

<https://doi.org/10.15388/vu.thesis.997>
<https://orcid.org/0000-0003-0919-7826>

VILNIUS UNIVERSITY

Vaida Paketurytė-Latvė

Structural Recognition in Carbonic Anhydrase–Sulfonamide Complexes: Isozyme and Ligand Effects on Binding Thermodynamics

DOCTORAL DISSERTATION

Natural Sciences,
Biochemistry (N 004)

VILNIUS 2026

This dissertation was written between 2019 and 2025 at Vilnius University.

Academic supervisor:

Prof. Dr. Daumantas Matulis (Vilnius University, Natural Sciences, Biochemistry – N 004).

Dissertation Defence Panel:

Chair – **Dr. Gintautas Tamulaitis** (Vilnius University, Natural Sciences, Biochemistry – N 004).

Members:

Prof. Dr. Dalia Barsyte-Lovejoy (University of Toronto, Canada, Natural Sciences, Biochemistry – N 004).

Dr. Laurynas Dagys (Vilnius University, Natural Sciences, Physics – N 002).

Dr. Kliment Olechnovič (Vilnius University, Natural Sciences, Biochemistry – N 004).

Dr. Miglė Tomkuvienė (Vilnius University, Natural Sciences, Biochemistry – N 004).

The dissertation shall be defended at a public meeting of the Dissertation Defence Panel at 2 p.m. on 29 October 2026 in room No. R401 of the Life Sciences Center, Vilnius University. Address: Saulėtekio av. 7, Vilnius, Lithuania; e-mail: vaida.paketuryte@gmc.vu.lt

<https://doi.org/10.15388/vu.thesis.997>
<https://orcid.org/0000-0003-0919-7826>

VILNIAUS UNIVERSITETAS

Vaida Paketurytė-Latvė

Struktūrinis atpažinimas
karboanhidrazių–sulfonamidų
kompleksuose: izofermentų ir
ligandų efektai jungimosi
termodinamikoje

DAKTARO DISERTACIJA

Gamtos mokslai,
biochemija (N 004)

VILNIUS 2026

Disertacija rengta 2019 – 2025 metais Vilniaus universitete.

Mokslinis vadovas –

prof. dr. Daumantas Matulis (Vilniaus universitetas, gamtos mokslai, biochemija – N 004).

Gynimo taryba:

Pirmininkas – **dr. Gintautas Tamulaitis** (Vilniaus universitetas, gamtos mokslai, biochemija – N 004).

Nariai:

prof. dr. Dalia Barsytė-Lovejoy (Toronto universitetas, Kanada, gamtos mokslai, biochemija – N 004).

dr. Laurynas Dagys (Vilniaus universitetas, gamtos mokslai, fizika – N 002).

dr. Kliment Olechnovič (Vilniaus universitetas, gamtos mokslai, biochemija – N 004).

dr. Miglė Tomkuvienė (Vilniaus universitetas, gamtos mokslai, biochemija – N 004).

Disertacija ginama viešame Gynimo tarybos posėdyje 2026 m. spalio 29 d. 14 val. Vilniaus universiteto Gyvybės mokslų centro R401 auditorijoje.

Adresas: Saulėtekio al. 7, Vilnius, Lietuva; el. paštas: vaida.paketuryte@gmc.vu.lt

Contents

Abbreviations and Variables	vii
Introduction	1
Scientific novelty	4
Defending statements	5
1 Literature Overview	6
1.1 Structure-Based Drug Design: Potential and Challenges .	6
1.2 Protein–Ligand Binding: Physical Basis and Thermodynamic Description	10
1.2.1 Interactions Governing Protein–Ligand Binding .	10
Interactions Governing Protein–Ligand Binding	10
1.2.2 Thermodynamic Description of Protein–Ligand Binding	16
Thermodynamic Description of Protein–Ligand Binding	16
1.3 Carbonic Anhydrases as a Model System	20
1.3.1 Structural Features of Carbonic Anhydrases . . .	20
1.3.2 Sulfonamides as Carbonic Anhydrase Inhibitors .	32
2 Methods	38
2.1 Materials	39
2.1.1 Carbonic Anhydrase Isozymes	39
2.1.2 Sulfonamide Inhibitors	40
2.2 Experimental Determination of Protein–Ligand Interactions	40
2.2.1 Fluorescent Thermal Shift Assay	41
2.2.2 Isothermal Titration Calorimetry	42
2.3 Determination of Protonation Parameters	42
2.3.1 Sulfonamide pK_a Measurements	42
2.3.2 Sulfonamide Protonation Enthalpy Measurements	43
2.4 Calculation of Intrinsic Thermodynamic Parameters . . .	43
2.4.1 Correction of Binding Constants	44
2.4.2 Correction of Binding Enthalpies and Entropies .	44
2.5 Structural Studies	45

3 Summary of Results	47
3.1 Determinants of Protein–Ligand Recognition	47
3.1.1 Enthalpy-Entropy Compensation	48
3.1.2 Affinity and Selectivity	50
3.1.3 Structure-Thermodynamics Relationships	53
3.2 Reliability of Thermodynamic Measurements	55
3.2.1 FTSA Standardization	55
3.2.2 Uncertainty Quantification	56
Discussion	59
Conclusions	63
Bibliography	64
Santrauka (Summary in Lithuanian)	81
Įvadas	82
Mokslinis naujumas	85
Ginamieji teiginiai	86
Literatūros apžvalga	86
Metodai	92
Rezultatai	100
Diskusija	111
Išvados	115
Acknowledgements	116
PUBLICATIONS INCLUDED IN THE DISSERTATION	118
Copies of publications	119
Enthalpy–Entropy Compensation (Publication 1)	120
Affinity and Selectivity (Publication 2)	146
Structure–Thermodynamics Relationships (Publication 3)	169
FTSA Standardization (Publication 4)	184
Uncertainty Quantification (Publication 5)	192
Publications not included in the dissertation	202
List of scientific events	207
Curriculum Vitae	210

Abbreviations and Variables

Abbreviations

- AZM — acetazolamide, a classical sulfonamide carbonic anhydrase inhibitor;
- CA — carbonic anhydrase. CA followed by a Roman numeral denotes a specific isozyme;
- DMSO — dimethyl sulfoxide;
- FTSA — fluorescence-based thermal shift assay (also called thermal shift assay, TSA), alternatively termed ThermoFluor® or differential scanning fluorimetry (DSF);
- GPI — glycosylphosphatidylinositol;
- IC — intracellular domain/tail;
- ITC — isothermal titration calorimetry;
- PDB ID — Protein Data Bank identifier;
- PG — proteoglycan-like domain;
- PLBD — Protein–Ligand Binding Database (plbd.org);
- TM — transmembrane region;
- TRIS — tris(hydroxymethyl)aminomethane;
- Zn^{II} — zinc ion;

Terminology

- Intrinsic* — refers to binding between the protein and ligand in their binding-competent protonation states, with contributions from linked protonation equilibria accounted for. Intrinsic thermodynamic quantities and equilibrium constants are denoted by the subscript “int”. See Figure 1.12 for an illustration of its meaning and practical application;
- Observed* — refers to binding as measured experimentally under the specified solution conditions and therefore includes contributions from linked protonation equilibria. Observed thermodynamic quantities and equilibrium constants are denoted by the subscript “obs”.

Thermodynamic quantities

- ΔG — standard Gibbs energy change of binding (equivalent to ΔG°);
 ΔG_{int} — intrinsic standard Gibbs energy change of binding;
 ΔG_{obs} — observed standard Gibbs energy change of binding;
 ΔH — standard enthalpy change of binding (equivalent to ΔH°);
 ΔH_{int} — intrinsic standard enthalpy change of binding;
 ΔH_{obs} — observed standard enthalpy change of binding;
 ΔS — standard entropy change of binding (equivalent to ΔS°);
 ΔS_{int} — intrinsic standard entropy change of binding;
 ΔS_{obs} — observed standard entropy change of binding;
 $\Delta H_{\text{pr,buf}}$ — standard enthalpy change of buffer protonation;
 $\Delta H_{\text{pr,CA}}$ — standard enthalpy change of CA–Zn^{II} bound hydroxide protonation;
 $\Delta H_{\text{pr,SA}}$ — standard enthalpy change of protonation of deprotonated sulfonamide amino group;

Equilibrium constants

- K_{b} — equilibrium binding constant;
 $K_{\text{b,int}}$ — intrinsic equilibrium binding constant;
 $K_{\text{b,obs}}$ — observed equilibrium binding constant;
 K_{d} — equilibrium dissociation constant;
 $K_{\text{d,int}}$ — intrinsic equilibrium dissociation constant;
 $K_{\text{d,obs}}$ — observed equilibrium dissociation constant;
 $\text{p}K_{\text{a}}$ — negative decimal logarithm of the acid dissociation constant;
 $\text{p}K_{\text{a,CA}}$ — negative decimal logarithm of the acid dissociation constant of the Zn^{II}-bound water molecule in carbonic anhydrase;
 $\text{p}K_{\text{a,SA}}$ — negative decimal logarithm of the acid dissociation constant of the protonated sulfonamide group;

Other variables and constants

- R — gas constant (8.314 J mol^{−1} K^{−1});
 T — absolute temperature (K);

Introduction

Drug discovery relies on identifying molecules that bind selectively and strongly to their biological targets. Advances in structural biology, biophysics, and computational chemistry have greatly expanded our ability to characterize protein–ligand interactions and have become essential components of modern drug design [1, 2]. High-resolution structures, molecular simulations, and structure-based drug design approaches provide valuable insights into binding modes and guide the optimization of lead compounds [3]. Nevertheless, accurately predicting how structural modifications influence binding affinity and selectivity remains a major challenge [4].

Protein–ligand recognition arises from a complex interplay of intermolecular interactions, hydration, protonation equilibria, and conformational dynamics. Although these factors collectively determine binding thermodynamics, their individual contributions are often difficult to isolate and quantify. As a result, molecules with similar structures may exhibit unexpectedly different binding properties, whereas comparable affinities can arise through fundamentally different energetic mechanisms. Understanding these relationships remains an important objective for both experimental and computational studies of molecular recognition.

Several factors are particularly important for the interpretation and prediction of protein–ligand binding:

- **Enthalpy–Entropy Compensation and Hydration Networks:** Binding affinity is determined by the balance between enthalpic and entropic contributions, which frequently compensate each other [5]. Similar binding affinities can therefore arise from markedly different thermodynamic signatures. Such differences are often associated with changes in hydration networks, which remain challenging to characterize and predict quantitatively [6–9].

- **Conformational Effects:** Small chemical modifications can alter ligand geometry, flexibility, and intramolecular interactions, thereby influencing the energetic cost of adopting a binding-competent conformation. These effects may substantially influence protein–ligand interactions [6, 10].
- **Observed versus Intrinsic Thermodynamics:** Experimentally measured binding parameters often include contributions from linked protonation reactions and buffer-dependent processes [11, 12]. Consequently, observed thermodynamic quantities do not always directly reflect the intrinsic energetics of molecular recognition, complicating mechanistic interpretation and comparison between systems.

Together, these factors illustrate that molecular recognition cannot be fully understood from affinity measurements or static structures alone. A suitable model system is therefore required to systematically investigate how structural, thermodynamic, and environmental factors contribute to binding. Carbonic anhydrases (CA) provide an excellent model system for investigating these challenges [13]. Although their catalytic architecture is highly conserved, subtle differences in active-site composition and hydration-network organization can produce substantial differences in ligand recognition. Sulfonamide inhibitors are particularly useful probes because they share a common zinc-binding group while allowing systematic modification of substituents that interact with distinct regions of the active site. This combination enables detailed investigation of how structural changes influence binding thermodynamics, affinity, and selectivity.

Reliable interpretation of structure–thermodynamics relationships requires not only mechanistic understanding but also robust treatment of experimental uncertainty, particularly when comparing parameters derived from non-linear binding models. The central hypothesis of this dissertation is that meaningful relationships between protein–ligand structure and thermodynamics require the systematic separation and characterization of the underlying processes that contribute to molecular recognition.

Goal:

To identify and quantify the structural, thermodynamic, and environmental factors that govern protein–ligand recognition using carbonic anhydrases and primary sulfonamide inhibitors as a model system.

Objectives:

1. To determine how subtle differences in active-site residues, ligand structure, and hydration networks, modulate the balance between enthalpic and entropic contributions to binding. (Publication 1)
2. To investigate the impact of minor ligand modifications on conformational effects, including intramolecular hydrogen bonding and conformational energetics, and their role in determining binding affinity and selectivity. (Publication 2)
3. To systematically separate protonation-linked contributions from experimentally observed thermodynamic parameters to derive intrinsic thermodynamic parameters. (Publication 3)
4. To standardize FTSA experimental protocols for robust affinity determination and to implement asymmetric profile-likelihood confidence intervals for the accurate quantification and interpretation of uncertainty in non-linear binding models. (Publication 4 and Publication 5)

Scientific novelty

The scientific novelty of this dissertation lies in its integrated approach to investigating protein–ligand recognition by combining high-resolution structural analysis, intrinsic binding thermodynamics, and rigorous uncertainty evaluation. This work goes beyond conventional affinity measurements to demonstrate how ligand structure, protein environment, hydration networks, and proton-linked equilibria collectively shape the energetic landscape of molecular recognition.

Key novel contributions:

- **Identification of the structural basis for divergent energetics:** demonstrated that closely related CAs (CA II, CA XII, and CA XIII) achieve comparable binding affinities through fundamentally opposing enthalpic–entropic mechanisms. These distinct thermodynamic fingerprints are modulated by subtle differences in active-site residues and the reorganization of conserved hydration networks. (Publication 1)
- **Insight into the energetic cost of ligand conformation:** identified that a single-atom oxidation change (from sulfanyl to sulfinyl and sulfonyl) altered binding affinity by up to a million-fold. This drop in affinity is linked to the formation of unfavorable intramolecular hydrogen bonds that distort ligand conformation and create steric hindrance. (Publication 2)
- **Revealing the limitations of observed thermodynamic parameters:** demonstrated that binding-linked protonation equilibria substantially alter experimentally observed thermodynamic parameters. Dissection of these contributions enabled meaningful correlations between atomic-resolution structures and the underlying energies of molecular recognition. (Publication 3)
- **Consistent representation of uncertainty across thermodynamic quantities:** emphasized that conventional symmetric error estimates lead to misleading interpretations when binding affinities are transformed between dissociation constants and Gibbs energies. In contrast, asymmetric profile-likelihood confidence intervals preserved the underlying uncertainty structure and ensured consistent interpretation across different representations of binding affinity. (Publication 4 and Publication 5)

Defending statements

- Subtle structural differences in proteins and ligands can produce disproportionately large changes in protein–ligand recognition and selectivity because binding is governed by the entire molecular environment, including protonation equilibria, hydration, conformational effects, and intermolecular interactions. (Publication 1, Publication 2, and Publication 3)
- Protein–ligand recognition cannot be fully understood from structural information and observed thermodynamic parameters alone; reliable interpretation requires the integration of atomic-level structural features with intrinsic thermodynamic parameters. (Publication 1, Publication 2, and Publication 3)
- The representation of uncertainty influences the interpretation of protein–ligand binding measurements; asymmetric profile-likelihood confidence intervals provide a consistent description of uncertainty, while conventional symmetric standard errors are inappropriate. (Publication 4 and Publication 5)

Chapter 1

Literature Overview

1.1 Structure-Based Drug Design: Potential and Challenges

The discovery and development of new medicines remains a costly and high-risk process. Although estimates vary considerably, the median capitalized research and development investment per new therapeutic agent has been estimated at approximately US \$1 billion, including costs associated with failed development programs [14]. A major challenge at the discovery stage is the identification and optimization of compounds with sufficient potency, selectivity, and other properties required for further development. Large numbers of initially considered compounds are progressively reduced to only a few candidates that enter clinical development (Figure 1.1). Improving the efficiency of early compound identification and optimization is therefore an important objective of modern drug discovery.

Drug discovery encompasses both biological products and small molecules, with small molecules accounting for approximately 65 % of new drug approvals in 2025 [17]. Small molecules are amenable to systematic chemical modification and optimization against diverse molecular targets. Their discovery may proceed through several complementary strategies (Table 1.1), ranging from phenotypic and target-based screening to approaches exploiting prior knowledge of active ligands or three-dimensional target structures [18, 19]. The research presented in this dissertation is centered on structure-based drug design, utilizing high-resolution structures of carbonic anhydrases integrated with

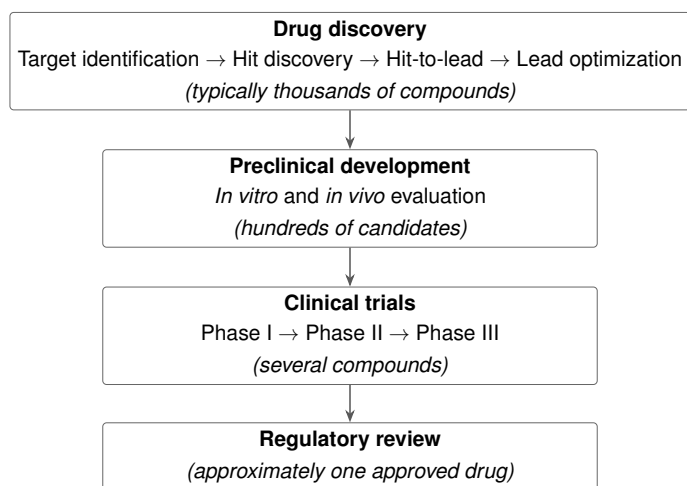


Figure 1.1: Simplified representation of the small-molecule drug discovery and development process, illustrating the substantial reduction in candidate numbers from thousands of compounds during discovery to approximately one approved drug over a typical period of 10–15 years [15]. Even after entry into clinical development, attrition remains substantial: an analysis of 3,999 drug candidates reported an overall probability of marketing approval of 12.8 % [16]. The research presented in this dissertation focuses on the first stage of this process (Drug discovery), specifically on hit-to-lead and lead optimization phases.

binding thermodynamics to guide the systematic optimization of ligand affinity and selectivity.

Structure-based drug design (SBDD) represents one such approach and exploits three-dimensional information about a biological target to guide ligand discovery and optimization. The importance of structural information in modern drug discovery is illustrated by an analysis of 34 low-molecular-weight, protein-targeted anticancer drugs approved by the US FDA between 2019 and 2023. Structural information contributed to the discovery or development of all 34 drugs, while structure-guided drug discovery contributed to more than 80 % of cases, either confirmed or probable [21]. Experimentally determined structures can reveal the architecture of binding sites, ligand binding modes, and specific protein–ligand interactions. Protein–ligand structures can therefore guide chemical modification by identifying interactions and binding-site regions that may be exploited during optimization. These experimental data can also be combined with computational approaches to screen compound libraries, generate or evaluate ligand poses, investigate molecular

Table 1.1: Major approaches to small-molecule discovery and design [20].

Approach	Starting point	Typical methods
Phenotypic screening	Desired biological phenotype	Cell- or organism-based screening
Target-based screening	Defined molecular target	High-throughput biochemical or biophysical screening
Ligand-based drug design	Known active ligands	Similarity searching, pharmacophore modeling, quantitative structure–activity relationships
Fragment-based drug discovery	Small, low-complexity fragments	Biophysical or structural screening followed by fragment growing, linking, or merging
Structure-based drug design	Three-dimensional target structure	Molecular docking, structure-based virtual screening, structure-guided ligand optimization

dynamics, and estimate the energetic consequences of chemical modifications [20, 22].

The particular value of SBDD lies in its ability to relate chemical modifications of a ligand to specific features of the target binding site, providing a structural basis for iterative ligand optimization. However, the presence of a favorable interaction in a protein–ligand structure does not directly establish its contribution to binding affinity. A chemical modification intended to introduce or strengthen a particular interaction may simultaneously alter ligand and protein conformation, desolvation, solvent organization, protonation equilibria, and the entropic cost of binding. The energetic consequences of structural modifications are therefore determined by multiple coupled processes and cannot generally be inferred from individual structural interactions alone.

Computational methods complement experimental structures by addressing different aspects of molecular recognition, ranging from ligand-pose prediction and virtual screening to molecular dynamics and binding-energy calculations (Table 1.2). However, these approaches differ substantially in physical detail, computational cost, and their

ability to predict binding affinity quantitatively [20].

Table 1.2: Major computational approaches used in protein–ligand modeling and binding-affinity prediction [20]. The diverse methods summarized here illustrate the various strategies employed to model molecular recognition, while their inherent limitations underscore the physical and computational challenges that complicate the accurate prediction of binding affinity in drug design.

Method	Main application	Major limitation
Molecular docking	Generation of plausible protein–ligand binding poses and scoring of protein–ligand interactions	Binding-affinity prediction depends strongly on the scoring function and treatment of the molecular system
Molecular dynamics	Sampling of protein–ligand conformations, solvent, and ion dynamics; estimation of binding free energies using additional sampling or analysis approaches	Conventional MD does not directly provide binding affinity; enhanced sampling may suffer from slow convergence and sampling limitations
Quantum mechanics	Estimation of protein–ligand interactions and binding affinities using electronic-structure calculations	High computational cost, particularly for large molecular systems
Machine learning	Prediction of binding poses and binding affinity from relationships learned from experimental and structural data	Performance and applicability depend on training-data quality and model validation

Despite major advances in structural biology and computational chemistry, accurate prediction of protein–ligand affinity and selectivity remains challenging [23–25]. A central goal of computational drug design is to predict binding affinity and selectivity rapidly and accurately, thereby enabling the identification of potent inhibitors while minimizing interactions with relevant off-targets. The research presented in this dissertation contributes to this goal by investigating the physical basis of protein–ligand recognition and relating molecular structure to binding thermodynamics, affinity, and selectivity.

1.2 Protein–Ligand Binding: Physical Basis and Thermodynamic Description

Protein–ligand binding arises from a complex interplay of hydrogen bonding, electrostatic interactions, hydrophobic effects, van der Waals contacts, conformational dynamics, and solvent-mediated processes. These contributions are highly interdependent, making molecular recognition difficult to interpret from static structural information alone. Consequently, thermodynamic and dynamic analyses are essential for understanding the energetic basis of binding [26, 27].

Human carbonic anhydrases and primary benzenesulfonamide inhibitors are used throughout this dissertation as a model system for investigating these relationships (Figure 1.2). The conserved binding of the primary sulfonamide group to the catalytic zinc ion provides a common structural anchor, while modifications of the benzenesulfonamide scaffold and differences between carbonic anhydrase isozymes allow the effects of relatively small structural changes on affinity, selectivity, and binding thermodynamics to be examined. Examples from this model system are therefore used throughout the following sections to illustrate general principles of protein–ligand recognition. The carbonic anhydrase family and its inhibition are described in detail in Section 1.3.

1.2.1 Interactions Governing Protein–Ligand Binding

Metal Coordination

Metal coordination is a distinct class of protein–ligand interaction in which a ligand directly interacts with a metal ion located within a protein binding site [28]. Compared with conventional noncovalent interactions, metal coordination exhibits mixed electrostatic and covalent character and is generally stronger and more directional than hydrogen bonds or van der Waals contacts [29, 30].

Catalytically active human carbonic anhydrases provide a well-characterized example of metal-assisted molecular recognition. Their active sites contain a catalytic Zn^{II} ion [13, 31]. Sulfonamide inhibitors bind through direct coordination of the deprotonated sulfonamide nitrogen to the Zn^{II} ion, displacing the zinc-bound water molecule and forming a tetrahedral complex. Because only the deprotonated sulfonamide

species can coordinate the catalytic metal ion, inhibitor binding is intrinsically linked to protonation equilibria [11]. Crystal structures reveal conserved Zn–N distances of approximately 1.9 Å to 2.2 Å and a hydrogen-bond network involving Thr199 [32–34].

The importance of this interaction is demonstrated by apo-carbonic anhydrases, where removal of the catalytic zinc ion decreases sulfonamide affinity by approximately $(10^4) - (10^6)$ -fold, corresponding to a loss of 25 kJ mol^{-1} to 35 kJ mol^{-1} in binding energy [35]. Thermodynamic analyses further suggest that 60 % to 65 % of the stabilization

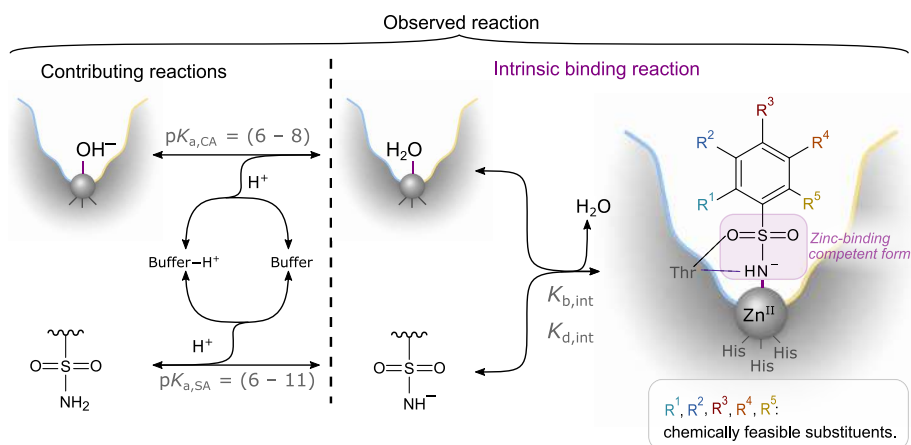


Figure 1.2: General scheme of carbonic anhydrase–primary sulfonamide recognition. The scheme summarizes the common features of primary sulfonamide (emphasis on benzenesulfonamides, which are the main focus of this work) binding to the catalytically active human carbonic anhydrase isozymes and introduces concepts discussed in the following sections. Binding is coupled to the protonation equilibria of both partners, with approximate pK_a values indicated. The conserved zinc-binding motif involves coordination of the sulfonamide nitrogen to the catalytic Zn^{II} ion and hydrogen-bonding interactions with Thr199. Located at the interface between the more hydrophobic and more hydrophilic sides of the active site (yellow and blue, respectively), this motif provides a common structural anchor. Beyond this motif, variation in sulfonamide substituents and the active-site environments of individual CA isozymes gives rise to diverse interaction patterns affecting affinity, selectivity, and binding thermodynamics. Experimentally observed binding additionally includes contributions from linked protonation equilibria; accounting for these enables determination of intrinsic binding parameters that directly reflect protein–ligand interactions.

energy of carbonic anhydrase–sulfonamide complexes originates from direct interaction with the catalytic Zn^{II} ion [13]. Thus, metal coordination serves as the primary anchoring interaction, while secondary interactions and solvent effects fine-tune affinity and selectivity.

Desolvation and Water-Network Reorganization

Protein–ligand binding occurs in aqueous solution and therefore involves not only direct protein–ligand interactions but also reorganization of surrounding solvent molecules and dissolved ions [7, 9, 36–38]. Both the protein and ligand are surrounded by hydration shells that stabilize exposed polar, charged, and hydrophobic groups. Complex formation requires partial removal and rearrangement of these solvent molecules, creating a desolvation penalty that must be compensated by favorable protein–ligand interactions [39]. Consequently, the energetic contribution of any protein–ligand interaction cannot be evaluated independently of the protein–water and ligand–water interactions displaced upon binding [37–39]. Water molecules within binding sites often differ substantially from bulk solvent: displacement of energetically unfavorable waters can enhance affinity, whereas retention of structurally important waters may stabilize hydrogen-bond networks and contribute to molecular recognition [38, 40, 41].

Carbonic anhydrases provide a notable example of the importance of solvent organization. Their active sites contain an extended hydrogen-bond network connecting the catalytic Zn^{II} ion, the zinc-bound water molecule, Thr199, His64, and several ordered water molecules involved in proton transfer [31, 42]. Binding of sulfonamide inhibitors displaces the zinc-bound water molecule and reorganizes the surrounding solvent network, contributing substantially to binding thermodynamics and enthalpy–entropy compensation [13, 43, 44]. Most current knowledge of active-site water organization in carbonic anhydrases originates from studies of CA II, which revealed a highly ordered network of water molecules involved in proton transfer and stabilization of the catalytic center [43, 45–51].

In addition to water, dissolved ions such as sodium, potassium, chloride, phosphate, and sulfate can influence molecular recognition through electrostatic screening and ion displacement. Although often absent from crystallographic models, these species contribute to experimentally measured binding thermodynamics and complicate interpretation of protein–ligand interactions [52, 53].

Hydrophobic Effect

The hydrophobic effect is one of the major driving forces of protein–ligand binding in aqueous environments. Unlike electrostatic interactions, hydrogen bonds, or metal coordination, it does not arise from a direct attractive force between nonpolar groups. Instead, it originates from the thermodynamic consequences of water reorganization around hydrophobic surfaces [37, 38, 54–56].

In aqueous solution, water molecules surrounding nonpolar groups adopt a more ordered arrangement than in bulk solvent. Upon ligand binding, burial of hydrophobic surfaces releases these constrained water molecules back into the bulk phase, resulting in a favorable increase in solvent entropy. Consequently, the hydrophobic effect often contributes substantially to the energy of binding and generally scales with the amount of nonpolar surface area buried upon complex formation [54, 57].

Hydrophobic interactions are closely coupled to desolvation and water-network reorganization. Therefore, their energetic contribution cannot always be separated from the thermodynamic effects associated with displacement of water molecules from the binding interface [37, 38]. In structure-based drug design, optimization of hydrophobic contacts is frequently used to enhance ligand affinity, although excessive hydrophobicity may compromise solubility and pharmacokinetic properties [30, 38].

In carbonic anhydrases, hydrophobic interactions contribute significantly to inhibitor affinity and isozyme selectivity. The hydrophobic region of the active site, formed primarily by residues Val121, Val143, Leu198, and Trp209 (CA II numbering), plays an important role in accommodation of aromatic inhibitor substituents. While the Zn^{II} coordination is highly conserved among isozymes, differences in hydrophobic regions surrounding the active-site cavity provide opportunities for selective recognition. Aromatic and fluorinated [58–60] substituents of sulfonamide inhibitors frequently occupy these regions, forming favorable contacts and displacing ordered water molecules, thereby enhancing binding affinity and selectivity [13, 34].

Hydrogen-Bond Networks

Hydrogen bonds are among the most important noncovalent interactions involved in protein–ligand recognition. Owing to their strong directionality, they contribute to ligand affinity, specificity, and orientation within the binding site [30, 61]. Although the energetic contribution of an individual hydrogen bond is typically estimated to range from 4 kJ mol^{-1} to 20 kJ mol^{-1} , its net stabilizing effect depends strongly on geometry, solvent exposure, and desolvation effects [8, 30].

Hydrogen bonds frequently form cooperative networks involving protein residues, the ligand, and ordered water molecules, thereby stabilizing protein–ligand complexes and maintaining binding-site architecture [13, 38].

In carbonic anhydrases, a conserved hydrogen-bond network surrounds the catalytic Zn^{II} ion and contributes to both catalysis and inhibitor binding. The residue Thr199 plays a central role by forming hydrogen bonds with the sulfonamide group of bound inhibitors while simultaneously linking the catalytic center to the surrounding solvent network [13, 34]. Together with Zn^{II} coordination, these conserved interactions are major determinants of the affinity of sulfonamide inhibitors.

Electrostatic Interactions

Electrostatic interactions arise from attractive or repulsive forces between charged and polar groups and represent one of the longest-range contributors to protein–ligand recognition. They include ion–ion, ion–dipole, and dipole–dipole interactions, all of which can influence ligand orientation and binding-site recognition before direct molecular contact is established [30, 61].

The strength of electrostatic interactions depends strongly on the local environment. In aqueous solution, solvent screening substantially weakens interactions between charged groups, making their energetic contribution highly context dependent and closely coupled to desolvation effects [37, 38]. In buried binding sites, however, electrostatic interactions can contribute significantly to binding affinity owing to reduced dielectric screening.

In carbonic anhydrases, electrostatic interactions contribute to both catalysis and inhibitor binding. The catalytic Zn^{II} ion generates a strong

positive electrostatic potential that stabilizes the zinc-bound hydroxide ion and promotes binding of negatively charged ligands [13, 62]. In addition, interactions between charged or polar ligand substituents and residues lining the active-site cavity can influence ligand orientation and contribute to isozyme selectivity [34].

van der Waals and Packing Interactions

van der Waals interactions arise from transient fluctuations in electron density that generate temporary dipoles between neighboring atoms. Although individually weak, they are ubiquitous in protein–ligand complexes and can contribute substantially to binding through their cumulative effect [30, 61]. Because these interactions operate only over short distances, their contribution depends strongly on shape complementarity and efficient packing between the ligand and the protein binding site.

Optimal packing maximizes favorable contacts while avoiding steric clashes. Consequently, small structural modifications can markedly influence binding affinity by improving or disrupting local van der Waals interactions [30, 38]. In many cases, affinity improvements achieved during lead optimization arise from more efficient occupation of available space rather than the introduction of additional hydrogen bonds.

In carbonic anhydrases, van der Waals interactions stabilize inhibitors after primary anchoring interactions, such as Zn^{II} coordination and hydrogen bonding, have been established. Aromatic, heterocyclic, and fluorinated substituents frequently occupy hydrophobic regions of the active-site cavity, where favorable packing interactions enhance binding affinity [34, 63]. Because the architecture of the active-site entrance differs among isozymes, optimization of packing interactions in these regions is often exploited to improve inhibitor selectivity [63].

Conformational Effects

Protein–ligand binding is frequently accompanied by conformational changes in both the protein [64–66] and the ligand [10, 67]. Rather than existing as rigid structures, proteins populate dynamic ensembles of interconverting conformational states that differ in their ability to recognize and bind ligands [68, 69]. Molecular recognition is therefore

often associated with structural rearrangements that optimize protein–ligand interactions [70].

Two complementary models are commonly used to describe these processes. In the induced-fit model, ligand binding promotes structural changes that generate a favorable binding geometry, whereas in the conformational-selection model the ligand preferentially binds a pre-existing conformation and shifts the equilibrium toward the bound state [68, 71]. In practice, both mechanisms may contribute to molecular recognition.

Conformational changes influence binding through both enthalpic and entropic contributions. Structural rearrangements can improve shape complementarity and intermolecular interactions but are often accompanied by a loss of conformational freedom upon complex formation [38, 72].

Although carbonic anhydrases are generally considered structurally conserved enzymes [73], recent experimental and computational studies have demonstrated conformational flexibility of active-site residues and surface loops [74–78]. These motions can influence ligand binding, solvent organization, and isozyme selectivity, particularly near the active-site entrance, highlighting the importance of protein dynamics in molecular recognition.

1.2.2 Thermodynamic Description of Protein–Ligand Binding

The thermodynamic behavior of molecular systems can be described by a network of interrelated state functions and response properties. These include Gibbs energy, enthalpy, entropy, heat capacity, volume, compressibility, and thermal expansion coefficients, all of which provide complementary information about molecular interactions and structural changes. Figure 1.3 summarizes the thermodynamic relationships between several quantities most frequently used to characterize protein–ligand interactions. Among these, Gibbs energy (ΔG), enthalpy (ΔH), and entropy (ΔS) are the principal descriptors of molecular recognition because they directly characterize binding affinity and the energetic and dynamic processes accompanying complex formation [8]. Additional parameters, such as changes in heat capacity (ΔC_p) and volume (ΔV), can provide further insight into solvation, conformational rearrangements, and molecular packing [79, 80].

$$\begin{array}{rcccl}
& 1^{st} \text{ derivative} & & 2^{nd} \text{ derivative} & \\
\mathbf{G} \rightarrow & \left(\frac{\partial G}{\partial T} \right)_p & \rightarrow & -\mathbf{S} & \\
& \left(\frac{\partial G}{\partial p} \right)_T & \rightarrow & V & \\
& - \left(\frac{\partial(G/T)}{\partial T} \right)_p T^2 & \rightarrow & \mathbf{H} & \\
& & & & -T \left(\frac{\partial S}{\partial T} \right)_p \searrow \\
& & & & C_p \\
& & & & \left(\frac{\partial H}{\partial T} \right)_p \nearrow
\end{array}$$

Figure 1.3: Derivative relationships between Gibbs energy (G), entropy (S), volume (V), enthalpy (H), and constant-pressure heat capacity (C_p). These quantities constitute the principal thermodynamic parameters used to describe protein–ligand interactions. Adapted from [79, 80].

Binding Equilibrium and Binding Constants

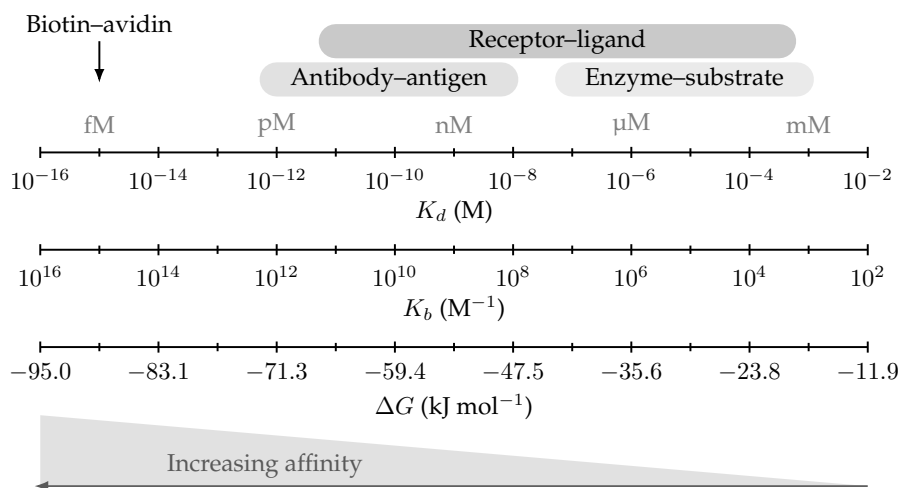


Figure 1.4: Relationship between the dissociation constant (K_d), binding constant (K_b), and Gibbs energy change (ΔG) for representative biomolecular interactions at 37 °C. Interaction classes are shown as approximate ranges. The figure was constructed based on data and concepts presented in [81].

Protein–ligand binding is typically described as a reversible equilibrium between the free protein (P), free ligand (L), and the protein–ligand complex (PL):



The position of this equilibrium is quantified by the association (binding) constant, K_b , or equivalently by the dissociation constant, K_d , defined as:

$$K_b = \frac{[PL]}{[P][L]} \quad \text{and} \quad K_d = \frac{[P][L]}{[PL]} = \frac{1}{K_b} \quad (1.2)$$

where $[P]$, $[L]$, and $[PL]$ represent the equilibrium concentrations of free protein, free ligand, and the protein–ligand complex, respectively. High-affinity interactions are characterized by large K_b values and correspondingly low K_d values. Because binding constants directly reflect the extent of complex formation at equilibrium, they provide the fundamental quantitative measure of molecular recognition and form the basis for the thermodynamic description of protein–ligand interactions.

Gibbs Energy of Binding

The Gibbs energy of binding (ΔG) is the primary thermodynamic parameter describing the spontaneity and favorability of protein–ligand interactions. It is directly related to the equilibrium binding constant (Figure 1.4) according to:

$$\Delta G = -RT \ln K_b \quad (1.3)$$

where R is the universal gas constant and T is the absolute temperature. More negative values of ΔG correspond to stronger binding affinities and greater stability of the protein–ligand complex. Because ΔG integrates all energetic contributions associated with binding, it provides a quantitative measure of the overall driving force for molecular recognition. Importantly, ligands with comparable binding affinities may display markedly different thermodynamic signatures, indicating that similar ΔG values can arise from different balances of intermolecular interactions, conformational changes, and solvent reorganization. The relationship between Gibbs energy, enthalpy, and entropy is described by:

$$\Delta G = \Delta H - T\Delta S \quad (1.4)$$

which forms the basis for thermodynamic analysis of protein–ligand binding.

Binding Enthalpy

Binding enthalpy (ΔH) reflects the energetic changes associated with formation and disruption of intermolecular interactions during protein–ligand complex formation. Favorable enthalpic contributions typically arise from hydrogen bonding, electrostatic interactions, metal coordination, van der Waals contacts, and other attractive interactions established between the protein and the ligand [82, 83]. However, the observed enthalpy change also includes contributions from desolvation, conformational rearrangements, and reorganization of solvent molecules surrounding the binding interface. Consequently, highly favorable binding enthalpies do not necessarily indicate a greater number of direct protein–ligand interactions, but rather reflect the overall energetic balance between interactions formed and interactions disrupted upon binding. In many systems, binding enthalpy provides valuable insight into the molecular determinants of ligand recognition and is frequently used to guide structure-based optimization of protein–ligand interactions.

Binding Entropy

Binding entropy (ΔS) reflects changes in molecular freedom and the distribution of accessible states accompanying protein–ligand complex formation. Binding is generally associated with unfavorable entropy due to restriction of translational, rotational, and internal conformational freedom of both the ligand and the protein. However, favorable entropy contributions may arise from desolvation and the release of ordered water molecules or ions from the binding interface into bulk solution [10, 64, 70]. Consequently, the observed entropy change reflects a balance between losses of molecular flexibility and gains in solvent entropy. Because many entropic contributions cannot be directly visualized in static structures, entropy provides important information about protein dynamics, conformational heterogeneity, and solvent reorganization that complements structural and enthalpic analyses of molecular recognition.

Enthalpy–Entropy Compensation

Enthalpy–entropy compensation refers to the tendency of favorable changes in binding enthalpy (ΔH) to be offset by less favorable entropy changes (ΔS), and vice versa [8, 36, 72]. Strong intermolecular interactions often increase ordering within the protein–ligand complex and reduce conformational freedom, leading to entropic penalties. Conversely, entropy-favorable processes such as desolvation and release of ordered water molecules may contribute less favorably to binding enthalpy. Consequently, large variations in ΔH and ΔS frequently result in relatively small changes in the overall Gibbs energy of binding (ΔG). This phenomenon explains why ligands with similar binding affinities can display markedly different thermodynamic signatures and highlights the importance of considering both enthalpic and entropic contributions during drug design.

1.3 Carbonic Anhydrases as a Model System

Carbonic anhydrases [84] (CAs; EC 4.2.1.1) constitute a family of metalloenzymes that catalyze the reversible hydration of carbon dioxide. These enzymes are widely distributed across all kingdoms of life and are classified into eight evolutionarily distinct families. They participate in a broad range of physiological processes, including pH regulation, respiration, and ion transport. [85, 86]

This thesis focuses on the human isozymes of the α -CA family, which includes 15 members, 12 of which are catalytically active. These isozymes display considerable variation in physiological roles, tissue distribution, expression levels, kinetic parameters, thermal stability, and inhibitor binding properties [13]. The subcellular localization of human CAs is presented in Figure 1.5. Three isozymes, CA VIII, CA X, and CA XI, are catalytically inactive due to the absence of zinc-coordinating histidine residues [87]. These are referred to as carbonic anhydrase-related proteins (CARPs); therefore, they are not considered further in this work.

1.3.1 Structural Features of Carbonic Anhydrases

Carbonic anhydrases belonging to the α -class share a highly conserved globular fold composed of a central twisted β -sheet surrounded

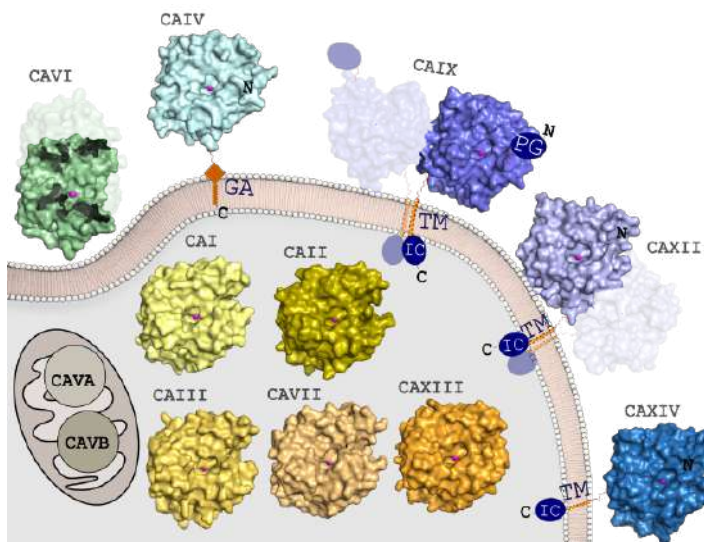


Figure 1.5: Schematic representation and cellular distribution of human carbonic anhydrases. Isozymes are cytosolic (CA I, CA II, CA III, CA VII, CA XIII), membrane-bound (CA IV, CA IX, CA XII, CA XIV), mitochondrial (CA VA, CA VB), or secreted (CA VI). Most are monomeric, whereas CA VI, CA IX, and CA XII form dimers with distinct interfaces. All contain a catalytic domain; CA IV additionally carries a glycosylphosphatidylinositol anchor (GA), CA XII and CA XIV contain a transmembrane region (TM) and intracellular tail (IC), and CA IX additionally contains a proteoglycan-like domain (PG). X-ray crystal structures of the catalytic domains are available for all shown isozymes except CA VA and CA VB. Structures with the following PDB IDs were used: CA I - 1AZM, CA II - 3HS4, CA III - 1Z93, CA IV - 5JN8, CA VI - 3FE4, CA VII - 3ML5, CA IX - 3IAI, CA XII - 1JD0, CA XIII - 3CZV, CA XIV - 4LU3.

by several α -helices and flexible loop regions. Despite substantial sequence diversity among individual isozymes, the overall catalytic architecture remains strongly conserved (Figure 1.6). The active site forms a conical cavity approximately 15 Å deep extending from the protein surface toward the catalytic Zn^{II} ion located at the bottom of the pocket. The zinc ion is tetrahedrally coordinated by three histidine residues (His94, His96, and His119) and a water molecule or hydroxide ion [88].

The active-site cavity contains distinct hydrophobic and hydrophilic regions that collectively regulate substrate binding, proton transfer, solvent organization, and inhibitor recognition. The hydrophobic region, formed primarily by residues Val121, Val143, Leu198, and Trp209, facili-

tates accommodation and orientation of CO₂ and aromatic fragments of inhibitors. The hydrophilic region contains an extended hydrogen-bonding network centered around Thr199, Glu106, and conserved ordered water molecules involved in stabilization of the zinc-bound solvent molecule, proton transfer and inhibitor-binding geometries.

Although the catalytic core remains strongly conserved among isozymes, subtle variations in amino acid composition surrounding the active-site entrance modify cavity geometry, hydrophobicity, solvent accessibility, and electrostatic surface properties. These differences generate distinct binding-site microenvironments that influence inhibitor affinity, selectivity, and binding thermodynamics.

Human CA II is used throughout as the reference structure for structural comparisons, and residue numbers refer to the corresponding CA II positions unless otherwise indicated. Figures 1.8 and 1.9 provide complementary comparisons of the active-site architectures of CAs, showing differences in amino acid composition relative to CA II and qualitative differences in active-site cavity shape and accessibility. The latter depend on the conformations of individual residues and should therefore not be interpreted as direct comparison of cavity size.

CA I

Human CA I adopts the conserved α -carbonic anhydrase fold and possesses an active-site architecture highly similar to that of CA II [89]. Despite strong conservation of the catalytic core, several amino acid substitutions near the active-site entrance distinguish CA I from CA II. In particular, Ala65, Asn67, and His200 replace Ser65, His67, and Thr200 of CA II, respectively. The Ser65→Ala substitution removes a hydroxyl group, His67→Asn replaces an imidazole side chain with a shorter amide, and Thr200→His introduces a bulkier imidazole group. Additional differences surrounding residues Phe91, Ala121, Leu131, and Ala135 modify the hydrophobic region of the active site. Consequently, CA I exhibits lower catalytic activity and generally weaker inhibitor affinities than CA II. Differences in the organization of ordered water molecules surrounding Thr199 and the zinc-bound solvent may further contribute to isozyme-dependent binding thermodynamics.

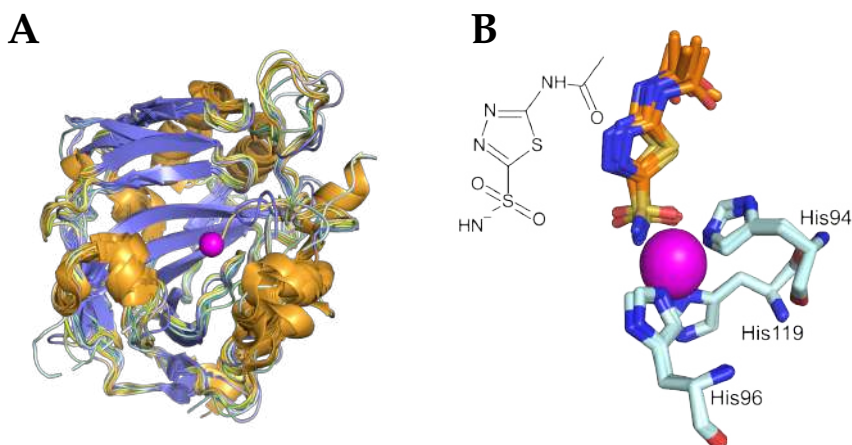


Figure 1.6: Structural conservation of human CAs. (A) Superposition of the catalytic domains of human carbonic anhydrases: CA I - 1AZM, CA II - 3HS4, CA III - 1Z93, CA IV - 5JN8, CA VI - 3FE4, CA VII - 3ML5, CA IX - 3IAI, CA XII - 1JD0, CA XIII - 3CZV, CA XIV - 4LU3. (B) Superposition of active-site regions from carbonic anhydrase isozymes for which crystal structures in complex with acetazolamide (AZM), a classical sulfonamide carbonic anhydrase inhibitor, are available in the Protein Data Bank. The catalytic zinc ion (magenta sphere), AZM, and the three zinc-coordinating histidine residues are shown. The high degree of conservation of the zinc-binding site and surrounding active-site architecture explains the broad inhibitory activity of primary sulfonamides.

CA II

Human CA II is the most extensively investigated member of the carbonic anhydrase family and one of the best structurally characterized metalloenzymes. Owing to its exceptionally high catalytic activity, stability, and the availability of abundant structural and kinetic data, CA II has become the principal model protein for studies of carbonic anhydrases, and structural as well as mechanistic features of other CA isozymes are commonly interpreted and compared relative to CA II. Research on its structural features has evolved from early crystallographic characterization toward detailed mechanistic analyses of catalysis, proton transfer, and active-site dynamics. The first crystal structure of CA II (historically termed carbonic anhydrase C) established the overall architecture of the enzyme and provided the initial description of the catalytic zinc environment and active-site organization [32]. High-resolution refinement studies later improved the accuracy of the active-site geometry

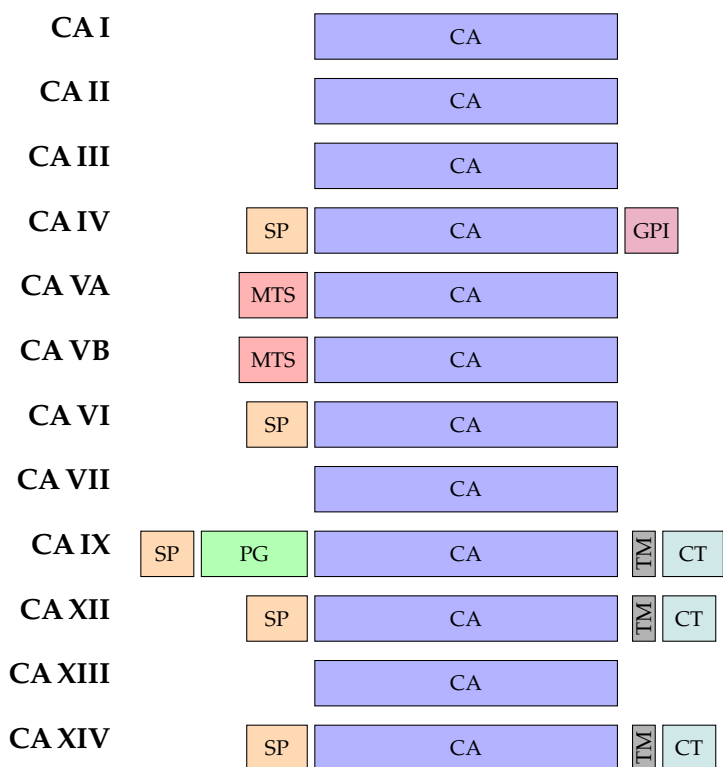


Figure 1.7: Domain organization of catalytically active human CAs. Catalytic carbonic anhydrase domains (CA) are shown together with signal peptides (SP), mitochondrial targeting sequences (MTS), the proteoglycan-like domain (PG), transmembrane helices (TM), cytoplasmic tails (CT), and the GPI-anchor region. Carbonic anhydrase-related proteins (CARPs) are not shown; despite retaining the carbonic anhydrase domain architecture, substitutions of residues required for catalytic zinc coordination render them catalytically inactive.

and solvent structure, providing the basis for mechanistic interpretation of catalysis [90]. Subsequent structural and kinetic investigations identified the central role of His64 as a proton shuttle and characterized the conserved hydrogen-bond network involved in proton transfer and stabilization of the zinc-bound solvent molecule [42, 91, 92]. Mutagenesis combined with crystallographic analysis further clarified the contributions of active-site residues to catalysis, inhibitor binding, and stabilization of the catalytic water network [93, 94]. More recent ultrahigh-resolution X-ray and neutron diffraction studies provided direct insight into protonation states, hydrogen-bonding patterns, and active-site water organization, substantially advancing understanding of

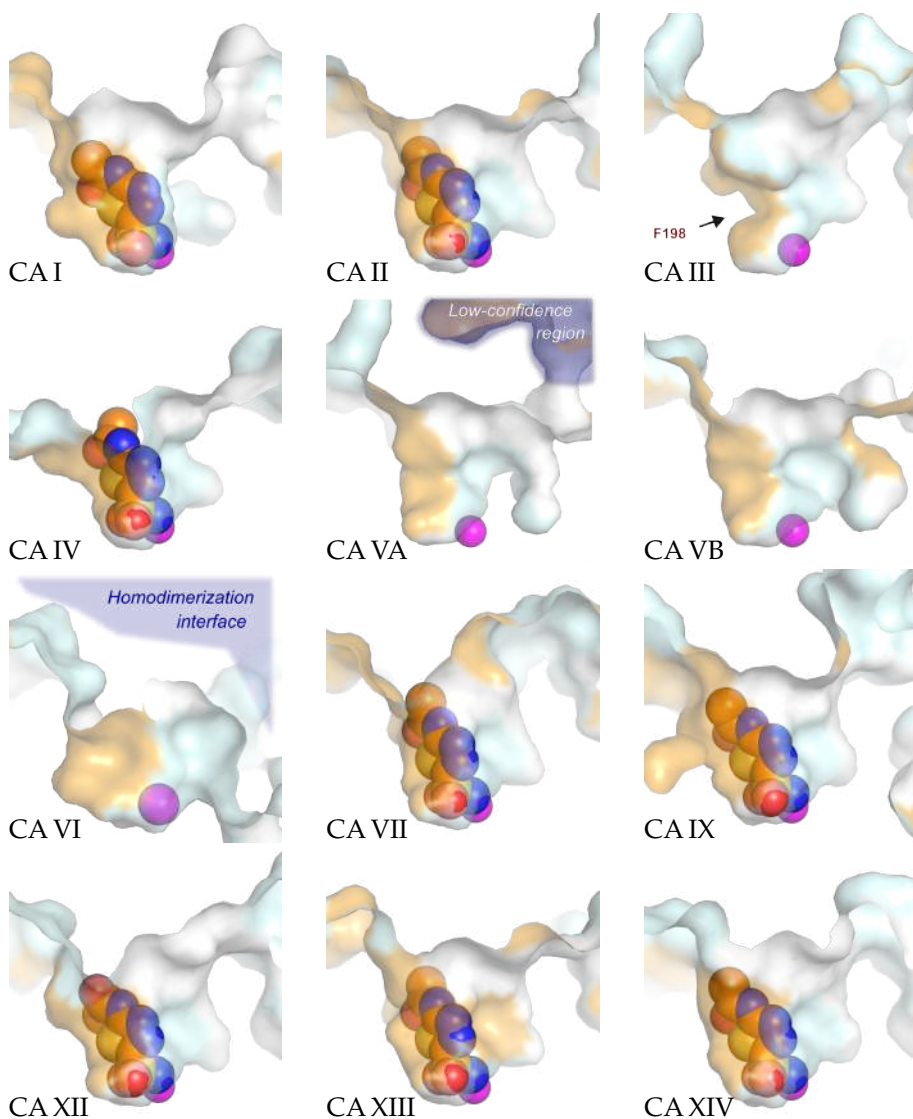
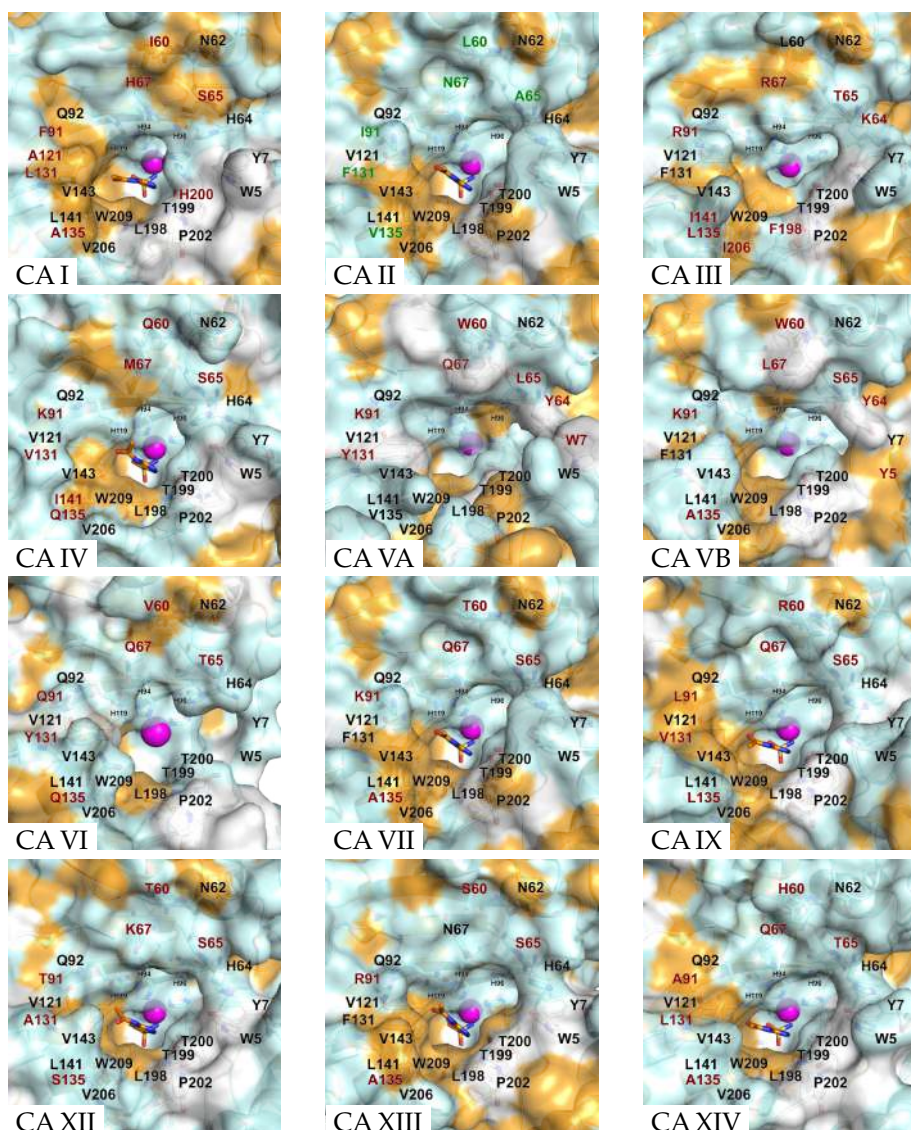


Figure 1.8: Surface representations of active-site cavities in CA isozymes (hydrophobic: orange, hydrophilic: cyan, zinc: magenta sphere). AZM, where available in the experimental structure, indicates the sulfonamide-binding region. In CA III, the bulky Phe198 residue partially occludes the active-site cavity and contributes to its reduced affinity for sulfonamide inhibitors. CA VI forms a homodimer with the dimer interface located adjacent to the active-site entrance, while the catalytic site remains accessible. Experimental structures and PDB IDs are the same as in Figure 1.5. Representations of CA VA and CA VB were generated from AlphaFold-predicted structures (AF-A0A3B3ITA6-F1 and AF-Q9Y2D0-F1, respectively); a low-confidence region of the CA VA model is marked, as it overlaps with the active-site entrance and may lead to a misleading interpretation of ligand access.



the catalytic mechanism and solvent dynamics in CA II [45, 95]. Recent work has additionally highlighted the importance of active-site water dynamics for efficient catalysis and product release [51].

CA III

Among the cytosolic carbonic anhydrases, CA III possesses one of the most distinctive active-site environments. The most characteristic substitution is Lys64 replacing the proton shuttle residue His64 of CA II, which substantially reduces proton-transfer efficiency and catalytic turnover [96]. Another important substitution is Phe198 replacing Leu198 near the active-site entrance. The bulky phenylalanine side chain partially restricts access to the active-site cavity and introduces steric hindrance that affects accommodation of sulfonamide inhibitors, contributing to weaker inhibitor affinities compared with CA II. In addition, CA III contains reactive surface-exposed cysteine residues, particularly Cys181 and Cys186, which are susceptible to oxidative modification and are implicated in redox regulation and protection against oxidative stress.

CA IV

CA IV is an extracellular membrane-associated isozyme attached to the cell surface through a glycosylphosphatidylinositol (GPI) anchor [97]. The catalytic domain is additionally stabilized by a disulfide bond between Cys6 and Cys11 and by several glycosylation sites that influence hydration properties near the active-site entrance. Compared with cytosolic isozymes, CA IV exhibits a narrower and less solvent-accessible catalytic cavity, with substitutions including Gln60, Met67, Lys91, Val131, Ile141, and Gln135 around the active-site cavity. Relative to CA II, substitutions around the active-site cavity include Leu60→Gln, Asn67→Met, Ile91→Lys, Phe131→Val, Leu141→Ile, and Ala135→Gln. Leu60→Gln and Ala135→Gln introduce polar amide groups, Asn67→Met removes an amide group and introduces a hydrophobic thioether side chain, and Ile91→Lys introduces a positively charged side chain. Phe131→Val removes a bulky aromatic ring while retaining hydrophobic character, whereas Leu141→Ile represents a comparatively conservative hydrophobic substitution.

CA VA and CA VB

No experimentally determined structures are currently available for the mitochondrial isozymes CA VA and CA VB. Structural insights are therefore derived from sequence analysis and homology models based on structurally characterized human carbonic anhydrases, particularly CA II [98]. These models indicate that both isozymes retain the catalytic zinc-coordinating residues His94, His96, and His119, suggesting conservation of the canonical catalytic mechanism. Despite the highly conserved catalytic core, several substitutions surrounding the active-site cavity distinguish CA VA and CA VB from CA II. In CA VA, Leu60→Trp, His64→Tyr, Ser65→Leu, Asn67→Gln, Ile91→Lys, Phe131→Tyr, and Ala135→Val are observed, whereas CA VB contains Leu60→Trp, His64→Tyr, Asn67→Leu, and Ile91→Lys, while Ser65, Phe131, and Ala135 are conserved relative to CA II. Leu60→Trp introduces a substantially bulkier aromatic side chain in both mitochondrial isozymes, while His64→Tyr replaces the proton-shuttling imidazole group with a phenolic aromatic side chain. Ile91→Lys replaces a hydrophobic side chain with a positively charged one. In CA VA, Ser65→Leu removes a hydroxyl group and introduces a larger hydrophobic side chain, Phe131→Tyr adds a hydroxyl group while retaining the aromatic ring, and Ala135→Val increases the size and hydrophobic character of the side chain. In CA VB, Asn67→Leu removes a polar amide group and introduces a hydrophobic aliphatic side chain. Both isozymes additionally contain N-terminal mitochondrial targeting sequences that are removed during mitochondrial import, yielding mature enzymes localized within the mitochondrial matrix.

CA VI

CA VI is the only secreted member of the human α -carbonic anhydrase family and forms a homodimer [99]. In addition, several N-glycosylation sites, including Asn67 and Asn256 (CA VI numbering), contribute to extracellular stability and hydration of the protein surface. Compared with CA II, the active-site entrance of CA VI is wider and more solvent exposed, altering the organization of active-site water molecules and accommodation of aromatic inhibitor fragments. The active-site cavity also contains several substitutions relative to CA II: Leu60→Val, Ser65→Thr, Asn67→Gln, Ile91→Gln, Phe131→Tyr, and Ala135→Gln. Leu60→Val represents a conservative hydrophobic substitution with a smaller side chain, whereas Ser65→Thr adds a

methyl group while retaining the hydroxyl functionality. Asn67→Gln extends the amide-containing side chain by one methylene group. More pronounced changes occur at positions 91 and 135, where Ile91→Gln and Ala135→Gln replace hydrophobic side chains with polar amide-containing groups. Phe131→Tyr retains the aromatic ring but introduces a hydroxyl group capable of hydrogen bonding.

CA VII

CA VII is among the most catalytically active cytosolic carbonic anhydrases and retains the proton shuttle residue His64 responsible for efficient proton transfer (Ugochukwu *et. al.*, PDB ID: 3MDZ). The active-site cavity contains an extended hydrophobic cluster formed by Phe131, Val143, Leu198, and Trp209, facilitating accommodation of both CO₂ and hydrophobic ligand fragments. Relative to CA II, the active-site region of CA VII contains the substitutions Leu60→Thr, Asn67→Gln, and Ile91→Lys, while Ala135 and Phe131 are conserved. Leu60→Thr replaces a hydrophobic side chain with a smaller hydroxyl-containing residue, whereas Asn67→Gln extends the amide-containing side chain by one methylene group. Ile91→Lys represents the most pronounced change, replacing a hydrophobic side chain with a positively charged lysine side chain. In addition, reactive cysteine residues, particularly Cys183 and Cys217, may undergo oxidative modifications that affect enzyme activity and structural stability.

CA IX

CA IX is a multidomain transmembrane carbonic anhydrase composed of an N-terminal proteoglycan-like (PG) domain, catalytic domain, transmembrane (TM) segment, and short intracellular (IC) tail [100, 101]. The PG domain forms a highly hydrated and negatively charged extracellular region that contributes to enzyme stability and interactions with the tumor microenvironment, while the extracellular catalytic domains form a homodimer [100, 101]. Crystal structures of the catalytic domain revealed several active-site substitutions relative to CA II, including Leu60→Arg, Ile91→Leu, Phe131→Val, and Ala135→Leu. Leu60→Arg replaces a hydrophobic side chain with a positively charged guanidinium-containing group. The Ile91→Leu substitution preserves a hydrophobic side chain but changes its branching

and local steric profile, whereas Phe131→Val removes the bulky aromatic phenyl group while retaining hydrophobic character. The latter substitution therefore increases the available space adjacent to position 131. Ala135→Leu introduces a larger hydrophobic side chain at the cavity wall. These substitutions generate a more hydrophobic and sterically accessible selective pocket capable of accommodating bulky tail-containing sulfonamide inhibitors and represent key determinants of CA IX selectivity [63]. CA IX additionally contains an intramolecular disulfide bond (Cys119–Cys299) and an N-glycosylation site at Asn309 that contribute to extracellular stability and hydration. The TM segment anchors the enzyme to the plasma membrane, whereas the IC tail contains phosphorylation sites implicated in regulation of enzyme activity and cellular signaling under hypoxic conditions.

CA XII

CA XII is a membrane-associated carbonic anhydrase that forms extracellular homodimers and contains a catalytic domain connected to a single-pass transmembrane segment through a flexible stalk region [98, 102]. Compared with CA IX, the active-site cavity of CA XII is narrower and more hydrophilic. Relative to CA II, the active-site region contains the substitutions Asn67→Lys, Ile91→Thr, Phe131→Ala, and Ala135→Ser. Asn67→Lys replaces a neutral amide-containing side chain with a positively charged amino group, while Ile91→Thr replaces a hydrophobic side chain with a smaller hydroxyl-containing residue. Phe131→Ala removes the bulky aromatic phenyl group and substantially reduces side-chain volume, whereas Ala135→Ser introduces a hydroxyl group with only a small increase in side-chain size. Consequently, the selective pocket of CA XII is less hydrophobic than that of CA IX, contributing to differences in inhibitor specificity and binding thermodynamics [102]. Several extracellular glycosylation sites, including Asn28 and Asn80, contribute to protein stability and hydration, while the short intracellular tail contains phosphorylation sites implicated in regulation of catalytic activity and membrane-associated signaling.

CA XIII

CA XIII exhibits active-site properties intermediate between highly active CA II and low-activity CA III [103]. Unlike CA III, it retains the

proton shuttle residue His64, allowing efficient proton-transfer activity. Relative to CA II, the active-site region contains the substitutions Leu60→Ser, Asn67→Ser, and Ile91→Arg, while Ser65 is conserved. Leu60→Ser replaces a hydrophobic side chain with a smaller hydroxyl-containing residue, and Asn67→Ser replaces an amide group with a shorter hydroxyl-containing side chain. Ile91→Arg represents the most pronounced change, replacing a hydrophobic side chain with a larger, positively charged guanidinium-containing side chain. Compared with CA II, the active-site cavity of CA XIII is slightly narrower and exhibits altered surface electrostatics near the active-site entrance. Differences in the arrangement of ordered water molecules surrounding Thr199 and the zinc-bound solvent further influence inhibitor orientation and binding thermodynamics, contributing to catalytic and inhibitor-binding properties distinct from those of both CA II and CA III [103].

CA XIV

CA XIV is a transmembrane carbonic anhydrase with a compact extracellular catalytic domain [104]. Compared with CA IX and CA XII, its active-site cavity is narrower, more sterically constrained, and less solvent accessible. Relative to CA II, the active-site region contains the substitutions Leu60→His, Ser65→Thr, Asn67→Gln, Ile91→Ala, and Phe131→Leu, while Ala135 is conserved. Leu60→His replaces a hydrophobic side chain with an imidazole-containing residue, whereas Ser65→Thr adds a methyl group while retaining the hydroxyl functionality. Asn67→Gln extends the amide-containing side chain by one methylene group. Ile91→Ala substantially reduces the size of the hydrophobic side chain, while Phe131→Leu removes the aromatic phenyl group but retains a hydrophobic side chain. In addition, extracellular glycosylation sites contribute to surface hydration and structural stability. Differences in loop organization and electrostatic surface properties surrounding the active-site cavity further contribute to isozyme-dependent inhibitor selectivity and binding thermodynamics [104].

Catalytic Mechanism

The active-site environment lowers the pK_a of the zinc-bound water molecule, favoring formation of the catalytically active hydroxide ion under physiological conditions (Figure 1.10). Classical mechanistic

studies [105] highlighted the importance of both hydration/dehydration reaction steps and proton-transfer processes.

The catalytic cycle is generally described as a two-step ping-pong mechanism [62]: (i) **Hydration of CO₂**: CO₂ binds within a hydrophobic pocket near the zinc-bound hydroxide ion. The hydroxide ion attacks the CO₂, forming a zinc-bound bicarbonate, which is subsequently displaced by a water molecule. (ii) **Proton transfer (enzyme regeneration)**: the zinc-bound water molecule must be deprotonated to regenerate the catalytically active hydroxide ion. This proton transfer step is generally considered rate-limiting.

The residue His64 (in CA II) plays a central role as a proton shuttle. Its imidazole side chain alternates between inward and outward conformations, thereby transferring protons between the active site and the bulk solvent. Mutagenesis and structural studies demonstrated that substitution of His64 significantly reduces catalytic efficiency.

Additionally, networks of hydrogen-bonded water molecules in the active-site cavity connect the zinc-bound water to His64. These ordered networks are stabilized by surrounding hydrophilic residues and serve as proton transfer pathways (Figure 1.10). Neutron crystallography and computational simulations have provided further insights into protonation states, hydrogen atom positions, and alternative proton-transfer mechanisms, emphasizing the importance of protein dynamics in catalysis [106, 107].

1.3.2 Sulfonamides as Carbonic Anhydrase Inhibitors

Among the diverse classes of CA inhibitors, sulfonamides remain the most thoroughly investigated and clinically relevant [110–113]. Primary sulfonamides contain a sulfonamide group ($\text{-SO}_2\text{NH}_2$) that, upon deprotonation, coordinates directly to the catalytic zinc ion. The sulfonamide nitrogen displaces the zinc-bound water molecule, forming a tetrahedral complex that prevents substrate access and effectively halts catalysis (Figure 1.10). The primary sulfonamide group therefore provides a conserved zinc-binding anchor, whereas the remainder of the molecule can be modified to establish additional interactions within the active-site cavity. These interactions can modulate both binding affinity and isozyme selectivity.

Clinically used CA inhibitors containing sulfonamide or related zinc-binding groups encompass considerable structural diversity, including

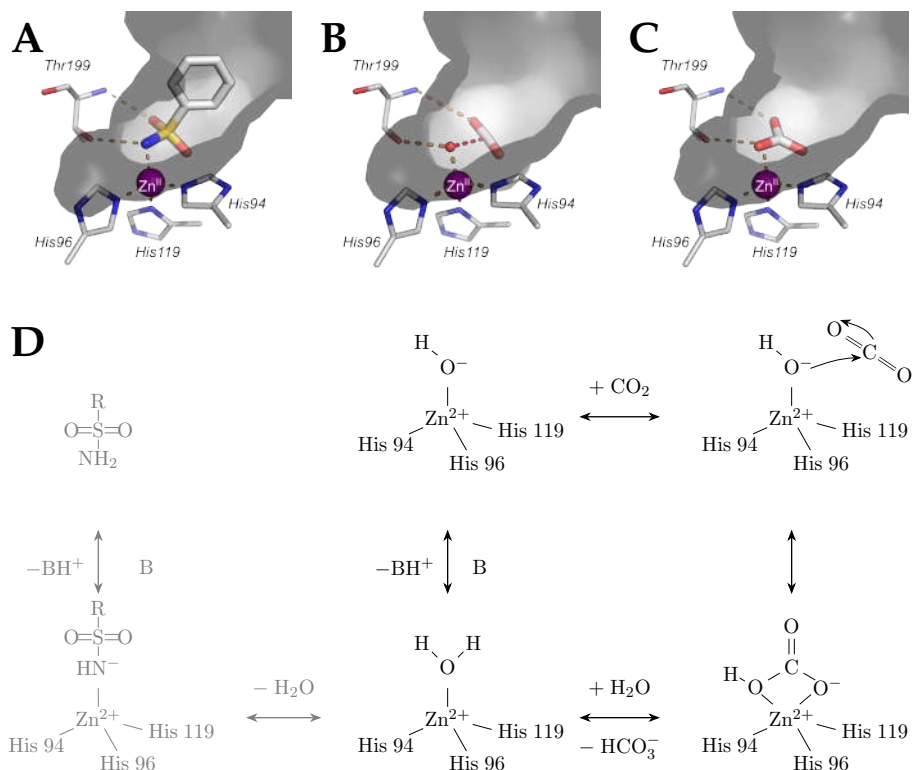


Figure 1.10: Structural and mechanistic views of catalysis and inhibition in human CA II. The catalytic Zn^{II} ion (magenta) is coordinated by His94, His96, and His119; Thr199 participates in the interaction network. Dashed lines indicate zinc coordination and H-bonds. (A) Benzenesulfonamide bound to Zn^{II} (PDB ID: 6GDC; closely mimics the transition state of the substrate hydration reaction [108]). (B) CO₂ adjacent to zinc-bound hydroxide (PDB ID: 2VVA). (C) Bicarbonate bound to Zn^{II} (PDB ID: 2VVB). (D) The catalytic cycle shows CO₂ hydration and enzyme regeneration via proton transfer. Inhibition (grey) illustrates sulfonamide coordination to Zn^{II} and the associated proton-linked equilibria. These proton-linked events (B/BH⁺) must be accounted for to determine intrinsic thermodynamic parameters that reflect protein–ligand interaction energetics.

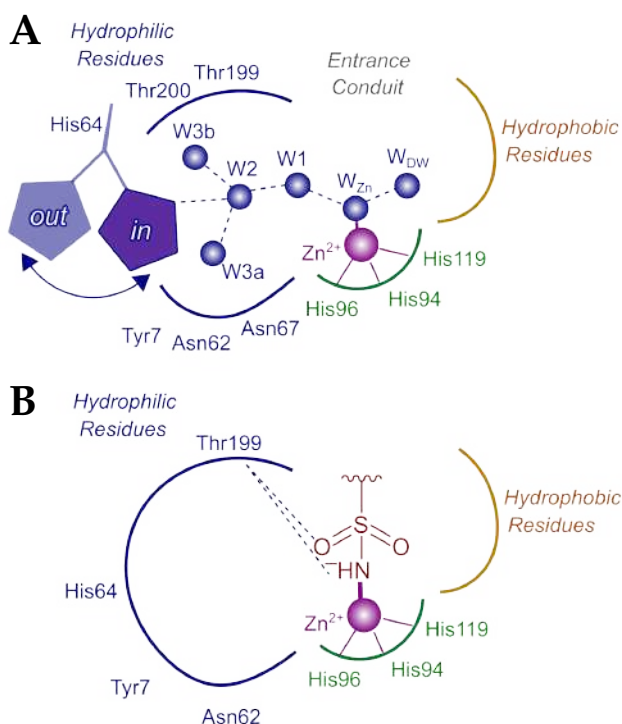


Figure 1.11: Active-site water network and general sulfonamide binding in carbonic anhydrases. (A) Schematic representation of the active-site water network in CA II, based on recent structural studies of CA II water dynamics [51]. The catalytic Zn^{II} ion is coordinated by His94, His96, and His119, while an ordered network of water molecules (W_{Zn}, W1, W2, W3a, and W3b) connects the catalytic center to the proton-shuttling residue His64 through a hydrophilic region formed by Tyr7, Asn62, Asn67, Thr199, and Thr200. This water network plays a central role in proton transfer and catalysis. Similar water-mediated architectures likely exist in other α -carbonic anhydrases, but none have been characterized in comparable detail to CA II. (B) General binding mode of sulfonamide inhibitors in carbonic anhydrases. The sulfonamide zinc-binding group forms conserved interactions with Zn^{II} and Thr199 and occupies a highly conserved position. In contrast, the attached substituents can adopt markedly different orientations and interaction patterns depending on their size, flexibility, and chemical composition, often extending toward distinct regions of the active-site cavity.

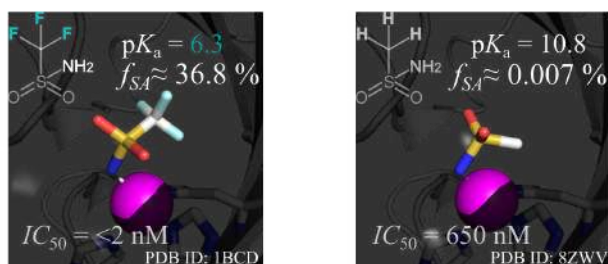


Figure 1.12: Influence of sulfonamide acidity on carbonic anhydrase observed inhibition and binding. Comparison of trifluoromethanesulfonamide (TFMSA, left) and methanesulfonamide (MSA, right) bound to human CA II. The two inhibitors adopt similar binding modes, coordinating the catalytic Zn^{II} ion through the deprotonated sulfonamide nitrogen. Replacement of the methyl group with the strongly electron-withdrawing trifluoromethyl group lowers the sulfonamide pK_{a,SA} from 10.8 to 6.3, increasing the fraction of the zinc-binding sulfonamide anion (f_{SA}) at pH 7.0 from approximately 0.007 % to 36.8 %. Consequently, the markedly improved inhibitory potency of TFMSA (IC₅₀ < 2 nM) relative to MSA (IC₅₀ = 650 nM) arises primarily from a greater population of the active, deprotonated zinc-binding species rather than from substantial differences in protein–ligand interactions. Structures are shown from PDB entries 1BCD (TFMSA) and 8ZWV (MSA); pK_a and IC₅₀ values were taken from the literature [109].

compounds developed specifically as CA inhibitors as well as compounds developed primarily for other therapeutic targets (Figure 1.13). Despite their structural diversity, these compounds generally show limited selectivity among individual CA isozymes.

An important consequence of this binding mechanism is its dependence on protonation equilibria. Because binding of sulfonamide inhibitors to carbonic anhydrases is coupled to protonation and deprotonation equilibria, experimentally observed binding parameters do not directly reflect the intrinsic protein–ligand interaction strength. The equations used to eliminate the contributions of linked protonation reactions and calculate intrinsic binding parameters are provided in the Methods section. Figure 1.12 illustrates one of the most striking examples of the importance of considering intrinsic rather than observed parameters: despite nearly identical binding modes, methanesulfonamide and trifluoromethanesulfonamide differ dramatically in apparent potency owing primarily to differences in sulfonamide acidity and the resulting population of the zinc-binding anionic species. This example

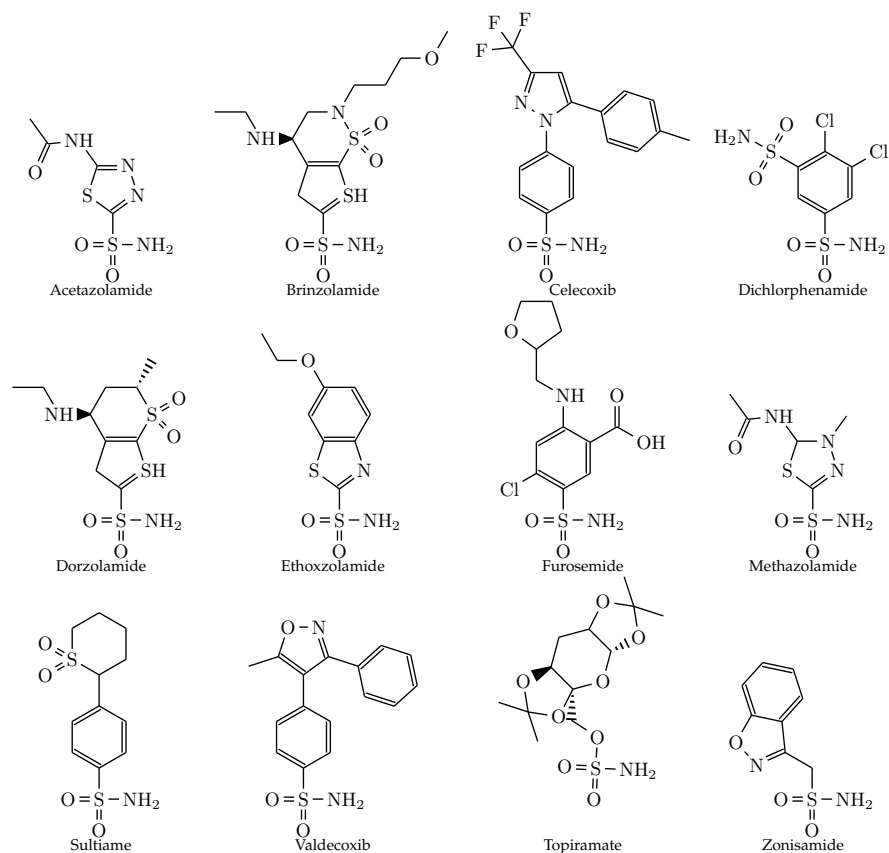


Figure 1.13: Chemical structures of selected clinically relevant sulfonamide and sulfamate carbonic anhydrase inhibitors. A more comprehensive review of CA inhibitors was provided by Kojis *et al.* [114]. Acetazolamide, one of the earliest systemic CA inhibitors, was introduced into clinical practice in the 1950s and remains in use for several indications. Methazolamide is also a systemic CA inhibitor, whereas dorzolamide and brinzolamide were developed for topical treatment of glaucoma. Dichlorphenamide, another early systemic CA inhibitor, is currently used for the treatment of primary periodic paralysis, while ethoxzolamide has little current clinical use. Sulthiame, topiramate, and zonisamide are antiepileptic drugs with CA-inhibitory activity, whereas furosemide, celecoxib, and valdecoxib were developed primarily for pharmacological targets other than CAs but also inhibit CA isozymes. Together, these compounds illustrate the structural and pharmacological diversity of CA-inhibiting drugs, many of which show limited selectivity among individual CA isozymes.

highlights the importance of intrinsic binding parameters for correctly assessing binding strength and guiding compound design.

In addition to direct coordination, sulfonamide inhibitors frequently engage in secondary hydrophobic and hydrogen-bonding interactions with residues lining the active site. These interactions substantially enhance both binding affinity and isozyme selectivity. Structural modifications of the aromatic or heterocyclic substituents have been shown to influence selectivity, which remains a critical challenge in the development of clinically useful CA inhibitors.

The therapeutic applications of sulfonamide-based CA inhibitors are extensive and include the treatment of glaucoma, epilepsy, edema, while inhibition of tumor-associated CA isozymes, particularly CA IX and CA XII, is being investigated as a therapeutic strategy for hypoxic solid tumors [86]. For example, inhibition of CA isozymes in the ciliary body of the eye reduces aqueous humor secretion, thereby lowering intraocular pressure in glaucoma patients. A detailed overview of clinically used CA inhibitors, their therapeutic applications, and emerging CA-targeted drugs is provided elsewhere [114].

Acetazolamide (AZM) is the classical and most extensively studied sulfonamide carbonic anhydrase inhibitor and has been widely used both clinically and as a model compound in structural and mechanistic studies of CA inhibition. Although AZM exhibits high affinity toward carbonic anhydrases and is often described as a selective CA inhibitor relative to other classes of proteins, it displays limited selectivity among CA isozymes, inhibiting a broad range of isozymes with comparable potency. Consequently, AZM is frequently employed as a reference inhibitor in structure-based drug design, serving as a starting point for the development of next-generation inhibitors with improved isozyme selectivity. The combination of a conserved sulfonamide–zinc binding mode and variable interactions with the surrounding active-site environment makes CA–sulfonamide complexes a particularly useful model system for investigating the structural determinants of ligand affinity, selectivity, and binding thermodynamics.

Chapter 2

Methods

Figure 2.1 summarizes the experimental and analytical workflow used in this dissertation. Recombinant carbonic anhydrases and sulfonamide inhibitors constituted the primary study materials.

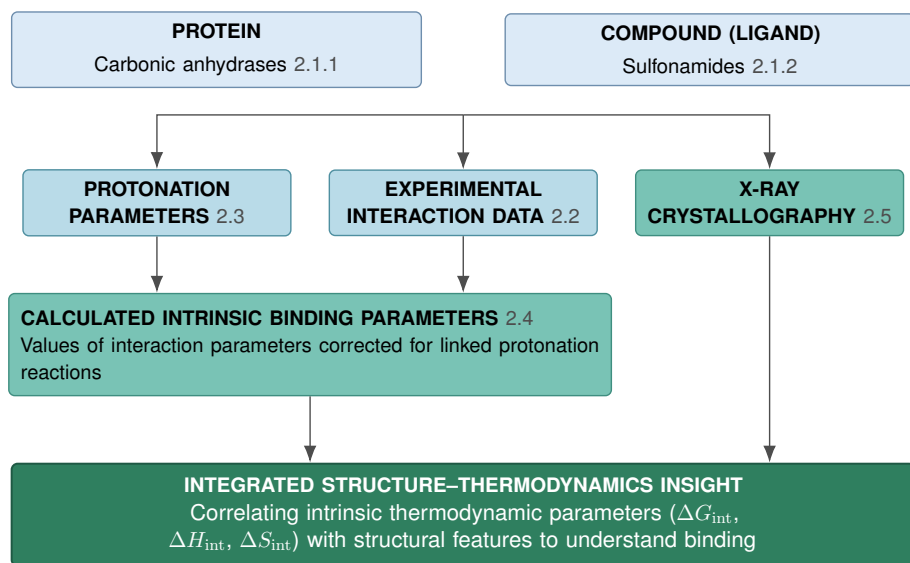


Figure 2.1: Workflow of the main experimental and computational approaches used in this dissertation. Observed protein–ligand interaction data and protonation parameters were combined to calculate intrinsic thermodynamic parameters, which were subsequently integrated with structural information obtained by X-ray crystallography.

Protein–ligand interactions were characterized using biophysical

methods, primarily FTSA and ITC, yielding observed binding constants and thermodynamic parameters. In parallel, protonation parameters of the ligands were determined experimentally, while corresponding protonation parameters of carbonic anhydrase isozymes were obtained from previous studies.

Observed binding parameters were subsequently corrected for linked protonation reactions to derive intrinsic thermodynamic parameters that are independent of experimental conditions and therefore more suitable for comparison across systems. Structural information on protein–ligand complexes was obtained by X-ray crystallography. Finally, intrinsic thermodynamic parameters were integrated with structural data to identify correlations between binding energetics, and molecular recognition, providing insights relevant to inhibitor design. Together, these approaches enabled analysis of relationships between protein–ligand binding thermodynamics and structural features of carbonic anhydrase–inhibitor complexes.

2.1 Materials

2.1.1 Carbonic Anhydrase Isozymes

The CA expression and purification platform used in this work was developed and maintained through the contributions of numerous laboratory members and collaborators over many years. The author gratefully acknowledges these contributions and the protein resources made available for this study.

Twelve CA isoforms were produced recombinantly as full-length proteins or truncated constructs retaining the catalytic domain. Most isoforms were expressed in *Escherichia coli*, using expression strains selected according to the requirements of individual proteins, including BL21(DE3), Rosetta 2(DE3), and Origami B(DE3) [115]. Because stable recombinant CA IX could not be obtained in bacterial cells, its catalytic domain was produced using mammalian expression systems (FreeStyle 293-F or Expi293F) [115]. Proteins were purified using isozyme-specific combinations of Ni²⁺-chelation, ion-exchange, and sulfonamide-based ligand-affinity chromatography. Protein purity, molecular mass, and catalytic activity were assessed by SDS-PAGE, high-resolution mass spectrometry (HR-MS), and stopped-flow CO₂ hydration assays, respectively. Detailed construct sequences, expression conditions, and

purification procedures are provided in [115].

Information on the recombinant proteins can be also accessed through the PLBD: <https://plbd.org/db/main/protein>. Experimental datasets and the corresponding protein batches used in the experiments are linked and described within the database.

Protein samples were stored at -80°C . Prior to experiments, proteins were buffer-exchanged into the corresponding experimental buffers when necessary. The concentrations of CA isozymes were determined spectrophotometrically using their respective molar extinction coefficients at 280 nm.

2.1.2 Sulfonamide Inhibitors

Sulfonamide compounds investigated in this work were synthesized in-house by laboratory collaborators. Different synthetic strategies were employed depending on the compound series, including nucleophilic substitution, alkylation, amidation, and oxidation reactions. Compound identity and purity were confirmed using appropriate analytical methods, primarily NMR spectroscopy and mass spectrometry. Detailed synthesis, purification, and characterization procedures are described in the corresponding publications on which this dissertation is based. Information on the synthesized compounds, including their laboratory identifiers, is available through PLBD (<https://plbd.org/db/main/compound>).

Stock solutions of all compounds were prepared in DMSO at concentrations of 10 mM to 50 mM and subsequently diluted in the appropriate experimental buffers prior to use.

2.2 Experimental Determination of Protein–Ligand Interactions

A combination of biophysical, biochemical, and structural methods was employed to characterize carbonic anhydrase–ligand interactions and to establish relationships between binding energetics and molecular recognition. The overall strategy of this dissertation was to obtain experimental binding parameters, correct them for linked protonation reactions to derive intrinsic thermodynamic parameters, and subsequently

integrate these data with structural information obtained by X-ray crystallography (Section 2.5). These included fluorescent thermal shift assay (FTSA) [116, 117], isothermal titration calorimetry (ITC) [118], X-ray crystallography [119–121], stopped-flow measurements of CO₂ hydration activity [122], and other complementary approaches described in the corresponding publications included in the dissertation.

The Methods section summarizes the experimental approaches and procedures relevant to the interpretation of the results presented in this dissertation. Complete experimental protocols, study-specific conditions, instrumentation, and data analysis procedures are reported in the corresponding publications included. Further methodological details and comparisons between the experimental approaches are provided in the corresponding reviews and standard operating procedures (SOPs) [20, 26, 123, 124].

To ensure data traceability and reproducibility, experimental datasets and associated metadata were recorded in the Protein–Ligand Binding Database (PLBD; <https://plbd.org>) [125]. Structural data obtained by X-ray crystallography were deposited in the Protein Data Bank (PDB).

2.2.1 Fluorescent Thermal Shift Assay

FTSA served as the primary method for determining observed binding constants for large numbers of carbonic anhydrase–inhibitor complexes included in this dissertation. The resulting observed binding constants were subsequently corrected for linked protonation reactions to obtain intrinsic binding constants, as described in Section 2.4.

Protein thermal stability was monitored using FTSA, which detects heat-induced protein unfolding using environmentally sensitive fluorescent dyes. As the protein unfolds, hydrophobic regions become exposed and interact with the fluorescent dye, resulting in changes in fluorescence intensity that can be monitored as a function of temperature [126, 127].

Ligand binding alters protein thermal stability and shifts the protein melting temperature (T_m). Protein–ligand binding constants were derived from ligand-induced thermal stabilization curves using thermodynamic models implemented in Microsoft Excel and later in the Thermott web server (thermott.com) [128] developed in our laboratory.

2.2.2 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) was used to directly measure the heat released or absorbed during protein–ligand binding [129]. The method provided observed binding enthalpies, binding constants, and binding stoichiometries, from which Gibbs energy and entropy changes were calculated. These observed thermodynamic parameters were subsequently corrected for linked protonation reactions to obtain intrinsic binding parameters (see Section 2.4).

To evaluate linked protonation reactions accompanying ligand binding, selected experiments were performed in buffers with different ionization enthalpies [11]. The resulting thermodynamic parameters were subsequently used for the calculation of intrinsic binding constants and intrinsic thermodynamic parameters.

Analysis of ITC datasets involved baseline determination, peak integration, model fitting, and thermodynamic parameter extraction. Depending on the study, calorimetric data were analyzed using manufacturer-provided software and/or specialized ITC software packages, including NITPIC [130, 131], SEDPHAT [131, 132], and GUSSI [133]. Particular attention was devoted to the quality assessment and fitting of calorimetric data, as accurate determination of binding enthalpies was essential for subsequent correction of linked protonation reactions.

2.3 Determination of Protonation Parameters

2.3.1 Sulfonamide pK_a Measurements

The acid dissociation constants (pK_a) of sulfonamide inhibitors were determined experimentally using UV–Vis spectrophotometric titration [56]. The measurements rely on changes in the absorption spectra accompanying protonation and deprotonation of the sulfonamide group. Spectra were recorded at different pH values, and the resulting pH-dependent absorbance changes were fitted to obtain pK_a values.

The experimentally determined pK_a values were subsequently used to calculate the fractions of binding-ready ligand species required for correction of observed thermodynamic parameters and to derive intrinsic thermodynamic parameters of carbonic anhydrase–sulfonamide binding.

2.3.2 Sulfonamide Protonation Enthalpy Measurements

The protonation enthalpies of sulfonamide inhibitors were determined by isothermal titration calorimetry (ITC). The measurements were based on the heat associated with protonation of the deprotonated sulfonamide group during titration with a strong acid. The resulting calorimetric data were analyzed to obtain the protonation enthalpy ($\Delta H_{\text{pr,SA}}$) of each compound.

The experimentally determined protonation enthalpies were subsequently used to correct observed binding enthalpies and calculate intrinsic thermodynamic parameters of carbonic anhydrase–sulfonamide interactions.

2.4 Calculation of Intrinsic Thermodynamic Parameters

Intrinsic thermodynamic parameters reflect the energetic contributions of the protein–ligand interaction itself and therefore provide a more suitable basis for comparison of structurally diverse complexes than the experimentally observed parameters.

Structural studies have demonstrated that sulfonamide inhibitors bind in their deprotonated form by coordinating the active-site Zn^{II} ion and displacing the zinc-bound water molecule [107]. Because only a fraction of both the ligand and the enzyme is present in the binding-competent form under a given set of experimental conditions, the measured binding constants and enthalpies represent observed thermodynamic parameters that depend on pH, buffer composition, and protonation equilibria [11, 12].

To enable direct comparison of thermodynamic and structural data, observed binding parameters were converted into intrinsic thermodynamic parameters by correcting for linked protonation reactions. Figure 1.12 provides one of the most striking examples of the importance of intrinsic over observed parameters in carbonic anhydrase inhibition and binding.

2.4.1 Correction of Binding Constants

The intrinsic binding constant, $K_{b,int}$, was calculated from the observed binding constant, $K_{b,obs}$, and the fractions of the binding-competent forms of the protein (f_{CA}) and ligand (f_{SA}):

$$K_{b,int} = \frac{K_{b,obs}}{f_{CA} \cdot f_{SA}} \quad (2.1)$$

The fraction of the protonated enzyme was calculated as

$$f_{CA} = \frac{10^{pK_{a,CA}-pH}}{1 + 10^{pK_{a,CA}-pH}} \quad (2.2)$$

and the fraction of the deprotonated sulfonamide inhibitor as

$$f_{SA} = \frac{10^{pH-pK_{a,SA}}}{1 + 10^{pH-pK_{a,SA}}} \quad (2.3)$$

Here, $pK_{a,CA}$ denotes the acid dissociation constant of the zinc-bound water molecule in the enzyme active site and $pK_{a,SA}$ the acid dissociation constant of the sulfonamide group.

The intrinsic Gibbs energy of binding, ΔG_{int} , was then calculated from $K_{b,int}$ using the standard thermodynamic relationship

$$\Delta G_{int} = -RT \ln(K_{b,int}). \quad (2.4)$$

2.4.2 Correction of Binding Enthalpies and Entropies

The intrinsic binding enthalpy, ΔH_{int} , was obtained by correcting the observed binding enthalpy, ΔH_{obs} , for proton transfer reactions associated with ligand deprotonation, deprotonation of the zinc-bound water molecule, and protonation of the experimental buffer:

$$\Delta H_{int} = \Delta H_{obs} - n_1 \Delta H_{pr,SA} - n_2 \Delta H_{pr,CA} + (n_1 + n_2) \Delta H_{pr,buf} \quad (2.5)$$

where

$$n_1 = f_{\text{SA}} - 1 \quad (2.6)$$

$$n_2 = 1 - f_{\text{CA}}. \quad (2.7)$$

The parameters n_1 and n_2 correspond to the numbers of protons released or accepted during the binding process.

Values of $\text{p}K_{\text{a,CA}}$ and $\Delta H_{\text{pr,CA}}$ for individual carbonic anhydrase isozymes were determined previously in our laboratory and published in [134]. Buffer protonation enthalpies ($\Delta H_{\text{pr,buf}}$) were obtained from published thermodynamic data [135]. The values of $\text{p}K_{\text{a,SA}}$ and $\Delta H_{\text{pr,SA}}$ were determined experimentally as described in Section 2.3.

Observed binding constants were determined by fluorescent thermal shift assay, whereas observed binding enthalpies were measured by isothermal titration calorimetry. These parameters were subsequently used to calculate intrinsic binding constants and thermodynamic parameters that are independent of linked protonation equilibria and therefore more suitable for comparison with structural data.

Intrinsic entropy changes were calculated from intrinsic Gibbs energies and intrinsic enthalpies according to Eq. S.8.

$$\Delta S_{\text{int}} = \frac{\Delta H_{\text{int}} - \Delta G_{\text{int}}}{T}. \quad (2.8)$$

2.5 Structural Studies

Structural information for carbonic anhydrase–ligand complexes was obtained by X-ray crystallography. Crystals of CA isozymes were grown at room temperature using the sitting-drop vapor diffusion method. Recombinant protein samples were concentrated via ultrafiltration to final concentrations ranging from 15 mg mL^{-1} to 50 mg mL^{-1} . Crystallization drops ($0.1 \text{ }\mu\text{L}$ to $5 \text{ }\mu\text{L}$) were prepared by mixing the protein stock with the reservoir solution and equilibrated against 0.1 mL to 0.5 mL of the same reservoir solution. Optimized crystallization conditions typically involved 0.1 M Tris-HCl ($\text{pH } 8.5$) with 0.2 M salts (NaCl or ammonium sulfate) and 20% to 28% PEG (4000, 3350, or 8000) for CA I; 0.1 M sodium

bicine (pH 9.0) and 2.0 M sodium malonate (pH 7.0) for CA II; 0.1 M ammonium citrate (pH 7.0–7.2) and 26 % to 30 % PEG 4000 for CA XII; and sodium citrate or acetate buffers (pH 4.5–5.5) with 26 % to 30 % PEG 4000 for CA XIII. Protein–ligand complexes were primarily obtained by soaking pre-grown crystals in reservoir solutions supplemented with 0.5 mM to 1.0 mM ligand (diluted from 50 mM DMSO stocks) for up to one week, while co-crystallization was specifically employed for CA I complexes. CA I and CA II crystals were flash-cooled directly from the mother liquor, while CA XII and CA XIII crystals were transferred to a cryoprotectant solution containing 15 % to 20 % ethylene glycol.

Diffraction data were collected either at the macromolecular crystallography beamlines P13 and P14 at the DESY synchrotron (Hamburg, Germany) or using a Rigaku MM007-HF diffractometer at the Life Sciences Center of Vilnius University. Diffraction datasets were processed using standard crystallographic procedures, and structures were solved by molecular replacement using previously determined carbonic anhydrase structures as search models. Iterative cycles of model building and refinement were performed until satisfactory agreement between the atomic model and experimental electron density was achieved.

The majority of crystal structures included in this dissertation were determined by the author. A subset of CA IX and CA XII structures was determined by collaborators Dr. Jānis Leitāns and Prof. Dr. Kaspars Tars (Latvian Biomedical Research and Study Centre, Latvia). All structures included in this dissertation were subsequently analyzed by the author to identify protein-ligand interactions, binding-site hydration patterns, and structural features relevant to thermodynamic binding parameters.

Detailed crystallization conditions, data collection parameters, refinement statistics, structure validation metrics, and software used for structure determination are reported in the corresponding publications included in this dissertation.

Chapter 3

Summary of Results

3.1 Determinants of Protein–Ligand Recognition

The studies presented in this section investigate the molecular determinants governing protein–ligand recognition in carbonic anhydrases. Particular emphasis is placed on understanding how thermodynamic mechanisms, ligand structure, and protein environment contribute to binding affinity and isozyme selectivity. First, the phenomenon of enthalpy–entropy compensation is examined to reveal how similar binding affinities can arise through distinct energetic mechanisms. Subsequently, the influence of minor ligand modifications on affinity and selectivity is investigated. Finally, structural and thermodynamic data are integrated to establish relationships between protein–ligand structure and binding energy.

3.1.1 Enthalpy-Entropy Compensation

This subsection summarizes the results of the publication Paketuryte-Latve et al. 2026:

Vaida Paketurytė-Latvė, Audrius Zakšauskas, Visvaldas Kairys, Edita Čapkauskaitė, Martynas Bagdonas, Alexey Smirnov, Elena Manakova, Saulius Gražulis, Daumantas Matulis

Enthalpy–Entropy Compensation in Inhibitor Binding to Carbonic Anhydrase Isozymes

Manuscript under review.

My contribution included development of the study concept, experimental determination of binding thermodynamics by ITC, calculation and analysis of intrinsic thermodynamic parameters, X-ray crystallographic studies, interpretation of enthalpy–entropy compensation phenomena, and manuscript preparation.

To investigate the molecular origins of enthalpy–entropy compensation, a matched series of sulfonamide inhibitors (specifically, the series of six *ortho*, *meta*-disubstituted benzenesulfonamides 4a–4d, 4h, and 7h) was characterized in complex with CA II, CA XII, and CA XIII. Binding affinities were determined by FTSA and ITC, while ITC additionally provided binding enthalpies. High-resolution X-ray structures were determined for selected protein–ligand complexes yielding seven atomic-resolution structures of complexes with CA II (PDB: 3OEN, 3OEO, 3OIC), CA XII (PDB: 3OFA, 3OFB), and CA XIII (PDB: 3OFP, 3OFN) to relate the measured thermodynamic differences to specific features of the active sites.

Observed thermodynamic parameters were corrected for proton-linked equilibria involving sulfonamide deprotonation, protonation of the zinc-bound hydroxide, and buffer ionization. This deconvolution relied on experimentally determined ligand-scaffold parameters, including a $pK_{a,SA}$ of 9.8 and a protonation enthalpy ($\Delta H_{pr,SA}$) of $-35.2 \text{ kJ mol}^{-1}$ for compounds 4a–4d and 4h, and a $pK_{a,SA}$ of 8.9 and $\Delta H_{pr,SA}$ of $-30.4 \text{ kJ mol}^{-1}$ for compound 7h, using TRIS as the measurement buffer ($\Delta H_{pr,buf} = -46.56 \text{ kJ mol}^{-1}$). The resulting intrinsic Gibbs energies were generally similar among the three isozymes, whereas their enthalpic and entropic contributions differed substantially. For example, compound 4a exhibited intrinsic Gibbs energies

of $-59.3 \text{ kJ mol}^{-1}$ $-61.4 \text{ kJ mol}^{-1}$ for CA II and CA XIII, respectively, while the corresponding intrinsic enthalpies were $-26.8 \text{ kJ mol}^{-1}$ and $-64.0 \text{ kJ mol}^{-1}$. Thus, a difference of only 2.1 kJ mol^{-1} in binding Gibbs energy was accompanied by a difference of 37.2 kJ mol^{-1} in enthalpy, compensated by an opposing change in the entropic contribution (with $-T\Delta S_{\text{int}}$ shifting from a highly favorable $-32.5 \text{ kJ mol}^{-1}$ in CA II to an unfavorable 2.6 kJ mol^{-1} in CA XIII).

This behavior was observed across the investigated series. Binding to CA II was generally more entropy-favored, whereas binding to CA XIII was more enthalpy-favored, with CA XII displaying intermediate thermodynamic profiles. Specifically, the average relative partitioning of the binding Gibbs energy across the compounds demonstrated that binding to CA II was 70 % entropy-favored (30 % enthalpy), CA XII was intermediate (70 % enthalpy, 30 % entropy), whereas binding to CA XIII was entirely enthalpy-driven (100 % enthalpy, 0 % entropy). Differences between isozymes were generally larger than those produced by modifications within the ligand series. Correction for proton-linked equilibria was essential for this comparison, as the observed thermodynamic profiles differed substantially from the intrinsic profiles.

Despite these pronounced thermodynamic differences, the crystallographic binding modes were highly conserved. In all structures, the deprotonated sulfonamide nitrogen coordinated the catalytic Zn^{II} ion, and the sulfonamide group formed conserved hydrogen bonds with Thr199. Differences were observed farther from this conserved anchoring region. At position 62, Asn in CA II and CA XII could form a direct hydrogen bond with the ligand, whereas the shorter Ser62 in CA XIII resulted in indirect, water-mediated interactions. At position 131, CA II and CA XIII contain Phe, while CA XII contains Ala; Phe131 showed conformational variability in response to different ligands. Another notable difference occurred at position 200, which is occupied by Thr in CA II and CA XII but by Val in CA XIII, altering the local hydration environment adjacent to the benzenesulfonamide group.

Comparison of the structural and thermodynamic data therefore showed that similar ligand poses and binding Gibbs energies do not imply similar enthalpic and entropic contributions. The largest differences were associated with the protein isozyme rather than with relatively small modifications of the ligand. These differences are consistent with isozyme-specific active-site hydration, where differences in water organization and the energetic cost of dehydration upon ligand binding may contribute to the distinct thermodynamic profiles. Specifically, in CA II, Thr200 stabilizes the ordered W1 water molecule via a short

hydrogen bond, and its displacement upon ligand binding yields a highly favorable solvent-entropy contribution. In contrast, Val200 in CA XIII cannot form this stabilizing hydrogen bond, leading to a less ordered, unconstrained active-site water network that provides a negligible solvent-entropy change during ligand binding, thereby shifting the thermodynamic profile to be entirely enthalpy-driven.

3.1.2 Affinity and Selectivity

This subsection summarizes the results of the publication [136]:

Audrius Zakšauskas, **Vaida Paketurytė-Latvė**, Alberta Jankūnaitė, Edita Čapkauskaitė, Yann Becart, Alexey Smirnov, Klára Pospíšilová, Janis Leitans, Jiří Brynda, Andris Kazaks, Lina Baranauskienė, Elena Manakova, Saulius Gražulis, Visvaldas Kairys, Kaspars Tars, Pavlína Řezáčová, Daumantas Matulis

Affinity and Selectivity of Protein-Ligand Recognition: A Minor Chemical Modification Changes Carbonic Anhydrase Binding Profile

J. Med. Chem., 2025, 68 (16), 17752–17773.

DOI: 10.1021/acs.jmedchem.5c01421

My contribution included experimental determination of binding affinities by FTSA, analysis of intrinsic binding thermodynamics, X-ray crystallographic studies, interpretation of selectivity trends, data analysis, and manuscript preparation. A significant contribution to this publication was also provided by the organic synthesis of the studied compounds performed by Dr. A. Zakšauskas, Dr. A. Jankūnaitė, Dr. E. Čapkauskaitė, and Yann Becart. It is important to note that independent enzyme inhibition experiments were carried out by K. Pospíšilová. This enabled the demonstration of a strong correlation between independently obtained binding (FTSA) and inhibition (SFA) data.

This study systematically investigated how minor chemical modifications of *ortho*-substituted benzenesulfonamides dramatically alter binding affinity and selectivity toward human carbonic anhydrase (CA) isozymes. Nearly 50 compounds were synthesized, representing seven chemical classes based on sulfur oxidation state and linker heteroatoms (including 4-sulfanyl, 4-sulfinyl, and 4-sulfonyl esters and benzamides, along with 4-amino benzamides). Binding affinities were quantified for

all 12 catalytically active human CA isozymes using fluorescence-based thermal shift assay (FTSA), complemented by stopped-flow enzymatic inhibition assays and structural characterization by X-ray crystallography.

Small structural changes produced affinity differences spanning several orders of magnitude. In particular, 4-sulfanyl-substituted benzenesulfonamides exhibited exceptionally strong and selective binding to CA IX, with intrinsic dissociation constants reaching the picomolar range. Specifically, the cyclohexylsulfanyl ester 4b displayed an outstanding intrinsic K_d of 1 pM (0.0010 nM) for CA IX, while binding hundreds of times more weakly to the off-target isozyme CA II ($K_{d,int} = 0.49$ nM). Furthermore, while the bulky adamantylsulfanyl ester 4g achieved a similar CA IX affinity ($K_{d,int} = 3$ pM) with several hundred-fold selectivity over CA II, the smaller methylsulfanyl analogue 4h lost this selectivity. In contrast, oxidation of the sulfanyl group to sulfinyl or sulfonyl markedly reduced affinity and largely abolished CA IX selectivity; for instance, the sulfinyl derivative 7b and sulfonyl derivative 9b displayed dramatically reduced CA IX affinities of 28 nM and 320 nM, respectively, despite minimal changes to the overall molecular scaffold.

Intrinsic binding constants were derived from observed affinities by correcting for pH-dependent protonation equilibria of both the sulfonamide group and the zinc-bound water molecule. This analysis demonstrated that enhanced binding was not solely attributable to favorable pK_a shifts, but instead reflected genuine improvements in protein-ligand interactions. This correction is crucial because, although the strong electron-withdrawing effect of sulfinyl and sulfonyl groups lowered the sulfonamide $pK_{a,SA}$ by nearly a full pH unit (from 9.8 for sulfanyl compounds to 8.9 for sulfoxides) and increased the fraction of deprotonated ligand ready to coordinate the zinc ion at pH 7.0, their intrinsic binding affinities were actually hundreds of thousands of times weaker. These findings emphasize the importance of intrinsic thermodynamic parameters for reliable structure–affinity relationships in CA inhibitor design.

Comparison between ester and benzamide analogues showed that benzamides generally bound more weakly to CA isozymes, although several retained notable selectivity for CA IX since most benzamides exhibited no detectable binding against off-target CAs. Substituent size and hydrophobicity played a critical role: bulky hydrophobic groups optimized interactions within the enlarged CA IX active site while causing steric clashes in smaller isozymes, thereby driving selectivity. Additionally, comparing sulfanyl linkers to amino linkers showed similar

potencies, where the cyclooctyl-substituted 4-amino-benzamide 6e selectively bound CA IX ($K_{d,int} = 0.53$ nM) and showed no binding to smaller active-site isozymes.

X-ray crystal structures and molecular docking provided structural explanations for the affinity trends. High-affinity sulfanyl-substituted ligands adopted binding poses that optimized hydrophobic packing and water-mediated hydrogen bonding networks; specifically, the nine newly determined structures (including CA II-4c, CA IX-4d, and CA XIII-4d; PDB IDs: 9FPQ, 9R8X, 9FPW, respectively) showed that the ligand ester group formed a direct hydrogen bond with Asn62 in CA II, but transitioned to a water-mediated hydrogen bond network with Gln70/Asn65 in CA IX and Ser64/Asn69 in CA XIII. Furthermore, crystallographic analysis of compound 4h bound to CA II (PDB ID: 9FPS) revealed an atypical binding mode containing two ligand molecules, where the second molecule bound non-classically at the active-site periphery via hydrophobic and water-mediated interactions. In contrast sulfinyl and sulfonyl derivatives exhibited unfavorable intramolecular clashes that distorted binding geometries. Density Functional Theory (DFT) calculations revealed that the docked conformer of sulfonyl derivative 9a undergoes a severe intramolecular oxygen-oxygen clash, resulting in a conformational energy penalty of 53.8 kJ mol^{-1} relative to its low-energy solution state conformation, whereas the sulfanyl analogue 4a remains highly stable and free of clashes. Molecular docking reproduced several experimental trends for sulfanyl compounds but failed to account for the behavior of sulfonyl derivatives due to the scoring function's inability to calculate ligand conformational strain.

Overall, the results demonstrate that the oxidation state and conformational flexibility of a single *ortho* substituent can dictate CA isozyme affinity and selectivity over several orders of magnitude. This work establishes a clear mechanistic link between subtle chemical modifications, ligand conformational energetics, and selective recognition of CA IX, providing valuable guidance for the rational design of highly selective carbonic anhydrase inhibitors.

The pronounced effects of minor structural modifications prompted a broader investigation into how structural information and thermodynamic parameters can be integrated to explain protein–ligand recognition.

3.1.3 Structure-Thermodynamics Relationships

This subsection summarizes the results of the publication[137]:

Vaida Paketurytė-Latvė, Alexey Smirnov, Elena Manakova, Lina Baranauskienė, Vytautas Petrauskas, Asta Zubrienė, Jurgita Matulienė, Virginija Dudutienė, Edita Čapkauskaitė, Audrius Zakšauskas, Janis Leitans, Saulius Gražulis, Kaspars Tars, Daumantas Matulis

From X-ray crystallographic structure to intrinsic thermodynamics of protein-ligand binding using carbonic anhydrase isozymes as a model system

IUCrJ, 2024, 11 (4), 556–569.

DOI: 10.1107/S2052252524004627

My contribution included conceiving the publication idea, thermodynamic analysis of protein–ligand interactions, integration of structural and thermodynamic data, and preparation of the manuscript. The majority of the binding thermodynamic parameters were determined by Dr. A. Zubrienė and Dr. L. Baranauskienė, while some of the crystal structures cited in the article were solved by Dr. A. Smirnov, Dr. E. Manakova, and Dr. J. Leitans. The cell-based assays were carried out by Dr. J. Matulienė.

To establish correlations between protein–ligand structure and binding thermodynamics, X-ray crystallographic data for carbonic anhydrase–sulfonamide complexes were combined with binding measurements obtained by FTSA, ITC, enzymatic inhibition, and cell-based assays. The analysis focused on structurally related sulfonamide inhibitors (specifically the fluorinated and chlorinated benzenesulfonamide series) exhibiting markedly different affinities and selectivities toward human carbonic anhydrase isozymes, particularly CA IX. To overcome the crystallization challenges associated with native CA IX, a chimeric CA IX (chCA IX) was utilized as a highly accurate mimic, containing six active-site mutations (A65S, N67Q, I91L, F130V, V134L, L203A) introduced into the CA II.

Crystal structures revealed a conserved binding mode for sulfonamide inhibitors, with coordination of the catalytic zinc ion via the negatively charged deprotonated nitrogen and formation of similar interactions within the active site. However, ligands exhibiting nearly identical binding poses often differed in affinity by several orders of

magnitude. For instance, structurally highly similar inhibitors EA3-3 and EA5-3 in complexes with chCA IX and CA II (both crystallizing in the space group $P2_1$ with nearly identical unit-cell parameters) adopted conserved binding conformations but demonstrated vastly divergent affinities. These observations demonstrated that structural information alone is insufficient to explain binding affinity or isozyme selectivity.

Comparison of different experimental approaches highlighted the strengths and limitations of individual binding assays. FTSA provided the broadest affinity range, extending from millimolar to picomolar dissociation constants, whereas enzymatic inhibition assays and ITC were restricted to narrower measurement windows. The use of multiple orthogonal methods enabled more reliable characterization of protein–ligand interactions.

A central finding of this work was the importance of intrinsic thermodynamic parameters. Correction of observed binding data for linked protonation equilibria yielded intrinsic Gibbs energies, enthalpies, and entropies of binding. This deconvolution accounted for the simultaneous protonation of the zinc-bound hydroxide in the enzyme active site (with a $pK_{a,CA}$ of 6–7, or up to 8 for CA I and CA XIII) and the deprotonation of the benzenesulfonamide group (with a $pK_{a,SA}$ typically between 8 and 11). These intrinsic parameters more accurately reflected the energetic consequences of protein–ligand interactions and enabled meaningful structure–thermodynamics correlations.

The combined analysis showed that ligand affinity and selectivity arise from a complex interplay of structural, solvation, and thermodynamic effects. Minor chemical modifications often produced substantial changes in binding energetics despite only subtle differences in crystallographic structures. Specifically, substituting a tail methyl group (VD12-05) with a hydroxy tail (VD10-35) resulted in a systematic loss of 7.5–10.7 kJ mol^{−1} of intrinsic Gibbs energy across CA isozymes, as the polar hydroxy group preferentially formed hydrogen bonds with bulk solvent rather than the protein. Conversely, adding a cyclooctyl amino ring at the *meta*-position (VD11-4-2) caused steric clashes that reduced CA I affinity by 18.0 kJ mol^{−1} (a 1000-fold decrease), but achieved a 21.9 kJ mol^{−1} gain (4000-fold increase) for CA IX and 14.9 kJ mol^{−1} for CA XII, driving a 100,000-fold selectivity for CA IX over CA I. Furthermore, cyclohexane substitution in compound EA2-3 optimized hydrophobic interactions within the enlarged CA IX pocket, yielding an exceptionally strong intrinsic Gibbs energy of −75.0 kJ mol^{−1} (femtomolar range). These results emphasize that reliable interpretation of protein–ligand recognition requires integration of structural information

with intrinsic thermodynamic analysis.

Overall, this work emphasized a framework for connecting X-ray crystallographic structures with intrinsic thermodynamic parameters and demonstrated how such integration can improve the understanding of molecular recognition and support rational drug design. This is clearly illustrated by the opposing thermodynamic driving forces governing potent binders: while the binding of VD10-35 to CA I is purely enthalpy-driven ($\Delta H_{\text{int}} = -76.8 \text{ kJ mol}^{-1}$), the selective binding of VD11-4-2 to CA IX is heavily entropy-driven ($-T\Delta S_{\text{int}} = -47.5 \text{ kJ mol}^{-1}$) despite its lower enthalpic contribution ($\Delta H_{\text{int}} = -24.3 \text{ kJ mol}^{-1}$), together forming an exceptionally strong complex ($\Delta G_{\text{int}} = -71.8 \text{ kJ mol}^{-1}$).

3.2 Reliability of Thermodynamic Measurements

Reliable interpretation of protein–ligand recognition depends not only on the quality of structural and thermodynamic data, but also on the methods used to obtain and analyze these data. Experimental limitations, parameter correlations, and statistical uncertainty can substantially influence the accuracy and interpretation of binding constants and thermodynamic parameters. The studies presented in this section address two complementary aspects of this problem: standardization of fluorescent thermal shift assay measurements and rigorous quantification of uncertainty in protein–ligand binding parameters.

3.2.1 FTSA Standardization

This subsection summarizes the results of the publication[116]:

Egidijus Kazlauskas, Vytautas Petrauskas, **Vaida Paketurytė**,
Daumantas Matulis

Standard Operating Procedure for Fluorescent Thermal Shift Assay (FTSA) for Determination of Protein-Ligand Binding and Protein Stability

Eur. Biophys. J., 2021, 50 (5), 707–714.

DOI: 10.1007/s00249-021-01537-1

My contribution included the analysis and interpretation of data, the preparation of figures and manuscript text focused on the FTSA standard operating procedure for protein stability and binding.

This standard operating procedure (SOP) demonstrated that FTSA provides reproducible, high-quality melting curves suitable for determining both protein stability and ligand binding affinities across a wide dynamic range (10^3 to 10^{13} M⁻¹). The melting temperature (T_m) is determined at the inflection point of the sigmoidal denaturation transition, where ligand-induced thermal shifts (ΔT_m) across varied concentrations are fitted to calculate the equilibrium binding constant (K_b). Using CA II (specifically, the binding of acetazolamide up to 200 μ M with 5 μ M CA II, 2 % v/v DMSO, and 50 μ M ANS) and Hsp90 α N-terminal domain (screened across pH 5.0 to 10.5 in a multicomponent acetate-borate-phosphate buffer) as examples on qPCR equipment (relying on centrifugal rotary spectrofluorimeters and 10 μ L-20 μ L volumes for optimal temperature uniformity), the assay reliably detected ligand-induced thermal shifts and pH-dependent stability changes. The work also showed that assay quality depends on temperature-insensitive buffer composition (utilizing phosphate or HEPES instead of TRIS to avoid temperature-dependent pH shifts), DMSO levels (safe at 2 % v/v, but destabilizing at ≥ 4 %), protein structure (favoring mono-domain proteins over complex multi-domain assemblies that yield overlapping peaks), and reporter dye selection (recommending ANS with $\lambda_{EX}/_{EM} \approx 350$ nm/470 nm over SYPRO Orange due to weaker baseline binding, or utilizing label-free tryptophan fluorescence). Overall, the results validated FTSA as a robust, low sample, high throughput method for routine protein-ligand characterization.

3.2.2 Uncertainty Quantification

This subsection summarizes the results of the publication [138]:

Vaida Paketurytė, Vytautas Petrauskas, Asta Zubrienė, Olga Abian, Margarida Bastos, Wen-Yih Chen, Maria João Moreno, Georg Krainer, Vaida Linkuvienė, Arthur Sedivy, Adrian Velazquez-Campoy, Mark A. Williams, Daumantas Matulis

Uncertainty in protein-ligand binding constants: asymmetric

confidence intervals versus standard errors

Eur. Biophys. J., 2021, 50 (3–4), 661–670.

DOI: 10.1007/s00249-021-01518-4

This work was a task of the COST Action CA15126 (MOBIEU), emerging from collaborative discussions within the European biophysics community to improve uncertainty reporting in molecular interactions. My contribution to this effort involved participation in scientific discussions, experimental planning, data analysis, and the preparation of the manuscript.

The study demonstrates that symmetric standard errors, commonly reported as $\pm x$, substantially underestimate the true uncertainty of protein–ligand binding constants. Through the analysis of simulated and experimental datasets, the work shows that the distributions of the equilibrium binding constant (K_b), dissociation constant (K_d), and Gibbs energy change (ΔG) are inherently asymmetric due to their logarithmic, exponential, and reciprocal mathematical relationships. Specifically, in Case 1, assuming a normal distribution for K_b with a mean of $2.0 \times 10^7 \text{ M}^{-1}$ and a standard deviation of $0.6 \times 10^7 \text{ M}^{-1}$ leads, upon transformation to K_d , to a strongly skewed distribution (skewness 4.3). Moreover, because a normal distribution has unbounded support, it assigns a nonzero probability to physically impossible negative values of K_b . In contrast, if normality is assumed for ΔG , with a mean of $-41.67 \text{ kJ mol}^{-1}$ at 25°C and a standard deviation of 0.77 kJ mol^{-1} , the symmetric 95 % confidence interval is $-43.18 \text{ kJ mol}^{-1}$ to $-40.16 \text{ kJ mol}^{-1}$. Direct transformation of these limits using the thermodynamic relation between ΔG and K_b preserves the corresponding uncertainty structure and gives the asymmetric K_b 95 % confidence interval $[1.09, 3.68] \times 10^7 \text{ M}^{-1}$. Transformations between these parameters amplify skewness, making symmetric error reporting statistically inappropriate and potentially leading to physically impossible results, such as negative binding constants. Monte Carlo simulations of 1:1 binding titrations reveal that fitted K_b values become increasingly positively skewed at higher affinities, whereas ΔG remains closer to a normal distribution only under reasonably optimal experimental conditions. To evaluate experimentally realistic titrations, Case 2 modeled VP-ITC experiments consisting of $20 \times 15 \mu\text{L}$ injections of $200 \mu\text{M}$ ligand into $20 \mu\text{M}$ protein, with $\Delta H = -10 \text{ kJ mol}^{-1}$ and a constant injection-heat noise of $3 \mu\text{J}$. Under these conditions, increasing the true K_b from $2 \times 10^5 \text{ M}^{-1}$ (Wiseman parameter $c = 4$) to $2 \times 10^7 \text{ M}^{-1}$ ($c = 400$) markedly increased the skewness of the fitted K_b distribution from 1.3 ($\text{CI}_{95} = [0.78, 5.8] \times 10^5 \text{ M}^{-1}$) to 23 ($\text{CI}_{95} = [0.51, 22] \times 10^7 \text{ M}^{-1}$). At the high- c value, only 3–4 data points sample the transition region

of the binding isotherm. Consequently, the curvature that constrains K_b is poorly resolved, making its fitted value highly sensitive to experimental noise and producing a strongly asymmetric distribution with a pronounced upper tail.

We conclude that profile likelihood-based 95 % confidence intervals ($CI_{P,95}$) offer a more consistent and inter-convertible method for representing uncertainty across all thermodynamic quantities. These asymmetric intervals are calculated using the Fisher-Snedecor F -distribution by systematically stepping a selected parameter through fixed values while minimizing the residual sum of squares (RSS) over the remaining free parameters. Unlike standard errors, this approach ensures a perfect correspondence between the upper and lower uncertainty limits when transforming between K_b and ΔG , preserving all information regarding experimental variability. For example, for a simulated titration with a true $K_b = 2 \times 10^6, \text{M}^{-1}$ and $\Delta H = 41.8, \text{kJ mol}^{-1}$, the profile likelihood method ($CI_{P,95}$) produces identical corresponding intervals $([1.44, 3.30] \times 10^6, \text{M}^{-1}$ for K_b and $[-37.2, -35.1], \text{kJ mol}^{-1}$ for ΔG) regardless of which parameter is directly fitted. In contrast, standard asymptotic-symmetric intervals ($CI_{S,95}$) fail to maintain mathematical correspondence after parameter transformation. Ultimately, the results establish that asymmetric confidence intervals are essential for the accurate interpretation and reusability of biophysical data and should replace standard symmetric errors in protein-ligand thermodynamic studies.

Discussion

The ultimate goal of structure-based drug design is to predict how chemical modifications influence affinity and selectivity and to use this knowledge to design compounds that bind strongly to a desired target while avoiding interactions with other proteins. Despite major advances in structural biology, computational chemistry, and biophysics, this objective remains difficult to achieve. The results presented in this dissertation highlight several factors that complicate the interpretation and prediction of protein–ligand recognition.

The results presented in Section 3.1.1 Enthalpy-Entropy Compensation provide additional insight into the molecular origins of enthalpy–entropy compensation. Previous studies by Whitesides and co-workers [9, 139] demonstrated that relatively small ligand modifications may produce large changes in enthalpic and entropic contributions while having only modest effects on overall affinity for CA II, highlighting the importance of hydration and solvent reorganization in molecular recognition. The present results show that substantial differences in thermodynamic partitioning can also arise between closely related carbonic anhydrase isozymes despite highly conserved active-site architectures and similar ligand binding modes. This observation suggests that enthalpy–entropy compensation cannot be attributed solely to ligand-dependent effects but is strongly influenced by the local protein environment. Even subtle differences in active-site residues, hydration networks, or protein dynamics are sufficient to alter the balance between enthalpic and entropic contributions. Consequently, the thermodynamic signature of binding should be regarded as an emergent property of the entire protein–ligand system rather than a direct consequence of individual molecular interactions.

From a practical perspective, these observations also illustrate both the strengths and limitations of thermodynamic analysis. While decomposition of binding energy into enthalpic and entropic contributions

can reveal differences between protein–ligand complexes that appear equivalent when evaluated solely by affinity, interpretation of these contributions remains challenging because enthalpy and entropy reflect the combined effects of numerous coupled processes. As discussed by Klebe and others [140], assignment of individual thermodynamic contributions to specific molecular interactions is often not straightforward because changes in intermolecular contacts, hydration, conformational flexibility, and dynamics occur simultaneously. Nevertheless, the ability of thermodynamic signatures to distinguish between systems with similar structures and affinities demonstrates their value as sensitive probes of molecular recognition and highlights the complementary information they provide beyond conventional affinity measurements.

The findings described in Section 3.1.2 Affinity and Selectivity further emphasize the importance of conformational effects in protein–ligand recognition [141]. Numerous studies have demonstrated that relatively small structural modifications may alter ligand flexibility, conformational preferences, and protein dynamics, thereby producing unexpectedly large effects on binding. Gilson and co-workers [3, 10] have highlighted the importance of conformational entropy in molecular recognition and the challenges associated with accurately quantifying its contribution to binding, while Wand [142] and Akke [6, 70] and their colleagues have demonstrated through NMR studies that changes in protein or ligand dynamics can contribute substantially to binding thermodynamics. More recently, Wankowicz and co-workers [66] showed that conformational heterogeneity inferred also from X-ray crystallographic data can provide valuable insight into the entropic consequences of ligand binding. However, despite these advances, experimental characterization of conformational entropy remains considerably more demanding than determination of overall binding thermodynamics. In routine ITC experiments, only the total entropy change is obtained indirectly from measured Gibbs energies and enthalpies, whereas the individual contributions arising from protein conformational changes, ligand flexibility, solvent reorganization, and other processes generally cannot be separated. Consequently, the present observations illustrate both the importance of conformational effects and the current limitations in their characterization. The pronounced affinity losses observed upon oxidation of sulfur-containing substituents in the present study illustrate how subtle changes in ligand conformation may outweigh the expected increase in affinity associated with the introduction of electron-withdrawing substituents.

The results presented in Section 3.1.3 Structure-Thermodynamics Re-

relationships highlight the importance of distinguishing between observed and intrinsic thermodynamic parameters when interpreting protein–ligand interactions. The influence of linked protonation equilibria on binding thermodynamics has been recognized for decades, and several studies by Baker and Murphy [11], Whitesides and co-workers [143], Klebe and co-workers [140], as well as researchers in our laboratory [134] have emphasized that observed thermodynamic parameters may not directly reflect the energetics of molecular recognition. Nevertheless, intrinsic thermodynamic analysis remains relatively uncommon in contemporary protein–ligand studies because determination of the required protonation parameters is experimentally demanding and such information is unavailable for many protein–ligand systems [140, 144]. As a result, observed thermodynamic parameters are frequently interpreted as though they directly describe protein–ligand interactions. The present results demonstrate that protonation-linked contributions can substantially alter both the magnitude and interpretation of binding thermodynamics, suggesting that meaningful comparisons between ligands, proteins, and binding mechanisms often require consideration of intrinsic rather than observed parameters. More broadly, these findings illustrate that accurate structure–thermodynamics relationships depend not only on high-resolution structural information but also on rigorous separation of the underlying processes contributing to the measured thermodynamic signal.

The methodological studies described in Sections 3.2.1 FTSA Standardization and 3.2.2 Uncertainty Quantification emphasize the importance of experimental design and uncertainty quantification for reliable interpretation of protein–ligand binding data. Analysis of inhibition values reported in the literature [145] revealed substantial variability between independent studies, highlighting the challenges associated with comparing data generated using different experimental protocols and conditions. Such variability is particularly important for computational approaches, which frequently rely on published affinity data for model development, validation, and benchmarking. Inaccurate or inconsistent experimental measurements may therefore limit not only experimental interpretation but also predictive modeling efforts. Models trained on noisy experimental data cannot be expected to achieve predictive accuracy exceeding the quality of the underlying measurements. Similar concerns have been raised in the carbonic anhydrase field, where Liljas and co-workers [146] questioned the suitability of some widely cited inhibition datasets and argued that the precision and reliability of the reported measurements may have been overestimated. These observations demonstrate that meaningful structure–activity and

structure–thermodynamics relationships require not only high-quality experimental measurements but also realistic assessment of their uncertainty. While experimental uncertainty cannot be eliminated entirely [147–152], its rigorous quantification [153] is essential for distinguishing genuine biological effects from methodological variability and for enabling meaningful comparisons between studies.

Although the studies presented in this dissertation addressed different aspects of protein–ligand interactions, they collectively support the same central conclusion: binding affinity is a composite observable arising from multiple interconnected processes. Hydration, conformational changes, protonation equilibria, and experimental uncertainty can each substantially influence the interpretation of molecular recognition. Consequently, establishing meaningful structure–thermodynamics relationships requires more than structural information and affinity measurements alone; it also requires consideration of the processes that contribute to the observed experimental parameters. The results presented in this dissertation highlight both the potential and the limitations of structure-based drug design and demonstrate that accurate understanding of protein–ligand interactions depends on the ability to deconvolute observed thermodynamic parameters into their underlying physical contributions and relate them to structural information. Further progress in this field will depend on the ability to deconvolute observed thermodynamic parameters into the individual processes that govern binding, to understand their interplay, and to relate them to structural information.

Conclusions

1. Comparisons of closely related CAs revealed that minor differences in the active-site, including hydration and the presence of either a hydroxyl-bearing Thr or a methyl-bearing Val residue, were associated with the pronounced differences in ΔH and $T\Delta S$ contributions to binding ($\approx 40 \text{ kJ mol}^{-1}$ despite similar binding geometries and ΔG). *These findings highlight the sensitivity of binding thermodynamics to local protein environments and illustrate the limitations of interpreting molecular recognition from affinity or structure alone.* (Publication 1)
2. Small chemical modifications produced marked changes in affinity and selectivity across CAs. Structural and thermodynamic analyses indicated that oxidation of sulfur-containing substituents promoted intramolecular hydrogen-bond formation with the sulfonamide group, altering ligand conformation and reducing productive interactions. *These findings demonstrate the highly cooperative nature of protein–ligand recognition.* (Publication 2)
3. Correction for linked protonation equilibria produced intrinsic thermodynamic parameters that were affected considerably from the observed values, demonstrating that meaningful comparisons of binding energetics require separation of binding and protonation contributions. (Publication 3)
4. Systematic evaluation of FTSA conditions demonstrated the importance of protocol standardization for obtaining comparable protein–ligand binding data. Asymmetric profile-likelihood confidence intervals provided a statistically appropriate representation of uncertainty, whereas symmetric standard errors are misleading for quantities transformed between affinity and thermodynamic scales. (Publication 4 and Publication 5)

Bibliography

- [1] Gerhard Klebe. *Drug Design: From Structure and Mode-of-Action to Rational Design Concepts*. Springer Berlin Heidelberg, 2024. ISBN 978-3-662-68997-4 978-3-662-68998-1. doi: 10.1007/978-3-662-68998-1.
- [2] William L. Jorgensen. The many roles of computation in drug discovery. *Science*, 303(5665):1813–1818, 2004. ISSN 1095-9203. doi: 10.1126/science.1096361.
- [3] Huan-Xiang Zhou and Michael K. Gilson. Theory of free energy and entropy in noncovalent binding. *Chemical Reviews*, 109(9):4092–4107, 2009. ISSN 0009-2665, 1520-6890. doi: 10.1021/cr800551w.
- [4] David L. Mobley and Michael K. Gilson. Predicting binding free energies: Frontiers and benchmarks. *Annual Review of Biophysics*, 46:531–558, 2017. ISSN 1936-1238. doi: 10.1146/annurev-biophys-070816-033654.
- [5] Jack D. Dunitz. Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. *Chemistry & Biology*, 2(11):709–712, 1995. ISSN 1074-5521. doi: 10.1016/1074-5521(95)90097-7.
- [6] Johan Wallerstein, Vilhelm Ekberg, Majda Misini Ignjatović, Rohit Kumar, Octav Caldararu, Kristoffer Peterson, Sven Wernersson, Ulrika Brath, Hakon Leffler, Esko Oksanen, Derek T. Logan, Ulf J. Nilsson, Ulf Ryde, and Mikael Akke. Entropy–entropy compensation between the protein, ligand, and solvent degrees of freedom fine-tunes affinity in ligand binding to galectin-3c. *JACS Au*, 1(4): 484–500, 2021. doi: 10.1021/jacsau.0c00094.
- [7] J. E. Ladbury. Just add water! the effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. *Chemistry & Biology*, 3(12):973–980, 1996. ISSN 1074-5521.
- [8] Gerhard Klebe. Applying thermodynamic profiling in lead finding and optimization. *Nature Reviews Drug Discovery*, 14(2):95–110, 2015. ISSN 1474-1776. doi: 10.1038/nrd4486.
- [9] Benjamin Breiten, Matthew R. Lockett, Woody Sherman, Shuji

- Fujita, Mohammad Al-Sayah, Heiko Lange, Carleen M. Bowers, Annie Heroux, Goran Krilov, and George M. Whitesides. Water networks contribute to enthalpy/entropy compensation in protein-ligand binding. *Journal of the American Chemical Society*, 135 (41):15579–15584, 2013. ISSN 1520-5126. doi: 10.1021/ja4075776.
- [10] Chia-en A. Chang, Wei Chen, and Michael K. Gilson. Ligand configurational entropy and protein binding. *Proceedings of the National Academy of Sciences*, 104(5):1534–1539, 2007. doi: 10.1073/pnas.0610494104.
- [11] B M Baker and K P Murphy. Evaluation of linked protonation effects in protein binding reactions using isothermal titration calorimetry. *Biophysical Journal*, 71(4):2049–2055, 1996. ISSN 0006-3495.
- [12] Asta Zubrienė and Daumantas Matulis. Observed versus intrinsic thermodynamics of inhibitor binding to carbonic anhydrases. In Daumantas Matulis, editor, *Carbonic Anhydrase as Drug Target: Thermodynamics and Structure of Inhibitor Binding*, pages 107–123. Springer International Publishing, 2019. ISBN 978-3-030-12780-0. doi: 10.1007/978-3-030-12780-0_8.
- [13] Vijay M. Krishnamurthy, George K. Kaufman, Adam R. Urbach, Irina Gitlin, Katherine L. Gudiksen, Douglas B. Weibel, and George 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, 2008. ISSN 0009-2665. doi: 10.1021/cr050262p.
- [14] Olivier J. Wouters, Martin McKee, and Jeroen Luyten. Estimated research and development investment needed to bring a new medicine to market, 2009-2018. 323(9):844–853, 2020. ISSN 0098-7484. doi: 10.1001/jama.2020.1166.
- [15] Colin G. Rousseaux, William M. Bracken, and Silvia Guionaud. Chapter 1 - overview of drug development. In Wanda M. Haschek, Colin G. Rousseaux, Matthew A. Wallig, and Brad Bolon, editors, *Haschek and Rousseaux's Handbook of Toxicologic Pathology (Fourth Edition)*, pages 3–48. Academic Press, 2023. ISBN 978-0-12-821047-5. doi: 10.1016/B978-0-12-821047-5.00016-6.
- [16] Approval success rates of drug candidates based on target, action, modality, application, and their combinations. 14(3):1113–1122, 2021. ISSN 1752-8054. doi: 10.1111/cts.12980.
- [17] A year in pharmacology: new drugs approved by the US food and drug administration in 2025. 399:11157–11179, 2026. ISSN 0028-1298. doi: 10.1007/s00210-026-05193-0.
- [18] How were new medicines discovered? 10:507–519, 2011. ISSN 1474-1784. doi: 10.1038/nrd3480.

- [19] The discovery of first-in-class drugs: origins and evolution. 13(8): 577–587, 2014. ISSN 1474-1784. doi: 10.1038/nrd4336.
- [20] Visvaldas Kairys, Lina Baranauskiene, Migle Kazlauskiene, Asta Zubrienė, Vytautas Petrauskas, Daumantas Matulis, and Egidijus Kazlauskas. Recent advances in computational and experimental protein-ligand affinity determination techniques. *Expert Opinion on Drug Discovery*, 19(6):649–670, 2024. ISSN 1746-0441. doi: 10.1080/17460441.2024.2349169.
- [21] Stephen K. Burley, Amy Wu-Wu, Shuchismita Dutta, Shridar Ganesan, and Steven X. F. Zheng. Impact of structural biology and the protein data bank on us fda new drug approvals of low molecular weight antineoplastic agents 2019–2023. 43(29):2229–2243, 2024. ISSN 1476-5594. doi: 10.1038/s41388-024-03077-2.
- [22] Caught between a ROCK and a hard place: current challenges in structure-based drug design. 29(9):104106, 2024. ISSN 1359-6446. doi: 10.1016/j.drudis.2024.104106.
- [23] A critical assessment of docking programs and scoring functions. 49:5912–5931, 2005. ISSN 0022-2623. doi: 10.1021/jm050362n.
- [24] Comparative assessment of scoring functions: The CASF-2016 update. 59(2):895–913, 2018. ISSN 1549-9596. doi: 10.1021/acs.jcim.8b00545.
- [25] The last mile problem: A critical assessment of physics-based and AI tools for small molecule binding prediction in virtual screening. 66(2026):5964–5977. ISSN 1549-960X. doi: 10.1021/acs.jcim.6c00942.
- [26] Jean-Paul Renaud, Chun-wa Chung, U. Helena Danielson, Ursula Egner, Michael Hennig, Roderick E. Hubbard, and Herbert Nar. Biophysics in drug discovery: Impact, challenges and opportunities. *Nature Reviews Drug Discovery*, 15(10):679–698, 2016. ISSN 1474-1776. doi: 10.1038/nrd.2016.123.
- [27] Remo Perozzo, Gerd Folkers, and Leonardo Scapozza. Thermodynamics of protein-ligand interactions: History, presence, and future aspects. *Journal of Receptor and Signal Transduction Research*, 24(1):1–52, 2004. ISSN 1079-9893.
- [28] Bert L. Vallee and David S. Auld. Zinc coordination, function, and structure of zinc enzymes and other proteins. *Biochemistry*, 29(24): 5647–5659, 1990. ISSN 0006-2960. doi: 10.1021/bi00476a001.
- [29] Protein-ligand interactions: Hydrodynamics and calorimetry, 2001.
- [30] Caterina Bissantz, Bernd Kuhn, and Martin Stahl. A medicinal chemist’s guide to molecular interactions. *Journal of Medicinal Chemistry*, 53(14):5061–5084, 2010. ISSN 0022-2623, 1520-4804. doi: 10.1021/jm100112j.

- [31] S. Lindskog. Structure and mechanism of carbonic anhydrase. *Pharmacology & Therapeutics*, 74(1):1–20, 1997. ISSN 0163-7258. doi: 10.1016/s0163-7258(96)00198-2.
- [32] A. Liljas, K. K. Kannan, P. C. Bergsten, I. Waara, K. Fridborg, B. Strandberg, U. Carlbom, L. Jarup, S. Lovgren, and M. Petef. Crystal structure of human carbonic anhydrase c. *Nat. New. Biol.*, 235(57):131–137, 1972. doi: 10/gjv9sf.
- [33] Claudiu T. Supuran. Carbonic anhydrases—an overview. *Current Pharmaceutical Design*, 14(7):603–614, 2008. ISSN 1873-4286. doi: 10.2174/138161208783877884.
- [34] Vincenzo Alterio, Anna Di Fiore, Katia D’Ambrosio, Claudiu T. Supuran, and Giuseppina De Simone. Multiple binding modes of inhibitors to carbonic anhydrases: How to design specific drugs targeting 15 different isoforms? *Chemical Reviews*, 112(8):4421–4468, 2012. ISSN 0009-2665, 1520-6890. doi: 10.1021/cr200176r.
- [35] Denis Baronas, Birutė Knašienė, Aurelija Mickevičiūtė, Jelena Jachno, Evaldas Naujalis, Asta Zubrienė, and Daumantas Matulis. Inhibitor binding to metal-substituted metalloenzyme: Sulfonamide affinity for carbonic anhydrase IX. *Journal of Inorganic Biochemistry*, 256:112547, 2024. ISSN 0162-0134. doi: 10.1016/j.jinorgbio.2024.112547.
- [36] Tjelvar S. G. Olsson, Mark A. Williams, William R. Pitt, and John E. Ladbury. The thermodynamics of protein-ligand interaction and solvation: insights for ligand design. *J. Mol. Biol.*, 384(4):1002–1017, 2008. ISSN 1089-8638. doi: 10.1016/j.jmb.2008.09.073.
- [37] George M. Whitesides and Vijay M. Krishnamurthy. Designing ligands to bind proteins. *Quarterly Reviews of Biophysics*, 38(4): 385–395, 2005. ISSN 0033-5835. doi: 10.1017/S0033583506004240.
- [38] Gerhard Klebe. The use of thermodynamic and kinetic data in drug discovery: decisive insight or increasing the puzzlement? *ChemMedChem*, 10(2):229–231, 2015. ISSN 1860-7187. doi: 10.1002/cmdc.201402521.
- [39] Steve W. Homans. Water, water everywhere — except where it matters? *Drug Discovery Today*, 12(13):534–539, 2007. ISSN 1359-6446. doi: 10.1016/j.drudis.2007.05.004.
- [40] Tom Young, Robert Abel, Byungchan Kim, Bruce J. Berne, and Richard A. Friesner. Motifs for molecular recognition exploiting hydrophobic enclosure in protein–ligand binding. *Proceedings of the National Academy of Sciences*, 104(3):808–813, 2007. doi: 10.1073/pnas.0610202104.
- [41] Robert Abel, Tom Young, Ramy Farid, Bruce J. Berne, and Richard A. Friesner. Role of the active-site solvent in the thermodynamics of factor xa ligand binding. *Journal of the American*

- Chemical Society*, 130(9):2817–2831, 2008. ISSN 0002-7863. doi: 10.1021/ja0771033.
- [42] D. N. Silverman and S. Lindskog. The catalytic mechanism of carbonic anhydrase: implications of a rate-limiting proteolysis of water. *Accounts of Chemical Research*, 21:30–36, 1988. doi: doi.org/10.1021/ar00145a005.
 - [43] Balendu Sankara Avvaru, Chae Un Kim, Katherine H Sippel, Sol M Gruner, Mavis Agbandje-McKenna, David N Silverman, and Robert McKenna. A short, strong hydrogen bond in the active site of human carbonic anhydrase II. *Biochemistry*, 49(2):249–251, 2010. doi: 10.1021/bi902007b.
 - [44] Melissa A. Pinard, Brian Mahon, and Robert McKenna. Probing the surface of human carbonic anhydrase for clues towards the design of isoform specific inhibitors. *BioMed Research International*, 2015:e453543, 2015. ISSN 2314-6133. doi: 10.1155/2015/453543.
 - [45] Zoë Fisher, Andrey Y. Kovalevsky, Marat Mustyakimov, David N. Silverman, Robert McKenna, and Paul Langan. Neutron structure of human carbonic anhydrase II: a hydrogen-bonded water network "switch" is observed between pH 7.8 and 10.0. *Biochemistry*, 50(44):9421–9423, 2011. doi: 10.1021/bi201487b.
 - [46] Ryszard Michalczyk, Clifford J Unkefer, John-Paul Bacik, Tobias E Schrader, Andreas Ostermann, Andrey Y Kovalevsky, Robert McKenna, and Suzanne Zoë Fisher. Joint neutron crystallographic and NMR solution studies of tyr residue ionization and hydrogen bonding: Implications for enzyme-mediated proton transfer. *Proc. Natl. Acad. Sci. U.S.A.*, 112(18):5673–5678, 2015. doi: 10.1073/pnas.1502255112.
 - [47] Chae Un Kim, HyoJin Song, Balendu Sankara Avvaru, Sol M. Gruner, SangYoun Park, and Robert McKenna. Tracking solvent and protein movement during CO₂ release in carbonic anhydrase II crystals. *Proceedings of the National academy of Sciences of the United States of America*, 113(19):5257–5262, 2016. doi: 10.1073/pnas.1520786113.
 - [48] J. K. Kim, C. L. Lomelino, B. S. Avvaru, B. P. Mahon, R. McKenna, S. Park, and C. U. Kim. Active-site solvent replenishment observed during human carbonic anhydrase II catalysis. *IUCrJ*, 5(1):93–102, 2018. ISSN 2052-2525. doi: 10.1107/S2052252517017626.
 - [49] Rose Mikulski, Dayne West, Katherine H. Sippel, Balendu Sankara Avvaru, Mayank Aggarwal, Chingkuang Tu, Robert McKenna, and David N. Silverman. Water networks in fast proton transfer during catalysis by human carbonic anhydrase II. *Biochemistry*, 52(1):125–131, 2013. ISSN 0006-2960, 1520-4995. doi: 10.1021/bi301099k.

- [50] David N. Silverman and Robert McKenna. Solvent-mediated proton transfer in catalysis by carbonic anhydrase. *Accounts of Chemical Research*, 40(8):669–675, 2007. ISSN 0001-4842. doi: 10.1021/ar7000588.
- [51] Jin Kyun Kim, Seon Woo Lim, Hannah Jeong, Cheol Lee, Seoyoon Kim, Dong Woo Son, Rajeev Kumar, Jacob T. Andring, Carrie Lomelino, Jennifer L. Wierman, Aina E. Cohen, Tae Joo Shin, Cheol-Min Ghim, Robert McKenna, Byung Hoon Jo, Duyoung Min, Jeong-Mo Choi, and Chae Un Kim. Fast product release requires active-site water dynamics in carbonic anhydrase. *Nature Communications*, 16(1):4404, 2025. ISSN 2041-1723. doi: 10.1038/s41467-025-59645-x.
- [52] M. T. Record, C. F. Anderson, and T. M. Lohman. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. *Quarterly Reviews of Biophysics*, 11(2):103–178, 1978. ISSN 0033-5835. doi: 10.1017/s003358350000202x.
- [53] Kim A. Sharp. Polyelectrolyte electrostatics: Salt dependence, entropic, and enthalpic contributions to free energy in the nonlinear poisson-boltzmann model. *Biopolymers*, 36(2):227–243, 1995. ISSN 1097-0282. doi: 10.1002/bip.360360210.
- [54] Ken A. Dill. Dominant forces in protein folding. *Biochemistry*, 29(31):7133–7155, 1990. ISSN 0006-2960. doi: 10.1021/bi00483a001.
- [55] Adam Biela, Nader N. Nasief, Michael Betz, Andreas Heine, David Hangauer, and Gerhard Klebe. Dissecting the hydrophobic effect on the molecular level: the role of water, enthalpy, and entropy in ligand binding to thermolysin. *Angewandte Chemie (International Ed. in English)*, 52(6):1822–1828, 2013. ISSN 1521-3773. doi: 10.1002/anie.201208561.
- [56] Phillip W. Snyder, Jasmin Mecinović, Demetri T. Moustakas, Samuel W. Thomas, Michael Harder, Eric T. Mack, Matthew R. Lockett, Annie Héroux, Woody Sherman, and George M. Whitesides. Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase. *Proceedings of the National Academy of Sciences of the United States of America*, 108(44):17889–17894, 2011. ISSN 0027-8424. doi: 10.1073/pnas.1114107108.
- [57] David Chandler. Interfaces and the driving force of hydrophobic assembly. *Nature*, 437(7059):640–647, 2005. ISSN 1476-4687. doi: 10.1038/nature04162.
- [58] Steffen Glöckner, Khang Ngo, Björn Wagner, Andreas Heine, and Gerhard Klebe. The influence of varying fluorination patterns on

- the thermodynamics and kinetics of benzenesulfonamide binding to human carbonic anhydrase II. *Biomolecules*, 10(4):509, 2020. doi: 10.3390/biom10040509.
- [59] Sophie Purser, Peter R. Moore, Steve Swallow, and Véronique Gouverneur. Fluorine in medicinal chemistry. *Chemical Society Reviews*, 37(2):320–330, 2008. ISSN 0306-0012. doi: 10.1039/b610213c.
 - [60] Miguel Bernús, Gonzalo D. Núñez, Will C. Hartley, Marc Guasch, Jordi Mestre, Maria Besora, Jorge J. Carbó, and Omar Boutureira. Impact of fluorine pattern on lipophilicity and acid–base properties of 2-(thiofluoroalkyl)pyridines: Insights from experiments and statistical modeling. *Journal of Medicinal Chemistry*, 68(4):4787–4800, 2025. ISSN 0022-2623. doi: 10.1021/acs.jmedchem.4c03045.
 - [61] Xing Du, Yi Li, Yuan-Ling Xia, Shi-Meng Ai, Jing Liang, Peng Sang, Xing-Lai Ji, and Shu-Qun Liu. Insights into Protein–Ligand Interactions: Mechanisms, Models, and Methods. *International Journal of Molecular Sciences*, 17(2), 2016. ISSN 1422-0067. doi: 10.3390/ijms17020144.
 - [62] Sven Lindskog and David N. Silverman. The catalytic mechanism of mammalian carbonic anhydrases. In Prof Dr W. Richard Chegwidden, Nicholas D. Carter, and Yvonne H. Edwards, editors, *The Carbonic Anhydrases*, number 90 in EXS 90, pages 175–195. Birkhäuser Basel, 2000. ISBN 978-3-0348-9570-5 978-3-0348-8446-4.
 - [63] Virginija Dudutienė, Asta Zubrienė, Visvaldas Kairys, Alexey Smirnov, Joana Smirnovienė, Janis Leitans, Andris Kazaks, Kaspars Tars, Lena Manakova, Saulius Gražulis, and Daumantas Matulis. Isoform-selective enzyme inhibitors by exploring pocket size according to the lock-and-key principle. *Biophysical Journal*, 119: 1513–1524, 2020. ISSN 1542-0086. doi: 10.1016/j.bpj.2020.08.037.
 - [64] Kendra King Frederick, Michael S. Marlow, Kathleen G. Valentine, and A. Joshua Wand. Conformational entropy in molecular recognition by proteins. *Nature*, 448(7151):325–329, 2007. ISSN 1476-4687. doi: 10.1038/nature05959.
 - [65] A Joshua Wand. The dark energy of proteins comes to light: conformational entropy and its role in protein function revealed by NMR relaxation. *Current Opinion in Structural Biology*, 23(1): 75–81, 2013. ISSN 0959-440X. doi: 10.1016/j.sbi.2012.11.005.
 - [66] Charlotte A Miller and Stephanie A Wankowicz. Binding entropy can be predicted by crystallographic ensembles.
 - [67] Stefan Senger, Chuen Chan, Máire A. Convery, Julia A. Hubbard, Gita P. Shah, Nigel S. Watson, and Robert J. Young. Sulfonamide-related conformational effects and their importance in structure-based design. *Bioorganic & Medicinal Chemistry Letters*, 17(10): 2931–2934, 2007. ISSN 0960-894X. doi: 10.1016/j.bmcl.2007.02.034.

- [68] David D Boehr, Ruth Nussinov, and Peter E Wright. The role of dynamic conformational ensembles in biomolecular recognition. *Nature Chemical Biology*, 5(11):789–796, 2009. ISSN 1552-4450, 1552-4469. doi: 10.1038/nchembio.232.
- [69] Ruth Nussinov, Buyong Ma, and Chung-Jung Tsai. Multiple conformational selection and induced fit events take place in allosteric propagation. *Biophysical chemistry*, 186:22–30, 2014. ISSN 0301-4622. doi: 10.1016/j.bpc.2013.10.002.
- [70] Maria Luisa Verteramo, Olof Stenström, Majda Misini Ignjatovic, Octav Caldararu, Martin A. Olsson, Francesco Manzoni, Hakon Leffler, Esko Oksanen, Derek T. Logan, Ulf J. Nilsson, Ulf Ryde, and Mikael Akke. Interplay between conformational entropy and solvation entropy in protein–ligand binding. *Journal of the American Chemical Society*, 2019. ISSN 0002-7863. doi: 10.1021/jacs.8b11099.
- [71] D. E. Koshland. Application of a theory of enzyme specificity to protein synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 44(2):98–104, 1958. ISSN 0027-8424. doi: 10.1073/pnas.44.2.98.
- [72] Ernesto Freire. Do enthalpy and entropy distinguish first in class from best in class? *Drug Discovery Today*, 13(19):869–874, 2008. ISSN 1359-6446. doi: 10.1016/j.drudis.2008.07.005.
- [73] Craig A. Behnke, Isolde Le Trong, Jeff W. Godden, Ethan A. Merritt, David C. Teller, Jürgen Bajorath, and Ronald E. Stenkamp. Atomic resolution studies of carbonic anhydrase II. *Acta Crystallogr. D*, 66(5):616–627, 2010. doi: 10.1107/S0907444910006554.
- [74] Himanshu Singh, Chandan K. Das, Suresh K. Vasa, Kristof Grohe, Lars V. Schäfer, and Rasmus Linser. The active site of a prototypical “rigid” drug target is marked by extensive conformational dynamics. *Angewandte Chemie International Edition*, 2020. ISSN 1433-7851, 1521-3773. doi: 10.1002/anie.202009348.
- [75] Himanshu Singh, Suresh K. Vasa, Harish Jangra, Petra Rovó, Christopher Päslock, Chandan K. Das, Hendrik Zipse, Lars V. Schäfer, and Rasmus Linser. Fast microsecond dynamics of the protein–water network in the active site of human carbonic anhydrase II studied by solid-state NMR spectroscopy. *Journal of the American Chemical Society*, 141(49):19276–19288, 2019. ISSN 0002-7863. doi: 10.1021/jacs.9b05311.
- [76] Suresh K. Vasa, Himanshu Singh, Kristof Grohe, and Rasmus Linser. Assessment of a large enzyme–drug complex by proton-detected solid-state NMR spectroscopy without deuteration. *Angewandte Chemie International Edition*, 58(17):5758–5762, 2019. ISSN 1521-3773. doi: 10.1002/anie.201811714.

- [77] Divya Rai, Satyajit Khatua, and Srabani Taraphder. Structure and dynamics of the isozymes II and IX of human carbonic anhydrase. *ACS Omega*, 7(35):31149–31166, 2022. doi: 10.1021/acsomega.2c03356.
- [78] Joshua A. Hull, Cheol Lee, Jin Kyun Kim, Seon Woo Lim, Jaehyun Park, Sehan Park, Sang Jae Lee, Gisu Park, Intae Eom, Minseok Kim, HyoJung Hyun, Jacob E. Combs, Jacob T. Andring, Carrie Lomelino, Chae Un Kim, and Robert McKenna. XFEL structure of carbonic anhydrase II: a comparative study of XFEL, NMR, x-ray and neutron structures. *Acta Crystallographica. Section D, Structural Biology*, 80:194–202, 2024. ISSN 2059-7983. doi: 10.1107/S2059798324000482.
- [79] Vytautas Petrauskas, Lina Baranauskienė, Asta Zubrienė, and Daumantas Matulis. Isothermal Titration Calorimetry and Fluorescent Thermal and Pressure Shift Assays in Protein-Ligand Interactions. In *Biocalorimetry*, pages 261–280. CRC Press, 2016. doi: 10.1201/b20161-18.
- [80] Gediminas Skvarnavičius, Daumantas Matulis, and Vytautas Petrauskas. Change in Volume Upon Inhibitor Binding to Carbonic Anhydrases by Fluorescent Pressure Shift Assay. In Daumantas Matulis, editor, *Carbonic Anhydrase as Drug Target: Thermodynamics and Structure of Inhibitor Binding*, pages 97–106. Springer International Publishing, Cham, 2019. ISBN 978-3-030-12780-0.
- [81] Lehninger, Albert L, David L. Nelson, and Michael M. Cox. *Lehninger Principles of Biochemistry*. Worth Publishers, New York, 5th edition, 2000.
- [82] John E. Ladbury. Calorimetry as a tool for understanding biomolecular interactions and an aid to drug design. *Biochemical Society Transactions*, 38(4):888–893, 2010. ISSN 1470-8752. doi: 10.1042/BST0380888.
- [83] Jerome M Fox, Mengxia Zhao, Michael J Fink, Kyungtae Kang, and George M Whitesides. The Molecular Origin of Enthalpy /Entropy Compensation in Biomolecular Recognition. page 31, 2018.
- [84] N. U. Meldrum and F. J. W. Roughton. Carbonic anhydrase. Its preparation and properties. *The Journal of Physiology*, 80(2):113–142, 1933. ISSN 1469-7793. doi: 10.1113/jphysiol.1933.sp003077.
- [85] Carbonic anhydrase: Mechanism, regulation, links to disease, and industrial applications, 2014.
- [86] The carbonic anhydrases: Current and emerging therapeutic targets, 2021.
- [87] Ashok Aspatwar and Seppo Parkkila. Carbonic anhydrase-related proteins VIII, α , and XI: Evolutionary divergence and emerging biological roles. In *The Enzymes*, volume 59, pages 177–194. Elsevier,

2026. ISBN 978-0-443-47086-8. doi: 10.1016/bs.enz.2026.05.004.
- [88] Carrie L. Lomelino, Jacob T. Andring, and Robert McKenna. Crystallography and Its Impact on Carbonic Anhydrase Research. *International Journal of Medicinal Chemistry*, 2018:9419521, 2018. ISSN 2090-2069. doi: 10.1155/2018/9419521.
 - [89] K. K. Kannan, B. Notstrand, K. Fridborg, S. Lovgren, A. Ohlsson, and M. Petef. Crystal structure of human erythrocyte carbonic anhydrase b. three-dimensional structure at a nominal 2.2-Å resolution. *Proc. Natl. Acad. Sci. U. S. A.*, 72(1):51–55, 1975.
 - [90] A. E. Eriksson, T. A. Jones, and A. Liljas. Refined structure of human carbonic anhydrase II at 2.0 Å resolution. *Proteins*, 4(4): 274–282, 1988. doi: 10.1002/prot.340040406.
 - [91] 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, 1993.
 - [92] Chingkuang Tu, Minzhang Qian, Haiqian An, Nina R. Wadhwa, David Duda, Craig Yoshioka, Yashash Pathak, Robert McKenna, Philip J. Laipis, and David N. Silverman. Kinetic Analysis of Multiple Proton Shuttles in the Active Site of Human Carbonic Anhydrase. *J. Biol. Chem.*, 277(41):38870–38876, 2002. ISSN 0021-9258, 1083-351X. doi: 10.1074/jbc.M20E5791200.
 - [93] R. S. Alexander, L. L. Kiefer, C. A. Fierke, and D. W. Christianson. Engineering the zinc binding site of human carbonic anhydrase II: structure of the his-94→cys apoenzyme in a new crystalline form. *Biochemistry*, 32(6):1510–1518, 1993.
 - [94] Zoë Fisher, Jose A. Hernandez Prada, Chingkuang Tu, David Duda, Craig Yoshioka, Haiqian An, Lakshmanan Govindasamy, David N. Silverman, and Robert McKenna. Structural and kinetic characterization of active-site histidine as a proton shuttle in catalysis by human carbonic anhydrase II. *Biochemistry*, 44(4): 1097–1105, 2005. doi: 10.1021/bi0480279.
 - [95] John F. Domsic, Balendu Sankara Avvaru, Chae Un Kim, Sol M. Gruner, Mavis Agbandje-McKenna, David N. Silverman, and Robert McKenna. Entrapment of carbon dioxide in the active site of carbonic anhydrase II. *J. Biol. Chem.*, 283(45):30766–30771, 2008. doi: 10.1074/jbc.M805353200.
 - [96] A. E. Eriksson and A. Liljas. Refined structure of bovine carbonic anhydrase III at 2.0 Å resolution. *Proteins*, 16(1):29–42, 1993. doi: 10.1002/prot.340160104.
 - [97] T. Stams, S. K. Nair, T. Okuyama, A. Waheed, W. S. Sly, and D. W. Christianson. Crystal structure of the secretory form of membrane-associated human carbonic anhydrase IV at 2.8-Å resolution. *Proc.*

- Natl. Acad. Sci. U. S. A.*, 93(24):13589–13594, 1996.
- [98] Mayank Aggarwal, Christopher D. Boone, Bhargav Kondeti, and Robert McKenna. Structural annotation of human carbonic anhydrases. *J. Enzyme Inhib. Med. Chem.*, 28(2):267–277, 2013. ISSN 1475-6366, 1475-6374. doi: 10.3109/14756366.2012.737323.
 - [99] Ewa S Pilka, Grazyna Kochan, Udo Oppermann, and Wyatt W Yue. Crystal structure of the secretory isozyme of mammalian carbonic anhydrases CA VI: implications for biological assembly and inhibitor development. *Biochem. Biophys. Res. Commun.*, 419(3):485–489, 2012. doi: 10.1016/j.bbrc.2012.02.038.
 - [100] Vincenzo Alterio, Mika Hilvo, Anna Di Fiore, Claudiu T. Supuran, Peiwen Pan, Seppo Parkkila, Andrea Scaloni, Jaromir Pastorek, Silvia Pastorekova, Carlo Pedone, and others. Crystal structure of the catalytic domain of the tumor-associated human carbonic anhydrase IX. *Proc. Natl. Acad. Sci. U. S. A.*, 106(38):16233–16238, 2009.
 - [101] Katarina Koruza, A. Briana Murray, Brian P. Mahon, Jesse B. Hopkins, Wolfgang Knecht, Robert McKenna, and S. Zoë Fisher. Biophysical characterization of cancer-related carbonic anhydrase IX. *International Journal of Molecular Sciences*, 21(15):5277, 2020. ISSN 1422-0067. doi: 10.3390/ijms21155277.
 - [102] D. A. Whittington, A. Waheed, B. Ulmasov, G. N. Shah, J. H. Grubb, W. S. Sly, and D. W. Christianson. Crystal structure of the dimeric extracellular domain of human carbonic anhydrase XII, a bitopic membrane protein overexpressed in certain cancer tumor cells. *Proc. Natl. Acad. Sci. U.S.A.*, 98(17):9545–9550, 2001. doi: 10.1073/pnas.161301298.
 - [103] Anna Di Fiore, Simona Maria Monti, Mika Hilvo, Seppo Parkkila, Vincenza Romano, Andrea Scaloni, Carlo Pedone, Andrea Scozzafava, Claudiu T. Supuran, and Giuseppina De Simone. Crystal structure of human carbonic anhydrase XIII and its complex with the inhibitor acetazolamide. *Proteins*, 74(1):164–175, 2009. doi: 10.1002/prot.22144.
 - [104] Vincenzo Alterio, Peiwen Pan, Seppo Parkkila, Martina Buonanno, Claudiu T. Supuran, Simona M. Monti, and Giuseppina De Simone. The structural comparison between membrane-associated human carbonic anhydrases provides insights into drug design of selective inhibitors. *Biopolymers*, 101(7):769–778, 2014. ISSN 1097-0282. doi: 10.1002/bip.22456.
 - [105] S. Lindskog and J. E. Coleman. The catalytic mechanism of carbonic anhydrase. *Proc Natl Acad Sci U S A*, 70(9):2505–8, 1973.
 - [106] S. Zoë Fisher, Andrey Y. Kovalevsky, John F. Domsic, Marat Mustyakimov, Robert McKenna, David N. Silverman, and Paul A.

- Langan. Neutron Structure of Human Carbonic Anhydrase II: Implications for Proton Transfer. *Biochemistry*, 49(3):415–421, 2010. ISSN 0006-2960. doi: 10.1021/bi901995n.
- [107] S. Zoë Fisher, Mayank Aggarwal, Andrey Y. Kovalevsky, David N. Silverman, and Robert McKenna. Neutron diffraction of acetazolamide-bound human carbonic anhydrase II reveals atomic details of drug binding. *Journal of the American Chemical Society*, 134(36):14726–14729, 2012. ISSN 1520-5126. doi: 10.1021/ja3068098.
- [108] T. Mann and D. Keilin. Sulphanilamide as a Specific Inhibitor of Carbonic Anhydrase. *Nature*, 146(3692):164, 1940. ISSN 1476-4687. doi: 10.1038/146164a0.
- [109] Nicholas A. Meanwell. Fluorine and fluorinated motifs in the design and application of bioisosteres for drug design. *Journal of Medicinal Chemistry*, 61(14):5822–5880, 2018. ISSN 1520-4804. doi: 10.1021/acs.jmedchem.7b01788.
- [110] Fabrizio Carta, Andrea Scozzafava, and Claudiu T. Supuran. Sulfonamides: A patent review (2008 – 2012). *Expert Opinion on Therapeutic Patents*, 22(7):747–758, 2012. ISSN 1354-3776. doi: 10.1517/13543776.2012.698264.
- [111] Claudiu T. Supuran. How many carbonic anhydrase inhibition mechanisms exist? *Journal of Enzyme Inhibition and Medicinal Chemistry*, 31(3):345–360, 2016. ISSN 1475-6366, 1475-6374. doi: 10.3109/14756366.2015.1122001.
- [112] Katia D’Ambrosio, Anna Di Fiore, Vincenzo Alterio, Emma Langella, Simona Maria Monti, Claudiu T. Supuran, and Giuseppina De Simone. Multiple binding modes of inhibitors to human carbonic anhydrases: An update on the design of isoform-specific modulators of activity. *Chemical Reviews*, 125(1):150–222, 2025. ISSN 0009-2665. doi: 10.1021/acs.chemrev.4c00278.
- [113] Sarfraz Ahmad, Muhammad Usman Mirza, Mamoon Nazir, Lee Yean Kee, Noorsaadah Binti Abdrahman, Iskandar Abdullah, and John F. Trant. A comprehensive review of the structural basis for the selectivity of sulfonamide-based inhibitors for carbonic anhydrase isoforms IX and XII over carbonic anhydrase II based on the available high quality crystallographic data. *Archives of Biochemistry and Biophysics*, 777:110682, 2026. ISSN 0003-9861. doi: 10.1016/j.abb.2025.110682.
- [114] Tautvydas Kojis, Helgi B. Schiöth, Lina Baranauskiene, and Daurantas Matulis. Trends in drug discovery for carbonic anhydrases: FDA approvals, clinical trials, molecular functions, and therapeutic potential. *Pharmacological Research*, 228:108210, 2026. ISSN 10436618. doi: 10.1016/j.phrs.2026.108210.
- [115] Aurelija Mickevičiūtė, Vaida Juozapaitienė, Vilma Michailovienė,

- Jelena Jachno, Jurgita Matulienė, and Daumantas Matulis. Recombinant production of 12 catalytically active human CA isoforms. In Daumantas Matulis, editor, *Carbonic Anhydrase as Drug Target: Thermodynamics and Structure of Inhibitor Binding*, pages 15–37. Springer International Publishing, 2019. ISBN 978-3-030-12780-0. doi: 10.1007/978-3-030-12780-0_2.
- [116] Egidijus Kazlauskas, Vytautas Petrauskas, Vaida Paketurytė, and Daumantas Matulis. Standard operating procedure for fluorescent thermal shift assay (FTSA) for determination of protein–ligand binding and protein stability. *European Biophysics Journal*, 50(3):373–379, 2021. ISSN 0175-7571, 1432-1017. doi: 10.1007/s00249-021-01537-1.
- [117] Vytautas Petrauskas, Egidijus Kazlauskas, Marius Gedgaudas, Lina Baranauskienė, Asta Zubrienė, and Daumantas Matulis. Thermal shift assay for protein–ligand dissociation constant determination. *TrAC Trends in Analytical Chemistry*, 170:117417, 2024. ISSN 0165-9936. doi: 10.1016/j.trac.2023.117417.
- [118] Asta Zubrienė, Egidijus Kazlauskas, Lina Baranauskienė, Vytautas Petrauskas, and Daumantas Matulis. Isothermal titration calorimetry and thermal shift assay in drug design. *European Pharmaceutical Review*, 16:56–59, 2011.
- [119] Andrea Ilari and Carmelinda Savino. Protein structure determination by x-ray crystallography. *Methods in Molecular Biology (Clifton, N.J.)*, 452:63–87, 2008. ISSN 1064-3745. doi: 10.1007/978-1-60327-159-2_3.
- [120] Alexey Smirnov, Elena Manakova, Saulius Gražulis, Robert McKenna, and Daumantas Matulis. Structures of human carbonic anhydrases and their complexes with inhibitors. In Daumantas Matulis, editor, *Carbonic Anhydrase as Drug Target: Thermodynamics and Structure of Inhibitor Binding*, pages 179–202. Springer International Publishing, 2019. ISBN 978-3-030-12780-0. doi: 10.1007/978-3-030-12780-0_13.
- [121] Kaspars Tars and Daumantas Matulis. X-ray crystallographic structures of high-affinity and high-selectivity inhibitor complexes with CA IX that plays a special role in cancer. In Daumantas Matulis, editor, *Carbonic Anhydrase as Drug Target: Thermodynamics and Structure of Inhibitor Binding*, pages 203–213. Springer International Publishing, 2019. ISBN 978-3-030-12780-0. doi: 10.1007/978-3-030-12780-0_14.
- [122] Joana Smirnovienė, Lina Baranauskienė, Asta Zubrienė, and Daumantas Matulis. A standard operating procedure for an enzymatic activity inhibition assay. *European Biophysics Journal*, 50(3):345–352, 2021. ISSN 1432-1017. doi: 10.1007/s00249-021-01530-8.

- [123] Visvaldas Kairys, Lina Baranauskiene, Migle Kazlauskiene, Daumantas Matulis, and Egidijus Kazlauskas. Binding affinity in drug design: experimental and computational techniques. *Expert Opinion on Drug Discovery*, pages 1–14, 2019. ISSN 1746-0441, 1746-045X. doi: 10.1080/17460441.2019.1623202.
- [124] Joana Smirnovienė, Vytautas Smirnovas, and Daumantas Matulis. Picomolar inhibitors of carbonic anhydrase: Importance of inhibition and binding assays. *Anal. Biochem.*, 522:61–72, 2017. ISSN 0003-2697. doi: 10.1016/j.ab.2017.01.022.
- [125] Darius Lingė, Marius Gedgaudas, Andrius Merkys, Vytautas Petrauskas, Antanas Vaitkus, Algirdas Grybauskas, Vaida Paketurytė, Asta Zubrienė, Audrius Zakšauskas, Aurelija Mickevičiūtė, Joana Smirnovienė, Lina Baranauskienė, Edit Čapkauskaitė, Virginija Dudutienė, Ernestas Urniežius, Aleksandras Konovalovas, Egidijus Kazlauskas, Kirill Shubin, Helgi B Schiöth, Wen-Yih Chen, John E Ladbury, Saulius Gražulis, and Daumantas Matulis. PLBD: protein–ligand binding database of thermodynamic and kinetic intrinsic parameters. *Database*, 2023:baad040, 2023. ISSN 1758-0463. doi: 10.1093/database/baad040.
- [126] Daumantas Matulis and Rex Lovrien. 1-Anilino-8-Naphthalene Sulfonate Anion-Protein Binding Depends Primarily on Ion Pair Formation. *Biophys. J.*, 74(1):422–429, 1998. ISSN 0006-3495. doi: 10.1016/S0006-3495(98)77799-9.
- [127] Piotras Cimmperman, Lina Baranauskienė, Simona Jachimovičiūtė, Jelena Jachno, Jolanta Torresan, Vilma Michailovienė, Jurgita Matulienė, Jolanta Sereikaitė, Vladas Bumelis, and Daumantas Matulis. A Quantitative Model of Thermal Stabilization and Destabilization of Proteins by Ligands. *Biophysical Journal*, 95(7):3222–3231, 2008. ISSN 00063495. doi: 10.1529/biophysj.108.134973.
- [128] Marius Gedgaudas, Denis Baronas, Egidijus Kazlauskas, Vytautas Petrauskas, and Daumantas Matulis. Thermott: A comprehensive online tool for protein–ligand binding constant determination. *Drug Discovery Today*, 27(8):2076–2079, 2022. ISSN 1359-6446. doi: 10.1016/j.drudis.2022.05.008.
- [129] T. Wiseman, S. Williston, J. F. Brandts, and L. N. Lin. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Analytical Biochemistry*, 179(1):131–137, 1989. ISSN 0003-2697.
- [130] Thomas H. Scheuermann and Chad A. Brautigam. High-precision, automated integration of multiple isothermal titration calorimetric thermograms: new features of NITPIC. *Methods (San Diego, Calif.)*, 76:87–98, 2015. ISSN 1095-9130. doi: 10.1016/j.ymeth.2014.11.024.

- [131] Peter Schuck, Huaying Zhao, and Chad A. Brautigam. Chapter 7 - unbiased baseline determination, global analysis, and multimethod analysis of titration isotherms in SEDPHAT and NITPIC. In Munishwar Nath Gupta, Anthony Mittermaier, and Justin Di Trani, editors, *Isothermal Titration Calorimetry in Enzymology*, Foundations and Frontiers in Enzymology, pages 119–140. Academic Press, 2026. ISBN 978-0-443-21848-4. doi: 10.1016/B978-0-443-21848-4.00009-1.
- [132] Huaying Zhao, Grzegorz Piszczek, and Peter Schuck. SEDPHAT – a platform for global ITC analysis and global multi-method analysis of molecular interactions. *Methods*, 76:137–148, 2015. ISSN 10462023. doi: 10.1016/j.ymeth.2014.11.012.
- [133] Chad A. Brautigam. Calculations and publication-quality illustrations for analytical ultracentrifugation data. *Methods in Enzymology*, 562:109–133, 2015. ISSN 1557-7988. doi: 10.1016/bs.mie.2015.05.001.
- [134] Vaida Linkuvienė, Asta Zubrienė, Elena Manakova, Vytautas Petrauskas, Lina Baranauskienė, Audrius Zakšauskas, Alexey Smirnov, Saulius Gražulis, John E. Ladbury, and Daumantas Matulis. Thermodynamic, kinetic, and structural parameterization of human carbonic anhydrase interactions toward enhanced inhibitor design. *Quarterly Reviews of Biophysics*, 51, 2018. ISSN 0033-5835, 1469-8994. doi: 10.1017/S0033583518000082.
- [135] Robert N. Goldberg, Nand Kishore, and Rebecca M. Lennen. Thermodynamic quantities for the ionization reactions of buffers. *J. Phys. Chem. Ref. Data*, 31(2):231–370, 2002. ISSN 0047-2689. doi: 10.1063/1.1416902.
- [136] Audrius Zakšauskas, Vaida Paketurytė-Latvė, Alberta Jankūnaitė, Edita Čapkauskaitė, Yann Becart, Alexey Smirnov, Klára Pospíšilová, Janis Leitans, Jiří Brynda, Andris Kazaks, Lina Baranauskienė, Elena Manakova, Saulius Gražulis, Visvaldas Kairys, Kaspars Tars, Pavlína Řezáčová, and Daumantas Matulis. Affinity and selectivity of protein–ligand recognition: A minor chemical modification changes carbonic anhydrase binding profile. *Journal of Medicinal Chemistry*, 2025. ISSN 0022-2623. doi: 10.1021/acs.jmedchem.5c01421.
- [137] V. Paketurytė-Latvė, A. Smirnov, E. Manakova, L. Baranauskiene, V. Petrauskas, A. Zubrienė, J. Matulienė, V. Dudutienė, E. Čapkauskaitė, A. Zakšauskas, J. Leitans, S. Gražulis, K. Tars, and D. Matulis. From x-ray crystallographic structure to intrinsic thermodynamics of protein–ligand binding using carbonic anhydrase isozymes as a model system. *IUCr*, 11(4):556–569, 2024. ISSN 2052-2525. doi: 10.1107/S2052252524004627.

- [138] Vaida Paketurytė, Vytautas Petrauskas, Asta Zubrienė, Olga Abian, Margarida Bastos, Wen-Yih Chen, Maria João Moreno, Georg Krainer, Vaida Linkuvienė, Arthur Sedivy, Adrian Velazquez-Campoy, Mark A. Williams, and Daumantas Matulis. Uncertainty in protein-ligand binding constants: asymmetric confidence intervals versus standard errors. *European biophysics journal: EBJ*, 2021. ISSN 1432-1017. doi: 10/gjr8hp.
- [139] Jerome M. Fox, Kyungtae Kang, Madhavi Sastry, Woody Sherman, Banumathi Sankaran, Peter H. Zwart, and George M. Whitesides. Water-Restructuring Mutations Can Reverse the Thermodynamic Signature of Ligand Binding to Human Carbonic Anhydrase. *Angewandte Chemie (International Ed. in English)*, 2017. ISSN 1521-3773. doi: 10.1002/anie.201609409.
- [140] Gerhard Klebe. Thermodynamic data remain a hot tip for decoding binding affinity and water impact on protein–ligand complex formation to assist lead optimization. *Journal of Medicinal Chemistry*, 69(4):3719–3730, 2026. ISSN 0022-2623. doi: 10.1021/acs.jmedchem.5c03100.
- [141] J. Boström, P. O. Norrby, and T. Liljefors. Conformational energy penalties of protein-bound ligands. *Journal of Computer-Aided Molecular Design*, 12(4):383–396, 1998. ISSN 0920-654X. doi: 10.1023/a:1008007507641.
- [142] A. Joshua Wand and Kim A. Sharp. Measuring Entropy in Molecular Recognition by Proteins. *Annual Review of Biophysics*, 47(1): 41–61, 2018. doi: 10.1146/annurev-biophys-060414-034042.
- [143] Vijay M. Krishnamurthy, Brooks R. Bohall, Chu-Young Kim, Demetri T. Moustakas, David W. Christianson, and George M. Whitesides. Thermodynamic parameters for the association of fluorinated benzenesulfonamides with bovine carbonic anhydrase II. *Chemistry, an Asian Journal*, 2(1):94–105, 2007. ISSN 1861-471X. doi: 10.1002/asia.200600360.
- [144] Helge Vatheuer, Oscar Palomino-Hernández, Janis Müller, Phillip Galonska, Serghei Glinca, and Paul Czodrowski. Protonation effects in protein-ligand complexes – a case study of endothiapepsin and pepstatin a with computational and experimental methods. *ChemMedChem*, 20(8):e202400953, 2025. ISSN 1860-7187. doi: 10.1002/cmdc.202400953.
- [145] Gregory A. Landrum and Sereina Riniker. Combining IC₅₀ or k_i values from different sources is a source of significant noise. *Journal of Chemical Information and Modeling*, 64(5):1560–1567, 2024. ISSN 1549-9596. doi: 10.1021/acs.jcim.4c00049.
- [146] Bengt-Harald Jonsson and Anders Liljas. Perspectives on the classical enzyme carbonic anhydrase and the search for inhibitors.

- Biophysical Journal*, 119(7):1275–1280, 2020. ISSN 0006-3495. doi: 10.1016/j.bpj.2020.08.020.
- [147] Joel Tellinghuisen and John D. Chodera. Systematic errors in isothermal titration calorimetry: Concentrations and baselines. *Analytical Biochemistry*, 414(2):297–299, 2011. ISSN 00032697. doi: 10.1016/j.ab.2011.03.024.
 - [148] Joel Tellinghuisen. Isothermal titration calorimetry at very low c. *Analytical Biochemistry*, 373(2):395–397, 2008. ISSN 00032697. doi: 10.1016/j.ab.2007.08.039.
 - [149] Joel Tellinghuisen. Optimizing experimental parameters in isothermal titration calorimetry. *The Journal of Physical Chemistry B*, 109(42):20027–20035, 2005. ISSN 1520-6106. doi: 10.1021/jp053550y.
 - [150] Joel Tellinghuisen. Optimizing experimental parameters in isothermal titration calorimetry: Variable volume procedures. *The Journal of Physical Chemistry B*, 111(39):11531–11537, 2007. ISSN 1520-6106, 1520-5207. doi: 10.1021/jp074515p.
 - [151] Joel Tellinghuisen. Optimizing isothermal titration calorimetry protocols for the study of 1:1 binding: Keeping it simple. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1860(5):861–867, 2016. ISSN 03044165. doi: 10.1016/j.bbagen.2015.10.011.
 - [152] Žiga Medoš, Marija Bešter-Rogač, Epameinondas Leontidis, and Joel Tellinghuisen. Systematic errors in isothermal titration calorimetry: The role of feedback power and effects of mixing and diffusion on concentrations. *Analytical Biochemistry*, page 116010, 2025. ISSN 0003-2697. doi: 10.1016/j.ab.2025.116010.
 - [153] Joel Tellinghuisen. Can you trust the parametric standard errors in nonlinear least squares? yes, with provisos. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 2017. ISSN 0304-4165. doi: 10.1016/j.bbagen.2017.12.016.

Santrauka

(Summary in Lithuanian)

Terminų žodynelis

Kai kuriems disertacijoje vartojamiems terminams nėra nusistovėjusių lietuviškų atitikmenų, todėl šiame žodynelyje pateikiami terminai ir jų apibrėžtys, kurie yra vartojami šiame darbe.

Jungimosi stiprumas (angl. *affinity*) – ligando ir baltymo sąveikos stiprumą apibūdinanti charakteristika (kiekybiškai išreiškiama parametrais K_b , K_d arba ΔG).

Profilinis tikėtinumas (angl. *profile likelihood*) – statistinis metodas pasikliautiniams intervalams nustatyti netiesiniuose modeliuose, geriau atspindintis parametru (pvz., K_b) paklaidų asimetriją.

Slopiklis (angl. *inhibitor*) – cheminis junginys, kuris jungiasi prie fermento ir sumažina arba visiškai sustabdo jo katalizinį aktyvumą.

Stebimasis parametras (angl. *observed parameter*) – eksperimentiškai išmatuota vertė (žymima indeksu *obs*), kuri apima ne tik tiesioginę sąveiką, bet ir susietas protonizacijos reakcijas.

Tikrasis parametras (angl. *intrinsic parameter*) – termodinaminis dydis (žymimas indeksu *int*), apibūdinantis baltymo ir ligando sąveikos energiją, atsižvelgus į susietų protonacijos pusiausvyrų indėlį.

Pagrindinės santrumpos

Santraukoje vartojamos tos pačios santrumpos ir kintamieji kaip pagrindinėje disertacijos dalyje. Vartojamos šios pagrindinės santrumpos: CA – karboanhidrazė; AZM – acetazolamidas; FTSA – fluorescencinis terminio poslinkio metodas; ITC – izoterminė titravimo kalorimetriją; PDB – Baltymų duomenų bankas (angl. *Protein Data Bank*); PLBD – Baltymų–ligandų jungimosi duomenų bazė (angl. *Protein–Ligand Binding Database* (plbd.org)).

Įvadas

Vienas svarbiausių vaistų kūrimo uždavinių yra junginių, gebančių selektyviai ir stipriai jungtis prie biologinių taikinių, paieška ir optimizavimas. Struktūrinės biologijos, biofizikos ir kompiuterinės chemijos pažanga reikšmingai išplėtė galimybes tirti baltymų ir ligandų sąveikas, todėl šios disciplinos tapo neatsiejama šiuolaikinio racionalaus vaistų projektavimo dalimi [1, 2]. Aukštos skiriamosios gebos struktūros, molekulinės simuliacijos ir struktūra grįsto vaistų projektavimo metodai suteikia vertingos informacijos apie sąveikas bei padeda optimizuoti perspektyvius junginius [3]. Nepaisant to, tiksliai prognozuoti, kaip struktūriniai pokyčiai paveiks jungimosi stiprumą ir selektyvumą, išlieka didelis ir sunkus iššūkis [4].

Baltymų ir ligandų atpažinimą lemia sudėtinga tarpmolekuliinių sąveikų, hidratacijos, protonacijos pusiausvyrų ir konformacinės dinamikos visuma. Nors šie veiksniai kartu nulemia jungimosi termodinamiką, jų individualų indėlį dažnai sudėtinga atskirti ir kiekybiškai įvertinti. Dėl šios priežasties struktūriškai panašūs junginiai gali pasižymėti netikėtai skirtingomis jungimosi savybėmis, o panašus jungimosi stiprumas gali būti pasiekiamas iš esmės skirtingais mechanizmais. Šių ryšių supratimas išlieka svarbus tiek eksperimentinių, tiek kompiuterinių molekulinio atpažinimo tyrimų tikslas.

Baltymų ir ligandų sąveikų interpretaciją bei prognozavimą ypač apsunkina keli veiksniai:

- **Entalpijos ir entropijos kompensacija bei hidratacijos tinklai.** Jungimosi stiprumą lemia entalpiinių ir entropinių indėlių pusiausvyra, o šie indėliai dažnai vienas kitą kompensuoja [5]. Todėl panašus jungimosi stiprumas gali būti pasiekiamas skirtingais termodinaminiais mechanizmais. Tokie skirtumai dažnai siejami su hidratacijos tinklų persitvarkymu, kurį vis dar sudėtinga tiksliai apibūdinti ir prognozuoti [6–9].
- **Konformaciniai efektai.** Nedideli cheminės struktūros pokyčiai gali pakeisti ligando geometriją, lankstumą ir vidumolekulinę sąveiką, taip paveikdami energijos sąnaudas, reikalingas jungimuisi palankiai konformacijai įgyti. Tokie pokyčiai gali reikšmingai paveikti jungimąsi, net jei tiesioginės baltymo ir ligando sąveikos išlieka beveik nepakitusios [6, 10].
- **Stebimoji ir tikroji termodinamika.** Eksperimentais nustatyti jungimosi parametrai dažnai apima su protonizacijos procesais ir

buferinio tirpalo savybėmis susijusius indėlius [11, 12]. Todėl stebimi termodinaminiai parametrai ne visuomet tiesiogiai atspindi tikrąją molekulinio atpažinimo energiją, o tai apsunkina mechanistinę rezultatų interpretaciją ir skirtingų sistemų palyginimą.

Šie veiksniai rodo, kad molekulinio atpažinimo negalima visapusiškai suprasti remiantis vien suminio jungimosi stiprumo matavimais ar statinėmis struktūromis. Todėl būtina pasirinkti tinkamą modelinę sistemą, leidžiančią sistemingai tirti struktūrinių, termodinaminių ir aplinkos veiksnių įtaką jungimuisi. Karboanhidrazės (CA) yra tinkama modelinė sistema šiems klausimams nagrinėti [13]. Nors jų katalizinio centro architektūra yra labai konservatyvi, nedideli aktyviojo centro aminorūgščių ir hidratacijos tinklų organizacijos skirtumai gali lemti reikšmingus ligandų atpažinimo skirtumus. Sulfonamidiniai slopikliai šiems tyrimams ypač tinkami, nes turi bendrą karboanhidrazės cinko joną koordinuojančią funkcinę grupę, tačiau leidžia sistemiškai modifikuoti pakaitus, sąveikaujančius su skirtingomis aktyviojo centro sritimis. Tai suteikia galimybę detalai tirti, kaip struktūriniai pokyčiai veikia jungimosi termodinamiką, stiprumą ir selektyvumą.

Patikima struktūros ir termodinamikos ryšių interpretacija reikalauja ne tik mechanistinio supratimo, bet ir tinkamo eksperimentinio neapibrėžtumo įvertinimo, ypač lyginant parametrus, gaunamus taikant netiesinius jungimosi modelius. Pagrindinė šios disertacijos hipotezė yra ta, kad prasmingi ryšiai tarp baltymo-ligando struktūros ir termodinamikos gali būti nustatyti tik sistemingai įvertinus atskirų molekulinį atpažinimą lemiančių procesų indėlį ir eksperimentinį neapibrėžtumą.

Tikslas:

Tirti ir kiekybiškai įvertinti struktūrinius, termodinaminius ir molekulinės aplinkos veiksnius, lemiančius baltymų ir ligandų molekulinį atpažinimą, naudojant karboanhidrazes ir pirminius sulfonamidus kaip modelinę sistemą.

Uždaviniai:

1. Nustatyti, kaip nedideli baltymo aktyviojo centro aminorūgščių, ligando struktūros ir hidratacijos tinklų skirtumai veikia entalpiinių ir entropinių indėlių pusiausvyrą baltymo ir ligando sąveikos metu. (Publikacija 1)
2. Tirti nedidelių ligando cheminių modifikacijų poveikį ligando konformacijai, įskaitant vidumolekulinių vandenilinių ryšių susidarymą ir konformacinę energiją, bei jų reikšmę sąveikos stiprumui ir selektyvumui. (Publikacija 2)
3. Atskirti su baltymo ir ligando jungimusi susietų protonizacijos procesų indėlius nuo stebimųjų eksperimentinių parametrų ir apskaičiuoti tikruosius termodinaminius parametrus, nepriklausančius nuo protonizacijos. (Publikacija 3)
4. Standartizuoti FTSA eksperimentinius protokolus patikimam sąveikos stiprumo nustatymui bei naudoti asimetrinius profilinio tikėtinumumo (angl. *profile likelihood*) pasikliautinuosius intervalus, siekiant tiksliai įvertinti ir interpretuoti neapibrėžtį netiesiniuose sąveikos modeliuose. (Publikacija 4 ir Publikacija 5)

Mokslinis naujumas

Šios disertacijos mokslinis naujumas pagrįstas integruotu požiūriu į baltymo ir ligando atpažinimo tyrimus, derinant didelės skiriamosios gebos struktūrinę analizę, tikrąją jungimosi termodinamiką ir griežtą neapibrėžčių vertinimą. Šiame darbe neapsiribojama vien jungimosi stiprumo matavimais; parodoma, kaip ligando struktūra, baltymo aplinka, hidratacijos tinklai bei su protonacija susietos pusiausvyros kartu formuoja molekulinį atpažinimą.

Pagrindiniai nauji rezultatai:

- **Struktūrinių skirtumų vaidmuo formuojant skirtingus termodinaminius mechanizmus:** parodyta, kad izofermentai (CA II, CA XII ir CA XIII) pasiekia panašų jungimosi stiprumą per iš esmės priešingus entalpinius-entropinius mechanizmus. Šiuos termodinaminius skirtumus lemia subtilūs aktyviojo centro liekanų skirtumai ir konservatyvių hidratacijos tinklų organizacijos ypatumai. (Publikacija 1)
- **Ižvalgos apie ligando konformacijos energijos kainą:** nustatyta, kad sieros atomo oksidacijos laipsnio pakeitimas (iš sulfanilo į sulfinilą ir sulfonilą) pakeitė jungimosi stiprumą iki milijono kartų. Šis stiprumo sumažėjimas susijęs su nepalankių vidumolekulinių vandenilinių ryšių susidarymu, kurie lemia nepalankią ligando konformaciją ir sukuria sterines kliūtis sąveikai su karboanhidrazėmis. (Publikacija 2)
- **Tikrųjų termodinaminių parametrų atskyrimas nuo stebimųjų:** parodyta, kad su jungimusi susietos protonacijos pusiausvyros reikšmingai veikia eksperimentiškai nustatomus stebimuosius termodinaminius parametrus. Šių indėlių atskyrimas leido nustatyti prasmingas sąsajas tarp didelės skiriamosios gebos struktūrinių duomenų ir tikrųjų molekulinio atpažinimo energijų. (Publikacija 3)
- **Nuoseklus neapibrėžties pateikimas tarp termodinaminių dydžių:** pabrėžta, kad įprasti simetriniai paklaidų įverčiai veda prie klaidingų interpretacijų, kai jungimosi stiprumas perskaičiuojamas iš disociacijos konstantų į Gibso energijas. Priešingai, asimetriniai profilinio tikėtinumo pasikliautiniai intervalai (angl. *profile-likelihood confidence intervals*) išsaugojo tikrąją neapibrėžties struktūrą ir užtikrino nuoseklią interpretaciją skirtingose jungimosi stiprumo išraiškose (Publikacija 4 ir Publikacija 5)

Ginamieji teiginiai

- Nedideli baltymų ir ligandų struktūriniai skirtumai gali lemti neproporcingai didelius molekulinio atpažinimo ir selektyvumo pokyčius, nes jungimasi lemia visa molekulinė aplinka, įskaitant protonacijos pusiausvyras, hidrataciją, konformacinius efektus ir tarpmolekulinės sąveikas. (Publikacija 1, Publikacija 2 ir Publikacija 3)
- Molekulinis baltymų ir ligandų atpažinimas negali būti visapusiškai suprastas remiantis vien struktūrine informacija ir stebimaisiais termodinaminiais parametrais; patikimai interpretacijai būtina integruoti atominės raiškos struktūrinius duomenis ir tikruosius termodinامينius parametrus. (Publikacija 1, Publikacija 2 ir Publikacija 3)
- Neapibrėžties pateikimo būdas daro įtaką baltymo ir ligando jungimosi matavimų interpretacijai; asimetriniai profilinio tikėtumo pasikliautiniai intervalai (angl. *profile-likelihood confidence intervals*) užtikrina nuoseklų neapibrėžties aprašymą, o įprastos simetrinės standartinės paklaidos yra tam netinkamos. (Publikacija 4 ir Publikacija 5)

Literatūros apžvalgos apibendrinimas

Struktūra grįstas vaistų kūrimas: galimybės ir iššūkiai

Naujų vaistinių preparatų kūrimas ir vystymas yra itin brangus ir rizikingas procesas, reikalaujantis didelių finansinių investicijų [14]. Šis procesas paprastai trunka apie 10-15 metų [15], o kandidatų atmetimo rizika išlieka milžiniška - net ir patekus į klinikinių tyrimų stadiją, bendra patvirtinimo tikimybė tėra apie 12,8 % [16]. Nepaisant to, struktūrinė informacija yra kritiškai svarbi: pavyzdžiui, 2019-2023 m. patvirtintų 34 priešvėžinių mažos molekulinės masės vaistų vystymui struktūrinė biologija turėjo įtakos visais atvejais, o daugiau nei 80 % atvejų vaistai sukurti taikant struktūrą grįstą projektavimą [21].

Svarbiausias uždavinys ankstyvosiose vaistų paieškos stadijose (ieškant pirminių junginių-lyderių bei optimizuojant jų struktūrą) yra sukurti molekules, pasižyminčias dideliu sąveikos stiprumu ir selektyvumu tiksliniam baltymui. Struktūrą grįstas vaistų kūrimas (angl.

Structure-Based Drug Design) naudoja trimačius taikinio struktūros duomenis tam, kad ligandų optimizavimas būtų kuo racionalesnis. Nors rentgeno kristalografija atskleidžia atskirų atomų sąveikas komplekse, ji tiesiogiai nenusako paties prisijungimo stiprumo. Racionalų vaistų kūrimą apsunkina tai, kad bet kokia cheminė ligando modifikacija vienu metu veikia desolvataciją, konformacinius efektus, protonizacijos pusiausvyrą bei entropiją. Todėl eksperimentinėse struktūrose stebima sąveika dažnai tiesiogiai nenurodo jos įnašo į bendrą sąveikos energiją.

Kompiuteriniai modeliavimo metodai [20], skirti kiekybiškai prognozuoti sąveikas, vis dar turi fizikinio detalumo bei tikslumo ribotumą.

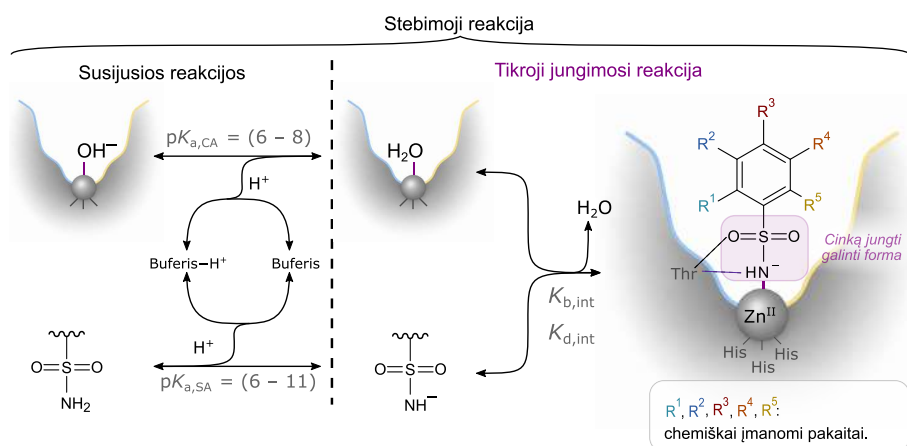
- Molekulinis dokingas (angl. *molecular docking*) įvertio funkcijomis padeda nuspėti ir įvertinti jungimosi pozas (angl. *scoring functions*), tačiau sąveikos stiprumo prognozavimas išlieka labai netikslus ir stipriai priklauso nuo modelio pasirinkimo.
- Molekulinė dinamika (angl. *molecular dynamics*, MD) leidžia tirti molekulių struktūros pokyčius, tirpiklio ir jonų judėjimą. Tačiau įprasti MD skaičiavimai neleidžia tiesiogiai įvertinti skirtingų molekulės būsenų stabilumo. Taip pat dažnai yra reikalingi labai ilgi skaičiavimai, kad būtų gauti patikimi rezultatai.
- Kvantinė mechanika (angl. *quantum mechanics*, QM) leidžia įvertinti sąveikas naudojant elektroninės struktūros skaičiavimus, tačiau reikalauja didelių skaičiavimo išteklių.
- Mašininis mokymasis (angl. *machine learning*, ML) nuspėja pozas ir sąveikos stiprumą iš eksperimentinių bei struktūrinių duomenų, tačiau jo tikslumas visiškai priklauso nuo mokymo duomenų kokybės bei modelio validavimo.

Todėl išlieka svarbu integruoti atominės raiškos struktūrinius duomenis su išsamiais eksperimentiniais termodinaminiais tyrimais, siekiant aiškintis veiksnius, lemiančius baltymų ir ligandų sąveiką.

Baltymo ir ligando sąveikas lemiantys veiksniai

Baltymų ir ligandų sąveika lemia biologinių sistemų funkcionavimą ir vaistinių junginių veiksmingumą. Ligando prisijungimą prie

baltymo lemia įvairių tarpmolekulinių sąveikų visuma, įskaitant metalo koordinacinius ryšius, vandenilinius ryšius, elektrostatines sąveikas, hidrofobinį efektą, van der Waalso kontaktus bei kitas sąveikas. Šiame darbe kaip modelinė sistema yra naudojamos karboanhidrazės ir benzensulfonamidai. Biofiziniai tyrimai rodo, kad apie 60 %–65 % karboanhidrazių ir sulfonamidų komplekso stabilizavimo energijos kyla iš tiesioginės koordinacinės sąveikos su kataliziniu Zn^{II} [35], kuri veikia kaip pagrindinis inkaras, o kitos nekovalentinės sąveikos veikia papildomai jungimosi stiprumą bei selektyvumą. Šių sąveikų indėlis



S.1 pav.: Bendra karboanhidrazių ir pirminių sulfonamidų atpažinimo schema. Schemoje apibendrinami pirminių sulfonamidų prisijungimo prie kataliziškai aktyvių žmogaus karboanhidrazių izofermentų ypatumai (ypatingą dėmesį skiriant benzensulfonamidams, kurie yra pagrindinis šio darbo objektas) ir pristatomos sąvokos, aptariamose vėlesniuose skyriuose. Prisijungimas yra susietas su abiejų partnerių (de)protonizacijos pusiausvyromis, nurodant apytiksles pK_a vertes. Konservatyvų cinką rišantį motyvą sudaro sulfonamido azoto koordinacija su Zn^{II} jonu ir vandeniliniai ryšiai su Thr199. Šis motyvas, esantis sandūroje tarp hidrofobiškesnės ir hidrofilinės aktyviojo centro pusių (atitinkamai geltonos ir mėlynos spalvos), sudaro bendrą struktūrinį inkarą. Be šio motyvo, sulfonamidų pakaitų įvairovė ir skirtinga individualių izofermentų aktyviojo centro aplinka lemia įvairius sąveikos ypatumus, turinčius įtakos sąveikos stiprumui, selektyvumui ir prisijungimo termodinamikai. Eksperimentiškai stebimas prisijungimas papildomai apima susijusias protonizacijos reakcijas; į jas atsižvelgus galima nustatyti tikruosius prisijungimo parametrus, kurie tiesiogiai atspindi baltymo ir ligando sąveiką.

priklauso ne tik nuo tiesioginių baltymo ir ligando kontaktų, bet ir nuo aplinkos savybių, ypač vandens molekulių tinklų organizacijos bei protonacijos pusiausvyrų (Pav. S.1).

Baltymo–ligando jungimasi apibūdina pagrindiniai termodinaminiai parametrai: Gibso energijos pokytis (ΔG), entalpijos pokytis (ΔH) ir entropijos pokytis (ΔS). Nors jungimosi stiprumą nusako ΔG , jo vertę dažnai lemia skirtingas entalpijų ir entropijų veiksnių balansas. Dažnai stebimas entalpijos–entropijos kompensacijos reiškinys, kai palankesni entalpiniai pokyčiai kompensuojami mažiau palankiais entropiniais pokyčiais arba atvirkščiai, todėl panašų jungimosi stiprumą gali lemti skirtingi molekuliniai mechanizmai [8].

Svarbų vaidmenį taip pat atlieka vandens molekulių reorganizacija ir desolvatacija, nes ligandui jungiantis dalis aktyviojo centro vandens molekulių pašalinama arba persitvarko, reikšmingai paveikdama jungimosi termodinamiką [7, 9, 36–38]. Eksperimentiniai tyrimai rodo, kad vandens molekulių išlaisvinimas iš aktyviojo centro hidrofobinių sričių gali suteikti palankų entropinį įnašą, o stipriai struktūrizuotų vandens molekulių tinklo išlaikymas arba naujų vandenilinių ryšių suformavimas gali prisidėti prie palankaus entalpinio indėlio. Tyrimų, detaliai analizuojančių šiuos reiškinius, vis dar yra nedaug, o kiekvienas baltymo ir ligando sąveikos atvejis pasižymi unikaliomis savybėmis. Dėl šios priežasties nėra bendrų, visoms sistemoms tinkamų gairių, o tai apsunkina itin tikslų prognozavimo modelių sukūrimą.

Žmogaus karboanhidrazės ir pirminiai sulfonamidai kaip modelinė sistema

Karboanhidrazės [84] (CAs; EC 4.2.1.1) yra metalofermentų šeima, katalizuojanti grįžtamąją anglies dioksido hidratacijos reakciją. Žmogaus α klasės karboanhidrazių šeimą sudaro 15 izofermentų, iš kurių 12 yra kataliziškai aktyvūs [13]. Trys izofermentai - CA VIII, CA X ir CA XI - kataliziškai neaktyvūs (vadinami su karboanhidrazėmis susijusiais baltymais), kadangi juose nėra cinką koordinuojančių histidinų [87] ir šiame darbe jie nėra nagrinėjami. Nepaisant didelio struktūrinio panašumo, aktyvūs izofermentai pasižymi skirtinga lokalizacija, fiziologinėmis funkcijomis ir slopiklių atpažinimo savybėmis.

Visų kataliziškai aktyvių žmogaus karboanhidrazių aktyviajame centre yra Zn^{II} jonas, koordinuotas trijų histidino liekanų (His94, His96,

His119) ir vandens molekulės arba hidroksido jono [88]. Aktyvusis centras sudarytas iš:

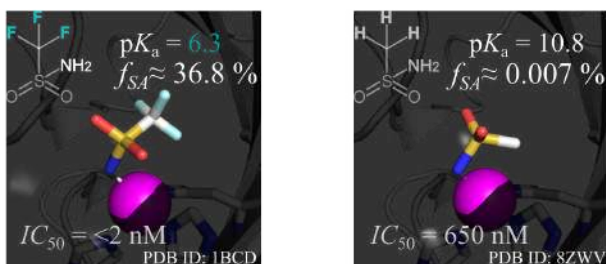
- labiau hidrofobinės srities, kurią formuoja Val121, Val143, Leu198 ir Trp209 bei kitos aminorūgštys.
- labiau hidrofilinės srities, kurioje vandenilinių ryšių tinklas yra išsidėstęs aplink Thr199 ir Glu106 bei konservatyvius vandens molekules.
- kataliziškai svarbaus protonų pernešėjo His64, kuris keičia konformacijas tarp vidinės (angl. *inward*) ir išorinės (angl. *outward*), perduodamas protonus iš aktyviojo centro į tirpiklį per vandenilniais ryšiais susietą vandens molekulių tinklą.

Nors katalizinė architektūra tarp izofermentų išlieka labai panaši, nedideli aminorūgščių skirtumai aktyviojo centro prieigoje keičia ertmės geometriją, paviršiaus elektrostazines savybes ir vandens tinklą organizaciją, todėl gali lemti reikšmingus jungimosi stiprumo ir selektyvumo skirtumus. Šie skirtumai puikiai iliustruojami tirtų izofermentų pavyzdžiais (aminorūgščių skirtumai lyginami su modeliniu baltymu vadinama CA II, naudojant jos aminorūgščių numeraciją):

- CA IV yra ekstraląstelinis, su membrana per GPI (angl. *glycosylphosphatidylinositol*) inkarą susietas izofermentas, stabilizuotas disulfidiniu ryšiu (Cys6 ir Cys11) ir glikoziliniu [97]. Jo aktyviojo centro ertmė yra siauresnė ir mažiau prieinama tirpikliui dėl tokių prieigos skirtumų kaip Leu60→Gln, Asn67→Met, Ile91→Lys, Phe131→Val, Leu141→Ile ir Ala135→Gln.
- CA IX yra tiesiogiai su navikų hipoksija susijęs transmembraninis izofermentas, kurio katalizinis domenas sudaro homodimerą, stabilizuotą disulfidine jungtimi (Cys119–Cys299) ir *N*-glikoziliniu ties Asn309 [100, 101]. Jo aktyviajame centre esantys skirtumai Leu60→Arg, Ile91→Leu, Phe131→Val (pašalinta reikšminga sterinė kliūtis) bei Ala135→Leu sukuria erdvesnę ir labiau hidrofobinę kišenę, puikiai pritaikytą didesniems ir selektyviems slopikliams nei CA II ir likusioms karboanhidrazėms.
- CA XIII būdingi tokie skirtumai kaip Leu60→Ser, Asn67→Ser ir Ile91→Arg, o 200-ojoje pozicijoje esantis Val200 (skirtingai nei Thr200 CA II) [103] keičia vandens molekulės W_1 padėtį ir stabilizavimą. CA II izofermente Thr200 stabilizuoja W_1 vandeniliniu

ryšiu, o CA XIII esantis Val200 tokio ryšio sudaryti negali. Šis skirtumas gali prisidėti prie izofermentams būdingų aktyviojo centro hidratacijos ir jungimosi termodinamikos skirtumų, kaip aptarta disertacijos rezultatų dalyje.

Plačiausiai ištirta karboanhidrazių slopiklių klasė yra sulfonamidai [110–113]. Jų deprotonuota sulfonamido grupė tiesiogiai koordinuojasi prie katalizinio cinko jono, išstumdama su juo susijusią vandens molekulę (Pav. S.1). Be šios pagrindinės sąveikos, slopiklių jungimosi stiprumą ir selektyvumą lemia papildomi vandeniliniai ryšiai, hidrofobiniai kontaktai bei sąveikos su aktyviojo centro vandens molekulė tinklais.



S.2 pav.: Sulfonamido pK_a efektas stebimai sąveikai su karboanhidrazėmis. Trifluormetansulfonamido (TFMSA, kairėje) ir metansulfonamido (MSA, dešinėje), susijungusių su žmogaus CA II, palyginimas. Abu slopikliai pasižymi panašiu jungimusi - jie koordinuoja katalizinį Zn^{II} per deprotonuotą sulfonamido azoto atomą. Metilo grupę pakeitus stipriai elektronus traukiančia trifluormetilo grupe, sulfonamido pK_a sumažėja nuo 10,8 iki 6,3 ir todėl, esant pH 7,0, su cinku besijungiančio sulfonamido anijono dalis (f_{SA}) padidėja nuo maždaug 0,007 % iki 36,8 %. Dėl šios priežasties gerokai didesnę TFMSA slopinamąją aktyvumą ($IC_{50} < 2$ nM), palyginti su MSA ($IC_{50} = 650$ nM), lemia didesnė aktyvios, deprotonuotos, su cinku besijungiančios formos dalis, o ne reikšmingi baltymo ir ligando sąveikų skirtumai. Struktūros pateiktos pagal PDB įrašus 1BCD (TFMSA) ir 8ZWV (MSA); pK_a ir IC_{50} vertės paimtos iš literatūros [109].

Kadangi slopikliai jungiasi tik deprotonuotos anijoninės formos pavidalu, o pats fermentas turi būti Zn^{II} -H₂O formos, šis procesas yra stipriai susietas su (de)protonizacijos pusiausvyromis tirpale [11, 12]. Tai reiškia, kad eksperimentiškai stebimas (angl. *observed*) jungimosi stiprumas yra visuomet mažesnis už tikrąjį (angl. *intrinsic*), o viena geriausių šios įtakos iliustracijų yra metansulfonamido (MSA) ir trifluormetansulfonamido (TFMSA) palyginimas (Figure S.2). Abu slopikliai

su CA II jungiasi beveik identišškai koordinuodami cinko joną, tačiau stipriai elektronus traukianti CF_3 grupė sumažina TFM-SA pK_a iki 6,3 (MSA pK_a yra 10,8). Dėl šio rūgštingumo skirtumo, esant pH 7,0, TFM-SA deprotonizuotos formos tirpale yra net 36,8 %, o MSA - vos 0,007 %, kas nulemia drastišką stebimo aktyvumo skirtumą tirpale ($IC_{50} < 2 \text{ nM}$ prieš 650 nM) [109], nors jų tikroji sąveikos energija su baltymu išlieka panaši. Todėl būtent tikrieji jungimosi parametrai turėtų būti naudojami modeliavimo etapuose.

Karboanhidrazės turi įvairių kliniškai naudojamų sulfonamidų ir jų darinių pagrindo slopiklių [114]. Acetazolamidas (AZM) išlieka klasiki- niu ir plačiausiai ištirtu etaloniniu slopikliu, pasižyminčiu dideliu sąvei- kos stiprumu, tačiau kaip ir daugelis - ribotu selektyvumu tarp atskirų izofermentų. Būtent selektyvumo paieškos išlieka vienas svarbiausių uždavinių.

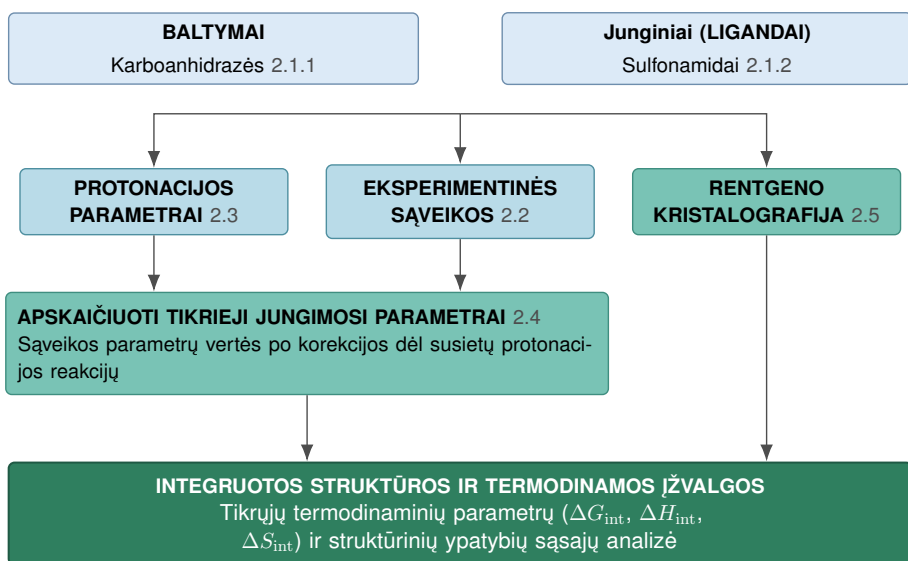
Nepaisant gausių struktūrinių ir termodinaminių duomenų apie karboanhidrazių ir sulfonamidų sąveikas, vis dar sudėtinga susieti ne- didelius baltymo ir ligando struktūros skirtumus su jų poveikiu jungi- mosi termodinamai ir selektyvumui. Karboanhidrazių ir sulfonamidų kompleksai yra tinkama modelinė sistema šiems ryšiams tirti, nes sulfo- namidams būdingas konservatyvus jungimosi prie cinko jono būdas, o jų sąveikos su skirtingų izofermentų aktyviojo centro aplinka skiriasi [13]. Tai leidžia sistemiškai tirti, kaip baltymo ir ligando struktūra bei molekulinė aplinka lemia jungimosi termodinamiką ir selektyvumą [1].

Metodų apibendrinimas

Šioje disertacijoje taikytų pagrindinių metodų seka pateikta S.3 pav. Tyrimuose buvo naudojamos rekombinantinės žmogaus karboanhidrazės ir sulfonamidiniai slopikliai.

Baltymo ir ligando sąveikos buvo charakterizuojamos biofizikiniais metodais, daugiausia fluorescenciniu terminio poslinkio metodu ir izoterminio titravimo kalorimetrija, nustatant stebimąsias jungimosi konstantas ir termodinaminius parametrus. Lygiagrečiai eksperimentiškai buvo nustatyti ligandų protonizacijos parametrai, o atitinkami karboanhidrazės izofermentų protonizacijos parametrai buvo paimti iš ankstesnių tyrimų.

Vėliau stebimieji jungimosi parametrai buvo koreguojami atsižvel- giant į su jungimusi susietas protonizacijos reakcijas, siekiant nustatyti



S.3 pav.: Disertacijoje taikytų eksperimentinių ir analizės metodų schema. Baltymų ir ligandų sąveikos duomenys, protonizacijos parametrai ir rentgeno spindulių kristalografijos struktūros buvo integruoti siekiant apskaičiuoti tikruosius termodinaminius parametrus ir ištirti struktūros bei termodinamikos sąsajas karboanhidrazių ir jų slopiklių sistemose.

tikruosius termodinaminius parametrus, nepriklausančius nuo eksperimentinių sąlygų (pH bei buferinio tirpalo) ir todėl tinkamus skirtingoms sistemoms palyginti. Struktūrinė informacija apie baltymo ir ligando kompleksus buvo gauta rentgeno spindulių kristalografijos metodu. Galiausiai tikrieji termodinaminiai parametrai buvo susieti su struktūriniais duomenimis. Šių metodų visuma leido išanalizuoti ryšius tarp baltymo ir ligando jungimosi termodinamikos bei karboanhidrazės ir slopiklių kompleksų struktūrinių ypatybių, siekiant gauti slopiklių kūrimui svarbių įžvalgų.

Išsamūs eksperimentinių procedūrų ir duomenų analizės metodų aprašymai pateikti publikacijoje, sudarančioje šios disertacijos pagrindą.

Medžiagos

Karboanhidrazių izofermentai

Šiame darbe naudota CA raiškos ir gryninimo platforma buvo sukurta ir daugelį metų plėtojama daugelio laboratorijos narių bei bendradarbių

pastangomis. Autorė nuoširdžiai dėkoja už jų indėlį ir šiam tyrimui suteiktus baltymų išteklius.

Dvylika CA izofermentų buvo gauti rekombinantiniu būdu kaip viso ilgio baltymai arba sutrumpintos jų konstrukcijos, išlaikančios katalizinį domeną. Daugumos izofermentų raiška buvo vykdoma *Escherichia coli* ląstelėse, naudojant pagal konkrečių baltymų poreikius parinktas padermes, pvz., BL21(DE3), Rosetta 2(DE3) ir Origami B(DE3) [115]. Kadangi bakterijų ląstelėse nepavyko gauti stabilios rekombinantinės CA IX, jos katalizinis domenas buvo gautas naudojant žinduolių ląstelių raiškos sistemas (FreeStyle 293-F arba Expi293F) [115]. Baltymai buvo gryninami taikant kiekvienam izofermentui pritaikytus Ni^{2+} chelatinimo, jonų mainų ir sulfonamidų ligandais pagrįstos chromatografijos metodų derinius. Baltymų grynumas, molekulinė masė ir katalizinis aktyvumas buvo vertinami atitinkamai SDS-PAGE, didelės skiriamosios gebos masių spektrometrijos (HR-MS) ir sustabdytos srovės (angl. *stopped-flow*) CO_2 hidratacijos metodais. Išsamios baltymų konstrukcijų sekos, raiškos sąlygos ir gryninimo procedūros pateiktos [115].

Informacija apie rekombinantinius baltymus taip pat pateikiama PLBD duomenų bazėje: <https://plbd.org/db/main/protein>. Eksperimentinių duomenų rinkiniai ir atitinkamos eksperimentuose naudotos baltymų partijos yra susieti ir aprašyti šioje duomenų bazėje.

Baltymų mėginiai buvo laikomi -80°C temperatūroje. Prieš eksperimentus, kuomet būtina, baltymų laikymo buferinis tirpalas buvo pakeičiamas į atitinkamam eksperimentui naudojamą buferinį tirpalą. Baltymų koncentracijos buvo nustatomos spektrofotometriškai, naudojant atitinkamus jų molinius ekstinkcijos koeficientus ties 280 nm.

Sulfonamidiniai slopikliai

Šiame darbe tirti sulfonamidiniai junginiai buvo susintetinti laboratorijoje bendradarbių. Atsižvelgiant į junginių seriją, buvo taikomos skirtingos sintezės strategijos, įskaitant nukleofilinio pakeitimo, alkilini-
mo, amidinimo ir oksidacijos reakcijas. Junginių tapatumas ir grynumas buvo patvirtinti tinkamais analizės metodais, daugiausia branduolių magnetinio rezonanso spektroskopija ir masių spektrometrija. Išsamios sintezės, gryninimo ir charakterizavimo procedūros aprašytos atitinkamose publikacijose, kuriomis grindžiama ši disertacija. Informacija apie susintetintus junginius, įskaitant jų laboratorinius identifikatorius, pateikiama PLBD duomenų bazėje (<https://plbd.org/db/main/compound>).

Visų junginių pradiniai tirpalai buvo ruošiami DMSO, įprastos koncentracijos 10 mM to (numerical range) 50 mM, o prieš naudojimą jie buvo skiedžiami atitinkamais eksperimentiniais buferiniais tirpalais.

Eksperimentinis baltymo ir ligando sąveikų nustatymas

Šioje disertacijoje naudotų biofizikinių ir biocheminių metodų išsamias apžvalgas galima rasti nurodytuose literatūros šaltiniuose: fluorescencinis terminio poslinkio metodas (FTSA) [116, 117], izoterminė titravimo kalorimetrija (ITC) [118], rentgeno spindulių kristalografija [119–121], sustabdyto srauto (angl. *stopped-flow*) CO₂ hidratacijos aktyvumo metodas [122] ir kitus papildomus metodus, nurodytus atitinkamose į disertaciją įtrauktose publikacijose.

Metodų skyriuje apibendrinami eksperimentiniai metodai ir procedūros, svarbūs šioje disertacijoje pateiktų rezultatų interpretacijai. Išsamūs eksperimentų protokolai, konkrečioms tyrimams taikytos sąlygos, naudota įranga ir duomenų analizės procedūros pateiktos atitinkamose į disertaciją įtrauktose publikacijose.

Siekiant užtikrinti duomenų atsekamumą ir atkuriamumą, eksperimentinių duomenų rinkiniai ir su jais susiję metaduomenys buvo registruojami duomenų bazėje (PLBD; <https://plbd.org>) [125]. Rentgeno spindulių kristalografijos būdu gauti struktūriniai duomenys registruoti duomenų banke (angl. *Protein Data Bank*, PDB).

Fluorescencinis terminio poslinkio metodas

Fluorescencinis terminio poslinkio metodas (FTSA) buvo pagrindinis metodas, taikytas stebimosioms jungimosi konstantoms nustatyti daugeliui šioje disertacijoje tirtų sąveikų. Gautos konstantos vėliau buvo koreguojamos atsižvelgiant į su jungimusi susietas protonizacijos reakcijas ir taip apskaičiuojamos tikrosios jungimosi konstantos, kaip aprašyta skyriuje 3.2.2.

Tyrimo metu baltymo terminis išsivyniojimas stebimas naudojant aplinkai jautrius fluorescencinius dažus, kurie jungiasi prie atsivėrusių hidrofobinių sričių ir keičia fluorescencijos intensyvumą [126, 127]. Ligando jungimasis keičia baltymo terminį stabilumą ir jo lydymosi temperatūrą (T_m). Baltymo ir ligando jungimosi konstantos buvo nustatomos

iš ligando sukeltų terminio stabilizavimo kreivių, taikant termodinaminius modelius, įdiegtus Microsoft Excel, o vėliau mūsų laboratorijoje sukurtoje Thermott internetinėje platformoje (thermott.com) [128].

Izoterminė titravimo kalorimetrija

Izoterminė titravimo kalorimetrija (ITC) buvo naudojama tiesiogiai matuoti baltymo ir ligando jungimosi metu išsiskiriančią arba sugeriamą šilumą [129]. Šiuo metodu buvo nustatomos stebimosios jungimosi entalpijos, jungimosi konstantos ir stochiometrija, o iš jų apskaičiuojami Gibso energijos ir entropijos pokyčiai. Stebimieji termodinaminiai parametrai vėliau buvo koreguojami atsižvelgiant į susietas protonizacijos reakcijas (3.2.2 skyrius).

Siekiant įvertinti su jungimusi susietas protonizacijos reakcijas, dalis eksperimentų buvo atliekama skirtingos jonizacijos entalpijos buferiuose [11]. ITC duomenų analizė apėmė bazinės linijos nustatymą, smailių integravimą, modelio pritaikymą ir termodinaminių parametų nustatymą. Duomenys buvo analizuojami naudojant gamintojo programinę įrangą ir specializuotas programas NITPIC [130, 131], SEDPHAT [131, 132] bei GUSSI [133].

Protonizacijos parametų nustatymas

Sulfonamidų pK_a nustatymas

Sulfonamidinių slopiklių pK_a buvo nustatytos UV–Vis spektrofotometrinio titravimo metodu [56]. Metodas pagrįstas sulfonamido grupės protonizaciją ir deprotonizaciją lydinčiais sugerties spektro pokyčiais. Spektrai buvo registruojami esant skirtingoms pH vėrtėms, o iš absorbancijos priklausomybės nuo pH nustatytos pK_a vėrtės. Jos buvo naudojamos jungimuisi tinkamos ligando formos daliai apskaičiuoti ir stebimiesiems termodinaminiams parametrams koreguoti.

Sulfonamidų protonizacijos entalpijos nustatymas

Sulfonamidinių slopiklių protonizacijos entalpijos buvo nustatytos izoterminės titravimo kalorimetrijos (ITC) metodu, matuojant šilumą,

išsiskiriančią arba sugeriamą deprotonuotą sulfonamido grupę titruojant stipria rūgštimi. Iš kalorimetrinių duomenų nustatyta kiekvieno junginio protonizacijos entalpija ($\Delta H_{pr,SA}$), kuri buvo naudojama stebimosioms jungimosi entalpijoms koreguoti.

Tikrųjų termodinaminių parametrų apskaičiavimas

Tikrieji termodinaminiai parametrai apibūdina paties baltymo ir ligando sąveiką, todėl yra tinkami objektyviam skirtingų kompleksų palyginimui, priešingai nei eksperimentiškai nustatomi stebimieji parametrai.

Sulfonamidai jungiasi deprotonuota forma, koordinuodami aktyviojo centro Zn^{II} joną ir išstumdami su juo koordinuotą vandens molekulę [107]. Kadangi eksperimentinėmis sąlygomis tik dalis ligando ir fermento yra jungimuisi tinkamos formos, stebimosios jungimosi konstantos ir entalpijos priklauso nuo pH, buferio ir protonizacijos pusiausvyrų [11, 12].

Siekiant tiesiogiai palyginti termodinaminius ir struktūrinius duomenis, stebimieji jungimosi parametrai buvo koreguojami atsižvelgiant į susietas protonizacijos reakcijas ir perskaičiuojami į tikruosius termodinaminius parametrus. Paveikslas S.2 iliustruoja šios korekcijos svarbą vertinant karboanhidrazių ir sulfonamidų jungimąsi.

Jungimosi konstantų koregavimas

Tikroji jungimosi konstanta $K_{b,int}$ buvo apskaičiuota iš stebimosios jungimosi konstantos $K_{b,obs}$, atsižvelgiant į jungimuisi tinkamų baltymo (f_{CA}) ir ligando (f_{SA}) formų dalis:

$$K_{b,int} = \frac{K_{b,obs}}{f_{CA} \cdot f_{SA}} \quad (S.1)$$

Protonuotos fermento formos dalis apskaičiuota pagal lygtį

$$f_{CA} = \frac{10^{pK_{a,CA}-pH}}{1 + 10^{pK_{a,CA}-pH}} \quad (S.2)$$

o deprotonuotos sulfonamido formos dalis:

$$f_{SA} = \frac{10^{pH-pK_{a,SA}}}{1 + 10^{pH-pK_{a,SA}}} \quad (S.3)$$

Čia $pK_{a,CA}$ yra fermento aktyviajame centre su cinko jonų koordinuotos vandens molekulės rūgšties disociacijos konstanta, o $pK_{a,SA}$ – sulfonamido grupės rūgšties disociacijos konstanta.

Tikroji jungimosi Gibso energija ΔG_{int} apskaičiuota pagal standartinę termodinaminę lygtį:

$$\Delta G_{int} = -RT \ln(K_{b,int}). \quad (S.4)$$

Jungimosi entalpijos ir entropijos koregavimas

Tikroji jungimosi entalpija ΔH_{int} apskaičiuota koreguojant stebimąją jungimosi entalpiją ΔH_{obs} pagal protonų pernašos reakcijas, susijusias su sulfonamido ir su cinku koordinuotos vandens molekulės deprotonizacija bei buferio protonizacija:

$$\Delta H_{int} = \Delta H_{obs} - n_1 \Delta H_{pr,SA} - n_2 \Delta H_{pr,CA} + (n_1 + n_2) \Delta H_{pr,buf} \quad (S.5)$$

kur

$$n_1 = f_{SA} - 1 \quad (S.6)$$

$$n_2 = 1 - f_{CA}. \quad (S.7)$$

Parametrai n_1 ir n_2 apibūdina jungimosi metu perduodamų protonų skaičių.

Atskirų karboanhidrazės izofermentų $pK_{a,CA}$ ir $\Delta H_{pr,CA}$ vertės buvo anksčiau nustatytos mūsų laboratorijoje [134], o buferių protonizacijos entalpijos ($\Delta H_{pr,buf}$) paimitos iš publikuotų termodinaminių duomenų [135]. $pK_{a,SA}$ ir $\Delta H_{pr,SA}$ vertės buvo nustatytos eksperimentiškai, kaip aprašyta aukščiau.

Tikrasis jungimosi entropijos pokytis apskaičiuotas iš tikrosios Gibso energijos ir entalpijos:

$$\Delta S_{\text{int}} = \frac{\Delta H_{\text{int}} - \Delta G_{\text{int}}}{T}. \quad (\text{S.8})$$

Struktūriniai tyrimai

Karboanhidrazių ir ligandų kompleksų struktūros buvo nustatytos rentgeno spindulių kristalografijos metodu. CA izofermentų kristalai buvo auginami įprastai kambario temperatūroje sėdinčio lašo garų difuzijos metodu. Rekombinantiniai baltymai ultrafiltracijos būdu buvo koncentruojami iki 15 mg mL^{-1} to (numerical range) 50 mg mL^{-1} , o kristalizacijos lašai ($0,1 \mu\text{L}$ to (numerical range) $5 \mu\text{L}$) ruošiami sumaišant baltymo ir rezervuaro tirpalus.

Optimizuotos kristalizacijos sąlygos priklausė nuo CA izofermento. CA I kristalizacijai dažniausiai naudotas $0,1 \text{ M}$ Tris-HCl (pH 8,5), $0,2 \text{ M}$ NaCl arba amonio sulfatas ir 20% to (numerical range) 28% PEG; CA II – $0,1 \text{ M}$ natrio bicinas (pH 9,0) ir $2,0 \text{ M}$ natrio malonatas (pH 7,0); CA XII – $0,1 \text{ M}$ amonio citratas (pH 7,0–7,2) ir 26% to (numerical range) 30% PEG 4000; CA XIII – natrio citrato arba acetato buferiai (pH 4,5–5,5) su 26% to (numerical range) 30% PEG 4000.

Baltymo ir ligando kompleksai daugiausia buvo gaunami išaugintus kristalus iki vienos savaitės mirkant rezervuaro tirpale su $0,5 \text{ mM}$ to (numerical range) $1,0 \text{ mM}$ ligando. CA I kompleksams buvo taikoma kokristalizacija. CA I ir CA II kristalai buvo greitai užšaldomi tiesiogiai iš kristalizacijos tirpalo, o CA XII ir CA XIII kristalams naudotas 15% to (numerical range) 20% etilenglikolio krioprotektoriaus tirpalas.

Rentgeno spindulių difrakcijos duomenys buvo renkami DESY sinchrotrono P13 ir P14 makromolekulinės kristalografijos stotyse (Hamburgas, Vokietija) arba Vilniaus universiteto Gyvybės mokslų centre naudojant Rigaku MM007-HF difraktometrą. Struktūros buvo nustatomos molekulinio pakeitimo metodu, kaip paieškos modelius naudojant anksčiau nustatytas karboanhidrazių struktūras. Modeliai buvo iteratyviai konstruojami ir tikslinami pagal eksperimentinį elektronų tankį.

Daugumą šioje disertacijoje pateiktų kristalinių struktūrų nustatė disertacijos autorė. Dalį CA IX ir CA XII struktūrų nustatė bendradarbiai Dr. Jānis Leitāns ir Prof. Dr. Kaspars Tars (Latvijos biomedicininii

tyrimų ir studijų centras, Latvija). Visas į disertaciją įtrauktas struktūros autorė analizavo vertindama baltymo ir ligando sąveikas, jungimosi centro hidrataciją ir struktūrinius požymius, susijusius su termodinaminiais jungimosi parametrais.

Išsamios kristalizacijos sąlygos, difrakcijos duomenų rinkimo parametrai, struktūrų tikslinimo ir validavimo rodikliai bei naudota programinė įranga pateikti atitinkamose į šią disertaciją įtrauktose publikacijose.

Rezultatai

Baltymo ir ligando atpažinimą lemiantys veiksniai

Šiame skyriuje pristatomi tyrimai nagrinėja molekulinis veiksniai, lemiančius baltymo ir ligando atpažinimą, naudojant karboanhidrazių ir sulfonamidų sistemą. Išskirtinis dėmesys skiriamas tam, kaip termodinaminiai mechanizmai, ligando struktūra ir baltymo aplinka prisideda prie baltymų ir ligandų jungimosi stiprumo bei selektyvumo. Pirmiausia tiriamas entalpijos ir entropijos kompensacijos reiškinys, siekiant atskleisti, kaip panašus jungimosi stiprumas gali susidaryti dėl iš esmės skirtingų energetinių mechanizmų. Vėliau nagrinėjama nedidelių ligando modifikacijų įtaka jungimosi stiprumui ir selektyvumui. Galiausiai, struktūriniai ir termodinaminiai duomenys integruojami siekiant nustatyti sąsajas tarp baltymo ir ligando struktūros bei jungimosi energijos.

Entalpijos ir entropijos kompensavimas

Šioje dalyje apibendrinami publikacijos Paketurytė-Latvė ir kt. 2026 m. rezultatai:

Vaida Paketurytė-Latvė, Audrius Zakšauskas, Visvaldas Kairys, Edita Čapkauskaitė, Martynas Bagdonas, Alexey Smirnov, Elena Manakova, Saulius Gražulis, Daumantas Matulis

Enthalpy–Entropy Compensation in Inhibitor Binding to Carbonic Anhydrase Isozymes
Recenzuojama.

Mano indėlių sudarė tyrimo koncepcijos kūrimas, eksperimentinis jungimosi termodinamikos nustatymas ITC metodu, tikrųjų termodinaminių parametrų skaičiavimas ir analizė, rentgeno kristalografinė analizė, entalpijos ir entropijos kompensacijos reiškinų interpretavimas bei rankraščio rengimas.

Siekiant ištirti molekulinės entalpijos ir entropijos kompensacijos priežastis, buvo tiriama sulfonamidinių slopiklių serija (konkrečiai – šešių *ortho*, *meta*-dipakeistų benzensulfonamidų serija, junginiai 4a–4d, 4h ir 7h) trijų žmogaus karboanhidrazių izofermentų – CA II, CA XII ir CA XIII – atžvilgiu. Nors šių fermentų bendrosios struktūros yra labai konservatyvios ir jie tos pačios klasės slopiklius jungia panašiu stiprumu, jų aktyvieji centrai skiriasi tik keliomis aminorūgščių pozicijomis. Ši sistema sudarė sąlygas nustatyti, kaip nedideli baltymo aplinkos skirtumai veikia termodinaminius mechanizmus, lemiančius ligando atpažinimą. Jungimosi stiprumai ir kiti termodinaminiai parametrai buvo nustatyti naudojant FTSA ir ITC (šis tyrimas papildomai leido nustatyti jungimosi entalpijas) metodus, o rentgenostruktūrinė analizė suteikė informaciją apie atitinkamų baltymo ir ligando kompleksų struktūras. Buvo nustatytos septynios kompleksų su CA II (PDB kodai: 3OEN, 3OEO, 3OIC), CA XII (PDB: 3OFA, 3OFB) ir CA XIII (PDB: 3OFF, 3OFN) kristalinės struktūros.

Eksperimentiškai nustatyti (stebimieji) termodinaminiai parametrai buvo pakoreguoti atsižvelgiant į visas su sulfonamidų jungimusi susietas protonacijos pusiausvyras, įskaitant ligando deprotonaciją, cinko jono koordinuotos hidroksido grupės protonaciją bei buferio jonizaciją. Šis skaičiavimas rėmėsi eksperimentiškai nustatytais sulfonamidų pK_a : junginiams 4a–4d ir 4h nustatyta $pK_{a,SA} = 9,8$ bei protonacijos entalpija $\Delta H_{pr,SA} = -35,2 \text{ kJ mol}^{-1}$, o junginiui 7h - $pK_{a,SA} = 8,9$ ir $\Delta H_{pr,SA} = -30,4 \text{ kJ mol}^{-1}$ (naudojant TRIS buferį, kurio $\Delta H_{pr,buf} = -46,56 \text{ kJ mol}^{-1}$). Atlikus šią analizę, buvo gauti tikrieji termodinaminiai parametrai (ΔG_{int} , ΔH_{int} ir $-T\Delta S_{int}$), atspindintys tik pačią baltymo ir ligando sąveiką. Tikrosios Gibso energijos lyginant jungimąsi tarp izofermentų kito nedaug, tuo tarpu tikrosios entalpijos ir entropijos skyrėsi dešimtimis kJ mol^{-1} . Pavyzdžiui, junginio 4a tikroji Gibso energija jungiantis prie CA II ir CA XIII skyrėsi vos $2,1 \text{ kJ mol}^{-1}$ (atitinkamai $-59,3 \text{ kJ mol}^{-1}$ ir $-61,4 \text{ kJ mol}^{-1}$), tačiau atitinkamos tikrosios entalpijos skirtumas siekė net $37,2 \text{ kJ mol}^{-1}$ (atitinkamai $-26,8 \text{ kJ mol}^{-1}$ ir $-64,0 \text{ kJ mol}^{-1}$). Šį skirtumą visiškai kompensavo priešingas entropinio indėlio pokytis (čia $-T\Delta S_{int}$ pakito nuo labai palankaus $-32,5 \text{ kJ mol}^{-1}$ sistemoje su CA II iki nepalankaus $2,6 \text{ kJ mol}^{-1}$

su CA XIII). Šie stebėjimai įrodo, kad panašus jungimosi stiprumas gali susidaryti dėl iš esmės skirtingų energetinių mechanizmų balansavimo.

Nustatytas ryškus, nuo izofermento priklausantis entalpijos ir entropijos kompensacijos pobūdis, stebimas visoje tirtoje junginių serijoje. Jungimasi prie CA II lėmė palanki entropija, o jungimasi prie CA XIII - entalpija, kol CA XII pasižymėjo tarpiniais termodinaminiais profiliais. Apibendrinant, vidutinis santykinis jungimosi Gibso energijos pasiskirstymas tarp tirtų junginių parodė, kad visų reakcijų suminis jungimasis prie CA II buvo ~70 % nulemtas entropijos (~30 % entalpijos), CA XII buvo tarpinis (~70 % entalpijos, ~30 % entropijos), o jungimasis prie CA XIII buvo beveik visiškai nulemtas entalpijos (~100 % entalpijos, 0 % entropijos). Svarbu tai, kad šie skirtumai tarp izofermentų buvo gerokai didesni už tuos, kuriuos sukėlė paties ligando struktūros modifikacijos; tai rodo, kad šiuo atveju pagrindinis termodinaminio skirtumo veiksnys yra baltymo aplinka, o ne ligandas. Rezultatai dar kartą pabrėžė, kad vien stebimaisiais termodinaminiais parametrais grįsta analizė gali būti klaidinanti, nes su protonacija susiję procesai gali reikšmingai iškraipyti tiek entalpijos, tiek entropijos indėlius, todėl šiam palyginimui buvo būtina atlikti su protonizacija susijusių pusiausvyrų korekciją.

Didelės raiškos kristalinės struktūros atskleidė, kad nepaisant ryškių termodinaminių skirtumų, visų slopiklių bendri jungimosi būdai yra labai panašūs: jie koordinuoja katalizinį cinko joną per deprotonuotą sulfonamidinę grupę ir visuose kompleksuose formuoja panašius vandenilinius ryšius su Thr199. Nepaisant to, nustatyti keli subtilūs aktyviojo centro architektūros skirtumai toliau nuo šios konservatyvios jungimosi srities, kurie yra pagrindiniai stebimos termodinaminės elgsenos veiksniai. Konkrečiai, ties 62 pozicija esantis Asn izofermentuose CA II ir CA XII gali suformuoti tiesioginį vandenilinį ryšį su ligandu, tuo tarpu trumpesnis Ser62 izofermente CA XIII lemia netiesioginį vandenilinį ryšį per vandens molekulę. Ties 131 pozicija CA II ir CA XIII turi Phe, o CA XII turi Ala aminorūgštį (be to, Phe131 pasižymėjo neįprastu ir retu konformaciniu paslankumu reaguodamas į skirtingus ligandus). Kitas pastebimas skirtumas nustatytas ties 200 pozicija, kurią užima Thr izofermentuose CA II ir CA XII, tačiau Val - izofermente CA XIII, o tai reikšmingai keičia vietinę hidratacijos aplinką šalia benzensulfonamido grupės. Šių struktūrinių variacijų pakako, kad būtų pakeista pusiausvyrą tarp entalpijų ir entropijų varomųjų jėgų, kartu išlaikant bendrą jungimosi stiprumą.

Šie rezultatai rodo, kad panašūs jungimosi stiprumai gali atsirasti dėl iš esmės skirtingų energetinių mechanizmų, o subtilūs aktyviojo centro architektūros pokyčiai gali reikšmingai pakeisti entalpijų ir entropijų

indėlių balansą nedarydami įtakos bendram stiprumui. Didžiausi skirtumai buvo susiję su baltymo izofermentu, o ligando struktūros modifikacijomis turėjo gerokai mažesnį efektą šiuo atveju. Šie skirtumai dera su izofermentams būdinga aktyviojo centro hidratacija, kur vandens organizacijos skirtumai ir dehidratacijos energetiniai kaštai jungiantis ligandui lemia skirtingus termodinaminius profilius. Konkrečiai, CA II izofermente Thr200 trumpu vandeniliniu ryšiu stabilizuoja tvarkingą vandens molekulę W1, kurios išstūmimas prisijungiant ligandui sukuria labai palankų tirpiklio entropijos indėlį. Priešingai, Val200 izofermente CA XIII negali suformuoti šio stabilizuojančio vandenilinio ryšio, todėl ten esantis aktyviojo centro vandens tinklas yra mažiau tvarkingas ir laisvesnis, o jo reorganizacija suteikia nežymų entropinį pokytį, taip nulemdama dominuojantį entalpinį jungimosi pobūdį. Integruojant tikrųjų termodinaminių parametrų analizę su didelės raiškos struktūriniu apibūdinimu, tyrime sukuriamas išsamus sulfonamidų atpažinimo struktūros ir termodinamikos modelis. Šie rezultatai suteikia svarbių įžvalgų apie molekulinės entalpijos ir entropijos kompensacijos priežastis bei pabrėžia būtinybę atsižvelgti tiek į struktūrinius, tiek į termodinaminius veiksnius racionaliai projektuojant selektyvius ligandus.

Jungimosi stiprumas ir selektyvumas

Šioje dalyje apibendrinami publikacijos [136] rezultatai:

Audrius Zakšauskas, **Vaida Paketurytė-Latvė**, Alberta Jankūnaitė, Edita Čapkauskaitė, Yann Becart, Alexey Smirnov, Klára Pospíšilová, Janis Leitans, Jiří Brynda, Andris Kazaks, Lina Baranauskienė, Elena Manakova, Saulius Gražulis, Visvaldas Kairys, Kaspars Tars, Pavlína Řezáčová, Daumantas Matulis

Affinity and Selectivity of Protein-Ligand Recognition: A Minor Chemical Modification Changes Carbonic Anhydrase Binding Profile

J. Med. Chem., 2025, 68 (16), 17752–17773.

DOI: 10.1021/acs.jmedchem.5c01421

Mano indėlį sudarė eksperimentinis jungimosi stiprumo nustatymas FTSA metodu, tikrosios jungimosi termodinamikos analizė, rentgeno struktūrinė kristalografinė analizė, selektyvumo tendencijų interpretavi-

mas bei duomenų analizė ir rankraščio rengimas. Prie šios publikacijos svariai prisidėjo Dr. A. Zakšausko, Dr. A. Jankūnaitės, Dr. E. Čapkauskaitės bei Yann Becart atlikta tiriamų junginių organinė sintezė. Svarbu pažymėti, kad nepriklausomus fermentų slopinimo eksperimentus atliko K. Pospíšilová. Tai leido parodyti stiprią koreliaciją tarp nepriklausomai atliktų jungimosi (FTSA) ir slopinimo (SFA) duomenų.

Šiame tyrime sistemingai nagrinėta, kaip nedidelės *orto*-pakeistų benzensulfonamidų cheminės modifikacijos iš esmės keičia jungimosi stiprumą ir selektyvumą žmogaus karboanhidrazės (CA) izofermentų atžvilgiu. Buvo susintetinta apie 50 naujų junginių, atstovaujančių septynias chemines klases pagal sieros oksidacijos laipsnį ir jungiančiojo tilto heteroatomus (įskaitant 4-sulfanil-, 4-sulfinil- ir 4-sulfonil-esterius bei benzamidus, kartu su 4-amino-benzamidais). Jungimosi stiprumas visiems 12 kataliziškai aktyvių žmogaus CA izofermentų buvo nustatytas naudojant FTSA, kurią papildė „stopped-flow“ fermentinis slopinimas bei struktūrinė rentgenostruktūrinė analizė.

Maži struktūriniai pokyčiai lėmė jungimosi stiprumo skirtumus, siekiančius kelias eiles. Konkrečiai, 4-sulfanil-pakeisti benzensulfonamidai pasižymėjo išskirtinai stipriu ir selektyviu jungimusi prie CA IX, o jų tikrosios disociacijos konstantos pasiekė pikomolinį diapazoną. Pavyzdžiui, cikloheksilsulfanilesteris 4b pasižymėjo išskirtine $K_{d,int}$, lygia 1 pM (0,0010 nM) taikinio CA IX atžvilgiu, o jungimasis su CA II šimtus kartų silpnesnis ($K_{d,int} = 0,49$ nM). Be to, nors stambus adamantilsulfanilesteris 4g pasiekė panašų stiprumą CA IX ($K_{d,int} = 3$ pM) su keliais šimtus kartų didesniu selektyvumu prieš CA II, mažesnis metilsulfanilo analogas 4h ši selektyvumą prarado. Priešingai, sulfanilo grupės oksidacija iki sulfinilo ar sulfonilo žymiai sumažino sąveiką ir beveik panaikino selektyvumą CA IX izofermentui, nepaisant minimalių viso molekulinio karkaso pokyčių. Pavyzdžiui, sulfinilo darinys 7b ir sulfonilo darinys 9b rodė stipriai sumažėjusį ar dingusį jungimąsi CA IX (atitinkamai 28 nM ir 320 nM)

$K_{d,int}$ buvo gautos iš stebimųjų disociacijos konstantų, įvertinus nuo pH priklausančias tiek sulfonamido grupės, tiek prie cinko koordinuotos vandens molekulės protonacijos pusiausvyras. Ši analizė parodė, kad jungimasis nebuvo susijęs vien tik su pK_a poslinkiais, bet atspindėjo tikrąją baltymo ir ligando sąveiką. Ši korekcija yra kritinė, nes nors stiprus elektronų sutraukiantis sulfinilo ir sulfonilo grupių efektas sumažino sulfonamido $pK_{a,SA}$ beveik visu pH vienetu (nuo 9,8 sulfanilo junginiams iki 8,9 sulfoksidams) ir padidino deprotonuoto ligando dalį, paruoštą koordinuoti cinko joną prie pH 7,0, jų tikrosios jungimosi

konstantos iš tikrųjų buvo šimtus tūkstančių kartų silpnesnės. Tai pa-
brėžia ΔG_{int} ir kitų tikrųjų parametrų svarbą patikimai analizuojant CA
slopiklių struktūros ir jungimosi stiprumo sąsajas.

Esterių ir benzamidų analogų palyginimas parodė, kad benzamidai
prie CA izofermentų paprastai jungėsi silpniau, nors keli iš jų išlaikė
pastebimą selektyvumą CA IX atžvilgiu. Pakaito dydis ir hidrofobišku-
mas buvo lemiami veiksniai: didelės hidrofobinės grupės optimizavo
sąveikas didesniame (negu kituose izofermentuose) CA IX aktyviajame
centre, tačiau sukėlė sterines kliūtis mažesnių aktyviųjų centrų izofer-
mentuose, taip nulemdamos selektyvumą.

Konformacijų analizė suteikė struktūrinį jungimosi stiprumo tendencijų
paaiškinimą. Didelio jungimosi stiprumo sulfanil-pakeisti ligandai
jungėsi su karboanhidrazėmis tokiu būdu, kuris optimizavo hidrofobinį
efektą ir vandenilinių ryšių tinklus bei svarbiausia sudarydami koordina-
cinį ryšį su cinko jonu. Tiksliau, devynios naujai nustatytos struktūros
(įskaitant CA II-4c, CA IX-4d ir CA XIII-4d; PDB ID: 9FPQ, 9R8X, 9FPW
atitinkamai) parodė, kad ligando esterio grupė sudaro tiesioginį vandenilinių
ryšių su Asn62 izofermente CA II, tačiau pereina į vandens tarpininkau-
jamą vandenilinių ryšių tinklą su Gln70/Asn65 izofermente CA IX bei
Ser64/Asn69 izofermente CA XIII. Be to, junginio 4h kristalografinė
analizė komplekse su CA II (PDB ID: 9FPS) atskleidė netipinį jungimo-
si būdą su dviem ligando molekulėmis, kur antroji molekulė jungiasi
neklasikiniu būdu aktyviojo centro periferijoje per hidrofobines ir van-
dens tarpininkaujamas sąveikas. Tuo tarpu sulfinilo ir sulfonilo dariniai
pasižymėjo nepalankiais vidumolekuliniais vandenilinių ryšiais, lemian-
čius jungimuisi su karboanhidrazėmis netinkamą konformaciją. Mole-
kulinis dokinimas sėkmingai atkūrė eksperimentines sulfanil-pakeistų
junginių tendencijas, tačiau negebėjo paaiškinti sulfonilo darinių elgse-
nos, kuri buvo nulemta vidumolekulinių ryšių energijos.

Apibendrinant, rezultatai rodo, kad vieno *orto* pakaito oksidacijos
laipsnis ir konformacinis lankstumas gali nulemti jungimosi stiprumą
CA izofermentams ir selektyvumą kelių eilių diapazone. Šis darbas
nustato aiškų mechanistinį ryšį tarp subtilių cheminių modifikacijų,
ligando konformacinės energetikos ir selektyvaus CA IX atpažinimo,
suteikdamas vertingų gairių racionaliam karboanhidrazinių slopiklių pro-
jektavimui. Ryškus nedidelių struktūrinių modifikacijų poveikis pas-
katino platesnį tyrimą, siekiant nustatyti, kaip struktūrinė informacija
ir termodinaminiai parametrai gali būti integruoti aiškinant baltymo ir
ligando atpažinimą.

Šioje dalyje apibendrinami publikacijos [137] rezultatai:

Vaida Paketurytė-Latvė, Alexey Smirnov, Elena Manakova, Lina Baranauskienė, Vytautas Petrauskas, Asta Zubrienė, Jurgita Matulienė, Virginija Dudutienė, Edita Čapkauskaitė, Audrius Zakšauskas, Janis Leitans, Saulius Gražulis, Kaspars Tars, Daumantas Matulis

From X-ray crystallographic structure to intrinsic thermodynamics of protein-ligand binding using carbonic anhydrase isozymes as a model system

IUCrJ, 2024, 11 (4), 556–569.

DOI: 10.1107/S2052252524004627

Mano indėlį sudarė publikacijos idėjos sukūrimas, baltymo ir ligando sąveikos termodinaminė analizė, struktūrinių ir termodinaminių duomenų integravimas bei rankraščio rengimas. Svarbu pabrėžti, kad daugumą jungimosi termodinaminių parametrų nustatymų atliko Dr. A. Zubrienė ir Dr. L. Baranauskienė, o dalį straipsnyje cituojamų kristalinių struktūrų nustatė Dr. A. Smirnov, Dr. E. Manakova ir Dr. J. Leitans. Ląstelių tyrimus atliko Dr. J. Matulienė.

Siekiant nustatyti koreliacijas tarp baltymo ir ligando struktūros bei jungimosi termodinamikos, rentgenostruktūrinės analizės duomenys buvo derinami su matavimais, gautais naudojant FTSA, ITC, fermentinio slopinimo bei metodus ląstelėse. Tyrime analizuota struktūriškai panašių sulfonamidinių slopiklių (daugiausiai fluorintų ir chlorintų benzensulfonamidų serijos) sąveika, pasižyminti itin skirtingu jungimosi stiprumu bei selektyvumu žmogaus karboanhidrazių izofermentams, didžiausią dėmesį skiriant su kietaisiais hipoksiniais navikais susijusiai CA IX. Siekiant įveikti su CA IX kristalizacija susijusius iššūkius, buvo naudota chimerinė CA IX (chCA IX) kaip itin tikslus atitikmuo, turintis šešias aktyviojo centro mutacijas (A65S, N67Q, I91L, F130V, V134L, L203A), įvestas į CA II.

Kristalinės struktūros atskleidė konservatyvų sulfonamidinių slopiklių jungimosi būdą: jie koordinuoja katalizinį Zn^{II} ir formuoja panašias sąveikas aktyviajame centre. Nepaisant to, ligandai, pasižymintys beveik identiškomis jungimosi pozomis, dažnai skyrėsi savo jungimosi su karboanhidrazėmis stiprumu keliomis eilėmis. Pavyzdžiui, struktūriškai itin panašūs slopikliai EA3-3 ir EA5-3 kompleksuose su

chCA IX ir CA II (abu kristalizuojasi $P2_1$ erdvinėje grupėje su beveik identiškais elementariosios narvelio parametrais) pasižymėjo panašiomis jungimosi konformacijomis, tačiau demonstravo visiškai skirtingą jungimosi stiprumą. Šie stebėjimai įrodė, kad vien struktūrinės informacijos nepakanka norint paaiškinti jungimosi stiprumą ar izofermentinį selektyvumą.

Skirtingų eksperimentinių metodų palyginimas išryškino individualių jungimosi analizių privalumus ir apribojimus. FTSA metodas pasižymėjo plačiausiu nustatymo diapazonu - nuo milimolinių iki pikomolinių K_d verčių, tuo tarpu fermentinio slopinimo analizės bei ITC metodas apsiribojo siauresniais matavimo intervalais. Kelių ortogonalinių metodų taikymas leido patikimiau apibūdinti baltymo ir ligando sąveikas.

Pagrindinis šio darbo rezultatas - tikrųjų termodinaminių parametru nustatymas. Įvertinus su jungimusi susietas protonacijos pusiausvyras, stebimieji jungimosi duomenys buvo perskaičiuoti į ΔG_{int} , ΔH_{int} ir $-T\Delta S_{\text{int}}$. Tam reikėjo įvertinti cinko koordinuotos hidroksido grupės protonaciją aktyviajame centre (fermento $pK_{a,CA}$ siekia 6–7, o CA I ir CA XIII atveju – iki 8) bei benzensulfonamido grupės deprotonaciją ($pK_{a,SA}$ paprastai yra tarp 8 ir 11). Šie tikrieji parametrai tiksliau atspindėjo baltymo ir ligando sąveikos pasekmes ir sudarė sąlygas nustatyti prasmingas struktūros ir termodinamikos sąsajas.

Kompleksinė analizė parodė, kad ligandų jungimosi stiprumą ir selektyvumą lemia sudėtinga struktūrinių, solvatacijos bei termodinaminių veiksnių sąveika. Nedidelės cheminės modifikacijos dažnai sukeldavo esminius tikrosios jungimosi energijos pokyčius, nepaisant tik subtilių skirtumų kristalografinėse struktūrose. Konkrečiai, uodegos metilo grupės (VD12-05) pakeitimas hidroksilo uodega (VD10-35) lėmė sistemingą 7, 5–10, 7 kJ mol⁻¹ tikrosios Gibso energijos praradimą skirtinguose CA izofermentuose, nes polinė hidroksigrupė buvo linkusi formuoti vandenilinius ryšius su tirpikliu, o ne su baltymu. Priešingai, ciklooktilamino žiedo įvedimas *meta*-padėtyje (VD11-4-2) sukėlė sterinių trukdžių, kurie sumažino CA I jungimosi stiprumą 18,0 kJ mol⁻¹ (1000 kartų), tačiau užtikrino 21,9 kJ mol⁻¹ padidėjimą (4000 kartų) CA IX ir 14,9 kJ mol⁻¹ CA XII izofermentui, taip nulemdamas net 100 000 kartų selektyvumą CA IX naudai prieš CA I. Be to, cikloheksano pakaitas junginyje EA2-3 optimizavo hidrofobines sąveikas padidintoje CA IX kišenėje, suteikdamas išskirtinai didelę tikrąją Gibso energiją, lygią –75,0 kJ mol⁻¹ (fentomoliniame diapazone). Šie rezultatai pabrėžia, kad patikimai baltymo ir ligando atpažinimo interpretacijai būtina integruoti struktūrinę informaciją su tikrųjų termodinaminių parametru analize.

Apibendrinant, šiame darbe buvo sukurtas modelis, susiejantis rentgenostruktūrinės analizės duomenis su tikraisiais termodinaminiais parametrais, bei parodyta, kaip ši integracija gali pagerinti molekulinio atpažinimo supratimą ir padėti racionaliai projektuoti vaistus. Ši principą aiškiai iliustruoja priešingos varomi mechanizmai, lemiantys stiprių slopiklių jungimąsi: nors VD10-35 jungimasis su CA I yra dominuojančiai entalpinis ($\Delta H_{\text{int}} = -76,8 \text{ kJ mol}^{-1}$), selektyvus VD11-4-2 jungimasis su CA IX yra daugiausiai nulemtas entropijos ($-T\Delta S_{\text{int}} = -47,5 \text{ kJ mol}^{-1}$), nepaisant mažesnio entalpinio indėlio ($\Delta H_{\text{int}} = -24,3 \text{ kJ mol}^{-1}$), o abu kartu jie suformuoja itin stiprų kompleksą ($\Delta G_{\text{int}} = -71,8 \text{ kJ mol}^{-1}$).

Termodinaminių matavimų patikimumas

Patikima baltymo ir ligando atpažinimo interpretacija priklauso ne tik nuo struktūrinių ir termodinaminių duomenų kokybės, bet ir nuo šių duomenų gavimo bei analizės metodų. Eksperimentiniai apribojimai, parametų tarpusavio priklausomybė (koreliacijos) ir statistinė neapibrėžtis gali daryti didelę įtaką jungimosi konstantų bei termodinaminių parametų tikslumui ir interpretacijai. Šiame skyriuje pristatomi tyrimai nagrinėja du vienas kitą papildančius šios problemos aspektus: fluorescentinio terminio poslinkio metodo (FTSA) matavimų standartizavimą bei baltymo ir ligando jungimosi parametų neapibrėžties vertinimą.

FTSA matavimų standartizavimas

Šioje dalyje apibendrinami publikacijos [116] rezultatai:

Egidijus Kazlauskas, Vytautas Petrauskas, **Vaida Paketurytė**,
Daumantas Matulis

Standard Operating Procedure for Fluorescent Thermal Shift Assay (FTSA) for Determination of Protein-Ligand Binding and Protein Stability

Eur. Biophys. J., 2021, 50 (5), 707–714.

DOI: 10.1007/s00249-021-01537-1

Mano indėlį sudarė duomenų analizė ir interpretavimas, iliustracijų bei rankraščio teksto rengimas.

Tyrimas parodė, kad FTSA užtikrina atkuriamas, aukštos kokybės lydimosi kreives, tinkamas tiek baltymų stabilumui, tiek ligandų jungimosi stiprumui nustatyti plačiame diapazone (nuo 10^3 iki 10^{13}M^{-1}). Lydimosi temperatūra (T_m) nustatoma sigmoidinio denatūracijos virsmo vidurio taške, o eksperimentiniai temperatūriniai poslinkiai (ΔT_m) pritaikomi pusiausvyrinei jungimosi konstantai (K_b) skaičiuoti. Naudojant CA II (konkrečiai – acetazolamido iki $200 \mu\text{M}$ koncentracijos jungimasi su $5 \mu\text{M}$ CA II, 2% v/v DMSO ir $50 \mu\text{M}$ ANS) ir Hsp90 α N-galinį domeną (ištirtą pH 5,0–10,5 diapazone daugiakomponentiniame acetato-borato-fosfato buferyje) kaip pavyzdžius qPCR aparate (naudojant išcentrinis rotorinius spektrofluorimetrus ir $10 \mu\text{L}$ – $20 \mu\text{L}$ tūrius optimaliam temperatūros tolygumui), šis metodas patikimai nustatė ligandų sukeltus temperatūrinius poslinkius bei nuo pH priklausančius stabilumo pokyčius, patvirtindamas savo jautrumą tiek jungimosi, tiek tirpalo sąlygoms. Ligandų sukelti temperatūriniai poslinkiai leido nustatyti jungimosi konstantas plačiame sąveikų diapazone. Darbe taip pat atskleista, kad metodo veiksmingumas stipriai priklauso nuo dažų koncentracijos, baltymo grynumo, buferio sudėties bei DMSO koncentracijos, taip išryškinant pagrindinius veiksnius, lemiančius duomenų kokybę. Konkrečiai, sėkmingas matavimas priklauso nuo temperatūrai nejautraus buferio sudėties (naudojant fosfatinį arba HEPES vietoje TRIS, siekiant išvengti nuo temperatūros priklausomo pH poslinkio), DMSO koncentracijos (saugi ties 2% v/v, tačiau destabilizuojanti esant $\geq 4\%$), baltymo struktūros (teikiant pirmenybę monomeriniams baltymams prieš sudėtingus oligomerinius darinius, kurie lemia persidengiančius efektų signalus), bei dažų parinkimo (rekomenduojant ANS su $\lambda_{EX}/EM \approx 350 \text{ nm}/470 \text{ nm}$ vietoje SYPRO Orange dėl silpnescio pradinio fono, arba naudojant triptofano fluorescenciją). Apibendrinant, gauti rezultatai patvirtino, kad FTSA yra patikimas, mažas mėginių sąnaudas naudojantis ir našus metodas, tinkamas rutiniškam baltymų ir ligandų sąveikų apibūdinimui bei stabilumo atrankai.

Neapibrėžties vertinimas

Šioje dalyje apibendrinami publikacijos [138] rezultatai:

Vaida Paketurytė, Vytautas Petrauskas, Asta Zubrienė, Olga Abian, Margarida Bastos, Wen-Yih Chen, Maria João Moreno, Georg Krainer, Vaida Linkuvienė, Arthur Sedivy, Adrian Velazquez-Campoy, Mark A. Williams, Daumantas Matulis

Uncertainty in protein-ligand binding constants: asymmetric confidence intervals versus standard errors

Eur. Biophys. J., 2021, 50 (3–4), 661–670.

DOI: 10.1007/s00249-021-01518-4

Šis darbas buvo COST veiklos CA15126 (MOBIEU) užduotis, kilusi iš bendrų Europos biofizikų bendruomenės diskusijų, siekiant patobulinti molekulinį sąveikų neapibrėžties pateikimą. Mano indėlis į šį tyrimą apėmė dalyvavimą mokslinėse diskusijose, eksperimentų planavimą, duomenų analizę ir rankraščio rengimą.

Šis tyrimas parodo, kad simetrinės standartinės paklaidos, dažniausiai pateikiamos kaip $\pm x$, reikšmingai nuvertina tikrąjį baltymų–ligandų sąveikos konstantų neapibrėžtumą. Analizuojant tiek modeliuotus, tiek eksperimentinius duomenų rinkinius nustatyta, kad pusiausvyros jungimosi konstanta (K_b), disociacijos konstanta (K_d) ir Gibso energijos pokytis (ΔG) pasižymi iš esmės asimetriškais pasiskirstymais dėl juos siejančių logaritminių, eksponentinių ir atvirkštinių matematinių priklausomybių. Šių parametrų tarpusavio transformacijos dar labiau sustiprina pasiskirstymų asimetriją, todėl simetrinis paklaidų pateikimas yra statistškai netinkamas ir gali lemti fiziškai neįmanomus rezultatus, pavyzdžiui, neigiamas jungimosi konstantų reikšmės. Pavyzdžiui, 1-uoju atveju (Case 1 publikacijoje), darant prielaidą, kad K_b pasiskirstymas yra normalusis, kurio vidurkis lygus $2,0 \cdot 10^7 \text{ M}^{-1}$, o standartinis nuokrypis $-0,6 \cdot 10^7 \text{ M}^{-1}$, transformavus į K_d gaunamas stipriai asimetriškas pasiskirstymas (asimetrijos koeficientas siekia 4,3), be to, normalusis pasiskirstymas priskiria teigiamą tikimybę fiziškai neįmanomoms neigiamoms K_b reikšmėms. Priešingai, jei normalumas taikomas ΔG su vidurkiu $-41,67 \text{ kJ mol}^{-1}$ prie 25°C ir standartiiniu nuokrypiu $0,77 \text{ kJ mol}^{-1}$, gaunamas simetrinis 95 % pasikliautinis intervalas $-43,18 \text{ kJ mol}^{-1}$ to (numerical range) $-40,16 \text{ kJ mol}^{-1}$, kurį tiesiogiai transformavus gaunamas prasmingas asimetris K_b 95 % pasikliautinis intervalas $[1, 09; 3, 68] \times 10^7 \text{ M}^{-1}$.

Monte Karlo simuliacijos, atliktos 1:1 jungimosi titravimams, parodė, kad apskaičiuotos K_b reikšmės tampa vis labiau teigiamai asimetriškos didėjant jungimosi stiprumui, tuo tarpu ΔG išlieka artimesnis normaliam pasiskirstymui esant pakankamai optimalioms eksperimentinėms sąlygoms. Siekiant įvertinti eksperimentiškai realistišką situaciją, 2-uoju atveju (Case 2 publikacijoje) sumodeliuotas VP-ITC eksperimentas, kurį sudarė $20 \times 15 \mu\text{L}$ $200 \mu\text{M}$ ligando įšvirkštimai į $20 \mu\text{M}$ baltymo tirpalą, esant ($\Delta H = -10 \text{ kJ mol}^{-1}$) ir pastoviam įšvirkštimo šilumos triu-

kšmui $3 \mu\text{J}$. Tokiomis sąlygomis, padidinus tikrąją K_b reikšmę nuo $2 \cdot 10^5 \text{ M}^{-1}$ (Wiseman parametras $c = 4$) iki $2 \cdot 10^7 \text{ M}^{-1}$ ($c = 400$), gautos K_b reikšmės pasiskirstymo asimetrijos koeficientas išaugo nuo 1,3 ($(CI_{95}=[0,78;5,8] \times 10^5 \text{ M}^{-1})$) iki net 23 ($(CI_{95}=[0,51;22] \times 10^7 \text{ M}^{-1})$). Esant tokiai didelio jungimosi stiprumo c reikšmei, tik 3–4 duomenų taškai patenka į jungimosi izotermos pereinamąją sritį, todėl kreivės linkis, apribojantis K_b , yra prastai išreikštas, o gautos reikšmės tampa itin jautrios triukšmui ir suformuoja asimetrišką skirstinį su ilga viršutine uodega.

Su bendraautoriais priėjome išvadą, kad profilio tikėtinumą (angl. *profile likelihood*) metodu apskaičiuoti 95 % pasikliautinieji intervalai ($CI_{P,95}$) yra nuoseklesnis ir tarpusavyje transformuojamas neapibrėžtumo vaizdavimo būdas visiems termodinaminiams dydžiams. Asimetriniai intervalai yra apskaičiuojami naudojant Fisherio–Snedecoro F pasiskirstymą, sistemingai keičiant tiriamą parametą fiksuotais žingsniais ir minimizuojant liekamųjų kvadratų sumą kitų laisvųjų parametų atžvilgiu. Skirtingai nei standartinės paklaidos, šis metodas užtikrina visišką viršutinių ir apatinių neapibrėžtumo ribų atitikimą transformuojant tarp K_b ir ΔG , išsaugant visą informaciją apie eksperimentinį kintamumą. Pavyzdžiui, sumodeliuotam titravimui su tikrosiomis reikšmėmis ($K_b = 2 \times 10^6 \text{ M}^{-1}$) ir ($\Delta H = 41,8 \text{ kJ mol}^{-1}$), profilio tikėtinumą metodu ($CI_{P,95}$) duoda identiškus atitinkančius intervalus ($[1,44;3,30] \times 10^6 \text{ M}^{-1}$) parametrui (K_b) ir $[-37,2;-35,1] \text{ kJ mol}^{-1}$ parametrui (ΔG)), nepriklausomai nuo to, kuris iš jų buvo tiesiogiai optimizuojamas, tuo tarpu standartiniai asimptotiniai simetriniai intervalai ($CI_{S,95}$) šio sąryšio po transformacijos neišlaiko. Tad rezultatai rodo, kad asimetriniai pasikliautinieji intervalai yra būtini siekiant tiksliai interpretuoti ir pakartotinai naudoti biofizikinius duomenis, todėl baltymų–ligandų sąveikos tyrimuose jie turėtų pakeisti nusistovėjusias simetrines standartines paklaidas.

Diskusija

Pagrindinis struktūra paremto vaistų kūrimo tikslas yra numatyti, kaip cheminės modifikacijos veikia jungimosi stiprumą ir selektyvumą, bei panaudoti šias žinias kuriant junginius, kurie stipriai jungiasi prie norimo taikinio ir kartu nesąveikauja (arba silpnai sąveikauja) su kitais baltymais. Nepaisant didelės pažangos struktūrinės biologijos, kompiuterinės chemijos ir biofizikos srityse, šį tikslą pasiekti vis dar sudėtinga.

Šioje disertacijoje pateikti rezultatai išryškina keletą veiksnių, kurie apsunkina baltymų–ligandų atpažinimo interpretavimą ir prognozavimą.

Skyriuje Entalpijos–entropijos kompensavimas pateikti rezultatai suteikia papildomų įžvalgų apie molekulinės entalpijos–entropijos kompensavimo priežastis. Ankstesni Whitesides ir bendradarbių darbai [9, 139] parodė, kad net ir nedidelės ligandų modifikacijos gali lemti didelius entalpijos ir entropijos pokyčius, nors bendras jungimosi stiprumas su CA II kinta tik nežymiai. Tai pabrėžia hidratacijos ir tirpiklio persitvarkymo svarbą molekuliniam atpažinime. Šiame darbe gauti rezultatai rodo, kad reikšmingi termodinaminių parametrų skirtumai gali atsirasti ir tarp karboanhidrazės izofermentų, nepaisant labai konservatyvios aktyviojo centro architektūros ir panašios ligandų jungimosi geometrijos. Tai rodo, kad entalpijos–entropijos kompensavimą lemia ne tik paties ligando savybės, bet ir aktyviojo centro aplinka bei kiti baltymo ypatumai. Net ir nedideli aktyviojo centro aminorūgščių liekanų, hidratacijos tinklų ar baltymo dinamikos skirtumai yra pakankami pakeisti entalpijos ir entropijos pusiausvyrą. Todėl jungimosi termodinaminės savybės turėtų būti vertinamos kaip visos baltymo–ligando sistemos savybė, o ne kaip tiesioginiai baltymo–ligando tarpmolekuliniai ryšiai.

Praktiniu požiūriu šie rezultatai atskleidžia tiek termodinaminės analizės galimybes, tiek jos ribotumus. Nors jungimosi energijos išskaidymas į entalpinės ir entropinės dedamąsias leidžia atskirti baltymo–ligando kompleksus, kurie, vertinant vien pagal jungimosi stiprumą, atrodo labai panašūs, šių dedamųjų interpretavimas išlieka sudėtingas, nes entalpija ir entropija atspindi daugelio tarpusavyje susijusių procesų bendrą poveikį. Kaip pažymėjo Klebe ir kiti autoriai [140], atskirų termodinaminių indėlių priskyrimas konkrečioms molekulinėms sąveikoms dažnai nėra tiesioginis, nes tarpmolekuliniai kontaktai, hidratacija, konformacinis lankstumas ir dinamika kinta vienu metu. Vis dėlto termodinaminių parametrų gebėjimas atskirti sistemas, pasižyminčias panašiomis struktūromis ir jungimosi stiprumais, rodo jų vertę tiriant molekulinio atpažinimo procesus ir pabrėžia papildomą informaciją, kurią jie suteikia greta įprastinių jungimosi stiprumo matavimų.

Skyriuje Jungimosi stiprumas ir selektyvumas aprašyti rezultatai dar kartą pabrėžia konformacinių veiksnių svarbą baltymų–ligandų atpažinimui. Daugybė tyrimų parodė, kad net ir nedidelės struktūrinės modifikacijos gali pakeisti ligando lankstumą, konformacines preferencijas ir baltymo dinamiką, taip sukeldamos netikėtai didelius jungimosi pokyčius. Gilson ir bendradarbiai [3, 10] pabrėžė konformacinės entropijos svarbą molekuliniam atpažinime bei sunkumus tiksliai įvertinant jos indėlį į jungimąsi, o Wand [142], Akke [6, 70] ir jų kolegos, rem-

damiesi NMR tyrimais, parodė, kad baltymo ar ligando dinamikos pokyčiai gali reikšmingai prisidėti prie jungimosi termodinamikos. Neseniai Wankowicz ir bendradarbiai [66] parodė, kad net iš rentgeno kristalografijos duomenų nustatytas konformacinis heterogeniškumas gali suteikti vertingų įžvalgų apie entropines ligando jungimosi pasekmes. Tačiau, nepaisant šios pažangos, eksperimentinis konformacinės entropijos charakterizavimas išlieka gerokai sudėtingesnis nei bendrų jungimosi termodinaminių parametru nustatymas. Įprastiniuose ITC eksperimentuose bendras entropijos pokytis apskaičiuojamas netiesiogiai iš išmatuotų Gibso energijos ir entalpijos verčių. Tuo tarpu atskiri baltymo konformacinių pokyčių, ligando lankstumo, tirpiklio persitvarkymo ir kitų procesų indėliai dažniausiai negali būti išskirti. Todėl šiame darbe gauti rezultatai atskleidžia tiek konformacinių veiksnių svarbą, tiek dabartinius jų charakterizavimo ribotumus. Šiame tyrime stebėti ryškūs jungimosi stiprumo sumažėjimai oksiduojant sierą turinčius pakaitus rodo, kad ligando konformacijos pokyčiai gali nusverti tikėtiną jungimosi stiprėjimą, susijusį su elektronus ištraukiančių grupių įvedimu.

Skyriuje Sąsajos tarp struktūros ir termodinamikos pateikti rezultatai pabrėžia stebimų ir tikrųjų termodinaminių parametru atskyrimo svarbą interpretuojant baltymų–ligandų sąveikas. Susietų protonacijos pusiausvyrų įtaka jungimosi termodinamikai žinoma jau kelis dešimtmečius, o Baker ir Murphy [11], Whitesides ir bendradarbių [143], Klebe ir bendradarbių [140] darbai bei mūsų laboratorijos tyrimai [134] pabrėžė, kad stebimi termodinaminiai parametrai nebūtinai tiesiogiai atspindi molekulinio atpažinimo energiją. Nepaisant to, tikrųjų termodinaminių parametru analizė šiuolaikiniuose baltymų–ligandų tyrimuose vis dar naudojama palyginti retai, nes reikalingų protonacijos parametru nustatymas yra eksperimentiškai sudėtingas, o daugeliui sistemų tokie duomenys apskritai nėra prieinami [140]. Dėl šios priežasties stebimi termodinaminiai parametrai dažnai interpretuojami taip, tarsi jie tiesiogiai apibūdintų baltymo–ligando sąveikas. Šiame darbe gauti rezultatai rodo, kad su protonacija susiję indėliai gali reikšmingai pakeisti tiek jungimosi termodinamikos dydį, tiek jos interpretaciją, todėl prasmingam ligandų, baltymų ir jungimosi mechanizmų palyginimui dažnai būtina remtis tikraisiais, o ne stebimais parametrais. Plačiau žvelgiant, šie rezultatai rodo, kad tikslioms struktūros ir termodinamikos sąsajoms nustatyti reikalinga ne tik aukštos raiškos struktūrinė informacija, bet ir griežtas procesų, lemiančių išmatuotus termodinaminius parametrus, atskyrimas.

Skyriuose FTSA standartizavimas ir Neapibrėžties vertinimas aprašy-

ti metodologiniai tyrimai pabrėžia eksperimentų dizaino ir neapibrėžties vertinimo svarbą patikimai baltymų–ligandų jungimosi duomenų interpretacijai. Literatūroje publikuotų slopinimo verčių palyginimo analizė [145] atskleidė reikšmingą ir didelį rezultatų kintamumą tarp nepriklausomų tyrimų ir išryškino sunkumus, susijusius su skirtingomis eksperimentinėmis metodikomis bei sąlygomis gautų duomenų palyginimu. Toks kintamumas ypač svarbus skaičiavimo metodams, kurie modelių kūrimui, validavimui ir palyginimui dažnai remiasi publikuotais jungimosi duomenimis. Netikslūs ar tarpusavyje nesuderinami eksperimentiniai matavimai gali riboti ne tik eksperimentinių rezultatų interpretaciją, bet ir prognozuojančių modelių kūrimą. Negalima tikėtis, kad modeliai, apmokyti naudojant triukšmingus eksperimentinius duomenis, pasieks didesnę prognozavimo tikslumą nei leidžia pačių matavimų kokybė. Panašūs klausimai buvo keliami ir karboanhidrazijų tyrimų srityje, kur Liljas ir bendradarbiai [146] suabejojo kai kurių plačiai cituojamų slopinimo duomenų rinkinių tinkamumu ir teigė, kad jų tikslumas bei patikimumas galėjo būti pervertinti. Šie pastebėjimai rodo, kad prasmingoms struktūros–aktyvumo ir struktūros–termodinamikos sąsajoms nustatyti reikalingi ne tik aukštos kokybės eksperimentiniai matavimai, bet ir realistiškas jų neapibrėžties įvertinimas. Nors eksperimentinės neapibrėžties visiškai pašalinti neįmanoma, jos griežtas įvertinimas yra būtinas norint atskirti tikrus biologinius efektus nuo metodologinio kintamumo ir užtikrinti prasmingą skirtingų tyrimų palyginimą.

Nors šiuose darbuose nagrinėti skirtingi baltymų–ligandų sąveikų aspektai, visi jie pagrindžia tą pačią pagrindinę išvadą – jungimosi stiprumas yra sudėtinis dydis, kurį lemia daugelis tarpusavyje susijusių procesų. Hidratacija, konformaciniai pokyčiai, protonacijos pusiausvyros ir eksperimentinė neapibrėžtis gali reikšmingai paveikti molekulinio atpažinimo supratimą. Todėl siekiant nustatyti prasmingas struktūros ir termodinamikos sąsajas nepakanka vien struktūrinės informacijos ar jungimosi stiprumo matavimų – būtina įvertinti ir procesus, kurie lemia stebimus eksperimentinius parametrus. Šioje disertacijoje pateikti rezultatai atskleidžia tiek struktūra paremtą vaistų kūrimo galimybes, tiek jo ribotumus ir pabrėžia, kad tikslus baltymų–ligandų sąveikų supratimas priklauso nuo gebėjimo išskaidyti stebimus termodinaminius parametrus į atskirus juos formuojančius procesus bei susieti juos su struktūrine informacija. Tolesnė pažanga šioje srityje priklausys nuo vis detalesnio jungimąsi lemiančių procesų apibūdinimo, jų tarpusavio sąveikos supratimo ir jų indėlio į molekulinį atpažinimą įvertinimo.

Išvados

- Karboanhidrazų izofermentų palyginimas parodė, kad nedideli aktyviojo centro skirtumai, įskaitant hidrataciją bei Thr arba Val liekanos buvimą aktyviajame centre, buvo susiję su reikšmingais ΔH bei $T\Delta S$ skirtumais (siekusiais apie 40 kJ mol^{-1} nepaisant panašios prisijungimo geometrijos ir ΔG reikšmių). *Šie rezultatai pabrėžia tarpmolekulinės sąveikos termodinamikos jautrumą baltymo aktyviojo centro aplinkai ir atskleidžia molekulinio atpažinimo interpretavimo vien pagal sąveikos stiprumą ar struktūrą ribotumus.* (Publikacija 1)
- Nedidelės cheminės modifikacijos lėmė reikšmingus sąveikos stiprumo ir selektyvumo pokyčius įvairių CA izofermentų atžvilgiu. Analizė parodė, kad sieros turinčių pakaitų oksidacija skatino vidumolekulinių vandenilinių ryšių su sulfonamido grupe susidarymą, keitė ligando konformaciją ir silpnino baltymo ir ligando sąveiką. *Šie rezultatai atskleidžia, kad baltymo ir ligando atpažinimą lemia glaudžiai tarpusavyje susijusių veiksnių visuma.* (Publikacija 2)
- Atsižvelgus į su jungimusi susietas protonacijos pusiausvyras, buvo apskaičiuoti tikrieji termodinaminiai parametrai, kurie reikšmingai skyrėsi nuo stebimųjų reikšmių dėl protonacijos poveikio. Tai parodė, kad prasmingam sąveikos palyginimui būtina atskirti baltymo ir ligando jungimosi bei protonacijos indėlius į bendrą sąveikos termodinamiką. (Publikacija 3)
- Sisteminis FTSA eksperimentų sąlygų vertinimas parodė protokolų standartizavimo svarbą siekiant gauti tarpusavyje palyginamus baltymų ir ligandų jungimosi duomenis. Asimetriniai profilinio tikėtinumo (angl. *profile likelihood*) pasikliautinieji intervalai suteikė pagrįstą neapibrėžties įvertinimą, tuo tarpu simetrinės standartinės paklaidos yra klaidinančios dydžiams, transformuojamiems netiesiniuose modeliuose. (Publikacija 4 ir Publikacija 5)

Acknowledgements

I would like to express my gratitude to Prof. Daumantas Matulis for his guidance and continuous support throughout my studies. His scientific insight and dedication to understanding the fundamental principles of molecular recognition have profoundly influenced the way I approach scientific research.

I am sincerely grateful to Dr. Asta Zubrienė for introducing me to experimental research on protein–ligand interactions and for her support during the early stages of my scientific career; to Dr. Alexey Smirnov for sharing his expertise in X-ray crystallography and structural biology; and to Dr. Vytautas Petrauskas for shaping my approach to data analysis, scientific visualization, and problem-solving. I am also grateful to Dr. Visvaldas Kairys for valuable discussions on computational modeling and for his insights into the interpretation of ligand conformational rearrangements, including rotameric states and intermolecular hydrogen-bonding interactions. I also gratefully acknowledge Dr. Elena Manakova and her colleagues for their assistance with synchrotron diffraction data collection, as well as Dr. Saulius Gražulis for developing and maintaining the computational infrastructure and software tools used in crystallographic structure determination and analysis.

I am especially thankful to the organic synthesis experts Dr. Audrius Zakšauskas, Dr. Edita Čapkauskaitė, and Dr. Alberta Jankūnaitė for synthesizing and providing compounds that were essential for this research. I would also like to acknowledge the many colleagues whose work over the years established and maintained the protein and compound collections. I am especially grateful to Vilma Michailovienė (1968–2022), whose dedication to protein purification contributed substantially to the carbonic anhydrase platform used throughout this work. I also thank Aurelija Mickevičiūtė, Dr. Vaida Juozapaitienė, Dr. Jurgita Matulienė, Marius Gedgaudas, Jelena Jachno, Dr. Justina Kazokaitė-Adomaitienė, Dr. Virginija Dudutienė, Dr. Aivaras Vaškevičius, Dr. Benas Balandis, Dr. Irena Vaškevičienė, and many other colleagues and students whose

contributions are not individually acknowledged here but were nevertheless important to the work presented in this dissertation. I also thank Darius Lingė and Ernestas Urniežius for their assistance with the deposition of structural and binding data to the Protein–Ligand Binding Database (PLBD). I am thankful to Tomas Šimkūnas and Kristina Gluščiukaitė, whose collaboration provided me with valuable opportunities to develop my mentoring and scientific supervision skills.

I am grateful to all current and former members of the Department of Biothermodynamics and Drug Design for scientific discussions and for creating a supportive research environment. Over the years, the laboratory has been a place of scientific discovery and professional growth, where colleagues generously shared their knowledge, discussed ideas, helped solve experimental challenges, and provided encouragement throughout my doctoral journey. I would also like to thank Skaidrius Jukša and Jokūbas Preikša for their contributions to our ongoing computational modeling project and for many valuable discussions that helped shape future directions of my work. I am also thankful to all collaborators who contributed to the projects that shaped this dissertation. I am grateful to Dr. Joana Smirnovienė, Dr. Vytautas Petrauskas, and Prof. Daumantas Matulis for reviewing this dissertation and providing valuable comments and suggestions.

I gratefully acknowledge the financial support provided by the Research Council of Lithuania, Vilnius University, and other funding sources that supported my doctoral studies. Their support enabled the execution of the research presented in this dissertation and participation in scientific conferences, workshops, training schools, and collaborative activities in Lithuania and abroad. I would particularly like to acknowledge support from the Research Council of Lithuania through the project "Development and Application of a Protein–Ligand Interaction Database for Artificial Intelligence Solutions" (S-ITP-25-10).

Finally, I would like to thank my family. To my husband and son, thank you for your patience, understanding, and unwavering support. I also thank my father and sister for their support and belief in me. I am especially grateful to my mother, whose influence helped shape the person I am today. Although she is no longer here to see this achievement, her influence has accompanied me throughout this journey, and I know she would have been proud of this milestone.

Vaida Paketurytė-Latvė
Vilnius
September 16, 2026

PUBLICATIONS INCLUDED IN THE DISSERTATION

1. **Paketurytė-Latvė, V.; Zakšauskas, A.; Kairys, V.; Čapkauskaitė, E.; Bagdonas, M.; Smirnov, A.; Manakova, E.; Gražulis, S.; Matulis, D. Enthalpy–Entropy Compensation in Inhibitor Binding to Carbonic Anhydrase Isozymes.** *Manuscript under review.*
2. Zakšauskas, A.; **Paketurytė-Latvė, V.**; Jankūnaitė, A.; Čapkauskaitė, E.; Becart, Y.; Smirnov, A.; Pospíšilová, K.; Leitans, J.; Brynda, J.; Kazaks, A.; Baranauskienė, L.; Manakova, E.; Gražulis, S.; Kairys, V.; Tars, K.; Řezáčová, P.; Matulis, D. **Affinity and Selectivity of Protein-Ligand Recognition: A Minor Chemical Modification Changes Carbonic Anhydrase Binding Profile.** *J. Med. Chem.* 2025, 68 (16), 17752–17773. DOI: 10.1021/acs.jmedchem.5c01421
3. **Paketurytė-Latvė, V.**; Smirnov, A.; Manakova, E.; Baranauskienė, L.; Petrauskas, V.; Zubrienė, A.; Matulienė, J.; Dudutienė, V.; Čapkauskaitė, E.; Zakšauskas, A.; Leitans, J.; Gražulis, S.; Tars, K.; Matulis, D. **From X-Ray Crystallographic Structure to Intrinsic Thermodynamics of Protein-Ligand Binding Using Carbonic Anhydrase Isozymes as a Model System.** *IUCrJ* 2024, 11 (4), 556–569. DOI: 10.1107/S2052252524004627
4. Kazlauskas, E.; Petrauskas, V.; **Paketurytė, V.**; Matulis, D. **Standard Operating Procedure for Fluorescent Thermal Shift Assay (FTSA) for Determination of Protein-Ligand Binding and Protein Stability.** *Eur. Biophys. J.* 2021, 50 (5), 707–714. DOI: 10.1007/s00249-021-01537-1
5. **Paketurytė, V.**; Petrauskas, V.; Zubrienė, A.; Abian, O.; Bastos, M.; Chen, W.-Y.; Moreno, M. J.; Krainer, G.; Linkuvienė, V.; Sedivy, A.; Velazquez-Campoy, A.; Williams, M. A.; Matulis, D. **Uncertainty in Protein–Ligand Binding Constants: Asymmetric Confidence Intervals versus Standard Errors.** *Eur. Biophys. J.* 2021, 50 (3-4), 661–670. DOI: 10.1007/s00249-021-01518-4

Publication 1

Enthalpy–Entropy Compensation in Inhibitor Binding to Carbonic Anhydrase Isozymes

Vaida Paketurytė-Latvė, Audrius Zakšauskas, Visvaldas Kairys, Edita Čapkauskaitė, Martynas Bagdonas, Alexey Smirnov, Elena Manakova, Saulius Gražulis, Daumantas Matulis

Under review.

Enthalpy–Entropy Compensation in Inhibitor Binding to Carbonic Anhydrase Isozymes

Vaida Paketurytė-Latvė,[†] Audrius Zakšauskas,[†] Visvaldas Kairys,[‡] Editā Čapkauskaitė,[†] Martynas Bagdonas,[†] Alexey Smirnov,[†] Elena Manakova,[¶] Saulius Gražulis,[§] and Daumantas Matulis^{*,†}

[†]*Department of Biothermodynamics and Drug Design, Institute of Biotechnology, Life Sciences Center, Vilnius University, Saulėtekio 7, Vilnius, LT-10257, Lithuania*

[‡]*Department of Bioinformatics, Institute of Biotechnology, Life Sciences Center, Vilnius University, Saulėtekio 7, Vilnius, LT-10257, Lithuania*

[¶]*Department of Protein – DNA Interactions, Institute of Biotechnology, Life Sciences Center, Vilnius University, Saulėtekio 7, Vilnius, LT-10257, Lithuania*

[§]*Sector of Crystallography and Chemoinformatics, Institute of Biotechnology, Life Sciences Center, Vilnius University, Saulėtekio 7, Vilnius, LT-10257, Lithuania*

E-mail: daumantas.matulis@bti.vu.lt

SIGNIFICANCE Enthalpy–entropy compensation is widely observed in protein–ligand binding, yet its molecular origins remain difficult to resolve. Here, we combine binding thermodynamics with high-resolution X-ray crystallography to investigate a matched series of sulfonamide inhibitors across closely related human carbonic anhydrase isozymes. By correcting experimentally observed binding parameters for proton-linked equilibria, we show that similar binding affinities can arise from markedly different enthalpic and entropic contributions.

Abstract

Enthalpy–entropy compensation is a widely observed feature of protein–ligand recognition, yet its structural origins remain difficult to resolve. Here, we investigate a matched series of sulfonamide inhibitors binding to three closely related human carbonic anhydrase isozymes (CA II, CA XII, and CA XIII) using isothermal titration calorimetry, fluorescence thermal shift assays, and high-resolution X-ray crystallography. Although these isozymes share nearly identical overall folds and exhibit highly similar *intrinsic* Gibbs energies of binding, their enthalpic and entropic contributions diverge markedly. Correction of all observed thermodynamic quantities for proton-linked equilibria reveals that this divergence originates from subtle differences in the active-site microenviron-

ment. In particular, single-residue variations—including the Thr/Thr/Val substitution at position 200, the Phe/Ala/Phe variation at position 131, and the Asn/Asn/Ser difference at position 62—modulate hydrogen-bond geometry, steric complementarity, and local solvation, thereby reshaping the balance between enthalpic and entropic driving forces. These results demonstrate how closely related proteins can achieve similar binding affinities through distinct energetic strategies and establish a coherent structure–thermodynamics framework for sulfonamide recognition.

Abbreviations

CA – carbonic anhydrase;

FTSA – fluorescence-based thermal shift assay (also called thermal shift assay, TSA), alternatively

termed ThermoFluor® or differential scanning fluorimetry (DSF);
ITC – isothermal titration calorimetry;
PLBD – Protein-Ligand Binding Database (plbd.org);
PDB – Protein Data Bank;
TRIS – tris(hydroxymethyl)aminomethane buffer;
Zn^{II} – zinc ion;

Introduction

Small-molecule engineering to achieve high-affinity binding to target proteins is a central challenge in drug discovery and design. Truly rational drug design requires a detailed understanding of how the chemical structure of a small molecule recognizes and interacts with specific surface regions of a protein. In addition, the thermodynamics of the binding process must be thoroughly characterized, along with the underlying physical forces among all interacting components – not only the protein and the ligand, but also the surrounding network of water molecules, which remains difficult to model and simulate^{1,2}.

Protein-ligand binding is governed by a balance of enthalpic and entropic contributions arising from intermolecular interactions, dehydration, and conformational reorganization^{3–5}. Chemically similar ligands often display markedly different enthalpies and entropies upon binding to the same protein; however, their sum – the standard Gibbs energy of binding – tends to remain nearly constant, giving rise to the widely observed phenomenon of enthalpy-entropy compensation^{6–13}. This compensation complicates mechanistic interpretation: ligands with comparable affinities may bind through distinct combinations of hydrogen bonding, hydrophobic contacts, and solvent rearrangements, while even minor chemical modifications can dramatically shift this balance^{14,15}. Understanding how structural features translate into specific thermodynamic signatures, therefore, remains a challenge in quantitative ligand design.

A recurring concern in studies of enthalpy-entropy compensation is that the phenomenon may be non-physical and instead arise from ex-

perimental error. In typical experiments, the binding enthalpy (ΔH) is measured directly, whereas the entropy change (ΔS) is not obtained independently. There is no direct experimental method to determine the standard entropy change upon binding^{16,17}; instead, binding affinity (K_d or ΔG) is measured – often by techniques such as ITC – and ΔS is subsequently calculated from the relationship $\Delta G = \Delta H - T\Delta S$. Consequently, if both ΔH and ΔG carry significant experimental errors, the apparent enthalpy-entropy compensation may simply reflect propagated uncertainty rather than a true physical effect. It is therefore essential to carefully assess the accuracy and precision of these measurements^{17,18}.

The interaction of carbonic anhydrase (CA) with sulfonamide ligands has been widely employed as a model system for investigating enthalpy-entropy compensation, supported by high-quality thermodynamic measurements and X-ray crystallographic data. In one study, a series of fluorinated benzothiazole sulfonamides was designed, and their binding to human CA II was characterized using isothermal titration calorimetry (ITC)¹¹. The results suggested that water networks play a partial or even dominant role in driving the observed compensation effects. In a related study, conformational variations in alkyl chains were shown to modulate the thermodynamics of sulfonamide binding to CA II, highlighting the importance of ligand flexibility in shaping enthalpic and entropic contributions¹⁹.

However, the human body contains twelve catalytically active carbonic anhydrase (CA) isozymes²⁰. All catalyze the same CO₂ hydration reaction and play important roles in carbon metabolism and pH homeostasis. In addition, there are three inactive but highly homologous carbonic anhydrase-related proteins (CARPs), which lack Zn^{II} due to mutations of the histidine residues responsible for metal coordination. Sequence identity among these proteins ranges from approximately 30% to 70%, yet their overall fold is highly conserved. In the active site, the differences are largely confined to a small number of amino acids, making this family a convenient system for studying how

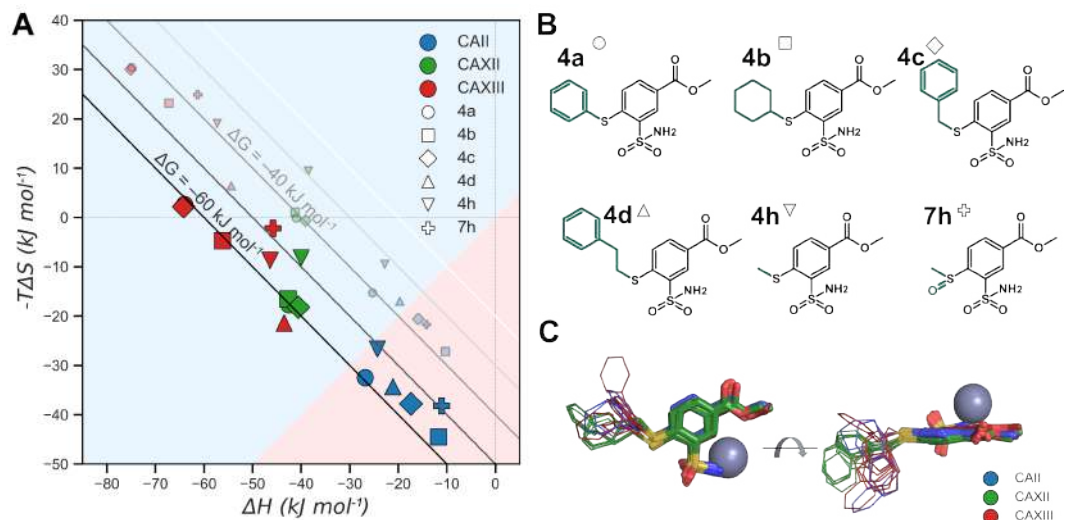


Figure 1: (A) Enthalpy-entropy compensation governing sulfonamide binding to human carbonic anhydrases: CA II (blue), CA XII (green), and CA XIII (red). Distinct thermodynamic signatures reflect isozyme-dependent binding mechanisms. Individual compounds are represented by different symbols. Intrinsic thermodynamic parameters are shown as the largest data points, while observed parameters are displayed as smaller, semi-transparent symbols. Diagonal lines indicate Gibbs energies in 10 kJ mol⁻¹ increments. Intrinsic Gibbs energies cluster around -60 kJ mol⁻¹, whereas observed Gibbs energies are centered near -40 kJ mol⁻¹. The pastel blue background highlights enthalpy-favored binding regimes, while the pastel red background corresponds to the entropy-favored interactions. (B) Chemical structures of the sulfonamide inhibitors examined in this study. (C) Representative X-ray crystal structures of superimposed sulfonamides bound in the active sites of CA II (blue; PDB IDs: [30EN](#), [30EO](#), [9FPQ](#), [9FPR](#)), CA XII (green; PDB IDs: [30FA](#), [30FB](#)), and CA XIII (red; PDB IDs: [30FP](#), [30FN](#), [9FPV](#), [9FPW](#)), illustrating isozyme-specific ligand orientations. The gray sphere represents the catalytic Zn^{II} ion.

specific residues influence small-molecule recognition.

For many small-molecule inhibitors, including sulfonamides, the situation is further complicated by proton-linked equilibria^{21,22}. Binding is accompanied by deprotonation of the sulfonamide group and protonation of the metal-bound water molecule in the active site, such that the experimentally measured enthalpy and affinity reflect not only ligand-protein interactions but also the energetics of these coupled protonation events. As a result, parameters obtained from ITC and FTSA cannot be directly interpreted in structural terms. The extraction of intrinsic thermodynamic parameters—

corrected for all protonation-linked processes—is therefore essential to isolate the true energetic contributions of binding. Despite their importance, such intrinsic parameters have rarely been determined systematically across related protein-ligand systems.

Structural determinants of binding energetics are equally challenging to resolve without high-resolution structural information. Even subtle changes in ligand structure or in the microenvironment of the binding site can modify hydrogen-bond geometry, steric complementarity, and local solvation patterns, leading to substantial shifts in ΔH and $T\Delta S$ ^{11,16,23}. However, in the absence of paired crystallo-

Table 1: Thermodynamic parameters of investigated compounds binding to CA II, CA XII, and CA XIII at 37 °C. (A) Observed enthalpy (ΔH_{obs} , kJ mol⁻¹) and Gibbs energy (ΔG_{obs} , kJ mol⁻¹) changes determined by ITC in TRIS at pH 7.0 (values represent the mean of two independent experiments; individual measurements are provided in the Supporting Information; detailed experimental conditions are provided in the Methods) and by FTSA (values taken from²⁴). (B) Intrinsic enthalpy (ΔH_{int} , kJ mol⁻¹) and Gibbs energy (ΔG_{int} , kJ mol⁻¹) changes calculated from the observed thermodynamic parameters using the equations defined in Figure 2. ND – not determined.

A Observed thermodynamic binding parameters										
Compound	ID	CA II			CA XII			CA XIII		
		ΔG_{ITC}	ΔG_{FTSA}	ΔH_{ITC}	ΔG_{ITC}	ΔG_{FTSA}	ΔH_{ITC}	ΔG_{ITC}	ΔG_{FTSA}	ΔH_{ITC}
4a	(JA17-1-1)	-40.6	-41.9	-25.3	-41.1	-39.1	-41.0	-44.5	-47.0	-74.9
4b	(JA17-1-2)	-37.5	-37.1	-10.3	-40.1	-38.9	-41.2	-44.0	-43.2	-67.2
4c	(JA17-1-3)	-36.5	-37.0	-15.9	-39.7	-38.3	-39.1	-45.1	-43.2	-75.1
4d	(JA17-1-4)	-36.7	-37.0	-19.7	ND	-37.7	ND	-48.0	-44.9	-54.4
4h	(JA19-17)	-32.3	-33.5	-22.8	-29.1	-30.8	-38.5	-38.2	-37.1	-57.3
7h	(JA19-19)	-35.9	-31.9	-14.2	ND	-28.6	ND	-36.4	-36.5	-61.3

B Intrinsic thermodynamic binding parameters										
Compound	ID	CA II			CA XII			CA XIII		
		ΔG_{ITC}	ΔG_{FTSA}	ΔH_{ITC}	ΔG_{ITC}	ΔG_{FTSA}	ΔH_{ITC}	ΔG_{ITC}	ΔG_{FTSA}	ΔH_{ITC}
4a	(JA17-1-1)	-59.3	-60.7	-26.8	-60.2	-58.2	-42.5	-61.4	-63.9	-64.0
4b	(JA17-1-2)	-56.2	-55.9	-11.7	-59.2	-57.9	-42.7	-60.9	-60.1	-56.2
4c	(JA17-1-3)	-55.2	-55.7	-17.4	-58.8	-57.4	-40.6	-62.0	-60.1	-64.2
4d	(JA17-1-4)	-55.4	-55.8	-21.1	ND	-56.8	ND	-64.9	-61.8	-43.5
4h	(JA19-17)	-51.0	-52.3	-24.3	-48.2	-49.9	-40.0	-55.1	-54.0	-46.4
7h	(JA19-19)	-49.3	-45.3	-11.1	ND	-42.4	ND	-48.0	-48.0	-45.8

graphic data, such thermodynamic differences are difficult to attribute to specific molecular interactions. Establishing a coherent structure-thermodynamics framework therefore requires datasets in which both observed and intrinsic binding parameters are directly correlated with experimentally determined binding modes.

There is a notable shortage of datasets of this kind^{25,26}. Only a limited number of systematic studies have examined a series of inhibitors across multiple, closely related protein environments, while simultaneously determining the observed and intrinsic thermodynamic parameters alongside corresponding X-ray structures. This hinders a detailed understanding of how subtle chemical modifications influence binding energetics and give rise to different thermodynamic profiles²⁴. It also hides the structural origins of enthalpy-entropy compensation in proton-linked binding processes.

In this work, we investigate a matched series of sulfonamide inhibitors interacting with

three closely related proteins. Using isothermal titration calorimetry and fluorescent thermal shift assay, we determine both the observed and intrinsic Gibbs energies, enthalpies, and entropic contributions for each ligand-protein pair. High-resolution X-ray crystal structures provide direct insight into the molecular interactions underlying these thermodynamic signatures. By integrating these complementary approaches, we delineate how subtle structural variations in both the ligand and the binding site give rise to specific energetic fingerprints.

Results

Determination of thermodynamic binding parameters

A series of *ortho*, *meta*-disubstituted benzene-sulfonamides (Figure 1) were previously designed and synthesized as potential binders of human carbonic anhydrase isozymes²⁴. Their

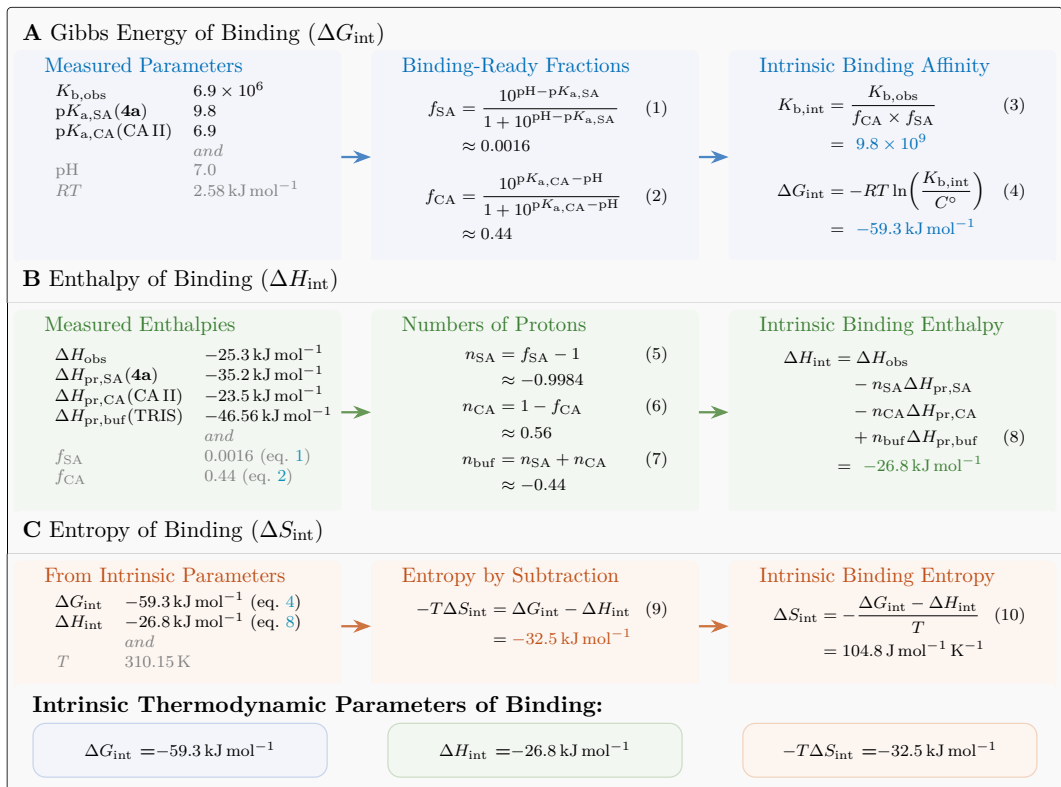


Figure 2: Representative example of the stepwise conversion of observed thermodynamic quantities into intrinsic Gibbs energy, enthalpy and entropy for the interaction between CA II and compound **4a**. Here, $C^\circ = 1 \text{ M}$ is the standard-state concentration, used to make the argument of the logarithm dimensionless.

interaction was determined by two independent techniques, isothermal titration calorimetry (ITC) and thermal shift assay (FTSA). ITC determines the enthalpy of binding (ΔH) by measuring the heat, and affinity (K_{b} , ΔG) from the shape of the titration curve, and the entropy (ΔS) is then calculated by subtraction. Due to the limitation of ITC in the picomolar K_{d} range, it is important to confirm the affinity by an independent technique and the FTSA is a straightforward technique to determine the affinity by heat-denaturation at various added compound concentrations. Experimental values obtained by both techniques of compounds binding to CA II, CA XII, and CA XIII are summarized in Table 1. Both ITC and FTSA

yielded ΔG values that matched each other within a few kJ mol^{-1} . This variation indicates that the measured binding affinities are reliable and agree within the experimental uncertainty.

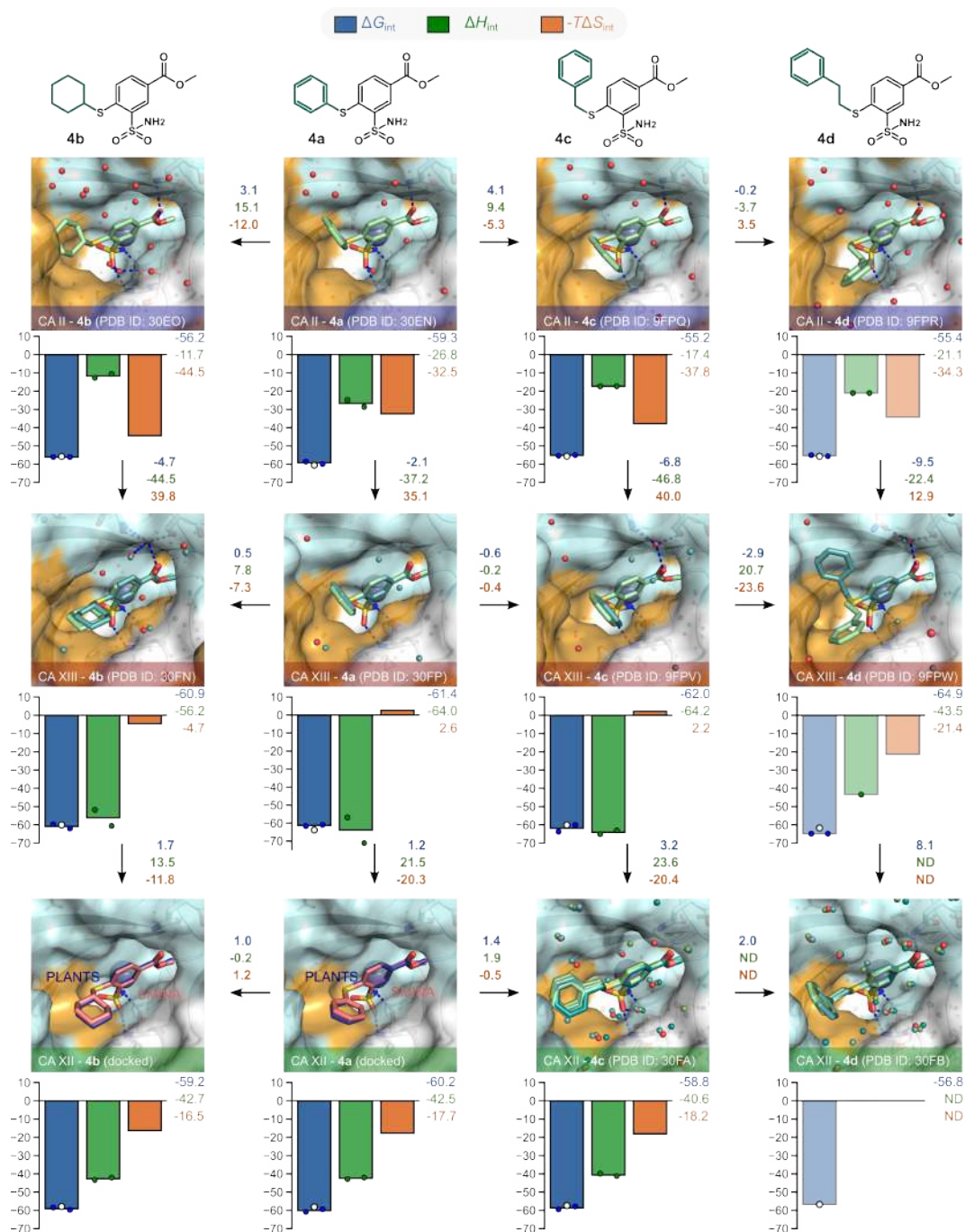


Figure 3: Intrinsic binding thermodynamic parameters for sulfonamide interactions with CA II (upper panels), CA XIII (middle panels), and CA XII (lower panels). Gibbs energy (ΔG_{int} , blue), enthalpy (ΔH_{int} , green), and entropic contribution ($-T\Delta S_{\text{int}}$, orange) are shown in kJ mol^{-1} . (caption continued on next page)

Figure 3: (caption continued from previous page) For each ligand–isozyme complex, intrinsic thermodynamic parameters derived from ITC are displayed adjacent to the corresponding X-ray crystal structure or docked. Bar graphs represent mean values, colored data points indicate individual ITC replicates, and the white data point associated with ΔG_{int} denotes values independently determined by FTSA. Numerical values for each thermodynamic parameter are shown next to the bars. Bars corresponding to compound **4d** are shown with reduced opacity to indicate that its slightly poorer solubility affected the quality of the ITC measurements. Arrows indicate differences in thermodynamic contributions, highlighting how subtle chemical modifications translate into changes in binding energetics across CA isozymes.

The ΔH s were determined only by ITC, since this is the only technique to directly measure the enthalpies. Their analysis is significantly complicated by the ligand-binding-linked protonation-deprotonation reactions. The sulfonamide group must deprotonate upon binding, while the Zn^{II} -bound hydroxide in active site of CAs must protonate to form water molecule that can be displaced by the deprotonated sulfonamide. Therefore, the ΔH s obtained from ITC are distinguished as *observed* values (ΔH_{obs}) and contain the contributions of buffer, ligand, and protein protonation. These contributions must be dissected to obtain the *intrinsic* ΔH_{int} . A worked example of this correction procedure is provided in Figure 2.

Intrinsic thermodynamic parameters for sulfonamide binding to carbonic anhydrases were obtained by correcting the observed FTSA- and ITC-derived quantities for all protonation equilibria coupled to complex formation, as described in detail previously^{21,27}. This approach separates the intrinsic binding energetics from contributions arising from ligand deprotonation, protonation of the Zn^{II} -bound hydroxide, and compensatory buffer ionization.

All corrections were performed at pH 7.0 and 310.15 K (37°C), corresponding to the experimental conditions. The fraction of inhibitor present in its deprotonated sulfonamide form (eq. 1) was calculated from its acidity constant ($\text{p}K_{\text{a,SA}}$), while the fraction of enzyme active sites containing the protonated Zn^{II} -bound water species ($\text{Zn}^{\text{II}}-\text{H}_2\text{O}$; eq. 2) was determined using the corresponding $\text{p}K_{\text{a,CA}}$ value (see Materials and Methods). These fractions were

used to correct the observed association constant ($K_{\text{b,obs}}$) to obtain the intrinsic association constant ($K_{\text{b,int}}$, eq. (3)), which reflects binding between RSO_2NH^- and $\text{CA}-\text{Zn}^{\text{II}}-\text{H}_2\text{O}$. The intrinsic Gibbs energy change was calculated from $K_{\text{b,int}}$ using standard thermodynamic relationships (eq. (4)).

Intrinsic binding enthalpies (eq. (8)) were obtained by subtracting the heats of all protonation-linked processes from the observed ITC enthalpy, including contributions from sulfonamide deprotonation, protonation of the Zn^{II} -bound hydroxide, and buffer protonation (see Materials and Methods). The corresponding stoichiometric coefficients describing proton uptake or release were derived from the calculated species fractions (eqs. (5) to (7)). The intrinsic entropy (eq. (9)) contribution was then obtained from the Gibbs energy-enthalpy relationship. ΔG_{int} and ΔH_{int} values are listed in Table 1.

Enthalpy-Entropy Compensation

The enthalpies and entropies of compounds binding are plotted in Figure 1. Both intrinsic enthalpy (ΔH) and the entropic contribution ($-T\Delta S$) favored binding, but some processes were mostly enthalpy-favored, while others were mostly entropy-favored resulting in relatively similar ΔG_{int} values. The ligand structure appeared to play a smaller role than differences among the protein isozymes. Ligand binding to CA II was dominated by entropy, CA XII was more in the middle, while CA XIII binding was enthalpy-dominated.

The large symbols in the plot show the intrinsic enthalpies and entropies, while the smaller semi-transparent symbols show the observed values in TRIS buffer at pH 7.0. It is important to correct for buffer because its enthalpy of protonation could completely distort the observed picture. However, it is worth depicting the observed values that do not involve correction for binding-linked protonation of ligand and protein. This is true for both ΔG and ΔH . Here we see, how different the picture is if we study observed versus intrinsic enthalpies and entropies (also see in SI).

All three CA isozymes are highly homologous with essentially the same overall fold. They differ only in several amino acid side chains located in the binding site. Structural arrangement of the ligand in the crystal structures (discussed in detail below) of bound complexes is highly similar, with only minor changes in the location of the benzenesulfonamide ring (Figure 1C), and somewhat more different positions of the second ring. Therefore, there is no obvious reason for such significant differences between the enthalpy-entropy contributions to the binding thermodynamics.

X-Ray Crystallography

In an attempt to gain the understanding of the structure-thermodynamics correlation of the protein-ligand binding system, many of the compound-isozyme pairs were crystallized and their structures determined to atomic resolution. Table 2 lists 7 crystal structures of compounds in complex with CA II, CA XII, and CA XIII. The structures were well resolved and provided detailed structural information about protein-ligand interactions. In all structures, both the ones that have been described previously²⁴ for this group of compounds and here (all PDB IDs are listed in Figure 3 and 6), the amino group of sulfonamide formed a coordination bond with the Zn^{II} in the active site of the protein. Residue numbering follows the CA II sequence throughout.

Figure 4 shows the binding modes and closest interactions of compounds in the CA isozyme active sites. The compounds are primar-

ily oriented by a strong coordination bond with the catalytic Zn^{II} . The sulfonamide group is oriented identically by forming hydrogen bonds with Thr199 in all three studied CA isozymes, a binding mode that is characteristic of zinc-bound sulfonamides across the human CAs. The remainder of the molecule is oriented in a way to maximize the hydrophobic effect and hydrogen bonds. Only the ring that is furthest from the sulfonamide group occupied variable positions depending on exact compound structure and CA isozyme.

The binding sites in three CA isozymes are highly similar. One difference is in CA XIII which contains Ser62 instead of Asn62 in CA II and CA XII. Ser cannot fully reach the compound, it makes a water-molecule-mediated hydrogen bond with the compound, while the Asn amino group forms a direct hydrogen bond.

The most notable difference among the hydrophobic amino acids, is the replacement of Phe131 present in CA II and CA XIII with Ala131 in CA XII. This Phe131 is observed in crystal structures in several orientations. Some inhibitors appear to induce an alternative Phe131 conformation.

Two structural features warranted separate discussion due to their relevance for inhibitor binding and modeling: the conformational variability of Phe131 in CA II, and the stereochemical ambiguity of the sulfoxide in compound **7h**.

Conformational Variability of Phe131 in CA II

Figure 5 compares the crystal structure determined in this study of CA II - compound **4b** (PDB ID: 30EO) with four previously reported CA II structures in which Phe131 also adopts a non-canonical orientation. Analysis of Phe131 conformations across CA II structures shows that, although the vast majority of more than a thousand PDB entries adopt the canonical *t* (trans) rotamer (shown in gray), some structures display alternative orientations (shown in red shades). These non-canonical conformations arise when bound ligands sterically displace Phe131. In the 1CA2 structure, Phe131 adopts a χ_1 angle of 175.4°. Rotamer

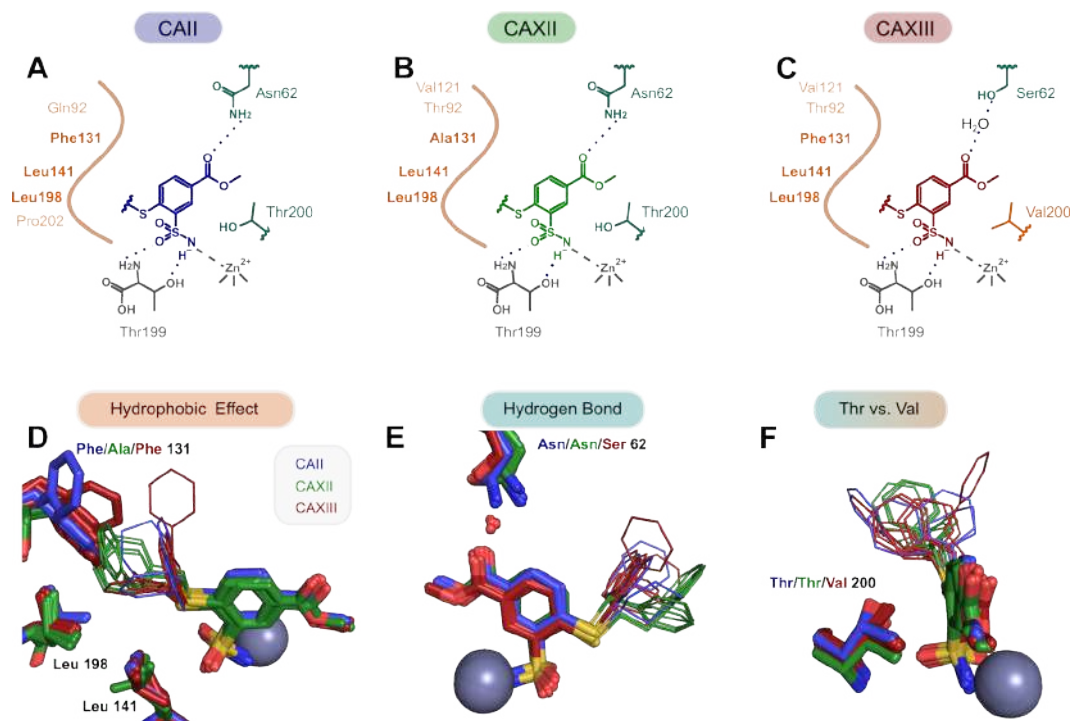


Figure 4: Binding modes of investigated sulfonamide inhibitors to CA II (A), CA XII (B), and CA XIII (C) as determined by X-ray crystallography. Panels A–C illustrate the principal interaction patterns in each isozyme, including coordination of the sulfonamide nitrogen to the catalytic Zn^{II} ion (dashed lines), hydrogen-bonding interactions (dotted lines), and the distribution of hydrophobic (orange) and hydrophilic (electric blue) residues surrounding the ligand. Panels D–F highlight interaction features that are conserved across the three isozymes: (D) hydrophobic contacts between the aromatic moiety of the sulfonamides and the hydrophobic wall of the active site; (E) residues involved in direct or water-mediated hydrogen bonding; and (F) comparison of residue 200 adjacent to the benzenesulfonamide, which is Thr in CA II and CA XII but Val in CA XIII. CA II is shown in blue (PDB IDs: 30EN, 30EO, 9FPQ, 9FPR); CA XII in green (PDB IDs: 30FA, 30FB); and CA XIII in red (PDB IDs: 30FP, 30FN, 9FPV, 9FPW). Zn^{II} is shown as a gray sphere. CA II residue numbering is used throughout, and all chains present in the asymmetric unit were included in the structural comparison.

analysis performed in ChimeraX²⁸ places this angle in the *t* rotamer bin centered at 176.9°, which has a relative population of 0.871 in the Dunbrack backbone-dependent rotamer library²⁹. In contrast, the *g*⁺ (gauche-plus) conformation observed in structures such as 2AW1 (bin centered at 73.9°, actual angle 84.8°) has a population of only 0.000199. Thus, the canonical *t* rotamer is approximately 4400-fold

more populated than the *g*⁺ rotamer, corresponding to an energy difference of $\Delta G = -RT \ln K \approx 21 \text{ kJ mol}^{-1}$. Although rare, these ligand-induced Phe131 rearrangements can significantly impact structure-based modeling, as they reshape the active-site pocket, alter hydrophobic and electrostatic contacts, and may bias docking, virtual screening, or energy calculations if not properly accounted for.

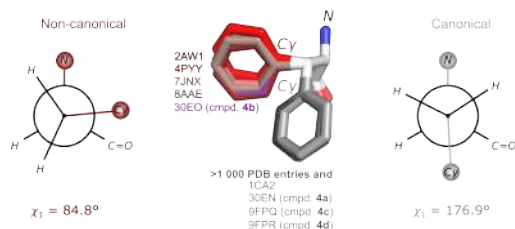


Figure 5: Canonical and non-canonical Phe131 conformations observed in human CAII. The canonical rotamer of Phe131 (represented by PDB ID [1CA2](#)) is adopted by the vast majority of CAII structures, including all ligand-protein complexes determined in this study. In contrast, a rotated, non-canonical χ_1 conformation is rare and has been previously observed only in a few PDB entries, including [2AW1](#), [4PYY](#), [7JNX](#), and [8AAE](#). The structure [3OEO](#) reported here provides an additional example of this non-canonical state. Such ligand-induced deviations from the canonical rotamer may reflect energetically costly conformational adjustments required to accommodate steric or interaction-favored constraints imposed by the bound inhibitor.

The available structures do not establish why particular ligands induce this alternative Phe131 conformation or whether it contributes measurably to their binding thermodynamics. Rather, the observation demonstrates that the CA active site can accommodate ligands through local protein rearrangements, emphasizing the importance of considering receptor flexibility in structure-based modeling.

Sulfoxide Stereochemistry in Compound **7h**

The synthesis of compound **7h** (methyl 4-methylsulfinyl-3-sulfamoyl-benzoate)²⁴ is not stereoselective, and both enantiomers are therefore expected to be present. Sulfoxides of this type typically exhibit high configurational stability³⁰, with inversion barriers on the order of approximately 150 kJ mol⁻¹ to 200 kJ mol⁻¹^{31,32}, indicating that racemization

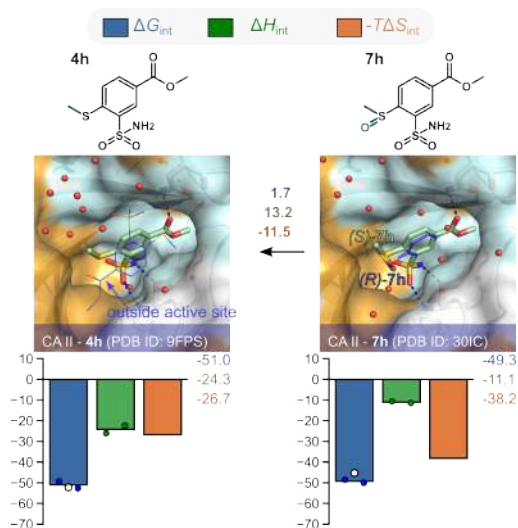


Figure 6: Correlation between intrinsic binding thermodynamic parameters and binding modes of sulfonamide ligands in CAII. (ΔG_{int} - blue; ΔH_{int} - green; $-T\Delta S_{\text{int}}$ - orange; all in kJ mol⁻¹ alongside the corresponding X-ray crystal structures. Left panels: Structure of ligand **4h** (PDB ID: [9FPS](#)) reveals two ligand molecules, one bound in the active site coordinating the Zn^{II}, and a second molecule located outside the catalytic pocket. The latter position likely reflects a crystallization-related arrangement rather than a functionally relevant binding mode and has been discussed previously²⁴. Right panels: Structure of ligand **7h** (PDB ID: [3OIC](#)) shows two possible enantiomeric binding modes within the active site. The associated bar graph summarizes intrinsic thermodynamic parameters, with individual ITC replicates shown as colored points and the independently determined ΔG_{int} (FTSA) indicated in white. Arrow indicates differences in thermodynamic contributions.

is unlikely under the reaction and solution conditions employed.

In the X-ray structure of **7h** bound to CAII, the electron density around the S-O/S-C region does not allow unambiguous assignment of the absolute configuration: the substituent is small, and the scattering contrast between oxygen and carbon is insufficient for

reliable discrimination (Figure 9). Refinement of the ligand as either enantiomer yields comparable global statistics, although the (*R*) model converges to atypical S–O and S–C distances, whereas the (*S*) model refines to values consistent with standard sulfoxide geometry (S–O $\approx$ 1.5 Å, S–C $\approx$ 1.8 Å³³). However, in the absence of a definitive stereochemical indicator, the deposited structure (PDB ID: 30IC) includes both enantiomers (Figure 6). Consistent with this interpretation, ITC measurements of the racemate yielded a binding stoichiometry close to unity ($N \approx 1$), suggesting that both enantiomers participate in binding and that no major inactive fraction is present in solution.

Molecular Docking into the CA XII Active Site

To explore the interactions of investigated ligands within the binding site of CA XII, compounds **4a–d** were docked into a CA XII receptor from its complex with **4c** (PDB ID 30FA) using Smina³⁴ and PLANTS³⁵ docking software. Positional restraint on the sulfonamide nitrogen during docking was used in both programs. Smina docked poses obtained via constrained docking protocol were afterwards rescored using Vinardo scoring function³⁶ in order to remove the constraining potential (See Methods). No such rescoring was necessary for PLANTS. The docking pose accuracy was validated by comparing the docked poses of **4c** and **4d** with the experimental X-ray structures (30FA and 30FB, respectively). The best scoring poses for docked **4c** and **4d** had relatively low Root Mean Square Deviations (RMSDs) with respect to the X-ray structures: **4c** exhibited 0.693 Å and 0.698 Å RMSDs for PLANTS and Smina docking, respectively, while the corresponding RMSDs for **4d** were 0.993 Å and 1.050 Å (Figure S5 in SI). An RMSD value below 2 Å is often considered as a threshold for the correct prediction of the ligand native pose³⁷. Therefore it is reasonable to expect that docked poses of compounds **4a–b** should also have a reasonable accuracy. The best ranked poses of **4a** and **4b** are shown in Figure 3. The

RMSD between the best Smina and PLANTS docked poses for **4a** were 0.207 Å, and for **4b** 0.167 Å, which signifies that these conformations could be close to the ones observed in Nature. The spatial positioning of the phenyl and cyclohexyl rings (in **4a** and **4b**, correspondingly) are, therefore, very close in dockings of both Smina and PLANTS programs. These hydrophobic rings rest in a subpocket in CA XII lined with the sidechains of residues Ala131, Ser135, Leu141, Leu198, and Pro202.

Table S4 in SI lists the Voronoi interface contact areas that **4c** and **4d** ligands show in their X-ray structures, and for **4a** and **4b**, because of the lack of the experimental structures, in PLANTS docked best scoring poses. PLANTS docked poses were chosen over Smina’s because of the slightly better docked pose accuracy. From Table S4 it can be observed how interaction with Ala131 sidechain becomes more prominent in larger hydrophobic groups.

In the X-ray structure of CA XII–**4d**, the ligand’s carbonyl group forms a hydrogen bond ($\approx$ 3.1 Å) with Asn62 (CA II numbering), whereas in CA XII–**4c** this interaction is not observed (4.1 Å). In contrast, CA II complexes of **4a–d** consistently display this hydrogen bond, while in CA XIII the analogous interaction occurs via a bridging water molecule with Ser62 (except for **4a**, where no crystal water was detected at this position). The CA XII–**4c** structure used for docking was obtained by ligand soaking, and the loop containing Asn62 is likely flexible; therefore, the absence of the hydrogen bond in the docked pose ($\approx$ 3.8 Å) does not necessarily preclude its formation in solution. Considering the intrinsic mobility of this loop, such an interaction could plausibly occur under physiological conditions.

Correlation of Structure with Binding Thermodynamics

In an attempt to understand the structural reasons that determine compound binding enthalpic versus entropic dominating origins, the X-ray crystal structures are put next to the thermodynamic binding profiles (Figure 3). From top to bottom, the structures and ther-

modynamics of binding to CA II, CA XIII, and CA XII is shown. Each column represents one of the four chemical compound binding, showing the chemical structure on the top, position in the crystal structure in the middle, and the binding thermodynamics on the bottom. Gibbs energies of binding (ΔG_{int} , blue) are split into enthalpic (ΔH_{int} , green) and entropic ($T\Delta S_{\text{int}}$, orange) contributions.

The standard intrinsic thermodynamic parameter numbers are listed on the right of the bar chart. In addition, arrows connect adjacent conditions and the numbers next to arrows show the differences in thermodynamic parameters between the two conditions. For example, going from compound **4a** (second column) binding to CA II to compound **4b** (first column), the difference in Gibbs energy ($\Delta\Delta G_{\text{int}}$) was 3.1 kJ mol^{-1} . Thus the compound **4a** bound stronger than compound **4b**. This means that the cyclohexyl moiety makes slightly more favorable interactions than the benzene ring of the compound.

Similarly, the second number (green) shows the difference in enthalpies, equal to 15.1 kJ mol^{-1} . This means that **4b** bound less exothermically by 15.1 kJ mol^{-1} than **4a**. The large loss in enthalpy ($\Delta\Delta H_{\text{int}}$) was compensated by the large gain in entropy ($-\Delta(T\Delta S_{\text{int}})$) equal to -12.0 kJ mol^{-1} .

The most important observation here is that despite highly similar compound chemical structures and highly similar binding affinities (ΔG_{int}), the enthalpy-entropy contributions are highly variable. In some cases, the binding process is nearly entirely favored by enthalpy, while in other cases it is favored largely by entropy. Second, the differences between compounds were smaller than the differences between CA II and CA XIII. Binding to CA II of all four compounds was dominated by favorable entropy, with a smaller favorable enthalpy contribution. However, binding to CA XIII was strongly dominated by enthalpy with a near-zero entropy contribution for three compounds, except **4d**, where favorable entropic contribution was appreciable.

As we see from Figure 4, there are only several structural differences between CA II

and CA XIII that could be responsible for the large differences in binding thermodynamics. First, the Val200 present in CA XIII instead of Thr200 in CA II. Second, the presence of Ser62 in CA XIII instead of Asn62 in CA II, causing a shorter side chain needing the water molecule presence to connect with the ligand. We propose that these residue-level differences offer a mechanistic explanation for the distinct enthalpy-entropy compensation observed for compound binding to CA XIII versus CA II.

Materials and Methods

Proteins and Inhibitors

Synthesized Sulfonamide Compounds

Sulfonamide ligands **4a**, **4b**, **4c**, **4d**, **4h**, and **7h** were synthesized according to previously reported procedures²⁴. The compound numbering used here follows the nomenclature of the original publication. Stock solutions were prepared in DMSO; the exact concentrations used for ITC measurements and for ligand soaking in crystallographic experiments are specified in the corresponding sections below.

Human Carbonic Anhydrases Preparation

Recombinant human carbonic anhydrases CA II, CA XII, and CA XIII were expressed and purified following established protocols³⁸ and stored at -80°C . Prior to ITC experiments, proteins were dialyzed and buffer-exchanged into a buffer containing 50 mM TRIS, pH 7.0 and 100 mM NaCl and performed at 37°C . Protein concentrations were determined by UV absorbance at 280 nm using extinction coefficients calculated from the amino acid sequence.

Protein-Ligand Binding Thermodynamics

Affinity of the sulfonamide ligands for the carbonic anhydrases was previously quantified using a fluorescent thermal shift assay²⁴, providing an initial assessment of binding strength.

Table 2: X-ray diffraction data collection and refinement statistics.

PDB id							
Isozyme	CA II	CA II	CA II	CA XII	CA XII	CA XIII	CA XIII
Compound	4a	4b	7h	4c	4d	4a	4b
PDB ID	30EN	30EO	30IC	30FA	30FB	30FP	30FN
Data-collection statistics							
Space group	P12 ₁ 1	P12 ₁ 1	P12 ₁ 1	P12 ₁ 1	P12 ₁ 1	P212121	P212121
Unit-cell parameters:							
a (Å)	42.27	42.16	42.21	77.22	77.34	54.99	54.94
b (Å)	41.30	41.09	41.06	74.07	74.22	58.48	57.72
c (Å)	71.90	71.70	71.91	91.65	91.69	161.40	159.46
α (°)	90.00	90.00	90.00	90.00	90.00	90.00	90.00
β (°)	104.40	104.35	104.31	108.68	108.71	90.00	90.00
γ (°)	90.00	90.00	90	90.00	90.00	90.00	90.00
Resolution range (Å)	1.42-40.94	1.59-40.84	1.41-35.40	1.13-86.82	1.23-73.25	2.25-54.98	1.38-51.95
Unique reflections number, overall (outer shell)	43783 (6249)	31958 (4610)	45706 (6601)	354567 (43345)	273829 (32817)	25479 (3638)	105144 (15161)
R_{merge} , overall (outer shell)	0.040 (0.376)	0.060 (0.401)	0.034 (0.393)	0.089 (0.294)	0.079 (0.393)	0.139 (0.414)	0.066 (0.312)
I/σ , overall (outer shell)	23.0 (5.1)	15.5 (4.6)	25.4 (4.6)	12.0 (3.5)	10.8 (2.6)	12.3 (6.4)	19.7 (7.3)
Multiplicity, overall (outer shell)	7.1 (7.2)	6.9 (6.8)	6.9 (6.8)	6.4 (4.5)	6.5 (4.5)	11.3 (12.0)	13.1 (12.5)
Completeness (%), overall (outer shell)	96.4 (95.0)	99.3 (99.2)	98.9 (98.9)	97.3 (81.7)	97.1 (80.0)	99.9 (99.9)	100 (100)
Wilson B-factor	15.8	18.9	17.1	11.7	12.7	21.5	12.7
Refinement statistics							
R_{work}	0.1480	0.1676	0.1691	0.1569	0.1446	0.1753	0.1218
R_{free}	0.2056	0.2141	0.2029	0.1863	0.1798	0.2349	0.1549
RMSD bond lengths, (Å)	0.0129	0.0099	0.0125	0.0164	0.0156	0.0073	0.0133
RMSD bond angles (°)	1.9851	1.8928	2.0895	2.0592	2.0399	1.6600	1.9764
Average B factors (Å ²)							
all	19.3	20.3	19.9	19.8	20.7	21.0	18.3
main-chain	15.8	17.3	16.8	16.4	17.5	19.0	14.6
side-chain	19.2	20.1	19.5	20.4	21.6	22.6	18.9
inhibitors	15.9	16.6	15.3 ^a , 17.7 ^b	17.2	17.4	20.0	18.2
waters	31.4	31.2	30.6	29.1	30.0	25.1	30.8
zinc	11.6	12.1	11.7	9.1	10.1	9.3	9.7
other molecules	42.3	34.8	36.1	25.1	24.6	none	29.5
Number of atoms							
protein	2085	2078	2119	8583	8539	4111	4201
inhibitor	21	21	17 ^a /17 ^b	88	92	42	42
water	311	337	340	1099	1105	260	547
zinc	1	1	1	4	4	2	2
other molecules	16	4	26	42	28	0	21
Ramachandran statistics (%)							
most favored regions	97	96	97	98	98	96	97
additionally allowed regions	3	4	3	2	2	4	3

Enantiomers of compound **7h** (methyl 4-methylsulfinyl-3-sulfamoyl-benzoate):

^a (*S*)-**7h**

^b (*R*)-**7h**

In the present study, these measurements were complemented with isothermal titration calorimetry, which enabled a more detailed thermodynamic characterization of the interactions. Exemplary ITC and FTSA datasets for ligand **4a** binding to CA II are shown in Figure 7.

To obtain intrinsic, pH-independent binding parameters,^{21,27} we additionally used the sulfonamide pK_a values from²⁴ and determined the enthalpies of protonation (Figure 8). These experimentally measured quantities allowed correction of the observed thermodynamic data for all protonation equilibria coupled to complex formation, thereby isolating

the energetics of the binding event itself.

Isothermal Titration Calorimetry

ITC experiments were carried out using a MicroCal PEAQ-ITC calorimeter (Malvern Panalytical Ltd). The protein solution in the cell contained 20 μ M CA, while the compound concentration in the syringe was 200 μ M. Both cell and syringe solutions were prepared in buffer: 50 mM TRIS at pH 7.0 containing 100 mM sodium chloride and 2.0 % (v/v) of DMSO. Before titration, all samples were centrifuged to minimize baseline noise and bubble formation.

A typical experiment consisted of 19 injec-

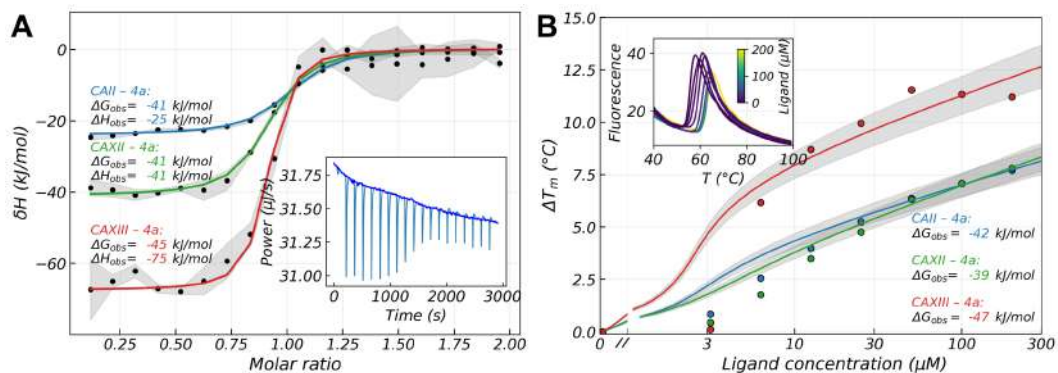


Figure 7: Thermodynamic characterization of ligand **4a** binding to carbonic anhydrase isoforms by ITC and FTSA. (A) ITC binding isotherms for CA II, CA XII, and CA XIII titrated with ligand **4a**, together with the corresponding binding affinities and enthalpy changes. The inset shows a thermogram for the CA II-**4a** titration. (B) FTSA data showing the melting temperature as a function of ligand concentration for CA II, CA XII, and CA XIII. The inset displays raw fluorescence traces for the CA II-**4a** interaction. FTSA data were taken from²⁴.

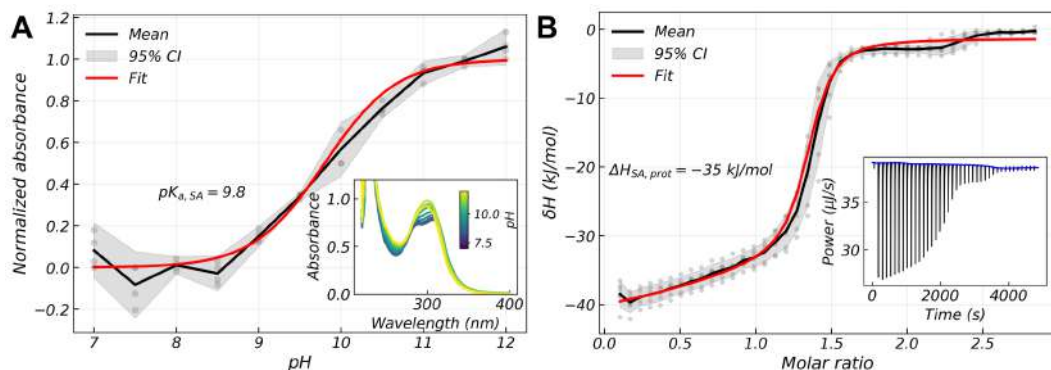


Figure 8: Determination of sulfonamide protonation parameters. (A) pH-dependent UV-visible absorbance profiles of sulfonamides **4a**, **4b**, **4c** **4d**, with the fitted Henderson-Hasselbalch model used to determine the sulfonamide $pK_{a,SA}$ value ($pK_{a,SA} = 9.8$). The inset shows representative absorbance spectra of compound **4a** recorded across the pH range. (B) Isothermal titration calorimetry of compounds **4a**, **4b**, **4c**, **4d** dissolved in 1.5 equivalent NaOH, titrated with nitric acid to determine the enthalpy of protonation of the sulfonamide amino group ($\Delta H_{pr,SA} = -35.2 \text{ kJ mol}^{-1}$ at 37°C). The inset shows a representative thermogram of compound **4a**.

tions with 150 s spacing between injections; the first injection was 0.4 μL (duration 0.8 s) and was excluded from analysis, followed by 2.0 μL injections (duration 4.0 s). All experiments were performed at 37°C with a reference power of $4.0 \mu\text{cal s}^{-1}$.

Thermograms were integrated using NITPIC

2.0.7³⁹, and binding isotherms were fitted in SEDPHAT 15-2b⁴⁰ using a one-site binding model. Final figures were prepared in GUSI 1.4.2⁴¹ (see Supporting Information).

Determination of pK_a

The pK_a values of the Zn^{II} -bound water molecule in the enzyme active site of CAs ($pK_{a,CA}$) were taken from²⁷: 6.9 for CA II, 6.8 for CA XII, and 8.0 for CA XIII at 37 °C.

Notably, the pK_a of the Zn^{II} -bound water differs substantially between CA II/CA XII (6.9/6.8) and CA XIII (8.0). These differences affect the observed thermodynamic parameters but are accounted for in the calculation of intrinsic parameters and therefore do not contribute to the isozyme-dependent differences discussed here.

The pK_a values of the sulfonamide inhibitors ($pK_{a,SA}$) were determined previously²⁴. The measurement reliability was within 0.2 pK_a units; therefore, the median value was used to represent compounds sharing the same scaffold. Compounds **4a**, **4b**, **4c**, **4d**, and **4h** were assigned a pK_a value of 9.8 (Figure 8A), whereas compound **7h** exhibited a lower pK_a of 8.9, consistent with the stronger electron-withdrawing character of its substituent.

Determination of Protonation Enthalpy

The enthalpy of protonation of the Zn^{II} -OH⁻ species in the active sites of carbonic anhydrases, $\Delta H_{pr,CA}$, was taken from²⁷. Reported values at 37 °C are: CA II, -23.5 kJ/mol; CA XII, -25.5 kJ mol⁻¹; and CA XIII, -41.5 kJ mol⁻¹.

The enthalpy of protonation of the sulfonamide amino group, $\Delta H_{pr,SA}$, was determined experimentally by ITC. Solutions containing 0.25 mM sulfonamide and 0.375 mM sodium hydroxide were titrated with 2.5 mM nitric acid using a MicroCal PEAQ-ITC calorimeter. The DMSO content in both syringe and cell solutions was 2% (v/v). Titrations consisted of 40 injections of 0.9 μ L each, with 120 s spacing between injections, at 37 °C. Representative thermogram and integrated heats are shown in Figure 8B.

The reaction enthalpy for the sulfonamides **4a**, **4b**, **4c**, **4d**, and **4h** was experimentally determined or adjusted based on structural similarity as $\Delta H_{pr,SA} = -35.2$ kJ mol⁻¹ and for **7h** as $\Delta H_{pr,SA} = -30.4$ kJ mol⁻¹ (37 °C). The used

enthalpy of TRIS buffer protonation at 37 °C was -46.56 kJ mol⁻¹⁴².

X-Ray Crystallography

Crystallization

Crystals of CA II, CA XII, and CA XIII (Figure 9) were obtained at room temperature using the sitting-drop vapor diffusion method. Protein samples were concentrated by ultrafiltration to approximately 20 mg mL⁻¹ to 30 mg mL⁻¹. Crystallization conditions were as follows:

CA II – 0.1 M sodium bicine (pH 9.0), 0.1 M sodium malonate (pH 7.0), and 0.2 M ammonium sulfate.

CA XII – 0.1 M ammonium citrate (pH 7.0), 0.2 M ammonium sulfate, and 26% (w/v) PEG 4000.

CA XIII – 0.1 M sodium citrate (pH 5.5), 0.1 M sodium acetate (pH 4.5), and 26% (w/v) PEG 4000.

Ligand soaking

Ligand-protein complexes of CA II, CA XII and CA XIII were prepared by soaking. Stock solutions of each ligand were prepared at 50 mM in 100% DMSO. One microlitre of the ligand stock was diluted into 50 μ L of the corresponding reservoir solution used for crystallization. Crystals were transferred into the soaking solution and incubated for up to one week before data collection.

Data Collection and Structure Determination

For data collection, CA XII and CA XIII crystals were transferred into a cryoprotectant consisting of the reservoir solution supplemented with 15% to 20% ethylene glycol. After several minutes of incubation, crystals were flash-cooled in liquid nitrogen. CA II crystals were flash-cooled directly from the mother liquor.

Diffraction data were collected at the PETRA III synchrotron (DESY, Hamburg) on EMBL beamlines P13 (MX1) and P14 (MX2). Data

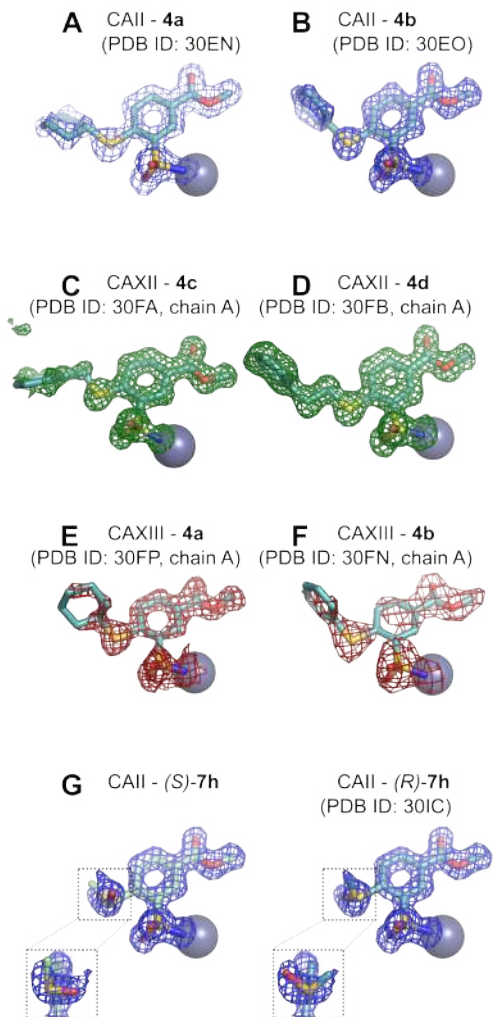


Figure 9: $|F_o - F_c|$ omit maps contoured at 3σ for the investigated ligands bound in the active sites of carbonic anhydrases. (A) CA II-**4a**, (B) CA II-**4b**, (C) CA XII-**4c**, (D) CA XII-**4d**, (E) CA XIII-**4a**, (F) CA XIII-**4b**, (G) CA II-(*S*)-**7h** and CA II-(*R*)-**7h** (The two enantiomers of **7h** are shown separately for visualization purposes; however, both configurations were modeled into the same binding position). Omit maps were generated from refinement runs in which the ligand was removed from the final model. Zinc ions are shown as gray spheres.

were processed and scaled using XDS⁴³, MOS-

FLM⁴⁴, and SCALA⁴⁵. Structures were solved by molecular replacement with MOLREP⁴⁶. Refinement was performed with REFMAC⁴⁷, and manual model building was carried out in COOT⁴⁸. Ligand coordinates were generated in AVOGADRO⁴⁹, and ligand restraint files were prepared using LIBCHECK⁵⁰.

Final refinement statistics and PDB accession codes are provided in Table 2.

Molecular Docking

Docking was performed using Smina³⁴ and PLANTS³⁵ software. The X-ray structure of **4c** in the complex with CA XII (PDB ID 30FA) was used as the receptor for the docking. The receptor structure was pre-processed using ChimeraX version 1.10⁵¹ software and converted to the pdbqt format using a command-line tool prepare_receptor4.py, a part of MGLTools version 1.5.7⁵². Most other conversions between various molecular formats were done using OpenBabel version 3.1.1⁵³. The ligand geometries were optimized using MMFF94s force field⁵⁴ and saved in pdbqt format using Avogadro version 1.2.0 software⁴⁹. The Smina docking protocol was similar to the one used previously²⁴. In brief, the sulfonamide nitrogen of the docked ligand was attracted to an artificial quadratic potential centered on a dummy atom with weight equal to 10 and cut-off radius 2.2 Å. The center of the docking box of the cubic shape with a side equal to 24 Å was located in the center of the **4c** ligand in its complex with the receptor (in other words, with the coordinates $x = x_{\min} + \frac{x_{\max} - x_{\min}}{2}$, $y = y_{\min} + \frac{y_{\max} - y_{\min}}{2}$, $z = z_{\min} + \frac{z_{\max} - z_{\min}}{2}$, where $q_{\max}$ and $q_{\min}$ denote the maximum and minimum values of the coordinate q over all atoms of ligand **4c**. Docking exhaustiveness was set to 100, the hit energy range to 10 kcal mol⁻¹, and the random seed to 0 (for reproducibility). The hits generated by a constrained Smina run were afterwards rescored using Vinardo scoring function³⁶.

Additional docking was performed using PLANTS version 1.2³⁵ docking software. The receptor and ligand mol2 files were generated using SPORES version 1.3 program⁵⁵ from the

best scoring pose for the Smina docked pose bound to the receptor. SPORES deprotonated sulfonamide nitrogen based on its proximity to zinc atom in the receptor. The sulfonamide atom was kept fixed in its X-ray position during docking via the "fixed scaffold" option in PLANTS. The docking box center was set to the center of the **4c** ligand, similarly to Smina docking, and the binding site radius was set to 24 Å. The default ChemPLP scoring function³⁵ was used for scoring.

The Voronoi ligand-receptor interface areas and nearest distances were calculated using Voronota software⁵⁶ expansion Voronota-JS (v. 1.26.3717) via the voronota-js-fast-iface-contacts command line tool.

Discussion

Understanding of the structural determinants of the protein–ligand binding thermodynamics remains limited. Exact contributions of enthalpy and entropy to the standard Gibbs energy of binding cannot yet be accurately predicted computationally, and the physical origins of the enthalpy–entropy compensation effect remain only partially understood¹. Consequently, experimentally tractable model systems remain essential for dissecting the molecular determinants governing thermodynamic partitioning in biomolecular recognition.

The CA II isozyme has served as a model protein in molecular biophysics for decades²². The human carbonic anhydrase family binding by primary sulfonamide ligands represents a particularly convenient and informative model system. The convenience of this system arises from the relatively strong affinity, which falls within the optimal range for accurate ITC measurements. Consequently, reliable Gibbs energies and enthalpies of binding can be obtained experimentally. Furthermore, the sulfonamide–Zn^{II} coordination bond ensures a well-defined 1:1 binding stoichiometry, simplifying thermodynamic analysis.

Twelve catalytically active human CA isozymes contain the active-site Zn^{II} ion and bind sulfonamide inhibitors. Although the

overall protein folds and active-site architectures of these isozymes are highly conserved, several amino acids differ within the ligand-binding cavity. This provides a unique opportunity to investigate how subtle changes in the local protein environment influence ligand recognition and the balance between enthalpic and entropic contributions to binding.

The thermodynamics of sulfonamide binding to carbonic anhydrases are additionally complicated by two coupled protonation/deprotonation reactions. First, the sulfonamide amino group typically possesses a pK_a in the range of 9–11 and therefore exists predominantly in the protonated form at neutral pH. However, upon binding, the sulfonamide must deprotonate because only the negatively charged form coordinates the Zn^{II} ion. Second, the CA active site contains a Zn^{II}-bound water molecule with a pK_a near 7. During catalysis, this species functions as a hydroxide ion, whereas during inhibitor binding the hydroxide must first become protonated before displacement by the negatively charged sulfonamide ligand can occur. Consequently, both binding-linked protonation reactions must be dissected and corrected to obtain the intrinsic thermodynamic parameters of ligand binding^{21,27}.

The principal finding of this study is the observation that highly similar protein–ligand systems can exhibit nearly diametrically opposing thermodynamic signatures, ranging from predominantly enthalpy-favored to predominantly entropy-favored binding, while maintaining relatively similar overall Gibbs energies of binding. Such behavior represents a characteristic manifestation of the Enthalpy–Entropy Compensation in protein–ligand interactions^{1,57}. Importantly, this trend was consistently observed across a series of closely related compounds.

Structurally, the binding orientation and position of the ligands within the active sites of the studied CA isozymes were nearly identical, with only minor local differences such as single amino-acid substitutions or subtle variations in active-site hydration. No major structural rearrangement was observed that could directly explain the transition between enthalpy- and entropy-dominated binding, sug-

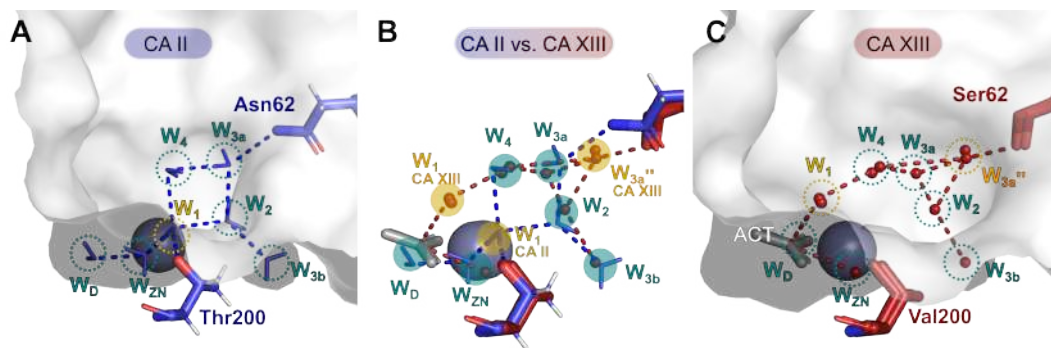


Figure 10: (A) Crystal structure of CA II (PDB ID: [3KKX](#)) highlighting the ordered active-site water network. The Zn^{II} -bound (W_{ZN}) and the so-called deep water (W_D) are clearly resolved, together with the proton-transfer water chain (W_1, W_2, W_{3a}, W_{3b}). This hydrogen-bonded network connects the Zn-bound water to the proton shuttle residue His64 and is essential for catalytic proton transfer. W_1 is stabilized by Thr199/Thr200, while W_2 forms a branching point toward W_{3a} and W_{3b} , which are further coordinated by residues such as Tyr7 and Asn62, respectively, forming a conserved proton wire in CA II. (B) Superposition of CA II (blue) and CA XIII (red) active sites. Conserved water positions are shown in blue, indicating a largely preserved proton-transfer network. Differences between the isoforms are highlighted in yellow, including variations of W_1 (adjacent to Thr200 in CA II and Val200 in CA XIII; CA II numbering) and the presence of an additional water position (denoted $W_{3a''}$) in CA XIII. These differences are attributed to residue substitutions (Asn62 in CA II vs Ser62 in CA XIII), which alter local hydrogen-bonding patterns and solvent organization. (C) Crystal structure of CA XIII (PDB ID: [3D0N](#)) in complex with acetate. The acetate ligand occupies the deep water position and perturbs the Zn^{II} -bound water position. The surrounding water network is reorganized relative to CA II, particularly near positions W_1 and W_{3a} , reflecting altered hydrogen-bonding interactions due to residue substitutions. Together, these structures illustrate both the conserved nature of the CA proton-transfer water network and isoform-specific differences in solvent organization that may influence catalysis and ligand binding. Because no ligand-free (apo) structure of CA XIII has been determined, the native water network of CA XIII remains inferred from ligand-bound structures and comparative analysis rather than directly observed.

gesting that subtle differences in hydration and solvent organization may dominate the thermodynamic signature. The observed thermodynamic differences were substantially larger between CA isozymes than between the closely related chemical compounds themselves, suggesting that the primary determinant originates from differences in the protein environment rather than ligand structure. Approximate averaging of the thermodynamic contributions across the studied compounds suggests the following relative partitioning of the Gibbs energy:

Isozyme	Enthalpy	Entropy
CA II	30 %	70 %
CA XII	70 %	30 %
CA XIII	100 %	0 %

Attempts to rationalize these differences by analyzing Figure 4 suggest that several structural features are unlikely to represent the dominant determinants. First, the effect cannot primarily arise from the rotation of Phe131 (CA II numbering) because both CA II and CA XIII contain this residue despite exhibiting nearly opposite thermodynamic profiles, whereas CA XII, lacking Phe131, displays an intermediate thermodynamic behavior. Con-

sistent with the X-ray crystal structures, the bulky Phe131 present in CA II and CA XIII is associated with greater variability in the conformations of the ligand substituents (Figure 4), which may contribute to the larger thermodynamic scatter through differences in hydrophobic contacts and solvent reorganization. In CA XII, the smaller Ala at this position allows more similar ligand conformations, potentially contributing to the smaller thermodynamic scatter across the series. Second, the difference is unlikely to originate primarily from the residue 62 (Asn in CA II and CA XII, Ser in CA XIII), because previous studies demonstrated similar thermodynamic trends even when this residue was not directly involved in ligand interactions⁵⁸.

A more plausible explanation arises from differences in active-site hydration. Comparison of the ordered water networks in CA II and CA XIII (Figure 10) indicates that the principal distinction originates from the W_1 water molecule located near residue 200. In CA II, Thr200 stabilizes W_1 through a short hydrogen bond, whereas in CA XIII, substitution by Val200 alters the position and stabilization of this water molecule. Although apo structures of CA XIII are currently unavailable and the comparison was therefore performed using the acetate-bound structure, this interpretation is supported by the CA II T200V mutant structure, which exhibits a water network highly similar to that observed in CA XIII⁵⁹. This Thr-to-Val variation reflects different requirements for catalytic efficiency and water organization among the isozymes, as CA II is approximately 14-fold more catalytically efficient than CA XIII, with the Thr200/Val200 difference contributing significantly to this difference in activity^{59,60}. At the same time, the resulting enthalpy-entropy balance depends on the particular ligand and its binding mode.

Previous thermodynamic studies similarly demonstrated that CA XIII binding is substantially more enthalpy-favored, whereas CA II binding exhibits a larger entropy contribution despite conservation of the overall active-site architecture⁵⁸. Together, these observations suggest that subtle differences in hydration

and water organization, rather than direct ligand contacts alone, dominate the thermodynamic signature of binding. Structured water molecules are well known to contribute significantly to enthalpy-entropy compensation effects in biomolecular recognition^{61,62}.

Mechanistically, the stronger stabilization of W_1 by Thr200 in CA II generate a more ordered and constrained local hydration network. Upon ligand binding, reorganization of this solvent environment produce a favorable entropy contribution associated with increased solvent freedom, while simultaneously incurring an enthalpic penalty due to disruption of hydrogen-bonding interactions. This interpretation is consistent with the predominantly entropy-favored binding profile observed for CA II.

In contrast, Val200 in CA XIII cannot directly stabilize W_1 through hydrogen bonding, likely resulting in a less constrained hydration environment and consequently a smaller solvent-entropy contribution upon ligand binding. Under these conditions, the binding thermodynamics may appear more enthalpy-favored because a smaller contribution arises from solvent reorganization.

Conclusion

Despite high sequence similarity and nearly superimposable active-site architectures, CA II, CA XII, and CA XIII display distinct thermodynamic signatures for binding of closely related sulfonamide inhibitors. After correcting all observed quantities for proton-linked equilibria, the *intrinsic* Gibbs energies (ΔG_{int}) fall within a narrow range, whereas the underlying enthalpic (ΔH_{int}) and entropic ($-T\Delta S_{\text{int}}$) contributions differ substantially among the three isozymes. Thus, comparable binding affinities arise from distinct balances of enthalpic and entropic driving forces.

High-resolution crystal structures indicate that this divergence originates from subtle microenvironmental variations within the binding pocket. In particular, single-residue substitutions at positions 200 (Thr/Thr/Val), 131

(Phe/Ala/Phe), and 62 (Asn/Asn/Ser) modulate hydrogen-bond geometry, steric complementarity, and the organization of nearby water molecules. Notably, the Thr→Val substitution at position 200 (CA II numbering) in CA XIII disrupts the conserved water network observed in CA II, providing a structural basis for the pronounced enthalpy–entropy shift relative to CA II. These localized perturbations reshape the energetic landscape without substantially altering the overall binding mode.

Together, these results demonstrate how closely related proteins can generate distinct thermodynamic fingerprints through minimal, localized structural changes. By integrating intrinsic thermodynamic analysis with crystallographic data, this work establishes a coherent structure–thermodynamics framework for inhibitor recognition and provides a strategy for dissecting enthalpy–entropy compensation in protonation-linked protein–ligand systems.

Supporting Information Available

- ITC Raw Data (Figure S1) and Replicate Analyses (Table S1-S3)
- ITC Determination of Sulfonamide Protonation Enthalpy (Figure S2)
- Observed vs. Buffer-Corrected vs. Intrinsic Thermodynamics (Figure S3)
- Phe131 Conformations in CA II (Figure S4)
- Docking Validation against X-Ray Structures (Figure S5)
- Interface Contact Areas (Table S4)

Data Availability

Structural data have been deposited in the Protein Data Bank (PDB) under the accession numbers 30EN, 30EO, 9FPQ, 9FPR, 9FPS, 30IC, 30FA, 30FB, 30FP, 30FN, 9FPV, and 9FPW. Thermodynamic data have been deposited in the Protein-Ligand Binding Database (PLBD) under the individual accession numbers provided in the table below. All deposited data are publicly available as of the article publication date.

Compound	PDB ID	PLBD label ^a
CA II		
4a (<i>JA17-1-1</i>)	30EN	ITC-VA0021, ITC-VA0022
4b (<i>JA17-1-2</i>)	30EO	ITC-VA0023, ITC-VA0024
4c (<i>JA17-1-3</i>)	9FPQ ^b	ITC-VA0025, ITC-VA0026
4d (<i>JA17-1-4</i>)	9FPR ^b	ITC-VA0027, ITC-VA0028
4h (<i>JA19-17</i>)	9FPS ^b	ITC-VA0029, ITC-VA0030
7h (<i>JA19-19</i>)	30IC	ITC-VA0031, ITC-VA0032
CA XII		
4a (<i>JA17-1-1</i>)	ND	ITC-VA0033, ITC-VA0034
4b (<i>JA17-1-2</i>)	ND	ITC-VA0035, ITC-VA0036
4c (<i>JA17-1-3</i>)	30FA	ITC-VA0037, ITC-VA0038
4d (<i>JA17-1-4</i>)	30FB	ND
4h (<i>JA19-17</i>)	ND	ITC-VA0039, ITC-VA0040
7h (<i>JA19-19</i>)	ND	ND
CA XIII		
4a (<i>JA17-1-1</i>)	30FP	ITC-VA0041, ITC-VA0042
4b (<i>JA17-1-2</i>)	30FN	ITC-VA0043, ITC-VA0044
4c (<i>JA17-1-3</i>)	9FPV ^b	ITC-VA0045, ITC-VA0046
4d (<i>JA17-1-4</i>)	9FPW ^b	ITC-VA0047
4h (<i>JA19-17</i>)	ND	ITC-VA0048, ITC-VA0049
7h (<i>JA19-19</i>)	ND	ITC-VA0050, ITC-VA0051

^a ITC experiments deposited in the Protein-Ligand Binding Database.

^b Published previously in ²⁴.

Variables

ΔG	standard Gibbs energy change of binding (equivalent to ΔG°)
ΔG_{int}	intrinsic standard Gibbs energy change of binding
ΔG_{obs}	observed standard Gibbs energy change of binding
ΔH	standard enthalpy change of binding (equivalent to ΔH°)
ΔH_{int}	intrinsic standard enthalpy change of binding
ΔH_{obs}	observed standard enthalpy change of binding
ΔS	standard entropy change of binding (equivalent to ΔS°)
ΔS_{int}	intrinsic standard entropy change of binding
ΔS_{obs}	observed standard entropy change of binding
$\Delta H_{\text{pr,buf}}$	standard enthalpy change of buffer protonation
$\Delta H_{\text{pr,CA}}$	standard enthalpy change of CA–Zn ^{II} bound hydroxide protonation
$\Delta H_{\text{pr,SA}}$	standard enthalpy change of protonation of deprotonated sulfonamide amino group
K_{b}	equilibrium binding constant
$K_{\text{b,int}}$	intrinsic equilibrium binding constant
$K_{\text{b,obs}}$	observed equilibrium binding constant
K_{d}	equilibrium dissociation constant
$K_{\text{d,int}}$	intrinsic equilibrium dissociation constant
$K_{\text{d,obs}}$	observed equilibrium dissociation constant
$\text{p}K_{\text{a}}$	negative logarithm of the dissociation constant of acid protons
$\text{p}K_{\text{a,CA}}$	negative logarithm of the dissociation constant of acid protons from CA–Zn ^{II} -bound water molecule
$\text{p}K_{\text{a,SA}}$	negative logarithm of the dissociation constant of acid protons from the protonated sulfonamide amino group
R	gas constant (8.314 J mol ^{−1} K ^{−1})
T	absolute temperature (K)

Author Contributions

Designed research: V.P., D.M.; performed research: V.P., A.Ž., V.K., E.Č., M.B., E.M.; contributed analytic tools: V.P., A.S., S.G.; analyzed data: V.P., D.M.; wrote the paper: V.P.,

V.K., D.M.;

Declaration of Interest

Some of the authors (V.P., A.Ž., E.Č., D.M.) are inventors on patents and/or patent applications related to carbonic anhydrase inhibitors. The authors declare no other competing interests.

References

- (1) Klebe, G. *Drug Design: From Structure and Mode-of-Action to Rational Design Concepts*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2024.
- (2) Minetti, C. A.; Remeta, D. P. Forces Driving a Magic Bullet to Its Target: Revisiting the Role of Thermodynamics in Drug Design, Development, and Optimization. *Life* **2022**, *12*, 1438.
- (3) Fox, J. M.; Zhao, M.; Fink, M. J.; Kang, K.; Whitesides, G. M. The Molecular Origin of Enthalpy/Entropy Compensation in Biomolecular Recognition. *Annual Review of Biophysics* **2018**, *47*, 223–250.
- (4) Klebe, G. Thermodynamic Data Remain a Hot Tip for Decoding Binding Affinity and Water Impact on Protein–Ligand Complex Formation to Assist Lead Optimization. *Journal of Medicinal Chemistry* **2026**, *69*, 3719–3730.
- (5) Ahmad, M.; Helms, V.; Lengauer, T.; Kalinina, O. V. Enthalpy–Entropy Compensation upon Molecular Conformational Changes. *Journal of Chemical Theory and Computation* **2015**, *11*, 1410–1418.
- (6) Sharp, K. Entropy–Enthalpy Compensation: Fact or Artifact? *Protein Science* **2001**, *10*, 661–667.
- (7) Ryde, U. A Fundamental View of Enthalpy–Entropy Compensation. *MedChemComm* **2014**, *5*, 1324–1336.

- (8) Eftink, M. R.; Anusiem, A. C.; Biltonen, R. L. Enthalpy-Entropy Compensation and Heat Capacity Changes for Protein-Ligand Interactions: General Thermodynamic Models and Data for the Binding of Nucleotides to Ribonuclease A. *Biochemistry* **1983**, *22*, 3884–3896.
- (9) Chodera, J. D.; Mobley, D. L. Entropy-Enthalpy Compensation: Role and Ramifications in Biomolecular Ligand Recognition and Design. *Annual Review of Biophysics* **2013**, *42*, 121–142.
- (10) Olsson, T. S. G.; Ladbury, J. E.; Pitt, W. R.; Williams, M. A. Extent of Enthalpy-Entropy Compensation in Protein-Ligand Interactions. *Protein Science: A Publication of the Protein Society* **2011**, *20*, 1607–1618.
- (11) Breiten, B.; Lockett, M. R.; Sherman, W.; Fujita, S.; Al-Sayah, M.; Lange, H.; Bowers, C. M.; Heroux, A.; Krilov, G.; Whitesides, G. M. Water Networks Contribute to Enthalpy/Entropy Compensation in Protein-Ligand Binding. *Journal of the American Chemical Society* **2013**, *135*, 15579–15584.
- (12) Jencks, W. P. On the Attribution and Additivity of Binding Energies. *Proceedings of the National Academy of Sciences* **1981**, *78*, 4046–4050.
- (13) Dragan, A. I.; Read, C. M.; Crane-Robinson, C. Enthalpy-Entropy Compensation: The Role of Solvation. *Eur. Biophys. J.* **2017**, *46*, 301–308.
- (14) Ladbury, J. E.; Klebe, G.; Freire, E. Adding Calorimetric Data to Decision Making in Lead Discovery: A Hot Tip. *Nature Reviews Drug Discovery* **2010**, *9*, 23–27.
- (15) Klebe, G. Broad-scale analysis of thermodynamic signatures in medicinal chemistry: are enthalpy-favored binders the better development option? *Drug Discovery Today* **2019**, *24*, 943–948.
- (16) Freire, E. A thermodynamic approach to the affinity optimization of drug candidates. *Chemical Biology & Drug Design* **2009**, *74*, 468–472.
- (17) Krimmer, S. G.; Klebe, G. Thermodynamics of Protein–Ligand Interactions as a Reference for Computational Analysis: How to Assess Accuracy, Reliability and Relevance of Experimental Data. *J. Comput. Aided Mol. Des.* **2015**, *29*, 867–883.
- (18) Perez-Benito, J. F.; Mulero-Raichs, M. Enthalpy-Entropy Compensation Effect in Chemical Kinetics and Experimental Errors: A Numerical Simulation Approach. *The Journal of Physical Chemistry. A* **2016**, *120*, 7598–7609.
- (19) Glöckner, S.; Klebe, G. Simultaneous Determination of Thermodynamic and Kinetic Data by Isothermal Titration Calorimetry. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2021**, *1865*, 129772.
- (20) Frost, S. C. Physiological functions of the alpha class of carbonic anhydrases. *Subcell Biochem* **2014**, *75*, 9–30.
- (21) Baker, B. M.; Murphy, K. P. Evaluation of Linked Protonation Effects in Protein Binding Reactions Using Isothermal Titration Calorimetry. *Biophysical Journal* **1996**, *71*, 2049–2055.
- (22) Krishnamurthy, V. M.; Kaufman, G. K.; Urbach, A. R.; Gitlin, I.; Gudiksen, K. L.; Weibel, D. B.; Whitesides, G. M. Carbonic Anhydrase as a Model for Biophysical and Physical-Organic Studies of Proteins and Protein-Ligand Binding. *Chem. Rev.* **2008**, *108*, 946–1051.
- (23) Snyder, P. W.; Lockett, M. R.; Moustakas, D. T.; Whitesides, G. M. Is it the shape of the cavity, or the shape of the water in the cavity? *The European Physical Journal Special Topics* **2013**, *223*, 853–891.

- (24) Zakšauskas, A.; Paketurytė-Latvė, V.; Jankūnaitė, A.; Čapkauskaitė, E.; Becart, Y.; Smirnov, A.; Pospíšilová, K.; Leitans, J.; Brynda, J.; Kazaks, A. *et al.* Affinity and Selectivity of Protein–Ligand Recognition: A Minor Chemical Modification Changes Carbonic Anhydrase Binding Profile. *Journal of Medicinal Chemistry* **2025**, *68*, 17752–17773.
- (25) Bissantz, C.; Kuhn, B.; Stahl, M. A Medicinal Chemist’s Guide to Molecular Interactions. *Journal of Medicinal Chemistry* **2010**, *53*, 5061–5084.
- (26) Klebe, G. Applying thermodynamic profiling in lead finding and optimization. *Nature Reviews Drug Discovery* **2015**, *14*, 95–110.
- (27) Zubrienė, A.; Matulis, D. In *Carbonic Anhydrase as Drug Target: Thermodynamics and Structure of Inhibitor Binding*; Matulis, D., Ed.; Springer International Publishing, 2019; pp 107–123.
- (28) Meng, E. C.; Goddard, T. D.; Petersen, E. F.; Couch, G. S.; Pearson, Z. J.; Morris, J. H.; Ferrin, T. E. UCSF ChimeraX: Tools for structure building and analysis. *Protein Science* **2023**, *32*, e4792.
- (29) Shapovalov, M. V.; Dunbrack, R. L. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure* **2011**, *19*, 844–858.
- (30) Axelrod, M.; Bickart, P.; Goldstein, M. L.; Green, M. M.; Kjær, A.; Mislow, K. Absolute configuration and optical rotatory dispersion of methyl alkyl sulfoxides. *Tetrahedron Letters* **1968**, *9*, 3249–3252.
- (31) Rayner, D. R.; Gordon, A. J.; Mislow, K. Thermal racemization of diaryl, alkyl aryl, and dialkyl sulfoxides by pyramidal inversion. *Journal of the American Chemical Society* **1968**, *90*, 4854–4860.
- (32) Fernández, I.; Khiar, N. Recent Developments in the Synthesis and Utilization of Chiral Sulfoxides. *Chemical Reviews* **2003**, *103*, 3651–3706.
- (33) Haynes, W. M., Lide, D. R., Bruno, T. J., Eds. *CRC Handbook of Chemistry and Physics - 2017*, 97th ed.; CRC Press.
- (34) Koes, D. R.; Baumgartner, M. P.; Camacho, C. J. Lessons Learned in Empirical Scoring with smina from the CSAR 2011 Benchmarking Exercise. *Journal of Chemical Information and Modeling* **2013**, *53*, 1893–1904.
- (35) Korb, O.; Stützle, T.; Exner, T. E. An ant colony optimization approach to flexible protein–ligand docking. *Swarm Intelligence* **2007**, *1*, 115–134.
- (36) Quiroga, R.; Villarreal, M. A. Vinardo: A Scoring Function Based on Autodock Vina Improves Scoring, Docking, and Virtual Screening. *PLOS ONE* **2016**, *11*, e0155183.
- (37) Su, M.; Yang, Q.; Du, Y.; Feng, G.; Liu, Z.; Li, Y.; Wang, R. Comparative Assessment of Scoring Functions: The CASF-2016 Update. *Journal of Chemical Information and Modeling* **2019**, *59*, 895–913.
- (38) Mickevičiūtė, A.; Juozapaitienė, V.; Michailovienė, V.; Jachno, J.; Matulienė, J.; Matulis, D. In *Carbonic Anhydrase as Drug Target: Thermodynamics and Structure of Inhibitor Binding*; Matulis, D., Ed.; Springer International Publishing, 2019; pp 15–37.
- (39) Scheuermann, T. H.; Brautigam, C. A. High-Precision, Automated Integration of Multiple Isothermal Titration Calorimetric Thermograms: New Features of NIT-PIC. *Methods (San Diego, Calif.)* **2015**, *76*, 87–98.
- (40) Zhao, H.; Piszczek, G.; Schuck, P. SEDPHAT – A Platform for Global ITC Analysis and Global Multi-Method Analysis of

Molecular Interactions. *Methods* **2015**, *76*, 137–148.

- (41) Brautigam, C. A. Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. *Methods in Enzymology* **2015**, *562*, 109–133.
- (42) Goldberg, R. N.; Kishore, N.; Lennen, R. M. Thermodynamic Quantities for the Ionization Reactions of Buffers. *J. Phys. Chem. Ref. Data* **2002**, *31*, 231–370.
- (43) Kabsch, W. XDS. *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66*, 125–132.
- (44) Battye, T. G. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. W. iMOSFLM: A New Graphical Interface for Diffraction-Image Processing with MOSFLM. *Acta Crystallographica Section D: Biological Crystallography* **2011**, *67*, 271–281.
- (45) Evans, P. Scaling and Assessment of Data Quality. *Acta Crystallographica. Section D, Biological Crystallography* **2006**, *62*, 72–82.
- (46) Vagin, A.; Teplyakov, A. Molecular Replacement with MOLREP. *Acta Crystallographica. Section D, Biological Crystallography* **2010**, *66*, 22–25.
- (47) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallographica. Section D, Biological Crystallography* **2011**, *67*, 355–367.
- (48) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Ict Coot. *Acta Crystallographica Section D* **2010**, *66*, 486–501.
- (49) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *Journal of Cheminformatics* **2012**, *4*, 17.
- (50) Lebedev, A. A.; Young, P.; Isupov, M. N.; Moroz, O. V.; Vagin, A. A.; Murshudov, G. N. JLLigand: A Graphical Tool for the CCP4 Template-Restraint Library. *Acta Crystallographica. Section D, Biological Crystallography* **2012**, *68*, 431–440.
- (51) Goddard, T. D.; Huang, C. C.; Meng, E. C.; Pettersen, E. F.; Couch, G. S.; Morris, J. H.; Ferrin, T. E. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science: A Publication of the Protein Society* **2018**, *27*, 14–25.
- (52) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry* **2009**, *30*, 2785–2791.
- (53) O’Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *Journal of Cheminformatics* **2011**, *3*, 33.
- (54) Halgren, T. A. MMFF VI. MMFF94s option for energy minimization studies. *Journal of Computational Chemistry* **1999**, *20*, 720–729.
- (55) ten Brink, T.; Exner, T. E. Influence of Protonation, Tautomeric, and Stereoisomeric States on Protein-Ligand Docking Results. *Journal of Chemical Information and Modeling* **2009**, *49*, 1535–1546.
- (56) Olechnovič, K.; Venclovás, Č. Voronota: A fast and reliable tool for computing the vertices of the Voronoi diagram of atomic balls. *Journal of Computational Chemistry* **2014**, *35*, 672–681.

- (57) Jiménez, J. S.; Benítez, M. J. Gibbs Free Energy and Enthalpy–Entropy Compensation in Protein–Ligand Interactions. *Biophysica* **2024**, *4*, 298–309.
- (58) Zakšauskas, A.; Čapkauskaitė, E.; Jezepčikas, L.; Linkuvienė, V.; Kišonaitė, M.; Smirnov, A.; Manakova, E.; Gražulis, S.; Matulis, D. Design of two-tail compounds with rotationally fixed benzenesulfonamide ring as inhibitors of carbonic anhydrases. *European Journal of Medicinal Chemistry* **2018**, *156*, 61–78.
- (59) Kim, J. K.; Lim, S. W.; Jeong, H.; Lee, C.; Kim, S.; Son, D. W.; Kumar, R.; Andring, J. T.; Lomelino, C.; Wierman, J. L. *et al.* Fast product release requires active-site water dynamics in carbonic anhydrase. *Nature Communications* **2025**, *16*, 4404.
- (60) Simone, G. D.; Fiore, A. D.; Truppo, E.; Langella, E.; Vullo, D.; Supuran, C. T.; Monti, S. M. Exploration of the residues modulating the catalytic features of human carbonic anhydrase XIII by a site-specific mutagenesis approach. **2019**, *34*, 1506–1510.
- (61) Fox, J. M.; Kang, K.; Sastry, M.; Sherman, W.; Sankaran, B.; Zwart, P. H.; Whitesides, G. M. Water-Restructuring Mutations Can Reverse the Thermodynamic Signature of Ligand Binding to Human Carbonic Anhydrase. **2017**, *56*, 3833–3837.
- (62) Ladbury, J. E. Just add water! The effect of water on the specificity of protein–ligand binding sites and its potential application to drug design. *Chemistry & Biology* **1996**, *3*, 973–980.

Publication 2

Affinity and Selectivity of Protein-Ligand Recognition: A Minor Chemical Modification Changes Carbonic Anhydrase Binding Profile

Audrius Zakšauskas, **Vaida Paketurytė-Latvė**, Alberta Jankūnaitė, Edita Čapkauskaitė, Yann Becart, Alexey Smirnov, Klára Pospíšilová, Janis Leitans, Jiří Brynda, Andris Kazaks, Lina Baranauskienė, Elena Manakova, Saulius Gražulis, Visvaldas Kairys, Kaspars Tars, Pavlína Řezáčová, Daumantas Matulis

J. Med. Chem., 2025, 68 (16), 17752–17773

DOI: 10.1021/acs.jmedchem.5c01421

Affinity and Selectivity of Protein–Ligand Recognition: A Minor Chemical Modification Changes Carbonic Anhydrase Binding Profile

Audrius Zakšauskas, Vaida Paketurytė-Latvė, Alberta Jankūnaitė, Edita Čapkauskaitė, Yann Becart, Alexey Smirnov, Klára Pospíšilová, Janis Leitans, Jiří Brynda, Andris Kazaks, Lina Baranauskienė, Elena Manakova, Saulius Gražulis, Visvaldas Kairys, Kaspars Tars, Pavlína Rezáčová, and Daumantas Matulis*



Cite This: <https://doi.org/10.1021/acs.jmedchem.5c01421>



Read Online

ACCESS |

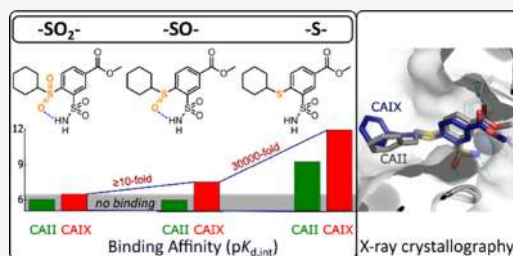
Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Discovery of small-molecule drugs relies on their strong binding affinity compared to nontarget proteins, thus possessing selectivity. Minor chemical structure changes usually exhibit little change in the compound efficacy, with rare exceptions. We developed a series of nearly 50 *ortho*-substituted benzene-sulfonamides and experimentally measured their interactions with the 12 catalytically active human carbonic anhydrase (CA) isozymes. Inhibitors were designed using seven different substituent groups, including 4-sulfanyl-substituted 3-sulfamoyl benzoates and benzamides, 4-sulfinyl-substituted 3-sulfamoyl benzoates and benzamides, 4-sulfonyl-substituted 3-sulfamoyl benzoates and benzamides, and 4-amino-substituted benzamides.

The oxidation state of sulfur at the *ortho* position significantly influenced the compound's affinity for CAIX, a target for anticancer drugs, demonstrating affinities hundreds of thousands of times stronger than related compounds. Coupled with X-ray crystal structures and molecular docking, the relationship between structure and thermodynamics offers insights into how small changes in the structure lead to significant changes in affinity for drug design.



INTRODUCTION

A detailed understanding of protein–ligand recognition is an essential goal in small-molecule drug discovery.¹ Any drug-like chemical compound should bind the disease-related target protein with sufficiently high affinity. Furthermore, the compound must bind with high selectivity and thus not interact strongly with other nontarget proteins whose inhibition or binding could cause undesired side effects.² Rational design of such compounds is complex because of the limited understanding of the underlying energies of binding and how a compound recognizes and binds to the target protein.

As a model system, we study sulfonamide compound binding to human carbonic anhydrases (CA), zinc-containing enzymes.³ Humans have 12 catalytically active CA isozymes (EC 4.2.1.1).^{4–6} The enzyme catalyzes the reversible hydration of carbon dioxide and has many essential physiological functions. Since these enzymes function in pH and electrolyte homeostasis and regulation, many drugs target CA isozymes to treat diseases like glaucoma, edema, obesity, epilepsy, infertility, and cancer.^{7–9} Primary sulfonamides are the most investigated CA inhibitors.^{10–12} Their amino group binds directly to the catalytic zinc in the active site by forming a coordination bond and inhibits the activity of all CA isozymes. However, the binding affinity may be low or high and depends on small details of each

compound arrangement on the protein surface, possible steric hindrances, or attraction due to hydrogen bonds or hydrophobic interactions.¹³

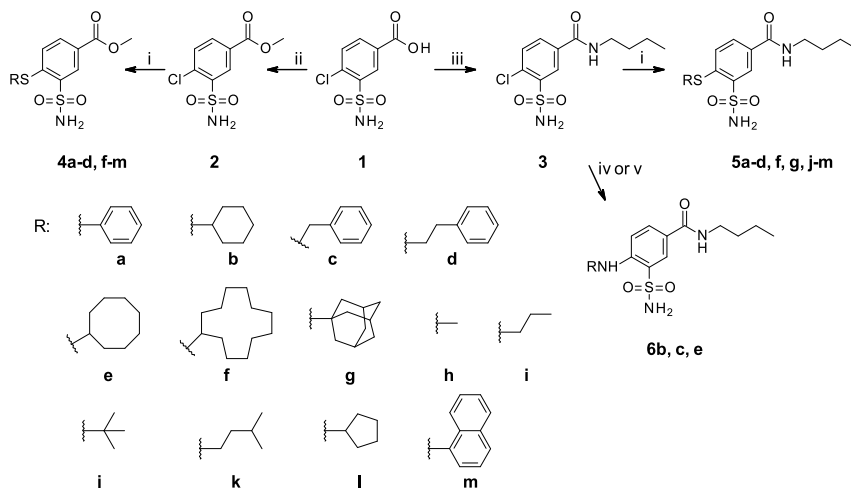
The 12 catalytically active human CA isozymes have nearly identical beta-sheet folds. Their active sites are highly similar in shape, but several amino acids in the active site vary among the isozymes.^{14–16} Because the differences in the active site amino acid composition are small, it is difficult to design compounds that would bind one isozyme with high affinity while all other isozymes with low affinity, thus leading toward high selectivity for only one isozyme. The active site of CAs is funnel-shaped and has hydrophobic and hydrophilic sides for some isozymes. Differences in a few amino acids determine the selectivity of inhibitors for particular isozymes. Introducing various scaffolds on the aromatic sulfonamide ring targets unique residues in the active site.¹⁷ In most studies, the tails are relatively distant from

Received: May 26, 2025

Revised: June 27, 2025

Accepted: July 8, 2025

Scheme 1. Synthesis of Methyl 4-Substituted-3-Sulfamoylbenzoates 4a–d, f–m, 4-Substituted 3-Sulfamoylbenzamides 5a–d, f, g, j–m, and 4-Amino-Substituted Benzamides 6b, c, e^a



^aReagents and conditions: (i) RSH, K₂CO₃, DMF, 80 °C; (ii) H₂SO₄, MeOH, Δ; (iii) (a) SOCl₂, toluene, Δ, (b) RNH₂, THF, 0 °C; (iv) RNH₂, 130 °C; (v) RNH₂, TEA, toluene, Δ.

the sulfonamide group, resulting in weak interactions with peripheral amino acids. The affinity profile depended mostly on the substituents located close to the sulfonamide head-group.^{18–20} This is also dependent on the flexibility of the substituents, which help adjust to the protein shape.

Depending on the variation of the substituents while investigating doubly substituted compounds,^{20–22} several high-affinity compounds exhibiting picomolar *K_d* were obtained for CAVII, CAIX, CAXII, and CAXIV. A significant achievement was a *LJJ15-12* compound that exhibited 0.08 pM intrinsic *K_d* for CAIX.²⁰ In this study, we investigate the functional groups in the *ortho* position by varying 13 substituents from small methyl to bulky adamantyl. Sulfinyl and amino substituents were also synthesized to examine the influence of the linking atoms. Interestingly, the substitution of only one atom, an addition of an oxygen atom in the *ortho* position, decreased compound affinity for nontarget isozymes by a million-fold, significantly improving the selectivity, which is one of the main goals in small-molecule drug discovery.

RESULTS

Organic Synthesis of Designed Compounds. The synthesis of 2,5-disubstituted benzenesulfonamides was performed starting from 4-chloro-3-sulfamoylbenzoic acid **1** (Scheme 1). Methyl ester **2** was obtained from 4-chloro-3-sulfamoylbenzoic acid **1** by reflux in methanol in the presence of a catalytic amount of sulfuric acid. Amide **3** was synthesized by refluxing acid **1** with thionyl chloride in toluene and subsequent treatment, the resulting anhydride with an appropriate amine, using as the base excess of amine according to the procedure reported in ref 21.

The synthesis of 4-substituted sulfamoylbenzoic acid derivatives **4a–d, f–m** was achieved by aromatic nucleophilic substitution of the chlorine substituent with various thiols under an inert atmosphere in dimethylformamide using potassium carbonate as the base (Scheme 1). All thiols were commercially

available except cyclododecylthiol, which was synthesized by subjecting cyclododecanone to reaction with 1,2-ethanedithiol and subsequent reduction of intermediate dithiolane with *n*-butyllithium.²³

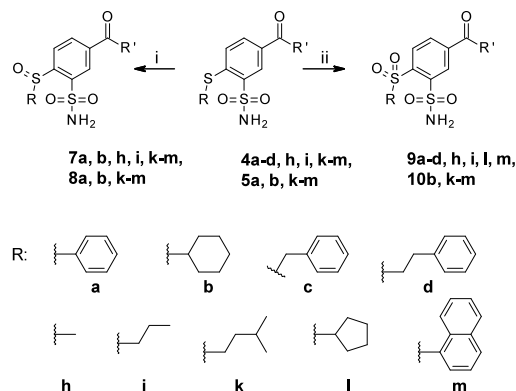
The 4-substituted 3-sulfamoylbenzamides **5a–d, f, g, j–m** were synthesized using the same reaction conditions as 4-sulfanyl-substituted esters **4a–d, f–m**. Substitution of the chlorine group with amines required harsher reaction conditions than thiols. Therefore, 4-amino-substituted benzamides **6b** and **c** were synthesized by heating amide **3** over the appropriate amine at 130 °C. Compound **6e** was synthesized by boiling amide **3** in toluene with 2 equiv of cyclooctylamine and 2 equiv of triethylamine.

The oxidation of esters **4a–d, h, i, k–m** and benzamides **5a, b, k–m** to the sulfinyl and sulfonyl compounds was performed using in situ generated peracetic acid (Scheme 2). The reaction was carried out at room temperature and produced 4-sulfinyl compounds **7a, b, h, i, k–m, 8a, b, k–m**, and heating the reaction mixture at 70 °C yielded the corresponding 4-sulfonyl compounds **9a–d, h, i, l, m, 10b**, and **k–m**.

Compound Binding to CA Isozymes. All synthesized compounds were divided into 7 groups: **4(a–d, f–m)**, **5(a–d, f, g, j–m)**, **6(b, c, e)**, **7(a, b, h, i, k–m)**, **8(a, k–m)**, **9(a–d, h, i, l, m)**, and **10(b, k–m)** (Figure 1). Compounds that started with numbers **4, 7, and 9** were 4-sulfanyl-, 4-sulfinyl-, and 4-sulfonyl-substituted esters, respectively. Compounds starting with the numbers **5, 8, and 10** were 4-sulfanyl-, 4-sulfinyl-, and 4-sulfonyl-substituted benzamides analogous to the previous series. Compounds starting with the number **6** were 4-amino-substituted benzamides. Various linear, branched, and cyclic aliphatic and aromatic substituents at the *ortho* position were tested to assess whether size, flexibility, and hydrophobicity affect affinity. Compound affinities for the 12 catalytically active human CA isozymes are listed in Table 1.

As one of the important findings in this manuscript, Figure 2 arranges the compounds in the order of their affinity for CAIX

Scheme 2. Synthesis of Methyl 4-Sulfinyl-Substituted-3-Sulfamoyl Benzoates **7a, b, h, i, k–m**, *N*-Butyl-4-sulfinyl-Substituted-3-Sulfamoylbenzamides **8a, b, k–m**, Methyl 4-Sulfonyl-Substituted-3-Sulfamoylbenzoates **9a–d, h, i, l, m**, and *N*-Butyl-4-sulfonyl-Substituted-3-Sulfamoylbenzamides **10b, k–m**^a



R': OMe (**4, 7, 9**); NH-nBu (**5, 8, 10**)

^aReagents and conditions: (i) 30% H₂O₂, AcOH, r. t.; (ii) 30% H₂O₂, AcOH, 70 °C.

and compared to undesired target CAII. A high affinity for CAIX is desired, because CAIX is implicated in various types of cancer.²⁴ However, CAII is abundant in erythrocytes, thus an off-target for anticancer inhibitors.

Compound affinities for human CA isozymes were analyzed using fluorescence-based thermal shift assay (FTSA) and the enzymatic activity stopped-flow-based inhibition assay (SFA) (Figure 3). The FTSA determined the observed dissociation constants for all compounds with all 12 CA isozymes (Table 1 and Supplementary Figure S1). From these experimentally determined $K_{d,obs}$, the intrinsic dissociation constants $K_{d,int}$ were calculated, and the results are primarily focused on them. Figure 3B,C shows two compounds with strong and weak affinity for CAII and CAIX by FTSA. A single oxygen atom strongly diminished affinity for both isozymes. The stopped-flow assay of CO₂ hydration enzymatic activity inhibition confirmed that the compounds inhibited CAIX and CAII (Figure 3D,E). Inhibition K_i values are presented in Table 2 and Figure S2. The FTSA is a

convenient technique covering a significantly wider K_d range than the SFA.²⁵

The experimental conditions slightly differed between FTSA and SFA, and it is therefore not appropriate to directly compare the $K_{d,obs}$ and K_i . However, the measured affinities were similar in both techniques. Further analysis is based on FTSA data due to its wider limits in identifying strong binders. Note that the enzyme concentration limits the SFA's ability to determine high-affinity binders. For example, if we use a 10 nM concentration of a CA isozyme, the lowest IC₅₀ is 5 nM, half of the protein concentration. Any compound with an IC₅₀ stronger than 5 nM would exhibit a dosing curve that appears as an IC₅₀ of 5 nM. Thus, compounds that possess single-digit nM or picomolar K_d , cannot be distinguished by SFA.

Sulfonamide binding of CA is a pH-dependent reaction.^{6,26,27} The water molecule bound to the CA zinc ion is replaced by the deprotonated form of sulfonamide upon binding.^{28,29} The protonation forms required for the interaction exist at different pHs: the CA-Zn(II)-H₂O has the largest fraction at acidic pH and the deprotonated sulfonamide has the largest fraction at alkaline pH. Therefore, the measured affinity is always lower than the intrinsic affinity. The intrinsic parameters are calculated (see equations in the Experimental section) by knowing the pK_a of the CA zinc-bound water and the pK_a of the sulfonamide group (Table 1 and Figure S3). Experimental data of the sulfonamide group pK_a determination for compound **4b** are shown in Figure 4. Figure S3 shows the graphs of pK_a determination for the remaining compounds.

Intrinsic affinities are especially important in rational drug design. It is not rare when a stronger affinity is observed not because of the formed bonds but because of the substitutions that lower the pK_a of the sulfonamide group and thus increase the fraction of the ready-to-bind form.³⁰ Intrinsic parameters are used to avoid misleading conclusions when comparing the affinity of compounds for CAs.

Esters vs Benzamides. 4-sulfinyl-substituted esters **4a–d, f–m** were the largest (12 compounds) studied group of compounds with the same framework. Compounds of this group, even almost independently of the substituent, interacted strongly and selectively with several CAs. $K_{d,int}$ of **4b** for CAIX was 0.0010 nM and it was the strongest interaction measured in this study, CAXIII—0.090 nM and CAXIV—0.020 nM, with others interacting much weaker. The size of the substituent was critical in this interaction. CAIX has a larger active site than the rest of the CAs.³¹ It was interesting that compound **4g** (adamantyl), which has a similar affinity for CAIX ($K_{d,int}$ =

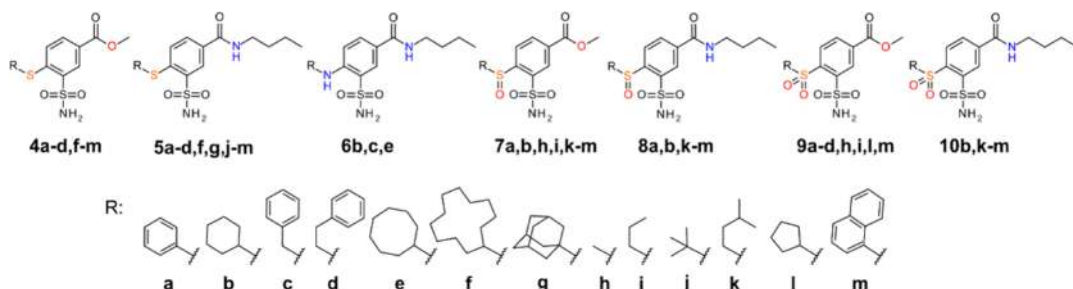


Figure 1. Chemical structures of the compounds synthesized and investigated in this study. Compounds **1, 2**, and **3** are the starting compounds of the synthesis shown in Scheme 1.

Table 1. Observed and Intrinsic ($K_{d,obs}$ and $K_{d,int}$) Dissociation Constants (in nM Units) of Investigated Compounds to All Catalytically Active Human CAs at 37 °C Obtained by Fluorescent Thermal Shift Assay (FTSA)^a

Compound ID, name	A framework of the structure	$pK_{a,obs}$	fraction at pH 7.0	pH _{CA,obs}											
				CAI	CAII	CAIII	CAIV	CAVA	CAVB	CAVI	CAVII	CAIX	CAXI	CA XII	CA XIV
				8.1	6.9	6.5	6.6	7.3	7.0	6.0	6.8	6.6	6.8	8.0	6.8
				92.6%	44.3%	24.0%	28.5%	66.6%	50.0%	9.1%	38.7%	28.5%	38.7%	90.9%	38.7%
2 JA17-1		9.2	0.63%	$K_{d,obs}$	21000	460	≥200000	500	770	42	300	240	240	92	220
				$K_{d,int}$	120	1.3	≥300	0.90	3.2	0.13	0.17	0.58	0.43	0.22	1.3
3 JA17-3		9.1	0.79%	$K_{d,obs}$	36000	140	44000	99	1800	70	340	94	190	10	880
				$K_{d,int}$	260	0.50	83	0.22	10	0.28	0.24	0.29	0.43	0.030	6.3
4a JA17-1-1		9.8	0.16%	$K_{d,obs}$	≥200000	87	≥200000	460	≥200000	120	430	140	4.6	260	12
				$K_{d,int}$	≥290	0.060	≥76	0.21	≥210	0.10	0.060	0.090	0.0020	0.16	0.030
4b JA17-1-2		9.7	0.20%	$K_{d,obs}$	≥200000	560	≥200000	810	≥200000	440	710	680	1.8	280	52
				$K_{d,int}$	≥370	0.49	≥96	0.46	≥270	0.44	0.13	0.52	0.0010	0.22	0.090
4c JA17-1-3		9.8	0.16%	$K_{d,obs}$	≥200000	590	≥200000	110	140000	930	660	960	10	350	52
				$K_{d,int}$	≥290	0.42	≥76	0.050	150	0.70	0.10	0.60	0.0050	0.21	0.080
4d JA17-1-4		9.6	0.25%	$K_{d,obs}$	160000	580	≥200000	1100	97000	190	1300	360	19	450	27
				$K_{d,int}$	380	0.64	≥120	0.79	160	0.23	0.29	0.34	0.010	0.43	0.060
4f JA18-29		9.7*	0.20%	$K_{d,obs}$	≥200000	1000	≥200000	2100	≥200000	1800	6500	18000	11	68	100
				$K_{d,int}$	≥370	0.90	≥96	1.2	≥270	1.8	1.2	14	0.0060	0.050	0.18
4g E19-3		9.7*	0.20%	$K_{d,obs}$	≥200000	3300	≥200000	100000	≥200000	≥200000	13000	96000	4.8	1500	1200
				$K_{d,int}$	≥370	2.9	≥96	57	≥270	≥200	2.3	74	0.0030	1.1	2.2
4h JA19-17		9.8	0.16%	$K_{d,obs}$	≥200000	2200	≥200000	960	120000	680	500	1000	650	6500	570
				$K_{d,int}$	≥290	1.6	≥76	0.43	130	0.54	0.070	0.61	0.29	4.0	0.80
4i JA18-30		9.6	0.25%	$K_{d,obs}$	≥200000	1000	≥200000	400	91000	990	700	490	23	1800	98
				$K_{d,int}$	≥460	1.1	≥120	0.28	150	1.2	0.16	0.48	0.020	1.7	0.22
4j JA18-33		9.6	0.25%	$K_{d,obs}$	≥200000	2400	≥200000	10000	≥200000	4900	4600	15000	290	120	220
				$K_{d,int}$	≥460	2.7	≥120	7.4	≥330	6.1	1.0	14	0.21	0.12	0.51
4k JA18-32		9.8	0.16%	$K_{d,obs}$	40000	370	≥200000	160	85000	310	300	210	2.7	190	40
				$K_{d,int}$	58	0.26	≥76	0.070	89	0.25	0.040	0.13	0.0010	0.11	0.060
4l E19-1		9.6	0.25%	$K_{d,obs}$	≥200000	670	≥200000	1000	180000	6700	560	180	3.3	670	110
				$K_{d,int}$	≥460	0.74	≥120	0.71	300	8.4	0.13	0.18	0.0020	0.65	0.25
4m E19-2		9.7	0.20%	$K_{d,obs}$	≥200000	50	11000	2000	≥200000	1100	500	8.7	2.5	500	1.4
				$K_{d,int}$	≥370	0.040	5.3	1.1	≥270	1.2	0.090	0.0070	0.0010	0.39	0.0030
5a JA17-3-1		9.7	0.20%	$K_{d,obs}$	≥200000	16000	≥200000	≥200000	120000	9000	19000	53000	360	17000	17000
				$K_{d,int}$	≥370	14	≥96	≥110	160	9	3.5	41	0.20	13	30
5b JA17-3-2		9.8	0.16%	$K_{d,obs}$	≥200000	180000	100000	≥200000	≥200000	33000	≥200000	≥200000	450	140000	150000
				$K_{d,int}$	≥290	120	38	≥90	≥210	26	≥29	≥120	0.20	83	220
5c JA17-3-3		9.8	0.16%	$K_{d,obs}$	≥200000	12000	≥200000	40000	68000	33000	43000	77000	2100	8600	43000
				$K_{d,int}$	≥290	8.1	≥76	18	72	26	6.2	47	0.93	5.3	62
5d JA17-3-4		9.9	0.13%	$K_{d,obs}$	≥200000	58000	≥200000	≥200000	110000	17000	98000	≥200000	3200	46000	110000
				$K_{d,int}$	≥230	32	≥60	≥72	95	10	11	≥97	1.2	22	130
5f JA19-16		9.8*	0.16%	$K_{d,obs}$	≥200000	100000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	840	≥200000	≥200000
				$K_{d,int}$	≥290	70	≥76	≥90	≥210	≥160	≥29	≥120	0.38	≥120	≥290
5g YB19-4		9.8*	0.16%	$K_{d,obs}$	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	2000	≥200000	≥200000
				$K_{d,int}$	≥290	≥140	≥76	≥90	≥210	≥160	≥29	≥120	0.90	≥120	≥290
5j YB19-3		9.8	0.16%	$K_{d,obs}$	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	25000	≥200000	≥200000
				$K_{d,int}$	≥290	≥140	≥76	≥90	≥210	≥160	≥29	≥120	11	≥120	≥290

Table 1. continued

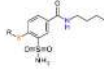
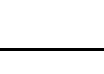
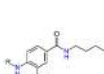
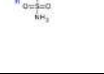
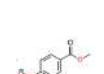
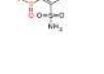
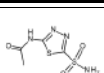
		p <i>K</i> _a														
		fraction at pH														
		CAI	CAII	CAIII	CAIV	CAVA	CAVB	CAVI	CAVII	CAIX	CAXII	CAXIII	CAXIV			
Compound lab. name	A framework of the structure	p <i>K</i> _{a,50}	7.0	8.1	6.9	6.5	6.6	7.3	7.0	6.0	6.8	6.6	6.8	8.0	6.8	
5k YB19-2		9.7	0.20%	<i>K</i> _{a,obs}	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	670	≥200000	≥200000	≥200000	
5l YB19-5				<i>K</i> _{a,est}	≥370	≥180	≥96	≥110	≥270	≥200	≥36	≥150	0.38	≥150	≥360	≥150
5m YB19-1				<i>K</i> _{a,est}	≥290	≥140	≥76	≥90	≥210	≥160	≥29	≥120	0.36	≥120	≥290	≥120
6b JA17-3-11		9.9	0.13%	<i>K</i> _{a,obs}	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	500	≥200000	10000	≥200000	
6c JA17-3-9				<i>K</i> _{a,est}	≥230	≥110	≥60	≥72	≥170	≥130	≥23	≥97	0.18	≥97	11	≥97
6d JA17-3-10				<i>K</i> _{a,est}	≥200000	22000	≥200000	63000	83000	10000	25000	62000	1500	6500	32000	33000
7a JA19-11		8.9	1.24%	<i>K</i> _{a,obs}	≥200000	110000	≥200000	180000	≥200000	≥200000	≥200000	1200	64000	150000	20000	
7b JA19-10				<i>K</i> _{a,est}	≥290	79	≥76	≥90	200	≥160	≥29	≥120	0.53	39	220	12
7c JA19-19				<i>K</i> _{a,est}	≥200000	4300	≥200000	5500	≥200000	33000	1200	18000	5200	15000	720	3300
7d JA19-8		8.9	1.24%	<i>K</i> _{a,obs}	≥200000	24000	≥200000	9000	≥200000	14000	13000	9300	140	7600	400	
7e E19-6				<i>K</i> _{a,est}	≥2300	24	≥600	19	≥1700	200	1.3	87	18	74	8.1	16
7f E19-7				<i>K</i> _{a,est}	≥200000	24000	≥200000	9000	≥200000	14000	13000	9300	140	7600	400	2000
8a JA19-15		9.0	0.99%	<i>K</i> _{a,obs}	≥200000	160000	≥200000	≥200000	≥200000	≥200000	≥200000	3300	≥200000	180000	≥200000	
8b YB19-10				<i>K</i> _{a,est}	≥1800	710	≥480	≥560	≥1300	≥990	≥180	≥770	9	≥770	1600	≥770
8c YB19-11				<i>K</i> _{a,est}	≥200000	24000	≥200000	9000	≥200000	14000	13000	9300	140	7600	400	2000
8d YB19-7		8.9	1.24%	<i>K</i> _{a,obs}	≥200000	141800	≥200000	≥200000	≥200000	≥200000	≥200000	4600	≥200000	146300	≥200000	
8e JA17-1-5				<i>K</i> _{a,est}	≥2300	780	≥600	≥710	≥1700	≥1200	≥230	≥960	16.0	≥960	1700	≥960
8f JA17-1-7				<i>K</i> _{a,est}	≥200000	56000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000
8g JA17-1-8		9.0	0.99%	<i>K</i> _{a,obs}	≥200000	160000	≥200000	≥200000	≥200000	≥200000	≥200000	3300	≥200000	180000	≥200000	
8h JA17-1-6				<i>K</i> _{a,est}	≥2300	≥1100	≥600	≥240	≥1700	≥1200	≥230	≥340	37	≥960	≥640	≥1400
8i JA19-20				<i>K</i> _{a,est}	≥200000	24000	≥200000	9000	≥200000	14000	13000	9300	140	7600	400	2000
9a JA19-20		8.8	1.56%	<i>K</i> _{a,obs}	≥200000	82000	≥200000	89000	140000	≥200000	160000	≥200000	25000	≥200000	59000	
				<i>K</i> _{a,est}	≥2900	570	≥750	400	1400	≥1600	91	960	320	1100	310	390
				<i>K</i> _{a,est}	≥200000	82000	≥200000	89000	140000	≥200000	160000	≥200000	25000	≥200000	59000	

Table 1. continued

Compound lab. name	A framework of the structure	pK _{a,CA}	fraction at pH 7.0	pK _{a,CA}											
				CAI	CAII	CAIII	CAIV	CAVA	CAVB	CAVI	CAVII	CAIX	CAXI	CAXII	CAXIII
9i JA19-2		8.9	1.24%	K _{d,int}	≥200000	1700	≥200000	6200	≥200000	9000	1500	≥200000	900	16000	250
9i E19-4				K _{d,int}	≥200000	71000	≥200000	190000	≥200000	≥200000	≥200000	9100	≥200000	22000	20000
9m E19-5				K _{d,int}	≥200000	5900	≥200000	≥200000	≥200000	≥200000	35000	6800	4000	≥200000	630
10b JA19-14				K _{d,int}	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000
10k YB19-8				K _{d,int}	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	160000
10l YB19-9				K _{d,int}	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000
10m YB19-6				K _{d,int}	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000	≥200000
AZM		7.0	50%	K _{d,int}	2400	46	40000	87	840	140	220	13	21	130	79
				K _{d,int}	1100	9.8	4600	12	270	34	9.8	2.4	2.9	24	35

“Observed values were determined using 50 mM sodium phosphate buffer at pH 7.0 containing 100 mM sodium chloride, 50 μM ANS dye, and 2% (v/v) DMSO. $K_{d,int}$ s were calculated according to eq 1 (see the Experimental Section). Values with a “≥” sign show that they are at the detection limit of ≥200 000 nM of $K_{d,obs}$ according to the highest used ligand concentration. The intrinsic value limits vary for different CAs and compounds due to differences in pK_a , $pK_{a,SA}$ – pK_a of the sulfonamide group; $pK_{a,CA}$ – pK_a value of water molecule bound to Zn(II) in the active site of CA; AZM, acetazolamide (a standard inhibitor).¹ Not determined due to solubility issues or low intensity of the spectrum curves. The $pK_{a,SA}$ value was assigned based on similarities in chemical structure.

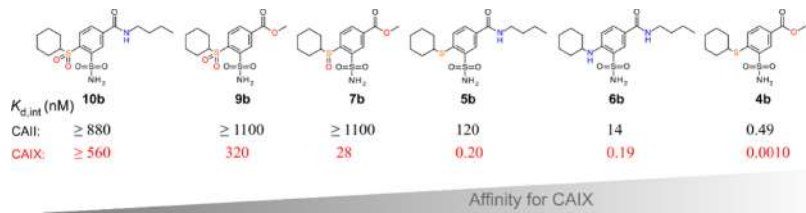


Figure 2. Compounds are arranged in the order of increasing affinity for CAIX. Compound 4b had the highest affinity and selectivity for CAIX. The intrinsic dissociation constants (in nM units) are compared for CAII and CAIX, while values for the remaining CAs are listed in Table 1.

0.0030 nM) to most of the compounds in this group, is a few hundred times more selective for CAIX, whereas, for example, 4b (cyclohexyl) is only a few tens of times more selective for CAIX and can be considered a strong binder to several undesired CAs. On the other hand, the 4-methylsulfanyl-substituted compound 4h is no longer selective. Thus, the hydrophobic interaction made a significant contribution to affinity for CAIX, and selectivity was mainly obtained by the size of the substituent when the binding to CAIX was optimal and binding to other CAs was limited by steric interference. Analogous benzamides were much weaker binders of CAs. Most benzamides exhibited no interaction with CAs. Nevertheless, all compounds in this series bound to CAIX and were, in most cases, at least several dozen times more selective for it than for other CAs.

Sulfanyl vs Sulfenyl vs Sulfonyl Compounds. Different forms of sulfur oxidation led to drastically different affinities for CAs. A higher degree of oxidation in these compounds weakened the

affinity. However, in our opinion, it was not the oxidation itself that had the main influence, but rather the conformation of the compound. The affinity of all sulfenyl- compounds with all CAs was significantly weaker than analogous sulfanyl- compounds. The decrease in affinity varied depending on the CA isozyme and the substituents. Therefore, no generalized observations could be made. For example, compound 7a did not bind to CAI and CAVA, but the weak affinity was determined for CAIII, with all other CAs the interactions were similar and did not exceed more than 10-fold in most cases and there was no selectivity for CAIX. Compounds 7b, 7h, and 7i did not bind or bind weakly and nonselectively to all CAs. Except for 7b, it is bound only to CAIX with 28 nM. The $K_{d,int}$ of compounds 7k, 7l, and 7m were 0.63, 0.72, and 0.24 nM, respectively. Also, there was a similar affinity for CAXIII and CAXIV and weaker for the other CAs. From compounds 5(a–d, f, g, j–m) to 8(a, k–m) decreased affinity for all CAs.

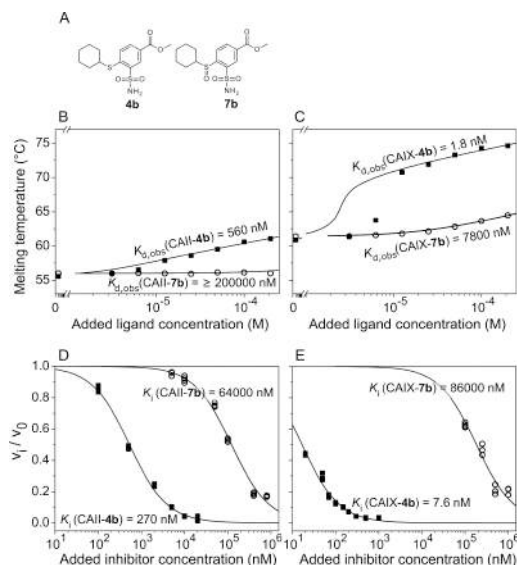


Figure 3. Compound affinity was determined by two assays in this study. (A) Chemical structures of two compounds whose binding data are shown below. (B) and (C) Fluorescent thermal shift assay data (FTSA) of compounds **4b** (closed squares) and **7b** (open circles) binding to CAII and CAIX, respectively. (D) and (E) Stopped-flow carbon dioxide hydration assay (SFA) data of compounds **4b** (closed squares) and **7b** (open circles) inhibition of CAII and CAIX, respectively. The dissociation constants (K_d) or inhibition constants (K_i) are given next to the corresponding curves. It is important to note that the experimental conditions of the methods were slightly different: FTSA, pH 7.0, 37 °C, while for SFA, pH 7.5, 25 °C.

Switching to sulfonyl compounds reduced the affinity and abolished the selectivity for CAIX. **9b** only bound to CAVII, CAIX, CAXIII, and CAXIV with $K_{d,int}$ 130, 320, 1100, and 680 nM, respectively. The most strongly interacting compound in this series was **9i**, e.g., $K_{d,int}$ of CAVI—1.7 nM, CAIX—3.2 nM but also bound to other CAs quite similarly. Meanwhile, compounds **10(b, k–m)** did not bind to any isozymes, only a couple of measurements showed a weak interaction.

4-Sulfonyl- vs 4-Amino-Substituted. Comparing **5b** vs. **6b** and **5c** vs. **6c**, in most cases, the dissociation constants differed

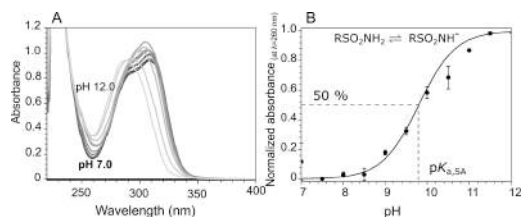


Figure 4. Spectrophotometric determination of the sulfonamide group deprotonation pK_a for compound **4b**. (A) Absorption spectra of compound **4b** in solutions at various pH at intervals of 0.5 pH units at 37 °C. (B) Normalized absorbance at 260 nm was plotted as a function of pH, and the pK_a value was determined as a midpoint of the curve. The data points are the mean points of two repeats ((A) shows one repeat for simplicity), with standard deviations. The pK_a value of compound **4b** was 9.74 ± 0.13 ($\pm 1.3\%$) with a confidence interval [9.62–9.87] of 95%.

only a few times, and the constants for CAIX did not exceed the margin of error. The S or N atom in the same position of the compound almost did not change the affinity for all isozymes. Compound **6e** (cyclooctyl-substituted) selectively interacted with CAIX, $K_{d,int}$ of CAIX—0.53 nM, CAXII—39 nM, and CAXIV—12 nM. It did not bind to other CAs. Most likely the selectivity was due to the size of the active site pocket.

X-ray Crystal Structures of Compound Binding to CAII and CAXIII. Nine crystal structures were determined by X-ray crystallography, the complexes of CAII with compounds **2**, **3**, **4c**, **4d**, **4h**, CAIX with **4d** and **5b**, and CAXIII with **4c** and **4d**. Table 3 lists the data collection and refinement statistics. Figure 5 shows the electron density maps of these compounds in the active site of CAs. Two molecules of compound **4h** were identified in the active site of CAII, one conventionally formed a coordination bond between the sulfonamide group and zinc, and the other was independently located near the periphery of the active site. The localization of both separated molecules relative to the same zinc is shown in Figure 5E,F. The main highlights of the identified protein–ligand interactions are described below and illustrated in Figure 6.

Structures of 2 vs 3 Bound to CAII. Starting compounds **2** and **3** used in the synthesis differed by one substitution at the *meta* position relative to the sulfonamide group, ester vs. amide. Figure 6A,B shows the position of these compounds in the active site of CAII. Both compounds retained the same position of the sulfonamide group and the chlorine atom in the active site of

Table 2. Inhibition Constants and IC_{50} (in nM Units) of CAII and CAIX with Compounds Obtained by Stopped-Flow Carbon Dioxide Hydration Assay (SFA) at 25 °C^a

compound	K_i (nM)		IC_{50} (nM)	
	CAII	CAIX	CAII	CAIX
4a	25 $\pm$ 1.3	5.4 $\pm$ 0.5	47 $\pm$ 2.6	12 $\pm$ 0.5
4b	270 $\pm$ 20	7.6 $\pm$ 0.5	530 $\pm$ 40	16 $\pm$ 1.3
4c	210 $\pm$ 30	37 $\pm$ 3.5	400 $\pm$ 50	79 $\pm$ 7
4d	120 $\pm$ 10	51 $\pm$ 6	230 $\pm$ 20	110 $\pm$ 12
4h	9200 $\pm$ 500	1300 $\pm$ 200	18,000 $\pm$ 1000	2700 $\pm$ 400
7a	63 $\pm$ 6	360 $\pm$ 30	120 $\pm$ 12	760 $\pm$ 60
7b	64,000 $\pm$ 7500	86,000 $\pm$ 16,000	100,000 $\pm$ 10,000	200,000 $\pm$ 30,000
7h	3400 $\pm$ 200	7600 $\pm$ 700	6500 $\pm$ 300	16,000 $\pm$ 1600
9a	34,000 $\pm$ 4000	3700 $\pm$ 600	65,000 $\pm$ 7000	7800 $\pm$ 1300

^aExperiments were performed using 20 mM HEPES Na at pH 7.5, 20 mM Na₂SO₄, and 0.2 mM Phenol Red.

Table 3. X-ray Diffraction Data Collection and Refinement Statistics

isozyme–ligand	CAI—2	CAI—3	CAI—4c	CAI—4d	9FPV	CAI—4h	9R8X	CAIX—5b	CAIX—4c	CAIX—4d
PDB ID	9FPT	9FPU	9FPQ	9FPR	9FPS	9R8Y	9R8X	9R8Y	9FPV	9FPW
space group	P12 ₁ 1	P12 ₁ 1	P12 ₁ 1	P12 ₁ 1	P12 ₁ 1	H3	H3	H3	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
unit-cell parameters (<i>a</i> , <i>b</i> , <i>c</i> (Å); α , β , γ (°))	<i>a</i> = 42.1, <i>b</i> = 40.9, <i>c</i> = 71.7, α = γ = 90, β = 104.0	<i>a</i> = 42.0, <i>b</i> = 41.1, <i>c</i> = 71.9, α = γ = 90, β = 104.2	<i>a</i> = 42.3, <i>b</i> = 41.3, <i>c</i> = 72.2, α = γ = 90, β = 104.2	<i>a</i> = 42.3, <i>b</i> = 41.3, <i>c</i> = 72.1, α = γ = 90, β = 104.5	<i>a</i> = 42.3, <i>b</i> = 41.2, <i>c</i> = 72.0, α = γ = 90, β = 104.4	<i>a</i> = <i>b</i> = 151.9, <i>c</i> = 173.7, α = β = 90, γ = 120	<i>a</i> = <i>b</i> = 152.5, <i>c</i> = 172.4, α = β = 90, γ = 120	<i>a</i> = <i>b</i> = 151.9, <i>c</i> = 173.7, α = β = 90, γ = 120	<i>a</i> = 55.1, <i>b</i> = 58.0, <i>c</i> = 160.2, α = β = γ = 90	<i>a</i> = 55.5, <i>b</i> = 58.4, <i>c</i> = 160.7, α = β = γ = 90
resolution range (Å)	69.6–1.2	40.7–1.1	70.0–1.5	40.0–1.5	39.9–1.4	47.8–1.95	47.95–2.0	47.8–1.95	80.1–1.7	40.2–1.9
unique reflections number	75321	87594	38824	41447	47284	108927	101012	108927	57402	42102
<i>R</i> _{merge} , overall (outer shell)	0.04 (0.38)	0.05 (0.31)	0.05 (0.25)	0.05 (0.39)	0.04 (0.38)	0.21 (1.30)	0.08 (1.18)	0.21 (1.30)	0.09 (0.32)	0.12 (0.31)
<i>I</i> / σ overall (outer shell)	16.9 (3.5)	13.7 (3.8)	15.9 (4.9)	19.6 (4.6)	24.2 (4.7)	7.1 (1.5)	21.1 (2.3)	7.1 (1.5)	16.8 (7.4)	13.0 (7.6)
multiplicity overall (outer shell)	6.8 (6.7)	6.6 (6.0)	6.6 (6.4)	7.0 (6.9)	7.0 (6.8)	10.4 (9.9)	10.4 (10.5)	10.4 (9.9)	13.2 (13.1)	13.1 (13.4)
completeness (%) overall (outer shell)	96.7 (94.0)	96.0 (91.3)	99.6 (99.2)	98.7 (97.7)	97.6 (94.9)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100 (100)	100 (100)
Wilson B-factor	12.8	11.4	17.8	15.5	17.0	22.4	28.0	22.4	16.9	22.5
<i>R</i> _{work}	0.14	0.15	0.17	0.14	0.14	0.21	0.17	0.21	0.16	0.19
<i>R</i> _{free}	0.18	0.17	0.20	0.21	0.19	0.25	0.21	0.25	0.19	0.23
RMSD bond lengths (Å)/ angles (°)	0.02/2.13	0.02/2.26	0.01/1.80	0.01/2.04	0.01/2.02	0.01/1.73	0.01/1.81	0.01/1.73	0.01/1.90	0.01/1.57
average B factors (Å ²): all atoms/inhibitors	22.7/17.2	19.6/16.9	21.6/21.9	20.1/15.7	21.8/21.7	30.7/39.4	31.9/50.4	30.7/39.4	21.6/32.5	25.9/35.6
Ramachandran statistics (%) favored/allowed/outliers	97/3/0	97/3/0	97/3/0	96/4/0	96/4/0	94/5/0.32	96/4/0.11	94/5/0.32	chain A: 98/2/0 chain B: 97/3/0	chain A: 97/3/0 chain B: 97/3/0

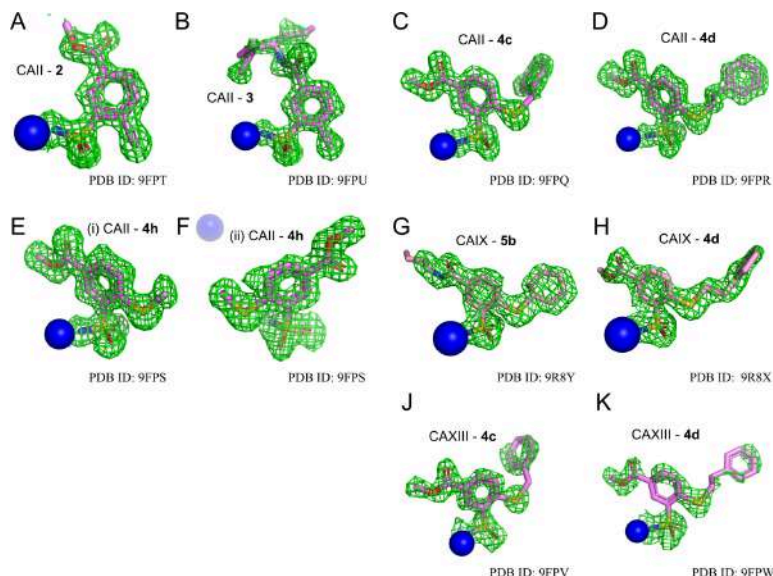


Figure 5. $2F_o - F_c$ omit maps at 3σ in green for the investigated ligands in the active site of CAs. (A) CAII—2 (PDB ID: 9FPT), (B) CAII—3 (PDB ID: 9FPU), (C) CAII—4c (PDB ID: 9FPQ), (D) CAII—4d (PDB ID: 9FPR), (E) and (F) CAII—4h (PDB ID: 9FPS; two ligand molecules were identified, one bound directly to Zn in a conventional position, while the second was seen located nearby toward the edge of the active site), (G) CAIX—5b (PDB ID: 9R8Y), (H) CAIX—4d (PDB ID: 9R8X), (J) CAXIII—4c (PDB ID: 9FPV), (K) CAXIII—4d (PDB ID: 9FPW). Omit maps were taken from a refinement run of the final model without the ligand. Zinc is shown in blue.

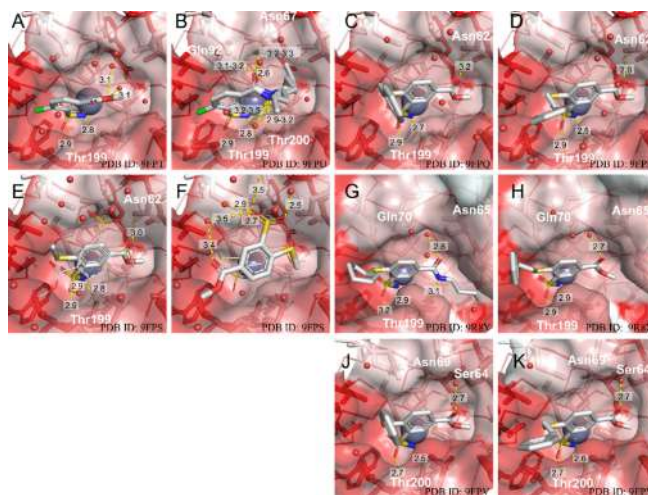


Figure 6. X-ray crystal structures of (A) CAII—2 (PDB ID: 9FPT), (B) CAII—3 (PDB ID: 9FPU), (C) CAII—4c (PDB ID: 9FPQ), (D) CAII—4d (PDB ID: 9FPR), (E) and (F) CAII—4h (PDB ID: 9FPS; two ligand molecules were identified, one bound directly (E) to Zn in a conventional position, while the second is seen located nearby toward the edge (F) of the active site), (G) CAIX—5b (PDB ID: 9R8Y), (H) CAIX—4d (PDB ID: 9R8X), (J) CAXIII—4c (PDB ID: 9FPV), (K) CAXIII—4d (PDB ID: 9FPW). The yellow dashed line represents the hydrogen bond; the distances are given in angstroms. The amino acids directly involved in hydrogen bond formation are labeled. Amino acids are colored according to hydrophobicity:³² the most intense red color represents the most hydrophobic amino acids.

CAII. The amide substituent formed multiple hydrogen bonds with the amino acid side groups, while the ester was stabilized by a network of hydrogen bonds through water molecules. Presumably, a different network of hydrogen bonds pulled the

entire molecule slightly, so a partially rotated benzene ring was observed in the crystal structure. No significant conformational changes were observed in the amino acid chains of the active site of CAII. However, the amide substituent of the compound was

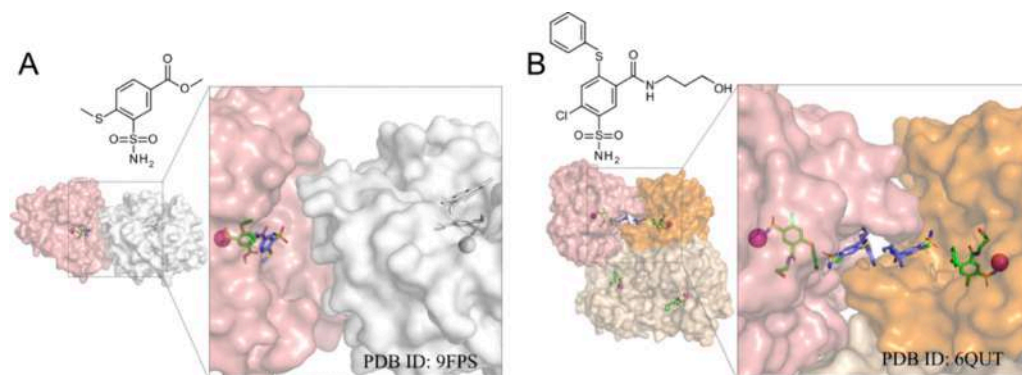


Figure 7. Unusual ligand positions in crystal structures of CAII (PDB ID: 9FPS) and CAIX (PDB ID: 6QUT, published previously²). The chemical structures of the ligands are shown above the crystal structures. The crystal structures represent a view of the monomers and a zoomed-in view of the active site. Zn(II) is shown as a pink sphere. Bound to Zn(II) inhibitor molecules are green and others are blue. (A) Monomer of CAII is shown in surface mode and colored salmon. The symmetric chain is colored white. (B) In the case of CAIX, only the asymmetric unit is represented. Six inhibitor molecules are bound to 4 CAIX chains. Chain C (orange) and chain D (salmon) have two inhibitor molecules, and chains A and B (beige) have one molecule. In this case, the interaction of ligands between separate chains is visible.

not in one fixed position. Instead, 3 alternative conformations of the substituent were identified in the structure. However, it should be noted that this substituent has a poor electron density, likely due to its high flexibility. Consequently, its exact arrangement cannot be determined.

Structures of CAII with Chlorine-Substituted (2, 3) vs Sulfanyl-Substituted (4c, 4d, and 4h). The position of *ortho*-sulfanyl and chlorine-substituted compounds in the CAII active site differed significantly. The entire sulfanyl-substituted molecule was shifted to avoid steric interference but maintained a similar distance between the sulfonamide group's nitrogen and the enzyme's zinc compared to the chlorine-substituted compound. Meanwhile, the studied *ortho*-substituted compounds occupied similar positions, and the ester groups of all three compounds formed a hydrogen bond with Asn62. Notably, the electron density of all inhibitor molecules is well-defined.

Structures of CAII vs CAIX and CAXIII with Bound Sulfanyl-Substituted (4c and 4d or 5b). Figure 6 shows the interactions between CAII with compounds 4c and 4d (Figure 6C,D), CAIX with 4d (Figure 6H), and CAXIII with compounds 4c and 4d (Figure 6J,K). One of the main differences was in the formed hydrogen bonds. The oxygen of the ester group of the compound formed a direct hydrogen bond with Asn62 in the active site of CAII. Meanwhile, a hydrogen bond is formed in the active site of CAIX and CAXIII through a water molecule with Gln70 and Asn65 or Ser64 and Asn69, respectively. Asn62 in CAII, Asn65 in CAIX, and Asn69 in CAXIII differ only by the numbering but correspond to the same position.

Incidentally, compound 5b CAIX also forms the same hydrogen bond through a water molecule, and the entire molecule adopts a similar conformation. Notably, the electron density of the 17 amino acids at the N-terminus of both X-ray crystal structures of CAIX is not defined as expected. These amino acids fold in their 3D structure to form the active site of the protein. This side of the active site is called the hydrophilic part. This suggests that the bound ligand pushes these amino acids to fit fully into the active site. When comparing these structures (PDB ID: 9R8X and 9R8Y) with those existing in the

PDB (e.g., PDB ID: 3iai), the clash between ligand and Tyr7 is seen without changing their arrangement.

It is important to emphasize that both 4c and 4d have very poor electron density in the CAXIII active center, so the arrangement of the compound should be evaluated more cautiously. On the other hand, a comparison of the structures shows that compound 4d is more similarly located in the active sites of CAII and CAXIII, while in CAIX, it is shifted toward the hydrophilic side, which, as mentioned, is not fully visible in the X-ray structure.

Atypical Binding Position. The structure 9FPS was unique because two ligand molecules were identified as bound in the active site. One molecule was bound classically and formed a coordination bond with the zinc ion. The second formed a hydrophobic interaction with the first molecule, and hydrogen bonds with water molecules located toward the edge of the active site. This could be a crystallographic artifact or a secondary interaction with the protein.³³ A previous publication² identified a similar case with CAIX (PDB ID: 6QUT) (Figure 7). In that case, the molecules interacted in the active site and among themselves between the chains in an asymmetric unit. This case with CAII was different because the two chains had no interaction between ligand molecules. The active sites of symmetric chains were not oriented face-to-face. The atoms of the second nonclassical ligand were further than 5 Å from the amino acids of the symmetric chain.

Molecular Docking. To understand the differences in the binding affinities of series 4, 7, and 9 ligands, they were docked into CAII, except for 4e and 4f, containing large, flexible rings that are challenging to dock. Series 4 ligands were also docked into CAIX and CAXIII, to compare with crystal structures and thus assess the docking accuracy. To reduce bias, different receptors (PDB IDs – CAII: 3HS4; CAIX: 6G9U, chain A; CAXIII: 4KNN, chain A) were chosen instead of the new X-ray structures presented in this paper.

To mimic the donor–acceptor bond between the zinc ion and the sulfonamide nitrogen of the ligand, we employed constrained docking using the Smina program.³⁴ The constraint forced the sulfonamide nitrogen to maintain its original position in the X-ray structure. The generated poses were afterward

rescored using the Vinardo scoring function.³⁵ Vina³⁶ and GNINA³⁷ Machine-learning-based scoring functions were found to be inferior to Vinardo for this system, and therefore only the latter was employed for further analysis. The pose with the best rescored affinity that matched the X-ray conformation of the benzenesulfonamide moiety was then chosen as the representative best docked structure (in all cases, for ligand 4 series, it was ranked 1 by Vinardo, except for 4j, where the correct benzenesulfonamide conformation was ranked 3). The best docked poses for CAII are shown in Figure 8A. The Pearson

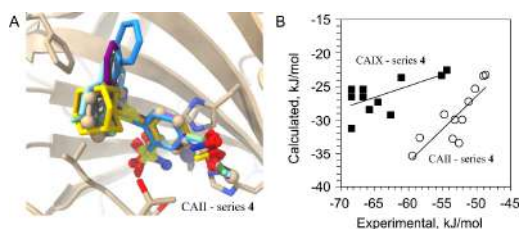


Figure 8. (A) Docked series 4 ligands (see main text for details) superposed onto the X-ray structure of 4c (PDB ID: 9FPQ; rendered using ball-and-stick representation) in a complex with CAII. Note that hydrophobic substituents stack against the Phe131 side chain. (B) Computed ligand binding affinities to CAII (circles) and CAIX (squares) using Vinardo scoring function plotted against the experimental binding affinities—intrinsic Gibbs energy changes. The Pearson correlations R of the plotted points are 0.87 for CAII and 0.66 for CAIX. The large difference between the computed and experimental binding affinities is because the Vinardo scoring function does not take into account the interaction with zinc. However, it can be assumed to be approximately equal for all ligands.

correlation coefficient R with the experimental *intrinsic* binding affinities to CAII for the best docked poses was 0.87. Figure 8B plots the corresponding computed and experimental binding affinities for CAII and CAIX (values listed in Table S1 in the Supporting Information). Notably, for CAII, the scoring function correctly predicts the best binder, 4m, and the four worst binders (4g–j). The panel also shows the correspondence between the experimental and predicted binding affinities for CAIX, with $R = 0.66$.

Validation of the Docking Protocol via Comparison with the X-ray Structures. The proposed docking and scoring protocol was further validated by comparing the predicted docked poses of some of the ligands with their conformations in the newly reported X-ray structures. Figure S4A–C shows the ranked poses of ligands 4c, 4d, and 4h compared against their X-ray conformations in the complex with CAII. The heavy atom Root Mean Square Deviation (RMSD) between the docked and X-ray conformation is 2.03, 3.13, and 0.46 Å, respectively. For the first two ligands, however, the rank 2 conformations are close to the X-ray conformation (4c: 0.61 Å; 4d: 0.67 Å, see Figure S4A,B).

To further validate the chosen docking protocol, 4d and 5b were docked into CAIX (PDB ID: 6G9U, chain A). The best-scored poses after reranking using Vinardo are shown in Figure S5. The corresponding heavy atom RMSDs for 4d and 5b are 2.18 and 0.92 Å, respectively, and despite the RMSD for 4d being above 2 Å, the binding mode is captured by the docking reasonably well, especially since the receptor used for docking was originally bound to a ligand from a chemically different series.

Compounds 4c and 4d were also docked into the CAXIII receptor (PDB ID: 4KNN, chain A). Similarly to CAIX, the docking protocol picked the docked pose with the approximately correct binding mode (with RMSDs equal to 1.90 and 0.95 Å, respectively) (Figure S6).

Docking of Series 7 and 9 Compounds. Since series 4 compounds using the docking protocol described above can reproduce the scaffold rather well and generally seem to stack the hydrophobic substituents reasonably, we applied a similar procedure by docking series 9 ligands into CAII as a model protein, in hopes that it will help to explain a difference between the binding affinities of series 4 and 9. Docking using the same protocol led to comparable binding affinities between the two series (not shown). However, experiments indicate that 9 binds worse by several orders of magnitude, and rescored constrained pose affinities for this series exhibit practically no correlation with the experiment ($R \cong 0.12$). A careful examination of the docking results for series 9 revealed the source of the poor binding affinities and the poor performance of the docking score. Figure 9A displays the best ranking pose of 9a. While the phenyl substituent of 9a stacks well with Phe131 side chain, the two sulfonyl moieties exhibit an apparent clash against each other.

To further explore this clash, we ran simulations of a simple compound 2-(methylsulfonyl)benzenesulfonamide containing sulfonamide and methylsulfonyl groups at the *ortho*-positions on a benzene ring. The lowest energy conformer for this compound was generated using CREST software³⁸ using semiempirical GFN2-xTB wave function,³⁹ and afterward reoptimized in the implicit solvent using the Density Functional Theory (DFT) approach with the GAMESS-US program.⁴⁰ In the lowest energy conformer, sulfonamide nitrogen forms a hydrogen bond with the sulfonyl oxygen at the *ortho* position (Figure 9B, bottom). The approximate conformational energy of the docked conformation was also computed using DFT, keeping frozen the sulfonamide S–N and methylsulfonyl S–C bond torsional angles for the benzene ring, and optimizing the rest of the 2-methylsulfonylbenzenesulfonamide molecule (Figure 9B, top). The freezing of the select torsions was necessary because the DFT calculation lacks the protein environment that keeps them at certain values seen in docking. The energy difference between the lowest energy and the docked conformers in Figure 9B was 53.8 kJ/mol. It became clear that the used scoring function did not report a correct binding energy of the conformation in Figure 9A because the Vinardo scoring function neither takes into account the change of the ligand conformation when going from the solution into the receptor, nor the ligand intramolecular energy (in fact, none of the several built-in scoring functions in Smina do). This means that in the conformation shown in Figure 9A the 2-methylsulfonyl group must rotate to avoid the rather severe clash with the sulfonamide oxygens (Figure 9B, top), and by doing so the hydrophobic substituent rotates away from Phe131, potentially yielding other clashes with the protein and leading to the poor binding energies.

Series 7 compounds are more complex to investigate because two enantiomers of the S atom of sulfinyl exist (Figure 9C,D). We will explore the behavior of sulfinyl-containing series 7 using compound 7a as an example. One of its docked enantiomers, (S)-7a forms sulfonyl-sulfinyl clash (Figure 9C and the top left part of Figure 9D) similar to what we found for 9a. The docked (R)-7a enantiomer is about 1.7 times more stable compared to the (S)- enantiomer because of the lack of the clash (Figure 9D,F). Calculations also show that in the absence of the protein,

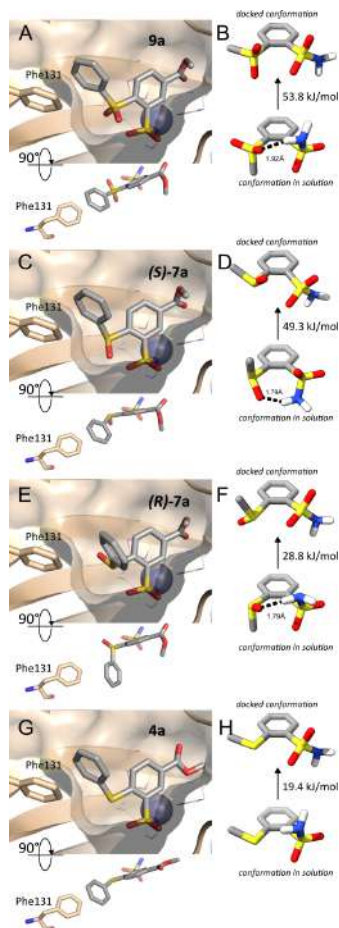


Figure 9. Best ranking docked pose of **9a**, (*S*)-**7a**, (*R*)-**7a**, and **4a** in the active site of CAII as a model protein. (A) Best-ranking docked pose of **9a**. While the phenyl substituent stacks well with the Phe131 side chain, the two SO₂ moieties presumably clash. (B) Two conformations of 2-(methylsulfonyl)benzenesulfonamide: Top: conformation with two benzene-to-S bond dihedrals constrained so that it is similar to the docked conformation in (A); bottom: lowest energy conformation with an intramolecular H-bond that is likely to be found in solution. The energy difference between the two rotamers computed using density functional theory is 53.8 kJ/mol. The fact that the used scoring function does not take into account this difference explains why the calculated affinities (not shown) do not match the experimental values. (C), (E) Best-ranked docked conformations of (*S*)- and (*R*)-**7a**, correspondingly. (D), (F) Top: the conformations 2-(methylsulfonyl)benzenesulfonamide, matching the docked geometries in (C) and (E). Bottom: The lowest energy conformer with an intramolecular H-bond, presumably existing in solution. The docked (*R*)- enantiomer is much more stable due to the lack of clashes present in the docked (*S*)-enantiomer. (G) Best-ranking docked pose of **4a**. (H) Two conformations of 2-methylsulfonylbenzenesulfonamide: Top: the optimized with constraints conformation matching the docked conformation; bottom: The lowest energy conformer. Compared with the sulfinyl- and sulfonyl-forms, the docked conformation for the sulfinyl analog is the most stable, which is also reflected in the binding affinities.

the preferred conformation **7a** (of either chirality) has an intramolecular H-bond (Figure 9D,F, bottom).

For comparison, we also used DFT calculations of the 2-methylsulfonylbenzenesulfonamide molecule to estimate the stability of the docked **4a** (Figure 9G), which does not seem to exhibit clashes. Comparison of Figure 9B,D,F,H shows that the docked conformer **4a** is ~ 1.5 times more stable than the most stable docked conformer of **7a** and nearly 3 times more stable than docked **9a**, corresponding to the experimentally determined binding affinities that are best in series **4**, followed by **7** and **9**. Interestingly, calculations suggest that one reason for the relative stability of docked **4** is the lack of the intramolecular hydrogen bond in solution (Figure 9H, bottom), and this observation could be potentially used in designing better binders for carbonic anhydrases, or even other receptors.

DISCUSSION

Recognition of the pockets on the protein surface by small-molecule ligands is still rather poorly understood, and it is not possible to accurately calculate the thermodynamic binding parameters.⁴¹ Primary sulfonamide compounds are known to inhibit CA isozymes by binding to the catalytic Zn(II). Chemical variation of the remaining molecule often has a limited influence on binding affinity and selectivity. However, some chemical changes made near the sulfonamide group have a significant impact and may change the behavior from strong binder to completely undetectable binding. To investigate this phenomenon, we synthesized *ortho*-substituted benzenesulfonamides, determined their binding affinity to all catalytically active isoforms of human carbonic anhydrases, obtained crystal structures with CAII, CAIX, and CAXIII, and performed molecular docking. Different oxidation forms of the linker at the *ortho* position were used: $-S-$, $-SO-$, and $-SO_2-$. Compounds with $-SO-$ or $-SO_2-$ linker were much weaker binders to any CA isozyme than compounds with the $-S-$ linker. Several reasons could cause this. First, a substituent with a higher oxidation state exhibits a stronger electron-withdrawing effect, thus lowering the p*K*_a of the amine of the sulfonamide group and causing stronger binding. Second, the linker oxygen atoms may prevent the easier rotation of the ligand molecule bonds and prevent conformational changes needed for the ligand to adapt to the protein surface. Third, the oxygen atoms take additional space, and the steric hindrance could be an essential factor in diminishing the binding affinity.

To test the electron-withdrawing effect, the p*K*_a values of compounds were determined experimentally. The p*K*_a of compounds containing $-S-$ linker was almost an entire pH unit higher than compounds with linkers $-SO-$ or $-SO_2-$. Sulfonamides bind to CA in their negatively charged deprotonated form.⁴² Therefore, the lowering of sulfonamide p*K*_a increases the fraction of the binding-ready deprotonated sulfonamide and thus increases the observed affinity. This effect is simply the effect of compound availability in the proper form. To eliminate this misleading increase in affinity, we subtract the fraction effects and calculate the *intrinsic* affinity. The intrinsic dissociation constants of all compounds are provided next to the observed values in the table. In all cases, the intrinsic affinities are higher than the observed ones. The intrinsic values show the 'real' affinities between the binding components in the binding-ready protonation state. These values should be used in drug design to explain the structure–function relationship of the compound effects and not the experimentally observed ones. However, the experimentally observed affinity values show the

affinities observed by any experimental technique, and the values are biologically relevant. They should be used to calculate the bound fractions and drug effect at particular conditions.

Furthermore, the compounds with a higher oxidation state (7a, b, h, i, k–m and 9a–d, h, i, l, m) did not bind or bind with lower affinity than compounds 4a–d, f–m. Even the compounds 7h and 9h with the smallest methyl substituent did not match the sulfanyl compound 4h in affinity. The main reason for this was the allowed conformations of the compounds, which were calculated using quantum Density Functional Theory (DFT). The oxygen atoms limited both the flexibility of the molecule in finding the optimal position and acted as a steric hindrance. The present study shows that the *ortho* modifications had a more significant effect than the *para* variations.²¹

The highest affinity for CAIX in this study was exhibited by compound 4b (methyl 4-cyclohexylsulfanyl-3-sulfamoylbenzoate). Previously we have designed similar compounds bearing 2-chloro or 2-bromo substituents that exhibited slightly higher affinity for CAIX.²⁰ Here, we have intentionally omitted the halogen atoms to synthesize the *ortho*-substituted compounds more easily and explore the binding without halogens. It was also easier to synthesize compounds with an amide substituent in the *meta* position. The length of these amide-substituted compounds was of limited importance in the previous study but had a significant influence on this study, likely due to reasons of steric hindrance.

Sulfonamide compounds usually bind to CA isozymes by forming a coordination bond between the Zn(II) and sulfonamide amino groups. This bond significantly increases the affinity, but is not necessary for binding to occur.⁴³ Removal of the metal or change of the metal with another demonstrated that the coordination bond contribution is additive and metal-dependent. Strongly binding compounds like brinzolamide to CAII also bind to the Zn-free apoCAII. The energy contribution of the coordination bond could be determined by using a metal-exchange approach.

In the crystal structure of compound 4h bound to CAII, two well-resolved compound molecules were bound in the active site. The first was bound in a conventional way forming a coordination bond with the Zn(II), but the second was bound to the protein residues without forming a coordination bond with the Zn(II). This second compound molecule did not bind solely due to crystal-forming effects because it did not bind to the second protein molecule. Therefore, the binding of the second molecule is likely not an artifact. However, the presence of the second molecule in the crystal structure does not mean that it is bound as strongly, nor that we can measure its binding affinity experimentally. Binding assays showed that the stoichiometry here was 1:1, and the affinity of the second molecule was likely weak compared to the first one. Compound concentration in the crystallization experiments was relatively high, millimolar, thus the second molecule could be seen in the crystal structure even if the K_d was in the millimolar range and therefore not interfering in any binding assays. This also means that the second molecule is biologically irrelevant and would not play an essential role in drug design.

CONCLUSIONS

The strategy to acidify the pK_a of *ortho* sulfanyl-substituted benzenesulfonamides by oxidation, with the goal of increased affinity to CA, yielded an unexpected drop of affinity in the order of hundreds of thousands of times. Small changes in chemical

structures influenced the flexibility of the molecule's substituents and steric restrictions on interactions with proteins. Furthermore, some minor changes led to the discovery of novel CAIX inhibitors with high affinity and selectivity.

EXPERIMENTAL SECTION

Organic Synthesis. All starting materials and reagents were commercial products used without further purification. Melting points of the compounds were determined in open capillaries on a Thermo Scientific 9100 Series and are uncorrected. ¹H and ¹³C NMR spectra (Figure S7) were recorded on a (400 and 100 MHz, respectively) spectrometer in DMSO-*d*₆ using residual DMSO signals (2.52 and 40.21 ppm for ¹H and ¹³C NMR spectra, respectively) as the internal standard. TLC was performed with silica gel 60 F254 aluminum plates (Merck) and visualized with UV light. Column chromatography used silica gel 60 (0.040–0.063 mm, Merck). High-resolution mass spectra (HRMS) were recorded by an Agilent TOF 6230 equipped with an Agilent Infinity 1260 HPLC system (Agilent Technologies). HPLC verified the purity of final compounds to be >95% (Figure S8) using the Agilent Infinity 1260 instrument with a ZORBAX Eclipse Plus C18 Rapid Resolution 4.6 × 100 mm 3.5 μm column, ZORBAX Eclipse Plus-C18 4.6 × 12.5 mm 5.0 μm analytical guard column, eluents A – 20 mM ammonium acetate (pH = 6.9 of unadjusted solution) and B – 100% MeOH was used. HPLC gradient elution with a flow rate of 0.80 mL/min was used, B gradient: 0–10 min 55–95%, 10–14 min 95%. UV detection was recorded at 254 nm. Figure S9 contains ESI-MS spectra of representative compounds.

Methyl 4-Chloro-3-sulfamoylbenzoate (2). 4-chloro-3-sulfamoylbenzoic acid 1 (1.78 g, 7.55 mmol, Sigma-Aldrich) was refluxed in MeOH (30 mL) with concentrated H₂SO₄ (0.3 mL) for 8 h. The reaction mixture was concentrated under reduced pressure. Residue was filtered, washed with H₂O and crystallized from H₂O:MeOH(4:1). Yield: 1.70 g, 91%, mp 129–130 °C (Literature⁴⁴ mp 124–125 °C).

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 3.91 (s, 3H, CH₃O), 7.80 (d, *J* = 8.4 Hz, 1H, ArH), 7.84 (s, 2H, SO₂NH₂), 8.14 (dd, *J* = 8.4 Hz, *J* = 2.1 Hz, 1H, ArH), 8.52 (d, *J* = 2.1 Hz, 1H, ArH).

***N*-Butyl-4-chloro-3-sulfamoylbenzamide (3).** The mixture of 4-chloro-3-sulfamoylbenzoic acid 1 (500 mg, 2.12 mmol), SOCl₂ (0.616 mL, 8.48 mmol), and one drop DMF in toluene (6.0 mL) was refluxed for four h. Excess SOCl₂ and toluene were removed by distillation under reduced pressure. The crude acid chloride was dissolved in THF (20 mL) and added dropwise to a solution of *N*-butylamine (0.591 mL, 6.0 mmol) in THF (20 mL) at 0 °C and allowed stirring for one h. The mixture was warmed to room temperature and stirred for another 4 h. THF was removed under reduced pressure. The crude product was crystallized from H₂O. Yield: 492 mg, 80%, mp 172–173 °C (lit.⁴⁵ mp 171–172 °C).

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 0.92 (t, *J* = 7.3 Hz, 3H, CH₃CH₂), 1.37 (sextet, *J* = 7.3 Hz, 2H, CH₂CH₃), 1.54 (quint, *J* = 7.3 Hz, 2H, CH₂CH₂CH₃), 3.29 (q, *J* = 6.8 Hz, 2H, CH₂NH), 7.72 (s, 2H, SO₂NH₂), 7.77 (d, *J* = 8.2 Hz, 1H, ArH), 8.04 (dd, *J* = 8.2 Hz, *J* = 2.0 Hz, 1H, ArH), 8.45 (d, *J* = 2.0 Hz, 1H, ArH), 8.77 (t, *J* = 5.5 Hz, 1H, NH).

General Procedure for the syntheses of 4a–d, f, g, i–m. The mixture of methyl 4-chloro-3-sulfamoylbenzoate 2 (285 mg, 1.14 mmol), DMF (4.0 mL), appropriate thiol (1.25 mmol), and K₂CO₃ (630 mg, 4.56 mmol) was heated at 80 °C for 4–6 h in an inert atmosphere (argon). The mixture was cooled to room temperature, and 10 mL of H₂O was added. The product was extracted with EtOAc (3 × 8 mL). The organic layer was washed with H₂O, dried over anhydrous MgSO₄, filtered, and concentrated under reduced pressure.

Methyl 4-Phenylsulfanyl-3-sulfamoylbenzoate (4a). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 6:1). Yield: 303 mg, 82%, mp 155–156 °C. (lit.⁴⁶ mp 154–157).

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 3.86 (s, 3H, CH₃O), 7.00 (d, *J* = 8.4 Hz, 1H, ArH), 7.55–7.62 (m, 5H, ArH), 7.74 (s, 2H, SO₂NH₂), 7.93 (dd, *J* = 8.4 Hz, *J* = 1.9 Hz, 1H, ArH), 8.46 (t, *J* = 1.9 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 52.9, 126.8, 129.0, 129.2, 130.5, 130.8, 131.0, 132.6, 135.5, 140.8, 144.2, 165.4.

HRMS calcd for C₁₄H₁₃NO₄S₂ [(M+H)⁺]: 324.0359, found: 324.0354.

Methyl 4-Cyclohexylsulfanyl-3-sulfamoyl-benzoate (4b). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 7:1). Yield: 315 mg, 84%, mp 119–120 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 1.26–1.99 (m, 10H, CH cyclohexyl), 3.61 (m, 1H, CHS), 3.88 (s, 3H, CH₃O), 7.47 (s, 2H, SO₂NH₂), 7.74 (d, *J* = 8.4 Hz, 1H, ArH), 8.04 (dd, *J* = 8.3 Hz, *J* = 1.9 Hz, 1H, ArH), 8.43 (d, *J* = 1.9 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 25.7, 25.8, 32.5, 44.4, 52.9, 100.0, 125.9, 129.1, 132.3, 141.7, 142.8, 165.6.

HRMS calcd for C₁₄H₁₉NO₄S₂ [(M+H)⁺]: 330.0828, found: 330.0833.

Methyl 4-Benzylsulfanyl-3-sulfamoyl-benzoate (4c). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 6:1). Yield: 211 mg, 55%, mp 135–136 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 3.88 (s, 3H, CH₃O), 4.43 (s, 2H, CH₂S), 7.29–7.37 (m, 3H, ArH), 7.51–7.53 (m, 2H, ArH), 7.58 (s, 2H, SO₂NH₂), 7.74 (d, *J* = 8.3 Hz, 1H, ArH), 8.02 (dt, *J* = 8.3 Hz, *J* = 1.7 Hz, 1H, ArH), 8.42 (t, *J* = 2.2 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 36.4, 52.9, 125.9, 127.7, 128.0, 128.9, 129.0, 129.8, 132.2, 136.0, 140.7, 143.7, 165.6.

HRMS calcd for C₁₅H₁₅NO₄S₂ [(M+H)⁺]: 338.0515, found: 338.0517.

Methyl 4-(2-Phenylethylsulfanyl)-3-sulfamoyl-benzoate (4d). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 5:1). Yield: 332 mg, 83%, mp 146–147 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 2.99 (t, *J* = 7.5 Hz, 2H, CH₂Ph), 3.41 (t, *J* = 7.5 Hz, 2H, CH₂S), 3.88 (s, 3H, CH₃O), 7.23–7.34 (m, 5H, ArH), 7.51 (s, 2H, SO₂NH₂), 7.73 (d, *J* = 8.4 Hz, 1H, ArH), 8.05 (dd, *J* = 8.4 Hz, *J* = 1.8 Hz, 1H, ArH), 8.43 (d, *J* = 1.8 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 33.4, 34.3, 52.9, 125.8, 126.9, 127.8, 128.8, 128.9, 129.1, 132.4, 140.2, 141.0, 143.6, 165.6.

HRMS calcd for C₁₆H₁₇NO₄S₂ [(M+H)⁺]: 352.0672, found: 352.0675.

Methyl 4-Cyclododecylsulfanyl-3-sulfamoyl-benzoate (4f). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 10:1). Yield: 203 mg, 43%, mp 167–168 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 1.32–1.59 (m, 20H, CH cyclododecyl), 1.76 (m, 2H, CH cyclododecyl), 3.66 (m, 1H, CHS), 3.88 (s, 3H, CH₃O), 7.47 (s, 2H, SO₂NH₂), 7.70 (d, *J* = 8.4 Hz, 1H, ArH), 8.07 (dd, *J* = 8.4 Hz, *J* = 2.0 Hz, 1H, ArH), 8.44 (d, *J* = 2.0 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 22.0, 23.2, 23.3, 24.0, 24.2, 29.2, 43.5, 52.9, 125.9, 129.0, 129.2, 132.3, 142.0, 143.2, 165.6.

HRMS calcd for C₂₀H₃₁NO₄S₂ [(M+H)⁺]: 414.1767, found: 414.1773.

Methyl 4-(1-Adamantylsulfanyl)-3-sulfamoyl-benzoate (4g). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 10:1). Yield: 173 mg, 40%, mp 189–190 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 1.63 (br s, 6H, CH adamantanyl), 1.98 (br s, 6H, CH adamantanyl), 2.00 (br s, 3H, CH adamantanyl), 3.90 (s, 3H, OCH₃), 7.40 (s, 2H, SO₂NH₂), 7.87 (d, *J* = 8.0 Hz, 1H, ArH), 8.08 (dd, *J* = 8.0 Hz, *J* = 2.0 Hz, 1H, ArH), 8.53 (d, *J* = 2.0 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 30.1, 35.9, 43.8, 52.2, 53.0, 128.8, 128.9, 131.8, 137.7, 137.9, 147.0, 165.5.

HRMS calcd for C₁₈H₂₃NO₄S₂ [(M+H)⁺]: 382.1141, found: 382.1144.

Methyl 4-Propylsulfanyl-3-sulfamoyl-benzoate (4i). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 5:1). Yield: 202 mg, 61%, mp 111–112 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 1.05 (t, *J* = 7.3 Hz, 3H, CH₃CH₂), 1.71 (quint, *J* = 7.3 Hz, 2H, CH₂CH₂), 3.11 (t, *J* = 7.3 Hz, 2H, CH₂S), 3.88 (s, 3H, CH₃O), 7.54 (s, 2H, SO₂NH₂), 7.65 (d, *J* = 8.3 Hz, 1H, ArH), 8.04 (dd, *J* = 8.3 Hz, *J* = 1.9 Hz, 1H, ArH), 8.43 (d, *J* = 1.9 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 13.8, 21.7, 33.9, 52.8, 125.6, 127.5, 128.9, 132.3, 140.9, 144.1, 165.6.

HRMS calcd for C₁₁H₁₅NO₄S₂ [(M+H)⁺]: 290.0515, found: 290.0519.

Methyl 4-tert-Butylsulfanyl-3-sulfamoyl-benzoate (4j). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 5:1). Yield: 214 mg, 62%, mp 124–125 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 1.43 (s, 9H, CH₃), 3.90 (s, 3H, CH₃O), 7.43 (s, 2H, SO₂NH₂), 7.91 (d, *J* = 8.2 Hz, 1H, ArH), 8.11 (dd, *J* = 8.2 Hz, *J* = 2.0 Hz, 1H, ArH), 8.52 (d, *J* = 2.0 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 31.6, 49.6, 53.0, 128.4, 128.9, 132.0, 136.2, 140.0, 146.1, 165.5.

HRMS calcd for C₁₂H₁₇NO₄S₂ [(M+H)⁺]: 304.0672, found: 304.0671.

Methyl 4-Isopentylsulfanyl-3-sulfamoyl-benzoate (4k). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 5:1). Yield: 265 mg, 74%, mp 92–93 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 0.94 (d, *J* = 6.6 Hz, 6H, CH₃), 1.57 (td, *J* = 7.8 Hz, *J* = 6.8 Hz, 2H, CH₂CH), 1.76 (m, 1H, CH), 3.11 (t, *J* = 7.8 Hz, 2H, CH₂S), 3.88 (s, 3H, CH₃O), 7.52 (s, 2H, SO₂NH₂), 7.67 (d, *J* = 8.4 Hz, ArH), 8.05 (dd, *J* = 8.4 Hz, *J* = 2.0 Hz, 1H, ArH), 8.42 (d, *J* = 2.0 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 22.6, 27.5, 30.2, 36.9, 52.8, 125.6, 127.5, 128.9, 132.3, 140.9, 144.1, 165.6.

HRMS calcd for C₁₃H₁₉NO₄S₂ [(M+H)⁺]: 318.0828, found: 318.0833.

Methyl 4-Cyclopentylsulfanyl-3-sulfamoyl-benzoate (4l). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 8:1). Yield: 259 mg, 72%, mp 106–107 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 1.55–1.65 (m, 4H, CH cyclopentyl), 1.73–1.78 (m, 2H, CH cyclopentyl), 2.18–2.23 (m, 4H, CH cyclopentyl), 3.84–3.92 (m, 4H, CH₃O and CHS), 7.50 (s, 2H, SO₂NH₂), 7.72 (d, *J* = 8.4 Hz, 1H, ArH), 8.04 (dd, *J* = 8.4 Hz, *J* = 2.0 Hz, 1H, ArH), 8.42 (d, *J* = 2.0 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 25.1, 33.3, 43.9, 52.8, 125.5, 128.4, 128.9, 132.3, 140.8, 144.7, 165.6.

HRMS calcd for C₁₂H₁₇NO₄S₂ [(M+H)⁺]: 316.0672, found: 316.0677.

Methyl 4-(1-Naphthylsulfanyl)-3-sulfamoyl-benzoate (4m). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 8:1). Yield: 281 mg, 66%, mp 187–188 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 3.83 (s, 3H, CH₃O), 6.64 (d, *J* = 8.4 Hz, 1H, ArH), 7.54–7.64 (m, 2H, ArH), 7.69 (m, 1H, ArH), 7.74 (dd, *J* = 8.4 Hz, *J* = 2.0 Hz, 1H, ArH), 7.93 (s, 2H, SO₂NH₂), 8.04–8.15 (m, 3H, ArH), 8.22 (d, *J* = 8.3 Hz, 1H, ArH), 8.49 (d, *J* = 2.0 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 52.9, 125.6, 126.7, 126.9, 127.3, 127.5, 128.4, 128.5, 129.1, 129.6, 132.2, 132.5, 133.8, 134.8, 136.8, 140.4, 143.9, 165.4.

HRMS calcd for C₁₈H₁₅NO₄S₂ [(M+H)⁺]: 374.0515, found: 374.0511.

Methyl 4-Methylsulfanyl-3-sulfamoyl-benzoate (4h). The mixture of methyl 4-chloro-3-sulfamoylbenzoate **2** (308 mg, 1.23 mmol),

DMSO (4.0 mL), sodium methanethiolate (259 mg, 3.69 mmol), and K₂CO₃ (509 mg, 3.69 mmol) was heated at 80 °C temperature for 12 h in inert atmosphere (argon). The mixture was cooled to room temperature and 20 mL brine was added. The product was extracted with EtOAc (3 × 10 mL). The organic layer was washed with H₂O, dried over anhydrous MgSO₄, filtered, and concentrated under reduced pressure. The product was crystallized from H₂O:MeOH(4:1). Yield: 203 mg, 64%, mp 157–158 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 2.59 (s, 3H, CH₃S), 3.89 (s, 3H, CH₃O), 7.54 (s, 2H, SO₂NH₂), 7.61 (d, *J* = 8.3 Hz, 1H, ArH), 8.07 (dd, *J* = 8.3 Hz, *J* = 1.4 Hz, 1H, ArH), 8.42 (d, *J* = 1.4 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 15.5, 52.9, 125.4, 126.6, 128.8, 132.4, 140.4, 145.1, 165.7.

HRMS calcd for C₉H₁₁NO₄S₂ [(M+H)⁺]: 262.0202, found: 262.0199.

General Procedure for the Syntheses of 5a–d, f–g, j–m. The mixture of *N*-butyl-4-chloro-3-sulfamoylbenzamide **3** (111 mg, 0.381

mmol), DMF (4.0 mL), appropriate thiol (0.572 mmol), and K_2CO_3 (210 mg, 1.52 mmol) was heated at 80 °C temperature for 12 h in an inert atmosphere (argon). The mixture was cooled to room temperature and 20 mL H_2O was added. The product was extracted with EtOAc (3 × 8 mL). The organic layer was washed with H_2O , dried over anhydrous $MgSO_4$, filtered, and concentrated under reduced pressure.

N-Butyl-4-phenylsulfanyl-3-sulfamoyl-benzamide (5a). The product was purified by flash chromatography on silica gel ($CHCl_3$:EtOAc, 1:1). Yield: 71 mg, 51%, mp 109–110 °C.

1H NMR (400 MHz, $DMSO-d_6$) δ ppm: 0.90 (t, J = 7.3 Hz, 3H, CH_3), 1.33 (sext, J = 7.3 Hz, 2H, $CH_2CH_2CH_3$), 1.52 (quint, J = 7.3 Hz, 2H, $CH_2CH_2CH_3$), 3.27 (dt, J = 6.9 Hz, J = 5.7 Hz, 2H, CH_2NH), 6.99 (d, J = 8.3 Hz, 1H, ArH), 7.51–7.58 (m, 5H, ArH), 7.60 (s, 2H, SO_2NH_2), 7.82 (dd, J = 8.3 Hz, J = 2.0 Hz, 1H, ArH), 8.38 (d, J = 2.0 Hz, 1H, ArH), 8.59 (t, J = 5.7 Hz, 1H, NH).

^{13}C NMR (100 MHz, $DMSO-d_6$) δ ppm: 14.2, 20.1, 31.6, 39.4, 127.5, 129.6, 129.9, 130.5, 130.6, 132.3, 132.4, 134.9, 140.7, 141.2, 164.9.

HRMS calcd for $C_{17}H_{20}N_2O_3S_2$ [(M+H) $^+$]: 365.0988, found: 365.0985.

N-Butyl-4-cyclohexylsulfanyl-3-sulfamoyl-benzamide (5b). The product was purified by flash chromatography on silica gel ($CHCl_3$:EtOAc, 1:1). Yield: 70 mg, 50%, mp 114–115 °C.

1H NMR (400 MHz, $DMSO-d_6$) δ ppm: 0.92 (t, J = 7.3 Hz, 3H, CH_3), 1.23–1.63 (m, 10H, CH cyclohexyl and butyl), 1.72 (m, 2H, CH cyclohexyl), 1.90–1.96 (m, 2H, CH cyclohexyl), 3.29 (q, J = 6.6 Hz, 2H, CH_2NH), 3.55–3.60 (m, 1H, CHS), 7.34 (s, 2H, SO_2NH_2), 7.70 (d, J = 8.2 Hz, 1H, ArH), 7.96 (dd, J = 8.2 Hz, J = 1.2 Hz, 1H, ArH), 8.36 (d, J = 1.2 Hz, 1H, ArH), 8.64 (t, J = 5.3 Hz, 1H, NH).

^{13}C NMR (100 MHz, $DMSO-d_6$) δ ppm: 14.2, 20.1, 25.7, 25.8, 31.6, 32.5, 39.4, 44.8, 127.6, 129.7, 130.3, 131.6, 139.0, 142.1, 165.1.

HRMS calcd for $C_{17}H_{26}N_2O_3S_2$ [(M+H) $^+$]: 371.1458, found: 371.1463.

4-Benzylsulfanyl-N-butyl-3-sulfamoyl-benzamide (5c). The product was purified by flash chromatography on silica gel ($CHCl_3$:EtOAc, 1:1). Yield: 65 mg, 45%, mp 185–186 °C.

1H NMR (400 MHz, $DMSO-d_6$) δ ppm: 0.91 (t, J = 7.2 Hz, 3H, CH_3), 1.36 (sext, J = 7.2 Hz, 2H, CH_2CH_3), 1.53 (quint, J = 7.2 Hz, 2H, $CH_2CH_2CH_3$), 3.28 (q, J = 6.9 Hz, 2H, CH_2NH), 4.40 (s, 2H, CH_2S), 7.29 (t, J = 7.1 Hz, 1H, ArH), 7.36 (t, J = 7.1 Hz, 2H, ArH), 7.44 (s, 2H, SO_2NH_2), 7.52 (d, J = 7.1 Hz, 2H, ArH), 7.66 (d, J = 8.3 Hz, 1H, ArH), 7.94 (dd, J = 8.3 Hz, J = 1.9 Hz, 1H, ArH), 8.36 (d, J = 1.9 Hz, 1H, ArH), 8.62 (t, J = 5.6 Hz, 1H, NH).

^{13}C NMR (100 MHz, $DMSO-d_6$) δ ppm: 14.2, 20.1, 31.6, 36.5, 39.4, 127.6, 127.7, 127.8, 128.9, 129.7, 130.2, 131.2, 136.5, 140.2, 140.7, 165.0.

HRMS calcd for $C_{18}H_{22}N_2O_3S_2$ [(M+H) $^+$]: 379.1145, found: 379.1140.

N-Butyl-4-(2-phenylethylsulfanyl)-3-sulfamoyl-benzamide (5d). The product was purified by flash chromatography on silica gel ($CHCl_3$:EtOAc, 1:1). Yield: 95 mg, 63%, mp 116–117 °C.

1H NMR (400 MHz, $DMSO-d_6$) δ ppm: 0.92 (t, J = 7.3 Hz, 3H, CH_3), 1.37 (sext, J = 7.3 Hz, 2H, CH_2CH_3), 1.55 (quint, J = 7.3 Hz, 2H, $CH_2CH_2CH_3$), 3.28 (q, J = 7.3 Hz, 2H, CH_2NH), 3.28 (q, J = 6.6 Hz, 2H, CH_2NH), 3.34–3.38 (m, 2H, CH_2S), 7.25–7.27 (m, 1H, ArH), 7.31–7.33 (m, 4H, ArH), 7.35 (s, 2H, SO_2NH_2), 7.69 (d, J = 8.3 Hz, 1H, ArH), 7.99 (d, J = 8.3 Hz, 1H, ArH), 8.38 (s, 1H, ArH), 8.65 (t, J = 5.4 Hz, 1H, NH).

^{13}C NMR (100 MHz, $DMSO-d_6$) δ ppm: 14.2, 20.1, 31.6, 33.7, 34.4, 39.4, 126.9, 127.5, 127.9, 128.9, 129.0, 130.4, 131.3, 140.1, 140.3, 141.2, 165.1.

HRMS calcd for $C_{19}H_{24}N_2O_3S_2$ [(M+H) $^+$]: 393.1301, found: 393.1297.

N-Butyl-4-cyclododecylsulfanyl-3-sulfamoyl-benzamide (5f). The product was purified by flash chromatography on silica gel ($CHCl_3$:EtOAc, 3:1). Yield: 76 mg, 44%, mp 166–168 °C.

1H NMR (400 MHz, $DMSO-d_6$) δ ppm: 0.93 (t, J = 7.3 Hz, 3H, CH_3), 1.24–1.59 (m, 24H, $CH_2CH_2CH_3$ and CH cyclododecyl), 1.72–1.75 (m, 2H, CH cyclododecyl), 3.29 (dt, J = 6.8 Hz, J = 5.6 Hz, 2H, CH_2NH), 3.68 (s, 1H, CHS), 7.33 (s, 2H, SO_2NH_2), 7.66 (d, J =

8.3 Hz, 1H, ArH), 7.99 (dd, J = 8.3 Hz, J = 2.0 Hz, 1H, ArH), 8.37 (d, J = 2.0 Hz, 1H, ArH), 8.64 (t, J = 5.6 Hz, 1H, NH).

^{13}C NMR (100 MHz, $DMSO-d_6$) δ ppm: 14.2, 20.1, 22.1, 23.4, 23.8, 23.9, 24.2, 29.4, 31.6, 39.4, 43.7, 127.6, 129.5, 130.3, 131.6, 139.7, 142.3, 165.1.

HRMS calcd for $C_{23}H_{38}N_2O_3S_2$ [(M+H) $^+$]: 455.2397, found: 455.2399.

4-(1-Adamantylsulfanyl)-N-butyl-3-sulfamoyl-benzamide (5g). The product was purified by flash chromatography on silica gel ($CHCl_3$:EtOAc, 1:1). Yield: 64 mg, 40%, mp 187–189 °C.

1H NMR (400 MHz, $DMSO-d_6$) δ ppm: 0.91 (t, J = 6.8 Hz, 3H, CH_3), 1.34 (sext, J = 7.2 Hz, 2H, CH_2CH_3), 1.52 (quint, J = 7.2 Hz, 2H, $NHCH_2CH_2$), 1.63 (br s, 6H, CH adamantanyl), 1.95 (br s, 6H, CH adamantanyl), 2.01 (br s, 3H, CH adamantanyl), 3.28 (q, J = 6.8 Hz, 2H, $NHCH_2$), 7.28 (s, 2H, SO_2NH_2), 7.78 (d, J = 8.0 Hz, 1H, ArH), 7.98 (dd, J = 8.0 Hz, J = 2.0 Hz, 1H, ArH), 8.44 (d, J = 2.0 Hz, 1H, ArH), 8.71 (t, J = 5.6 Hz, 1H, NH).

^{13}C NMR (100 MHz, $DMSO-d_6$) δ ppm: 14.2, 20.1, 30.1, 31.6, 36.0, 39.5, 43.9, 51.8, 127.4, 129.8, 134.0, 134.5, 138.2, 147.3, 165.2.

HRMS calcd for $C_{27}H_{30}N_2O_3S_2$ [(M+H) $^+$]: 423.1771, found: 423.1777.

N-Butyl-4-tert-butylsulfanyl-3-sulfamoyl-benzamide (5j). The product was purified by flash chromatography on silica gel ($CHCl_3$:EtOAc, 1:1). Yield: 55 mg, 42%, mp 150–152 °C.

1H NMR (400 MHz, $DMSO-d_6$) δ ppm: 0.91 (t, J = 7.6 Hz, 3H, CH_3), 1.30–1.37 (m, 2H, $CH_2CH_2CH_3$), 1.41 (s, 9H, C(CH_3) $_3$), 1.52 (quint, J = 7.2 Hz, 2H, $NHCH_2CH_2$), 3.28 (q, J = 6.0 Hz, 2H, $NHCH_2$), 7.31 (s, 2H, SO_2NH_2), 7.82 (d, J = 7.6 Hz, 1H, ArH), 7.99 (dd, J = 8.0 Hz, J = 1.6 Hz, 1H, ArH), 8.44 (d, J = 2.0 Hz, 1H, ArH), 8.71 (t, J = 5.8 Hz, 1H, NH).

^{13}C NMR (100 MHz, $DMSO-d_6$) δ ppm: 14.2, 20.1, 31.6, 31.7, 39.5, 49.4, 127.5, 130.1, 134.3, 136.2, 137.0, 146.7, 165.2.

HRMS calcd for $C_{15}H_{24}N_2O_3S_2$ [(M+H) $^+$]: 345.1301, found: 345.1295.

N-Butyl-4-isopentylsulfanyl-3-sulfamoyl-benzamide (5k). The product was purified by flash chromatography on silica gel ($CHCl_3$:EtOAc, 1:1). Yield: 98 mg, 72%, mp 130–132 °C.

1H NMR (400 MHz, $DMSO-d_6$) δ ppm: 0.89–0.94 (m, 9H, CH_3), 1.33 (sext, J = 7.2 Hz, 2H, $CH_2CH_2CH_3$), 1.48–1.57 (m, 4H, $NHCH_2CH_2$ and SCH_2CH_3), 1.75 (hept, J = 7.6 Hz, 1H, CH(CH_3) $_2$), 3.09 (t, J = 7.8 Hz, 2H, SCH_2), 3.30 (q, J = 6.4 Hz, 2H, $NHCH_2$), 7.38 (s, 2H, SO_2NH_2), 7.61 (d, J = 8.4 Hz, 1H, ArH), 7.97 (dd, J = 8.3 Hz, J = 1.6 Hz, 1H, ArH), 8.36 (d, J = 1.7 Hz, 1H, ArH), 8.62 (t, J = 5.3 Hz, 1H, NH).

^{13}C NMR (100 MHz, $DMSO-d_6$) δ ppm: 14.2, 20.1, 22.6, 27.5, 30.4, 31.7, 37.2, 39.4, 127.6(2C), 130.4, 131.1, 140.6, 141.0, 165.1.

HRMS calcd for $C_{16}H_{26}N_2O_3S_2$ [(M+H) $^+$]: 359.1458, found: 359.1464.

N-Butyl-4-cyclopentylsulfanyl-3-sulfamoyl-benzamide (5l). The product was purified by flash chromatography on silica gel ($CHCl_3$:EtOAc, 1:1). Yield: 84 mg, 62%, mp 122–124 °C.

1H NMR (400 MHz, $DMSO-d_6$) δ ppm: 0.91 (t, J = 7.2 Hz, 3H, CH_3), 1.33 (sext, J = 7.2 Hz, 2H, $CH_2CH_2CH_3$), 1.48–1.67 (m, 6H, $NHCH_2CH_2$ and CH cyclopentyl), 1.71–1.81 (m, 2H, CH cyclopentyl), 2.12–2.21 (m, 2H, CH cyclopentyl), 3.28 (q, J = 6.4 Hz, 2H, $NHCH_2$), 3.92 (quint, J = 6.0 Hz, 1H, SCH), 7.36 (s, 2H, SO_2NH_2), 7.66 (d, J = 8.4 Hz, 1H, ArH), 7.96 (d, J = 7.6 Hz, 1H, ArH), 8.36 (s, 1H, ArH), 8.61 (t, J = 5.2 Hz, 1H, NH).

^{13}C NMR (100 MHz, $DMSO-d_6$) δ ppm: 14.2, 20.1, 25.1, 31.7, 33.3, 39.4, 44.2, 127.6, 128.5, 130.3, 131.0, 141.0, 141.2, 165.1.

HRMS calcd for $C_{16}H_{24}N_2O_3S_2$ [(M+H) $^+$]: 357.1301, found: 357.1307.

N-Butyl-4-(1-naphthylsulfanyl)-3-sulfamoyl-benzamide (5m). The product was purified by flash chromatography on silica gel ($CHCl_3$:EtOAc, 1:1). Yield: 96 mg, 61%, mp 216–218 °C.

1H NMR (400 MHz, $DMSO-d_6$) δ ppm: (DMSO- D_6) 0.86 (t, J = 7.6 Hz, 3H, CH_3), 1.28 (sext, J = 7.6 Hz, 2H, $CH_2CH_2CH_3$), 1.45 (quint, J = 7.2 Hz, 2H, $NHCH_2CH_2$), 3.21 (q, J = 6.8 Hz, 2H, $NHCH_2$), 6.57 (d, J = 8.4 Hz, 1H, ArH), 7.53–7.63 (m, 3H, ArH and naphthyl-H), 7.68 (t, J = 8.0 Hz, 1H, naphthyl-H), 7.83 (s, 2H,

SO_2NH_2), 8.03 (d, $J = 6.8$ Hz, 1H, naphthyl-H), 8.08 (d, $J = 8.0$ Hz, 1H, naphthyl-H), 8.15 (d, $J = 8.4$ Hz, 1H, naphthyl-H), 8.19 (d, $J = 8.4$ Hz, 1H, naphthyl-H), 8.41 (d, $J = 2.0$ Hz, 1H, ArH), 8.50 (t, $J = 5.6$ Hz, 1H, NH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 14.1, 20.0, 31.6, 39.4, 125.8, 126.9, 127.4, 127.7, 128.1, 128.2(2C), 129.5, 130.4, 131.8, 132.1, 133.7, 134.8, 136.5, 140.4, 140.7, 165.1.

HRMS calcd for $\text{C}_{21}\text{H}_{22}\text{N}_2\text{O}_3\text{S}_2$ [(M+H) $^+$]: 415.1145, found: 415.1151.

General Procedure for the Syntheses of 6b, c, e. The mixture of *N*-butyl-4-chloro-3-sulfamoyl-benzamide **3** (100 mg, 0.344 mmol) and appropriate amine (2.0 mL) was heated at 130 °C for 24 h in an inert atmosphere. The mixture was cooled to room temperature and 2 N HCl(aq) (5 mL) was added. The product was extracted with EtOAc (3 $\times$ 5 mL). The organic layer was washed with H_2O , dried over anhydrous MgSO_4 , filtered, and concentrated under reduced pressure.

***N*-Butyl-4-(cyclohexylamino)-3-sulfamoyl-benzamide (6b).** The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 1:1). Yield: 64 mg, 53%, mp 89–91 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 0.91 (t, $J = 7.3$ Hz, 3H, CH_3), 1.23–1.60 (m, 10H, $\text{CH}_2\text{CH}_2\text{CH}_3$ and CH cyclohexyl), 1.66–1.69 (m, 2H, CH cyclohexyl), 1.90–1.93 (m, 2H, CH cyclohexyl), 3.25 (dt, $J = 6.8$ Hz, $J = 5.7$ Hz, 2H, CH_2NH), 3.49–3.52 (m, 1H, CHNH), 6.17 (d, $J = 7.6$ Hz, 1H, NH-cyclohexyl), 6.84 (d, $J = 8.8$ Hz, 1H, ArH), 7.43 (s, 2H, SO_2NH_2), 7.85 (dd, $J = 8.8$ Hz, $J = 2.1$ Hz, 1H, ArH), 8.21 (d, $J = 2.1$ Hz, 1H, ArH), 8.25 (t, $J = 5.7$ Hz, 1H, NH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 14.2, 20.1, 24.4, 25.8, 31.9, 32.3, 39.2, 50.5, 111.8, 120.9, 124.5, 129.2, 132.3, 146.3, 165.4.

HRMS calcd for $\text{C}_{17}\text{H}_{27}\text{N}_3\text{O}_3\text{S}_2$ [(M+H) $^+$]: 354.1846, found: 354.1840.

4-(Benzylamino)-*N*-butyl-3-sulfamoyl-benzamide (6c). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 1:1). Yield: 40 mg, 32%, mp 132–133 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 0.88 (t, $J = 7.2$ Hz, 3H, CH_3), 1.35 (sext, $J = 7.2$ Hz, 2H, CH_2CH_3), 1.49 (quint, $J = 7.2$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{CH}_3$), 3.23 (dt, $J = 6.8$ Hz, $J = 5.9$ Hz, 2H, CH_2NH), 4.55 (d, $J = 5.8$ Hz, 2H, PhCH_2NH), 6.69 (d, $J = 8.8$ Hz, 1H, ArH), 6.82 (t, $J = 5.8$ Hz, 1H, NH-benzyl), 7.23–7.39 (m, 5H, ArH), 7.49 (s, 2H, SO_2NH_2), 7.77 (dd, $J = 8.8$ Hz, $J = 2.0$ Hz, 1H, ArH), 8.23 (m, 2H, ArH and NH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 14.2, 20.1, 31.8, 39.2, 46.5, 111.9, 121.5, 125.1, 127.4, 128.9, 128.9, 132.1, 139.2, 146.8, 165.4.

HRMS calcd for $\text{C}_{18}\text{H}_{23}\text{N}_3\text{O}_3\text{S}_2$ [(M+H) $^+$]: 362.1533, found: 362.1539.

***N*-Butyl-4-(cyclooctylamino)-3-sulfamoyl-benzamide (6e).** The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 1:1). Yield: 55 mg, 42%, brownish oily residue.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 0.91 (t, $J = 7.3$ Hz, 3H, CH_3), 1.35 (sext, $J = 7.3$ Hz, 2H, CH_2CH_3), 1.44–1.73 (m, 14H, CH cyclooctyl and $\text{CH}_2\text{CH}_2\text{CH}_3$), 1.81–1.86 (m, 2H, CH cyclooctyl), 3.25 (dt, $J = 6.9$ Hz, $J = 5.9$ Hz, 2H, CH_2NH), 3.70–3.73 (m, 1H, CHNH cyclooctyl), 6.19 (d, $J = 7.6$ Hz, 1H, NH-cyclooctyl), 6.75 (d, $J = 8.8$ Hz, 1H, ArH), 7.43 (s, 2H, SO_2NH_2), 7.86 (dd, $J = 8.8$ Hz, $J = 2.1$ Hz, 1H, ArH), 8.19 (d, $J = 2.1$ Hz, 1H, ArH), 8.24 (t, $J = 5.9$ Hz, 1H, NH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 14.2, 20.1, 23.6, 25.5, 27.2, 31.7, 31.9, 39.2, 51.9, 111.9, 120.8, 124.6, 129.3, 132.4, 146.1, 165.5.

HRMS calcd for $\text{C}_{19}\text{H}_{23}\text{N}_3\text{O}_3\text{S}_2$ [(M+H) $^+$]: 382.2159, found: 382.2155.

General Procedure for the Syntheses of 7(a, b, h, i, k–m), 8(a, b, k–m). The mixture of appropriate methyl 4-sulfanilsubstituted-3-sulfamoyl-benzoate (**4a**, **b**, **h**, **i**, **k–m**) or appropriate *N*-butyl-4-sulfanilsubstituted-3-sulfamoyl-benzamide (**8a**, **b**, **k–m**) (0.200 mmol), acetic acid (2.0 mL) and 30% H_2O_2 (0.090 mL, 1.15 mmol) was stirred for 12 h. Then 1% Na_2SO_3 (aq) (6.0 mL) was added to the mixture and the product was extracted with EtOAc (3 $\times$ 5 mL). The organic layer was washed with H_2O , dried over anhydrous MgSO_4 , filtered, and concentrated under reduced pressure.

Methyl 4-(Benzenesulfinyl)-3-sulfamoyl-benzoate (7a). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 3:1). Yield: 56 mg, 82%, mp 107–108 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 3.91 (s, 3H, CH_3O), 7.48–7.53 (m, 3H, ArH), 7.75–7.78 (m, 2H, ArH), 8.09 (s, 2H, SO_2NH_2), 8.32 (d, $J = 8.2$ Hz, 1H, ArH), 8.37 (dd, $J = 8.2$ Hz, $J = 1.7$ Hz, 1H, ArH), 8.47 (d, $J = 1.7$ Hz, 1H, ArH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 53.3, 125.6, 126.9, 128.5, 129.9, 131.6, 132.8, 133.9, 142.3, 145.9, 149.7, 164.9.

HRMS calcd for $\text{C}_{14}\text{H}_{13}\text{NO}_5\text{S}_2$ [(M+H) $^+$]: 340.0308, found: 340.0311.

Methyl 4-Cyclohexylsulfinyl-3-sulfamoyl-benzoate (7b). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 3:1). Yield: 52 mg, 75%, mp 188–189 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 1.14–2.07 (m, 10H, CH_2 cyclohexyl), 2.96–3.04 (m, 1H, CHSO), 3.94 (s, 3H, CH_3O), 7.97 (s, 2H, SO_2NH_2), 8.11 (d, $J = 8.2$ Hz, 1H, ArH), 8.38 (dd, $J = 8.2$ Hz, $J = 1.7$ Hz, 1H, ArH), 8.51 (d, $J = 1.7$ Hz, 1H, ArH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 20.9, 25.1, 25.4, 26.1, 28.6, 53.3, 60.9, 127.3, 129.1, 132.5, 132.6, 142.1, 146.5, 165.1.

HRMS calcd for $\text{C}_{14}\text{H}_{19}\text{NO}_5\text{S}_2$ [(M+H) $^+$]: 346.0777, found: 346.0783.

Methyl 4-Methylsulfinyl-3-sulfamoyl-benzoate (7h). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 5:1). Yield: 50 mg, 90%, mp 258–259 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 2.81 (s, 3H, CH_3SO), 3.95 (s, 3H, CH_3O), 7.97 (s, 2H, SO_2NH_2), 8.34 (d, $J = 8.2$ Hz, 1H, ArH), 8.43 (dd, $J = 8.2$ Hz, $J = 1.6$ Hz, 1H, ArH), 8.47 (d, $J = 1.6$ Hz, 1H, ArH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 44.6, 53.3, 125.7, 128.5, 132.6, 133.8, 141.5, 151.2, 165.1.

HRMS calcd for $\text{C}_{10}\text{H}_{11}\text{NO}_5\text{S}_2$ [(M+H) $^+$]: 278.0151, found: 278.0155.

Methyl 4-Propylsulfinyl-3-sulfamoyl-benzoate (7i). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 3:1). Yield: 40 mg, 66%, mp 144–145 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 0.99 (t, $J = 7.4$ Hz, 3H, CH_2CH_3), 1.59–1.62 (m, 1H, $\text{CH}_2\text{H}_\text{B}\text{CH}_3$), 1.77–1.81 (m, 1H, $\text{CH}_2\text{H}_\text{A}\text{CH}_3$), 2.66–2.69 (m, 1H, $\text{CH}_2\text{H}_\text{B}\text{SO}$), 3.10–3.13 (m, 1H, $\text{CH}_2\text{H}_\text{A}\text{SO}$), 3.93 (s, 3H, CH_3O), 7.97 (s, 2H, SO_2NH_2), 8.26 (d, $J = 8.2$ Hz, 1H, ArH), 8.41 (dd, $J = 8.2$ Hz, $J = 1.7$ Hz, 1H, ArH), 8.48 (d, $J = 1.7$ Hz, 1H, ArH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 13.2, 16.5, 53.3, 58.9, 126.3, 128.7, 132.5, 133.4, 141.7, 149.2, 165.1.

HRMS calcd for $\text{C}_{11}\text{H}_{15}\text{NO}_5\text{S}_2$ [(M+H) $^+$]: 306.0464, found: 306.0470.

Methyl 4-Isopentylsulfinyl-3-sulfamoyl-benzoate (7k). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 5:1). Yield: 37 mg, 55%, mp 98–99 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 0.86 (d, $J = 6.5$ Hz, 3H, CH_3), 0.89 (d, $J = 6.5$ Hz, 3H, CH_3), 1.40–1.45 (m, 1H, $\text{CH}_2\text{H}_\text{B}\text{CH}_3$), 1.63–1.68 (m, 2H, CHCH_3 and $\text{CH}_2\text{H}_\text{B}\text{CH}_3$), 2.66–2.69 (m, 1H, $\text{CH}_2\text{H}_\text{B}\text{SO}$), 3.15–3.19 (m, 1H, $\text{CH}_2\text{H}_\text{A}\text{SO}$), 3.94 (s, 3H, CH_3O), 7.97 (s, 2H, SO_2NH_2), 8.26 (d, $J = 8.2$ Hz, 1H, ArH), 8.41 (dd, $J = 8.2$ Hz, $J = 1.7$ Hz, 1H, ArH), 8.49 (d, $J = 1.7$ Hz, 1H, ArH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 22.4, 22.8, 27.2, 31.4, 53.3, 55.1, 126.4, 128.7, 132.5, 133.4, 141.7, 149.1, 165.1.

HRMS calcd for $\text{C}_{13}\text{H}_{19}\text{NO}_5\text{S}_2$ [(M+H) $^+$]: 334.0777, found: 334.0771.

Methyl 4-Cyclopentylsulfinyl-3-sulfamoyl-benzoate (7l). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 4:1). Yield: 51 mg, 77%, mp 109–111 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 1.08–1.18 (m, 1H, CH cyclopentyl), 1.43–1.53 (m, 1H, CH cyclopentyl), 1.54–1.66 (m, 3H, CH cyclopentyl), 1.80–1.94 (m, 2H, CH cyclopentyl), 2.00–2.11 (m, 1H, CH cyclopentyl), 3.53 (quint, $J = 7.6$ Hz, 1H, CH cyclopentyl), 3.94 (s, 3H, OCH_3), 7.97 (s, 2H, SO_2NH_2), 8.16 (d, $J = 8.4$ Hz, 1H, ArH), 8.36 (dd, $J = 8.0$ Hz, $J = 1.6$ Hz, 1H, ArH), 8.49 (d, $J = 1.3$ Hz, 1H, ArH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 21.3, 25.7, 26.3, 29.6, 53.3, 61.7, 126.6, 128.9, 132.4, 132.9, 141.7, 148.0, 165.1.

HRMS calcd for $\text{C}_{13}\text{H}_{17}\text{NO}_5\text{S}_2$ [(M+H) $^+$]: 332.0621, found: 332.0615.

Methyl 4-(1-Naphthylsulfinyl)-3-sulfamoyl-benzoate (7m). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 1:1). Yield: 47 mg, 60%, mp 125–127 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 3.93 (s, 3H, CH_3), 7.61–7.68 (m, 4H, naphthyl-H), 8.04–8.08 (m, 1H, naphthyl-H), 8.11 (s, 2H, SO_2NH_2), 8.14 (d, J = 7.6 Hz, 1H, naphthyl-H), 8.22 (d, J = 8.0 Hz, 1H, ArH), 8.39 (dd, J = 1.6 Hz, J = 8.0 Hz, 1H, ArH), 8.49–8.52 (m, 1H, naphthyl-H), 8.54 (d, 1H, J = 1.6 Hz, ArH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 53.4, 123.5, 125.5, 126.3, 127.5, 128.0, 128.7, 128.8, 129.2, 129.8, 132.5, 133.2, 133.8 (2C), 141.9, 143.1, 146.7, 165.0.

HRMS calcd for $\text{C}_{18}\text{H}_{15}\text{NO}_5\text{S}_2$ [(M+H) $^+$]: 390.0464, found: 390.0460.

4-(Benzenesulfinyl)-N-butyl-3-sulfamoyl-benzamide (8a). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 1:1). Yield: 61 mg, 80%, mp 168–171 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 0.89 (t, J = 7.4 Hz, 3H, CH_3), 1.32 (sext, J = 7.8 Hz, 2H, $\text{NH}(\text{CH}_2)_2\text{CH}_2\text{CH}_3$), 1.50 (quint, J = 7.2 Hz, 2H, NHCH_2CH_2), 3.27 (q, J = 7.2 Hz, 2H, NHCH_2), 7.47–7.53 (m, 3H, ArH), 7.75 (d, J = 6.4 Hz, 2H, ArH), 7.97 (s, 2H, SO_2NH_2), 8.20 (s, 2H, ArH), 8.35 (s, 1H, ArH), 8.77 (t, J = 5.4 Hz, 1H, NH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 14.1, 20.1, 31.5, 39.6, 125.5, 126.4, 127.2, 129.8, 131.5, 131.7, 138.0, 142.0, 146.3, 147.3, 164.8.

HRMS calcd for $\text{C}_{17}\text{H}_{20}\text{N}_2\text{O}_4\text{S}_2$ [(M+H) $^+$]: 381.0937, found: 381.0933.

N-Butyl-4-cyclohexylsulfinyl-3-sulfamoyl-benzamide (8b). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 1:1). Yield: 63 mg, 82%, mp 195–196 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 0.93 (t, J = 7.3 Hz, 3H, CH_3), 1.07–1.18 (m, 3H, CH cyclohexyl), 1.29–1.43 (m, 4H, CH cyclohexyl and CH_2CH_3), 1.49–1.60 (m, 4H, CH cyclohexyl and $\text{CH}_2\text{CH}_2\text{CH}_2$), 1.71–1.74 (m, 1H, CH cyclohexyl), 1.81–1.87 (m, 1H, CH cyclohexyl), 2.01–2.05 (m, 1H, CH cyclohexyl), 3.01 (tt, J = 12.3 Hz, J = 3.6 Hz, 1H, CHSO), 3.24–3.29 (m, 2H, CH_2NH), 7.84 (s, 2H, SO_2NH_2), 8.02 (d, J = 8.2 Hz, 1H, ArH), 8.24 (dd, J = 8.2 Hz, J = 1.7 Hz, 1H, ArH), 8.41 (d, J = 1.7 Hz, 1H, ArH), 8.82 (t, J = 5.7 Hz, 1H, NH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 14.2, 20.1, 21.0, 25.1, 25.4, 26.1, 28.5, 31.6, 39.5, 60.9, 126.6, 127.9, 130.3, 137.7, 141.7, 143.8, 164.9.

HRMS calcd for $\text{C}_{17}\text{H}_{26}\text{N}_2\text{O}_4\text{S}_2$ [(M+H) $^+$]: 387.1407, found: 387.1401.

N-Butyl-4-isopentylsulfinyl-3-sulfamoyl-benzamide (8k). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 1:1). Yield: 55 mg, 74%, mp 121–124 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 0.83 (d, J = 6.4 Hz, 6H, $\text{S}(\text{CH}_2)_2\text{CH}(\text{CH}_3)_2$), 0.92 (t, J = 7.2 Hz, 3H, $\text{NH}(\text{CH}_2)_3\text{CH}_3$), 1.35 (sext, J = 7.2 Hz, 2H, $\text{NHCH}_2\text{CH}_2\text{CH}_3$), 1.45–1.66 (m, 5H, NHCH_2CH_2 , $\text{SCH}_2\text{CH}_2\text{CH}$), 3.31 (q, J = 6.8 Hz, 2H, NHCH_2), 3.66–3.70 (m, 2H, SCH_2), 7.33 (s, 2H, SO_2NH_2), 8.25 (d, J = 8.0 Hz, 1H, ArH), 8.30 (dd, J = 8.0 Hz, J = 1.6 Hz, 1H, ArH), 8.60 (d, J = 1.6 Hz, 1H, ArH), 8.96 (t, J = 5.6 Hz, 1H, NH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 14.2, 20.1, 22.3, 27.0, 30.7, 31.5, 39.6, 53.4, 129.6, 131.5, 133.8, 138.1, 140.4, 143.1, 164.2.

HRMS calcd for $\text{C}_{16}\text{H}_{26}\text{N}_2\text{O}_4\text{S}_2$ [(M+H) $^+$]: 375.1407, found: 375.1411.

N-Butyl-4-cyclopentylsulfinyl-3-sulfamoyl-benzamide (8l). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 1:1). Yield: 47 mg, 63%, mp 103–107 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 0.92 (t, J = 7.2 Hz, 3H, CH_3), 1.09–1.17 (m, 1H, CH cyclopentyl), 1.35 (sext, J = 7.2 Hz, 2H, $\text{NH}(\text{CH}_2)_2\text{CH}_2\text{CH}_3$), 1.45–1.63 (m, 6H, NHCH_2CH_2 , CH cyclopentyl), 1.82–1.94 (m, 2H, CH cyclopentyl), 2.01–2.10 (m, 1H, CH cyclopentyl), 3.26–3.33 (m, 2H, NHCH_2), 3.52 (quint, J = 7.6 Hz, 1H, CH cyclopentyl), 7.85 (s, 2H, SO_2NH_2), 8.07 (d, J = 8.4 Hz, 1H, ArH), 8.23 (dd, J = 8.0 Hz, J = 1.6 Hz, 1H, ArH), 8.39 (d, J = 1.6 Hz, 1H, ArH), 8.80 (t, J = 5.6 Hz, 1H, NH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 14.2, 20.1, 21.4, 25.8, 26.3, 29.6, 31.5, 39.6, 61.7, 125.9, 127.7, 130.6, 137.7, 141.3, 145.4, 164.9.

HRMS calcd for $\text{C}_{16}\text{H}_{24}\text{N}_2\text{O}_4\text{S}_2$ [(M+H) $^+$]: 373.1250, found: 373.1255.

N-Butyl-4-(1-naphthylsulfinyl)-3-sulfamoyl-benzamide (8m). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 1:1). Yield: 28 mg, 33%, mp 232–235 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 0.95 (t, J = 7.2 Hz, 3H, CH_3), 1.38 (sext, J = 7.2 Hz, 2H, $\text{NH}(\text{CH}_2)_2\text{CH}_2\text{CH}_3$), 1.56 (quint, J = 7.2 Hz, 2H, NHCH_2CH_2), 3.28–3.37 (m, 2H, NHCH_2), 7.63–7.73 (m, 4H, naphthyl-H), 8.08 (s, 2H, SO_2NH_2), 8.11–8.13 (m, 2H, naphthyl-H, ArH), 8.18–8.22 (m, 1H, naphthyl-H), 8.26 (dd, J = 1.6 Hz, J = 8.0 Hz, 1H, ArH), 8.47 (d, 1H, J = 1.6 Hz, ArH), 8.52–8.54 (m, 1H, naphthyl-H), 8.86 (t, J = 5.6 Hz, 1H, NH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 14.2, 20.1, 31.5, 39.6, 123.5, 125.3, 126.2, 127.4, 127.5, 128.0, 128.3, 129.2, 129.7, 131.5, 132.3, 133.8, 138.4, 142.0, 142.8, 144.3, 164.8.

HRMS calcd for $\text{C}_{21}\text{H}_{22}\text{N}_2\text{O}_4\text{S}_2$ [(M+H) $^+$]: 431.1094, found: 431.1088.

General Procedure for the Syntheses of 9(a–d, h, i, l, m) and 10(b, k–m). The 30% H_2O_2 (aq) (1.08 mmol, 0.110 mL) was added in small portions over 3 h to a solution of appropriate methyl 4-sulfanilsubstituted-3-sulfamoyl-benzoate (4a–d, h, i, l, m) or appropriate N-butyl-4-sulfanilsubstituted-3-sulfamoyl-benzamide (10b, k–m) (0.200 mmol) and acetic acid (2.0 mL) at 70 °C and allowed stirring for 6–8 h. The solvent was removed under reduced pressure and the resultant precipitate was filtered, and washed with H_2O .

Methyl 4-(Benzenesulfonyl)-3-sulfamoyl-benzoate (9a). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 5:1). Yield: 57 mg, 80%, mp 212–213 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 3.94 (s, 3H, CH_3O), 7.55 (s, 2H, SO_2NH_2), 7.61 (t, J = 7.4 Hz, 2H, ArH), 7.70 (t, J = 7.4 Hz, 1H, ArH), 7.95 (d, J = 7.4 Hz, 2H, ArH), 8.43 (dd, J = 8.2 Hz, J = 1.7 Hz, 1H, ArH), 8.61 (d, J = 8.2 Hz, 1H, ArH), 8.64 (d, J = 1.7 Hz, 1H, ArH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 53.6, 128.4, 129.6, 130.7, 133.9, 134.3, 134.4, 135.1, 140.6, 141.6, 143.4, 164.5.

HRMS calcd for $\text{C}_{14}\text{H}_{13}\text{NO}_6\text{S}_2$ [(M+H) $^+$]: 356.0257, found: 356.0266.

Methyl 4-Cyclohexylsulfonyl-3-sulfamoyl-benzoate (9b). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 5:1). Yield: 61 mg, 85%, mp 155–156 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 1.17–1.81 (m, 10H, CH_2 cyclohexyl), 3.86–3.91 (m, 1H, CHSO_2), 3.95 (s, 3H, CH_3O), 7.42 (s, 2H, SO_2NH_2), 8.25 (d, J = 8.1 Hz, 1H, ArH), 8.42 (dd, J = 8.1 Hz, J = 1.7 Hz, 1H, ArH), 8.70 (d, J = 1.7 Hz, 1H, ArH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 24.8, 24.9, 25.1, 53.6, 61.8, 130.9, 133.6, 134.9, 135.2, 138.6, 143.8, 164.6.

HRMS calcd for $\text{C}_{14}\text{H}_{19}\text{NO}_6\text{S}_2$ [(M+H) $^+$]: 362.0727, found: 362.0721.

Methyl 4-Benzylsulfonyl-3-sulfamoyl-benzoate (9c). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 5:1). Yield: 66 mg, 89%, mp 189–190 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 3.94 (s, 3H, CH_3O), 5.03 (s, 2H, CH_2SO_2), 7.19–7.21 (m, 2H, ArH), 7.30–7.34 (m, 3H, ArH), 7.53 (s, 2H, SO_2NH_2), 7.88 (d, J = 8.0 Hz, 1H, ArH), 8.25 (dd, J = 8.0 Hz, J = 1.8 Hz, 1H, ArH), 8.70 (d, J = 1.8 Hz, 1H, ArH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 53.6, 60.9, 127.9, 129.0, 129.3, 130.6, 131.2, 133.4, 134.5, 135.1, 139.4, 143.7, 164.5.

HRMS calcd for $\text{C}_{15}\text{H}_{15}\text{NO}_6\text{S}_2$ [(M+H) $^+$]: 370.0414, found: 370.0419.

Methyl 4-(2-Phenylethylsulfonyl)-3-sulfamoyl-benzoate (9d). The product was purified by flash chromatography on silica gel (CHCl_3 :EtOAc, 5:1). Yield: 64 mg, 83%, mp 186–187 °C.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm: 2.98 (t, J = 8.0 Hz, 2H, CH_2Ph), 3.94 (s, 3H, CH_3O), 4.02 (t, J = 8.0 Hz, 2H, CH_2SO_2), 7.18–7.27 (m, 5H, ArH), 7.44 (s, 2H, SO_2NH_2), 8.30 (d, J = 8.0 Hz, 1H, ArH), 8.40 (d, J = 8.0 Hz, 1H, ArH), 8.67 (s, 1H, ArH).

^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ ppm: 28.1, 53.6, 55.9, 127.1, 128.9, 129.0, 130.7, 133.9, 134.3, 135.1, 137.9, 140.0, 143.5, 164.5.

HRMS calcd for $C_{16}H_{17}NO_6S_2$ [(M+H)⁺]: 384.0570, found: 384.0578.

Methyl 4-Methylsulfonyl-3-sulfamoyl-benzoate (9h). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 5:1). Yield: 55 mg, 95%, mp 187–188 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 3.51 (s, 3H, CH₃SO₂), 3.96 (s, 3H, CH₃O), 7.42 (s, 2H, SO₂NH₂), 8.36 (d, *J* = 8.2 Hz, 1H, ArH), 8.45 (dd, *J* = 8.2 Hz, *J* = 1.6 Hz, 1H, ArH), 8.67 (d, *J* = 1.6 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 44.1, 53.6, 130.5, 133.4, 134.1, 135.1, 141.6, 143.2, 164.6.

HRMS calcd for $C_{15}H_{11}NO_6S_2$ [(M+H)⁺]: 294.0101, found: 294.0106.

Methyl 4-Propylsulfonyl-3-sulfamoyl-benzoate (9i). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 5:1). Yield: 57 mg, 88%, mp 169–170 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 0.94 (t, *J* = 7.4 Hz, 3H, CH₂CH₃), 1.63–1.66 (m, 2H, CH₂CH₃), 3.66 (t, *J* = 7.7 Hz, 2H, CH₂SO₂), 3.96 (s, 3H, CH₃O), 7.43 (s, 2H, SO₂NH₂), 8.33 (d, *J* = 8.1 Hz, 1H, ArH), 8.44 (dd, *J* = 8.1 Hz, *J* = 1.6 Hz, 1H, ArH), 8.69 (d, *J* = 1.6 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 13.1, 16.2, 53.6, 56.7, 130.7, 133.9, 134.2, 135.2, 140.2, 143.5, 164.6.

HRMS calcd for $C_{17}H_{15}NO_6S_2$ [(M+H)⁺]: 322.0414, found: 322.0422.

Methyl 4-Cyclopentylsulfonyl-3-sulfamoyl-benzoate (9l). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 5:1). Yield: 63 mg, 90%, mp 199–200 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 1.55–1.65 (m, 2H, CH cyclopentyl), 1.68–1.76 (m, 2H, CH cyclopentyl), 1.77–1.86 (m, 2H, CH cyclopentyl), 1.88–1.96 (m, 2H, CH cyclopentyl), 3.96 (s, 3H, OCH₃), 4.44 (quint, *J* = 7.2 Hz, 1H, CH cyclopentyl), 7.41 (s, 2H, SO₂NH₂), 8.34 (d, *J* = 8.0 Hz, 1H, ArH), 8.42 (dd, *J* = 8.0 Hz, *J* = 1.6 Hz, 1H, ArH), 8.70 (d, *J* = 1.6 Hz, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 26.0, 27.0, 53.6, 62.7, 130.9, 133.8, 134.6, 135.1, 139.7, 143.6, 164.6.

HRMS calcd for $C_{15}H_{17}NO_6S_2$ [(M+H)⁺]: 348.0570, found: 348.0560.

Methyl 4-(1-Naphthylsulfonyl)-3-sulfamoyl-benzoate (9m). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 8:1). Yield: 63 mg, 78%, mp 223–224 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 3.92 (s, 3H, OCH₃), 7.63–7.67 (m, 4H, naphthylH and SO₂NH₂), 7.74 (t, *J* = 8.0 Hz, 1H, ArH), 8.12–8.14 (m, 1H, naphthylH), 8.25 (d, *J* = 8.0 Hz, 1H, ArH), 8.31–8.37 (m, 4H, naphthylH), 8.70 (s, 1H, ArH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 53.6, 123.9, 125.0, 127.5, 127.8, 129.1, 129.9, 130.8, 131.0, 133.0, 133.8, 134.2, 134.9, 135.2, 135.9, 141.8, 143.5, 164.5.

HRMS calcd for $C_{18}H_{15}NO_6S_2$ [(M+H)⁺]: 406.0414, found: 406.0423.

N-Butyl-4-cyclohexylsulfonyl-3-sulfamoyl-benzamide (10b). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 3:1). Yield: 68 mg, 84%, mp 155–156 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 0.93 (t, *J* = 7.3 Hz, 3H, CH₃), 1.17–1.81 (m, 14H, CH₂CH₂CH₃ and CH cyclohexyl), 3.30–3.32 (m, 2H, CH₂NH), 3.90–3.93 (m, 1H, CHSO₂), 7.33 (s, 2H, SO₂NH₂), 8.19 (d, *J* = 8.1 Hz, 1H, ArH), 8.29 (d, *J* = 8.1 Hz, 1H, ArH), 8.61 (s, 1H, ArH), 8.96 (t, *J* = 5.3 Hz, 1H, NH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 14.1, 20.1, 24.9, 24.9, 25.1, 31.5, 39.6, 61.8, 129.8, 131.2, 134.4, 136.5, 140.4, 143.4, 164.2.

HRMS calcd for $C_{17}H_{26}N_2O_5S_2$ [(M+H)⁺]: 403.1356, found: 403.1351.

N-Butyl-4-isopentylsulfonyl-3-sulfamoyl-benzamide (10k). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 1:1). Yield: 53 mg, 68%, mp 160–162 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 0.83 (d, *J* = 6.4 Hz, 6H, S(CH₂)₂CH(CH₃)₂), 0.92 (t, *J* = 7.6 Hz, 3H, NH(CH₂)₃CH₃), 1.35 (sext, *J* = 7.2 Hz, 2H, NHCH₂CH₂CH₃), 1.46–1.66 (m, 5H, NHCH₂CH₂, SCH₂CH₂CH), 3.31 (q, *J* = 6.8 Hz, 2H, NHCH₂), 3.66–3.70 (m, 2H, SCH₂), 7.33 (s, 2H, SO₂NH₂), 8.25 (d, *J* = 8.0 Hz,

1H, ArH), 8.30 (dd, *J* = 8.0 Hz, *J* = 1.6 Hz, 1H, ArH), 8.60 (s, 1H, ArH), 8.97 (t, *J* = 5.2 Hz, 1H, NH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 14.2, 20.1, 22.3, 27.0, 30.7, 31.5, 39.7, 53.4, 129.6, 131.5, 133.8, 138.1, 140.4, 143.1, 164.2.

HRMS calcd for $C_{16}H_{26}N_2O_5S_2$ [(M+H)⁺]: 391.1356, found: 391.1359.

N-Butyl-4-cyclopentylsulfonyl-3-sulfamoyl-benzamide (10l). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 1:1). Yield: 57 mg, 73%, mp 143–145 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 0.92 (t, *J* = 7.6 Hz, 3H, CH₂CH₃), 1.35 (sext, *J* = 7.6 Hz, 2H, CH₂CH₂CH₃), 1.54 (quint, *J* = 7.2 Hz, 2H, NHCH₂CH₂), 1.57–1.65 (m, 2H, CH cyclopentyl), 1.68–1.85 (m, 4H, CH cyclopentyl), 1.88–1.96 (m, 2H, CH cyclopentyl), 3.31 (q, *J* = 6.8 Hz, 2H, NHCH₂), 4.44 (quint, *J* = 7.2 Hz, 1H, SO₂CH cyclopentyl), 7.32 (br. s, 2H, SO₂NH₂), 8.26–8.31 (m, 2H, ArH), 8.62 (s, 1H, ArH), 8.96 (t, *J* = 5.6 Hz, 1H, NH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 14.2, 20.1, 26.0, 27.0, 31.5, 39.6, 62.6, 129.7, 131.4, 134.1, 137.7, 140.3, 143.2, 164.2.

HRMS calcd for $C_{16}H_{24}N_2O_5S_2$ [(M+H)⁺]: 389.1199, found: 389.1205.

N-Butyl-4-(1-naphthylsulfonyl)-3-sulfamoyl-benzamide (10m). The product was purified by flash chromatography on silica gel (CHCl₃:EtOAc, 1:1). Yield: 22 mg, 24%, mp 266–268 °C.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm: 0.89 (t, *J* = 7.6 Hz, 3H, CH₂CH₃), 1.32 (sext, *J* = 7.2 Hz, 2H, CH₂CH₂CH₃), 1.51 (quint, *J* = 7.2 Hz, 2H, NHCH₂CH₂), 3.29 (q, *J* = 6.4 Hz, 2H, NHCH₂), 7.52 (s, 2H, SO₂NH₂), 7.62–7.68 (m, 2H, naphthylH), 7.74 (t, *J* = 8.0 Hz, 1H, naphthylH), 8.13–8.15 (m, 1H, naphthylH), 8.19–8.23 (m, 2H, ArH), 8.30 (d, *J* = 8.4 Hz, 1H, naphthylH), 8.34–8.39 (m, 2H, naphthylH), 8.60 (d, *J* = 1.6 Hz, 1H, ArH), 8.92 (t, *J* = 5.2 Hz, 1H, NH).

¹³C NMR (100 MHz, DMSO-*d*₆) δ ppm: 14.1, 20.1, 31.4, 39.6, 124.0, 125.0, 127.5, 127.8, 129.0, 129.8, 129.9, 130.7, 131.4, 132.7, 134.2, 135.5, 135.8, 139.8, 140.2, 143.1, 164.3.

HRMS calcd for $C_{21}H_{22}N_2O_5S_2$ [(M+H)⁺]: 447.1043, found: 447.1048.

Protein Preparation. Recombinant human carbonic anhydrases (CAI, CAII, CAIII, CAIV, CAVA, CAVB, CAVI, CAVII, CAIX, CAXII, CAXIII, CAXIV) were expressed and chromatographically purified according to previously published protocols⁴⁷ and were used for FTSA experiments. Proteins (CAII and CAXIII) used for crystallization were additionally purified by affinity chromatography and concentrated. CAIX used for crystallization were expressed and purified according to this protocol.⁴⁸ Production of recombinant CAII and the extracellular part of CAIX comprising PG and CA domains (residues 38–391) used in SFA experiments is described in ref 49.

Determination of Observed Binding Parameters. Fluorescent Thermal Shift Assay (FTSA). Dissociation constants, $K_{d,obs}$ (listed in Table 1) of the compounds binding to CAs were determined by the fluorescent thermal shift assay using a QIAGEN's real-time PCR cycler the "Rotor-Gene Q" and Rotor-Gene Style 4-strip tubes from STARLAB. Ligands were dissolved in DMSO stock solutions to concentrations of 10 mM or 20 mM and used for the serial dilution of the dilution factor 2 in DMSO. These samples were diluted with buffer solution and mixed with a prepared protein solution, consisting of protein stock, buffer solution, and solvatochromic dye 8-anilino-1-naphthalenesulfonate (ANS). All final samples typically contained up to 10 μM CA, compound solutions of serial dilution from 0 μM to 200 μM at 8 different concentrations differing by 2 times, 50 μM ANS, 50 mM sodium phosphate buffer at pH 7.0, 100 mM sodium chloride, and 2.0% (v/v) DMSO. Samples preparation is explained in detail in.⁵⁰ The excitation and emission wavelengths of ANS were 365 ± 20 and 460 ± 15 nm. Samples were heated from 25 to 99 °C at the rate of 1 °C/min. The curve-fitting procedure was performed by Thermott⁵¹ at 37 °C. Data are deposited in the public database: Protein–Ligand Binding Database⁵² (Database URL: <https://pldb.org/>).

Stopped-Flow Carbon Dioxide Hydration Assay. Recombinant CAII and the extracellular part of CAIX comprising PG and CA domains (residues 38–391) were used in inhibition assays. A stopped-flow instrument (BioLogic) was used for measuring the CA-catalyzed CO₂ hydration activity in the presence of inhibitors.⁵³ The assay buffer

Table 4. Crystallization Conditions Used to Grow Protein Crystals in This Study

PDB ID	isozyme—compound	crystallization buffer	sitting drop
9FPT	CAII—2	0.1 M sodium bicine (pH 9.0) and 2 M sodium malonate (pH 7.0)	2 μ L of 41 mg/mL CAII solution and 2 μ L of crystallization buffer
9FPU	CAII—3	0.1 M sodium bicine (pH 9.0) and 2 M sodium malonate (pH 7.0)	2 μ L of 41 mg/mL CAII solution and 2 μ L of crystallization buffer
9FPQ	CAII—4c	0.1 M sodium bicine (pH 9.0), 0.2 M ammonium sulfate, and 2 M sodium malonate (pH 7.0)	2 μ L of 41 mg/mL CAII solution and 2.5 μ L of crystallization buffer
9FPR	CAII—4d	0.1 M sodium bicine (pH 9.0), 0.2 M ammonium sulfate, and 2 M sodium malonate (pH 7.0)	3 μ L of 12 mg/mL CAII solution and 3 μ L of crystallization buffer
9FPS	CAII—4h	0.1 M sodium bicine (pH 9.0) and 2 M sodium malonate (pH 7.0)	3 μ L of 12 mg/mL CAII solution and 3 μ L of crystallization buffer
9R8X	CAIX—4d	1.0 M diammonium hydrogen phosphate and 0.1 M sodium acetate (pH 4.5)	10 mg/mL CAIX solution
9R8Y	CAIX—5b	1.0 M diammonium hydrogen phosphate and 0.1 M sodium acetate (pH 4.5)	10 mg/mL CAIX solution
9FPV	CAXIII—4c	0.1 M sodium citrate (pH 5.5), 0.1 M sodium acetate (pH 4.5), and 26% of PEG4000	1 μ L of 23 mg/mL CAXIII solution and 0.4 μ L of crystallization buffer
9FPW	CAXIII—4d	0.1 M sodium citrate (pH 5.5), 0.1 M sodium acetate (pH 4.5), and 26% of PEG4000	1 μ L of 23 mg/mL CAXIII solution and 0.4 μ L of crystallization buffer

consisted of 0.2 mM Phenol Red (pH indicator used in absorbance maximum of 557 nm), 20 mM HEPES Na (pH 7.5), and 20 mM Na₂SO₄. The concentration of CAII and CAIX in the enzyme assay was 4 nM and 1 nM, respectively. To stabilize CAIX during the measurements, 0.0025% dodecyl- β -D-maltopyranoside (DDM, Anatrace, and Anagrade purity) was included in the reaction mixture.

The substrate (CO₂) concentration in the reaction was 8.5 mM. Rates of the CA-catalyzed CO₂ hydration reaction were followed for 30 s at 25 °C. Four traces of substrate conversion in the reaction were fitted by the exponential function to determine the rate for each inhibitor concentration. The uncatalyzed rates were determined in the same manner and subtracted from the total observed rates. Stock solutions of inhibitors (100 mM) were prepared in dimethyl sulfoxide, and dilutions of up to 100 nM were made thereafter in DMSO. Apparent K_i' values were obtained from dose–response curves recorded for at least six different concentrations of the test compound by the nonlinear least-squares method using an Excel spreadsheet fitting the Williams–Morrison equation.⁵⁴ K_i values were then derived using the Cheng–Prusoff equation.⁵⁵ The K_m values used in the Cheng–Prusoff equation were 9.3 mM for CAII and 7.5 mM for CAIX. Inhibition data are provided in Figure S2.

Calculation of the Intrinsic Binding Parameters. The detailed description of the importance and calculation of the intrinsic values have been previously described.²⁷ For the calculation of the intrinsic dissociation constants, the experimentally measured observed dissociation constants determined by the FTSA, the pK_a of the sulfonamide group of the compound, and the pK_a of the water molecule bound to the zinc cation by the CA were used.

The intrinsic dissociation constant, $K_{d,int}$, is equal to

$$K_{d,int} = K_{d,obs} \times f_{RSO_2NH^-} \times f_{CAZnH_2O} \quad (1)$$

The fraction of deprotonated sulfonamide:

$$f_{RSO_2NH^-} = \frac{10^{pH-pK_{a,SA}}}{1 + 10^{pH-pK_{a,SA}}} \quad (2)$$

The fraction of Zn(II)-bound water form of CA:

$$f_{CAZnH_2O} = \frac{10^{pK_{a,CA}-pH}}{1 + 10^{pK_{a,CA}-pH}} \quad (3)$$

- $K_{d,obs}$ – observed dissociation constant;
- $f_{RSO_2NH^-}$ and f_{CAZnH_2O} – fractions of deprotonated sulfonamide and Zn(II)-bound water molecule;
- $pK_{a,SA}$ – pK_a of the sulfonamide group;
- $pK_{a,CA}$ – pK_a value of water molecule bound to Zn(II) in the active site of CA.

In this study, the pH value was equal to 7.0.

Determination of pK_a Values of the Compound Sulfonamide Group. The pK_a values of the water molecule bound to Zn²⁺ in the active site of CAs, $pK_{a,CA}$, were taken from ref 42 and of compounds, $pK_{a,SA}$ (Figure S3) were experimentally determined as described in ref 56.

We used a constant concentration of sulfonamide (25–400 μ M) and 2.0% (v/v) or 20% (v/v; but only for very poorly soluble ones) of DMSO in universal buffer (50 mM sodium acetate, 25 mM sodium borate, and 50 mM sodium phosphate) at different pH values (in the range from pH 6.0 to 12.0 with 0.5 pH increment). UV/vis spectra of the compound solution were recorded at 37 °C using BMG Labtech CLARIOstarPlus plate reading spectrophotometer. The pK_a values were calculated by normalizing the absorbance and plotting it as a function of pH, then fitting it to the Henderson–Hasselbalch equation using the least-square method. The midpoint of this fitted curve is equal to the sulfonamide group $pK_{a,SA}$.

X-ray Crystallography. Crystallization. Crystals of CAII, CAIX, and CAXIII were obtained using the sitting-drop vapor diffusion method at room temperature. Table 4 lists the concentrations of proteins and buffers used for crystallization.

Ligand Soaking. The crystal structures of CAII and CAXIII with ligands were obtained by soaking. A 50 mM solution of each ligand was prepared in dimethyl sulfoxide. One μ L of this solution was then diluted using 50 μ L of matching reservoir solution corresponding to the conditions under which the crystal was formed. Crystals were incubated for up to 1 week in the soaking solution.

Cocrystallization. The crystal structures of CAIX with ligands were obtained by cocrystallization. Table 4 lists crystallization conditions. The ligand used for cocrystallization was in 5–10 mM concentration, while the stock solution contained 100 mM ligand dissolved in dimethyl sulfoxide.

Data Collection and Structure Determination. Data for CAII and CAXII were collected at PETRA III BEAMLINE P13 (MX1) and for CAIX at BESSY II beamline 14.1.

The data were processed and scaled using XDS,⁵⁸ MOSFLM,^{59,60} and SCALA.⁶¹ The crystal structures were solved by molecular replacement using MOLREP.⁶² The initial model for molecular replacement—CAII: 3HLJ; CAIX: 8QJ8,⁶³ CAXIII: 2NNO. The structure was refined by REFMAC⁶⁴ and remodeled using COOT.⁵⁷ The 3D models of compounds were constructed by the AVOGADRO⁶⁵ program and ligand parameter files were created using LIBCHECK.^{66,67}

The data diffraction and final model refinement statistics and PDB IDs are summarized in Table 3. All graphics were created using PyMOL Molecular Graphics System. Authors will release the atomic coordinates upon article publication.

Molecular Docking. The following receptors PDB IDs were chosen for docking: CAII: 3HS4; CAIX: 6G9U, chain A; CAXIII: 4KNN, chain A. The receptors selected from the Protein Data Bank

differed from the new X-ray structures presented in this Paper to decrease bias. The proteins were prepared for docking using ChimeraX (version 1.9).^{68–70} The ligands were optimized using the MMFF94s force field^{71–76} within Avogadro molecular viewer (version 1.2.0).⁶⁵ For series 7 ligands, two enantiomers of the chiral sulfur were created and docked separately. The format conversions were performed using OpenBabel (The Open Babel Package, version 3.1.1, <http://openbabel.org>).⁷⁷ The docking was performed using the Smina program (version master:dc3dfab+).³⁴ Smina is based on Autodock Vina.³⁶ The constrained optimization was done using Smina, using a custom scoring function with a quadratic bias function with weight $w=-10$ added to the default Vina scoring function.³⁶ The quadratic constraint forced sulfonamide nitrogen to adhere to its position in the X-ray structure. A cubic docking box of size (24 Å),³ exhaustiveness 100, and energy range 10 kcal/mol was set as docking parameters. The resulting poses were rescored with Smina using the Vinardo scoring function³⁵ without the constraint. Only one of the symmetry equivalent poses (e.g., phenyl ring flip) was included when evaluating pose ranking after docking. Heavy atom Root Mean Square Deviations (RMSD) were calculated using DockRMSD software (version 1.1).⁷⁸

The quantum Density Functional Theory (DFT) optimizations were performed using GAMESS-US (version 2019.R2)⁴⁰ using DFT functional ω B97X-D⁷⁹ with the cc-pVDZ basis set^{80,81} and the C-PCM implicit solvation model for water.⁸² The conformational search was performed using CREST software (version 3.0.2)³⁸ using xTB (version 6.7.1)⁸³ computational engine, employing the GFN2-xTB semiempirical tight binding method³⁹ and the ALPB implicit solvation model for water.⁸⁴

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jmedchem.5c01421>.

FTSA data of compound binding to CA isozymes; SFA data of CA isozyme inhibition by compounds; determination of pK_a values of compound – RSO_2NH_2 ; NMR spectra of compounds; HPLC chromatograms of representative compounds; ESI-MS spectra of representative compounds; and comparison of docking and crystallographic binding poses (PDF)

SMILE strings of the synthesized compound and corresponding biological activity: *observed and intrinsic* ($K_{d,obs}$ and $K_{d,int}$) dissociation constants (in nM units) of investigated compounds to all catalytically active human CAs at 37 °C obtained by FTSA (CSV)

■ AUTHOR INFORMATION

Corresponding Author

Daumantas Matulis – Department of Biothermodynamics and Drug Design, Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius LT-10257, Lithuania;
✉ orcid.org/0000-0002-6178-6276;
Email: daumantas.matulis@bti.vu.lt

Authors

Audrius Zakšauskas – Department of Biothermodynamics and Drug Design, Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius LT-10257, Lithuania

Vaida Paketurytė-Latvė – Department of Biothermodynamics and Drug Design, Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius LT-10257, Lithuania;
✉ orcid.org/0000-0003-0919-7826

Alberta Jankūnaitė – Department of Biothermodynamics and Drug Design, Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius LT-10257, Lithuania

Edita Čapkauskaitė – Department of Biothermodynamics and Drug Design, Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius LT-10257, Lithuania

Yann Becart – Department of Biothermodynamics and Drug Design, Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius LT-10257, Lithuania

Alexey Smirnov – Department of Biothermodynamics and Drug Design, Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius LT-10257, Lithuania

Klára Pospíšilová – Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague 6 16610, Czech Republic

Janis Leitans – Latvian Biomedical Research and Study Centre, Riga LV-1067, Latvia

Jiří Brynda – Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague 6 16610, Czech Republic

Andris Kazaks – Latvian Biomedical Research and Study Centre, Riga LV-1067, Latvia

Lina Baranauskienė – Department of Biothermodynamics and Drug Design, Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius LT-10257, Lithuania;
✉ orcid.org/0000-0002-9924-9177

Elena Manakova – Department of Protein–DNA Interactions, Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius LT-10257, Lithuania

Saulius Gražulis – Sector of Crystallography and Cheminformatics, Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius LT-10257, Lithuania

Visvaldas Kairys – Department of Bioinformatics, Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius LT-10257, Lithuania; ✉ orcid.org/0000-0002-5427-0175

Kaspars Tars – Latvian Biomedical Research and Study Centre, Riga LV-1067, Latvia; ✉ orcid.org/0000-0001-8421-9023

Pavlna Rezáčová – Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague 6 16610, Czech Republic

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acs.jmedchem.5c01421>

Author Contributions

A.Z. and V.P.-L. contributed equally to this paper.

Notes

The authors declare the following competing financial interest(s): Authors declare that they have patents and patent applications pending on CA inhibitors.

■ ACKNOWLEDGMENTS

This research was supported by (i) a grant from the Research Council of Lithuania (No. S-MIP-22-35); (ii) the National Institute for Cancer Research (Programme EXCELES, ID Project No. LX22NPO5102), Funded by the European Union, Next Generation EU awarded to PR; and (iii) the European Union's Recovery and Resilience Mechanism project (No.5.2.1.i.i.0/2/24/I/CFLA/001) "Consolidation of the Latvian Institute of Organic Synthesis and the Latvian Biomedical Research and Study Centre". The authors also thank Martynas Malikėnas for performing HPLC analysis of compounds. We thank Vilma Michailovienė¹, who died in 2022, for performing numerous protein affinity purification procedures.

■ ABBREVIATIONS

CA, human carbonic anhydrase; FTSA, fluorescent thermal shift assay (or differential scanning fluorimetry DSF); int, intrinsic; $K_{d,int}$, intrinsic equilibrium dissociation constant; $K_{d,obs}$, observed equilibrium dissociation constant; obs, observed; $pK_{a,CA}$, pK_a value of water molecule bound to Zn(II) in the active site of CA; $pK_{a,SA}$, pK_a of the sulfonamide group; SFA, stopped-flow carbon dioxide hydration assay

■ REFERENCES

- (1) Whitesides, G. M.; Krishnamurthy, V. M. Designing ligands to bind proteins. *Q. Rev. Biophys.* **2005**, *38*, 385–395.
- (2) Maren, T. H. Use of inhibitors in physiological studies of carbonic anhydrase. *Am. J. Physiol.-Ren. Physiol.* **1977**, *232*, F291–F297.
- (3) Meldrum, N. U.; Roughton, F. J. Some properties of carbonic anhydrase, the CO₂ enzyme present in blood. *J. Physiol* **1932**, *75*, 15P–16P.
- (4) *The Carbonic Anhydrases*; Springer US: Boston, MA, 1991.
- (5) *The Carbonic Anhydrases*; Birkhäuser Basel: Basel, 2000.
- (6) Krishnamurthy, V. M.; et al. Carbonic Anhydrase as a Model for Biophysical and Physical-Organic Studies of Proteins and Protein–Ligand Binding. *Chem. Rev.* **2008**, *108*, 946–1051.
- (7) *The Carbonic Anhydrases: Current and Emerging Therapeutic Targets*; Springer International Publishing: Cham, 2021; Vol. 75.
- (8) Zamanova, S.; Shabana, A. M.; Mondal, U. K.; Ilies, M. A. Carbonic anhydrases as disease markers. *Expert Opin. Ther. Pat.* **2019**, *29*, 509–533.
- (9) *Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications*; Springer: Netherlands, Dordrecht, 2014; Vol. 75.
- (10) Abbate, F.; et al. Nonaromatic sulfonamide group as an ideal anchor for potent human carbonic anhydrase inhibitors: role of hydrogen-bonding networks in ligand binding and drug design. *J. Med. Chem.* **2002**, *45*, 3583–3587.
- (11) Martin, D. P.; Hann, Z. S.; Cohen, S. M. Metalloprotein–Inhibitor Binding: Human Carbonic Anhydrase II as a Model for Probing Metal–Ligand Interactions in a Metalloprotein Active Site. *Inorg. Chem.* **2013**, *52*, 12207–12215.
- (12) Carta, F.; Supuran, C. T.; Scozzafava, A. Sulfonamides and their isomers as carbonic anhydrase inhibitors. *Future Med. Chem.* **2014**, *6*, 1149–1165.
- (13) Fox, J. M.; Zhao, M.; Fink, M. J.; Kang, K.; Whitesides, G. M. The Molecular Origin of Enthalpy/Entropy Compensation in Biomolecular Recognition. *Annu. Rev. Biophys.* **2018**, *47*, 223–250.
- (14) Aggarwal, M.; Boone, C. D.; Kondeti, B.; McKenna, R. Structural annotation of human carbonic anhydrases. *J. Enzyme Inhib. Med. Chem.* **2013**, *28*, 267–277.
- (15) Supuran, C. T. Structure and function of carbonic anhydrases. *Biochem. J.* **2016**, *473*, 2023–2032.
- (16) Lomelino, C. L.; Andring, J. T.; McKenna, R. Crystallography and Its Impact on Carbonic Anhydrase Research. *Int. J. Med. Chem.* **2018**, *2018*, No. 9419521.
- (17) Smirnovienė, J.; et al. Switching the Inhibitor-Enzyme Recognition Profile via Chimeric Carbonic Anhydrase XII. *ChemistryOpen* **2021**, *10*, 567–580.
- (18) Scott, A. D.; et al. Thermodynamic Optimisation in Drug Discovery: A Case Study using Carbonic Anhydrase Inhibitors. *ChemMedChem.* **2009**, *4*, 1985–1989.
- (19) Glöckner, S.; Ngo, K.; Wagner, B.; Heine, A.; Klebe, G. The Influence of Varying Fluorination Patterns on the Thermodynamics and Kinetics of Benzenesulfonamide Binding to Human Carbonic Anhydrase II. *Biomolecules* **2020**, *10*, 509.
- (20) Zakšauskas, A.; et al. Methyl 2-Halo-4-Substituted-5-Sulfamoyl-Benzoates as High Affinity and Selective Inhibitors of Carbonic Anhydrase IX. *Int. J. Mol. Sci.* **2021**, *23*, 130.
- (21) Zakšauskas, A.; et al. Design of two-tail compounds with rotationally fixed benzenesulfonamide ring as inhibitors of carbonic anhydrases. *Eur. J. Med. Chem.* **2018**, *156*, 61–78.
- (22) Zakšauskas, A.; et al. Halogenated and di-substituted benzenesulfonamides as selective inhibitors of carbonic anhydrase isoforms. *Eur. J. Med. Chem.* **2020**, *185*, No. 111825.
- (23) MERCAPTANS FROM THIOKETALS: CYCLODODECYL MERCAPTAN. *Org. Synth.* **61**, 74 (1983).
- (24) Mahalingam, S. M.; Chu, H.; Liu, X.; Leamon, C. P.; Low, P. S. Carbonic Anhydrase IX-Targeted Near-Infrared Dye for Fluorescence Imaging of Hypoxic Tumors. *Bioconjugate Chem.* **2018**, *29*, 3320–3331.
- (25) Smirnovienė, J.; Smirnovas, V.; Matulis, D. Picomolar inhibitors of carbonic anhydrase: Importance of inhibition and binding assays. *Anal. Biochem.* **2017**, *522*, 61–72.
- (26) Taylor, P. W.; King, R. W.; Burgen, A. S. V. Influence of pH on the kinetics of complex formation between aromatic sulfonamides and human carbonic anhydrase. *Biochemistry* **1970**, *9*, 3894–3902.
- (27) Zubrienė, A.; Matulis, D. Observed Versus Intrinsic Thermodynamics of Inhibitor Binding to Carbonic Anhydrases. In *Carbonic Anhydrase as Drug Target: Thermodynamics and Structure of Inhibitor Binding*; Matulis, D., Ed.; Springer International Publishing: Cham, 2019; pp 107–123.
- (28) Fisher, S. Z.; Aggarwal, M.; Kovalevsky, A. Y.; Silverman, D. N.; McKenna, R. Neutron Diffraction of Acetazolamide-Bound Human Carbonic Anhydrase II Reveals Atomic Details of Drug Binding. *J. Am. Chem. Soc.* **2012**, *134*, 14726–14729.
- (29) Aggarwal, M.; et al. Neutron structure of human carbonic anhydrase II in complex with methazolamide: mapping the solvent and hydrogen-bonding patterns of an effective clinical drug. *IUCr.* **2016**, *3*, 319–325.
- (30) Zubrienė, A.; et al. Intrinsic thermodynamics of 4-substituted-2,3,5,6-tetrafluorobenzenesulfonamide binding to carbonic anhydrases by isothermal titration calorimetry. *Biophys. Chem.* **2015**, *205*, 51–65.
- (31) Dudutienė, V.; et al. Isoform-Selective Enzyme Inhibitors by Exploring Pocket Size According to the Lock-and-Key Principle. *Biophys. J.* **2020**, *119*, 1513–1524.
- (32) Eisenberg, D.; Schwarz, E.; Komaromy, M.; Wall, R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **1984**, *179*, 125–142.
- (33) Supuran, C. T. How many carbonic anhydrase inhibition mechanisms exist? *J. Enzyme Inhib. Med. Chem.* **2016**, *31*, 345–360.
- (34) Koes, D. R.; Baumgartner, M. P.; Camacho, C. J. Lessons Learned in Empirical Scoring with smina from the CSAR 2011 Benchmarking Exercise. *J. Chem. Inf. Model.* **2013**, *53*, 1893–1904.
- (35) Quiroga, R.; Villarreal, M. A. Vinardo: A Scoring Function Based on Autodock Vina Improves Scoring, Docking, and Virtual Screening. *PLoS One* **2016**, *11*, No. e0155183.
- (36) Trott, O.; Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.
- (37) McNutt, A. T.; et al. GNINA 1.0: molecular docking with deep learning. *J. Cheminf.* **2021**, *13*, 43.
- (38) Pracht, P.; et al. CREST-A program for the exploration of low-energy molecular chemical space. *J. Chem. Phys.* **2024**, *160*, 114110.
- (39) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.
- (40) Schmidt, M. W.; et al. General atomic and molecular electronic structure system. *J. Comput. Chem.* **1993**, *14*, 1347–1363.
- (41) Summa, C. M.; Langford, D. P.; Dinshaw, S. H.; Webb, J.; Rick, S. W. Calculations of Absolute Free Energies, Enthalpies, and Entropies for Drug Binding. *J. Chem. Theory Comput.* **2024**, *20*, 2812–2819.
- (42) Linkuvienė, V.; et al. Thermodynamic, kinetic, and structural parameterization of human carbonic anhydrase interactions toward enhanced inhibitor design. *Q. Rev. Biophys.* **2018**, *51*, No. e10.
- (43) Baronas, D.; et al. Inhibitor binding to metal-substituted metalloenzyme: Sulfonamide affinity for carbonic anhydrase IX. *J. Inorg. Biochem.* **2024**, *256*, No. 112547.

- (44) Jackman, G. B.; Petrow, V.; Stephenson, O.; Wild, A. M. Studies in the Field of Diuretic Agents. Part VI Some Sulphamoylbenzoic Acids. *J. Pharm. Pharmacol.* **1962**, *14*, 679–686.
- (45) Jackman, G. B.; Petrow, V.; Stephenson, O.; Wild, A. M. Studies in the Field of Diuretic Agents. *J. Pharm. Pharmacol.* **1963**, *15*, 202–211.
- (46) Englert, H. C.; Lang, H.-J.; Linz, W.; Schoelkens, B.; Scholz, W. Benzoylguanidines, a Process for Their Preparation, Their Use as Medicaments and Medicaments Containing Them; 1992.
- (47) Mickevičiūtė, A. et al. Recombinant Production of 12 Catalytically Active Human CA Isoforms. In *Carbonic Anhydrase as Drug Target: Thermodynamics and Structure of Inhibitor Binding*; Matulis, D., Ed.; Springer International Publishing: Cham, 2019; pp 15–37.
- (48) Leitans, J.; et al. Efficient Expression and Crystallization System of Cancer-Associated Carbonic Anhydrase Isoform IX. *J. Med. Chem.* **2015**, *58*, 9004–9009.
- (49) Grüner, B.; et al. Metallocarborane Sulfamides: Unconventional, Specific, and Highly Selective Inhibitors of Carbonic Anhydrase IX. *J. Med. Chem.* **2019**, *62*, 9560–9575.
- (50) Kazlauskas, E.; Petrauskas, V.; Paketytytė, V.; Matulis, D. Standard operating procedure for fluorescent thermal shift assay (FTSA) for determination of protein–ligand binding and protein stability. *Eur. Biophys. J.* **2021**, *50*, 373–379.
- (51) Gedgaudas, M.; Baronas, D.; Kazlauskas, E.; Petrauskas, V.; Matulis, D. Thermott: A comprehensive online tool for protein–ligand binding constant determination. *Drug Discovery Today* **2022**, *27*, 2076–2079.
- (52) Lingé, D.; et al. PLBD: protein–ligand binding database of thermodynamic and kinetic intrinsic parameters. *Database* **2023**, *2023*, baad040.
- (53) Khalifah, R. G. The carbon dioxide hydration activity of carbonic anhydrase. I. Stop-flow kinetic studies on the native human isoenzymes B and C. *J. Biol. Chem.* **1971**, *246*, 2561–2573.
- (54) Williams, J. W.; Morrison, J. F. The kinetics of reversible tight-binding inhibition. In *Methods in Enzymology*; Academic Press: 1979; Vol. 63, pp 437–467.
- (55) Cheng, Y.-C.; Prusoff, W. H. Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50% inhibition (I_{50}) of an enzymatic reaction. *Biochem. Pharmacol.* **1973**, *22*, 3099–3108.
- (56) Snyder, P. W.; et al. Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 17889–17894.
- (57) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Biol. Crystallogr.* **2004**, *60*, 2126–2132.
- (58) Kabsch, W. XDS. *Acta Crystallogr. Sect. D* **2010**, *D66*, 125–132.
- (59) Battice, T. G. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. W. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 271–281.
- (60) Leslie, A. G. W. The integration of macromolecular diffraction data. *Acta Crystallogr. Biol. Crystallogr.* **2006**, *62*, 48–57.
- (61) Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 72–82.
- (62) Vagin, A.; Teplyakov, A. MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **1997**, *30*, 1022–1025.
- (63) Leitans, J.; et al. Structural Basis of Saccharin Derivative Inhibition of Carbonic Anhydrase IX. *ChemMedChem* **2023**, *18*, No. e202300454.
- (64) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Biol. Crystallogr.* **1997**, *53*, 240–255.
- (65) Hanwell, M. D.; et al. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminf.* **2012**, *4*, 17.
- (66) Vagin, A. A.; et al. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. Biol. Crystallogr.* **2004**, *60*, 2184–2195.
- (67) Lebedev, A. A.; et al. JLigand: a graphical tool for the CCP4 template-restraint library. *Acta Crystallogr. D Biol. Crystallogr.* **2012**, *68*, 431–440.
- (68) Meng, E. C.; et al. UCSF CHIMERAx: Tools for structure building and analysis. *Protein Sci.* **2023**, *32*, No. e4792.
- (69) Pettersen, E. F.; et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci. Publ. Protein Soc.* **2021**, *30*, 70–82.
- (70) Goddard, T. D.; et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci. Publ. Protein Soc.* **2018**, *27*, 14–25.
- (71) Halgren, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490–519.
- (72) Halgren, T. A. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *J. Comput. Chem.* **1996**, *17*, 520–552.
- (73) Halgren, T. A. Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94. *J. Comput. Chem.* **1996**, *17*, 553–586.
- (74) Halgren, T. A.; Nachbar, R. B. Merck molecular force field. IV. conformational energies and geometries for MMFF94. *J. Comput. Chem.* **1996**, *17*, 587–615.
- (75) Halgren, T. A. Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules. *J. Comput. Chem.* **1996**, *17*, 616–641.
- (76) Halgren, T. A. MMFF VI. MMFF94s option for energy minimization studies. *J. Comput. Chem.* **1999**, *20*, 720–729.
- (77) O’Boyle, N. M.; et al. Open Babel: An open chemical toolbox. *J. Cheminf.* **2011**, *3*, 33.
- (78) Bell, E. W.; Zhang, Y. DockRMSD: an open-source tool for atom mapping and RMSD calculation of symmetric molecules through graph isomorphism. *J. Cheminf.* **2019**, *11*, 40.
- (79) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
- (80) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (81) Woon, D. E.; Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.
- (82) Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998**, *102*, 1995–2001.
- (83) Bannwarth, C.; et al. Extended tight-binding quantum chemistry methods. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2021**, *11*, No. e1493.
- (84) Ehlert, S.; Stahn, M.; Spicher, S.; Grimme, S. Robust and Efficient Implicit Solvation Model for Fast Semiempirical Methods. *J. Chem. Theory Comput.* **2021**, *17*, 4250–4261.

Publication 3

From X-ray crystallographic structure to intrinsic thermodynamics of protein-ligand binding using carbonic anhydrase isozymes as a model system

Vaida Paketurytė-Latvė, Alexey Smirnov, Elena Manakova, Lina Baranauskienė, Vytautas Petrauskas, Asta Zubrienė, Jurgita Matulienė, Virginija Dudutienė, Edita Čapkauskaitė, Audrius Zakšauskas, Janis Leitans, Saulius Gražulis, Kaspars Tars, Daumantas Matulis

IUCrJ., 2024, 11 (4), 556–569

DOI: 10.1107/S2052252524004627



From X-ray crystallographic structure to intrinsic thermodynamics of protein–ligand binding using carbonic anhydrase isozymes as a model system

Vaida Paketurytė-Latvė,^a Alexey Smirnov,^a Elena Manakova,^b Lina Baranauskienė,^a Vytautas Petrauskas,^a Asta Zubrienė,^a Jurgita Matulienė,^a Virginija Dudutienė,^a Edita Čapkauskaitė,^a Audrius Zakšauskas,^a Janis Leitans,^c Saulius Gražulis,^d Kaspars Tars^c and Daumantas Matulis^{a*}

Received 31 January 2024

Accepted 17 May 2024

Edited by Z.-J. Liu, Chinese Academy of Sciences, China

This article is part of a collection of articles from the IUCr 2023 Congress in Melbourne, Australia, and commemorates the 75th anniversary of the IUCr.

Keywords: drug discovery; protein structure; molecular recognition; X-ray crystallography; intermolecular interactions; carbonic anhydrase isozymes; protein–ligand binding; intrinsic thermodynamics; binding assays.

PDB references: chimeric human carbonic anhydrase IX with EA3-3, 8rbp; chimeric human carbonic anhydrase IX with EA5-3, 8rar; carbonic anhydrase II with EA5-3, 8rj2

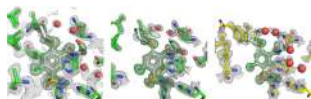
Supporting information: this article has supporting information at www.iucrj.org

^aDepartment of Biothermodynamics and Drug Design, Institute of Biotechnology, Life Sciences Center, Vilnius University, Saulėtekio 7, 10257 Vilnius, Lithuania, ^bDepartment of Protein–DNA Interactions, Institute of Biotechnology, Life Sciences Center, Vilnius University, Saulėtekio 7, 10257 Vilnius, Lithuania, ^cLatvian Biomedical Research and Study Centre, Ratsupites 1 k-1, 1067 Riga, Latvia, and ^dSector of Crystallography and Chemical Informatics, Institute of Biotechnology, Life Sciences Center, Vilnius University, Saulėtekio 7, 10257 Vilnius, Lithuania. *Correspondence e-mail: daumantas.matulis@bti.vu.lt

Carbonic anhydrase (CA) was among the first proteins whose X-ray crystal structure was solved to atomic resolution. CA proteins have essentially the same fold and similar active centers that differ in only several amino acids. Primary sulfonamides are well defined, strong and specific binders of CA. However, minor variations in chemical structure can significantly alter their binding properties. Over 1000 sulfonamides have been designed, synthesized and evaluated to understand the correlations between the structure and thermodynamics of their binding to the human CA isozyme family. Compound binding was determined by several binding assays: fluorescence-based thermal shift assay, stopped-flow enzyme activity inhibition assay, isothermal titration calorimetry and competition assay for enzyme expressed on cancer cell surfaces. All assays have advantages and limitations but are necessary for deeper characterization of these protein–ligand interactions. Here, the concept and importance of intrinsic binding thermodynamics is emphasized and the role of structure–thermodynamics correlations for the novel inhibitors of CA IX is discussed – an isozyme that is overexpressed in solid hypoxic tumors, and thus these inhibitors may serve as anticancer drugs. The abundant structural and thermodynamic data are assembled into the Protein–Ligand Binding Database to understand general protein–ligand recognition principles that could be used in drug discovery.

1. Introduction

Structural biology plays a central role in drug discovery (Renaud *et al.*, 2016; Nasim & Qureshi, 2022). Small-molecular-weight organic compounds could be biologically active due to binding to proteins and other macromolecules. Such compounds are often found in nature or could be rationally designed and synthesized. According to the drug-discovery paradigm, the first hypothesis points to a protein involved in a disease progression (Huang *et al.*, 2012; Liu *et al.*, 2022) and an assumption that an inhibitor could alter the disease. A compound is then designed to bind the protein and possibly inhibit or activate its enzymatic action (Minetti & Remeta, 2022). A crucial step to understand the mode of ligand binding and further optimize the interaction is to determine the atomic resolution structure of the protein in complex with the bound compound. This identifies the mode of interaction and the functional groups that determine the affinity of the compound



Published under a CC BY 4.0 licence

binding to the protein target (Timothy & Sanjeev, 2007; Maveyraud & Mourey, 2020). Although the complex structure is usually determined by X-ray crystallography, this method does not directly provide information on compound affinity (Schlichting, 2005; Danley, 2006). Biophysical assays are necessary to determine thermodynamic and kinetic parameters of protein–ligand interaction.

Carbon dioxide is a cellular metabolism product that must be eliminated from the body. Due to the limited carbon dioxide solubility, bicarbonate serves as the primary mode of CO₂ transport in the blood (Alka & Casey, 2014; Occhipinti & Boron, 2019). Therefore, rapid conversion between these two forms is critical (Casey, 2006). Although these conversions can occur spontaneously, their natural rate is insufficient to meet the needs of an organism. The discovery of the catalyst responsible for the reaction between carbon dioxide and bicarbonate came with the isolation of carbonic anhydrase (CA, EC4.2.1.1.) from bovine blood in 1932 (Meldrum & Roughton, 1933). This enzyme is essential for accelerating the conversion of carbon dioxide to bicarbonate ions and protons, increasing the rate up to a million-fold.

It took about 40 years of CA research before Liljas *et al.* (1972) determined the first crystal structure of CA II, now considered a model protein, to a resolution of 2.0 Å. This and later crystallographic studies revealed that the CA II molecule has an ellipsoidal shape, measuring ≈ 40 Å $\times$ 40 Å $\times$ 40 Å. The cone-shaped active site has a depth of about 15 Å, with the catalytic zinc located at the base and coordinated by three histidine residues. The enzyme core comprises a ten-stranded β sheet surrounded by seven right-handed α helices. The active site of CA II has two sides of different polarity: one with hydrophobic residues (Ile91, Val121, Phe131, Val135, Leu141, Val143, Leu198, Pro202, Leu204, Val207 and Trp209) and the other with hydrophilic residues (Asn62, His64, Asn67, Gln92, Thr199 and Thr200). The hydrophobic side hosts the substrate CO₂ binding site and the hydrophilic residues help stabilize the ordered water network essential for proton transfer (Lomelino *et al.*, 2018).

CA occurs in various structural forms in different organisms. To date, researchers have identified eight structural classes of CAs, designated by Greek letters: α , β , γ , δ , ζ , η , θ and ι (Jensen *et al.*, 2019). Humans have 15 α -CA isozymes that regulate pH and fluid balance, and they are involved in various physiological and pathological processes (Imtaiyaz Hassan *et al.*, 2013; Frost, 2014). The 15 isozymes differ in tissue distribution and have various catalytic activities and cellular localization. However, they share a similar overall structure, with only minor differences in their active site (Lomelino *et al.*, 2018). Several human CA isozymes are recognized as therapeutic targets in glaucoma, edema, epilepsy, obesity and cancer (Cicone *et al.*, 2021; Mboge *et al.*, 2018; Lomelino & McKenna, 2016), but the design of selective inhibitors for specific isozymes remains challenging due to the high structural similarity.

The first-generation drug to treat/control disease using CA inhibition was acetazolamide, which entered medical use in 1953 and is still used as a systemic medication for glaucoma,

altitude sickness, epilepsy, idiopathic intracranial hypertension and other conditions (Rankin, 2007; Scott & Njardarson, 2018). Over the next 70 years, more than 20 CA inhibitors entered clinical practice. However, none of them show significant CA isozyme selectivity. Off-target binding can lead to undesired effects since CAs participate in crucial metabolic processes. Furthermore, off-target binding decreases the available drug fraction, thus necessitating a higher dose. Although many chemical classes of compounds have been shown to inhibit CA activity (Lomelino *et al.*, 2016; Singh *et al.*, 2018; Kumar *et al.*, 2021), all clinically used CA inhibitors have either a sulfonamide or its bioisostere group that directly interacts with the catalytic zinc and displaces the zinc-bound water molecule (Swenson, 2014).

In this study, we describe thermodynamic analysis of CA IX-selective compounds binding to their target and off-target isozymes. CA IX is the CA isozyme that has direct relevance to cancer – its expression is induced in hypoxia under the control of HIF-1 (Pastorek *et al.*, 1994; Saarnio *et al.*, 1998). In tumors, CA IX helps cancer cells survive under hypoxic conditions, while in healthy tissues, its expression is limited to the gastrointestinal tract (Becker, 2020).

Numerous assays may be applied to measure protein–ligand binding interactions, described for CA II isozyme as a model protein (Krishnamurthy *et al.*, 2008). We primarily use fluorescence-based thermal-shift assay (TSA or FTSA) and consider it the most robust and accurate method to determine the K_d of protein–ligand interactions because it uniformly covers a wide range from millimolar to picomolar affinities. We described TSA previously and discussed its advantages and limitations over the other orthogonal binding assays (Matulis *et al.*, 2005; Cimmperman *et al.*, 2008; Petrauskas *et al.*, 2024). We have made a web-based service called *Thermott* that analyses user-supplied thermal-shift data, determines the binding parameters, and thus simplifies TSA data analysis (Gedgaudas *et al.*, 2022). In the literature, enzyme activity inhibition is the primary assay to describe sulfonamide binding to CA isozymes. However, this assay has limitations that prevent the discovery of tight picomolar inhibitors, as discussed below.

Protein–ligand interactions often occur simultaneously with the linked protonation–deprotonation reactions (Bahr *et al.*, 2023; Huang *et al.*, 2012; Hognon *et al.*, 2022). Such reactions hide the actual thermodynamic parameters of interaction and may be misinterpreted when studying the underlying forces in the protein–ligand complex (Bradshaw & Waksman, 1998). To dissect binding-energy contributions from these linked protonation–deprotonation reactions, we emphasize the importance of the intrinsic energetics. Fig. 1 shows a simplified concept to understand why systematic errors often occur in the thermodynamic interpretation of the binding data.

The example in Fig. 1 shows that any experimental technique will determine the K_d (dissociation constant) equal to 100 nM, a relatively weak affinity for the drug candidate. However, the user incorrectly assumed that the entire 1 μ M of inhibitor, applied in the assay, participated in the reaction. In reality, only 1 nM of the compound was in the binding-ready

form. If the concentration of the binding-ready form was 1 μM , then the K_d would be 0.1 nM and the inhibitor would be considered as a high-affinity inhibitor of the enzyme. Only the intrinsic values should be used for structure–thermodynamics correlations (Gomez & Freire, 1997; Krishnamurthy *et al.*, 2007, 2008).

The example in Fig. 1 is simplified, excluding the effect of protein protonation on the observed binding parameters. It was previously shown by applying biophysical techniques that the binding of sulfonamides to CA isozymes has a U-shape dependence on pH. Later, by applying neutron-diffraction crystallography (Fisher *et al.*, 2012), it was confirmed that sulfonamide binds in deprotonated anionic form to the Zn-bound water form of CA. The pK_a of the Zn-bound water in the active site of 12 human CA isozymes is between 6 and 7, and only the pK_a s of CA I and CA XIII are 8 (Linkuvienė *et al.*, 2018b). Thus, the majority fraction of CA I and CA XIII is in its protonated form, capable of binding the deprotonated sulfonamide. In contrast, benzenesulfonamides have pK_a s between 8 and 11. Therefore, at pH 7 only a small fraction is in the anionic form, capable of binding CA, and the observed affinity values mostly depend on the sulfonamide protonation.

To make it easier to analyze abundant data scattered among numerous publications, we assembled a database of the X-ray crystal structures of CA–ligand complexes and the intrinsic thermodynamics of binding (Lingė *et al.*, 2023). It is the only database that systematically analyzes intrinsic thermodynamics and kinetics of protein–ligand binding. The freely accessible database could be used when analyzing protein–ligand recognition principles and applied for *in silico* drug discovery.

In this article, we performed correlation analysis between several new crystal structures and the thermodynamics of binding of two groups of compounds, fluorinated (VD series) and chlorinated (EA series) benzenesulfonamides. Their binding to 12 human CA isozymes was studied with the goal to determine the factors that are important for the affinity and selectivity towards particular CA isozymes, especially the cancer-associated CA IX.

2. Results

2.1. X-ray crystal structures of sulfonamide binding to carbonic anhydrases

Hundreds of X-ray crystallographic structures of free or inhibitor-bound CAs from humans and other organisms have been solved at atomic resolution. The main class of CA inhibitors is the compounds bearing the primary sulfonamide functional group. Such compounds bind to CA isozymes by making a coordination bond between the negatively charged sulfonamide amino group nitrogen and the Zn^{2+} ion. Our research group has solved 147 X-ray crystal structures of various sulfonamide compounds bound to different human CA isozymes and deposited them in the Protein Data Bank (PDB). The sulfonamide amino group made the coordination bond with the Zn^{2+} ion in all cases. However, the compounds bound the proteins with highly variable affinities, ranging from barely detectable millimolar affinity to extremely high picomolar affinity. In extreme cases, some compounds exhibited billion-fold differences in the value of the dissociation constant K_d . Therefore, despite the relatively similar recog-

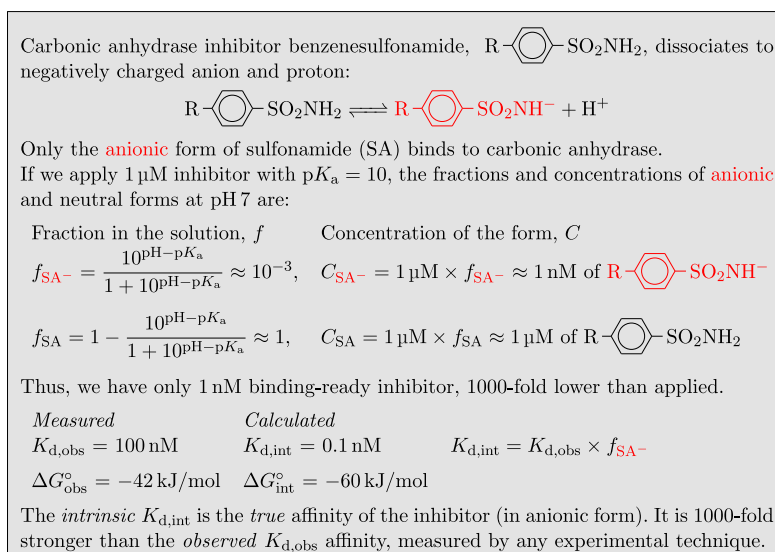


Figure 1

The concept of *intrinsic* and *observed* dissociation constant explained using benzenesulfonamide deprotonization reaction and fractions of binding-ready species.

Table 1
X-ray crystallographic data-collection and refinement statistics.

All datasets were collected at 100 K and the test set size was 10%.

Isozyme–ligand PDB ID	chCA IX–EA3-3 8rbp	chCA IX–EA5-3 8rar	CA II–EA5-3 8rj2
Data-collection statistics			
Space group	$P2_1$	$P2_1$	$P2_1$
Unit-cell parameters (Å)	$a = 42.12, b = 41.31, c = 72.12, \alpha = \gamma = 90, \beta = 104.3$	$a = 42.13, b = 41.27, c = 72.40, \alpha = \gamma = 90, \beta = 104.3$	$a = 42.07, b = 41.15, c = 72.03, \alpha = \gamma = 90, \beta = 104.2$
Resolution range (Å)	69.88–1.15	70.15–1.15	40.79–1.12
Wavelength (Å)	0.97552	0.97552	0.97552
Radiation source	EMBL beamline P14 at PETRA III storage ring (DESY, Hamburg, DE)	EMBL beamline P14 at PETRA III storage ring (DESY, Hamburg, DE)	EMBL beamline P14 at PETRA III storage ring (DESY, Hamburg, DE)
Unique reflections, overall (outer shell)	81171 (9330)	81933 (9111)	81344 (7994)
R_{merge} , overall (outer shell)	0.049 (0.237)	0.118 (1.335)	0.046 (0.145)
$\langle I/\sigma \rangle$, overall (outer shell)	17.4 (5.6)	5.8 (1.1)	25.1 (10.6)
Multiplicity, overall (outer shell)	6.5 (5.2)	6.5 (4.8)	6.9 (6.2)
Completeness (%), overall (outer shell)	95.1 (75.2)	94.7 (72.7)	89.3 (60.4)
Wilson B factor	10.863	10.741	9.087
Refinement statistics			
R_{work}	0.11768	0.14114	0.14893
R_{free}	0.14572	0.18011	0.17601
RMSD bond lengths (Å)	0.013	0.013	0.017
RMSD bond angles (°)	1.857	1.855	2.272
B factors			
Atoms in the structure			
All	19.99	21.97	17.58
Main chain	15.71	17.59	14.43
Side chain	20.44	23.18	17.88
Inhibitor	17.34	20.65	25.20
Waters	33.23	36.79	28.5
Zinc	8.45	10.3	7.27
Other molecules	31.05	26.13	29.36
Atoms in the structure (non- hydrogen)			
All	2583	2455	2459
Protein	2227	2189	2152
Inhibitor	24	25	25
Water	311	236	266
Zinc	1	1	1
Other molecules	20	4	15
Ramachandran statistics			
Favored regions (%)	95	97	97
Allowed regions (%)	5	3	3
Outliers	0	0	0

nitron mode via the coordination bond, the compound’s exact chemical structure significantly influenced the affinity.

Here, we describe CA II and CA IX recognition details of two structurally highly similar sulfonamide compounds, EA3-3 and EA5-3. Crystallization of CA IX is rather challenging and thus we used a mimic of CA IX, called chimeric CA IX (chCA IX). The chimeric protein is a multiple mutant of CA II, where six amino acids in the active site of CA II are replaced by the amino acids present in CA IX at the same positions. We have previously demonstrated that such chimeric protein mimics actual CA IX to a rather good approximation (Dudutienė *et al.*, 2014).

Table 1 shows the refinement statistics of the three crystal structures, while the structural arrangement of the compounds

in the active sites and the electron-density maps are shown in Fig. 2. The sulfonamide amino group formed the coordination bond with the Zn²⁺ ion in all three structures. The orientations of compounds were similar in CA II and chCA IX, but the position of cyclohexane was essential for the recognition of CA IX. The crystal structures provided positions of all atoms in the complex. However, they did not provide information on the affinity or other thermodynamic parameters of binding, which will be described in the following sections.

2.2. Assays to measure protein–ligand binding *in vitro*

Numerous techniques could be used to determine protein–ligand binding affinity, or in other words, the change in the

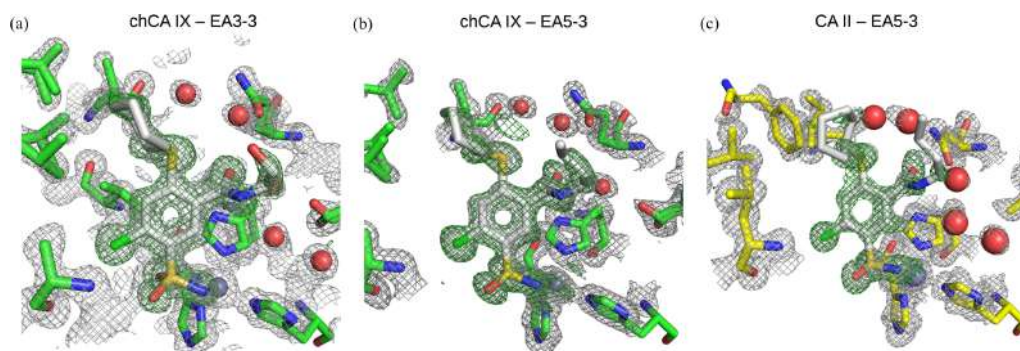


Figure 2

X-ray crystallographic structure determination of inhibitors EA3-3 and EA5-3 binding to CA II and chimeric CA IX (a multiple amino acid mutant of CA II mimicking amino acids of the CA IX active site). The electron density of the inhibitors is shown in green. Amino acids that were mutated in CA II to reproduce the CA IX active site are also shown, as well as histidine residues coordinating the Zn^{2+} ion. Electron-density maps were calculated in the absence of inhibitors. The $2F_o - F_c$ map contoured at 1σ is gray, while the difference map $F_o - F_c$ contoured at 3σ is green. Residues of chCA IX are colored green and CA II is yellow. The inhibitors are light gray. The Zn^{2+} ion is shown as a gray sphere and water molecules are shown as red spheres.

standard Gibbs energy upon ligand binding to protein. Other thermodynamic parameters – such as binding enthalpy, entropy, heat capacity, volume and compressibility – may also provide insights into the interaction mechanism. In addition to thermodynamic parameters, the binding kinetics may also be crucial for the understanding of recognition principles, as they show the reaction order, the approach of the on-rate to the diffusion-limited rates, and the stability of the protein–ligand complex (the residence time or off-rate).

To determine the structure–thermodynamics correlations, we primarily perform the TSA, the stopped-flow assay of enzymatic activity inhibition (SFA) and isothermal titration calorimetry (ITC). We consider the TSA as the most robust and easy-to-use assay, spanning the largest dynamic range of affinities among all mentioned assays. TSA provides accurate affinities from millimolar to picomolar K_d . However, TSA does not provide information on whether the compound inhibits the enzymatic activity. The SFA is the most applicable for this purpose, but it spans a lower affinity range than the TSA. ITC is indispensable when we need to determine the enthalpy, entropy and heat capacity of binding.

To determine the affinity of sulfonamide binding to CA IX expressed in cancer cells, we developed an assay where we applied the competitive binding model where a fluorescently labeled high-affinity compound competes with an unlabeled one for the same binding site of CA IX. Fig. 3 shows the determination of fluorinated compound VD10-13 [2,3,5,6-tetrafluoro-4-[(2-hydroxyethyl)thio]benzenesulfonamide] binding to CA IX by using all four methods (TSA, SFA, ITC and the competition assay).

2.2.1. Fluorescence-based thermal-shift assay. The TSA is based on the observation that the bound ligand increases the protein's melting temperature in a dose-dependent manner. A purified protein has a relatively well defined temperature at which it unfolds and denatures. The addition of a stoichiometric amount of high-affinity ligand stabilizes the complex

and increases the melting temperature. However, a tenfold increase in ligand concentration above the stoichiometric ratio will further increase the melting temperature. Therefore, the protein–ligand complex does not possess a well-defined melting temperature (Fukada *et al.*, 1983; Matulis *et al.*, 2005).

Fig. 3(a) shows how the affinity of the VD10-13 compound for purified recombinant CA IX was determined by the TSA method. The inset shows the fluorescence dependencies on the temperature at various added compound concentrations. There is a shift of T_m upon an increase in ligand concentration. There is no need to obtain datapoints at ligand concentrations below the protein concentration. Only the datapoints that determine the shift are important for the K_d determination. The TSA resultant K_d was equal to 25 nM.

2.2.2. Assay of inhibition of the enzymatic activity. There are several variations of the enzymatic activity inhibition assay of CAs (Khalifah, 1971; Colombo & Pinelli, 1981; Barauskienė & Matulis, 2019). CA enzyme catalyzes CO_2 hydration to bicarbonate and protons, but different reactions, such as hydrolysis of *para*-nitrophenylacetate ester, are often followed due to the assay's simplicity. However, we prefer an SFA of the inhibition of CA enzymatic activity of CO_2 hydration (Khalifah, 1971; Smirnovienė *et al.*, 2017). This assay is relatively robust and could be used to quantify the K_d of the studied compound binding to a particular CA isozyme. However, the assay has a significant disadvantage due to a varied limit of K_d determination for different CA isozymes. This limit depends on the enzyme concentration used in the assay. For example, to measure CA I enzyme activity, we need to use a minimum 0.1 μM of this isozyme. Therefore, the highest-affinity IC_{50} that could be measured at this concentration is 0.1 μM , the same as enzyme concentration. Thus, inhibitor affinities that are stronger than 0.1 μM would show as equal to 0.1 μM . All inhibitors that possess higher affinities would not be discovered. For CA IX, the situation is slightly different because this enzyme has greater specific activity than

CA I, and the assay can be performed at 10 nM enzyme concentration. Thus, the IC_{50} limit of determination is around 10 nM, tenfold lower than CA IX. Some equipment and assay optimization allowed enzyme concentration to be further diminished to ≈ 1 nM and thus reach greater affinities (Kugler *et al.*, 2020). This difference among CA isozymes may lead to misinterpretation of compound selectivities and affinities.

Fig. 3(b) shows a typical dosing curve and the underlying raw data (in the inset) to determine the IC_{50} of the compound inhibition effect. The value may be transformed to the K_i inhibition constant if the Michaelis–Menten constant K_M and the substrate concentration $[S]$ are known, by using the Cheng–Prusoff equation:

$$K_i = \frac{IC_{50}}{1 + ([S]/K_M)}. \quad (1)$$

The inhibition constant K_i is not entirely equivalent to the dissociation constant K_d but their values could be compared. This assay proves that a compound is an enzyme inhibitor, not just a random protein site binder. Other assays, such as TSA or ITC, cannot make this distinction.

2.2.3. Isothermal titration calorimetry. ITC is often considered a central assay to determine compound affinities for proteins. The main advantage of ITC is that it is the only assay that measures directly the heat effects and, therefore, it determines the enthalpy if heat is measured at a constant

pressure. No other assay can do this directly. The enthalpy could also be determined by performing other assays at different temperatures and then calculating the enthalpy by applying the van't Hoff equation. However, this indirect determination of the enthalpy is much less accurate than direct determination by ITC (Klebe, 2015).

If the affinity K_d of a compound is around 0.1 μM to 1 μM , the ITC yields highly accurate K_d values. However, ITC has significant limitation in the narrow dynamic range of K_d values. Since the best results are obtained when the Wiseman parameter is within the range from 5 to 500 (Wiseman *et al.*, 1989; Velazquez-Campoy, 2015), thus, at 10 μM enzyme concentration required for the assay, the obtainable K_d range is 2 μM –20 nM. Further range extension is possible only by performing the competition assay with a strong binder (Krainer & Keller, 2015), complicating the data analysis.

Fig. 3(c) shows a typical dosing curve obtained by ITC, with the raw data curve shown in the inset. The curve yielded a K_d value of 74 nM. This value is slightly higher but still comparable to the values obtained by other assays. The standard error of all K_d measurements is approximately equal to twofold in the K_d value (Paketurytė *et al.*, 2021). Thus, a value of 74 nM is within a range of 37 nM to 150 nM, while a value of 25 nM is within a range of 12 nM to 50 nM. Therefore, the values of 74 and 25 nM are within the error margin of both determinations.

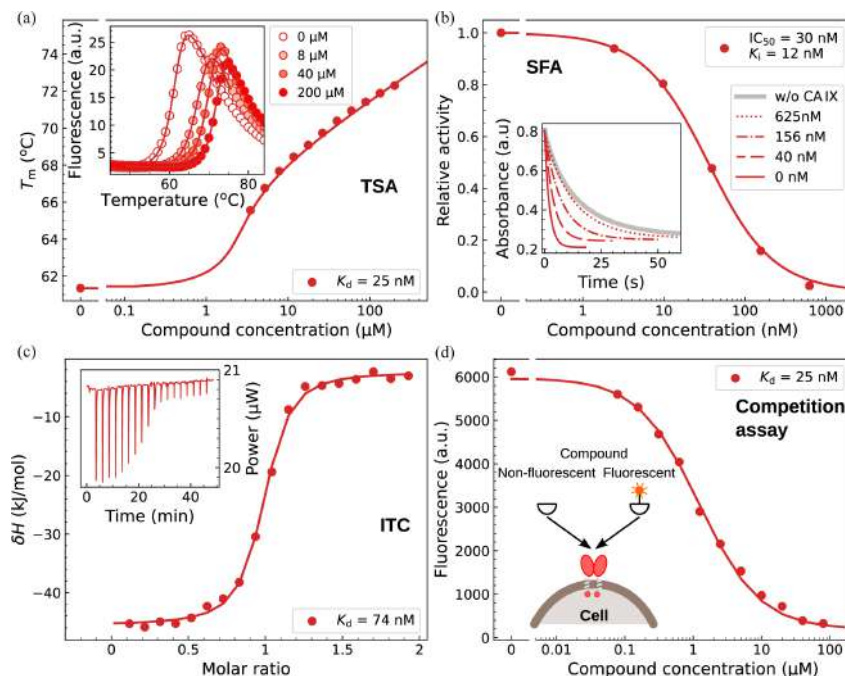


Figure 3

Affinity measurements of benzenesulfonamide-bearing compound VD10-13 binding to CA IX by using four different techniques: (a) FTSA, (b) SFA of CO_2 hydration enzymatic activity inhibition, (c) ITC and (d) competition assay for cell-expressed CA IX. The insets show raw data curves, while the main panels show processed data dosing curves used for dissociation constant (K_d) calculation.

The ITC assay determines the observed enthalpy. This value for the data in Fig. 3(c) is equal to -45 kJ mol^{-1} . However, this value contains linked protonation reactions that could contribute comparable enthalpies due to buffer protonation. Therefore, it is essential to dissect all contributing reaction enthalpies and calculate the intrinsic enthalpy (Linkuvienė *et al.*, 2018b).

2.2.4. Competition assay in live cancer cells. When we designed a series of high-affinity compounds for CA IX as potential anticancer agents, it was necessary to demonstrate that these compounds recognize and bind exclusively to CA IX, whose expression increases upon growing cancer-cell cultures under hypoxia. It was essential to demonstrate whether the compounds bind (1) with high affinity, and (2) specifically and selectively to CA IX. For this purpose, we designed a compound GZ19-32 that bears a fluorescein group and possesses high-affinity sulfonamide-based moiety for CA IX recognition (K_d , determined by FTSA of GZ19-32 binding to recombinant CA IX, equal to 10 pM). Simultaneous addition of the fluorescent compound and the test compound competed for the same CA IX binding sites on the cell surface in a dose-dependent manner. Affinities of the tested compounds matched those determined for purified CA IX enzymes (Matulienė *et al.*, 2022). Thus, the assay recognized CA IX and not other isozymes that may also be expressed on the cell surface, especially CA XII and CA XIV.

The competition assay was also helpful in following the expression of CA IX, its concentration, and amount dependence on the time of expression. No CA IX was detectable after 24-hour expression, but the numbers increased over the next three days. The number of CA IX residues exceeded one million per cell, mostly visible at the invadopodia, where the invasion of adjacent tissues is orchestrated. The concentration of CA IX reached $\approx 10 \text{ nM}$. Fig. 3(d) shows the VD10-13 compound dosing curve while keeping the concentration of the fluorescein-labeled compound constant (10 nM). Higher concentrations of the test compound out-competed the fluorescent compound. The midpoint of the dosing curve does not match the K_d but is approximately 100-fold higher because the midpoint here does not match the 50%-bound value.

The competitive binding assay confirmed that our designed compounds possessed extremely high picomolar affinity, as determined by the TSA (Matulienė *et al.*, 2022). Furthermore, the assay demonstrated that the compounds selectively recognized CA IX in the presence of other CA isozymes. Therefore, the compounds possessed high affinity and selectivity for CA IX.

2.3. Structure–thermodynamics correlations of protein–ligand binding

After crystal structural characterization of compound binding and thermodynamic measurements by the assays described above, we attempted to perform two kinds of correlations. First, we correlated compound chemical structural features, such as the presence or position of various functional groups, with the thermodynamic parameters of

binding. Second, we correlated the compound–CA isozyme complex crystal structures with the thermodynamics of the compound binding to the protein. Both correlations are essential to fully understand the underlying forces that hold the complex together.

Fig. 4 shows both correlations – here we compare three compound (VD11-4-2, VD10-35 and VD12-05) binding modes to five CA isozymes (CA I, CA II, CA IX, CA XII and CA XIII). These 15 binding pairs were characterized crystallographically (except VD10-35 and VD12-05 binding to CA IX) and thermodynamically. Intrinsic thermodynamic parameters of all pairs were determined by TSA and ITC (for calculations we used Gibbs energy obtained by TSA and enthalpy from ITC measurements).

First, we analyzed the Gibbs energy contributions that indicate compound affinities and selectivities. Comparing compounds VD12-05 and VD10-35, where the first contains the methyl group on the tail while the second contains the hydroxy group capable of making a hydrogen bond with bulk water or the protein, we see that all five CA isozymes lost $7.5\text{--}10.7 \text{ kJ mol}^{-1}$ of Gibbs energy when changing the methyl group to the hydroxy group. A loss of 6 kJ mol^{-1} is equivalent to a $10\times$ loss in K_d value. This result indicates that the hydroxy group preferentially made the hydrogen bond with water molecules not the protein molecule. Instead, the hydrophobic methyl group probably bound preferentially to the protein through non-specific hydrophobic interaction.

Another example is the addition of a cyclooctyl amino ring to the *meta* position relative to the sulfonamide group (going from the middle line to the top line in the figure). Here, the situation is quite different from the first example. The addition of the ring decreased compound affinity for CA I (by 18.0 kJ mol^{-1} , equivalent to a $1000\times$ loss in K_d affinity), but gained significant affinity or Gibbs energy for CA IX, $-21.9 \text{ kJ mol}^{-1}$, or $-14.9 \text{ kJ mol}^{-1}$ for CA XII. The gain of $-21.9 \text{ kJ mol}^{-1}$ is equivalent to $4000\times$ gain in affinity K_d . Therefore, while the compound VD10-35 was a stronger binder of CA I than of CA IX, the VD11-4-2 compound is a more than $100\,000\times$ stronger binder of CA IX than of CA I. Thus, the compound is highly selective for CA IX over CA I.

Second, we analyzed the enthalpies of interaction, shown as blue values in Fig. 4, and see no correlation between the Gibbs energies and the enthalpies of binding. The enthalpies show information on the contact between the compound and the protein. In contrast, the Gibbs energies show the sum of enthalpies and entropies (which provide information about the degrees of freedom in the free and bound states). Enthalpically, most exothermic reaction was for VD10-35 binding ($-76.8 \text{ kJ mol}^{-1}$) and VD12-05 binding ($-74.9 \text{ kJ mol}^{-1}$) to CA I. Both these reactions were driven by enthalpy with a minor opposition by entropy. In contrast, the VD11-4-2 compound binding to CA IX was largely driven by entropy ($+47.5 \text{ kJ mol}^{-1}$) and to a lesser extent by enthalpy ($-24.3 \text{ kJ mol}^{-1}$). Such balance of these forces led to the strongest binder (Gibbs energy of $-71.8 \text{ kJ mol}^{-1}$).

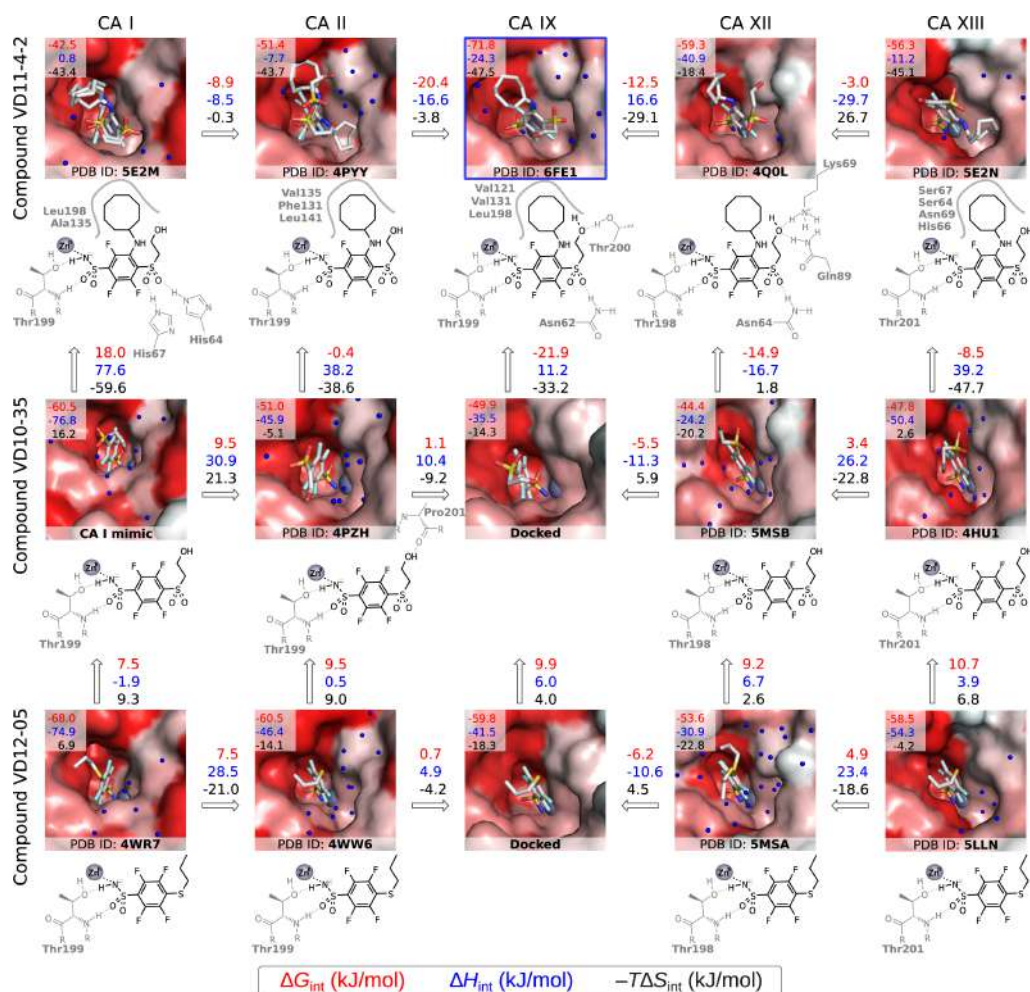


Figure 4

A map of correlation between three-dimensional X-ray crystallographic structure of enzyme-inhibitor complexes and binding thermodynamics. The three rows represent three compounds (VD11-4-2, VD10-35 and VD12-05) while the five columns represent CA isozymes (CA I, CA II, CA IX, CA XII and CA XIII). The three numbers on the upper-left corner of each crystal structure list the intrinsic standard Gibbs energy, enthalpy and negative entropy (multiplied by the absolute temperature) change upon binding to the CA isozyme (at 37°C). The PDB ID is listed at the bottom of each structure (except for two unavailable structures that are docked models). Chemical structure diagrams below each structure show interactions in the crystal structures. Arrows connect neighboring structures, and the numbers next to each arrow show the energy difference between binding processes when comparing different isozymes (horizontal arrows) or compounds (vertical arrows). The map shows binding energies and suggests structural reasons behind the energy changes.

Schemes below the crystal structures show crystallographically observed interactions, hydrogen bonds and potential hydrophobic interactions. However, the exact assignment of the energetic contributions to every interatomic contact is still impossible, primarily due to a limited understanding of the thermodynamics and structure of a compound while free in solution.

Fig. 5 shows another example of the structure-thermodynamics correlation map for a series of compounds that do

not contain fluorine atoms but instead have chlorine atoms. Unlike the previous figure, we do not show information on the enthalpy and entropy of interaction here. Instead, we show the binding Gibbs energies to all 12 catalytically active human CA isozymes. Having affinities for all 12 CA isozymes makes it easy to judge whether the compound will possess significant selectivity toward one or several isozymes. Numbers next to the structures in Fig. 5 show affinities for isozymes, while the numbers next to the arrows show the differences between the

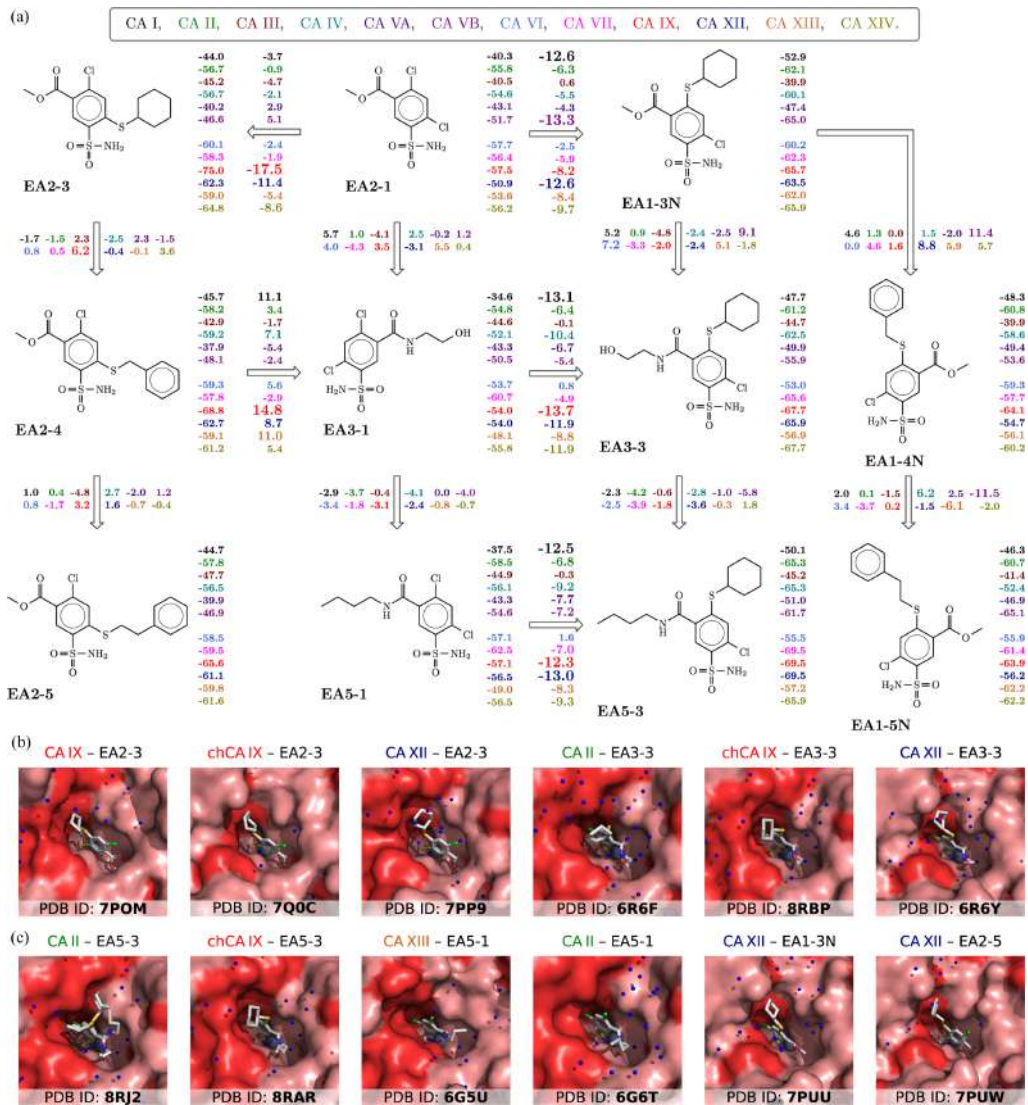


Figure 5
(a) A map of correlation between the chemical structures of compounds and their intrinsic affinities (changes in standard Gibbs energies upon binding at 37°C). The 12 numbers next to each compound structure show the intrinsic standard Gibbs energies of binding to each catalytically active CA isozyme. The 12 numbers next to each arrow connecting structurally related compounds show differences in the intrinsic standard Gibbs energy between the compounds. The most significant gains (negative numbers) or losses (positive numbers) in affinity are listed in larger font. Panels (b) and (c) show available X-ray crystallographic structures for several compounds binding to several CA isozymes. The map shows the path toward the most potent binder of CA IX, compound EA2-3.

affinities of adjacent compounds. Thus, we can judge a contribution in affinity due to the chemical change of the compound structure.

In this series of EA compounds, the most notable is the addition of the cyclohexane group to compound EA2-1, leading to compound EA2-3. The change in structure led to a

gain in Gibbs energy for CA IX of $-17.5 \text{ kJ mol}^{-1}$, equivalent to nearly a 1000× gain in affinity. This gain was observed primarily in CA IX and to a lesser extent in CA XII. The EA2-3 compound exhibited a binding Gibbs energy of $-75.0 \text{ kJ mol}^{-1}$, the highest affinity among the studied compounds, reaching the femtomolar range. However, this number shows

the intrinsic affinity and would not be observed experimentally by applying any technique. The observed affinity is always lower than the intrinsic one.

3. Methods

3.1. Chemical compound synthesis

EA compounds were synthesized and analytically characterized as described by Zakšauskas *et al.* (2020, 2022). The synthesis of the VD series of compounds was described by Dudutienė *et al.* (2013, 2014). The synthesis of the fluorescein-labeled compound GZ19-32 was described by Matulienė *et al.* (2022).

3.2. CA isozyme preparation

Recombinant human CAs were expressed in *Escherichia coli* [CA I, CA II, chCA IX (mutant CA II A65S, N67Q, I91L, F130V, V134L, L203A), CA III, CA IV, CA VA, CA VB, CA VII, CA XII, CA XIII, CA XIV], yeast [CA IX for crystallization experiments (Leitans *et al.*, 2015)] or mammalian cells (CA VI and CA IX used for binding affinity experiments). Protein expression and purification protocols are described in: CA I (Baranauskienė *et al.*, 2009); CA II (Cimpmperman *et al.*, 2008); CA IV (Mickevičiūtė *et al.*, 2018); CA VI (Kazokaitė *et al.*, 2015); chCA IX (CA II A65S, N67Q, I91L, F130V, V134L, L203A) CA III, CA VA, CA VB, CA IX, CA XIV (Dudutienė *et al.*, 2014); CA VII, CA XIII (Sūdžius *et al.*, 2010); and CA XII (Jogaitė *et al.*, 2013). The protein stock solutions were stored at -80°C . The protein concentration was determined by UV absorption at 280 nm.

3.3. Crystallization and structure determination

Crystals of chimeric CA IX protein (mutant CA II A65S, N67Q, I91L, F130V, V134L, L203A) reproducing CA IX active sites were prepared like crystals of CA II by the sitting-drop method. The crystallization solution contained 0.2 M ammonium sulfate, 0.1 M sodium-bicine (pH 9.0) and 2 M sodium-malonate (pH 7.0). The protein was concentrated to 29 mg ml $^{-1}$ and mixed with an equal volume of reservoir solution. Crystals were soaked with EA3-3 and EA5-3 inhibitors and added to the crystallization solution at a final concentration of 1 mM for several days. Diffraction data were collected at the P14 EMBL beamline at the PETRA III storage ring (DESY, Hamburg, Germany). Data reduction was performed by *XDS* (Kabsch, 2010). *COMBAT* v. 7.1.002, *SCALA* v. 3.3.22, *TRUNCATE* v. 7.1.002 and other *CCP4* v. 7.1.002 tools were used for data processing (Agirre *et al.*, 2023). The structure was solved by molecular replacement using *MOLREP* (Vagin & Teplyakov, 2010) with a protein chain of PDB ID 3hlj (Baranauskienė *et al.*, 2010) as an initial model. The model was refined by *REFMAC* v. 5.8.0258 (Murshudov *et al.*, 2011) and rebuilt using *Coot* v. 0.9 (Emsley *et al.*, 2010). Inhibitor models were created and minimized by *AVOGADRO* v. 1.2.0 (Hanwell *et al.*, 2012).

3.4. Competition model for the determination of VD10-13 compound K_d values for cell-expressed CA IX

HeLa cell culture was grown in 12-well plates under hypoxia (1% O $_2$) or normoxia (21% O $_2$) for 70–76 h. The medium was removed and then 200 μl of a two compounds mixture containing 10 nM of fluorescent compound GZ19-32 and different concentrations of ligand VD10-13 from 0 to 80 μM (by performing 12 serial twofold dilutions with the last sample concentration of 0 nM) in the FB medium was added and incubated at 37°C under normoxia in the CO $_2$ incubator for 20 min. Then, the compound solution was removed, and the cells were washed three times for 1.5–2 min with 400 μl of PBS (phosphate-buffered saline) at room temperature. Then 180 μl of TrypLE express enzyme (ThermoFisher) was added to each well and incubated in the CO $_2$ incubator for 5–10 min until the cells were detached from the surface. After the addition of 20 μl of the Defined Trypsin Inhibitor solution (ThermoFisher), cells were resuspended by pipetting, and 150 μl of the suspension from each well was transferred to black Thermo Scientific Nunc MicroWell 96-Well Optical-Bottom Plates for fluorescence and absorbance measurements. The fluorescence was recorded at 485 nm excitation and 520 nm emission wavelengths in the plate reader (Synergy HTX, BioTek) (Matulienė *et al.*, 2022).

A ligand binding to a protein in the presence of a competing compound was described by Wang (1995). The fraction of fluorescently labeled GZ19-32 (ligand B), α_B , bound to CA IX in the presence of VD10-13 (ligand A) can be calculated using the following set of equations:

$$\alpha_B = \frac{1}{P_t} \frac{L_{t,B}[2(a^2 - 3b)^{1/2} \cos(\theta/3) - a]}{3K_{d,B} + [2(a^2 - 3b)^{1/2} \cos(\theta/3) - a]}, \quad (2)$$

$$\theta = \arccos \frac{-2a^3 + 9ab - 27c}{2[(a^2 - 3b)^3]^{1/2}}, \quad (3)$$

$$a = K_{d,A} + K_{d,B} + L_{t,A} + L_{t,B} - P_t, \quad (4)$$

$$b = K_{d,A}(L_{t,B} - P_t) + K_{d,B}(L_{t,A} - P_t) + K_{d,B}K_{d,B} \quad (5)$$

and

$$c = -K_{d,A}K_{d,B}P_t, \quad (6)$$

where P_t is the concentration of total added protein, and $L_{t,A}$ and $L_{t,B}$ are the molar concentrations of added ligands A and B, respectively. The $K_{d,A}$ and $K_{d,B}$ terms are the equilibrium dissociation constants of ligands A and B, respectively.

3.5. Stopped-flow assay of CO $_2$ hydration

Recombinant human CA IX activity was determined by measuring the hydration rates of CO $_2$ using an Applied Photophysics SX.18MV-R stopped-flow spectrometer. Reaction velocities were measured by recording the absorbance of pH indicator phenol red (final dye concentration of 30 μM) at 557 nm. The pH changes occur due to protons formed in the enzymatic or spontaneous CO $_2$ hydration reaction. A saturated CO $_2$ solution was prepared by bubbling CO $_2$ in Milli-Q

water at 25°C for 1 h. The samples consisted of 10 nM CA IX containing 0, 2.4, 9.8, 39, 156 or 625 nM inhibitor VD10-13 (final concentration of DMSO $\leq 0.2\%$) in 25 mM HEPES buffer containing 50 mM NaCl, pH 7.4. The K_d value was determined using the Morrison equation:

$$\text{Active CA IX fraction} = \frac{K_d + P_t + L_t - [(K_d + P_t + L_t)^2 - 4P_t L_t]^{1/2}}{2P_t}, \quad (7)$$

where P_t denotes the total concentration of CA IX, L_t is the total inhibitor VD10-13 concentration and K_d is the protein–ligand dissociation constant.

3.6. Estimation of binding affinity by fluorescent thermal-shift assay

5 μM CA IX solutions containing different concentrations of VD10-13 ranging from 0–200 μM (1.5-fold dilutions) and 100-fold diluted solvatochromic dye Glomelt (Biotium) were prepared in 50 mM sodium phosphate buffer (pH 7.0) containing 100 mM NaCl with a final 2% (v/v) concentration of DMSO. FTSA experiments were performed with real-time PCR instrument QIAGEN Rotor-Gene Q. The samples were heated from 25 to 95°C, increasing the temperature by 1°C min^{-1} . Fluorescence measurements were taken at each temperature in the green channel (Ex/Em 470/510 nm). Protein-unfolding data analysis was performed, and compound binding constants were determined using the *Thermott* web server as described by Gedgaudas *et al.* (2022).

3.7. Estimation of binding affinity by isothermal titration calorimetry

100 μM solution of the inhibitor VD10-13 or VD12-05 was prepared in 50 mM Tris buffer (pH 7.0) containing 100 mM NaCl (final DMSO concentration in solution 1%). A solution of 10 μM CA IX was prepared in 50 mM Tris buffer (pH 7.0) containing 100 mM NaCl and 1% DMSO solvent. The cell of the calorimeter PEAQ-ITC (Malvern Panalytical) was filled with the protein solution, and the syringe was filled with the inhibitor solution. The experiment was run at 37°C. Titrations consisted of a 1.2 μl first injection followed by 18 subsequent 2 μl additions every 150 s. Thermodynamic binding parameters were determined by fitting the data to a single binding-site mode using the *Microcal Origin* program (OriginLab, USA).

4. Discussion

Crystallography and other structural techniques, such as nuclear magnetic resonance or cryo-electron microscopy, have contributed significantly to structure-based drug discovery (Nasim & Qureshi, 2022). Insight into protein structure, especially of protein–lead compound complexes, could provide crucial information for successfully improving the designed compound, both in its increased affinity toward the

disease target protein and its decreased affinity for all remaining proteins to avoid possible toxic side effects.

The members of the CA family have been drug targets for a while. However, to this day, two human mitochondrial isozymes, CA VA and CA VB, still do not have crystal structures solved and deposited to the PDB. All remaining members of the catalytically active CA isozymes have their structures solved, but the numbers of solved structures are uneven for various isozymes. Over 1000 structures of CA II are available, but only one for CA VI. Our research group has significantly contributed to providing X-ray crystallographic structures of various CA isozymes bound with inhibitors (Fig. 6). We have deposited 147 structures to the PDB, especially a large share of CA IX, CA XII and CA XIII structures.

Crystallography has helped us significantly in the design of high-affinity compounds for CA IX, such as VD11-4-2 (Dudutienė *et al.*, 2014), GZ18-23 (Matulienė *et al.*, 2022) and EA2-3 (Zakšauskas *et al.*, 2022). These compounds possess affinities with $K_{d,\text{obs}}$ in the range 20–300 pM. Furthermore, their selectivity toward CA IX over remaining isozymes reaches a hundred- to a million-fold depending on the isozyme. Thus, the compounds could be highly suitable for development as anticancer drugs that target CA IX. However, crystallography does not provide information on the compound affinity for proteins (Schlichting, 2005; Danley, 2006). Millimolar concentrations of inhibitors are used in crystallography. Thus, compounds possessing millimolar K_d would be visible in the structure and could be assumed as strong binders. In reality, such compounds could equally possess millimolar or picomolar K_d that would not be distinguishable by crystallography. Crystal structures often give an insight into the flexibility of inhibitors bound in the active sites (high B factors, partial occupancy, electron density, alternate

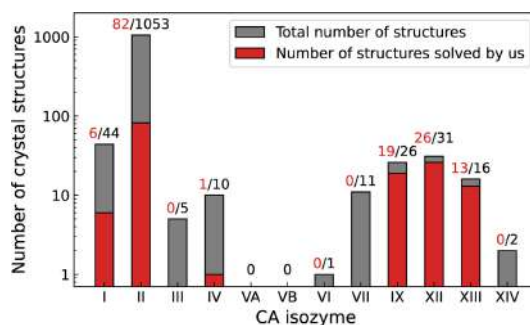


Figure 6

Numbers of X-ray crystallographic structures of catalytically active CA isozymes in the PDB. Red numbers show how many structures were solved and deposited by our research group. Most structures in the PDB are of CA II, a total of 1053, of which 82 are from our group. We have deposited a large part of the available structures of CA IX, CA XII and CA XIII. The availability of CA isozyme structures varies considerably; there are no structures available of both mitochondrial isozymes CA VA and CA VB, and only one structure is available of CA VI, two of CA XIV, and five of CA III. Values were obtained from the PDB on December 14, 2023.

conformations). Significant variability could be found in the position of ligands between several subunits.

Thermodynamic and kinetic assays must be performed to determine the binding between a purified protein and a pure compound. Various techniques are used for this purpose and have been extensively reviewed by Krishnamurthy *et al.* (2008). We have been using several orthogonal techniques, which are very important and significantly contribute to the overall picture of protein–ligand interaction.

The affinities are best determined by the TSA (Dudutienė *et al.*, 2020; Smirnovienė *et al.*, 2017). This assay is highly robust and easy to perform, and provides the most accurate affinities between stable globular proteins and synthetic small-molecule ligands (Linkuvienė *et al.*, 2018*b*). This assay has the widest dynamic range for the K_d values, spanning from millimolar to picomolar. However, this assay has been underused, probably due to complex thermodynamic equations, which need to be solved for K_d determination. To ease this task, we have established a web-based service where a user can plug in their TSA data and obtain a suggestion of necessary thermodynamic parameters for the K_d determination. The freely available web application is called *Thermott* (Gedgudas *et al.*, 2022).

Interaction kinetics is as important as thermodynamics. The on- and off-rates provide significant insight into the structural stability of the complex. We have also demonstrated that it is important to distinguish intrinsic from observed kinetics, similar to its importance in thermodynamics (Linkuvienė *et al.*, 2018*a,b*). ITC provided us with the enthalpies and the ITC measurements at several temperatures – the heat capacities of ligand binding, which have led us to a more complete understanding of the intrinsic binding principles (Linkuvienė *et al.*, 2018*b*).

The enzyme-inhibition assay is the most abundant in the literature but is sometimes misused and misinterpreted. As explained in the literature (Copeland, 2013), such assays are unsuitable for tight-binding inhibitors and are limited to cases where the affinity of the inhibitor, K_d , does not go below the enzyme concentration used in the assay. For CA inhibitors, where the affinity often goes to the picomolar range, it is not possible to distinguish very potent inhibitors and accurately determine the K_d value by the enzyme-inhibition assay.

It is insufficient to measure the affinity by one technique, and orthogonal assays should be used. Unfortunately, the use of a sole technique rather often leads to erroneous conclusions. We have stumbled upon this several times and suggest that others perform several independent assays. Therefore, since many of the measurements of CA IX potent inhibitors have been obtained solely by TSA, we searched for an independent method to confirm the affinity values and discovered that the competition model for binding to cell-expressed CA IX is an excellent technique to confirm binding of the compound to the protein (Matulienė *et al.*, 2022). This technique could be extended to other CA isozymes, especially those expressed on the cell membrane.

The abundance of data that contained numerous measurements between hundreds of compounds and 12 CA isozymes

led to the establishment of a web-based database. The data were also scattered among numerous publications, and it took much work to compare them. Therefore, we compiled all data on the freely available web-based database (Lingė *et al.*, 2023). The database enables analysis of all accessible data in one place. There are quite a few databases available, but, to the best of our knowledge, our database is the only one that distinguishes intrinsic thermodynamics by dissecting linked protonation that accompanies most protein–ligand interactions. Dissection of these reactions, as explained in the *Introduction*, led us to significantly different explanations of compound affinities and selectivities based on observed versus intrinsic data.

The concept of intrinsic parameters emphasizes the need to dissect all possible binding-linked reactions. These reactions could be protonation, but also could be linked to protein conformation, salt effects and many other reactions. To dissect these reactions, one needs to understand the binding mechanism. Subtraction of the binding-linked reactions would yield more accurate thermodynamic and kinetic binding parameters. These intrinsic parameters must be used in drug design to help rank compounds in their affinity, binding enthalpy, entropy or off-rate. However, the compound physiological effect should be evaluated according to the experimentally observed value. For example, the K_d value at pH 7.0 is not the intrinsic value, but the experimentally observed value will determine actual affinity at particular conditions because compound binding will be accompanied by all binding-linked reactions that will weaken the affinity. Therefore, the use of intrinsic parameters is not always necessary and depends on the task. Furthermore, in CA–sulfonamide binding, the mechanism involves two linked reactions, but in other protein–ligand binding cases, the mechanism might be quite different. Sometimes there may be no obvious linked reactions, but in our experience, there are more such reactions involved than appreciated by the scientific community.

While studying the family of CAs and their interaction with sulfonamide compounds, we have learned that the family is a convenient object to study general principles of compound–protein recognition, which is a still poorly understood phenomenon, and, therefore, underused in rational drug design. First, the CA isozymes have nearly identical folds, and the protein structures differ little in the shape of their active sites. Thus, studying the contribution of particular amino acids to the binding reaction is convenient. Second, the coordination bond between the sulfonamide nitrogen atom and the zinc of the protein strengthens the binding by more than 1000-fold. Thus, these compounds, even the simplest benzenesulfonamide, are strong binders of CA isozymes, which makes it easier to rationally synthesize a series of compounds where every compound will possess measurable binding affinity, enabling more efficient use of synthetic capabilities and more data for a deeper understanding of the recognition principles.

Despite abundant data, it is still rather discussable and controversial to state whether sulfonamide inhibitors bind and dynamically change the shape of the CA enzyme, as has been

concluded by computational studies (Ma *et al.*, 2017), or whether the enzyme does not undergo substantial conformational changes upon ligand binding and the binding reaction is more like a 'lock-and-key' mechanism, as we have determined previously using crystallography (Dudutienė *et al.*, 2020). We think the truth lies between these extreme cases, and both mechanisms are possible.

The overall goal of rational drug design is to predict the chemical structure of a compound that would bind with great affinity to the disease-target protein and possess great selectivity by not binding to any of the 20 000+ proteins present in any cell to avoid possible toxic side effects. To achieve this goal, we must design a high-affinity compound *in silico* and demonstrate that the compound does not substantially bind to any other protein. This is not experimentally feasible but will probably be achieved by computer modeling in the future. When the design of such compounds is possible, the goal of rational drug design will be accomplished.

Acknowledgements

The synchrotron data were collected at beamline P14, operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). We would like to thank G. Bourenkov for the assistance in using the beamline.

References

Agirre, J., Atanasova, M., Bagdonas, H., Ballard, C. B., Baslé, A., Beilstein-Edmands, J., Borges, R. J., Brown, D. G., Burgos-Mármol, J. J., Berrisford, J. M., Bond, P. S., Caballero, I., Catapano, L., Chojnowski, G., Cook, A. G., Cowtan, K. D., Croll, T. I., Debrezen, J. É., Devenish, N. E., Dodson, E. J., Drevon, T. R., Emsley, P., Evans, G., Evans, P. R., Fando, M., Foadi, J., Fuentes-Montero, L., Garman, E. F., Gerstel, M., Gildea, R. J., Hatti, K., Hekkelman, M. L., Heuser, P., Hoh, S. W., Hough, M. A., Jenkins, H. T., Jiménez, E., Joosten, R. P., Keegan, R. M., Keep, N., Krissinel, E. B., Kolenko, P., Kovalevskiy, O., Lamzin, V. S., Lawson, D. M., Lebedev, A. A., Leslie, A. G. W., Lohkamp, B., Long, F., Malý, M., McCoy, A. J., McNicholas, S. J., Medina, A., Millán, C., Murray, J. W., Murshudov, G. N., Nicholls, R. A., Noble, M. E. M., Oeffner, R., Pannu, N. S., Parkhurst, J. M., Pearce, N., Pereira, J., Perrakis, A., Powell, H. R., Read, R. J., Rigden, D. J., Rochira, W., Sammito, M., Sánchez Rodríguez, F., Sheldrick, G. M., Shelley, K. L., Simkovic, F., Simpkin, A. J., Skubak, P., Sobolev, E., Steiner, R. A., Stevenson, K., Tews, I., Thomas, J. M. H., Thorn, A., Valls, J. T., Uski, V., Usón, I., Vagin, A., Velankar, S., Vollmar, M., Walden, H., Waterman, D., Wilson, K. S., Winn, M. D., Winter, G., Wojdyr, M. & Yamashita, K. (2023). *Acta Cryst.* **D79**, 449–461.

Alka, K. & Casey, J. R. (2014). *IUBMB Life*, **66**, 596–615.

Bahr, G., Tomatis, P. E. & Vila, A. J. (2023). *Compr. Inorg. Chem.* **III**, pp. 231–267. Elsevier.

Baranauskienė, L., Hilvo, M., Matulienė, J., Golovenko, D., Manakova, E., Dudutienė, V., Michailovienė, V., Torresan, J., Jachno, J., Parkkila, S., Maresca, A., Supuran, C. T., Gražulis, S. & Matulis, D. (2010). *J. Enzyme Inhib. Med. Chem.* **25**, 863–870.

Baranauskienė, L. & Matulis, D. (2019). *Carbonic Anhydrase as Drug Target: Thermodynamics and Structure of Inhibitor Binding*, edited by D. Matulis, pp. 39–49. Cham: Springer International Publishing.

Baranauskienė, L., Petrikaitė, V., Matulienė, J. & Matulis, D. (2009). *Int. J. Mol. Sci.* **10**, 2752–2762.

Becker, H. M. (2020). *Br. J. Cancer*, **122**, 157–167.

Bradshaw, J. M. & Waksman, G. (1998). *Biochemistry*, **37**, 15400–15407.

Casey, J. R. (2006). *Biochem. Cell Biol.* **84**, 930–939.

Ciccone, L., Cerri, C., Nencetti, S. & Orlandini, E. (2021). *Molecules*, **26**, 6380.

Cimpmperman, P., Baranauskienė, L., Jachimovičiūtė, S., Jachno, J., Torresan, J., Michailovienė, V., Matulienė, J., Sereikaitė, J., Bumelis, V. & Matulis, D. (2008). *Biophys. J.* **95**, 3222–3231.

Colombo, R. & Pinelli, A. (1981). *J. Pharmacol. Methods*, **6**, 177–191.

Copeland, R. A. (2013). *Evaluation of Enzyme Inhibitors in Drug Discovery*, ch. 7, pp. 245–285. John Wiley & Sons, Ltd.

Danley, D. E. (2006). *Acta Cryst.* **D62**, 569–575.

Dudutienė, V., Matulienė, J., Smirnov, A., Timm, D. D., Zubrienė, A., Baranauskienė, L., Morkūnaitė, V., Smirnovienė, J., Michailovienė, V., Juozapaitienė, V., Mickeviciūtė, A., Kazokaite, J., Bakšytė, S., Kasiliauskaitė, A., Jachno, J., Revuckienė, J., Kišonaitė, M., Piliuitytė, 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. (2014). *J. Med. Chem.* **57**, 9435–9446.

Dudutienė, V., Zubrienė, A., Kairys, V., Smirnov, A., Smirnovienė, J., Leitans, J., Kazaks, A., Tars, K., Manakova, L., Gražulis, S. & Matulis, D. (2020). *Biophys. J.* **119**, 1513–1524.

Dudutienė, V., Zubrienė, A., Smirnov, A., Glylytė, J., Timm, D., Manakova, E., Gražulis, S. & Matulis, D. (2013). *Bioorg. Med. Chem.* **21**, 2093–2106.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.

Fisher, Z., Boone, C. D., Biswas, S. M., Venkatakrishnan, B., Aggarwal, M., Tu, C., Agbandje-McKenna, M., Silverman, D. & McKenna, R. (2012). *Protein Eng. Des. Sel.* **25**, 347–355.

Frost, S. C. (2014). *Subcell. Biochem.* **75**, 9–30.

Fukada, H., Sturtevant, J. M. & Quiocho, F. A. (1983). *J. Biol. Chem.* **258**, 13193–13198.

Gedgaudas, M., Baronas, D., Kazlauskas, E., Petrauskas, V. & Matulis, D. (2022). *Drug Discovery Today*, **27**, 2076–2079.

Gomez, J. & Freire, E. (1997). *Structure-Based Drug Design: Thermodynamics, Modeling and Strategy*, pp. 111–141. New York: Springer.

Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E. & Hutchison, G. R. (2012). *J. Cheminform.* **4**, 17.

Hoggon, C., Marazzi, M. & García-Iriepa, C. (2022). *Int. J. Mol. Sci.* **23**, 5855.

Huang, J. X., Cooper, M. A., Baker, M. A., Azad, M. A. K., Nation, R. L., Li, J. & Velkov, T. (2012). *J. Mol. Recognit.* **25**, 642–656.

Intaiyaz Hassan, M., Shajee, B., Waheed, A., Ahmad, F. & Sly, W. S. (2013). *Bioorg. Med. Chem.* **21**, 1570–1582.

Jensen, E. L., Clement, R., Kosta, A., Maberly, S. C. & Gontero, B. (2019). *ISME J.* **13**, 2094–2106.

Jogaitė, V., Zubrienė, A., Michailovienė, V., Glylytė, J., Morkūnaitė, V. & Matulis, D. (2013). *Bioorg. Med. Chem.* **21**, 1431–1436.

Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.

Kazokaite, J., Milinavičiūtė, G., Smirnovienė, J., Matulienė, J. & Matulis, D. (2015). *FEBS J.* **282**, 972–983.

Khalifah, R. (1971). *J. Biol. Chem.* **246**, 2561–2573.

Klebe, G. (2015). *Nat. Rev. Drug Discov.* **14**, 95–110.

Krainer, G. & Keller, S. (2015). *Methods*, **76**, 116–123.

Krishnamurthy, V. M., Bohall, B. R., Kim, C.-Y., Moustakas, D. T., Christianson, D. W. & Whitesides, G. M. (2007). *Chem. Asian J.* **2**, 94–105.

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

Kugler, M., Holub, J., Brynda, J., Pospíšilová, K., Anwar, S. E., Bovol, D., Havránek, M., Král, V., Fábry, M., Grünert, B. & Rezáčová, P. (2020). *J. Enzyme Inhib. Med. Chem.* **35**, 1800–1810.

- Kumar, S., Rulhania, S., Jaswal, S. & Monga, V. (2021). *Eur. J. Med. Chem.* **209**, 112923.
- Leitans, J., Kazaks, A., Balode, A., Ivanova, J., Zalubovskis, R., Supuran, C. T. & Tars, K. (2015). *J. Med. Chem.* **58**, 9004–9009.
- Liljas, A., Kannan, K. K., Bergstén, P. C., Waara, I., Fridborg, K., Strandberg, B., Carlbom, U., Järup, L., Lövgren, S. & Petef, M. (1972). *Nat. New. Biol.* **235**, 131–137.
- Lingé, D., Gedgaudas, M., Merkys, A., Petrauskas, V., Vaitkus, A., Grybauskas, A., Paketurytė, V., Zubrienė, A., Zakšauskas, A., Mickevičiūtė, A., Smirnovienė, J., Baranauskienė, L., Čapkauskaitė, E., Dudutienė, V., Urniežius, E., Kononovas, A., Kazlauskas, E., Shubin, K., Schiöth, H. B., Chen, W.-Y., Ladbury, J. E., Gražulis, S. & Matulis, D. (2023). *Database*, **2023**, baad040.
- Linkuvienė, V., Talibov, V. O., Danielson, U. H. & Matulis, D. (2018a). *J. Med. Chem.* **61**, 2292–2302.
- Linkuvienė, V., Zubrienė, A., Manakova, E., Petrauskas, V., Baranauskienė, L., Zakšauskas, A., Smirnov, A., Gražulis, S., Ladbury, J. E. & Matulis, D. (2018b). *Q. Rev. Biophys.* **51**, 1–48.
- Liu, Z., Hu, M., Yang, Y., Du, C., Zhou, H., Liu, C., Chen, Y., Fan, L., Ma, H., Gong, Y. & Xie, Y. (2022). *Mol. Biomed.* **3**, 46.
- Lomelino, C. & McKenna, R. (2016). *Expert Opin. Ther. Pat.* **26**, 947–956.
- Lomelino, C., Šupuran, C. & McKenna, R. (2016). *Int. J. Mol. Sci.* **17**, 1150.
- Lomelino, C. L., Andring, J. T. & McKenna, R. (2018). *Int. J. Med. Chem.* **2018**, 9419521.
- Ma, H., Li, A. & Gao, K. (2017). *ACS Omega*, **2**, 8414–8420.
- Matulienė, J., Žvinys, G., Petrauskas, V., Kvietkauskaitė, A., Zakšauskas, A., Shubin, K., Zubrienė, A., Baranauskienė, L., Kačenauskaitė, L., Kopanchuk, S., Veiksina, S., Paketurytė-Latvė, V., Smirnovienė, J., Juozapaitienė, V., Mickevičiūtė, A., Michailovienė, V., Jachno, J., Stravinskienė, D., Šližienė, A., Petrošiūtė, A., Becker, H. M., Kazokaitė-Adomaitienė, J., Yaromina, A., Čapkauskaitė, E., Rinken, A., Dudutienė, V., Dubois, L. J. & Matulis, D. (2022). *Sci. Rep.* **12**, 17644.
- Matulis, D., Kranz, J. K., Salemme, F. R. & Todd, M. J. (2005). *Biochemistry*, **44**, 5258–5266.
- Maveyraud, L. & Mourey, L. (2020). *Molecules*, **25**, 1030.
- Mboge, M. Y., Mahon, B. P., McKenna, R. & Frost, S. C. (2018). *Metabolites*, **8**, 19.
- Meldrum, N. U. & Roughton, F. J. W. (1933). *J. Physiol.* **80**, 113–142.
- Mickevičiūtė, A., Timm, D. D., 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., Di Cera, E., Sly, W. S. & Matulis, D. (2018). *Eur. Biophys. J.* **47**, 271–290.
- Minetti, C. A. & Remeta, D. P. (2022). *Life*, **12**, 1438.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst. D* **67**, 355–367.
- Nasim, F. & Qureshi, I. A. (2022). *Advances in Protein Molecular and Structural Biology Methods*, pp. 357–371. Elsevier.
- Occhipinti, R. & Boron, W. F. (2019). *Int. J. Mol. Sci.* **20**, 3841.
- Paketurytė, V., Petrauskas, V., Zubrienė, A., Abian, O., Bastos, M., Chen, W.-Y., Moreno, M. J., Krainer, G., Linkuvienė, V., Sedivy, A., Velazquez-Campoy, A., Williams, M. A. & Matulis, D. (2021). *Eur. Biophys. J.* **50**, 661–670.
- Pastorek, J., Pastoreková, S., Callebaut, I., Mornon, J. P., Zelník, V., Opavský, R., Zát'ovicová, M., Liao, S., Portetelle, D., Stanbridge, E. J., et al. (1994). *Oncogene*, **9**, 2877–2888.
- Petrauskas, V., Kazlauskas, E., Gedgaudas, M., Baranauskienė, L., Zubrienė, A. & Matulis, D. (2024). *TrAC Trends Anal. Chem.* **170**, 117417.
- Rankin, G. O. (2007). *xPharm: The Comprehensive Pharmacology Reference*, edited by S. J. Enna & D. B. Bylund, pp. 1–5. New York: Elsevier.
- Renaud, J.-P., Chung, C., Danielson, U. H., Egner, U., Hennig, M., Hubbard, R. E. & Nar, H. (2016). *Nat. Rev. Drug Discov.* **15**, 679–698.
- Saarnio, J., Parkkila, S., Parkkila, A.-K., Haukipuro, K., Pastoreková, S., Pastorek, J., Kairaluoma, M. I. & Karttunen, T. J. (1998). *Am. J. Pathol.* **153**, 279–285.
- Schlichting, I. (2005). *Protein–Ligand Interactions*, edited by G. Ulrich Nienhaus, Vol. 305, pp. 155–165. Totowa: Humana Press.
- Scott, K. A. & Njardarson, J. T. (2018). *Top. Curr. Chem. (Z.)*, **376**, 5.
- Singh, S., Lomelino, C., Mboge, M., Frost, S. & McKenna, R. (2018). *Molecules*, **23**, 1045.
- Smirnovienė, J., Smirnovas, V. & Matulis, D. (2017). *Anal. Biochem.* **522**, 61–72.
- Sūdžius, J., Baranauskienė, L., Golovenko, D., Matulienė, J., Michailovienė, V., Torresan, J., Jachno, J., Sukackaitė, R., Manakova, E., Gražulis, S., Tumkevičius, S. & Matulis, D. (2010). *Bioorg. Med. Chem.* **18**, 7413–7421.
- Swenson, E. R. (2014). *Expert Opin. Drug Saf.* **13**, 459–472.
- Timothy, A. & Sanjeev, M. (2007). *Eur. Pharm. Rev.* pp. 5.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst. D* **66**, 22–25.
- Velazquez-Campoy, A. (2015). *J. Therm. Anal. Calorim.* **122**, 1477–1483.
- Wang, Z.-X. (1995). *FEBS Lett.* **360**, 111–114.
- Wiseman, T., Williston, S., Brandts, J. F. & Lin, L. N. (1989). *Anal. Biochem.* **179**, 131–137.
- Zakšauskas, A., Čapkauskaitė, E., Jezepčikas, L., Linkuvienė, V., Paketurytė, V., Smirnov, A., Leitans, J., Kazaks, A., Dvinskis, E., Manakova, E., Gražulis, S., Tars, K. & Matulis, D. (2020). *Eur. J. Med. Chem.* **185**, 111825.
- Zakšauskas, A., Čapkauskaitė, E., Paketurytė-Latvė, V., Smirnov, A., Leitans, J., Dvinskis, E., Stančaitis, L., Mickevičiūtė, A., Jachno, J., Jezepčikas, L., Linkuvienė, V., Sakalauskas, A., Manakova, E., Gražulis, S., Matulienė, J., Tars, K. & Matulis, D. (2022). *Int. J. Mol. Sci.* **27**, 130.

Publication 4

Standard Operating Procedure for Fluorescent Thermal Shift Assay (FTSA) for Determination of Protein-Ligand Binding and Protein Stability

Egidijus Kazlauskas, Vytautas Petrauskas, **Vaida Paketurytė**,
Daumantas Matulis

Eur. Biophys. J., 2021, 50 (5), 707–714.

DOI: 10.1007/s00249-021-01537-1



Standard operating procedure for fluorescent thermal shift assay (FTSA) for determination of protein–ligand binding and protein stability

Egidijus Kazlauskas¹  · Vytautas Petrauskas¹  · Vaida Paketurytė¹  · Daumantas Matulis¹ 

Received: 23 December 2020 / Revised: 31 March 2021 / Accepted: 9 April 2021 / Published online: 29 April 2021
© European Biophysical Societies' Association 2021

Abstract

A standard operating procedure for a fluorescence-based thermal shift assay (FTSA) is provided describing its typical applications, advantages and limitations. FTSA is a simple, robust, universal and quick assay to determine protein–ligand binding affinities and protein stabilities in the presence of various excipients and solution conditions. Therefore, the assay is very useful for the straightforward characterization of new recombinantly produced proteins. The assay has a wide dynamic range enabling simultaneous determination of affinities in the millimolar to picomolar range. The assay could be used for essentially any protein that is sufficiently soluble and stable in the studied aqueous solution. Here we provide examples and typical experimental protocols for both affinity and stability determinations.

Keywords Thermal shift assay · Protein–ligand binding · Protein stability

Introduction

Fluorescent thermal shift assay (FTSA; also referred to as TSA, DSF – differential scanning fluorimetry, and Thermo-Fluor®) has been developed as a high throughput method to study protein–ligand (small molecule) interactions and to characterize newly produced recombinant proteins (Semisotnov et al. 1991). Similarly to differential scanning calorimetry (DSC) (Brandts and Lin 1990), FTSA relies on subjecting a protein to gradually increasing temperature and evaluating thermal stability. As the name suggests, FTSA employs fluorescence readout to track protein unfolding. This can be the innate tryptophan fluorescence or, more often, a signal from solvatochromic dyes whose fluorescence intensity changes upon their binding to the gradually unfolding protein. FTSA is primarily used for assaying protein thermal stability and measuring ligand affinity (K_b or K_d). Such features of FTSA

as low sample consumption, relatively common real-time PCR equipment, and ability to measure up to several hundred samples simultaneously make this technique widely applicable both in academia and industry. Recent advances and highlights of this technique have been reviewed (Scott et al. 2016; Gao et al. 2020).

The fluorescence dependence on temperature has a distinct curvature (Fig. 1b) indicating protein transition from native to unfolded state. The midpoint value of the curve corresponds to the melting temperature, T_m , a good indicator of protein stability at the particular solution conditions. This allows for a relatively easy screening of various salts, buffers, and other additives to find suitable conditions for a chosen protein application (Vedadi et al. 2006). Therefore, FTSA is widely used for optimizing composition to improve the purification of challenging proteins, their storage and downstream applications (from biophysical or biochemical assays to structural biology experiments).

It has been observed that usually target-specific ligands stabilize the target protein upon binding and thus cause a favorable up-shift in T_m . Such ligand-induced thermal shift (ΔT_m) can be correlated with its binding affinity towards the target protein (K_b) (Matulis et al. 2005; Cimperman et al. 2008; Cimperman and Matulis 2011). If compared to isothermal titration calorimetry (ITC), FTSA is less time and sample consuming while yielding as reliable

Special Issue: COST Action CA15126, MOBIEU: between atom and cell.

✉ Daumantas Matulis
daumantas.matulis@bti.vu.lt; matulis@ibt.lt

¹ Department of Biothermodynamics and Drug Design, Institute of Biotechnology, Life Sciences Center, Vilnius University, Sauletekio 7, 10257 Vilnius, Lithuania

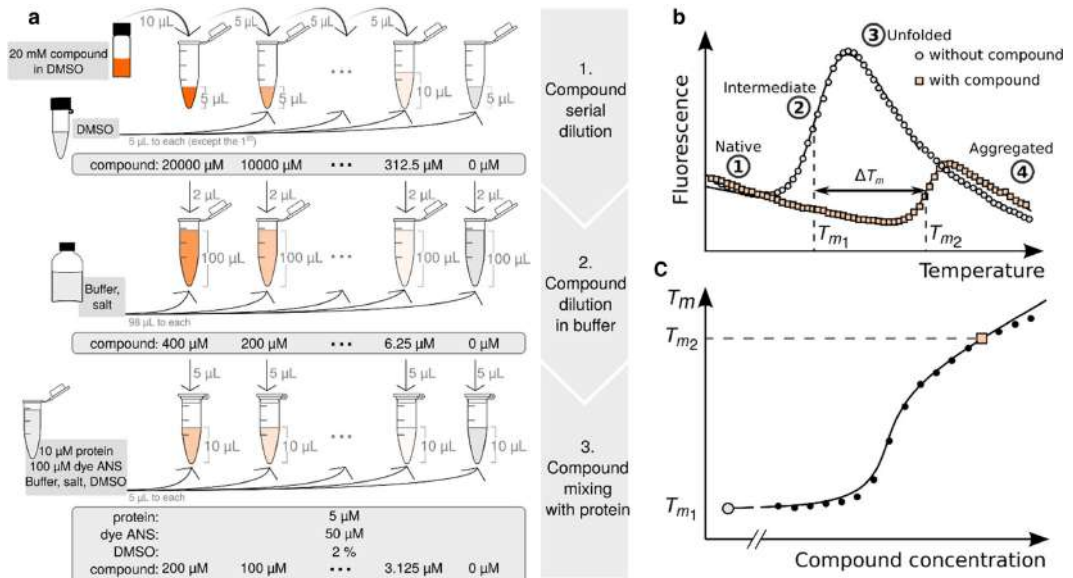


Fig. 1 Typical FTSA experiment for ligand binding measurements. **a** Schematic representation of sample preparation for the example experiment. First, the ligand acetazolamide (AZM) is diluted in DMSO. The dilution is performed in series, and no ligand is transferred into the last tube. Next, each DMSO dilution is further diluted in the buffer. Finally, samples from these dilutions are mixed with the prepared solution containing the protein carbonic anhydrase II (CA II) and the solvatochromic dye 1,8-anilinonaphthalene sulfonate (ANS) in the tubes, in which the experiment will be performed. At each step, each sample is thoroughly mixed. **b** Raw FTSA data –

ANS fluorescence intensity as a function of temperature for protein without ligand (grey circles) and for a certain tight-binding ligand concentration (peach squares). These curves represent the denaturation profile of the protein from native to unfolded state, as indicated by four stages next to the grey curve. The inflection point of this transition curve corresponds to the melting temperature (T_m) of the protein. **c** Determined T_m values from curves in part **b** are then plotted as a function of added ligand concentration (filled circles). The line is drawn according to the model, enabling the affinity (K_d) calculation

and precise K_b over a wide range of K_b , 10^3 – 10^{13} M $^{-1}$ (Linkuvienė et al. 2018; Kairys et al. 2019). For these reasons, FTSA is widely employed for such applications as drug-candidate screening (Pantoliano et al. 2001; Yanchunas et al. 2005; McDonnell et al. 2009). FTSA is easiest to apply for single-domain and single-binding site systems. Different domains within a multi-domain protein can unfold independently at different temperatures, and thus yield multiple denaturation signals that can be difficult to interpret. Therefore, complex multi-domain proteins and their assemblies are a challenge for study by FTSA, but not impossible targets for it (Baranauskienė et al. 2010; Layton and Hellinga 2011).

While FTSA is a powerful low-cost high-throughput technique for a variety of applications, as with any other method, it is only as reliable as the laboratory practice and data analysis behind the experiments. Therefore, here we present the standard operating procedures for protein–ligand binding measurements and stability evaluation using FTSA.

Protocols

The equipment required for FTSA is a real-time thermocycler, which is routinely used for qPCR experiments. It should be able to gradually increase the temperature of the sample and monitor its fluorescence. Thermocyclers typically have a plate-based or a centrifugal rotary design. The latter offers higher precision by ensuring consistent temperature across all the samples. Appropriate 0.1 mL or 0.2 mL sized tubes or plates with similarly sized wells should be used. The volumes used in the experiment depend on the thermocycler and labware — the light beam must pass through the solution. Even smaller volumes can be analyzed on specialized devices using capillaries.

A critical component is the reporter dye — a molecule whose fluorescence indicates protein unfolding. The most common reporter dyes are ANS (1-anilinonaphthalene-8-sulfonic acid, or its derivative bis-ANS) (Anderson and Weber 1969; Slavik et al. 1982; Matulis and Lovrien 1998;

Matulis et al. 1999; Cimperman and Matulis 2011) and SYPRO Orange (Lo et al. 2004; Niesen et al. 2007). In our experience, ANS exhibits more favorable properties for FTSA than SYPRO Orange because ANS binds weaker to proteins than SYPRO Orange. However, the excitation/emission wavelength profile of SYPRO Orange (peaking at ~ 470 nm excitation and ~ 570 nm emission) is more compatible with a wider range of real-time thermocyclers. Therefore, if the instrument is equipped for wavelengths required for ANS ($\lambda_{\text{EX}} \sim 350$ nm, $\lambda_{\text{EM}} \sim 470$ nm for ANS; $\lambda_{\text{EX}} \sim 390$ nm and $\lambda_{\text{EM}} \sim 490$ nm for bis-ANS), it is advised to use this dye. Otherwise, SYPRO Orange is a suitable alternative. The following protocol is written for ANS but can also be applied to SYPRO Orange, provided that the amount of the dye added to the protein solution is adjusted accordingly. The 10 mM ANS stock should be prepared in DMSO, aliquoted, and stored in the dark, in an airtight container. ANS can be exposed to light for short periods of time, such as when preparing for the experiment. For some proteins, the reporter dye is not required and the fluorescence of innate tryptophan ($\lambda_{\text{EX}} \sim 295$ nm, $\lambda_{\text{EM}} \sim 350$ nm) can be monitored instead. However, it cannot be applied to all proteins and the resulting signal is significantly lower than that of the reporter dyes.

Protocol for stability evaluation

This protocol should be used to screen various conditions to identify optimal solutions for protein storage, subsequent FTSA ligand binding assays or other downstream applications — biophysical, biochemical, or structural biology assays. It can also be applied to identify protein refolding conditions (Biter et al. 2016; Wang et al. 2017). A large variety of conditions — pH, buffer, ionic strength, additives — can be assayed (see Sect. "Optimization" for more details). Commercial screening stocks are available but home-made ones can be designed and tailored to specific needs.

1. Decide how many different samples and of what volume are needed, what concentrations of protein (typically within 2.5–20 μM range) and dye (typically 50 μM ANS) to use (see also Sects. "Optimization" and "Considerations" for optimization and considerations).
2. Prepare the solutions you want to screen and transfer to the final assay tubes or plate (e.g. 17 μL per tube or well if the final volume would be 20 μL and the protein stock will be tenfold from the final concentration).
3. Prepare protein stock at ~ 100 μM concentration in a minimal buffer.

Note: If possible, it is even better to have a more concentrated protein so that the assayed solutions would be less diluted in the final sample. However, handling small volumes might decrease pipetting precision. If the origi-

nal protein stock concentration is low and there is a risk of the protein storage buffer interfering with the assay, it is advisable to dialyze it against a minimal buffer in advance.

4. Distribute the protein stock, prepared in step 3, to the tubes or wells with the assayed solutions, prepared in step 2.
5. Dilute the ANS stock to 1 mM in high-quality water (e.g., Millipore standard) and distribute to the tubes or wells with the assayed solutions, prepared in step 4. Mix well and make sure the tubes and plates are well sealed.
6. Insert the tubes or plates into the qPCR machine, and run the following program:
 - 1) Pre-incubate for 3–5 min at the starting temperature (e.g. 25 $^{\circ}\text{C}$).
 - 2) Increase the temperature gradually from 25 to 99 $^{\circ}\text{C}$, at the rate of 1 $^{\circ}\text{C}$ per min. Measure the fluorescence every 30 s using the appropriate filter. If possible, measure at several different sensitivity settings (gains).
7. Analyze the data to choose the optimal conditions.
 - a) Analyze the raw data (fluorescence intensity as a function of temperature) to identify the melting temperatures as described in (Layton and Hellinga 2010; Petrauskas et al. 2019). They correspond to the inflection points of the native-to-unfolded transition slopes (Fig. 1, panel b). The higher the T_m value, the more stable is the protein under the corresponding conditions.
 - b) In cases when the data is a superposition of several reactions and thus it is difficult to determine T_m values accurately, a rough estimate of melting temperatures can be obtained from the peak positions of the first derivatives of the melting curves.

Example experiment: pH-dependence of human Hsp90 α N-terminal domain stability

Materials:

ANS, 10 mM.

array of buffers (50 mM CH_3COONa , 25 mM $\text{Na}_2\text{B}_4\text{O}_7$, 100 mM NaCl, and 50 mM of phosphates — H_3PO_4 , Na_2HPO_4 , NaH_2PO_4 , and Na_3PO_4 — at a ratio required to achieve the selected pH: 5.0–10.5).

protein dilution buffer (2.5 mM HEPES, pH 7.5, 100 mM NaCl).

Hsp90 α N, 200 μM in protein dilution buffer.

1. The plan is to test 12 different buffers in 10 μL samples containing 5 μM protein and 50 μM ANS.
2. Put 8 μL of each buffer into separate 0.1 mL tube.

3. Prepare protein stock (at 50 μM concentration) by mixing 3.5 μL of the 200 μM Hsp90 α N with 10.5 μL of the protein dilution buffer.
4. Distribute 1 μL of the 50 μM protein stock, prepared in step 3, to each of the tubes with the assayed solutions, prepared in step 2.
5. Dilute the ANS stock to 0.5 mM in high-quality water (e.g., Millipore standard) and distribute 1 μL of it to the tubes with the assayed solutions, prepared in step 4. Mix well and make sure the tubes are closed properly.
6. Insert the tubes into the Corbett Rotor-Gene 6000 (Qia-gen Rotor-Gene Q) spectrofluorimeter, and run the following program:
 - 1) Pre-incubate for 3–5 min at 25 $^{\circ}\text{C}$.
 - 2) Increase the temperature gradually from 25 to 99 $^{\circ}\text{C}$, at the rate of 1 $^{\circ}\text{C}$ per min. Measure the fluorescence every 30 s using the Blue channel ($\lambda_{\text{EX}} \sim 365$ nm, $\lambda_{\text{EM}} \sim 460$ nm) at 3 different gains.
7. Analyze the data to choose the optimal conditions (Fig. 2).

Protocol for ligand binding measurements

This protocol should be used to evaluate the protein–ligand interaction by obtaining the binding constant (K_b). Protein thermal stability is assayed at different concentrations of the ligand, and K_b can be calculated from the resulting dosing curve (Fig. 1, panel c).

Tubes (such as the 100 μL or 200 μL sized) and rotary real-time thermocyclers are more suitable than plates for precise measurements. Regarding reaction volume, refer to Sect. "Considerations".

1. Decide how many different samples (i.e. different ligand concentrations) are needed.

Typically, 8 samples (including the no-ligand control) are being analyzed but this number can be adjusted as needed: increased up to 16 samples to get a more precise curve fit or reduced if the available amount of protein and/or ligand is limited.

2. Prepare the ligand stock at 20 mM or the highest possible concentration in DMSO.

Note: DMSO is commonly required to ensure solubility of most ligands being assayed; ligands that are highly soluble in water could be dissolved in water or appropriate buffer. The steps 3 and 4 should be skipped for aqueous ligand stocks, and instead a 0.4 mM solution and its subsequent serial dilutions should be prepared directly in the buffer of the protein solution.

3. Generate twofold serial dilutions of the ligand stock in DMSO (Fig. 1, panel a). Typically, 8 samples are being prepared and analyzed, including the stock concentration and a sample containing only DMSO and no ligand.

Note: The dilution range can be adjusted as needed.

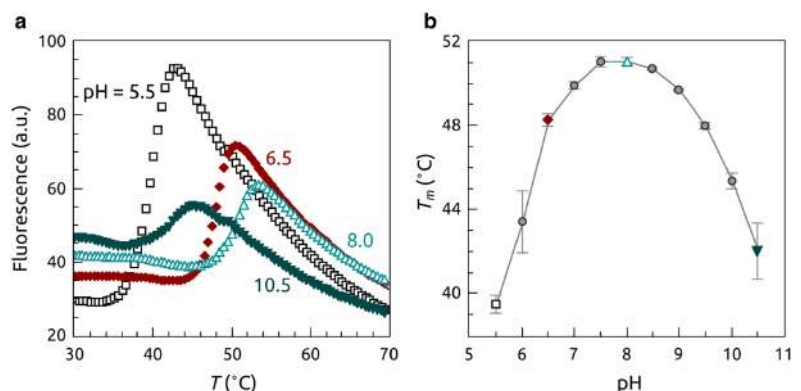
4. Dilute the prepared samples 50-fold with the buffer – so that the final DMSO concentration would be 2% v/v. Mix them well by pipetting.
5. Prepare protein stock at 10 μM by diluting the protein in the buffer, supplemented with 100 μM ANS and DMSO so that final DMSO concentration would be 2%.

Note: Protein and dye concentrations can be adjusted depending on the system (see Sect. "Optimization" for details), commonly within ranges of 2.5–20 μM of protein and 50–200 μM dye.

If the original protein stock concentration is low and there is a risk of the protein storage buffer interfering with the assay, it is advisable to dialyze it against the buffer in advance.

6. Distribute the protein stock to the tubes and add the ligand dilutions (from step 4) at a 1:1 ratio. Mix carefully by pipetting, make sure not to leave bubbles that could interfere with the measurements.

Fig. 2 Application of FTSA for recombinant protein (human Hsp90 α N-terminal domain) characterization; determination of its stability dependence on pH. **a** Raw FTSA data — ANS fluorescence intensities as a function of temperature at different pH. **b** The plot of determined T_m values as a function of pH. The protein is most stable at approximately pH 8.0



7. Insert the tubes into the real-time thermocycler, and run the following program:

- 1) Pre-incubate for 3–5 min at the starting temperature (e.g. 25 °C).
- 2) Increase the temperature gradually from 25 °C to 99 °C, at the rate of 1 °C per min. Measure the fluorescence every 30 s at the appropriate wavelength. If possible, measure at several different sensitivity levels (also referred to as gains).
- 3) (Optional) Cool down the device to 25 °C for 5 min to expedite subsequent experiment.

Note: If you are familiar with the protein, the temperature range can be narrower. As a rule of thumb, the starting temperature should be at least 10 °C lower than the base of the protein transition curve, and the gradient should end at a temperature which is at least 20 °C higher than the expected highest maximum of the protein transition curve (which will increase when protein is stabilized by the bound ligand).

8. Analyze the data to calculate K_b :

- 1) Analyze the raw data (fluorescence intensity as a function of temperature) to identify the melting temperatures as described in (Layton and Hellinga 2010; Petrauskas et al. 2019). They correspond to the inflection points of the native-to-unfolded transition slopes (Fig. 1, panel b).
- 2) Plot these T_m values as a function of ligand concentration to generate the dosing curve (Fig. 1, panel c).
- 3) Fit the model to the dosing curve to obtain the K_b as previously described (Petrauskas et al. 2019).

Example experiment: acetazolamide binding to human carbonic anhydrase II

Materials:

ANS, 10 mM in DMSO.
 Acetazolamide (AZM).
 Buffer A (50 mM HEPES, pH 7.5, 100 mM NaCl).
 Carbonic anhydrase II (CAII), 180 μ M in buffer A.
 DMSO

1. The goal is to assay 8 different concentrations of AZM (0, 3.13, 6.25, 12.5, 25, 50, 100, and 200 μ M) with 5 μ M CAII in buffer A, with a total 2% DMSO and 50 μ M ANS. Sample volume: 10 μ L.
2. Prepare a 20 mM acetazolamide solution in DMSO.
3. Generate twofold serial dilutions of the 20 mM AZM solution in DMSO (Fig. 1, panel a). For this, add 5 μ L DMSO to all tubes except the first, to which add 10 μ L of the 20 mM acetazolamide solution. Transfer 5 μ L

from the first tube to the next one, mix, then transfer 5 μ L of this mix to the third tube with a new pipet tip, and so on. Continue until you reach the 7th tube. The last tube should contain only DMSO and no ligand should be transferred there.

4. Mix 2 μ L of the samples from step 3 with 98 μ L of buffer A.
5. Prepare a 10 μ M protein stock by mixing 41.6 μ L buffer A with 0.45 μ L DMSO, 0.45 μ L 10 mM ANS, and 2.5 μ L of the 180 μ M CAII.
6. Distribute 5 μ L of the protein stock from step 5 to each of the 8 tubes that are compatible with the thermocycler. To each, add 5 μ L of the corresponding ligand dilutions (from step 4). Mix carefully.
7. Insert the tubes into the Corbett Rotor-Gene 6000 (Qia-gen Rotor-Gene Q) spectrofluorimeter, and run the following program:
 - 1) Pre-incubate for 3 min at 25 °C.
 - 2) Increase the temperature gradually from 25 to 99 °C, at the rate of 1 °C per min. Measure the fluorescence every 30 s using the Blue channel ($\lambda_{EX} \sim 365$ nm, $\lambda_{EM} \sim 460$ nm).
8. Analyze the data (Fig. 1, panels b and c) to calculate K_b .

Optimization

When working with a new protein, it is advisable to first perform a screening experiment to check the quality of the fluorescence readout. Example FTSA curves can be seen in Figs. 1b and 2a. It is recommended to adjust the experimental setup so that the fluorescence peak would be at $\sim 80\%$ of the arbitrary fluorescence units. A higher peak can easily go over the quantifiable detection limit while a low fluorescence signal tends to be noisy and thus less accurate. Additionally, a clear contrast between the peak and the initial baseline is needed for good data interpretation. While the overall signal strength can be adjusted by changing the sensitivity level in the software settings, this would not improve the ratio between the peak and the initial fluorescence. To solve this problem, try to increase the concentration of the reporter dye (ANS concentration can be increased up to 100 μ M), or, if that does not help sufficiently, adjust the protein concentration. Finally, try assaying different buffer solutions (see protocol in Sect. "Protocol for stability evaluation") in case the protein stability is impaired in the assay buffer.

For initial screening of a new protein for ligand binding measurements, 4 samples can be sufficient: containing the highest ligand concentration, tenfold dilution of the ligand, and two no-ligand controls. Prepare the samples in a fashion similar to that described in Sect. "Protocol for ligand binding measurements". Even when screening different protein or ANS concentrations, be sure to have at least 2

replicates of each condition. Moreover, measure such samples at different sensitivity settings since the fluorescence signal level varies a lot in response to the concentrations of reporter dye and protein.

In cases when the melting temperature of the protein is extremely high ($> 90^{\circ}\text{C}$), addition of destabilizing agents (such as guanidinium chloride) can aid to compare conditions in the stability assay. To evaluate ligand binding, try to reach destabilizing levels of solution pH. However, deviation from physiological conditions should be taken into consideration when assessing the binding data.

Considerations

Assay volumes

The assay volume depends on the tube/plate and the instrument — the beam must pass through the solution during measurement. Typically, 10 and 20 μL samples are analyzed in 0.1 and 0.2 mL tubes, respectively, in the rotating thermocyclers. In the case of plates, the samples tend to evaporate significantly, so proper sealing or addition of mineral oil may be required. Abnormal fluorescence signal curve (especially with fluorescence suddenly spiking up or plummeting down) can indicate that the beam is passing through the air instead of sample due to evaporation or air bubbles in the sample or that the transparency of the sample is compromised.

Protein

The protein sample should be as pure as possible. The protein also has to be structured, stable and soluble under the starting experimental conditions, and should not interfere with the fluorescence in the detection range. It is estimated that 15–25% of recombinant proteins provide unsuitable denaturation curves (Crowther et al. 2010), most often due to intrinsic disorder, lack of stability at room temperature, or lack of inner hydrophobic patches that would be exposed to the dye once the protein is unfolded. Therefore, when working with a new protein, it is suggested to first conduct a minimal set of experiments to check that the protein is suitable for the technique. These cases can usually be identified by very high fluorescence at the initial temperatures and/or lack of a sigmoidal transition to the unfolded state, which cannot be rectified by optimization procedures (Sect. "Optimization"). For the calculation of ligand binding constants, mono-domain protein constructs are preferred. Denaturation of proteins with multiple domains could result in multiple overlapping peaks that significantly complicate the binding data analysis.

Ligands and additives

Other reaction components that exhibit intrinsic fluorescence within the detection range can also hinder the measurements. As some ligands increase fluorescence, it is advisable to repeat the experiment without protein if the denaturation profile looks atypical. If a critical component interferes with the monitored fluorescence, consider switching to a different reporter dye. It should be noted that some compounds tend to stabilize proteins while others tend to destabilize them. For example, from our experience usually 2% DMSO concentration is safe to use, while 4% and higher concentrations can significantly affect protein stability.

Buffer

The solution overall has to be transparent, so that the light can pass efficiently through. Temperature-induced changes in buffer pH should also be taken into consideration, which is especially critical for ligand binding measurements. Therefore, less temperature-sensitive buffers, such as phosphate and HEPES, are used much more often than, e.g., Tris buffer, which changes its pH significantly in response to changes in temperature.

Acknowledgements Authors acknowledge the ARBRE network and the COST project MOBIEU CA15126.

References

- Anderson SR, Weber G (1969) Fluorescence polarization of the complexes of 1-anilino-8-naphthalenesulfonate with bovine serum albumin. Evidence for preferential orientation of the ligand. *Biochemistry* 8:371–377
- Baranauskienė L, Hilvo M, Matulienė J et al (2010) Inhibition and binding studies of carbonic anhydrase isozymes I, II and IX with benzimidazo[1,2-c][1,2,3]thiadiazole-7-sulphonamides. *J Enzyme Inhib Med Chem* 25:863–870. <https://doi.org/10.3109/14756360903571685>
- Biter AB, de la Peña AH, Thapar R et al (2016) DSF guided refolding as a novel method of protein production. *Sci Rep* 6:18906. <https://doi.org/10.1038/srep18906>
- Brandts JF, Lin LN (1990) Study of strong to ultratight protein interactions using differential scanning calorimetry. *Biochemistry* 29:6927–6940
- Cimmperman P, Matulis D (2011) Protein thermal denaturation measurements via a fluorescent dye. In: Podjarny A, Dejaegere AP, Kieffer B (eds) *RSC biomolecular sciences*. Royal Society of Chemistry, Cambridge, pp 247–274
- Cimmperman P, Baranauskienė L, Jachimovičiūtė S et al (2008) A quantitative model of thermal stabilization and destabilization of proteins by ligands. *Biophys J* 95:3222–3231. <https://doi.org/10.1529/biophysj.108.134973>
- Crowther GJ, He P, Rodenbough PP et al (2010) Use of thermal melt curves to assess the quality of enzyme preparations. *Anal Biochem* 399:268–275. <https://doi.org/10.1016/j.ab.2009.12.018>
- Gao K, Oerlemans R, Groves MR (2020) Theory and applications of differential scanning fluorimetry in early-stage drug

- discovery. *Biophys Rev* 12:85–104. <https://doi.org/10.1007/s12551-020-00619-2>
- Kairys V, Baranauskienė L, Kazlauskienė M et al (2019) Binding affinity in drug design: experimental and computational techniques. *Expert Opin Drug Discov*. <https://doi.org/10.1080/17460441.2019.1623202>
- Layton CJ, Hellinga HW (2010) Thermodynamic analysis of ligand-induced changes in protein thermal unfolding applied to high-throughput determination of ligand affinities with extrinsic fluorescent dyes. *Biochemistry* 49:10831–10841. <https://doi.org/10.1021/bi101414z>
- Layton CJ, Hellinga HW (2011) Quantitation of protein–protein interactions by thermal stability shift analysis. *Protein Sci* 20:1439–1450. <https://doi.org/10.1002/pro.674>
- Linkuvienė V, Zubrienė A, Manakova E et al (2018) Thermodynamic, kinetic, and structural parameterization of human carbonic anhydrase interactions toward enhanced inhibitor design. *Q Rev Biophys*. <https://doi.org/10.1017/S0033583518000082>
- Lo M-C, Aulabaugh A, Jin G et al (2004) Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal Biochem* 332:153–159. <https://doi.org/10.1016/j.ab.2004.04.031>
- Matulis D, Lovrien R (1998) 1-anilino-8-naphthalene sulfonate anion-protein binding depends primarily on ion pair formation. *Biophys J* 74:422–429. [https://doi.org/10.1016/S0006-3495\(98\)77799-9](https://doi.org/10.1016/S0006-3495(98)77799-9)
- Matulis D, Baumann CG, Bloomfield VA, Lovrien RE (1999) 1-Anilino-8-naphthalene sulfonate as a protein conformational tightening agent. *Biopolymers* 49:451–458. [https://doi.org/10.1002/\(SICI\)1097-0282\(199905\)49:6%3c451::AID-BIP3%3e3.0.CO;2-6](https://doi.org/10.1002/(SICI)1097-0282(199905)49:6%3c451::AID-BIP3%3e3.0.CO;2-6)
- Matulis D, Kranz JK, Salemme FR, Todd MJ (2005) Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor. *Biochemistry* 44:5258–5266. <https://doi.org/10.1021/bi048135v>
- McDonnell PA, Yanchunas J, Newitt JA et al (2009) Assessing compound binding to the Eg5 motor domain using a thermal shift assay. *Anal Biochem* 392:59–69. <https://doi.org/10.1016/j.ab.2009.05.044>
- Niesen FH, Berglund H, Vedadi M (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Prot* 2:2212–2221. <https://doi.org/10.1038/nprot.2007.321>
- Pantoliano MW, Petrella EC, Kwasnoski JD et al (2001) High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J Biomol Screen* 6:429–440. <https://doi.org/10.1177/108705710100600609>
- Petrauskas V, Zubrienė A, Todd MJ, Matulis D (2019) Inhibitor binding to carbonic anhydrases by fluorescent thermal shift assay. In: Matulis D (ed) *Carbonic anhydrase as drug target: thermodynamics and structure of inhibitor binding*. Springer International Publishing, Cham, pp 63–78
- Scott DE, Spry C, Abell C (2016) Differential scanning fluorimetry as part of a biophysical screening cascade. In: Erlanson DA, Jahnke W (eds) *Methods and principles in medicinal chemistry*. Wiley-VCH Verlag GmbH & Co KGaA, Weinheim, pp 139–172
- Semisotnov GV, Rodionova NA, Razgulyaev OI et al (1991) Study of the “molten globule” intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers* 31:119–128
- Slavik J, Horak J, Rihova L, Kotyk A (1982) Anilino-naphthalene sulfonate fluorescence and amino acid transport in yeast. *J Membr Biol* 64:175–179
- Vedadi M, Niesen FH, Allali-Hassani A et al (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci* 103:15835–15840. <https://doi.org/10.1073/pnas.0605224103>
- Wang Y, van Oosterwijk N, Ali AM et al (2017) A systematic protein refolding screen method using the DGR approach reveals that time and secondary TSA are essential variables. *Sci Rep* 7:9355. <https://doi.org/10.1038/s41598-017-09687-z>
- Yanchunas J, Langley DR, Tao L et al (2005) Molecular basis for increased susceptibility of isolates with atazanavir resistance-conferring substitution I50L to other protease inhibitors. *Antimicrob Agents Chemother* 49:3825–3832. <https://doi.org/10.1128/AAC.49.9.3825-3832.2005>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Publication 5

Uncertainty in protein–ligand binding constants: asymmetric confidence intervals versus standard errors

Vaida Paketurytė, Vytautas Petrauskas, Asta Zubrienė, Olga Abian,
Margarida Bastos, Wen-Yih Chen, Maria João Moreno, Georg Krainer,
Vaida Linkuvienė, Arthur Sedivy, Adrian Velazquez-Campoy, Mark A
Williams, Daumantas Matulis

Eur. Biophys. J., 2021, 50, 3–4, 661–670

DOI: 10.1007/s00249-021-01518-4



ORIGINAL ARTICLE

Uncertainty in protein–ligand binding constants: asymmetric confidence intervals versus standard errors

Vaida Paketurytė¹  · Vytautas Petrauskas¹  · Asta Zubrienė¹  · Olga Abian^{2,3,4,5,6}  · Margarida Bastos⁷  · Wen-Yih Chen⁸  · Maria João Moreno⁹  · Georg Krainer¹⁰  · Vaida Linkuvienė¹  · Arthur Sedivy¹¹  · Adrian Velazquez-Campoy^{2,3,4,5,12}  · Mark A. Williams¹³  · Daumantas Matulis¹ 

Received: 2 October 2020 / Revised: 5 February 2021 / Accepted: 13 March 2021 / Published online: 10 April 2021
© European Biophysical Societies' Association 2021

Abstract

Equilibrium binding constants (K_b) between chemical compounds and target proteins or between interacting proteins provide a quantitative understanding of biological interaction mechanisms. Reported uncertainties of measured experimental parameters are critical for decision-making in many scientific areas, e.g., in lead compound discovery processes and in comparing computational predictions with experimental results. Uncertainties in measured K_b values are commonly represented by a symmetric normal distribution, often quoted in terms of the experimental value plus–minus the standard deviation. However, in general, the distributions of measured K_b (and equivalent K_d) values and the corresponding free energy change ΔG_b are all asymmetric to varying degree. Here, using a simulation approach, we illustrate the effect of asymmetric K_b distributions within the realm of isothermal titration calorimetry (ITC) experiments. Further we illustrate the known, but perhaps not widely appreciated, fact that when distributions of any of K_b , K_d and ΔG_b are transformed into each other, their degree of asymmetry is changed. Consequently, we recommend that a more accurate way of expressing the uncertainties of K_b , K_d , and ΔG_b values is to consistently report 95% confidence intervals, in line with other authors' suggestions. The ways to obtain such error ranges are discussed in detail and exemplified for a binding reaction obtained by ITC.

Keywords Isothermal titration calorimetry · Confidence intervals · Standard error · Log-normal distribution · Dissociation constant · Binding constant

Introduction

Interactions between biomolecules are central to many areas of biomedicine. Protein–protein interactions (Typas and Sourjik 2015; Pierce et al. 1999) are important, e.g., in immunological antibody–antigen binding reactions (Dam et al. 2008), or gene regulatory protein–nucleic acid interactions (Wells et al. 1980; Buurma and Haq 2007; Salim and Feig 2009). Furthermore, interactions of small molecular weight compound with proteins are fundamental to the action of many metabolic enzymes and their regulators, and in drug discovery during the search for lead compounds, as well as in the final characterization of promising therapeutic drugs (Geschwindner et al. 2015; Renaud et al. 2016; Ladbury et al. 2010).

Numerous techniques are used to determine biomolecular interactions (Renaud et al. 2016; Ciulli 2013), such as the inhibition of enzymatic activity (Šmirnovienė et al. 2017), surface plasmon resonance (Myszka and Rich 2000;

Special Issue: COST Action CA15126, MOBIEU: Between atom and cell.

✉ Margarida Bastos
mbastos@fc.up.pt

✉ Arthur Sedivy
arthur.sedivy@vbcf.ac.at

✉ Adrian Velazquez-Campoy
adrianvc@unizar.es

✉ Mark A. Williams
ma.williams@bbk.ac.uk

✉ Daumantas Matulis
daumantas.matulis@bti.vu.lt; matulis@ibt.lt

Extended author information available on the last page of the article

Patching 2014; Olaru et al. 2015), isothermal titration calorimetry (ITC) (Krimmer and Klebe 2015; Callies and Daranas 2016; Falconer 2016; Vega et al. 2016; Chaires 2008; Leavitt and Freire 2001), thermal shift assay (differential scanning fluorimetry) (Pantoliano et al. 2001; McDonnell et al. 2009; Yanchunas et al. 2005; Cimperman and Matulis 2011; Cimperman et al. 2008) and numerous others (Renaud et al. 2016). All these techniques are expected to provide comparable values of intermolecular interaction affinities provided that measurements are feasible at nearly identical conditions.

Reporting the errors of association reactions must be done consistently and accurately to enable reliable interpretation and reuse of results (Jarmoskaite et al. 2020). In most scientific literature, the uncertainty (i.e., the repeatability) of a binding equilibrium measurement is expressed as $\pm x$ (or $x\%$), accounting for the standard deviation or error of the measured value, with the underlying assumption that the measured affinity values are distributed randomly according to a normal (i.e., Gaussian) statistical distribution (Krimmer and Klebe 2015; Schnapp et al. 2016; Ladbury and Doyle 2004; Lafont et al. 2007; Dullweber et al. 2001; Rühmann et al. 2015; Gaspari et al. 2016; Pulido et al. 2015; Hörtnner et al. 2007; Rechlin et al. 2017; Cheng et al. 2017; Huschmann et al. 2016; Guan et al. 2013; Ren et al. 2014; Krishnamurthy et al. 2007). For a normal distribution, the reported $\pm x$ -value typically corresponds to the 68.3% symmetric confidence interval ($CI_{s,68.3}$). In some cases, the reported symmetric uncertainty is estimated from analysis of several repeated measurements via the appropriate Student's t -distribution. However, the uncertainty is frequently retrieved only from the fitting program used to estimate the parameter from a single set of experimental data (Brautigam et al. 2016).

The equilibrium binding constant K_b (also referred to as the association or affinity constant) is inversely related to the dissociation equilibrium constant $K_d = 1/K_b$ and is related to the change in the standard Gibbs energy upon binding, ΔG_b , (often written with the naught symbol, ΔG_b^0 , which is omitted here for terms of simplicity) by the equations: $\Delta G_b = -RT \ln K_b$, or $K_b = e^{-\Delta G_b/RT}$. These logarithmic, exponential, and reciprocal relationships do not preserve the shape of distribution of values. For example, a continuous probability distribution of a random variable, whose logarithm is normally distributed follows a log-normal distribution, taking only positive, real values. Therefore, if a normal distribution is assumed for ΔG_b , K_b must have a log-normal distribution.

The inherent asymmetry in parameter distribution has consequences for the reporting of errors. Symmetric $\pm x$ -value is not an appropriate way of accurately expressing the uncertainty for ΔG_b , K_b and related values. Asymmetric F -statistics-associated CIs have been introduced as

an elegant approach to report error ranges and the propagation of CIs in non-linear, asymmetrical variance spaces, accounting for non-normal distributed parameter distributions, including ΔG_b , K_b and K_d (Kemmer and Keller 2010; Krainer and Keller 2015; Krainer et al. 2012; Broecker et al. 2011; Johnson 1983).

In this work, we explore in detail the unrealistic nature of the assumption of a normal distribution of K_b , K_d and other thermodynamic quantities through various examples and evaluate a practical approach, based on previously established procedures (Kemmer and Keller 2010; Krainer and Keller 2015; Krainer et al. 2012; Broecker et al. 2011), to obtain estimates of the asymmetric F -statistics-associated CIs that are both simpler to treat consistently in transforming between different expressions for the affinity and better reflect the real uncertainties in the data.

Results and discussion

We use simulated distribution curves to show how the mathematical relationships between the parameters K_b , K_d , and ΔG_b affect the asymmetry of their distributions.

Case 1: Transformation of the distribution of K_b and ΔG_b values assuming that each has a normal distribution

Here, we show how transformations of K_b and ΔG_b values results in an asymmetry of their distributions. For this, we assumed that either K_b or ΔG_b values are normally distributed and generated a set of 10,000 random K_b 's with mean value $2 \times 10^7 \text{ M}^{-1}$ and standard deviation of $\pm 0.6 \times 10^7 \text{ M}^{-1}$ (Fig. 1a) and a set of 10,000 random ΔG_b 's with mean value -41.67 kJ/mol and standard deviation of $\pm 0.77 \text{ kJ/mol}$ (Fig. 1d). Note that $K_b = 2 \times 10^7 \text{ M}^{-1}$ corresponds to $\Delta G_b = -41.67 \text{ kJ/mol}$ at 25°C . Each set of values is then converted to the other representation through the appropriate logarithmic or exponential formula (Fig. 1b, e), and in both cases to the dissociation constant K_d (Fig. 1c, f). These transformations show that the Gibbs energy change, and the binding and dissociation constants do not follow a normal distribution at the same time. The transformation between K_b and ΔG_b in either direction introduces a similar amount of skew into the distribution, and that the transformation between K_b and K_d may impart even greater skew.

When considering the kinetics of the interactions, inaccuracies in the reported parameters (most probable value and uncertainty) may be even more severe. The rate constants are related with the Gibbs energy variation between the reactants and the transition state through an exponential function, and this energy variation is usually larger than that observed between the reactants and the products of the

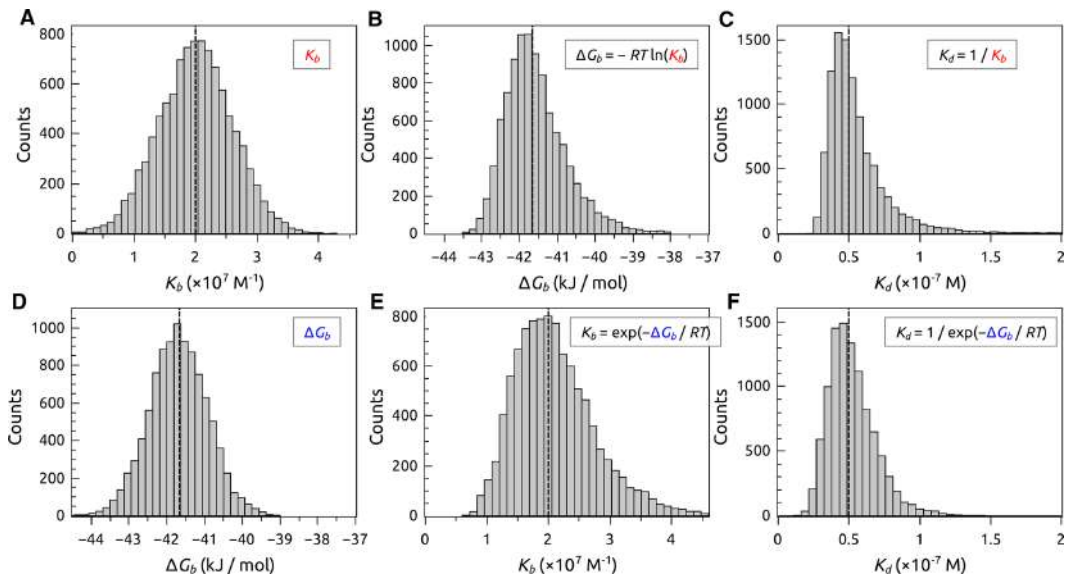


Fig. 1 Illustrating the effect of transformation of K_b and ΔG_b values into each other and into K_d distributions of protein–ligand binding. **a** Distribution of 10,000 random values of K_b generated for a normal distribution with mean $2 \times 10^7 \text{ M}^{-1}$ and standard deviation of $\pm 0.6 \times 10^7 \text{ M}^{-1}$. **b** Distribution of ΔG_b values calculated from the 10,000 K_b values in (a). **c** The distribution of dissociation constants K_d corresponding to data in (a). **d** Distribution of 10,000 random values of the Gibbs energy of binding generated for a normal distribu-

tion with mean -41.67 kJ/mol and a standard deviation of 0.77 kJ/mol . **e** Distribution of K_b values calculated from the ΔG_b values in (d). **f** The dissociation constant calculated from (d). The logarithmic transformation introduces a skew of 1.3 (b) and the exponential transformation a skew of 0.9 (e) from the normal distributions. The greatest skewness arises from the transformation of the normally distributed K_b to K_d [skew is equal to 4.3 in (c)]

transformation. This situation is analyzed in the Supplementary Material.

Confidence intervals are a consistent and inter-convertible way to accurately represent measurement uncertainty

The previous illustration shows that it is not warranted to use a symmetric $\pm x$ -value as an accurate expression of the uncertainty for all three quantities ΔG_b , K_b and K_d describing the same physical equilibrium. For consistent reporting of uncertainties in experimental data, it is always necessary to state the range of uncertainty $[x_{\text{low}}, x_{\text{high}}]$ which makes apparent any degree of asymmetry. Asymmetry in any skewed distribution is more evident toward its extremes; thus, the use of a central 68.3% confidence interval does not effectively describe the asymmetry. Consequently, it is advantageous for clear description to report the larger 95% confidence interval (CI_{95}), which in any case typically better represents the range in which the true value of the parameter is likely to be found.

Given that a normally distributed ΔG_b is transformed into a log-normal distribution for K_b , one could imagine that given some CI for ΔG (symmetrical), it might be required to calculate the 95% confidence interval of this log-normal distribution to obtain uncertainties in K_b . However, as the individual values for the quantities are correctly transformed by the exponential, logarithmic and reciprocal relationships, it is simple to directly convert the lower and upper values of any confidence interval between representations (this is rigorously true for one-to-one mono-parametric conversions). For example, using the $\Delta G_b = -41.67 \pm 0.77 \text{ kJ/mol}$ from Case 1 gives a $\text{CI}_{95} = [-43.18, -40.16]$ for ΔG_b (as 95% CI limits are $1.96 \times \sigma$ for a normal distribution in the limit of a large number of data points). The upper limit of the CI interval for $K_b = \exp(-\Delta G_{b,\text{lower}}/RT) = \exp(43.18/2.47896) = 3.68 \times 10^7 \text{ M}^{-1}$. Similar transformation for the lower limit gives a $\text{CI}_{95} = [1.09, 3.68] \times 10^7 \text{ M}^{-1}$ (which can be seen to match with the distribution in Fig. 1e for which the CI_{95} is $[1.09, 3.69] \times 10^7 \text{ M}^{-1}$). Thus, consistently reporting CI_{95} makes it easy to transform between representations preserving all information regarding uncertainties in experimental values. This procedure also eliminates the possibility of

getting error intervals with negative values for the equilibrium constants.

The logarithmic relationship between K_b and ΔG_b , is also important for other practical purposes such as calculating averages from a set of determinations. The mean of several ΔG_b values can be calculated using the arithmetic average, while the mean of several K_b values should be calculated using the geometric average. Alternatively, the value to report for K_b can be calculated from the obtained arithmetic average of ΔG_b values. That way, the correspondence between ΔG_b and K_b averages is maintained in the same way that the correspondence between confidence interval limits is maintained.

Case 2: Simulated error distributions for ITC measurements of 1:1 binding

The preceding hypothetical illustrations (Case 1) leave open the question of whether the uncertainty in any parameter might be expected to have a strongly asymmetric or near normal distribution, and, if so, under what circumstances. To investigate these issues, we simulated the impact of binding affinity changes on the variability of ITC experiments.

In common with many methods for obtaining binding constants, ITC experiments are performed as a titration where successive injections of one reactant species into a fixed amount (or concentration) of the other species leads to progressive saturation of a binding site. A signal monitors formation of the bound form and the values for the signal (transformed into heat, Q) are fit to an equation describing the relationship between the heat released or absorbed and the thermodynamic parameters K_b (or ΔG_b), the enthalpy change ΔH_b and the apparent stoichiometry, that are, thus, determined as parameters of the fit. The measurement errors are typically propagated to the fitted parameters in a way that depends on the equation describing the titration and a range of experimental variables, e.g., the number of injections, the final degree of saturation and concentrations of reactants.

The measurement's variability for the heat of each injection is random normally distributed (as it arises from the combined random effects of mechanical variability of the injection volume, electrical noise and how these impact upon the software process for integrating the heat signal for each injection).

To simulate the effect of measurement variation, random values from an appropriate normal distribution are added to the theoretically expected injection heat values for an ideal 1:1 binding reaction. The simulations used here model a VP-ITC instrument (MicroCal/Malvern). Tellinghuisen has determined that for this instrument (when set for high maximum injection heats) the standard deviation of measurements (when including the effect of subtracting blanks or heats of dilution) is approximately constant at 3 μJ for

injections of heat $< 500 \mu\text{J}$ (Tellinghuisen 2005). Recently, alternative error models with a smaller constant standard deviation of 0.5–0.9 μJ plus an injection-heat-dependent term of 0.002–0.01 μJ per μJ (not including effects of subtraction) have been proposed (Tellinghuisen, 2017,2018). We use the original injection-heat-independent error model here, but note that it may overestimate the error for an optimally set up instrument for a low-heat biochemical reaction; however, the absolute value of the error is not critical as it is the proportional error (the signal to noise) that influences the shape of the parameter error distribution. The variation of the fitted parameters, K_b (and the corresponding ΔG_b) and ΔH_b across 10,000 Monte Carlo simulated 1:1 binding reactions of fixed stoichiometry are shown in Fig. 2. The simulations performed here calculate the dilution effect on reactant concentrations occurring during the titration and modification of injection heats due to the volume displaced from the reaction cell following a discrete inject step model (or instantaneous injection model) (Tellinghuisen 2003). The simulated experiments are then fitted using an unweighted least-squares fit to the Wiseman equation (Wiseman et al. 1989) following typical experimental data analysis practice.

The simulations show that for all three true K_b values, the expected variation in the observed value is asymmetrically distributed. Furthermore, increasing the magnitude of K_b increases both the range of variation and the positive skew of the distribution of the K_b values (Table 1), i.e., there is an increasing tendency to observe more frequently higher than lower affinity values as K_b increases; thus, the higher K_b values become more likely to occur. For this particular experimental scenario, we find that there is an 8% probability of observing a K_b less than half, and a 9% chance of more than twice the true value when $K_b = 2 \times 10^5 \text{ M}^{-1}$. Those probabilities, respectively, increase to 18% and 24% when the assumed $K_b = 2 \times 10^7 \text{ M}^{-1}$ (and the probability of an observation more than 4 $\times$ the true value is 12%).

The simulated scenario in Fig. 2 has relatively low injection heats of approximately 1/3rd of the average for protein–ligand interactions in a recent large-scale study (Scheuermann and Brautigam 2015) and with high proportional measurement errors. These values have been chosen to clearly illustrate the various effects on the distributions of increasing affinity. Since the asymmetry increases with proportional measurement error (Tellinghuisen 2017), observed asymmetries would be smaller for most experimental cases where those errors are smaller. Even in this simulated experimental scenario, the resulting ΔG_b distribution has little asymmetry except at the highest affinity and conversely the enthalpy change is only appreciably asymmetric and has greater uncertainty at the lowest affinity. In particular, we note that in this experimental scenario even with its quite large proportional measurement ΔG_b is near normally distributed for reasonably optimal experiments

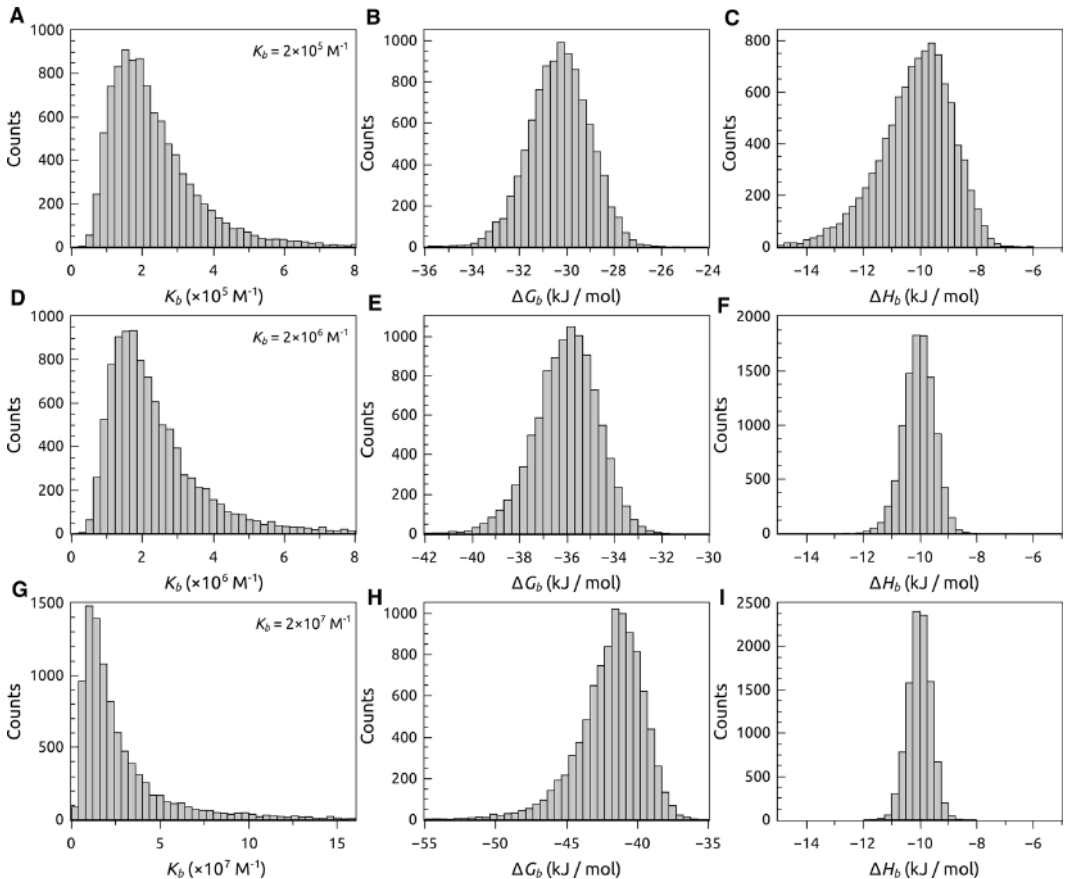


Fig. 2 Monte Carlo simulation modeling of variation in the fitted parameters of ITC experiments for three different binding affinities—**a–c** $2 \times 10^5 \text{ M}^{-1}$, **d–f** $2 \times 10^6 \text{ M}^{-1}$ and **g–i** $2 \times 10^7 \text{ M}^{-1}$. For each binding affinity, 10,000 simulated data sets were created. All simulations were for a 1:1 binding reaction with molar enthalpy change

$\Delta H_b = -10 \text{ kJ/mol}$, $20 \times 15 \text{ }\mu\text{L}$ injections of $200 \text{ }\mu\text{M}$ ligand into $20 \text{ }\mu\text{M}$ protein leading to a final ligand:protein ratio of 2.5:1. The simulations are modelling a VP-ITC instrument and a measurement error of standard deviation of $3 \text{ }\mu\text{J}$ per injection. Simulated datasets where a satisfactory fit could not be obtained are excluded

Table 1 Statistical parameters for simulated ITC data at three different binding affinities

Starting $K_b \text{ (M}^{-1}\text{)}$	K_b Skew	$K_b \text{ (M}^{-1}\text{)}$ CI_{95}	$\Delta G_b \text{ (kJ/mol)}$	ΔG_b Skew	$\Delta G_b \text{ (kJ/mol)}$ CI_{95}	ΔH_b Skew	$\Delta H_b \text{ (kJ/mol)}$ CI_{95}
2×10^5	1.3	$[0.78, 5.8] \times 10^5$	− 30.26	− 0.2	[− 32.9, − 27.9]	− 0.7	[− 13.2, − 8.0]
2×10^6	1.5	$[0.78, 6.5] \times 10^6$	− 35.97	− 0.4	[− 38.9, − 33.6]	− 0.2	[− 11.1, − 9.1]
2×10^7	23	$[0.51, 22] \times 10^7$	− 41.67	− 1.1	[− 47.6, − 38.3]	− 0.1	[− 10.8, − 9.3]

Data correspond to the distributions in Fig. 2. The 95% confidence intervals CI_{95} are determined directly from central 95% of values of the distributions

(Wiseman parameters C of 4 and 40, respectively for the first two simulations) for less optimal experiments (Wiseman parameter of 400 for the last simulation) skewness can

be readily seen. This can most likely be attributed to the fact that for a 20-point ITC experiment with the parameters used for the simulation essentially only 3–4 datapoints fall

into the transition region and any noisy data here will have a bigger influence leading to larger deviations and long tailing to higher K_b values. Also seen is an inverse relationship between the variations in measured affinity and enthalpy, this is expected due to the shape of the titration curve changing appreciably with binding affinity under the experimental conditions we have simulated with fixed concentration. At high C , almost all the ligand binds in the early injections of a titration giving several injection heats with high signal to noise and, thus, lowering the observed variation in the enthalpy measurement. An inverse relationship between the uncertainties in affinity and enthalpy is generally expected for ITC data. Under the experimental conditions that we have simulated with fixed concentrations, the shape of the titration curve changes appreciably with binding affinity.

Error estimations for ΔG_b and K_b using asymmetric profile likelihood Confidence Intervals

Because the calculation is built into most analysis software, most researchers (Salim and Feig 2009; Geschwindner et al. 2015; Renaud et al. 2016; Ladbury et al. 2010; Ciulli 2013; Smirnovienė et al. 2017; Myszkowski and Rich 2000; Patching 2014; Olaru and Bala 2015; Krimmer and Klebe 2015; Callies and Daranas 2016; Falconer 2016; Vega et al. 2016; Chaires 2008; Leavitt and Freire 2001) report the precision of the affinities between molecules using asymptotic-symmetric confidence intervals ($CI_{S,\alpha}$) calculated using the standard deviation for each parameter, as estimated from the sum of squared residuals, RSS, and the covariance matrix from the fitting analysis, and a chosen confidence level α (typically 68.3%) of a t -Student distribution. The experimental and simulated results have shown that this symmetric approach does not communicate and accurately preserve information about uncertainty, and that use of asymmetric F -statistics-associated (or profile likelihood) CI_{95} is both simple and preferable. How then can we estimate CI_{95} in practice, e.g., from single experiments? We will analyze this case here in detail, as it is the simplest one. In the case of several experiments, the reasoning will be basically the same—we suggest to perform a global fit with all individual datasets and get the global confidence intervals.

A possible approach is to fit for ΔG_b and use the capabilities of different software to obtain the asymptotic-symmetric error $CI_{S,95}$ for the Gibbs energy change. Asymmetric confidence intervals for K_b or K_d can then be obtained by simply transformation of the upper and lower limits of the ΔG_b confidence interval. As we have seen that for titration data ΔG_b is less affected by asymmetry, consequently this approach will be effective in many circumstances and is certainly better than current practice. However, it also fails in some circumstances, e.g., for high affinity interactions as shown in Case 2 or where measurement errors

are large [as extensively studied previously (Tellinghuisen 2017)]; so, prior knowledge of the theoretically expected behavior of ΔG_b uncertainty for the experiment scenario is required to apply it. Consequently, this approach cannot be generally recommended.

A second possible approach that is applicable in all circumstances is to use a ‘bootstrap’ procedure in which values of the residuals of the best fit are randomly selected and added to the fitted value at each titration point and the new set of data points so created refitted. Repeating this re-sampling procedure many times (> 1000) yields a distribution of values for each fitted parameter from which the confidence intervals can be obtained. This approach is applicable where the measurement variability is constant throughout the titration and produces $CI_{S,95}$ ranges that are only slightly larger than the true values. Unfortunately, the requisite re-sampling procedure is not widely available in commercial software.

We believe that the most practicable way to express the repeatability for K_b and ΔG_b is through calculation of the profile likelihood confidence intervals $CI_{P,\alpha}$ (at statistical significance level α) (Kemmer and Keller 2010; Krainer and Keller 2015; Krainer et al. 2012; Broecker et al. 2011; Johnson 1983), which can be determined by an extension of the typical fitting approach. Once the non-linear least squares regression analysis of N experimental points has been performed with a model with P parameters, the best estimates for the P parameters are obtained with an associated residual sum of squares RSS_0 . Ideally, the P parameters could be systematically varied to get a P -dimensional contour fulfilling the expression:

$$RSS = RSS_0 \left(1 + \frac{P}{N-P} F_{P,N-P}(\alpha) \right),$$

where $F_{n,m}$ is the Fisher–Snedecor distribution with $n=P$ and $m=N-P$ degrees of freedom, and the α is the chosen confidence level (Motulsky and Christopoulos 2004). Within that P -dimensional contour, the different possible sets of P parameters provide RSS values that are not statistically different (at a confidence level α) from RSS_0 . Then, by projecting the P -dimensional contour onto the different P axes, the confidence interval for each parameter can be determined. However, this procedure is not practical if there are more than two fitting parameters. Therefore, very often marginal confidence intervals are determined by varying just one parameter at a time (Kemmer and Keller 2010; Bates and Watts 2007). Thus, a given parameter, p , is selected and kept fixed at different values, while the RSS is minimized over the remaining free parameters, an RSS (p) curve (RSS as a function of p) with the resulting minimized RSS values (see Fig. 3, right). The two limiting values for the given parameter p defining its profile confidence interval $CI_{P,\alpha}$ will fulfill the expression:

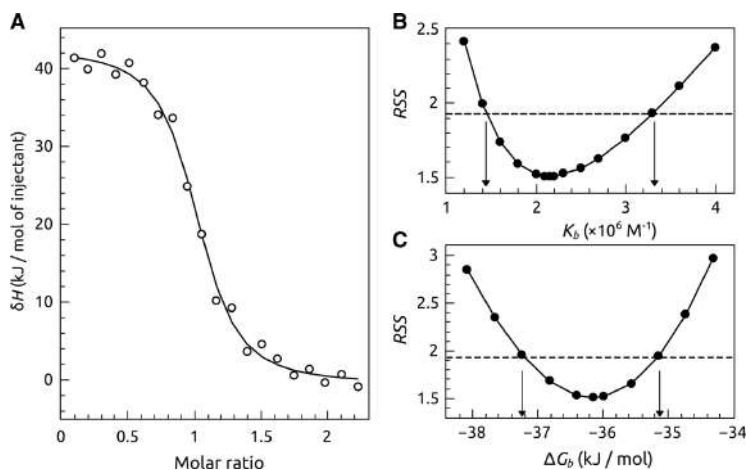


Fig. 3 **a** A simulated ITC titration curve corresponding to the $K_b = 2 \times 10^6 \text{ M}^{-1}$, $\Delta H = 41.8 \text{ kJ/mol}$, and 1:1 binding stoichiometry with random measurement error for each injection. The continuous line corresponds to the best fit considering either K_b or ΔG_b as fitting parameters. **b** RSS dependence on K_b . **c** RSS dependence on ΔG_b . Either K_b or ΔG_b are varied systematically (stepped through fixed values below and above their best estimate) and the remaining parameters are freely adjusted to re-minimize the RSS; then, that minimum RSS is plotted as a function of either K_b or ΔG_b . The minimum in

RSS (i.e., RSS_0) corresponds to the best fit (all parameters freely varying, including K_b or ΔG_b). The horizontal dotted line is the limit reference value for RSS with 95% confidence, equal to $\text{RSS}_0 (1 + (1/(N-P))F_{1,N-P}(0.95))$. The intercepts between the RSS curve and the reference RSS value (indicated by arrows) provide the limits for the confidence intervals. The way to get these CIs with the Origin software is shown in Supplementary Material as an example (2. Quick calculation of profile likelihood asymmetric confidence intervals $\text{CI}_{\text{P},95}$.)

$$\text{RSS}(p) = \text{RSS}_0 \left(1 + \frac{1}{N-P} F_{1,N-P}(\alpha) \right)$$

These two limiting values define the interval in which the parameter p provides RSS values that are not statistically different (at a confidence level α) from RSS_0 . The process can be repeated for each of the other parameters, and all P marginal confidence intervals estimated.

In general, RSS is not a symmetric function of the studied parameter with respect to the minimum value RSS_0 , and the two limiting values satisfying the previous equation define an asymmetric confidence interval for each parameter. The asymmetry degree and the size of the confidence interval depends on the nature of the parameter considered and the sensitivity of RSS to that parameter.

The procedure is illustrated by analyzing an isothermal calorimetric titration simulated (with $K_b = 2 \times 10^6 \text{ M}^{-1}$, $\Delta H = 41.8 \text{ kJ/mol}$, and $n = 1$) with a chosen noise level (Fig. 3) and shown in detail for the Origin software as an example in Supplementary Material (2. Quick calculation of profile likelihood asymmetric confidence intervals $\text{CI}_{\text{P},95}$).

The standard output for the best fit to the titration in Fig. 3 has the following estimated parameters: $K_b = (2.2 \pm 0.4) \times 10^6 \text{ M}^{-1}$, $\Delta H = 43.3 \pm 1.2 \text{ kJ/mol}$, and $n = 0.99 \pm 0.01$. If the Gibbs energy of

interaction is considered as a fitting parameter instead of K_b the best fit provides the following estimated parameters: $\Delta G = -36.2 \pm 0.4 \text{ kJ/mol}$, $\Delta H = 43.3 \pm 1.3 \text{ kJ/mol}$, and $n = 0.99 \pm 0.01$.

These uncertainties correspond to the typical standard errors $\text{CI}_{\text{S},68.3}$ for the fitting parameters generated from the covariance matrix. The asymptotic-symmetric confidence intervals for a confidence level of $\alpha = 95\%$ $\text{CI}_{\text{S},95}$ are calculated as a multiple of these values determined by the t -Student distribution. The asymptotic-symmetric and asymmetric profile likelihood 95% confidence intervals are shown in Table 2.

As expected from the previous results, the uncertainty in K_b is revealed to be asymmetric by the profile likelihood confidence interval $\text{CI}_{\text{P},95}$. Also as expected from the Monte Carlo simulations (in Case 2) at this binding affinity, the $\text{CI}_{\text{P},95}$ for ΔG_b is almost symmetric. Indeed, as discussed above in this case of an actual symmetric uncertainty in ΔG_b , both the asymptotic-symmetric and profile likelihood methods give the same confidence intervals for both parameters provided that ΔG_b is fitted and then transformed to K_b . However, only the profile likelihood approach produces the same results if K_b is fitted. When using $\text{CI}_{\text{P},95}$, there is a perfect correspondence between estimated values and uncertainty interval limits for K_b and ΔG_b ; thus, no matter which parameter is employed as a fitting parameter, the

Table 2 Best fit parameters and uncertainties determined as asymptotic-symmetric $CI_{S,95}$ and profile likelihood $CI_{P,95}$ confidence intervals

	K_b (2.2) ^a ($\times 10^6$ M ⁻¹)	ΔG_b (− 36.2) ^b (kJ/mol)	ΔH_b (kJ/mol)	N
$CI_{S,95}$ ^a	[1.29, 3.03]	[− 37.0, − 34.9]	[40.8, 45.9]	[0.96, 1.02]
$CI_{S,95}$ ^b	[1.44, 3.30]	[− 37.2, − 35.1]	[40.7, 46.0]	[0.96, 1.02]
$CI_{P,95}$ ^a	[1.44, 3.30]	[− 37.2, − 35.1]	[40.9, 46.1]	[0.96, 1.03]
$CI_{P,95}$ ^b	[1.44, 3.30]	[− 37.2, − 35.1]	[40.9, 46.1]	[0.96, 1.03]

The directly fitted values are shown in bold

^aFitting of K_b and calculation of its $CI_{P,95}$ with subsequent transformation to ΔG_b values

^bFitting of ΔG_b and subsequent calculation of K_b

other parameter and its confidence interval can be readily calculated through their mathematical relationship. However, when using $CI_{S,95}$, the estimated values for K_b and ΔG_b are in correspondence, but the uncertainty intervals are not. Furthermore, in some experimental scenarios the upper limit of the affinity may be undefined, e.g., for a very steep titration with few data points contributing to the fit. The profile likelihood confidence intervals will only produce a lower limit in such cases, a highly advantageous presentation over the use of symmetrical intervals. The profile likelihood approach is, thus, more robust in estimating uncertainties and, consequently, the recommended approach.

Conclusions

Although the inherent asymmetry that appears as a result of the measurement error propagating through the data analysis process has been illustrated here with only ITC simulations, they are true for all binding experiments (i.e., to an extent that depends on the nature of the measurement errors and the equations used to analyze each particular method). The presented concepts and procedures for dealing with this asymmetry to determine accurate uncertainties can be extended to any experimental technique used to determine binding affinities or any related quantities (see the distribution of the kinetic rate constant in the supplementary material). The calculation of the confidence intervals can be performed manually following the step-by-step procedure explained above. These can be conveniently carried out in Excel (Kemmer and Keller 2010), but fortunately, commercially available software packages (e.g., Origin, GraphPad, Sedphat) also provide profile likelihood confidence intervals, for all fitting parameters at any confidence level, in a user friendly manner, without the need for complicated calculations. Once the best fit is achieved (with all parameters freely varying), the profile likelihood confidence intervals (at a certain confidence level) for all fitting parameters are readily calculated within just a single step as explained in the Supplementary materials.

Considering the variations in the shape and inherent difference of the uncertainty distribution of the two

thermodynamically related parameters K_b and ΔG_b , reporting a symmetric error appears not to be the scientifically correct way. Reporting 95% confidence intervals removes the artificial restriction of symmetry and enables more accurate reporting of uncertainty. The statistically sound construction of the profile likelihood confidence intervals, and the perfect agreement shown of CIs obtained when the fitting of ITC data was performed for K_b or ΔG_b and ΔH_b , shows that profile likelihood confidence intervals can be used to report the repeatability of K_b and ΔH_b as retrieved from ITC, or for binding affinities determined by any other method.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00249-021-01518-4>.

Acknowledgements This research was funded by grant no. S-LLT-20-2 from the Research Council of Lithuania (DM). MB and MJM acknowledge Fundação para a Ciência e Tecnologia (FCT), Portugal, for the financial support to Projects UIDB/00081/2020 and UIDB/00313/2020, respectively. The authors also acknowledge the COST action ARBRE-MOBIEU CA15126 supported by COST (European Cooperation in Science and Technology).

References

- Bates DM, Watts DG (2007) Nonlinear regression analysis and its applications. Wiley, Hoboken
- Brautigam CA, Zhao H, Vargas C, Keller S, Schuck P (2016) Integration and global analysis of isothermal titration calorimetry data for studying macromolecular interactions. *Nat Protoc* 11:882–894
- Broecker J, Vargas C, Keller S (2011) Revisiting the optimal cvalue for isothermal titration calorimetry. *Anal Biochem* 418:307–309
- Buurma NJ, Haq I (2007) Advances in the analysis of isothermal titration calorimetry data for ligand–DNA interactions. *Methods* 42:162–172
- Callies O, Daranas AH (2016) Application of isothermal titration calorimetry as a tool to study natural product interactions. *Nat Prod Rep* 33:881–904
- Chaires JB (2008) Calorimetry and thermodynamics in drug design. *Annu Rev Biophys* 37:135–151
- Cheng RKY et al (2017) Structural insight into allosteric modulation of protease-activated receptor 2. *Nature* 545:112–115
- Cimpmperman P, Matulis D (2011) Protein thermal denaturation measurements via a fluorescent dye. In: Podjarny A, Dejaegere AP, Kieffer B (eds) RSC biomolecular sciences. Royal Society of Chemistry, pp 247–274

- Cimpmperman P et al (2008) A quantitative model of thermal stabilization and destabilization of proteins by ligands. *Biophys J* 95:3222–3231
- Ciulli A (2013) Biophysical screening for the discovery of small-molecule ligands. *Methods Mol Biol Clifton NJ* 1008:357–388
- Dam TK, Torres M, Brewer CF, Casadevall A (2008) Isothermal titration calorimetry reveals differential binding thermodynamics of variable region-identical antibodies differing in constant region for a univalent ligand. *J Biol Chem* 283:31366–31370
- Dullweber F, Stubbs MT, Musil D, Stürzebecher J, Klebe G (2001) Factorising ligand affinity: a combined thermodynamic and crystallographic study of trypsin and thrombin inhibition†. *J Mol Biol* 313:593–614
- Falconer RJ (2016) Applications of isothermal titration calorimetry—the research and technical developments from 2011 to 2015: review of Isothermal Titration Calorimetry from 2011 to 2015. *J Mol Recognit* 29:504–515
- Gaspari R et al (2016) Kinetic and structural insights into the mechanism of binding of sulfonamides to human carbonic anhydrase by computational and experimental studies. *J Med Chem* 59:4245–4256
- Geschwindner S, Ulander J, Johansson P (2015) Ligand binding thermodynamics in drug discovery: still a hot tip? *J Med Chem* 58:6321–6335
- Guan R, Tyler PC, Evans GB, Schramm VL (2013) Thermodynamic analysis of transition-state features in picomolar inhibitors of human 5′-methylthioadenosine phosphorylase. *Biochemistry* 52:8313–8322
- Hörtnér SR et al (2007) Potent inhibitors of tRNA-guanine transglycosylase, an enzyme linked to the pathogenicity of the *Shigella* Bacterium: charge-assisted hydrogen bonding. *Angew Chem Int Ed* 46:8266–8269
- Huschmann FU et al (2016) Structures of endotheiapepsin-fragment complexes from crystallographic fragment screening using a novel, diverse and affordable 96-compound fragment library. *Acta Crystallogr Sect F Struct Biol Commun* 72:346–355
- Jarmoskaite I, AlSadhan I, Vaidyanathan PP, Herschlag D (2020) How to measure and evaluate binding affinities. *Elife*. <https://doi.org/10.7554/eLife.57264>
- Johnson ML (1983) Evaluation and propagation of confidence intervals in nonlinear, asymmetrical variance spaces. Analysis of ligand-binding data. *Biophys J* 44:101–106
- Kemmer G, Keller S (2010) Nonlinear least-squares data fitting in Excel spreadsheets. *Nat Protoc* 5:267–281
- Krainer G, Keller S (2015) Single-experiment displacement assay for quantifying high-affinity binding by isothermal titration calorimetry. *Methods* 76:116–123
- Krainer G, Broecker J, Vargas C, Fanghänel J, Keller S (2012) Quantifying high-affinity binding of hydrophobic ligands by isothermal titration calorimetry. *Anal Chem* 84:10715–10722
- Krimmer SG, Klebe G (2015) Thermodynamics of protein–ligand interactions as a reference for computational analysis: how to assess accuracy, reliability and relevance of experimental data. *J Comput Aided Mol Des* 29:867–883
- Krishnamurthy VM et al (2007) Thermodynamic parameters for the association of fluorinated benzenesulfonamides with bovine carbonic anhydrase II. *Chem Asian J* 2:94–105
- Ladbury JE, Doyle ML (2004) *Biocalorimetry 2: applications of calorimetry in the biological sciences*. Wiley, Hoboken
- Ladbury JE, Klebe G, Freire E (2010) Adding calorimetric data to decision making in lead discovery: a hot tip. *Nat Rev Drug Discov* 9:23–27
- Lafont V et al (2007) Compensating enthalpic and entropic changes hinder binding affinity optimization. *Chem Biol Drug Des* 69:413–422
- Leavitt S, Freire E (2001) Direct measurement of protein binding energetics by isothermal titration calorimetry. *Curr Opin Struct Biol* 11:560–566
- McDonnell PA et al (2009) Assessing compound binding to the Eg5 motor domain using a thermal shift assay. *Anal Biochem* 392:59–69
- Motulsky H, Christopoulos A (2004) *Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting*. Oxford University Press, Oxford
- Myszka DG, Rich RL (2000) Implementing surface plasmon resonance biosensors in drug discovery. *Pharm Sci Technol Today* 3:310–317
- Olaru A, Bala C, Jaffrezic-Renault N, Aboul-Enein HY (2015) Surface plasmon resonance (SPR) biosensors in pharmaceutical analysis. *Crit Rev Anal Chem* 45:97–105
- Pantoliano MW et al (2001) High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J Biomol Screen* 6:429–440
- Patching SG (2014) Surface plasmon resonance spectroscopy for characterisation of membrane protein–ligand interactions and its potential for drug discovery. *Biochim Biophys Acta Biomembr* 1838:43–55
- Pierce MM, Raman CS, Nall BT (1999) Isothermal titration calorimetry of protein–protein interactions. *Methods* 19:213–221
- Pulido NO et al (2015) On the molecular basis of the high affinity binding of basic amino acids to LAOBP, a periplasmic binding protein from *Salmonella typhimurium*: ENERGETIC BASIS OF LAOBP'S LIGAND RECOGNITION. *J Mol Recognit* 28:108–116
- Rechlin C et al (2017) Price for opening the transient specificity pocket in human aldose reductase upon ligand binding: structural, thermodynamic, kinetic, and computational analysis. *ACS Chem Biol*. <https://doi.org/10.1021/acscchembio.7b00062>
- Ren J et al (2014) Thermodynamic and structural characterization of halogen bonding in protein–ligand interactions: a case study of PDE5 and its inhibitors. *J Med Chem* 57:3588–3593
- Renaud J-P et al (2016) Biophysics in drug discovery: impact, challenges and opportunities. *Nat Rev Drug Discov* 15:679–698
- Rühmann E, Betz M, Heine A, Klebe G (2015) Fragment binding can be either more enthalpy-driven or entropy-driven: crystal structures and residual hydration patterns suggest why. *J Med Chem* 58:6960–6971
- Salim NN, Feig AL (2009) Isothermal titration calorimetry of RNA. *Methods* 47:198–205
- Scheuermann TH, Brautigam CA (2015) High-precision, automated integration of multiple isothermal titration calorimetric thermograms: new features of NITPIC. *Methods San Diego Calif* 76:87–98
- Schnapp G, Klein T, Hoevels Y, Bakker RA, Nar H (2016) Comparative analysis of binding kinetics and thermodynamics of dipeptidyl peptidase-4 inhibitors and their relationship to structure. *J Med Chem* 59:7466–7477
- Smirnovienė J, Smirnovas V, Matulis D (2017) Picomolar inhibitors of carbonic anhydrase: importance of inhibition and binding assays. *Anal Biochem* 522:61–72
- Tellinghuisen J (2003) A study of statistical error in isothermal titration calorimetry. *Anal Biochem* 321:79–88
- Tellinghuisen J (2005) Statistical error in isothermal titration calorimetry: variance function estimation from generalized least squares. *Anal Biochem* 343:106–115
- Tellinghuisen J (2017) Can you trust the parametric standard errors in nonlinear least squares? Yes, with provisos. *Biochim Biophys Acta BBA Gen Subj*. <https://doi.org/10.1016/j.bbagen.2017.12.016>
- Tellinghuisen J (2018) Critique of methods for estimating heats in isothermal titration calorimetry. *Anal Biochem* 563:79–86

- Typas A, Sourjik V (2015) Bacterial protein networks: properties and functions. *Nat Rev Microbiol* 13:559–572
- Vega S, Abian O, Velazquez-Campoy A (2016) On the link between conformational changes, ligand binding and heat capacity. *Biochim Biophys Acta Gen Sub* 1860:868–878
- Wells RD et al (1980) DNA structure and gene regulation. *Prog Nucleic Acid Res Mol Biol* 24:167–267
- Wiseman T, Williston S, Brandts JF, Lin LN (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem* 179:131–137
- Yanchunas J et al (2005) Molecular basis for increased susceptibility of isolates with atazanavir resistance-conferring substitution I50L to other protease inhibitors. *Antimicrob Agents Chemother* 49:3825–3832

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Vaida Paketurytė¹  · Vytautas Petrauskas¹  · Asta Zubrienė¹  · Olga Abian^{2,3,4,5,6}  · Margarida Bastos⁷  · Wen-Yih Chen⁸  · Maria João Moreno⁹  · Georg Krainer¹⁰  · Vaida Linkuvienė¹  · Arthur Sedivy¹¹  · Adrian Velazquez-Campoy^{2,3,4,5,12}  · Mark A. Williams¹³  · Daumantas Matulis¹ 

¹ Department of Biothermodynamics and Drug Design, Institute of Biotechnology, Life Sciences Center, Vilnius University, Saulėtekio 7, 10257 Vilnius, Lithuania

² Institute of Biocomputation and Physics of Complex Systems (BIFI), Joint Units IQFR-CSIC-BIFI, and GBsC-CSIC-BIFI, Universidad de Zaragoza, Zaragoza, Spain

³ Aragon Institute for Health Research (IIS Aragon), Zaragoza, Spain

⁴ Department of Biochemistry and Molecular and Cell Biology, Universidad de Zaragoza, Zaragoza, Spain

⁵ Centro de Investigación Biomédica en Red en el Área Temática de Enfermedades Hepáticas y Digestivas (CIBERehd), Barcelona, Spain

⁶ Instituto Aragonés de Ciencias de la Salud (IACS), Zaragoza, Spain

⁷ Department of Chemistry and Biochemistry, Faculty of Sciences, CIQ-UP, University of Porto, Porto, Portugal

⁸ Department of Chemical and Materials Engineering, National Central University, Taoyüan, Taiwan

⁹ Chemistry Department, Faculty of Sciences and Technology, Coimbra Chemistry Center, University of Coimbra, 3004-535 Coimbra, Portugal

¹⁰ Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge, UK

¹¹ Protein Technologies, Vienna Biocenter Core Facilities GmbH, Vienna, Austria

¹² Fundacion ARAID, Government of Aragon, Zaragoza, Spain

¹³ Department of Biological Sciences, Institute for Structural and Molecular Biology, Birkbeck, University of London, London, UK

PUBLICATIONS NOT INCLUDED IN THE DISSERTATION

During Doctoral Studies (2019–2026)

1. Žvirblis, M.; **Paketurytė-Latvė, V.**; Gedgaudas, M.; Lingė, D.; Urniežius, E.; Kojis, T.; Stančaitis, L.; Zakšauskas, A.; Mickevičiūtė, A.; Jachno, J.; Juozapaitienė, V.; Dudutienė, V.; Zubrienė, A.; Matulis, D. **Design of ortho/meta-substituted selective carbonic anhydrase IX inhibitors.** *Chemija* 2026, 37 (1), 29–42. DOI: 10.6001/chemija.2026.37.1.4
2. Zubrienė, A.; Kurtenoka, M.; **Paketurytė-Latvė, V.**; Leitans, J.; Manakova, E.; Žvirblis, M.; Kazaks, A.; Eimonta, V.; Tars, K.; Gražulis, S.; Petrauskas, V.; Matulienė, J.; Dudutienė, V.; Shubin, K.; Matulis, D. **Achieving femtomolar affinities in structure-based drug design.** *Eur. Biophys. J.* 2026, 55 (1), 55–62. DOI: 10.1007/s00249-025-01812-5
3. Vaškevičius, A.; Žvirblis, M.; Kurtenoka, M.; Leitans, J.; Manakova, E.; **Paketurytė-Latvė, V.**; Kvietkauskaitė, A.; Kazaks, A.; Eimonta, V.; Čerepenkaitė, K.; Kazokaitė-Adomaitienė, J.; Mickevičiūtė, A.; Juozapaitienė, V.; Tars, K.; Gražulis, S.; Matulienė, J.; Dudutienė, V.; Shubin, K.; Matulis, D.; Zubrienė, A. **Di-Meta-Substituted Fluorinated Benzenesulfonamides as Potent and Selective Anticancer Inhibitors of Carbonic Anhydrase IX and XII.** *J. Med. Chem.* 2025. DOI: 10.1021/acs.jmedchem.5c01142
4. Vaškevičius, A.; Trumpickaitė, G.; Paražanovič, E.; Manakova, E.; Mickevičiūtė, A.; Gedgaudas, M.; Kojis, T.; **Paketurytė-Latvė, V.**; Smirnov, A.; Baranauskienė, L.; Gražulis, S.; Zubrienė, A.; Dudutienė, V.; Matulis, D. **Design of Rigid Compounds to Enhance Selectivity for Carbonic Anhydrase IX.** *Chem. Weinheim. Bergstr. Ger.* 2025, 31 (18), e202404409. DOI: 10.1002/chem.202404409

5. Lingè, D.; Gedgaudas, M.; Merkys, A.; Petrauskas, V.; Vaitkus, A.; Grybauskas, A.; **Paketurytė, V.**; Zubrienė, A.; Zakšauskas, A.; Mickėvičiūtė, A.; Smirnovienė, J.; Baranauskienė, L.; Čapkauskaitė, E.; Dudutienė, V.; Urniežius, E.; Konovalovas, A.; Kazlauskas, E.; Shubin, K.; Schiöth, H. B.; Chen, W.-Y.; Ladbury, J. E.; Gražulis, S.; Matulis, D. **PLBD: Protein–Ligand Binding Database of Thermodynamic and Kinetic Intrinsic Parameters**. *Database* 2023, 2023, baad040. DOI: 10.1093/database/baad040
6. Zakšauskas, A.; Čapkauskaitė, E.; **Paketurytė-Latvė, V.**; Smirnov, A.; Leitans, J.; Kazaks, A.; Dvinskis, E.; Stančaitis, L.; Mickevičiūtė, A.; Jachno, J.; Jezepčikas, L.; Linkuvienė, V.; Sakalauskas, A.; Manakova, E.; Gražulis, S.; Matulienė, J.; Tars, K.; Matulis, D. **Methyl 2-Halo-4-Substituted-5-Sulfamoyl-Benzoates as High Affinity and Selective Inhibitors of Carbonic Anhydrase IX**. *Int. J. Mol. Sci.* 2022, 23 (1), 130. DOI: 10.3390/ijms23010130
7. Matulienė, J.; Žvinys, G.; Petrauskas, V.; Kvietkauskaitė, A.; Zakšauskas, A.; Shubin, K.; Zubrienė, A.; Baranauskienė, L.; Kačėnauskaitė, L.; Kopanchuk, S.; Veiksina, S.; **Paketurytė-Latvė, V.**; Smirnovienė, J.; Juozapaitienė, V.; Mickevičiūtė, A.; Michailovienė, V.; Jachno, J.; Stravinskienė, D.; Sližienė, A.; Petrošiūtė, A.; Becker, H. M.; Kazokaitė-Adomaitienė, J.; Yaromina, A.; Čapkauskaitė, E.; Rinken, A.; Dudutienė, V.; Dubois, L. J.; Matulis, D. **Picomolar Fluorescent Probes for Compound Affinity Determination to Carbonic Anhydrase IX Expressed in Live Cancer Cells**. *Sci. Rep.* 2022, 12 (1), 17644. DOI: 10.1038/s41598-022-22436-1
8. Balandis, B.; Šimkūnas, T.; **Paketurytė-Latvė, V.**; Michailovienė, V.; Mickevičiūtė, A.; Manakova, E.; Gražulis, S.; Belyakov, S.; Kairys, V.; Mickevičius, V.; Zubrienė, A.; Matulis, D. **Beta and Gamma Amino Acid-Substituted Benzenesulfonamides as Inhibitors of Human Carbonic Anhydrases**. *Pharmaceuticals* 2022, 15 (4), 477. DOI: 10.3390/ph15040477
9. Baronas, D.; Dudutienė, V.; **Paketurytė, V.**; Kairys, V.; Smirnov, A.; Juozapaitienė, V.; Vaškevičius, A.; Manakova, E.; Gražulis, S.; Zubrienė, A.; Matulis, D. **Structure and Mechanism of Secondary Sulfonamide Binding to Carbonic Anhydrases**. *Eur. Biophys. J.* 2021, 50 (7), 993–1011. DOI: 10.1007/s00249-021-01561-1
10. Zakšauskas, A.; Čapkauskaitė, E.; Jezepčikas, L.; Linkuvienė, V.; **Paketurytė, V.**; Smirnov, A.; Leitans, J.; Kazaks, A.; Dvinskis, E.; Manakova, E.; Gražulis, S.; Tars, K.; Matulis, D. **Halogenated and Di-Substituted Benzenesulfonamides as Selective Inhibitors of Carbonic Anhydrase Isoforms**. *Eur. J. Med. Chem.* 2020, 185, 111825.

DOI: 10.1016/j.ejmech.2019.111825

Publications Prior to Doctoral Studies (2016–2019)

1. Vaškevičienė, I.; **Paketurytė, V.**; Pajanok, N.; Žukauskas, Š.; Sapijanskaitė, B.; Kantiminienė, K.; Mickevičius, V.; Zubrienė, A.; Matulis, D. **Pyrrolidinone-bearing methylated and halogenated benzenesulfonamides as inhibitors of carbonic anhydrases.** *Bioorg. Med. Chem.* 2019, 27, 322–337. DOI: 10.1016/j.bmc.2018.12.011
2. **Paketurytė, V.**; Linkuvienė, V.; Krainer, G.; Chen, W-Y.; Matulis, D. **Repeatability, precision, and accuracy of the enthalpies and Gibbs energies of a protein-ligand binding reaction measured by isothermal titration calorimetry.** *Eur. Biophys. J.* 2018, 48 (2), 139–152. DOI: 10.1007/s00249-018-1341-z
3. Čapkauskaitė, E.; Zubrienė, A.; **Paketurytė, V.**; Timm, D.; Tumkevičius, S.; Matulis, D. **Thiazole-substituted benzenesulfonamides as inhibitors of 12 human carbonic anhydrases.** *Bioorg. Chem.* 2018, 77, 534–541. DOI: 10.1016/j.bioorg.2018.02.004
4. Čapkauskaitė, E.; Zakšauskas, A.; Ruibys, V.; Linkuvienė, V.; **Paketurytė, V.**; Gedgaudas, M.; Kairys, V.; Matulis, D. **Benzimidazole design, synthesis, and docking to build selective carbonic anhydrase VA inhibitors.** *Bioorg. Med. Chem.* 2018, 26 (3), 675–687. DOI: 10.1016/j.bmc.2017.12.035
5. Vaškevičienė, I.; **Paketurytė, V.**; Zubrienė, A.; Kantiminienė, K.; Mickevičius, V.; Matulis, D. **N-Sulfamoylphenyl- and N-Sulfamoylphenyl-N-thiazolyl- β -alanines and their derivatives as inhibitors of human carbonic anhydrases.** *Bioorg. Chem.* 2017, 75, 16–29. DOI: 10.1016/j.bioorg.2017.08.017
6. Paliulis, E.; **Paketurytė, V.**; Matulis, D. **Protease, Amylase and Lactase Enzyme Stability in Gastroval® Capsules after Incubation at Acidic pH and Elevated Temperature.** *Biochem. Physiol.: Open Access.* 2016. DOI: 10.4172/2168-9652.1000211

Book Chapters

1. **Paketytė, V.; Zubrienė, A.; Ladbury, J. E.; Matulis, D. Intrinsic Thermodynamics of Protein-Ligand Binding by Isothermal Titration Calorimetry as Aid to Drug Design.** *Microcalorimetry of Biological Molecules*. New York: Springer. Methods Mol. Biol. 1964, 61–74. 2019. DOI: 10.1007/978-1-4939-9179-2_5
2. **Paketytė, V.; Zubrienė, A.; Chen, W.-Y.; Keller, S.; Bastos, M.; Todd, M. J.; Ladbury, J. E.; Matulis, D. Inhibitor Binding to Carbonic Anhydrases by Isothermal Titration Calorimetry.** *Carbonic Anhydrase as Drug Target: Thermodynamics and Structure of Inhibitor Binding*. Cham: Springer International Publishing. 2019, 79–95. DOI: 10.1007/978-3-030-12780-0_6

Scientific Events and Courses Attended by the Author

This section includes only scientific events attended by the author and presentations personally delivered by the author; presentations given solely by co-authors are not listed.

During Doctoral Studies (2019–2026)

Scientific Conferences

1. **Paketurytė-Latvė, V.**; Smirnov, A.; Dudutienė, V.; Zubrienė, A.; Matulis, D. "Understanding Carbonic Anhydrase Inhibition: Structural and Thermodynamic Insights". *4th MOSBRI Scientific Conference*. Paris, France. 2025 06 10–13.

–Oral presentation

2. **Paketurytė-Latvė, V.**; Petrauskas, V.; Zubrienė, A.; Matulis, D. "Constant Uncertainty in Protein-Ligand Binding: Asymmetric Confidence Intervals vs Standard Errors". *XXI Conference of the International Society for Biological Calorimetry ISBC 2022*. Vilnius, Lithuania. 2022 06 08–10.

–Poster presentation

3. **Paketurytė, V.**; Smirnov, A.; Manakova, E.; Gražulis, S.; Matulis, D. "Crystallography of Human Carbonic Anhydrases with Inhibitors for Drug Design". *PDB50*. Online. 2021 05 04–05.

–Poster and video presentation

4. **Paketurytė, V.**; Smirnov, A.; Jankūnaitė, A.; Zakšauskas, A.; Čapkauskaitė, E.; Matulis, D. "Correlations between Thermodynamics and Structure of Carbonic Anhydrase-Inhibitor Binding". *64th Annual*

Meeting of the Biophysical Society. San Diego, California, USA. 2020 02 15–19. (Funding: Research Council of Lithuania support for doctoral students' academic visits)

–Poster presentation

5. **Pakuturytė, V.**; Smirnov, A.; Jankūnaitė, A.; Zakšauskas, A.; Čapkauskaitė, E.; Matulis, D. "Thermodynamics-Structure Correlations of Carbonic Anhydrase-Inhibitor Binding". *FEBS INSTRUCT MOBIEU Practical Course: HyThabio*. Grenoble, France. 2020 01 26–31.

–Poster presentation

Courses

1. Training school (2026 02 22–03 29) "*HERCULES SCHOOL: Neutrons & Synchrotron Radiation for Science*"
2. Practical course, Grenoble, France (2020 03 26–31) *FEBS INSTRUCT MOBIEU Practical course "HyThabio": Hydrodynamic and thermodynamic analysis of biological macromolecules and their interactions*.
3. Practical course, Online (2021 02 01–03 10) "*Introduction to Molecular Modeling in Drug Discovery*". Organizer - Schrödinger, Inc.
4. Practical course, Online (2021 05 24–27) "*Instruct-ERIC course: Biophysics, structural mass spectrometry, and X-ray techniques*". Organizer - Institute of Biotechnology of the Czech Academy of Sciences.
5. Practical course, Online (2021 08 30–10 31) "*Pharmaceutical Bioinformatics*". Organizer - Uppsala University.

Prior to Doctoral Studies (2016–2019)

Scientific Conferences

1. **Pakuturytė, V.**; Matulis, D. "Precision of ITC Measurements For Target-Based Drug Design". *The 6th NovAliX Conference – Biophysics in Drug Discovery*. Cannes, France. 2019 03 20–22.

–Poster presentation

2. **Pakuturytė, V.**; Matulis, D. "Intrinsic Affinity of N-Substituted Benzenesulfonamides to Carbonic Anhydrases". *Molecular Biophysics – ABC of the puzzle of Life (The third COST-sponsored ARBRE-MOBIEU plenary meeting)*. Zagreb, Croatia. 2019 03 18–20.

–Poster presentation

3. **Paketyrytė, V.**; Matulis, D. "Intrinsic Affinity of N-Substituted Benzenesulfonamides to Carbonic Anhydrases". *The COINS*. Vilnius, Lithuania. 2019 02 26–28.

–Poster presentation

4. **Paketyrytė, V.**; Jankūnaitė, A.; Smirnov, A.; Zakšauskas, A.; Zubrienė, A.; Matulis, D. "Intrinsic Thermodynamics of o-S-Substituted Benzenesulfonamide Binding to Carbonic Anhydrases". *XVth Conference of the Lithuanian Biochemical Society*. Dubingiai, Lithuania. 2018 06 26.

–Oral presentation

5. **Paketyrytė, V.**; Čapkauskaitė, E.; Zubrienė, A.; Matulis, D. "Binding and inhibition of catalytically active human carbonic anhydrases by thiazole-substituted benzenesulfonamides". *Open Readings*. Vilnius, Lithuania. 2018 03 23.

–Poster presentation

6. **Paketyrytė, V.**; Vaškevičienė, I.; Zubrienė, A.; Mickevičius, V.; Matulis, D. "Thermodynamic Analysis of Benzenesulfonamides Bearing Oxopyrrolidine Moiety Binding to Human Carbonic Anhydrases". *The COINS*. Vilnius, Lithuania. 2018 03 02.

–Poster presentation

7. **Paketyrytė, V.**; Linkuvienė, V.; Matulis, D. "Repeatability of enthalpies and Gibbs energies of a protein–ligand binding reaction measured by ITC". *62nd Annual Meeting of the Biophysical Society*. San Francisco, California, USA. 2018 02 18.

–Oral presentation

8. **Paketyrytė, V.**; Vaškevičienė, I.; Zubrienė, A.; Mickevičius, V.; Matulis, D. "Selective inhibitors of human carbonic anhydrase isoform VB". *VitaScientia*. Vilnius, Lithuania. 2018 01 03.

–Poster presentation

9. **Paketyrytė, V.**; Paliulis, E.; Matulis, D. "Enzyme survivability at acidic pH and high temperature". *European Biotechnology Congress*. Riga, Latvia. 2016 05 04–07.

–Poster presentation

Internship

1. Internship. Malvern Instruments, Sweden (2017 04) - ITC precision measurements on PEAQ-ITC. *Advisor*: Dr. Natalia Markova

Vaida PAKETURYTĖ-LATVĖ

BIOCHEMIST | PROTEIN-LIGAND INTERACTIONS & DRUG DESIGN

vaida.paketuryte@gmc.vu.lt

ORCID: 0000-0003-0919-7826 | Web of Science: NDS-0906-2025 | ResearchGate: Vaida Paketuryte-Latve

SCIENTIFIC PROFILE

Biochemist specializing in protein-ligand interactions, binding thermodynamics, and structural studies of carbonic anhydrases. Research focuses on understanding molecular recognition through the integration of structural and thermodynamic approaches and their application to drug design.

RESEARCH AREAS

- Protein-ligand interactions
- Binding thermodynamics
- Molecular recognition
- Carbonic anhydrases (CA)
- Structure-based drug design

PATENTS

Co-inventor of carbonic anhydrase inhibitors synthesized on interconnecting linker chains:

- US2024000946A1
- EP4255886A1

HONORS & AWARDS

- President Kazys Grinius Memorial Scholarship for Academic Achievement (2018)
- Lithuanian Academy of Sciences Award for University Students' Research (2020)
- Vilnius University Academic Excellence Scholarship (2018, 2019, 2020)
- FEBS Youth Travel Fund Award (2020)
- COST MOBIEU Conference Award (2018) and STSM Fellowship (2017)

EDUCATION

2019–2026

(Leave: 2022–2024)

Doctoral studies in Biochemistry at Vilnius University, Life Sciences Center; Supervisor: Prof. Daumantas Matulis

2017–2019

M.Sc. in Biochemistry (*Magna Cum Laude*) at Vilnius University, Life Sciences Center

2013–2017

B.Sc. in Biochemistry at Vilnius University, Faculty of Chemistry

RESEARCH EXPERIENCE

2019–present

Junior Scientist at Department of Biothermodynamics and Drug Design, Institute of Biotechnology, Vilnius University

2017–2022

Researcher at ThermoPharma Baltic

2016–2019

Laboratory Assistant at Department of Biothermodynamics and Drug Design, Vilnius University

SUPERVISION

2022

Scientific supervisor of BSc thesis (T. Šimkūnas)

2020

Scientific consultant of BSc thesis (K. Gluščiukaitė)

RESEARCH PROJECTS

2017–2029

Participation in 7 research projects, from student research initiatives to national and EU-funded programs.

PROFESSIONAL MEMBERSHIPS

Biophysical Society; ARBRE-MOBIEU; Lithuanian Biochemical Society (co-organizer of a national conference in 2018);

SCIENTIFIC OUTCOMES

2

Book chapters
in scientific volumes

20

Publications
in international journals

14

Presentations
oral and poster

40+

PDB structures
of CA-ligand complexes

1500+

PLBD entries
protein-ligand binding data

Vaida Paketurytė-Latvė

Structural Recognition in Carbonic Anhydrase–Sulfonamide
Complexes: Isozyme and Ligand Effects on Binding Thermodynamics

Doctoral Dissertation
Natural Sciences
Biochemistry (N 004)

Struktūrinis atpažinimas karboanhidrazių–sulfonamidų kompleksuose:
izofermentų ir ligandų efektai jungimosi termodinamikoje

Daktaro disertacija
Gamtos mokslai
Biochemija (N 004)

Vilnius University Press
9 Saulėtekio Ave., Building III, LT-10222 Vilnius
Email: info@leidykla.vu.lt, www.leidykla.vu.lt
bookshop.vu.lt, journals.vu.lt
Print run of 15 copies