum of 'voronota' contacts [25] (BSA) and the solvent accessible surface (ASA). The thermodynamic parameters are shown on y-axis. Colored lines connect two compounds in the pair with compound number at the end of lines. Due to the enthalpy-entropy compensation effect described for similar binders (Fig. 2.2A-I, pairs **S1-S9**), the intrinsic entropies mirrored the enthalpies as emphasized by the imaginary mirror plane (dashed line) for the similar binders in **A**.

Intrinsic Gibbs energy comparison for each molecular pair of group **S1-9** (Fig. 2.3A) showed that there was a relatively small increase in affinity with an increase in the accessible and buried surface area. Additional hydrophobic surface seemed to not have a large impact on binding affinities. However, the intrinsic enthalpies of binding in all similar molecular pairs were significantly less exothermic for compounds bearing larger hydrophobic substituents. Additional contacts between the protein and the ligand did not make the enthalpy more favorable. Instead, the intrinsic enthalpies of binding became less favorable. The

entropies of binding mirrored the enthalpies – the enthalpy-entropy (H/S) compensation was observed (see Table 2.3). Similar behavior of binding thermodynamics may be explained by a change in the water structure around the compounds and in the active site.

Table 2.3: The changes of thermodynamic parameters and molecular surfaces (ASA, BSA) of different substituents between the compounds in *similar* molecular pair.

CA isoform	Ligands	$\Delta\Delta G$, (kJ mol^{-1})	$\Delta\Delta H$, (kJ mol^{-1})	$-\Delta T\Delta S$, (kJ mol^{-1})
CA II: (**S2**)	**3** vs **2**	-0.1	21.3	-21.4
CA II: (**S4**)	**7** vs **6**	3.1	13.6	-10.5
CA II: (**S9**)	**9** vs **8**	-0.8	6.5	-7.3
CA II: (**S5**)	**13** vs **14**	0.3	26.3	-26
CA I: (**S3**)	**3** vs **5**	-0.9	14.7	-15.6
CA XIII: (**S6**)	**9** vs **8**	-1.4	3.2	-4.6
CA II: (**S7**)	**10** vs **12**	-2.3	10.3	-12.6
CA II: (**S8**)	**3** vs **4**	0.2	8.7	-8.6
CA XIII: (**S1**)	**10** vs **11**	-8.8	39	-47.9

The pair **S1** (compounds **10** and **11**) is an exception from similar pairs exhibiting 30-fold increase (8.9 nM vs 0.29 nM) in binding affinity to CA XIII upon the addition of cyclooctyl group to **10**. This pair has an interesting feature of binding which is not present in other *similar* pairs. The access of water molecules to sulfonamide did not change after the increase of hydrophobic surface in pairs **S2-S9** (Fig. 2.2). Only in the pair **S1** there are changes in water molecules present near the sulfonamide and Zn(II). Several possibilities to displace water molecules upon compound binding are schematically drawn in Fig. 2.4. The dissociation of **10** from CA XIII (pair **S1**) could be described by the mechanism shown in Fig. 2.4A, while the dissociation of **11** – by the mechanism shown in Fig. 2.4B. The complicated path for water molecule access to displace the sulfonamide is a likely reason for the lowering of the measured off-rate k_d (k_d of SPR: $7.9{*}10^{-3}$ (**11**) vs $1.1{*}10^{-2}$ (**10**)).

We propose that similar binders can have the similar $\Delta H_{interactions}$ values due to the same binding mode to target in each pair. $\Delta H_{exchange}$ gain (from buffer and pH) is estimated and dissected from the observed values. The $\Delta H_{conformational}$ gain can also be similar due to **1)** identical binding mode of both ligands to CA II and **2)** similarity between these compounds. Finally, the terms $\Delta H_{desolvation}$ and $-T\Delta S_{desolvation}$ could be considered as main reason for changes between binding thermodynamics in pairs of similar binders (pairs **S1-9**).

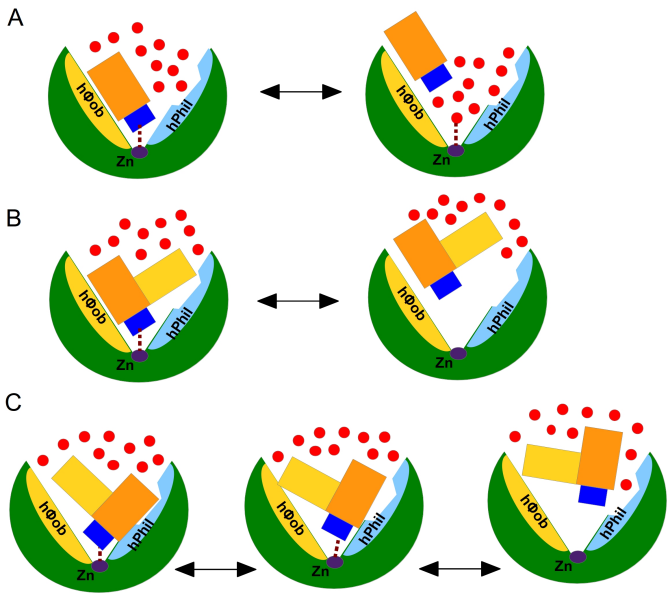

Figure 2.4: Proposed mechanisms of non-classical impact of inhibitor hydrophobic group to the binding affinity. The CA active site is shown as green arc-shaped object where Zn(II) ion is a small violet circle. The schematic compound consists of the sulfonamide (blue rectangle), hydrophobic substituent (yellow rectangle) and benzene ring with *para*-group (orange rectangle). The small red circles represent the water molecules. The h-phob and h-phil labels mean the hydrophobic and hydrophilic parts of schematic active site. **(A)** Water molecules and sulfonamide group of inhibitor freely compete for the binding to zinc ion. **(B)** Larger hydrophobic substituent of the inhibitor obstructed an access of water molecules to zinc, the dissociation of ligand is hindered and depends on the conformation of the ligand. **(C)** The possible dual mechanism of the ligand dissociation (for pairs **D1-4**): **1**) a bulky hydrophobic substituent needs a conformational changes in the protein side chains when bound to zinc **2**) does not allow water molecules to displace the sulfonamide group from coordination with zinc.

2.1.2 Dissimilar binders

The thermodynamic parameters of the *dissimilar* inhibitor binding as a function of ligand substituent surface are shown in Fig. 2.3B. The dissimilar ligands are bound to CA XII that has a more spacious active site due to the absence of phenylalanine side chain (in comparison to CA II). For the *dissimilar* binders, there is a significant increase of the intrinsic affinity between compounds of matched molecular pair (e.g. 318 times for pair **D1**) with an increase in the accessible and buried surface area. The intrinsic enthalpies of binding in all 4 pairs of *dissimilar* binders were significantly more exothermic for compounds with larger hydrophobic substituents. Additional contacts between the protein and the ligand made the enthalpy more favorable for this group. The entropies, however, showed different behavior – significantly favorable for pair **D3**, significantly unfavorable for **D4**, and the changes were negligible for **D1** and **D2**.

Among *dissimilar* binders, the matched molecular pair **D1** (Fig. 2.5), where compounds **11** and **12** are bound to CA XII (**11** bound CA XII 318× weaker than **12**, 35 nM for **11** and 0.11 nM for **12**), is interesting. Compound **11** interacted with the particular part of the active site while the other part is filled by water molecules similar to compound **11** in pair **S3** as shown in Fig. 2.2 pair **S3**. However, the binding kinetics of compound **12** to CA XII could not be measured by SPR due to an extremely low dissociation rate (too slow dissociation, $k_d < 10^{-3}$ s^{-1}).

In the crystal structure of complex CA XII with **12**, the *para*-group is located under the side chains of N64 and K69 and such binding mode may explain the lower k_d. Moreover, **12** effectively displaced water molecules. The compound can exit the active site only after conformational changes of the *meta*-cyclooctyl group that is likely responsible for pushing of the *para*-group under the N64 and K69 side chains. The cyclooctyl group also creates a barrier for water molecules to enter the active site.

Proposed mechanism explaining the high affinity of **12** binding to CA XII is schematically represented in Fig. 2.4C. The bulky hydrophobic group pushes the *para*-substituent of **12** under the side chains of amino acids forming the active site cavity. It is possible that the process of dissociation of compound **12** could begin from the conformational changes of the large flexible hydrophobic group in the *meta*-position which can induce the escape of the *para*-group from under protein side chains. The bulky hydrophobic substituent can serve like a temporary barrier that prevents the entry of water molecules during the process of dissociation. We note that the favorable changes of the binding enthalpies of the *dissimilar* binders could be explained by the favorable location of the hydrophilic *para*-group which is pushed into the hydrophilic part of the active site where the formation of additional hydrogen bonds are more possible.

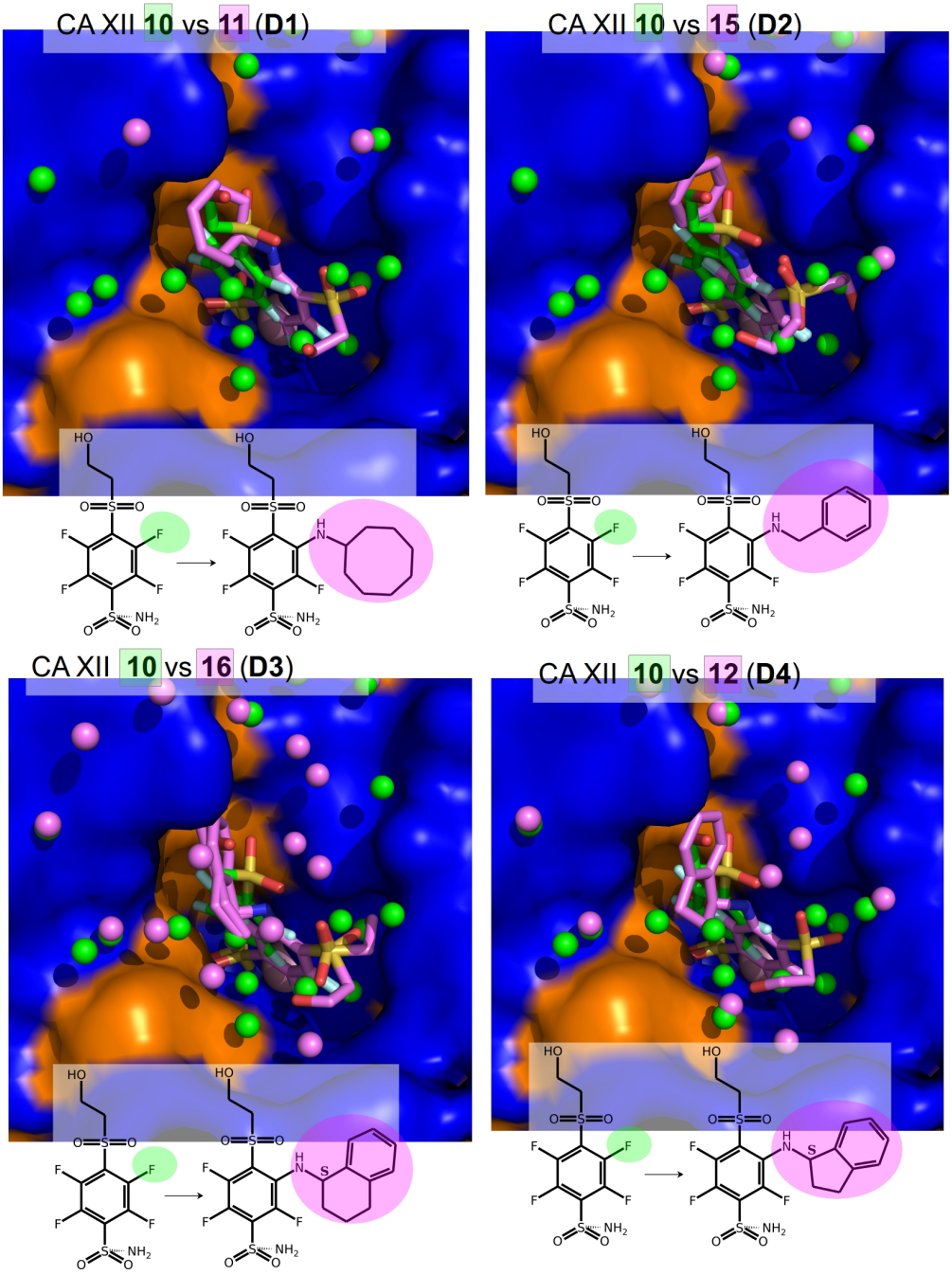

Figure 2.5: "Dissimilar binders": matched pairs of crystal structures. Inhibitors bound to CAs are presented in the same orientation. The structural differences of the inhibitors in pairs are shaded by green or pink. The ligands in crystal structures are colored accordingly. Water molecules found in crystal structures are shown as green and pink spheres. Zn(II) ion is shown as a magenta sphere. The protein surface of CA active site is colored orange for hydrophobic residues (Val, Ile, Leu, Phe, Met, Ala, Gly, and Pro) and blue for the residues with charged and polar side chains (Arg, Asp, Asn, Glu, Gln, His, Lys, Ser, Thr, Tyr, Trp, and Cys). **(pair D1)** Compounds **10** (green, PDB ID 5MSB) and **11** (pink, PDB ID 4Q0L) are bound to CA XII. **(pair D2)** Compounds **10** (green, PDB ID 5MSB) and **15** (pink, PDB ID 4QJW) bound to CA XII. **(pair D3)** Compounds **10** (green, PDB ID 5MSB) and **16** (pink, PDB ID 5LLP) bound to CA XII. **(pair D4)** Compounds **10** (green, PDB ID 5MSB) and **12** (pink, PDB ID 5LLO) bound to CA XII.

The compounds that are more hydrophobic in these pairs (**15**, **16**, **12**, and **11**) differ by the *meta*-

substituent from the reference compound **10** and have the same location of the fluorinated benzene ring in the CA XII active site. The *meta*-substituents are located in the hydrophobic environment and sterically dislocate the remaining part of the compound that is bound in the hydrophilic concave of CA XII active site (**Fig. 2.5**). The favorable changes of the binding enthalpies of the *dissimilar* binders **D1-4** could be explained by the favorable location of the hydrophilic *para*-group which is dislocated into the hydrophilic part of the active site where the formation of additional hydrogen bonds is more likely.

It is interesting to compare the compounds **12** (pair **D4**) and **16** (pair **D3**) that differ by only one bridged methyl group in the *meta*-substituent. The difference in the binding energy between these compounds (**12** vs **16**) is significant: $\Delta\Delta G$ = 1, $\Delta\Delta H$ = 17.2 (favorable), $\Delta T\Delta S$ = -16.2 (unfavorable) (kJ mol^{-1}). As we have mentioned previously these compounds are bound to CA XII in the identical mode (compare Fig. 2.5D3, D4). This pair can also be added to the group of *similar* binders. Difference in the binding energy between **12** and **16** also could be explained by two hydrogen bonds formed by the *para*-group of **12** with residues Asn64 and Pro200. Hydrogen bonds fix the conformation of the *para*-substituent with respect to CA XII active site. The *para*-group of **16 (VD12-23)** was found in the alternate conformation, but in the compound **12** bound to CA XII both substituent groups were found in a single conformation (Fig. 2.5D3 and D4).

2.1.3 Aliphatic-aromatic stacking interactions in the CA-inhibitor complexes

Compound **1 (VD12-10)** has the identical orientation of the benzenesulfonamide ring in the binding to CA II and CA XII as compounds of the pair **S4** (Fig. 2.6). As seen in both crystal structures, the ligand **1** makes the tight aromatic-aliphatic interaction between the benzene ring and a conservative leucine side chain (Leu198 in CA II and Leu197 in CA XII). The compound **1** has a relatively simple structure and its binding is enthalpy-driven (Fig. 2.6(C1)) exhibiting very high intrinsic affinity to both analyzed CA isoforms, 7.4 pM to CA II and 67 pM to CA XII (9× higher affinity to CA II). An important difference between CA II and CA XII is the substitution of Phe131 in CA II with the alanine side chain in CA XII. A loss of the exothermic enthalpy ($\Delta\Delta G$ = 5.7, $\Delta\Delta H$ = 12.6, $\Delta T\Delta S$ = -6.9 (kJ mol^{-1}) was observed for the binding of **1** to CA XII as compared to CA II due to the possible absence of additional interaction of *para*-substituent of **1** with Phe131 in CA II. In the crystal structures of **1** in CA II and CA XII, the loss of additional surface of the contact with protein (Ala instead of Phe) changes the location of the propyl group. In CA XII this group is directed perpendicularly to the plane of aromatic-aliphatic interaction between Leu198 and benzene ring of **1**. The propyl group of the compound **1** bound to CA II is located in the hydrophobic cavity (formed by Phe131 and Pro202) and possibly increases the binding affinity due to same direction of interactions between **1)** propyl group and hydrophobic cavity and **2)** benzene ring of compound and leucine side chain.

The pair **S4** (CA II, **7** vs **6**, Fig. 2.2) has the decrease of the binding affinity towards CA II due to the uncompensated enthalpic loss after increase of the hydrophobic surface (two methyl groups are added to pyrimidine in **6**). The K_d values of **7** and **6** bound to CA II are 25 pM and 80 pM, respectively. The more hydrophobic compound **6** binds only 3× weaker to CA II, from the other hand such values are on the border of the inherent uncertainty of FTSA method. It is interesting to understand why these compounds bind to CA II so strongly. As seen in both crystal structures shown in Fig. 2.6(S4), the ligands are likely

to bind so strongly due to the tight aromatic-aliphatic interaction between the benzene ring and a leucine side chain (Leu198 in CA II). Methylpyrimidine is larger than pyrimidine and could not occupy the same deeper position as pyrimidine between side chains of Pro204 and Phe133 in CA II. The aromatic-aliphatic interaction looks like a possible reason of the favorable enthalpic gain (**7**: ΔH = -62.0 kJ/mol and **6**: ΔH = -48.4 kJ mol^{-1}) of which a major part is due to ΔH interactions.

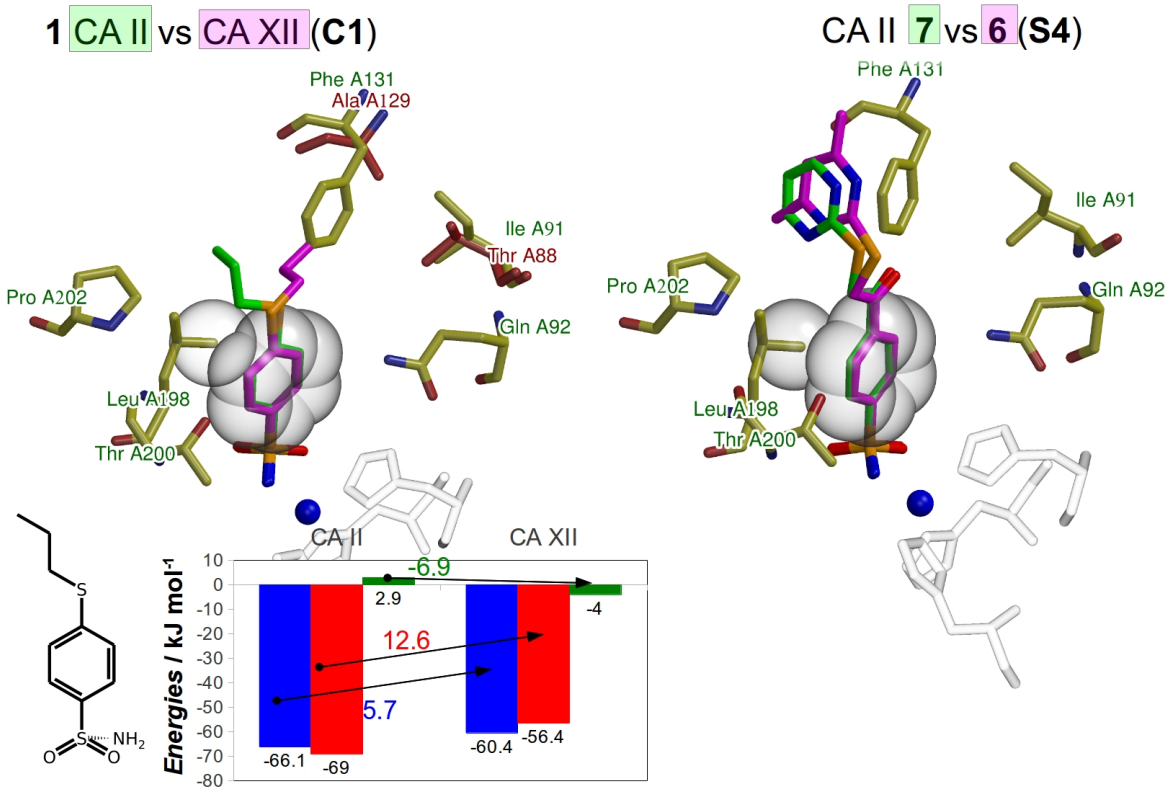

Figure 2.6: The compound **1** bound to CA II and CA XII (left panel) and compounds of pair **S4** (right panel). The histidine side chains holding the active site Zn(II) (blue sphere) ion are transparent. The Leu198 conservative between isoforms CA II and XII and benzene rings of compounds are shown in CPK representation which marked the aliphatic-aromatic interaction. **(C1)** Comparison of the compound **1** position bound to CA II (**1** is green, CA II is colored yellow (green labels), PDB ID 5LLG) and CA XII (**1** is pink, CA XII side chains are red (red labels), PDB ID 4WW8) and corresponding binding thermodynamics. The histogram in the insert shows the comparison of binding thermodynamics: ΔG is blue, ΔH – red and -$T\Delta S$ – green. **(S4)** Compounds **7** (green, PDB ID 3SBI) and **6** (pink, PDB ID 3SBH) bound to CA II (colored yellow and green labels).

The compound **1** in binding to CA II and CA XIII and the compounds of the pair **S4** (to CA II) are enthalpy-driven compounds:

1) CA II-**7**: (ΔG = -63, ΔH = -62 (kJ mol^{-1})), CA II-**6**: (ΔG = -59.9, ΔH = -48.4);

2) CA II-**1 (VD12-10)**: (ΔG = -66.1, ΔH = -69), CA XII-**1**: (ΔG = -60.4, ΔH = -56.4).

2.2 Selective inhibitors of CA IX

As it was mentioned in Table 1.1, our crystals of CA IX were small and diffracted to very low resolution (9 Å). We tried to obtain crystals of this isoform because its overexpression is connected with certain tumors [34]. It was suggested to design several point mutations into easily crystallizable CA II active site in order to mimic the active site of CA IX. This approach was introduced with two mutations [10, 31] and was developed further up to seven mutations [28] and finally to 10 mutations [23]. Mutant proteins (chimeric CA IX) could be easily crystallized like CA II and contained active site residues characteristic for CA IX. The mutant (chimeric CA IX, chCA IX) form of CA II with six mutations (S65A, Q67N, L91I, V130F, L134V, A203L) was designed in the Department of Biothermodynamics and Drug Design ([5] or Table 2.4.

The mutations were supposed to better resemble the environment of the active site of CA IX. Unfortunately, such an approach cannot be considered as fully replacing the crystallization of CA IX itself. The binding data of chCA IX did not precisely resemble the binding to CA IX (Table 2.5). However, this method was the best available at that time to enable structural visualization of the mode of binding of various CA IX inhibitors. Furthermore, this approach worked excellently with CA XII and chCA XII, for which we determined the crystal structures of both chimeric chCA XII and CA XII itself. Both the binding data and the crystal structures confirmed that chCA XII is an excellent model of CA XII (see Table 2.5).

Table 2.4: Point mutations introduced into CA II in order to resemble the active sites of CA IX and CA XII.

Chimeric CA	Mutations of CA II active site
chCA IX (6 mutations)	A65S, N67Q, I91L, F130V, V134L, L203A
chCA XII (6 mutations)	A65S, N67K, I91L, F130A, V134S, L203N

Table 2.5: Compound dissociation constants ($K_{d,observed}$ (nM), FTSA) of the 6 compounds for CA II, CA IX, CA IX (Cys41Ser), CA XII, CA XIII, chCA IX, and chCA XII. Data are taken from article [5].

CA	**VD12-09**	**PG7**	**VD11-4-2**	**VD10-13**	**VD10-35**	**AZM**
CA I	50 000	>200 000	710	0.11	0.20	1400
CA II	1300	>200 000	60	6.7	17	38
CA IX	1.1	9.5	0.050	32	50	20
CA IX (Cys41Ser)	2.0	33	0.050	-	25	33
chCA IX (6 mutations)	25	630	2.0	63	83	50
CA XII	330	>200 000	3.3	220	250	130
chCA XII (6 mutations)	500	>200 000	6.7	310	250	330
CA XIII	140	1700	3.6	8.3	29	50

The chemical structures of CA inhibitors discussed here are shown in the Fig. 2.7. Acetazolamide (AZM) is commonly used as a control inhibitor of CAs.

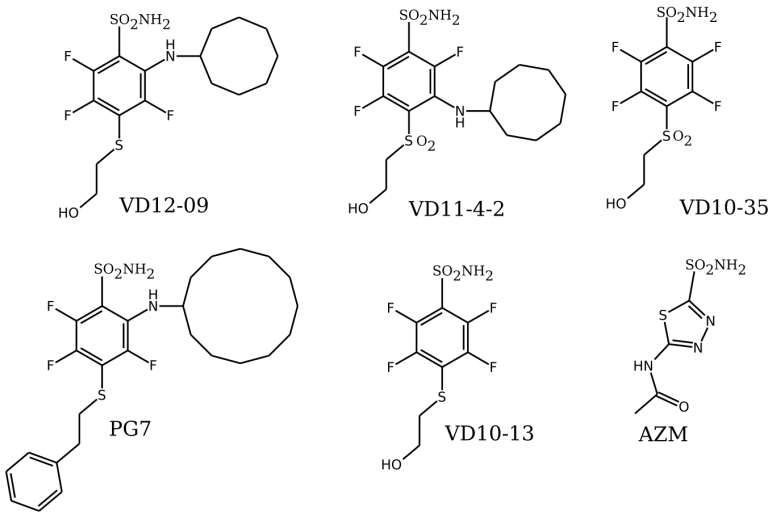

Figure 2.7: Chemical structures of CA inhibitors discussed in this section. Compounds **VD11-4-2**, **VD12-09**, and **PG7** are the inhibitors selective to CA IX.

The compounds **VD10-13** and **VD10-35** are *para*-substituted benzensulfonamides which bound CA IX relatively weakly and exhibited essentially no selectivity toward CA IX (Table 2.5). The compounds **VD11-4-2**, **VD12-09**, and **PG7** are selective inhibitors of CA IX isoform. Thus, these ligands bind CA IX with $K_{d,observed}$ that were 1180×, 1200×, and 21000× higher than the corresponding values for CA II.

Fig. 2.8A illustrates the binding of compound **VD10-35** to CA II and CA XIII active sites. This compound occupies the same part of active site, whereas the other part of active site contains water molecules (see yellow and orange spheres). The active sites of CA II and CA XIII are too spacious for this compound. We propose that it could explain the unselective binding of the *para*-substituted benzenesulfonamides to CA IX. The CA IX active site is larger than in CA II and CA XIII.

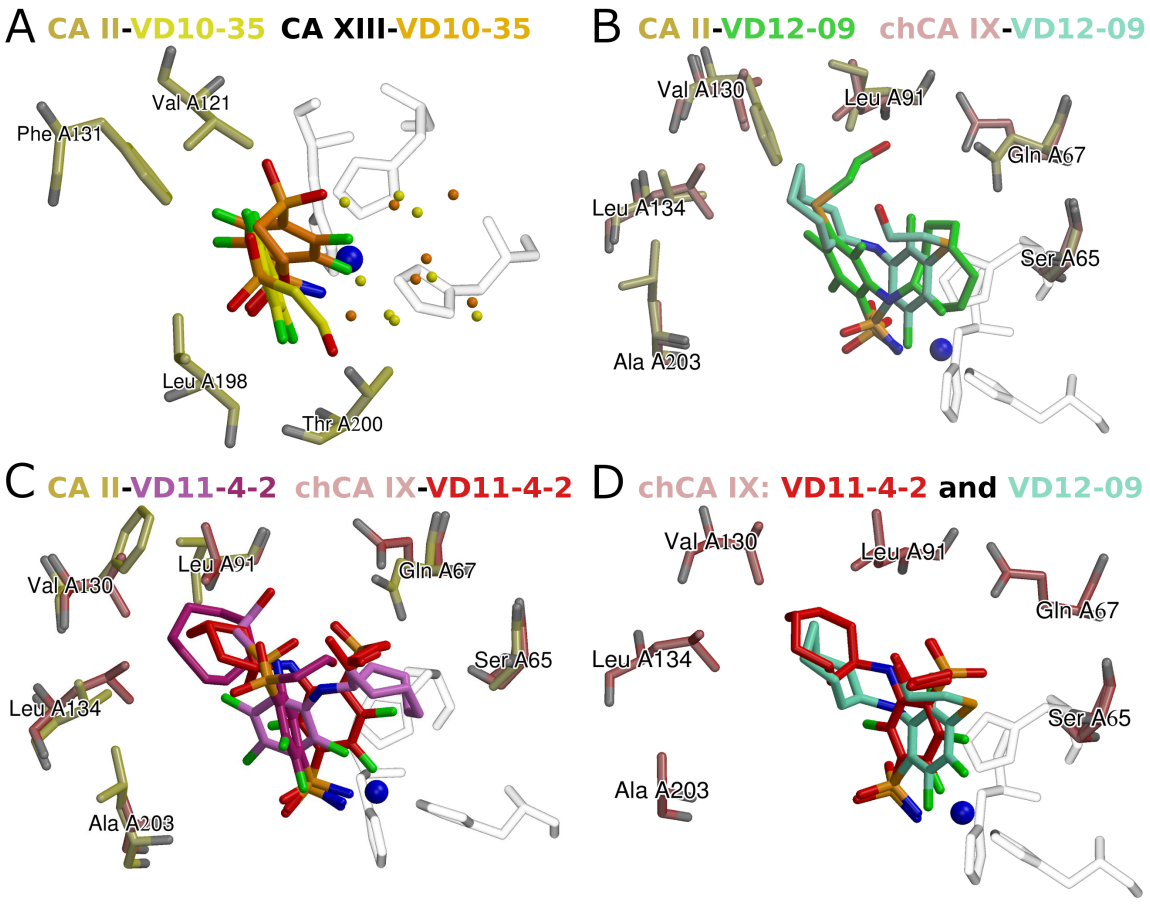

Figure 2.8: The binding modes of inhibitors **VD10-35**, **VD12-09** and **VD11-4-2** bound to the active sites of CA isoforms. The surrounding amino acids of the CA II are colored yellow and chCA IX - crystalline red. The zinc ion is shown as a blue sphere, and the histidine residues holding the zinc are transparent. Residue labeling corresponds to chCA IX. (A) Compound **VD10-35** bound to CA II (yellow; PDB ID 4PZH) and CA XIII (orange; PDB ID 4HU1) exhibits a significantly different rotated position of the benzene ring. The active sites also contain numerous water molecules bound deeply in the active site (shown as small spheres: yellow in CA II and orange in CA XIII). The presence of these water molecules shows how much space is left unoccupied by ligand in the active site. These water molecules are absent in the structures with compounds **VD12-09** and **VD10-35**. (B) Compound **VD12-09** bound to CA II (green; PDB ID 4PYX) and chCA IX (aquamarine; PDB ID 4Q06). In CA II the cyclooctyl group is pushed by Phe131 into the opposite direction as compared with chCA IX. (C) Compound **VD11-4-2** bound to CA II (there are two inhibitor orientations shown in violet and violet-red; PDB ID 4PYY) and chCA IX (red; PDB ID 4Q07). Two orientations of the compound in CA II indicate the reason why it is bound much more weakly to CA II than to CA IX. (D) Comparison of the binding modes of compounds **VD12-09** (aquamarine; PDB ID 4Q06) and **VD11-4-2** (red; PDB ID 4Q07) bound to chCA IX. The overall positions of both compounds are similar.

Figure 2.8B compares the binding of compound **VD12-09** to CA II and chCA IX. The chCA IX (as well as CA IX, as judged from the crystal structure PDB ID 3IAI) contains a deeper hydrophobic pocket than that of CA II, because the bulky Phe131 in CA II occupies part of the pocket. Using this difference between CA II and CA IX, rather bulky cyclooctyl group has been designed to fit the newly accessible pocket formed by residues Val130, Leu134 and Leu91. The ligand **VD12-09** adopts the opposite orientation in CA II, and makes poor connections with the protein. The cyclooctyl group is located in the hydrophilic part of CA II

active site.

Compound **VD11-4-2**, bearing the cyclooctyl group in the *meta*-position, exhibited some of the highest affinities to CA IX ever observed. Compound **VD11-4-2** bound CA II with 60 nM and CA IX with 50 pM affinities. Thus, there is 1000× selectivity toward CA IX. The compound **VD11-4-2** when bound to CA II exhibited two oppositely facing positions (Figure 2.8C, violet and violet-red). Interestingly, the cyclooctyl in the conformation colored violet-red changed the position of the side chain of Phe131. The cyclooctyl group of the second molecule displaces the water molecules from the hydrophilic part of CA active site (Figure 2.8C). The compound **VD11-4-2** bound chCA IX quite similarly to **VD12-09** (Figure 2.8D). The overall position of the cyclooctyl group was also similar in both structures, but in **VD11-4-2**, the hydrophobic group located deeper in the hydrophobic pocket.

As we have mentioned before, the binding data (Table 2.5) and the crystal structures confirmed that chCA XII is an excellent model of CA XII. The compound **VD11-4-2** was bound to both active sites of chCA XII and CA XII in the identical binding mode (Figure 2.9A). The Figure 2.9B contains the comparison of the binding modes of compound **VD11-4-2** to chCA IX and CA XII active sites. The compound **VD11-4-2** has identical binding mode in the active sites of CA XII, chCA XII and chCA IX. Moreover, from the personal communication (K. Tars) it is known that the compounds **VD11-4-2** and **VD12-09** have the same position in active site of CA IX as in the crystal structures of the complexes chCA IX-**VD11-4-2** and chCA IX-**VD12-09**, respectively: the fluorinated rings are perfectly matched.

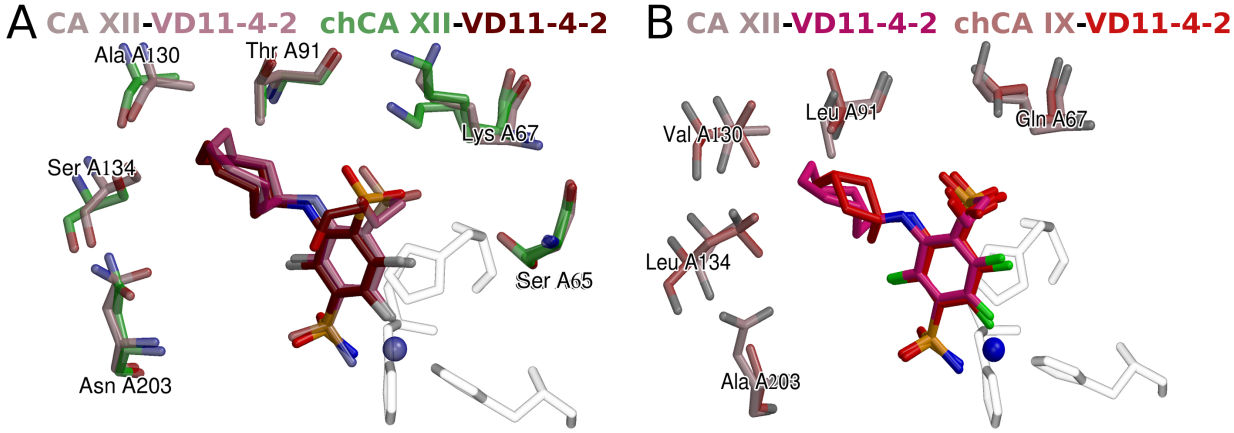

Figure 2.9: X-ray crystallographic structures comparing the binding modes of inhibitor **VD11-4-2** bound to the active sites of CA XII, chCA IX and CA IX. Inhibitor **VD11-4-2** is shown in the red colors. The surrounding amino acids of the CA isoforms are colored red (chCA IX), pink (CA XII) and green (chCA XII). The zinc ion is shown as a blue sphere, and the histidine residues holding the Zn are transparent. (A) Compound **VD11-4-2** bound to CA XII (deep pink; PDB ID 4Q0L) and chCA XII (dark-red; PDB ID 4Q09) in almost the same orientation. Residue labeling corresponds to chCA XII. (B) The positions of **VD11-4-2** bound to chCA IX (red; PDB ID 4Q07) and CA XII (pink; PDB ID 4Q0L) were also essentially similar. Residue labeling corresponds to chCA IX.

The compounds **VD10-35** and **VD11-4-2** were already analyzed in the previous section (*Dissimilar* binders) as the molecular pair **D1** in the binding to CA XII. The *intrinsic* binding affinity to CA XII within this pair increased 318× after the increase of hydrophobic surface. If we analyze the observed binding data, the change of binding affinities of these compounds to CA XII is 75× and to CA IX – 500-1000×.

There is no crystal structure of compound **VD10-35** bound to chCA IX, but analysis of the known crystal structures of similar compounds showed that *para*-substituted fluorinated benzensulfonamides in the active sites of CA II, CA XII, and CA XIII are located in same part CA active sites (e.g. **VD10-35**: CA II (PDB ID: 4PZH), CA XII (PDB ID: 4HU1), and CA XIII (PDB ID: 5MSB), Figures 2.8B and 2.5). We proposed that **VD10-35** is also located there in CA IX active site.

We also proposed that the stronger binding of compound **VD11-4-2** than **VD10-35** to CA IX depends on the same binding mechanism as in the case of CA XII. The *meta*-substituents are located in the hydrophobic environment and sterically dislocate the remaining part of the compound in the hydrophilic part of CA IX active site. The hydrophilic *para*-group, which is sterically dislocated into the hydrophilic part of the active site, establishes the hydrogen bonds with the hydrophilic part of CA IX active site (Figure 2.9A). The active site of CA IX is divided to same hydrophobic and hydrophilic part like the active sites of CA II, CA XII, and CA XIII [29]. Moreover, the *para*-group of compound **VD11-4-2** is more hydrophilic than **VD12-09**: -SO_2- instead -S-. Thus, the net of hydrogen bonds between the fragment of *para*-group -SO_2 (compound **VD11-4-2**) and hydrophilic part of CA active sites could be a major reason for the better binding of compound **VD11-4-2** to CA IX and CA XII in comparison with compound **VD12-09** (-S-): to CA IX – 22-40×.

We can try to explain the better binding of the compounds **VD11-4-2** to CA IX vs. CA XII using our crystal structures. The fluorinated benzene ring of compound **VD11-4-2** in CA IX (K. Tars, personal communication) and chCA IX active sites occupies the identical position in the hydrophilic cavities of the active sites. But in the CA IX active site the cyclooctyl group occupies a slightly different position in comparison with the position in chCA IX. As previously stated, the binding data of chCA IX did not precisely resemble the binding to CA IX. **VD11-4-2** did not exhibit similar binding affinities to CA IX and chCA IX due to the variation in shape of the hydrophobic part of the active site. Residues in the hydrophobic part of chCA IX are shifted ≈ 2 Å) in comparison with the positions of same residues of CA IX active site (Fig. 2.10). The chCA XII and CA XII exhibited similar binding because the residues of chCA XII and CA XII active sites (Fig. 2.9) are perfectly matched. Possibly, the variation in hydrophobic part between CA XII and CA IX active sites is a major reason of binding differences of compound **VD11-4-2** to these isoforms (**VD11-4-2** bound the CA IX 66× better).

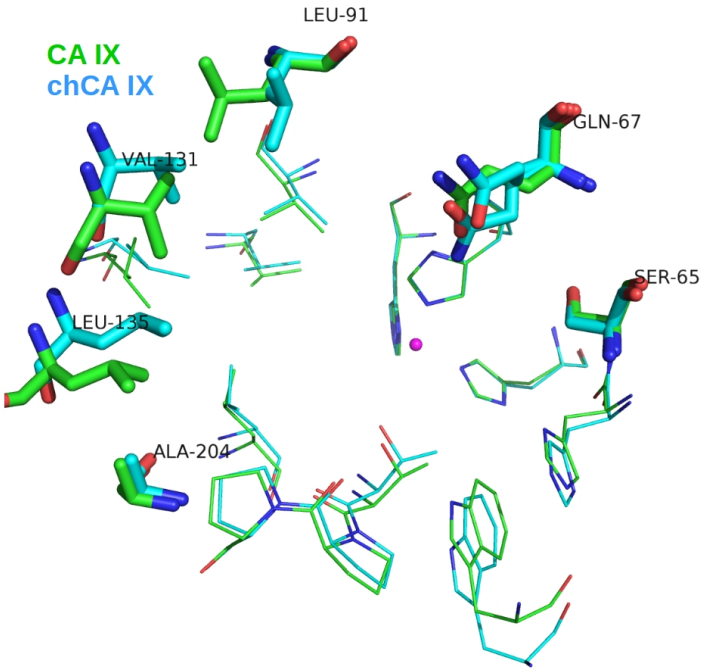

Figure 2.10: Active site residues of chCA IX superimposed on CA IX (PDB ID: IAI3). The amino acids of the chCA IX are colored cyan and CA IX – green. Most residues have identical positions between chCA IX and CA IX active sites (including the mutated amino acids Gln67, Ser65, and Ala204). Conformations of Leu91 are different, while for Val131 and Leu135 there is a 1,5-2,0 Å shift in the positions od Cα atoms.

As we can see from the data in the Table 2.5, compound **VD11-4-2** bound in CA I and CA II isoforms 14200× and 1200× weaker (by observed data) in CA IX. The structural analysis of the binding of compound **VD11-4-2** to CA I was performed in the article [38]. We have shown there that the active site of CA I is narrow for **VD11-4-2**, since CA I contains in active site two histidine residues (His67 and His200) that makes it narrow. The low affinity of the compound towards ubiquitous isoforms CA I and CA II makes it a very interesting lead compound, because these isoforms are important for the respiration and transport of CO_2/bicarbonate. The compound **VD11-4-2** bound CA XIII active site in the similar mode as CA II active site (see (CA II) Fig. 2.8C (violet molecule of inhibitor) and (CA XIII) 2.2 (S1)).

The compound **VD12-09** cannot efficiently bind the CA I active site due to the narrowness of active site and *ortho*-substitution of the compound. **VD12-09** bound CA I and CA II isoforms 45000× and 1000× weaker than CA IX. The compound **PG7** exhibited the good selectivity, but it is a strongly hydrophobic compound and has low solubility in water. The hydrophobic compounds are often toxic and have strong side effects.

The mechanisms of binding selectivity of our lead compound **VD11-4-2** to CA IX vs CA I, CA II, CA XIII were explained in detail using crystallographic data. **VD11-4-2** compound exhibited the different binding mode to the CA IX vs CA I, CA II, and CA XIII due to the shape of the hydrophobic pocket which is available only in the CA IX isoform.

Conclusions

(1) Crystallization conditions for CA IV, CA VI, and CA IX were determined experimentally. The CA IX crystals diffracted to the resolution less than 9 Å, and were not suitable for the crystallographic analysis. The diffraction properties of CA VI crystals were good (1.6-2.0 Å). In the CA VI crystal structure, the protein crystallized as a dimer with the entrances to the active sites obstructed and therefore always without the inhibitors. The CA IV easily crystallized and the obtained crystals were of medium resolution (2.5 Å).

(2) Out of 181 solved crystal structures, 61 structures of CA-ligand complexes were deposited to the PDB. The binding modes of newly synthesized aromatic sulfonamides in the active site of five CA isoforms (CA I, CA II, CA IV, CA XII, and CA XIII) and several CA II that mimic active sites of CA IX and CA XII were determined. The crystals of studied proteins diffracted up to the resolution of 1.1-2.0 Å, suitable to establish the detailed position of ligands in the active sites. No significant conformational changes of the protein main-chain in the CA structures before and after binding of studied ligands were detected. The binding energies depended on the interactions between protein and ligand, but not on the protein conformational changes during the binding event. A group of compounds were described bound to several CA isoforms providing important information for the studies of binding selectivity between CA isoforms.

(3) We have explained the selectivity of the lead compounds **VD11-4-2** and **VD12-09** to CA IX isoform using crystal structures of CA isoforms I, II, XII, and XIII (PDB ID: 5E2M, 4PYY, 4Q0L, 4Q09, 4Q07, 5E2N) and chimeric CA IX and CA XII.

(4) We have analyzed the relationships between the *intrinsic* binding thermodynamics and the structural data of CA-ligand complexes. After the addition of hydrophobic group, the modified ligand usually became more entropy- and less enthalpy-driven than the primary ligand. Introduction of additional hydrophobic surface often did not change the binding affinity to the protein when the modified and primary ligands were found in the identical positions in the active site. In the other cases, the addition of hydrophobic group significantly affected the binding affinity and the modified compound was bound in the active site in other position and the thermodynamic results varied.

List of publications

Publications included in this thesis:

(1) **Smirnov A.**, Zubrienė A., Manakova E., Gražulis S., Matulis D. 2018. Crystal structure correlations with the intrinsic thermodynamics of human carbonic anhydrase inhibitor binding. *PeerJ*, 6:e4412, Feb. 2018. ISSN 2167-8359. doi: 10.7717/peerj.4412.

I have performed the comprehensive analysis of ligand binding modes in the crystal structures and their binding thermodynamics. I wrote the manuscript. I have solved and deposited new crystal structures into PDB database (5LLH, 5LLC, 5LLE, 5LLG, 5MSB, 5LLO, ir 5LLP).

(2) Dudutienė V, Matulienė J, **Smirnov A**, Timm DD, Zubrienė A, Baranauskienė L, Morkūnaitė V, Smirnovienė J, Michailovienė V, Juozapaitienė V, Mickevičiūtė A, Kazokaitė J, Bakšytė S, Kasiliauskaitė A, Jachno J, Revuckienė J, Kišonaitė M, Pilipuitytė V, Ivanauskaitė E, Milinavičiūtė G, Smirnovas V, Petrikaitė V, Kairys V, Petrauskas V, Norvaišas P, Lingė D, Gibieža P, Capkauskaitė E, Zakšauskas A, Kazlauskas E, Manakova E, Gražulis S, Ladbury JE and Matulis D. Discovery and characterization of novel selective inhibitors of carbonic anhydrase IX. *J Med Chem* 2014, 57: 9435-9446.

I have crystallyzed the CA proteins, solved the crystal structures and deposited their into PDB (4PYX, 4PYY, 4Q0L, 4PZH, 4Q06, 4Q08, 4Q09, 4Q07). I have participated in the manuscript writting.

(3) Zubrienė A., **Smirnov A.**, Dudutienė V., Timm DD., Matulienė J., Michailovienė V., Zakšauskas A., Manakova E., Gražulis S., Matulis D. 2017. Intrinsic Thermodynamics and Structures of 2,4- and 3,4-Substituted Fluorinated Benzenesulfonamides Binding to Carbonic Anhydrases. *ChemMedChem* 12:161–176. DOI: 10.1002/cmdc.201600509.

I have collected the diffraction datasets from CA crystals, solved their structures and deposited their into PDB (5E2N, 5DOG, 5DOH, 5DRS, 5E2M, ir 5EHE). I have participated in the manuscript writting and data analysis.

(4) Zubrienė A., Smirnovienė J., **Smirnov A.**, Vaida Morkūnaitė, Michailovienė V., Jachno J., Juozapaitienė V., Norvaišas P., Manakova E., Gražulis S., Matulis D. 2015. Intrinsic thermodynamics of 4-substituted-2,3,5,6-tetrafluorobenzenesulfonamide binding to carbonic anhydrases by isothermal titration calorimetry. *Biophys Chem* 205:51–65. DOI: 10.1016/j.bpc.2015.05.009.

I have collected the diffraction datasets from CA crystals, solved their structures and deposit their into PDB (4WR7, 4WUP, 4WUQ, 4WW6, ir 4WW8). I have participated in the manuscript writting and data analysis.

(5) Dudutienė V, Zubrienė A, **Smirnov A**, Timm DD, Smirnovienė J, Kazokaitė J, Michailovienė V, Zakšauskas A, Manakova E, Gražulis S and Matulis D. Functionalization of Fluorinated Benzenesulfonamides and Their Inhibitory Properties toward Carbonic Anhydrases. *ChemMedChem* 2015, 10: 662-687.

The CA II, CA XII, and CA XIII isoforms crystallization, the collection of diffraction datasets and

their processing. The crystal structures of good quality were deposited into PDB database (4QIY, 4QIZ, 4QJ0, 4QJM, 4QJO, 4QJP, 4QJW, 4QJX, 4QTL). I have participated in the manuscript writting and data analysis.

Publications not included in this thesis:

(1) Čapkauskaitė E., Linkuvienė V., **Smirnov A.**, Milinavičiūtė G., Timm DD., Kasiliauskaitė A., Manakova E., Gražulis S., Matulis D. 2017. Combinatorial Design of Isoform-Selective N-Alkylated Benzimidazole-Based Inhibitors of Carbonic Anhydrases. *ChemistrySelect* 2:5360–5371. DOI: 10.1002/slct.201700531.

(2) Mickevičiūtė A., Timm DD., Gedgaudas M., Linkuvienė V., Chen Z., Waheed A., Michailovienė V., Zubrienė A., **Smirnov A.**, Čapkauskaitė E., Baranauskienė L., Jachno J., Revuckienė J., Manakova E., Gražulis S., Matulienė J., Cera ED., Sly WS., Matulis D. 2017. Intrinsic thermodynamics of high affinity inhibitor binding to recombinant human carbonic anhydrase IV. *Eur. Biophys. J.* DOI: 10.1007/s00249-017-1256-0.

(3) Kišonaitė M, Zubrienė A, Capkauskaitė E, **Smirnov A**, Smirnovienė J, Kairys V, Michailovienė V, Manakova E, Gražulis S and Matulis D. Intrinsic Thermodynamics and Structure Correlation of Benzenesulfonamides with a Pyrimidine Moiety Binding to Carbonic Anhydrases I, II, VII, XII, and XIII. *PLoS One* 2014, 9: e114106.

(4) Rutkauskas K, Zubrienė A, Tumosienė I, Kantminienė K, Kažemėkaitė M, **Smirnov A**, Kazokaitė J, Morkūnaitė V, Capkauskaitė E, Manakova E, Gražulis S, Beresnevičius ZJ and Matulis D. 4-Amino-substituted Benzenesulfonamides as Inhibitors of Human Carbonic Anhydrases. *Molecules* 2014, 19: 17356-17380.

(5) Capkauskaitė E, Zubrienė A, **Smirnov A**, Torresan J, Kišonaitė M, Kazokaitė J, Gylytė J, Michailovienė V, Jogaitė V, Manakova E, Gražulis S, Tumkevičius S and Matulis D. Benzenesulfonamides with pyrimidine moiety as inhibitors of human carbonic anhydrases I, II, VI, VII, XII, and XIII. *Bioorg Med Chem* 2013, 21: 6937-6947.

(6) Dudutienė V, Zubrienė A, **Smirnov A**, Gylytė J, Timm D, Manakova E, Gražulis S and Matulis D. 4-Substituted-2,3,5,6-tetrafluorobenzenesulfonamides as inhibitors of carbonic anhydrases I, II, VII, XII and XIII. *Bioorg Med Chem* 2013, 21: 2093-2106.

List of conferences:

(1) **Smirnov, A.**, Manakova, E., Zubrienė, A., Čapkauskaitė, E., Dudutienė, V., Matulis, D. "Correlation of binding thermodynamics with crystal structures in drug design". RICT 2017 - 53rd International Conference on Medicinal Chemistry - Drug Discovery & Selection. Toulouse, France. 2017 07 5 – 7. (poster presentation).

(2) **Smirnov**, A., Manakova, E., Zubrienė, A., Čapkauskaitė, E., Dudutienė, V., Matulis, D. "Crystallography and thermodynamics of primary sulfonamide inhibitors complexed with human carbonic anhydrases I, II, XII and XIII". XIVth International Conference of Lithuanian Biochemical Society, Druskininkai, Lithuania. 2016 06 27 – 30 (poster presentation).

(3) Linkuvienė, V., Zubrienė, A., Paketurytė, V., **Smirnov, A.**, Petrauskas, V., Matulis, D. Database of CA Protein-Ligand Binding Gibbs Energies, Enthalpies, Entropies and Crystal Structures. "ARBRE-MOBIEU ant COST Action CA 15126". Warsaw, Poland. 2018 03 19-21.

(4) Dudutienė, V., Matulienė, J., **Smirnov, A.**, Timm, D., Zubrienė, A., Baranauskienė, L., Morkunaitė, V., Smirnovienė, J., Michailovienė, V., Juozapaitienė, V., Mickevičiūtė, A., Kazokaitė, J., Bakšytė, S., Kasiliauskaitė, A., Jachno, J., Revuckienė, J., Kišonaitė, M., Pilipuitytė, V., Ivanauskaitė, E., Milinavičiūtė, G., Smirnovas, V., Petrikaitė, V., Kairys, V., Petrauskas, V., Norvaišas, P., Lingė, D., Gibieža, P., Čapkauskaitė, E., Zakšauskas, A., Kazlauskas, E., Manakova, E., Gražulis, S., Ladbury, J.E., Matulis, D. "Discovery and Characterization of Novel Selective Inhibitors of Carbonic Anhydrase IX". "10th International Carbonic Anhydrase Conference". Maastricht, Netherlands. 2015 04 19-22.

(5) Matulis, D., Zubrienė, A., Baranauskienė, L., **Smirnov, A.**, Morkunaitė, V., Smirnovienė, J., Kišonaitė, M., Norvaišas, P., Timm, D. „Intrinsic thermodynamics-structure correlation of carbonic anhydrase inhibitors". „Biophysical Society 59th Annual Meeting". Baltimore, USA. 2015 02 06-12.

(6) Gylytė, J., Zubrienė, A., Dudutienė, **V., Smirnov, A.**, Timm, D., Manakova, E., Gražulis, S., Matulis, D. „Intrinsic thermodynamics-structure correlations of fluorinated benzensulfonamides as inhibitors of human carbonic anhydrases". „Biophysical Society 58th Annual Meeting". San Francisco, USA. 2014 02 14-25.

(7) Gylytė, J., Zubrienė, A., Dudutienė, V., **Smirnov, A.**, Timm, D., Manakova, E., Gražulis, S., Matulis, D. „Intrinsic thermodynamics-structure correlations of fluorinated benzensulfonamides as inhibitors of human carbonic anhydrases". „The Twenty-seventh Annual Gibbs Conference on Biothermodynamics". Carbondale, USA. 2013 10 05-08.

(8) Justina Kazokaitė, Goda Milinavičiūtė, Joana Gylytė, Virginija Dudutienė, **Alexey Smirnov**, Jurgita Matulienė, and Daumantas Matulis „Differences in Stability Profiles and Thermodynamics of Inhibitor Binding to Target Protein Purified from E. coli, Mammalian Cells, and Human Saliva". European Biotehnology Congress 2014, Leece, Italy.

Santrauka

Vaistų kūrimas negali apsieiti be šiuolaikinių biofizikinių metodų, kurie padeda detaliai analizuoti sąveikas tarp baltymų ir ligandų. Baltymų kristalografija yra vienas iš plačiausiai naudojamų struktūrinių metodų, kurių pagalba galima nustatyti atominę baltymo makromolekulės struktūrą ir gauti baltymo-ligando komplekso struktūros trijų matmenų (3D) vaizdą. Nustatytos struktūros yra įrodymas, kad tiriamas ligandas jungiasi baltymo aktyviajame centre. Be to, ligando ir baltymo sąveikos struktūrinė informacija yra labai svarbi atrankaus jungimosi mechanizmo nustatymui.

Šio darbo tikslas buvo **1)** nustatyti keleto CA izoformų kompleksų su eile sulfonamidinių slopiklių kristalines struktūras, naudojant Rentgeno struktūrinės kristalografijos metodą, **2)** išspręstose kompleksų struktūrose išnagrinėti ligandų sąveikas su CA baltymu ir nustatyti sąveikos mechanizmus, kurie lemia atrankų bei stiprų jungimąsi prie CA izoformų bei **3)** ieškoti jungimosi struktūros-termodinaminių parametrų koreliacijų.

Buvo spręstos CA izoformų ir CA II mutantinių baltymų kompleksų su naujai susintetintais aromatiniais sulfonamidais kristalinės struktūros (iš viso 181). Šiame darbe buvo publikuota 61 kristalinė struktūra PDB duomenų bazėje. Gautos kompleksų kristalinės struktūros buvo detaliai ištirtos. Nustatyti sąveikų tipai ir tiriamų ligandų jungimosi CA aktyviajame centre būdai. Pasiūlyti ligandų sąveikos su taikiniu mechanizmai, kurie paaiškina jungimosi giminingumą.

Tiriant jų kompleksų kristalines struktūras, buvo paaiškintas atrankių CA IX ligandų jungimasis. Atrankaus CA IX ligando **VD11-4-2** *meta*-padėties ciklooktilo žiedas sudaro sąveiką su CA IX aktyviojo centro hidrofobine dalimi, kas CA I, CA II ir CA XIII aktyviuosiuose centruose nėra įmanoma dėl struktūrinių skirtumų tarp aktyviųjų centrų.

Buvo nustatytos ir aprašytos jungimosi termodinaminių parametrų pokyčių tendencijos struktūriškai panašiems junginiams. Tai yra svarbu naujų junginių su norimomis savybėmis kūrimui.

Curriculum Vitae

BIOGRAPHICAL SKETCH

Name **Alexey Smirnov**

Department of Biothermodynamics and Drug Design
Institute of Biotechnology
Vilnius university
Sauletekio 7, LT-10257, Vilnius

alexeyus1@gmail.com

EDUCATION

INSTITUTION	DEGREE	YEAR(s)	FIELD OF STUDY
Institute of Biotechnology Vilnius University	Ph.D. Student	2013 - 2017	Chemical Engineering
Vilnius University	M.S.	2011-2013	Biochemistry
Vilnius University	B.S.	2004-2008	Biochemistry

Thesis title of Bachelor degree in Biochemistry : Assessment of Micromycetes Resistance to Heavy Metals and Investigation of their Accumulation Abilities.
Thesis title of Master degree in Biochemistry : X-ray Crystallographic Studies of Human Carbonic Anhydrase Isoforms II, XII and XIII in Complex with Inhibitors.

PROFESSIONAL EXPERIENCE

2008.01 – 2008.10, Institute of Botany, research assistant
2013.03 – 2015.09, Institute of Biotechnology, Vilnius university, research assistant
2017.01 – present, Institute of Biotechnology, Vilnius university, biology research assistant.

PROFESSIONAL SKILLS AND EXPERTISE

Applied X-ray Crystallography (the author of 57 crystal structures of human carbonic anhydrase isoforms bound with various inhibitors in the Protein DataBank: 4QIY,4QIZ,4QJ0,4QJM, 4QJO,4QJP,4QJW,4QJX,4QTL,4QSA,4QSB,4QSI,4QSJ,4PYX,4PYY,4Q0L,4PZH,4Q06,4Q08,4Q09, 4Q07,4Q6D,4Q6E,4LHI,4KNJ,4KNN,4KP5,4KNI,4KNM,4KP8,4HT2,4HU1,4HT0,4WR7,4WUP, 4WUQ,4WW6,4WW8,5E2N,5DOG,5DOH,5DRS,5E2M,5EHE,5IPZ,5LLH,5LLC,5LLE,5LLG, 5LLN,5LLP,5LLO,5MSA,5MSB,5LL4,5LL5,5LL9,5LLA).

SCIENTIFIC PUBLICATION (reverse chronological order, for period 2012-2018)

1. **Smirnov A.**, Zubrienė A., Manakova E., Gražulis S., Matulis D. 2018. Crystal structure correlations with the intrinsic thermodynamics of human carbonic anhydrase inhibitor binding. *PeerJ*.

2. Čapkauskaitė E., Linkuvienė V., **Smirnov A.**, Milinavičiūtė G., Timm DD., Kasiliauskaitė A., Manakova E., Gražulis S., Matulis D. 2017. Combinatorial Design of Isoform-Selective N-Alkylated Benzimidazole-Based Inhibitors of Carbonic Anhydrases. *ChemistrySelect* 2:5360–5371. DOI: 10.1002/slct.201700531.

3. Mickevičiūtė A., Timm DD., Gedgaudas M., Linkuvienė V., Chen Z., Waheed A., Michailovienė V., Zubrienė A., **Smirnov A**., Čapkauskaitė E., Baranauskienė L., Jachno J., Revuckienė J., Manakova E., Gražulis S., Matulienė J., Cera ED., Sly WS., Matulis D. 2017. Intrinsic thermodynamics of high affinity inhibitor binding to recombinant human carbonic anhydrase IV. *Eur. Biophys. J.* DOI: 10.1007/s00249-017-1256-0.

4. Zubrienė A., Smirnov A., Dudutienė V., Timm DD., Matulienė J., Michailovienė V., Zakšauskas A., Manakova E., Gražulis S., Matulis D. 2017. Intrinsic Thermodynamics and Structures of 2,4- and 3,4-Substituted Fluorinated Benzenesulfonamides Binding to Carbonic Anhydrases. *ChemMedChem* 12:161–176. DOI: 10.1002/cmdc.201600509.

5. Zubrienė A., Smirnovienė J., **Smirnov A**., Vaida Morkūnaitė, Michailovienė V., Jachno J., Juozapaitienė V., Norvaišas P., Manakova E., Gražulis S., Matulis D. 2015. Intrinsic thermodynamics of 4-substituted-2,3,5,6-tetrafluorobenzenesulfonamide binding to carbonic anhydrases by isothermal titration calorimetry. *Biophys Chem* 205:51–65. DOI: 10.1016/j.bpc.2015.05.009.

6. Dudutienė V, Zubrienė A, **Smirnov A**, Timm DD, Smirnovienė J, Kazokaitė J, Michailovienė V, Zakšauskas A, Manakova E, Gražulis S and Matulis D. Functionalization of Fluorinated Benzenesulfonamides and Their Inhibitory Properties toward Carbonic Anhydrases. *ChemMedChem* 2015, 10: 662-687.

7. Kišonaitė M, Zubrienė A, Capkauskaitė E, **Smirnov A**, Smirnovienė J, Kairys V, Michailovienė V, Manakova E, Gražulis S and Matulis D. Intrinsic Thermodynamics and Structure Correlation of Benzenesulfonamides with a Pyrimidine Moiety Binding to Carbonic Anhydrases I, II, VII, XII, and XIII. *PLoS One* 2014, 9: e114106.

8. Dudutienė V, Matulienė J, **Smirnov A**, Timm DD, Zubrienė A, Baranauskienė L, Morkūnaitė V, Smirnovienė J, Michailovienė V, Juozapaitienė V, Mickevičiūtė A, Kazokaitė J, Bakšytė S, Kasiliauskaitė A, Jachno J, Revuckienė J, Kišonaitė M, Pilipuitytė V, Ivanauskaitė E, Milinavičiūtė G, Smirnovas V, Petrikaitė V, Kairys V, Petrauskas V, Norvaišas P, Lingė D, Gibieža P, Capkauskaitė E, Zakšauskas A, Kazlauskas E, Manakova E, Gražulis S, Ladbury JE and Matulis D. Discovery and characterization of novel selective inhibitors of carbonic anhydrase IX. *J Med Chem* 2014, 57: 9435-9446.

9. Rutkauskas K, Zubrienė A, Tumosienė I, Kantminienė K, Kažemėkaitė M, **Smirnov A**, Kazokaitė J, Morkūnaitė V, Capkauskaitė E, Manakova E, Gražulis S, Beresnevičius ZJ and Matulis D. 4-Amino-substituted Benzenesulfonamides as Inhibitors of Human Carbonic Anhydrases. *Molecules* 2014, 19: 17356-17380.

10. Capkauskaitė E, Zubrienė A, **Smirnov A**, Torresan J, Kišonaitė M, Kazokaitė J, Gylytė J, Michailovienė V, Jogaitė V, Manakova E, Gražulis S, Tumkevičius S and Matulis D. Benzenesulfonamides with pyrimidine moiety as inhibitors of human carbonic anhydrases I, II, VI, VII, XII, and XIII. *Bioorg Med Chem* 2013, 21: 6937-6947.

11. Dudutienė V, Zubrienė A, **Smirnov A**, Gylytė J, Timm D, Manakova E, Gražulis S and Matulis D. 4-Substituted-2,3,5,6-tetrafluorobenzenesulfonamides as inhibitors of carbonic anhydrases I, II, VII, XII and XIII. *Bioorg Med Chem* 2013, 21: 2093-2106.

POSTER PRESENTATIONS AT MEETINGS AND CONFERENCES:

1. Smirnov, A., Manakova, E., Zubrienė, A., Čapkauskaitė, E., Dudutienė, V., Matulis, D. "*Correlation of binding thermodynamics with crystal structures in drug design*". RICT 2017 - 53rd International Conference on Medicinal Chemistry - Drug Discovery & Selection. Toulouse, France. 2017 07 5 – 7. (poster).

2. Smirnov, A., Manakova, E., Zubrienė, A., Čapkauskaitė, E., Dudutienė, V., Matulis, D. "Crystallography and thermodynamics of primary sulfonamide inhibitors complexed with human carbonic anhydrases I, II, XII and XIII". XIVth International Conference of Lithuanian Biochemical Society, Druskininkai, Lithuania. 2016 06 27 – 30 (poster).

References

[1] L. Baranauskienė and D. Matulis. Intrinsic thermodynamics of ethoxzolamide inhibitor binding to human carbonic anhydrase XIII. *BMC biophys.*, 5(1):12, 2012.

[2] T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, and A. G. W. Leslie. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica. Section D, Biological Crystallography*, 67(Pt 4):271–281, Apr. 2011.

[3] E. Čapkauskaitė, V. Linkuvienė, A. Smirnov, G. Milinavičiūtė, D. D. Timm, A. Kasiliauskaitė, E. Manakova, S. Gražulis, and D. Matulis. Combinatorial Design of Isoform-Selective N-Alkylated Benzimidazole-Based Inhibitors of Carbonic Anhydrases. *ChemistrySelect*, 2(19):5360–5371, July 2017.

[4] E. Čapkauskaitė, A. Zubrienė, A. Smirnov, J. Torresan, M. Kišonaitė, J. Kazokaitė, J. Gylytė, V. Michailovienė, V. Jogaitė, E. Manakova, S. Gražulis, S. Tumkevičius, and D. Matulis. Benzenesulfonamides with pyrimidine moiety as inhibitors of human carbonic anhydrases I, II, VI, VII, XII, and XIII. *Bioorganic & Medicinal Chemistry*, 21(22):6937–6947, Nov. 2013. 00007.

[5] V. Dudutienė, J. Matulienė, A. Smirnov, D. D. Timm, A. Zubrienė, L. Baranauskienė, V. Morkūnaite, J. Smirnovienė, V. Michailovienė, V. Juozapaitienė, A. Mickevičiūtė, J. Kazokaitė, S. Bakšytė, A. Kasiliauskaitė, J. Jachno, J. Revuckienė, M. Kišonaitė, V. Pilipuitytė, E. Ivanauskaitė, G. Milinavičiūtė, V. Smirnovas, V. Petrikaitė, V. Kairys, V. Petrauskas, P. Norvaišas, D. Lingė, P. Gibieža, E. Capkauskaitė, A. Zakšauskas, E. Kazlauskas, E. Manakova, S. Gražulis, J. E. Ladbury, and D. Matulis. Discovery and characterization of novel selective inhibitors of carbonic anhydrase IX. *Journal of Medicinal Chemistry*, 57(22):9435–9446, Nov. 2014.

[6] V. Dudutienė, A. Zubrienė, A. Smirnov, J. Gylytė, D. Timm, E. Manakova, S. Gražulis, and D. Matulis. 4-Substituted-2,3,5,6-tetrafluorobenzenesulfonamides as inhibitors of carbonic anhydrases I, II, VII, XII, and XIII. *Bioorganic & Medicinal Chemistry*, 21(7):2093–2106, Apr. 2013. 00010.

[7] V. Dudutienė, A. Zubrienė, A. Smirnov, D. D. Timm, J. Smirnovienė, J. Kazokaitė, V. Michailovienė, A. Zakšauskas, E. Manakova, S. Gražulis, and D. Matulis. Functionalization of Fluorinated Benzenesulfonamides and Their Inhibitory Properties toward Carbonic Anhydrases. *ChemMedChem*, 10(4):662–687, Apr. 2015. 00000.

[8] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan. Features and development of it Coot. *Acta Crystallographica Section D*, 66:486–501, 2010.

[9] J. M. Fox, K. Kang, M. Sastry, W. Sherman, B. Sankaran, P. H. Zwart, and G. M. Whitesides. Water-Restructuring Mutations Can Reverse the Thermodynamic Signature of Ligand Binding to Human Carbonic Anhydrase. *Angewandte Chemie (International Ed. in English)*, Mar. 2017.

[10] C. Genis, K. H. Sippel, N. Case, Wengang Cao, B. S. Avvaru, L. J. Tartaglia, Lakshmanan Govindasamy, C. Tu, M. Agbandje-McKenna, D. N. Silverman, C. J. Rosser, and R. McKenna. Design of a carbonic anhydrase IX active-site mimic to screen inhibitors for possible anticancer properties. *Biochemistry*, 48(6):1322–1331, Feb. 2009.

[11] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, and G. R. Hutchison. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J Cheminform*, 4(1):17, Aug. 2012.

[12] V. Jogaitė, A. Zubrienė, V. Michailovienė, J. Gylytė, V. Morkūnaitė, and D. Matulis. Characterization of human carbonic anhydrase XII stability and inhibitor binding. *Bioorganic & Medicinal Chemistry*, 21(6):1431–1436, Mar. 2013. 00014.

[13] W. Kabsch. XDS. *Acta Crystallographica Section D*, D66:125–132, 2010.

[14] M. Kišonaitė, A. Zubrienė, E. Čapkauskaitė, A. Smirnov, J. Smirnovienė, V. Kairys, V. Michailovienė, E. Manakova, S. Gražulis, and D. Matulis. Intrinsic Thermodynamics and Structure Correlation of Benzenesulfonamides with a Pyrimidine Moiety Binding to Carbonic Anhydrases I, II, VII, XII, and XIII. *PLoS ONE*, 9(12):e114106, Dec. 2014. 00000.

[15] G. Klebe. Applying thermodynamic profiling in lead finding and optimization. *Nature Reviews Drug Discovery*, 14(2):95–110, 2015.

[16] J. F. Krebs, C. A. Fierke, R. S. Alexander, and D. W. Christianson. Conformational mobility of His-64 in the Thr-200—-Ser mutant of human carbonic anhydrase II. *Biochemistry*, 30(38):9153–9160, Sept. 1991. PDB ID: 5CA2.

[17] J. F. Krebs, J. A. Ippolito, D. W. Christianson, and C. A. Fierke. Structural and functional importance of a conserved hydrogen bond network in human carbonic anhydrase II. *J. Biol. Chem.*, 268(36):27458–27466, Dec. 1993. PDB ID: 1CVA, 1CVB.

[18] V. M. Krishnamurthy, G. K. Kaufman, A. R. Urbach, I. Gitlin, K. L. Gudiksen, D. B. Weibel, and G. M. Whitesides. Carbonic Anhydrase as a Model for Biophysical and Physical-Organic Studies of Proteins and Protein-Ligand Binding. *Chem. Rev.*, 108(3):946–1051, Mar. 2008.

[19] Y. Lou, P. C. McDonald, A. Oloumi, S. Chia, C. Ostlund, A. Ahmadi, A. Kyle, U. auf dem Keller, S. Leung, D. Huntsman, B. Clarke, B. W. Sutherland, D. Waterhouse, M. Bally, C. Roskelley, C. M. Overall, A. Minchinton, F. Pacchiano, F. Carta, A. Scozzafava, N. Touisni, J.-Y. Winum, C. T. Supuran, and S. Dedhar. Targeting Tumor Hypoxia: Suppression of Breast Tumor Growth and Metastasis by Novel Carbonic Anhydrase IX Inhibitors. *Cancer Res.*, 71(9):3364–3376, May 2011.

[20] E. Masini, F. Carta, A. Scozzafava, and C. T. Supuran. Antiglaucoma carbonic anhydrase inhibitors: A patent review. *Expert Opin Ther Pat*, 23(6):705–716, 2013.

[21] A. Mickevičiūtė, D. D. Timm, M. Gedgaudas, V. Linkuvienė, Z. Chen, A. Waheed, V. Michailovienė, A. Zubrienė, A. Smirnov, E. Čapkauskaitė, L. Baranauskienė, J. Jachno, J. Revuckienė, E. Manakova, S. Gražulis, J. Matulienė, E. D. Cera, W. S. Sly, and D. Matulis. Intrinsic thermodynamics of high affinity inhibitor binding to recombinant human carbonic anhydrase IV. *Eur. Biophys. J.*, Oct. 2017.

[22] V. Morkūnaitė, J. Gylytė, A. Zubrienė, L. Baranauskienė, M. Kišonaitė, V. Michailovienė, V. Juozapaitienė, M. J. Todd, and D. Matulis. Intrinsic thermodynamics of sulfonamide inhibitor binding to human

carbonic anhydrases I and II. *Journal of Enzyme Inhibition and Medicinal Chemistry*, 30(2):204–211, Apr. 2015. 00000.

[23] P. Mujumdar, K. Teruya, K. F. Tonissen, D. Vullo, C. T. Supuran, T. S. Peat, and S.-A. Poulsen. An Unusual Natural Product Primary Sulfonamide: Synthesis, Carbonic Anhydrase Inhibition, and Protein X-ray Structures of Psammaplin C. *Journal of Medicinal Chemistry*, 59(11):5462–5470, June 2016.

[24] G. N. Murshudov, A. A. Vagin, and E. J. Dodson. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*, 53(Pt 3):240–55, May 1997.

[25] K. Olechnovič and Č. Venclovas. Voronota: A Fast and Reliable Tool for Computing the Vertices of the Voronoi Diagram of Atomic Balls. *Journal of Computational Chemistry*, 35:672–681, 2014.

[26] E. O. Pettersen, P. Ebbesen, R. G. Gieling, K. J. Williams, L. Dubois, P. Lambin, C. Ward, J. Meehan, I. H. Kunkler, S. P. Langdon, A. H. Ree, K. Flatmark, H. Lyng, M. J. Calzada, L. D. Peso, M. O. Landazuri, A. Görlach, H. Flamm, J. Kieninger, G. Urban, A. Weltin, D. C. Singleton, S. Haider, F. M. Buffa, A. L. Harris, A. Scozzafava, C. T. Supuran, I. Moser, G. Jobst, M. Busk, K. Toustrup, J. Overgaard, J. Alsner, J. Pouyssegur, J. Chiche, N. Mazure, I. Marchiq, S. Parks, A. Ahmed, M. Ashcroft, S. Pastorekova, Y. Cao, K. M. Rouschop, B. G. Wouters, M. Koritzinsky, H. Mujcic, and D. Cojocari. Targeting tumour hypoxia to prevent cancer metastasis. From biology, biosensing and technology to drug development: The METOXIA consortium. *J Enzyme Inhib Med Chem*, pages 1–33, Oct. 2014.

[27] V. Pilipuitytė and D. Matulis. Intrinsic thermodynamics of trifluoromethanesulfonamide and ethoxzolamide binding to human carbonic anhydrase VII: THERMODYNAMICS OF TFS AND EZA BINDING TO CA VII. *Journal of Molecular Recognition*, 28(3):166–172, Mar. 2015. 00000.

[28] M. A. Pinard, C. D. Boone, B. D. Rife, C. T. Supuran, and R. McKenna. Structural study of interaction between brinzolamide and dorzolamide inhibition of human carbonic anhydrases. *Bioorg Med Chem*, 21(22):7210–7215, Nov. 2013. PDB ID: 4M2R, 4M2U, 4M2W, 4M2V.

[29] M. A. Pinard, B. Mahon, and R. McKenna. Probing the Surface of Human Carbonic Anhydrase for Clues towards the Design of Isoform Specific Inhibitors. *BioMed Research International*, vol. 2015:1–15, 2015.

[30] K. Rutkauskas, A. Zubrienė, I. Tumosienė, K. Kantminienė, M. Kažemėkaitė, A. Smirnov, J. Kazokaitė, V. Morkūnaitė, E. Čapkauskaitė, E. Manakova, S. Gražulis, Z. Beresnevičius, and D. Matulis. 4-Amino-substituted Benzenesulfonamides as Inhibitors of Human Carbonic Anhydrases. *Molecules*, 19(11):17356–17380, Oct. 2014. 00000.

[31] K. H. Sippel, A. Stander, C. Tu, B. Venkatakrishnan, A. H. Robbins, M. Agbandje-McKenna, J. Fourie, A. M. Joubert, and R. McKenna. Characterization of Carbonic Anhydrase Isozyme Specific Inhibition by Sulfamated 2-Ethylestra Compounds. *Lett. Drug Des. Discovery*, 8:1–25, 2011. PDB ID: 3OIK, 3OIL, 3OIM, 3OKU, 3OKV.

[32] A. Smirnov, A. Zubrienė, E. Manakova, S. Gražulis, and D. Matulis. Crystal structure correlations with the intrinsic thermodynamics of human carbonic anhydrase inhibitor binding. *PeerJ*, 6:e4412, Feb. 2018.

[33] P. W. Snyder, M. R. Lockett, D. T. Moustakas, and G. M. Whitesides. Is it the shape of the cavity, or the shape of the water in the cavity? *The European Physical Journal Special Topics*, 223:853–891, 2013.

[34] E. Svastova and S. Pastorekova. Carbonic anhydrase IX: A hypoxia-controlled "catalyst" of cell migration. *Cell Adh Migr*, 7(2):226–231, 2013.

[35] P. W. Taylor, R. W. King, and A. S. V. Burgen. Influence of pH on the kinetics of complex formation between aromatic sulfonamides and human carbonic anhydrase. *Biochemistry*, 9(20):3894–3902, 1970.

[36] A. A. Vagin, R. A. Steiner, A. A. Lebedev, L. Potterton, S. McNicholas, F. Long, and G. N. Murshudov. REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr D Biol Crystallogr*, 60(Pt 12 Pt 1):2184–95, Dec. 2004.

[37] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, and K. S. Wilson. Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography*, 67(Pt 4):235–242, Apr. 2011.

[38] A. Zubrienė, A. Smirnov, V. Dudutienė, D. D. Timm, J. Matulienė, V. Michailovienė, A. Zakšauskas, E. Manakova, S. Gražulis, and D. Matulis. Intrinsic Thermodynamics and Structures of 2,4- and 3,4-Substituted Fluorinated Benzenesulfonamides Binding to Carbonic Anhydrases. *ChemMedChem*, 12(2):161–176, Jan. 2017.

[39] A. Zubrienė, J. Smirnovienė, A. Smirnov, Vaida Morkūnaitė, V. Michailovienė, J. Jachno, V. Juozapaitienė, P. Norvaišas, E. Manakova, S. Gražulis, and D. Matulis. Intrinsic thermodynamics of 4-substituted-2,3,5,6-tetrafluorobenzenesulfonamide binding to carbonic anhydrases by isothermal titration calorimetry. *Biophys Chem*, 205:51–65, June 2015.