Thermodynamic, Kinetic and Crystallographic Investigations of Benzenesulfonamides as Ligands of Human Carbonic Anhydrase II

Dissertation

Zur Erlangung des Grades

Doktor der Naturwissenschaften

(Dr. rer. nat.)

dem Fachbereich Pharmazie

der Philipps-Universität Marburg

vorgelegt von

Steffen Glöckner

aus

Usingen

Marburg/Lahn 2020

Die vorliegende Dissertation wurde am Institut für Pharmazeutische Chemie der Philipps-Universität Marburg auf Anregung von Prof. Dr. Gerhard Klebe im Zeitraum von Juni 2016 bis Januar 2020 angefertigt.

Erstgutachter: Prof. Dr. Gerhard Klebe

Zweitgutachter: Prof. Dr. Andreas Heine

Eingereicht am: 03.01.2020

Tag der mündlichen Prüfung: 14.02.2020

Hochschulkennziffer: 1180

Table of Contents

Abbreviations

Ampicillin-Na Amipicillin sodium salt

bCAII bovine Carbonic Anhydrase II

BL beamline

BSA benzenesulfonamide

c *Wiseman* parameter

c_L ligand concentration

c_m mass concentration

c_n substance concentration

c_P protein concentration

CA Carbonic Anhydrase

CAI Carbonic Anhydrase inhibitor

CAII Carbonic Anhydrase II

CSD Cambridge Structural Database

cryo-EM cryogenic electron microscopy

χ electronegativity

χ^2 goodness of fit

$\vec{d}$ deviation vector

$d_{C_\delta-Ph}$ distance between Leu198–C_δ and the centroid of the ligand's phenyl ring

d_{hb} hydrogen-bond distance

DMSO dimethyl sulfoxide

DNA deoxyribonucleic acid

Δ difference in travelled distance

Abbreviations

ΔG difference in *Gibbs* free energy

ΔG$_{solv}$ solvation free energy

ΔH difference in enthalpy

ΔH$_{ion}$ enthalpy of ionization

ΔS difference in entropy

EEC enthalpy-entropy compensation

EGFR epidermal growth factor receptor

ETC equilibration time curve

F hydrophobicity-based intermediate

FBDD fragment-based drug design

FGFR1 fibroblast growth factor receptor 1

GoF goodness of fit

GST glutathione-S-transferase

h total heat

hCAII human Carbonic Anhydrase II

hCAII–1^{kin} kinetically favored hCAII–**1** complex

hCAII–1^{td} thermodynamically favored hCAII–**1** complex

hCAIX human Carbonic Anhydrase IX

HEPES 2-(4-(2-hydroxyethyl)piperazin-1-yl)ethane-1-sulfonic acid

HIV human immunodeficiency virus

HSP90 heat shock protein 90

HTS high-throughput screening

I$_{hkl}$ intensity

ITC isothermal titration calorimetry

k$_{-1}$ dissociation rate constant

k$_{-2}$ reverse isomerization rate constant

k_1 association rate constant

k_2 forward isomerization rate constant

k_3 final forward rate constant

K_a equilibrium association constant

K_d equilibrium dissociation constant

K_i inhibitory constant

k_{off} overall dissociation rate constant

k_{on} overall association rate constant

L ligand

LB lysogeny broth

MD molecular dynamics

MicroED microcrystal electron diffraction

MW molecular weight

MX macromolecular X-ray crystallography

$\vec{n}$ normal vector

n diffraction order

n_s stoichiometry of binding

N_i number of injections

NMR nuclear magnetic resonance

OD optical density

$\vec{p}$ directional vector of phenyl ring plane

P protein

P partition coefficient

PanDDA Pan-Dataset Density Analysis

PBS phosphate buffered saline

PCMB *para*-chloromercuribenzoic acid

PDB Protein Data Bank

PEG polyethylene glycol

pH *potentia hydrogenii*

pK_a acidity constant

PKA protein kinase A

P_Q power

PL protein-ligand complex

PL* high-affinity complex

$\vec{q}$ directional vector of phenyl ring plane

Q heat of injection

QSAR quantitative structure-activity relationship

R universal gas constant

RMSD root mean square deviation

RO3 Rule of Three

r_m molar ratio

RO5 Rule of Five

R_{free} free *R*-factor

r_{vdW} *van der Waals* radius

S advanced intermediate

SAR structure-activity relationship

SASA solvent accessible surface area

SAD single-wavelength anomalous dispersion

SPR surface plasmon resonance

SBDD structure-based drug design

SRE secondary recognition element

SESA solvent excluded surface area

t time

T temperature

t_i injection time

t_s delay between injections

TRIS Tris(hydroxymethyl)-aminomethane

τ_{eq} equilibration time

τ_{exp} experiment response time

τ_{ITC} instrument response time

τ_{rd} residence time

v_i injection rate

V_i injection volume

v_r stirring speed

V_t total volume

XFEL X-ray free-electron laser

XRD X-ray diffraction

X

List of Figures

List of Schemes

List of Tables

Objective

The aim of this work was the establishment of quantitative structure-activity relationships for series of systematically altered benzenesulfonamide ligands of human Carbonic Anhydrase II by macromolecular X-ray crystallography and kinetic isothermal titration calorimetry. Two series of alkyl and alkyloxy substituted compounds were to be investigated to investigate the influence of different alkyl chain geometries on the thermodynamics and kinetics of binding and furthermore compare the kinetic results of the alkyl series with kinetic data from surface plasmon resonance experiments from a previous study for the same compounds. Furthermore, the influence of varying fluorination patterns on the one hand and the influence of structurally similar *para*-substituents with different effects on ligand acidities on the other hand were to be exmained with respect to their influence on the thermodynamics and kinetics of ligand binding. Given the relative novelty of the concept of kinetic isothermal titration calorimetry, a measurement protocol that provides for the robust determination of both thermodynamic and kinetic parameters from one experiment, a novel titration protocol for microcalorimetric measurements that aims at simultaneously providing these at minimal input of protein and ligand material and high experimental efficiency was to be established and validated. Furthermore, an already known, highly potent inhibitor of human Carbonic Anhydrase II was to be examined to investigate the thermodynamic and kinetic driving forces for its high potency. Additionally, separated enantiomers of a hetereocyclic Carbonic Anhydrase inhibitor were investigated both crystallographically and by means of microcalorimetry in order to assess, whether the crystallographically usually not observed enantiomer shows inhibitory against human Carbonic

Anhydrase II and whether the compound racemizes under experimental conditions. A further study was aimed at the concentration-dependend occupancy of a Carbonic Anhydrase inhibitor in a second binding site, that could be potentially useful for the *in silico* predictions. A previously started proof-of-concept fragment-screening campaign was to be continued. Moreover, one aim of this thesis was to obtain neutron diffraction data for human Carbonic Anhydrase II in complex with the artificial sweetener saccharin in order to elucidate the protonation state of saccharin in the bound form.

Summary

The data presented here show that different geometries of *para*-alkyl substituents of a benzene-sulfonamide scaffold can have a significant influence on the thermodynamic and kinetic binding parameters and suggest, that the fine-tuning of the ligand geometry to match the geometry of the active site leads to a reduced dissociation rate of the protein-ligand complex and thus to a prolonged binding in the active site. In addition, it was found that the kinetic data extracted by isothermal titration calorimetry correlate with data from experiments with surface plasmon resonance for a subset of compounds, despite a difference of one order of magnitude that could not be explained so far. Furthermore, the investigation of fluorinated benzenesulfonamide ligands revealed complex structural-thermodynamic and structural-kinetic relationships, which could not be fully elucidated herein, but suggest that fluorination of a *meta*-position of the aromatic part of the benzenesulfonamide backbone favors high association as well as small dissociation rates. In addition, it has been shown that a higher degree of fluorination does not necessarily have an advantage for the thermodynamic or kinetic binding profile or affinity itself. Furthermore, structurally similar ligands with only one *para*-substituent but different acidities indicate that the association rate of complex formation benefits from increased acidity. A new measurement protocol for microcalorimetric measurements was analyzed for its precision and compared with other possible measurement protocols and showed that it is best suited for the reliable simultaneous determination of thermodynamic and kinetic data by microcalorimetry. The microcalorimetric investigation of an already known and supposedly very potent Carbonic Anhydrase inhibitor showed an unexpected two-step binding behavior, which, however, excludes

a 2 : 1 binding of inhibitor and protein. Time-dependent experiments suggest that the binding process between protein and ligand leads to a thermodynamically and kinetically preferred complex. The investigation of a chiral heterocyclic Carbonic Anhydrase inhibitor could not conclusively clarify, whether the crystallographically apparently non-binding enantiomer also shows inhibitory activity. However, microcalorimetric investigations allow the conclusion that racemization occurs under experimental conditions, which makes the clarification of the original question considerably more difficult. Soaking of Carbonic Anhydrase crystals in solutions of different concentrations of a ligand showed, that the occupancy of a second binding site was too weak to allow modeling of the ligand, which made it impossible to investigate the concentration dependence of the occupancy in this binding site. Concentration differences, however, resulted in different occupancies of the active site. A crystallographic model based on neutron diffraction data of human Carbonic Anhydrase II in complex with saccharin was not obtained. However, an attempt was made to assess the protonation state of saccharin in the active site of the enzyme by experimental phasing using high-resolution X-ray diffraction data. The continuation of a fragment screening showed the binding of further small molecules in Carbonic Anhydrase crystals. An analysis with the Pan-Dataset Density Analysis program suggests, that electron density maps are strongly dependent on the time of soaking of the crystals.

Zusammenfassung

Die hierin vorgestellten Daten zeigen, dass verschiedene Geometrien eines *para*-Alkylsubstituenten eines Benzolsulfonamidgerüsts einen deutlichen Einfluss auf die thermodynamischen und kinetischen Bindungsparameter haben können und deuten darauf hin, dass die Feinabstimmung der Komplementarität der Molekülgeometrie des Liganden auf die des aktiven Zentrums des Proteins zu einer verlangsamten Dissoziationskinetik des Protein-Ligand Komplexes und somit zu einer längeren Blockade des aktiven Zentrums führen kann. Zusätzlich wurde festgestellt, dass die mit isothermer Titrationskalorimetrie extrahierten kinetischen Daten mit Daten aus Experimenten mit Oberflächenplasmonenresonanz für eine Teilmenge von Verbindungen korrelieren, trotz eines bisher nicht zu erklärenden Unterschieds von einer Größenordnung. Darüber hinaus zeigte die Untersuchung von fluorierten Benzolsulfonamidliganden komplexe strukturthermodynamische und strukturkinetische Beziehungen, die hierin nicht vollständig aufgeklärt werden konnten, aber darauf hindeuten, dass die selektive Fluorierung einer *meta*-Position des aromatischen Rings hohe Assoziations- als auch niedrige Dissoziationsraten begünstigt. Außerdem konnte gezeigt werden, dass hochgradiger fluorierte Analoga nicht notwendigerweise einen Vorteil für das thermodynamische oder kinetische Bindungsprofil oder die Affinität selbst haben. Darüber hinaus deuten strukturell ähnliche Liganden mit nur einem *para*-Substituenten, aber unterschiedlichen Aciditäten darauf hin, dass eine erhöhte Acidität des Liganden erhöhte Assoziationsraten begünstigt. Ein neues Messprotokoll für mikrokalorimetrische Messungen wurde auf seine Präzision hin analysiert und mit anderen möglichen Messprotokollen verglichen und zeigte, dass es sich am besten für

die zuverlässige gleichzeitige Ermittlung von thermodynamischen und kinetischen Daten mittels Mikrokalorimetrie eignet. Die mikrokalorimetrische Untersuchung eines bereits bekannten und vermeintlich sehr potenten Carboanhydraseinhibitors zeigte ein unerwartetes zweistufiges Bindungsverhalten, das eine reine 2 : 1 Bindung von Inhibitor und Protein jedoch ausschließt. Zeitabhängige Experimente lassen den Schluss zu, dass der Bindungsvorgang zwischen Protein und Ligand zu einem thermodynamisch und einem kinetisch bevorzugten Komplex führt. Die Untersuchung eines chiralen heterocyclischen Carboanhydraseinhibitors konnte nicht abschließend klären, ob auch das kristallografisch scheinbar nicht bindende Enantiomer auch inhibitorische Wirkung zeigt. Mikrokalorimetrische Untersuchungen weisen jedoch darauf hin, dass unter den experimentellen Bedingungen Racemisierung Auftritt, was die Klärung der ursprünglichen Frage deutlich erschwert. Das Tränken von Carboanhydrasekristallen in Lösungen unterschiedlicher Konzentrationen eines Liganden zeigte, dass die Besetzung einer zweiten Bindestelle zu schwach war, um kristallographisch modelliert zu werden, was die angestrebte Untersuchung der Konzentrationsabhängigkeit der Besetzung unmöglich machte. Konzentrationsunterschiede sorgten jedoch für unterschiedliche Besetzungen des aktiven Zentrums. Ein kristallografisches Modell, das auf Neutronenbeugungsdaten der humanen Carboanhydrase II im Komplex mit Saccharin basiert, wurde nicht erhalten. Es wurde jedoch versucht, den Protonierungszustand von Saccharin an der aktiven Stelle des Enzyms mit Hilfe von hochaufgelösten Röntgenbeugungsdaten durch experimentelle Phasierung zu beurteilen. Die Fortsetzung eines Fragment-Screenings zeigte die Bindung weiterer Kleinmoleküle in Carboanhydrasekristallen. Eine Analyse mit dem Pan-Dataset Density Analysis Programm lässt darauf schließen, dass Elektronendichtekarten stark von der Zeit der Tränkung des Kristalls abhängig sind.

Part I

Structure-Activity Relationships of Diversely Substituted Benzenesulfonamides as Ligands of Human Carbonic Anhydrase II

1 Introduction

1.1 Rational Drug Design

Rational Drug Design aims at the improvement of binding affinity of a small-molecule with a given scaffold for a protein that is of interest for the treatment of a certain disease, based on the understanding of the assembly and disassembly of protein-ligand complexes.[1] In the early days of drug development, serendipity was the determining factor for success, as shown by the example of acetylsalicylic acid. It had been synthesized for the first time in 1879, but its mechanism of action was not elucidated until 1971.[2,3] Since then, the development of novel drugs has increasingly emancipated itself from serendipity-based discoveries and high-throughput screenings of millions of compounds toward the knowledge-based approach of structure-based drug design (SBDD).[1,4] A fundamental requirement for SBDD is the knowledge about the three-dimensional shape and physicochemical properties of a given receptor, ideally with information about the binding of organic molecules that act as probe heads and can serve as potential starting points for the design of a lead-like molecular scaffold.[1,5] Macromolecular X-ray crystallography (MX) is the method of choice to obtain these information, which can also be complemented by neutron diffraction data to obtain information about protonation states or sites of a protein that are amenable to proton exchange.[6] Cryogenic electron microscopy (cryo-EM) and microcrystal electron diffraction (MicroED) become increasingly important, the former one especially for large protein complexes that are difficult to crystallize, the latter as a powerful tool for the determination of both macromolecular structures as well as the compositions of small-

molecule crystals.[7–9] Furthermore, X-ray free-electron lasers (XFELs) enable the time-resolved observation of biomolecular recognition processes.[10] Also, nuclear magnetic resonance (NMR) spectroscopy is a powerful tool to follow a binding event in solution, whereas the previously mentioned methods investigate molecules in the crystalline state.[11,12] Structural data from the above methods are then combined with biophysical data about the binding between target and ligand molecules in so called quantitative structure-activity relationships (QSARs) to assess the potential of molecular scaffolds as new lead structures and furthermore elucidate the fashion in which affinity of a ligand for a target is conveyed.[13,14] *Fox et al.* differentiate at least nine individual processes, e.g. desolvation and conformational changes, that are necessary for the overall process of biomolecular recognition.[15] These processes can occur after one another or partly simultaneously and thereby render the task of separating their individual contributions to the overall binding process experimentally quite complex, if not impossible.[15] If QSAR studies are to derive universally valid and applicable coherences, it becomes exceedingly important to keep the influence of these processes constant to empirically observe as few superimposed effects as possible, ideally only one.[15,16] This is usually achieved by employing a suitable model system for the respective question.[15] Quite intuitively, the more systematically a QSAR study is planned, i.e. the smaller the chemical modifications of a lead scaffold are within the series of investigated congeneric molecules, the higher is the chance of increasing the number of processes whose influence on the measured parameters will remain constant across all derivatives in a molecular series.[15] Iterative, highly systematic chemical changes can then not only disclose general trends that might also be applicable to other protein-ligand systems but may also pinpoint optimal performance as well as pitfalls in the molecular series and thereby enable deduction of the responsible characteristic as described for example by *Krimmer et al.* and *Cramer et al.* in the investigation of water networks surrounding protein-ligand complexes.[17–19] Furthermore, data for highly systematic molecular series may also be used to parameterize computer programs which are then capable of predicting the properties of protein-ligand systems that are difficult to

observe experimentally.[20] However, this approach comes at the expense of a comparably slow exploration of conceivable chemical space and is thereby more time and resource consuming compared to studies that are less systematically designed. This might be a probable reason for the relative scarcity of such systematic studies.[21] But despite a certain probability that trends derived from such studies will partly adhere to the studied systems themselves without wider applicability, there is also the chance that a higher number of such highly systematic studies will enable mutual comparisons, which will eventually lead to fundamentally applicable principles for the design of novel drug molecules.[21]

1.2 Thermodynamics and Kinetics in Drug Design

At the heart of every QSAR study is the question for the strength, i.e. the affinity, with which different molecules bind to a receptor. Affinity can be derived from the chemical equilibrium as shown in reaction (1) for a simple one to one binding reaction between protein (P) and ligand (L) to form a protein-ligand complex (PL).[22,23]

$$P + L \xrightleftharpoons[K_d, k_{-1}]{K_a, k_1} PL \qquad (1)$$

Mathematically, the equilibrium is described by the equilibrium association constant (K_a), the equilibrium dissociation constant (K_d), the association rate constant (k_1) and the dissociation rate constant (k_{-1}). Affinity, usually expressed as K_a, is the fundamental quantity in drug design, as it provides information about the strength of the interaction between protein and ligand without the necessity of further knowledge about the contributing factors that determine it.[22,23] However, K_d is usually the preferred quantity to compare the aptitude of different ligands to bind to a protein. From the law of mass action (2) corresponding to reaction (1) it is deducible that, for a one to one interaction, K_d bears the unit of concentration and corresponds to that equilibrium concentration of free ligand, at which half of the initial amount of free protein is

converted into the protein-ligand complex. This renders K_d a more intuitive and thus convenient quantity with which to work.

$$K_a = \frac{1}{K_d} = \frac{[PL]}{[P] \cdot [L]} \tag{2}$$

Affinity *per se* only gives information about the strength of binding of a ligand as a whole to a protein. It does not, however, provide evidence about the fashion in which affinity is conveyed, i.e. in which manner chemical changes on a lead scaffold influence the thermodynamic and kinetic parameters that govern the bound state and can shed light on the details of the molecular processes involved.[16] Thus, it is important to partition affinity into its contributors ot obtain further insights.[21]

1.2.1 Thermodynamic Partioning

A spontaneous chemical binding reaction between protein and ligand is characterized by a negative (i.e. favorable) value of difference in *Gibbs* free energy (ΔG), which can be obtained *via* equation (3) including the universal gas constant (R) and temperature (T).[22] ΔG in turn comprises enthalpic (difference in enthalpy (ΔH)) and entropic (difference in entropy (ΔS)) contributions as shown by the *Gibbs-Helmholtz* equation (4).[22]

$$\Delta G = R \cdot T \cdot \ln K_d \tag{3}$$

$$\Delta G = \Delta H - T \cdot \Delta S \tag{4}$$

Under the prerequisite of systematically designed QSAR studies, ΔH can be used as a measure for the change in the numbers and/or strengths of non-covalent interactions formed between protein and ligand during assembly of the complex.[24] Furthermore, it can be used to assess the quality of the geometric orientation of functional groups with respect to their contribution to ΔG.[24]

With increasing improvements in the applicability of microcalorimetry for the investigation of biomolecular recognition events, the assumption was put forward that strong enthalpic binders make for better drug candidates compared to binders with weaker enthalpic contributions. The terms 'enthalpically driven binder' and 'entropically driven binder' were coined, referring to that quantity of a ligand's thermodynamic binding signature that predominantly contributes to affinity.[25] In addition to the possibility of directly engineering individual protein-ligand contacts, the knowledge about enthalpic and entropic contributions can be used to rationalize crystallographic observations like highly disordered or flexible portions of a ligand molecule and a simultaneously comparably large entropic contribution to binding.[26–28] Furthermore, even the implication of crystallographically resolved water molecules that interact with the ligand as well as protons that are expelled or taken up upon complex formation can be elucidated.[29–34] Despite many reasonable correlations between thermodynamic and structural data, more recent QSARs demonstrated that knowledge about thermodynamics and structure alone does not necessarily yield conclusive insights. This can be the case if a ligand is preorganized, i.e. if its favored conformation in solution closely resembles the bioactive conformation in the bound state, a scenario that is structurally difficult to access, but often requires the support of MD simulations.[35,36] This circumstance can become even more cumbersome to elucidate, if the preferred conformation in solution comprises the entrapment of a water molecule, which is released before binding to the protein. This finding for the most flexible in a series of congeneric ligands of protein kinase A (PKA) caused a counterintuitively favorable entropic signature of binding, which refuted the traditional view that highly flexible ligands usually bind entropically penalized due to the immobilization of rotatable bonds.[37] Additionally, thermodynamics can also help to rationalize findings that lie beyond the physicochemical properties, e.g. the formation of mutation-induced resistances by identifying positions on a molecular scaffold that allow for enough flexibility of the scaffold to enable binding to the target despite mutations.[38,39] Moreover, a strong enthalpic signature has also been associated with enhanced selectivity.[25] Importantly,

the thermodynamic partitioning of ΔG enables the elucidation of the origin of enthalpy-entropy compensation (subsection 1.2.4), a phenomenon in which enthalpic and entropic contributions developing along a series of congeneric molecules cancel each other out and which, if actually present in the studied system and not caused by inaccuracies in the measurements, can seriously impede the aspired increase in affinity during lead optimization.[15,40] The only method capable of directly providing ΔH for a given binding event is isothermal titration calorimetry (ITC), which measures the apparent heat signal of the overall binding process including, importantly, all associated sub-processes as addressed in section 1.1, which makes the choice of a suitable model system important in order to circumvent the overlay of multiple effects.[15,21] But despite clear empirical correlations between structure and thermodynamic binding signatures, it is reasonable to assume that even the most thoroughly designed series of congeneric molecules will always bear a residual risk of additive or compensating effects which render the studied system more complex than initially intended.[41] Therefore, it has been argued that the mere enthalpic contribution of the interaction between protein and ligand is not at all accessible experimentally and that experimental values can only aid in the overall understanding of ligand binding and should preferably be complemented with computer simulations to take into account as much as possible of the complexity of ligand binding.[41] Recently, it was revealed that especially the effect of enthalpy-entropy compensation was often prematurely assumed to be present in certain studies without taking into account deficiencies in the experimental setups of microcalorimetric measurements, an issue that has increased awareness for the need of high microcalorimetric data quality.[40] Furthermore, it inspired the reasoning to desist from absolute terms such as 'enthalpically driven' and 'entropically driven', as the presence of systematic errors may not impact the general trends derived from individual experimental series, but does complicate the absolute comparison of experimental results of different protein-ligand systems on an absolute scale.[20,21] Additionally, despite examples for the correlation of favorable properties of a molecule with strong enthalpic binding, the increasing appreciation for the complex nature

of ligand binding has rendered the view, that simply a large enthalpic contribution makes for a better drug candidate, an outdated one.[16,21,24,25,41] But although it has been argued that computer-assisted design aimed directly at increasing affinity without initial consideration of individual thermodynamic contributions is putatively more productive due to the lack of universally applicable correlations between structure and the thermodynamic contributors to ΔG, the ability to partition affinity into its contributors remains an indispensable tool, as it has the power to disclose effects and mechanisms that cannot be elucidated easily otherwise and has the advantage of being experimentally evidenced.[21,40,42]

1.2.2 Kinetic Partitioning

While the initial perception of the significance of enthalpy has been relativized, the kinetic contributions governing affinity have gained in attention. Since the works of *Swinney et al.* and the introduction of the concept of residence time (τ_{rd}) by *Copeland*, the importance of the kinetics of biomolecular recognition has gained in appreciation.[23,43] For a simple one to one interaction, K_d can be calculated by the association rate constant (k_1) and dissociation rate constant (k_{-1}), which, in this scenario, equal the overall association rate constant (k_{on}) and overall dissociation rate constant (k_{off}, equation (5)).

$$K_d = \frac{k_{-1}}{k_1} = \frac{k_{off}}{k_{on}} \tag{5}$$

As τ_{rd} is the inverse of k_{off}, it can be directly derived from experiments providing the latter.[23] The importance of τ_{rd}, which is measured *in vivo* lies in the fact that it often resembles a ligand's *in vivo* efficacy a lot better than affinity *per se*.[23,44] An impressive example for this phenomenon was given by *Gooljarsingh et al.*, who showed that, despite low micromolar *in vitro* affinities, inhibitors of the heat shock protein 90 (HSP90) exerted a distinct antiproliferative effect in cell cultures, which was attributable to a pronouncedly long residence time.[45] The rationale for the validity of the residence time concept lies in the fundamental difference between an equilibrium

state under which many *in vitro* experiments are carried out as opposed to the open system of a human body.[44] In the latter case, the drug molecules are exposed to the pharmacokinetic processes of absorption, distribution, metabolism and excretion, which prevent the formation of an equilibrium state upon which the definition of K_d is based.[44] A prolonged residence time provides for a longer target occupation and thereby prolonged action at the receptor, irregardless of the concentration of free inhibitor molecules in the vicinity.[23,44] The importance of this concept becomes even more distinct, if the presence of an induced fit (section 1.3) is considered.[46] In this scenario, the initially formed complex or intermediate state PL, kinetically characterized by k_1 and k_{-1}, is considered loosely bound, and successively isomerizes with the forward isomerization rate constant (k_2) and reverse isomerization rate constant (k_{-2}) to form a high-affinity complex (PL*, equation (6)).[47]

$$P + L \; \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \; PL \; \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \; PL^* \tag{6}$$
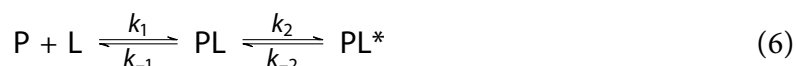
The binding model of induced fit leads to the more complex dependence of k_{off} shown in equation (7) and can be utilized for the design of k_{off}, if the structural features of protein isomerization and the means to modulate these are known.[23]

$$k_{\text{off}} = \frac{k_{-1} \cdot k_{-2}}{k_{-1} + k_2 + k_{-2}} \tag{7}$$

Exemplarily, *Cramer et al.* succeeded in demonstrating the paramount importance of an interaction between the side chain of Asn112 of the model metalloprotease Thermolysin and a bound ligand. Movement of this side chain determines the open and closed state of the protease and fixing it in the closed state by hydrogen bonding to a bound ligand was suggested to be a crucial determinant for a prolonged residence time.[48] Moreover, additional blockage of the movement of this side chain by the known ligand ZF^{P}LA was suggested to majorly contribute to the extraordinarily long residence time of 168 d by impeding the transition from the closed to the open state.[48] A further example for the relevance of the residence time concept is the case of

inhibitors of the epidermal growth factor receptor (EGFR) tyrosine kinase. An inhibitor with a prolonged residence time suppressed EGFR tyrosine phosphorylation and thereby activity significantly longer and to a larger extent than inhibitors with pronouncedly shorter target occupancies.[49] Furthermore, k_{off} can be used to manipulate selectivity profiles, e.g. in the case of isozyme-selective inhibition of cyclooxygenases, or indicate the tendency of molecules to induce resistances, as shown for marketed drugs targeting the human immunodeficiency virus (HIV)-1 protease.[50–52] Despite the plausible argumentation for residence time, it should be noted that the design efforts aiming at the prolongation of residence time are futile in cases in which residence time does not exceed the pharmacokinetic lifetime of a drug molecule.[53] Despite the more intuitive character of τ_{rd} or the argument that the maximum value achievable for k_{on} is the rate of diffusion, k_{on} can still be utilized in the concept of rebinding.[53] It relies on an exceedingly large association rate constant which provides for a rapid renewed engagement of the same target or other target molecules within the same effect compartment after dissociation from its initial target and can thereby overcompensate an insufficient residence time.[23,53] *Sykes et al*. demonstrated that the extrapyramidal side effects of atypical antipsychotic drugs correlate with k_{on} rather than k_{off}, a finding that was rationalized by a model of rebinding.[54] A practical obstacle to the rational design of overall association rate constants is the fact that processes like desolvation and diffusion, which are necessary prerequisites for the eventual complex formation, are non-trivial to approach as there is little systematic knowledge about the correlation between structure and kinetics to guide organic synthetic efforts.[23] A contribution in this regard was made by *Schuetz et al*. who demonstrated the deliberate deceleration of inhibitors of HSP90 by increasing the energetic barrier of desolvation.[55]

1.2.3 The Hydrophobic Effect

The effect of spontaneous self-assembly of hydrophobic molecules or hydrophobic portions of amphiphilic molecules in water has long-since been accepted as vitally important for the formation of three-dimensional biological structures such as proteins or membranes and biomolecular

recognition events such as binding between a drug and its target.[56–59] The driving force for this is assumed to be the discrepancy between the strength of interaction between the hydrophobic solute and water molecules on the one hand, and the strong bonding between water molecules. The interactions formed between a hydrophobic solute and water are relatively weak, as they occur between induced dipoles of the solute molecule and the permanent dipoles of water molecules, while the hydrogen bonding between water molecules is solely based on the interaction of permanent dipoles. Based on the 'iceberg' model postulated by *Frank* and *Evans*, water molecules in the first hydration shell surrounding the solute are considered to be ordered in hydrogen-bonding networks which are similar to those formed in ice in geometry and strength and more highly ordered than in bulk water.[60] This view has been relativized to some degree with respect to the extent at which order is increased in the first hydration layer.[58,60] As a consequence of this non-optimal binding situation of the water molecules, the hydrophobic solute is released from its water envelope to assemble with another solute molecule of its kind. Thus, the assembly of hydrophobic solutes seems to be mainly driven by the strong attraction between water molecules.[61] Simultaneously, the number of water molecules required to cover the hydrophobic assembly and thereby involved in weaker interactions with the solute is reduced compared to the number required to cover the individual hydrophobic entities.[56,62] Based on this model, the 'classical' hydrophobic effect predicts that the release of water molecules resulting from the binding between two hydrophobic molecules will entail a favorable entropic contribution to the difference in *Gibbs* free energy as the released water molecules will form an increased number of hydrogen bonds with the surrounding bulk water.[61,62] Nowadays, however, the view of a single hydrophobic effect is outdated, as an additional observation can be made for certain host-guest interactions. The 'non-classical' hydrophobic effect includes the existence of highly mobile high-energy ('frustrated') water molecules present in hydrophobic binding pockets.[63] Displacement of these water molecules into the bulk solvent will result in hydrogen bonds stronger than those formed around the hydrophobic surface, which in turn results in an

enthalpic gain.[62] The differentiation is complicated by the circumstance that the shape of the hydrophobic solute influences the energetic state of the water molecules surrounding it.[59,64] Thus, it can be assumed that differentiating only these two possible extremes will not give sufficient consideration to the issue. The nature of the hydrophobic effect will always depend on the system under investigation, especially with respect to the geometric characteristics of the investigated system, which strongly influence the nature of water molecules that are assembled around these regions.[62] In consequence, the complex role that water plays in protein-ligand binding is difficult to grasp, especially because detailed systematic studies addressing the issue will always bear a context-based bias.[59,65,66]

1.2.4 Enthalpy-Entropy Compensation

The effect of mutual compensation of enthalpy and entropy, termed enthalpy-entropy compensation (EEC), which entails a minimal value of ΔG for a spontaneous reaction despite strong enthalpic and entropic signatures, has long-since been recognized as a widely occurring feature of biochemical processes such as protein folding.[67] Furthermore, the formation of protein-ligand complexes investigated for the establishment of QSARs can be subject to the phenomenon, which can seriously impede the effort to improve a given lead scaffold and has been shown to occurr to various extents by several studies.[26,33,68,69] The most intuitive reason for the occurrence of this phenomenon is the opposed nature of motion and bonding.[42] Stronger, i.e. more enthalpic bonding reduces mobility and therefore entropy, while weaker, i.e. less enthalpic bonding, maintains a higher mobility and higher entropy.[20,42] Furthermore, the rearrangement of solvent molecules can be the source of this effect.[40] However, the origin and detailed characteristics of compensation are usually system-dependent and the generalization of universally applicable rules is therefore difficult.[42] Despite the evidence for the phenomenon of EEC, it is also known that the derivation of thermodynamic data by ITC is not without flaw. It has been suggested that deficient experimental setups, especially with respect to reactant concentrations, are the major source of measurement errors in the directly measured quantities of K_a and ΔH, which

consequentially translate into an error in ΔS as shown in subsection 1.2.1.[40,70] Therefore, EEC is often likely to be an attribute of erroneous measurements, rather than a genuine feature of the investigated sytem.[20,40]

1.3 Main Types of Interactions in Protein-Ligand Binding

The complementary nature of an enzyme's binding pocket and a ligand, termed 'lock-and-key' principle by *Fischer*, can be seen as the initial, intuitive approach for the design of a ligand.[71] If a ligand is to bind to a target, a basic requirement is a certain degree of shape complementarity between ligand and target. This concept was expanded by *Koshland* to the theory of 'induced fit', which takes into account the flexibility of molecules and the resulting possibility of a conformational adaptation of protein and ligand upon binding.[46] In addition to the indispensable knowledge about the three-dimensional shape of a protein binding site that enables assumptions about the required three-dimensional shape of a putative inhibitor, the directed design of small molecule inhibitors also requires information about the physicochemical properties of the binding pocket. The knowledge about the orientation of polar functional groups and the locations of hydrophobic as well as aromatic hot spots is necessary to devise the interaction pattern to be emulated by a ligand in order to potently and selectively bind to its target.[25,46,72,73] A recent systematic investigation of protein-ligand complexes deposited in the Protein Data Bank (PDB) provides insight into the relative abundance of different types of intermolecular interactions between small-molecule ligands and proteins.[72] Hydrophobic interactions were found with the highest abundance, hydrogen bonds constitute the second most abundant kind of interaction.[72] The fact that hydrophobic interactions occupy the first position in this ranking corroborates the medicinal chemistry dogma of the relative simplicity of the installation of hydrophobic groups on a molecular scaffold. The fact that the desolvation of apolar residues requires less energy than that of polar moieties as well as the direction-independent nature of hydrophobic interactions renders the installation of apolar functionalities on a molecular scaffold the most prominent

and convenient, because comparably simple, starting point for a design effort.[72] However, an increased hydrophobic character inevitably entails a reduced bioavailability due to reduced solubility in an aqueous environment.[72] Moreover, due to the lack of polar functionalities that interact in a spatially specific manner, an enhanced hydrophobic character is usually accompanied by a decrease in selectivity and thereby an increased off-target toxicity.[73] In contrast, polar interactions such as hydrogen bonds underlie high geometric restraints. Hydrogen bonds were found to be the second most abundant form of interactions between proteins and ligand molecules. They resemble the main form of directed interactions in biological systems and the determining factor in biomolecular recognition, impressively demonstrated by the pairwise assembly of guanine and cytosine on the one hand and adenine and thymine on the other hand in the formation of the double helical DNA strand.[72,74,75] As useful as they are, the nature of polar chemical functionalities bears an inherent dilemma which is especially detrimental to drug design. Not only is it comparably difficult to install hydrogen bond donors and acceptors on a molecular scaffold in a way that satisfies the geometric requirements of the binding pocket.[72] Polar groups also come at a relatively high energetic cost of desolvation. Therefore, in order to contribute favorably to the overall difference in *Gibbs* free energy, the bond formed between the polar functionality and receptor must be sufficiently strong to overcome the energy required for desolvation.[55] The magnitude of the problem is additionally increased, if a charge-assisted hydrogen bond is to be designed. A charged polar functionality on the ligand scaffold can form a salt bridge with a charged residue on the protein, but it will also increase the energetic cost of desolvation. Furthermore, it is necessary to distinguish the location of the hydrogen bond, as a higher degree of solvent exposure was found to correlate with a lower contribution to the overall affinity, compared to that of hydrogen bonds buried deeply within the binding site.[72] Naturally, hydrophobic interactions and hydrogen bonds are not the only interactions relevant to biomolecular recognition, but clearly the most abundant.[72]

1.3.1 Fluorine in Medicinal Chemistry

A common substitution in lead-like molecules is that of a hydrogen atom for a fluorine atom. But despite the similar steric demands, which entails considerable electronic consequences.[76] Fluorine is the element with the highest electronegativity electronegativity (χ). As a consequence of the differences in χ, the charge distribution changes distinctly upon exchange of H for F. While the C−H bond has a distinctly apolar character ($\Delta\chi_{C-H}$ = 0.4), the C−F bond is notably polarized, with fluorine bearing the negative and carbon bearing the positive partial charge ($\Delta\chi_{C-F}$ = −1.5).[76] The C−F bond thus has a partly ionic character and dipole moment and therefore can be interpreted as a strong electrostatic interaction between $C^{\delta+}$ and $F^{\delta-}$.[76] This is exemplarily reflected in the successive shortening of all C−F bonds upon the iterative exchange of one H-atom for one F-atom at a time in the transition from CH_4 to CF_4, which correlates with the increasing positive polarization of the C atom with an increasing degree of fluorination.[76] Furthermore, due to the combination of its *van der Waals* radius (r_{vdW}) of 1.47 Å and nine electrons, fluorine bears a much denser, and thus less polarizable, electron cloud than a hydrogen atom (r_{vdW} = 1.2 Å, one electron) which assigns fluorine a distinct role in the fine-tuning of physicochemical properties in organic molecules.[76] The manner in which these are altered, however, is context dependent. The perfluorination of hydrocarbons leads to fluorocarbons. Due to the increased size of F, the arrangement of the CF_2 units follows a helical fashion, opposed to the straight but alternating (zig-zag type) arrangement of CH_2 groups and, in addition to low polarizability, accounts for the increased hydrophobicity as a result of the enlarged cross-section area of the fluorocarbons.[77,78] Fig. 1 shows the structural difference between *n*-decanoate and *n*-perfluorodecanoate from Cambridge Structural Database (CSD) entries LUNZUP and GUSTIZ, respectively.[79,80]

Futhermore, fluorocarbons feature the unique property of lipophobicity, a term that expresses the tendency of fluorocarbons to form a phase of their own instead of mixing with hydrocarbons.[77] Moreover, the highly polarized C−F bond leads to a stabilization of the C−C bond by a reduction
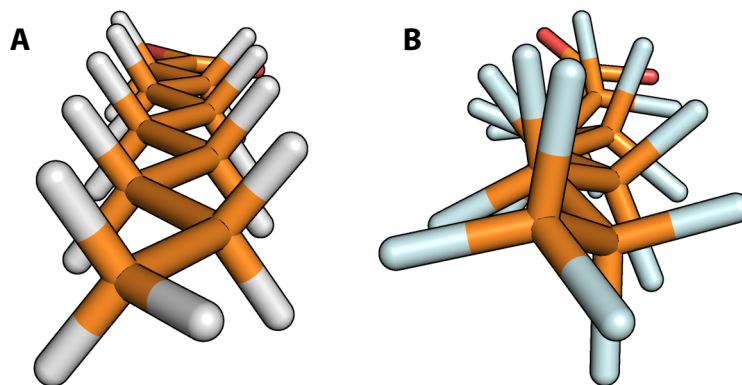
Fig. 1: Structural consequences of perfluorination of an alkyl chain. **A** Stick model of *n*-decanoate from CSD entry LUNZUP. **B** Stick model of *n*-perfluorodecanoate from CSD entry GUSTIZ.

of the size of the carbon orbitals, which furthermore entails a decrease in reactivity due to a lack of energetically suitable orbitals.[77] Furthermore, the slightly larger fluorine atoms cover the carbon backbone more efficiently than hydrogen, which, in combination with reduced orbital energies, renders fluorocarbons one of the most inert groups of chemicals known to date.[77] In contrast, in a lead development effort, usually not all C−H units are converted into C−F bonds. Rather, specific sites of the molecular scaffold are modified by the introduction of a single fluorine atom or fluorinated structural motifs. The interest in selectively installing fluorine atoms in various types of organic molecules is reflected in the vast number of fluorination agents and strategies and the plethora of fluorine-containing structural motifs.[81,82] Reasons for the incorporation of fluorine in drug molecules are manifold. With respect to the conformational properties of fluorocarbons, the selective incorporation of single fluorine atoms can be applied to fine-tune conformational preferences of alkyl chains, which can be important for the design of macrocyclic drug molecules to preorganize them so as to enforce the conformation of the bound state in the molecule in solution.[82,83] Furthermore, fluorine can alter the pharmacokinetic stability of a molecule by affecting rate and path of metabolism with respect to different fluorination sites and alter the distribution of a drug in tissue.[84] In this latter respect, fluorine is advantageous in so far, as the tissue distribution of the only natural isotope ^{19}F can be traced nonivasively by magnetic resonance spectroscopy.[84] This principle has also been employed in NMR-based

fragment screening.[85] Furthermore, fluorine can have an impact on the nature of protein-ligand binding itself as described e.g. for the first approved fluorinated drug fludrocortisone.[82,86] As opposed to the perfluorination-induced non-polarizable, hydrophobic and lipophobic nature of fluorocarbons, the incorporation of a single fluorine atom or a fluorine-containing structural motif can distinctly influence the charge distribution of a molecule. A prominent example for the utilization of fluorine is the systematic replacement of $C-OH$ bonds for $C-F$ bonds because of the similarly polar character of both and the lack of the hydrogen bond donor ability of the latter. The probing of the binding site of a carbohydrate-binding protein with fluorinated carbohydrate analogues demonstrated that the $C-F$ bond formidably mimics the polar character of the $C-OH$ bond, and, due to the lack of the hydrogen bond donor ability, revealed positions of the carbohydrate scaffold at which this ability was required for binding.[87] Furthermore, the investigation of collagen stability by the substitution of 4*R*-hydroxyproline for 4*R*-fluoroproline showed that the polar character of the $C-OH$ bond is necessary for helix stability, rather than the hydrogen bond donor functionality.[88] To sum up this combination of the practically non-polarizable nature of a fluorine atom and its strongly polarizing effect, the term 'polar hydrophobicity' was coined.[89] Given the low polarizability of fluorine itself, the introduction of fluorine-bearing motifs intuitively leads to an increased hydrophobicity of the molecule.[82] However, such groups are often associated with an increased polarity due to their strong electron withdrawing nature.[82] This shift in charge distribution has consequences for proximal functional groups. The 2-monofluorination or the 2,6-difluorination of a phenol distinctly decreases the acidity constant (pK_a) of the OH group, which can be used to increase binding affinity or install the difluorinated motif as a bioisostere for a carboxylic acid function.[90-94] An often employed strategy is the replacement of a functional group with fluorinated groups that not only mimic the geometry, but also the dipole moment of the respective functionality. E.g., CF_2H can be used isosterically for OH, as the fluorine atoms mimic the lone pairs of the oxygen atoms and furthermore render the alkyl proton comparably acidic and thus enable it to act as

hydrogen-bond donor.[82] In addition to the implications for specific interaction patterns formed between protein and ligand, fluorination can also increase membrane permeability and thereby bioavailability due to an increased hydrophobic character of either the fluorinated motif itself, or e.g. a proximal basic functionality with reduced basicity due to the inductive effect of fluorine. A fluorinated motif can nonetheless bear a strong dipole moment itself and thereby maintain the ability to form polar interactions (compare OH and CF_2H).[82,95] Another example is the CF_3 group. Intuitively, it seems reasonable to assume that the non-polarizable fluorine atom and isotropically fluorinated structural motifs such as CF_3 often occupy hydrophobic pockets.[82] This is reflected in the fact that the interaction between a receptor and either a fluorine or chlorine atom of a ligand is the second most abundant hydrophobic interaction in the PDB.[72] Nonetheless, CF_3 is also capable of interacting with a carboxylate function due to the high positive polarization of the carbon atom ('tetrel bonding').[82] Another important aspect is the fluorination of aromats, which does not only modulate metabolism or membrane permeability, but distinctly changes the aromatic quadrupole moment and in consequence distinctly alters aromatic π-π interactions. This principle was used exemplarily to fine-tune π-stacking interactions for ligands of Carbonic Anhydrase (CA), the protease human Cathepsin L and the bacterial adhesin FimH.[96–98] Furthermore, various fluorination patterns of the phenyl core of BSA inhibitors of CA showed an increased affinity compared to their non-fluorinated counterparts.[99–101] Beside the mere modulation of overall molecular properties, fluorine is capable of forming bonds to a macromolecular target. An approach to assess the fluorophilicity of a target's binding site is the systematic variation of the fluorination pattern of a phenyl ring ('fluorine scan').[102] Fluorine is known to engage in other polar interactions, albeit in putatively weak ones, for example the interaction with the carbonyl group of a backbone amide bond in a *Bürgi-Dunitz* type trajectory.[76,103,104] Whether ligand-bound fluorine forms hydrogen bonds with a macromolecule is still matter of debate.[76] However, there is crystallographic precedence for the involvement of fluorine in hydrogen bonds, most simply with itself in an HF-trimer motif , even though

these are assumed to be weak compared to more commonly found hydrogen bond motifs.[76,105] Furthermore, due to its small size and the resulting low polarizability, fluorine differs significantly from its higher homologues. Unlike chlorine, bromine and iodine, fluorine is capable of weakly engaging in 'side-on' interactions with its lone-pairs, but does not engage in 'head-on' halogen bonding, which refers to the interaction between the positively polarized cap of a halogen atom ('σ-hole') with an aromatic π-system or a lone pair.[106] Fluorine only shows a σ-hole in extreme cases with little relevance to the design of drugs.[106]

1.4 Alkyl Chains in QSAR Studies

Due to the relative chemical simplicity of installing alkyl chains on a molecular scaffold and the strictly systematic fashion in which their attachment can produce chemical variety, they were employed in many studies aimed at establishing QSARs. *Malham et al.* found a favorable development of standard free energies of binding along with an enhancement of the standard enthalpy of binding throughout a series of increasingly hydrophobic primary alcohols as inhibitors of the major urinary protein.[107] The entropic contribution became increasingly unfavorable with increasing chain length, which could be traced back to dispersive protein-ligand interactions, while the classical hydrophobic effect fell short of explaining these findings.[107] The expansion of cyclic hydrocarbons as subunits at the center of a Grb2-SH2 domain inhibitor scaffold was observed to entail an increase in affinity based on the burial of hydrophobic surface area and, contrarily to the classical hydrophobic effect, with increasingly favorable enthalpic contributions by *Myslinski et al.*.[108] Furthermore, it was found, that a gain in the standard enthalpy of binding was compensated by a decreasing contribution of the entropic term by aliphatic derivatives of this ligand series.[109] *Krimmer et al.* succeeded in the design of water networks around solvent-exposed parts that contribute favorably affinity to the overall binding affinity, by decorating a thermolysine inhibitor scaffold with hydrophobic moieties of various lengths and degrees of branching.[17,18] *Cramer et al.* demonstrated, that it can even be detrimental for affinity

to expose polar groups to the bulk solvent, rather than hydrophobic ones, due to solvation phenomena.[19] *Jain et al.* and *Krishnamurthy et al.* investigated hydrophobic tails of CAIs with oligoethylene glycol-, oligoglycine- and oligosarcosine tails, by which the presence of EEC could be demonstrated for these systems, whereas growing alkyl- and perfluoroalkyl tails were shown to increase affinity for CA.[68,110,111]

1.5 Isothermal Titration Calorimetry

ITC is the method of choice for the determination of the thermodynamic parameters governing a (bio)molecular recognition process. The reacting species in the sample cell is titrated step-wise with a reactant from a syringe under constant stirring. A thermal element maintains the temperature of the reference cell. A heat sensor reports the temperature difference between sample cell and reference cell caused by the reaction heat (Fig. 2). The power (P_Q) required by or withdrawn from the thermal element in order to keep the temperature of the sample cell at that of the reference cell is recorded and generates the final thermogram (Fig. 3).[20,112]
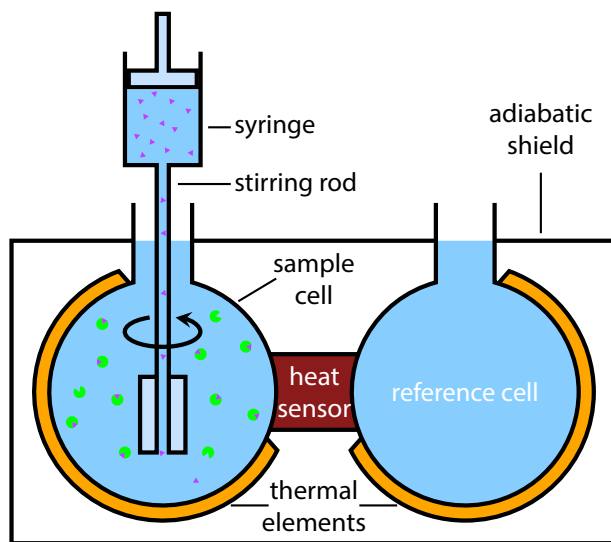
Fig. 2: Schematic representation of a microcalorimeter, adapted from references [112],[22] and [20].

The signals of the resulting thermogram are integrated and plotted against the molar ratio of titrant to titrand, which enables the extraction of ΔH, K_a and stoichiometry of binding

(n_s) (Fig. 3).[22] Thus, by applying equations (2), (3) and (4), ITC provides all thermodynamic parameters characterizing the binding reaction. More recently, the scope of extractable data was augmented by the method of kinITC, which provides information about k_{on} and k_{off}.[113,114] Thereby, a full characterization of a binding event under one set of experimental conditions becomes possible without the need to resort to other techniques with different experimental prerequisites or the need for the labelling of samples. For this purpose, the equilibration time (τ_{eq}), i.e. the time the reaction solution in the sample cell requires to equilibrate (visible by the return of the signal to baseline level in a thermogram) is calculated with k_{off} as fit parameter.[115] Hence, k_{on} is accessible *via* equation (5). The presence of kinetic information in the experimental data becomes apparent in an equilibration time curve (ETC), which depicts the dependence of an injection's equilibration time on the molar ratio (r_m). Fig. 3 shows the connection between an experimental thermogram, the integrated heat values and the occurrence of kinetic information in an ITC experiment.
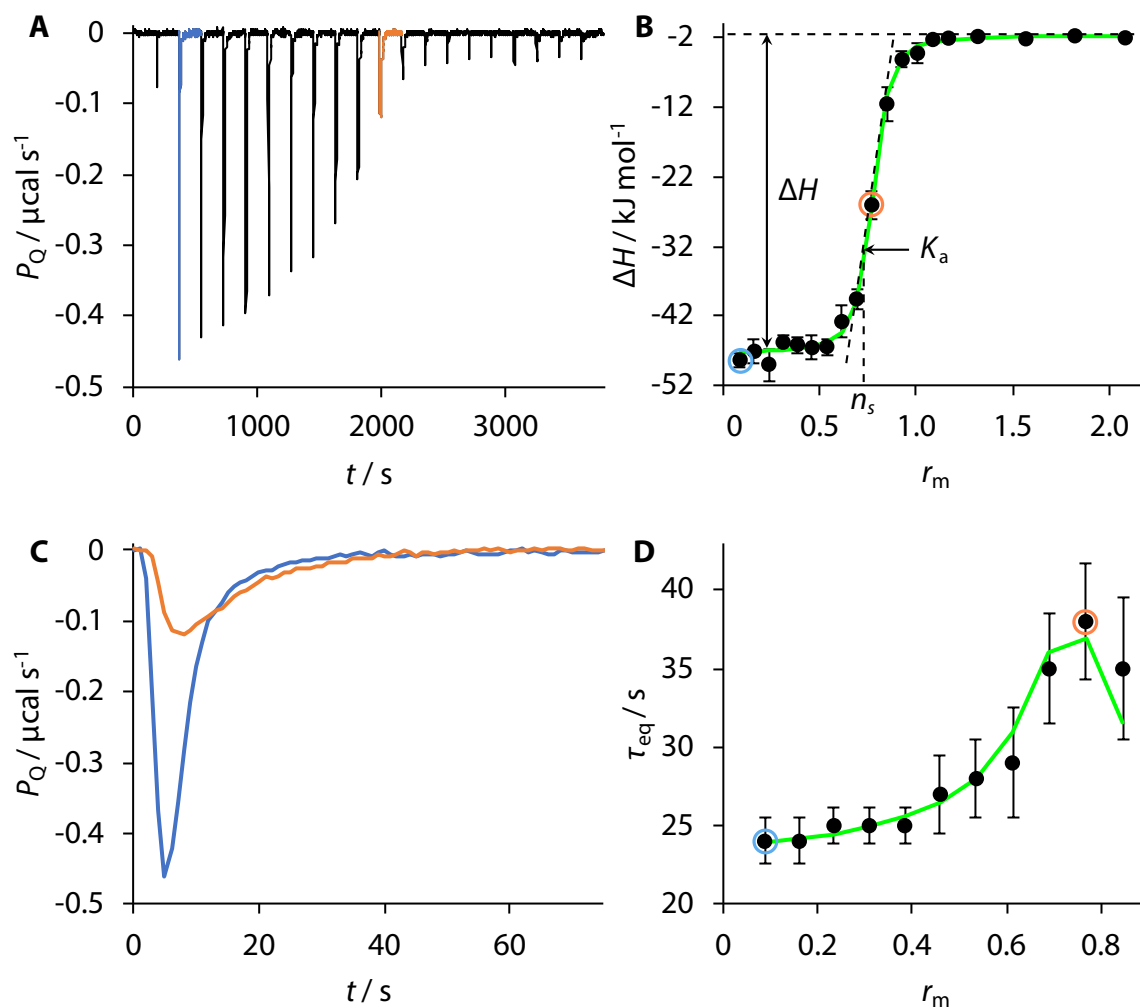
Fig. 3: A Thermogram measured for the interaction between hCAII and 4-nitrobenzenesulfonamide with an early injection colored in blue and the injection closest to the inflection point colored in orange. P_Q denotes the applied power. **B** Integrated heat values from the thermogram in **A** with a green fit function and marked data points representing the marked injections in **A**. ΔH is represented by the difference between both plateaus, n_s is the value of r_m at the inflection point. **C** Close-up view of the superimposed marked injections in **A**. A broadening of the signal with proceeding titration is visible. **D** ETC derived from the broadening of injections with proceeding titration.

1.6 Macromolecular X-ray Crystallography

X-rays interact with matter by scattering on the electron clouds of atoms.[116] Crystalline material

is constructed periodically, i.e. composed from a multitude of unit cells in all three spatial

directions.[116] The reflection of X-rays from a crystal in an X-ray diffraction (XRD) experiment

occurs on sets of parallel planes that intersect the crystal (Fig. 4).[117]

A

B
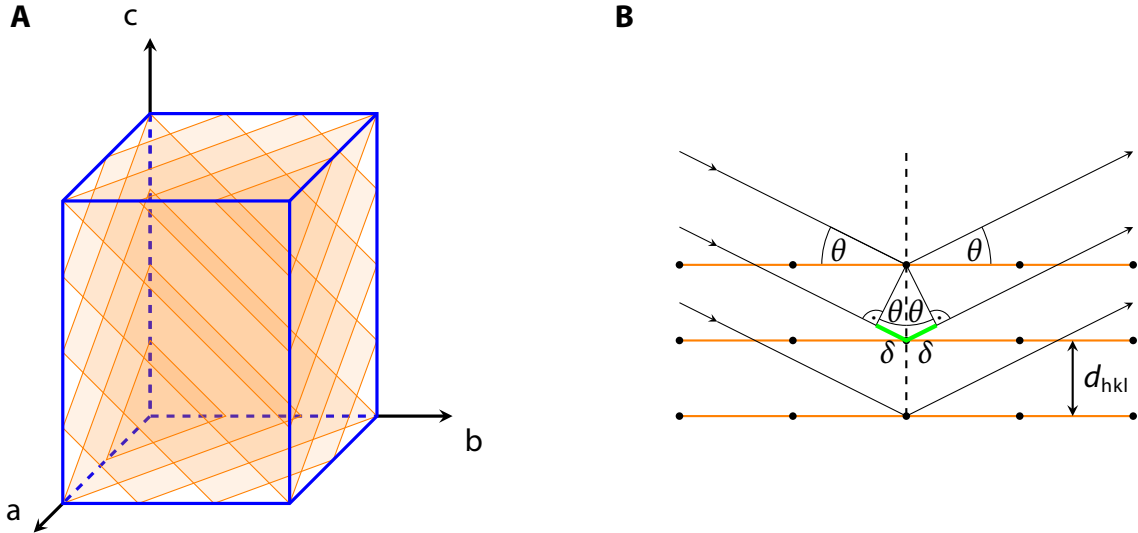
Fig. 4: A Unit cell with arbitrary dimensions a, b and c and the semi-transparent intersecting (234) planes in orange. **B** Side-on view of a set of planes in orange with scattering atoms in black. The incident beams are diffracted and enclose the *Bragg* angle θ with the respective plane. As a consequence of geometrical considerations, θ is also present between the dashed line orthogonal to the set of planes and lines between a scattering atom in the upper plane, which intercepts the middle incident and reflected wave orthogonally. The green lines represent the offset $\Delta = 2 \cdot \delta$. By *Bragg*'s law, constructive interference, and thus a reflection, occurs, if Δ is an integral multiple of λ.

From Fig. 4**B** it becomes evident, that the middle ray travels an additional distance of 2δ compared to the uppermost ray. It can furthermore be deduced that δ can be calculated by equation (8).

$$\delta = d_{hkl} \cdot \sin \theta \qquad (8)$$

Father and Son *Bragg* showed, that in order for a reflection to occur, θ must fulfill the following condition, known as *Bragg*'s law:

$$2 \cdot d_{hkl} \cdot \sin \theta = n \cdot \lambda \qquad (9)$$

in which θ represents the incident angle of the impinging ray, λ the wavelength and n an integer, known as the diffraction order. Constructive interference between rays reflected from different planes within a set, and thereby eventually a recordable reflection, can only occur for values of θ that enable the difference in travelled distance $\Delta = 2 \cdot \delta$ between two rays to be an integral multiple of λ.[118] Rotation of the crystal through the X-ray beam arranges several sets of planes in an orientation that fulfills equation (9) and the recording of one image per

orientation thereby produces a dataset that contains information about the three-dimensional distribution of electrons in a unit cell.[118] The dataset can thus be used to produce an electron density map, that, combined with chemical knowledge about the crystal's components, can be used to assign atomic coordinates and thereby build a molecular model.[118,119] Fig. 5 shows a diffraction image recorded at beamline (BL) 14.1 at the BESSY II electron storage ring operated by the Helmholtz-Zentrum Berlin and a snapshot of the crystallographic modelling software Coot that shows an electron density map with modelled protein and ligand, produced from the respective dataset.[120,121]
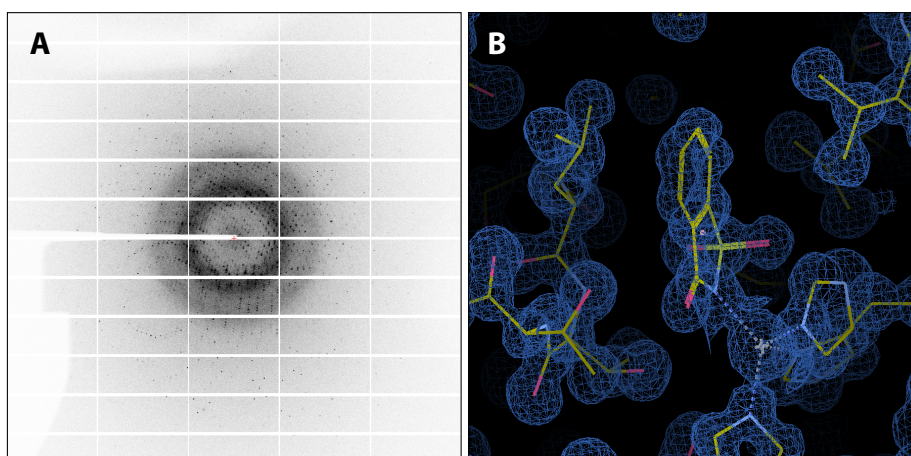
Fig. 5: A Diffraction image of an hCAII crystal. **B** Electron density map with modelled protein and ligand generated from an XRD dataset of an hCAII crystal soaked with saccharin sodium salt.

The investigation of crystalline materials by XRD has profoundly shaped the understanding of the spatial organization of matter in that it provides information about the positions of atoms relative to one another and hence enables the deduction of atomic radii, bond lengths, the influence of the chemical environment on these, or the characteristics of various types of noncovalent interactions between chemical entities.[72,118,122] Scattering of X-rays occurs on the electrons surrounding an atom. Hence, the result of a successful XRD experiment is a three-dimensional electron density map which resembles the shape of the chemical matter responsible for the scattering.[116] Assuming that the unit cell, from which the crystal can be constructed by repeated addition of unit cells in all three spatial directions, does not have strictly the same

orientation and thermal motion in every asymmetric unit, the final electron density map will always depict the average of electron distribution over all the asymmetric units in the whole crystal.[116] Thus, it is intuitively understandable, that protein crystals usually feature a higher degree of disorder than small molecules or peptide crystals, given the considerably larger size of proteins.[123] The derivation of atomic properties such as radii and bond lengths are not as exactly determinable from protein crystals as from small molecule crystals, especially not at resolutions lower than common bond lengths in small molecules, and thus the subsequent building and refinement of a molecular protein model highly depends on restraints and constraints derived from small molecule structures.[116] After crystallization and incubation with a ligand of interest and X-ray data collection, the main problem in the process of structure determination is the step of phasing. The XRD experiment delivers the intensities of the recorded reflections from which the structure factors (i.e. the amplitudes of the detected X-rays) can be calculated. However, the phases of the recorded X-rays bear a considerably higher information content, but are lost in the process of detection and thus need to be recovered subsequently.[117] This circumstance was termed 'phase problem'.[117] As amplitudes and phases are not linked by a formal relationship, their only mutual connection is the molecular structure or electron density.[117] Thus, if there is some initial knowledge about either of these, it can be used for the determination of phase values.[117] Under the prerequisite of atomic resolutions < 1.2 Å, direct methods mostly used for small molecule crystals, e.g. the utilization of known heavy atom sites, can lead to a solution to the phase problem.[117] However, resolutions in this range are rarely observed for protein crystals, except for crystals of intensively studied model proteins such as CA. Molecular replacement with a structurally similar model is a convenient, because simple and fast, method for phasing, but comes at the disadvantage of model bias.[117] Furthermore, isomorphism of crystals (i.e. the circumstance that both crystals comprise the same unit cell), can be used. Derivatizing one crystal by soaking it in a solution of a heavy metal compound can alter the intensities of the reflections compared to those of the non-derivatized crystal. The differences in intensities between both datasets

can be used to infer the positions of the heavy atoms.[117] A disadvantage of this approach is, that metal species can putatively change the local environment in which they bind and hence reduce isomorphism in the derivatized crystal. Furthermore, the phenomenon of anomalous dispersion can be employed to derive the coordinates of the so called anomalous scatterers in experiments such as single-wavelength anomalous dispersion (SAD).[117] This approach exploits the inherent ability of atoms to absorb X-rays of specific wavelengths dependent on the atomic number.[118] The absorbed photon excites an electron, which, upon falling back to its ground state, emits a photon with a phase different from that of the incident photon.[118] Although an inversion center is not necessarily present in a crystal, every diffraction pattern is centrosymmetric.[122] *Friedel*'s law states, that the intensity (I_{hkl}) is the same for centrosymmetrically related reflections (equation (10)).[122]

$$I_{hkl} = I_{\bar{h}\bar{k}\bar{l}} \tag{10}$$

In the presence of anomalous scatteres, however, this parity is broken due to the phase difference between absorbed and emitted photon and the resulting disparity can be used to determine the locations of the anomalous scatterers.[122]

1.7 Carbonic Anhydrase

The enzyme family of Carbonic Anhydrases (CAs) has long since been subject of intensive research. Since the discovery of the inhibitory acitivity of sulfanilamide towards CA and the clinical investigation of the first marketed CAI acetazolamide in 1954, CAs have been under extensive investigation for the treatment of various diseases.[101,124–126] Of the seven genetically different CA families (α-, β-, γ-, δ-, ζ-, η-, θ-CAs) known to date, α-CAs represent the only family present in humans. It comprises 15 isoforms with different organ and tissue distributions, catalytic activities and susceptibilities to different inhibitor classes.[126,127] The main function

of α-CAs is the interconversion of carbon dioxide and water and bicarbonate and a proton (reaction (11)).

$$CO_2 + H_2O \rightleftharpoons HCO_3^- + H^+ \qquad (11)$$

Consequentially, α-CAs are involved in pH regulation and more generally, electrolyte secretion. Futhermore they are linked to several vital biological processes such as the synthesis of D-glucose or lipids and the metabolism of bones but also in the pathogenicity of tumors and neurodegeneration.[127,128] Thus, CAs offer a broad range for the treatment of various diseases by activation or inhibition. The second reason for the high profile of CAs in biophysical and biomedical research are the physicochemical properties, which render especially the best-known isoform II a highly convenient model for studies of the underlying principles of ligand binding.[59,129] CAII is a comparably rigid protein, is usually assumed to remain virtually unchanged upon ligand binding.[15] This is especially convenient for the study of protein-ligand binding as differences in binding behavior must originate from the structure of the ligands, rather than from conformational changes of the protein.[15] Furthermore, the conically shaped binding site comprises two halves, one characterized by hydrophobic amino acid side chains ('hydrophobic wall'), the other by hydrophilic amino acid side chains.[59] Furthermore, the binding of sulfonamides with the catalytically active Zn^{II} ion is well understood.[59] Additionally, CAII is easily expressed and crystallizes readily.[59]

1.7.1 Sulfonamides as Inhibitors of human Carbonic Anhydrase II

Aromatic sulfonamides are the classical pharmacophore motif for the inhibition of CA since the discovery of the inhibitor sulfanilamide.[125,130,131] The significance of the sulfonamide motif lies in mimicry of the transition state of CO_2 hydration (reaction (11)).[132] A transition state analog is generally assumed to be bound more tightly than either substrates or products of the enzymatic reaction.[133–136] Therefore, the design of transition-state analogs aims at creating a stable chemical

moiety with a *van der Waals* geometry and electrostatic surface potential as close as possible to that of the transition state to be mimicked.[136] The neutral nature of the Zn^{II}-bicarbonate complex is maintained by the circumstance, that sulfonamides coordinate the Zn^{II} cofactor anionically.[137,138] Fig. 6 depicts the similarity between a Zn^{II}-bound bicarbonate ion resulting from CO_2 hydration and the Zn^{II}-bound sulfonamide.
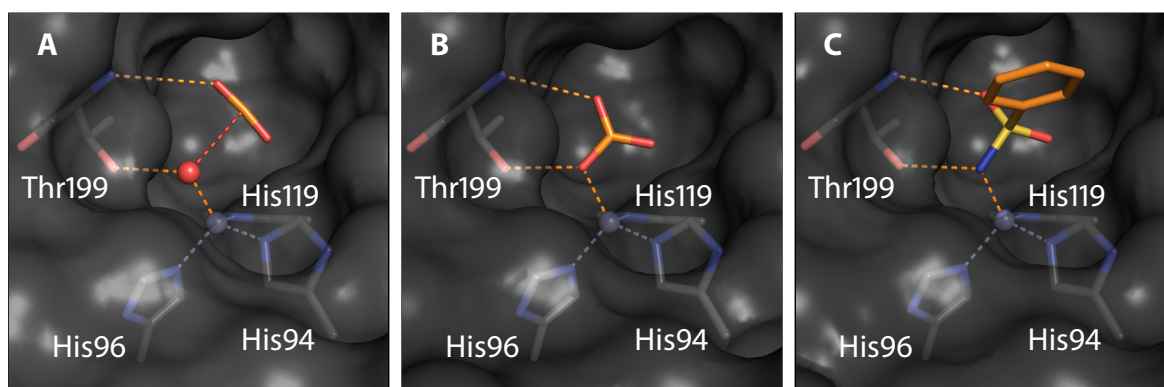
Fig. 6: CO_2 (2VVA,[139] **A**), bicarbonate (2VVB,[139] **B**) and benzenefsulfonamide (6GDC,[140] **C**) bound in the active site of hCAII. Carbon atoms are shown in orange, oxygen atoms in red, nitrogen atoms in blue and the sulfur atom in yellow. The Zn^{II} cofactor is shown as grey sphere, the coordinating hydroxide ion/water molecule as red sphere. Dashed lines represent the coordinating bonds between the Zn^{II} cofactor (grey sphere) and the side chains of His94, His96 and 119 (grey), coordinating bonds between the Zn^{II} cofactor and the respective ligand as well as hydrogen bonds between ligand and Thr199 (orange) and the bond to be formed between the Zn^{II} bound water molecule/hydroxide ion (red sphere) and CO_2 (red). Unless otherwise stated, this coloring code applies to all images.

The sulfonamide group mimics the geometry and polar interaction pattern of a bicarbonate ion. In addition to the coordination to the Zn^{II} cofactor, the coordinating oxygen or nitrogen atom donates a hydrogen bond to the side chain of Thr199 while an oxygen atom of either molecule accepts a hydrogen bond from the nitrogen atom of said protein residue.[129] A determining feature for the binding of arylsulfonamides is the additional hydrophobic interaction between the aromatic core of the inhibitor and the side chain of Leu198.[141]

1.7.1.1 The General Structure of Carbonic Anhydrase Inhibitors

The general composition of classical sulfonamidic CAIs comprises a primary sulfonamide head group, an aromatic ring, a tail region and a secondary recognition element (SRE).[129] *Grzybowski et al.* designed ligand **1** (Fig. 7) which features all four design elements of the general structure of CAIs for the binding site of hCAII in a computer-assisted approach and showed a K_d of 30 pM in

a competitive binding assay with dansylamide as reference inhibitor.[142] The variations of the ring
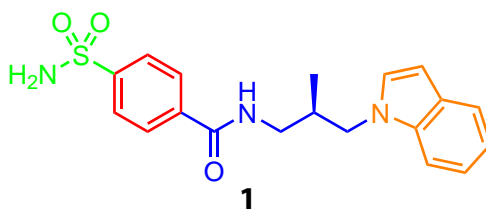
Fig. 7: CAI **1** developed *in silico* containing all four classical design elements of BSA CAIs. The sulfonamide head group is shown in green, the aromatic ring in red, the tail in blue and the SRE in orange.

and tail portion, termed 'ring approach' and 'tail approach', have led to a multitude of CAIs.[127] The ring approach entailed a wide variety of different ring systems that were investigated for their usefulness in CAIs and refuted the former dogma that aromatic rings are essential for the inhibitory performance of CAIs.[143] This became especially useful, as it was found that CAIs incorporating saturated ring systems strongly inhibited the supposedly sulfonamide-resistant human isoform III. Ultimately, the ring approach led to the development of the marketed drugs dorzolamide and brinzolamide.[143] However, a large number of the resulting compounds did show inhibitory power against the ubiquitously expressed isoform II in the human body, but were also potent allergens, and thereby unsuited for the use as therapeutics.[143] The tail approach, introcuded by *Supuran*, increased the chemical variability of CAIs in a convenient way, seeing that the attachment of electrophilic functionalities to nucleophilic moieties on the ring scaffold is chemically simpler than the construction of ring systems. The tail approach aims at fine-tuning physicochemical properties of already known CAIs by the introduction of residues that are ionized under physiological conditions and thereby improve solubility as well as therapeutic performance and reduce side effects.[143]

1.7.1.2 Thermodynamics of the Carbonic Anhydrase II – Benzenesulfonamide System

Taking into account the relatively high rigidity of Carbonic Anhydrase II (CAII), *Krishnamurthy et al.* sought to separate the individual contributions of individual parts of CAIs to the overall free energy of binding.[129] Based on the assumption, that the electrostatic interactions mainly depend

on the ligand's pK_a and that the dispersive interactions between aromatic ring and protein are merely a function of hydrophobicity (measured by the logarithmic partition coefficient (log$P$)), semiquantitative QSAR were established for a series of diversely fluorinated BSAs.[144] The findings allowed the conclusion, that electrostatic interactions are the main driving force for binding with a contribution of 65 % by the interaction between the Zn^{II} cofactor and the sulfonamide anion, and 10 % added by the hydrogen network formed between the sulfonamide inhibitor and Thr199 in the active site.[144] Hydrophobic interactions between the aromatic ring and the protein account for the residual 25 %.[144] Despite the relative physicochemical simplicity of the concept of acidity, it entails many effects that are difficult to separate: An increased acidity of the ligand provides for a higher fraction of anionic sulfonamide for binding, but also reduces the anion's *Lewis*, which counterbalances the former effect.[144] Thus, it was argued that the optimal pK_a value for a BSA CAI is equal to the pH value of the surrounding medium.[144] In addition, the authors also issued a note of caution with respect to the general applicability of these values, seeing that the investigated ligand molecules were comparably small, and thus should not be considered a valid model for larger BSA ligands which, e.g., bear a large hydrophobic tail.[144] In general, an enhanced ability to form hydrophobic contacts with the protein, usually achieved by the attachment of hydrophobic tails, correlates with an increase in affinity. However, a non-optimally designed hydrophobic tail can frustrate an increase in affinity due to steric repulsion.[129] Apart from either favorable or unfavorable contributions, non-optimally designed tails can furthermore entail EEC (subsection 1.2.4), as demonstrated by *Krishnamurthy et al.* by the use of BSA inhibitors with oligoglycine, -sarcosine and -ethylene glycol tails.[68] Microcalorimetry revealed indifference of ΔG to the tail lengths as enthalpic and entropic contributions cancelled each other out.[68] A later study by *Mecinović et al.* demonstrated that alkyl- and perfluoroalkyl chains did indeed show an increase in ΔG with increasing tail length.[145] Interestingly, ligands investigated in the latter study show shape complementarity between the tail and the hydrophobic wall. In combination with the work by *Krishnamurthy et al.* this gives rise to the assumption, that not merely the

accumulation of hydrophobic matter on a hydrophobic surface is of importance to binding, but that shape complementarity between the interacting entities plays a distinct role.

1.7.1.3 Kinetics of the Carbonic Anhydrase II - Benzenesulfonamide System

Not only has CAII been employed as a model system for the investigation of protein-ligand binding thermodynamics but also for the establishment of kinetic selectivity profiles of different BSAs for CA isoforms and the demonstration of the concept of intrinsic kinetics.[146,147] The latter aimes at the determination of kinetics without superimposed protonation effects and reveals the inherent kinetic properties of a ligand molecule.[147] Earlier studies aimed at elucidating the fundamental mechanics underlying the association and dissociation processes and thereby the determinants of k_{on} and k_{off}. In general, the following observations were made with respect to the kinetics governing the CA – BSA complex formation and disassembly:

a) k_{on} is pH dependent[137]

b) k_{off} is pH independent[137]

c) BSAs with large hydrophobic substituents can have faster association rates than smaller ligands.[129]

Especially the mechanism of association has been subject to debate for a long time. In general, it is necessary to differentiate observations a) and c) in the argument. With respect to the chemical species it was postulated, that association between CA and a BSA can occur *via* either of the following reactions.[129]

$$ArSO_2NH_2 + CA-Zn^{II}-OH_2^+ \rightleftharpoons ArSO_2NH-Zn^{II}-CA + H_2O + H^+ \qquad (12)$$

$$ArSO_2NH_2 + CA-Zn^{II}-OH \rightleftharpoons ArSO_2NH-Zn^{II}-CA + H_2O \qquad (13)$$

$$ArSO_2NH^- + CA-Zn^{II}-OH_2^+ \rightleftharpoons ArSO_2NH-Zn^{II}-CA + H_2O \qquad (14)$$

$$ArSO_2NH^- + CA-Zn^{II}-OH \rightleftharpoons ArSO_2NH-Zn^{II}-CA + OH^- \qquad (15)$$
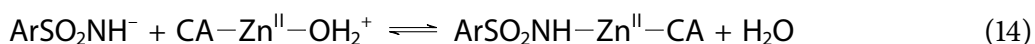
It was found that only reactions (13) and (14) can reproduce the experimental data based on the fractions of reactive species considered in the respective equation.[129] Already in the 1970s, *King et al.* argued, based on experimental evidence, that an encounter between the uncharged BSA and CA species is the likely premise for the association process. Later, although not experimentally evidenced, *Krishnamurthy et al.* assumed the charged species to be the intially reactive forms.[129,144,148,149] Based on the experiment-derived assumption that the initial approach occurs between the neutral species, it was argued that the mechanism of association must comprise more than one step, given the necessity to release one sulfonamidic proton and that, furthermore, a preequilibrium governed by hydrophobic interactions is likely to account for the increasing k_{on} along a series of BSAs with increasing hydrophobicities.[148,149] This is especially important to consider with respect to observation c). Based on the reasoning of *Taylor et al.*, *Gaspari et al.* confirmed the kinetic signatures of binding of a subset of increasingly hydrophobic compounds already investigated by *Taylor et al.* in a complementary approach comprising surface plasmon resonance (SPR), MX and MD simulations. The latter comprised a four-state model, that includes two intermediate states (reaction (16)).[150]

$$P + L \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} F \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} S \overset{k_3}{\longrightarrow} PL \qquad (16)$$
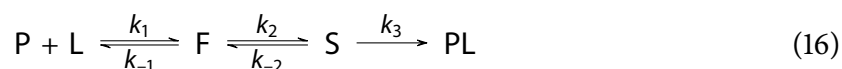
The hydrophobicity-based intermediate (F) represents a state, in which protein and ligand interact merely hydrophobically. It is then commuted into the advanced intermediate (S), which describes a polar coordination of the ligand with polar functionalities in the active site and the Zn^{II}-bound hydroxide ion prior to the final, crystallographically observable binding pose, into which it is converted with the final forward rate constant (k_3). This step furthermore includes the transfer of the sulfonamide proton to the hydroxide group, which supports reaction 14.[150] Unlike reaction (6), k_2 and k_{-2} here refer to the isomerization of the ligand's state rather than that of the receptor. With respect to the question of proton transfer, it was furthermore suggested that deprotonation of the sulfonamide function is unlikely to occur before approaching the active site,

given that the pK_a only changes negligibly along the path of association.[150] This raises a doubt about the theory of the interplay between the fraction of free sulfonamide anion and its *Lewis* basicity. Based on the insights provided by *Gaspari et al.*, the free species of ligand should not be of concern for the rate of interaction, seeing that initially the uncharged sulfonamide approaches the active site. While these findings paint a plausible picture of the association mechanism, they do not consider the circumstance, that it was shown by means of neutron diffraction, that the oxygen species coordinating the Zn^{II} cofactor is a water molecule in the crystalline state, even at pH 9.[151] Thus, the association mechanism, despite many efforts to elucidate it, is still not doubtlessly explicable, while the dissociation mechanism remains even more obscure.

2 The Search for a New Cryoprotectant

The storage of protein crystals and the execution of XRD experiments at cryogenic temperatures has a distinct impact on the outcome of XRD experiments.[152] An early perception of the influence of X-rays on protein crystals was, that the chemical matter within the crystal can be, to some extent, radiolyzed. The products of these reactions are free radicals and ionic species, which themselves are capable of initiating chemical reactions, that ultimately lead to chemical modifications of the crystallized molecules, and thereby putatively the disintegration of the crystal itself.[152] Measurement at cryogenic temperatures (often 100 K in practice) distinctly reduces the occurrence of radiolysis events and furthermore slows down the motion of their reaction products through the crystal.[152] An inherent problem of the freezing of protein crystals, usually grown in aqueous media, is the formation of ice. The density of water ice is lower than that of water, which leads to an increase in volume of water aggregates within the solvent channels.[152] This will either lead to a change of the crystal's lattice due to the displacement of its molecules, or may even cause a breaking of the crystal to release the mechanical strain imposed by ice formation.[152] Ice crystals formed on the surface of the protein crystal can have similar effects.[152] Furthermore, ice crystals themselves diffract X-rays, and thereby cause a powder diffraction pattern of rings, that will impede the analysis of *Bragg* reflections that happen to be located close to or underneath these rings.[152] A suitable cryoprotectant forestalls the formation of crystalline ice by promoting vitrifaction of the sample.[152] If the crystallization or soaking conditions do not already provide for sufficient cryoprotection, the choice of the cryoprotectant is of paramount importance for the success of the planned XRD experiment. There is, however, a putative disad-

vantage to the use of cryoprotectants, if crystals are to be soaked with small organic molecules. Given that many cryoprotectants are small organic molecules themselves, they can also bind in a protein's active site and thereby change, impede or even prevent the binding of the small molecule of interest. This is possibly the case with hCAII and glycerol. Glycerol is a common cryoprotectant for the hCAII–BSA system, and is found in numerous of such complexes in the PDB, e.g. in entry 5LL8 (Fig. 8).
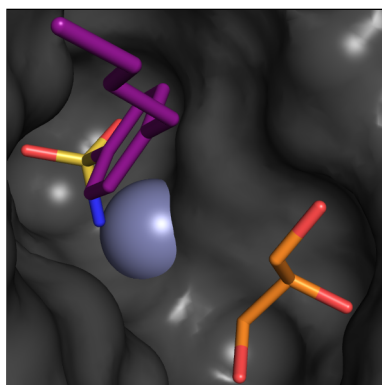
Fig. 8: Glycerol in the active site of hCAII in orange with an adjacent BSA ligand in purple.

To enable soaking experiments without the influence of glycerol, a new cryoprotectant was sought. Tab. 1 lists the investigated solutions.

Tab. 1: Investigated cryoprotectants.

Investigated compound	Content in final solution
1-amino-propan-2-ol[a]	30 % v/v
Betaine·HCl[a]	1.0 M
Glycine[a]	1.0 M
Sarcosine[a]	1.0 M
L-Proline[a]	3.0 M
D-Fructose (anhydrous)[a]	25 % w/v
D-Glucose (anhydrous)[a]	25 % w/v
D-sucrose (anhydrous)[a]	25 % w/v
D-Sorbitol (anhydrous)[a]	25 % w/v
Trisodium citrate[b]	1.6 M
Potassium sodium tartrate[b]	2.5 M

[a] Dissolved in a solution containing 3.0 M ammonium sulfate, 0.1 M TRIS at pH 7.8. [b] Dissolved in demineralized water.

Crystals of similar size and grown from the same batch of protein in 2.7 M ammonium sulfate, 0.1 M TRIS at pH 7.8 were incubated in the solutions listed in Tab. 1 for 30 min and flash-frozen in liquid nitrogen. Images were recorded on a MAR345 image plate detector with Cu-K_α radiation. Fig. 9 shows diffraction images for all tested conditions.
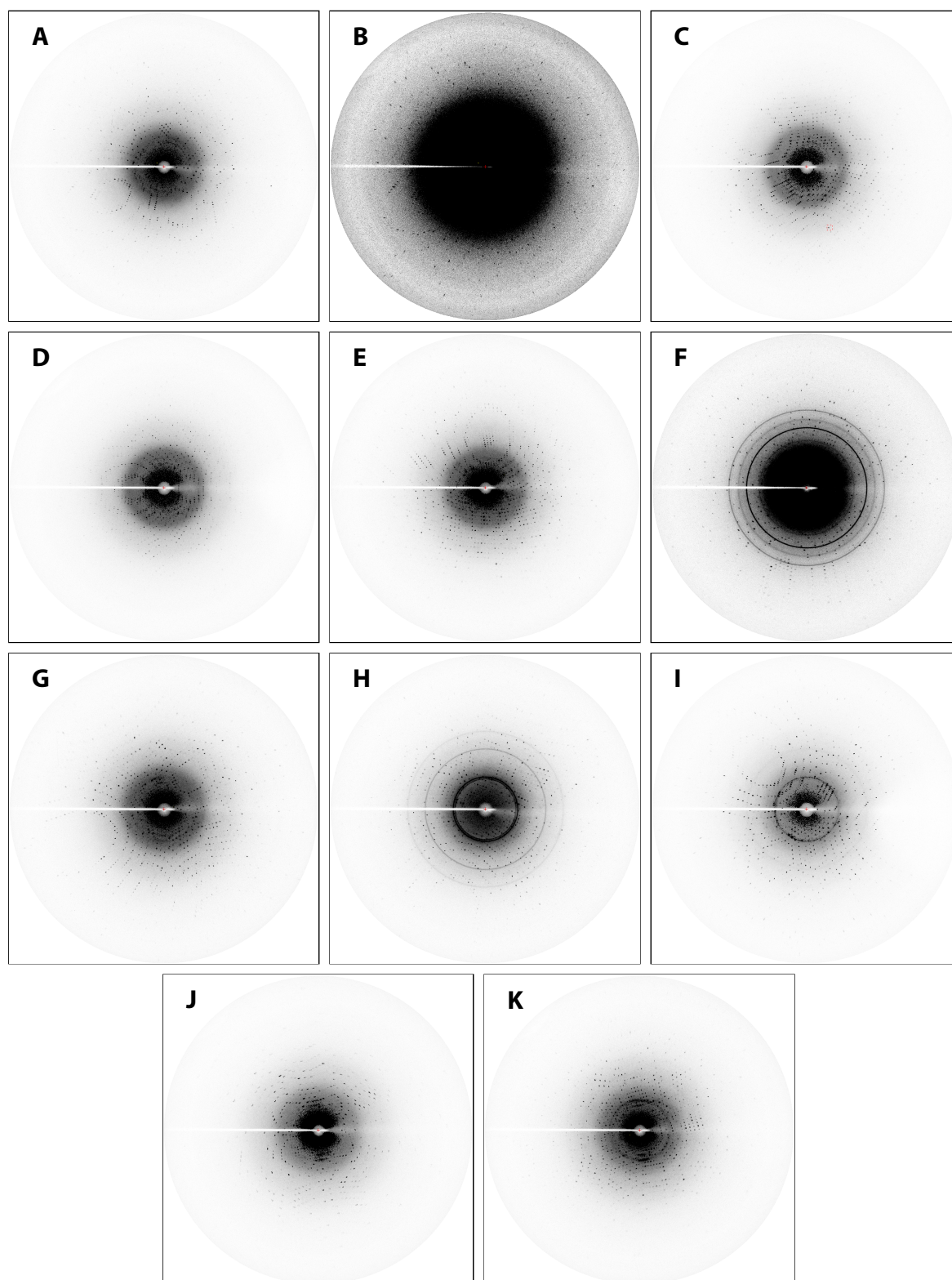
Fig. 9: Diffraction images of hCAII crystals incubated with different cryoprotectants. **A** *rac*-1-aminopropan-2-ol. **B** Betaine hydrochloride. **C** Glycine. **D** L-Proline. **E** Sarcosine. **F** D-Fructose. **G** D-Glucose. **H** D-Sucrose. **I** D-Sorbitol. **J** Trisodium citrate. **K** Potassium sodium tartrate.

Trisodium citrate solution caused streaked reflections and the incubation with D-fructose, aqueous potassium sodium tartrate, D-sorbitol and D-sucrose entailed ice rings. Based on visual inspection, many conditions yielded similarly strong diffraction, e.g. *rac*-1-aminopropan-2-ol and D-glucose. It was found, however, that the (*R*)-enantiomer of the former coordinates to the Zn^{II} cofactor (Fig. 10), similar to already reported CA-inhibiting polyamines spermine and spermidine.[153] D-glucose was not found near the active site, but only in three remote binding sites between crystal mates (Fig. 10), and was thus chosen as cryoprotectant for the preparation of most of the complexes reported herein.
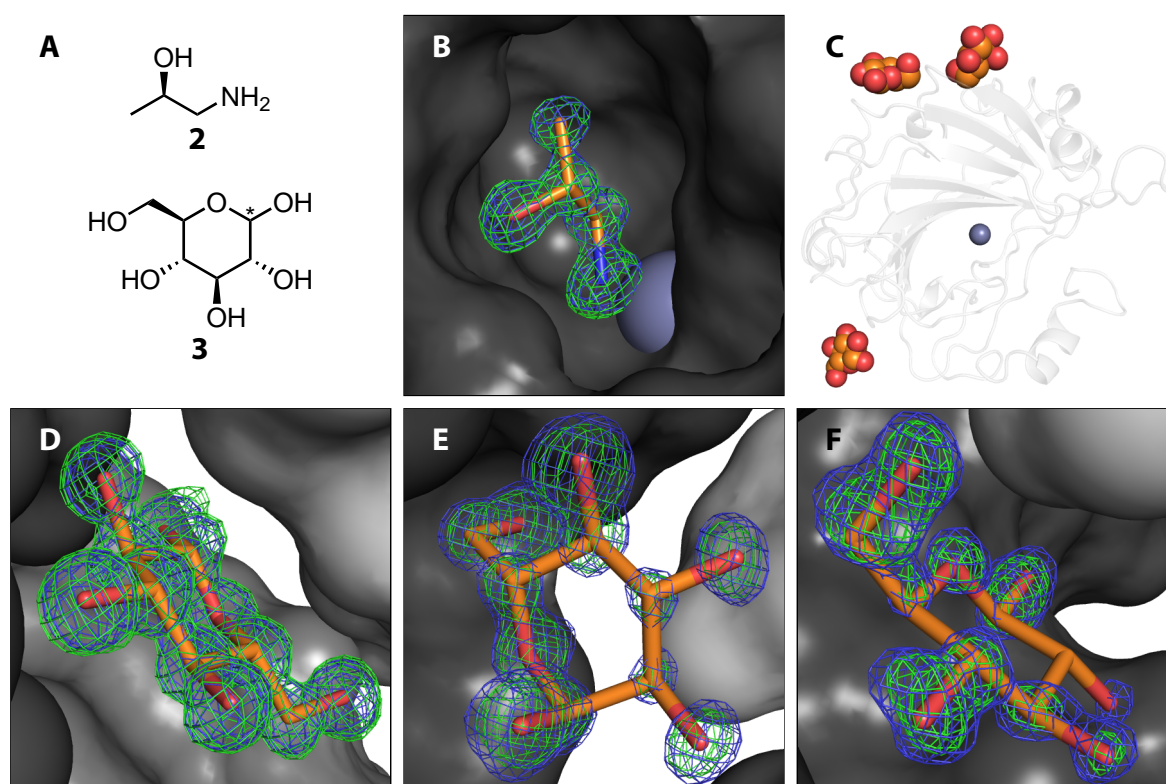
Fig. 10: (*R*)-1-aminopropan-2-ol (**2**) in the acitve site of hCAII and D-glucose (**3**) in the hCAII crystal packing. Omit composite electron density maps are shown at 1σ in blue and difference electron density maps at 3σ in green. **A** Valence bond formula of **2** and **3**. The anomeric carbon atom of the latter is marked with an asterisk. **B** (*R*)-1-aminopropan-2-ol (**2**) in the hCAII active site. **C** Overview of D-glucose binding sites on the hCAII asymmetric unit. **D** Highly populated binding site of β-D-glucose molecule, which interacts with the side chains of Lys39, Gln255 and Lys257. **E** α-D-glucose binds to the backbone carbonyl atom of Thr177. **F** β-D-glucose binds to the backbone carbonyl atom of Thr177.

3 Investigation of Varying Alkyl Tail Lengths and Geometries

A question that has not yet been answered with respect to the hCAII–BSA system is the influence of different geometries of the alkyl chain relative to the phenyl core. Therefore, *n*-alkyl and *n*-alkyloxy substituted BSAs **4** – **13** (Tab. 2) were investigated by means of MX and ITC to elucidate the influence of different tail geometries on the thermodynamics and kinetics of binding.

Tab. 2: Valence bond formulae of investigated *n*-alkyl and *n*-alkyloxy BSAs and the respective PDB entry of the deposited hCAII–ligand complex.

Compound	R	PDB	Compound	R	PDB
4	H	6GDC	**10**	Me	6I0W
5	Me	6GM9	**11**	Et	6I1U
6	Et	6HQX	**12**	Pr	6I2F
7	Pr	6HR3	**13**	Bu	6I3E
8	Bu	6HXD			
9	Pen	6SBH			

3.1 Crystallographic Results

The necessary prerequisite for the detailed investigation and comparison of the influence of the *n*-alkyl and *n*-alkyloxy tails is that the non-tail moieties of the molecules maintain the binding pose of the unsubstituted BSA **4**, which is shown in Fig. 11.
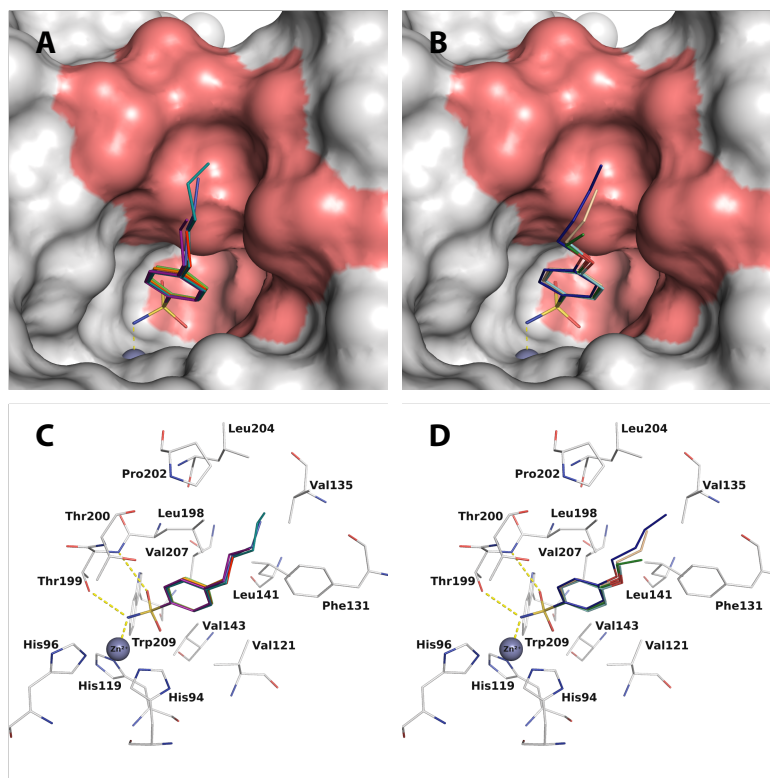
Fig. 11: A Superposition of *n*-alkyl derivatives (**4** light green, **5** orange, **6** red, **7** violet, **8** light blue, **9** cyan) in the hCAII active site with solvent excluded surface representation. Side chain atoms forming the hydrophobic patch are colored in light red. The Zn^{II} cofactor is shown as grey sphere. **B** Superposition of *n*-alkyloxy derivatives (**10** light cyan, **11** green, **12** beige, **13** blue) according to **A**. **C** *n*-alkyl derivatives in the active site in stick represension. Residues forming polar interactions with either ligands or Zn^{II} cofactor are shown. **D** *n*-alkyloxy derivatives in the active site in stick represension. Residues forming polar interactions with either ligands or Zn^{II} cofactor are shown.

Omit electron density maps of all ligands in Fig. 11 are shown in Fig. 12. All ligands refined to an occupancy of 1.0. The alkyl tails of compounds **4 – 8** are accommodated on the hydrophobic wall in an all-*trans* conformation. Ligand **9**, deviantly, adopts a *gauche* conformation. The change in composition of both molecular series entails a distinct structural effect: Whereas the torsion angle between the phenyl-bound ethylene unit and the aromatic ring approximates 60° for *n*-alkyl derivatives, it only slightly diverts from 0° for the *n*-alkyloxy derivatives.
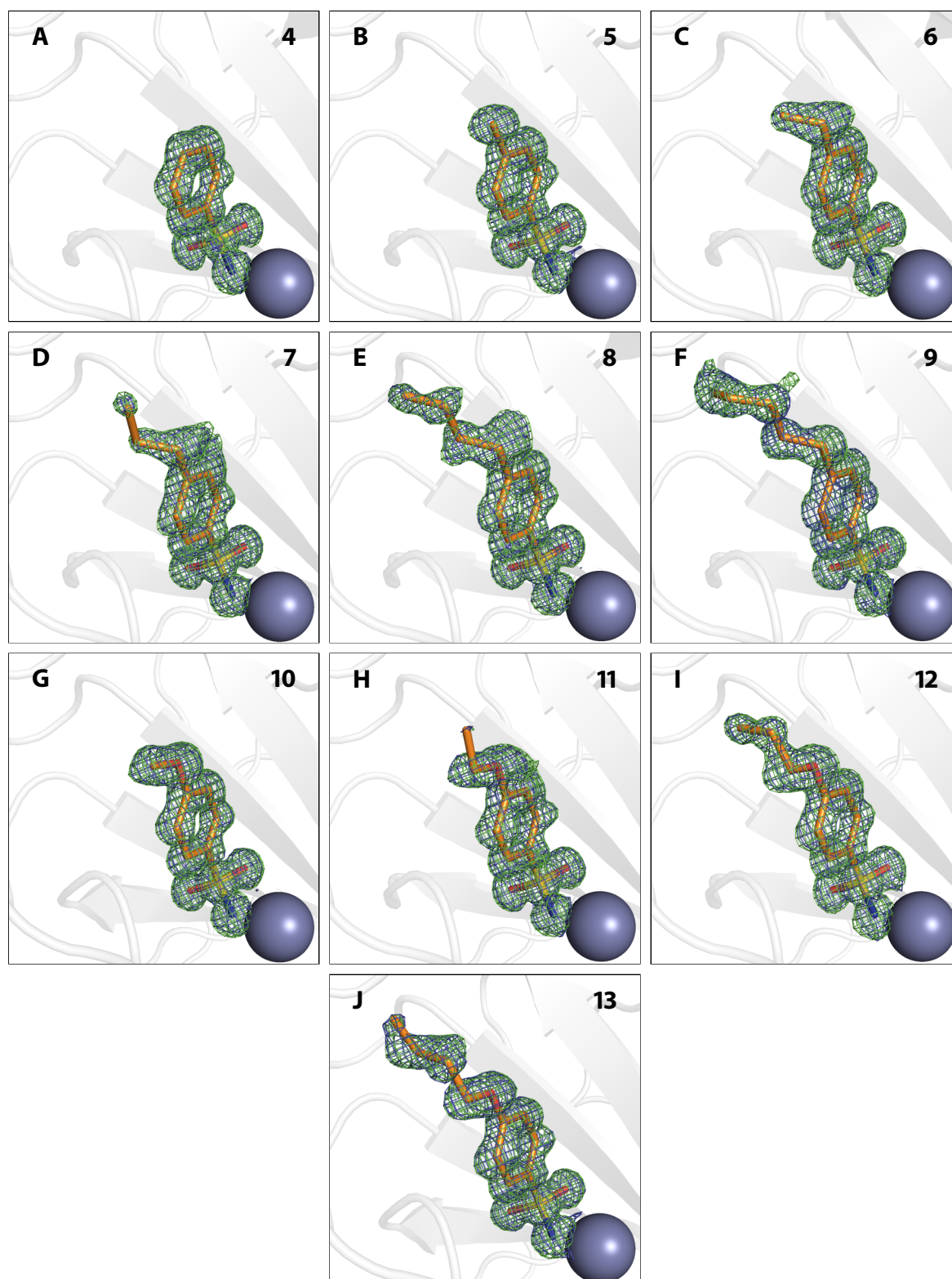
Fig. 12: *n*-alkyl and *n*-alkyloxy substituted BSAs bound in the active site of hCAII. **4** (**A**), **5** (**B**), **6** (**C**), **7** (**D**), **8** (**E**), **9** (**F**), **10** (**G**), **11** (**H**), **12** (**I**), **13** (**J**). Omit 2mF_o–DF_c and mF_o–DF_c maps are contoured in blue at 1 σ and green at 3 σ, respectively.

The structural difference is a consequence of steric and electronic effects of the aliphatic carbon or oxygen atom bound to the phenyl moiety. It is thus only possible for molecules **12** and **13** to interact extensively with the hydrophobic patch, whereas this is possible 'earlier' within the *n*-alkyl series for specimens with more than one tail atom (**6 – 9**). A quantification of these characteristics was approached by the calculation of solvent excluded surface areas (SESAs), which are better suited for the correlation with thermodynamic paramters than solvent accessible surface areas (SASAs).[154] Solvent excluded surface areas were estimated by equation (17).

$$SESA = \frac{SESA_P + SESA_L - SESA_{PL}}{2} \tag{17}$$

Tab. 3 shows SESAs of all compounds in their bound conformation, calculated contact area with the protein, the increments in contact areas of ligands with successive chain lengths and the ratio of contact area to the surface of the respective ligand.

<div align="center">Tab. 3: Investigated <i>n</i>-alkyl and <i>n</i>-alkyloxy BSAs.</div>

Compound	Chain length	Ligand / Å^2	Contact / Å^2	Increment^a	Ratio^b
4	n = 0	151.3	127.9		0.85
5	n = 1	162.6	130.9	3.0	0.81
6	n = 2	178.8	139.3	8.4	0.78
7	n = 3	196.0	144.4	5.1	0.74
8	n = 4	213.2	160.2	15.8	0.75
9	n = 5	229.2	168.8	8.6	0.74
10	n = 1	178.0	137.1		0.77
11	n = 2	196.1	140.6	3.5	0.72
12	n = 3	213.1	162.4	21.8	0.76
13	n = 4	224.1	165.9	3.5	0.74

^{a} Increments refer to the difference between ligands of successive lengths. ^{b} The Ratio column provides the fraction of the contact area of a ligand with the protein in comparison to the ligand's surface area.

Notably, ligands **7** and **9** exert a lower efficiency at burying surface area in the *n*-alkyl series. The geometric setting of the *n*-alkyloxy series caused by the conjugation between an oxygen lone-pair and the aromatic π-system entails a pairwise development of contact area within the ether series. The interface area is similar for compounds **10** and **11** on the one hand, and for **12** and **13** on the other, whereas there is a pronounced difference between ligands **11** and **12**. The

notion to bury surface area recedes within the *n*-alkyl series until compound **7**, increases for **8** and recedes again for **9**.

3.2 Isothermal Titration Calorimetry

A novel titration protocol was applied to enable reliable extraction of both thermodynamic and kinetic information from ITC data. It is based on the variation of injection volumes to shift injections to the first half of an ITC thermogram (before the inflection point) and a reduction in the number of injections in the saturation region.

3.2.1 Thermodynamic Results

For binding of the above ligand series to hCAII, the following thermodynamic signatures were found (Fig. 13).
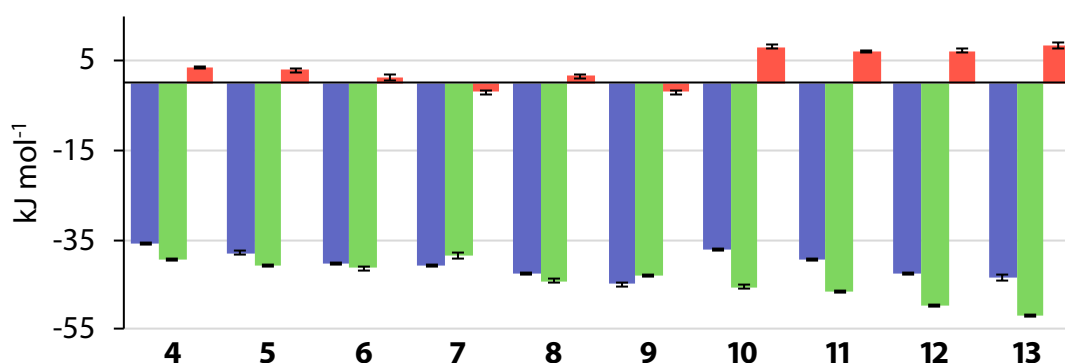
Fig. 13: Globally fitted values of $\Delta H°$(green), values for $\Delta G°$ (blue) calculated from the globally determined values of $K_a$and values for $-T\Delta S°$(red) calculated from the former two. Standard errors of three individual measurements are given to present experimental uncertainty. As **9** was characterized in a displacement setup against 4-carboxybenzenesulfonamide (**20**), error bars resemble the fitting error for the global fitting of $\Delta H°$ and the propagated errors of $\Delta G°$ and $-T\Delta S°$.

3.2.1.1 *n*-Alkyl Series

A linear increase in the free energy of binding can be observed from **4** to **6**. Notably, **7** seems to represent a 'pit' of EEC:[40,62] $\Delta H°$ recedes from the trend of linear growth, whereas the decrease of the unfavorable entropic contribution culminates in a virtually favorable contribution and

thereby prevents a loss in $\Delta G°$ compared to **6**. Compound **8** shows a renewed increase in $\Delta G°$, which results from a distinct increase in the enthalpic contribution, as the entropic contribution becomes distinctly unfavorable again, opposed to the increasingly favorable enthalpic and decreasingly unfavorable entropic developments for ligands **4 – 6**. Notably, compound **9** exerts a similar thermodynamic behavior as **7** compared to the respective preceding compound, as its binding is enthalpically weaker, while the entropic term becomes favorable. This coincides with the same efficiency at burying surface area of **7** and **9**. However, compound **9** still gains in the overall free energy of binding, which is easily explained by the mere burial of surface area. However, residual mobilities derived from high-resolution MX models can partly shed light on the issue. Fig. 14 shows the numbered model of **9** taken from the respective MX model and plots of normalized B-factors as function of the atom number for ligands in the n-alkyl series.
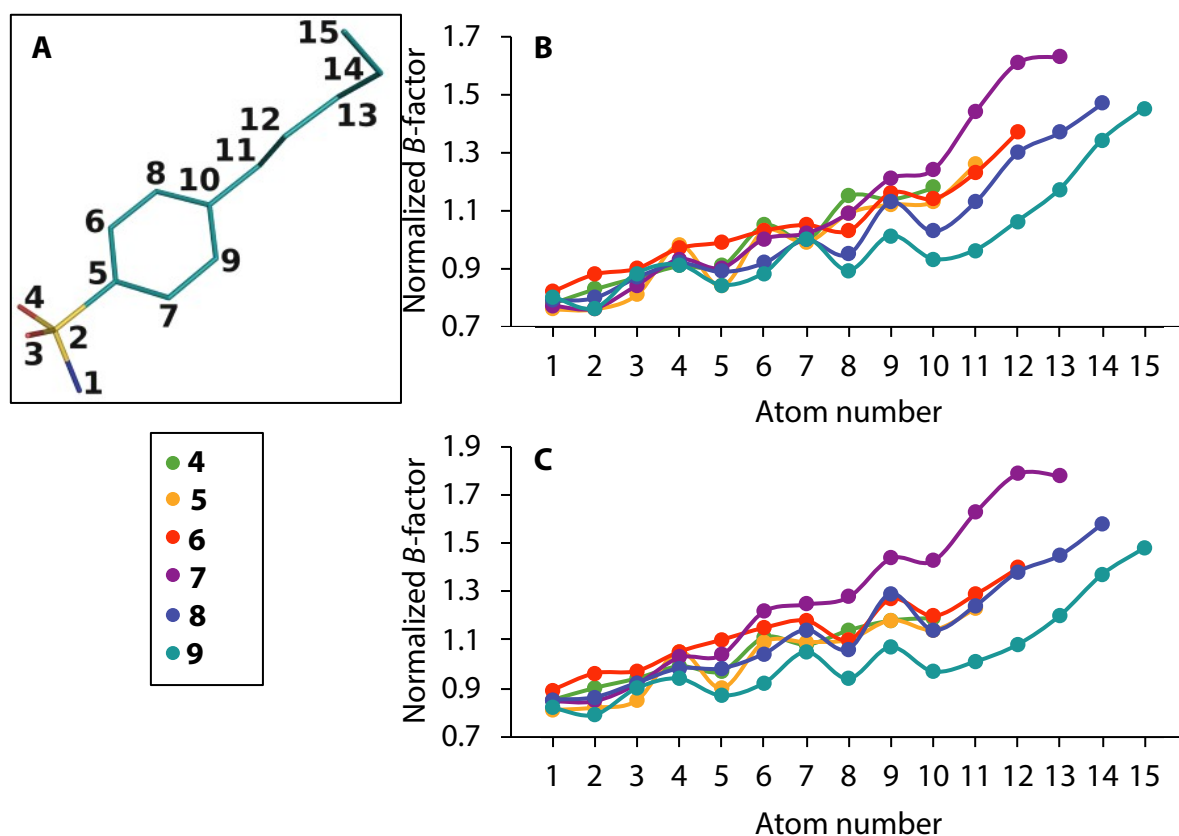
Fig. 14: A Numbered model of **9**. **B** Atomic B-factors normalized on the average B-factor of all atoms (including protein and other ligand atoms) within a radius of 6.5 Å around the respective atom. **C** Like **B**, but only taking into account protein atoms for normalization.

4 – 6 are similar in this respect, pronounced differences become apparent for compounds **7** and **9**. Notably, the relation between these ligands is the same with both methods applied for normalization. Ligand **7** is the ligand with the highest residual mobility, which coincides with its entropically favorable contribution to binding, whereas the loosening of the interface area is reflected in the receding enthalpic term. Intuitively, the renewed entropic loss and enthalpic gain of **8** can be seen as a result of its stronger immobilization. Importantly, this reasoning does not apply to the thermodynamic development from **8** to **9**, as **9** is the most strongly immobilized ligand in the series. The binding of **9**, however, entails the formation of a *gauche* conformation about the bond between atoms 13 and 14, which appears to be necessary for an optimal burial of surface area, but simultaneously counterbalances the enthalpic gain resulting thereof. The thermodynamic similarity of **7** and **9** thus seems to have different origins due to the changing topography of the protein surface.

3.2.1.2 *n*-Alkyloxy Series

As with the pairwise development of interface areas, also a pairwise development of the thermodynamic signatures was observed. The difference between **11** and **12** ($\Delta\Delta G° = -3.1 \, \text{kJ mol}^{-1}$) is larger than either of the differences for the pairs **10/11** ($\Delta\Delta G° = -2.2 \, \text{kJ mol}^{-1}$) and **12/13** ($\Delta\Delta G° = -1.0 \, \text{kJ mol}^{-1}$). A similar development was found for the difference in $\Delta H°$. It differs by smaller amounts between the pairs **10/11** ($\Delta\Delta G° = -2.0 \, \text{kJ mol}^{-1}$) and **12/13** ($\Delta\Delta G° = -1.0 \, \text{kJ mol}^{-1}$) than between compounds **11** and **12** ($\Delta\Delta G° = -3.6 \, \text{kJ mol}^{-1}$), which is reflected in the larger difference in interface areas of compounds compounds **11** and **12** (Tab. 3). Opposed to the *n*-alkyl series, the interface area formed between ligands and hCAII does not correlate with the entropic development for the *n*-alkyloxy derivatives, as **13** is entropically more heavily penalized than **12**. In this case, however, it needs to be considered, that **13** is structurally different in so far, as it is the only compound whose last chain atom protrudes from the binding pocket, a circumstance which might give rise to this finding. Interestingly, also residual mobilities reflect the pairwise geometric development already explained (Fig. 14). The tails of the larger

compound of each pair (**11** and **13** show a higher residual mobility than those of the respective smaller compounds, although a larger amount of surface area is buried by the larger ligands. This resembles the relationship between compounds **7** and **5/6** and contrasts that between molecules **7** and **8/9**. Compounds **7**, **11** and **13** all orient the terminal methyl group toward the entrance of the binding site. Opposed to this arrangement, all other compounds except for the shortest tail-bearing compounds in each series (**5** and **10**) accommodate the terminal ethyl moiety on the hydrophobic wall in a vertical fashion and, despite a smaller amount of buried surface area, exert a reduced residual mobility.
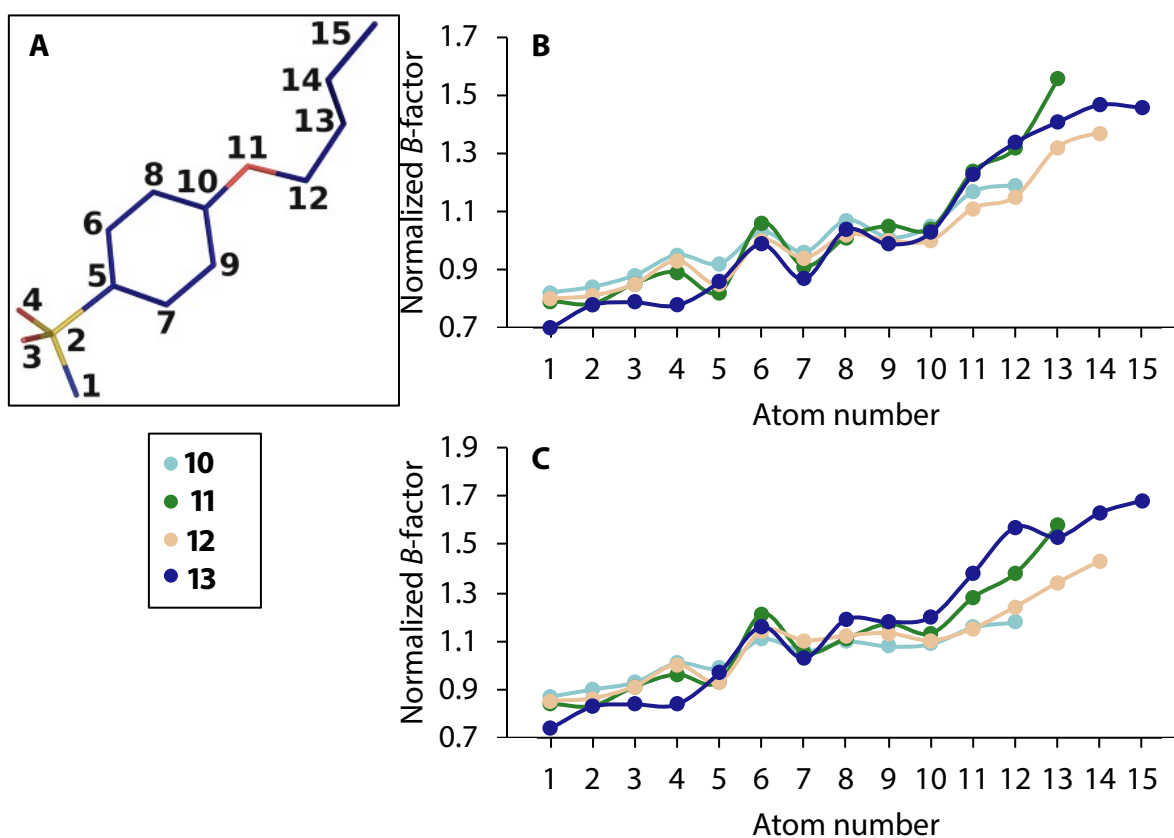
Fig. 15: *B*-factor analysis for the *n*-alkyloxy series analogous to Fig. 14.

Interestingly, both series differ overall in the partitioning of Δ*G*° into Δ*H*° and Δ*S*°. Binding of compounds in the *n*-alkyl is enthalpically distinctly stronger than binding of compounds in the *n*-alkyl series, but is also more heavily penalized entropically. A reason for this might the arrangement of water molecules in surrounding the complex. Water molecules are hardly

discernible above the active site, which renders an argument based on the formation of water networks hypothetical in this case. Desolvation of small-molecule inhibitor and protein is a process that must necessarily occur before the actual binding between both, but is difficult to grasp experimentally. Fig. 16 depicts the development of the calculated solvation free energy (ΔG_{solv}, bearing the inverse sign of the desolvation process) for both series, computed quantum mechanically using two different basis sets.
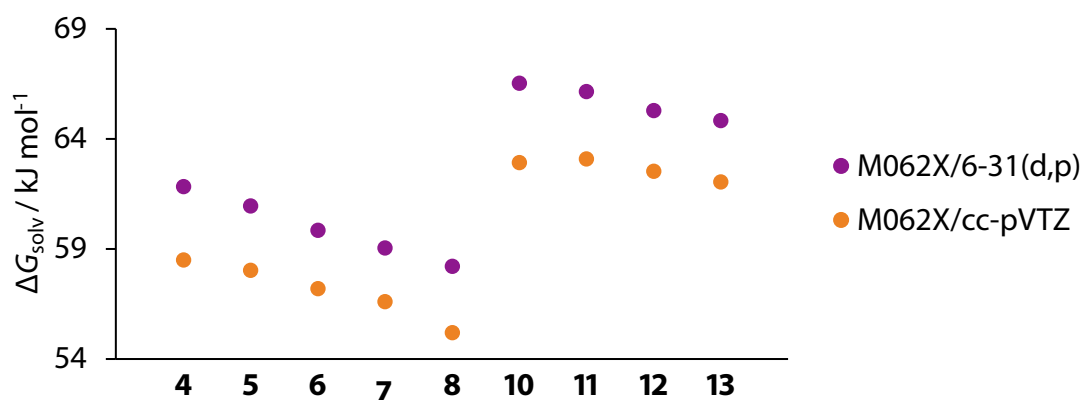
Fig. 16: Solvation free energies for both molecular series (not including compound **9**).

Intuitively, ΔG_{solv} recedes with increasing hydrophobicity and reflects the increasingly favorable affinity for the *n*-alkyl derivatives. The special position of **7** within the series, however, cannot be explained. Neither can solvation (and thereby desolvation) account for the pairwise development within the *n*-alkyloxy series, as ΔG_{solv} also develops in a linear fashion in this series. However, the distinctly larger values of ΔG_{solv} for the *n*-alkyloxy compounds may conribute to the enthalpic and entropic discrepancies between both series.

As the changing physicochemical characteristic in both series is an increasing hydrophobic surface, it can be reasonably assumed that the hydrophobic effect is responsible for the thermo-dynamic findings.[62] Given the paired enthalpic development in the *n*-alkyloxy series, however, which correlates with the paired development in surface burial, it seems more likely that the ther-modynamic findings are attributable to the dispersive interactions between ligands and protein, rather than the hydrophobic effect. It is therefore also more likely, that the pit of compensation

in the *n*-alkyl series is a consequence of the nature of the ligands, rather than a consequence of the hydrophobic effect.

As shown in Tab. 3, the ligands **7** to **9** bury successively more surface area on the protein, which becomes visible in the close-up of the tail regions of these compounds (Fig. 17).
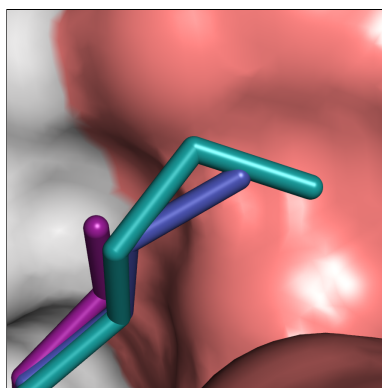
Fig. 17: Close-up of the tail regions of compounds **7** (violet), **8** (light blue) and **9** (cyan).

It can be inferred, that the pit of compensation seems to be reflected in the residual mobilities of atoms. **7** shows distinctly larger normalized *B*-factors than any of the preceding *n*-alkyl derivatives, which can be seen as a hint to the origin of the compensatory thermodynamic behavior and the favorable entropic contribution. It is furthermore worth noting, that the tails of **8** and **9** both show decreased residual mobility compared to **7**. This raises the assumption, as for the *n*-alkyloxy derivatives, that the transition from three to four tail atoms resembles a partitioning into two groups of molecules within the *n*-alkyl series, considering the ability to bury surface area. Contrarily to the above, a correlation between the efficiency at surface burial and decreasing entropic contribution, does not hold in this case. The fashion, in which the tail is accommodated on the hydrophobic patch, is changed after compound **8**. Compound **9** deviates from its predecessors in the series in that it adopts a *gauche* conformation instead of an all-*trans* conformation, which seems to be reflected in the thermodynamic data. The again favorable entropic contribution of **9**, however, cannot be easily explained.

3.2.2 Kinetic Results

With the method of kinITC it was possible to extract the kinetic parameters for both molecular series shown in Fig. 18.
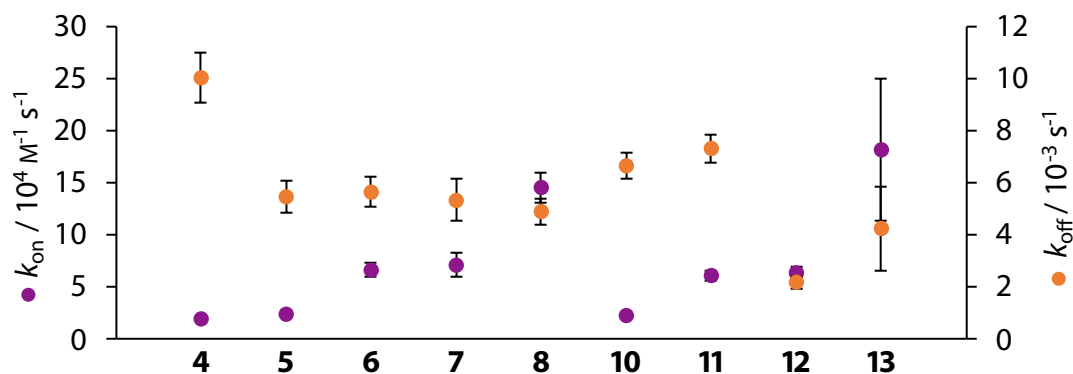
Fig. 18: Kinetic data for both molecular series except compound **9** recorded by kinITC. Error bars represent the standard error for three measurements.

3.2.2.1 *n*-Alkyl Series

Association rates become increasingly higher with growing chain lengths for compounds **4 – 8**. However, pairs of ligands with similar k_{on} values were observed (**4/5**, **6/7**). Dissociation rates develop differently. Compound **4** differs distinctly from **5**, which in turn is not discernible from the succeeding molecules in the series. Compound **9** was characterized in a displacement setup, which does not allow for the extraction of kinetic information.

3.2.2.2 *n*-Alkyloxy Series

An increase in the association rates was found across the series, while the dissociation rates develop in a pairwise recession. Compound **10** and **11** dissociate distinctly faster than compounds **12** and **13**, which follows the pairwise trend in which these compounds bury surface area on the protein.

A revealing comparison is that between **8** and **12** as both compounds have the same tail length and bind to hCAII with virtually the same free energy of binding of $\Delta G° = -42.6\,\text{kJ}\,\text{mol}^{-1}$.

However, the kinetic signatures of both compounds differ distinctly. **8** binds with a higher velocity (k_{on} = 1.46 · 10⁵ M⁻¹ s⁻¹) than **12** (k_{on} = 6.40 · 10⁴ M⁻¹ s⁻¹), but the latter is the more slowly dissociating compound (k_{off} = 2.2 · 10⁻³ M⁻¹ s⁻¹) compared to the former (k_{off} = 4.9 · 10⁻³ M⁻¹ s⁻¹). The rationale of *Gaspar et al.*for the hCAII–BSA system, that increased hydrophobicity accelerates the association step between ligand and protein, can be reasonably applied in this case, assuming that the ether oxygen atom renders **12** more polar than **8**.[150] The reason for the reduced dissociation rate of **12** can be explained by the a comparison of the burial of surface area. At almost the same surface area, **8** buries less surface than **12** due to an increased distance between the terminal methyl group and the hydrophobic wall, which is depicted in Fig. 19
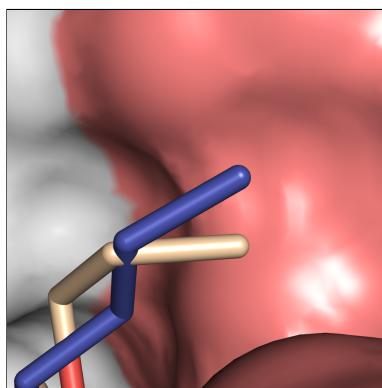
Fig. 19: Close-up of the tail regions of compounds **8** (light blue), and **12** (beige).

The increased proximity between the tail of **12** and the hydrophobic wall compared to **8** provides for stronger hydrophobic interactions and thereby impedes the dissociation, rendering **12** a more 'shape-efficient' binder than **8**. This circumstance furthermore demonstrates, that a minute chemical alteration can be used to shed light on the determinants of binding kinetics. Interestingly, the increasingly less unfavorable desolvation for both series agrees with the continous increase in k_{on}. Desolvation is a necessary prerequisite to complex formation between protein and ligand, and needs to be taken into account as contributing factor. This furthermore supports the mechanism of association elucidated by *Gaspari et al.*(subsubsection 1.7.1.3).[150]

3.2.3 Comparison with Surface Plasmon Resonance

A noteworthy observation in the course of this work is the observation, that the kinetic values of the *n*-alkyl series show the same trends as the data recorded by SPR for compounds **4 – 7** in an earlier study by *Gaspari et al.*(Fig. 20).
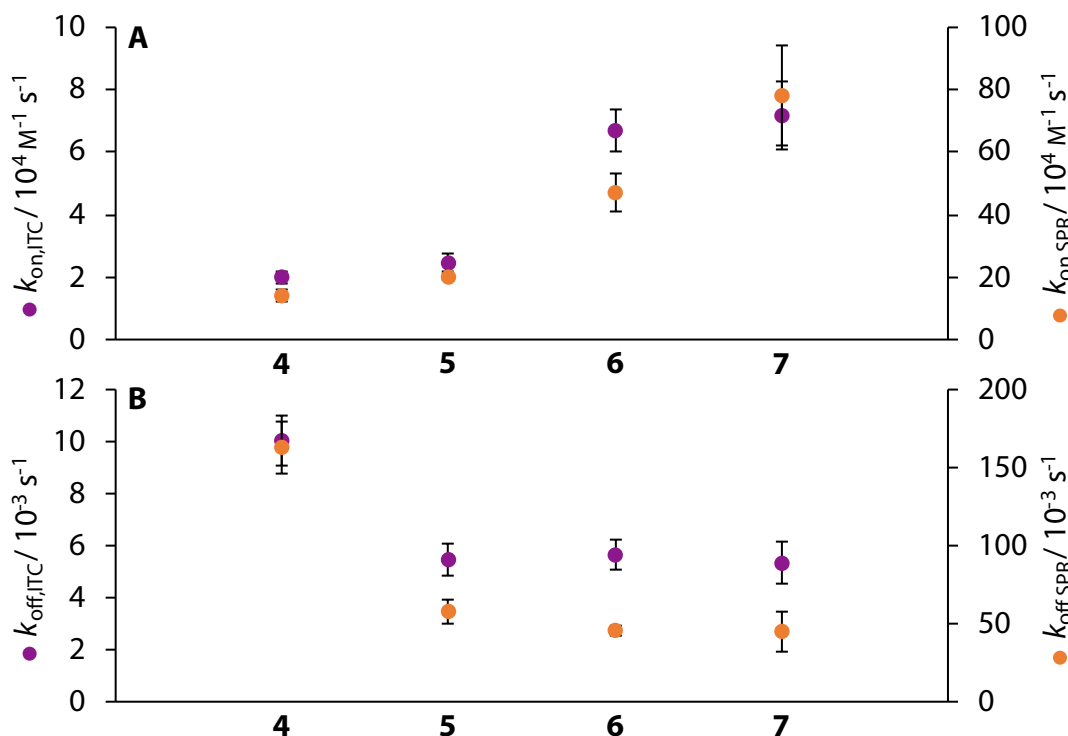
Fig. 20: Comparison of kinetic data from ITC experiments (left ordinates) with SPR data from an earlier study (right ordinates).

Despite the resemblance of trends, the values for k_{on} and k_{off} obtained from ITC experiments are approximately one order of magnitude lower than those derived from SPR experiments. It can be reasonably assumed, that protein molecules in solution show a difference in ligand binding behavior compared to protein molecules that are immobilized on a chip as in the case of SPR. Molecules in solution are unhindered in their mobility. On an SPR chip, different orientations fo the opening of the active site entrance relative to the immobilization matrix can be conceived. Depending on the amino acid residue that participates in immobilization by chemical coupling, the binding site may be oriented either toward the matrix or the solvent. Either 'specimen' can be assumed to show different ligand binding behavior. SPR thus delivers information about an

'average' binding site. Furthermore, ITC is an equilibrium-based method, whereas SPR records data based on a steady state in a flow chamber. Despite these putatively crucial differences between both methods, the reason for the finding that ITC records lower kinetic rate constants for these series of molecules must remain unresolved herein.

3.2.4 Further Investigated Molecules

In addition to the characterization the compounds discussed in this chapter, additional compounds were investigated both crystallographically and by means of ITC. Within the *n*-alkyl series, these compounds were the hexyl **14**, 6SBL), heptyl (**15**), ethylhydroxy (**16**, 6RG5) and butylhydroxy (**17**, 6SEY) derivatives. Pentyl (**18**, 6SBM) and hexyl (**19**) derivatives were excluded from the *n*-alkyloxy series. Fig. 21 shows the crystallographic models of hCAII in complex with these compounds.
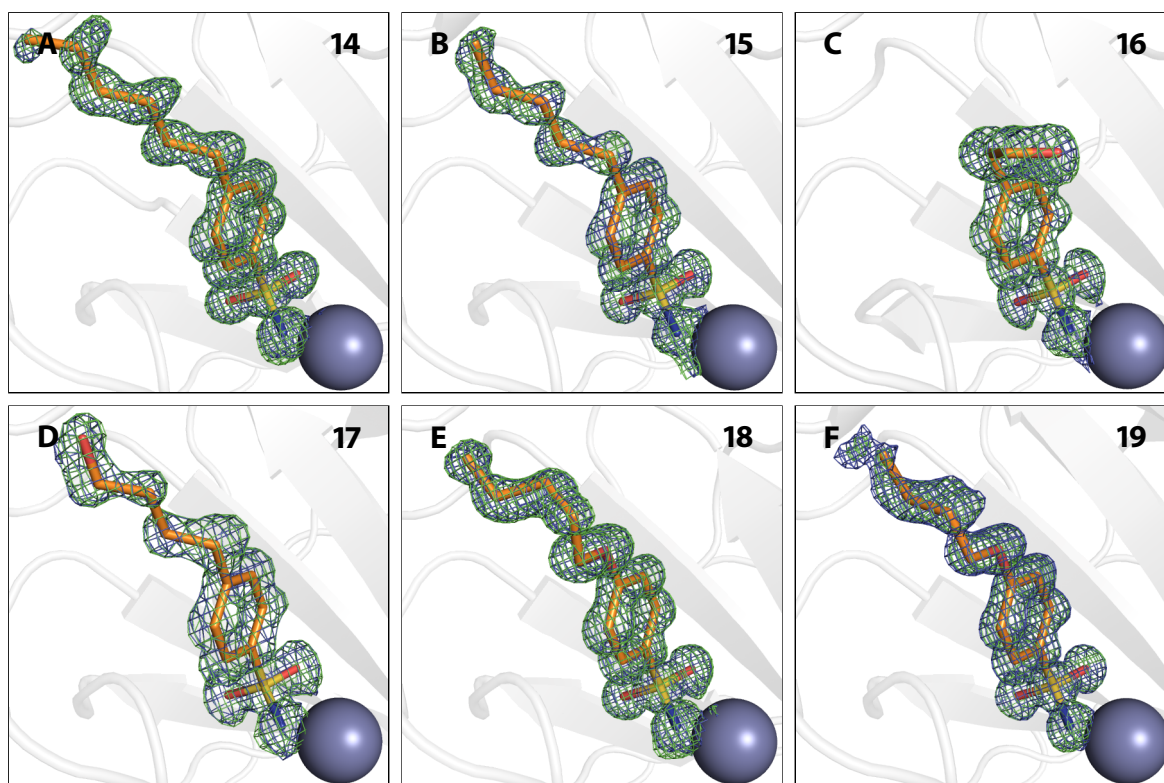
Fig. 21: Further *n*-alkyl and *n*-alkyloxy BSA CAIs **14** (**A**), **15** (**B**), **16** (**C**), **17** (**D**), **18** (**E**) and **19** (**F**) with omit electron density maps.

Models of hCAII in complex with compounds **14**, **16**, **17** and **18** were deposited in the PDB. Complexes hCAII–**15** and hCAII–**19** were not deposited, as the last two tail atoms of either ligand were not resolved in the electron density and could not be modeled (Fig. 22).
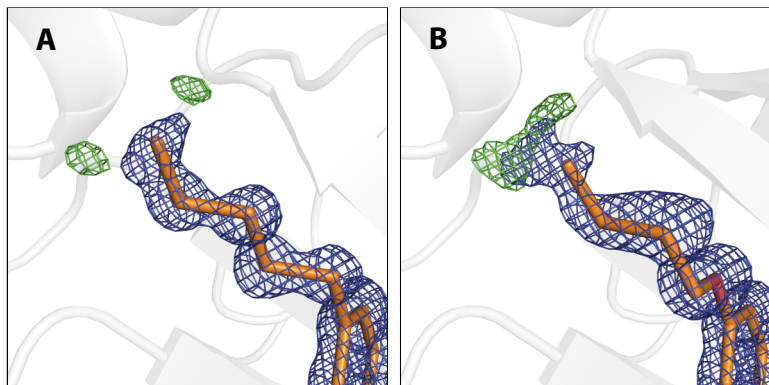
Fig. 22: Close-ups of the tail region of compounds **15** (**A**) and **19** (**B**) with $2mF_o - DF_c$ density at $1\,\sigma$ in blue and $mF_o - DF_c$ density in green at $3\,\sigma$ after refinement. The latter indicates one further tail atom for each ligand, but is to weak to be modeled.

ITC measurements revealed a comparably high level of noise and occurrence of artifacts for ligands **16** and **17**, which is especially detrimental to kinetic analysis (Fig. 23). Compounds **14**, **15**, **18** and **19** had to be measured in a displacement setup due to their high potencies. However, analysis of the data showed comparably high uncertainties that did not allow for structure-correlated intrepretation. Furthermore dilution titrations suggested interactions between the ligands of interest with the respective reference ligand.
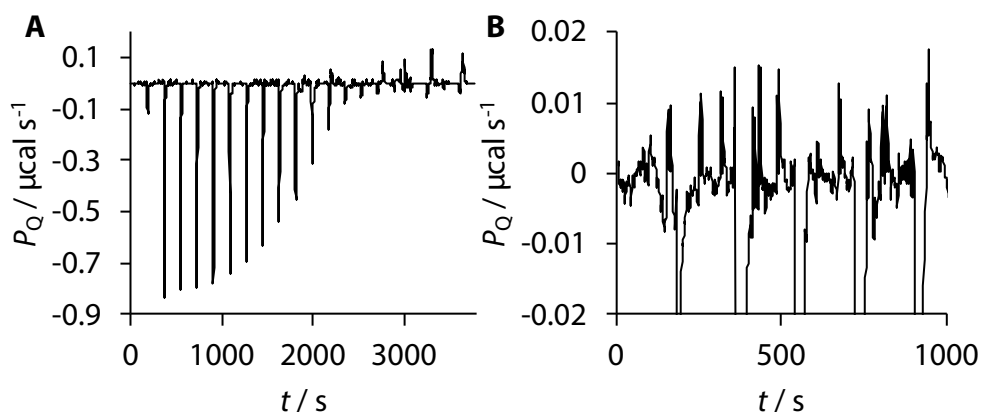
Fig. 23: **A** Exemplary processed thermogram of a titration of compound **16** into hCAII. A distinctly uneven baseline and artifact signals are visible. **B** Close-up of **A** for the first 1000 s of the titration experiment.

3.3 Conclusions

This study shows the importance of minute systematic chemical alterations in the investigation of molecular series for the establishment of detailed QSARs. Despite the argument, that studies of this kind produce results which are unique and only relevant for the investigated system, the concepts they deliver are needed, in combination with the results of similarly detailed studies, to incorporate them in computer programs that can be applied to rationalize findings in systems that have not yet been examined in a strictly systematic fashion.[21] Moreover, the results presented herein demonstrate, that the stringent development of a chemical scaffold need not necessarily exert linear thermodynamic or kinetic binding profiles. It additionally reveals, how the change of a ligand's preferred conformation by a minute chemical change can manipulate thermodynamic and kinetic signatures of binding, which gives rise to the finding, that the *n*-alkyl and *n*-alkyloxy analogs with identical tail lengths swap their binding kinetic properties at an unchanged binding affinity. It furthermore demonstrates, that a novel ITC measurement protocol yields both thermodynamic and kinetic data, the latter of which shows the same trends as data derived from SPR, the current gold standard for the retrieval of kinetic information.

4 A New Measurement Protocol for the Reliable Extraction of Thermodynamic and Kinetic Data from Isothermal Titration Calorimetry Experiments

Many adjustable parameters are available in the design of an ITC experiment. The applied concentrations of protein (c_P) and ligand (c_L), which determine the molar ratio (r_m) at every step of the titration experiment, that value of r_m, that resembles the stoichiometry of binding (n_s) of the complex formed between protein and ligand, the number of injections (N_i), the injection volume (V_i), the injection time (t_i) (with the injection rate (v_i) resulting of the former two), the delay between injections (t_s, which most strongly influences experimental runtime), the stirring speed (v_r), the *Wiseman* parameter (c), the heat of injection (Q) and the parameter for the total heat (h), introduced by *Tellinghuisen*.[112,155]

Considerable theoretical and experimental effort has gone into the fine-tuning of ITC experiments to ensure reliable thermodynamic results at minimally required experimental effort.[20,156] Less know-how is available for the application of kinITC. In addition to a plot of integrated heat values it provides the ETC, which shows the dependency of the equilibration time (τ_{eq}) on the ratio of injected species and the species in the cell.[113,114] Beginning from the so called 'standard protocol', the aim of this chapter is to provide a manual for the execution of high-quality kinITC experiments for the reliable extration of both thermodynamic and kinetic parameters.[156]

While the standard protocol will yield the thermodynamic information about a binding event, data precision can be improved with a smaller N_i and a larger V_i, which is particularly important for the comparison of congeneric ligand series with only minimal changes of the lead scaffold, and was substantiated by theoretical and experimental studies.[18,19,156] This approach becomes even more important, if subsequent computational studies rely on the experimental values.[20] Fig. 24 shows thermograms, integrated heat values and ETCs for different simulated experimental setups for a one to one binding reaction with the following constants: K_d = 620 nM, $\Delta H°$ = 15 kcal mol^{-1}, c_P = 20 μM, c_L = 200 μM and t_s = 180 s. Fig. 24**A** show the outcome of an experiment following the standard protocol (typically 20 – 30 injections).[156] It yields a thermogram with sufficiently strong injection heats and allows the determination of $\Delta H°$ and K_d. These data were simulated under the assumption of 100 % active protein sample, which yields a stoichiometry of n_s = 1.0. In practice, however, n_s = 1.0 is rarely found. Because of partial degradation and denaturation of the protein, n_s < 1.0 is the more commonly found scenario. Under the assumption, that the purity of the small molecule sample is exactly known, the stoichiometry of the binding event provides information about the fraction of active protein (i.e. protein competent to bind the ligand) present in the whole sample. Fig. 24**B** shows values for a titration experiment with a fraction of 70 % of active protein, which leads to an inflection point of the isotherm at n_s = 0.7. Notably, the peak of the ETC is also shifted to the same stoichiometry. This point is worth pondering before spending valuable protein and ligand material for a kinITC experiment. The significance of the differences between ETCs **A** and **B** lies in the number of kinetically analyzable data points (i.e. injections). The ETC's left part (smaller r_m-values before the peak is reached) is of particular importance for kinetic analysis. Injections in this range of the curve differ significantly from the instrument noise. As the injection signals determining the right part of the ETC differ much less from noise, they bear a comparably high uncertainty. Given this aspect, it becomes obvious that titration of a protein sample with a low fraction of binding-competent protein material with the standard protocol will lead to less data points in the kinetically crucial region of the ETC, while

most data points fall in the ETC's right half. This scenario culminates in the application of a thermodynamically optimized protocol with few large injections.[155] Fig. 24**C** demonstrates, that in this scenario the ETC contains almost no information in the region prior to mid-titration, and even less if an imperfect protein sample is used. Fig. 24**D** shows the outcome of a simulated ITC experiment optimized to measure both thermodynamic and kinetic values despite the practical reality of an imperfect protein sample. For kinetic analysis, as many injetions as possible are desirable before the mid-titration injection (for which τ_{eq} is the highest). Injection protocols in Fig. 24**A** – **C** will thus not provide a solution to the problem. Accordingly, it is necessary to relocate data points to the region prior to mid-titration by varying the injection volumes, while keeping the same (or a similar) number of injections to maintain the same overall run time as a standard protocol and thereby efficiency. Consequently, the corresponding injections need to be smaller in volume than injections defining the same region of thermogram 1**A**. This in turn requires larger injection volumes for the fewer final injections that define the saturation region in the right part. Both, an increase of the species in the cell at a given ligand concentration in the syringe or a decrease of the syringe concentration at a given cell concentration would shift more injections to the left part before mid-titration is reached, but at the same time prevent saturation, which is an important determinant for thermodynamic data quality.[156] One could account for the inactive fraction of protein by adjusting its concentration in a way that the solution in the cell will contain an apparent 100 % active protein. However, this will additionally increase the amount of exposed inactive protein, and the danger exists that unspecific interactions with the ligand occur and will definitely require more protein material. The experiment in Fig. 24**D** can thus be seen as a combination of titration protocols in **A** and **C** to ensure best possible titration data for thermodynamic and kinetic information despite the practical obstacle of a partially inactive protein sample.
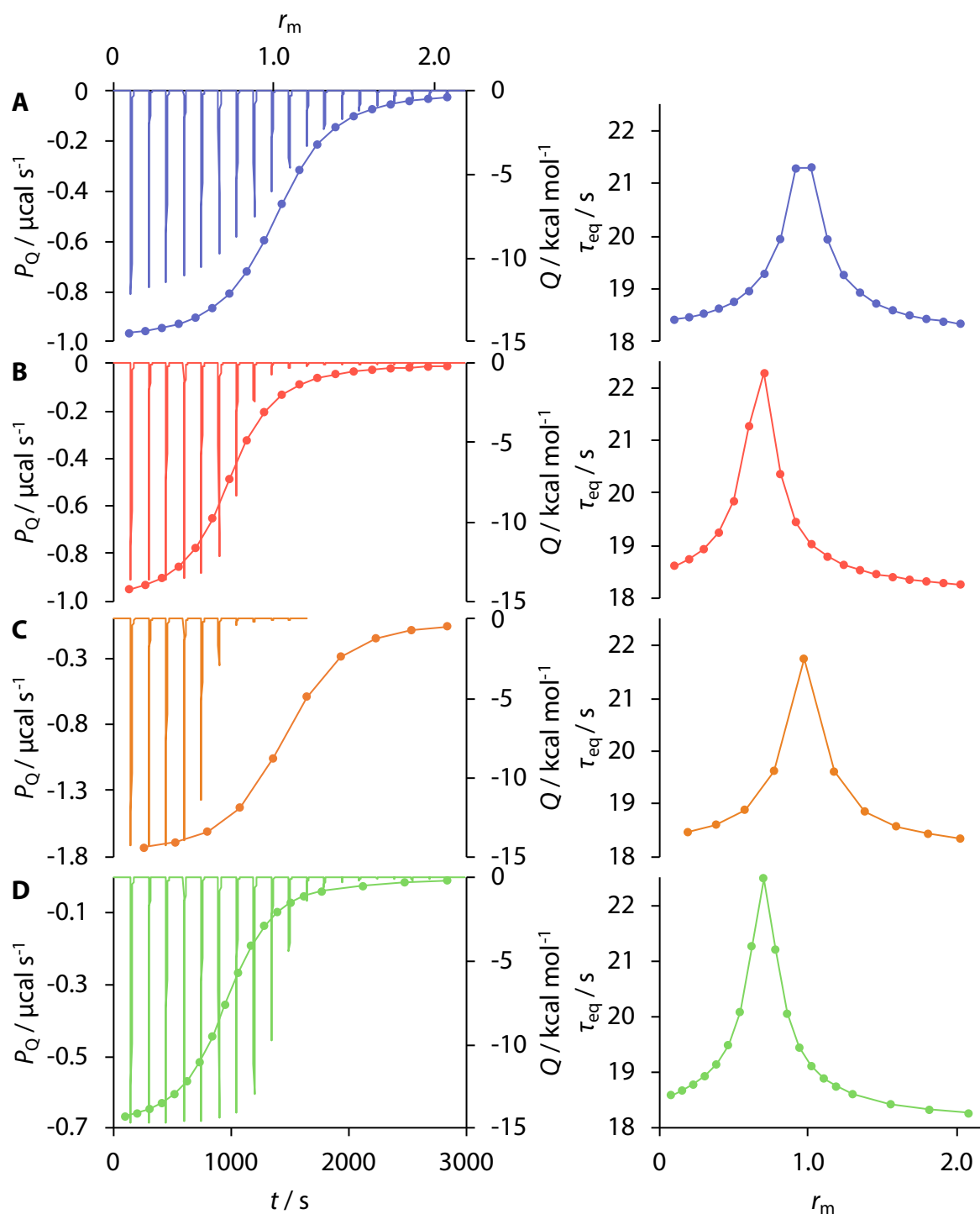
Fig. 24: Plots for simulated titration experiments. Plots on the left contain thermograms (left ordinate and abscissa in **D**) and the resulting integrated heat values as dots (right ordinate, abscissa in **A**). Plots on the right show the corresponding ETCs (abscissa in **D**). **A** 'Standard protocol' with $V_i = 2.0$ µL and $n_s = 1.0$. **B** Like **A**, but with $n_s = 0.7$. **C** Thermodynamically optimized protocol with $V_i = 3.8$ µL and $n_s = 1.0$. **D** Kinetically optimized protocol with varying V_i and $n_s = 0.7$.

To validate the approach, ITC experiments were conducted, following the different titration protocols with the model enzyme and drug target hCAII and the ligand 4-carboxybenzenesulfonamide (**20**, Fig. 25).
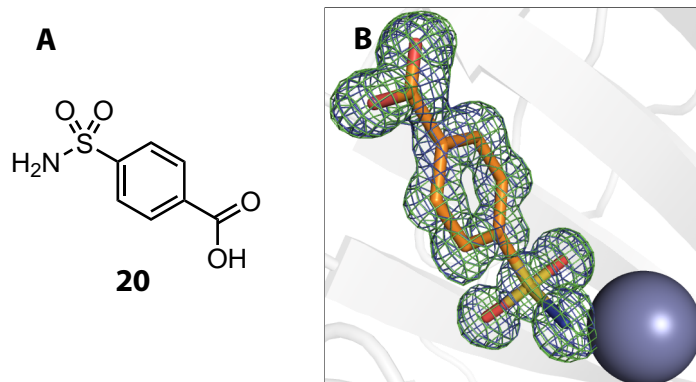
Fig. 25: A Valence bond formula of compound **20**. **B 20** bound in the active site of hCAII. Omit electron density maps are shown in blue ($2mF_o-DF_c$, 1σ) and green (mF_o-DF_c, 3σ).

To maintain comparability among protocols, the same concentrations of protein and ligand were used throughout all protocols to maintain the quantities c, h and r_m, which predominantly determine the thermodynamic readout of ITC experiments.[155]

4.1 Experimental Design

Experiments were carried out based on the following quantities: c_P = 20 µM, c_L = 210 µM (considering the purity of the ligand), t_s = 180 s, T = 298.15 K, v_r = 1000 rpm, initial differential power P_Q = 5.0 µcal s^{-1} and a filter period of 1. Additionally, a protocol was tested, which combines V_i of the standard protocol with t_i from the kinetically improved protocol to investigate the impact of the injection speed on the results. The applied injection schemes are shown in Tab. 4.

Tab. 4: Applied injection schemes for the assessment of thermodynamic and kinetic data quality.

Protocol[a]	V_i / µL[b]	t_i / s	v_i / µL s^{-1}
Standard	19×2.0	4.0	0.5
Kinetically optimized	15×1.5	3.0	0.5
	1×2.0	4.0	0.5
	3×4.5	9.0	0.5
Mixed	15×2.0	3.0	0.7
	1×2.0	4.0	0.5
	3×2.0	9.0	0.2
Thermodynamically optimized	10×3.8	7.6	0.5

[a] Color coding for the following plots. [b] An additional injection of 0.3 µL preceded every experimental run.

Titrations were carried out in degassed 40 mM HEPES buffer at pH 7.8. The purity of compound **20** was determined to 95.2 % by analytical HPLC by *Khang Ngo*. The concentration of the DMSO stocks of **20** used for the preparation of the syringe solution were corrected for this impurity so that the final concentration of **20** in the syringe was c_L = 210 µM. Measurements for each protocol were carried out fivefold. Raw and processed thermograms, integrated heat values and ETCs are provided in the experimental part.

4.2 Results

4.2.1 Thermodynamic Results

Fig. 26 shows thermodynamic results for titration experiments of the hCAII–**20** system with the above protocols.

All protocols yield comparable affinity values within the margin of error although the mixed protocol suggests a deviation from the other protocols. While the mixed protocol shows a notable difference between average and globally fitted value, these are virtually the same for the other protocols, respectively.

Furthermore, the standard protocol shows a larger error margin than the thermodynamically optimized one. Notably, the error margin of the kinetically optimized protocol lies in between those of standard and thermodynamically optimized protocols. The same applies to the values
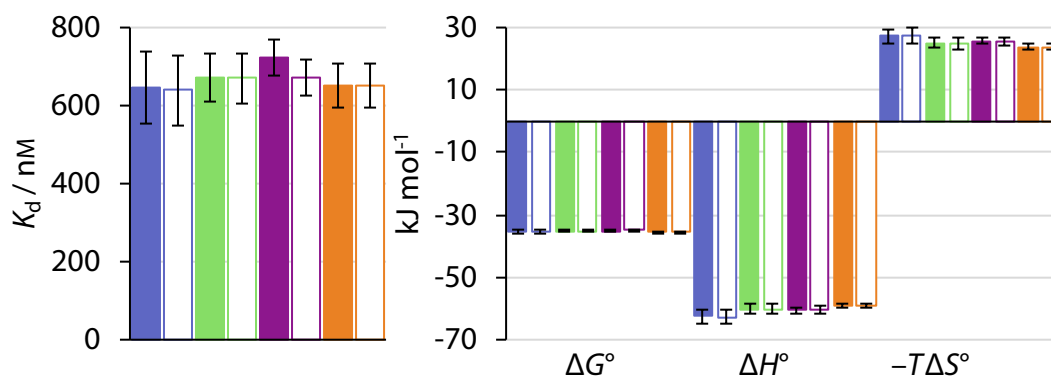
Fig. 26: Thermodynamic data of the hCAII–**20** complex obtained with different titration protocols. The color coding is adapted from Tab. 4. Filled bars represent the sample mean values, empty bars represent the globally fitted value. Error bars represent the extremes of a confidence interval of 95 % for five measurements under the assumption, that the experimental values are T-distributed.

of $\Delta H°$ and $-T\Delta S°$, showing that the kinetically optimized protocol even provides a benefit for the collection of thermodynamic data compared to the standard protocol. The mixed protocol virtually yields the same thermodynamic signature as the kinetically optimized protocol, however with even lower error margins. It also seems to produce reliable thermodynamic data.

Naturally, the previously published systematic work on the CAII–**20** system needs to be acknowledged, especially the different experimental outcomes between the present and previous studies. While the different protocols applied here all yield affinity values between 640 nM and 680 nM, an exhaustive study reported values around 690 nM and 1400 nM.[70,157,158] Furthermore, $\Delta H°$ was communicated to be between $-18\,\text{kJ mol}^{-1}$ and $-60\,\text{kJ mol}^{-1}$ while the measurements herein present an enthalpic signature of $-60\pm1\,\text{kJ mol}^{-1}$. The most prominent differences between those studies and the present one are that different enzymes were used (bovine Carbonic Anhydrase II (bCAII) instead of the human isoform herein) and different measurement buffers (phosphate buffered saline (PBS) as opposed to HEPES herein). Whether these are the main reasons for the observed deviations is difficult to answer without further independent experiments.

4.2.2 Kinetic Results

An essential quantity for the evaluation of the kinetic data is the instrument's response time (τ_{ITC}). The τ_{exp} is a fit parameter and needs to fall in close proximity of the actual physical,

calorimeter-specific value (which strongly depends on the cleanness of the cell).[159] The response time of the ITC200 was determined with ethanol in triplicate as described in the experimental part to $\tau_{ITC} = 4.36 \pm 0.09$ s. Fig. 27 shows the results for the fitted response time values for all titration protocols.
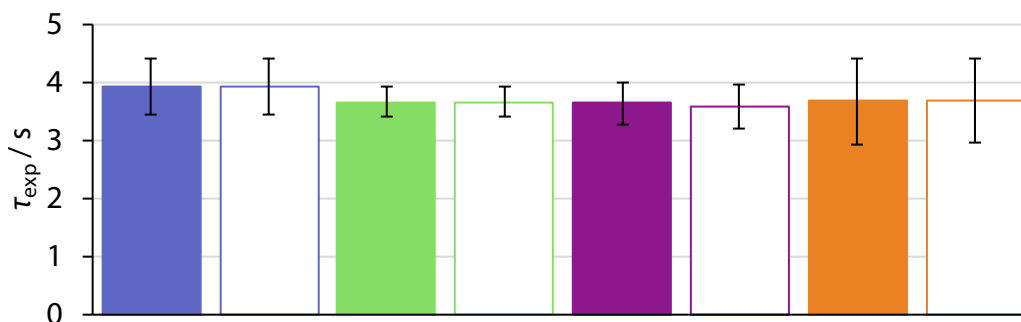
Fig. 27: Mean values (filled columns) and values after consideration of the globally determined thermodynamic parameters (empty columns) for τ_{exp} in s. Color coding as in Tab. 4.

All protocols seem to enable reasonable fitting of τ_{exp} with a deviation from τ_{ITC} of approximately 0.5 s. Notably, the kinetically optimized protocol (green) provides the smallest error margin. Kinetic parameters for all protocols are shown in Fig. 28. The respective ETCs are shown in the Appendix.
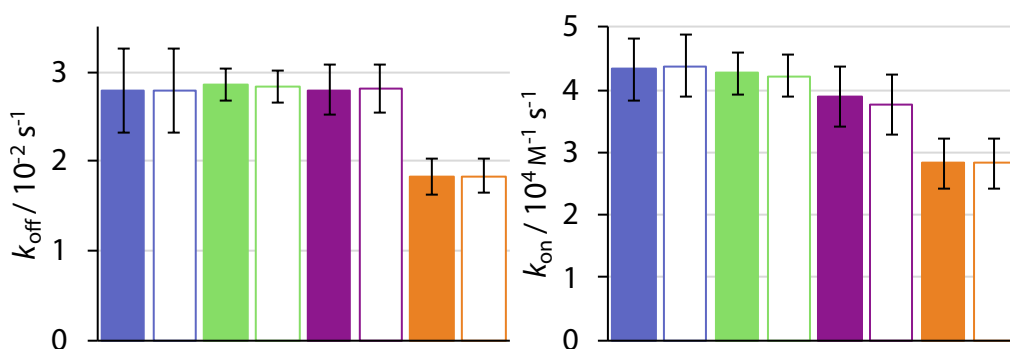
Fig. 28: Globally fitted and mean values for k_{off} and k_{on} for the applied titration protocols with the same coloring schemes as in Tab. 4.

Expectedly, the kinetic parameters derived from the thermodynamically optimized protocol differ significantly from those extracted by the other three protocols. While k_{off} is in good agreement among the latter, k_{on} agrees between standard and kinetic protocol by the margin of error, but the mixed protocol results in a deviation. As only k_{off} is fitted for the determination

of the equilibration time but not k_{on}, the difference must result from the difference in K_d. k_{on} results from the quotient of k_{off} and K_d (equation 5, p. 15).[113] The difference between the three protocols lies in the mid-titration region, from which K_d is determined, and where different injection rates are applied. Whether they are responsible for the deviation in K_d and thus in k_{on} cannot be ultimately concluded herein. However, caution is advisable for the comparison of ITC-derived kinetic data obtained at varying injection rates. One should not be misled by the comparably sized error values for the thermodynamically optimized protocol. The kinetic values are well comparable between different experiments within the series of experiments with this protocol. However, the χ^2 values between 1.5 and 3.4 for the fit of the ETC show, that kinetic data cannot be reliably extracted from such a short protocol. Notably, the kinetically optimized protocol yields the smallest error margins of all protocols and seems to be a legitimate approach to obtain both thermodynamic and kinetic data from the same ITC experiment.

4.2.3 Comparison with Surface Plasmon Resonance

Nowadays, surface plasmon resonance (SPR) is the method of choice for the measurement of biomolecular reaction kinetics. Thermodynamically, ITC and SPR seem to provide the same values for a given protein-ligand system using the same materials as demonstrated in a study with bCAII and **20**.[158] However, ITC and SPR differ fundamentally with respect to the state of the protein, which is either freely floating in solution or immobilized on an SPR chip *via* a linker. As *Zihlmann et al.* showed, kinetic data from ITC and SPR can be in good agreement, albeit with a rather constant offset between both methods.[160] Whether a linear correlation between both methods is always given or depends on the nature of the studied protein-ligand system remains to be seen, as kinetic data from ITC experiments are scarce at the moment. However, it is not unreasonable to assume that the immobilization of the protein on a chip surface can significantly influence its ability to bind a ligand, as is described in subsection 3.2.3. With this argument about the comparison between ITC and SPR in mind, it is reasonable to at least question whether kinetic data from ITC and SPR do necessarily fall within the same range or

even have to show the same trends within congeneric series of molecules, or if the inherent properties of the protein-ligand system bear differently on the experimental outcome of both methods due to the different measurement conditions. Therefore, it is necessary to ensure robust data collection for kinetic data analysis with ITC to enable reliable comparison with SPR data. In a previous study from our group, **20** was used as a reference compound for SPR measurements.[150] Kinetic values extracted from these experiments were K_d = 0.89 μM, k_{on} = $6.5 \cdot 10^4$ M^{-1} s^{-1} and k_{off} = $5.8 \cdot 10^{-2}$ s^{-1}. Despite falling within the same range, they do differ by the values presented herein by 30 % (k_{on}) and 50 % (k_{off}). This discrepancy is distinct, but not as large as a difference of one order of magnitude for the molecular series investigated in chapter 3, which underlines the validity and importance of differentiating the different experimental conditions of ITC and SPR.

4.3 Conclusions

It was shown that a combination of the standard protocol and a thermodynamically optimized protocol with varying injection volumes can be used to extract both thermodynamic and kinetic information for 1:1 protein-ligand reactions from a single titration with the method of kinITC with lower uncertainty in the thermodynamic and kinetic quantities than with the standard protocol. We found putative evidence for an influence of the injection rate on K_d and thereby k_{on}. Obviously, the combination of standard protocol and thermodynamically optimized protocol will double experimental runtime and increase the error in the thermodynamic parameters compared to the latter. This is, however, a necessary compromise, seeing that the kinITC method relies on data points before and next to the mid-titration region of an ITC experiment. As demonstrated herein, a sufficient number of data points in this region can be difficult or even impossible to achieve with either standard or thermodynamically optimized protocol, especially considering that protein samples are seldom perfect and contain 100 % of active protein.

While many theoretical considerations have advanced the field of microcalorimetry, only practical aspects and empirical findings can be described herein. For future work with the method of kinITC, mathematical models that predict putative dependencies of the kinetic parameters on experimental settings would be desirable.

5 The Influence of Varying Fluorination Patterns

Herein, a thermodynamic and kinetic fluorine-scan of the active site of hCAII, similar to the one carried out by *Olsen et al.* for the serine protease thrombin, is described.[102] For this purpose, variously fluorinated and *para*-substituted BSA ligands were investigated. Fig. 29 shows the general structure of these compounds, as well as the definition of certain geometric measures, that quantitatively describe their binding in the active site of hCAII. The respective data for these compounds are listed in Tab. 5
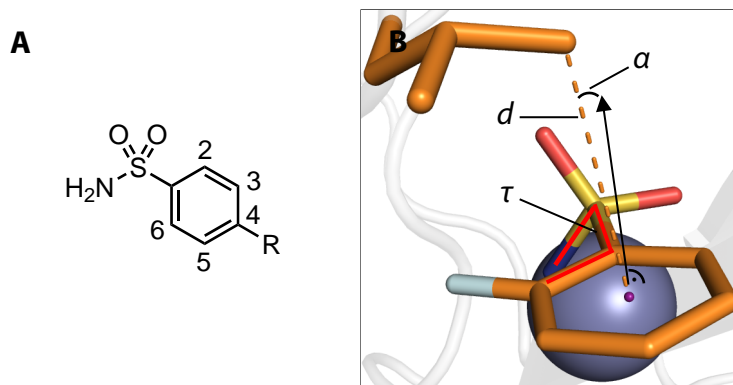
Fig. 29: A General structure of investigated compounds. Numbers refer to the fluorinated positions. **B** Definiton of geometric measures listed in Tab. 5

In 2007, *Krishnamurthy et al.* investigated BSAs **4** (chapter 3) and **21**, **22**, **23**, **26**, **27** and **32** (Tab. 5) as CAI by ITC to assess the individual thermodynamic contributions of the different kinds of interactions of these ligands to the overall difference in *Gibbs* free energy of binding to bCAII (structurally closely related to hCAII) and concluded, that 65 % were contributed by the

Tab. 5: Investigated compounds with geometric measures derived from the respective crystallographic models. Definitions of α, d_{C_6-Ph} and τ are shown in Fig. 29.

Compound	Fluorinated Positions	R	PDB Entry	Occupancy[b]	α / °[c]	Torsion angle τ / °[d]	d_{C_6-Ph} / Å	RMSD Thr200[e]	Interface area m-F / Å²[f]	Interface area o-F / Å²[g]
4[a]	—	H	6GDC[a]	1.0	0.9	49.2	3.5	—	—	—
5[a]	—	Me	6GM9[a]	1.0	2.4	50.9	3.5	0	—	—
21	2	H	6RIT	1.0	5.0	55.7	3.6	0.183	—	—
22	3	H	6RQI	1.0	2.9	49.1	3.6	0.072	127.4	—
23	4	F	6RKN	1.0	1.6	50.2	3.5	0.022	—	—
24	2,3	H	6RJI	1.0	40.1	0.5	4	0.086, 0.221	161.1	140.3
25	2,5	H	6RNP	1.0	4.7	54.8	3.5	0.061, 0.205	114.34	—
26	2,6	H	6ROE	0.66, 0.34	4.7	2.1, 56.9	3.6	0.272	—	138.8
27	3,5	H	6RRG	0.69, 0.31	2.0	2.3, 54.7	3.6	0.07	157.3	—
28	2,3,5,6	H	6RRI	1.0	42.2	2.0	4.0	0.077	142	138.3
29	2,3,5,6	Me	6RS5	0.35, 0.65	5.3	0.1, 58.2	3.7	0.068, 0.298	157.4	137
30	2,3,5,6	Et	6RSZ	1.0	42	1.2	4.0	0.065	165.6	139.9
31	2,3,5,6	Pr	6S9G	1.0	41.2	2.1	4.0	0.097	159.9	138.9
32	2,3,4,5,6	F	6SD7	1.0	41.8	4.2	4.1	0.059	155.7	138.5
34	—	CN	6ROB	1.0	2.4	48.8	3.6	0.039	—	—
35	—	NO_2	6RH4	1.0	3.2	49.2	3.6	0.038	—	—
36	—	NH_2^+	6RL9	1.0	2.2	50.1	3.6	0.015	—	—

[a] Previously published in [140]. [b] If two values are given, the first refers to the reference conformation, the second to the alternative conformation. [c] Measured for the reference conformation of Leu198 using PyMOL with the NumPy library for vector construction and the vg library for vector operations.[161–163] If two ligand conformations exist, the value refers to the alternative conformation, as this conformation is always correlated with the reference conformation of Leu198. [d] The *ortho* carbon atom closer to Thr200 was used. [e] Measured using the function 'rms' from PyMOL including all atoms of Thr200 except riding hydrogen atoms relative to the reference conformation of Leu198. [f] Determined using the program dr_sasa in mode 4, using the *meta*-fluorine atom accommodated in the hydrophobic pocket formed by the side chains of residues Val121, Phe131, Leu141 and Leu198 as ligand and only side chain atoms of these residues as receptor. [g] Determined using the program dr_sasa in mode 4, using the *ortho*-fluorine atom accommodated in the hydrophobic pocket formed by the side chains of residues Val121, Phe131, Val143 and Leu198 as ligand and only side chain atoms of these residues as receptor.[164] Calculated only for the reference orientation of Leu198 and the corresponding binding mode.

interaction between the sulfonamide anion and Zn^{II} cofactor, 10 % by the hydrogen bond network established between ligand and protein and 25 % by hydrophobic interactions between aromatic ring and protein.[129,144] *Scott et al.* later used molecules **4**, **21**, **22** and **23** among other compounds to establish a thermodynamic QSAR of CAIs.[99]

Considering the circumstance, that comparably little kinetic data derived from ITC experiments has been published so far, it seems appropriate to provide further validation for the method, especially under the prerequisite of an adjusted measurement protocol for the reliable extraction of both thermodynamic and kinetic data (chapter 4). Hence, the above compounds are convenient in this respect as thermodynamic data for subsets of these are available from the previously mentioned studies.

5.1 Crystallographic Results

Crystallographic models of compounds **21**, **22**, **23**, **26** and **27** were already available in the PDB under the accession codes 2WEG, 2WEO, 1IF4, 1IF5 and 1IF6. However, structures for all compounds in Tab. 5 were made to increase resolution and maintain the same conditions for all complexes. In general, two different conformations of fluorinated compounds in the hCAII active site were observed, which orient the phenyl ring differently, as depicted in Fig. 30. The first orientation is adopted by compounds **21**, **22**, **23**, **26**, **27** and **29** (mean value of torsion angle τ_{mean} = 54.2° with a standard deviation of ±3.5°)and mimics that of non-fluorinated compound **4**, which will be referred to as 'reference conformation'. The deviating orientation adopted by compounds **24** and **26** – **32** (τ_{mean} = 2.2° ± 1.9°) will be referred to as 'alternative conformation'. It is worth noting, that the reference conformation enables the key interaction between the aromatic portion and the side chain of Leu198.[141] Tab. 5 lists the torsion angle (τ) between the nitrogen atom, coordinated to the Zn^{II} cofactor, sulfur, the *ipso* carbon atom C1 and C2. The contact between the aromatic ring and Leu198 results in a short distance to one of the terminal C_δ methyl groups, the distance between C_δ and the centroid of the ring is listed in

Tab. 5, together with the deviation from the direction of the normal vector of the phenyl ring plane. In the reference conformation, the angular deviation is comparably small. In the second alternative conformation, it amounts to approximately 55°.

At first, monofluorinated compounds **21** – **23** will be considered. Comparable to **4** and **5**, compound **23** merely orients its *para*-fluoro substituent toward the entrance of the active site funnel, virtually adopting the same geometry as **4** and **5** with respect to both the ligand pose and the orientation of the adjacent residue Thr200. The latter residue is identically oriented as in the *apo*–hCAII structure (PDB entry 3KS3[165]). Compounds **21** and **22** (Fig. 30) orient the fluorine atoms in opposite directions. Notably, the fluorine atom of **21** is located within a distance of the hydroxy function of Thr200 that allows for the inference of a hydrogen bond, as already pointed out by *Scott et al.*[99] Compound **22** positions the fluorine atom toward the rim of the hydrophobic wall. This seems reasonable to expect with respect to the preference of fluorine atoms and fluorinated motifs to occupy hydrophobic pockets.[82] With this binding pose, **22** adopts a conformation very similar to **4** and **5** and Thr200 also remains in an unchanged orientation, whereas **21** pushes Thr200 slightly out of position (RMSD = 0.2 Å), likely due to steric repulsion with the *ortho* fluorine substituent. Interestingly, the combination of **21** and **22** with respect to the fluorination patterns to reveal the difluoro derivatives **24** and **25** results in different orientations of the latter two compounds. Whereas **25** is virtually a superposition of the binding modes of **21** and **22**, **24** adopts the alternative orientation of the phenyl ring. Supposedly, this shift results from steric repulsion with the terminal methyl group of Val121. With respect to Thr200, **25** induces the same geometry already observed for **21**. The binding of **24**, however, entails two different movements of Thr200, one of which can be described as toward the ligand, which can be caused by an attractive interaction. Notably, compounds **26**, **27** and **29** adopt both orientations of the phenyl ring in the crystal structure. Importantly, the alternate conformations found for compounds **26** and **27** merit the renewed production of crystal structures herein,
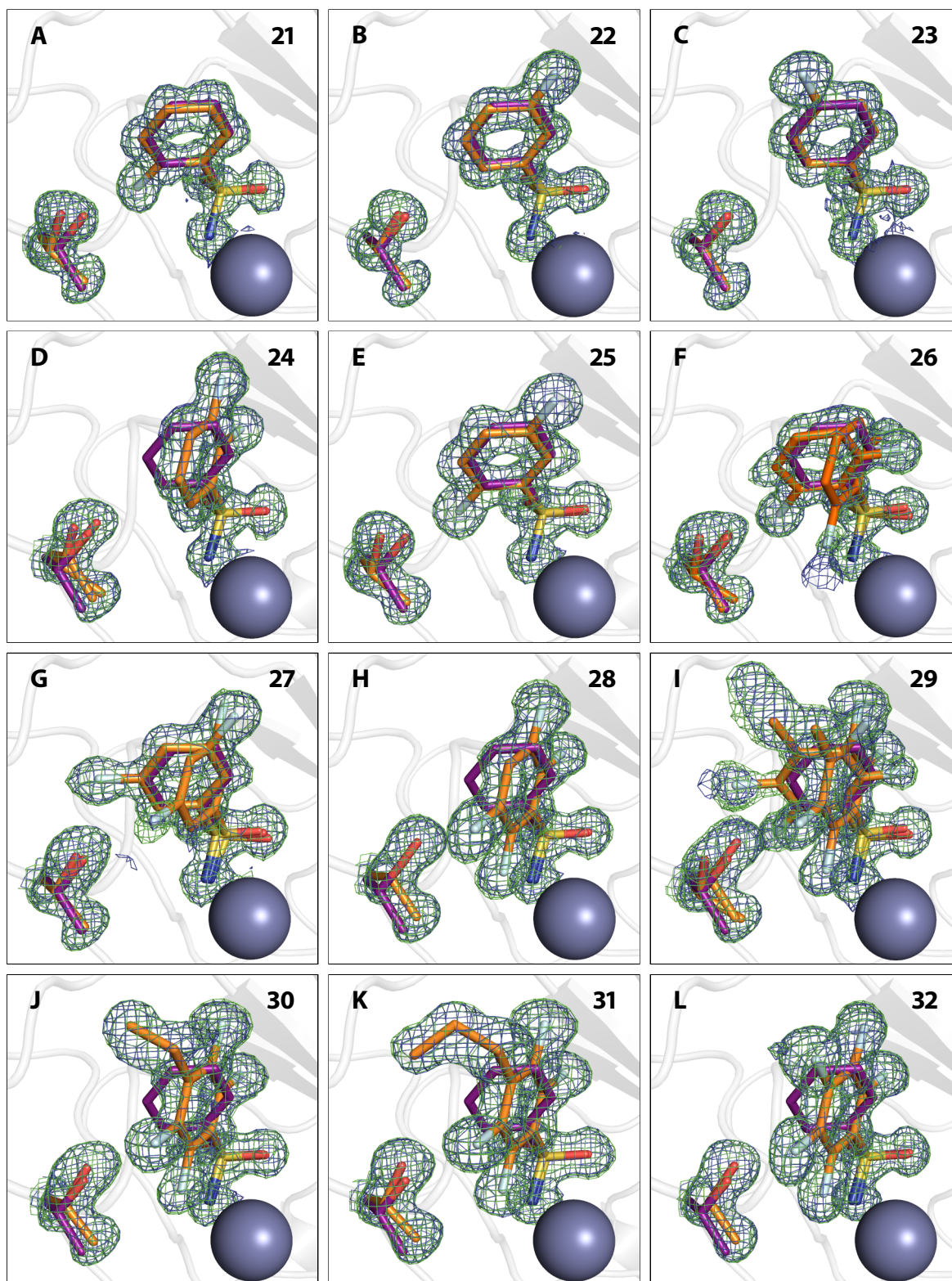
Fig. 30: Binding modes of fluorinated compounds **21** (**A**), **22** (**B**), **23** (**C**), **24** (**D**), **25** (**E**), **26** (**F**), **27** (**G**), **28** (**H**), **29** (**I**), **30** (**J**), **31** (**K**) and **32** (**L**) and the respective Thr200 residue in orange. Fluorine atoms are shown in light blue, the Zn^{II} cofactor as gray sphere. **4** and the according Thr200 residue are shown in purple. Omit $2mF_o–DF_c$ and $mF_o–DF_c$ maps are contoured in blue at 1 σ and green at 3 σ, respectively.

seeing that no alternate conformations could be resolved in the already deposited models 1IF5 (**26**) and 1IF6 (**27**, Fig. 31).
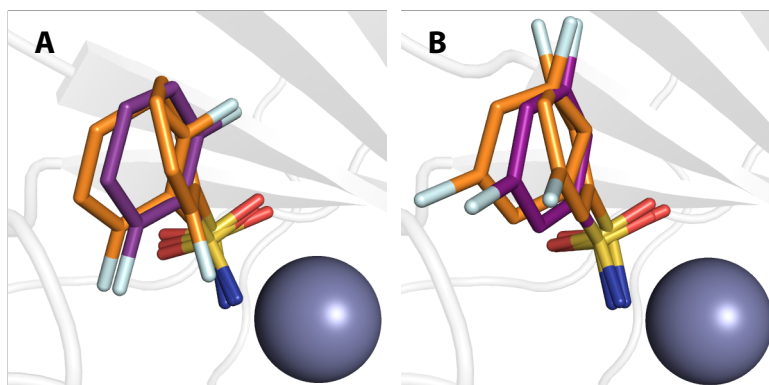
Fig. 31: Compounds **26** (**A**) and **27** (**B**) in orange in comparison with ligand conformations observed in PDB entries 1IF5 and 1IF6.

Furthermore, difference electron density of the final models indicates, that a second binding conformation seems possible also for compounds **21**, **22**, **28** and **30**, although the difference density is too weak to properly allow modeling of the second, definitely minor populated arrangement. Reconsidering the interpretation of the density for **21** and **22**, a second putative binding conformation appears visible in the positive difference density, negative difference density indicates overpopulation for the modeled conformations of compounds **28** and **30** (Fig. 32).
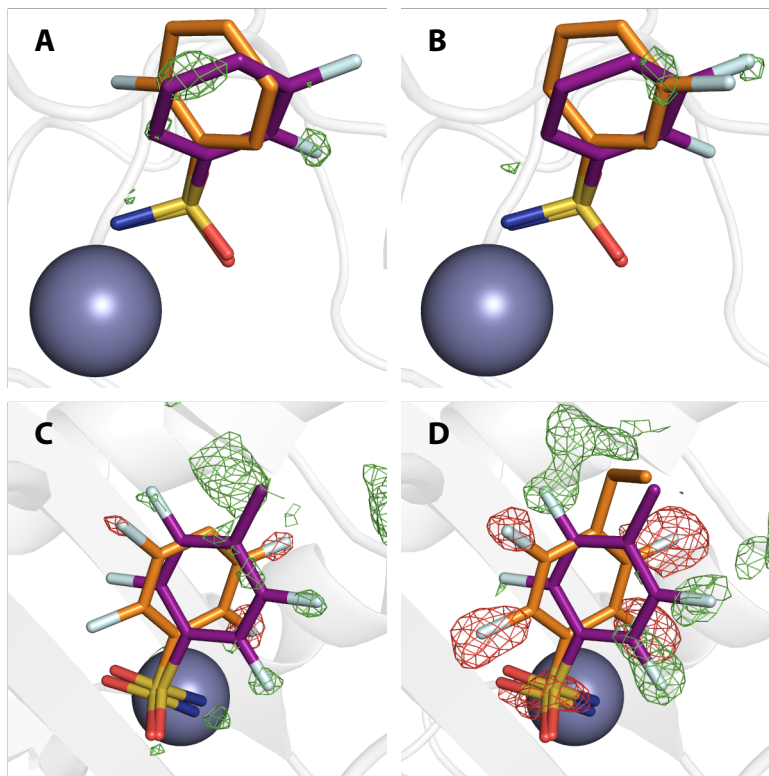
Fig. 32: Compounds **21**, **22**, **28** and **30** in orange with putative additional binding conformations. Compounds **21** (**A**) and **22** (**B**) are shown in comparison with compound **24** in purple to indicate the second conformation. Compounds **28** (**C**) and **30** (**D**) are shown in comparison with compound **29** in purple to indicate the second conformation.

5.1.1 Movement of Thr200

The notion, that hCAII has a highly rigid binding site which is not influenced structurally by the binding of ligands, needs to be relativized, seeing that almost all of the diversely fluorinated BSAs examined herein cause the movement of Thr200.[129] The orientation of Thr200 in the hCAII–**4** complex and in the unliganded protein will also be referred to as the reference state in this respect. It was found that ligands distinctly displace Thr200 with respect to its position in the reference state in two manners (Fig. 30). A quantification by RMSD analysis relative to the hCAII–**4** complex can be found in Tab. 5. The first one can be described as a movement away from the Zn^{II} cofactor, the second one as a movement toward it. Binding of compounds **21**, **22**, **25**, **26** and **27** entails movement away from the zinc ion. Compounds **28**, **29**, **30**, **31** and **32** cause a movement toward the zinc ion. Given the comparably low degree to which **22** displaces the

amino acid, it is still reasonable to assume the same geometry as in the reference structure. The same holds for **23**. In case the *ortho* position is occupied by a fluorine atom, the slightly larger size of F compared to H results in a shift of Thr200 (**21** and **25**) away from the Zn^{II} cofactor. Obviously, an additional *o*-fluorine atom attached to the monofluoro-**22** results in a strong rotation of the phenyl ring and the *o*-fluorine atom is accommodated in a small niche next to Val121. Thr200 is shifted away from the zinc ion. The di-*ortho* and di-*meta* substituted derivatives **26** and **27** bind with two conformers simultaneously. In the di-*ortho* case, Thr200 is shifted, whereas the di-*meta* derivative binds with Thr200 in nearly unchanged orientation. The series of **28** – **32** share the 2,3,5,6-tetrafluoro pattern. Apart from **31** and **32**, they all share two binding poses. However, these are differently populated. The *p*-Me derivative **29** shows both orientations, interestingly with reversed occupancies compared to **26** and **27**. In the case of **28** and **30**, the second orientation is only found with minor occurrence (Fig. 32). Notably, **31** shows only one orientation of the phenyl ring. Possibly, this is caused by the circumstance that the attached *n*-propyl group at C4 adopts a *gauche* conformation and occupies some space required for the placement of the ligand in the alternative conformation. Given the observation, that ligands with only one modeled orientation can have a putative second binding orientation (Fig. 32), this can be seen as a hint for a putative second orientation of Thr200 that might be caused by the indicated, but not modeled, second orientation of the ligand. Although the RMSD values for the movements of Thr200 in the hCAII–ligand complexes are small, it is probable that these movements are a feature of ligand binding, as the structures were solved at very high resolutions around 1 Å.

5.1.2 Accommodation of a Fluorine Atom in a Hydrophobic Pocket

As reasoned above, the different orientation of **24** compared to **25** is likely the avoidance of steric repulsion between the *o*-fluorine substituent and the side chain of Val121. The movement of the ligand furthermore entails a deeper burial of the *m*-fluorine substituent in the cavity bounded by the side chains of residues Val121, Phe131, Leu141, Val143 and Leu198 (Fig. 33).
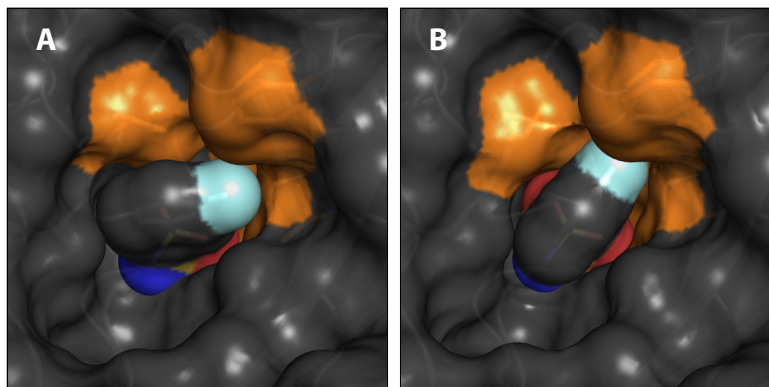
Fig. 33: Hydrophobic pocket in orange, that accommodates *meta* and *ortho* fluorine substituents, exemplarily depicted with compounds **22** (**A**) and **24** (**B**), which adopt either the reference ($\tau \sim 55°$) or alternative conformation ($\tau \sim 0°$).

This suggestion is supported by the circumstance, that also compound **26**, which bears no *meta* substituent, partially adopts the alternative orientation. Tab. 5 shows values for the interface area between the fluorine atom accommodated in the hydrophobic pocket defined above. It is noteworthy, that for the alternative orientation (**24**, **27** – **32**) the interface values for the accommodation of either *ortho* or *meta* substituent are larger than for the standard orientation (**22**, **25**) and in a similar range across either the *meta* or *ortho* substituted molecules. Although the elucidation of universal structural rules is impeded by the presence of an alternative binding mode, the binding of compounds **21**, **22**, **24** and **25** reveals, that the preference of a fluorine atom in *meta* position to bind to the hydrophobic pocket exceeds that of a fluorine atom in *ortho* position to be oriented toward Thr200.

The comparison of compounds **24** and **25** furthermore allows for the conclusion, that the preference for a fluorine atom in *meta* position to bind to the hydrophobic pocket is stronger than that of a fluorine atom in *ortho* position to be oriented toward Thr200.

Given the complex picture painted by various substitution patterns, it was furthermore sought to establish structures of hCAII complexes of merely *para*-substituted BSAs **34** – **37**, which would putatively maintain the same binding pose as unsubstituted BSA **4**, and furthermore not displace Thr200 (Fig. 34). However, given that compound **37** is protolytically active at the measurement pH of 7.8 with the phenolic OH functionality, it was not taken into account for thermodynamic

and kinetic analysis. The same was argument was considered for compound **20**. Compound **36** was characterized thermodynamically, but did not yield reliable kinetic parameters.
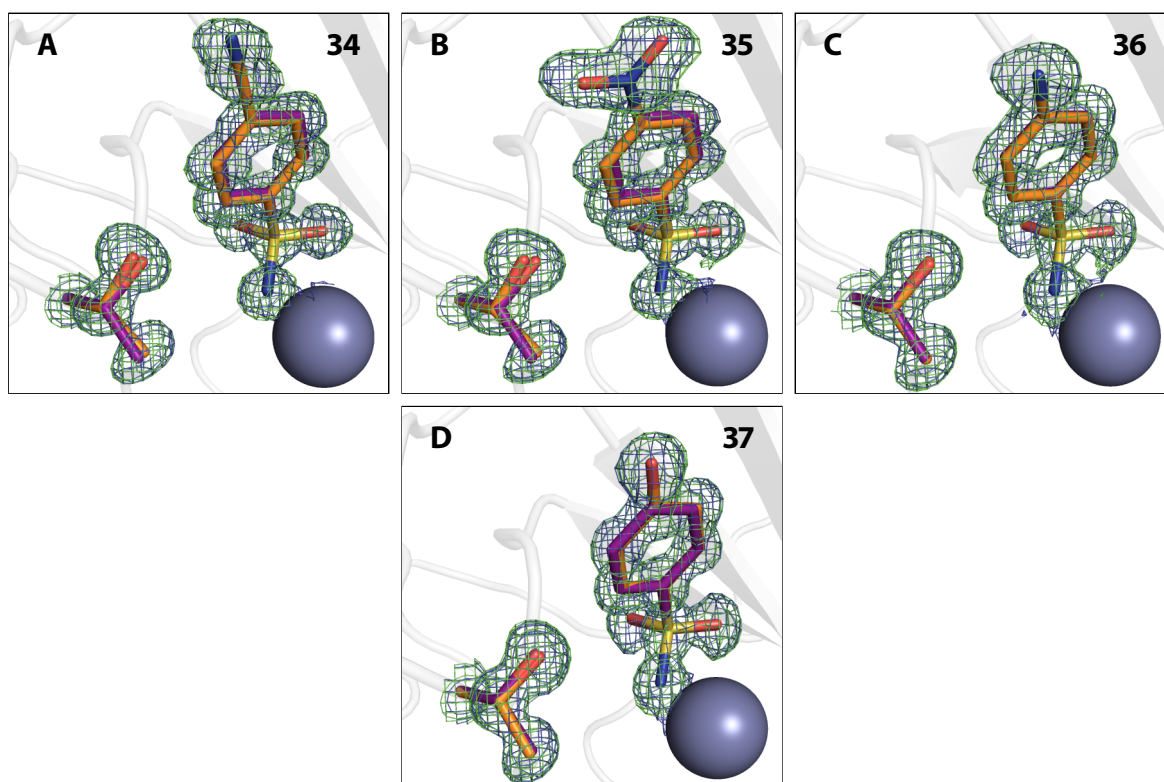
Fig. 34: *para*-substituted BSAs **34** (**A**), **35** (**B**), **36** (**C**) and **37** (**D**) in orange with omit electron density maps in blue at 1σ ($2mF_o - DF_c$) and 3σ ($mF_o - DF_c$) and the respective side chain of Thr200. Compound **4** is shown as reference in purple.

5.1.3 An Unexpected Dimerization Product

Electron density for hCAII in complex with compound **32** revealed, that not only the catalytic center next to the zinc ion is occupied by a ligand but that a further intriguing molecule, that represents a dimer of **32** is found in the crystal structure. It features intramolecular edge-to-face π-stacking interactions. Fig. 35 shows dimer **33** and its crystallographic model geometry suggested by omit electron density maps.

Considering the *van der Waals* radii of a carbon atom in benzene ($r_{\text{vdw}}(C_{\text{ar}}) = 1.77\,\text{Å}$) and a phenyl bound fluorine atom ($r_{\text{vdw}}(F_{\text{ar}}) = 1.47\,\text{Å}$), the sum of 3.24 Å is larger than the intramolecular distance of 3.0 Å found for the binding mode of **33**.[166] It depicts the electron deficient nature
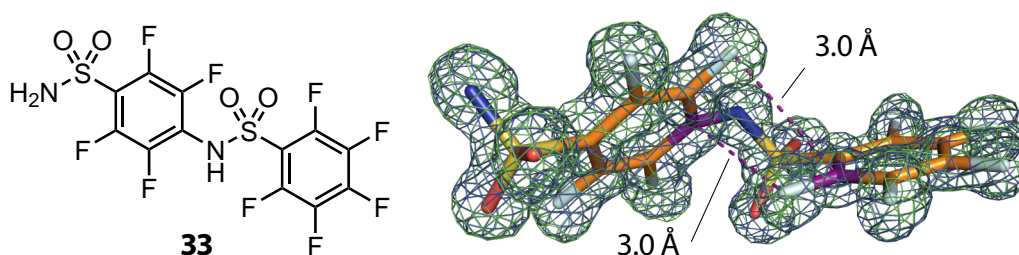
Fig. 35: Valence bond formula of dimerization product **33** (left) and crystallographic model from the hCAII complex with $2mF_o-DF_c$ map contoured at 1σ and mF_o-DF_c map in green and 3σ (right). Carbon atoms are shown in orange, fluorine atoms in light blue. Carbon atoms used to display intramolecular π-stacking are shown in purple. π-stacking interactions are indicated as purple dashed lines with the respective distances between atomic dispositions of F and C.

of the aromatic moieties, which leads to a decreased electron density around the carbon atoms. Compound **33** binds in a second, surface-exposed binding site, covered by a second crystal mate. There, it engages in both, classical hydrogen bonds and interactions often observed between fluorine atoms and proteins. Fig. 36 shows an overview of two hCAII symmetry mates clamping molecule **33** and a close-up view of the binding site. With respect to the catalytic center, a fair amount of difference density indicates binding of a ligand. The dimer **33**, however, could not be modeled in the active site, although pronounced residual mF_o-DF_c density above the modeled monomer **32** suggests, that **33** also populates the active, albeit with a very low occupancy.
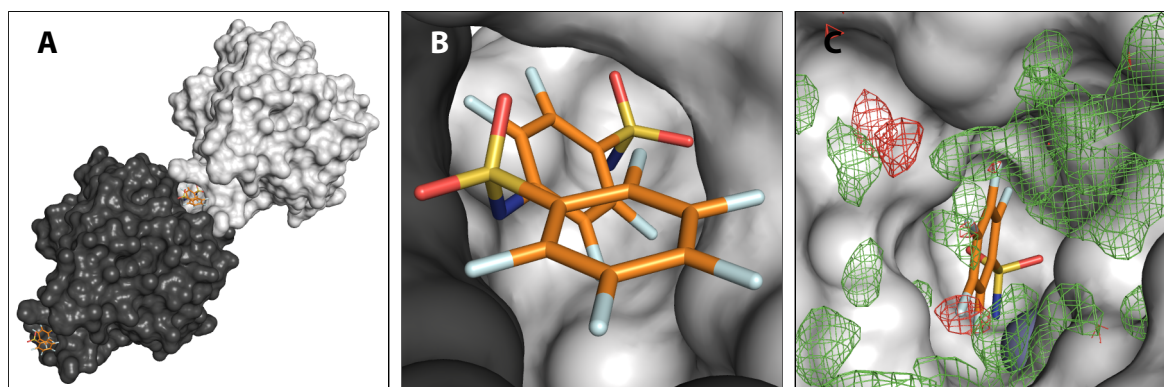
Fig. 36: A Surface representation of two hCAII crystal mates, which enclose dimer **33** between them. **B** Closeup of the binding site of **33** between two hCAII crystal mates. **C** Residual mF_o-DF_c density at 3σ (green) and −3σ (red) above the active site with modeled monomer **32** (for difference electron density of **32**, s. Fig. 30**L**).

85

It was not clear, whether dimer **33** was formed in the soaking drop or had already been present in the commercially available solid. Consequentially, it was desisted from further usage of **32** for ITC experiments, given the unclear quality and purity of the sample.

5.2 Thermodynamic Results

Titrations of hCAII with a different batch of **4** resulted in a molar amount of transferred protons of n_p = −0.1 mol L^{-1} (subsection 14.10), meaning that on molar scale, 0.1 protons are transferred to the surrounding buffer per formed hCAII–**4** complex. As sulfonamides are known to bind hCAII as anions, a proton from the sulfonamide group must be released into the surrounding buffer. Under the assumption, that the ligand associates with the protein in the uncharged state as described by *Gaspari et al.*, the proton needs to be transferred to the surrounding medium from within the active site. The reason for the small detected amount of 0.1 moles L^{-1} supports the assumption, that, prior to sulfonamide binding, the fourth vertex of the tetrahedral Zn^{II} complex is occupied by a hydroxide ion. The hydroxide ion can react with the proton to form water, which does not react with a buffer molecule, and thus does not alter the overall reaction heat in a buffer-dependent manner. Thus, the thermodynamic signatures, depicted in Fig. 37, which describe the entire complex formation process of the investigated compounds, can be assumed to not be overlaid by a protonation step.

The addition of a fluorine atom in *para* position of **4** entails a moderate increase in binding free energy ($\Delta\Delta G°_{4\to23}$ = −1.3 kJ mol^{-1}), whereas the addition of a fluorine atom in *ortho* position leading to compound **21** entails a distinct change in the thermodynamic binding profile compared to that of **4**, with a distinctly stronger enthalpic advantage and entropic penalty, which lead to a slightly larger gain in $\Delta G°$ ($\Delta\Delta G°_{4\to21}$ = −2.2 kJ mol^{-1}). A similar development holds true for *meta*-fluorinated compound **22** to a lesser extent in the enthalpic and entropic parameters, but a further increase in overall affinity ($\Delta\Delta G°_{4\to22}$ = −4.2 kJ mol^{-1}). Interestingly, the combination of **21** (2-F) and **22** (3-F) to reveal **25** (2,5-F), which adopts the same conformation ($\tau \sim 0°$) and
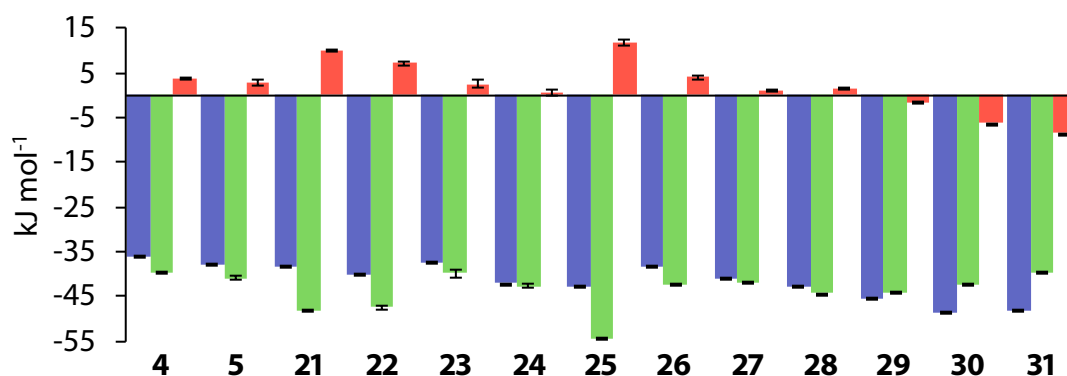
Fig. 37: Thermodynamic signatures of compounds **4**, **5** and **21** – **31** measured in 10 mM HEPES buffer at pH 7.8 with $\Delta G°$ in blue, $\Delta H°$ in green and $-T\Delta S°$ in red. The values for **4** and **5** are taken from chapter 3. For **4**, **5** and **21** – **31**, globally fitted values of three direct titrations are given. Error bars represent the standard error of measurement for these compounds, based on three measurements. Compounds **29** – **31** were characterized by a displacement experiment with subsequent global fitting, as they could not be characterized directly due to their high affinities. For these compounds, values of K_a and $\Delta H°$ from globally fitted displacement experiments and the values of $\Delta G°$ and $-T\Delta S°$ calculated from the former two are given. Error bars represent the error of global fitting for K_a and $\Delta H°$. Error values for $\Delta G°$ and $-T\Delta S°$ were calculated from the former two under assumption of propagation.

binding pose as the former three (**21**, **22** and **23**) and orients both fluorine atoms accordingly, amounts to more than the sum of its parts in the thermodynamic signatures. Both enthalpic benefit and and entropic penalty are larger than for either of the former three compounds and yield an overall gain in binding free energy ($\Delta\Delta G°_{4\to25} = -6.8\,\text{kJ mol}^{-1}$). Mono-*ortho* and mono-*meta* fluorination as well as 2,5-difluorination seem to be the enthalpically favored substitution patterns in the investigated series, whereas the *para* fluorination in **23** has no enthalpic advantage. However, the alternative combination of **21** and **22** to form the *ortho-meta* **24** entails a distinct diversion from the enthalpic and entropic signatures of the parent compounds. Hardly any entropic contritbution is observable in HEPES buffer, rendering **24** a merely enthalpic binder within the investigated series of molecules, with a similar value of $\Delta G°$ as **25**. Notably, a second *ortho* fluorine atom as in compound **26** reduces the entropic penalty of compound **21**, but also the enthalpic benefit, which leaves the free energy of binding virtually unaltered compared to **21**. Di-*meta* fluorination in **27** has a similar effect, given the different enthalpic and entropic signatures of monofluorinated **22** and difluorinated **27** and their highly similar values of $\Delta G°$ ($\Delta\Delta G°_{22\to27} = -0.7\,\text{kJ mol}^{-1}$). Tetrafluorination of **4** to afford **28** does not provide a distinct increase in affinity over related compounds, seeing that the addition of two fluorine atoms to **24** yields an almost negligible increase in $\Delta G°$ of ($\Delta\Delta G°_{24\to28} = -0.5\,\text{kJ mol}^{-1}$). In this particular

case, it is obvious that the addition of two additional fluorine atoms is thermodynamically futile. The addition of a methyl (**29**) or ethyl group (**30**) in 4-position, however, provides an increase in affinity for the tetrafluorinated scaffold. Given the second possible binding mode of the ethyl derivative **30** (Fig. 32), however, it is questionable, whether an increased ability to form interactions with the hydrophobic wall, as shown in previous studies, is the only reason for the increase in affinity.[140,150] It must certainly play a role, however, given the circumstance, that compound **31**, which does not show a second binding mode that would allow for an interaction between the *n*-propyl chain and the hydrophobic wall, has a slightly decreased affinity compared to **30** ($\Delta\Delta G°_{31\rightarrow30} = -0.5\,\text{kJ mol}^{-1}$).

A comparison of *para*-substituted compounds **23** and **34** – **36** and unsubstituted prototype **4** shows, that for electron withdrawing groups, a gain in affinity is mainly enthalpy-driven, while an electron-donating group in **36** causes a loss in affinity due to an entropic penalty. *para*-substituted compounds **20** and **37** were not considered for this comparison, as they show protolytic activity at the measurement pH of 7.8.
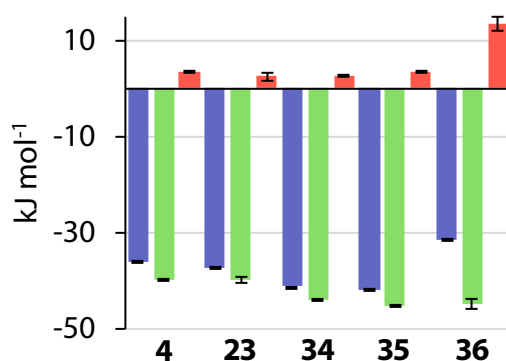
Fig. 38: Comparison of globally fitted thermodynamic signatures of compounds **4**, **23** and **34** – **36**. Error bars represent the standard error of measurement, based on three measurements.

5.3 Kinetic Results

The kinetic k_{on} and k_{off} rates were determined by the newly developed titration protocol using kinITC (chapter 4). The obtained data are summarized in the Appendix (p. 188). Overall, values of k_{on} scatter over four orders of magnitude ($\Delta k_{on} = 15.5 \cdot 10^4\,\text{M}^{-1}\,\text{s}^{-1}$) and k_{off} less with

$\Delta k_{\text{off}} = 1.7 \cdot 10^{-2} \text{ s}^{-1}$. Given the circumstance, that the variety of different substituents and substitution patterns investigated herein inevitably entail varying acidities of the BSA ligands, their pK_{a} values were measured by *Björn Wagner* from *Roche* (Tab. 6).

Tab. 6: pK_{a} values of fluorinated and *para*-substituted BSAs.

Compound	pK_{a}
4	9.96 ± 0.01
5	10.15 ± 0.01
21	9.38 ± 0.01
22	9.57 ± 0.01
23	9.84 ± 0.01
24	8.98 ± 0.01
25	8.99 ± 0.01
26	8.90 ± 0.01
27	9.18 ± 0.01
28	8.08 ± 0.01
29	8.20 ± 0.01
34	9.13 ± 0.01
35	9.06 ± 0.01
36	2.12 ± 0.01, 10.55 ± 0.01

Compared to the non-fluorinated ligands **4** and **5**, no clear-cut correlation is obvious, and a large scatter of k_{on} and k_{off} suggests a complex structure-kinetic correlation (Fig. 39)
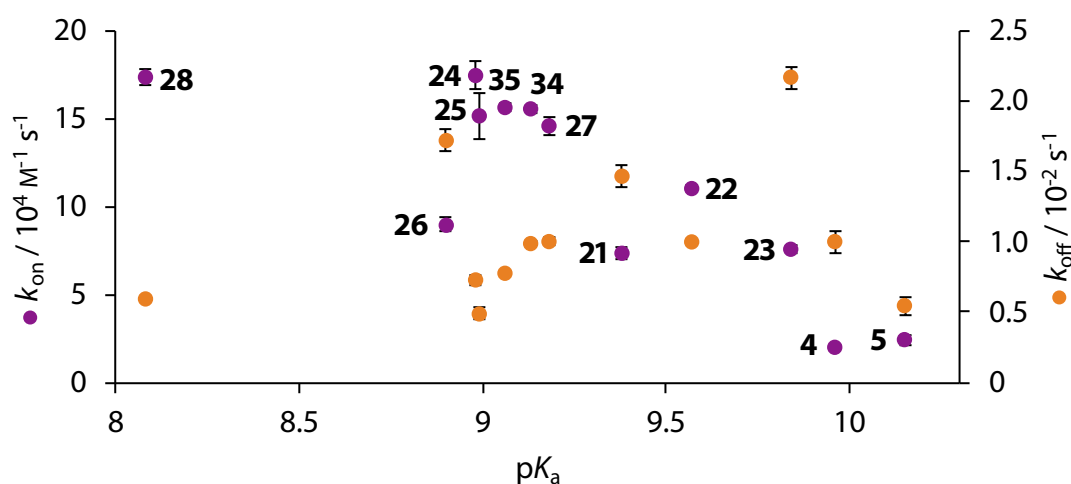
Fig. 39: Plot of k_{on} and k_{off} against the pK_{a} value for compounds **4**, **5**, **21 – 28**, **34** and **35**.

5.4 Discussion

5.4.1 Comparison of Thermodynamic Data with Earlier Studies

As mentioned above, ligands **4**, **21**, **22**, **23**, **26** and **27** had already been investigated previously by *Krishnamurthy et al.* and furthermore by *Scott et al.*, who included compounds **4**, **21**, **22** and **23** in a study of variously substituted CAIs.[99,144] Fig. 40 shows a comparison of the thermodynamic data of these studies with the respective data collected in the course of this work.
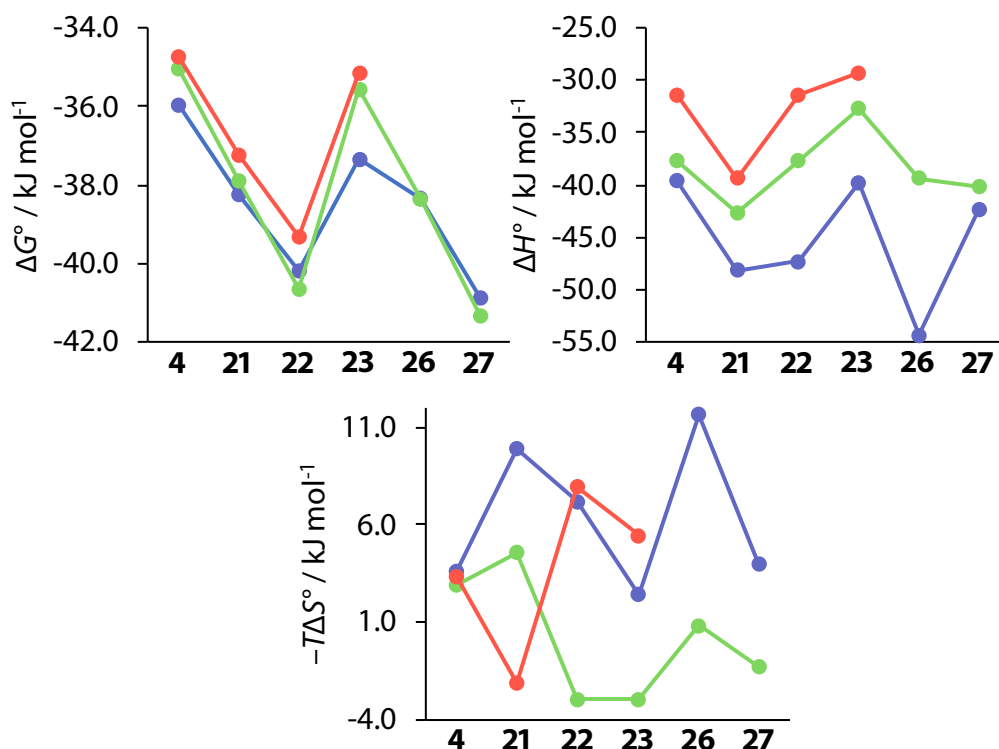
Fig. 40: Comparison of thermodynamic values recorded in this work (blue) with data from studies by *Krishnamurthy et al.* (green) and *Scott et al.* (red).

The values for $\Delta G°$ closely resemble each other with the largest difference of $\Delta\Delta G° = -2.1\,\text{kJ mol}^{-1}$ found for compound **23** between the value presented herein and that of *Scott et al.*[99] It needs to be noted here, that *Krishnamurthy et al.* used bCAII instead of hCAII, which are, at least structurally, largely similar.[129] Whether this similarity translates into the thermodynamic signatures is difficult to estimate. The enthalpic contributions differ by amounts up to almost $20\,\text{kJ mol}^{-1}$. This can be expected, considering that the studies were conducted in different buffers at different pH values,

and that the previously mentioned small amount of protons can be expected to be transferred between a CA–BSA complex and the surrounding buffer.[30,167] Generally, the enthalpic trends are roughly maintained. However, given that the offsets are not constant, other differences between the studies must be expected. The deviations between all three studies become even more pronounced in the entropic term, given that only some trends can be crudely reproduced. However, considering that the entropic term is calculated from the directly measured values of K_a and $\Delta H°$, errors in these quantities will strongly influence the entropic term. Seeing that the trends for the directly measured quantities K_a (and thereby $\Delta G°$) could be reproduced, this study is a further example for the validity of the measurement protocol, that was improved for kinetic analysis of ITC data. *Krishnamurthy et al.* investigated the above compounds in order to partition the contribution of electrostatic and hydrophobic interactions on the ligand's structural elements, arguing that this was possible due to the same orientation of the BSA cores of all ligands.[144] Given the finding, that **26** and **27** do, however, adopt two binding modes, and that a second conformation can also be inferred for **21** and **22**, it is questionable, whether a partitioning in this manner can be conducted this easily. Also, the argument does not take into account the special nature of fluorine, which greatly differs from that of hydrogen. The various orientations of fluorine atoms in the active site by different ligands enable different interaction patterns, and they cause structural changes of the protein and different binding modes of the ligands. It therefore seems questionable, that the overall free energy of binding can be simply partitioned on the individual structural motifs, given the large variety of interaction patterns across the ligand series.

5.4.2 Comparison of Crystallographic Results with Thermodynamic Data

Overall, two conformers are observed in the studied complexes. The phenyl moiety adopts an eclipsed arrangement with respect to the attached sulfonamide group, giving rise to two conformational families. Either the S–N bond ($\tau = 0°$) or one of the S=O bonds is in coplanar

arrangement with the phenyl ring ($\tau = 55°$). The two additional substituents at sulfur adopt an orientation with approximately 60° above and below the phenyl plane. The two binding-pose orientations correlate either with an occupancy of the fluorine substituent in a small hydrophobic niche next to Val121 and Leu141 or seemingly more deeply buried in a pocket below Phe131 and adjacent to Val143. The two orientations also correlate with an angular deviation of the terminal C_δ-methyl group of Leu198 from the normal vector perpendicular to the phenyl ring. The 4-substituent is oriented toward the opening of the funnel-shaped active-site pocket and is partly exposed to the solvent environment. Likely due to steric conflicts in some of the complexes, significant spatial shifts are recorded for the placement of Thr200, taking the *apo*-protein as reference.

Assuming that the crystallographically observed binding modes reflect geometries relevant for the thermodynamic signatures of the binding ligands, our data suggest that the occupancy of the hydrophobic niche next to Val121 by a fluorine substituent is enthalpically favored compared to the placement in the alternative pocket below Phe131. Opposing trends are found for the entropic signature, slightly compensating the enthalpic advantage. The seemingly deeper burial of the fluorine substituent may be compensated by a more costly steric placement with an increased contact surface to the protein (Tab.5). However, these opposing effects are difficult to translate directly into changes of the thermodynamic profiles. Remarkably, the placement of an *ortho*-fluoro substituent next to Thr200 results in a shift of the latter residue, however, without an enthalpically detrimental effect. This might be compensated for by a putative hydrogen bond between the fluorine atom and the hydroxy side chain function of Thr200, already suggested by *Scott et al.*

5.4.3 Comparison of Crystallographic Results with Kinetic Data

An inherent problem in the process of establishing structure-kinetic relationships is the contrary nature of structure on the one hand, and the processes of association and dissociation on the other. While structural knowledge is mostly elucidated in the solid state, the association and

dissociation of molecules comprises several subsequent states, whose time-dependency cannot be elucidated in the solid state. For the interpretation of kinetic data, structural information about the time-limiting step along the association and dissociation process must be known. While in the case of 4-alkyl substituted BSAs, a model for the association mechanism was derived by *in silico* methods, some deductions about the dissociation might be inferred from the bound state, given that it necessarily represents the starting point for the dissociation mechanism. For the 4-alkyl substituted BSAs, an intermediate pre-binding state governed by hydrophobic interactions is the initial and rate-limiting step in the association between hCAII and BSAs.[150] Subsequently, on the way to the final binding pose, the ligand forms hydrophobic interactions with the hydrophobic wall, but additionally engages in hydrogen bonding with Thr199, Thr200 and the Zn^{II}-bound hydroxide ion.[150] Given the increased hydrophobic character that is often entailed by fluorination, a similar scenario can be envisaged for the compounds investigated herein.[76] However, the rate-limiting step need not necessarily be the same in this case. To assess, whether hydrophobicity is important for binding kinetics, k_{on} and k_{off} were correlated with literature $logP$ values of the considered ligands (Fig. 41).[144] However, no clear-cut correlation was found.
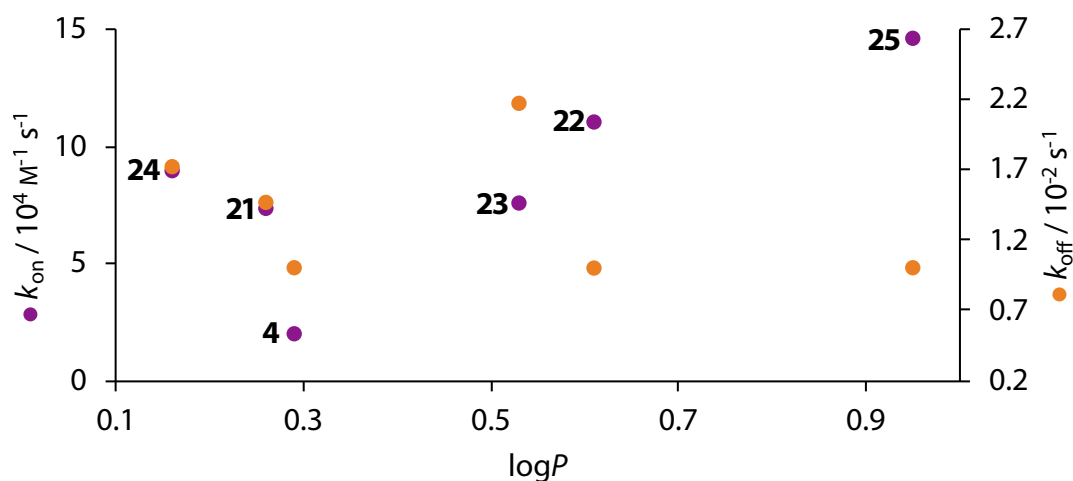
Fig. 41: A plot of k_{on} and k_{off} against $logP$ shows, that hydrophobicity is not a determining factor for the binding kinetics of fluorinated BSAs.

It is reasonable to assume, that the reprotonation of the ligand may play a role in the rate-limiting step of the dissociation process, given that BSAs bind as anions. In this scenario, the pK_a values should correlate with the dissociation rates (k_{off}). However, Fig. 39 definitely indicates a more complex correlation for k_{off}. Notably, for compounds **4**, **23**, **34** and **35**, which do not cause major structural perturbations in the active site or feature different substitution patterns, a correlation between the pK_a value and both $\Delta G°$ and k_{on} can be inferred, but not between the pK_a and k_{off} (Fig. 42).
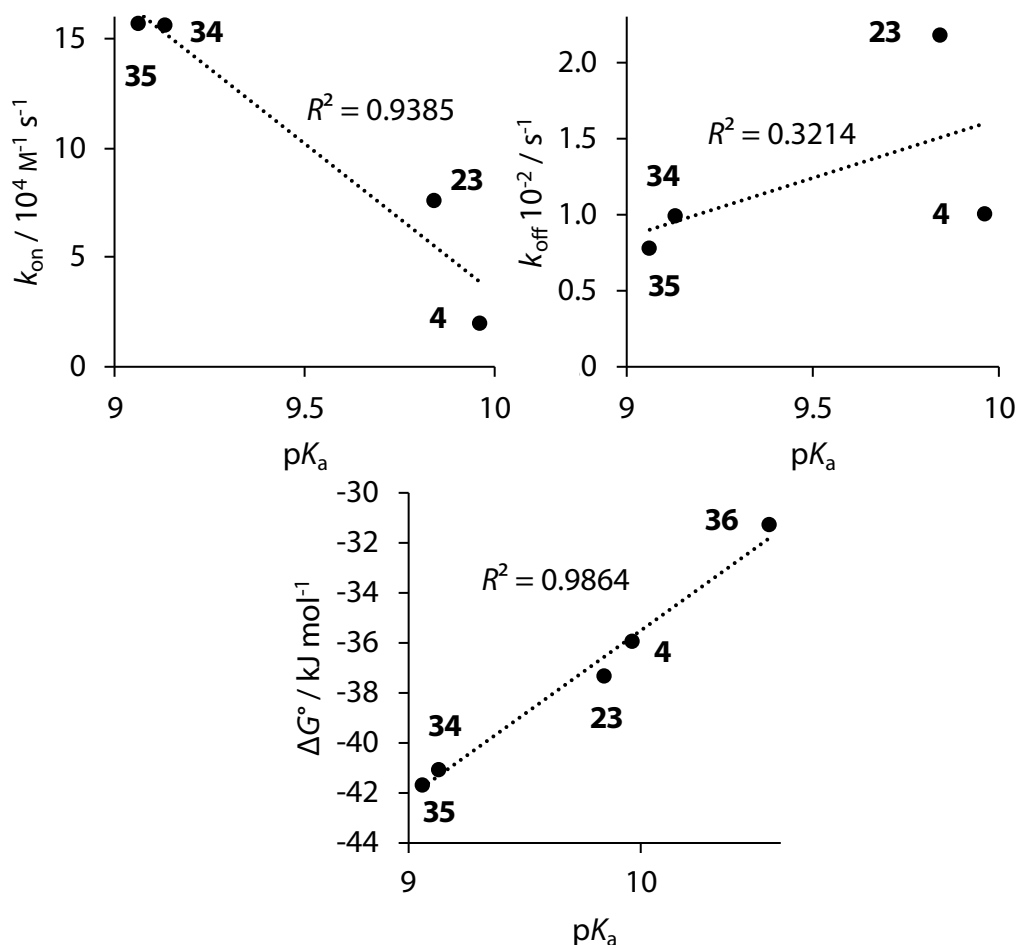
Fig. 42: Plots of k_{on} and k_{off} of compounds **4**, **23**, **34** and **35** and plot of $\Delta G°$ of compounds **4**, **23**, **34**, **35** and **36** against the pK_a value.

Opposed to the *para*-substituted derivatives, ligands bearing *ortho* or *meta* fluoro-substituents seem to contribute additional features that modulate the correlation of structural with biophysical data. Exemplarily, compounds **22**, **24**, **25** and **27**, which share the common feature of a

hydrophobically accommodated fluorine substituent in *meta* position along with solely one major binding pose of the phenly ring are comparably fast binders and simultaneously relatively slowly dissociating compounds. With some care, this finding seems to assign a fluorine atom in this position a special role in for the binding kinetics of the studied series. Moreover, the comparison of compounds **24** and **28** reveals, that, in addition to the thermodynamic similarity shown above, the kinetic signatures are almost identical. This provides further support for the assumption, that fluorination in *meta* position along with the accommodation of the fluorine substituent in a hydrophobic pocket is favorable for the binding to hCAII. This observation is further supported by the kinetic data of the fluorinated ligands **21** and **26**, which only feature an *ortho* fluoro substituent but lack a *meta* fluorine atom that could be favorably accommodated in the hydrophobic niche. Both, together with **23**, which lacks *ortho* and *meta* substituents, show higher dissociation rates and lower association rates compared to the remaining fluoro derivatives (Fig. 39).

5.5 Conclusions

Herein, high-resolution crystallographic and ITC data for a series of variously fluorinated and 4-substituted BSAs are presented. The applied substitution patterns have significant influence on the ligands' acidities (Tab. 6). While it is difficult to establish a straight-forward structure activity relationship because of the complex and multifaceted nature of the binding of these compounds, there is evidence that a *meta* fluorine substituent is favorable for both, overall affinity and furthermore high association and low dissociation rates. Moreover, it could be shown, that the incorporation of an additional second fluorine atom (compare compounds **21** and **26**) or two additional fluorine atoms (compare compounds **24** and **28**) can be futile, at least with respect to the biophysical properties.

A small set of chemically diverse, only *para*-substituted, BSAs shows, that an increasing acidity provides for a faster association process. As was demonstrated previously, the association

between para *n*-alkyl substituted BSAs benefits from a hydrophobic contact between the ligands and hCAII.[150] It can be expected, that the alkyl series features highly similar acidities across all compounds. In combination with the findings presented herein, it can be reasoned, that, as long as acidity is maintained for structurally similar *para*-substituted BSAs, the formation of a hydrophobic contact likely represents the rate-limiting step. However, if acidity is strongly modulated while structural similarty is largely preserved, the transfer of one of the sulfonamide protons to the Zn^{II}-bound hydroxide ion prior to active-site binding becomes the rate-limiting step. This finding is clearly indicated by the 4-substituted BSAs of deviating acidity (Fig. 42). This correlation still seems to hold for the entire series of studied ligands (Fig. 39), although deviations from this trend as for compounds **21** and **26** are difficult to explain structurally or eletronically on the basis of available material. Likely, a series of *ortho*-substituted ligands needs to be studied further elucidate the matter.

6 Renewed Investigation of a Potent Human Carbonic Anhydrase Inhibitor

In 2002, *Grzybowski et al.* devised CAI **1** (Fig. 43**A**) by computational means, which comprises the four classical design elements of BSA CAIs.[129,142] It showed an affinity of 30 pm in a competitive assay and was deemed the most potent CAI at the time of its development.[142] To obtain an estimate of the performance of **1** compared to alkylated compounds without SRE investigated in chapter 3, a new crystallographic model was produced to improve data quality, given that the previously available model of the hCAII–**1** complex (1IF7) dates back to the year 2002.[142] In both cases, compound **1** was soaked into the crystal (Fig. 43**B**).
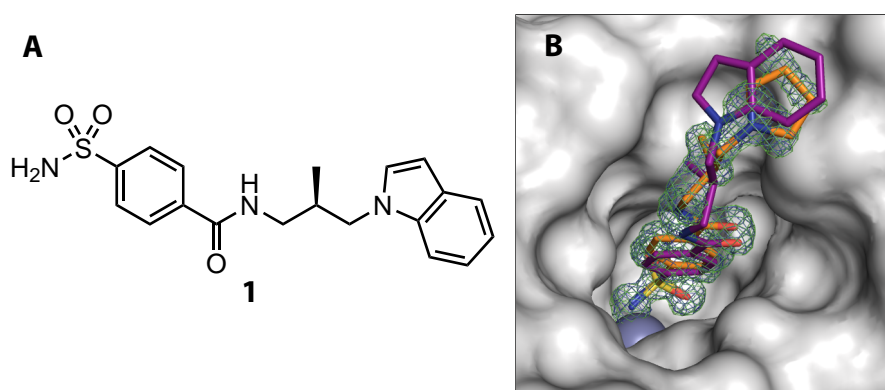
Fig. 43: A Valence bond formula of compound **1**. **B** Superimposition of the hCAII–**1** complex 6SDL (orange) with the previous model 1IF7 (purple) published by *Grzybowski et al.*, for which electron density is not available. Only the active site of 6SDL is shown. For model 6SDL, the omit $2mF_o$–DF_c map is shown at 1 σ in blue and the omit mF_o–DF_c map at 3 σ in green.

Furthermore, **1** was subjected to ITC measurements. Surprisingly, a titration of **1** into hCAII did not yield a curve characteristic for a 1 : 1 binding event, but revealed at least 2 events. To probe whether ligand degradation in the frozen DMSO stock solution was the cause, a new stock was

prepared and titrated into a solution of hCAII. Fig. 44**A** shows the processed thermograms (t denotes the time) and integrated heat values from titration experiments with different DMSO stocks of **1**.
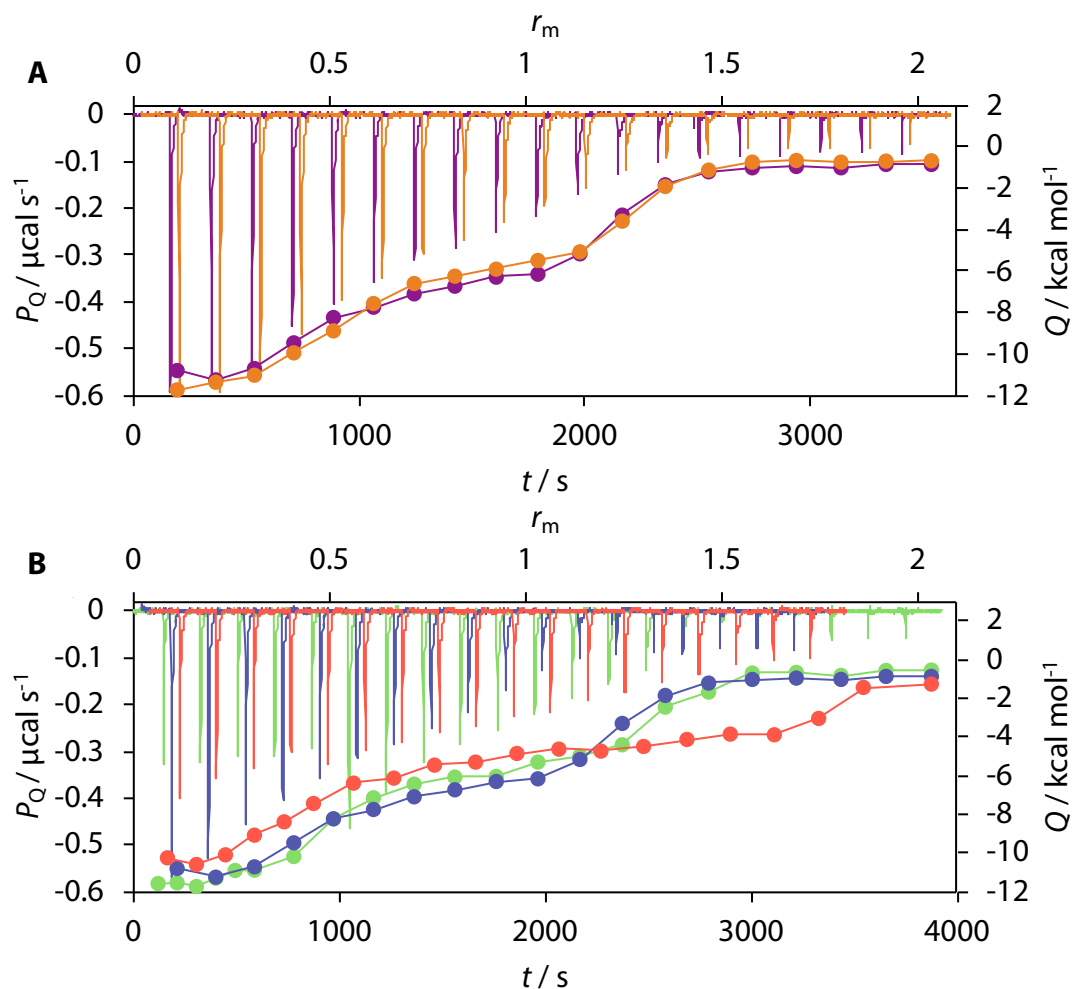
Fig. 44: A Comparison of ITC data for two stock solutions of **1** in purple (old stock) and orange (new stock). Thermograms are shown as lines with the power signal associated with an injection (P_Q) on the left ordinate and based on the experimental run time on the bottom abscissa. The orange thermogram is shifted by 40 s for clarity. Integrated heat values are represented as circles with the injection heat Q on the right ordinate and r_m on the top abscissa. Data points are connected for clarity. **B** Comparison of different injection protocols for three measurements with the older stock solution. The blue and red thermograms are shifted by 40 s and 80 s, respectively, and data points of integrated heat values are connected for clarity.

Importantly, the observed shape of the curves cannot be caused by a 2 : 1 binding stoichiometry of **1** to hCAII, as the two inflection points are found at values of r_m of 0.45 and 1.2, and are thus not integral multiples of each other, which would be expected for such a scenario. The high similarity between both thermograms and integrated heat values, respectively, proves, that the first DMSO

Tab. 7: Applied injection schemes to investigate compound **1**. Colors according to Fig. 44.

	blue	green	red
	$19 \times 2.0\,\mu L$	$6 \times 1.0\,\mu L$	$6 \times 1.5\,\mu L$
		$16 \times 2.0\,\mu L$	$13 \times 2.0\,\mu L$
			$1 \times 3.0\,\mu L$
V_t	$38.0\,\mu L$	$38.0\,\mu L$	$38.0\,\mu L$

stock solution (purple data), which had been prepared one year before the second (orange data), remained stable and that the second event visible in the integrated heat values does not arise as a result of decomposition of **1** in DMSO. Furthermore, a dilution titration of **1** into buffer suggests, that molecules of **1** seem to interact in solution, seeing that the resulting heat signals are not equally sized, as would be expected. However, subtraction of the dilution titration data from the titration data for the reaction between **1** and hCAII does not alter the progression of the resulting integrated heat values. An impact on the shape of the resulting integrated heat value curves, however, was noticed for different injection schemes with the same V_t of 38.0 μL to elucidate other putative events (Tab. 7). Given the circumstance that the shapes of the titration curves vary with the injection protocol (Fig. 44**B**), it was deduced that the visible events are time dependent, as the different titration protocols lead to different concentrations of L and consequentially PL at a specific instant during the experimental run. Thus, additional experiments with the same V_t, V_i and N_i with different t_s were carried out (Fig. 45**A**). Furthermore, the temperature dependence was investigated, given that the dependence of $\Delta H°$ is often observed (Fig. 45**B**).[22]
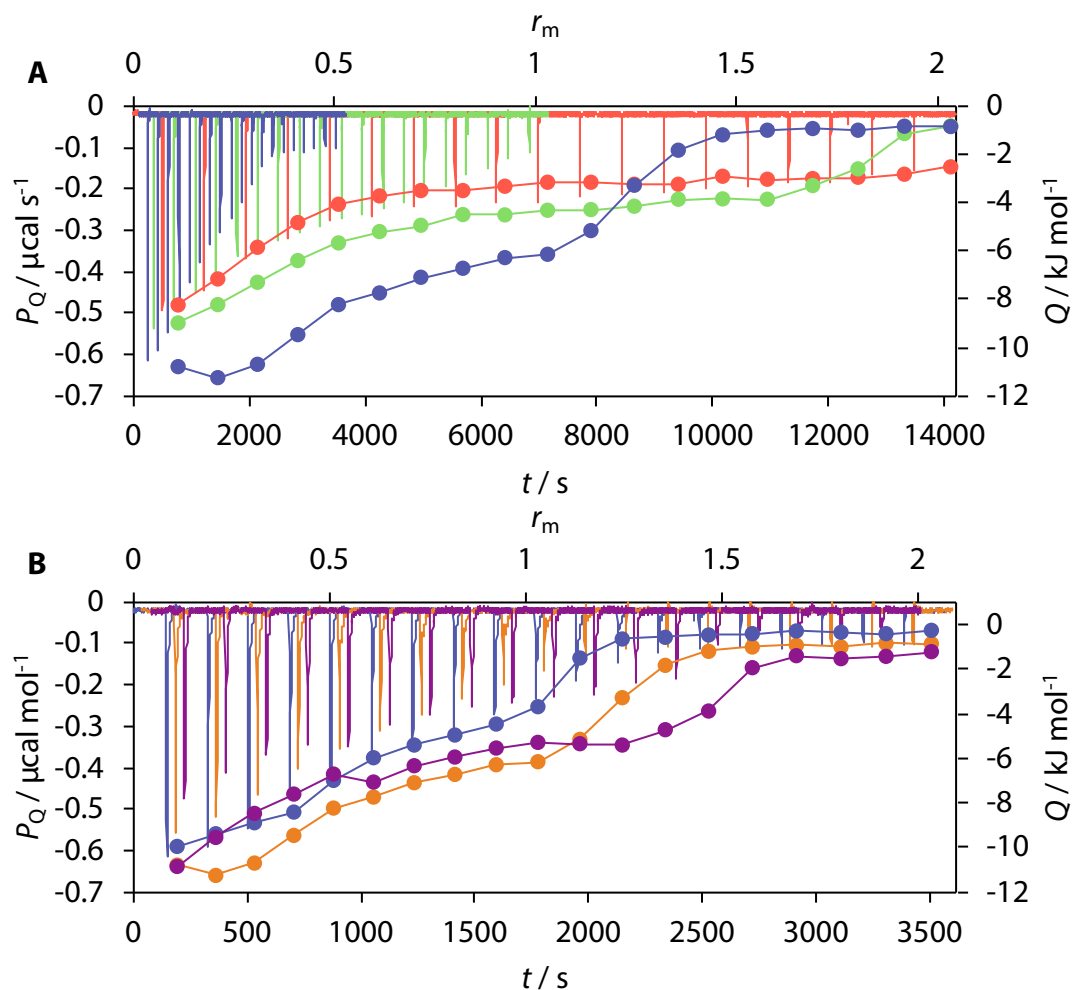
Fig. 45: A Comparison of different time intervals of 180 s (blue), 360 s (green) and 720 s (red) for the titration of **1** into hCAII. The blue and red thermograms are shift by 100 s and 200 s, respectively, for clarity. **B** Comparison of different temperatures of 25 °C (blue), 15 °C (orange) and 35 °C (purple) for the titration of **1** into hCAII. The orange and purple thermograms are shifted by 40 s and 80 s, respectively, for clarity.

Given the unusual, and in literature likely unprecedented observation, that the two inflection points found for the binding of **1** to hCAII are not integral multiples of each other, a binding model of two binding sites could be abandoned. Chemically, the sample was shown to be 99.5 % pure by quantitative NMR spectroscopy by *Hans-Dieter Gerber*, which rules out an impurity as a cause for the above. It was furthermore assumed, that the lack of coherence between the values of n_s of the inflection points is caused by a non-homogeneous ligand sample in the sense, that compound **1** adopts a second conformation in solution, which is also capable of binding to hCAII, which is symbolized by the superscripts 1 and 2 in equation (18). Furthermore, the

protein can adopt two different states that bind **1** differently (equations (19). The latter is unlikely, however, given the high structural rigidity of hCAII, but cannot be fully excluded because of the finding of the mobility of Thr200, which, however, only seems to occur after a ligand has bound, not before (chapter 5).

$$PL^1 \rightleftharpoons P + L^{1/2} \rightleftharpoons PL^2 \qquad (18)$$

$$P^1L \rightleftharpoons P^{1/2} + L \rightleftharpoons P^2L \qquad (19)$$
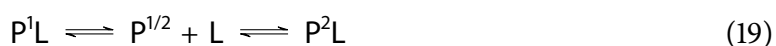
It was tried to fit models described either by equation (18) or (19) to the data under the assumption, that either the fractions of the respective ligand species add up to one (equation (20)), or that those of the respective protein species add up to 0.7 (equation (21)).

$$r_{L1} + r_{L2} = 1 \qquad (20)$$

$$r_{P1} + r_{P2} = 0.7 \qquad (21)$$

Condition (20) is based on the assumption, that the ligand has a purity of 100 % for simplicity. Condition (21) is based on the observation, that for 1 : 1 interactions reported herein, hCAII usually possessed an active fraction of approximately 0.7. Tab. 8 shows the thermodynamic parameters obtained from the fitting procedures.

Notably, both models enable a reasonably high goodness of fit (GoF) of 86.7 %. Moreover, the enthalpic values are in good agreement between both models, although $\Delta H°$ is relatively large for the first reaction of both models compared to other enthalpic signatures for the binding of BSAs reported herein. Also, the values of K_a of the respective first reactions, although not within error margin, are similar. The value of K_a for the respective second reaction shows, that it is thermodynamically less favored than the first reaction. Despite this difference, both models show

Tab. 8: Thermodynamic parameters for the fitting of models (18) and (19) to the titration data of the interaction between **1** and hCAII with errors of the fitting algorithm. Enthalpic values were converted into kJ mol^{-1}. The fractions of protein for model (19) do not add up to one.

Model	Fraction of protein	Fraction of ligand	Complex	K_a / M^{-1}	$\Delta H°$/kJ mol^{-1}
(18)	0.35	0.37	PL^1	$(5.52 \pm 0.85) \cdot 10^8$	-86.1 ± 4.9
	0.65	0.63	PL^2	$(7.99 \pm 2.93) \cdot 10^5$	-32.3 ± 2.3
(19)	0.40	0.43	P^{1}L	$(3.85 \pm 0.31) \cdot 10^8$	-91.7 ± 1.5
	0.63	0.57	P^{2}L	$(1.95 \pm 0.18) \cdot 10^7$	-35.1 ± 0.6

a partitioning of the formed two possible species of about 40 % and 60 %. The circumstance, that the more stable complex is formed to a lower extent than the less stable one, leads to the conclusion, that the latter must feature a higher association rate. Under this prerequisite, the change of titration profiles with an increased time span between injections (Fig. 45**A**) can be rationalized by the transformation of the kinetically favored hCAII–**1** complex (hCAII–**1**^{kin}) into the thermodynamically favored hCAII–**1** complex (hCAII–**1**^{td}). This is possible, theoretically, in two ways. Either the hCAII–**1**^{kin} species disassembles in favor of a subsequently formed hCAII–**1**^{td} species, or *via* conversion of the former into the latter without complete disassembly, although a rebinding scenario is also conceivable (subsection 1.2.2).

6.1 Conclusions

The findings in this chapter describe, to the best of my knowledge, the calorimetrically unprecedented observation, that two conformations of one ligand have different thermodynamic and kinetic binding signatures. Structurally, these findings will be difficult to elucidate on the basis of MX by either soaking or cocrystallization, given the usual time spans for crystallization of hCAII (1 d) and soaking (12 h) of hCAII crystals, which are long time spans compared to the longest ITC run performed herein with 720 s of spacing between injections. A more promising method in this respect might be NMR, as it is capable of providing information in the solvated state. Further

information might be gained by the investigation of the (*S*)-enantiomer. If the enantiomer shows similar integrated heat values as the (*R*)-enantiomer **1**, this could indicate an effect in solution to exert influence on the binding, as both enantiomers can be expected to show the same behavior in solution. Furthermore, the achiral derivative of **1** without the CH_3 group might shed light on whether the chiral center is a prerequisite to the assumed conformational characteristics of **1** and its complex with hCAII. Additionally, a small molecule crystal structure of **1** might provide insight into the preferred conformation for this ligand. While the preferential binding of bioactively preorganized ligands is a well-accepted principle in SBDD, it usually compares different chemical entities as opposed to the findings for compound **1** presented herein.[36]

7 A Chiral Zn^{II} Binder

In 2017, *Chrysanthopoulos et al.* reported on the discovery of an until then unknown Zn^{II}-coordinating non-BSA CAI with a 2,4-oxazolidinedione scaffold (Fig. 46).[168]
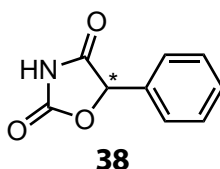
Fig. 46: Structure of 5-phenyloxazolidine-2,4-dione discovered by *Chrysanthopoulos et al.* as novel CAI motif.

Notably, only the (*R*)-enantiomer of **38** was found in the crystal structure.[168] Herein, it was tried to assess, whether the (*S*)-enantiomer also binds to hCAII both by crystallography as well as by means of ITC. Therefore, the racemic mixture of **38** synthesized by *Khang Ngo* was subjected to chiral HPLC by *Sebastian Hickert* from Merck KGaA, Darmstadt, to separate the enantiomers. The chromatogram (Fig. 47) shows apparent separation of the two enantiomers.
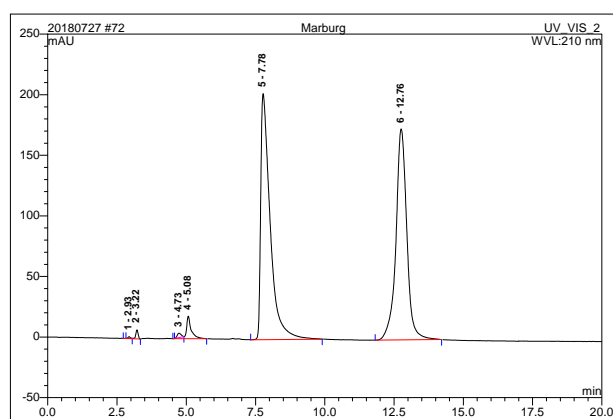
Fig. 47: Chromatogram of the separation of the racemic mixture of **38**.

Soaking of both enantiomers overnight, however, resulted in only the (*R*)-enantiomer being observed in the crystal structure, also in the crystallographic model of crystals soaked with the (*S*)-enantiomer. Considering that both structures showed an increased resolution compared to the already available crystallographic models 5TXY and 5TY8, the structure from the crystal soaked with the (*R*)-enantiomer was deposited in the PDB as entry 6SFQ. The observation, that only the (*R*)-enantiomer was found even for the crystal structure of the crystal soaked with the (*S*)-enantiomer gave rise to the assumption, that the binding to hCAII might promote epimerization. Thus, the concentrations of the enantiomers in the soaking drop were increased by increasing the amount of DMSO to decrease soaking time and thereby prevent putative racemize. However, even after 1 min of soaking, only the (*R*)-enantiomer was observed bound to the (*S*)-soaked crystal. Fig. 48 shows the comparison of the deposited model 6SFQ and the structures derived from crystals that were soaked shortly at high ligand concentrations of either (*R*)- or (*S*)-enantiomer.
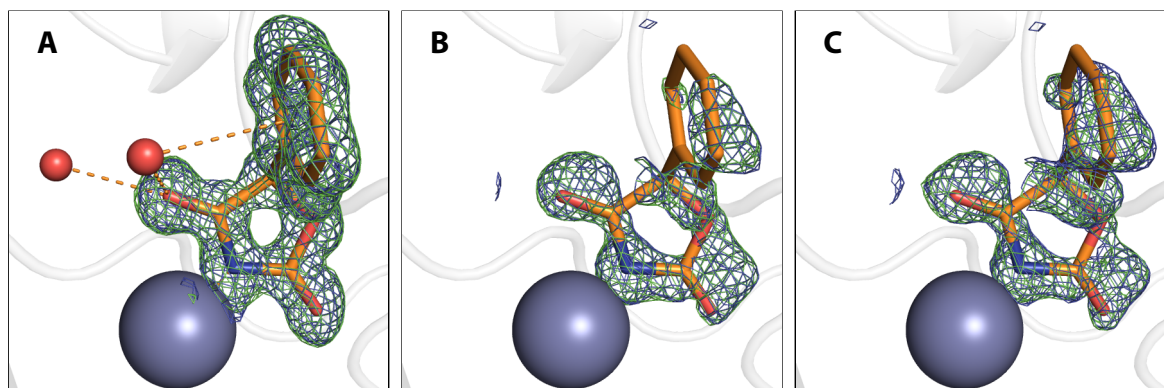
Fig. 48: Crystallographic model of compound **38** bound to hCAII. **A** Fully refined and deposited model 6SFQ with omit 2mF_o−DF_c map in blue at 1σ and mF_o−DF_c map at 3σ. 6SFQ is shown with electron density maps of the crystal soaked with (*R*)−**38** (**B**) and (*S*)−**38** (**C**).

Clearly, the crystal soaked with the (*S*)-enantiomer exhibits the (*R*)-enantiomer in the active site. It is unlikely, however, that racemization occurs in buffer at pH 7.8, given the pK_a of **38** of 11.96 ± 0.09. This allows for two possible scenarios, by which (*R*)−**38** finds its way into the crystal initially soaked with (*S*)−**38**. Either, the enantiomers were not fully separated by the HPLC run. Thereby, presuming a distinctly larger affinity of the (*R*)-enantiomer, even a small

impurity of (R)–**38** in the (S)–**38** sample can lead to the mere observation of the former bound in the active site of hCAII. The second possible scenario is the racemization of either enantiomer after binding to the Zn^{II} cofactor. The plot of integrated heat values from an ITC experiment with (R)–**38** provides putative evidence for the latter scenario (Fig. 49**A**).
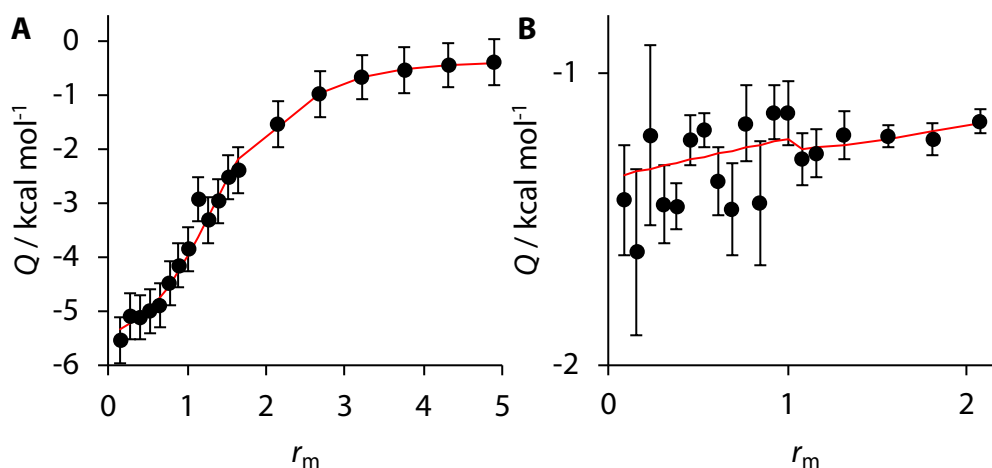
Fig. 49: ITC thermograms for the titrations of (R)-**38** (**A**) and (S)-**38** (**B**) into hCAII.

Fig. 49 **A** clearly shows binding between (R)-**38** and hCAII. Notably, however, the stoichiometry of the reaction amounts to $n_s = 1.4$. This is twice the approximate value of n_s usually found in the course of this work, and indicates an impurity of the ligand. This impurity might be, in fact, the non-binding (S), that is generated throughout the experiment from the (R)–**38**, which delays saturation of the protein. It is also conceivable, that the impurity was present in the ligand sample before the experiment. However, the chromatogram in Fig. 47 does not suggest a highly abundant impurity. Fig. 49**B** shows the titration of hCAII with (S)–**38**. Clearly, (S)–**38** does not bind to hCAII. Given the argument, that (R)–**38** racemizes within the active site, this is also conceivable for the (S)-enantiomer. However, seeing that the injection heat values in Fig. 49**B** do not distinctly change in magnitude, it is unlikely, that (S)–**38** racemizes rapidly to yield the more potent (R)–**38**. This can be seen as a further indicator for the theory, that only coordination to the Zn^{II} cofactor enables racemization, putatively because of increased acidity of the CH unit.

It was furthermore tried to assess the propensity of **38** to racemize in presence of hCAII by polarimetry. However, despite the degassing of the aqueous buffer that was used for the preparation of the experimental solution, formation of gas bubbles was still observed in the cuvette, which makes the investigation of the above problem infeasible under these conditions. The usage of a model system, such as a solution of a Zn^{II} salt and imidazol in a polar organic solvent as replacement for hCAII might be more preferable in this case.

8 Concentration-dependent Soaking to Assess the Occupancy of a Second Binding Site

The prediction of binding sites of proteins can facilitate the discovery of starting points for a drug design project, especially if the target of interest does not crystallize readily and renders an experimental structural approach difficult. In the group of *Antonella Di Pizio*, the prediction of a second binding site of a known CAI was envisaged. A complex of hCAII with compound **39** had previously been characterized crystallographically and deposited in the PDB as entry 4FVO with **39** occupying the active site as well as a remote site (Fig. 50).
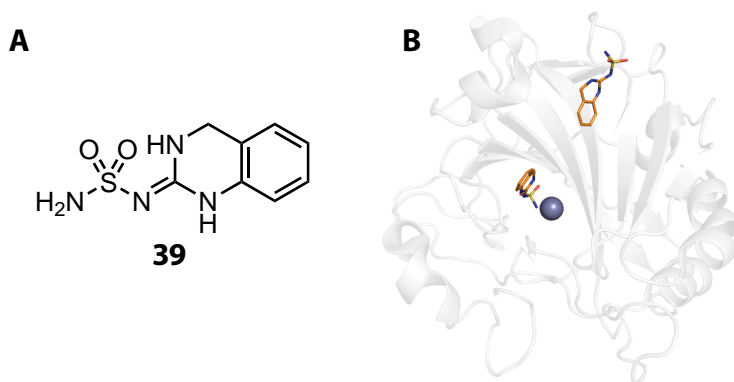
Fig. 50: A Structure of compound **39**. **B** Crystallographic model of the hCAII–**39** complex from PDB entry 4FVO with two occupied binding sites.

It was thus tried to assess the occupancy of the ligand in the second binding site as a function of concentration by soaking similarly sized crystals of hCAII over night in different concentrations of **39**. However, no concentration provided sufficient electron density to merit the modeling of

39 in the second binding site. Fig. 51 shows a close-up of the second binding site in PDB entry 4FVO with $2mF_o-DF_c$ maps and mF_o-DF_c maps from crystals soaked in solutions containing compound **39** in concentrations of 5 mM and 0.1 mM.
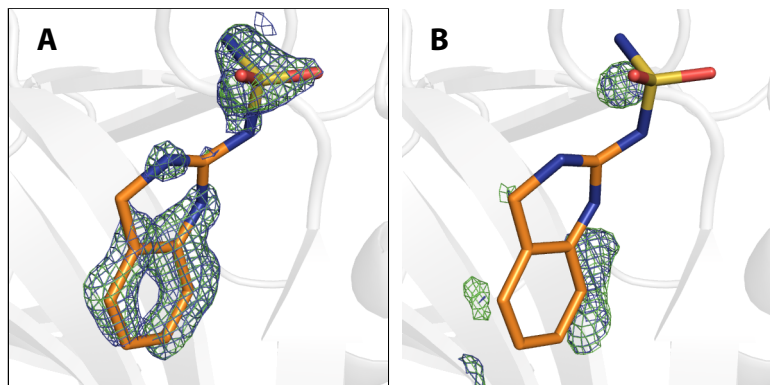
Fig. 51: Closeup of the second binding site of PDB entry 4FVO with electron densities of complexes of hCAII with compound **39** soaked at concentrations of 5.0 mM (**A**) and 0.1 mM (**B**).

Although different concentrations did not lead to results that could acutally be used for crystallographic modeling of the second binding site, differences in the occupation of the active site could be observed. Tab. 9 shows the applied concentrations and crystallographically observed occupancies of the second binding site.

Tab. 9: Concentrations and resulting crystallographically observed occupancies for hCAII–**39** complexes.

Concentration mM	Occupancy
0.1	0.63
0.5	0.91
1	0.84
5	0.87
10[a]	0.83

[a]Compound **39** recrystallized from the soaking drop at an initial concentration of 10 mM.

Given the small to negligible differences in occupancies for soaking concentrations of 0.5 to 5 mM, it was decided to deposit models originating from 5.0 mM and 0.1 mM in the PDB as entries 6SDH and 6SDI, respectively (Fig. 52).
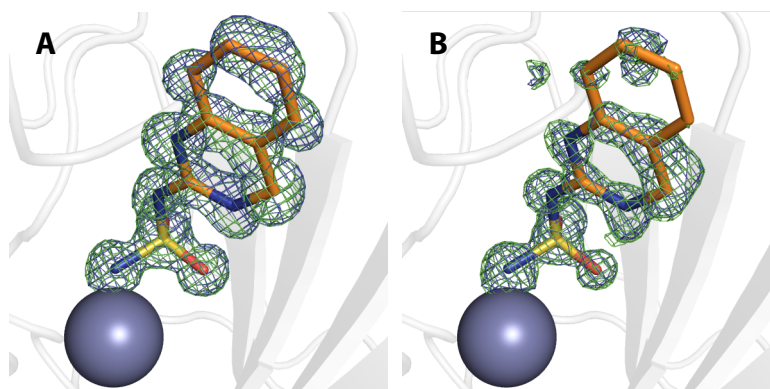
Fig. 52: Crystallographic models of the hCAII–**39** complex produced from soaking at concentrations of 5.0 mM (**A**, 6SDH) and 0.1 mM (**B**, 6SDI).

9 Investigation of Furosemide Analogs

The loop-diuretic furosemide (**40**) binds to the sodium-potassium-chloride transporter present in the thick ascending limb of the loop of *Henle* preventing the reabsorption of sodium, chloride and water.[169] It is known, that this sulfonamide-bearing drug also inhibits hCAII.[170] Because of its substitution pattern it adopts a different binding mode than unsubstituted BSA to interact with the sidechain of His64 and orient the chlorine substituent in a hydrophobic pocket (Fig. 53**A**). Given the distinct role of His64 as proton shuttle in the catalytic mechanism, it was hypothesized, that it might also be engaged in the dissociation of inhibitors, seeing that these need to be reprotonated after dissociation from the Zn^{II} cofactor.[171] For this purpose, it was tried to assess, whether simpler and commercially available derivatives of **40** could be used to assess the influence of His64 on the dissociation velocity. Therefore, complexes of hCAII with compounds **41** and **42** were produced to determine, whether the nitro functionality is capable of hydrogen bonding to the side chain of His64, and shows a distinct difference in the dissociation rate compared to **41**. The choice for a nitro compound was made to remove a protolytically active functionality, that would potentially influence thermodynamics and kinetics of binding to hCAII.[147] However, crystallography revealed, that the nitro moiety is accommodated on a part of the hydrophobic wall, rather than engaged in hydrogen bonding, and thus adopts the same binding mode as **41** (Fig. 53**B** and **C**).

Given these observations, compounds **41** and **42** are not suitable for the elucidation of the influence of His64 in the dissociation mechanism. *Scott et al.* already observed, that a chlorine substituent adjacent to the sulfonamide group leads to a binding mode similar to that of **40**.[99]
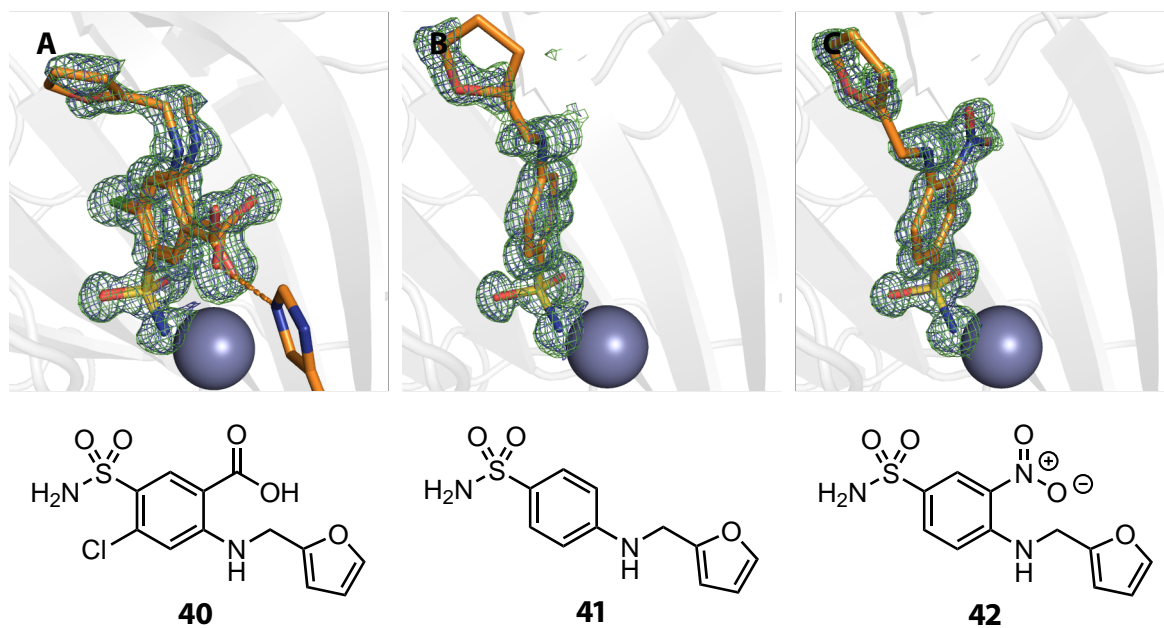
Fig. 53: A Complex of hCAII with furosemide (**40**). The carboxylate function interacts with the side chain of His64, the chlorine substituent points into a hydrophobic pocket. **B** hCAII in complex with furosemide derivative **41**. **C** hCAII in complex with compound **42**. The nitro moiety does not engage in hydrogen bonding to the side chain of His64.

Thus, the introduction of a chlorine substituent in compounds **41** and **42** in *ortho*-position to the sulfonamide, and in the case of **42** *para* to the nitro compound, is likely to induce binding geometries similar to that of **40**, which would then enable hydrogen bonding between hCAII and His64.

10 Investigation of the Effect of a Remote-Binding Sulfoxide on Active-Site Inhibitor Binding

D'Ambrosio et al. reported sulfoxide **43** to bind to the side chain of His64 outside the active site of hCAII and demonstrated, that it exerts inhibitory action (Fig. 54**A**).[172] Crystallography revealed, that, instead of coordinating to the Zn^{II} cofactor, it engages in a water-mediated contact with the side chain of His64 (Fig. 54**B**).
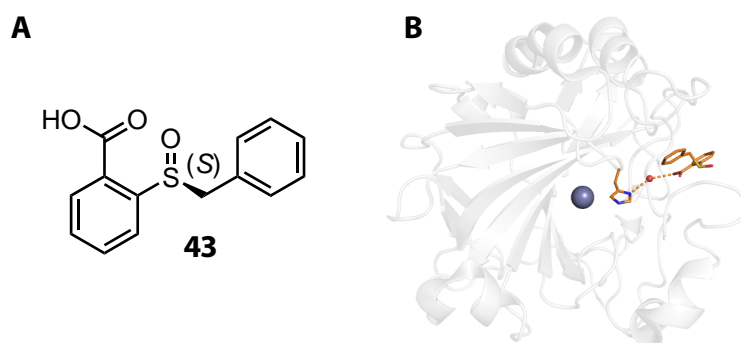
Fig. 54: A Remote-acting sulfoxide CAI **43**. **B** Binding of **43** to the side chain of His64, mediated by a water molecule (4QY3).

D'Ambrosio et al. reported an inhibitory constant for **43** of 150 nM. Based on this finding, it was tried to assess the affinity of **43** for hCAII by means of ITC. The resulting thermogram, however, did not show any binding event. Given that this can be caused by small heat signals (entropy-driven binding) and does not exclude binding *per se*, it was then tried to assess, whether **43** has an impact on the binding of the known ligand **37** (Chapter 5). The titration of **37** into **43**

did not show an interaction of a defined stoichiometry, which excluded interaction between both compounds, rendering them suitable for an experiment to test cooperativity. Thus, hCAII was mixed with an excess of **43** and titrated with **37**. Unexpectedly, the presence of **43** reduced the affinity of **37** (Tab. 10).

Tab. 10: Values of K_d and $\Delta H°$ for the characterization of the hCAII–**37** complex with and without sulfoxide **43**.

Compounds	K_d / nM	$\Delta H°$/ kJ mol^{-1}
37/43	790 ± 60.9	-53.9 ± 2.7
37	685 ± 40.7	-51.7 ± 0.8

Despite the circumstance, that no defined interaction between **37** and **43** could be observed, unspecific interactions might still be the cause for the reduced affinity of BSA **37**.

Part II

Growth of Crystals for Neutron Diffraction

11 Growth of Crystals Suitable for Neutron Diffraction

The artificial sweetener saccharin is a known inhibitor of hCAII (Fig. 55).[173]
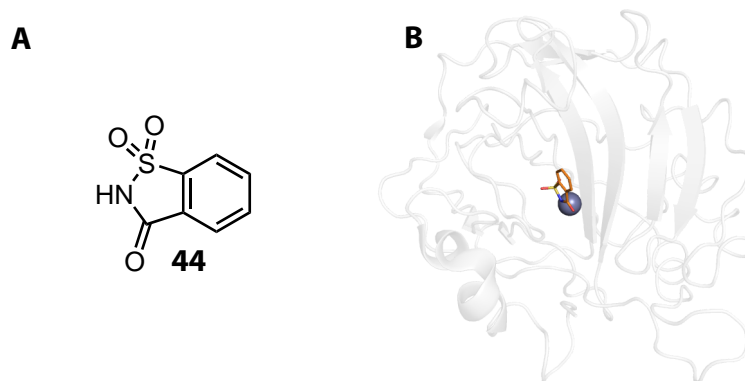
Fig. 55: A Structure of the artificial sweetener saccharin **44**. **B** Saccharin in the active site of hCAII (2Q1B).

As saccharin features an *N*-acyl sulfonamide function instead of a primary sulfonamide, the nitrogen atom has only one valence left to neutrally bind a hydrogen atom. Under the premise, that sulfonamides usually coordinate the Zn^{II} cofactor in the anionic state, saccharin likely looses a proton as well before complexation, given its comparably high acidity ($pK_a = 1.6$).[174] However, this does not have to hold true necessarily, considering the fact that the supposedly simple question of the protonation state of the Zn^{II}-bound oxygen species has not yet been doubtlessly answered. While most estimations conclude, that the pK_a value of the Zn^{II}-bound water molecule is 6.8, there is also experimental evidence that suggests a pK_a of 5.[129] This would entail that the Zn^{II}-bound water molecule is deprotonated in solution for pH > 5, or from pH 6.8 onwards. However, neutron diffraction studies of hCAII showed, that the fourth vertex of the

tetrahedral coordination sphere of the Zn^{II} species is, in fact, occupied by a water molecule in the crystalline state, even at pH 9.0.[151] It is thus not unlikely, that also saccharin, despite its comparably high acidity, binds to hCAII in the protonated state.

11.1 Screening of Already Known Crystallization Conditions

It was thus tried to obtain hCAII crystals sufficiently large for neutron diffraction, as both protons and deuterons scatter neutrons, and might thereby answer the question for saccharin's protonation state in complex with the Zn^{II} cofactor. This was to be done with perdeuterated hCAII provided by *Himanshu Singh* from the group of *Rasmus Linser*. To assess feasibility, already reported crystallization conditions based on trisodium citrate were investigated with non-deuterated hCAII.[175] Given that these conditions produced large crystals without the addition of PCMB in literature, it was desisted from its usage for the initial trials. Fig. 56 shows, however, that PCMB was necessary for the production of voluminous crystals with hCAII from own production, that have a clear expansion in all three spatial dimensions, as the lack of PCMB only produced shard-like two-dimensional crystals.
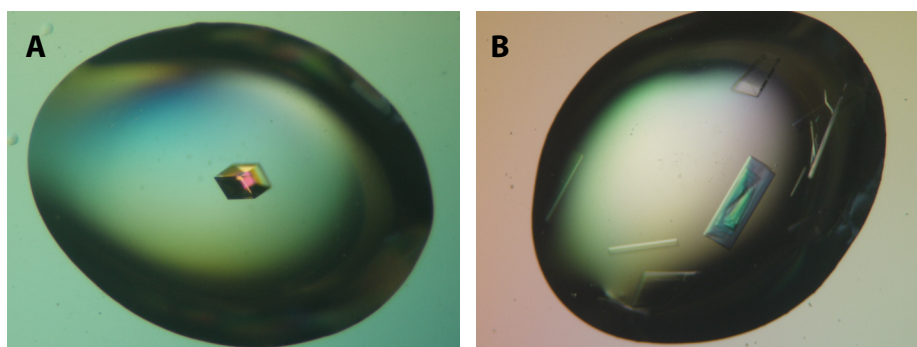
Fig. 56: HCAII crystals grown in 1.40 M trisodium citrate solution with PCMB (**A**) and without (**B**).

Notably, the crystallization with PCMB not only provided for voluminous three-dimensional crystal growth, but also enabled sufficiently fast crystallization, which prevented the formation of additional nuclei. Considering that larger crystals require a larger amount of protein than

crystals commonly used for XRD, the protein concentration was increased from $5\,\text{mg}\,\text{mL}^{-1}$ to $16.9\,\text{mg}\,\text{mL}^{-1}$ to provide more protein material for crystallization. Additionally, crystallization was carried out in both H_2O-based as well as D_2O-based buffer at various concentrations of trisodium citrate (Fig. 57).
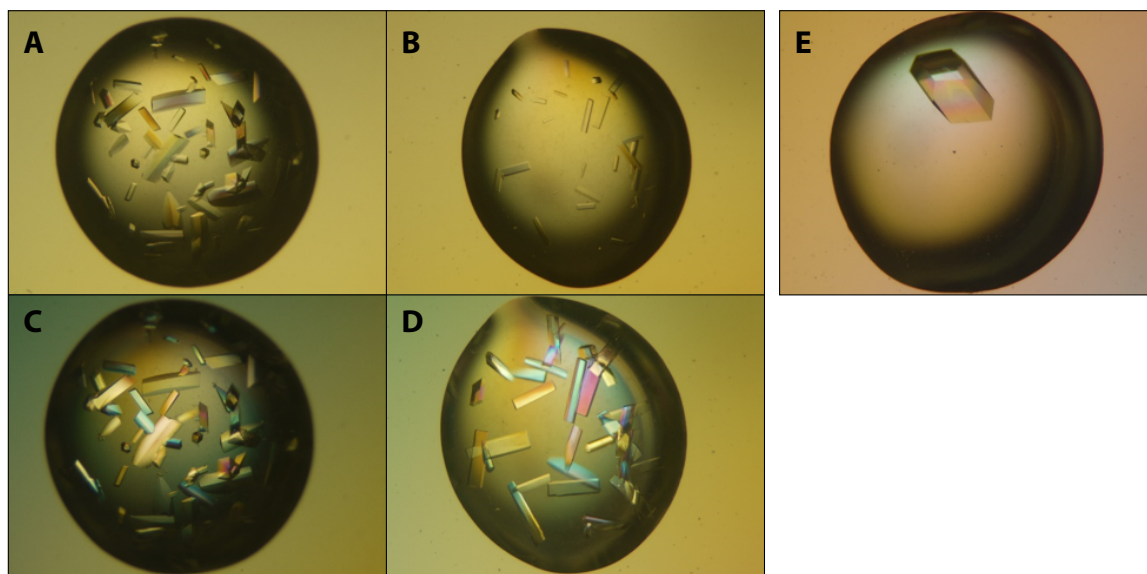
Fig. 57: Crystals of hCAII grown in different precipitant conditions. **A** 1.40 M trisodium citrate, H_2O after 24 h. **B** 1.40 M trisodium citrate, D_2O after 24 h. **C** 1.40 M trisodium citrate, H_2O after 48 h. **D** 1.40 M trisodium citrate, D_2O after 48 h. **E** 1.25 M trisodium citrate, H_2O after 6 d.

Due to the increased protein concentration, more crystals formed at a concentration of trisodium citrate of 1.40 M. This is counterbalanced by a decreased precipitant concentration of 1.25 M trisodium citrate, which yielded one large crystal. Notably, deuterated buffer decelerated crystal growth. Furthermore, the diffraction power had to be assessed. Fig. 58 shows diffraction images of crystals grown in H_2O-based and D_2O-based conditions recorded with Cu-K_α radiation on a MAR345 image plate detector.
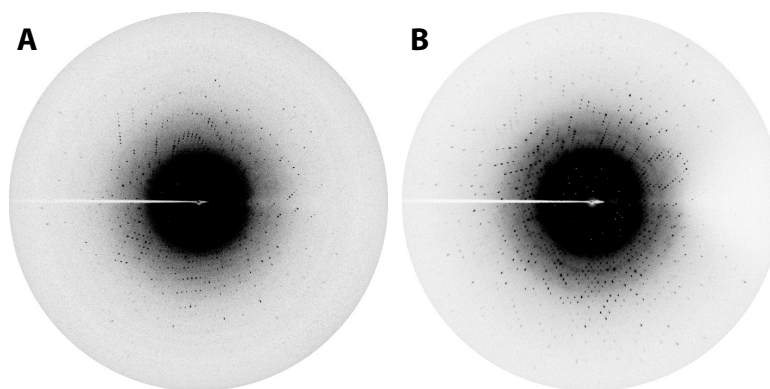
Fig. 58: A Diffraction image of an hCAII crystal grown in H_2O-based 1.25 M trisodium citrate, exposed for 3 min. **B** Diffraction image of an hCAII crystal grown in D_2O-based 1.40 M trisodium citrate, exposed for 5 min.

Due to the prolonged exposure time of the crystal grown in deuterated medium, the diffraction image shows a high number of overloaded spots, discernible as light spots in the bulk diffraction. It can therefore be assumed that deuteration does not distinctly alter the diffraction power. Subsequently, crystal growth was tried on a larger scale. Therefore, drops were set up in plastic culture dishes with a precipitant concentration of 1.35 M, a drop volume of 400 µL, a protein concentration of 5 mg mL^{-1} and a reservoir volume of 3.0 mL. This approach did not turn out satisfactory, however, given the large number and small size of crystals (Fig. 59**A** and **B**). Furthermore, crystallization was tried with protonated protein in a Hampton Research 9 well glass plate sandwich box setupwith a precipitant concentration of 1.30 M, with a drop volume of 400 µL, a protein concentration of 16 mg mL^{-1} and a reservoir volume of 25 mL, which yielded crystals with a length of up to 2 mm in the longest dimension (Fig. 59**C** and **D**).
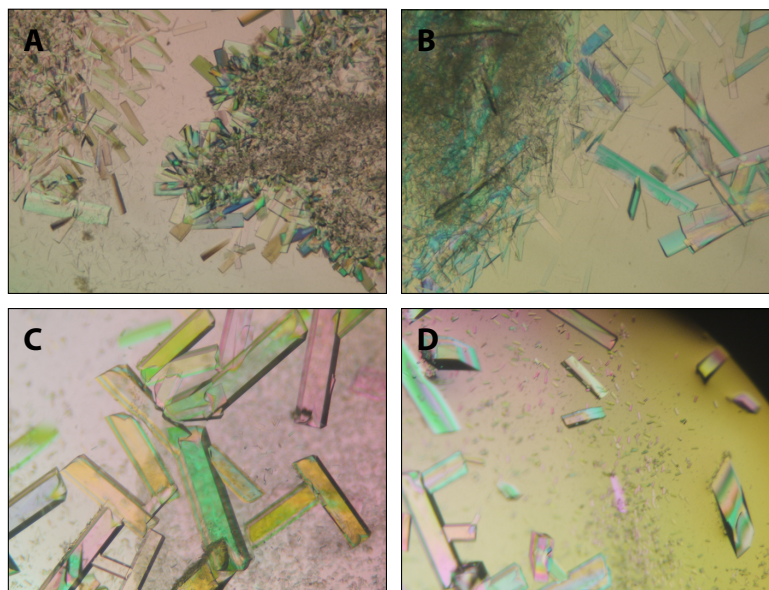
Fig. 59: Crystallization trials in **A** a culture dish in protonated medium and **B** deuterated medium, **C**, **D** in a sandwich box setup in protonated medium.

11.2 A Joint X-ray and Neutron Diffraction Structure of a Carbonic Anhydrase IX Mimic

In 2019, *Koruza et al.* published the crystallographic model of a mimic of the cancer-associated hCA isoform IX (hCAIX), which was constructed from hCAII by seven point mutations and was characterized by joint neutron and X-ray diffraction data.[176] Interestingly, also complexes of the hCAIX-chimera in complex with saccharin and a derivative were reported with a picture of the nuclear density of the active site, and it was argued, that saccharin coordinates the Zn^{II} cofactor in the deprotonated form.[176] However, considering the circumstance, that nuclear density of the remaining hydrogen atoms of saccharin is weak to non-existent, the argument for binding in the deprotonated form is not doubtlessly proved. Under the prerequisite, that for high-resolution MX models of hCAII hydrogen atoms can be observed in well-ordered regions, the solution of the phase problem was tried by SAD and directly at a resolution of 0.99 Å to obtain electron density maps without model-bias in order to find a clue for putative protonation of saccharin in complex with hCAII. Therefore, SAD phasing was conducted with the programs

SHELXC/D/E as combined in the hkl2map software suite and was additionally tried directly with the program ACORN from the CCP4i suite.[177,178] Fig. 60 shows the electron density maps from the experimentally determined phases of the SAD experiment (**A**) and from the phases determined directly from known starting coordinates (**B**).
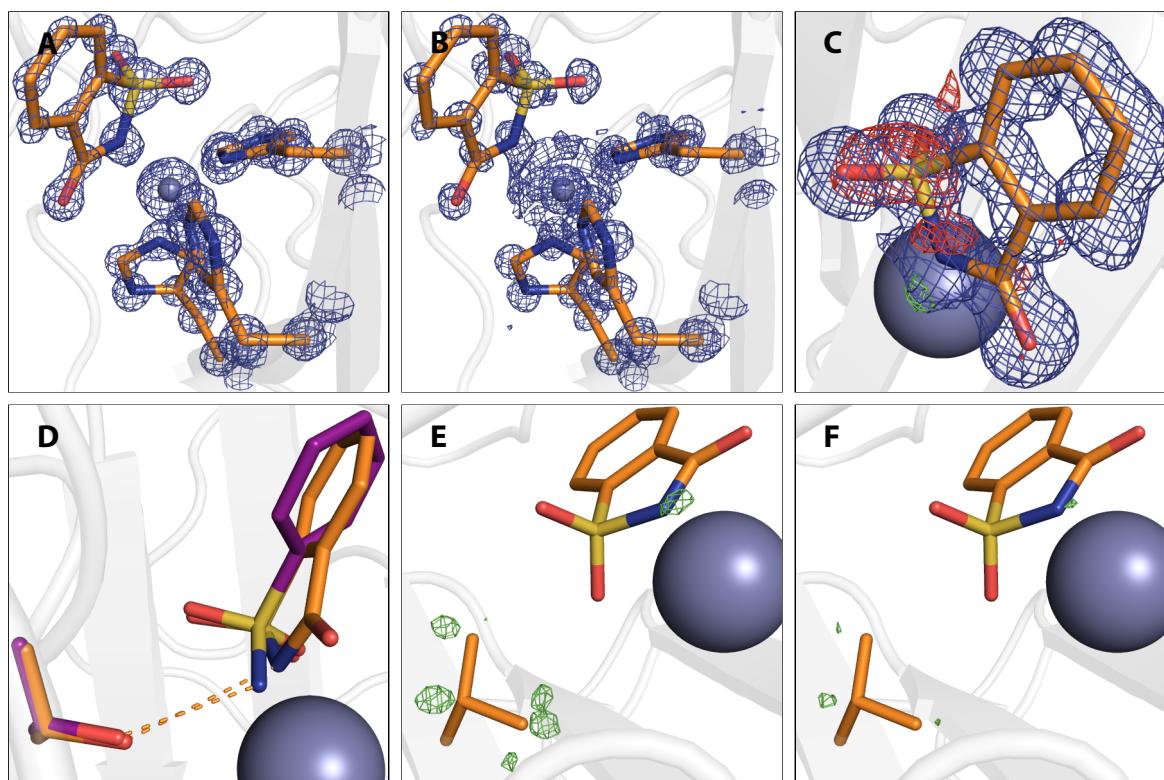
Fig. 60: A $2mF_o$–DF_c map as output by SHELXD at 3 σ. **B** $2mF_o$–DF_c map of the active site as output by ACORN at 3 σ. **C** Electron density maps after refinement. $2mF_o$–DF_c map in blue at 1 σ, mF_o–DF_c maps in green at 3 σ and in red at −3 σ. **D** Comparison of the active sites of the hCAII–**44** complex in orange and the hCAII–**4** complex (6GDC). (Putative) hydrogen bonds are indicated as dashed orange lines. **E** Saccharin and the side chain of Val143 shown as sticks with mF_o–DF_c density in at 3 σ. **F** Saccharin and the side chain of Val143 shown as sticks with mF_o–DF_c density at 3.6 σ.

Protons did indeed become visible in the active site during refinement. Fig. 60**C** shows electron density maps after the last refinement run without hydrogen atoms. A small peak of difference density was found at the position where a proton might be found. A comparison of the binding mode of saccharin with that of compound **4** (Fig. 60**D**) shows, that the hydrogen-bond distance (d_{hb}) between the nitrogen atom of the ligand and the hydrogen bond accepting hydroxy function of Thr199 is increased in the complex with saccharin (d_{hb} = 3.3 Å) compared to the hCAII–**4** complex (d_{hb} = 2.9 Å), which allows for the inference of a weak hydrogen bond in the

hCAII–**44** complex. Fig. 60**E** and **F** furthermore show, that the positive difference density peak close to the nitrogen atom of **44** has approximately the same magnitude as the peaks caused by hydrogen atoms of the side chain of the structurally well conserved residue Val143, which can be seen as an argument for the validity of the assumption, that this peak is indeed caused by a proton. This is furthermore complemented by the approximate distance of 0.9 Å between the density peak and the nitrogen atom of the saccharin molecule is close to the distance of 0.92 Å observed for the distance between the nitrogen hydrogen atoms in a small-molecule crystal structure of saccharin.[179]

Part III

Fragment Screening against Human Carbonic Anhydrase II

12 Introduction to Fragment-Based Drug Design

The initial step that has to be taken for the development of a small-molecule drug is the determination of a starting molecule, which can be elaborated into the eventual drug molecule. In the early 1990s, large chemical libraries were established for high-throughput screenings (HTSs) to accelerate this step by screening millions of drug-like compounds in biochemical and biophysical assays against a target of interest.[180] This resulted in marketed drugs against cancer, the HIV and diabetes, among others.[180] A shortcoming of HTS is the narrow applicability with respect to the nature of putative drug targets.[181] Another issue, potentially causative for the former, is the inefficient coverage of chemical space that is available for an HTS library of compounds comprising up to 30 atoms.[181,182] This resembles the typical size of drug-like molecules, and was estimated to enable more than 10^{60} conceivable compounds.[183] Fragment-based drug design (FBDD) relies on smaller molecules called fragments, despite their chemical integrity, as they can be seen as parts of drug-like molecules because of an atom count smaller than 20.[181] Similar to the statistical physicochemical description of drug molecules by *Lipinski*'s Rule of Five (RO5), early fragment libraries were assembled under adherence to the Astex Rule of Three (RO3).[184,185] It states that fragment-like molecules have a molecular weight (MW) < 300 Da, a number of hydrogen bond donors and acceptors ≤ 3, respectively, and clog$P$ ≤ 3.[185] Given the smaller chemical space of $166 \cdot 10^9$ molecules conceivable for molecules made up of 17 atoms of the elements C, N, O, S, F, Cl, Br, I and H, it is clear, that fragments cover the the chemical space more efficiently, which provides them with a wider applicability.[182,186] The rationale for the

superiority of fragments compared to drug-like molecules in the lead-finding process is based on the assumption that, based on the RO3, binding of a fragment to a target can only occur through a small number of interactions and thus with weak affinity.[182] However, the formed interactions are usually of high quality as the low spatial complexity of the fragments enables an optimal orientation within the binding site and thus enables supreme geometrical complementarity between the interacting functional groups of fragment and target while the chance of repulsive interactions of non-optimally placed moieties is reduced.[72,182,185] A fragment hit can then be used to establish a lead scaffold by the approaches of fragment growing, merging and linking.[182] The goal of fragment growing is to address additional functional groups in the binding pocket to enhance affinity by increasing molecular size.[182] Fragment merging and linking initially require at least two fragment hits. For merging, the two fragments need to address different parts of the binding site, but at the same time overlap with one portion. Linking requires two fragments to occupy different parts of the binding site without overlapping. The latter two methods by trend pose more of a challenge than the mere growth of the intial fragment scaffold, as the geometric requirements of both fragments need to be taken into account to maintain the efficient binding of both portions in the final follow-up molecular scaffold.[182] Especially the development of a suitable linker that does not impose an entropic penalty nor induce a hydrophobic collapse in solution with a distinct deviation from the biologically active conformation, is often challenging.[182] Although the approach of FBDD drastically reduces the number of compounds necessary to cover as much chemical space as possible, it still requires a high screening effort. The method of SAR by NMR and MX are the most prominent techniques used in this respect.[187–189] MX, despite the high demand of protein material, the necessity to grow suitable crystals and the dependence on the availability of measurement time at synchrotron facilities, is an especially powerful tool, as it is capable of detecting very weakly binding fragments and delivers information about the geometry of binding poses and the nature of the formed interactions.[182] In the laboratory of *Gerhard Klebe*, a 361-entry fragment library targeted at the model aspartic protease endothia-

pepsin was established, that does not strictly adhere to the RO3, but also considers the chemical feasibility of the generation of follow-up compounds from fragment hits.[190] Furthermore, this study revealed that a strict adherence to the RO3 would have excluded several hits.[190] This library was investigated again to compare the hit rates of different biophysical assays with MX and found, that MX elucidated binding events that would have been missed by other techniques, and advocates for crystallography as the prime, because the most promising, tool in FBDD.[191] Although the majority of clinical candidates still originates from the elaboration of already known binders and random HTS, the advantage of SBDD in combination with FBDD is apparent, given the relatively short period of six years it took to market the first fragment-based drug vemurafenib with antimelanomal activity. Fig. 61 shows the initial fragment-sized 7-azaindole (a prominent motif in kinase-targeted fragment libraries) derivatives **45** and **46** in complex with the Pim-1 kinase and the structurally related kinase domain of the fibroblast growth factor receptor 1 (FGFR1) as well as Vemurafenib in complex with its target B-Raf kinase.[192,193]
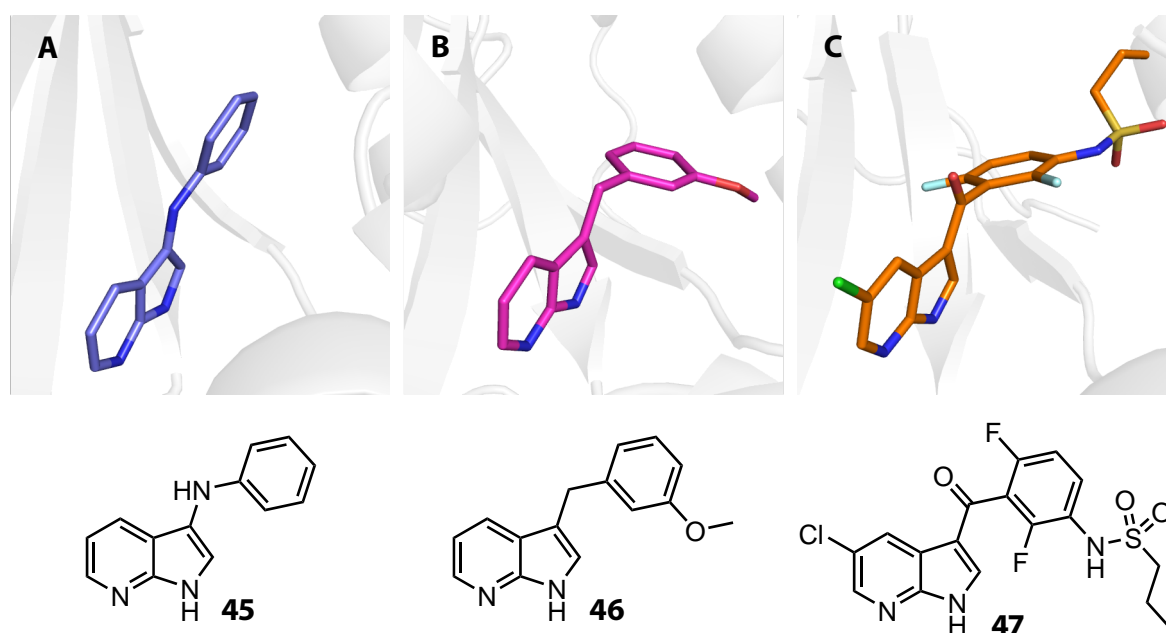
Fig. 61: Fragment hits in structurally related kinases that led to the development of Vemurafenib. **A** 7-azaindole derivative **45** in the active site of the Pim-1 kinase (3C4E). **B** 7-azaindole derivate **46** in the active site of the kinase domain of FGFR1 (3C4F). **C** Vemurafenib (**47**) in complex with B-Raf kinase (3C4C).

The inital lead discovery with fragments does not necessarily accelerate the entire process. The second marketed fragment-based drug Venetoclax entered the market after 20 years of development.[194] However, fragments are a promising approach to accelerate drug discovery, taking into account that for many HTS-derived drugs a development time around 10 years is common.[180]

13 Proof-of-Concept Fragment Screening of a 96-Entry Fragment Library

In an effort to make the power of fragments available to a broader public, a 96-entries fragment library was assembled in the laboratory of *Gerhard Klebe*. It comprises subsets of the 361-fragments library mentioned above, of a library of representative molecules from the PDBeChem database from the AG *Weiss* at the Helmholtz-Zentrum Berlin as well as 3D-enriched fragments.[190,195] Given the fact that approximately 50 % of the 96 compounds were taken from the aspartic-protease targeted 361-entry in-house library, it is valid to ask whether the assembly is broadly applicable in the lead discovery against targets of various functions. This question was addressed by screening the 96-entry assembly against various established model proteins of the *Klebe* laboratory, such as hCAII. Experimental details can be found in subsection 14.3.2.

13.1 Results

Out of 96 compounds, the eight compounds depicted in Fig. 62 were found in the crystal structure of hCAII. Four of these occupy the active site, one is found on the edge of the funnel-shaped active site. The remaining three fragments are positioned in between hCAII crystal mates, with one benefiting from a contact with the crystallization agent PCMB, the other two, surprisingly, attached covalently to the N-terminus. The binding of fragments **48**, **49**, **51** and

52 was described in a preceding Master thesis, but will be mentioned here for completeness.[196]

Fig. 62 gives an overview of the occupied sites.
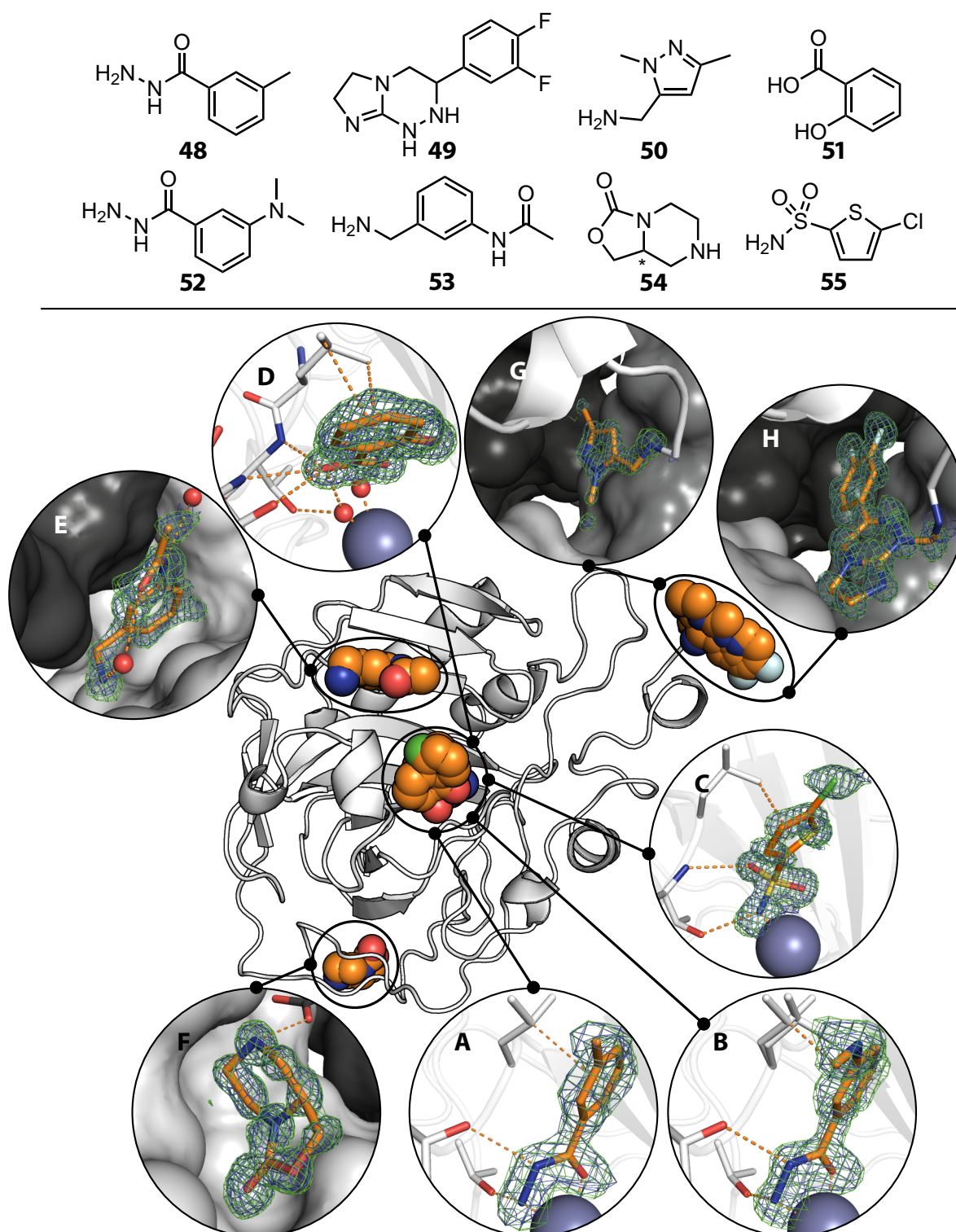
Fig. 62: Upper panel: Valence bond formulae of bound fragments. Lower panel: Overview of fragment binding sites. Fragments are shown in orange with omit 2mF_o–DF_c maps in blue at 1σ and mF_o–DF_c maps in green at 3σ. Active site binders are shown with the interacting amino acids in stick represension. Remotely binding fragments are shown on with the surface representation of adjacent protein molecules. Hydrazides **48** (**A**) and **52** (**B**) and sulfonamide **55** (**C**) coordinate to the Zn^{II} cofactor, salycilic acid (**51**, **D**) coordinates to the Zn^{II}-bound water molecule. Anilide **53** (**E**) is accommodated in a hydrophobic cavity on the rim of the funnel-shaped active site, that is covered by a proline residue of a symmtetry mate. Cyclic carbamate (*S*)-**54** (**F**) coordinates to the C-terminus and a Trp sidechain, and donates a hydrogen bond to the PCMB molecule of a symmetry mate. Pyrazole **50** (**G**) is directly covalently attached to the N-terminus, fluorinated compound **49** (**H**) *via* a linker.

135

Hydrazides **48** and **52** occupy the active site of hCAII and coordinate the Zn^{II} cofactor with the N_{β} and carbonyl O atom of the hydrazide function, which results in a pentacoordinated cofactor. The phenyl rings are involved in a π-interaction with the side chain of Leu198. Notably, the coordination geometry enables a hydrogen bond between N_{β} of the hydrazide group, and the side chain hydroxy function of Thr199. This entails a different orientation of the phenyl moiety compared to BSA (**4**), which triggers a flip of the sidechain of Leu198, that is also present in the complex of hCAII with *N*-hydroxybenzamide, discovered by *Di Fiore et al.* (Fig. 63).[197]
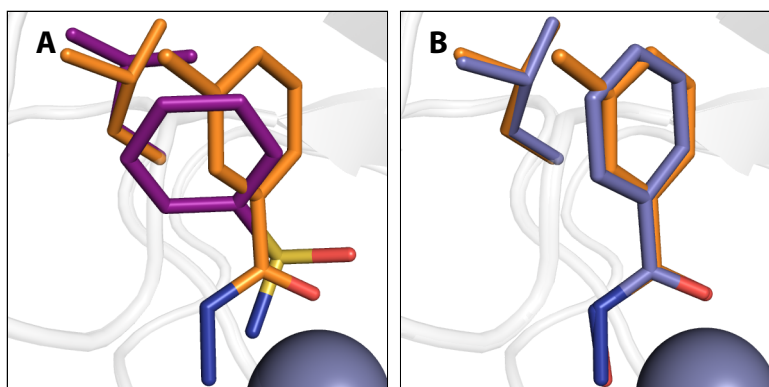
Fig. 63: A Comparison of hydrazide fragment **48** (orange) and BSA **4** (purple) and the respective side chains of Leu198. **B** Comparison of **48** and the analogous hydroxamate (blue) (4FL7).[197]

Salycilic acid **51** binds in a fashion that was observed previously for other hydroxybenzoic acids in complex with hCAII. The carboxylate function interacts *via* a hydrogen bond with the Zn^{II}-bound hydroxide ion/water molecule. The carboxylate function furthermore interacts with the side chain of Thr200, a state similar to the prebinding S-state defined by *Gaspari et al.* for the binding of BSAs and already observed by *Martin* and *Cohen* for variously hydroxylated benzoic acid derivatives.[198] Without surprise, sulfonamide **55** can be expected to bind to hCAII, and does so with the expected geometry. Primary amine **53** was found to occupy a hydrophobic cavity on the rim of the active site funnel. It is held in place by a charge-assisted hydrogen bond with the side chain of Asp72 and furthermore interacts with a water molecule, which mediates a contact between the anilide N-atom and the side chain of Glu69. Additionally, the phenyl interacts with the side chain of Ile91 *via* the π-system, which is furthermore involved in an interaction with

the sidechain of Pro237 of a symmetry mate. Moreover, an interaction between a lone pair of the O-atom of the symmetry mate's amide bond between Pro237 and Glu236 in a *Bürgi-Dunitz* type trajectory can be inferred.[103,104,199] Thus, the binding of **53** is likely to be an artifact enabled by the crystal packing, rather than by genuine affinity for the position on the rim of the hCAII active site itself. Fragment **54** was used as racemate. The (*S*) enantiomer was found to bridge the gap between C-terminus and the sidechain of Trp192. Furthermore, the secondary amine interacts with the carboxylate function of the crystallization agent PCMB of a symmetry mate. Whereas the binding of fragment **49** had been mentioned already in the preceding master thesis, the binding of **50** was only noticed after renewed inspection. Notably, both compounds are covalently attached to the N-terminus. In the case of **49** this occurs *via* a covalent linker of unknown identity. Given the circumstance that the soaking conditions contained polyethylene glycol (PEG) 400, it is reasonable to assume, that the linker is a methylene unit which was installed by incorporation of one molecule of formaldehyde, which is known to be present in solutions containing PEG 400.[200] A possible reaction is shown in Scheme 1.
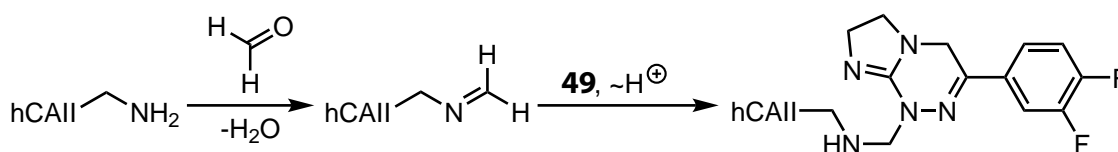
Scheme 1: Possible mechanism for the reaction between the N-terminus of the hCAII-expression tag protein with **49**. The N-terminal primary amine reacts with formaldehyde under loss of water to yield an imine. This in turn is nucleophilically attacked by **49** with concurrent proton transfer to yield the final protein-small-molecule hybrid.

Fragment **49** itself is involved in some interesting interactions. It is known, that the S-atom of a Met side chain is often found in the vicinity of π-donors, such as aromatic rings and amide bonds with which it engages mostly through dispersive interactions.[201] Especially the nucleobase adenine is often found in company of Met side chains in MX models. While the π-system of adenine is distributed over the whole molecular scaffold, this cannot be the case for compound **49**, given, its three CH_2 units. However, the hydrazinecarboximidamide substructure and the difluorophenly moiety, which are connected by a C_{sp^2} atom, provide an extended π system, which can be expected to act as binding partner for the Met side chain S-atom. Given the comparably

close distance of 3.3 Å between the $O_{\delta 1}$ atom of the side chain of Asp71 of a symmetry mate and the covalently linked N-atom of **49**, an interation can be inferred which resembles that between an Asp side chain and the positively charged aromatic scaffold of a false-positive fragment screening hit against endothiapepsin, discovered by *Schiebel et al.*[202] It is reasonable to assume that it is, in fact, a charged interaction, given that the hydrazinecarboximidamide substructure can be considered basic, thus bearing a positive charge that can be distributed over several atoms (scheme 2), and which will furthermore be beneficial for the interaction between **49** and the S-atom of Met1.
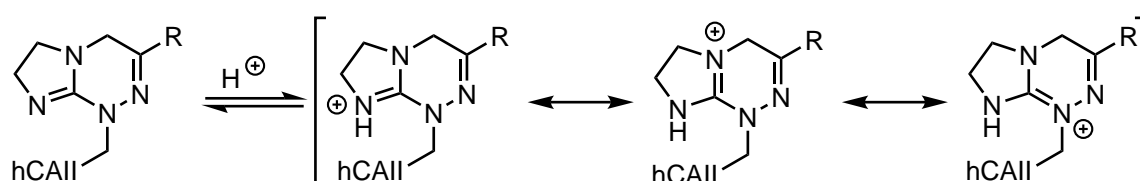
Scheme 2: Protonation of the small-molecule moiety of the hCAII–**49** hybrid and possible resonance structures. R represents the 3,4-difluorophenyl motif.

Furthermore, the side chain of Ser73 of the symmetry mate might interact with the N_3 atom of **49**, that is involved in the delocalization of a positive charge. An interaction that is often found for fluorinated compounds is the interaction between an F-atom and an amide bond.[102] The F-atom of **49** exerts such an interaction with the amide bond between Gln185 and Lys159 in a symmetry mate at a distance of 3.2 Å. Furthermore, the distance of 3.2 Å between the F_1 atom of **49** and the $C_{\delta 2}$-atom of Leu57 of a different symmetry mate is in the same range as the sum of the *van der Waals* radii of an aliphatic C-atom and an aromatically bound F-atom of 3.17 Å (see chapter 5).[76] It was tried to reproduce the binding of **49** in a 1.50 M trisodium citrate buffer, both with and without DMSO. Fig. 64 clearly shows, that molecule **49** is not found in the site described above in either condition.

Fragment **50** is also covalently attached to the N-terminus, but without a linker. Reaction between protein and fragment might have occurred through two possible reactions. Either, the unprotonated N-atom of the protein replaces the protonated primary amine function of **50** or the reaction proceeds *vice versa* (scheme 3). The presence of both **49** and **50** in this spot is an
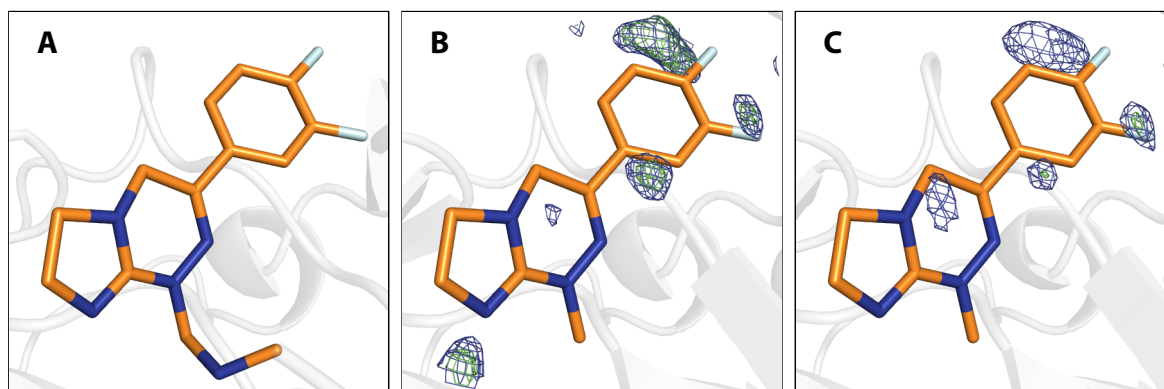
Fig. 64: Close-up of the binding site of **49**. **A** Model of covalently bound fragment **49**. **B** Model obtained from a crystal soaked in the above trisodium citrate conditions including 10 % DMSO after automated refinement superimposed with the binding pose of **49** from **A**. **C** Model obtained from a crystal soaked in the above trisodium citrate conditions including without DMSO after automated refinement superimposed with the binding pose of **49** from **A**. Electron density maps are shown at 1 σ (2mF_o–DF_c, blue) and 3 σ (mF_o–DF_c, green).
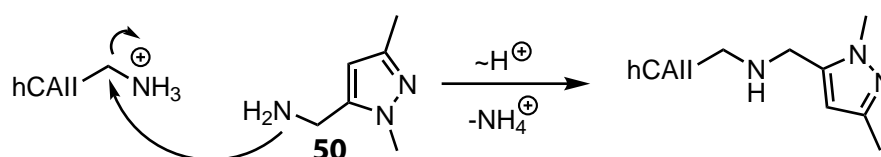
Scheme 3: Possible mechanism for the reaction between the N-terminus of the hCAII-expression tag protein with **50**. Under the assumption of a protonated N-terminal primary amine, the primary amine of compound **50** replaces the N-terminus, putatively in an S_N2 reaction given the electron withdrawing nature of the adjacent amide bond omitted here.

indicator for favorable accommodation of (aromatic) π-systems. Notably, the sidechain of Met1 is not visible in the presence of **50** and might suggest that the π-system of **56** is too small to interact favorably, or that it lacks a positive charge.

13.2 PanDDA Analysis

Weak population of binding sites results in weak density for the bound ligand, which renders its identification by visual inspection of the electron density maps firstly speculative, and secondly highly tedious, if not impossible. For a sufficiently large number of crystallographic datasets of the same protein, it is nowadays possible to unveil such binding events by the Pan-Dataset Density Analysis (PanDDA) method.[203] PanDDA, in short, relies on a density 'ground-state', which is generated as an average of the datasets included in the analysis.[203] A proportion of the ground-state is then substracted from each individual dataset to reveal 'changed states', such as

ligand binding.[203] The hCAII datasets collected in the course of the above fragment-screening campaign were subjected to PanDDA analysis. Surprisingly, only fragments **48**, **51** and **52** out of the eight identified hits and no additional binding events were found in this analysis, although visual inspection after a run of the in-house automated refinement pipeline by *Dr. Johannes Schiebel* clearly identified the residual five binders. A putative reason for the low success rate was assumed to be a non-optimal averaging of the ground-state, which was putatively caused by the fact, that soaking times ranged from 15 h at most to 3 min minimum. This became necessary, as it was found that the exposition of crystals to several fragments led to partial or even complete loss of diffraction power, sometimes without visible degradation of the crystal. Such a distinct difference in soaking time can be expected to cause differences, e.g. in the degree to which a binding site, that accommodates solvent molecules, is occupied. Therefore, a PanDDA run was carried out only with datasets from crystals, that had been submerged in the soaking drop for 15 h, which resulted in the detection of fragments **48**, **49**, **51**, **52**, **53** and **55**. This analysis still missed fragments **50** and **54**, but supports the theory of the influence of different soaking times on electron density. However, analysis of datasets only soaked for 3 min revealed not a single hit, but should at least have identified compound **57**. Fig. 65 shows the PanDDA event maps for the identified binding events.

13.3 Conclusions

The active site of hCAII is occupied by four out of 96 possible compounds, three of which directly coordinate to the Zn^{II} cofactor. Compounds **48** and **52** are, although not surprisingly found given the inhibitory action of the structurally highly similar *N*-hydroxybenzamide, to the best of my knowledge, the first hydrazide inhibitors of hCAII deposited in the PDB. Although probably not relevant for medicinal chemistry purposes, the binding of four additional fragments in various sites enabled by crystal packing partially revealed that this fragment library is also suited for binding sites that feature less common binding partners such as the sulfur atoms of Met
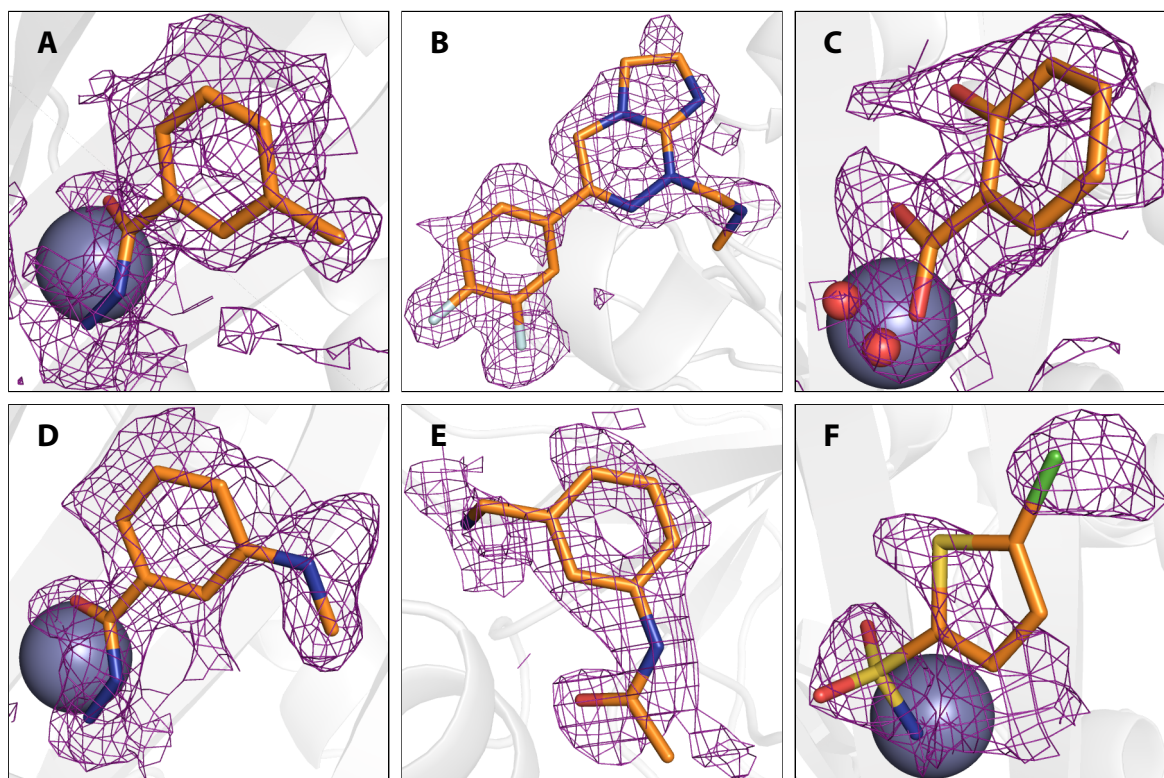
Fig. 65: PanDDA density maps for identified fragment binding events shown at 2 σ. **A 48**. **B 49**. **C 51**. **D 52**. **E 53**. **F 55**.

and Cys side chains and has the potential to reveal possibilities for interactions that might not be initially apparent, such as interactions between amide bonds as for **53** or the interactions between fluorine atoms and either amide bonds or aliphatic groups.

A PanDDA analysis of all 96 datasets showed a low success rate of only three identified events out of eight bound fragments, that were found by visual inspection of the electron density maps after performing automated refinement. A renewed PanDDA run with a subset of datasets obtained from crystals that were all soaked for 15 h showed an increased number of identified binding events of six bound fragments, and thereby supports the notion, that different soaking times bear differently on the electron density distribution in crystals. An explanation for this might be, that a soaking time of 15 h is sufficiently long to enable full equilibration of all unit cells with the surrounding medium. In this context, the unsuccessful recovery of a known binder from a subset of datasets, that were exposed to the soaking medium for a drastically shorter

period of only 3 min means, that equilibration of the crystal with the soaking drop had not been concluded. This can be expected to result in distinct differences in the distribution of electron density across unit cells, and is possibly the more pronounced the larger the crystal is. This last reasoning would furthermore entail not only drastic differences within one, but across several crystals and might well be an explanation for the unsuccessful PanDDA run with datasets obtained after a soaking period of 3 min.

14 Methods and Materials

14.1 Buffer Compositions

Tab. 11: Compositions of expression and purification buffers used in the course of this work.

	LB		PBS Expression		TRIS Expression		TRIS/NaCl Expression	
	c_m g L^{-1}	c_n mol L^{-1}	c_m g L^{-1}	c_n mol L^{-1}	c_m g L^{-1}	c_n mol L^{-1}	c_m g L^{-1}	c_n mol L^{-1}
KCl			0.2	$2.7 \cdot 10^{-3}$				
KH$_2$PO$_4$			0.25	$1.8 \cdot 10^{-3}$				
Na$_2$HPO$_4$ (anhydr.)			1.4	$10 \cdot 10^{-3}$				
NaCl	10	$17 \cdot 10^{-2}$	8.1	014			58.4	1.0
Pepton	10	–						
TRIS					6.06	0.05	12.1	0.1
Yeast	5	–						
ZnCl$_2$	$8 \cdot 10^{-3}$	$60 \cdot 10^{-6}$						
pH			7.3		7.8		7.4	

Tab. 12: Compositions of experiment buffers used in the course of this work.

	$(NH_4)_2SO_4$ Crystallization		Na_3-Citrate Crystallization		HEPES ITC	
	c_m g L^{-1}	c_n mol L^{-1}	c_m g L^{-1}	c_n mol L^{-1}	c_m g L^{-1}	c_n mol L^{-1}
$(NH_4)_2SO_4$	264.3/396.4	2.0/3.0				
HEPES					2.38	0.01
Na3-Citrate			294.1/411.8	1.0/1.5		
	7.8		7.3		7.8	

14.2 Protein Expression and Purification

14.2.1 Expression

HCAII was expressed in *Escherichia Coli* BL21 Codon plus cells. 100 µL ethanolic Chloramphenicol solution (34 mg mL^{-1}) and 100 µL aqeous Amipicillin sodium salt (Ampicillin-Na) solution (100 mg ml^{-1}) were added to 100 mL lysogeny broth (LB) medium, which was then incubated with cells carrying the pGEX-4T1 plasmid with the hCAII-GST fusion protein gene overnight at 37 °C. The cells were taken from a glycerol stock prepared by *Dr. Johannes Schiebel.* To 1 L of LB medium was added 1 mL of aqueous Amipicillin sodium salt (Ampicillin-Na) solution (100 mg ml^{-1}) and 20 mL of overnight culture. Expression culture growth was monitored with a NanoDrop$^{\text{TM}}$ 2000c. The expression culture was allowed to grow until an optical density (OD) of 0.6 at a wavelength of 600 nm was reached. Expression of the hCAII-GST fusion protein was induced by addition of 1 mL 1 M isopropyl-β-D-thiogalactopyranoside (IPTG) solution. 1 mL aqueous 0.94 M ZnCl$_2$ solution was added, following a procedure by *Cimmperman et al.*, and the expression was conducted for 4 h at 32 °C.[204]

14.2.2 Lysis

Cells were harvested by centrifugation (10 000 rpm, 10 min, 4 °C) and either stored at −80 °C overnight or directly used for further purification. Cell pellets were resuspended in 150 mL PBS containing one protease-inhibitor cocktail pill. Cell disruption was carried out in centrifugation vessels by sonication (Branson Sonifier 250). Every vessel was sonicated five times for 1 min at a duty cycle of 70 and an output control of 3. Cells were spun down by centrifugation (20 000 rpm, 45 min, 4 °C).

14.2.3 Purification

The lysis was loaded onto a Glutathione Sepharose$^{\text{TM}}$ High Performance column at a flow rate of 1 mL min^{-1} via peristaltic pump and afterwards washed with 400 mL of PBS at a flow rate of

3 mL min^{-1}. Cleavage of the GST tag was achieved overnight after the column was manually loaded by syringe with 10 mL PBS containing 500 IU of thrombin at an approximate flow rate of 1 mL min^{-1}. The protein was eluted in 25 mL of PBS via fast protein liquid chromatography (FPLC) at a flow rate of 1 mL min^{-1} and a pressure of 0.4 MPa. Removal of thrombin from the flow through was achieved by dilution with the same volume of 0.1 M TRIS buffer containing (1 M) NaCl at pH 7.4, and running the mixture through a 5 mL HiTrap™ Benzamidine FF(HS) column. The flow through was purified by size exclusion chromatography with a HiLoad™ 26/60 Superdex200 size exclusion column and the buffer exchanged to 0.05 M TRIS buffer, pH=7.8. Protein purity was monitored by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Fig. 66) as described by *Laemmli*.[205]
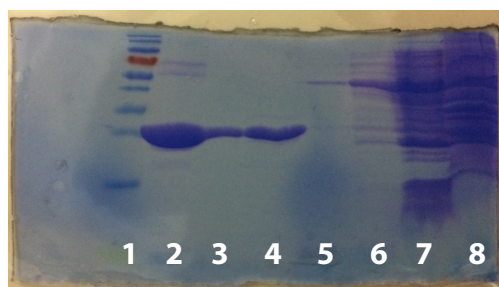
Fig. 66: SDS-PAGE gel for the purification of hCAII. **1** Ruler (molecular weight increases from top to bottom). **2** Eluate of SEC column. *3* Eluate of benzamidine column. *4* Eluate of GSH column. **5** Application of thrombin to loaded GSH column. **6** Flowthrough of washing GSH column. **7** Flowthrough of loading of GSH column. **8** Cell pellet after lysis.

14.3 Macromolecular Crystallography

14.3.1 Crystallization

Unless stated otherwise, crystals were grown in the following two conditions.

14.3.1.1 Ammonium Sulfate Conditions

Crystallization was carried out in hanging drop plates from Hampton Research in 24-well format at 18 °C. Crystallization drops were set up by mixing 2.0 μL of protein solution in the final expression buffer (c_P = 10 mg mL^{-1}) with 2.0 μL of a solution of $(NH_4)_2SO_4$ 2.25 – 2.70 M , TRIS

0.1 M, pH 7.8 and saturated with PCMB on siliconized 22 mm circular cover slides from Jena Bioscience. The reservoir contained 0.5 mL of the $(NH_4)_2SO_4$ solution used for the drop. Crystals grew overnight and reached their final size after one day.

14.3.1.2 Trisodium Citrate Conditions

Crystallization was carried out in hanging drop plates from Hampton Research in 24-well format at 18 °C. Crystallization drops were set up by mixing 2.0 µL of protein solution in the final expression buffer (c_P = 10 mg mL^{-1}) with 2.0 µL of a solution of trisodium citrate 1.40 M, TRIS 0.1 M, pH 7.8 and saturated with PCMB on siliconized 22 mm circular cover slides from Jena Bioscience. Crystals grew within one to two days and reached their final size after three days.

14.3.2 Soaking and Cryoprotection

Crystals grown in $(NH_4)_2SO_4$ conditions were soaked overnight in a solution of 3.0 M $(NH_4)_2SO_4$ TRIS 0.1 M, pH 7.8, D-glucose 25 % (w/v) saturated with the ligand of interest. Cryoprotection was conducted by submerging the crystals in a solution containing 3.0 M $(NH_4)_2SO_4$, TRIS 0.1 M, pH 7.8 and saturated with the ligand of interest for 30 s. Crystals grown in trisodium citrate conditions were soaked in a solution of trisodium citrate 1.50 M, TRIS 0.1 M, pH 7.8 and saturated with the ligand of interest, which was also suitable for cryoprotection. Crystals for fragment screening were soaked in conditions from *Dr. Kan Fu*'s PhD thesis with a high PEG content and 10 % DMSO for periods of 15 h or 3 min.[206]

14.3.3 Diffraction Data Collection, Processing and Refinement

Diffraction data were recorded on the macromolecular crystallography beamlines of BESSY II, ELETTRA and PETRA III.[120,207,208] Data indexing, integration and scaling were performed with the programs XDS and XDSAPP2.0.[209,210] A random selection of 5 % of all reflections was used for the calculation of R_{free}. Structure solution was done by molecular replacement using the program Phaser from the CCP4i suite with PDB entry 3KS3 as initial model, except for the hCAII–**44** com-

plex (Chapter 11).[165,178] Model building in Coot into $2mF_o-mF_c$ and mF_o-mF_c maps at 1 σ and 3 σ, respectively, alternating with subsequent refinement steps with the graphical user interface of different versions of Phenix.refine from the Phenix software suite was conducted until the convergence of R-factors.[121,211,212] Double conformations were modeled if appropriate. The refinement included atomic x, y and z coordinates as well as occupancies. Atomic B-factors were initially refined isotropically with subsequent anisotropic refinement of all atoms except water. Lastly, all atoms were refined anisotropically. Riding hydrogen atoms were added with the program ReadySet from the Phenix software suite.[211] The website https://www.molinspiration.com/ was used for SMILES-code generation and coordinate and restraint files for the modeled ligands were created with the Grade Web Server (http://grade.globalphasing.org/cgi-bin/grade/server.cgi). Coordinates and structure factor files were deposited in the PDB under the accession codes given in the crystallographic tables. Crystallographic images were created with the program PyMOL.[161] Atoms are colored according to the following color scheme: H – white, C – diverse, N – blue, O – red, F – light blue, S – yellow, Cl–green, Zn – gray.

14.3.4 SAD Phasing

SAD experiments were carried out on beamline 14.1 at the BESSY II electron storage ring operated by the Helmholtz-Zentrum Berlin. Two datasets were collected on a Dectris Pilatus 6M detector at a wavelength of 0.9794 Å at $\varkappa$-angles of 0° and 30°. Datasets were merged with the program XSCALE.[209] Phasing was performed with the progams SHELXC/D/E as combined in the HKL2MAP bundle.[177] The anomalous signal (d″/σ(d″) = 1.46 – 4.47) allowed the determination of 3 out of 3 possible binding sites of Hg and Zn by SHELXD (CC_{all} = 45.1). SAD phasing was carried out using SHELXE. Structure factor amplitudes from XSCALE and phases from SHELX were combined using the program CAD from the CCP4i suite.[178] An initial electron density map was calculated and a model (3 chains with 3, 74 and 172 residues, respecdtively) was generated *via* the program ARP/wARP as implemented in the CCP4i suite.[178,213] Further model building and refinement were carried out as described in subsection 14.3.3.

14.4 Isothermal Titration Calorimetry

14.4.1 Preparation of DMSO Stocks

30 mm stocks of all ligands used for ITC measurements were prepared following suggestions by *Boyce et al.*[70] The purity of compound **1** was determined by *Hans-Dieter Gerber* by quantitative NMR. *Khang Ngo* determined the purities of all other compounds by analytical HPLC. Purities of the solids were all larger than 95 % and taken into account for the preparation of the DMSO stocks.

14.4.2 General Sample and Experiment Preparation

Experimental solutions were prepared under the proposition of a total volume of 300 μL and a total DMSO content of 3 – 5 % (v/v). Protein was mixed with buffer and subsequently DMSO. The ligand DMSO stock was further diluted with DMSO for a final volume of 9 μL and mixed with 291 μL of buffer. The samples were vortexed shortly and centrifuged. After washing with buffer and drying of the measurement cell, the protein sample was loaded and the excess removed. The syringe was loaded with ligand solution, purged and refilled. A downward movement corresponding to 0.03 μL was performed.[214] Additionally, a smaller injection of 0.3 μL preceded the titration and was discarded before analysis.

14.4.3 *n*-Alkyl and *n*-Alkyloxy Compounds

Prior to ITC measurements, protein was transferred into the measurement buffer (10 mM HEPES, pH 7.8) by centrifugation in a Vivaspin 20 Centrifugal Concentrator containing a 10 kDa molecular-weight cutoff PSE membrane, concentrated to 15 mg mL^{-1}, aliquoted, frozen in liquid nitrogen and stored at −80 °C. Measurements were carried out on an ITC200 (GE, compounds **4** -**8** and **10** -**13**) and an ITC200 (Malvern, **9** and **20**) at 25 °C at a stirring speed of 750 rpm and 180 s of spacing between consecutive injections. For characterization of **9**, a displacement titration setup following suggestions by *Velazquez-Campoy* and *Freire* was applied.[215] HCAII

was incubated with **20** and titrated with **9**. Compounds **16** and **17** could not be characterized by ITC because of a comparably high noise level and artifact signals (p. 61). Compounds **14, 15, 18** and **19** could not be reliably characterized by ITC because of comparably high error values.

14.4.4 Fluorinated and Other Compounds

For compounds **1, 20 – 31, 34 – 36, 38** and **43**, the following procedure was applied. Prior to measurements, protein was dialyzed against 10 mM HEPES buffer at pH 7.8 at 26 °C overnight. The dialyzed protein solution was aliquoted and stored at −80 °C. The dialysis buffer was filtrated through a Thermo Scientific Nalgene Rapid-Flow PES Bottle Top Filter with a pore size of 0.2 μm and used for the measurements. Measurements were carried out on an ITC200 (Malvern) at 25 °C with a stirring speed of 1000 rpm and 180 s of spacing between consecutive injections. For characterization of compounds **29 – 31**, a displacement titration setup following suggestions by *Velazquez-Campoy* and *Freire* was applied.[215] hCAII was incubated with **20** and titrated with the respective ligand.

14.4.5 Data Analysis

ITC data were analyzed using the AFFINImeter software suite.[114] Raw data were processed and thermodynamic parameters determined with the global fit approach. Kinetic data were fitted anew by applying the globally fitted thermodynamic parameters to the raw data. Standard errors were calculated from the values of the individually fitted measurements. For the evaluation of displacement experiments, direct titrations of the ligand to be characterized, direct titrations of the reference ligand and titrations of the actual displacement titrations were included in the global analysis. Thermodynamic and kinetic data as well as thermograms, integrated heat values and equilibration time curves are located in the Appendix.

14.5 Compounds and Materials

Compounds **1**, **16** and **43** were synthesized by *Hans-Dieter Gerber*. Compound **39** was synthesized by Prof. Dr. Manfred Haake. Compounds **9**, **12**, **13**, **14**, **15**, **17**, **18**, **19**, **28**, **29**, **30**, **31** and **38** were synthesized by *Khang Ngo*. All other compounds were commercially available.

Tab. 13: Materials with suppliers used in the course of this work.

Substance	Supplier/ synthesized by	Substance	Supplier/ synthesized by
1	*Hans-Dieter Gerber*	**35**	Alfa Aesar
2	Sigma-Aldrich	**36**	Sigma-Aldrich
3	(anhydr.) Fluka	**37**	abcr
4	Sigma-Aldrich	**38**	*Khang Ngo*
5	Sigma-Aldrich	**39**	Prof. Dr. Manfred Haake
6	Apollo Scientific Ltd.	**40**	Sigma-Aldrich
7	Chempur	**41**	Enamine
8	abcr	**42**	Enamine
9	*Khang Ngo*	**43**	*Hans-Dieter Gerber*
10	abcr	**44**	Sigma-Aldrich
11	Enamine		
12	*Khang Ngo*	Ammoniumsulfate	Roth
13	*Khang Ngo*	Ampicillin-Na	Roth
14	*Khang Ngo*	Betaine-HCl	Sigma Aldrich
15	*Khang Ngo*	Chloramphenicol	Roth
16	*Hans-Dieter Gerber*	Protease Inhibitor cocktail	Roche
17	*Khang Ngo*	DMSO	Roth
18	*Khang Ngo*	D-Fructose	Sigma Aldrich
19	*Khang Ngo*	Glycine	Sigma Aldrich
20	Sigma-Aldrich	HEPES	Roth
21	Fluorochem	IPTG	Roth
22	Alfa Aesar	KCl	Sigma Aldrich
23	Alfa Aesar	K-Na-Tartrate	Grüssing
24	*Khang Ngo*	Na3-Citrate	Roth
25	Alfa Aesar	NaCl	Roth
26	Fluorochem	PCMB	Sigma Aldrich
27	Alfa Aesar	Peptone	Roth
28	*Khang Ngo*	L-Proline	Roth
29	*Khang Ngo*	Sarcosine	Alfa Aesar
30	*Khang Ngo*	D-Sorbitol	Sigma Aldrich
31	*Khang Ngo*	D-Sucrose	Sigma Aldrich
32	Fluorochem	TRIS	Roth
34	Fluorochem	ZnCl2	Sigma Aldrich

14.6 Calculation of Interface Areas

14.6.1 *n*-Alkyl- and *n*-Alkyloxy-Benzenesulfonamides

SES for the respective species were calculated with the program MSMS included in the Python

Molecular Viewer from the MGLTools suite version 1.5.6.[216] The protein part was reduced to

residues Gln92, His94, His96, His119, Val121, Phe131, Val135, Leu141, Val143, Leu198, Thr199, Thr200, Pro202, Leu204, Val207, Trp209 and the Zn^{II} cofactor.

14.6.2 Fluorinated Benzenesulfonamides

Interface areas were calculated with the Linux version of the program dr_sasa in mode 4 with the fluorine atom as ligand. For *meta*-fluorine substituents, the pocket was defined with residues Val121, Phe131, Leu141 and Leu198. For *ortho*-fluorine substituents, the pocket was defined with residues Val121, Phe131, Val143 and Leu198.

14.7 Normalization of *B*-factors

Atomic *B*-factors were calculated in two ways to exclude the putative influence of ligand atoms on one another. Hydrogen atoms on the protein were removed before the calculation. In general, the individual atomic *B*-factors were divided by the average *B*-factor of all atoms protein atoms surrounding the respective atom in a radius of 6.5 Å. One approach additionally involved all ligand atoms in the process, the other excluded those to remove putative bias due to the overall ligand mobility. All normalization operations were performed with the in-house python script normalize_B.py written by *Dr. Tobias Hüfner*, which can be obtained free of charge from https://github.com/wutobias/normalize_B. The python script uses the ParmEd library as of AmberTools17 for file parsing and the NumPy and SciPy libraries for numerical operations.[162,217,218]

14.8 *Ab initio* Solvation Calculations

Calculations were carried out by *Dr. Christoph P. Sager* by the following procedure.

Ab initio solvation free energies of molecules **4** – **8** and **10** - **13** were calculated after geometrical optimization using the density functional theory (DFT) with the M06-2X functional by the

Truhlar group and the 6-31G(d,p) basis-set implemented in Gaussian 16.[219–221] The SMD solvation mode was used with water ($\varepsilon = 78.3553$) as solvent to derive the presented ΔG of solvation.[222]

14.9 Calculation of α

The deviation angle from the normal vector ($\vec{n}$) of the phenyl ring plane (α) indicated by the deviation vector ($\vec{d}$) was measured for the reference conformation of Leu198 using PyMOL with the NumPy library for vector construction and the vg library for vector operations.[161–163] Directional vectors for the phenyl ring plane ($\vec{p}$, $\vec{q}$) were calculated by subtracting the position vector of the phenyl centroid from the respective position vectors of the *ortho* carbon atoms (C_{o1}, C_{o2}). $\vec{n}$ of the phenyl ring plane was determined *via* the cross product of the resulting vectors with the *cross* function from the *vg* library. $\vec{d}$ was calculated by subtraction of the position vector of the phenyl centroid from the position vector of the $C_{\delta2}$ atom of Leu198.
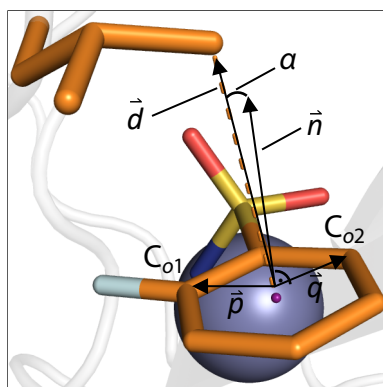
Fig. 67: Calculation of angle *α*.

14.10 pK_a Measurements
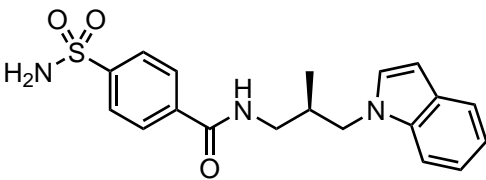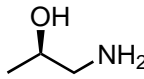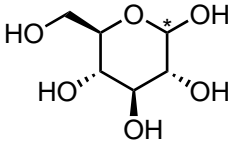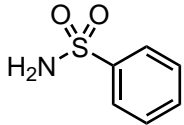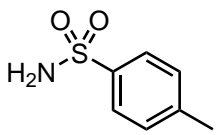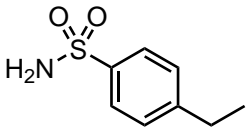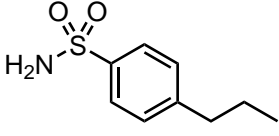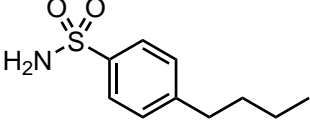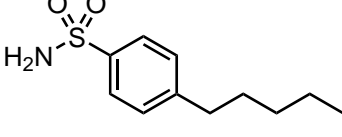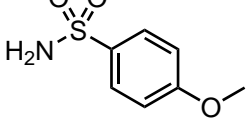
pK_a values were determined by *Björn Wagner* from *Roche* by the SiriusT3 Fast UV pK_a method, which is based on the spectrophotometric (UV-metric) titration method reported in reference [223].

Appendix
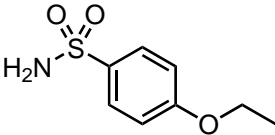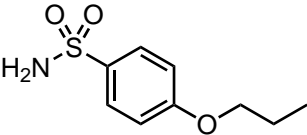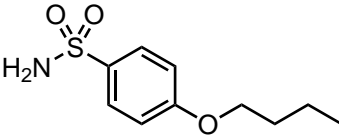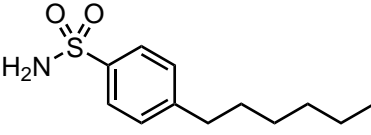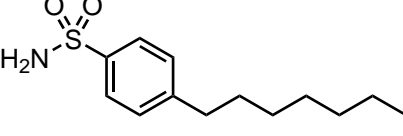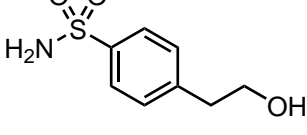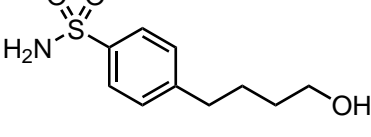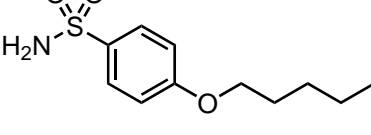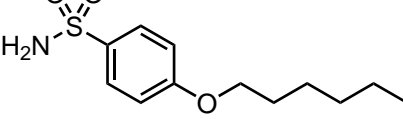
Overview of Compounds

The following tables show the compounds used experimentally in the course of this work.
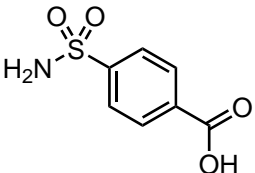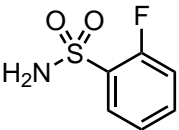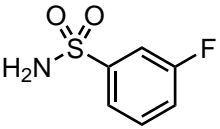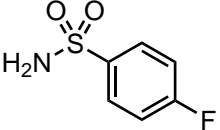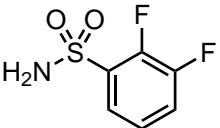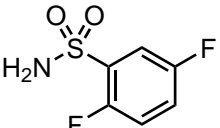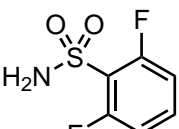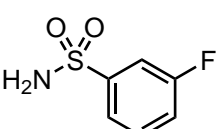
Tab. 14: Overview of compounds **1** – **10**.

Number	Compound	PDB entry	Ligand entry
1		6SDL	SBR
2		6SDJ	SBR
3		multiple	GLC, BGC
4		6GDC	FB2
5		6GM9	4J8
6		6HQX	4JC
7		6HR3	4JE
8		6HXD	6YP
9		6SBH	L4K
10		6I0W	M4S
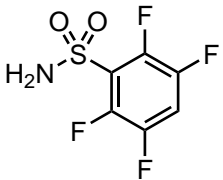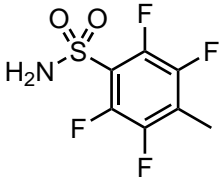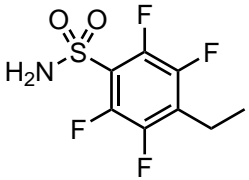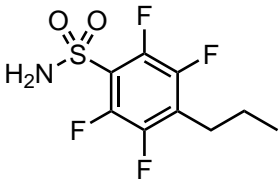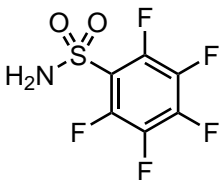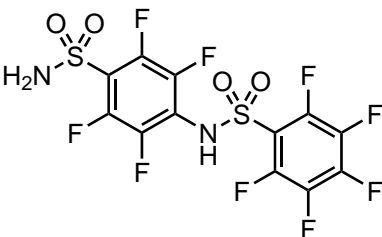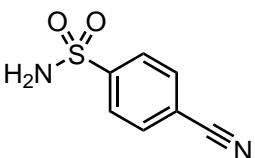
Tab. 15: Overview of compounds **11** – **19**.

Number	Compound	PDB entry	Ligand entry
11		6I1U	3W8
12		6I2F	3W6
13		6I3E	H1Z
14		6SBL	L4Q
15		not deposited	–
16		6RG5	S2O
17		6SEY	L9B
18		6SBM	L4W
19		not deposited	–
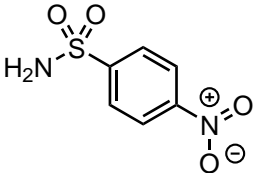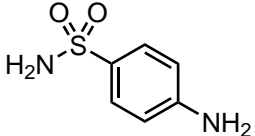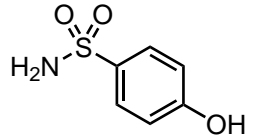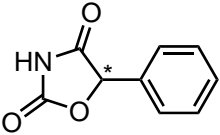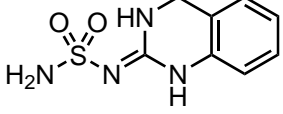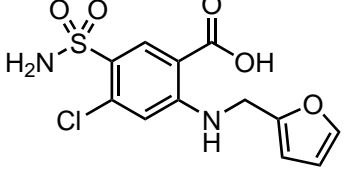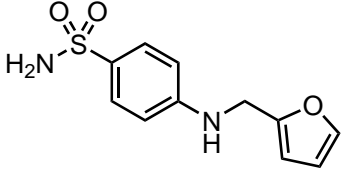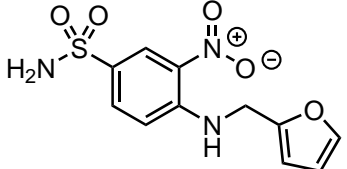
Tab. 16: Overview of compounds **20** – **27**.

Number	Compound	PDB entry	Ligand entry
20		6RFH	4SO
21		6RIT	FBV
22		6RQI	FBW
23		6RKN	FBS
24		6RJJ	K5W
25		6RNP	KBB
26		6ROE	FBT
27		6RGG	FBU
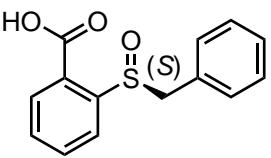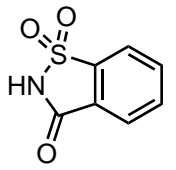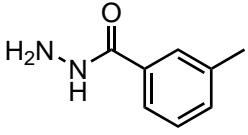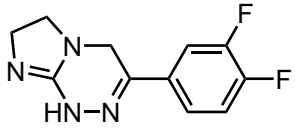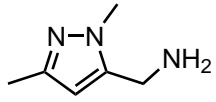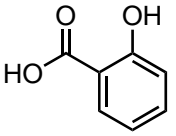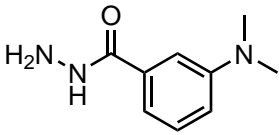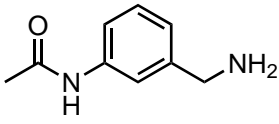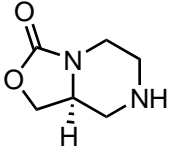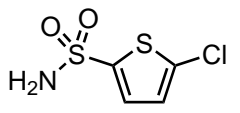
Tab. 17: Overview of compounds **28** – **34**.

Number	Compound	PDB entry	Ligand entry
28		6RRI	KFH
29		6RS5	KGK
30		6RSZ	KJ5
31		6S9G	L1Q
32		6SD7	L7K
33		6SD7	SXG
34		6ROB	KBZ

Tab. 18: Overview of compounds **35** – **42**.

Number	Compound	PDB entry	Ligand entry
35		6RH4	4NZ
36		6RL9	SAN
37		6RIG	K4N
38		6SFQ	7Q1
39		6SDH, 6SDI	0VV
40		6SG0	FUN
41		6SFU	LBQ
42		6SG6	LC8

Tab. 19: Overview of compounds **43**, **44** and **48** – **55**.

Number	Compound	PDB entry	Ligand entry
43		no structure prepared	3G1
44		not deposited	LSA
48		6RM1	45L
49		6S9Z	L2K
50		6SAC	47J
51		5M78	SAL
52		6RMP	47S
53		6SAS	463
54		6SAY	L3Q
55		6SB7	8K2

Crystallographic data

Crystallographic data were recorded at beamlines 14.1 and 14.2 at BESSY II, P13 at PETRA III and XRD1 at ELETTRA. Values in brackets refer to the highest resolution shell unless specified differently. The *Matthews* coefficient was calculated with the program Phaser Cell Content Analysis from the CCP4i suite.[178] Model validation was based on the free R-factor (R_{free}).[212] *Ramachandran* plots were calculated using the program PROCHECK.[224] Average B-factors were calculated with the program MOLEMAN.[225] 'n. d.' indicates that a structure has not been deposited in the PDB.

Tab. 20: Crystallographic and refinement data for hCAII in complex with compounds **1**, **2**, **4** and **5**.

	hCAII-**1** (6SDL)	hCAII-**2** (6SDJ)	hCAII-**4** (6GDC)	hCAII-**5** (6GM9)
Data collection and processing				
Wavelength / Å	0.9184	0.9184	0.9184	0.9184
Space group	P2$_1$	P2$_1$	P2$_1$	P2$_1$
a, b, c / Å	42.3, 41.7, 72.1	42.4, 41.1, 72.2	42.3, 41.4, 72.2	42.4, 41.5, 72.3
β / °	104.3	104.2	104.5	104.7
Matthews coefficient / Å^3 Da^{-1}	2.1	2.1	2.1	2.1
Solvent content / %	40.6	39.9	40.3	40.4
Diffraction data				
Resolution range / Å	41.7 – 1.08 (1.15 – 1.08)	41.1 – 1.02 (1.08 – 1.02)	41.4 – 1.08 (1.14 – 1.08)	41.5 – 1.09 (1.16 – 1.09)
Unique reflections	101798 (15740)	120252 (18157)	102486 (15973)	99517 (15508)
$CC_{1/2}$ / %	99.4 (81.5)	99.8 (96.0)	99.9 (96.6)	99.8 (95.6)
R_{sym} / %	8.9 (45.0)	4.8 (17.5)	4.2 (17.8)	4.5 (19.1)
Completeness / %	97.3 (93.6)	98.1 (92.1)	98.3 (95.1)	98.1 (95.0)
Wilson B-factor / Å^2	9.9	9.6	9.1	8.2
Multiplicity	3.7 (3.7)	3.5 (2.9)	3.6 (3.5)	3.6 (3.5)
$I/\sigma(I)$	7.3 (2.0)	13.5 (3.7)	15.9 (5.1)	15.4 (5.2)
Refinement				
Resolution range / Å	39.9 – 1.08	35.4 – 1.02	41.0 – 1.08	41.0 – 1.09
Reflections used in refinement (work/free)	101793 (96704/5089)	120242 (114229/6013)	102486 (97361/5125)	99517 (94541/4976)
Final R values for all reflections (work/free)	0.136/0.154	0.128/0.140	0.114/0.130	0.104/0.118
Protein residues	257	258	257	257
Inhibitor atoms	22	5	10/10/10	11/11
Water molecules	213	180	248	294
RMSD from ideality				
Bond lengths / Å	0.006	0.006	0.007	0.007
Bond angles / °	0.94	1.01	1.03	1.05
Ramachandran plot / %				
Residues in most favored regions	88.9	88.5	89.4	88.9
Residues in additionally allowed regions	11.1	11.5	10.2	10.6
Residues in generously allowed regions	0	0	0.5	0.5
Residues in disallowed regions	0	0	0	0
Mean B-factor / Å^2				
Protein non-hydrogen atoms	13.0	12.7	12.7	11.2
Inhibitor	17.1	9.0	9.1/14.6/16.2	8.3/14.1
Water molecules	24.8	24.0	24.3	23.7
Zinc ion	6.7	6.1	6.2	5.2

Tab. 21: Crystallographic and refinement data for hCAII in complex with compounds **6** – **9**.

	hCAII–**6** (6HQX)	hCAII–**7** (6HR3)	hCAII–**8** (6HXD)	hCAII–**9** (6SBH)
Data collection and processing				
Wavelength / Å	0.9184	0.9184	0.9184	0.9184
Space group	P2$_1$	P2$_1$	P2$_1$	P2$_1$
a, b, c / Å	42.3, 41.4, 72.2	42.3, 41.5, 72.3	42.2, 41.5, 72.1	42.3, 41.5, 72.2
β / °	104.7	104.6	104.5	104.5
Matthews coefficient / Å^3 Da^{-1}	2.1	2.1	2.1	2.1
Solvent content / %	40.2	40.4	40.2	40.3
Diffraction data				
Resolution range / Å	41.4 – 1.10 (1.10 – 1.09)	41.5 – 1.02 (1.08 – 1.02)	41.5 – 1.12 (1.19 – 1.12)	41.5 – 0.95 (1.01 – 0.95)
Unique reflections	97587 (15119)	121413 (19412)	91347 (14347)	142312 (21669)
$CC_{1/2}$ / %	99.9 (89.1)	99.9 (78.4)	99.9 (97.9)	99.6 (85.2)
R_{sym} / %	4.8 (33.9)	4.9 (46.2)	4.2 (13.7)	7.1 (36.1)
Completeness / %	98.6 (95.1)	98.2 (97.5)	97.7 (95.6)	93.2 (88.1)
Wilson$B$-factor / Å^2	10.7	9.8	9.3	8.1
Multiplicity	3.6 (3.6)	3.6 (3.5)	3.7 (3.7)	3.8 (3.6)
$I/\sigma(I)$	12.5 (2.9)	11.6 (2.0)	16.6 (6.4)	9.3 (2.1)
Refinement				
Resolution range / Å	40.9 – 1.09	40.0 – 1.02	40.9 – 1.12	34.9 – 0.95
Reflections used in refinement (work/free)	97587 (92707/4880)	121413 (115342/6071)	91339 (86772/4567)	142305 (135189/7116)
Final R values for all reflections (work/free)	0.122/0.139	0.122/0.140	0.115/0.131	0.120/0.134
Protein residues	257	257	257	257
Inhibitor atoms	12/11	13	14	15
Water molecules	275	243	252	227
RMSD from ideality				
Bond lengths / Å	0.008	0.008	0.008	0.008
Bond angles / °	1.10	1.06	1.10	1.06
Ramachandran plot / %				
Residues in most favored regions	89.4	89.4	89.4	90.7
Residues in additionally allowed regions	10.2	10.6	10.2	8.8
Residues in generously allowed regions	0.5	0	0.5	0.5
Residues in disallowed regions	0	0	0	0
Mean B-factor / Å^2				
Protein non-hydrogen atoms	13.2	12.4	11.7	10.6
Inhibitor	12.4/19.4	14.1	11.5	8.7
Water molecules	27.9	23.9	23.5	22.2
Zinc ion	7.8	6.9	6.12	5.3

Tab. 22: Crystallographic and refinement data for hCAII in complex with compounds 10 – 13.

	hCAII-10 (6IOW)	hCAII-11 (6IIU)	hCAII-12 (6I2F)	hCAII-13 (6I3E)
Data collection and processing				
Wavelength / Å	0.9184	0.9184	0.9184	0.9184
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
a, b, c / Å	42.4, 41.4, 72.3	42.3, 41.4, 72.2	42.2, 41.4, 72.0	42.5, 41.5, 71.9
β / °	104.6	104.5	104.3	104.2
Matthews coefficient / Å³ Da⁻¹	2.1	2.1	2.1	2.1
Solvent content / %	40.3	40.2	40.0	40.4
Diffraction data				
Resolution range / Å	41.4 – 1.04 (1.10 – 1.04)	41.4 – 1.08 (1.15 – 1.08)	41.4 – 1.19 (1.26 – 1.19)	41.5 – 1.07 (1.13 – 1.07)
Unique reflections	113777 (17379)	102490 (15912)	72714 (11290)	101318 (15148)
$CC_{1/2}$ / %	99.8 (75.1)	99.9 (94.2)	93.5 (90.1)	99.3 (81.7)
R_{sym} / %	5.7 (51.2)	4.4 (24.0)	4.24 (16.9)	8.8 (46.7)
Completeness / %	97.7 (93.1)	98.6 (95.1)	93.5 (90.1)	94.5 (87.7)
Wilson B-factor / Å²	9.2	9.7	9.5	9.2
Multiplicity	3.6 (3.3)	3.6 (3.6)	3.8 (3.7)	3.8 (3.9)
$I/\sigma(I)$	11.0 (2.0)	14.5 (4.1)	17.7 (6.3)	7.9 (2.3)
Refinement				
Resolution range / Å	35.0 – 1.04	40.9 – 1.08	40.9 – 1.19	41.2 – 1.07
Reflections used in refinement (work/free)	113766 (108078/5688)	102476 (97352/5124)	70978 (67429/3549)	101316 (96250/5066)
Final R values for all reflections (work/free)	0.119/0.132	0.116/0.133	0.116/0.139	0.129/0.147
Protein residues	257	257	257	257
Inhibitor atoms	12	13/12	14	15/14/12
Water molecules	266	255	282	234
RMSD from ideality				
Bond lengths / Å	0.008	0.008	0.010	0.007
Bond angles / °	1.08	1.09	1.16	1.01
Ramachandran plot / %				
Residues in most favored regions	88.9	89.8	89.4	88.9
Residues in additionally allowed regions	10.6	9.7	10.2	11.1
Residues in generously allowed regions	0.5	0.5	0.5	0
Residues in disallowed regions	0	0	0	0
Mean B-factor / Å²				
Protein non-hydrogen atoms	11.5	12.1	12.3	10.9
Inhibitor	9.6	11.1/20.2	10.7	11.7/20.3/18.5
Water molecules	23.2	23.5	23.7	24.4
Zinc ion	6.3	6.3	6.4	6.0

Tab. 23: Crystallographic and refinement data for hCAII in complex with compounds **14** – **17**.

	hCAII–**14** (6SBL)	hCAII–**15** (n. d.)	hCAII–**16** (6RG5)	hCAII–**17** (6SEY)
Data collection and processing				
Wavelength / Å	0.9184	0.9184	0.9184	0.9184
Space group	P2₁	P2₁	P2₁	P2₁
a, b, c / Å	42.3, 41.6, 72.0	42.3, 41.6, 72.1	42.5, 41.6, 72.3	42.7, 41.8, 72.8
β / °	104.3	104.3	104.6	104.7
Matthews coefficient / Å³ Da⁻¹	2.1	2.1	2.1	2.1
Solvent content / %	40.4	40.6	40.7	41.8
Diffraction data				
Resolution range / Å	41.6 – 0.94 (1.00 – 0.94)	41.6 – 1.02 (1.08 – 1.02)	41.6 – 1.09 (1.16 – 1.09)	41.8 – 1.23 (1.31 – 1.23)
Unique reflections	152258 (22843)	124150 (19214)	99707 (15418)	69652 (10672)
$CC_{1/2}$ / %	99.6 (87.1)	99.9 (81.9)	99.9 (98.5)	99.8 (98.4)
R_{sym} / %	6.1 (30.1)	4.9 (44.6)	4.1	4.6 (13.5)
Completeness / %	96.5 (90.1)	97.4 (96.0)	97.7 (94.2)	97.2 (92.7)
Wilson B-factor / Å²	8.3	9.8	7.9	8.3
Multiplicity	3.4 (2.9)	3.5 (3.5)	3.6 (3.7)	3.7 (3.8)
$I/\sigma(I)$	10.4 (2.3)	11.2 (2.1)	17.5 (7.5)	17.7 (8.2)
Refinement				
Resolution range / Å	41.0 – 0.94	41.1 – 1.02	41.1 – 1.09	41.3 – 1.23
Reflections used in refinement (work/free)	152252 (144639/7613)	120532 (1145056027)	99707 (94821/4986)	69646 (66163/3483)
Final R values for all reflections (work/free)	0.121/0.137	0.150/0.162	0.111/0.127	0.110/0.125
Protein residues	257	257	258	257
Inhibitor atoms	16	15	13/13/13/13	15/12
Water molecules	240	219	214	201
RMSD from ideality				
Bond lengths / Å	0.009	0.004	0.007	0.007
Bond angles / °	1.15	0.86	1.03	0.97
Ramachandran plot / %				
Residues in most favored regions	88.4	88.0	88.9	88.9
Residues in additionally allowed regions	11.1	11.1	11.1	10.6
Residues in generously allowed regions	0.5	0.9	0	0.5
Residues in disallowed regions	0	0	0	0
Mean B-factor / Å²				
Protein non-hydrogen atoms	11.1	12.9	10.1	10.8
Inhibitor	11.3	13.0	9.2/14.4/14.7/14.1	11.8/16.1
Water molecules	24.3	24.1	21.5	22.6
Zinc ion	5.6	7.0	4.8	5.1

Tab. 24: Crystallographic and refinement data for hCAII in complex with compounds **18 – 21**.

	hCAII-**18** (6SBM)	hCAII-**19** (n. d.)	hCAII-**20** (6RFH)	hCAII-**21** (6RIT)
Data collection and processing				
Wavelength / Å	0.9184	0.9184	0.9184	0.9184
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
a, b, c / Å	42.4, 41.6, 72.0	42.4, 41.6, 72.1	42.3, 41.5, 72.2	42.5, 41.6, 72.5
β / °	104.2	104.4	104.7	104.4
Matthews coefficient / Å³ Da⁻¹	2.1	2.1	2.1	2.1
Solvent content / %	40.4	40.4	40.2	41.1
Diffraction data				
Resolution range / Å	41.6 – 0.95 (1.01 – 0.95)	41.6 – 0.99 (1.05 – 0.99)	40.9 – 0.97 (1.03 – 0.97)	41.6 – 1.01 (1.07 – 1.01)
Unique reflections	147595 (22749)	131912 (20317)	137174 (21053)	126668 (19712)
$CC_{1/2}$ / %	99.8 (87.3)	99.9 (80.5)	99.8 (84.1)	99.9 (89.1)
R_{sym} / %	5.5 (33.3)	5.0 (40.4)	5.3 (43.4)	4.9 (39.4)
Completeness / %	96.4 (92.2)	97.2 (93.1)	95.6 (91.1)	97.5 (94.3)
Wilson B-factor / Å²	8.3	9.2	9.2	6.5
Multiplicity	4.0 (3.6)	3.6 (3.3)	3.68 (3.58)	3.6 (3.4)
$I/\sigma(I)$	12.1 (2.5)	11.6 (2.1)	11.3 (2.0)	13.7 (2.46)
Refinement				
Resolution range / Å	39.9 – 0.95	41.1 – 0.99	40.9 – 0.97	40.1 – 1.01
Reflections used in refinement (work/free)	147562 (140184/7378)	131911 (125315/6596)	137174 (130315/6859)	126658 (120324/6334)
Final R values for all reflections (work/free)	0.117/0.128	0.167/0.153	0.121/0.136	0.125/0.140
Protein residues	257	257	257	257
Inhibitor atoms	16/12	15	13/13/10	11/11
Water molecules	257	140	267	228
RMSD from ideality				
Bond lengths / Å	0.006	0.004	0.008	0.009
Bond angles / °	0.95	0.86	1.14	0.91
Ramachandran plot / %				
Residues in most favored regions	89.8	90.3	89.4	90.7
Residues in additionally allowed regions	10.2	9.3	10.2	8.8
Residues in generously allowed regions	0	0.5	0.5	0.5
Residues in disallowed regions	0	0	0	0
Mean B-factor / Å²				
Protein non-hydrogen atoms	11.0	11.9	11.8	9.3
Inhibitor	12.4/20.7	13.8	8.6/16.4/19.3	7.5
Water molecules	24.0	19.1	22.6	20.2
Zinc ion	5.5	6.4	6.3	4.8

Tab. 25: Crystallographic and refinement data for hCAII in complex with compounds **22** – **25**.

	hCAII–**22** (6RQI)	hCAII–**23** (6RKN)	hCAII–**24** (6RJJ)	hCAII–**25** (6RNP)
Data collection and processing				
Wavelength / Å	0.9184	0.9184	0.9184	0.9184
Space group	$P2_1$	P21	P21	P21
a, b, c / Å	42.4, 41.5, 72.3	42.3, 41.6, 72.2	42.4, 41.7, 72.6	42.3, 41.5, 72.3
β / °	104.6	104.6	104.6	104.5
Matthews coefficient / Å³ Da⁻¹	2.1	2.1	2.1	2.1
Solvent content / %	40.5	40.4	41.1	40.4
Diffraction data				
Resolution range / Å	41.5 – 0.95 (1.01 – 0.95)	41.6 – 0.96 (1.02 – 0.96)	41.7 – 1.06 (1.12 – 1.06)	41.5 – 1.07 (1.13 – 1.07)
Unique reflections	146227 (22470)	139058 (21525)	109461 (17106)	104839 (16391)
$CC_{1/2}$ / %	99.8 (83.3)	99.8 (93.4)	99.9 (84.2)	99.8 (83.1)
R_{sym} / %	5.1 (39.3)	5.3 (23.9)	5.1 (53.0)	6.0 (47.7)
Completeness / %	95.5 (91.1)	93.8 (90.2)	97.3 (94.3)	97.9 (95.2)
Wilson B-factor / Å²	8.1	7.5	8.6	8.3
Multiplicity	3.6 (3.3)	3.7 (3.5)	3.7 (3.6)	3.8 (3.6)
$I/\sigma(I)$	10.7 (2.2)	13.3 (3.9)	12.5 (2.0)	10.9 (2.1)
Refinement				
Resolution range / Å	41.0 – 0.95	41.0 – 0.96	41.0 – 1.06	41.0 – 1.07
Reflections used in refinement (work/free)	146215 (138903/7312)	139048 (132096/9652)	109451 (103978/5473)	104826 (99584/5242)
Final R values for all reflections (work/free)	0.116/0.130	0.108/0.124	0.123/0.138	0.119/0.136
Protein residues	257	257	257	257
Inhibitor atoms	11	11/11	12/12	12/12
Water molecules	206	248	263	272
RMSD from ideality				
Bond lengths / Å	0.006	0.007	0.009	0.007
Bond angles / °	0.99	1.07	1.11	1.02
Ramachandran plot / %				
Residues in most favored regions	89.8	90.3	89.8	88.9
Residues in additionally allowed regions	10.2	9.3	9.7	11.1
Residues in generously allowed regions	0	0.5	0.5	0
Residues in disallowed regions	0	0	0	0
Mean B-factor / Å²				
Protein non-hydrogen atoms	10.5	9.4	11.6	11.6
Inhibitor	8.5	8.1	9.9	8.6
Water molecules	22.3	21.7	23.7	23.7
Zinc ion	5.2	4.5	6.6	6.8

171

Tab. 26: Crystallographic and refinement data for hCAII in complex with compounds 26 – 29.

	hCAII–26 (6ROE)	hCAII–27 (6RGG)	hCAII–28 (6RRI)	hCAII–29 (6RS5)
Data collection and processing				
Wavelength / Å	0.9184	0.9184	0.9184	0.9184
Space group	P2₁	P2₁	P2₁	P2₁
a, b, c / Å	42.4, 41.4, 72.3	42.6, 41.7, 72.7	42.3, 41.4, 72.3	42.4, 41.5, 72.3
β / °	104.5	104.5	104.5	104.6
Matthews coefficient / Å³ Da⁻¹	2.1	2.1	2.1	2.1
Solvent content / %	40.3	41.4	40.2	40.5
Diffraction data				
Resolution range / Å	41.4 – 0.94 (1.00 – 0.94)	41.7 – 1.13 (1.20 – 1.13)	41.4 – 1.10 (1.16 – 1.10)	41.5 – 1.07 (1.13 – 1.07)
Unique reflections	152196 (23134)	91331 (14083)	97931 (15202)	103648 (16636)
$CC_{1/2}$ / %	99.9 (95.4)	99.8 (83.9)	98.8 (95.4)	99.8 (91.7)
R_{sym} / %	4.3 (18.7)	5.8 (49.2)	5.0 (43.5)	4.1 (23.4)
Completeness / %	96.5 (91.5)	97.9 (94.1)	99.9 (82.1)	96.5 (96.5)
Wilson B-factor / Å²	7.3	9.6	9.9	9.6
Multiplicity	3.5 (3.0)	3.7 (3.5)	3.6 (3.6)	2.5 (2.4)
$I/\sigma(I)$	14.8 (3.6)	11.3 (2.0)	12.3 (2.5)	11.8 (2.9)
Refinement				
Resolution range / Å	35.6 – 0.94	41.2 – 1.13	22.9 – 1.10	23.8 – 1.07
Reflections used in refinement (work/free)	152179 (144569/7610)	91307 (86740/4567)	97922 (93026/4896)	103624 (98443/5181)
Final R values for all reflections (work/free)	0.110/0.119	0.127/0.140	0.119/0.139	0.118/0.132
Protein residues	257	257	257	257
Inhibitor atoms	12/12/12	12/12	14/14	15/15/15
Water molecules	209	251	247	230
RMSD from ideality				
Bond lengths / Å	0.006	0007	0.006	0.007
Bond angles / °	0.97	0.97	0.97	1.00
Ramachandran plot / %				
Residues in most favored regions	88.9	88.0	89.4	89.8
Residues in additionally allowed regions	10.6	11.6	10.2	10.2
Residues in generously allowed regions	0.5	0.5	0.5	0
Residues in disallowed regions	0	0	0	0
Mean B-factor / Å²				
Protein non-hydrogen atoms	9.35	12.8	12.5	11.6
Inhibitor	7.6/8.7	9.3/9.8	11.9/18.1	9.3/10/19.9
Water molecules	20.0	24.5	25.4	22.3
Zinc ion	4.8	7.6	7.1	6.7

Tab. 27: Crystallographic and refinement data for hCAII in complex with compounds **30**, **31**, **32** and **34**.

	hCAII–**30** (6RSZ)	hCAII–**31** (6S9G)	hCAII–**32/33** (6SD7)	hCAII–**34** (6ROB)
Data collection and processing				
Wavelength / Å	0.9184	0.9184	0.9184	0.9184
Space group	P2$_1$	P2$_1$	P2$_1$	P2$_1$
a, b, c / Å	42.2, 41.4, 72.1	42.3, 41.5, 72.1	42.3, 41.4, 72.3	42.3, 41.5, 72.2
β / °	104.5	104.5	104.7	104.5
Matthews coefficient / Å^3 Da^{-1}	2.1	2.1	2.1	2.1
Solvent content / %	40.0	40.3	40.1	40.3
Diffraction data				
Resolution range / Å	41.5 – 1.09 (1.16 – 1.09)	41.5 – 1.14 (1.21 – 1.14)	41.4 – 1.05 (1.11 – 1.05)	41.5 – 0.93 (0.99 – 0.93)
Unique reflections	96745 (14752)	87774 (13942)	111421 (17610)	155258 (22819)
CC$_{1/2}$ / %	99.9 (99.0)	99.9 (80.2)	99.9 (91.3)	99.9 (97.2)
R_{sym} / %	3.5 (9.8)	5.3 (47.1)	4.5 (25.5)	4.0 (13.6)
Completeness / %	95.9 (91.0)	98.8 (97.5)	98.6 (96.6)	95.4 (86.9)
Wilson B-factor / Å^2	8.9	11.2	9.8	7.4
Multiplicity	3.7 (3.8)	3.3 (3.6)	3.5 (3.1)	3.5 (2.9)
$I/\sigma(I)$	20.9 (9.3)	11.2 (2.1)	13.6 (3.1)	17.0 (5.1)
Refinement				
Resolution range / Å	23.3 – 1.09	23.8 – 1.14	23.8 – 1.05	40.9 – 0.93
Reflections used in refinement (work/free)	96732 (91895/4837)	87761 (83373/4388)	111410 (105840/5570)	155254 (147491/7763)
Final *R* values for all reflections (work/free)	0.113/0.129	0.123/0.147	0.124/0.139	0.112/0.120
Protein residues	257	257	257	257
Inhibitor atoms	16	17/15	15/29	12/12
Water molecules	261	269	171	208
RMSD from ideality				
Bond lengths / Å	0.007	0.006	0.006	0.006
Bond angles / °	0.98	0.98	1.00	0.97
Ramachandran plot / %				
Residues in most favored regions	89.8	89.4	89.4	89.8
Residues in additionally allowed regions	9.7	9.7	10.6	9.7
Residues in generously allowed regions	0.5	0.9	0	0.5
Residues in disallowed regions	0	0	0	0
Mean *B*-factor / Å^2				
Protein non-hydrogen atoms	12.2	14.2	12.2	8.9
Inhibitor	11.6	12.2/27.2	13.9/14.8	8.5/13.5
Water molecules	24.3	28.8	22.0	20.0
Zinc ion	6.2	8.2	6.5	4.4

Tab. 28: Crystallographic and refinement data for hCAII in complex with compounds **35** – **37** and **48**.

	hCAII-**35** (6RH4)	hCAII-**36** (6RL9)	hCAII-**37** (6RIG)	hCAII-**48** (6RM1)
Data collection and processing				
Wavelength / Å	0.9184	0.9184	0.9184	1.0000
Space group	P2$_1$	P2$_1$	P2$_1$	P2$_1$
a, b, c / Å	42.4, 41.5, 72.3	42.5, 41.5, 72.0	42.4, 41.5, 72.1	42.5, 41.5, 72.0
β / °	104.7	104.7	104.6	104.7
Matthews coefficient / Å³ Da⁻¹	2.1	2.1	2.1	2.1
Solvent content / %	40.6	40.4	40.4	42.1
Diffraction data				
Resolution range / Å	41.5 – 0.95 (0.95 – 0.95)	41.0 – 1.00 (1.06 – 1.00)	41.1 – 1.00 (1.06 – 1.00)	41.5 – 1.68 (1.78 – 1.68)
Unique reflections	148154 (22694)	130686 (20986)	130452 (20940)	27053 (4228)
$CC_{1/2}$ / %	99.8 (90.9)	99.7 (82.5)	99.8 (82.0)	99.7 (99.2)
R_{sym} / %	5.3 (28.0)	6.6 (43.3)	5.4 (43.6)	5.1 (73)
Completeness / %	96.3 (91.5)	99.6 (99.2)	99.3 (98.9)	96.8 (94.5)
Wilson B-factor / Å²	6.8	8.3	8.3	8.7
Multiplicity	3.5 (3.2)	3.6 (3.3)	3.6 (3.3)	3.8 (3.6)
$I/\sigma(I)$	12.7 (3.3)	9.5 (2.1)	11.1 (2.1)	19.1 (12.1)
Refinement				
Resolution range / Å	41.0 – 0.95	32.0 – 1.00	41.0 – 1.00	35.6 – 1.68
Reflections used in refinement (work/free)	148150 (140742/7408)	130676 (124142/6534)	130446 (123923/6523)	27052 (25699/1353)
Final R values for all reflections (work/free)	0.103/0.117	0.119/0.130	0.119/0.136	0.153/0.181
Protein residues	257	258	257	257
Inhibitor atoms	13/13	11/11/11	11/11/11	11
Water molecules	253	220	219	166
RMSD from ideality				
Bond lengths / Å	0.008	0.008	0.008	0.008
Bond angles / °	1.07	1.08	0.91	1.00
Ramachandran plot / %				
Residues in most favored regions	89.4	88.9	88.9	88.0
Residues in additionally allowed regions	10.2	10.6	11.1	11.5
Residues in generously allowed regions	0.5	0.5	0	0.5
Residues in disallowed regions	0	0	0	0
Mean B-factor / Å²				
Protein non-hydrogen atoms	9.3	10.8	10.8	10.5
Inhibitor	9.5/16.5	8.4	8.1/15.9/28.9	8.3
Water molecules	21.7	22.9	22.7	18.3
Zinc ion	4.5	5.5	5.7	5.7

Tab. 29: Crystallographic and refinement data for hCAII in complex with compounds **49** – **52**.

	hCAII–**49** (6S9Z)	hCAII–**50** (6SAC)	hCAII–**51** (5M78)	hCAII–**52** (6RMP)
Data collection and processing				
Wavelength / Å	0.7999	0.9184	0.7999	0.7999
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
a, b, c / Å	42.4, 41.3, 72.4	42.4, 41.3, 72.3	42.4, 41.4, 72.4	42.4, 41.6, 72.1
β / °	104.7	104.8	104.8	104.7
Matthews coefficient / Å³ Da⁻¹	2.1	2.1	2.2	2.2
Solvent content / %	40.2	40.1	44.1	44.1
Diffraction data				
Resolution range / Å	41.3 – 0.95 (1.01 – 0.95)	41.0 – 1.02 (1.08 – 1.02)	41.4 – 1.08 (1.14 – 1.08)	41.6 – 1.22 (1.29 – 1.22)
Unique reflections	141261 (19435)	121056 (19033)	103053 (16045)	72050 (11150)
$CC_{1/2}$ / %	99.8 (79.8)	99.8 (78.0)	99.8 (84.6)	99.8 (85.3)
R_{sym} / %	6.3 (43.2)	5.3 (39.1)	6.6 (39.9)	6.8 (43.3)
Completeness / %	92.7 (79.0)	98.3 (96.0)	98.3 (95.0)	98.6 (94.8)
Wilson B-factor / Å²	6.9	8.5	8.9	9.9
Multiplicity	3.8 (3.8)	2.7 (2.5)	3.7 (3.6)	3.7 (3.7)
$I/\sigma(I)$	10.7 (2.3)	10.0 (2.1)	9.9 (2.4)	10.2 (2.4)
Refinement				
Resolution range / Å	35.0 – 0.95	32.0 – 1.02	35.6 – 1.08	34.9 – 1.22
Reflections used in refinement (work/free)	141257 (134196/7061)	121032 (114983/6049)	103048 (97895/5153)	72043 (68440/3603)
Final R values for all reflections (work/free)	0.119/0.137	0.126/0.142	0.125/0.144	0.132/0.162
Protein residues	265	265	259	259
Inhibitor atoms	17/1	9	10	12
Water molecules	241	199	244	208
RMSD from ideality				
Bond lengths / Å	0.006	0.008	0.00	0.007
Bond angles / °	0.97	1.07	0.97	1.00
Ramachandran plot / %				
Residues in most favored regions	88.3	87.4	89.9	88.5
Residues in additionally allowed regions	11.2	12.1	9.6	11.0
Residues in generously allowed regions	0.4	0.4	0.5	0.5
Residues in disallowed regions	0	0	0	0
Mean *B*-factor / Å²				
Protein non-hydrogen atoms	8.3	12.7	13.1	12.9
Inhibitor	6.5/7.1	29.0	17.0	15.5
Water molecules	20.4	22.5	24.1	24.5
Zinc ion	4.8	5.8	6.7	7.9

Tab. 30: Crystallographic and refinement data for hCAII in complex with compounds **53** – **55** and **38**.

	hCAII–**53** (6SAS)	hCAII–**54** (6SAY)	hCAII–**55** (6SB7)	hCAII–**38** (6SFQ)
Data collection and processing				
Wavelength / Å	0.9184	0.9184	0.9184	0.9184
Space group	P2$_1$	P2$_1$	P2$_1$	P2$_1$
a, b, c / Å	42.4, 41.5, 72.1	42.4, 41.5, 72.2	42.5, 41.7, 72.5	42.4, 41.6, 72.1
β / °	104.3	104.7	104.3	104.5
Matthews coefficient / Å^3 Da^{-1}	2.1	2.1	2.2	2.1
Solvent content / %	42.5	42.2	42.9	42.4
Diffraction data				
Resolution range / Å	41.5 – 1.10 (1.16 – 1.10)	41.5 – 0.95 (1.01 – 0.95)	41.7 – 1.09 (1.16 – 1.09)	41.6 – 1.00 (1.06 – 1.00)
Unique reflections	97293 (15236)	146380 (21821)	98467 (14806)	128491 (20448)
CC$_{1/2}$ / %	99.8 (78.9)	99.9 (79.5)	99.8 (93.0)	99.9 (86.5)
R_{sym} / %	6.7 (45.8)	5.2 (39.9)	5.2 (23.3)	4.5 (40.5)
Completeness / %	97.8 (95.4)	96.5 (89.1)	97.2 (91.1)	97.6 (96.2)
Wilson B-factor / Å^2	9.0	7.9	8.1	9.5
Multiplicity	3.6 (3.6)	3.5 (3.0)	3.6 (3.5)	3.6 (3.5)
$I/\sigma(I)$	9.9 (2.3)	11.4 (2.0)	13.5 (4.5)	12.8 (2.2)
Refinement				
Resolution range / Å	41.1 – 1.10	34.9 – 0.95	35.8 – 1.09	41.1 – 1.00
Reflections used in refinement (work/free)	97283 (92418/4865)	146362 (139043/7319)	98465 (93541/4924)	128473 (122049/6424)
Final R values for all reflections (work/free)	0.132/0.149	0.122/0.138	0.118/0.137	0.126/0.147
Protein residues	257	264	258	257
Inhibitor atoms	12	10	10	13
Water molecules	211	223	153	245
RMSD from ideality				
Bond lengths / Å	0.006	0.007	0.009	0.006
Bond angles / °	0.94	0.97	1.11	0.94
Ramachandran plot / %				
Residues in most favored regions	88.9	89.6	88.9	88.0
Residues in additionally allowed regions	10.6	9.5	11.1	12.0
Residues in generously allowed regions	0.5	0.9	0	0
Residues in disallowed regions	0	0	0	0
Mean B-factor / Å^2				
Protein non-hydrogen atoms	11.4	10.7	10.8	12.5
Inhibitor	23.5	18.5	14.0	10.2
Water molecules	22.8	20.5	20.7	24.5
Zinc ion	6.2	6.0	5.4	6.8

Tab. 31: Crystallographic and refinement data for hCAII in complex with compounds **39** – **41**.

	hCAII–**39** (6SDH)	hCAII–**39** (6SDI)	hCAII–**40** (6SG0)	hCAII–**41** (6SFU)
Data collection and processing				
Wavelength / Å	0.9763	0.9763	0.885	0.885
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
a, b, c / Å	42.3, 41.7, 72.3	42.3, 41.6, 72.3	42.4, 41.8, 72.2	42.5, 41.6, 72.4
β / °	104.1	104.1	104.5	104.6
Matthews coefficient / Å³ Da⁻¹	2.1	2.1	2.1	2.1
Solvent content / %	40.9	40.6	42.6	40.9
Diffraction data				
Resolution range / Å	41.7 – 1.04 (1.10 – 1.04)	41.6 – 1.03 (1.09 – 1.03)	41.8 – 1.13 (1.20 – 1.13)	41.6 – 1.04 (1.10 – 1.04)
Unique reflections	112924 (15400)	114294 (15542)	91927 (14411)	117443 (18662)
$CC_{1/2}$ / %	98.9 (96.7)	99.2 (94.5)	99.9 (92.6)	99.9 (90.2)
R_{sym} / %	8.3 (15.3)	7.3 (18.0)	6.1 (39.8)	7.3 (46.3)
Completeness / %	95.9 (81.1)	94.6 (80.0)	99.4 (97.0)	99.6 (98.4)
Wilson$B$-factor / Å²	10.8	9.8	10.3	7.6
Multiplicity	3.5 (2.9)	3.5 (2.4)	7.0 (6.8)	7.0 (6.9)
$I/\sigma(I)$	10.6 (4.3)	10.7 (3.0)	15.3 (3.6)	14.2 (3.1)
Refinement				
Resolution range / Å	39.9 – 1.04	21.4 – 1.03	40.0 – 1.13	29.2 – 1.04
Reflections used in refinement (work/free)	112922 (107275/5647)	114291 (108577/5714)	91919 (87322/4597)	117414 (111542/5872)
Final R values for all reflections (work/free)	0.131/0.148	0.125/0.141	0.126/0.145	0.126/0.138
Protein residues	257	257	257	257
Inhibitor atoms	15	15	21/16	17/12
Water molecules	162	174	211	231
RMSD from ideality				
Bond lengths / Å	0.006	0.006	0.006	0.006
Bond angles / °	1.02	1.03	0.93	0.93
Ramachandran plot / %				
Residues in most favored regions	89.4	89.8	89.8	90.7
Residues in additionally allowed regions	10.6	10.2	10.2	7.3
Residues in generously allowed regions	0	0	0	0
Residues in disallowed regions	0	0	0	0
Mean B-factor / Å²				
Protein non-hydrogen atoms	14.2	13.0	13.4	10.5
Inhibitor	12.1	14.6	13.5/9.4	15.5/13.7
Water molecules	25.5	23.6	24.2	21.6
Zinc ion	7.5	6.9	6.8	5.2

Tab. 32: Crystallographic and refinement data for hCAII in complex with compounds **42** and **44**.

	hCAII–**42** (6SG6)	hCAII–**44** (n. d.)
Data collection and processing		
Wavelength / Å	0.9184	0.9794
Space group	$P2_1$	$P2_1$
a, b, c / Å	42.3, 41.7, 72.2	42.4, 41.5, 72.1
β / °	104.4	104.4
Matthews coefficient / Å³ Da⁻¹	2.1	2.1
Solvent content / %	42.5	40.3
Diffraction data		
Resolution range / Å	41.7 – 0.98 (1.04 – 0.98)	41.5 – 0.99 (1.01 – 0.99)
Unique reflections	136715 (21571)	260462 (16576)
$CC_{1/2}$ / %	99.7 (85.3)	99.9 (96.0)
R_{sym} / %	5.7 (36.3)	4.0 (20.7)
Completeness / %	97.7 (95.5)	98.1 (84.5)
Wilson B-factor / Å²	9.3	8.2
Multiplicity	3.6 (3.3)	6.2 (4.0)
$I/\sigma(I)$	10.9 (2.4)	24.1 (5.9)
Refinement		
Resolution range / Å	40.0 – 0.98	35.7 – 0.99
Reflections used in refinement (work/free)	136703 (129868/6835)	260457 (257847/2610)
Final R values for all reflections (work/free)	0.122/0.135	0.118/0.132
Protein residues	257	257
Inhibitor atoms	20	12
Water molecules	226	244
RMSD from ideality		
Bond lengths / Å	0.006	0.006
Bond angles / °	0.96	0.97
Ramachandran plot / %		
Residues in most favored regions	88.4	88.4
Residues in additionally allowed regions	11.6	11.1
Residues in generously allowed regions	0	0.5
Residues in disallowed regions	0	0
Mean B-factor / Å²		
Protein non-hydrogen atoms	12.4	12.0
Inhibitor	15.4	8.8
Water molecules	23.6	23.4
Zinc ion	6.8	5.8

ITC Data

Thermodynamic and kinetic parameters for individual measurements as well as globally fitted values are provided. For the studies concerning the ITC measurement protocol, also mean values of the individual measurements are provided. '°' refers to the standard state. χ^2 denotes the goodness of fit.

n-Alkyl and *n*-Alkyloxy compounds

Tab. 33: Thermodynamic values for individual measurement for compounds **4 – 8** and **10 – 13**.

Compound	χ^2	$\Delta G°$ kJ mol^{-1}	$\Delta H°$ kJ mol^{-1}	$-T\Delta S°$ kJ mol^{-1}	K_a 10^6 M^{-1}	K_d nM^{-1}
4	0.12	-36.1	-39.5	3.4	2.15	465
	0.18	-35.8	-39.2	3.4	1.86	539
	0.51	-36.1	-39.9	3.8	2.08	479
5	0.17	-38.1	-41.0	2.9	4.76	210
	0.15	-37.6	-40.6	3.0	3.84	260
	0.71	-38.8	-40.2	1.4	6.2	161
6	0.68	-40.9	-41.1	0.2	14.71	68
	0.36	-40.2	-41.43	1.19	11.21	89.23
	0.52	-39.8	-42.2	2.3	9.46	105
7	0.26	-40.8	-38.0	-2.8	14.05	71
	0.28	-40.4	-39.5	-0.8	11.76	85
	0.43	-41.0	-39.2	-1.7	14.95	66
8	0.39	-42.2	-44.7	2.6	24.41	41.0
	0.28	-42.3	-43.4	1.0	26.04	38
	0.37	-43.4	-44.7	1.3	40.83	24
10	0.15	-37.3	-46.0	8.7	3.40	293
	0.27	-37.1	-45.2	8.0	3.20	312
	0.29	-37.6	-45.4	7.8	3.78	264
11	0.59	-39.6	-46.7	7.1	8.77	114
	0.25	-39.8	-46.4	6.6	9.22	108
	0.24	-39.1	-46.4	7.3	7.10	140
12	0.56	-42.1	-50.0	8.0	23.29	42
	0.54	-42.8	-49.3	6.5	31.24	32
	0.51	-43.4	-50.1	6.8	39.1	25
13	0.40	-44.5	-51.8	7.3	62.09	16
	0.30	-44.3	-51.5	7.2	57.81	17
	0.20	-42.9	-52.0	9.1	32.56	30

Tab. 34: Globally fitted thermodynamic values with standard errors from three individual ITC experiments for compounds **4 – 13**.

Compound	$\Delta G^{\circ a}$ kJ mol⁻¹	$\Delta H^{\circ b}$ kJ mol⁻¹	$-T\Delta S^{\circ c}$ kJ mol⁻¹	$K_a{}^b$ M⁻¹	Standard error[d] ΔG° kJ mol⁻¹	ΔH° kJ mol⁻¹	$-T\Delta S^\circ$ kJ mol⁻¹	K_a M⁻¹	Global fit error $\Delta G^{\circ e}$ kJ mol⁻¹	$\Delta H^{\circ f}$ kJ mol⁻¹	$-T\Delta S^{\circ e}$ kJ mol⁻¹	$K_a{}^f$ M⁻¹
4	-36.0	-39.6	3.6	$1.99 \cdot 10^6$	0.1	0.1	0.2	$8.89 \cdot 10^4$	0.02	0.03	0.04	$1.84 \cdot 10^4$
5	-38.0	-40.8	2.8	$4.46 \cdot 10^6$	0.3	0.2	0.5	$6.81 \cdot 10^5$	0.07	0.04	0.08	$1.23 \cdot 10^5$
6	-40.4	-41.5	1.1	$1.18 \cdot 10^7$	0.3	0.3	0.6	$1.54 \cdot 10^6$	0.04	0.05	0.06	$1.79 \cdot 10^5$
7	-40.7	-38.6	-2.1	$1.34 \cdot 10^7$	0.2	0.5	0.6	$9.50 \cdot 10^5$	0.05	0.06	0.06	$2.87 \cdot 10^5$
8	-42.6	-44.2	1.6	$2.97 \cdot 10^7$	0.4	0.5	0.5	$5.22 \cdot 10^6$	0.06	0.04	0.07	$6.76 \cdot 10^5$
9[d]	-45.0	-43.0	-2.0	$7.73 \cdot 10^7$	–	–	–	–	0.36	0.02	0.36	$1.11 \cdot 10^7$
10	-37.3	-45.5	8.2	$3.48 \cdot 10^6$	0.1	0.3	0.3	$1.70 \cdot 10^5$	0.02	0.03	0.03	$2.70 \cdot 10^4$
11	-39.5	-46.5	7.0	$8.41 \cdot 10^6$	0.2	0.1	0.2	$6.45 \cdot 10^5$	0.02	0.03	0.05	$1.21 \cdot 10^5$
12	-42.6	-49.8	7.2	$2.90 \cdot 10^7$	0.4	0.3	0.5	$4.56 \cdot 10^6$	0.04	0.06	0.08	$6.77 \cdot 10^5$
13[g]	-43.6	-51.8	8.3	$4.27 \cdot 10^7$	0.5	0.1	0.6	$9.21 \cdot 10^6$	0.12	0.05	0.13	$2.15 \cdot 10^6$

[a] Calculated using equations (2) and (3) with the globally fitted value of K_a. [b] Globally fitted value. [c] Calculated with equation 3. [d] Due to its high affinity, **9** could not be fitted individually and had to be characterized by a displacement setup as described in the Methods section, which was globally fitted. Standard errors can therefor not be provided. [e] Calculated from the global fit error of K_a assuming error propagation (equations (22) and (23)).[226] [f] Global fit error. [g] Values do not fully add up due to rounding.

$$e_{\Delta G^\circ} = \left| -R \cdot T \cdot \frac{e_{K_a}}{K_a} \right| \tag{22}$$

$$e_{-T\Delta S^\circ} = \sqrt{e_{\Delta G^\circ}^2 + e_{\Delta H^\circ}^2} \tag{23}$$

Tab. 35: Kinetic parameters for individual ITC experiments for compounds **4 – 8** and **10 – 13** before (left) and after (right) application of globally adjusted thermodynamic parameters.

Compound	Before global adjustment			After global adjustment		
	χ^2	k_{on} 10^4 M^{-1} s^{-1}	k_{off} 10^{-2} s^{-1}	χ^2	k_{on} 10^4 M^{-1} s^{-1}	k_{off} 10^{-2} s^{-1}
4	0.55	1.71	0.79	0.54	1.62	0.81
	0.95	2.08	1.12	1.39	2.17	1.09
	1.35	2.28	1.09	1.29	2.22	1.11
5	0.98	2.24	0.47	0.89	2.17	0.49
	0.84	2.84	0.74	1.06	3.02	0.68
	0.69	2.48	0.40	0.81	2.15	0.48
6	1.50	7.53	0.51	1.47	6.90	0.59
	2.48	5.27	0.47	2.54	5.41	0.46
	0.68	7.04	0.74	0.74	7.71	0.65
7	0.34	8.56	0.61	0.32	8.41	0.63
	0.81	4.82	0.41	0.84	5.00	0.37
	1.31	8.43	0.56	1.18	8.08	0.60
8	0.69	12.43	0.51	0.71	13.3	0.45
	1.33	16.74	0.64	1.36	17.4	0.59
	0.47	14.18	0.35	0.52	13.0	0.44
10	1.22	2.31	0.68	1.22	2.33	0.67
	2.01	1.91	0.60	2.15	2.02	0.58
	0.83	2.74	0.72	0.74	2.61	0.75
11	1.30	6.75	0.77	1.28	6.62	0.79
	3.93	5.65	0.61	4.60	5.27	0.63
	1.38	5.99	0.84	1.46	6.62	0.79
12	2.35	6.01	0.26	2.74	6.58	0.23
	2.50	7.76	0.25	2.38	7.47	0.26
	2.10	5.64	0.14	1.80	5.15	0.18
13	0.95	36.16	0.58	0.96	31.6	0.74
	1.59	10.90	0.19	1.47	9.67	0.23
	2.15	12.01	0.37	2.26	13.3	0.31

Tab. 36: Mean values for kinetic parameters from individual ITC experiments after application of globally fitted parameters for compounds **4** – **8** and **10** – **13** with standard errors for three experiments.

			Standard error	
Compound	$k_{on}{}^a$ 10^4 M^{-1} s^{-1}	$k_{off}{}^a$ 10^{-2} s^{-1}	$k_{on}{}^b$ 10^4 M^{-1} s^{-1}	$k_{off}{}^b$ 10^{-2} s^{-1}
4	2.00	1.00	0.19	0.10
5	2.45	0.55	0.29	0.06
6	6.68	0.57	0.67	0.06
7	7.16	0.53	1.08	0.08
8	14.6	0.49	1.43	0.05
10	2.32	0.67	1.72	0.05
11	6.17	0.73	0.45	0.05
12	6.40	0.22	0.68	0.02
13	18.2	0.43	6.81	0.16

[a] Kinetic data were obtained after renewed processing of the raw data with $\Delta H°$ and K_a values from global fitting. [b] Standard error of measurement after adjustment with global data.

ITC Protocol

Tab. 37: Thermodynamic values for individual measurements of compound **20** with different measurement protocols.

Protocol	χ^2	$\Delta G°$ kJ mol^{-1}	$\Delta H°$ kJ mol^{-1}	$-T\Delta S°$ kJ mol^{-1}	K_a 10^6 M^{-1}	K_d nM^{-1}
Standard	0.25	-35.0	-63.3	28.2	1.38	724
	0.20	-35.5	-61.3	25.7	1.69	594
	0.27	-35.7	-63.8	28.1	1.80	555
	0.21	-35.3	-59.8	24.5	1.54	649
	0.28	-35.1	-63.9	28.9	1.41	710
Kinetic	0.15	-35.5	-59.3	23.8	1.65	608
	0.18	-35.4	-59.0	23.7	1.56	641
	0.29	-35.2	-59.5	24.3	1.47	679
	0.34	-35.2	-61.2	26.1	1.44	694
	0.20	-35.0	-61.8	26.9	1.36	738
Mixed	0.17	-35.1	-61.0	25.9	1.39	719
	0.15	-35.1	-59.6	24.6	1.39	717
	0.42	-35.1	-61.9	26.8	1.40	714
	0.30	-35.2	-60.1	24.9	1.47	682
	0.49	-34.9	-60.3	25.4	1.28	784
Thermodynamic	0.09	-35.3	-59.6	24.3	1.52	659
	0.18	-35.6	-58.2	22.6	1.73	579
	0.05	-35.2	-59.0	23.8	1.47	679
	0.20	-35.4	-58.7	23.3	1.57	637
	0.06	-35.1	-59.5	24.4	1.44	697

Tab. 38: Mean and globally fitted values for the four measurement protocols.

Protocol	$\Delta G°$ kJ mol^{-1}	$\Delta H°$ kJ mol^{-1}	$-T\Delta S°$ kJ mol^{-1}	K_d M^{-1}
Standard, raw data mean	-35.3	-62.4	27.1	646
Standard, globally adjusted	-35.4	-62.6	27.2	638
Kinetic, raw data mean	-35.2	-60.2	24.9	672
Kinetic, globally adjusted	-35.2	-60.1	24.9	669
Mixed, raw data mean	-35.1	-60.6	25.5	723
Mixed, globally adjusted	-35.0	-60.2	25.3	673
Thermodynamic, raw data mean	-35.3	-59.0	23.7	650
Thermodynamic, globally adjusted	-35.3	-59.1	23.7	650

Tab. 39: Values for the left and right error margin of a 95 % confidence interval of the thermodynamic values for five experiments, based on the assumption of T-distributed individual experiments.

Protocol	$\Delta G°$ kJ mol^{-1}	$\Delta H°$ kJ mol^{-1}	$-T\Delta S°$ kJ mol^{-1}	K_d nM
Standard	0.4	2.2	2.3	90.5
Kinetic	0.2	1.6	1.8	61.9
Mixed	0.2	1.1	1.1	46.0
Thermodynamic	0.2	0.7	0.9	56.9

Tab. 40: Kinetic parameters for individual ITC experiments with different measurement protocols for compound **20** before (left) and after (right) application of globally adjusted thermodynamic parameters.

Compound	Before global adjustment			After global adjustment		
	χ^2	k_{on} 10^4 M^{-1} s^{-1}	k_{off} 10^{-2} s^{-1}	χ^2	k_{on} 10^4 M^{-1} s^{-1}	k_{off} 10^{-2} s^{-1}
Standard	0.79	4.35	3.2	0.81	4.87	3.1
	0.18	5.00	3.0	0.18	4.74	3.0
	0.49	3.92	2.2	0.49	3.51	2.2
	0.46	4.20	2.7	0.50	4.28	2.7
	0.93	4.18	3.0	0.87	4.54	2.9
Kinetic	0.36	4.58	2.8	0.34	4.19	2.8
	1.08	4.34	2.8	1.08	4.16	2.8
	0.44	4.40	3.0	0.44	4.43	3.0
	0.55	3.90	2.7	0.53	3.98	2.7
	0.88	4.10	3.0	0.84	4.36	2.9
Mixed	0.48	4.40	3.2	0.48	4.25	3.2
	0.68	3.57	2.6	0.73	3.45	2.6
	0.11	3.81	2.7	0.11	3.67	2.7
	0.22	4.14	2.8	0.24	3.86	2.9
	0.61	3.49	2.7	0.59	3.62	2.7
Thermodynamic	1.75	2.93	1.9	1.74	2.95	1.9
	3.08	3.09	1.8	3.15	2.84	1.8
	1.52	2.91	2.0	1.54	3.03	2.0
	3.33	2.94	1.9	3.37	2.91	1.9
	1.78	2.27	1.6	1.74	2.42	1.6

Tab. 41: Individual-measurement derived mean values and globally fitted/adjusted kinetic values for the four measurement protocols.

Protocol	k_{on} 10^4 M^{-1} s^{-1}	k_{off} 10^{-2} s^{-1}	τ_{ITC} s	τ_{rd} s
Standard, raw data mean	4.33	2.80	3.9	36.4
Standard, globally adjusted	4.39	2.80	3.9	36.2
Kinetic, raw data mean	4.26	2.86	3.7	35.1
Kinetic, globally adjusted	4.23	2.84	3.7	35.2
Mixed, raw data mean	3.88	2.80	3.6	35.9
Mixed, globally adjusted	3.77	2.82	3.6	35.7
Thermodynamic, raw data mean	2.83	1.83	3.7	55.0
Thermodynamic, globally adjusted	2.83	1.84	3.7	54.8

Tab. 42: Values for the left and right error margin of a 95 % confidence interval of the kinetic values for five experiments, based on the assumption of T-distributed individual experiments.

Protocol	k_{on} 10^4 M^{-1} s^{-1}	k_{off} 10^{-2} s^{-1}	τ_{ITC} s	τ_{rd} s
Standard	4.33	2.80	3.9	36.4
Kinetic	4.39	2.80	3.9	36.2
Mixed	4.26	2.86	3.7	35.1
Thermodynamic	4.23	2.84	3.7	35.2

Fluorinated and *para*-Substituted Compounds

Tab. 43: Thermodynamic values for individual measurements for compounds **21** – **28** and **34** – **36**.

Compound	χ^2	$\Delta G°$ kJ mol^{-1}	$\Delta H°$ kJ mol^{-1}	$-T\Delta S°$ kJ mol^{-1}	K_a 10^6 M^{-1}	K_d nM^{-1}
21	0.17	-37.6	-48.2	10.6	3.92	255
	0.28	-38.4	-48.2	9.8	5.28	189
	0.10	-38.6	-48.1	9.6	5.73	174
22	0.28	-40.1	-47.2	7.1	10.69	93
	0.20	-40.6	-46.8	6.2	12.81	78
	0.34	-40.0	-48.1	8.1	10.36	96
23	0.24	-36.9	-38.8	1.9	2.94	340
	0.58	-37.2	-41.8	4.5	3.33	300
	0.12	-37.3	-39.5	2.2	3.43	291
24	1.47	-41.8	-43.5	1.7	21.06	47
	0.18	-42.0	-42.9	0.9	23.04	43
	0.33	-42.5	-41.9	-0.6	28.09	35
25	0.43	-43.0	-54.6	11.6	33.53	29
	0.37	-42.0	-55.4	13.3	22.81	43
	0.59	-43.0	-53.9	10.9	34.16	29
26	0.19	-38.2	-43.1	4.9	4.91	203
	0.16	-38.1	-41.9	3.7	4.79	208
	0.50	-38.6	-42.1	3.5	5.68	176
27	0.19	-40.7	-41.5	0.7	13.63	73
	0.37	-41.1	-42.3	1.1	16.17	61
	0.19	-41.0	-41.7	0.7	15.42	64
28	0.46	-42.6	-44.0	1.4	29.59	33
	0.31	-42.6	-44.1	1.6	28.45	35
	0.25	-42.7	-44.5	1.8	30.41	32
34	0.71	-41.1	-44.2	3.1	15.96	62
	0.28	-41.0	-43.3	2.2	15.40	64
	0.77	-40.9	-43.9	3.0	14.69	68
35	0.29	-41.4	-45.1	3.8	17.59	56
	0.47	-41.9	-45.5	3.6	21.58	46
	0.30	-42.1	-44.9	2.8	24.02	41
36	0.13	-31.5	-41.8	10.3	0.33	2999
	0.20	-30.9	-46.9	16.0	0.26	3817
	0.45	-31.5	-45.5	14.1	0.33	3060

Tab. 44: Thermodynamic data for compounds **21** – **31**, **34** and **35**.

Compound	$\Delta G°$ kJ mol⁻¹	$\Delta H°$ kJ mol⁻¹	$-T\Delta S°$ kJ mol⁻¹	K_d nM	Error values $\Delta G°$ kJ mol⁻¹	$\Delta H°$ kJ mol⁻¹	$-T\Delta S°$ kJ mol⁻¹	K_d nM
21	-38.2	-48.2	9.9	200	0.3	0.04	0.3	24.8
22	-40.2	-47.4	7.2	90.7	0.2	0.4	0.5	5.7
23	-37.3	-39.8	2.5	287	0.1	0.9	0.8	14.9
24	-42.1	-42.7	0.6	41.7	0.2	0.5	0.7	3.5
25	-42.8	-54.4	11.7	32.1	0.3	0.4	0.7	4.8
26	-38.3	-42.3	4.0	192	0.1	0.4	0.4	10.2
27	-40.9	-41.9	1.0	68.6	0.1	0.2	0.4	3.5
28	-42.6	-44.2	1.6	34.2	0.07	0.2	0.1	0.9
29	-45.5	-44.0	-1.5	10.6	0.09	0.1	0.07	0.4
30	-48.6	-42.3	-6.3	3.1	0.08	0.1	0.09	0.1
31	-48.1	-39.5	-8.5	3.8	0.05	0.2	0.08	0.07
34	-41.1	-43.8	2.7	63.5	0.05	0.2	0.05	1.3
35	-41.7	-45.2	3.5	49.6	0.2	0.1	0.2	3.7
36	-31.3	-44.8	13.5	3280	0.2	1.2	1.4	215

Thermodynamic parameters from global fitting. Error values for compounds **21** – **28** represent the standard error of 3 measurements. K_d and $\Delta H°$ and the values of $\Delta G°$ and $\Delta S°$ calculated from the former two are given. Error bars represent the global fit error for Kd and $\Delta H°$ and the error calculated from those of the former two for $\Delta G°$ and $\Delta S°$.

Tab. 45: Kinetic parameters for individual ITC experiments for compounds **21** – **28**, **34** and **35** before (left) and after (right) application of globally adjusted thermodynamic parameters.

Compound	Before global adjustment			After global adjustment		
	χ^2	k_{on} $10^4\,M^{-1}\,s^{-1}$	k_{off} $10^{-2}\,s^{-1}$	χ^2	k_{on} $10^4\,M^{-1}\,s^{-1}$	k_{off} $10^{-2}\,s^{-1}$
21	0.74	6.84	1.7	0.85	8.10	1.6
	0.39	7.13	1.4	0.40	6.95	1.4
	0.36	7.54	1.3	0.37	7.04	1.4
22	0.74	11.07	1.0	0.65	11.25	1.0
	0.39	12.05	0.9	0.55	11.21	1.0
	0.36	10.50	1.0	0.55	10.72	1.0
23	0.74	6.20	2.1	0.44	7.05	2.0
	0.39	7.68	2.3	0.64	7.84	2.2
	0.36	7.80	2.3	0.39	7.90	2.3
24	0.74	15.22	0.7	1.70	16.15	0.7
	0.39	18.58	0.8	1.16	18.94	0.8
	0.36	18.50	0.7	0.49	17.45	0.7
25	0.74	16.13	0.5	0.37	15.49	0.5
	0.39	12.58	0.6	0.36	12.82	0.4
	0.36	18.01	0.5	0.26	17.33	0.6
26	0.74	7.95	1.6	0.47	8.22	1.6
	0.39	8.71	1.8	0.59	9.15	1.8
	0.36	9.98	1.8	0.65	9.55	1.8
27	0.74	14.66	1.1	0.51	15.30	1.1
	0.39	14.28	0.9	0.34	13.69	0.9
	0.36	15.29	1.0	0.55	14.92	1.0
28	0.74	17.43	0.6	0.52	17.41	0.6
	0.39	16.50	0.6	0.28	16.69	0.6
	0.36	18.56	0.6	0.43	18.15	0.6
34	0.74	15.20	1.0	0.72	15.08	1.0
	0.39	15.56	1.0	1.19	15.80	1.0
	0.36	15.43	1.1	0.75	15.91	1.0
35	0.74	14.57	0.8	0.71	15.45	0.8
	0.39	16.76	0.8	0.23	16.26	0.8
	0.36	16.56	0.7	1.04	15.34	0.8

Tab. 46: Kinetic data for compounds **21** – **31**, **34** and **35**.

Compound	k_{on} 10^5 M^{-1} s^{-1}	k_{off} 10^{-3} s^{-1}	Response time / s	k_{on} 10^5 M^{-1} s^{-1}	k_{off} 10^{-3} s^{-1}	Response time / s
				Standard error		
21	0.74	14.7	4.1	0.04	0.7	0.03
22	1.11	10.0	4.5	0.02	0.2	0.06
23	0.76	21.8	4.5	0.03	0.8	0.07
24	1.80	7.3	4.4	0.08	0.3	0.2
25	1.52	4.9	4.1	0.10	0.4	0.2
26	0.90	17.2	4.4	0.04	0.8	0.08
27	1.46	10.0	4.1	0.05	0.3	0.2
28	1.74	6.0	4.6	0.04	0.1	0.05
34	1.56	9.9	4.7	0.02	0.1	0.04
35	1.57	7.8	4.8	0.02	0.1	0.009

Protonation Dependence

Tab. 47: Mean values for thermodynamic parameters for the interaction of compound **4** with hCAII and standard error for three experiments.

Buffer	ΔH_{ion} kJ mol^{-1}	$\Delta G°$ kJ mol^{-1}	$\Delta H°$ kJ mol^{-1}	$-T\Delta S°$ kJ mol^{-1}	$\Delta G°$ kJ mol^{-1}	$\Delta H°$ kJ mol^{-1}	$-T\Delta S°$ kJ mol^{-1}
					Standard error		
HEPES	21.1	-35.9	-41.9	-6.1	0.1	0.3	0.3
TRICIN	32.0	-35.7	-43.5	-7.9	0.4	0.8	1.0
TRIS	48.1	-35.7	-45.1	-9.3	0.2	0.2	0.4
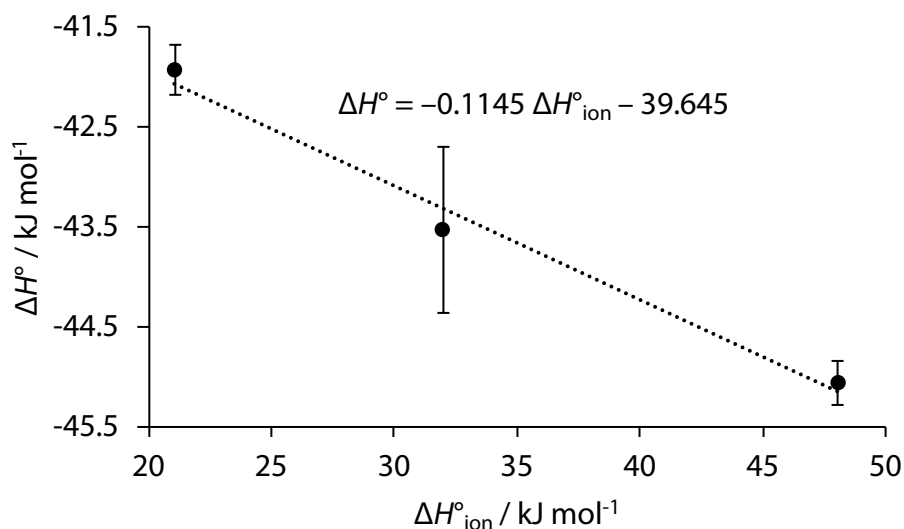
Fig. 68: Plot of $\Delta H°$ as a function of ΔH_{ion} for the titration of hCAII with compound **4** in HEPES, TRIS and TRICIN from Tab. 47.

Determination of the Instrument Response Time

2 µL of an aqueous solution of ethanol (p.a., 1 %, v/v) were titrated into demineralized water at an injection rate of 0.5 µL s^{-1}. The syringe plunger motor was primed with a preceding injection of 0.3 µL. The procedure was repeated three times. The resulting heat signals were plotted with QtiPlot 1.0.0-rc2. Equation (24) was fitted to the right half of the injection signal, beginning with the data point closest to half the amplitude as described by *Dumas et al.*[115] The constant C was added as signals were not shifted to $P_Q = 0$ before fitting.

$$P_m(t) = P_{max} \cdot e^{\frac{-t}{\tau_{ITC}}} + C \tag{24}$$

A

B

C
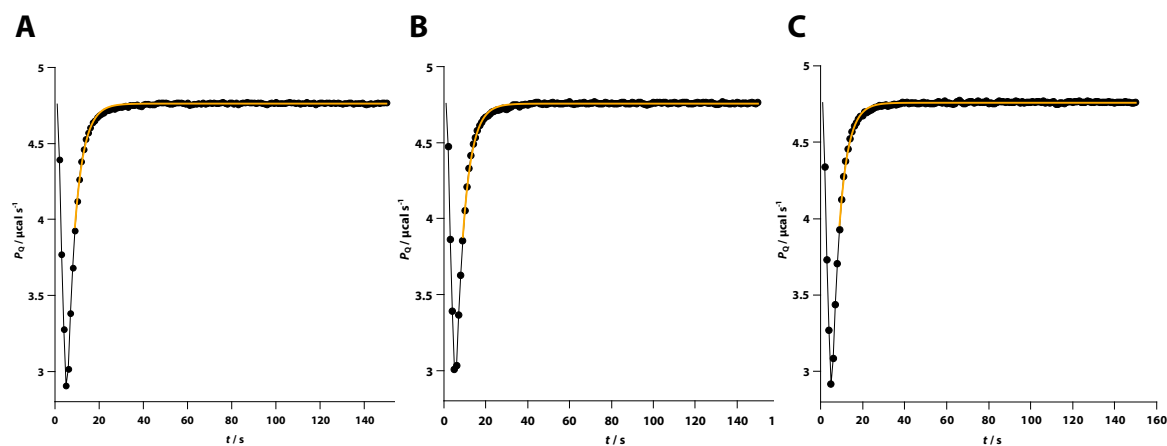
Fig. 69: Plotted heat signals with fitted equation 24 in orange. Data points are connected for clarity.

Tab. 48: Parameters for the fitting of equation (24) to heat signals in Fig. 69.

Measurement	P_{max}	τ_{ITC}	C
A	-6.61 ± 0.02	4.29 ± 0.06	4.76 ± 0.0008
B	-6.58 ± 0.20	4.47 ± 0.06	4.76 ± 0.0008
C	-6.44 ± 0.22	4.32 ± 0.08	4.76 ± 0.0008

Sulfoxide

Tab. 49: Thermodynamic parameters for individual measurements of compound **37** and **37** in combination with **43**.

Compound	χ^2	$\Delta G°$ kJ mol^{-1}	$\Delta H°$ kJ mol^{-1}	$-T\Delta S°$ kJ mol^{-1}	K_a 10^6 M^{-1}	K_d nM^{-1}
37	0.84	-35.5	-50.1	14.6	1.67	599
	0.17	-34.9	-53.2	18.2	1.32	756
	0.14	-35.0	-52.8	17.8	1.35	740
37/43	0.13	-35.0	-52.3	17.3	1.34	748
	0.30	-35.2	-54.9	19.7	1.45	689
	0.21	-34.4	-54.0	19.6	1.07	936

Tab. 50: Kinetic parameters for individual measurements of compound **37** and **37** in combination with **43** before (left) and after (right) global adjustment.

Compound	χ^2	Before global adjustment		χ^2	After global adjustment	
		k_{on} $10^4 \, \mathrm{M^{-1} \, s^{-1}}$	k_{off} $10^{-2} \, \mathrm{s^{-1}}$		k_{on} $10^4 \, \mathrm{M^{-1} \, s^{-1}}$	k_{off} $10^{-2} \, \mathrm{s^{-1}}$
37	0.41	5.61	4.2	0.46	5.44	4.3
	0.62	4.81	3.3	0.33	4.58	3.6
	1.09	4.10	3.8	1.11	4.94	3.9
37/43	0.97	3.77	2.3	0.95	3.46	2.4
	0.26	3.72	2.8	0.24	4.03	2.8
	0.45	3.65	2.7	0.45	3.90	2.7

Raw and Processed Thermograms, Integrated Heat Values and Equilibration Time Curves

Exemplary plots for the titration of every compound are provided. In addition to the raw data analysis, the integrated heat values and ETCs after application of the globally fitted parameters are shown. For compounds **9**, **29**, **30** and **31**, raw and processed thermograms and the globally fitted integrated heat values are shown. Thermograms for reference compound **20** are only provided with the data for compound **9**, as the same data were also used for compounds **29**, **30** and **31**.

n-Alkyl and *n*-Alkyloxy Compounds
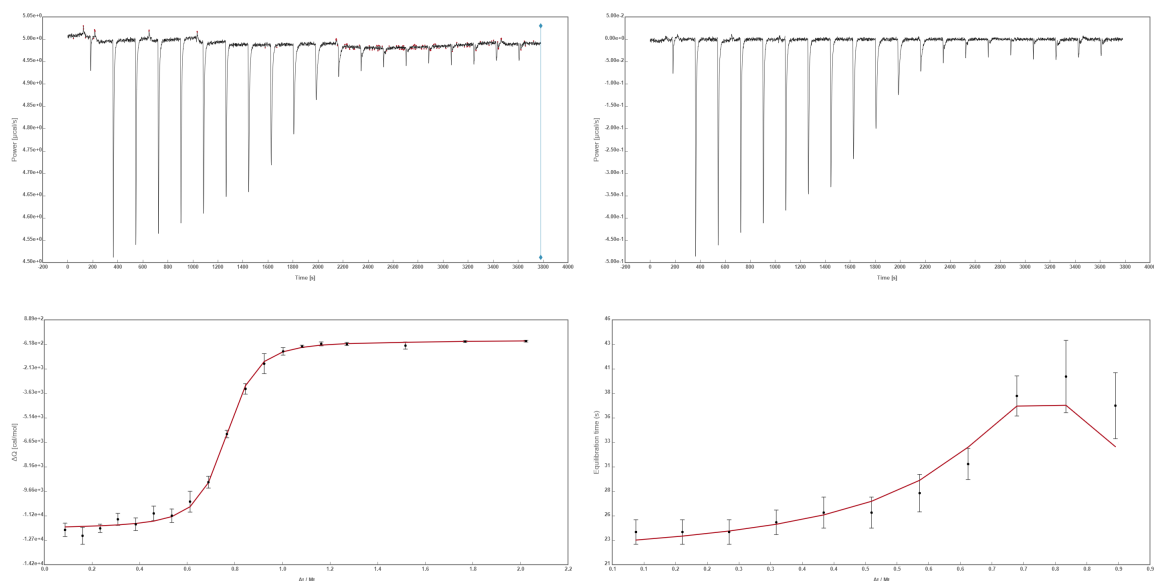
Compound 4

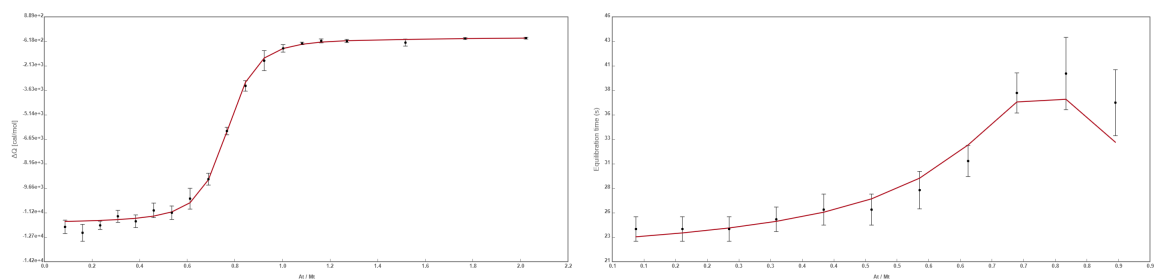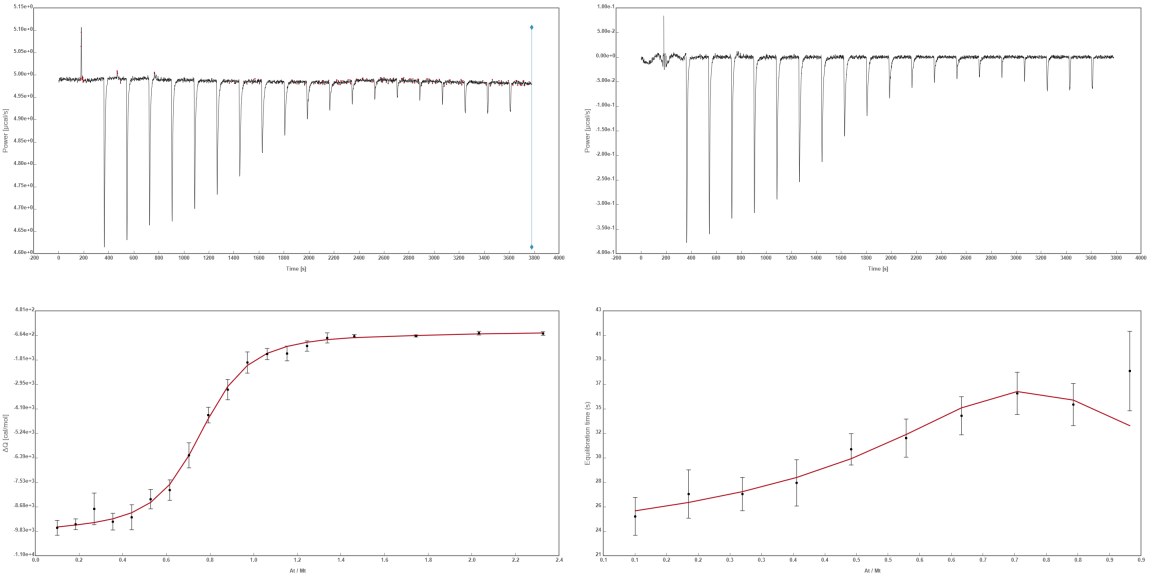### Raw data analysis
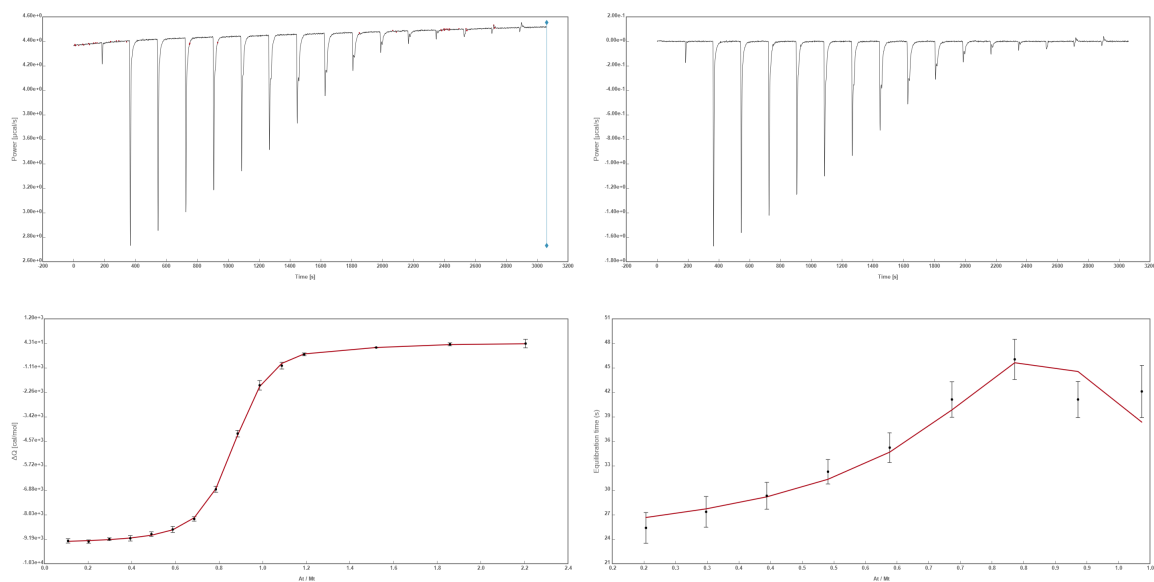
Adjustment with global parameters
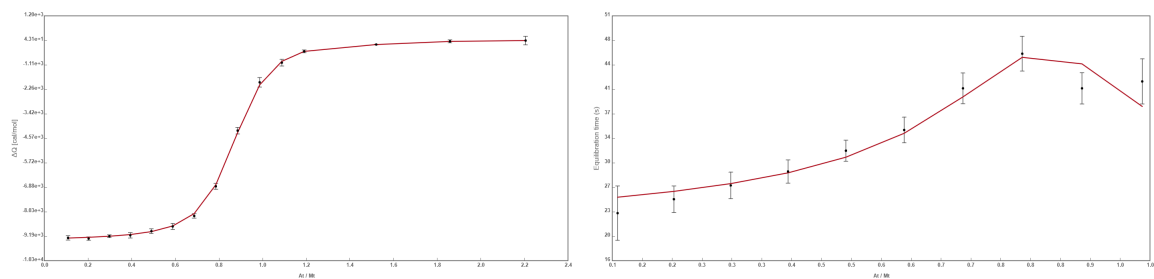
Compound 5

Raw data analysis
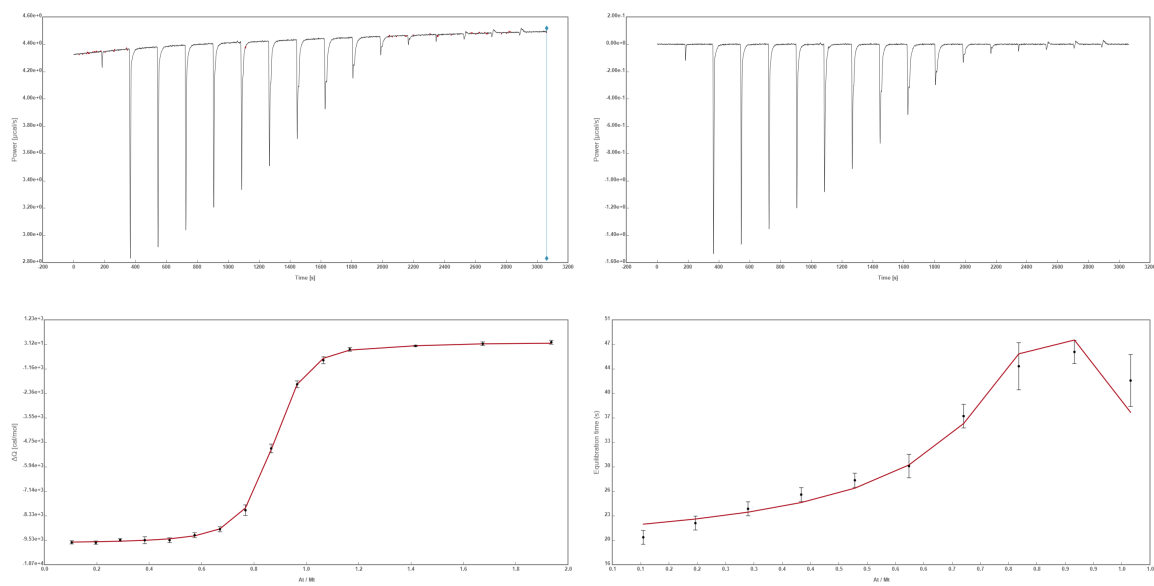
Adjustment with global parameters
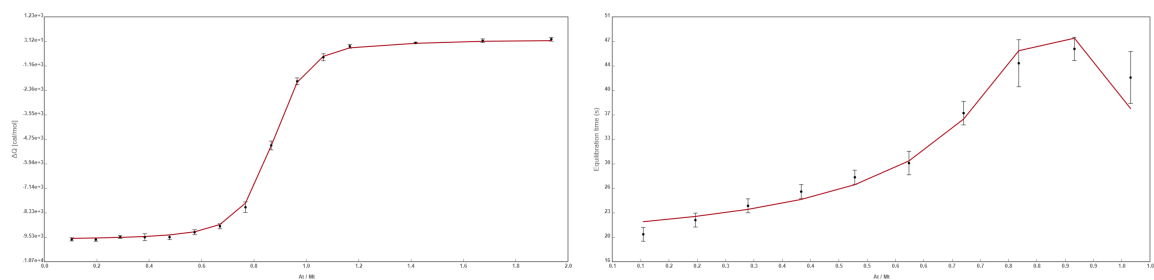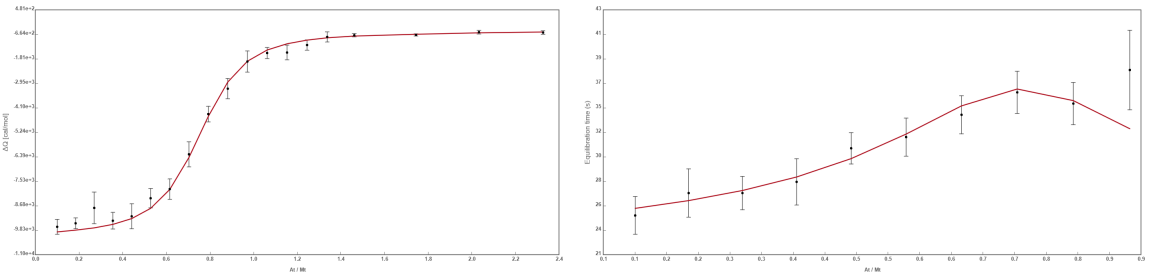
Compound 6

Raw data analysis
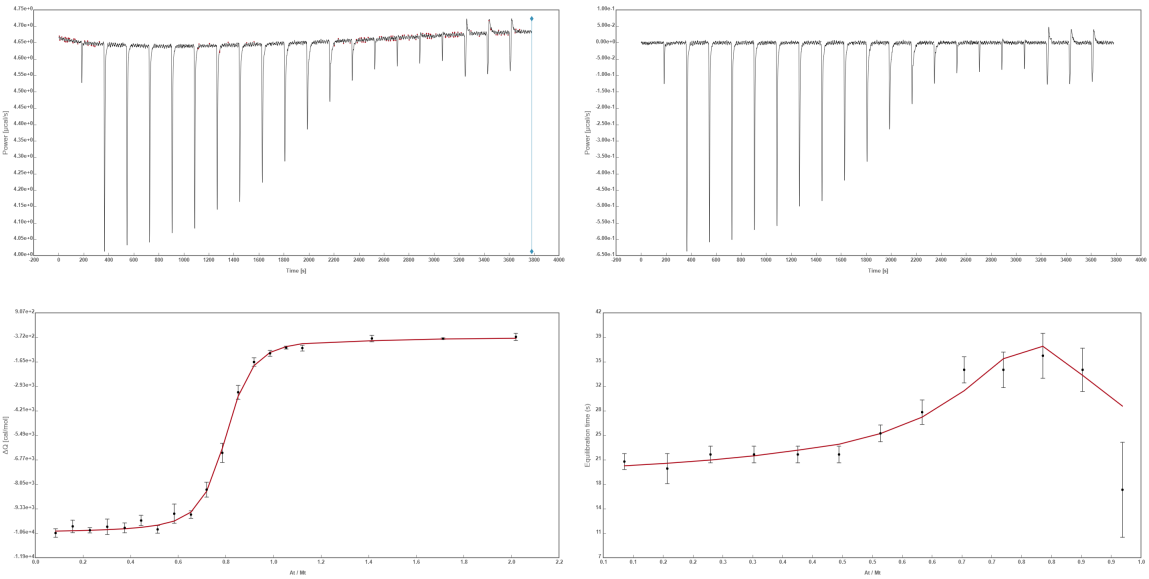
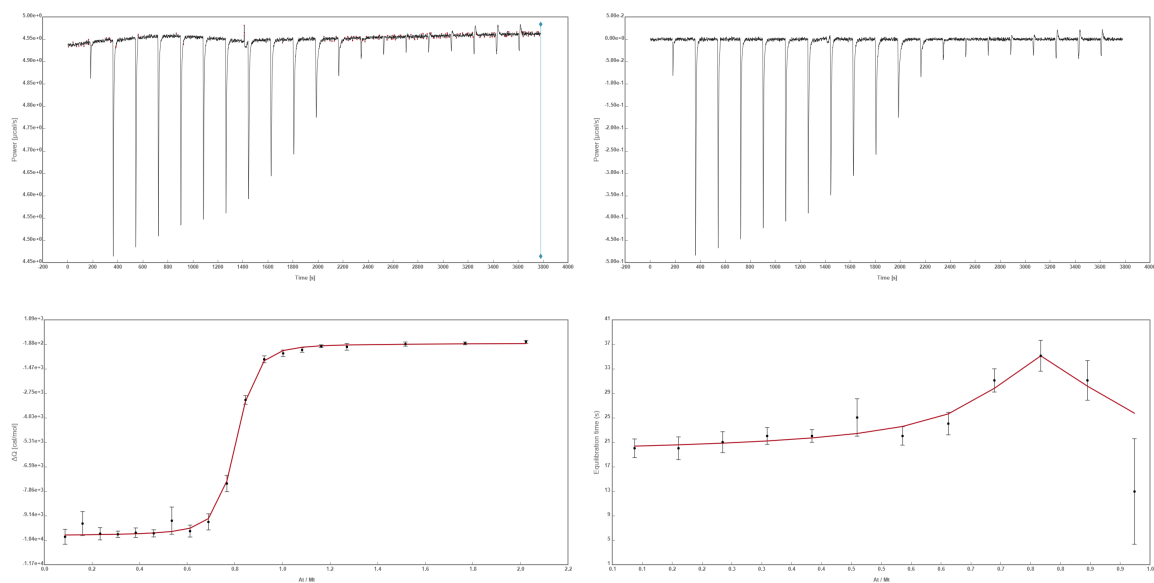### Adjustment with global parameters
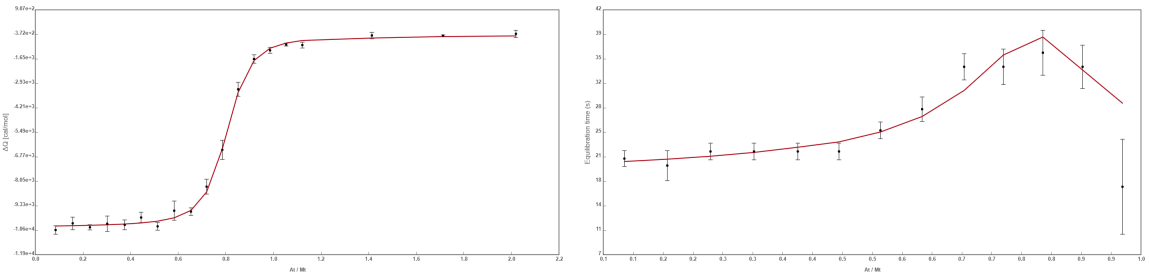
Compound 7

Raw data analysis
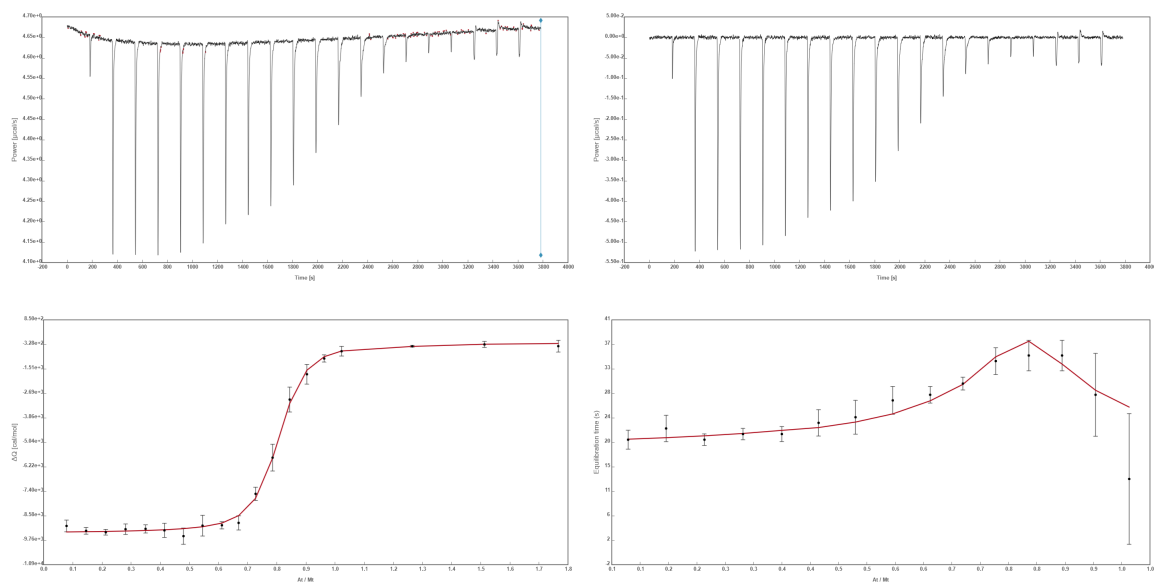
Adjustment with global parameters
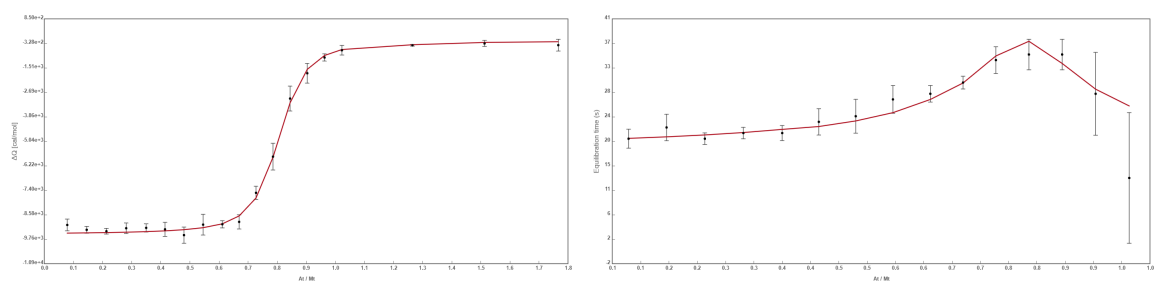
Compound 8

Raw data analysis
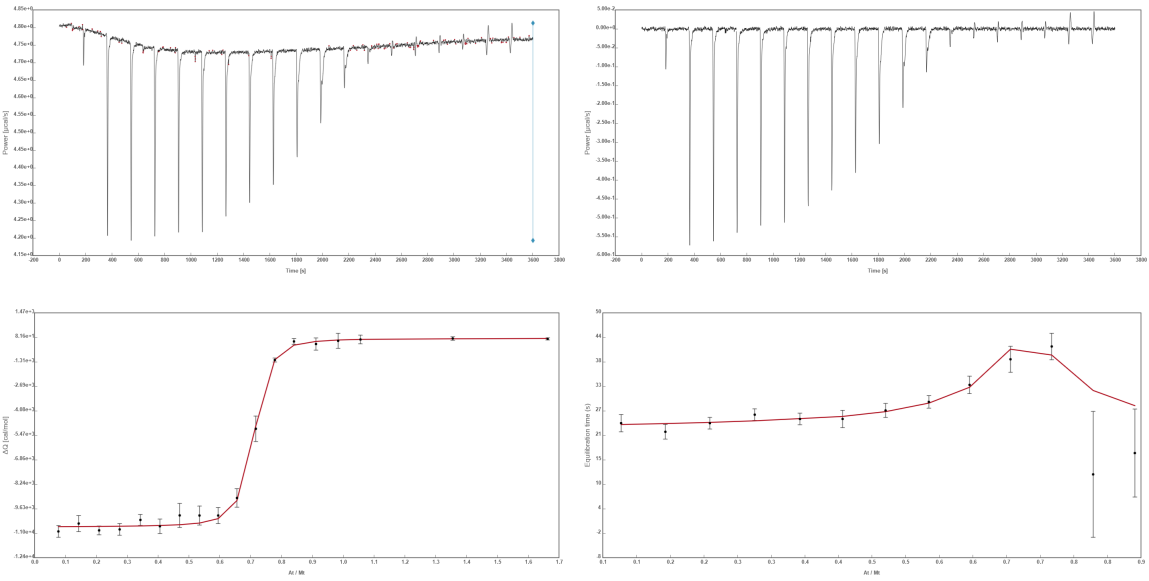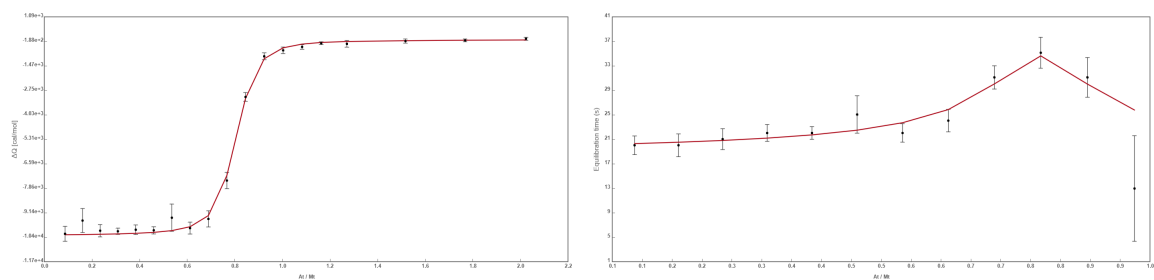

Adjustment with global parameters
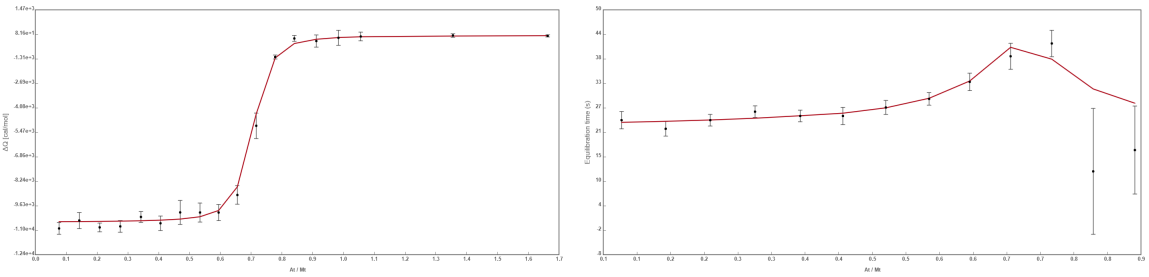
Compound 9

Raw data and global fit for direct titration
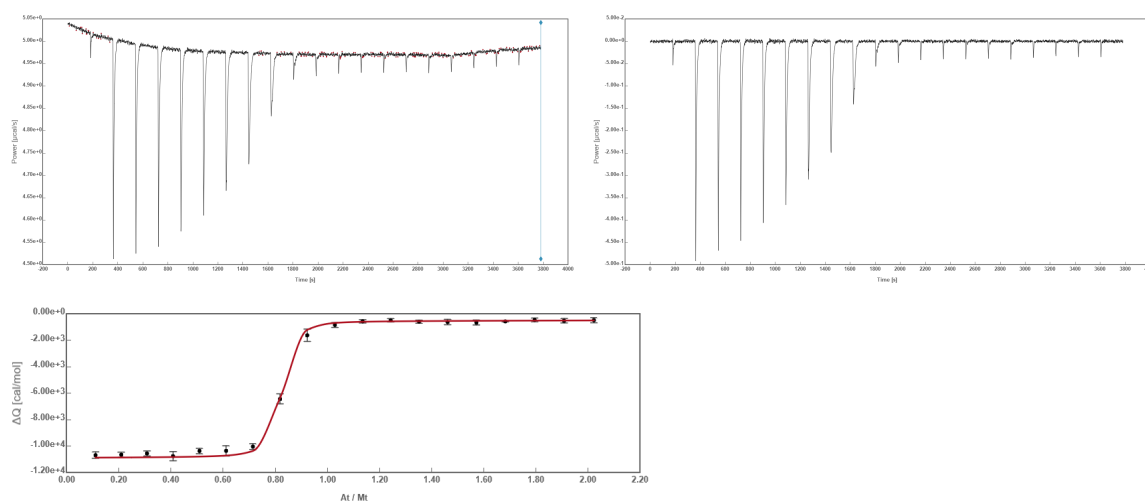
Raw data and global fit for reference compound **20**
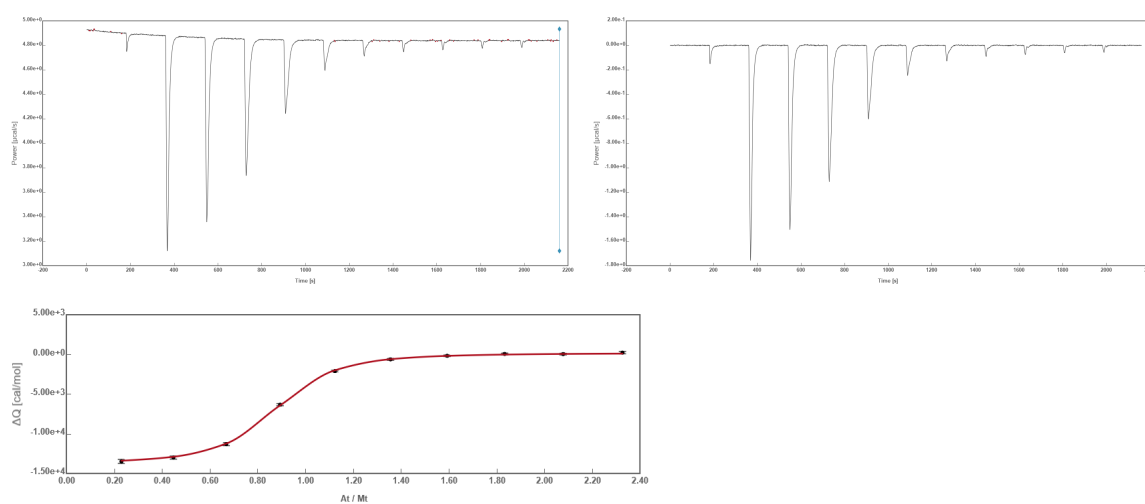
Raw data and global fit for displacement titration of compound **20** with **9**
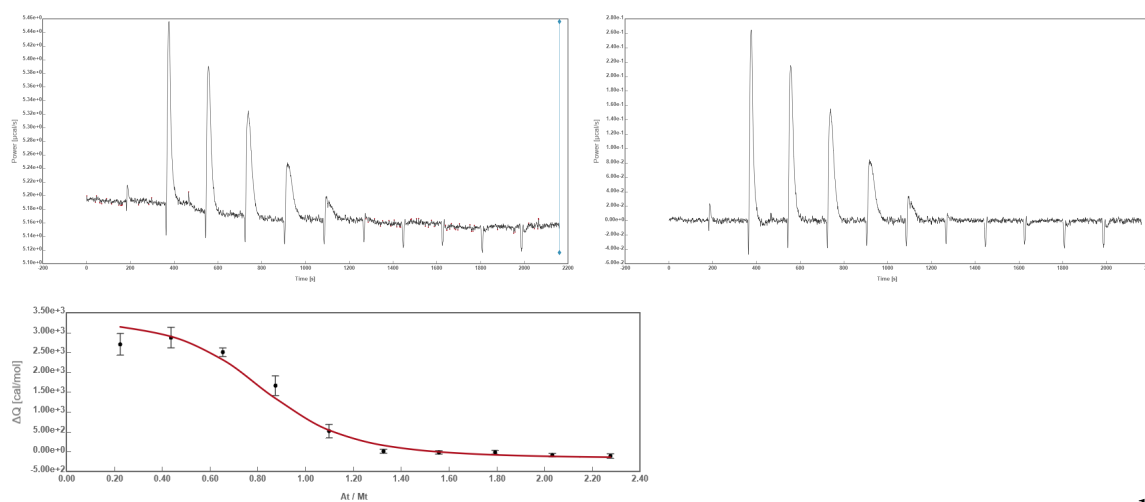
Compound 10

Raw data analysis
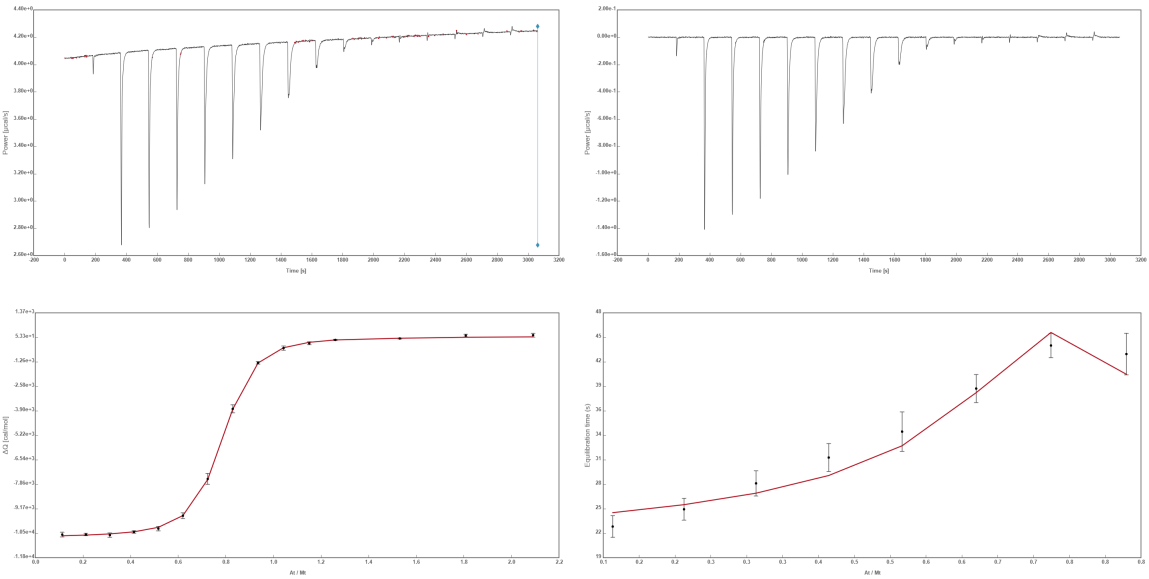
Adjustment with global parameters
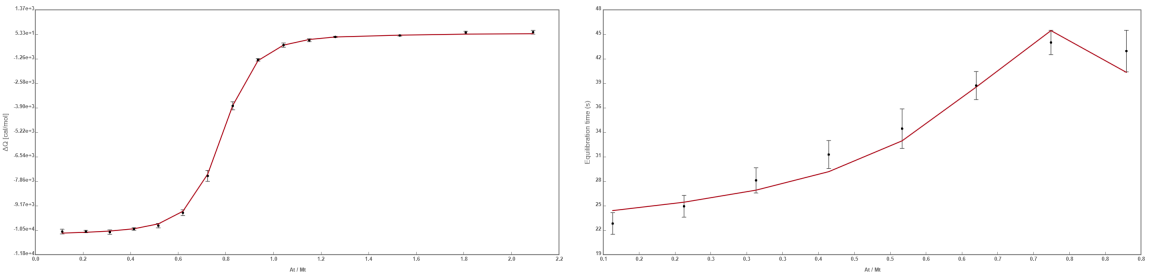
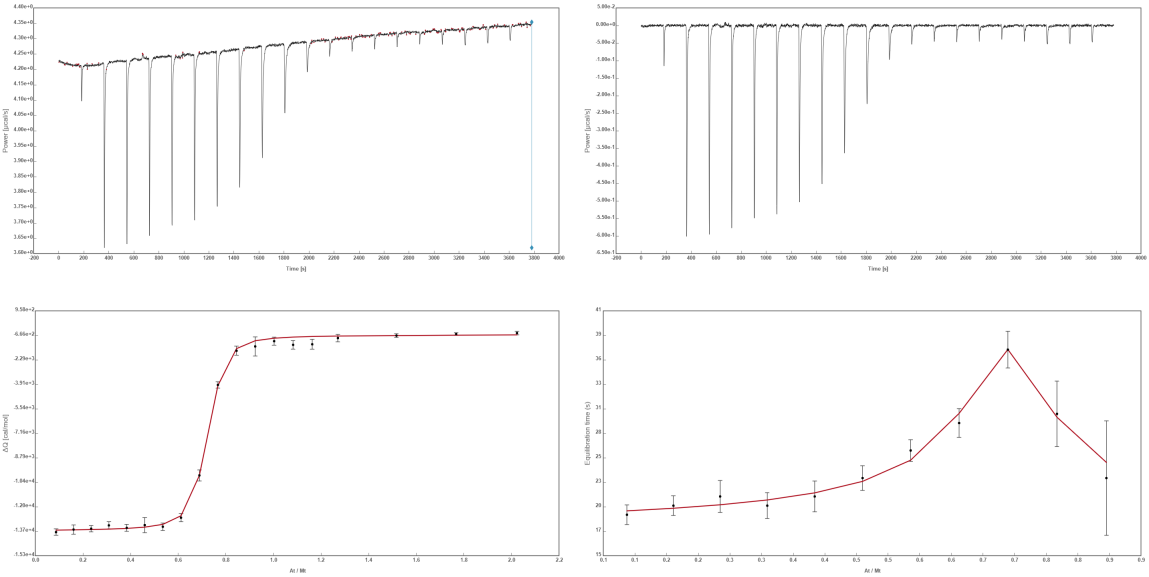## Compound 11

Raw data analysis
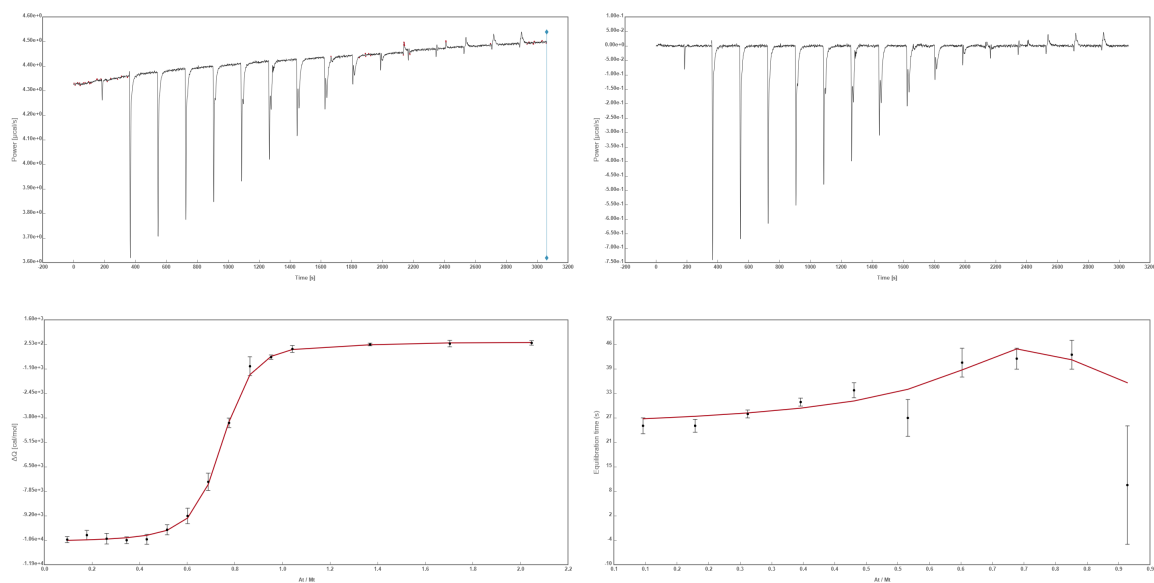
Adjustment with global parameters
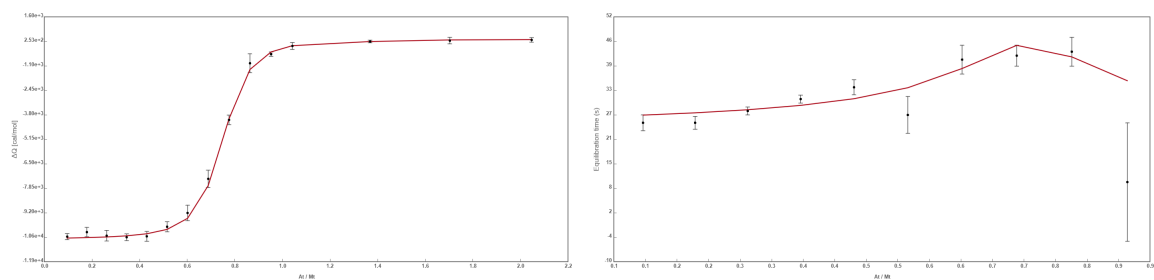
Compound 12

Raw data analysis
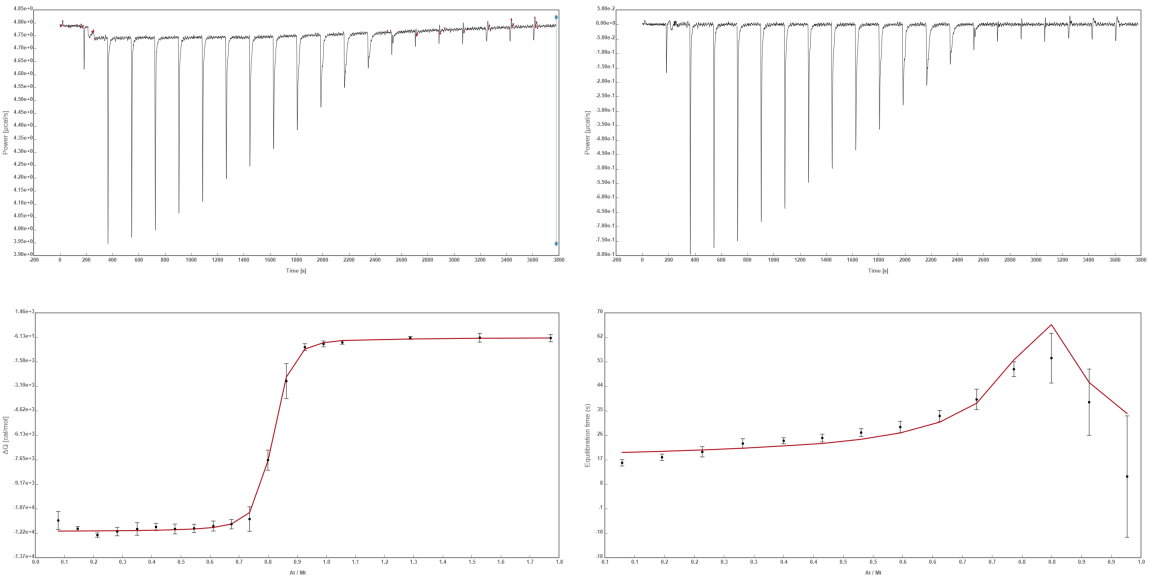
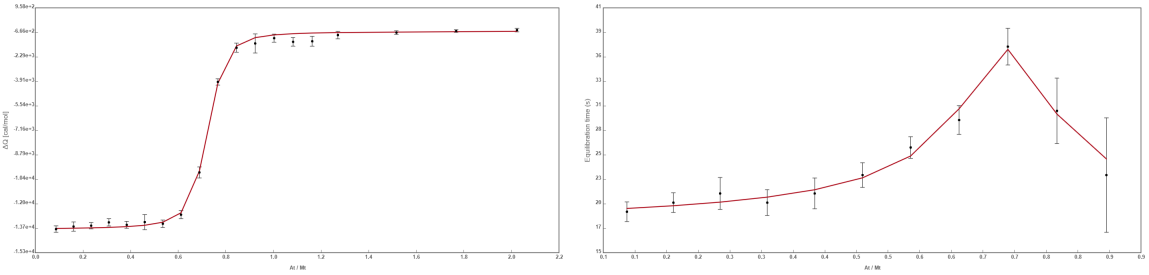

Adjustment with global parameters
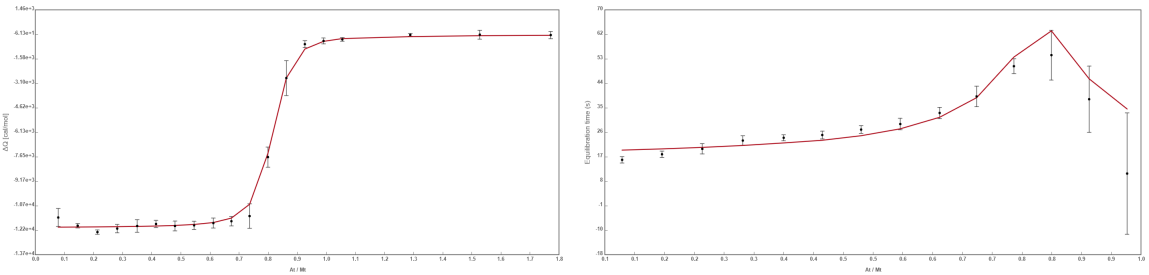
Compound 13

Raw data analysis
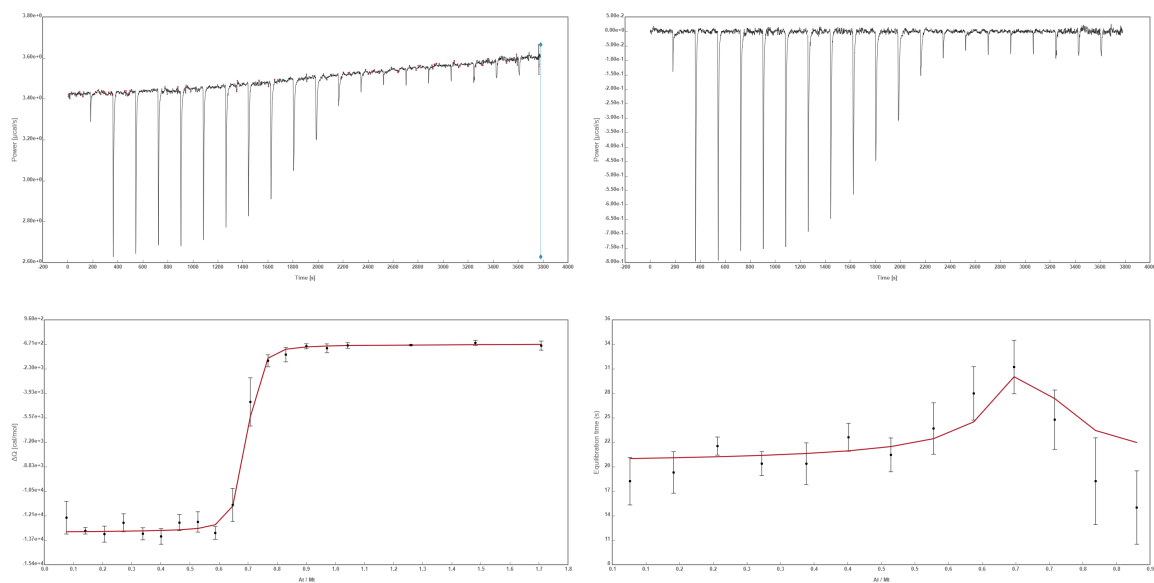
Adjustment with global parameters
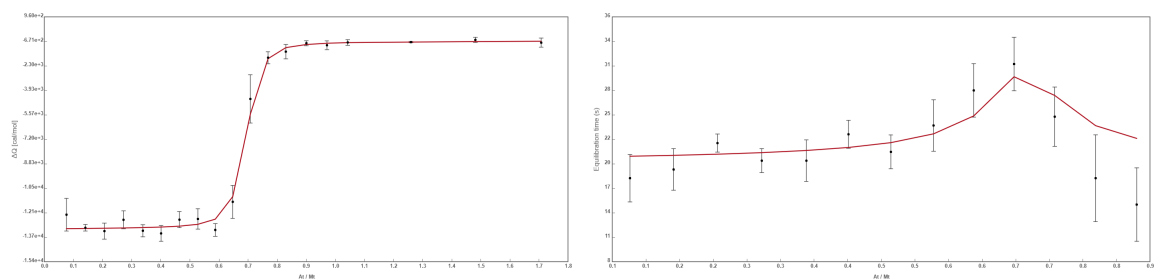
ITC Protocol

Standard protocol

Raw data analysis
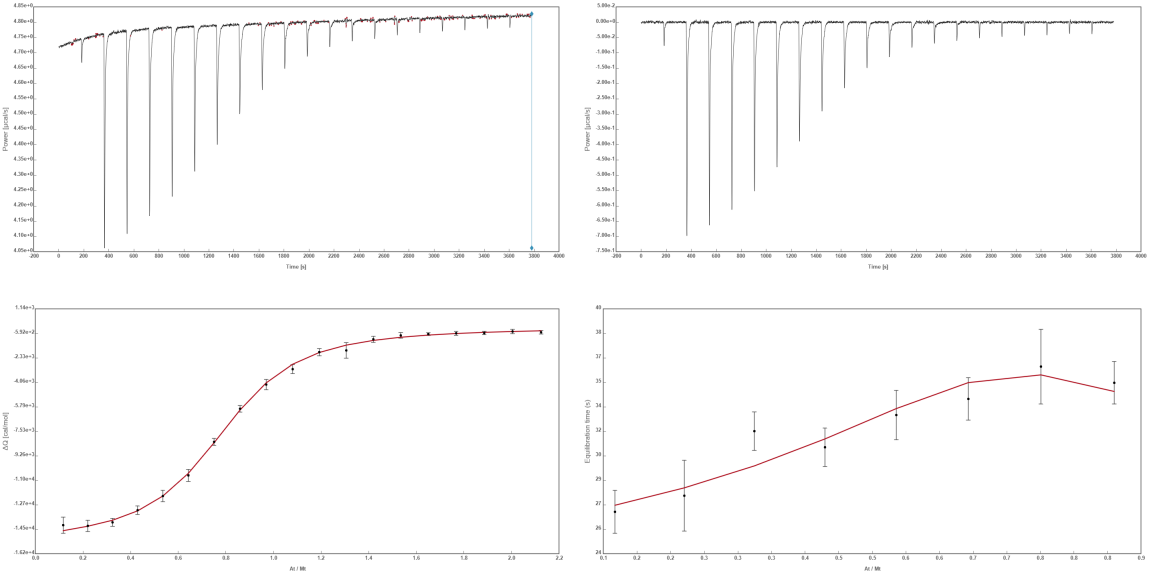
Adjustment with global parameters
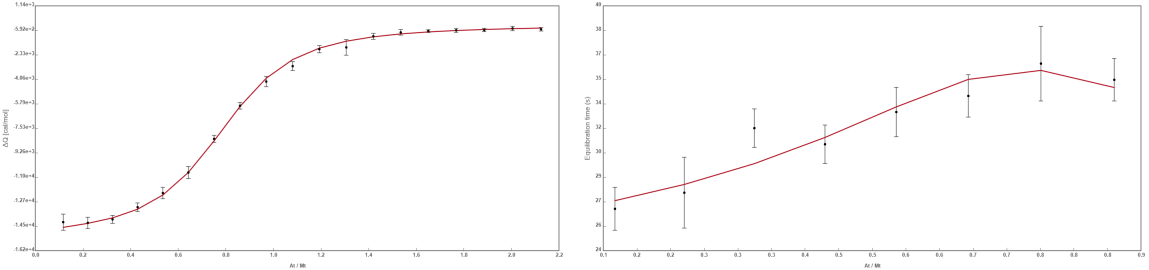
Kinetically improved protocol

Raw data analysis
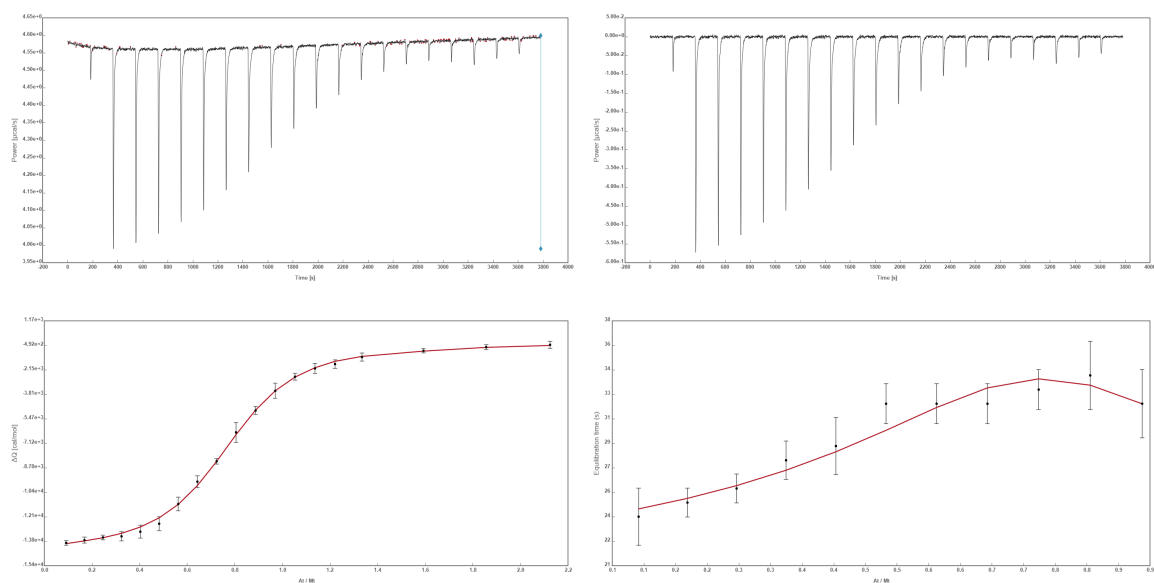
Adjustment with global parameters
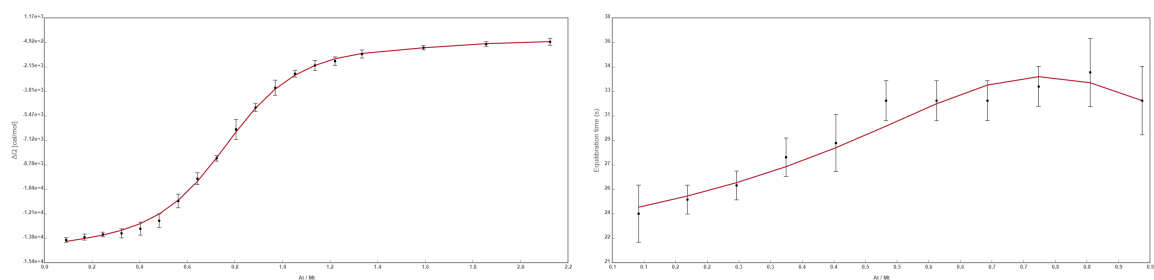
Mixed protocol

Raw data analysis
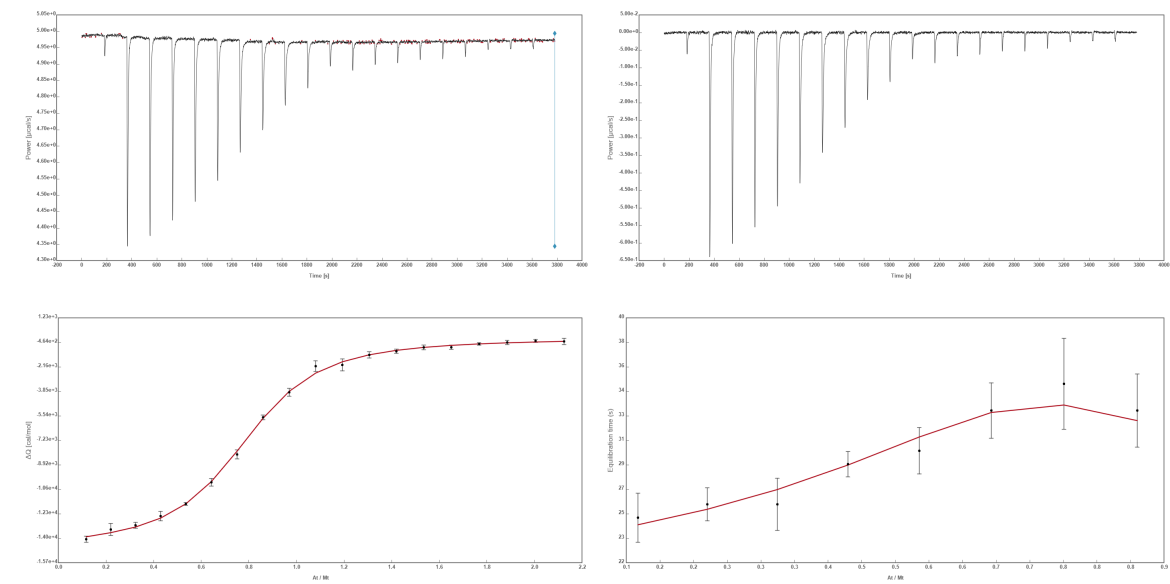
Adjustment with global parameters
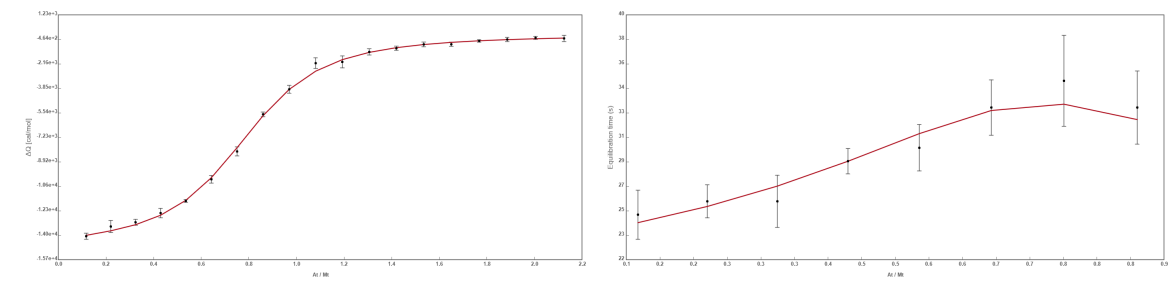
Thermodynamically optimized protocol

Raw data analysis
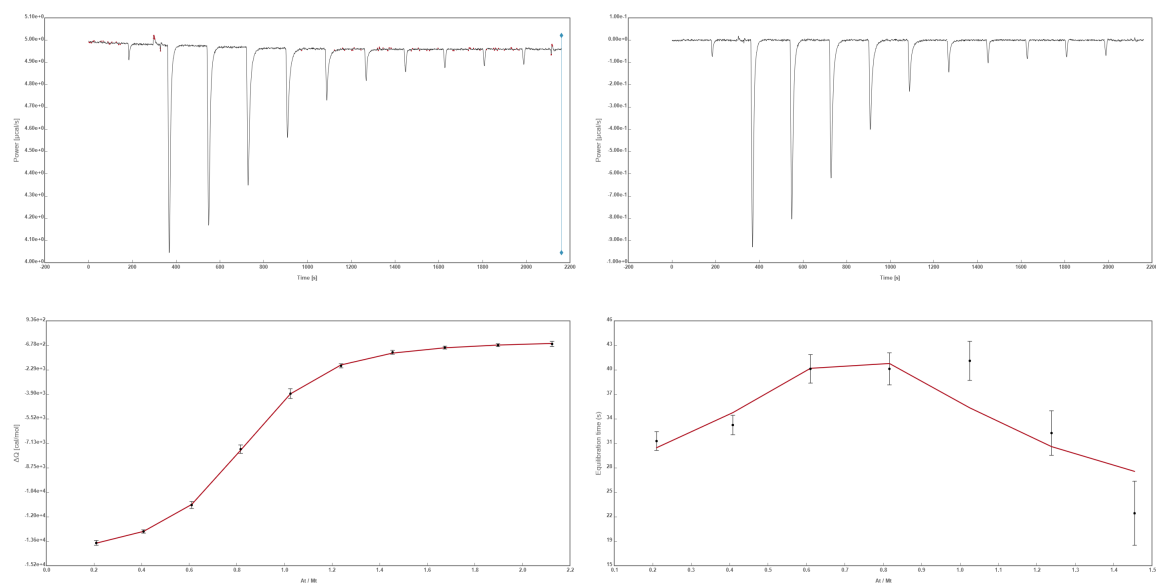
Adjustment with global parameters
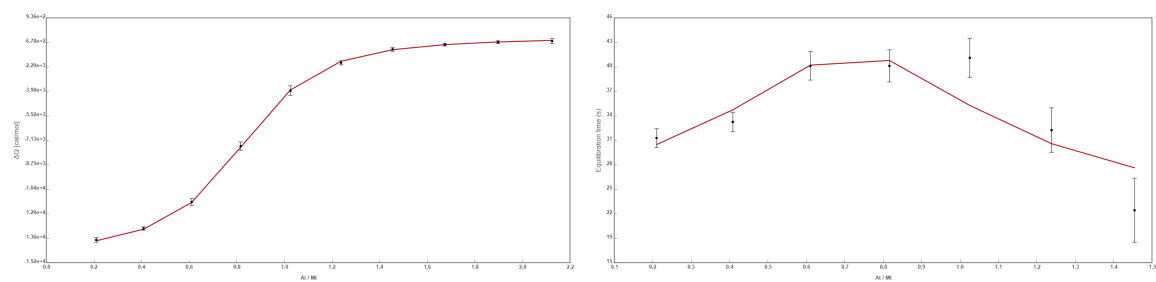
Fluorinated and *para*-substituted compounds

Compound 21

Raw data analysis
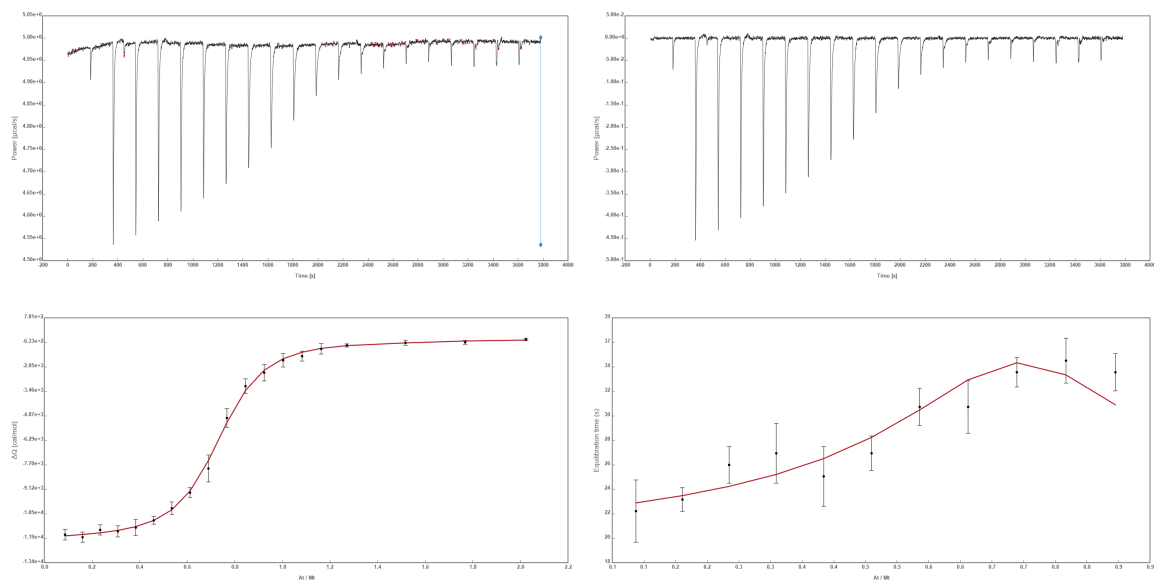
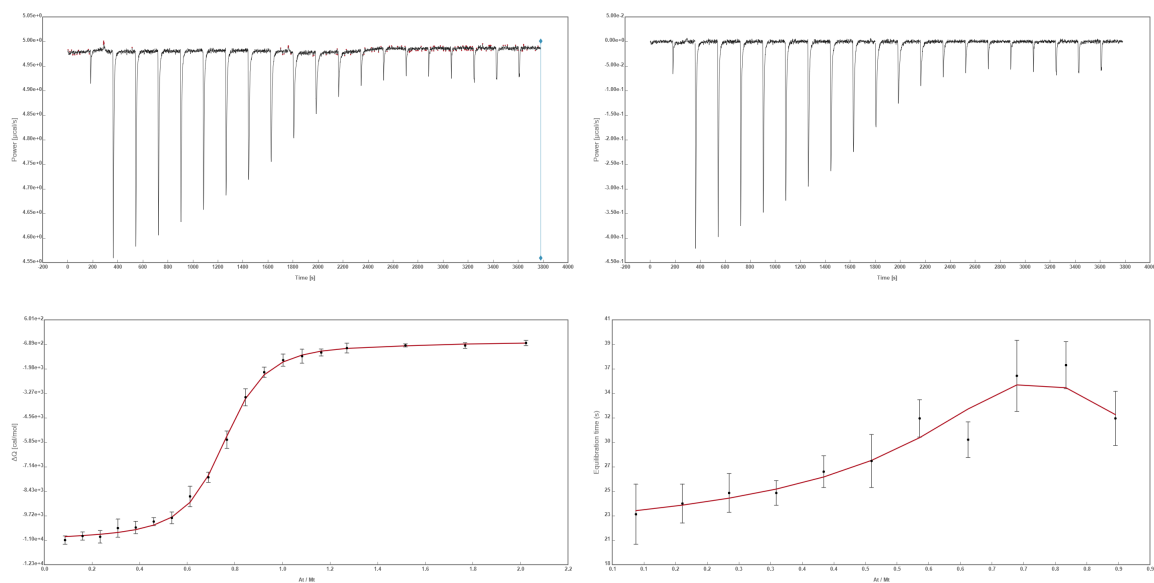### Adjustment with global parameters
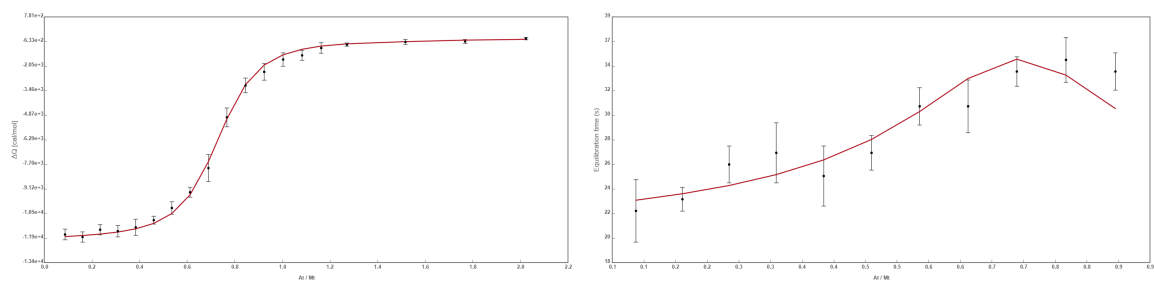
Compound 22
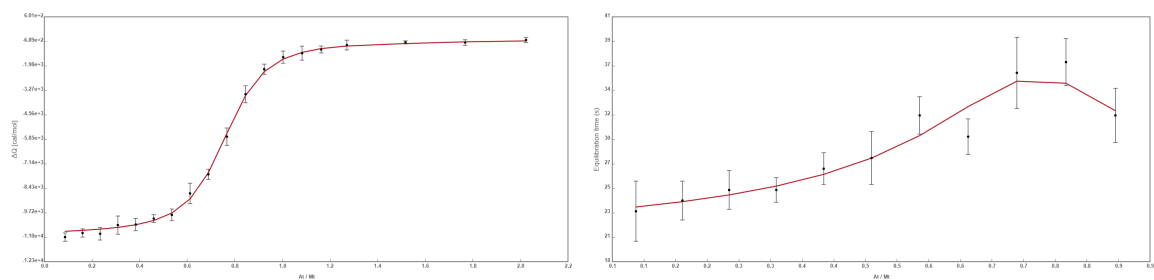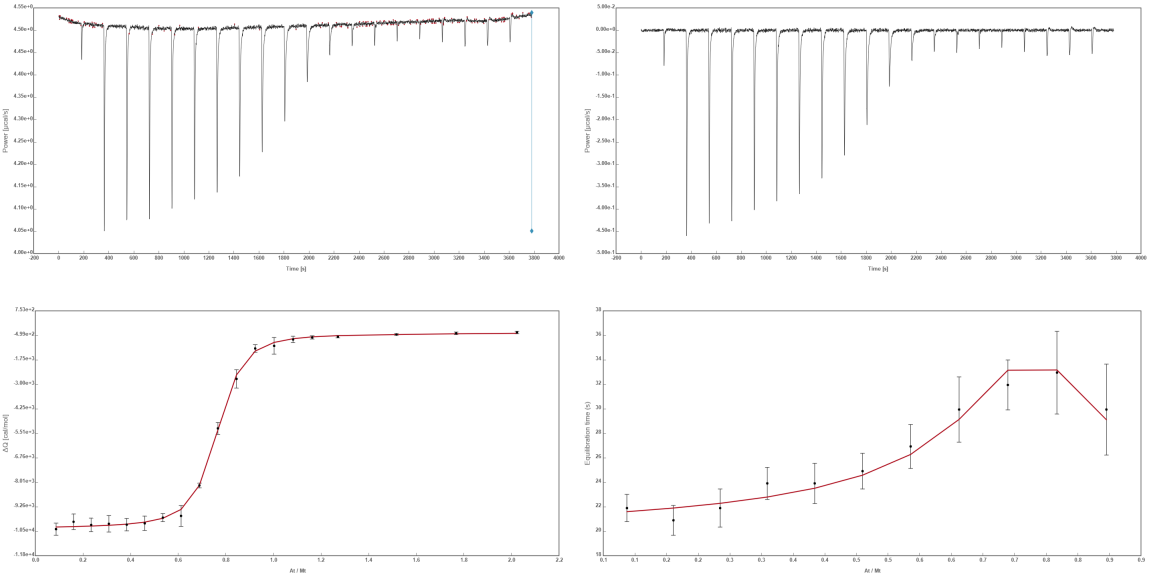
Raw data analysis

Adjustment with global parameters

Compound 23

Raw data analysis

Adjustment with global parameters

Compound 24

Raw data analysis

Adjustment with global parameters

Compound 25

Raw data analysis

Adjustment with global parameters

Compound 26

Raw data analysis

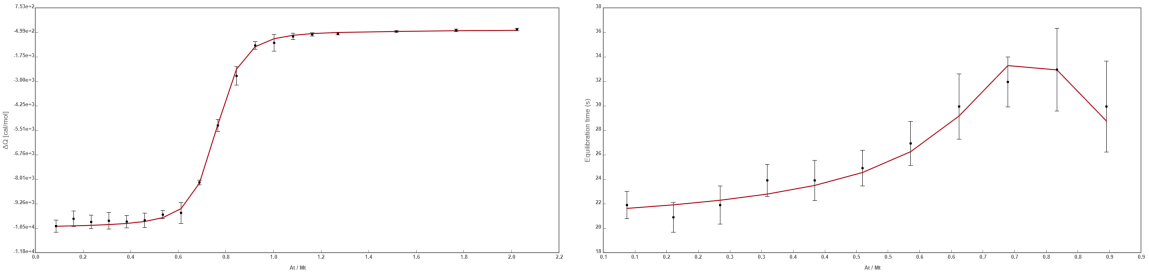

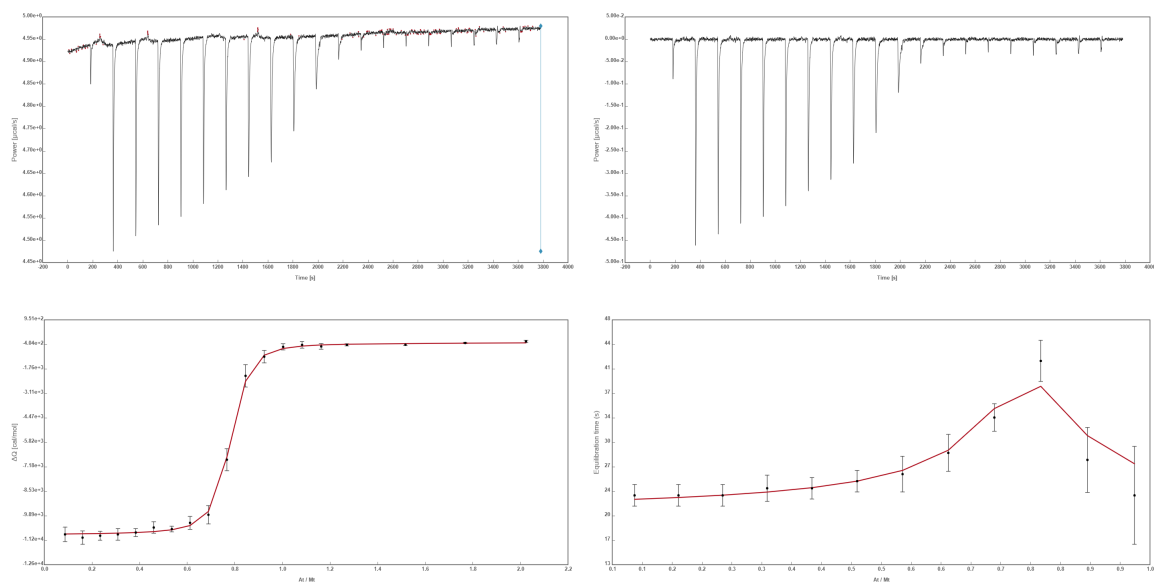### Adjustment with global parameters

Compound 27
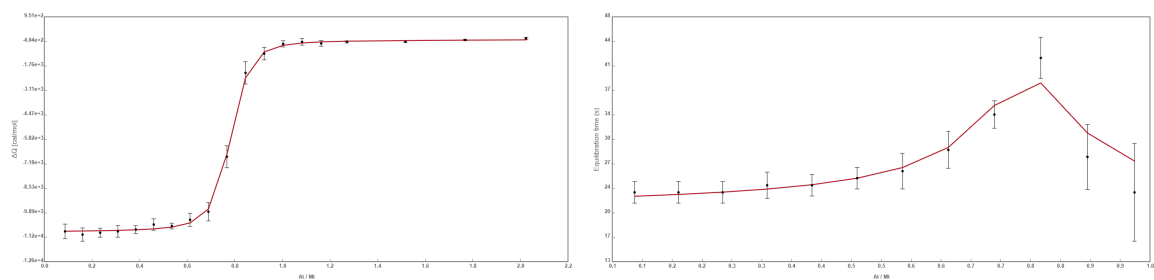
Raw data analysis

Adjustment with global parameters
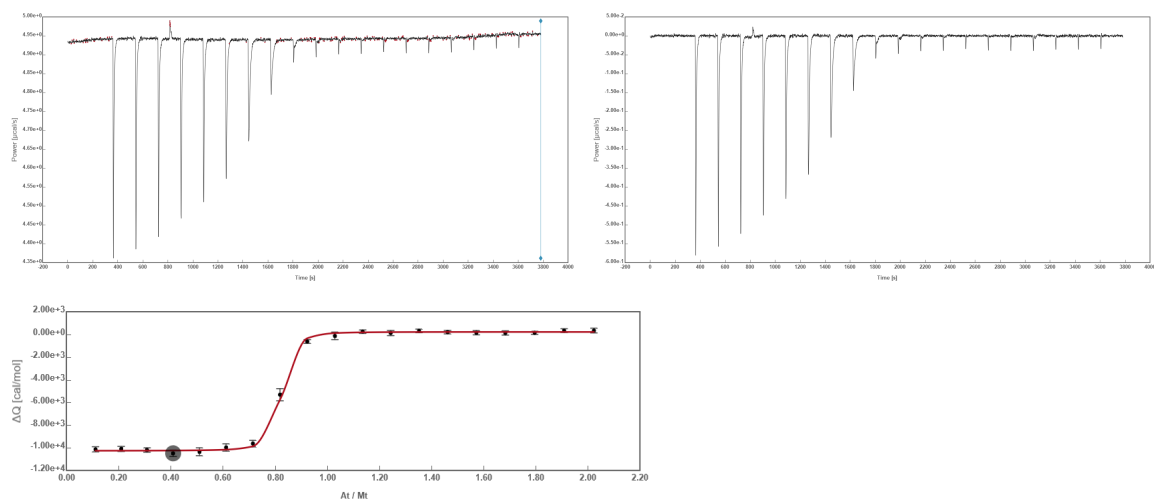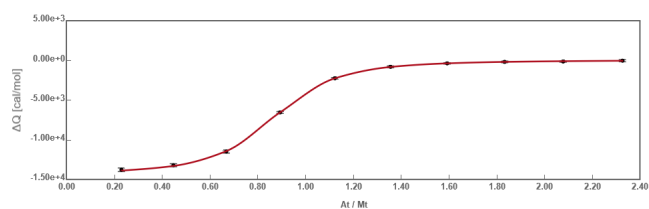
Compound 28

Raw data analysis

Adjustment with global parameters
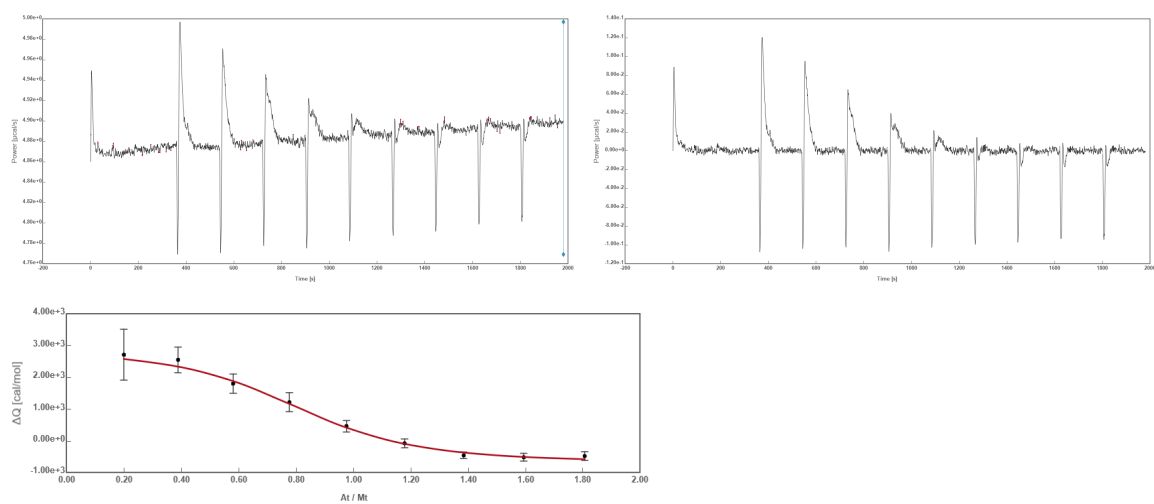
Compound 29

Raw data and global fit for direct titration

Global fit for reference compound **20**

Raw data and global fit for displacement titration of compound **20** with **29**

Compound 30

Raw data and global fit for direct titration

Global fit for reference compound **20**

Raw data and global fit for displacement titration of compound **20** with **30**
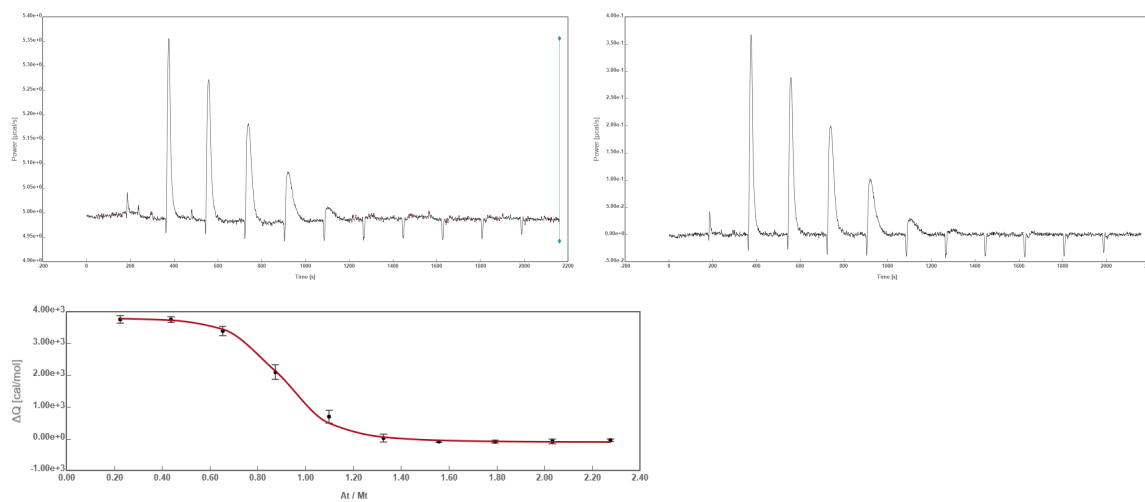
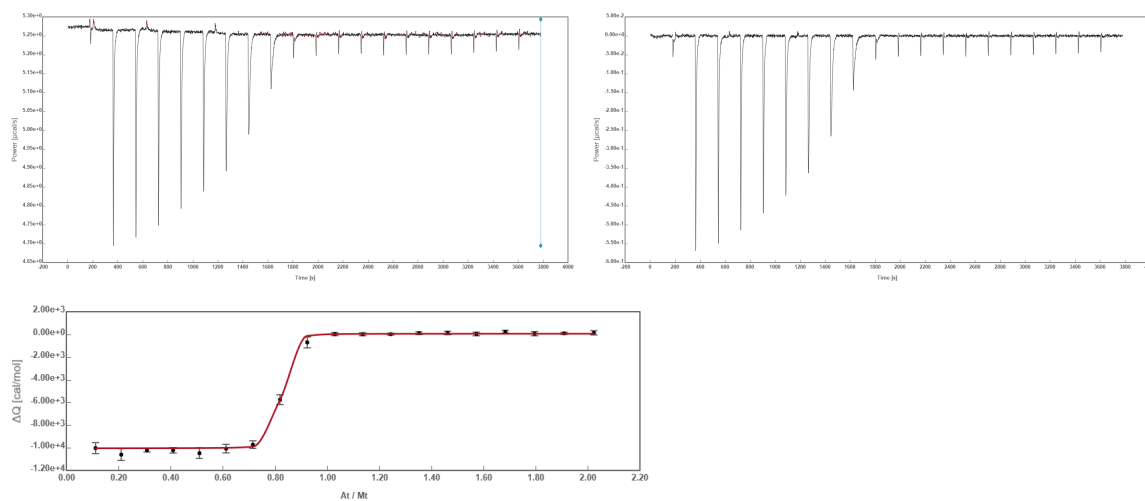## Compound 31

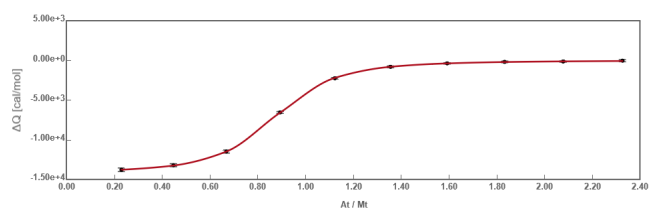Raw data and global fit for direct titration
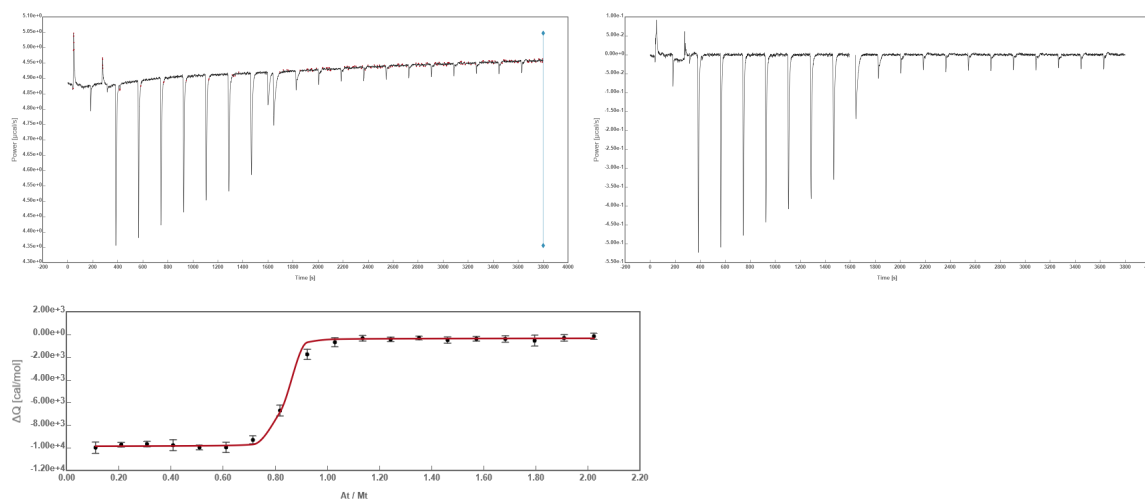
Global fit for reference compound **20**
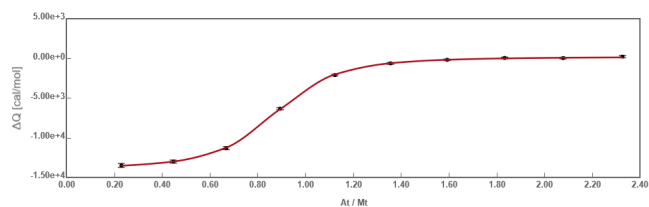
Raw data and global fit for displacement titration of compound **20** with **31**
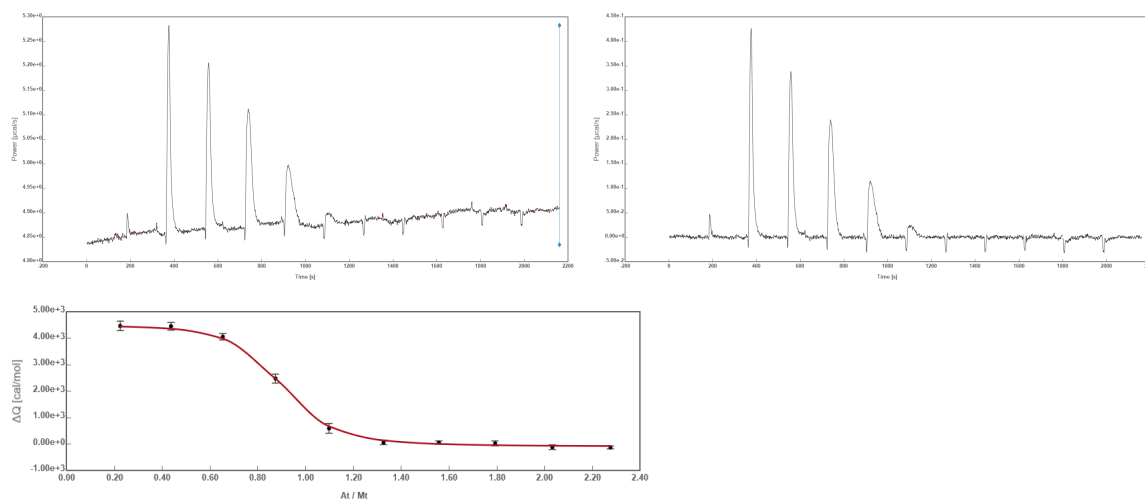
Compound 34

Raw data analysis
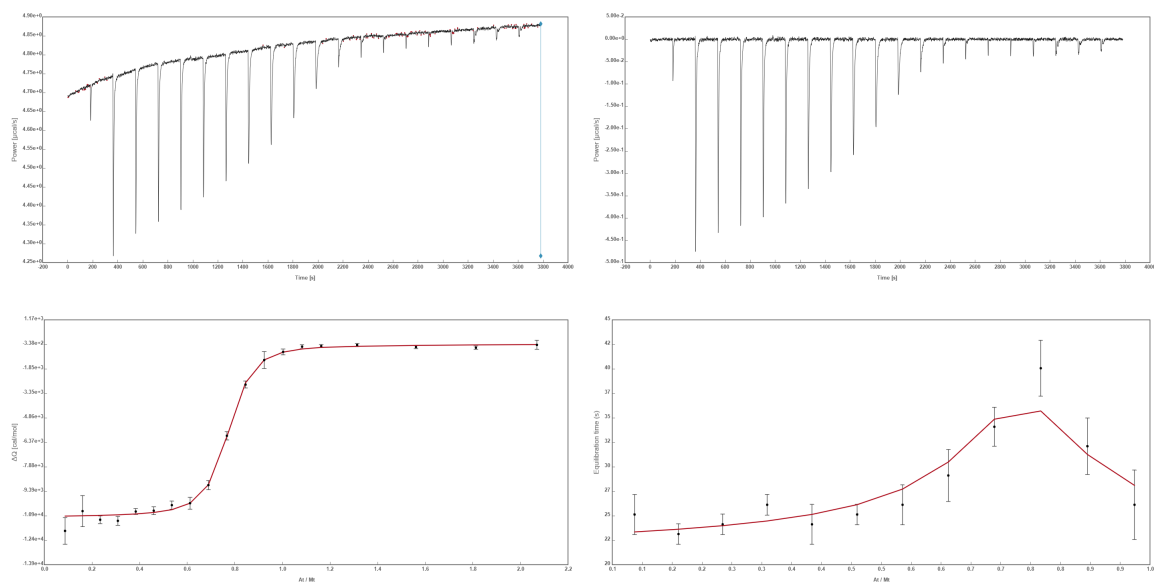
Adjustment with global parameters
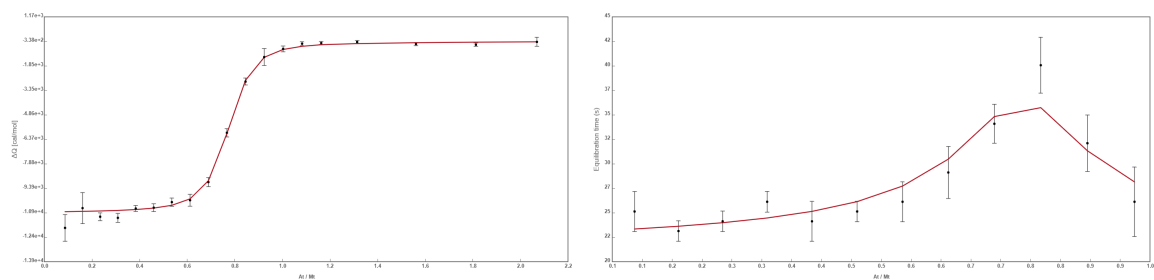
Compound 35

Raw data analysis
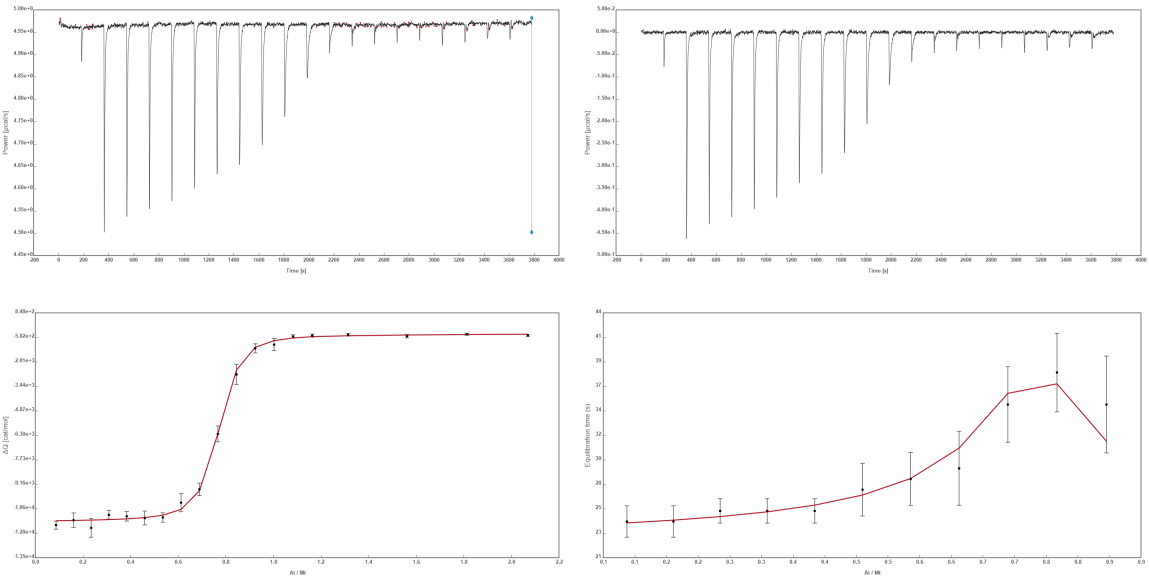
Adjustment with global parameters
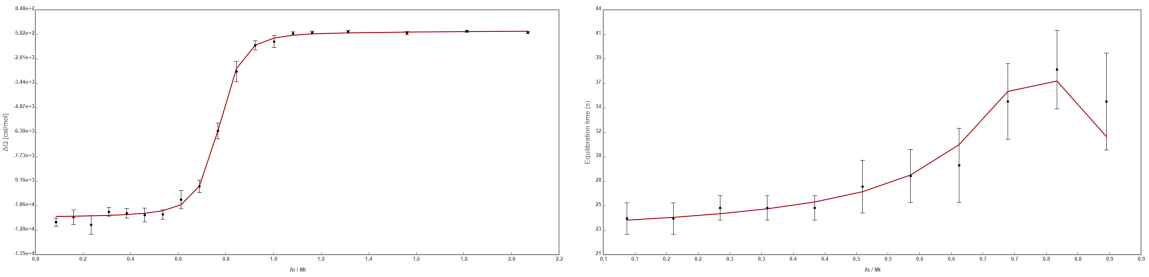
Compound 36

Raw data analysis
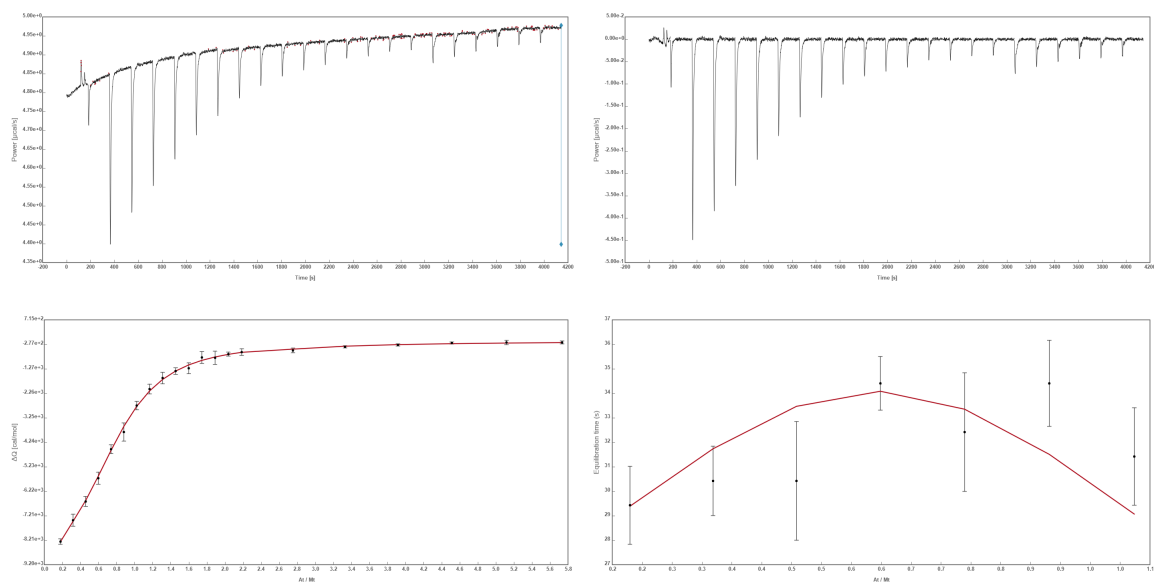
Adjustment with global parameters
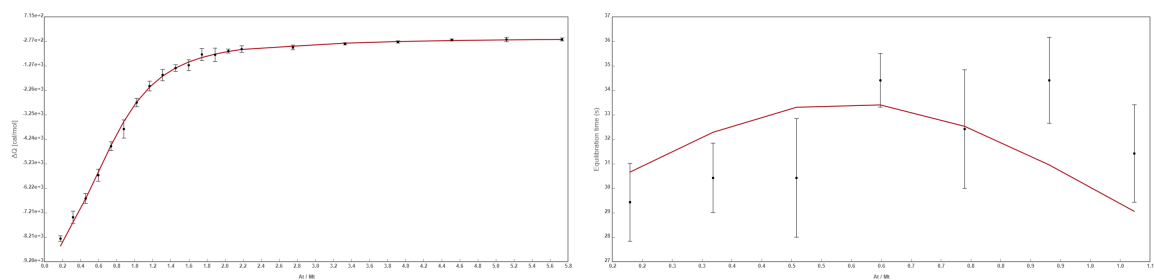
SRE-Bearing Benzenesulfonamide

Compound 1

Raw data analysis
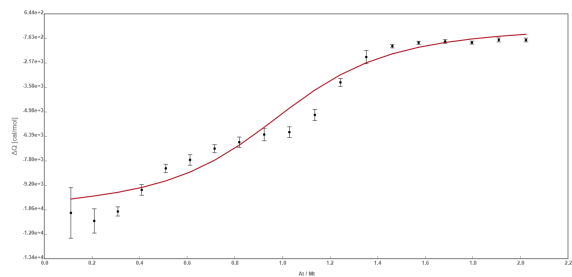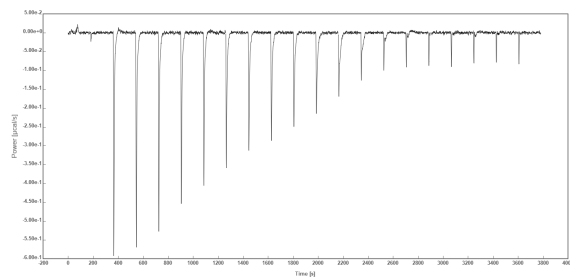
Dilution titration of **1** in buffer
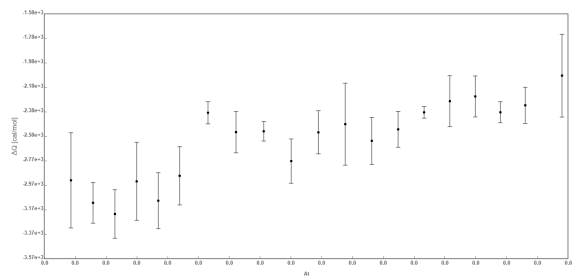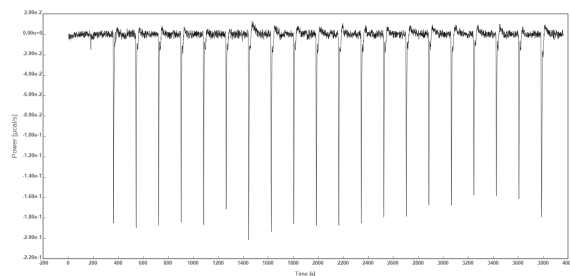
Oxazolidinedione

Compound 38

(*R*)–**38**
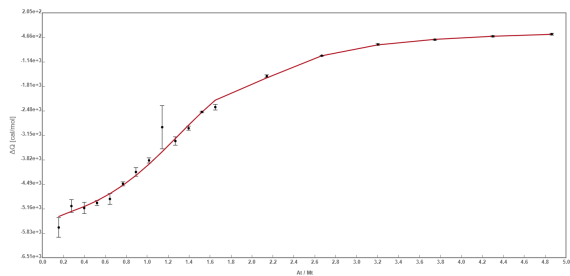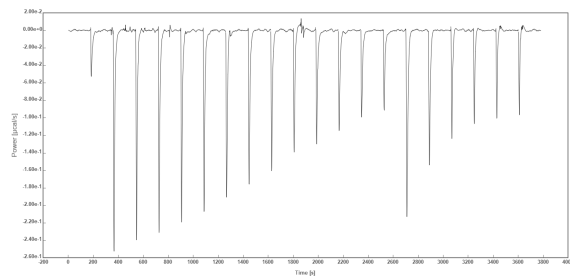
(*S*)–**38**
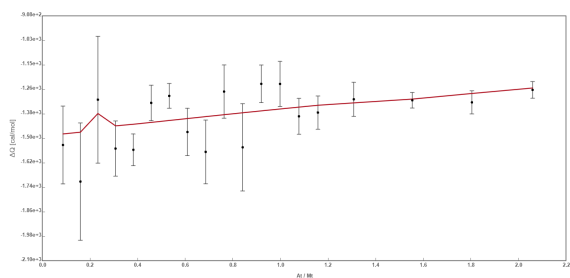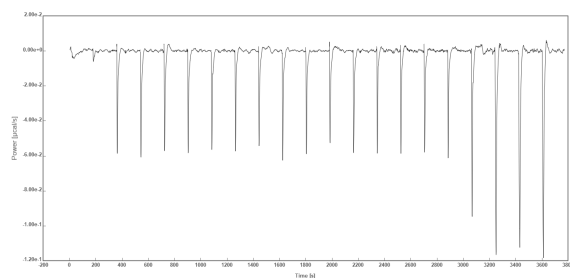
Remote Binder

Compound 37

Raw data analysis
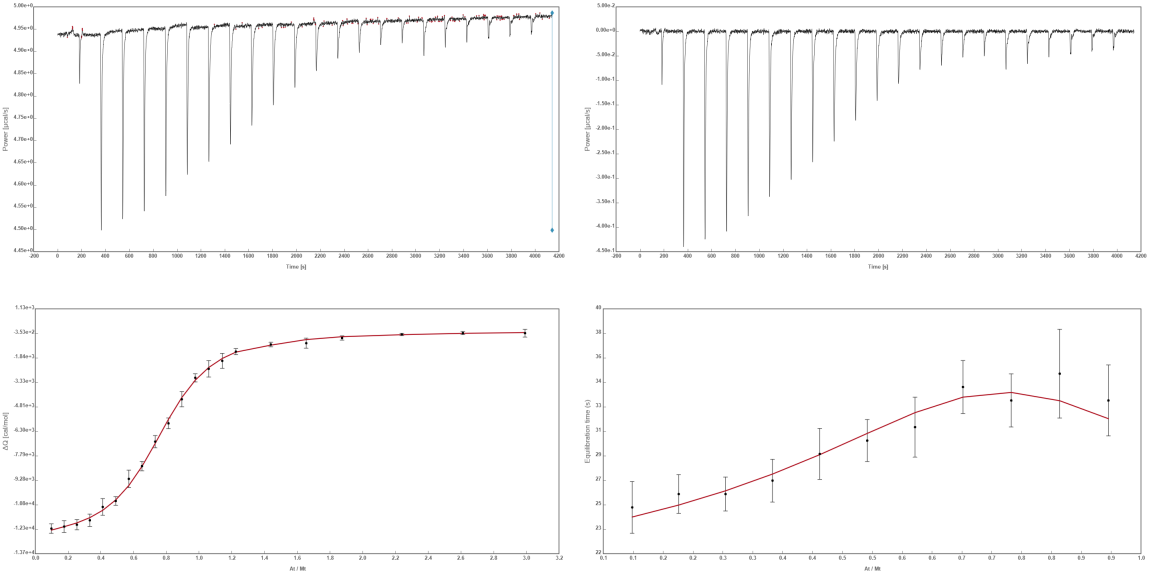
Adjustment with global parameters
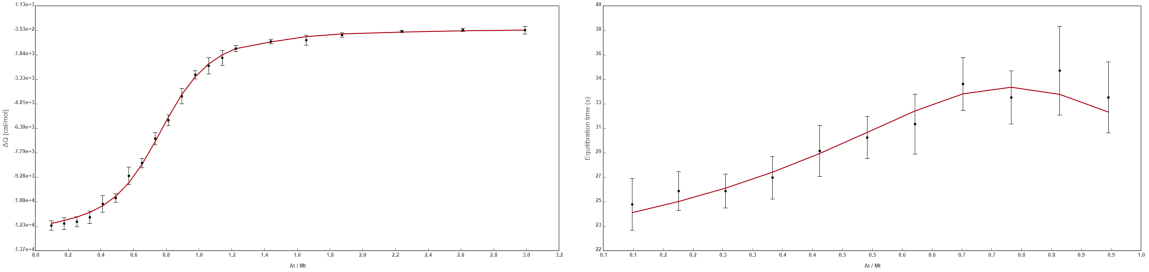
Compound 37/43

Raw data analysis
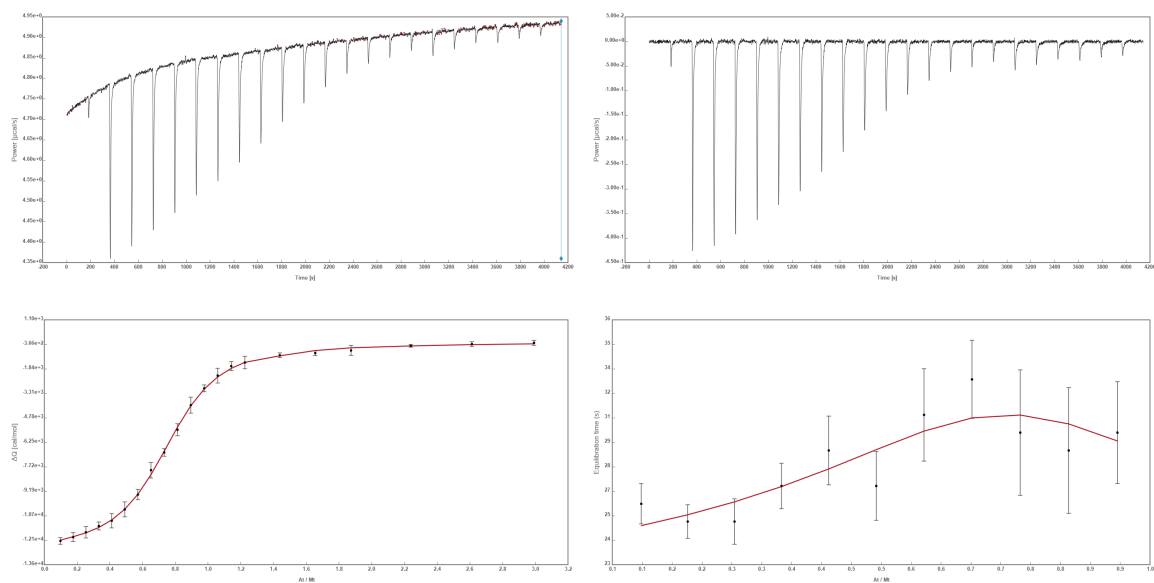
Adjustment with global parameters
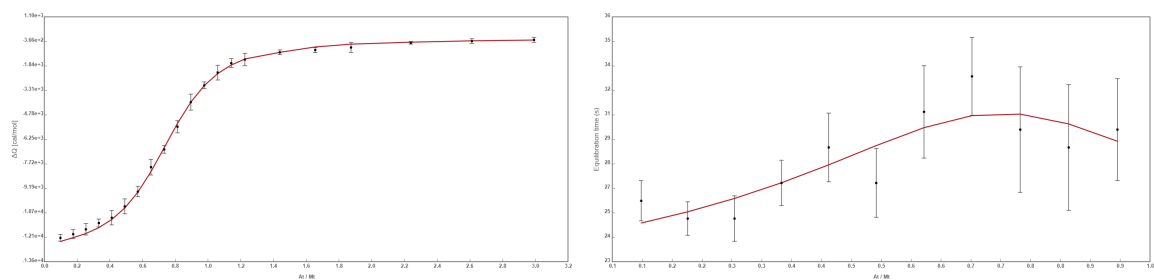
Dilution titration of **37** in **43**
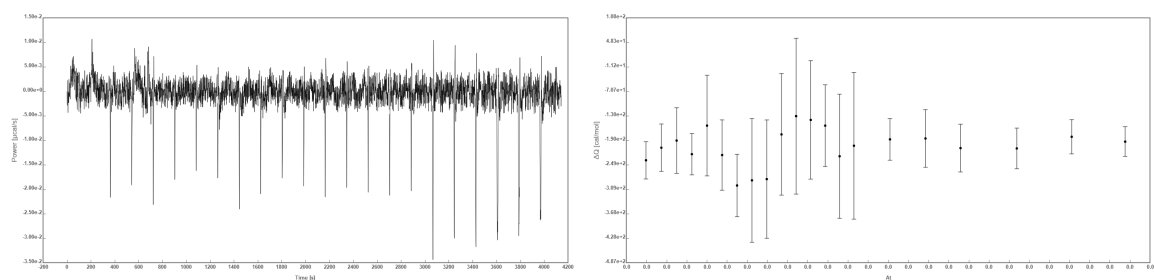
References

1. Klebe, G., Recent developments in structure-based drug design, *J. Mol. Med.* **2000**, *78*, 269–281.

2. Smith, J. B., Willis, A. L., Aspirin Selectively Inhibits Prostaglandin Production in Human Platelets, *Nat. New Biol.* **1971**, *231*, 235–237.

3. Blobaum, A. L., Marnett, L. J., Structural and Functional Basis of Cyclooxygenase Inhibition, *J. Med. Chem.* **2007**, *50*, 1425–1441.

4. Lounnas, V., Ritschel, T., Kelder, J., McGuire, R., Bywater, R. P., Foloppe, N., Current progress in structure-based rational drug design marks a new mindset in drug discovery, *Comput. Struct. Biotechnol. J.* **2013**, *5*, 1–14.

5. Van Montfort, R. L. M., Workman, P., Structure-based drug design: aiming for a perfect fit, *Essays Biochem.* **2017**, *61*, 431–437.

6. Fisher, S. Z., Maupin, C. M., Budayova-Spano, M., Govindasamy, L., Tu, C., Agbandje-McKenna, M., Silverman, D. N., Voth, G. A., McKenna, R., Atomic Crystal and Molecular Dynamics Simulation Structures of Human Carbonic Anhydrase II: Insights into the Proton Transfer Mechanism, *Biochemistry* **2007**, *46*, 2930–2937.

7. Boland, A., Chang, L., Barford, D., The potential of cryo-electron microscopy for structure-based drug design, *Essays Biochem.* **2017**, *61*, 543–560.

8. Nannenga, B. L., Gonen, T., The cryo-EM method microcrystal electron diffraction, *Nat. Methods* **2019**, *16*, 369–379.

9. Gruene, T., Wennmacher, J. T. C., Zaubitzer, C., Holstein, J. J., Heidler, J., Fecteau-Lefebvre, A., De Carlo, S., Müller, E., Goldie, K. N., Regeni, I., Li, T., Santiso-Quinones, G., Steinfeld, G., Handschin, S., Genderen, E. van, Bokhoven, J. A. van, Clever, G. H., Pantelic, R., Rapid Structure Determination of Microcrystalline Molecular Compounds Using Electron Diffraction, *Angew. Chem. Int. Ed.* **2018**, *57*, 16313–16317.

10. Wiedorn, M. O., Oberthür, D., Bean, R., Schubert, R., Werner, N., Abbey, B., Aepfelbacher, M., Adriano, L., Allahgholi, A., Al-Qudami, N., Andreasson, J., Aplin, S., Awel, S., Ayyer, K., Bajt, S., Barák, I., Bari, S., Bielecki, J., Botha, S., Boukhelef, D., Brehm, W., Brockhauser, S., Cheviakov, I., Coleman, M. A., Cruz-Mazo, F., Danilevski, C., Darmanin, C., Doak, R. B., Domaracky, M., Dörner, K., Du, Y., Fangohr, H., Fleckenstein, H., Frank, M., Fromme, P., Gañán-Calvo, A. M., Gevorkov, Y., Giewekemeyer, K., Ginn, H. M., Graafsma, H., Graceffa, R., Greiffenberg, D., Gumprecht, L., Göttlicher, P., Hajdu, J., Hauf, S., Heymann, M., Holmes, S., Horke, D. A., Hunter, M. S., Imlau, S., Kaukher, A., Kim, Y., Klyuev, A., Knoška, J., Kobe, B., Kuhn, M., Kupitz, C., Küpper, J., Lahey-Rudolph, J. M.,

Laurus, T., Le Cong, K., Letrun, R., Xavier, P. L., Maia, L., Maia, F. R., Mariani, V., Messerschmidt, M., Metz, M., Mezza, D., Michelat, T., Mills, G., Monteiro, D. C., Morgan, A., Mühlig, K., Munke, A., Münnich, A., Nette, J., Nugent, K. A., Nuguid, T., Orville, A. M., Pandey, S., Pena, G., Villanueva-Perez, P., Poehlsen, J., Previtali, G., Redecke, L., Riekehr, W. M., Rohde, H., Round, A., Safenreiter, T., Sarrou, I., Sato, T., Schmidt, M., Schmitt, B., Schönherr, R., Schulz, J., Sellberg, J. A., Seibert, M. M., Seuring, C., Shelby, M. L., Shoeman, R. L., Sikorski, M., Silenzi, A., Stan, C. A., Shi, X., Stern, S., Sztuk-Dambietz, J., Szuba, J., Tolstikova, A., Trebbin, M., Trunk, U., Vagovic, P., Ve, T., Weinhausen, B., White, T. A., Wrona, K., Xu, C., Yefanov, O., Zatsepin, N., Zhang, J., Perbandt, M., Mancuso, A. P., Betzel, C., Chapman, H., Barty, A., Megahertz serial crystallography, *Nat. Commun.* **2018**, *9*, 1–11.

11. Shuker, S. B., Hajduk, P. J., Meadows, R. P., Fesik, S. W., Discovering high-affinity ligands for proteins: SAR by NMR, *Science* **1996**, *274*, 1531–1534.

12. Carneiro, M. G., Ab, E., Theisgen, S., Siegal, G., NMR in structure-based drug design, *Essays Biochem.* **2017**, *61*, 485–493.

13. Hansch, C., McClarin, J., Klein, T., Langridge, R., A Quantitative Structure-Activity Relationship and Molecular Graphics Study of Carbonic Anhydrase Inhibitors, *Mol. Pharmacol.* **1985**, *27*, 493–498.

14. Patel, H. M., Noolvi, M. N., Sharma, P., Jaiswal, V., Bansal, S., Lohan, S., Kumar, S. S., Abbot, V., Dhiman, S., Bhardwaj, V., Quantitative structure-activity relationship (QSAR) studies as strategic approach in drug discovery, *Med. Chem. Res.* **2014**, *23*, 4991–5007.

15. 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.

16. Klebe, G., The use of thermodynamic and kinetic data in drug discovery: Decisive insight or increasing the puzzlement?, *ChemMedChem* **2015**, *10*, 229–231.

17. Krimmer, S. G., Betz, M., Heine, A., Klebe, G., Methyl, Ethyl, Propyl, Butyl: Futile But Not for Water, as the Correlation of Structure and Thermodynamic Signature Shows in a Congeneric Series of Thermolysin Inhibitors, *ChemMedChem* **2014**, *9*, 833–846.

18. Krimmer, S. G., Cramer, J., Betz, M., Fridh, V., Karlsson, R., Heine, A., Klebe, G., Rational Design of Thermodynamic and Kinetic Binding Profiles by Optimizing Surface Water Networks Coating Protein-Bound Ligands, *J. Med. Chem.* **2016**, *59*, 10530–10548.

19. Cramer, J., Krimmer, S. G., Heine, A., Klebe, G., Paying the Price of Desolvation in Solvent-Exposed Protein Pockets: Impact of Distal Solubilizing Groups on Affinity and Binding Thermodynamics in a Series of Thermolysin Inhibitors, *J. Med. Chem.* **2017**, *60*, 5791–5799.

20. 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.

21. 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.

22. Holdgate, G. A., Ward, W. H. J., Measurements of binding thermodynamics in drug discovery, *Drug Discovery Today* **2005**, *10*, 1543–1550.

23. Copeland, R. A., Pompliano, D. L., Meek, T. D., Drug-target residence time and its implications for lead optimization, *Nat. Rev. Drug Discov.* **2006**, *5*, 730–739.

24. Ladbury, J. E., Klebe, G., Freire, E., Adding calorimetric data to decision making in lead discovery: a hot tip, *Nat. Rev. Drug Discov.* **2010**, *9*, 23–27.

25. Freire, E., Do enthalpy and entropy distinguish first in class from best in class?, *Drug Discovery Today* **2008**, *13*, 869–874.

26. Gerlach, C., Smolinski, M., Steuber, H., Sotriffer, C. A., Heine, A., Hangauer, D. G., Klebe, G., Thermodynamic Inhibition Profile of a Cyclopentyl and a Cyclohexyl Derivative towards Thrombin: The Same but for Different Reasons, *Angew. Chem. Int. Ed.* **2007**, *46*, 8511–8514.

27. Lafont, V., Armstrong, A. A., Ohtaka, H., Kiso, Y., Mario Amzel, L., Freire, E., Compensating enthalpic and entropic changes hinder binding affinity optimization, *Chem. Biol. Drug Des.* **2007**, *69*, 413–422.

28. Brandt, T., Holzmann, N., Muley, L., Khayat, M., Wegscheid-Gerlach, C., Baum, B., Heine, A., Hangauer, D., Klebe, G., Congeneric but still distinct: How closely related trypsin ligands exhibit different thermodynamic and structural properties, *J. Mol. Biol.* **2011**, *405*, 1170–1187.

29. 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, *Chem. Biol.* **1996**, *3*, 973–980.

30. Baker, B. M., Murphy, K. P., Evaluation of Linked Protonation Effects in Protein Binding Reactions Using Isothermal Titration Calorimetry, *Biophys. J.* **1996**, *71*, 2049–2055.

31. Dullweber, F., Stubbs, M. T., Musil, Đ., Stürzebecher, J., Klebe, G., Factorising Ligand Affinity: A Combined Thermodynamic and Crystallographic Study of Trypsin and Thrombin Inhibition, *J. Mol. Biol.* **2001**, *313*, 593–614.

32. Steuber, H., Czodrowski, P., Sotriffer, C. A., Klebe, G., Tracing Changes in Protonation: A Prerequisite to Factorize Thermodynamic Data of Inhibitor Binding to Aldose Reductase, *J. Mol. Biol.* **2007**, *373*, 1305–1320.

33. Biela, A., Nasief, N. N., Betz, M., Heine, A., Hangauer, D., Klebe, G., Dissecting the hydrophobic effect on the molecular level: The role of water, enthalpy, and entropy in ligand binding to thermolysin, *Angew. Chem. Int. Ed.* **2013**, *52*, 1822–1828.

34. Neeb, M., Czodrowski, P., Heine, A., Barandun, L. J., Hohn, C., Diederich, F., Klebe, G., Chasing protons: How isothermal titration calorimetry, mutagenesis, and pK_a calculations trace the locus of charge in ligand binding to a tRNA-binding enzyme, *J. Med. Chem.* **2014**, *57*, 5554–5565.

35. Rühmann, E. H., Rupp, M., Betz, M., Heine, A., Klebe, G., Boosting Affinity by Correct Ligand Preorganization for the S2 Pocket of Thrombin: A Study by Isothermal Titration Calorimetry, Molecular Dynamics, and High-Resolution Crystal Structures, *ChemMedChem* **2016**, *11*, 309–319.

36. Sandner, A., Hüfner-Wulsdorf, T., Heine, A., Steinmetzer, T., Klebe, G., Strategies for Late-Stage Optimization: Profiling Thermodynamics by Preorganization and Salt Bridge Shielding, *J. Med. Chem.* **2019**, *62*, 9753–9771.

37. Wienen-Schmidt, B., Jonker, H. R., Wulsdorf, T., Gerber, H.-D., Saxena, K., Kudlinzki, D., Sreeramulu, S., Parigi, G., Luchinat, C., Heine, A., Schwalbe, H., Klebe, G., Paradoxically, Most Flexible Ligand Binds Most Entropy-Favored: Intriguing Impact of Ligand Flexibility and Solvation on Drug-Kinase Binding, *J. Med. Chem.* **2018**, *61*, 5922–5933.

38. Ohtaka, H., Freire, E., Adaptive inhibitors of the HIV-1 protease, *Prog. Biophys. Mol. Biol.* **2005**, *88*, 193–208.

39. Velazquez-Campoy, A., Kiso, Y., Freire, E., The binding energetics of first- and second-generation HIV-1 protease inhibitors: Implications for drug design, *Arch. Biochem. Biophys.* **2001**, *390*, 169–175.

40. Chodera, J. D., Mobley, D. L., Entropy-enthalpy Compensation: Role and Ramifications in Biomolecular Ligand Recognition and Design, *Annu. Rev. Biophys.* **2013**, *42*, 121–142.

41. Geschwindner, S., Ulander, J., Johansson, P., Ligand Binding Thermodynamics in Drug Discovery: Still a Hot Tip?, *J. Med. Chem.* **2015**, *58*, 6321–6335.

42. Bissantz, C., Kuhn, B., Stahl, M., A medicinal chemist's guide to molecular interactions, *J. Med. Chem.* **2010**, *53*, 5061–5084.

43. Swinney, D. C., Biochemical mechanisms of drug action: What does it take for success?, *Nat. Rev. Drug Discov.* **2004**, *3*, 801–808.

44. Copeland, R. A., The drug-target residence time model: a 10-year retrospective, *Nat. Rev. Drug Discov.* **2015**, *15*, 1–9.

45. Gooljarsingh, L. T., Fernandes, C., Yan, K., Zhang, H., Grooms, M., Johanson, K., Sinnamon, R. H., Kirkpatrick, R. B., Kerrigan, J., Lewis, T., Arnone, M., King, A. J., Lai, Z., Copeland, R. A., Tummino, P. J., A biochemical rationale for the anticancer effects of Hsp90 inhibitors: Slow, tight binding inhibition by geldanamycin and its analogues, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 7625–7630.

46. Koshland Jr., D. E., Application of a Theory of Enzyme Specificity to Protein Synthesis, *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 98–104.

47. Núñez, S., Venhorst, J., Kruse, C. G., Target-drug interactions: first principles and their application to drug discovery., *Drug Discovery Today* **2012**, *17*, 10–22.

48. Cramer, J., Krimmer, S. G., Fridh, V., Wulsdorf, T., Karlsson, R., Heine, A., Klebe, G., Elucidating the origin of long residence time binding for inhibitors of the metalloprotease thermolysin, *ACS Chem. Biol.* **2017**, *12*, 225–233.

49. Wood, E. R., Truesdale, A. T., McDonald, O. B., Yuan, D., Hassell, A., Dickerson, S. H., Ellis, B., Pennisi, C., Horne, E., Lackey, K., Alligood, K. J., Rusnak, D. W., Gilmer, T. M., Shewchuk, L., A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): Relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells, *Cancer Res.* **2004**, *64*, 6652–6659.

50. Copeland, R. A., Williams, J. M., Giannaras, J., Nurnberg, S., Covington, M., Pinto, D., Pick, S., Trzaskos, J. M., Mechanism of selective inhibition of the inducible isoform of prostaglandin G/H synthase, *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11202–11206.

51. Maschera, B., Darby, G., Palu, G., Wright, L. L., Tisdale, M., Myers, R., Blair, E. D., Furfine, E. S., Human Immunodeficiency Virus Mutations in the viral Protease that confer Resistance to Saquinavir increase the Dissociation Rate Constant of the Protease-Saquinavir Complex, *J. Biol. Chem.* **1996**, *271*, 33231–33235.

52. Shuman, C. F., Markgren, P. O., Hämäläinen, M., Danielson, U. H., Elucidation of HIV-1 protease resistance by characterization of interaction kinetics between inhibitors and enzyme variants, *Antiviral Res.* **2003**, *58*, 235–242.

53. Vauquelin, G., Effects of target binding kinetics on in vivo drug efficacy: koff, kon and rebinding, *Br. J. Pharmacol.* **2016**, *173*, 2319–2334.

54. Sykes, D. A., Moore, H., Stott, L., Holliday, N., Javitch, J. A., Robert Lane, J., Charlton, S. J., Extrapyramidal side effects of antipsychotics are linked to their association kinetics at dopamine D_2 receptors, *Nat. Commun.* **2017**, *8*, 1–11.

55. Schuetz, D. A., Richter, L., Amaral, M., Grandits, M., Grädler, U., Musil, D., Buchstaller, H.-P., Eggenweiler, H.-M., Frech, M., Ecker, G. F., Ligand Desolvation Steers On-Rate and Impacts Drug Residence Time of Heat Shock Protein 90 (Hsp90) Inhibitors, *J. Med. Chem.* **2018**, *61*, 4397–4411.

56. Tanford, C., The Hydrophobic Effect and the Organization of Living Matter, *Science* **1978**, *200*, 1012–1018.

57. Biela, A., Sielaff, F., Terwesten, F., Heine, A., Steinmetzer, T., Klebe, G., Ligand binding stepwise disrupts water network in thrombin: Enthalpic and entropic changes reveal classical hydrophobic effect, *J. Med. Chem.* **2012**, *55*, 6094–6110.

58. Galamba, N., Water's structure around hydrophobic solutes and the iceberg model, *J. Phys. Chem. B* **2013**, *117*, 2153–2159.

59. 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?, *Eur. Phys. J. Spec. Top.* **2013**, *223*, 853–891.

60. Frank, H. S., Evans, M. W., Free volume and entropy in condensed systems III. Entropy in binary liquid mixtures; Partial molal entropy in dilute solutions; Structure and thermo-dynamics in aqueous electrolytes, *J. Chem. Phys.* **1945**, *13*, 507–532.

61. Tanford, C., Interfacial free energy and the hydrophobic effect, *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 4175–4176.

62. Biedermann, F., Nau, W. M., Schneider, H.-J., The Hydrophobic Effect Revisited - Studies with Supramolecular Complexes Imply High-Energy Water as a Noncovalent Driving Force, *Angew. Chem. Int. Ed.* **2014**, *53*, 11158–11171.

63. Biedermann, F., Uzunova, V. D., Scherman, O. A., Nau, W. M., De Simone, A., Release of high-energy water as an essential driving force for the high-affinity binding of cucurbit[n]urils, *J. Am. Chem. Soc.* **2012**, *134*, 15318–15323.

64. Hillyer, M. B., Gibb, B. C., Moleular Shape and the Hydrophobic Effect, *Annu. Rev. Phys. Chem.* **2016**, *67*, 307–329.

65. Snyder, P. W., Mecinovic, J., Moustakas, D. T., Thomas, S. W., Harder, M., Mack, E. T., Lockett, M. R., Heroux, A., Sherman, W., Whitesides, G. M., Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase, *Proc. Natl. Acad. Sci.* **2011**, *108*, 17889–17894.

66. Darby, J. F., Hopkins, A. P., Shimizu, S., Roberts, S. M., Brannigan, J. A., Turkenburg, J. P., Thomas, G. H., Hubbard, R. E., Fischer, M., Water networks can determine the affinity of ligand binding to proteins, *J. Am. Chem. Soc.* **2019**, jacs.9b06275.

67. Dunitz, J. D., Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions, *Chem. Biol.* **1995**, *2*, 709–712.

68. Krishnamurthy, V. M., Bohall, B. R., Semetey, V., Whitesides, G. M., The paradoxical thermodynamic basis for the interaction of ethylene glycol, glycine, and sarcosine chains with bovine carbonic anhydrase II: An unexpected manifestation of enthalpy/entropy compensation, *J. Am. Chem. Soc.* **2006**, *128*, 5802–5812.

69. Olsson, T. S. G., Ladbury, J. E., Pitt, W. R., Williams, M. A., Extent of enthalpy-entropy compensation in protein-ligand interactions, *Protein Sci.* **2011**, *20*, 1607–1618.

70. Boyce, S. E., Tellinghuisen, J., Chodera, J. D., Avoiding accuracy-limiting pitfalls in the study of protein-ligand interactions with isothermal titration calorimetry, *bioRxiv* **2015**, 1–7.

71. Fischer, E., Einfluss der Configuration auf die Wirkung der Enzyme, *Ber. Dtsch. Chem. Ges.* **1894**, *27*, 2985–2993.

72. Ferreira De Freitas, R., Schapira, M., A systematic analysis of atomic protein-ligand interactions in the PDB, *Med. Chem. Commun.* **2017**, *8*, 1970–1981.

73. Chartier, M., Morency, L.-P., Zylber, M. I., Najmanovich, R. J., Large-scale detection of drug off-targets: Hypotheses for drug repurposing and understanding side-effects, *BMC Pharmacol. Toxicol.* **2017**, *18*, 1–16.

74. Huggins, M. L., 50 Jahre Theorie der Wasserstoffbrückenbindung, *Angew. Chem.* **1971**, *83*, 163–168.

75. Davis, A. M., Teague, S. J., Hydrogen Bonding, Hydrophobic Interactions, and Failure of the Rigid Receptor Hypothesis, *Angew. Chem. Int. Ed.* **1999**, *38*, 736–749.

76. O'Hagan, D., Understanding organofluorine chemistry. An introduction to the C-F bond, *Chem. Soc. Rev.* **2008**, *37*, 308–319.

77. Riess, J. G., Understanding the Fundamentals of Perfluorocarbons and Perfluorocarbon Emulsions Relevant to *In Vivo* Oxygen Delivery, *Artif. Cells. Blood Substit. Immobil. Biotechnol.* **2005**, *33*, 47–63.

78. Dalvi, V. H., Rossky, P. J., Molecular origins of fluorocarbon hydrophobicity, *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 13603–13607.

79. Bond, A. D., *In situ* co-crystallisation as a tool for low-temperature crystal engineering, *Chem. Commun.* **2003**, 250–251.

80. O'Neal, K. L., Geib, S., Weber, S. G., Extraction of Pyridines into Fluorous Solvents Based on Hydrogen Bond Complex Formation with Carboxylic Acid Receptors, *Anal. Chem.* **2007**, *79*, 3117–3125.

81. Champagne, P. A., Desroches, J., Hamel, J.-D., Vandamme, M., Paquin, J.-F., Monofluorination of Organic Compounds: 10 Years of Innovation, *Chem. Rev.* **2015**, *115*, 9073–9174.

82. Meanwell, N. A., Fluorine and Fluorinated Motifs in the Design and Application of Bioisosteres for Drug Design, *J. Med. Chem.* **2018**, *61*, 5822–5880.

83. Yudin, A. K., Macrocycles: lessons from the distant past, recent developments, and future directions, *Chem. Sci.* **2015**, *6*, 30–49.

84. Park, B. K., Kitteringham, N. R., O'Neill, P. M., Metabolism of Fluorine-Containing Drugs, *Annu. Rev. Pharmacol. Toxicol.* **2001**, *41*, 443–470.

85. Skora, L., Jahnke, W., ^{19}F-NMR-Based Dual-Site Reporter Assay for the Discovery and Distinction of Catalytic and Allosteric Kinase Inhibitors, *ACS Med. Chem. Lett.* **2017**, *8*, 632–635.

86. Fried, J., Sabo, E. F., 9α-Fluoro Derivatives of Cortisone and Hydrocortisone, *J. Am. Chem. Soc.* **1954**, *76*, 1455–1456.

87. Street, I. P., Armstrong, C. R., Withers, S. G., Hydrogen Bonding and Specificity: Fluorodeoxy Sugars as Probes of Hydrogen Bonding in the Glycogen Phosphorylase-Glucose Complex, *Biochemistry* **1986**, *25*, 6021–6027.

88. Holmgren, S. K., Taylor, K. M., Bretscher, L. E., Raines, R. T., Code for collagen's stability deciphered, *Nature* **1998**, *392*, 666–667.

89. Biffinger, J. C., Kim, H. W., Dimagno, S. G., The Polar Hydrophobicity of Fluorinated Compounds, *ChemBioChem* **2004**, *5*, 622–627.

90. Qiu, J., Stevenson, S. H., O'Beirne, M. J., Silverman, R. B., 2,6-Difluorophenol as a Bioisostere of a Carboxylic Acid: Bioisosteric Analogues of γ-Aminobutyric acid, *J. Med. Chem.* **1999**, *42*, 329–332.

91. Nicolaou, I., Zika, C., Demopoulos, V. J., [1-(3,5-Difluoro-4-hydroxyphenyl)-1*H*-pyrrol-3-yl]phenylmethanone as a Bioisostere of a Carboxylic Acid Aldose Reductase Inhibitor, *J. Med. Chem.* **2004**, *47*, 2706–2709.

92. Alexiou, P., Demopoulos, V. J., A Diverse Series of Substituted Benezensulfonamides as Aldose Reductase Inhibitors with Antioxidant Activity: Design, Synthesis, and in Vitro Activity, *J. Med. Chem.* **2010**, *53*, 7756–7766.

93. Kotsampasakou, E., Demopoulos, V. J., Synthesis of derivatives of the keto-pyrrolyl-difluorophenol scaffold: Some structural aspects for aldose reductase inhibitory activity and selectivity, *Bioorganic Med. Chem.* **2013**, *21*, 869–873.

94. Braun, F., Bertoletti, N., Möller, G., Adamski, J., Frotscher, M., Guragossian, N., Madeira Gírio, P. A., Le Borgne, M., Ettouati, L., Falson, P., Müller, S., Vollmer, G., Heine, A., Klebe, G., Marchais-Oberwinkler, S., Structure-Based Design and Profiling of Novel 17β-HSD14 Inhibitors, *Eur. J. Med. Chem.* **2018**, *155*, 61–76.

95. Müller, K., Faeh, C., Diederich, F., Fluorine in Pharmaceuticals: Looking Beyond Intuition, *Science* **2007**, *317*, 1881–1886.

96. Kim, C.-Y., Chang, J. S., Doyon, J. B., Baird Jr., T. T., Fierke, C. A., Jain, A., Christianson, D. W., Contribution of Fluorine to Protein - Ligand Affinity in the Binding of Fluoroaromatic Inhibitors to Carbonic Anhydrase II, *J. Am. Chem. Soc.* **2000**, *122*, 12125–12134.

97. Giroud, M., Harder, M., Kuhn, B., Haap, W., Trapp, N., Schweizer, W. B., Schirmeister, T., Diederich, F., Fluorine Scan of Inhibitors of the Cysteine Protease Human Cathepsin L: Dipolar and Quadrupolar Effects in the π-Stacking of Fluorinated Phenyl Rings on Peptide Amide Bonds, *ChemMedChem* **2016**, *11*, 1042–1047.

98. Schönemann, W., Cramer, J., Mühlethaler, T., Fiege, B., Silbermann, M., Rabbani, S., Dätwyler, P., Zihlmann, P., Jakob, R. P., Sager, C. P., Smieško, M., Schwardt, O., Maier, T., Ernst, B., Improvement of Aglycone π-Stacking Yields Nanomolar to Sub-nanomolar FimH Antagonists, *ChemMedChem* **2019**, *14*, 749–757.

99. Scott, A. D., Phillips, C., Alex, A., Flocco, M., Bent, A., Randall, A., O'Brien, R., Damian, L., Jones, L. H., Thermodynamic Optimisation in Drug Discovery: A Case Study using Carbonic Anhydrase Inhibitors, *ChemMedChem* **2009**, *4*, 1985–1989.

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

101. Dudutienė, V., Zubrienė, A., Smirnov, A., Timm, D. D., Smirnovienė, J., Kazokaitė, J., Michailovienė, V., Zakšauskas, A., Manakova, E., Gražulis, S., Matulis, D., Functionalization of fluorinated benzenesulfonamides and their inhibitory properties toward carbonic anhydrases, *ChemMedChem* **2015**, *10*, 662–687.

102. Olsen, J. A., Banner, D. W., Seiler, P., Sander, U. O., D'Arcy, A., Stihle, M., Müller, K., Diederich, F., A fluorine scan of thrombin inhibitors to map the fluorophilicity/fluorophobicity of an enzyme active site: Evidence for C-F···C=O interactions, *Angew. Chem. Int. Ed.* **2003**, *42*, 2507–2511.

103. Bürgi, H. B., Dunitz, J. D., Shefter, E., Geometrical Reaction Coordinates. II. Nucleophilic Addition to a Carbonyl Group, *J. Am. Chem. Soc.* **1973**, *95*, 5065–5067.

104. Bürgi, H. B., Dunitz, J. D., Lehn, J. M., Wipff, G., Stereochemistry of reaction paths at carbonyl centres, *Tetrahedron* **1974**, *30*, 1563–1572.

105. Liu, S.-Y., Michael, D. W., Dykstra, C. E., Lisy, J. M., The stabilities of the hydrogen fluoride trimer and tetramer, *J. Chem. Phys.* **1986**, *84*, 5032–5036.

106. Wilcken, R., Zimmermann, M. O., Lange, A., Joerger, A. C., Boeckler, F. M., Principles and applications of halogen bonding in medicinal chemistry and chemical biology, *J. Med. Chem.* **2013**, *56*, 1363–1388.

107. Malham, R., Johnstone, S., Bingham, R. J., Barratt, E., Phillips, S. E., Laughton, C. A., Homans, S. W., Strong solute-solute dispersive interactions in a protein-ligand complex, *J. Am. Chem. Soc.* **2005**, *127*, 17061–17067.

108. Myslinski, J. M., DeLorbe, J. E., Clements, J. H., Martin, S. F., Protein-ligand interactions: Thermodynamic effects associated with increasing nonpolar surface area, *J. Am. Chem. Soc.* **2011**, *133*, 18518–18521.

109. Myslinski, J. M., Clements, J. H., Delorbe, J. E., Martin, S. F., Protein-ligand interactions: Thermodynamic effects associated with increasing the length of an alkyl chain, *ACS Med. Chem. Lett.* **2013**, *4*, 1048–1053.

110. Jain, A., Huang, S. G., Whitesides, G. M., Lack of Effect of the Length of Oligoglycine- and Oligo(ethylene glycol)-Derived para-Substituents on the Affinity of Benzenesulfonamides for Carbonic Anhydrase II in Solution, *J. Am. Chem. Soc.* **1994**, *116*, 5057–5062.

111. Mecinović, J., Snyder, P. W., Mirica, K. A., Bai, S., Mack, E. T., Kwant, R. L., Moustakas, D. T., Héroux, A., Whitesides, G. M., Fluoroalkyl and Alkyl Chains Have Similar Hydrophobicities in Binding to the "Hydrophobic Wall" of Carbonic Anhydrase, *J. Am. Chem. Soc.* **2011**, *133*, 14017–14026.

112. Wiseman, T., Williston, S., Brandts, J. F., Lin, L. N., Rapid measurement of binding constants and heats of binding using a new titration calorimeter., *Anal. Biochem.* **1989**, *179*, 131–137.

113. Muñoz, E., Piñeiro, A., AFFINImeter Software: from its Beginnings to Future Trends - A Literature review, *J. Appl. Bioanal.* **2018**, *4*, 124–139.

114. Piñeiro, Á., Muñoz, E., Sabín, J., Costas, M., Bastos, M., Velázquez-Campoy, A., Garrido, P. F., Dumas, P., Ennifar, E., García-Río, L., Rial, J., Pérez, D., Fraga, P., Rodríguez, A., Cotelo, C., AFFINImeter: A software to analyze molecular recognition processes from experimental data, *Anal. Biochem.* **2019**, *577*, 117–134.

115. Dumas, P., Ennifar, E., Bec, G., Piñero, A., Sabín, J., Muñoz, E., Rial, J., Implementation of kinITC into AFFINImeter, tech. rep., **2015**.

116. Wlodawer, A., Minor, W., Dauter, Z., Jaskolski, M., Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures, *FEBS J.* **2008**, *275*, 1–21.

117. Taylor, G. L., Introduction to phasing, *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **2010**, *66*, 325–338.

118. Rhodes, G., *Crystallography Made Crystal Clear*, 3rd ed., Academic Press, Asterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney, Tokyo, **2006**.

119. Wlodawer, A., Minor, W., Dauter, Z., Jaskolski, M., Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination, *FEBS J.* **2013**, *280*, 5705–5736.

120. Mueller, U., Förster, R., Hellmig, M., Huschmann, F. U., Kastner, A., Malecki, P., Pühringer, S., Röwer, M., Sparta, K., Steffien, M., Ühlein, M., Wilk, P., Weiss, M. S., The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current status and perspectives, *Eur. Phys. J. Plus* **2015**, *130*, 141–150.

121. Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K., Features and development of Coot., *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **2010**, *66*, 486–501.

122. Massa, W., *Kristallstrukturbestimmung*, 8th ed., (Eds.: Heck, J., König, B., Winter, R.), Springer Fachmedien, Wiesbaden, **2015**.

123. Parois, P., Cooper, R. I., Thompson, A. L., Crystal structures of increasingly large molecules: Meeting the challenges with CRYSTALS software, *Chem. Cent. J.* **2015**, *9*, 1–14.

124. Mann, T., Keilin, D., Carbonic Anhydrase, *Nature* **1939**, *144*, 442–443.

125. Mann, T., Keilin, D., Sulphanilamide as a Specific Inhibitor of Carbonic Anhydrase, *Nature* **1940**, *146*, 164–165.

126. Supuran, C. T., Carbonic anhydrase inhibitors and their potential in a range of therapeutic areas, *Expert Opin. Ther. Pat.* **2018**, *28*, 709–712.

127. Alterio, V., Di Fiore, A., D'Ambrosio, K., Supuran, C. T., De Simone, G., Multiple Binding Modes of Inhibitors to Carbonic Anhydrases: How to Design Specific Drugs Targeting 15 Different Isoforms, *Chem. Rev.* **2012**, *112*, 4421–4468.

128. Supuran, C. T., Carbonic anhydrases: novel therapeutic applications for inhibitors and activators, *Nat. Rev. Drug. Discov.* **2008**, *7*, 168–181.

129. 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.

130. Krebs, H. A., Sykes, W. O., Bartley, W. C., Acetylation and Deacetylation of the *p*-Amino Group of Sulphonamide Drugs in Animal Tissues, *Biochem. J.* **1947**, *41*, 622–630.

131. Krebs, H. A., Inhibition of Carbonic Anhydrase by Sulphonamides, *Biochem. J.* **1948**, *43*, 525–528.

132. Kumar, K., King, R. W., Carey, P. R., Resonance Raman Studies on Some Carbonic Anhydrase-Aromatic Sulfonamide Complexes, *Biochemistry* **1976**, *15*, 2195–2202.

133. Wolfenden, R., Analog Approaches to the Structure of the Transition State in Enzyme Reactions, *Acc. Chem. Res.* **1972**, *5*, 10–18.

134. Wolfenden, R., Transition State Analogues for Enzyme Catalysis, *Nature* **1969**, *223*, 704–705.

135. Lienhard, G. E., Enzymatic Catalysis and Transition-State Theory, *Science* **1973**, *180*, 149–154.

136. Schramm, V. L., Enzymatic Transition States, Transition-State Analogs, Dynamics, Thermodynamics, and Lifetimes, *Annu. Rev. Biochem.* **2011**, *80*, 703–732.

137. King, R. W., Burgen, A. S. V., Sulphonamide Complexes of Human Carbonic Anhydrases Ultraviolett Difference Spectroscopy, *Biochim. Biophys. Acta* **1970**, *207*, 278–285.

138. Kanamori, K., Roberts, J. D., Nitrogen-15 Nuclear Magnetic Resonance Study of Benzenesulfonamide and Cyanate Binding to Carbonic Anhydrase, *Biochemistry* **1983**, *22*, 2658–2664.

139. Sjöblom, B., Polentarutti, M., Djinović-Carugo, K., Structural study of X-ray induced activation of carbonic anhydrase., *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 10609–10613.

140. Glöckner, S., Khang, N., Sager, C. P., Hüfner, T., Heine, A., Klebe, G., Conformational changes in alkyl chains determine the thermodynamic and kinetic binding profiles of Carbonic Anhydrase inhibitors and reveal effects of improved shape complementarity, *ACS Chem. Biol.*, *submitted*.

141. Schmid, M., Nogueira, E. S., Monnard, F. W., Ward, T. R., Meuwly, M., Arylsulfonamides as inhibitors for cabonic anhydrase: prediction & validation, *Chem. Sci.* **2012**, *3*, 690–700.

142. Grzybowski, B. A., Ishchenko, A. V., Kim, C.-Y., Topalov, G., Chapman, R., Christianson, D. W., Whitesides, G. M., Shakhnovich, E. I., Combinatorial computational method gives new picomolar ligands for a known enzyme, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1270–1273.

143. Supuran, C. T., Scozzafava, A., Casini, A., Carbonic anhydrase inhibitors, *Med. Res. Rev.* **2003**, *23*, 146–189.

144. Krishnamurthy, V. M., Bohall, B. R., Kim, C.-Y., Moustakas, D. T., Christianson, D. W., Whitesides, G. M., Thermodynamic Parameters for the Association of Fluorinated Benzenesulfonamides with Bovine Carbonic Anhydrase II, *Chem. Asian J.* **2007**, *2*, 94–105.

145. Mecinović, J., Snyder, P. W., Mirica, K. A., Bai, S., Mack, E. T., Kwant, R. L., Moustakas, D. T., Héroux, A., Whitesides, G. M., Fluoroalkyl and alkyl chains have similar hydrophobicities in binding to the "hydrophobic wall" of carbonic anhydrase, *J. Am. Chem. Soc.* **2011**, *133*, 14017–14026.

146. Talibov, V. O., Linkuvienė, V., Matulis, D., Danielson, U. H., Kinetically Selective Inhibitors of Human Carbonic Anhydrase Isozymes I, II, VII, IX, XII, and XIII, *J. Med. Chem.* **2016**, *59*, 2083–2093.

147. Linkuvienė, V., Talibov, V. O., Danielson, U. H., Matulis, D., Introduction of Intrinsic Kinetics of Protein-Ligand Interactions and Their Implications for Drug Design, *J. Med. Chem.* **2018**, *61*, 2292–2302.

148. Taylor, P. W., King, R. W., Burgen, A. S., Influence of pH on the Kinetics of Complex Formation between Aromatic Sulfonamides and Human Carbonic Anhydrase, *Biochemistry* **1970**, *9*, 3894–3902.

149. King, R. W., Burgen, A. S. V., Kinetic aspects of structure-activity relations: the binding of sulphonamides by carbonic anhydrase, *Proc. R. Soc. Lond. B Biol. Sci.* **1976**, *193*, 107–125.

150. Gaspari, R., Rechlin, C., Heine, A., Bottegoni, G., Rocchia, W., Schwarz, D., Bomke, J., Gerber, H.-D., Klebe, G., Cavalli, A., Kinetic and Structural Insights into the Mechanism of Binding of Sulfonamides to Human Carbonic Anhydrase by Computational and Experimental Studies, *J. Med. Chem.* **2016**, *59*, 4245–4256.

151. Zöe Fisher, S., Kovalevsky, A. Y., Domsic, J. F., Mustyakimov, M., McKenna, R., Silverman, D. N., Langan, P. A., Neutron structure of human carbonic anhydrase II: Implications for proton transfer, *Biochemistry* **2010**, *49*, 415–421.

152. Pflugrath, J. W., Practical macromolecular cryocrystallography, *Acta Crystallogr. Sect. F: Struct. Biol. Commun.* **2015**, *71*, 622–642.

153. Davis, R. A., Vullo, D., Supuran, C. T., Poulsen, S.-a., Natural Product Polyamines That Inhibit Human Carbonic Anhydrases, *Biomed Res. Int.* **2014**, *2014*, 1–6.

154. Tuñón, I., Silla, E., Pascual-Ahuir, J. L., Molecular surface area and hydrophobic effect, *Protein Eng.* **1992**, *5*, 715–716.

155. Tellinghuisen, J., Optimizing Experimental Parameters in Isothermal Titration Calorimetry, *J. Phys. Chem. B* **2005**, *109*, 20027–20035.

156. Tellinghuisen, J., Designing isothermal titration calorimetry experiments for the study of 1:1 binding: Problems with the "standard protocol", *Anal. Biochem.* **2012**, *424*, 211–220.

157. Myszka, D. G., Abdiche, Y. N., Arisaka, F., Byron, O., Eisenstein, E., Hensley, P., Thomson, J. A., Lombardo, C. R., Schwarz, F., Stafford, W., Doyle, M. L., The ABRF-MIRG'02 Study: Assembly State, Thermodynamic, and Kinetic Analysis of an Enzyme/Inhibitor Interaction, *J. Biomol. Tech.* **2003**, *14*, 247–269.

158. Navratilova, I., Papalia, G. A., Rich, R. L., Bedinger, D., Brophy, S., Condon, B., Deng, T., Emerick, A. W., Guan, H. W., Hayden, T., Heutmekers, T., Hoorelbeke, B., McCroskey, M. C., Murphy, M. M., Nakagawa, T., Parmeggiani, F., Qin, X., Rebe, S., Tomasevic, N., Tsang, T., Waddell, M. B., Zhang, F. F., Leavitt, S., Myszka, D. G., Thermodynamic benchmark study using Biacore technology, *Anal. Biochem.* **2007**, *364*, 67–77.

159. Dumas, P., Ennifar, E., Da Veiga, C., Bec, G., Palau, W., Di Primo, C., Piñeiro, A., Sabín, J., Muñoz, E., Rial, J. in *Methods Enzymol. Vol. 567*, **2016**, pp. 157–180.

160. Zihlmann, P., Silbermann, M., Sharpe, T., Jiang, X., Mühlethaler, T., Jakob, R. P., Rabbani, S., Sager, C. P., Frei, P., Pang, L., Maier, T., Ernst, B., KinITC-One Method Supports both Thermodynamic and Kinetic SARs as Exemplified on FimH Antagonists, *Chem. - A Eur. J.* **2018**, 1–10.

161. Delano, W. L., The PyMOL molecular graphics system on world wide web.

162. Van Der Walt, S., Colbert, S. C., Varoquaux, G., The NumPy array: A structure for efficient numerical computation, *Comput. Sci. Eng.* **2011**, *13*, 22–30.

163. vg v.1.6.1, https://pypi.org/project/vg/, accessed 05.12.2019.

164. Ribeiro, J., Ríos-Vera, C., Melo, F., Schüller, A., Calculation of accurate interatomic contact surface areas for the quantitative analysis of non-bonded molecular interactions, *Bioinformatics* **2019**, *35*, 3499–3501.

165. Avvaru, B. S., Kim, C. U., Sippel, K. H., Gruner, S. M., Agbandje-McKenna, M., Silverman, D. N., McKenna, R., A Short, Strong Hydrogen Bond in the Active Site of Human Carbonic Anhydrase II, *Biochemistry* **2010**, *49*, 249–251.

166. Bondi, A., Van der waals volumes and radii, *J. Phys. Chem.* **1964**, *68*, 441–451.

167. Brautigam, C. A., Zhao, H., Vargas, C., Keller, S., Schuck, P., Integration and global analysis of isothermal titration calorimetry data for studying macromolecular interactions., *Nat. Protoc.* **2016**, *11*, 882–94.

168. Chrysanthopoulos, P. K., Mujumdar, P., Woods, L. A., Dolezal, O., Ren, B., Peat, T. S., Poulsen, S.-A., Identification of a New Zinc Binding Chemotype by Fragment Screening, *J. Med. Chem.* **2017**, *60*, 7333–7349.

169. Wargo, K. A., Banta, W. M., A Comprehensive Review of the Loop Diuretics: Should Furosemide Be First Line, *Ann. Pharmacother.* **2009**, *43*, 1836–1847.

170. Supuran, C. T., Carbonic anhydrase inhibitors, *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3467–3474.

171. Tu, C., Silverman, D. N., Forsman, C., Jonsson, B.-H., Lindskog, S., Role of Histidine 64 in the Catalytic Mechanism of Human Carbonic Anhydrase II Studied with a Site-Specific Mutant, *Biochemistry* **1989**, *28*, 7913–7918.

172. D'Ambrosio, K., Carradori, S., Monti, S. M., Buonanno, M., Secci, D., Vullo, D., Supuran, C. T., De Simone, G., Out of the active site binding pocket for carbonic anhydrase inhibitors., *Chem. Commun. (Camb).* **2015**, *51*, 302–305.

173. Köhler, K., Hillebrecht, A., Schulze Wischeler, J., Innocenti, A., Heine, A., Supuran, C. T., Klebe, G., Saccharin inhibits carbonic anhydrases: Possible explanation for its unpleasant metallic aftertaste, *Angew. Chem. Int. Ed.* **2007**, *46*, 7697–7699.

174. Bell, R. P., Higginson, W. C. E., The catalyzed dehydration of acetaldehyde hydrate, and the effect of structure on the velocity of protolytic reactions, *Proc. R. Soc. London. Ser. A. Math. Phys. Sci.* **1949**, *197*, 141–159.

175. Fisher, S. Z., Kovalevsky, A. Y., Domsic, J. F., Mustyakimov, M., Silverman, D. N., McKenna, R., Langan, P., Preliminary joint neutron and X-ray crystallographic study of human carbonic anhydrase II, *Acta Crystallogr. Sect. F: Struct. Biol. Commun.* **2009**, *65*, 495–498.

176. Koruza, K., Mahon, B. P., Blakeley, M. P., Ostermann, A., Schrader, T. E., McKenna, R., Knecht, W., Fisher, S. Z., Using neutron crystallography to elucidate the basis of selective inhibition of carbonic anhydrase by saccharin and a derivative, *J. Struct. Biol.* **2019**, *205*, 147–154.

177. Pape, T., Schneider, T. R., HKL2MAP : a graphical user interface for macromolecular phasing with SHELX programs, *J. Appl. Crystallogr.* **2004**, *37*, 843–844.

178. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., Wilson, K. S., Overview of the CCP4 suite and current developments, *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **2011**, *67*, 235–242.

179. Okaya, Y., The Crystal Structure of Saccharin, *o*-Sulfobenzimide, $C_6H_4CO.NH.SO_2$, an Artificial Sweetening, *Acta Crystallogr. Sect. B: Struct. Crystallogr. Cryst. Chem.* **1969**, *25*, 2257–2263.

180. Macarron, R., Banks, M. N., Bojanic, D., Burns, D. J., Cirovic, D. A., Garyantes, T., Green, D. V., Hertzberg, R. P., Janzen, W. P., Paslay, J. W., Schopfer, U., Sittampalam, G. S., Impact of high-throughput screening in biomedical research, *Nat. Rev. Drug Discov.* **2011**, *10*, 188–195.

181. Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W., Jhoti, H., Twenty years on: the impact of fragments on drug discovery, *Nat. Rev. Drug Discov.* **2016**, *15*, 605–619.

182. Scott, D. E., Coyne, A. G., Hudson, S. A., Abell, C., Fragment-Based Approaches in Drug Discovery and Chemical Biology, *Biochemistry* **2012**, *51*, 4990–5003.

183. Bohacek, R. S., McMartin, C., Guida, W. C., The Art and Practice of Structure-Based Drug Design: A Molecular Modeling Perspective, *Med. Res. Rev.* **1996**, *16*, 3–50.

184. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Deliv. Rev.* **2001**, *46*, 3–26.

185. Congreve, M., Carr, R., Murray, C., Jhoti, H., A 'Rule of Three' for fragment-based lead discovery?, *Drug Discovery Today* **2003**, *8*, 876–877.

186. Ruddigkeit, L., Van Deursen, R., Blum, L. C., Reymond, J. L., Enumeration of 166 Billion Organic Small Molecules in the Chemical Universe Database GDB-17, *J. Chem. Inf. Model.* **2012**, *52*, 2864–2875.

187. Shuker, S. B., Hajduk, P. J., Meadows, R. P., Fesik, S. W., Discovering High-Affinity Ligands for Proteins: SAR by NMR, *Science* **1996**, *274*, 1531–1534.

188. Congreve, M., Murray, C. W., Blundell, T. L., Structural biology and drug discovery, *Drug Discovery Today* **2005**, *10*, 895–907.

189. Murray, C. W., Blundell, T. L., Structural biology in fragment-based drug design, *Curr. Opin. Struct. Biol.* **2010**, *20*, 497–507.

190. Köster, H., Craan, T., Brass, S., Herhaus, C., Zentgraf, M., Neumann, L., Heine, A., Klebe, G., A Small Nonrule of 3 Compatible Fragment Library Provides High Hit Rate of Endothiapepsin Crystal Structures with Various Fragment Chemotypes, *J. Med. Chem.* **2011**, *54*, 7784–7796.

191. Schiebel, J., Radeva, N., Krimmer, S. G., Wang, X., Stieler, M., Ehrmann, F. R., Fu, K., Metz, A., Huschmann, F. U., Weiss, M. S., Mueller, U., Heine, A., Klebe, G., Six Biophysical Screening Methods Miss a Large Proportion of Crystallographically Discovered Fragment Hits: A Case Study, *ACS Chem. Biol.* **2016**, *11*, 1693–1701.

192. Tsai, J., Lee, J. T., Wang, W., Zhang, J., Cho, H., Mamo, S., Bremer, R., Gillette, S., Kong, J., Haass, N. K., Sproesser, K., Li, L., Smalley, K. S. M., Fong, D., Zhu, Y.-L., Marimuthu, A., Nguyen, H., Lam, B., Liu, J., Cheung, I., Rice, J., Suzuki, Y., Luu, C., Settachatgul, C., Shellooe, R., Cantwell, J., Kim, S. H., Schlessinger, J., Zhang, K. Y., West, B. L., Powell, B., Habets, G., Zhang, C., Ibrahim, P. N., Hirth, P., Artis, D. R., Herlyn, M., Bollag, G., Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 3041–3046.

193. Mérour, J.-Y., Buron, F., Plé, K., Bonnet, P., Routier, S., The azaindole framework in the design of kinase inhibitors, *Molecules* **2014**, *19*, 19935–19979.

194. Souers, A. J., Leverson, J. D., Boghaert, E. R., Ackler, S. L., Catron, N. D., Chen, J., Dayton, B. D., Ding, H., Enschede, S. H., Fairbrother, W. J., Huang, D. C. S., Hymowitz, S. G., Jin, S., Khaw, S. L., Kovar, P. J., Lam, L. T., Lee, J., Maecker, H. L., Marsh, K. C., Mason, K. D., Mitten, M. J., Nimmer, P. M., Oleksijew, A., Park, C. H., Park, C.-M., Phillips, D. C., Roberts, A. W., Sampath, D., Seymour, J. F., Smith, M. L., Sullivan, G. M., Tahir, S. K., Tse, C., Wendt, M. D., Xiao, Y., Xue, J. C., Zhang, H., Humerickhouse, R. A., Rosenberg, S. H., Elmore, S. W., ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets, *Nat. Med.* **2013**, *19*, 202–208.

195. PDBeChem Database.

196. Glöckner, S., Masterthesis, Philipps-Universität Marburg, **2016**.

197. Di Fiore, A., Maresca, A., Supuran, C. T., De Simone, G., Hydroxamate represents a versatile zinc binding group for the development of new carbonic anhydrase inhibitors, *Chem. Commun.* **2012**, *48*, 8838–8840.

198. Martin, D. P., Cohen, S. M., Nucleophile recognition as an alternative inhibition mode for benzoic acid based carbonic anhydrase inhibitors, *Chem. Commun.* **2012**, *48*, 5259–5261.

199. Choudhary, A., Gandla, D., Krow, G. R., Raines, R. T., Nature of Amide Carbonyl-Carbonyl Interactions in Proteins, *J. Am. Chem. Soc.* **2009**, *131*, 7244–7246.

200. Hemenway, J. N., Carvalho, T. C., Rao, V. M., Wu, Y., Levons, J. K., Narang, A. S., Paruchuri, S. R., Stamato, H. J., Varia, S. A., Formation of Reactive Impurities in Aqueous and Neat Polyethylene Glycol 400 and Effects of Antioxidants and Oxidation Inducers, *J. Pharm. Sci.* **2012**, *101*, 3305–3318.

201. Salonen, L. M., Ellermann, M., Diederich, F., Aromatische Ringe in chemischer und biologischer Erkennung: Energien und Strukturen, *Angew. Chem.* **2011**, *123*, 4908–4944.

202. Cramer, J., Schiebel, J., Wulsdorf, T., Grohe, K., Najbauer, E. E., Ehrmann, F. R., Radeva, N., Zitzer, N., Linne, U., Linser, R., Heine, A., Klebe, G., A False-Positive Screening Hit in Fragment-Based Lead Discovery: Watch out for the Red Herring, *Angew. Chem. Int. Ed.* **2017**, *56*, 1908–1913.

203. Pearce, N. M., Krojer, T., Bradley, A. R., Collins, P., Nowak, R. P., Talon, R., Marsden, B. D., Kelm, S., Shi, J., Deane, C. M., Delft, F. von, A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density, *Nat. Commun.* **2017**, *8*, 15123.

204. Cimmperman, P., Baranauskienė, L., Jachimovičiūtė, Jachno, J., Torresan, J., Michailovienė, V., Matulienė, J., Sereikaitė, J., Bumelis, V., Matulis, D., A Quantitative Model of Thermal Stabilization and Destabilization of Protein by Ligands, *Biophys. J.* **2008**, *95*, 3222–3231.

205. Laemmli, U. K., Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4, *Nature* **1970**, *227*, 680–685.

206. Fu, K., PhD Thesis, Philipps-Universität Marburg, **2015**.

207. Lausi, A., Polentarutti, M., Onesti, S., Plaisier, J. R., Busetto, E., Bais, G., Barba, L., Cassetta, A., Campi, G., Lamba, D., Pifferi, A., Mande, S. C., Sarma, D. D., Sharma, S. M., Paolucci, G., Status of the crystallography beamlines at Elettra, *Eur. Phys. J. Plus* **2015**, *130*, 43.

208. Cianci, M., Bourenkov, G., Pompidor, G., Karpics, I., Kallio, J., Bento, I., Roessle, M., Cipriani, F., Fiedler, S., Schneider, T. R., P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing, *J. Synchrotron Radiat.* **2017**, *24*, 323–332.

209. Kabsch, W., XDS, *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **2010**, *66*, 125–132.

210. Sparta, K. M., Krug, M., Heinemann, U., Mueller, U., Weiss, M. S., XDSAPP2.0, *J. Appl. Crystallogr.* **2016**, *49*, 1085–1092.

211. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., Adams, P. D., Towards automated crystallographic structure refinement with phenix.refine, *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **2012**, *68*, 352–367.

212. Brünger, A. T., Free *R* value: a novel quantity for assessing the accuracy of crystal structures, *Nature* **1992**, *355*, 472–475.

213. Langer, G., Cohen, S. X., Lamzin, V. S., Perrakis, A., Automated macromolecular model building for x-ray crystallography using ARP/wARP version 7, *Nat. Protoc.* **2008**, *3*, 1171–1179.

214. Mizoue, L. S., Tellinghuisen, J., The role of backlash in the "first injection anomaly" in isothermal titration calorimetry, *Anal. Biochem.* **2004**, *326*, 125–127.

215. Velazquez-Campoy, A., Freire, E., Isothermal titration calorimetry to determine association constants for high-affinity ligands, *Nat. Protoc.* **2006**, *1*, 186–191.

216. Senner, M. F., Olson, A. J., Spehner, J.-C., Reduced Surface: An Efficient Way to Compute Molecular Surfaces, *Biopolymers* **1996**, *38*, 305–320.

217. Case, D. A., Cerutti, D. S., Cheatham III, T. E., Darden, T. A., Duke, T. E., Giese, T. J., Gohlke, H., Goetz, A. W., Greene, D., Homeyer, N., Izadi, S., Kovalenko, A., Lee, T. S., LeGrand, S., Li, P., Lin, C., Liu, J., Luchko, T., Luo, R., Mermelstein, D., Merz, K. M., Monard, G., Nguyen, H., Omelyan, I., Onufriev, A., Pan, F., Qi, R., Roe, D. R., Roitberg, A., Sagui, C., Simmerling, C. L., Botello-Smith, W. M., Swails, J., Walker, R. C., Wang, J., Wolf, R. M., Wu, X., Xiao, L., York, D. M., Kollman, P. A., AMBER 2017, University of California, San Francisco, **2017**.

218. Jones, E., Oliphant, E., Peterson, P., SciPy: Open Source Scientific Tools for Python, **2001**.

219. Zhao, Y., Truhlar, D. G., The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other function, *Theor. Chem. Acc.* **2008**, *120*, 215–241.

220. Zhao, Y., Truhlar, D. G., Density Functionals with Broad Applicability in Chemistry, *Acc. Chem. Res.* **2008**, *41*, 157–167.

221. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A., Nakatsuji, H., Caricato, M., Marenich, A. V., Bloino, J., Janesko, B. G., Gomperts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., Izmaylov, A. F., Sonnenberg, J. L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V. G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery Jr., J. A., Peralta, J. E., Ogliaro, F., Bearpark, M. J., Heyd, J. J., Brothers, E. N., Kudin, K. N., Staroverov, V. N., Keith, T. A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A. P., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Millam, J. M., Klene, M., Adamo, C., Cammi, R., Ochterski, J. W., Martin, R. L., Morokuma, K., Farkas, O., Foresman, J. B., Fox, D. J., Gaussian 16, Revision C.01, **2016**.

222. Marenich, A. V., Cramer, C. J., Truhlar, D. G., Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dieletric Constant and Atomic Surface Tensions, *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

223. Tam, K. Y., Takács-Novák, K., Multi-wavelength spectrophotometric determination of acid dissociation constants: A validation study, *Anal. Chim. Acta* **2001**, *434*, 157–167.

224. Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, J. M., PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.* **1993**, *26*, 283–291.

225. Kleywegt, G. J., Zou, J. Y., Kjeldgaard, M., Jones, T. A. in *Int. Tables Crystallogr. Vol. F. Crystallogr. Biol. Macromol.* (Eds.: Rossmann, M. G., Arnold, E.), Dordrecht: Kluwer Academic Publisher, The Netherlands, **2001**, Chapter 17.1, pp. 353–356, 366–367.

226. Harris, D. C., *Lehrbuch der Quantitativen Chemischen Analyse*, 8th ed., (Eds.: Werner, G., Werner, T.), Springer-Verlag, Berlin Heidelberg, **2014**.

References

244

Erklärung

Ich versichere, dass ich meine Dissertation

Thermodynamic, Kinetic and Crystallographic Investigations of Benzenesulfonamides as Ligands of Human Carbonic Anhydrase II

selbstständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe. Alle vollständig oder sinngemäß übernommenen Zitate sind als solche gekennzeichnet.

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

Marburg, den _____

Steffen Glöckner