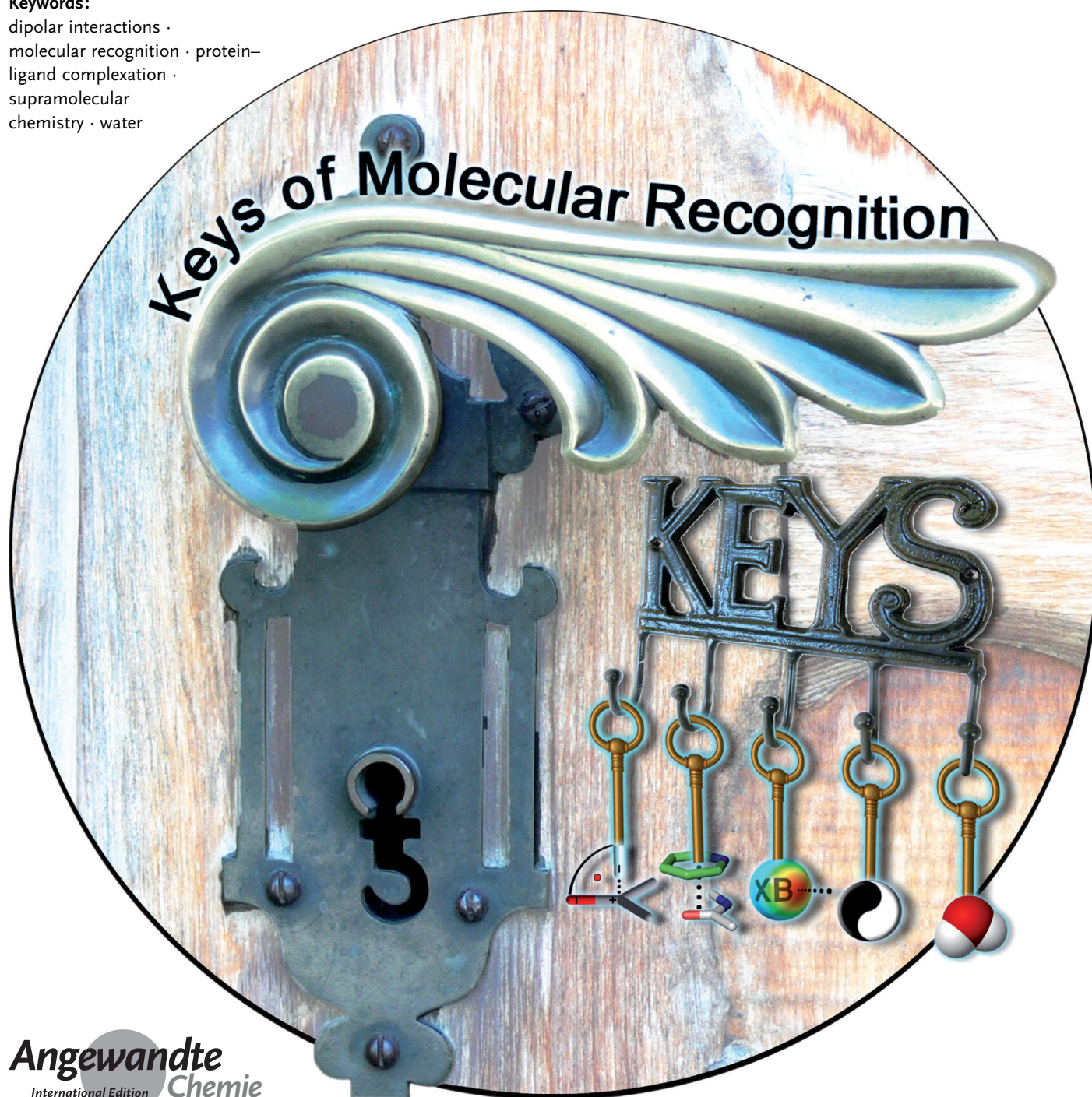


Molecular Recognition in Chemical and Biological Systems

*Elke Persch, Oliver Dumele, and François Diederich**

Keywords:

dipolar interactions ·
molecular recognition · protein–
ligand complexation ·
supramolecular
chemistry · water



Structure-based ligand design in medicinal chemistry and crop protection relies on the identification and quantification of weak noncovalent interactions and understanding the role of water. Small-molecule and protein structural database searches are important tools to retrieve existing knowledge. Thermodynamic profiling, combined with X-ray structural and computational studies, is the key to elucidate the energetics of the replacement of water by ligands. Biological receptor sites vary greatly in shape, conformational dynamics, and polarity, and require different ligand-design strategies, as shown for various case studies. Interactions between dipoles have become a central theme of molecular recognition. Orthogonal interactions, halogen bonding, and amide... π stacking provide new tools for innovative lead optimization. The combination of synthetic models and biological complexation studies is required to gather reliable information on weak noncovalent interactions and the role of water.

1. Introduction

There are close to 90 000 X-ray crystal structures of proteins, nucleic acids, and their complexes deposited in the Protein Data Bank (PDB), and more than 10 000 structures have been solved by NMR spectroscopic methods.^[1] These together with the more than 600 000 small-molecule crystal structures in the Cambridge Structural Database (CSD)^[2] result in rich structural information on molecular and supramolecular conformations^[3] and intermolecular interactions^[4] now being available for use in the rational optimization of biologically active ligands as leads for new pharmaceuticals and agrochemicals. Despite the increasing efficiency of modern structure-based design strategies,^[5–7] their outcome is not always as predicted. Molecular recognition in chemical and biological systems in aqueous solutions still holds many mysteries, in particular related to the role of water.^[8]

Our research group has been involved in molecular recognition studies^[9] since the beginning of the 1980s. A key topic in this early work was the investigation of host–guest complexation with synthetic cyclophane receptors in aqueous solution.^[10,11] Isothermal titration calorimetry (ITC)^[12] led to the surprising discovery that the tight complexation of apolar neutral aromatic guests in the preorganized narrow cyclophane cavities in water is not entropically driven, as expected from the classical hydrophobic effect,^[13] but rather has a strong favorable enthalpic signature, usually with a compensating unfavorable entropic term.^[14,15] The role of water and the energetics of its displacement in synthetic and biological ligand recognition events is today a vigorous topic of on-going experimental and computational research and will be the subject of case studies presented in Section 2.

After our move to ETH Zurich in 1992, collaborations with scientists at Roche Basel provided us with the unique opportunity to expand our molecular recognition studies to the exploration of proper cavity filling in biological receptors, in particular enzymes. Structure-based nonpeptidic ligand design was applied to various enzyme targets, starting with the serine protease thrombin from the blood coagulation cascade,

which features structurally well-defined hydrophobic subpockets in the active site.^[6b,16] More challenging targets were later addressed, with binding sites of higher conformational flexibility, highly hydrophilic character, or a very large size. Additionally, concepts such as bisubstrate inhibition were explored. Again, fertile collaborations on the biological and biostructural aspects with academic researchers as well as scientists from Roche Basel and other companies, such as BASF, made this research possible. This work was first reviewed in 2008, and Section 3 reports new case studies, in particular addressing enzymes with the potential to become new targets in the fight against malaria and human African trypanosomiasis (HAT, sleeping sickness).^[17,18] In parallel, we summarize our investigations of guest binding in the confined cavities of synthetic, container-type receptors, following the leads by Cram^[19] and Rebek,^[20] who showed that covalent and supramolecular capsules, respectively, are ideal to investigate—in detail and separated from any influence of the exterior bulk phase—individual intermolecular interactions and molecular conformations of guests trapped in their inner phase.^[21]

The collaborative development of nonpeptidic small-molecule inhibitors of thrombin led, in 2003, to the finding that C–F bonds of the ligands interact in an orthogonal fashion with peptide bond C=O groups of the enzyme^[22] and this raised our interest in the quantification of orthogonal

From the Contents

1. Introduction	3291
2. The Role of Water in Complexation at Synthetic and Biological Recognition Sites	3292
3. Ligands for Filling Enzyme Pockets of Different Shape, Conformational Dynamics, and Polarity	3298
4. Interactions between Dipoles: Orthogonal Interactions, Halogen Bonding, and Amide...π Stacking	3309
5. Summary and Conclusions	3318

[*] Dipl.-Chem. E. Persch,^[†] Dipl.-Chem. O. Dumele,^[‡]

Prof. Dr. F. Diederich

Laboratorium für Organische Chemie, Departement Chemie und Angewandte Biowissenschaften, ETH Zürich

Vladimir-Prelog-Weg 3, 8093 Zürich (Switzerland)

E-mail: diederich@org.chem.ethz.ch

[†] These authors contributed equally to this Review, their sequence was determined by the flip of a coin.



Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201408487>.

dipolar interactions^[23,24] and later of halogen bonding, again initially in collaboration with scientists at Roche.^[25] Investigations and quantification of dipolar interactions, both in protein–ligand complexes and in synthetic model systems, are the subject of Section 4 of this Review.

2. The Role of Water in Complexation at Synthetic and Biological Recognition Sites

Water is essential to all life on Earth. In fact, life without water seems hard to imagine: when scientists investigate the possibility for life on other planets in interstellar space, they are first looking to find traces of water. In this section we present convincing new studies by others that confirm the enthalpically driven (the “nonclassical hydrophobic effect”) complexation of apolar solutes in tight binding pockets in aqueous environments. This is followed by examples that show a large gain in binding free enthalpy on substitution of single water molecules in protein–ligand complexes and a case study investigating the energetics of water replacements in water clusters seen in protein cocrystal structures.

2.1. Enthalpically Driven Complexation in Tight Apolar Binding Sites in Water

The nonclassical hydrophobic effect^[26] measured by ITC for the complexation of aromatic solutes in apolar cyclophane cavities has been confirmed for many other synthetic host–guest systems.^[27,28] We assigned this enthalpic driving force to both a gain in van der Waals interactions in the tight complexes formed and to gains in solvent cohesive interactions, when the water molecules solvating the cyclophane cavity and the receptor are transferred into the bulk upon cyclophane–solute complexation.^[10b,14] A favorable enthalpic term had previously also been measured by ITC for the binding of apolar solutes in enzyme pockets and for inclusion complexes with cyclodextrins.^[14e] However, additional polar interactions (hydrogen bonding, ion pairing, dipolar interactions) of ligands bound to biological receptors or complexed to cyclodextrins can mask the nonclassical hydrophobic effect (favorable enthalpic term compensated by an unfavorable entropy change), as these interactions possess

different thermodynamic characteristics.^[11b,29] For example, ion pairing results in a larger number of water molecules from the ordered solvation shells around the charged species being released, which results in a favorable entropic change.

Bender and co-workers as well as Saenger were among the first to propose that water molecules inside a cavity—in this case the cyclodextrin torus—are of high energy (enthalpically frustrated), since they cannot undergo the favorable fourfold hydrogen bonding to four other water molecules that is characteristic for the bulk phase.^[30] This concept of enthalpically frustrated water molecules has found support in recent theoretical studies, which also revealed the role of the concave geometry of the apolar binding pocket and the nature and shape of the guest.^[17b,31]

Synthetic host–guest chemistry has currently undergone a major renaissance, to which the development of the chemistry and the investigation of the outstanding receptor properties of the cucurbiturils (CB n) make a bold contribution.^[32] These macrocycles are constructed from n glycoluril moieties, and the smaller and most investigated derivatives CB7, CB8, and CB9 (Figure 1), in particular, possess deep rigid cavities with pronounced hydrophobic character. The association constants for the 1:1 inclusion complexation of size-complementary guests such as neutral ferrocene, bicyclo-[2.2.2]octane, and adamantane derivatives towards CB7 (Figure 1a) are very high, with K_a values at 298 K between 10^9 and 10^{10} M^{-1} .^[33] The ITC data show for all the neutral guests studied that binding is strongly enthalpically driven, with some compensation by an unfavorable entropy term, which is illustrated in Figure 1 for adamantanol as the guest. Cucurbiturils are solvated in their interior cavity by enthalpically frustrated water molecules, which lack hydrogen-bonding partners. This was revealed by the crystal structure of inverted CB6 (in which one glycoluril moiety is inverted), which features three cavity-bound water molecules that lack tetra-coordination,^[28,34] as well as molecular dynamics simulations paired with ITC.^[14a,28,35] Upon guest complexation, the release of these high-energy water molecules from the cavity into the bulk, where they regain their favorable fourfold coordination, makes the major contribution to the strength of the host–guest association.^[36] Cucurbiturils have cavities of very low polarizability—lower than water—^[35a] and a gain in attractive dispersion interactions upon guest inclusion is not expected to be a driving force for inclusion complexation in this case. A



Oliver Dumele was born in 1987 in Ludwigshafen/Rhine, Germany, and studied Chemistry at the University of Mainz and UC Berkeley (with J. M. J. Freché). He joined the group of Prof. K. Müllen at the Max-Planck-Institut für Polymerforschung in Mainz, where he obtained his diploma in the subgroup of Prof. T. Weil. After an industrial internship at BASF Ludwigshafen and a research stay at NUS Singapore (Prof. C.-H. Tan), he joined the group of Prof. F. Diederich at ETH Zurich for PhD research on halogen bonding in model systems.



Elke Persch was born in Duisburg, Germany (1986), and studied Chemistry at the Universities of Heidelberg and Lund (2005–2010). She completed her diploma in the group of Prof. D. Menche in the field of natural products and subsequent to an internship at BASF, joined the group of Prof. F. Diederich at ETH Zurich. Currently, she is working on the development of trypanothione reductase inhibitors.

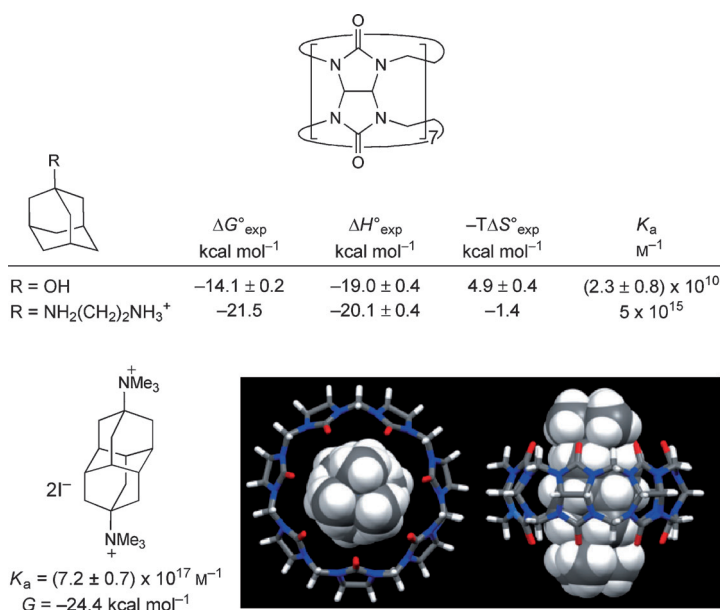


Figure 1. a) Chemical structure of cucurbit[7]uril (CB7) and binding characteristics of neutral and dicationic adamantyl derivatives as measured by ITC in H₂O at $T = 298 \text{ K}$. b) Diadamantyl diammonium guest with a record-high binding affinity in D₂O at 298 K . Top and side views of the X-ray crystal structure of diadamantyldiammonium diiodide bound to CB7. The crystal structure is reproduced with permission from the publisher.^[38]

recent review by Biederman, Nau, and Schneider beautifully extends this analysis of high-energy water in concave cavities and its displacement by guests to a wide class of synthetic receptors with cavities of diverse sizes and shapes.^[28]

Changing from a neutral alcohol to an adamantyldiammonium cation (Figure 1, or the corresponding ferrocene and bicyclo[2.2.2]octane derivatives) as guest,^[33] results in the enthalpic contribution to the binding free enthalpy hardly changing. However, the entropy term becomes much more favorable, as a consequence of the release of the water molecules, which solvate the ammonium ions and the convergent rim dipoles of the glycoluril carbonyl groups, with the latter undergoing tight ion–dipole interactions in the complexes.^[29,37] By both optimal cavity filling under displacement of high-energy water and replacing the ammonium ion solvating water molecules by the C=O dipoles at the host rim, Isaacs and co-workers recently reported a record-high

association constant of $K_a = 7.2 \times 10^{17} \text{ M}^{-1}$ ($\Delta G = -24.4 \text{ kcal mol}^{-1}$) for the complex of a diadamantane diammonium guest with CB7 (Figure 1 b).^[38]

This study once more confirms the value of exploring structurally well-characterized synthetic host–guest systems to decipher solvent effects on molecular recognition in water. It can be stated with confidence that apolar binding in tight concave pockets in biological receptor sites is driven by the same forces, namely a gain in dispersion interactions and replacement of high-energy water. However, the ligands applied to biomolecular recognition are most often much larger and engage in other interactions with the protein, such as hydrogen bonding or ion pairing, which makes it more difficult to decipher and dissect the contributions of apolar binding to the binding free enthalpy. Furthermore, the sizes and shapes of pockets in biological systems feature a much greater diversity, which also affects the thermodynamic signature.

The ITC results for cucurbituril binding convey an important message for ligand design in biological molecular recognition. A large amount of enthalpy can be gained by filling deep apolar protein pockets; however, with the cost of an unfavorable entropy change. Through judicious additional decoration of the ligand by groups, which can undergo polar interactions, such as ion pairing or ion–dipole interactions, the enthalpic gains from the nonclassical hydrophobic effect are maintained, while the entropic term also becomes more favorable. Water takes the reverse role in both processes: it contributes to a favorable enthalpy term in the hydrophobic binding in tight pockets and to a favorable entropy term in the polar interaction.

2.2. Replacement of Water Molecules Seen in X-Ray Crystal Structures of Proteins and their Complexes by Ligands

With the advent of cryogenic X-ray crystallography, abundant information on the hydration of proteins and their ligand complexes has become available.^[39] Hydration at binding sites can extend from single water molecules to entire water clusters, and conserved, ordered water molecules can be identified by superimposition of different cocrystal structures of a protein with the apo structure (if available). The next question of importance for further lead optimization is: can some of these water molecules be replaced by newly introduced ligand substituents with a gain of binding free enthalpy? Dunitz predicted that the entropic gains upon transferring a weakly bound water molecule from the interior of the protein into the bulk could generate a gain in free enthalpy of up to 2 kcal mol^{-1} at 300 K .^[40]

The past decade has seen a vigorous development of useful water scoring functions and computational approaches to distinguish between tightly bound and displaceable water molecules in protein–ligand cocrystal structures in an attempt to rationally guide ligand optimization by energetically favorable replacement of crystallographically ordered water



François Diederich was born in the Grand-Duchy of Luxembourg (1952) and studied Chemistry at the University of Heidelberg (1971–1977). He joined Prof. H. A. Staab at the Max-Planck-Institut für Medizinische Forschung in Heidelberg for his PhD (1977–1979). After postdoctoral studies with Prof. O. L. Chapman at UCLA (1979–1981), he returned to Heidelberg for his Habilitation (1981–1985). He then joined the Faculty in the Department of Chemistry and Biochemistry at UCLA, where he became Full Professor in 1989. Since 1992, he has been a Professor in the Laboratory of Organic Chemistry in the Department of Chemistry and Applied Biosciences at ETH Zurich.

molecules.^[41] These approaches certainly have contributed to the increasing number of energetically favorable water replacements by ligands in lead optimizations reported in the medicinal chemistry literature. As a first approximation rule of thumb, ordered water molecules that undergo four bonding interactions to the protein (hydrogen bonds, dipolar contacts) cannot be replaced by ligands. In the modeling studies, they are best treated in a similar manner as the protein surface. The replacement of water molecules with three close contacts to the protein is also very difficult and not recommended. Water molecules with two directional contacts to the protein are replaceable, but the gain in free enthalpy is usually small. On the other hand, water molecules with only one polar interaction to the protein are replaceable, often with a substantial gain in free enthalpy.

In the next sections we illustrate the energetically favorable replacement of single water molecules in two case studies and subsequently review the ongoing, challenging investigation of replacing water molecules in entire water clusters.

2.3. Single Water Replacements

As will be discussed in more detail in Section 3, we have pursued in a collaborative network the development of inhibitors against three of the seven enzymes in the non-mevalonate pathway of isoprenoid biosynthesis. During the search for new herbicides, whereby a BASF compound library was screened against the third enzyme from the pathway, IspD from *Arabidopsis thaliana* (At), the complexation of triazolopyrimidine **1** (Figure 2a) was observed with an IC_{50} value (median inhibitory concentration) of 140 nM.^[42] The X-ray cocrystal structure revealed that it was bound into a newly discovered allosteric pocket (Figure 2b). As the compound showed good herbicidal activity, it was further optimized by structure-based design. The cocrystal structure revealed the binding of an ordered water molecule to the side chain of Arg157 and to an N atom of the triazole ring of the ligand. The replacement of this water molecule was targeted; however, introducing a carboxylate in **2** to replace the water

molecule and interact with Arg157 was not successful. Although the crystal structure revealed the expected binding mode with the water replaced, compound **2** was found to be a very weak inhibitor (IC_{50} = 274 μ M), mainly because of the large costs for the desolvation of the carboxylate moiety upon binding. When a CN group was introduced in ligand **3** instead, a large gain in potency over the initial hit **1** was observed (IC_{50} = 35 nM). The cocrystal structure (Figure 2c) confirmed the replacement of the water molecule by this group, which forms a moderately strong hydrogen bond with the side chain of Arg157.^[42]

In another investigation in collaboration with Roche Basel, we developed bisubstrate inhibitors of the enzyme catechol-*O*-methyltransferase (COMT), which catalyzes the methylation of biologically active catechols in the presence of *S*-adenosinemethionine (SAM) and Mg^{2+} ions.^[43] Inhibition of this enzyme is important in several disorders of the central nervous system and finds application in the treatment of Parkinson patients by blocking undesired methylation of administered L-DOPA, thereby leading to enhanced dopamine levels in the brain.^[44]

When the cocrystal structure of COMT bound to bisubstrate inhibitor **4** (IC_{50} = 9 nM, Figure 3a) was solved, it became evident that a water molecule had been imported into the SAM nucleobase pocket, solvating N(6)H₂ and N(7) of adenine through moderately strong hydrogen bonding (Figure 3b).^[43b] This water molecule did not show any polar contact to the protein. In a new series of ligands,^[45] with a 4-fluorophenyl group replacing the NO₂ group on the catechol moiety, we introduced small alkyl substituents (Me, Et, Pr, cPr, CH₂CH₂OH) on N(6) and solved the cocrystal structures for four of these complexes.^[45a] In each case, the substituent at N(6) had taken the position of the water molecule imported by **4**, as shown for the complex of bound **5** (IC_{50} = 12 nM) in Figure 3c.

The new N-alkylated ligands showed IC_{50} values between 9 and 41 nM, which was unexpectedly in the same range as unsubstituted **6** (IC_{50} = 31 nM).^[45a] This was unexpected because in bulk aqueous solution, the free ligands prefer the *s-cis* conformation by at least $\Delta\Delta G$ = -1.8 kcal mol⁻¹ according to ¹H NMR studies; however, the N-alkylated ligands

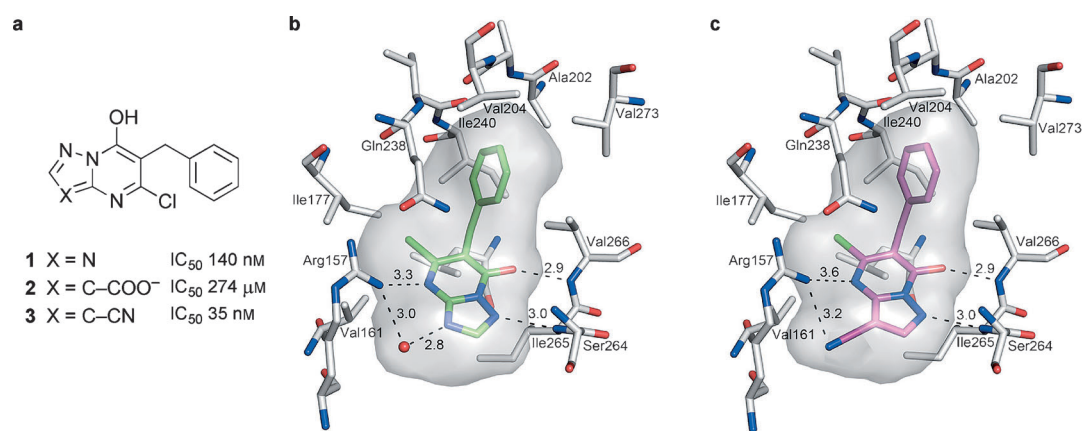


Figure 2. a) Allosteric inhibitors of the enzyme IspD and X-ray crystal structures of AtIspD cocrystallized with b) **1** (1.4 Å resolution, PDB ID: 2YC3) and c) **3** (1.6 Å resolution, PDB ID: 2YC5).^[42] Color code: C_{enzyme} gray, O red, N blue, Cl green, C_{ligand(1)} green, C_{ligand(3)} hot pink.

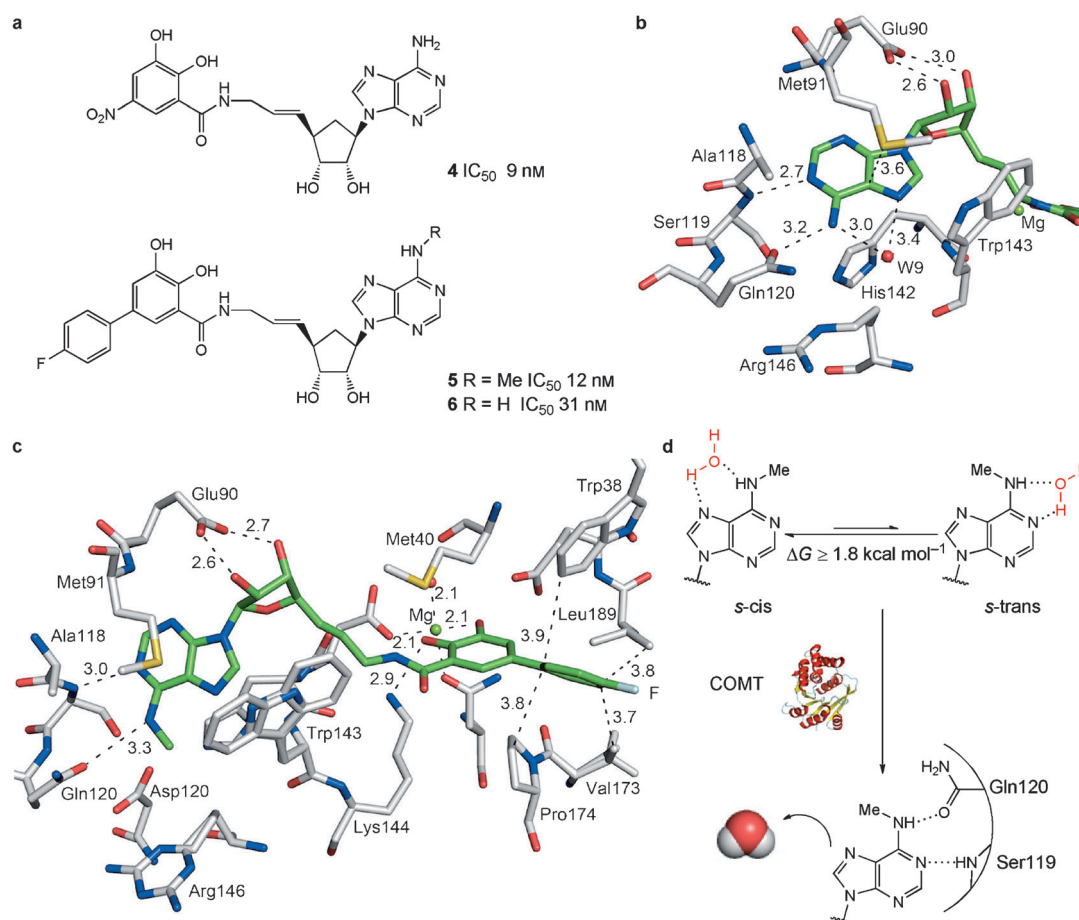


Figure 3. a) Bisubstrate inhibitors of COMT.^[43b, 45a] b) Structure of the adenosine binding site of the ternary complex of **4** with COMT and a Mg^{2+} ion (2.6 Å resolution, PDB ID: 1JR4)^[43b] showing the imported water molecule (W9). c) Binding mode of bisubstrate inhibitor **5** (1.30 Å resolution, PDB ID: 3HVH).^[45a] Color code: C_{enzyme} gray, O red, N blue, S yellow, F turquoise, Mg light green, C_{ligand} green. d) Summary of the process leading to the release of ligand-bound water.

bind in the unfavorable *s-trans* conformation (Figure 3d). This suggests that the displacement of the water molecule in the cocrystal structure contributes at least $-1.8 \text{ kcal mol}^{-1}$ to the measured binding affinity, thereby compensating similar energetic costs for adopting the *s-trans* ligand conformation (Figure 3d).

2.4. Addressing Water Clusters

In collaboration with the group of G. Klebe (Univ. Marburg), we have established that *lin*-benzoguanines (*linear* benzologue of guanine) and their 2- and 4-substituted derivatives are potent inhibitors of tRNA-guanine transglycosylase (TGT) from *Zymomonas mobilis* (Zm).^[46] TGT is a homodimeric enzyme, which catalyzes base exchange (guanine against preQ1) at the wobble position (G_{34}) in the anticodon of tRNA in prokaryotes, with preQ1 ultimately being transformed into the nucleobase queuine.^[47] Parasites lacking the TGT gene are apathogenic, thus validating the inhibition of TGT as a drug target against the highly contagious foodborne illness shigellosis. The human TGT

enzyme directly introduces queuine that is available from the food chain.

The *lin*-benzoguanines bind to the nucleobase (G_{34}) site, and without extended substituents, as in **7** ($K_i = 77 \text{ nM}$, Figure 4a), leave a pyramidal five-water cluster that solvates the side chains of the two catalytic Asp280 and Asp102 unaffected (Figure 4b).^[46c] It should be noted that this cluster differs from the so-called Walrafen-five-water cluster,^[48] in which one water molecule is tetrahedrally coordinated to four other water molecules, which is predominantly the situation in bulk water.^[30c] Near the tRNA-ribose-34 pocket is a shallow hydrophobic cleft (Val45, Val282, Leu68), which can be reached and filled with phenyl or cycloalkyl rings linked to position 4 of the ligand. The linker must penetrate through the water cluster and displace several of its components. Short lipophilic alkyl linkers provided the right directionality and enabled filling of the cleft, but were inappropriate, as they disrupted the required strong solvation of the two Asp side chains in an energetically highly unfavorable way.^[41i] However, when a protonated 2-aminoethyl substituent was attached to position 4 (**8**, $K_i = 55 \text{ nM}$), several water molecules from the cluster were displaced in an energetically neutral way (Figure 4c), when compared to **9** ($K_i = 58 \text{ nM}$) without

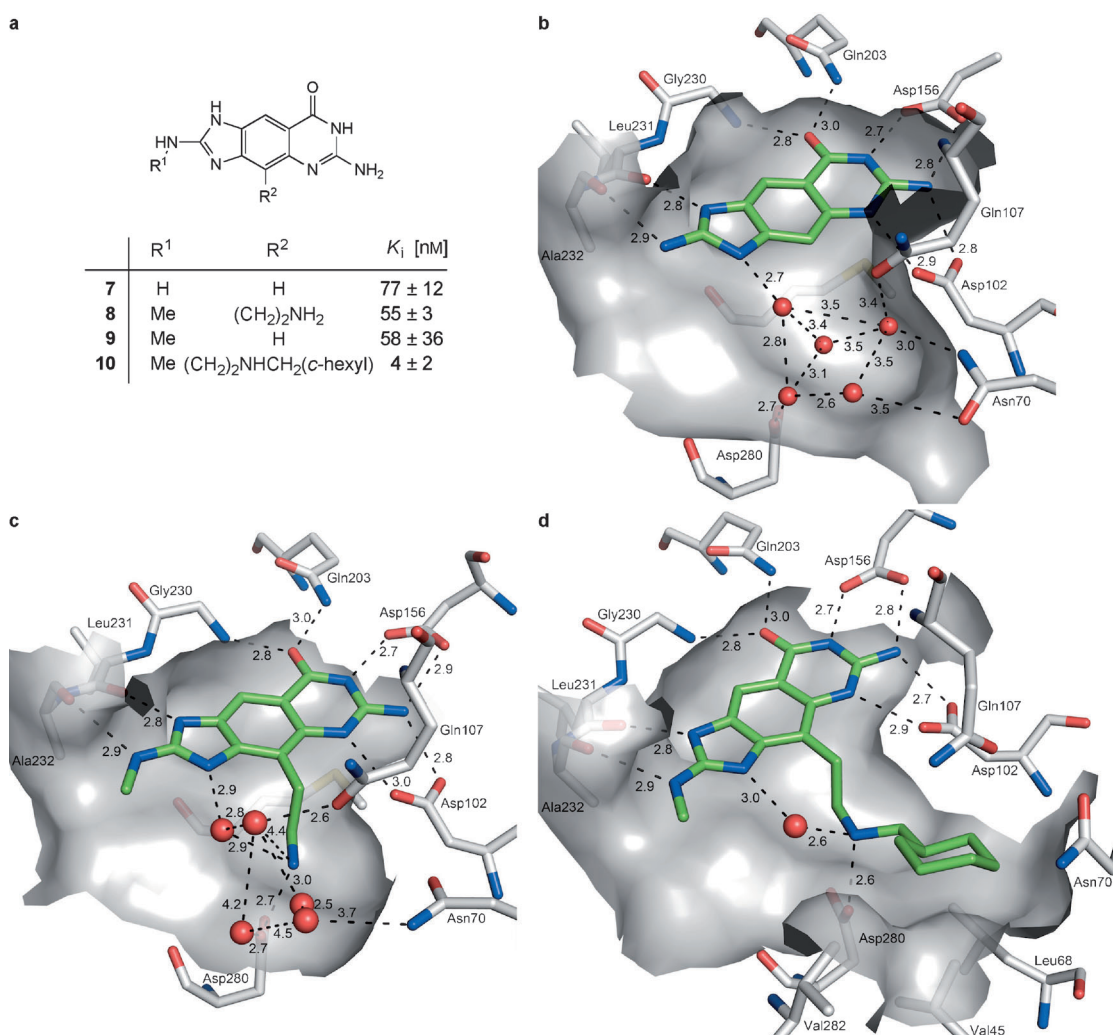


Figure 4. a) Selected *lin*-benzoguanine ligands for ZmTGT. b) Cocystal structure of ligand **7** bound to TGT (1.28 Å resolution, PDB ID: 2Z7K).^[46c] c) Cocystal structure of ligand **8** bound to TGT (1.4 Å resolution, PDB ID: 3GC5).^[49] d) Cocystal structure of ligand **10** bound to TGT (1.78 Å resolution, PDB ID: 3EOS).^[49] Color code: C_{enzyme} gray, O red, N blue, C_{ligand} green.

a C(4) substituent.^[49] Subsequent extension by sterically suitable hydrophobic residues provided highly potent ligands, such as the cyclohexylmethyl derivative **10** ($K_i = 4$ nM, Figure 4d). The combination of biological assay with X-ray cocystal structure analysis provided multiple evidence that the mentioned shallow hydrophobic cleft can be appropriately filled with substantial gain in the binding free enthalpy, provided secondary ammonium linkers are used to bridge across the original water cluster.^[50]

This study suggests that attempts aimed at passing through a stable water network, starting from one binding pocket to reach into a second, hydrophobic subsite require special attention to be energetically successful. Ligands need to ensure that the surrounding functional groups of the protein and residual water molecules of the initial cluster remain favorably solvated/networked in the presence of the linker. If the linker does not have the appropriate polarity, the free enthalpy gained by filling the second site, and even more, will all be paid as an energetic penalty for disrupting the initial water cluster.

A large decrease in inhibitory potency was found by the introduction of *lin*-hypoxanthines (Figure 5 a), which lack the exocyclic NH₂ group of the *lin*-benzoguanines, as ligands for TGT.^[50b,51] Three ionic hydrogen bonds are lost with ligand **11** ($K_i = 6500$ nM; compare to **9**: $K_i = 58$ nM, Figure 4 a): Asp156 only forms one such bond, while Asp102, which underwent two such contacts in the complex of **7** (Figure 4 b) and **9**, moves back to its position observed for the apo-enzyme structure. The pyrimidinone ring lacks the exocyclic amino group and is no longer sufficiently basic to be protonated and favorably interact with the carboxylate of Asp102.^[51] Asp102 now undergoes hydrogen bonding to the side chain NH₂ group of Asn70 and the backbone NH group of Thr71. As a result of this conformational change, a large void opens and a second, “six-water” cluster is imported, which is networked to the pyramidal “five-water” cluster near Asp280 (Figure 5 b). This additional solvation of the protein and ligand combined with the conformational changes in the protein cannot make up for the loss of three ionic hydrogen bonds, as reflected by the strongly reduced inhibitory potency. When the exocyclic NH₂ group of the *lin*-benzoguanines was

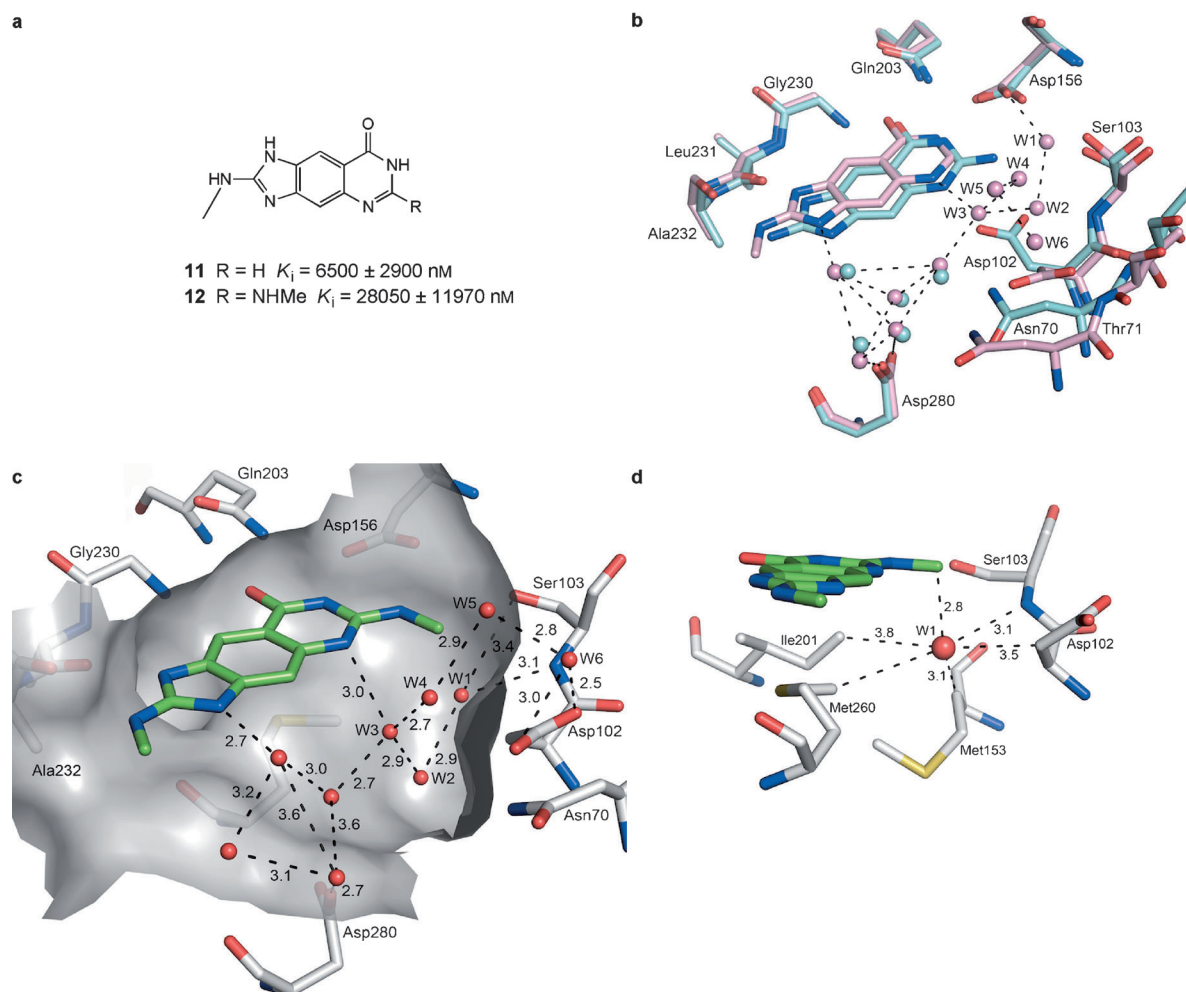


Figure 5. a) Examples of *lin*-benzohypoxanthine and *N*-alkylated *lin*-benzoguanine inhibitors of ZmTGT. b) Overlay of the cocrystal structure of ligands **7** (1.28 Å resolution, PDB ID: 2Z7K)^[46c] and **11** (1.82 Å resolution, PDB ID: 3S1G)^[51] bound to TGT. Color code: C_{ligand(7)} + C_{enzyme+water molecules(7)} cyan, C_{ligand(11)} + C_{enzyme+water molecules(11)} pink, O red, N blue. c) Cocrystal structure of ligand **12** bound to TGT (1.68 Å resolution, PDB ID: 3RR4)^[51] and d) water W1 of the newly imported water cluster in an energetically unfavorable hydrophobic environment.^[51] Color code: C_{ligand} green, C_{enzyme} gray, S yellow.

alkylated (Me, Et), a similar conformational shift of Asp102 occurred for steric reasons, and again a “six-water” network was imported (Figure 5c). However, the inhibitory affinity of the ligands was further decreased, as measured for *N*-methyl derivative **12** ($K_i = 28050$ nM). A more detailed examination of the interactions of the individual water molecules in the network showed that one of them (W1) becomes surrounded by the hydrophobic parts of the *N*-Me group of **12** and of Asp102, Ile201, and Met153 of the protein,^[51] which should be quite unfavorable (Figure 5d).

Clearly, the situation becomes quite complex with the conformational reorganization of the protein and the changes in the water network solvation occurring upon minor changes in the ligand structures.^[52] Without high-resolution structural information, a meaningful interpretation of the measured protein–ligand binding free enthalpies would not be possible, as Whalen and Spies also concluded from their work on glutamase racemase inhibitors.^[53]

Deeper insight into the role of water networks in contributing to the structure and thermodynamic signature

of protein–ligand complexes is increasingly obtained by application of multiple methods, including high-resolution cocrystallography, computational tools, pK_a determinations, mutations of key residues of the protein involved in ligand and water interactions, and ITC. This multidimensional approach has recently been applied to investigate in detail the complexation of *lin*-benzoguanines and *lin*-hypoxanthines to TGT.^[54] Whitesides and co-workers investigated the complexation of sulfonamide ligands to human carbonic anhydrase (HCA) as a model system and found that the binding free enthalpy of the protein–ligand complexes increases upon enlargement of the hydrophobic surface area of the ligand.^[55] However, differences in ligand affinity also stem from binding-induced changes in the number and organization of water molecules localized in the active site. Their study suggests that the water network in the active site has a major influence on the energetics of protein–ligand complexation. Different thermodynamic signatures for water replacement in the water networks are observed, depending on whether these solvent molecules solvate the free ligand,

the so-called “hydrophobic wall”, the enzyme surface, or the Zn^{2+} ion. The finding that the nature of the individual water molecules in the networks and their contribution to the overall thermodynamic signature of the binding process differs renders quantification of hydrophobic effects highly challenging.

The research groups of Klebe and Hangauer have for several years intensively applied a multidimensional analysis to the study of how ligand binding disrupts water networks in proteins, such as thrombin^[56] and thermolysin,^[57,58] and how this is reflected in the enthalpic and entropic changes in the complexation process. Overall, they came to similar conclusions about the challenging complexity of the protein–ligand–water systems as those drawn by the Whitesides group. In recent studies, they showed how subtle substrate changes—moving from methyl to butyl substituents on inhibitors of thermolysin—strongly affect the thermodynamic signature of the complexation process by affecting the surrounding water network.^[57c]

The comprehensive investigations on biological systems such as those by the Whitesides, Klebe, and Hangauer research groups illustrate that it is highly challenging to dissect contributions of individual water network changes to the overall thermodynamic signature of biological complexation process. Carefully designed synthetic host–guest studies, such as those described for cucurbit[7]uril (Section 2.1, Figure 1), certainly remain important for further elucidating and quantifying the nonclassical hydrophobic effect in chemical and biological systems.

3. Ligands for Filling Enzyme Pockets of Different Shape, Conformational Dynamics, and Polarity

Pockets at enzyme active sites and in proteins in general differ greatly in size, volume, degree of preorganization (conformational flexibility/rigidity), polarity, and polarizability.^[59] In this section, we report case studies on how ligands have been developed using structure-based design to fill different pockets with a substantial gain in binding affinity. Again, we start the section with synthetic model systems, since lessons learned from their study are directly applicable to biological systems.

3.1. Optimal Occupancy of Binding Pockets—Lessons from Model Systems

In his lock-and-key analogy, Emil Fischer already recognized in 1894 the importance of geometric complementarity between enzymes and their substrates.^[60] What is the ideal size and shape of a ligand for a given pocket in chemical or biological receptor systems?^[61] The precise determination of the volume of protein pockets is a demanding task due to the often ambiguous definition of their borders and the conformational flexibility of the surrounding protein, which additionally varies with the applied conditions. This is reflected in the large number of published methods for the identification and characterization of protein pockets.^[62] In

contrast, molecular capsules and containers offer geometrically well-defined cavities with known volumes for host–guest studies, whereby the thermodynamic profile of complexation can be determined by NMR binding studies or ITC. They are appropriate model systems to approach the question of the optimal ligand size. A study by Mecozzi and Rebek with several hydrogen-bonded supramolecular capsules of different internal cavity size led to the proposal of the 55 % rule, which states that the hydrophobic space of a host is energetically best filled with a lipophilic, shape-complementary guest at a packing coefficient (PC, ratio of guest volume to host volume) of 0.55 ± 0.09 .^[63] Interestingly, the packing density of many common organic liquids amounts to the same value.^[63] This finding was further corroborated by the complexation of longer *n*-alkanes and *n*-alkyl derivatives in the interior cavity of resorcin[4]arene cavitands, receptors introduced by Cram and co-workers in 1982.^[64,65] Although *n*-alkanes up to the length of hexadecane^[66] normally prefer a fully extended (all-*trans*) conformation, they adopt a helical form when they are bound to the cavitands, which causes a steric strain of 2–3 kcal mol^{−1}.^[67] The energetic costs for the coiled conformation of the alkyl chains are compensated by the proper filling of space (56 %), the burial of hydrophobic surface driven by dispersive interactions (in aqueous solution), and better C–H $\cdots\pi$ contacts.^[65] Recently, it was also demonstrated that longer (C_{13} and C_{14}) *n*-alkyl derivatives adopt a folded conformation to fit within a hydrophobic pocket of a water-soluble cavitand in water (D_2O).^[68]

Our investigations confirmed these findings for new container molecules. Resorcin[4]arene-based cavitands are highly dynamic systems, which can be switched by changes in temperature, pH value, Zn^{2+} ion concentration, or by redox processes between an open “kite” form, which is incapable of guest binding, and a closed “vase” form with an interior cavity for guest complexation.^[69] The introduction of diacetylene bridges of different length in **13a/b** (Figure 6a) between two diazaphthalimide wall flaps stabilizes the “vase” form and rigidifies the resulting molecular containers substantially relative to their open-top counterparts with four quinoxaline wall flaps, as visualized by molecular dynamics (MD) simulations (Figure 6b).^[70,71] The binding of cycloalkanes in the differently sized interior cavities of the two containers was greatly enhanced, with **13a** displaying the highest affinity for cyclopentane (PC 53 %, determined from MD simulations) and **13b** for cyclohexane (PC 54 %).

Mecozzi and Rebek earlier predicted that the optimal PC increases when polar contacts are established in confined environments.^[63] This was observed by ITC for the complexation of heteroalicyclic guests (a selection is shown in Figure 6c) by the *para*-xylylene-bridged cavitand **14** at 303 K in mesitylene, a solvent too large to compete for the interior cavity.^[72] The optimal PC for host–guest complexation increases to 0.63 ± 0.09 as a result of polar host–guest contacts, such as C–O \cdots C=O, N–H $\cdots\pi$, and S $\cdots\pi$, for which experimental evidence was obtained by rotating frame Overhauser effect spectroscopy (ROESY). The strengthening of C–H $\cdots\pi$ interactions as a consequence of the more polarized C–H bonds in the heterocyclic guests also contributes to the observed increase in PC.^[73] The strongest binding was

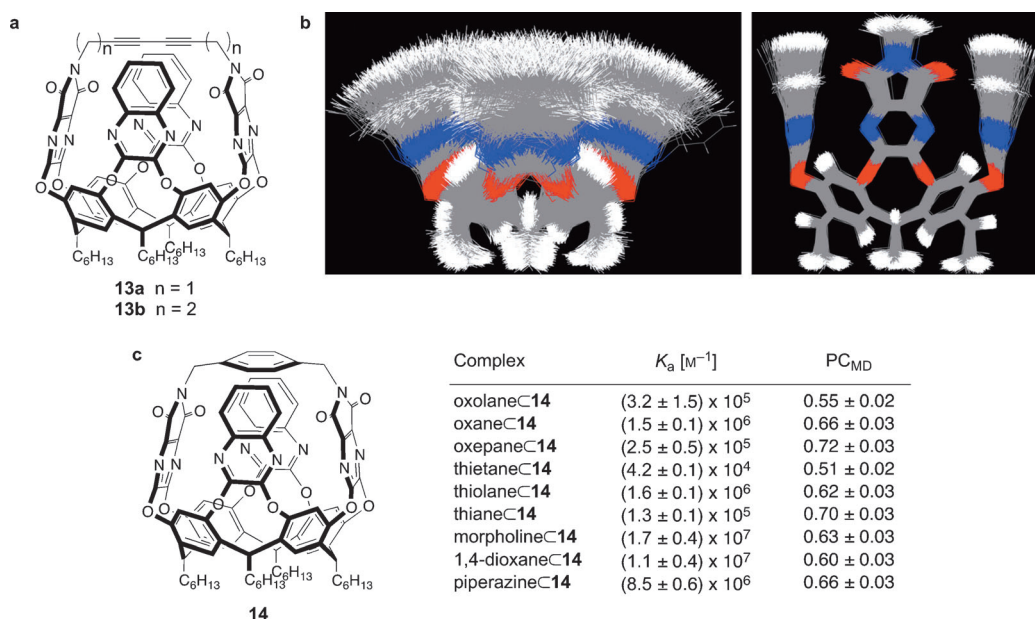


Figure 6. a) Chemical structures of diacetylene-bridged molecular container molecules **13 a/b**. b) MD simulations (1000 structures)^[71] showing the rigidification of cavitands by bridging. Left: “vase” form of an open resorcin[4]arene cavitant with four quinoxaline flaps. Right: Container **13 a**. In both cases, the hexyl substituents were substituted by methyl groups. c) *para*-Xylylene-bridged container **14**, association constants, and packing coefficients from MD simulations (PC_{MD}) of various heterocyclic guests, which reveals an optimal PC of 0.63 ± 0.09 . The image is reproduced with permission from the publisher.^[70b]

observed for six-membered ring heterocycles, with PC values between 0.60 and 0.66 (Figure 6c).^[72] Guest binding is promoted by both favorable enthalpic and entropic terms. The change in entropy is more positive for smaller guests, which lose less translational and rotational degrees of freedom upon complexation, whereas the enthalpic gains are larger for larger guests that can establish tighter contacts. The favorable entropic term was explained by the high degree of preorganization of the host, the gain in entropy of mixing upon complexation, and the ordered alignment of the unbound guests in the aromatic solvent mesitylene.^[72]

Completely surrounding the guests by the containers is essential for the strong binding affinity: a decrease in the association constants of up to five orders of magnitude was observed when the two quinoxaline flaps in **14** were removed^[74] or when the open-top cavitant with four quinoxaline flaps served as the host.^[75] Container **14** was rendered water soluble by the attachment of oligo(ethylene glycol) chains to the legs and the *para*-xylylene bridge, and the mentioned polar host–guest interactions were also observed with heterocyclic guests in D_2O/CD_3CN (2:1).^[74]

The studies of guest complexation in the highly confined space of supramolecular capsules and container molecules confirm that there is an optimal size for shape-complementary guests. For lipophilic guests, the optimal PC is around 55 %, but increases to around 63 % when additional polar interactions are established. Complete encapsulation is not only crucial for accurate determination of the cavity volume to calculate PC values, but also for high binding affinity. While the determination of pocket space and PC values is more

difficult in proteins, bell-shaped profiles^[72] obtained by plotting the binding free enthalpy against the calculated PC values provide valuable guidance to also reach optimal cavity occupancy in biological receptor systems.

3.2. Highly Polar Active Sites of the Enzymes IspE and IspF in the Non-Mevalonate Pathway of Isoprenoid Biosynthesis

The enzymes IspE, IspF, and the already mentioned IspD (Section 2.3) are catalysts in the non-mevalonate pathway of isopre-

noid biosynthesis (Figure 7).^[76] This pathway is exclusively used by eubacteria, such as *Mycobacterium tuberculosis* (Mt), and protozoa, such as *Plasmodium falciparum* (Pf), the causative agent of malaria. It is absent in humans, who use the mevalonate pathway for biosynthesis, while both pathways coexist compartmentalized in higher plants. The enzymes of the non-mevalonate pathway are recognized as promising targets in the development of new agents against infectious diseases.^[77] They have been validated as antimalarial targets through the finding that the antibiotic fosmidomycin and derivatives, which inhibit the second enzyme IspC, showed strong in vivo antimalarial activity.^[78] The substrates of the pathway are highly polar, with mono- and diphosphate groups. Accordingly, the respective enzymes feature highly polar active sites, which increase the challenge in designing potent inhibitors.

3.2.1. IspE—Boost in Potency by Filling a Small Hydrophobic Pocket

The kinase IspE catalyzes the phosphorylation of 4-diphosphocytidyl-2C-methyl-D-erythritol (CDP-ME, Figure 7), and its active site is divided into three main pockets: the cytidine binding pocket with the ribose sub-pocket, the methylerythritol (ME) binding site featuring a small hydrophobic cavity, and the ATP binding pocket with the glycine-rich loop for triphosphate recognition (P loop) (Figure 8a).^[79,80]

The proper filling of each of the three pockets proved to be challenging. While most kinase inhibitors bind to the

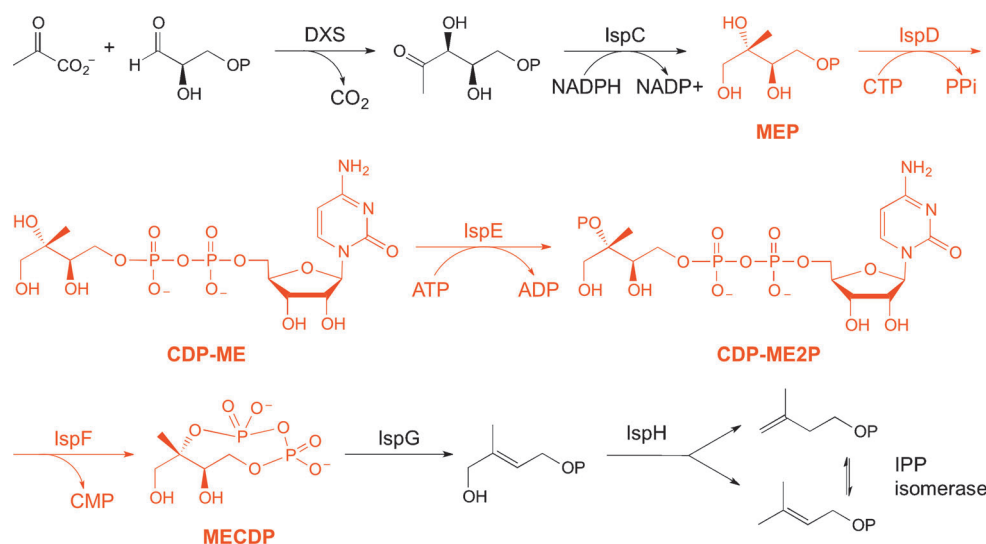


Figure 7. The non-mevalonate pathway of isoprenoid biosynthesis. Three of the seven enzymes involved are investigated by our research group (marked in red). IspD catalyzes the transfer of a diphosphocytidyl unit from cytidine triphosphate (CTP) to 2C-methyl-D-erythritol-4-phosphate (MEP), which results in 4-diphosphocytidyl-2C-methyl-D-erythritol (CDP-ME). CDP-ME is phosphorylated by IspE to 4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate (CDP-ME2P) and subsequently cyclized by IspF to 2C-methyl-D-erythritol 2,4-cyclodiphosphate (MECDP).

adenine pocket of the ATP binding site,^[81] this site is rather small in IspE, as ATP is bound in the unusual *syn* conformation (Figure 8a). Interest was, therefore, directed towards the cytidine binding pocket of the substrate. Whereas heterocyclic adenine replacements are diverse and abundant, as a result of the intense activity in the area of kinase inhibition, it is not the case for cytosine surrogates.^[82] This is despite the increasing interest in cytosine recognition originating from the important role of its methylation and demethylation at the 5-position in the field of epigenetics.^[83] We consistently observed a large loss in the activity of our ligands when we replaced cytosine by surrogates, such as 2-aminopyridines.^[84] Maintaining the cytosine ring and ligand design based on *Escherichia coli* IspE (*EcIspE*) in complex with CDP-ME and a stable ATP analogue (2.01 Å resolution, PDB ID: 1OJ4)^[80a] led to a series of potent inhibitors (±)-**15** to (±)-**23**, with activities (inhibitory constants K_i)^[85] reaching the upper nanomolar range (Figure 8b).^[86,87] The binding mode of the cytosine-based ligands was predicted by modeling using MOLOC^[88] (Figure 8c) and later confirmed by an X-ray cocrystal structure of *Aquifex aeolicus* IspE (*AaIspE*) bound to ligand **24**^[86b] (Figure 8d). This ligand featured an improved water solubility as a result of the introduction of the oxetanyl ring.^[89]

In the complex of *EcIspE* with (±)-**15**, the cytosine ring is sandwiched between the Tyr25 and Phe185 rings and undergoes favorable triple hydrogen bonding (DDA...AAD type; D = donor, A = acceptor)^[90] with the side chain and backbone N–H group and C=O group of His26 (Figure 8c).^[86a,b] The tetrahydrothiophene ring occupies the ribose subpocket and interacts with the Tyr25 ring. Control experiments showed that S... π interactions are not effective, but that favorable effects originate from the polarization of the C–H bonds by the heteroatom.^[84] Whereas replacement of the tetrahydro-

thiophene ring by a cyclopentyl ring analogue of (±)-**15** affords a K_i value of 1.5 μM , the tetrahydrofuran analogue ($K_i = 0.7 \mu\text{M}$) shows an affinity for *EcIspE* similar to that of (±)-**15** ($K_i = 0.29 \mu\text{M}$).^[84,86b]

The alkyne linker directs the sulfonamide moiety of (±)-**15** into the highly solvated region between the substrate and ATP binding sites, where it undergoes hydrogen bonding with the ion pair formed by Lys10 and Asp141. This is energetically favorable as N-methylation of the sulfonamide moiety in (±)-**22** ($K_i = 2.5 \mu\text{M}$) reduces the inhibitory activity by nearly a factor of 10 (Figure 8b). In its preferred conformation with the nitrogen lone pair of electrons bisecting the SO_2 group,^[3] which is

also adopted by the ligands in the free state,^[86a] the terminal alkyl residues are directed into a small, but tight and preorganized hydrophobic pocket lined by Phe185, Leu28, and Leu15. Only the entrance of this pocket is actually occupied by the methyl group of the methylerythritol moiety of the substrate (Figure 8a). The size of this pocket is about 100 Å³, and its occupancy exhibits a steep structure–activity relationship (SAR).^[86] Cyclopropyl derivative (±)-**15** is the best binder, and when the size of the alkyl residue is reduced in methyl derivative (±)-**21** or increased in cyclohexyl derivative (±)-**20**, nearly a factor of 10 is rapidly lost in the inhibitory affinity (Figure 8b). The 55 % rule by Mecozzi and Rebek^[63] (Section 3.1) also holds for this pocket, and a PC value of 56 % was calculated for the cyclopropyl residue.^[86b] A literature survey by Jorgensen and co-workers emphasizes the large gain in binding energy from filling small but confined lipophilic pockets with apolar substituents such as methyl groups.^[91]

Two additional findings should be mentioned. Energetically favorable filling of the highly water-exposed regions in the active site of IspE was revealed to be challenging. The ME-binding region and the cleft between the substrate and ATP sites are filled with a water cluster in both *EcIspE* and *AaIspE*. We prepared ligands similar to (±)-**15** but with sugar derivatives (ribose and mannose derivatives) oriented in that region to replace the water clusters, but a gain in binding free enthalpy was not achieved.^[92] Finally, we succeeded in preparing a second class of ligands that occupy the substrate binding site of *EcIspE*, with K_i values in the low micromolar range.^[86c] In compound (±)-**23** ($K_i = 6.9 \mu\text{M}$), an imidazole moiety directs the cyclohexyl ring into the discussed small hydrophobic pocket; the binding affinity is no longer measurable when the cyclohexyl ring is omitted from the imidazole ring.^[86c]

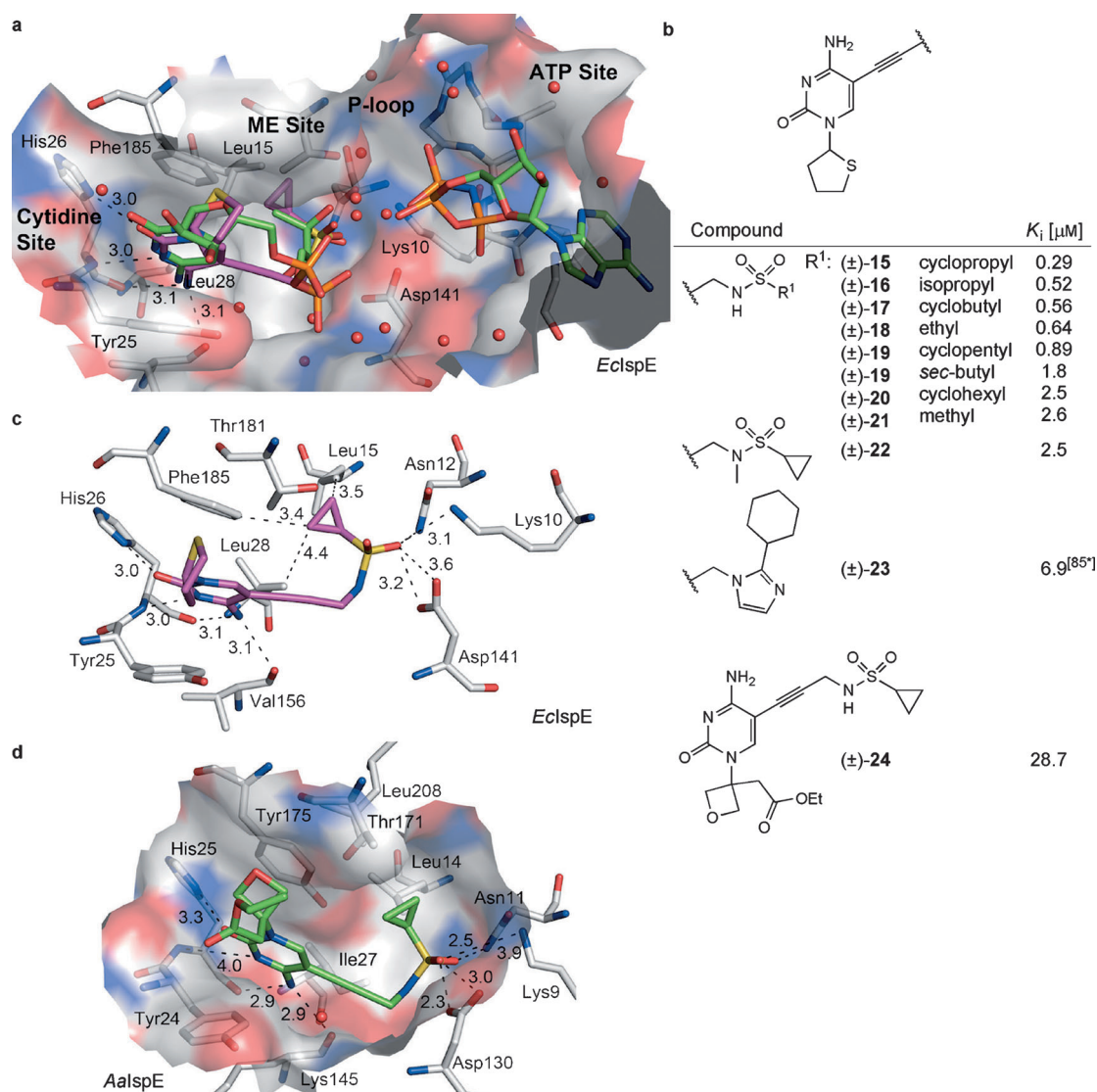


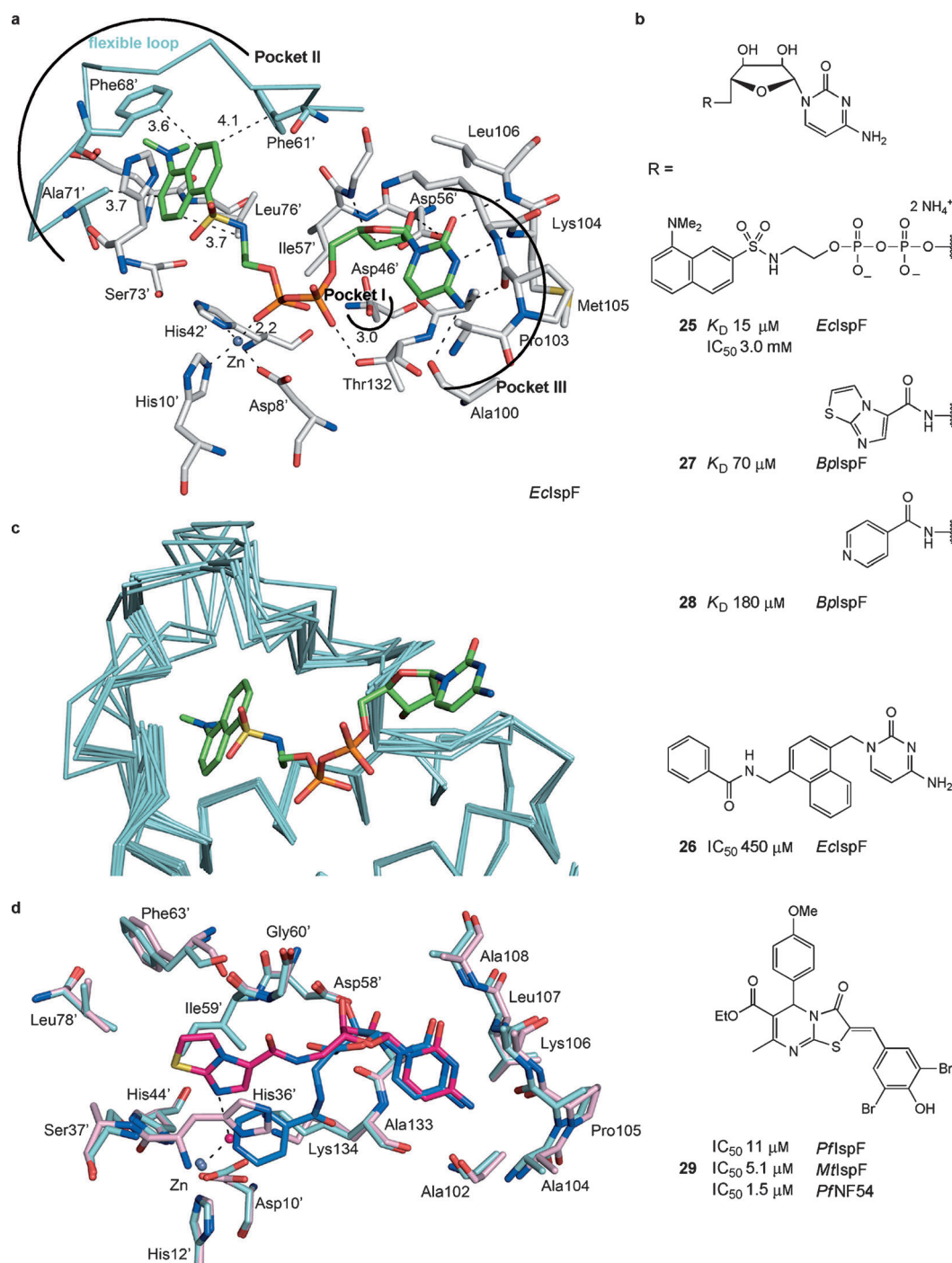
Figure 8. a) Active site of *Ecl*SpE in complex with CDP-ME and adenosine 5'-(β , γ -imino)triphosphate (AMP-PNP; 2.01 Å resolution, PDB ID: 1OJ4).^[80a] and overlap with the binding of (\pm)-**15** to *Ecl*SpE proposed by MOLOC. It can be seen that the substrate Me group does not properly fill the tight hydrophobic pocket, but the cyclopropyl group of the inhibitor does. Color code: C_{enzyme} gray, O red, N blue, C_{ligand(CDP-ME/AMP-PNP)} green, C_{ligand((\pm)-**15**)} pink. b) Selected *Ecl*SpE inhibitors. c) Binding mode of inhibitor (\pm)-**15** to *Ecl*SpE proposed by MOLOC (2.01 Å resolution, PDB ID: 1OJ4).^[80a] d) Binding mode of inhibitor (\pm)-**24** in complex with AalSpE, which occupies the hydrophobic subpocket in the ME-binding site (2.20 Å resolution, PDB ID: 2VF3).^[86b] Color code: C_{ligand((\pm)-**24**)} green.

The structure-based design of these first inhibitors of IspE showed that the small hydrophobic pocket, which hosts the methyl group of ME, is actually much larger than expected and that much binding free enthalpy is gained by its proper filling. The study also revealed the difficulty in replacing cytosine with energetic benefit, which is in contrast to the many adenine replacements in kinase inhibitors. Furthermore, the water cluster in the cleft between the substrate and ATP binding sites could not be properly addressed.

3.2.2. *IspF*—Lack of Energetic Gains when Binding to a Highly Flexible, Hydrophobic Enzyme Pocket

Targeting IspF, which catalyzes the cyclization of 4-diphosphocytidyl-2*C*-methyl-D-erythritol-2-phosphate (CDP-ME2P, Figure 7),^[76,93] proved to be even more chal-

lenging. The enzyme exists as a C_3 -symmetric homotrimer, and substrate binding occurs at the three protein interfaces involving amino acids of two neighboring proteins. We designed the first inhibitors, such as **25**, using the cocrystal structure of CDP and 2C-methyl-D-erythritol 2,4-cyclodiphosphate, each in complex with *Ec*IspF (PDB IDs: 1GX1 and 1JY8).^[94,95] The ligands, which we intended to use in the development of an assay, were substrate analogues^[96] and featured a cytidine diphosphate moiety, to bind to pocket III (Figure 9a), and a fluorescent probe, such as a dansyl [5-(dimethylamino)naphthalene-1-sulfonyl] residue (in **25**, Figure 9b) for occupancy of pocket II.^[97] Two cocrystal structures were obtained, which revealed that pocket II is highly flexible and can adopt very different conformations, which explains the limited gain in binding free enthalpy resulting from its occupancy. The exceptionally high conformational



flexibility of the loop shaping pocket II has also been observed in various other crystal structures,^[95] including the recent example of *PflspF*.^[98] An overlay of ten *EclspF* structures (Figure 9c) clearly shows the high conformational

flexibility of this loop.^[99] Occupation of such a pocket may yield enthalpic gains but are fully compensated by entropic losses resulting from freezing out one of the numerous pocket conformations. In contrast, pocket III is highly conserved in all structures. Much effort was invested in vain to improve the occupancy of pocket II.^[100] On the other hand, we showed that the ribose moiety^[101] and the diphosphate linker in **25** can be favorably replaced by a more hydrophobic naphthalene-1,4-diyl spacer, such as in ligand **26**. It is postulated that **26** binds with its termini in the two pockets II and III, with an IC_{50} value of 0.45 mM (*EclspF*; Figure 9b).^[100] This value is difficult to improve upon because of the conformational flexibility of pocket II and the lack of more potent cytosine surrogates.

Recent ligand developments in other laboratories have consequently taken a different approach, targeting the cytidine binding site (pocket III) and coordination to the conserved Zn^{2+} ion (Figure 9d), which is bound to three amino acid side chains (Asp10', His12', and His44') of the enzyme.^[102] Compounds **27** and **28** are examples of such ditopic binders, for which cocrystal structures with *IspF* from *Burkholderia pseudomallei* (*Bp*) were solved (Figure 9d). The binding of the cytidine moiety is identical in the two types of ligands

(pockets II/III and pocket III/ Zn^{2+} binders), with the cytosine ring undergoing triple hydrogen bonding to the conserved peptide sequence Met105 (N–H, *Ec* labeling)/Lys104 (backbone N–H)/Pro103 (C=O) (Figure 9a). The ribose interacts through its 2'-OH and 3'-OH groups with the side chain of Asp56'.

Binding affinity to the Zn^{2+} ion, which can be replaced by cations such as Mn^{2+} or Co^{2+} ,^[95] is best achieved by aromatic nitrogen donors, as revealed by a fragment-based approach.^[102b,c] The aromatic N-heterocycles in **27** ($K_D = 70 \mu\text{M}$) and **28** ($K_D = 180 \mu\text{M}$) give overall similar or weaker ligand affinities to *Bp*IspF^[102c] relative to CDP, which binds to pocket III of *Bp*IspF and the Zn^{2+} ion with a K_d value of $75 \mu\text{M}$.^[102a] Whereas compound **28** coordinates directly to the zinc(II) ion, ligand **27** binds to a water molecule at the fourth coordination site (Figure 9d).

Overall, effective solutions for binding to IspF enzymes have not yet been found using structure-based design. The cytidine binding site (pocket III) is the most reliable to have been addressed in terms of a gain in the binding free enthalpy, whereas the highly flexible pocket II is not very suitable for energetically favorable occupation. Binding to the Zn^{2+} ion

has been reliably established, but again without a breakthrough in gaining affinity. Large compound library screening seems to be highly desirable in the case of IspF inhibition. Indeed, screening a 40000 compound library identified thiazolo[3,2-*a*]pyrimidines, such as **29**, as inhibitors of the IspF enzymes of *Pf* and *Mt*, with IC_{50} values in the low micromolar range and good cell-based activity against *Pf* strains NF54 (Figure 9b).^[103] Their binding mode, however, remains unknown.

3.2.3. Very Large Enzyme Active Sites—Trypanothione Reductase

Protozoan parasites of the trypanosomatid family exhibit a special thiol redox metabolism that differs from the mammalian system and involves the essential flavoenzyme trypanothione reductase (TR).^[104] Members of this family are the causative agents of human African trypanosomiasis, Chagas disease, and the various forms of leishmaniasis, and compounds that inhibit the TR could be used as potential treatments with a new mechanism of action.^[105]

Trypanothione reductase is characterized by a large, negatively charged active site with dimensions of approxi-

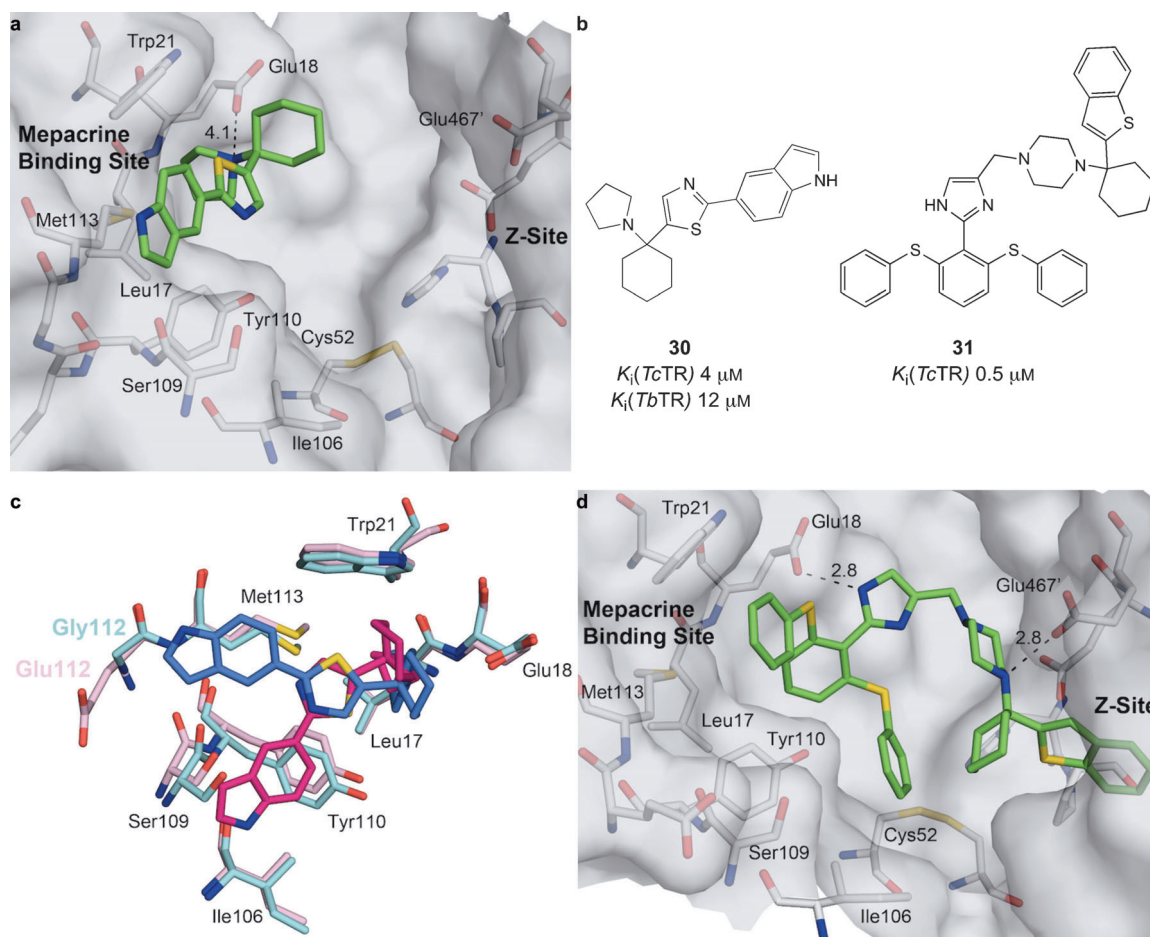


Figure 10. a) Overview of the large active site of *Tb*TR in complex with BTCP analogue **30** (2.50 Å resolution, PDB ID: 4NEV).^[109c] Color code: C_{enzyme} gray, O red, N blue, S yellow, C_{ligand} green. b) Chemical structure of BTCP analogue **30** and diaryl sulfide based ligand **31**. c) Overlay of the X-ray cocrystal structures of inhibitor **30** in complex with *Tb* and *Tc*TR (2.80 Å resolution, PDB ID: 4NEV).^[109c] Color code: *Tc*TR: C_{ligand} hot pink, C_{enzyme} pink; *Tb*TR: C_{ligand} dark blue, C_{enzyme} cyan. d) Binding mode of diaryl sulfide based ligand **31** to *Tc*TR, as modeled with MOLOC (2.40 Å resolution, PDB ID: 1BZL).^[107b]

mately $22 \times 20 \times 28 \text{ \AA}^3$ (Figure 10a).^[106] The negative charge is introduced by the essential side chain of Glu18 and has been revealed to be important for substrate and inhibitor recognition.^[107] Successful inhibition of TR is mostly achieved by ligands with a large molecular size, and the formation of protein–ligand complexes with a 1:2 stoichiometry is frequently observed.^[108] A high-throughput screening revealed that 1-(1-benzo[*b*]thiophen-2-yl-cyclohexyl)piperidine (BTCP) with a volume of about 366 \AA^3 and its analogues are potent, small-molecule TR inhibitors, which are capable of competing with the doubly large trypanothione disulfide substrate (TS_2 ; ca. 817 \AA^3).^[109] This lead has been the focus of further investigations.^[108d,109c]

Mutation studies and cocrystal structure analysis with inhibitor **30** showed that this ligand class binds to the hydrophobic wall at the mepacrine-binding site^[110] and that the protonated tertiary amine interacts with Glu18 through Coulombic interactions (Figure 10a,b). Their contribution to binding was estimated from mutation studies to be about $-0.7 \text{ kcal mol}^{-1}$; at an $\text{N}\cdots\text{O}$ distance of 4.1 \AA , a charge-assisted hydrogen bond is unlikely.^[111] Cocrystal structures of inhibitor **30** in complex with TR from the species *Trypanosoma brucei* (*Tb*) and *T. cruzi* (*Tc*) revealed slightly different binding modes ($K_i(\text{TcTR}) = 4 \text{ }\mu\text{M}$; $K_i(\text{TbTR}) = 12 \text{ }\mu\text{M}$; Figure 10b,c). The biaryl moiety takes up two different orientations: probably the mutation of Gly112 in *TbTR* to Glu112 in *TcTR* forces the indole ring into a new position remote from the highly solvated Glu112 side chain (Figure 10c). The key protein–ligand interactions in the *TbTR* structure, besides the already mentioned Coulombic ion pairing, are $\text{C}\cdots\text{H}\cdots\pi$ contacts between Trp21 and the cyclohexyl and pyrrolidine moieties, $\text{S}\cdots\pi$ ($d(\text{S}_{\text{Met113}}\cdots\text{thiazole}) = 4.0 \text{ \AA}$), $\text{SC}\cdots\text{H}\cdots\pi$ ($d(\text{SCH}_{2\text{Met113}}\cdots\text{indole}) = 3.7 \text{ \AA}$), and $\text{S}\cdots\text{S}$ ($d = 4.1 \text{ \AA}$) interactions (Figure 10c). In the *TcTR* structure the indole ring establishes dispersive contacts with Ile106 ($d(\text{C}\cdots\text{C}) = 3.9 \text{ \AA}$) and short $\text{O}\cdots\text{H}\cdots\pi$ interactions ($d(\text{O}\cdots\text{C}) = 2.8 \text{ \AA}$) with the side chain of Ser109. The large active site of TR allows a multitude of possible ligand orientations, not only for BTCP analogue inhibitors,^[108f] and therefore, the binding mode is susceptible to minor changes in both the protein and inhibitor. Noteworthy in this context is a study by Patterson et al., who showed by cocrystal structure analysis that a series of dihydroquinazoline inhibitors are capable of inducing a conformational change to the side chain of Met113 to open up a new pocket for the ligands.^[112] This movement of Met113 was not observed in our investigation or in any other crystal structure published so far.

Since the binding site of TR is a solvent-exposed groove rather than a cavity, the interaction surface between the protein and ligand is comparatively small. Although larger molecules are less “druglike”, they can establish more interactions with the protein surface and, as a result, their binding affinity is increased. Indeed, the conjugation of our diaryl sulfide based inhibitors^[107c,113] with the BTCP motif led to ligands with K_i values against *TcTR* in the sub-micromolar range (K_i (**31**; *TcTR*) = $0.5 \text{ }\mu\text{M}$; Figure 10b).^[108d] Figure 10d shows a proposal for the binding of **31** in the enzyme active site, originating from a modeling study using MOLOC and based on the cocrystal structure of TS_2 bound to *TcTR* (PDB

ID: 1BZL).^[107b] The ligand efficiency (LE) for **31** (LE $0.18 \text{ kcal mol}^{-1}/\text{non-hydrogen atom}$) decreased relative to **30** (LE $0.30 \text{ kcal mol}^{-1}/\text{non-hydrogen atom}$), which is a general trend observed for larger molecules; this value should be improved in future studies.^[114] Possibly even more important for drug development is the recently introduced ligand-lipophilicity efficiency (LLE;^[115] also known as lipophilic efficiency (LipE)),^[116] which accounts for the efficiency of lipophilic interactions and is often inversely proportional to pharmacokinetic properties such as clearance, solubility, toxicity, and side effects.

Our findings show that the possibility of various binding modes of ligands in a large enzyme active site should be considered carefully when structure-based design is applied. It might be inevitable that the contact area between the protein and ligand must be increased beyond the usual size of small-molecule leads to enhance the binding affinity.

3.2.4. Deep and Rigid Pockets of Reduced Polarity—Factor Xa

The well-studied enzyme factor Xa is a serine protease from the human blood coagulation cascade and the target of several anticoagulants in therapeutic use or clinical trials.^[117] The key pockets at the active site of this enzyme are the two orthogonally aligned S1 and S4 pockets, which are occupied by all the potent ligands, such as (\pm) -**32** ($K_i = 280 \text{ nM}$) and (\pm) -**33** ($K_i = 9 \text{ nM}$, Figure 11), prepared in our laboratory.^[118] These pockets are highly shape-persistent and, therefore, particularly well-suited for detailed molecular recognition studies (Figure 11a). The S1 pocket is lined by two planar peptide backbone segments (Trp215–Gly216 and Gln192–Cys191–Ala190), thus providing space for sandwiching flat, (hetero)aromatic ligand parts. The S4 pocket features an aromatic box, which is shaped by the side chains of Tyr99, Phe174, and Trp215; the fourth face of the box is open and solvent-exposed. Our ligands feature a tricyclic core^[6b] from which orthogonal vectors depart towards the S1 and S4 pockets (Figure 11b).

A large contribution to the binding affinity is gained from cation $\cdots\pi$ interactions in the aromatic box of the S4 pocket.^[11a,d,14e,18b] Replacement of the quaternary trimethylammonium cation in (\pm) -**32** and (\pm) -**33**, which binds in the center of the box (Figure 11a,c), by an isosteric *tert*-butyl group led to a decrease in the binding free enthalpy of $2.5\text{--}2.8 \text{ kcal mol}^{-1}$. These results quantify the energetics of quaternary ammonium cation $\cdots\pi$ interactions as $0.8\text{--}0.9 \text{ kcal mol}^{-1}$ per aromatic ring.^[118] Stepwise demethylation of (\pm) -**33** to the respective protonated tertiary, secondary, and primary ammonium ions results in the binding affinity being reduced by $1.2\text{--}1.8 \text{ kcal mol}^{-1}$ in each demethylation step.^[18b,118b] Similar results have been reported by Dougherty^[11d] as well as Waters and co-workers,^[120] who investigated cation $\cdots\pi$ interactions using cyclophanes and β -hairpin peptide model systems, respectively.

The concave subpocket S1 is a deep and rigid protein cavity and exhibits different surface polarities. The bottom of the pocket is characterized by the polar side chains of Asp189 and Tyr228 and the lipophilic residues Ala190, Val213, and Gly226 (Figure 11a,c). The phenylamminium motif in (\pm) -**32**

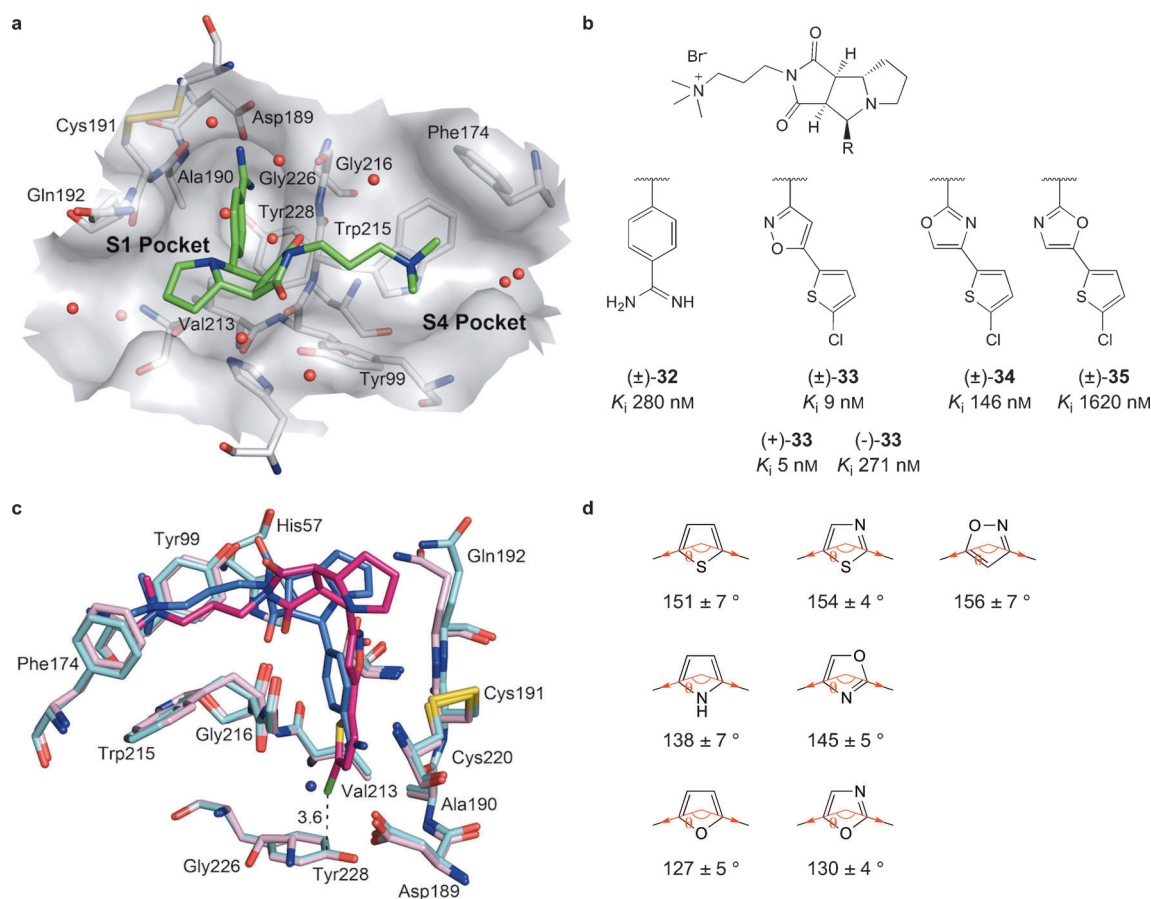


Figure 11. a) X-ray cocrystal structure of phenylamidinium-substituted ligand **32** in complex with factor Xa (1.64 Å resolution, PDB ID: 2BOK).^[118a] The ligand is L-shaped, occupies the S1 pocket with the phenylamidinium moiety, and binds with the ammonium ion in the aromatic box shaped by Tyr99, Phe174, and Trp215 in the S4 pocket. Color code: C_{enzyme} gray, O red, N blue, S yellow, C_{ligand} green. b) Selected inhibitors of factor Xa with different substituents for the deep and rigid S1 pocket. c) Overlay of the X-ray cocrystal structures of inhibitor **32** and **33** in complex with factor Xa (**33**: 1.25 Å resolution, PDB ID: 2JXH).^[119c] The chloro substituent of the heterobiaryl replaces the structural water at the bottom of the S1 pocket and establishes an orthogonal dipolar interaction with Tyr228. Color code: C_{ligand(33)} hot pink, C_{enzyme(33)} pink, C_{ligand(32)+water molecule} dark blue, C_{enzyme(32)} cyan. d) Means and standard deviation of a CSD search for the exit vectors of the most common five-membered heteroaromatic rings.

was utilized in earlier drug discovery research for binding in the S1 pocket through charged hydrogen-bonding interactions with the carboxylate group of Asp189.^[118a] It is now frequently substituted by halogenated heteroarenes, the hallmark of second-generation factor Xa inhibitors.^[121] We observed a gain in the binding affinity of $-\Delta\Delta G = 2.0$ kcal mol⁻¹ upon replacement of the phenylamidinium substituent of (±)-**32** by the chloroheterobiaryl motif in (±)-**33** (Figure 11b).^[118b]

The chloro substituent in (±)-**33** makes a substantial contribution to binding: the inhibitory potency is reduced by a factor of 68 upon its removal.^[119,122] Cocrystal structure analysis revealed multiple origins for the role of the Cl atom, which adopts a similar position with respect to Tyr228 in all the potent inhibitors with chloroheteroaryl substituents.^[118b,119] A comparison with the apo structure of factor Xa^[123] suggests that it replaces a water molecule to undergo a favorable orthogonal dipolar interaction with the phenolic C–O dipole of Tyr228 (Figure 11c).^[23] It also features close van der Waals contacts with the lipophilic amino acids

Ala190, Val213, and Gly226 that line the bottom of the pocket. The importance of the chloro substituent—similar in size to a water molecule—was also investigated in a study by Kuhn et al., who rationalized tight ligand binding through the establishment of multidimensional interaction networks.^[124] They implemented a “small world network” description^[125] into an empirical scoring function to reveal the effect of the local cooperativity of noncovalent interactions on the binding and described the binding affinity as not being simply a sum of single interactions. It was found that high network scores were often obtained for deeply buried ligand atoms that undergo several favorable interactions and that the omission or replacement of these atoms leads to a substantial decrease in ligand potency. Such a high network score was also obtained for the chlorine atom in the S1 pocket of factor Xa inhibitors, which emphasizes our experimental findings that this chloro substituent contributes to binding through several interactions with the protein.^[124]

Furthermore, inhibitors (±)-**33–35** gain binding energy by stacking through their heterobiaryl systems with the planar

peptide backbone segments Trp215-Gly216 and Gln192-Cys191-Ala190 in the S1 pocket (Figure 11c).^[119] Depending on the choice of the five-membered heteroaromatic linker between the tricyclic core and the chlorothienyl ring, K_i values are spread over more than two orders of magnitude. Different conformational preferences of the two heterocyclic rings as a result of interactions of the lone pairs of electrons certainly need to be taken into account.^[3,126] In addition, however, we discovered a remarkable difference in the angle between the two exit vectors in the five-membered heteroaromatic rings.^[119] A CSD search unexpectedly revealed that the mean values of the found exit vectors of the most common five-membered heteroaromatic rings vary between 127 and 156° (Figure 11d, see the Supporting Information, Figures S1–S7). Fujita and co-workers showed that these differences in exit vectors also have a considerable effect on the structural preference in metallosupramolecular self-assembly.^[127] In addition, the arrangement of the dipoles of the heteroarene and the stacking peptide amide bonds is of importance for binding strength. A computational study found that an antiparallel orientation of the dipoles of heteroarenes and stacking peptide bonds is clearly favored over a parallel arrangement (see Section 4.3).^[128] Dipolar interactions had previously been found to influence the stacking of substituted phenyl rings on flavins.^[129]

The investigation of the molecular recognition at the active site of factor Xa,^[119] with its two rigid and concave S1 and S4 pockets of reduced polarity provides good examples of how single-atom replacements can dramatically affect binding affinity. Changing from $\text{Me}_3\text{N}^+\text{CH}_2^-$ to $\text{Me}_3\text{CCH}_2^-$ in the S4 pocket substituent reduces the affinity by up to a factor of 100, while substitution of the ideally sized Cl atom by a H atom on the thienyl ring filling the S1 pocket leads to a reduction by up to a factor of 68. Furthermore, the investigations with our ligands revealed a surprisingly large difference in the exit vector angles of the five-membered heteroaromatic rings and initiated a computational search for the role of dipolar interactions in the π stacking between the heteroarene and peptide bond, which is now being followed up with experimental study.

3.2.5. Lipophilic Pockets in Aspartic Proteases—Plasmepsins

Aspartic proteases are key targets for the development of drugs for various illnesses.^[130] Inhibition of renin led to antithrombotic drugs,^[131] various HIV protease inhibitors are

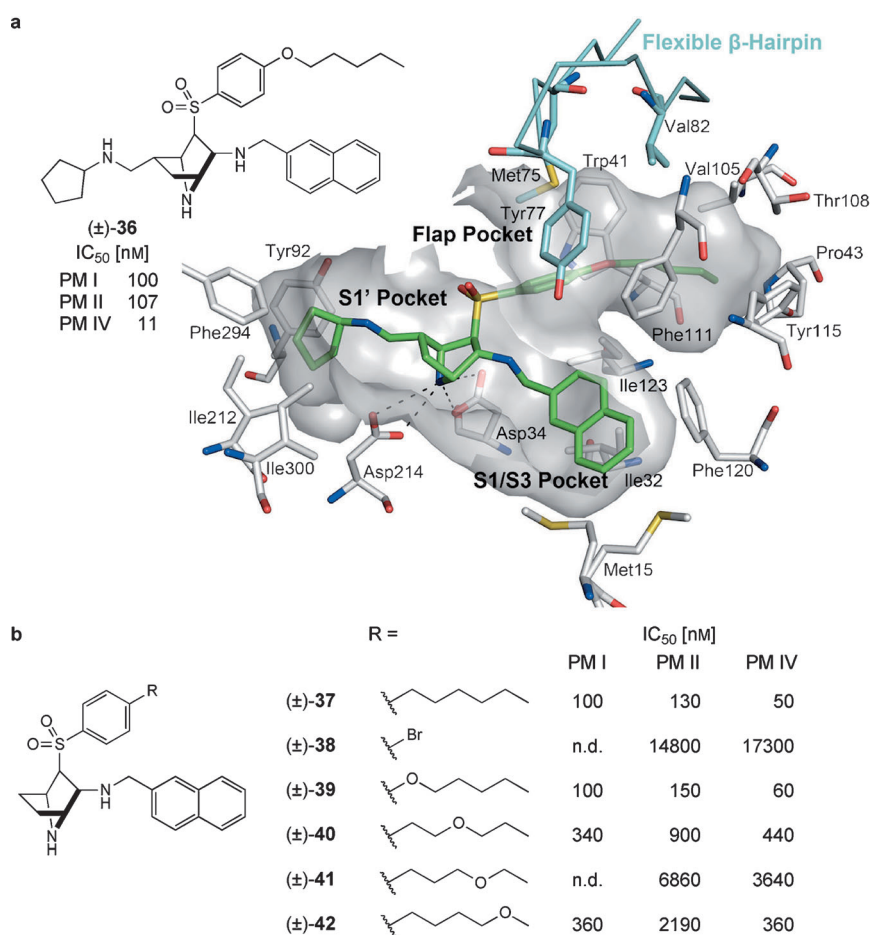


Figure 12. a) Binding mode of *exo*-3-amino-7-azabicyclo[2.2.1]heptane-based ligands of *Pf* PM II depicted with inhibitor **36**, as modeled with MOLOC (1.56 Å resolution, PDB ID: 2BJU).^[137] The monoprotonated 7-azabicyclo[2.2.1]heptane is recognized by the catalytic dyad (Asp34, Asp214) and offers three vectors to target the S1', S1/S3, and the flap pocket. Color code: C_{enzyme} gray, O red, N blue, S yellow, C_{ligand} green, C_{flexible loop} cyan. b) Selected “two vector ligands” of PM I, II, and IV with different substituents on the flap pocket. The introduction of an ether O atom into the alkyl chain decreases the binding affinity with a clear position dependency.

part of the multicomponent drug therapy against HIV infections,^[132] and β -secretase inhibitors have moved to clinical trials as anti-Alzheimer drugs.^[133] The aspartic proteases plasmepsins I, II, and IV have been recognized as potential drug targets in the fight against the malarial parasite *Pf*.^[134]

Together with cysteine proteases and the falcipains (see Section 3.2.6), plasmepsins (PMs) are involved in the degradation of human hemoglobin. As a result of overlapping substrate specificity, it is presumed that all of them need to be inhibited by a potential antimalarial drug.^[135] Among these aspartic proteases, PM II is the best studied enzyme by X-ray crystallography.^[136] The active site is characterized by the hydrophilic aspartic dyad Asp34-Asp214 and a rather large, hydrophobic site expanding from the S1/S3 pocket to the S1' pocket (Figure 12a). Commonly, inhibitors are anchored with a transition-state analogue or a protonated ammonium ion to the catalytic Asp dyad.^[135,137] Bur and co-workers showed that even two molecules with a C₃₇ backbone can bind together to the active site.^[137] Additionally, binding of appropriately

functionalized inhibitors can lead to the opening of a so-called flap pocket (Figure 12a).^[137,138] The ligands induce a substantial movement of the long and flexible β -hairpin consisting of Lys72-Phe85, and the rotation of Trp41 unlocks the entrance to the lengthy, apolar pocket (Figure 12a).

Our group introduced protonated, mildly basic *exo*-3-amino-4-*endo*-arylsulfonyl-7-azabicyclo[2.2.1]heptanes as the anchoring group to the catalytic dyad (Figure 12).^[138e] This central scaffold provides appropriate vectors for the attachment of substituents that target the flap and S1/S3 pockets. The resulting ligands show high potency in the nanomolar IC₅₀ range against PM I, PM II, and PM IV. Optical resolution of one of the racemic ligands led to impressive chiral recognition, with a difference in the IC₅₀ values of up to a factor of 3390, being observed for binding to PM IV.^[138e] Whereas various aromatic residues can fill the spacious S1/S3 pocket (an ideal occupancy has not yet been worked out), we focused our attention on the proper filling of the flap pocket. Optimal filling of this pocket obeys the 55 % rule of Mecozzi and Rebek,^[63,139] and *n*-pentyloxy, *n*-hexyl, or at best *n*-heptyl chains in the *para* position of the sulfonylaryl ring of the ligands give the highest binding affinity. These substituents bind strain-free in the all-*anti* conformation and fill the lipophilic pocket entirely.^[139] The search for the optimal volume occupancy of the flap pocket by the fully buried alkyl substituents and their conformation in the bound state was discussed in detail in a previous review,^[59] and here we only report new data obtained with ligands (\pm)-**36** to (\pm)-**42** (Figure 12).

Recently, we developed synthetic methods that enabled the introduction of an additional vector in the 5-*exo*-position of the azanorbornane scaffold to reach into the S1' subpocket of the active site.^[140] This is depicted in the binding model for ligand (\pm)-**36** developed on the basis of the X-ray cocrystal structure of PM II bound to a nonpeptidic ligand (1.56 Å resolution, PDB ID: 2BJU; Figure 12a).^[137] This smaller pocket is lined by Phe294, Tyr92, Ile212, and Ile300. Most of the members of the new "three-vector ligand" series, related to (\pm)-**36**, inhibit the three PM enzymes with potencies in the nanomolar range. The binding affinity to PM I and PM II was generally unchanged or only weakly enhanced relative to the "two-vector ligands" such as (\pm)-**37** and (\pm)-**39** (Figure 12b).^[139,140] In contrast, the inhibition of PM IV was substantially boosted, as illustrated by the comparison of the IC₅₀ values of (\pm)-**36** (11 nM) and (\pm)-**39** (60 nM). 5-*exo*-Cycloalkylaminomethyl substituents, as in (\pm)-**36**, were found to fill the S1' pocket best; X-ray structural information is, however, not yet available.

In another study with ligands (\pm)-**37** to (\pm)-**42**, we investigated the effects of placing polar atoms into the *n*-alkyl chains filling the lipophilic flap pocket.^[139b] Filling this pocket is essential for ligand potency. If the *n*-hexyl substituent of (\pm)-**37** is replaced by a bromine atom in (\pm)-**38**, binding affinity is decreased by a factor of about 110 for PM II and 350 for PM IV, which corresponds to a loss in binding free enthalpy of approximately 2.8 and 3.5 kcal mol⁻¹, respectively.^[138e] The insertion of an ether functionality at different positions of the alkyl chain led to a large decrease in activity in the series of ligands (\pm)-**39** to (\pm)-**42** and revealed

a clear position dependency (Figure 12b).^[139b] The reduced affinity can be explained by conformational effects, higher desolvation costs, and repulsion between the oxygen lone pairs of electrons and the π surface of Trp41. A similarly large decrease in affinity was observed when a hydroxy group was introduced at the end of the chain.^[139b] Furthermore, the introduction of a rigid linker segment led to a decreased binding affinity; similar observations upon stiffening ligand structures for biological complexation have been explained by enthalpy–entropy compensation effects.^[141]

Finally, in another approach, we showed that five- and six-membered ring alicyclic α,α -difluoroketone hydrates, decorated with flap and S1/S3 pocket-filling substituents similar to those in (\pm)-**37**, are good transition-state-analogue motifs for binding to the catalytic dyad of the plasmepsins, and presumably other polar enzyme active sites.^[142]

The studies on the plasmepsins have further validated the empirical rule by Mecozzi and Rebek for the optimal filling of confined lipophilic pockets, such as the ligand-induced flap pocket in these enzymes. The introduction of single ether O atoms into the alkyl chains that fill this pocket leads to a strong decrease in the inhibitory potency. The azanorbornane needle for binding to the catalytic Asp dyad has been further decorated by an additional vector to fill the S1' pocket, with substantial gain in the binding potency, in particular for PM IV. Clearly, designing nonpeptidic ligands based on crystal structures of PM II provides a valid route to generate ligands for the entire family of vacuolar plasmepsins, even for those where structural information in the flap-open conformation is not yet available.

3.2.6. Reversible Covalently Bound Inhibitors of Cysteine Proteases

The cysteine proteases rhodesain and falcipain-2 are potential targets against neglected diseases.^[143] Both proteins belong to the Cathepsin L-like subfamily and are utilized among others for degradation of the host protein by their organisms, the pathogen of human African trypanosomiasis, *Tb* (rhodesain), and the malarial parasite *Pf* (falcipain-2).^[144] The enzymes are essential to the parasites, and cysteine protease inhibitors have been shown to kill *Tb* both in cell culture and in animal models.^[145] These findings open up the possibility to develop chemotherapies for these parasitic diseases with novel modes of action.

The active sites of the two enzymes (rhodesain and falcipain-2) bear strong similarities, and most of our ligands are potent against both to a similar extent.^[146–148] We mainly report here our results for the inhibition of rhodesain. The active site of rhodesain can be best described as a solvent-exposed cleft with the three subpockets S1, S2, and S3 (Figure 13a). Most of the rhodesain inhibitors known from the literature undergo a reversible or irreversible covalent reaction with Cys25 of the catalytic dyad Cys25-His162 and thereby inhibit the enzyme.^[144a] We decided to employ in our ligands a nitrile group, which forms a reversible covalent thioimide with the side chain of Cys25, stabilized in the oxyanion hole (Gln19; Figure 13a). By utilizing a convenient synthesis and starting from trichloro-1,3,5-triazene, we first

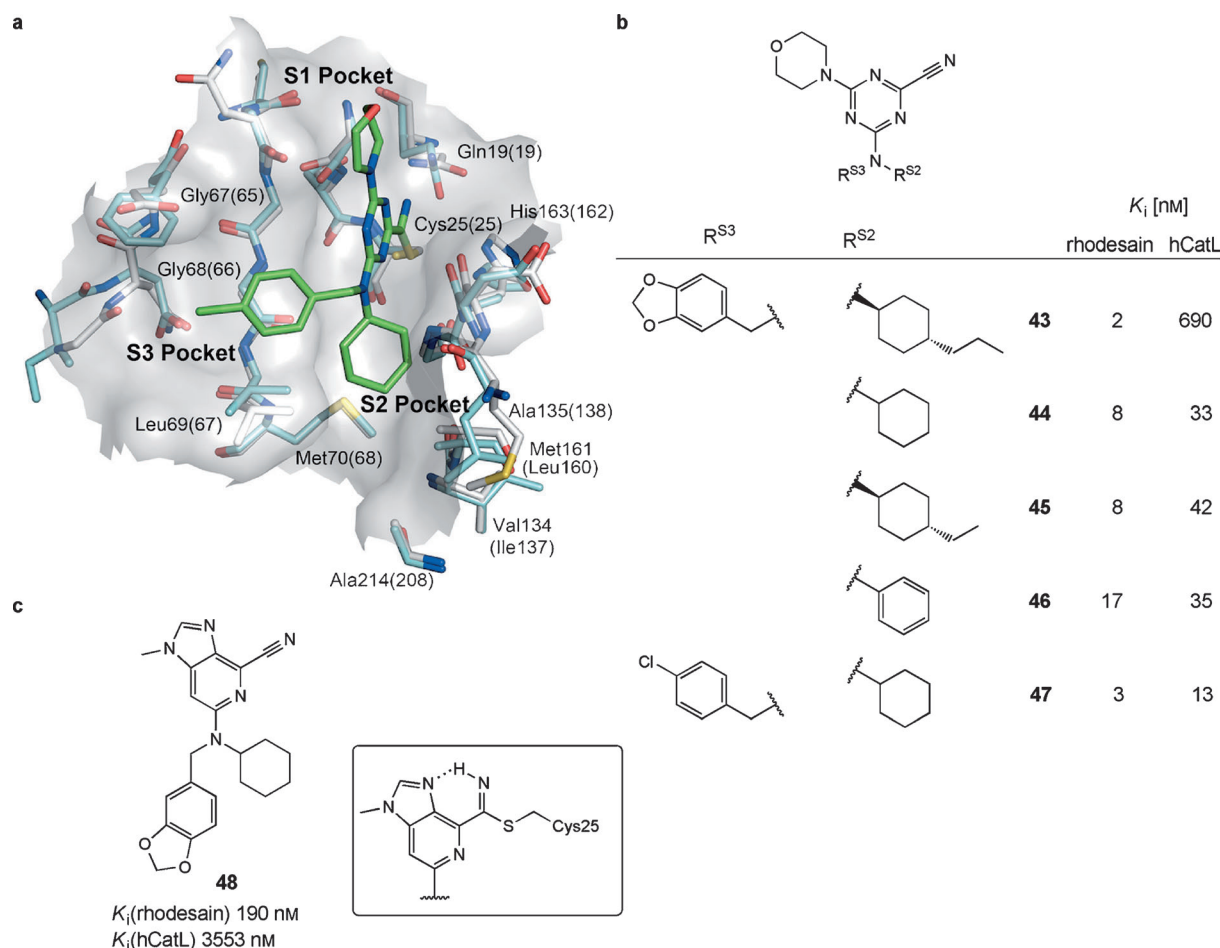


Figure 13. a) Overlay of the X-ray cocrystal structure of triazine nitrile **47** bound to hCatL (2.80 Å resolution, PDB ID: 4AXM)^[146d] and the X-ray crystal structure of *Tb* rhodesain (1.16 Å resolution, PDB ID: 2P86).^[147b] The ligand forms a reversible thioimide intermediate with Cys25 and gains selectivity for rhodesain over hCatL mainly through the proper filling of the S2 pocket. Color code: C_{enzyme}(hCatL) gray, O red, N blue, S yellow, C_{ligand} green, C_{enzyme}(rhodesain) cyan. Amino acid labels: hCatL(rhodesain). b) Selected triazine nitrile inhibitors of *Tb* rhodesain with various substituents targeting the S2 pocket. c) Chemical structure of imidazopyridine nitrile inhibitor **48** and illustration of the thioimide intermediate stabilized by an intramolecular hydrogen bond. The lower reactivity of the nitrile group in **48** reduces toxicity and off-targets effects.

assembled a series of triazenenitrile ligands such as **43–47** (Figure 13b). Although their potencies were outstanding, their cytotoxicity against human cell lines was significant, which we assigned to the high cross-reactivity of the activated nitrile group.^[146a,d] We subsequently conducted an experimental and computational investigation of a series of aryl and heteroaryl nitrile ligands and observed a clear dependency between the electrophilicity of the nitrile group and the binding affinity towards rhodesain, which varied over four orders of magnitude.^[146c] The cytotoxicity increases as the reactivity of the head group is enhanced, thus emphasizing the importance of its fine-tuning for a potential drug and the avoidance of over-activated head groups. A way to bypass cytotoxicity problems was demonstrated by the introduction of a 1-methyl-1*H*-imidazo[4,5-*c*]pyridine-4-carbonitrile head group as in **48** (Figure 13c).^[146d] Its nitrile group is significantly less reactive and the resultant covalent thioimide is stabilized by an intramolecular hydrogen bond. Most notably, this head group is stable towards the unspecific test nucleophile L-glutathione.^[146d] Nanomolar activity was retained for

the resulting ligand **48** (K_i = 190 nM), and the selectivity against the structurally related human Cathepsin L (hCatL) increased (selectivity ratios 19 (**48**) and 4 (**44**)). In addition, a fourfold decrease in the cytotoxicity for **48** relative to triazine **44** was observed in an assay with rat skeletal myoblasts.

Based on MOLOC modeling studies with the cocrystal structure of rhodesain and a vinyl sulfone inhibitor (PDB ID: 2P86),^[147b] the optimum filling of the subpockets was explored to enhance the biological activity of the ligands. Although no cocrystal structures of our inhibitors were obtained in complex with rhodesain, the predicted binding mode was confirmed by an X-ray cocrystal structure of **47** bound to hCatL.^[146d] The overlay of the cocrystal structure with the crystal structure of rhodesain (PDB ID: 2P86)^[147b] displays the high structural similarity of rhodesain and hCatL (Figure 13a).

The flat, solvent-exposed S1 pocket of rhodesain tolerates a variety of substituents without greatly affecting the binding affinity.^[146a,d] The wider S3 pocket is characterized by the flat

peptide segment of Gly65–Gly66 and shows a clear preference for aromatic substituents that establish amide $\cdots\pi$ stacking with this segment (see Section 4.3). The main contributions to the binding and selectivity are gained in the apolar S2 pocket, which is lined by Leu67, Met68, Ala138, Ile137, Leu160, and Ala208 (Figure 13a).^[149] This highly conserved hydrophobic cavity accommodates aromatic rings, as shown for ligand **46**, and cycloalkanes such as cyclohexyl (inhibitor **44**). Only two amino acids in the S2 pocket of rhodesain (Ile137, Leu160) are mutated in hCatL (Val134, Met161) and are important for gaining selectivity. An impressive example of the effect of properly filling the S2 pocket on the selectivity is the elongation of the 4-(*n*-ethyl) substituent of **45** to a 4-(*n*-propyl) unit in ligand **43**, which increased the binding affinity by a factor of 4 and at the same time the selectivity towards hCatL by a factor of 65 (Figure 13b).

Although the high affinity of the triazenitrile ligands is largely derived from the formation of the reversible covalent thioimide, additional proper filling of the S2 and S3 pockets yields inhibitors of rhodesain with K_i values in the single-digit nanomolar range. Targeting the S1 pocket affords little benefit. Imidazopyridine nitriles have been identified as less-reactive electrophiles that show reduced off-target effects. Combining the latter with optimized substituents for the S2 and S3 pocket in future studies should afford potent ligands with favorable physicochemical properties for further testing in *in vitro* and *in vivo* studies.

4. Interactions between Dipoles: Orthogonal Interactions, Halogen Bonding, and Amide $\cdots\pi$ Stacking

Identifying and understanding weak intermolecular interactions are key factors in successful structure-based drug design. They are occasionally overseen, underestimated, or simply not recognized in molecular modeling studies.^[150] The fundamental physical principal behind several of these weak interactions can be described by interactions between dipoles. Although dipolar interactions may originate from pure electrostatic attractive forces, complex quantum chemical effects very often play an additional role on closer look at the dipoles involved. Only through the combination of synthetic models, computational methods, and confirmation in biological case studies, can a comprehensive understanding of such interactions be gained. In this way, they become important and reliable tools for the design and optimization of leads in medicinal chemistry and crop science.

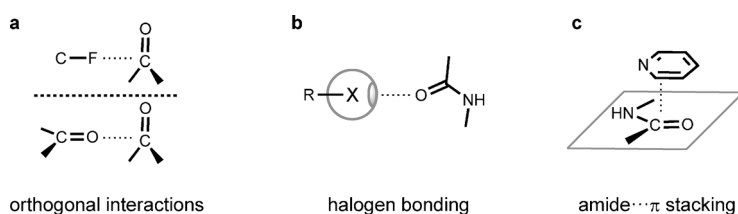


Figure 14. Different types of interactions involving dipoles reviewed in this section: a) orthogonal dipolar interactions, b) halogen bonding (XB), and c) amide $\cdots\pi$ stacking.

In this section, three types of interactions between dipoles are reviewed: orthogonal dipolar interactions, halogen bonding, and amide $\cdots\pi$ stacking (Figure 14). All of them appear in different characteristic alignments, which make each of them attractive and diverse subjects for detailed investigations. Investigation of these interactions has been a major part of our research program during the last decade.

4.1. Orthogonal Dipolar Interactions

Among the broad spectrum of intermolecular interactions, it is the weak interactions that remain a challenge for identification, quantification, and tailored applications.^[29,151] As such, the interaction of two dipoles aligned with an angle close to 90° has received detailed attention, as reviewed by us in 2005 and 2008.^[23,59] A survey of all the major interactions between bond dipoles in the CSD revealed that at close contact distances, in particular at or below van der Waals contacts, the orthogonal arrangement is the most abundant one, presumably for steric reasons.^[23] Two specific cases have been of particular interest in our studies: orthogonal C=O \cdots C=O interactions, and orthogonal C–X \cdots C=O interactions, both of which are relevant in protein–ligand complexation.

From a physical electrostatic point of view, no energetic gain arising from the orthogonal alignment of the dipoles is actually expected, as the potential energy for the interaction between aligned electrostatic dipole moments is dependent on the cosine of their angle (see the Supporting Information). Certainly, higher-order electrostatic and dispersion terms, also of quantum chemical nature, must be responsible for the attractiveness of the interaction.

Early small-molecule database searches^[152] provided clear indications for orthogonal C=O \cdots C=O interactions, which were first identified in X-ray crystal structures by Bolton,^[153] and later reviewed by Bent in 1968.^[154] Allen et al. re-examined C=O \cdots C=O interactions in a CSD search in 1998 and suggested their significance for stabilizing protein secondary structure motifs.^[152b] Intermolecular perturbation theory (IMPT) calculations led to the conclusion that the interaction energy at optimal distances for the perpendicular C=O \cdots C=O motif is similar to that of a sheared antiparallel motif and about a third of the interaction of the fully antiparallel motif.^[152b]

Recent CSD analyses by Raines and co-workers uncovered a small, but significant degree of pyramidalization out of the R₂C=O pseudotriangular plane when halide anions are below the van der Waals distance with the C=O carbonyl group (Figure 15).^[155] This change in geometry can be attributed to a Bürgi–Dunitz-type $n \rightarrow \pi^*$ interaction, where the lone pairs of electrons (n) of the halide anions overlap with the π^* orbital of the carbonyl group.^[156] Previously, the Raines research group confirmed in experimental studies and through databank searches the influence of such $n \rightarrow \pi^*$ interactions in amide C=O \cdots C=O arrays, which are omnipresent in protein secondary structures.^[157] Here too, they observed a Bürgi–Dunitz-type pyr-

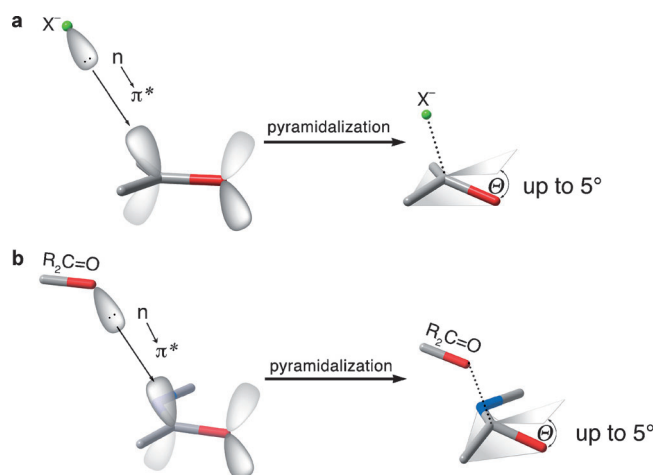


Figure 15. a) Halide anions and $R_2C=O$ carbonyl fragments undergo Bürgi–Dunitz-type $n \rightarrow \pi^*$ interactions below the sum of the van der Waals radii, which induces pyramidalization towards the halide anion, characterized by angles Θ up to 5° .^[155] b) Pyramidalization of the $C=O$ amide group in $C=O \cdots C=O$ interactions as a result of $n \rightarrow \pi^*$ interactions, with pyramidalization angles Θ reaching up to 5° .^[157]

amidization of the carbonyl group when the distance of the interacting atoms is less than the van der Waals distance.

Electrostatic dipolar and $n \rightarrow \pi^*$ interactions are not mutually exclusive, and their individual contribution to the overall interaction energy is certainly a matter of interacting distance and geometry. It can be expected that $n \rightarrow \pi^*$ interactions make important contributions at distances below the sum of the van der Waals radii, whereas electrostatic dipolar contributions dominate at distances close to or above the sum of the van der Waals radii. For orthogonal $C-F \cdots C=O$ interactions, we do not expect significant covalent bonding contributions even at short distances.

The quantification of orthogonal $C=O \cdots C=O$ and $C-F \cdots C=O$ interactions in solution was achieved by employing

“chemical double-mutant cycles”^[158] with unimolecular synthetic model systems. Molecular torsional balances developed by Wilcox and co-workers and based on the Tröger base scaffold^[159] were chosen as model systems, since they provide the perpendicular arrangement of the functional groups in the folded conformation, whereas their interaction is absent in the unfolded state (Figure 16).^[160] The $C_{sp^2}-F \cdots C=O$ (amide) interaction was quantified from a chemical double-mutant cycle using torsion balance **49** in a variety of apolar solvents (Figure 16a), and $\Delta(\Delta G)_{C-F \cdots C=O}$ was determined to be in the range of -0.2 to -0.3 kcal mol⁻¹,^[160] in agreement with earlier results.^[161] Quantification of the orthogonal amide $C=O \cdots C=O$ interaction in benzene and halogenated solvents, such as $CHCl_3$, was achieved by a double-mutant cycle employing balance **50** (Figure 16b).^[162] The free enthalpy increment $\Delta(\Delta G)_{amide\ C=O \cdots C=O}$ was experimentally determined as -0.7 kcal mol⁻¹ (benzene, 298 K), well within the range of aromatic stacking and edge-to-face interactions. The higher interaction free enthalpy determined experimentally for amide $C=O \cdots C=O$, compared to the $C_{sp^2}-F \cdots C=O$ (amide) interaction, presumably originates from the higher polarizability of the amide $C=O$ compared to the $C-F$ moiety; IMPT gas-phase calculations revealed a substantial contribution of dispersion forces to the orthogonal interaction between amide carbonyl groups. Although the attractive nature of orthogonal $C-F \cdots C=O$ and $C=O \cdots C=O$ interactions was for the first time confirmed in four different “chemical double-mutant cycle”, it should be kept in mind that the molecular balances are not static and that the gains in free energy ($\Delta(\Delta G)$) for the orthogonal interaction are in the range of the thermal energy kT at room temperature. Therefore, the measured free enthalpy increments do not correspond to a single interaction geometry, but rather to an ensemble in which the edge component wiggles on top of the face component.^[163]

Two orthogonal $C=O \cdots C=O$ interactions are indeed sufficient for defined self-association in the solid state and

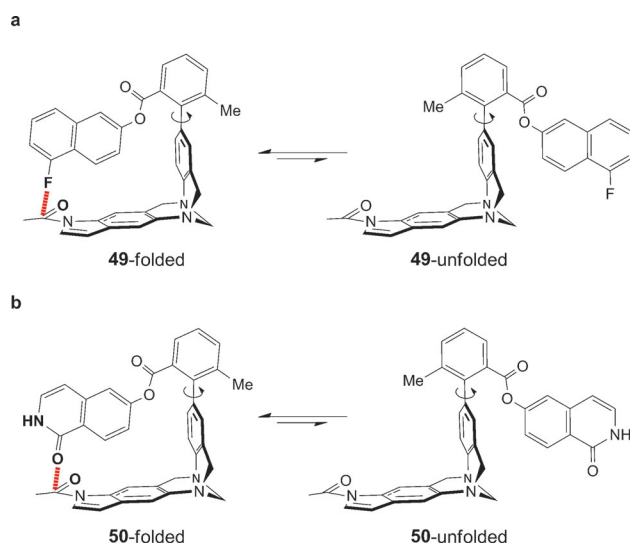


Figure 16. Wilcox-type torsion balances **49** and **50** used for the quantification of a) orthogonal $C-F \cdots C=O$ and b) $C=O \cdots C=O$ interactions by employing a double-mutant cycle.^[160, 162]

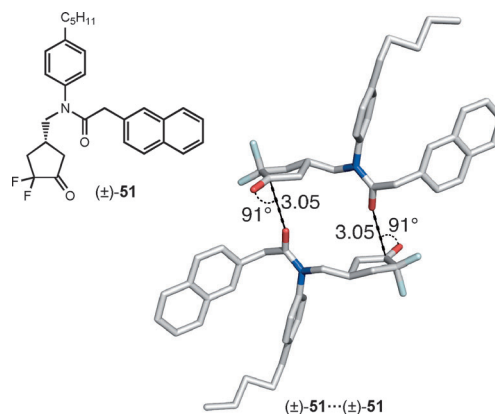


Figure 17. α, α -Difluorocyclopentanone (\pm)-**51** and its self-association in the crystalline state by two orthogonal $C=O \cdots C=O$ interactions of the motif $(R_2N)C=O \cdots C(CF_2)=O$. CSD code: 736727. Distances in Å. Self-dimerization of (\pm)-**51** in solution (perdeuterated cyclohexane (C_6D_{12}), 298 K) revealed a dimerization constant of $K_{dim} = 17.5$ M⁻¹ by a 1H NMR dilution study. Color code: C gray, O red, N blue, F turquoise.

most probably also in an apolar solvent.^[164] The crystal structure of α,α -difluorocyclopentanone (\pm)-**51** revealed a dimer held together by two such interactions at short contact (3.05 Å) and in a perpendicular arrangement (91°, Figure 17). This self-association also occurs with a similar geometry in solution, and the dimerization constant in perdeuterated cyclohexane (C_6D_{12} , 298 K) was estimated by NMR dilution studies as $K_{dim} = 17.5 M^{-1}$.

Besides the fact that organofluorine in drug lead compounds affects nearly all adsorption, distribution, metabolism, and excretion (ADME) properties, its contribution to enhanced affinity and selectivity through (ligand–F)⋯enzyme interactions has recently also been recognized.^[165] The significance of such interactions for biological systems was uncovered by a fluorine scan^[166] of inhibitors of thrombin, to map the fluorophilicity/fluorophobicity of the active site of this trypsin-like serine protease from the blood coagulation

cascade.^[22a,167] Cocrystal structure analysis had shown that the most potent ligands contained a 4-fluorophenyl ring to fill the hydrophobic D pocket and that the C–F group interacted with the backbone C=O group of Asn98 in a nearly orthogonal way, with an F⋯C=O angle of 96° at a distance of 3.5 Å. The fluorine scan of thrombin inhibitors has been extensively reviewed.^[59,165a]

A nice biological example, illustrating both orthogonal C–F⋯C=O and C=O⋯C=O interactions, is provided by complexes of a second-generation type II kinase inhibitor, nilotinib (trade name Tasigna, Figure 18a), which is applied to the treatment of Philadelphia chromosome positive chronic myeloid leukemia.^[168,169] This inhibitor binds to the Bcr-Abl tyrosine kinase and stabilizes its DFG (Asp381–Phe382–Gly383) loop in its “out conformation” (“DFG-out”). The cocrystal structure (2.21 Å resolution, PDB ID: 3CS9)^[170] revealed three types of protein–ligand interactions involving the DFG loop and its neighboring amino acid Ala380 (Figure 18b,c). The amide C=O group of the ligand undergoes “classical” hydrogen bonding to the backbone N–H group of Asp381. Two additional interactions are dipolar in nature. The backbone C=O group of Asp381 engages in a dipolar interaction ($d(O\cdots C)$ 3.4 Å, angle ($O\cdots C=O$) 85°) with the amide C=O group of the inhibitor. It can be assumed that there is cooperativity between the two different interactions (hydrogen bonding and dipolar), both involving the ligand C=O group. A third interaction is a short dipolar interaction of a C–F bond from the CF₃ group of the ligand with the backbone C=O group of Ala380 neighboring the DFG loop ($d(F\cdots C)$ 2.9 Å, angle ($F\cdots C=O$) 85°). This example illustrates once more the advantageous interplay of multiple molecular recognition events, which form a cooperative interaction network, as described by Kuhn et al.^[124] Given the numerous kinase targets pursued in the pharmaceutical industry, there is no doubt that many other examples exist in which amide bonds in DFG loops, but also in triphosphate-binding (of ATP) glycine-rich P loops,^[79] have been successfully addressed through orthogonal interactions with dipolar groups, such as C–X (X = F, Cl), CN, amide–C=O, and others.

4.2. Halogen Bonding

4.2.1. A Short Introduction to Halogen Bonding

Scatter plots for C–X⋯C=O interactions in the CSD are quite revealing.^[23] When the angle X⋯C=O is plotted against the X⋯C distance, a clear preference for orthogonal interactions at shorter contact distances is observed for X = F (see Section 4.1), but increasingly less for X = Cl, Br, and I. Furthermore, the percentage of close C–X⋯C=O interactions below the sum of the van der Waals radii increases with the increasing size of the halogen atom. Attractive interactions between the electronegative oxygen atom and the halogen atom come into play. The arrangement of C–X fragments pointing to the oxygen atom in carbonyl groups represents an important example of what has been called halogen bonding.^[171]

Halogen bonding, abbreviated as XB, describes the net attractive interaction between an electrophilic region associ-

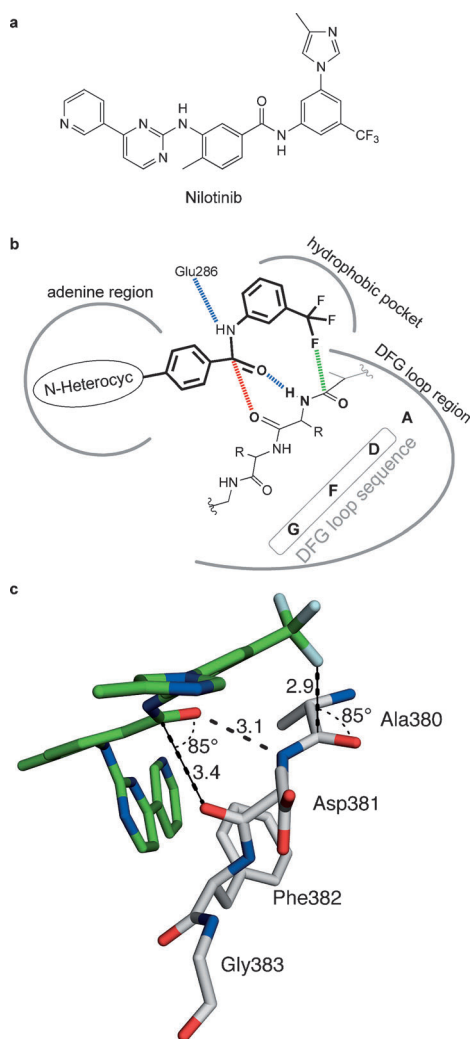


Figure 18. a) Structure of nilotinib.^[170] b) Interactions of nilotinib with the DFG loop.^[168b,170] Orthogonal dipolar C=O⋯C=O (red) and C–F⋯C=O (green) interactions, concomitant with hydrogen bonds (blue), build up the enzyme–ligand network for efficient binding. c) X-ray cocrystal structure of inhibitor nilotinib with Bcr-Abl tyrosine kinase, showing short contacts for the above discussed orthogonal dipolar interactions. Distances in Å. 2.21 Å resolution, PDB ID: 3CS9.^[170] Color code: C_{enzyme} gray, O red, N blue, F turquoise, C_{ligand} green.

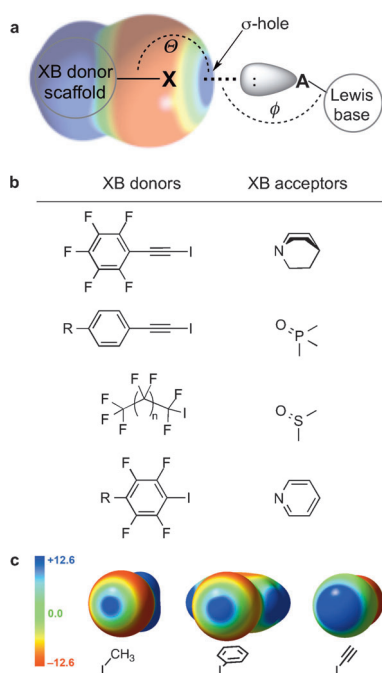


Figure 19. a) Schematic halogen bonding (XB) interaction with electrostatic potential surface, showing the σ -hole region at the outer cap of the halogen atom X (blue: positive, red: negative). b) Typical XB donor and XB acceptor motifs which have been investigated in solution studies.^[191a,b,192] R: different electron-donating and accepting groups. c) Molecular electrostatic potential (MEP) surface of iodomethane, iodobenzene, and iodoacetylene showing the increasing size of the σ hole by lowering the state of hybridization (scale in kcal mol⁻¹). Plotted at 0.001 electrons Bohr⁻³, calculated at the DFT:B3LYP/DGDZVP level of theory.

ated with a halogen atom in a molecular entity and a nucleophilic region in another, or the same, molecular entity, as defined by IUPAC (International Union of Pure and Applied Chemistry) in 2013.^[172] The halogenated part is defined as the XB donor, whereas the electron-donating molecule is referred to as the XB acceptor (in analogy to “hydrogen-bond acceptor”). This counterintuitive interaction, in which electron density approaches the electronegative halogen atom, can be rationalized by the anisotropic charge-density distribution around the heavier halogen atom, with a positive area (σ hole) of electrostatic potential opposite the C–X bond (Figure 19a).^[173] In a molecular orbital (MO) approach, the electrophilic region at the outer cap of the halogen atom corresponds to the antibonding σ^* orbital of the C–X bond.^[174] In this context, halogen bonding can be interpreted as an $n \rightarrow \sigma^*$ MO interaction, in which the lone pair of electrons (n) of a Lewis base interacts with the antibonding σ^* orbital along the C–X bond axis. Such an interpretation refers to the charge-transfer character of XB. The origin of XB, to what extent it is of electrostatic nature or orbital-driven, depends on the structures of the components involved in XB and is still under debate.^[175,176]

Halogen bonding is enhanced for the heavier halogens in the order $\text{Cl} < \text{Br} < \text{I}$, and is commonly not observed for fluorine. That trend can be rationalized by the increasing polarizability from F toward I within the periodic group.^[176a]

Electron-withdrawing substituents on the donor strengthen the halogen bonds, and this can be rationalized by an increase in the σ -hole potential.^[176b,c] Halogen bonding has been the subject of many reviews in recent years, as detailed in the following.

The first structural evidence for halogen bonding by X-ray crystallography was obtained by Hassel and Hvorslef in 1954 when examining cocrystals of diatomic bromine and 1,4-dioxane, which exhibit an unexpectedly short Br...O contact of 2.71 Å with a Br–Br...O angle of 178°.^[177] The significance of XB for molecular recognition in the field of crystal engineering was later systematically investigated by Metrangolo and Resnati,^[171c,178] and halogenated scaffolds with higher halogen substituents, interacting with atoms possessing lone pairs of electrons, were recognized by Desiraju as reliable synthons for predictable solid-state structures.^[179] Goroff et al. made use of the power of halogen bonding in an exemplary manner by designing a cocrystal structure, in which diiodobuta-1,3-diyne was aligned in a bis-(nitrile)oxalamide XB acceptor matrix and shown to undergo topochemical polymerization to form poly(diiododiacetylene) in a crystal-to-crystal transformation.^[180] Priimagi et al. demonstrated an application-oriented study on XB in amorphous solid-state materials by association of halogenated azobenzenes to Lewis-basic spin-coated polymers, such as poly(4-vinylpyridine).^[181] Subsequent surface-relief grating of these supramolecular polymer layers induces photoresponsive property changes, which correlates well with the XB donor ability of the halogenated azobenzenes.

Detailed studies from the field of computational chemistry are available for models describing halogen bonding,^[173–176,182] and the parameterization of XB for force fields is currently in development.^[183]

The application of halogen bonding to complex supramolecular systems and architectures has been of particular recent interest and has recently been reviewed.^[184] Rebek and co-workers presented one of the first XB-based supramolecular systems, consisting of a cleft-shaped receptor with a perylenediimide platform, which directs a higher halogen (Cl, Br) into the cleft. The system binds phenazine through N...X halogen bonding and π stacking, as monitored by UV/Vis binding titrations.^[185] The expected trend of the XB donor ability of Br being stronger than that of Cl is obeyed, as evidenced by a twofold increase in the association constant (benzene, 298 K). Halogen bonding has been implemented as an additional reliable recognition motif in the development of anion receptors. The negative charge of the anions is typically stabilized in proximity to multiple organohalogens, with their σ holes pointing towards the anion.^[186] Among supramolecular architectures with higher complexity, rotaxanes, catenanes, and macrocyclic hosts are the most elaborate structures in which XB contributes to molecular recognition and adds a templating effect to enhance synthesis.^[187] A recent report describes a selective fluorescence-quenching anion sensor, based on all-XB-recognition processes, in aqueous media.^[187d] XB electrophilic catalysis, pioneered by Huber and co-workers, is currently developing, and makes use of the σ -hole-assisted activation of substrates by a variety of halogenated catalysts.^[188]

XB strength has been quantified for several XB motifs in solution.^[189–192] However, systematic and comprehensive investigations are still rare.^[189b,c,191,192] Among these efforts, NMR binding titrations serve as a modern and precise technique, besides the recently less applied IR and UV/Vis spectroscopic methods. In 1983, Laurence et al. characterized the binding strength of several iodoalkynes to a variety of XB acceptors using IR spectroscopy and obtained good correlation with the electronic properties of the substituents directly attached to the C(2) atom on the 1-iodoethynyl moiety.^[189b,c] More-electron-withdrawing substituents resulted in higher association constants. Taylor and co-workers as well as Cabot and Hunter used NMR binding titrations to determine the association constants to various XB acceptors (Figure 19b), by employing perfluorinated iodoarenes and iodoalkanes as XB donors.^[191a,b] The binding free enthalpies for perfluorinated aromatic and aliphatic XB donors in apolar solvents, such as cyclohexane, range from -0.2 to -2.1 kcal mol⁻¹, depending on the XB acceptor ability. Figure 19b shows typical examples of XB donors and acceptor, which were the topic of these studies, with decreasing XB donor/acceptor strength from top to the bottom. Moreover, Taylor and co-workers gave valuable insight into the effect of competitive solvent on XB. It turned out that protic polar solvents (such as *i*PrOH) lowered the association by one magnitude relative to apolar solvents (e.g. cyclohexane, benzene).^[191b]

In a recent solution study, we examined (iodoethynyl)-benzenes as XB donors and revealed that their XB binding strength is as high as measured for perfluorinated organo-iodines (Figure 19b, top two XB pairs).^[192] Therefore, the σ -hole potential is tunable with the hybridization state of the C atom of the XB donor. Lower hybridization states n in C(sp^{*n*})-X result in stronger XB, as supported by our experimental findings and by modeling the electrostatic surface potential (Figure 19c). In this context, we showed that the formation of a halogen bond is enthalpy driven, with a large unfavorable entropic term, which remains constant when the XB donor strength is varied by different substitution at the *para* position of the aromatic ring. The strength of XB acceptors interacting with the (iodoethynyl)benzene derivatives is in the order pyridine < C=O < S=O < P=O < quinuclidine.^[192]

Efficient halogen bonding has several geometric prerequisites. The rigorous criteria for XB that the C–X···Y angle θ (see Figure 19a) must be close to linearity—ranging from 165° to ideally 180°—was demonstrated by CSD and PDB database analyses.^[171a,b,179,193] XB was found to be most effective if the distance between the halogen atom X and the Lewis-basic acceptor atom is below the sum of the van der Waals radii.^[175c,d] In the case of XB to carbonyl groups, the angle ϕ (X···O=C) (Figure 19a) can vary between 90 and 180°;^[175c,193] in other words, the carbonyl group interacts with halogen donors in a similar way as in the interaction with hydrogen-bond donors. In both cases, the approach pathways of the XB or hydrogen bond donor to the carbonyl oxygen atom define a cone shape pointing to the C=O bond and are not limited to an orientation towards the lone pairs of electrons.^[194]

4.2.2. Halogen Bonding in Biological Environments

The geometrical characteristics of XB in biological systems were analyzed by Auffinger, Ho et al. in an extensive PDB search, where they report the high abundance of C–X···O contacts of halogenated ligands to amide backbone carbonyl groups.^[193] As expected, all XB interaction criteria known from the field of organic crystal engineering, CSD searches, and computational studies were met in biological environments, although showing a slightly broader scattering. The role of halogen bonding for structure-based drug design was emphasized by additional systematic database searches, including the analysis of halogenated approved drugs in their binding modes, and accompanied with comprehensive theoretical descriptions for protein environments suitable for XB.^[195]

The first systematic investigation on halogen bonding in protein–ligand complexes was achieved by our group in cooperation with scientists from F. Hoffmann–la Roche.^[25] A larger ligand library (over 65 compounds) was employed to study the characteristics of halogen bonding to the backbone C=O group of Gly61 in the S3 pocket of human Cathepsin L (hCatL, see also Section 3.2.6).^[25a] These inhibitors undergo reversible covalent binding to the catalytic Cys25 residue in the S1 pocket under formation of thioimides, which are stabilized by the oxyanion hole of the protease (Figure 20a). Hydrogen bonds to the backbone NH group of Gly68 and C=O group of Asp162 are formed. Filling of the polar S3 pocket is achieved by amide··· π stacking of an aryl moiety on the flat peptide backbone Gly67–Gly68 (see Section 4.3) and halogen bonding to the backbone C=O group of Gly61 (red dashed line in Figure 20a). The biological affinity (IC₅₀) strongly increased upon changing the 4-aryl substituent from H (0.29 μ M) to Cl (0.022 μ M), Br (0.012 μ M), and I (0.0065 μ M; Table 1). In contrast, the IC₅₀ value remained essentially unchanged when moving from the unsubstituted inhibitor (–)-**52b** (X = H, 0.29 μ M) to fluorine-substituted (–)-**52a** (X = F, 0.34 μ M). In a control experiment with (–)-**52c** that has a methyl substituent (X = CH₃), which is most similar in size to Cl, binding affinity was not significantly enhanced. When the Cl substituent was installed at the 2- or 3-position of the aromatic ring or in a 2-thienyl ring *ortho* to the sulfur

Table 1: Inhibitors **52a–f** of hCatL bearing different substituents at the 4-position of the phenyl ring with measured inhibition (IC₅₀) and partition coefficient (logD). When X = Cl, Br, or I, XB increases the binding affinity to the enzyme.^[25a]

	Compound	X	IC ₅₀ [μ M]	logD	PDB ID	Res. [Å]
	(–)- 52a	F	0.34	2.36	2XU4	1.12
	(–)- 52b	H	0.29	2.11	–	–
	(–)- 52c	CH ₃	0.13	2.57	2XU5	1.60
	(–)- 52d	Cl	0.022	2.73	2YJC	1.14
	(+)- 52e	Br	0.012	2.96	2YJ2	1.15
	(+)- 52f	I	0.0065	3.23	2YJ8	1.30

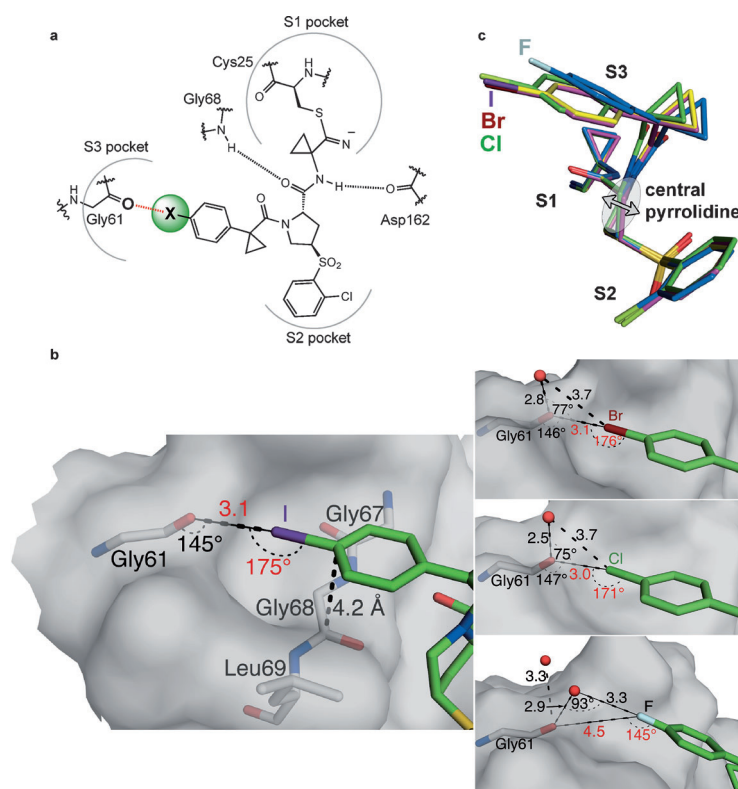


Figure 20. a) Binding mode of reversibly covalent inhibitors at the active site of hCatL. Halogen bonding of the 4-substituent to the C=O group of Gly61 in the S3 pocket is highlighted in red. b) Halogen bond geometries seen in the cocrystal structures of the I derivative (+)-**52f**, Cl derivative (–)-**52d**, and Br derivative (+)-**52e**. Also shown is the binding mode of F derivative (–)-**52a**. Distances are given in Å, XB distances X...O=C and XB angles C–X...O(=C) in red. Color code: C_{enzyme} gray, O red, N blue, I purple, Cl green, Br brown, F turquoise, C_{ligand} green. c) Overlay of the four halogenated ligands (–)-**52a**, (–)-**52d**, (+)-**52e**, and (+)-**52f** from their X-ray cocrystal structures with hCatL. Slight changes in the 5-ring pucker of the central pyrrolidine core (gray arrow) adjust the penetration of the *para*-substituted phenyl ring into the S3 pocket. Color code: C_{ligand}: depending on ligand, O red, N blue, I purple, Cl green, Br brown, F turquoise.

atom, the critical XB angle C–X...O (ideally 180°) was reduced to much smaller values (139° in the cocrystal structure of the thienyl ligand, 0.9 Å resolution, PDB ID: 2XU3).^[25a] A large loss in affinity of up to a factor of 28 was also evident.

X-ray cocrystal structures for the Cl- (–)-**52d**, Br- (+)-**52e**, and I-substituted ligand (+)-**52f**, all in the same space group *P*₂₁₂₁₂₁, confirmed both the binding mode of these inhibitors and the ranking of the measured IC₅₀ values (Table 1 and Figure 20b).^[25b] A clear C=O...X XB interaction is observed in all three cocrystals, as evidenced by short C=O...X contacts between 3.0 and 3.1 Å (89–92% sum of the van der Waals radii of O and X). The linear arrangement of the halogen bond is reflected by the O...X–C angles in the range 171–176°. In sharp contrast, the F atom in (–)-**52a** bound to hCatL is located at a F...O distance of 4.5 Å and bridged by a water molecule to the C=O group. This is indicative of an unfavorable interaction when organofluorine points closely to the O atom of a C=O group. However, the rest of the ligand skeleton of (–)-**52a** is maintained in a similar position and

geometry within the S1 and S2 pockets. Such precise adjustment of the halogenated phenyl ring into the S3 pocket is realized by small changes in the pucker of the pyrrolidine moiety at the center of the ligand scaffold (Figure 20c).

Additional water molecules, solvating the C=O group of Gly61 in the solvent-exposed S3 pocket, with nearly orthogonal (H₂O)...O(=C)...X angles of 75° and 77°, were observed in the complexes of the Cl ((–)-**52d**) and Br ((+)-**52e**) derivatives, respectively. In contrast, the I derivative (+)-**52f** exhibits no water molecules close to the XB motif, possibly because of the size of the iodine atom (van der Waals radius: 1.98 Å). It is remarkable that strong XB can be established in the optimal geometry, while at the same time the solvation of the XB-engaged amide C=O moiety is maintained in a similar geometry to that in the apo enzyme (1.92 Å resolution, PDB ID: 4AXL).^[146d] Such geometrically favored arrangements, with a hydrogen bond nearly perpendicular to the halogen bond, were found to be abundant in X-ray cocrystal structures of protein–ligand complexes, as revealed in a database search by Ho and co-workers.^[196] Protein–ligand structures, in which hydrogen and halogen bonds share a common oxygen atom contained in a peptide bond are rather narrowly distributed around an average D...O...X angle [D = electronegative hydrogen-bond donor atom (O, N, S), X = Cl, Br, I] of 88° (range ca. 75° to 90°, Figure 21). Moreover, ab initio calculations (B3LYP/6-31G) suggested that XB at amide carbonyl groups is energetically independent of the additional hydrogen bond at the shared carbonyl oxygen acceptor. The carbonyl group enjoys both interactions within the analogue classes of hydrogen and halogen bonding.

Although the S3 pocket of hCatL is polar and solvent-exposed, halogen bonding is also effective in the apolar back pocket in the active site of MEK1 kinase.^[25b,197] MEK1 kinase plays a major role in the extracellular signal-regulated kinase (Ras-Raf-MEK-ERK1/2) pathway of mutagen-activated protein (MAP) kinases, which is among the most frequently deregulated signaling

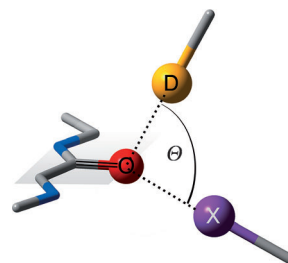
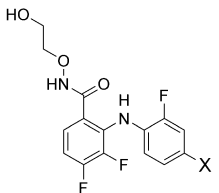


Figure 21. Orthogonal orientation of halogen bonds and hydrogen bonds to a peptide C=O moiety as observed in a PDB search for protein–ligand complexes. Typical X...O...D angles Θ at amide C=O carbonyl fragments with a shared interaction with D–H (hydrogen-bond donor) and X (X = Cl, Br, I) range from 75 to 95°, with the highest abundance at around 88°. ^[196]

pathways in human cancer.^[198] Therefore, MEK1 kinase is an attractive target in cancer therapy, with inhibitors currently being in phase III clinical trials.^[199]

MEK1 inhibitors with a substituted phenyl ring penetrate deeply into the hydrophobic back pocket surrounded by apolar residues (Met143, Leu118, Cys207, Phe209, Ile126, Val127). Binding affinity is increased by substitution at the 4-position in the sequence: $H < Cl < Br < I$ (Table 2), similar to

Table 2: Inhibitors **53a–e** of MEK1 kinase with increasing affinity (IC_{50}) for the higher halogens.^[197a]

	Compound	X	IC_{50} [nM] (MEK1)
	53a	F	120
	53b	H	52
	53c	Cl	26
	53d	Br	7.7
	53e	I	2.0

the hCatL inhibition discussed above.^[197b] X-ray cocrystal structures of halogen-substituted inhibitors **53c–e** bound to MEK1 confirmed XB to the backbone carbonyl group of Val127 (Figure 22, 2.7 Å resolution, PDB ID: 3DY7), with

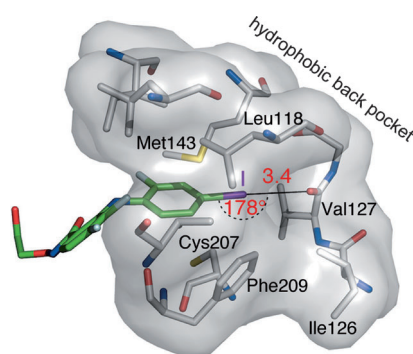


Figure 22. X-ray cocrystal structure of **53e** with MEK1 (resolution 2.7 Å, PDB ID: 3DY7).^[197c] One phenyl ring of the inhibitor (Table 2) reaches into the highly apolar back pocket of MEK1, gaining binding affinity from XB to the C=O group of Val127. Halogen bond angles and distances (in Å) are depicted in red. Hydrophobic amino acids of the MEK1 back pocket are labeled. Color code: C_{enzyme} gray, O red, S yellow, N blue, I purple, F turquoise, C_{ligand} green.

a short I...O contact of 3.4 Å with an C–I...O angle of 178° (IC_{50} = 2 nM) seen for **53e**. In another of the reported cocrystal structures with a closely related ligand^[197c] (2.3 Å resolution, PDB ID: 3DV3, not shown), the XB distance I...O is 3.1 Å, the C–I...O angle is 175°, and the C=O of Val127 is additionally solvated by a water molecule. Again, the angle ($H_2O \cdots O(=C) \cdots I$ is nearly orthogonal (98°, $d(H_2O \cdots O(=C))$ = 2.7 Å).

Taken together, the investigations on XB in the polar, solvent-exposed S3 pocket of hCatL and in the apolar back pocket of MEK1 show that the interaction is effective in

biological environments of different polarity. The studies with both enzymes also confirm that the amide C=O group can maintain solvation while engaged in XB. Halogen bonding is clearly an innovative opportunity for protein–ligand interaction that should not be overlooked in structure-based drug design.

Halogen-bond donors can be viewed as hydrophobic counterparts of hydrogen-bond donors, and less desolvation costs have to be paid upon complexation through XB. Such an example is discussed in the subsequent paragraph, where rather hydrophobic XB donor ligands bind to the polar hinge region in the ATP binding site of a kinase. Even moderately active aromatic XB donors instead of more costly perfluorinated ones are sufficient to enhance the binding affinity and selectivity in biological environment, in contrast to most 1:1 complexation events in solution. Ligand binding to the protein also occurs in the absence of XB, although to a weaker extent, but it pays most of the entropic penalty for complexation. Therefore, in biological complexation, the benefits of the enthalpically driven XB (down to ΔH = -9 kcal mol^{-1})^[192] are harvested to a large extent, in contrast to solution studies with small XB donors and acceptors, where strong enthalpy–entropy compensation is effective.

Kinase inhibition with ligands binding to the hinge at the ATP site is commonly realized through polar ligands featuring 1–3 hydrogen-bonding donor and acceptor groups.^[4] The alternative application of XB to establish binding to the polar hinge region was first exemplified in 2003 for CDK2 kinase (phospho-CDK2-cyclin A).^[200] Simple perhalogenated ligand scaffolds, such as 4,5,6,7-tetrabromobenzotriazole (**54**, IC_{50} = 1.6 μM), form two C–Br...O=C halogen bonds to the backbone carbonyl groups of Leu83 and Glu81, but there is no direct interaction with the NH group of Leu83, which usually is addressed by hydrogen-bond acceptors (Figure 23 a,b). A third halogen bond, although at a less favorable angle (145°), engages the C=O group of Ile10, and the fourth Br interacts favorably with the phenyl ring of Phe80.^[193]

In a more recent example by Johnson and co-workers, the same strategy was applied to gain kinase selectivity for CDK9 over CDK1 in an extraordinary molecular recognition process, all evidenced by X-ray cocrystal structures.^[201] CDK9, the kinase of the positive transcription elongation factor b, is required for processive transcription elongation by RNA polymerase II.^[202] As the inhibition of this kinase shows high antitumor activity, there was interest in the development of selective CDK9 inhibitors. Compared to the symmetric CDK2 inhibitor **54**, the scaffold of CDK9 inhibitor **55** (IC_{50} = 0.24 μM) was changed from tetra- to dihalogenation, and contains a ribose moiety attached to the central benzimidazole core (Figure 23 c). Again, XB to the convergent C=O groups of the hinge region of the ATP site accounts for the molecular recognition process in the binding mode of this inhibitor, in which the C=O group of Asp104 maintains solvation by a near perpendicular (104°) water molecule, in addition to XB, as discussed above (Figure 23 d). Part of the inhibition mechanism involves a conformational change in the glycine-rich P loop and the β3-αC loop induced by the inhibitor. This provides better shielding from solvent, which most probably adds to the stabilization of the CDK9 complex.

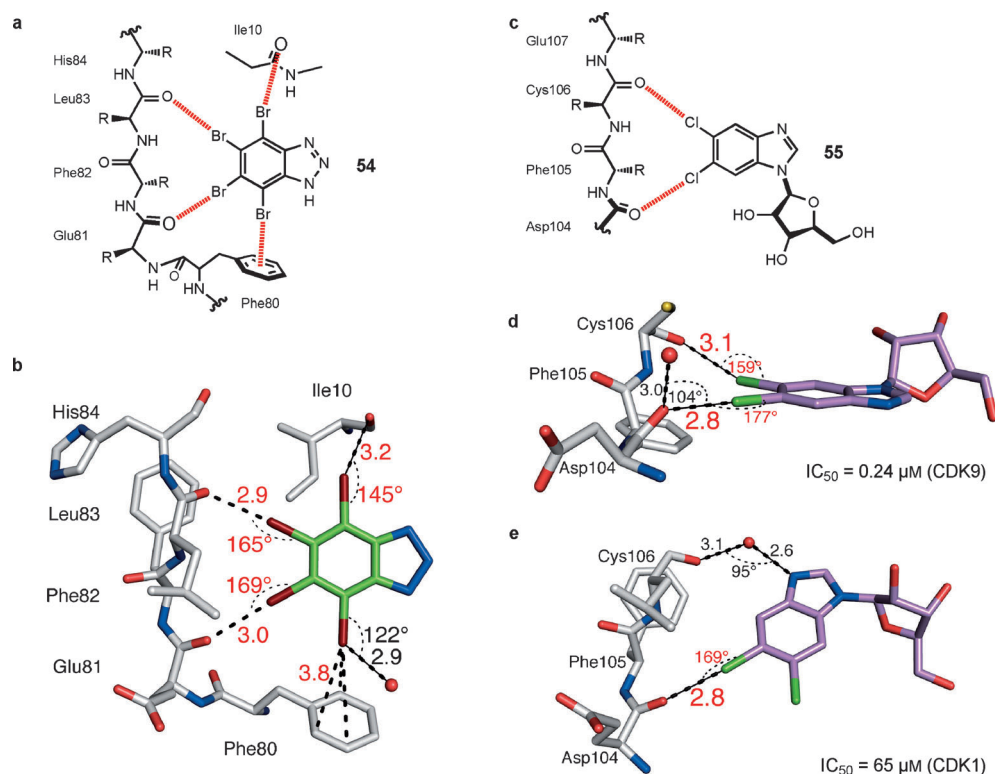


Figure 23. Halogen bonding of different kinase ligands binding to the hinge region of the ATP site. a) Schematic binding mode of 4,5,6,7-tetrabromobenzotriazole (**54**) to phospho-CDK2-cyclin A. b) X-ray cocrystal structure of CDK2 with inhibitor **54** (2.22 Å resolution, PDB ID: 1P5E).^[200b] c) Schematic binding mode of 5,6-dichlorobenzimidazole-1-β-D-ribofuranoside (**55**) to the hinge regions of d) CDK9 (2.80 Å resolution, PDB ID: 3MY1) and e) CDK1 (2.10 Å resolution, PDB ID: 5MY1).^[201] Halogen bond angles and distances (in Å) are depicted in red. Side chains of amino acids are omitted for clarity in (a) and (c). Color code: C_{enzyme} gray, O red, S yellow, N blue, Br brown, Cl green, C_{ligand} green (in b) or pink (in d and e).

In contrast, inhibition of CDK1 ($IC_{50} = 65 \mu M$) is 270-fold weaker. A different binding mode of inhibitor **55** occurs in the more restricted ATP binding site, and no significant ligand-induced conformational changes are observed.

The interplay of XB and metal-chelating interactions at the allosteric pocket of IspD^[42] (see Sections 2.3 and 3.2 as well as Figure 7) was uncovered in a collaborative study from a consortium of scientists from BASF and academia.^[203] The inhibition of enzymes from the non-mevalonate pathway of isoprenoid biosynthesis is of interest to both the development of new agents against infectious diseases (Section 3.2)^[77] and new herbicides.^[204] To improve the development of active compounds in both fields, lead compounds from agrochemical research have been examined against infectious germs.^[205] Compound library screening at BASF (about 100 000 compounds) yielded **56c** and **56d** as hits against *At*IspD. These lead compounds belong to the class of pseudilin alkaloids that were isolated in the 1960s from the marine bacterium *Pseudomonas bromoutilis* (Table 3).^[206] These highly halogenated marine natural products and other derivatives, such as **56a,b**, showed inhibition of *At*IspD with IC_{50} values around 50 μM (Table 3). The addition of divalent metal cations (Ca^{2+} , Cd^{2+} , Cu^{2+} , or Zn^{2+}) as assay buffer additives resulted in the affinity being enhanced to IC_{50} values in the low micromolar range. IC_{50} values around 50 μM were also measured against

Plasmodium vivax (*Pv*) IspD, but in this instance were not influenced by metal ions. EC_{50} values down to 1 μM were measured in cell-based assays against *Pf*NF54 strains.

Cocrystal structures, obtained by crystallization from buffers containing Cd^{2+} ions, shed light on the binding mode of the pseudilins to *At*IspD (Figure 24). The ligands complex in the newly discovered allosteric site of IspD^[42] but in a different binding mode than the previous ligands **1–3** (Section 2.3, Figure 2). The phenol and pyrrole rings of bound **56d** are in the *syn* conformation and complex a Cd^{2+} ion, which is also coordinated to the side chain of Gln238 of IspD and a water molecule. Besides this chelation, XB contributes to the binding affinity: the complex of **56d** displays a short Br...O contact of 3.25 Å with a C–Br...O angle of 164° to the

Table 3: Inhibitors of the pseudilin family of IspD from *Arabidopsis thaliana* (*At*IspD). Increased enzyme affinity (IC_{50}) is gained by bromine substitution at the R¹- and R²-position. $CdSO_4$ was added to the assay buffer, which was accompanied by up to a ninefold increase in inhibition.^[203]

Compound	R ¹	R ²	IC_{50} [μM] (<i>At</i> IspD)	
			without metal	40 μM Cd^{2+}
56a	H	Br	52 ± 6	19 ± 2
56b	F	F	79 ± 6	13 ± 2
56c	Cl	Cl	12 ± 1	2.2 ± 0.2
56d	Br	Br	13 ± 2	1.4 ± 0.2

backbone carbonyl oxygen atom of Val239 (Figure 24a). A similar XB geometry was crystallographically confirmed for the Cl-substituted inhibitor **56c** (1.8 Å resolution, PDB ID: 4NAL).

Ligand **56a** lacks a bromine substituent *para* to the phenolic OH group, and no XB was expected since metal-ion chelation should overwhelm the binding strength of XB and maintain the ligand in the Cd^{2+} -coordinated *syn* conformation. In fact, **56a** adopts two different conformations in the allosteric pocket. Only 70% of **56a** are in the expected *syn*

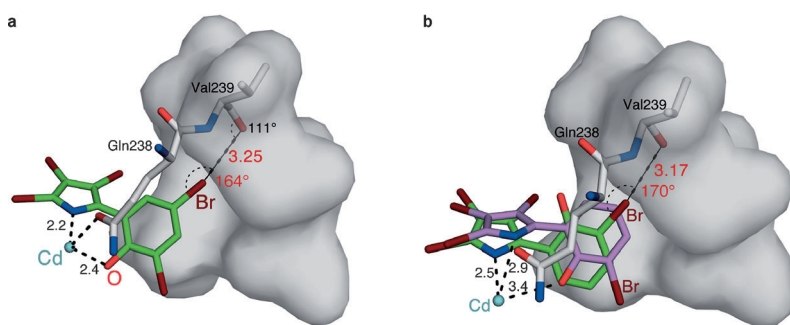


Figure 24. a) X-ray cocrystal structure of AtIspD with inhibitor **56d** (1.80 Å resolution, PDB ID: 4NAK).^[203] One Br atom of **56d** is engaged in XB with the C=O group of Val239. b) X-ray cocrystal structure of AtIspD with inhibitor **56a** (1.80 Å resolution, PDB ID: 4NAN). Ligand **56a** shows two (flipped) conformations in the X-ray structure: 70% *syn* conformation (pink ligand) with coordination of the phenolic group to the Cd²⁺ ion (cyan) and no XB of the bromine atom; 30% *anti* conformation (green ligand), exhibiting a preference for XB involving the Br atom *ortho* to the phenolic OH group over chelation. Halogen bond angles and distances (in Å) are depicted in red. Color code: C_{enzyme} gray, O red, N blue, Br brown, Cd turquoise, C_{ligand} green (*syn*, chelating) or pink (*anti*, XB).

conformation, whereas 30% are present as the *anti* conformer. This conformational flip of the phenol ring directs the Br atom *ortho* to the phenolic OH group toward the backbone carbonyl group of Val239 to establish a strong XB interaction [$d(\text{Br}\cdots\text{O}) = 3.17 \text{ Å}$, $\text{C}-\text{Br}\cdots\text{O}$ angle = 170° , Figure 24b]. Clearly, the energetics of metal-ion chelation and XB in the pseudilin complexes must be very similar.

Halogen bonding by the heavier organohalogen compounds provides a powerful new tool to the repertoire of intermolecular interactions for applications in supramolecular chemistry, catalysis, crystal engineering, advanced materials, and, as highlighted in this section, to structure-based hit generation and lead development. The criteria for efficient XB interactions are now well established and are identical for both synthetic and biological systems. Establishing a halogen bond might enhance protein–ligand interactions by as much as a factor of 74 which translates into a gain in free enthalpy of $\Delta\Delta G = -2.6 \text{ kcal mol}^{-1}$.^[25a] XB interactions can be favorably established in both a polar and apolar biological environment, and backbone amide C=O groups can maintain water solvation while undergoing halogen bonding. New opportunities for the design of innovative, active, and selective ATP-competitive kinase inhibitors arise, as illustrated by the study on CDK9 and CDK2.^[201] Perfluorination of the XB donor parts of the ligands is not necessary to harvest Gibbs energy from protein–ligand interactions, which simplifies ligand design and reduces costs. Enthalpically weaker XB interactions, such as between aryl halides (Cl, Br, I) and peptide C=O acceptors are sufficient, as the entropic costs of protein–ligand binding are largely paid by the complexation of the extended ligands, and the full additional enthalpic gains of XB are harvested without much additional loss in entropy.^[192]

4.3. Amide... π Stacking

The dipolar interaction of flat amide groups with aromatic π systems is closely related to the well-studied π -stacking

interactions between aromatic rings, which are of great importance in chemical and biomolecular recognition.^[14e,18b,c] In our research program toward the inhibition of factor Xa (see Section 3.2.4), we found a large difference in the affinity of ligands featuring constitutionally isomeric oxazole moieties to stack on the flat peptide bond Cys191–Gly192 lining the S1 pocket.^[119] The X-ray cocrystal structures of the two ligands (\pm)-**34** ($K_i = 1620 \text{ nM}$) and (\pm)-**35** ($K_i = 146 \text{ nM}$, Figure 25a) showed nearly identical binding modes, but an 11-fold difference in binding affinity. Besides differences in the angle between the two exit vectors on the oxazole rings (see Figure 11 in Section 3.2.4), we considered the different orientation of the dipole moments of the two oxazole rings as a possible main reason for the measured difference in binding affinity. In the complex of the better binder (\pm)-**35**, the local dipole moment vectors of the heterocycle and peptide amide

bond are aligned in an antiparallel way, whereas they are parallel in the less-stable complex of (\pm)-**34**. The observed difference in binding affinity, depending on the relative alignment of these two dipole moment vectors, either parallel or antiparallel, raised the general question about the influence of dipolar interactions on the stacking of amide π systems with aromatic moieties possessing a distinct dipolar moment.

The influence of dipole vector alignment on the energetics of amide bond...heteroaromatic π stacking was subsequently investigated in a computational study in collaboration with B. Kuhn (F. Hoffmann–la Roche)^[128] at the SCS-MP2 (spin component scaled Møller–Plesset 2) level of theory.^[207] Conducting a rotational scan with pyridine ($\mu_{\text{expt}} = 2.2 \text{ D}$), stacked parallel-displaced on the amide fragment *N*-methylacetamide (NMAC, $\mu_{\text{expt}} = 3.7 \text{ D}$), used as a model for the amide backbone fragment, revealed a clear preference for an antiparallel arrangement of the dipole moment vectors ($\alpha = 179^\circ$, $\Delta E = -2.5 \text{ kcal mol}^{-1}$) over the parallel arrangement ($\alpha = 2^\circ$, $\Delta E = -0.9 \text{ kcal mol}^{-1}$, Figure 25b).^[128]

We subsequently investigated the parallel-displaced stacking of a library of commonly used aromatic heterocycles featuring significantly different dipole moments on top of the flat NMAC amide fragment. In the optimized structures, the median angle between the dipole moment vectors was close to 161° , thus approaching the ideal antiparallel alignment. A linear trend of increased interaction energy $-\Delta E$ with increasing dipole moment strength with a correlation coefficient $R^2 = 0.84$ was observed for 17 aromatic heterocycles (Figure 25c).^[128]

Three main conclusions, useful for structure-based design, could be drawn from this theoretical study: 1) π stacking on peptide amide bonds becomes energetically more favorable as the π -electron density of the aromatic ring decreases (see benzene, pyrazine, triazine in Figure 25c), 2) the interaction strength increases as the dipole moment of the aromatic heterocycles increases, and 3) an antiparallel alignment of the dipole moment vectors of the amide and heterocycle is highly

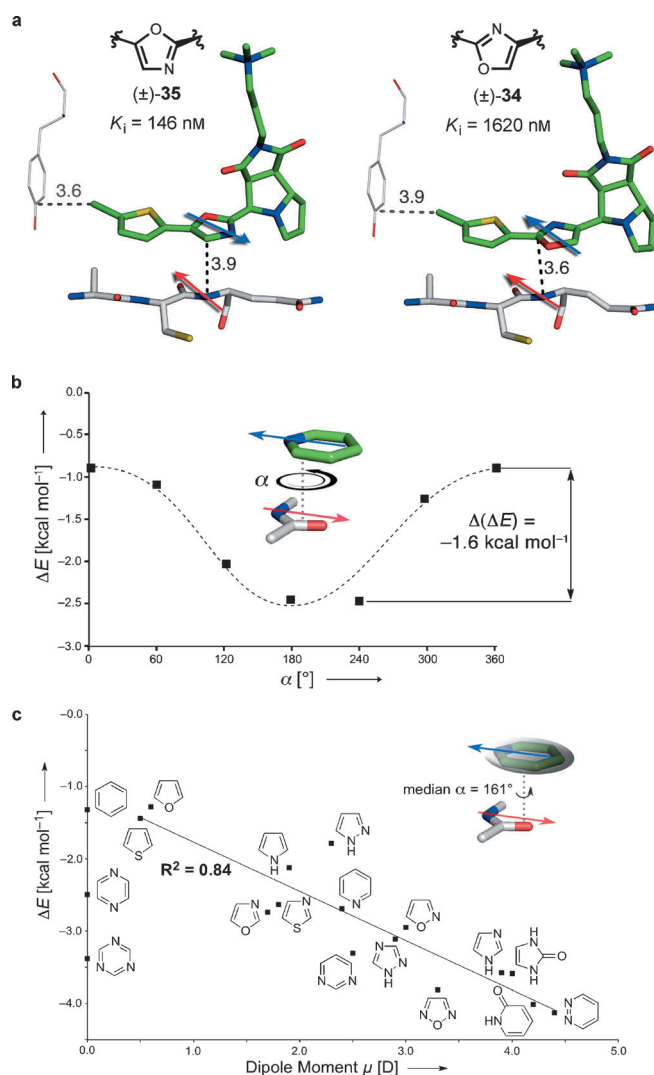


Figure 25. a) Dipolar amide- π stacking between inhibitors (±)-34 and (±)-35 and factor Xa (1.29 Å resolution, PDB ID: 2Y5G, and 1.33 Å resolution, PDB ID: 2Y5H).^[119] Ligand (±)-35 ($K_i = 146$ nM) exhibits favorable antiparallel dipolar stacking to the amide backbone, which translates into an affinity gain of a factor of 11 relative to ligand (±)-34 with parallel-aligned dipole moment vectors ($K_i = 1620$ nM). Color code: C_{enzyme} gray, O red, S yellow, N blue, Cl green, C_{ligand} green. b) Calculated energy profile for the rotational scan of the parallel-displaced NMAC-pyridine dimer with an optimized stacking arrangement of the two fragments. Antiparallel alignment ($\alpha = 179^\circ$) of the stacking dipole vectors is favored by 1.56 kcal mol⁻¹ compared to the parallel arrangement ($\alpha = 2^\circ$).^[128] Level of theory: SCS-MP2/aug-cc-pVDZ//LMP2/cc-pVDZ.^[207] c) Correlation between the calculated interaction energies ΔE of various NMAC-heteroarene dimers at the optimized geometry and the dipole moment μ of the heterocycle; level of theory as in (c). R^2 : Correlation coefficient; median angle between the dipole moment vectors $\alpha = 161^\circ$.^[128]

preferred. We are currently validating these theoretical predictions for amide $\cdots\pi$ stacking interactions both in model systems and in protein environments.

5. Summary and Conclusions

The various case studies presented in this Review on complexation processes, involving synthetic receptors and model systems as well as biological receptors, underline the challenges and opportunities for structure-based ligand design and optimization in medicinal chemistry and crop science. It has become clear over the past decades that the comprehensive study of an individual intermolecular interaction, separated from other influences, with biological receptors is difficult, given the complexity of the protein environment. It is a true “tour de force” to determine the role of water in biological complexation in a multidimensional approach including ITC, cryo-X-ray crystallography, high-level computing and molecular simulations, mutational studies, and other methods, as different regions in the protein binding site are characterized by different thermodynamic profiles. Complexation liberates water molecules that solvate the free binding partners, but this can occur in an entropically driven way (e.g. in ion pairing or in ion-dipole complexation) or with an enthalpic driving force (e.g. the displacement of high-energy water from apolar binding pockets). Here, the benefits of smaller, structurally more defined synthetic models and receptors come into play. The nonclassical enthalpically driven hydrophobic effect was discovered in host-guest complexation studies with cyclophane receptors featuring tight apolar binding sites and has now been fully validated in an elegant study of cucurbituril complexation. The quantification of individual intermolecular interactions, such as orthogonal dipolar interactions and halogen bonding would also not have been possible without model studies, outside the more complex protein environment. It does not come as a surprise that synthetic host-guest chemistry is currently seeing a renaissance, in particular with the work on cucurbituril receptors.

However, to push progress in small-molecule ligand design aimed at new drugs and agrochemicals, the in-detail exploration of biological complexation using today’s sophisticated experimental—analytical and structural—and computational methods repertoire is even more in demand. It would be impossible to investigate the role of water clusters seen in protein active sites with synthetic model systems. Gratifyingly, computational tools for identifying those individual water molecules in a water cluster, which can be displaced with energetic benefits by ligand parts, have been much refined. It is also predictable that the different thermodynamic profiles of different types of water molecules at active sites will become understood in the future, thereby enhancing the success probability of structure-based design, for which water poses a great challenge.

High-resolution protein X-ray crystallography provides the basis for structure-based ligand design. This is illustrated in the various case studies described in the Review, which address the diversity of biological receptor sites of different size and shape, conformational dynamics, and polarity. Enzyme active sites of low conformational dynamics and deep pockets of reduced polarity are the most rewarding targets, such as in Factor Xa, and biological activities in the low nanomolar range (IC_{50} values) are usually reached in

a reasonable time frame. This contrasts with the difficulties to develop potent inhibitors for very large pockets (such as in trypanothione reductase), highly polar active sites (as in IspE and IspF), or conformationally very flexible pockets (IspF). Structure-based design is a superior approach to lead generation and optimization, but it will not in the near future replace experimental high-throughput screening (HTS) or the increasingly applied virtual screening of large compound libraries. Without HTS, the potent hits for IspD would not have been found, as the allosteric pocket, which hosts the hits and leads, had not been recognized on the protein surface.

X-ray structure determination is almost “flooding” us today with structural information that can be applied to a new ligand-design process. A prerequisite to success, however, is the intelligent use of the vast information in the PDB and in-house databases, and that this knowledge is made available to the scientists on projects without the need for demanding computing. It is important to retrieve and analyze existing knowledge and not to simply “consume” the structural information that is abundantly collected around active projects, often amounting to 100 or more cocrystal structures. Being a more easily accessible database, the CSD has become an essential tool for avoiding conformational errors in ligand design. Furthermore, small-molecule networks in the crystal provide an invaluable tool to identify new intermolecular interactions. Examples are orthogonal dipolar $C=O \cdots C=O$ interactions and halogen bonding, which were structurally first discovered in crystal structures. It is certainly worthwhile to continue exploring the CSD to identify short intermolecular contacts that are statistically relevant and to try to understand their origin.

A topic not covered in this Review is the physicochemical and ADME/pharmacokinetic (pK) properties. As many of the physicochemical properties can now be determined in silico, such as solubility, metabolic stability, safety data, pK_a , logD, logP, or membrane permeability, it is desirable that these data become immediately available during structure-based design. This information can guide scientists on projects in the multidimensional optimization required to prepare the most promising leads for later development. User-friendly computational organization and analysis of “big data” has the highest potential to boost future drug discovery research. Computational approaches will identify the most promising hits from in silico libraries, provide improved automated structure-based design software, and identify suitable protein pockets for ligand binding. They will define conformational preferences of ligands based on the wealth of information in the CSD, recognize potential cross-reactivity based on protein database comparisons, and implement tools that give facile access to the existing structural, physicochemical, ADME information. Intelligent user-friendly computing will enable a true multidimensional approach to lead optimization through structure-based design in the future.

However, here the circle closes again: until recently the force fields in the computational design programs did not “know” about aromatic–aromatic interactions and their substituent effects, orthogonal dipolar interactions, halogen bonding, amide $\cdots\pi$ stacking, or about the energetics of

displacement of water molecules. The same holds on the conformational side, where until recently the distinct geometric preferences of sulfonamides and aryl sulfones were not parameterized in the modeling software. The experimental discovery and quantification of these processes were prerequisites for implementing their parameters into computer software. Molecular recognition studies with synthetic and biological receptors are, therefore, not only expanding our fundamental understanding of life and matter, but will continue to make important contributions to speeding up the successful generation and optimization of leads with potential for development.

Molecular recognition research in the Diederich group has continuously been funded by the ETH research council, and F. Hoffmann–La Roche, Basel, with additional support from Chugai and Genentech. The German Fonds der Chemischen Industrie is acknowledged for Kekulé doctoral fellowships to Michael Harder, Veronika Ehmke, and Oliver Dumele. This work could not possibly have been done solely in the chemistry environment at ETH, and our group has heavily relied on many stimulating and efficient cross-disciplinary outside collaborations—academic and industrial—over the past decades. The collaborators made available the proteins under investigation, provided most of the biostructure and bioassaying data, and participated in a most fertile way in project discussions and reviews. In fact, entire networks expanding to up to five groups in the IspD project were recently created to collect the biological data reported in part in this Review. The corresponding author is most grateful to the scientists at Roche Basel; it is difficult to single out individuals, but he wishes to thank in particular Prof. Hans-Joachim Böhm, Prof. Klaus Müller, Dr. Martin Stahl, Dr. David W. Banner, Dr. Bernd Kuhn, Dr. Wolfgang Haap, and Dr. Markus Rudolph for numerous inspiring discussions and for their support. He also wishes to thank Dr. Matthias Witschel, Dr. Wolfgang Höffken, and Thomas Mietzner (BASF), Dr. Jeff Blaney (Genentech), Dr. Daniel Bur (Actelion), as well the scientists at Chugai for their continuing great interest in molecular recognition phenomena. Since the early 2000s, we have been maintaining fertile academic collaborations with Prof. Gerhard Klebe (Univ. Marburg), Prof. Adelbert Bacher (TU Munich), Prof. Markus Fischer (Univ. Hamburg), Prof. Luise Krauth-Siegel (Univ. Heidelberg), and have expanded them in recent years to include Prof. Tanja Schirmeister (Univ. Mainz), Prof. Michael Groll (TU Munich), and Prof. Emil Pai (Univ. Toronto). The names of their co-workers on the projects are included in the references. At ETH, discussions on molecular recognition with Prof. Donald Hilvert and Prof. Jack D. Dunitz, and on computational developments with Prof. Gisbert Schneider are greatly acknowledged. All cell-based assays and cytotoxicity studies in tropical disease projects were done by Prof. Reto Brun, Dr. Marcel Kaiser, and Dr. Matthias Rottmann at the Swiss Tropical and Public Health Institute (SPHI) in Basel. Finally, the corresponding author thanks all his co-workers over the past three decades who pushed this work during doctoral or postdoctoral research in his laboratory; again their names are given in the references. We acknowledge proof-

reading of the manuscript by Dr. Michael Harder and Dr. Tristan Reekie (both ETH).

Received: September 10, 2014

Published online: January 28, 2015

- [1] <http://www.rcsb.org/pdb/statistics/holdings.do> (April 15, 2014).
- [2] <http://www.ccdc.cam.ac.uk/Solutions/CSDSystem/Pages/CSD.aspx> (April 22, 2014).
- [3] K. A. Brameld, B. Kuhn, D. C. Reuter, M. Stahl, *J. Chem. Inf. Model.* **2008**, *48*, 1–24.
- [4] C. Bissantz, B. Kuhn, M. Stahl, *J. Med. Chem.* **2010**, *53*, 5061–5084.
- [5] a) G. Klebe, *Wirkstoffdesign. Entwurf und Wirkung von Arzneimitteln*, 2nd ed., Spektrum, Heidelberg, **2009**; b) G. Klebe, *Drug Design*, Springer, Berlin, **2013**; c) *Structure-Based Drug Discovery* (Ed. L. W. Tari), Humana, Springer, New York, **2012**; d) *De Novo Molecular Design* (Ed.: G. Schneider), Wiley-VCH, Weinheim, New York, **2014**; e) A. K. Ghosh, S. Gemma, *Structure-Based Design of Drugs and Other Bioactive Molecules*, Wiley-VCH, Weinheim, **2014**.
- [6] For early examples, see a) P. Y. S. Lam, P. K. Jahdavi, C. J. Eyermann, C. N. Hodge, Y. Ru, L. T. Bachelier, J. L. Meek, M. J. Otto, M. M. Rayner, Y. N. Wong, C.-H. Chang, P. C. Weber, D. A. Jackson, T. R. Sharpe, S. Erickson-Viitanen, *Science* **1994**, *263*, 380–384; b) U. Obst, V. Gramlich, F. Diederich, L. Weber, D. W. Banner, *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1739–1742; *Angew. Chem.* **1995**, *107*, 1874–1877.
- [7] For recent reviews, see a) A. M. Davis, S. A. St-Gallay, G. J. Kleywegt, *Drug Discovery Today* **2008**, *13*, 831–841; b) M. Congreve, C. J. Langmead, J. S. Mason, F. H. Marshall, *J. Med. Chem.* **2011**, *54*, 4283–4311; c) D. J. Huggins, W. Sherman, B. Tidor, *J. Med. Chem.* **2012**, *55*, 1424–1444; d) A. A. Alex, D. S. Millan in *Drug Design Strategies: Quantitative Approaches, RSC Drug Discovery Series No. 13* (Eds.: D. J. Livingstone, A. M. Davis), Royal Society of Chemistry, Cambridge, **2012**, pp. 108–163; e) J. Michel, *Phys. Chem. Chem. Phys.* **2014**, *16*, 4465–4477.
- [8] a) J. E. Ladbury, *Chem. Biol.* **1996**, *3*, 973–980; b) R. Ludwig, *Angew. Chem. Int. Ed.* **2001**, *40*, 1808–1827; *Angew. Chem.* **2001**, *113*, 1856–1876; c) S. W. Homans, *Drug Discovery Today* **2007**, *12*, 534–539; d) P. Ball, *Chem. Rev.* **2008**, *108*, 74–108.
- [9] a) J.-M. Lehn, *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 89–112; *Angew. Chem.* **1988**, *100*, 91–116; b) D. J. Cram, *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 1009–1020; *Angew. Chem.* **1988**, *100*, 1041–1052; c) C. J. Pedersen, *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 1021–1027; *Angew. Chem.* **1988**, *100*, 1053–1059; d) for a historical survey on supramolecular chemistry from the perspective of this journal, see F. Diederich, *Angew. Chem. Int. Ed.* **2013**, *52*, 2714–2742; *Angew. Chem.* **2013**, *125*, 2778–2807.
- [10] a) F. Diederich, *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 362–386; *Angew. Chem.* **1988**, *100*, 372–396; b) F. Diederich, *Cyclophanes*, Royal Society of Chemistry, London, **1991**; c) F. Diederich in *Modern Cyclophane Chemistry* (Eds.: H. Hopf, R. Gleiter), Wiley-VCH, Weinheim, **2004**, pp. 519–546.
- [11] a) D. A. Dougherty, D. A. Stauffer, *Science* **1990**, *250*, 1558–1560; b) H.-J. Schneider, *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 1417–1436; *Angew. Chem.* **1991**, *103*, 1419–1439; c) *Supramolecular Chemistry: From Molecules to Nanomaterials, (Molecular Recognition), Vol. 3* (Eds.: P. A. Gale, J. W. Steed), Wiley, Chichester, **2012**; d) D. A. Dougherty, *Acc. Chem. Res.* **2013**, *46*, 885–893.
- [12] F. P. Schmidtchen in *Supramolecular Chemistry: From Molecules to Nanomaterials, (Techniques), Vol. 2* (Eds.: P. A. Gale, J. W. Steed), Wiley, Chichester, **2012**, pp. 275–296.
- [13] a) W. Kauzmann, *Adv. Prot. Chem.* **1959**, *14*, 1–63; b) C. Tanford, *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, 2nd ed., Wiley, New York, **1980**; c) W. Blokzijl, J. B. F. N. Engberts, *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 1545–1579; *Angew. Chem.* **1993**, *105*, 1610–1650.
- [14] a) S. B. Ferguson, E. M. Seward, F. Diederich, E. M. Sanford, A. Chou, P. Inocencio-Szweda, C. B. Knobler, *J. Org. Chem.* **1988**, *53*, 5593–5595. The crystal structure of this first cyclophane host, for which enthalpically driven complexation was demonstrated by ITC, contained two disordered high-energy water molecules in the cavity; b) D. B. Smithrud, F. Diederich, *J. Am. Chem. Soc.* **1990**, *112*, 339–343; c) D. B. Smithrud, T. B. Wyman, F. Diederich, *J. Am. Chem. Soc.* **1991**, *113*, 5420–5426; d) F. Diederich, D. B. Smithrud, E. M. Sanford, T. B. Wyman, S. B. Ferguson, D. R. Carcanague, I. Chao, K. N. Houk, *Acta Chem. Scand.* **1992**, *46*, 205–215; e) for a discussion, see also: E. A. Meyer, R. K. Castellano, F. Diederich, *Angew. Chem. Int. Ed.* **2003**, *42*, 1210–1250; *Angew. Chem.* **2003**, *115*, 1244–1287.
- [15] For enthalpy–entropy compensation, see a) Y. Inoue, Y. Liu, L.-H. Tong, B.-J. Shen, D.-S. Jin, *J. Am. Chem. Soc.* **1993**, *115*, 10637–10644; b) J. D. Dunitz, *Chem. Biol.* **1995**, *2*, 709–712; c) L. Liu, Q.-X. Guo, *Chem. Rev.* **2001**, *101*, 673–695; d) V. M. Krishnamurthy, B. R. Bohall, V. Semetey, G. M. Whitesides, *J. Am. Chem. Soc.* **2006**, *128*, 5802–5812; e) K. K. Frederick, M. S. Marlow, K. G. Valentine, A. J. Wand, *Nature* **2007**, *448*, 325–330; f) for enthalpy–entropy compensation in drug development, see E. Freire, *Drug Discovery Today* **2008**, *13*, 869–874; g) J. D. Chodera, D. L. Mobley, *Annu. Rev. Biophys.* **2013**, *42*, 121–142; h) J. M. Myslinski, J. H. Clements, J. E. DeLorbe, S. F. Martin, *ACS Med. Chem. Lett.* **2013**, *4*, 1048–1053; i) for an example of enthalpy–entropy compensation in multisubstrate binding to a large hydrophobic capsule, see A. Grego, A. Müller, I. A. Weinstock, *Angew. Chem. Int. Ed.* **2013**, *52*, 8358–8362; *Angew. Chem.* **2013**, *125*, 8516–8520.
- [16] U. Obst, D. W. Banner, L. Weber, F. Diederich, *Chem. Biol.* **1997**, *4*, 287–295.
- [17] For recent reviews and monographs on protein–ligand interactions, see a) V. Harmat, G. Náray-Szabó, *Croat. Chem. Acta* **2009**, *82*, 277–282; b) R. Baron, J. A. McCammon, *Annu. Rev. Phys. Chem.* **2013**, *64*, 151–175; c) *Protein-Ligand Interactions, Methods and Applications*, 2nd ed. (Eds.: M. A. Williams, T. Daviter), Humana, Springer, New York **2013**.
- [18] Interactions with aromatic rings in chemical and biological systems have been extensively reviewed: a) Ref. [14e]; b) L. M. Salonen, M. Ellermann, F. Diederich, *Angew. Chem. Int. Ed.* **2011**, *50*, 4808–4842; *Angew. Chem.* **2011**, *123*, 4908–4944; c) an entire issue of *Acc. Chem. Res.* has been devoted to this topic: *Aromatic Interactions in Chem. Biol., Acc. Chem. Res.* **2013**, *46*, 873–1050.
- [19] D. J. Cram, J. M. Cram, *Container Molecules and Their Guests*, Royal Society of Chemistry, Cambridge, **1997**.
- [20] a) J. Rebek, Jr., *Angew. Chem. Int. Ed.* **2005**, *44*, 2068–2078; *Angew. Chem.* **2005**, *117*, 2104–2115; b) S. M. Biros, J. Rebek, Jr., *Chem. Soc. Rev.* **2007**, *36*, 93–104; c) J. Rebek, Jr., *Acc. Chem. Res.* **2009**, *42*, 1660–1668; d) D. Ajami, J. Rebek, Jr., *Acc. Chem. Res.* **2013**, *46*, 990–999.
- [21] For selected reviews on further research on containers, see a) R. Warmuth, J. Yoon, *Acc. Chem. Res.* **2001**, *34*, 95–105; b) M. D. Pluth, K. N. Raymond, *Chem. Soc. Rev.* **2007**, *36*, 161–171; c) M. Yoshizawa, J. K. Klosterman, M. Fujita, *Angew. Chem. Int. Ed.* **2009**, *48*, 3418–3438; *Angew. Chem.* **2009**, *121*, 3470–3490; d) T. R. Cook, V. Vajpayee, M. H. Lee, P. J. Stang, K. W. Chi, *Acc. Chem. Res.* **2013**, *46*, 2464–2474; e) M. M. J. Smulders, I. A. Riddell, C. Browne, J. R. Nitschke, *Chem. Soc. Rev.* **2013**, *42*, 1728–1754.
- [22] a) J. A. Olsen, D. W. Banner, P. Seiler, U. O. Sander, A. D'Arcy, M. Stihle, K. Müller, F. Diederich, *Angew. Chem. Int. Ed.* **2003**,

- 42, 2507–2511; *Angew. Chem.* **2003**, *115*, 2611–2615; b) J. A. Olsen, D. W. Banner, P. Seiler, B. Wagner, T. Tschopp, U. Obst-Sander, M. Kansy, K. Müller, F. Diederich, *ChemBioChem* **2004**, *5*, 666–675.
- [23] R. Paulini, K. Müller, F. Diederich, *Angew. Chem. Int. Ed.* **2005**, *44*, 1788–1805; *Angew. Chem.* **2005**, *117*, 1820–1839.
- [24] A. Choudhary, D. Gandla, G. R. Krow, R. T. Raines, *J. Am. Chem. Soc.* **2009**, *131*, 7244–7246.
- [25] a) L. A. Hardegger, B. Kuhn, B. Spinnler, L. Anselm, R. Ecabert, M. Stihle, B. Gsell, R. Thoma, J. Diez, J. Benz, J.-M. Plancher, G. Hartmann, D. W. Banner, W. Haap, F. Diederich, *Angew. Chem. Int. Ed.* **2011**, *50*, 314–318; *Angew. Chem.* **2011**, *123*, 329–334; b) L. A. Hardegger, B. Kuhn, B. Spinnler, L. Anselm, R. Ecabert, M. Stihle, B. Gsell, R. Thoma, J. Diez, J. Benz, J.-M. Plancher, G. Hartmann, Y. Isshiki, K. Morikami, N. Shimma, W. Haap, D. W. Banner, F. Diederich, *ChemMedChem* **2011**, *6*, 2048–2054.
- [26] W. P. Jencks, *Catalysis in Chemistry and Enzymology*, McGraw Hill, New York, **1989**.
- [27] E. L. Piatnitski, R. A. Flowers II, K. Deshayes, *Chem. Eur. J.* **2000**, *6*, 999–1006.
- [28] F. Biedermann, W. M. Nau, H.-J. Schneider, *Angew. Chem. Int. Ed.* **2014**, *53*, 11158–11171; *Angew. Chem.* **2014**, *126*, 11338–11352.
- [29] H.-J. Schneider, *Angew. Chem. Int. Ed.* **2009**, *48*, 3924–3977; *Angew. Chem.* **2009**, *121*, 3982–4036.
- [30] a) R. L. VanEtten, J. F. Sebastian, G. A. Clowes, M. L. Bender, *J. Am. Chem. Soc.* **1967**, *89*, 3242–3253; b) W. Saenger, *Angew. Chem. Int. Ed. Engl.* **1980**, *19*, 344–362; *Angew. Chem.* **1980**, *92*, 343–361; c) T. D. Kühne, R. Z. Khaliullin, *Nat. Commun.* **2013**, *4*, 1450.
- [31] a) C. Carey, Y.-K. Cheng, P. J. Rossky, *Chem. Phys.* **2000**, *258*, 415–425; b) N. T. Southall, K. A. Dill, *J. Phys. Chem. B* **2000**, *104*, 1326–1331; c) T. Young, L. Hua, X. Huang, R. Abel, R. Friesner, B. J. Berne, *Proteins* **2010**, *78*, 1856–1869; d) P. Setny, R. Baron, J. A. McCammon, *J. Chem. Theory Comput.* **2010**, *6*, 2866–2871.
- [32] a) J. Lagona, P. Mukhopadhyay, S. Chakrabarti, L. Isaacs, *Angew. Chem. Int. Ed.* **2005**, *44*, 4844–4870; *Angew. Chem.* **2005**, *117*, 4922–4949; b) E. Masson, X. Ling, R. Joseph, L. Kyeremeh-Mensah, X. Lu, *RSC Adv.* **2012**, *2*, 1213–1247.
- [33] S. Moghaddam, C. Yang, M. Rekharsky, Y. H. Ko, K. Kim, Y. Inoue, M. K. Gilson, *J. Am. Chem. Soc.* **2011**, *133*, 3570–3581.
- [34] L. Isaacs, S.-K. Park, S. Liu, Y. H. Ko, N. Selvapalam, Y. Kim, H. Kim, P. Y. Zavalij, G.-H. Kim, H.-S. Lee, K. Kim, *J. Am. Chem. Soc.* **2005**, *127*, 18000–18001.
- [35] a) C. Marquez, W. M. Nau, *Angew. Chem. Int. Ed.* **2001**, *40*, 4387–4390; *Angew. Chem.* **2001**, *113*, 4515–4518; b) W. M. Nau, M. Florea, K. I. Assaf, *Isr. J. Chem.* **2011**, *51*, 559–577; c) F. Biedermann, V. D. Uzunova, O. A. Scherman, W. M. Nau, A. De Simone, *J. Am. Chem. Soc.* **2012**, *134*, 15318–15323; d) F. Biedermann, M. Vendruscolo, O. A. Scherman, A. De Simone, W. M. Nau, *J. Am. Chem. Soc.* **2013**, *135*, 14879–14888.
- [36] For a recent experimental study showing the importance of the gain in cohesive solvent–solvent interactions during molecular recognition processes, see L. Yang, C. Adam, G. S. Nichol, S. L. Cockcroft, *Nat. Chem.* **2013**, *5*, 1006–1010.
- [37] For an ITC study of the capsular binding of ammonium ions in water with a strong entropic driving force arising from the desolvation of the ions, see C. Sgarlata, J. S. Mugridge, M. D. Pluth, B. E. F. Tiedemann, V. Zito, G. Arena, K. N. Raymond, *J. Am. Chem. Soc.* **2010**, *132*, 1005–1009.
- [38] L. Cao, M. Šekutor, P. Y. Zavalij, K. Mlinarić-Majerski, R. Glaser, L. Isaacs, *Angew. Chem. Int. Ed.* **2014**, *53*, 988–993; *Angew. Chem.* **2014**, *126*, 1006–1011.
- [39] M. Nakasako, *Philos. Trans. R. Soc. London Ser. B* **2004**, *359*, 1191–1206.
- [40] J. D. Dunitz, *Science* **1994**, *264*, 670.
- [41] a) A. T. García-Sosa, R. L. Mancera, P. M. Dean, *J. Mol. Model.* **2003**, *9*, 172–182; b) M. L. Verdonk, G. Chessari, J. C. Cole, M. J. Hartshorn, C. W. Murray, J. W. M. Nissink, R. D. Taylor, R. Taylor, *J. Med. Chem.* **2005**, *48*, 6504–6515; c) N. Huang, B. K. Shoichet, *J. Med. Chem.* **2008**, *51*, 4862–4865; d) J. Michel, J. Tirado-Rives, W. L. Jorgensen, *J. Am. Chem. Soc.* **2009**, *131*, 15403–15411; e) S. B. A. de Beer, N. P. E. Vermeulen, C. Oostenbrink, *Curr. Top. Med. Chem.* **2010**, *10*, 55–66; f) R. Baron, P. Setny, J. A. McCammon, *J. Am. Chem. Soc.* **2010**, *132*, 12091–12097; g) Ref. [31c]; h) L. Wang, B. J. Berne, R. A. Friesner, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1326–1330; i) S. Riniker, L. J. Barandun, F. Diederich, O. Krämer, A. Steffen, W. F. van Gunsteren, *J. Comput.-Aided Mol. Des.* **2012**, *26*, 1293–1309; j) R. A. Pearlstein, W. Sherman, R. Abel, *Proteins* **2013**, *81*, 1509–1526; k) A. T. García-Sosa, *J. Chem. Inf. Model.* **2013**, *53*, 1388–1405; l) S. Barelier, S. E. Boyce, I. Fish, M. Fischer, D. B. Goodin, B. K. Shoichet, *PLoS One* **2013**, *8*, e69153; m) K. Haider, D. J. Huggins, *J. Chem. Inf. Model.* **2013**, *53*, 2571–2586; n) P. W. Snyder, M. R. Lockett, D. T. Moustakas, G. M. Whitesides, *Eur. Phys. J. Spec. Top.* **2014**, *223*, 853–891.
- [42] M. C. Witschel, H. W. Höffken, M. Seet, L. Parra, T. Mietzner, F. Thater, R. Niggeweg, F. Röhl, B. Illarionov, F. Rohdich, J. Kaiser, M. Fischer, A. Bacher, F. Diederich, *Angew. Chem. Int. Ed.* **2011**, *50*, 7931–7935; *Angew. Chem.* **2011**, *123*, 8077–8081.
- [43] a) B. Masjost, P. Ballmer, E. Borroni, G. Zürcher, F. K. Winkler, R. Jakob-Roetne, F. Diederich, *Chem. Eur. J.* **2000**, *6*, 971–982; b) C. Lerner, A. Ruf, V. Gramlich, B. Masjost, G. Zürcher, R. Jakob-Roetne, E. Borroni, F. Diederich, *Angew. Chem. Int. Ed.* **2001**, *40*, 4040–4042; *Angew. Chem.* **2001**, *113*, 4164–4166; c) C. Lerner, B. Masjost, A. Ruf, V. Gramlich, R. Jakob-Roetne, G. Zürcher, E. Borroni, F. Diederich, *Org. Biomol. Chem.* **2003**, *1*, 42–49.
- [44] a) J. Borgulya, H. Bruderer, K. Bernauer, G. Zürcher, M. Da Prada, *Helv. Chim. Acta* **1989**, *72*, 952–968; b) J. Vidgren, L. A. Svensson, A. Liljas, *Nature* **1994**, *368*, 354–358; c) P. T. Männistö, S. Kaakkola, *Pharmacol. Rev.* **1999**, *51*, 593–628; d) G. M. Keating, K. A. Lyseng-Williamson, *CNS Drugs* **2005**, *19*, 165–184.
- [45] a) M. Ellermann, R. Jakob-Roetne, C. Lerner, E. Borroni, D. Schlatter, D. Roth, A. Ehler, M. G. Rudolph, F. Diederich, *Angew. Chem. Int. Ed.* **2009**, *48*, 9092–9096; *Angew. Chem.* **2009**, *121*, 9256–9260; b) M. Ellermann, R. Paulini, R. Jakob-Roetne, C. Lerner, E. Borroni, D. Roth, A. Ehler, W. B. Schweizer, D. Schlatter, M. G. Rudolph, F. Diederich, *Chem. Eur. J.* **2011**, *17*, 6369–6381; c) M. Ellermann, C. Lerner, G. Burgy, A. Ehler, C. Bissantz, R. Jakob-Roetne, R. Paulini, O. Allemann, H. Tissot, D. Grünstein, M. Stihle, F. Diederich, M. G. Rudolph, *Acta Crystallogr. Sect. D* **2012**, *68*, 253–260; d) for a recent survey of crystal structures of apo-COMT, see A. Ehler, J. Benz, D. Schlatter, M. G. Rudolph, *Acta Crystallogr. Sect. D* **2014**, *70*, 2163–2174.
- [46] a) J. R. Hörtnner, T. Ritschel, B. Stengl, C. Kramer, W. B. Schweizer, B. Wagner, M. Kansy, G. Klebe, F. Diederich, *Angew. Chem. Int. Ed.* **2007**, *46*, 8266–8269; *Angew. Chem.* **2007**, *119*, 8414–8417; b) P. C. Kohler, T. Ritschel, W. B. Schweizer, G. Klebe, F. Diederich, *Chem. Eur. J.* **2009**, *15*, 10809–10817; c) T. Ritschel, S. Hoertner, A. Heine, F. Diederich, G. Klebe, *ChemBioChem* **2009**, *10*, 716–727.
- [47] U. Grädler, H.-D. Gerber, D. M. Goodenough-Lashua, G. A. Garcia, R. Ficner, K. Reuter, M. T. Stubbs, G. Klebe, *J. Mol. Biol.* **2001**, *306*, 455–467.
- [48] Q.-Q. Wang, V. W. Day, K. Bowman-James, *Angew. Chem. Int. Ed.* **2012**, *51*, 2119–2123; *Angew. Chem.* **2012**, *124*, 2161–2165.
- [49] T. Ritschel, P. C. Kohler, G. Neudert, A. Heine, F. Diederich, G. Klebe, *ChemMedChem* **2009**, *4*, 2012–2023.

- [50] a) F. Immekus, L. J. Barandun, M. Betz, F. Debaene, S. Petiot, S. Sanglier-Cianferani, K. Reuter, F. Diederich, G. Klebe, *ACS Chem. Biol.* **2013**, *8*, 1163–1178; b) L. J. Barandun, F. Immekus, P. C. Kohler, T. Ritschel, A. Heine, P. Orlando, G. Klebe, F. Diederich, *Acta Crystallogr. Sect. D* **2013**, *69*, 1798–1807.
- [51] L. J. Barandun, F. Immekus, P. C. Kohler, S. Tonazzi, B. Wagner, S. Wendelspiess, T. Ritschel, A. Heine, M. Kansy, G. Klebe, F. Diederich, *Chem. Eur. J.* **2012**, *18*, 9246–9257.
- [52] R. Brenk, M. T. Stubbs, A. Heine, K. Reuter, G. Klebe, *ChemBioChem* **2003**, *4*, 1066–1077.
- [53] a) K. L. Whalen, M. A. Spies, *J. Chem. Inf. Model.* **2013**, *53*, 2349–2359; b) M. A. Spies, *ACS Med. Chem. Lett.* **2013**, *4*, 895–897.
- [54] a) S. Grüner, M. Neeb, L. J. Barandun, F. Sielaff, C. Hohn, S. Kojima, T. Steinmetzer, F. Diederich, G. Klebe, *Biochim. Biophys. Acta Gen. Subj.* **2014**, *1840*, 2843–2850; b) M. Neeb, P. Czodrowski, A. Heine, L. J. Barandun, C. Hohn, F. Diederich, G. Klebe, *J. Med. Chem.* **2014**, *57*, 5554–5565; c) M. Neeb, M. Betz, A. Heine, L. J. Barandun, C. Hohn, F. Diederich, G. Klebe, *J. Med. Chem.* **2014**, *57*, 5566–5578.
- [55] a) V. M. Krishnamurthy, G. K. Kaufman, A. R. Urbach, I. Gitlin, K. L. Gudiksen, D. B. Weibel, G. M. Whitesides, *Chem. Rev.* **2008**, *108*, 946–1051; b) P. W. Snyder, J. Mecnović, D. T. Moustakas, S. W. Thomas III, M. Harder, E. T. Mack, M. R. Lockett, A. Héroux, W. Sherman, G. M. Whitesides, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 17889–17894; c) B. Breiten, M. R. Lockett, W. Sherman, S. Fujita, M. Al-Sayah, H. Lange, C. M. Bowers, A. Héroux, G. Krilov, G. M. Whitesides, *J. Am. Chem. Soc.* **2013**, *135*, 15579–15584; d) M. R. Lockett, H. Lange, B. Breiten, A. Héroux, W. Sherman, D. Rappoport, P. O. Yau, P. W. Snyder, G. M. Whitesides, *Angew. Chem. Int. Ed.* **2013**, *52*, 7714–7717; *Angew. Chem.* **2013**, *125*, 7868–7871.
- [56] a) A. Biela, F. Sielaff, F. Terwesten, A. Heine, T. Steinmetzer, G. Klebe, *J. Med. Chem.* **2012**, *55*, 6094–6110; b) A. Biela, M. Khayat, H. Tan, J. Kong, A. Heine, D. Hangauer, G. Klebe, *J. Mol. Biol.* **2012**, *418*, 350–366.
- [57] a) L. Englert, A. Biela, M. Zayed, A. Heine, D. Hangauer, G. Klebe, *Biochim. Biophys. Acta Gen. Subj.* **2010**, *1800*, 1192–1202; b) A. Biela, N. N. Nasief, M. Betz, A. Heine, D. Hangauer, G. Klebe, *Angew. Chem. Int. Ed.* **2013**, *52*, 1822–1828; *Angew. Chem.* **2013**, *125*, 1868–1876; c) S. G. Krimmer, M. Betz, A. Heine, G. Klebe, *ChemMedChem* **2014**, *9*, 833–846.
- [58] a) A. Biela, M. Betz, A. Heine, G. Klebe, *ChemMedChem* **2012**, *7*, 1423–1434; b) N. N. Nasief, H. Tan, J. Kong, D. Hangauer, *J. Med. Chem.* **2012**, *55*, 8283–8302; c) N. N. Nasief, D. Hangauer, *J. Med. Chem.* **2014**, *57*, 2315–2333.
- [59] M. Zürcher, F. Diederich, *J. Org. Chem.* **2008**, *73*, 4345–4361.
- [60] E. Fischer, *Chem. Ber.* **1894**, *27*, 2985–2993.
- [61] A. Nicholls, G. B. McGaughey, R. P. Sheridan, A. C. Good, G. Warren, M. Mathieu, S. W. Muchmore, S. P. Brown, J. A. Grant, J. A. Haigh, N. Nevins, A. N. Jain, B. Kelley, *J. Med. Chem.* **2010**, *53*, 3862–3886.
- [62] a) M. Weisel, E. Proschak, G. Schneider, *Chem. Cent. J.* **2007**, *1*, 7; b) S. Henrich, O. M. H. Salo-Ahen, B. Huang, F. Rippmann, G. Cruciani, R. C. Wade, *J. Mol. Recognit.* **2010**, *23*, 209–219; c) X. Zheng, L. Gan, E. Wang, J. Wang, *AAPS J.* **2013**, *15*, 228–241; d) *Focus on Structural Biology, Vol. 8* (Ed.: I. Roterman-Konieczna), Springer, Dordrecht, **2013**; e) D. Rognan in *Chemical Genomics and Proteomics* (Eds.: F. Dargatzas, A. Guttman, G. Dormán), CRC, Boca Raton, **2013**, pp. 173–204.
- [63] S. Mecozzi, J. Rebek, Jr., *Chem. Eur. J.* **1998**, *4*, 1016–1022.
- [64] a) J. R. Moran, S. Karbach, D. J. Cram, *J. Am. Chem. Soc.* **1982**, *104*, 5826–5828; b) D. J. Cram, *Science* **1983**, *219*, 1177–1183.
- [65] a) L. Trembleau, J. Rebek, Jr., *Science* **2003**, *301*, 1219–1220; b) A. Scarso, L. Trembleau, J. Rebek, Jr., *Angew. Chem. Int. Ed.* **2003**, *42*, 5499–5502; *Angew. Chem.* **2003**, *115*, 5657–5660; c) B. W. Purse, J. Rebek, Jr., *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 2530–2534.
- [66] a) N. O. B. Lüttschwager, T. N. Wassermann, R. A. Mata, M. A. Suhm, *Angew. Chem. Int. Ed.* **2013**, *52*, 463–466; *Angew. Chem.* **2013**, *125*, 482–485; b) J. N. Byrd, R. J. Bartlett, J. A. Montgomery, Jr., *J. Phys. Chem. A* **2014**, *118*, 1706–1712.
- [67] J. P. Lowe, *Prog. Phys. Org. Chem.* **1968**, *6*, 1–80.
- [68] K.-D. Zhang, D. Ajami, J. V. Gavette, J. Rebek, Jr., *J. Am. Chem. Soc.* **2014**, *136*, 5264–5266.
- [69] For reviews of the various switching processes, see a) V. A. Azov, A. Beeby, M. Cacciarini, A. G. Cheetham, F. Diederich, M. Frei, J. K. Gimzewski, V. Gramlich, B. Hecht, B. Jaun, T. Latychevskaia, A. Lieb, Y. Lill, F. Marotti, A. Schlegel, R. R. Schlittler, P. J. Skinner, P. Seiler, Y. Yamakoshi, *Adv. Funct. Mater.* **2006**, *16*, 147–156; b) I. Pochorovski, F. Diederich, *Isr. J. Chem.* **2012**, *52*, 20–29; c) I. Pochorovski, F. Diederich, *Acc. Chem. Res.* **2014**, *47*, 2096–2105.
- [70] a) T. Gottschalk, B. Jaun, F. Diederich, *Angew. Chem. Int. Ed.* **2007**, *46*, 260–264; *Angew. Chem.* **2007**, *119*, 264–268; b) T. Gottschalk, P. D. Jarowski, F. Diederich, *Tetrahedron* **2008**, *64*, 8307–8317.
- [71] 1000 ps simulation time with 1.0 fs time steps, MMFF94s force field, 300 K, GB/SA model for CHCl₃; MacroModel, version 9.5; Schrödinger LLC, New York, NY, USA, **2007**.
- [72] J. Hornung, D. Fankhauser, L. D. Shirtcliff, A. Praetorius, W. B. Schweizer, F. Diederich, *Chem. Eur. J.* **2011**, *17*, 12362–12371.
- [73] a) Ref. [18b]; b) S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, A. Fujii, *J. Phys. Chem. A* **2006**, *110*, 10163–10168; c) K. Shibasaki, A. Fujii, N. Mikami, S. Tsuzuki, *J. Phys. Chem. A* **2007**, *111*, 753–758; d) M. Nishio, Y. Umezawa, J. Fantini, M. S. Weiss, P. Chakrabarti, *Phys. Chem. Chem. Phys.* **2014**, *16*, 12648–12683.
- [74] D. Fankhauser, D. Kolarski, W. R. Grüning, F. Diederich, *Eur. J. Org. Chem.* **2014**, 3575–3583.
- [75] a) E. Dalcanele, P. Soncini, G. Bacchilega, F. Ugozzoli, *J. Chem. Soc. Chem. Commun.* **1989**, 500–502; b) P. Soncini, S. Bon-signore, E. Dalcanele, F. Ugozzoli, *J. Org. Chem.* **1992**, *57*, 4608–4612.
- [76] a) M. Rohmer, *Nat. Prod. Rep.* **1999**, *16*, 565–574; b) M. Rodríguez-Concepción, A. Boronat, *Plant Physiol.* **2002**, *130*, 1079–1089; c) W. Eisenreich, A. Bacher, D. Arigoni, F. Rohdich, *Cell. Mol. Life Sci.* **2004**, *61*, 1401–1426; d) T. Gräwert, M. Groll, F. Rohdich, A. Bacher, W. Eisenreich, *Cell. Mol. Life Sci.* **2011**, *68*, 3797–3814.
- [77] T. Masini, B. S. Kroezen, A. K. H. Hirsch, *Drug Discovery Today* **2013**, *18*, 1256–1262.
- [78] H. Joomaa, J. Wiesner, S. Sanderbrand, B. Altincicek, C. Weidemeyer, M. Hintz, J. Türbachova, M. Eberl, J. Zeidler, H. K. Lichtenthaler, D. Soldati, E. Beck, *Science* **1999**, *285*, 1573–1576.
- [79] A. K. H. Hirsch, F. R. Fischer, F. Diederich, *Angew. Chem. Int. Ed.* **2007**, *46*, 338–352; *Angew. Chem.* **2007**, *119*, 342–357.
- [80] a) L. Miallau, M. S. Alphey, L. E. Kemp, G. A. Leonard, S. M. McSwainey, S. Hecht, A. Bacher, W. Eisenreich, F. Rohdich, W. N. Hunter, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9173–9178 (PDB ID: 1OJ4); for other X-ray crystal structures of IspE, see b) J. Wada, T. Kuzuyama, S. Satoh, S. Kuramitsu, S. Yokoyama, S. Unzai, J. R. H. Tame, S.-Y. Park, *J. Biol. Chem.* **2003**, *278*, 30022–30027 (PDB ID: 1UEK); c) C. M. Crane, A. K. H. Hirsch, M. S. Alphey, T. Sgraja, S. Lauw, V. Illarionova, F. Rohdich, W. Eisenreich, W. N. Hunter, A. Bacher, F. Diederich, *ChemMedChem* **2008**, *3*, 91–101 (PDB ID: 2V2Q, 2V2V); d) T. Sgraja, M. S. Alphey, S. Ghilagaber, R. Marquez, M. N. Robertson, J. L. Hemmings, S. Lauw, F. Rohdich, A. Bacher, W. Eisenreich, V. Illarionova, W. N. Hunter, *FEBS J.* **2008**, *275*, 2779–2794 (PDB ID: 2V8P, 2V34, 2V2Z); e) J. Kalinowska-Tłuścik, L. Miallau, M. Gabrielsen, G. A. Leonard, S. M.

- McSweeney, W. N. Hunter, *Acta Crystallogr. Sect. F* **2010**, *66*, 237–241 (PDB ID: 2WW4); f) S. Shan, X. Chen, T. Liu, H. Zhao, Z. Rao, Z. Lou, *FASEB J.* **2011**, *25*, 1577–1584 (PDB ID: 3PYD, 3PYE, 3PYF, 3PYG); g) to be published: PDB ID: 4DXL, 4ED4, 4EMD.
- [81] a) A. C. Dar, K. M. Shokat, *Annu. Rev. Biochem.* **2011**, *80*, 769–795; b) H. Patterson, R. Nibbs, I. McInnes, S. Siebert, *Clin. Exp. Immunol.* **2014**, *176*, 1–10.
- [82] a) L. M. Wilhelmsson, *Q. Rev. Biophys.* **2010**, *43*, 159–183; b) A. A. Tanpure, M. G. Pawar, S. G. Srivatsan, *Isr. J. Chem.* **2013**, *53*, 366–378; c) M. Suchý, R. H. E. Hudson, *J. Org. Chem.* **2014**, *79*, 3336–3347; d) D. D. Haveliwala, N. R. Kamdar, P. T. Mistry, S. K. Patel, *Nucleosides Nucleotides Nucleic Acids* **2014**, *33*, 80–91.
- [83] a) P. A. Jones, D. Takai, *Science* **2001**, *293*, 1068–1070; b) S. Bareyt, T. Carell, *Angew. Chem. Int. Ed.* **2008**, *47*, 181–184; *Angew. Chem.* **2008**, *120*, 187–190; c) R. Bonasio, S. Tu, D. Reinberg, *Science* **2010**, *330*, 612–616; d) S. Feng, S. E. Jacobsen, W. Reik, *Science* **2010**, *330*, 622–627; e) M. Münzel, D. Globisch, T. Carell, *Angew. Chem. Int. Ed.* **2011**, *50*, 6460–6468; *Angew. Chem.* **2011**, *123*, 6588–6596.
- [84] A. P. Schütz, S. Osawa, J. Mathis, A. K. H. Hirsch, B. Bernet, B. Illarionov, M. Fischer, A. Bacher, F. Diederich, *Eur. J. Org. Chem.* **2012**, 3278–3287.
- [85] Some K_i values were calculated by the Cheng–Prusoff Equation: Y.-C. Cheng, W. H. Prusoff, *Biochem. Pharmacol.* **1973**, *22*, 3099–3108.
- [86] a) A. K. H. Hirsch, S. Lauw, P. Gersbach, W. B. Schweizer, F. Rohdich, W. Eisenreich, A. Bacher, F. Diederich, *ChemMedChem* **2007**, *2*, 806–810; b) A. K. H. Hirsch, M. S. Alpey, S. Lauw, M. Seet, L. Barandun, W. Eisenreich, F. Rohdich, W. N. Hunter, A. Bacher, F. Diederich, *Org. Biomol. Chem.* **2008**, *6*, 2719–2730; c) P. Mombelli, C. Le Chapelain, N. Munzinger, E. Joliat, B. Illarionov, W. B. Schweizer, A. K. H. Hirsch, M. Fischer, A. Bacher, F. Diederich, *Eur. J. Org. Chem.* **2013**, 1068–1079.
- [87] A recent report on 4-quinazolinone-based inhibitors of *EcIspE* with IC_{50} values in the low micromolar range could not be confirmed in our robust assay. We presume that the compounds bind to the auxiliary enzyme firefly luciferase rather than to *EcIspE*: N. Tidten-Luksch, R. Grimaldi, L. S. Torrie, J. A. Frearson, W. N. Hunter, R. Brenk, *PLoS One* **2012**, *7*, e35792.
- [88] P. R. Gerber, K. Müller, *J. Comput.-Aid. Mol. Des.* **1995**, *9*, 251–268.
- [89] a) G. Wuitschik, M. Rogers-Evans, A. Buckl, M. Bernasconi, M. Märki, T. Godel, H. Fischer, B. Wagner, I. Parrilla, F. Schuler, J. Schneider, A. Alker, W. B. Schweizer, K. Müller, E. M. Carreira, *Angew. Chem. Int. Ed.* **2008**, *47*, 4512–4515; *Angew. Chem.* **2008**, *120*, 4588–4591; b) G. Wuitschik, E. M. Carreira, B. Wagner, H. Fischer, I. Parrilla, F. Schuler, M. Rogers-Evans, K. Müller, *J. Med. Chem.* **2010**, *53*, 3227–3246.
- [90] a) W. L. Jorgensen, J. Pranata, *J. Am. Chem. Soc.* **1990**, *112*, 2008–2010; b) T. J. Murray, S. C. Zimmerman, *J. Am. Chem. Soc.* **1992**, *114*, 4010–4011; c) B. A. Blight, C. A. Hunter, D. A. Leigh, H. McNab, P. I. T. Thomson, *Nat. Chem.* **2011**, *3*, 244–248.
- [91] C. S. Leung, S. S. F. Leung, J. Tirado-Rives, W. L. Jorgensen, *J. Med. Chem.* **2012**, *55*, 4489–4500.
- [92] A. P. Schütz, S. Locher, B. Bernet, B. Illarionov, M. Fischer, A. Bacher, F. Diederich, *Eur. J. Org. Chem.* **2013**, 880–887.
- [93] I. Hale, P. M. O'Neill, N. G. Berry, A. Odom, R. Sharma, *MedChemComm* **2012**, *3*, 418–433.
- [94] a) L. E. Kemp, C. S. Bond, W. N. Hunter, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 6591–6596; b) S. Steinbacher, J. Kaiser, J. Wungstintaweekul, S. Hecht, W. Eisenreich, S. Gerhardt, A. Bacher, F. Rohdich, *J. Mol. Biol.* **2002**, *316*, 79–88.
- [95] A total of 56 X-ray crystal structures of IspF enzymes have been reported to date in the PDB (July 2014). Early structures are reviewed in W. N. Hunter, *J. Biol. Chem.* **2007**, *282*, 21573–21577.
- [96] C. M. Crane, J. Kaiser, N. L. Ramsden, S. Lauw, F. Rohdich, W. Eisenreich, W. N. Hunter, A. Bacher, F. Diederich, *Angew. Chem. Int. Ed.* **2006**, *45*, 1069–1074; *Angew. Chem.* **2006**, *118*, 1082–1087.
- [97] The small pocket I in *EcIspF* contains the side chain of Asp46'.
- [98] P. E. F. O'Rourke, K.-T. Justyna, P. K. Fyfe, A. Dawson, W. N. Hunter, *BMC Struct. Biol.* **2014**, *14*, 1–12.
- [99] Crystal data originate from: a) Ref. [94] (PDB ID: 1GX1, 1JY8, 1U3L, 1U3P, 1U40, 1U43); b) Ref. [96] (PDB ID: 2GZL); c) S. B. Richard, *J. Biol. Chem.* **2002**, *277*, 8667–8672 (PDB ID: 1KNJ, 1KNK); d) T. Sgraja, L. E. Kemp, N. Ramsden, W. N. Hunter, *Acta Crystallogr. Sect. F* **2005**, *61*, 625–629 (PDB ID: 1YQN).
- [100] a) C. Baumgartner, C. Eberle, F. Diederich, S. Lauw, F. Rohdich, W. Eisenreich, A. Bacher, *Helv. Chim. Acta* **2007**, *90*, 1043–1068; b) J. Geist, ETH dissertation No 19777, **2011**.
- [101] It is not always possible to replace a ribose moiety, which tightly interacts with an Asp or a Glu side chain, see R. Paulini, C. Trindler, C. Lerner, L. Brändli, W. B. Schweizer, R. Jakob-Roetne, G. Zürcher, E. Borroni, F. Diederich, *ChemMedChem* **2006**, *1*, 340–357.
- [102] a) N. L. Ramsden, L. Buetow, A. Dawson, L. A. Kemp, V. Ulaganathan, R. Brenk, G. Klebe, W. N. Hunter, *J. Med. Chem.* **2009**, *52*, 2531–2542; b) D. W. Begley, R. C. Hartley, D. R. Davies, T. E. Edwards, J. T. Leonard, J. Abendroth, C. A. Burris, J. Bhandari, P. J. Myler, B. L. Staker, L. J. Stewart, *J. Struct. Funct. Genomics* **2011**, *12*, 63–76; c) Z. Zhang, S. Jakkaraju, J. Blain, K. Gogol, L. Zhao, R. C. Hartley, C. A. Karlsson, B. L. Staker, T. E. Edwards, L. J. Stewart, P. J. Myler, M. Clare, D. W. Begley, J. R. Horn, T. J. Hagen, *Bioorg. Med. Chem. Lett.* **2013**, *23*, 6860–6863.
- [103] J. G. Geist, S. Lauw, V. Illarionova, B. Illarionov, M. Fischer, T. Gräwert, F. Rohdich, W. Eisenreich, J. Kaiser, M. Groll, C. Scheurer, S. Wittlin, J. L. Alonso-Gómez, W. B. Schweizer, A. Bacher, F. Diederich, *ChemMedChem* **2010**, *5*, 1092–1101.
- [104] A. H. Fairlamb, P. Blackburn, P. Ulrich, B. T. Chait, A. Cerami, *Science* **1985**, *227*, 1485–1487.
- [105] M. A. Comini, L. Flohé, *Trypanosomatid Dis. Mol. Routes Drug Discovery* **2013**, 167–199.
- [106] A search in the PDB for trypanothione reductase found 27 structures (July 2014).
- [107] a) C. H. Faerman, S. N. Savvides, C. Strickland, M. A. Breidenbach, J. A. Ponasik, B. Ganem, D. Ripoll, R. L. Krauth-Siegel, P. A. Karplus, *Bioorg. Med. Chem.* **1996**, *4*, 1247–1253; b) C. S. Bond, Y. Zhang, M. Berriman, M. L. Cunningham, A. H. Fairlamb, W. N. Hunter, *Structure* **1999**, *7*, 81–89; c) B. Stump, C. Eberle, M. Kaiser, R. Brun, R. L. Krauth-Siegel, F. Diederich, *Org. Biomol. Chem.* **2008**, *6*, 3935–3947.
- [108] a) S. Bonse, C. Santelli-Rouvier, J. Barbe, R. L. Krauth-Siegel, *J. Med. Chem.* **1999**, *42*, 5448–5454; b) B. Chitkul, M. Bradley, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2367–2369; c) A. Saravanamuthu, T. J. Vickers, C. S. Bond, M. R. Peterson, W. N. Hunter, A. H. Fairlamb, *J. Biol. Chem.* **2004**, *279*, 29493–29500; d) C. Eberle, B. S. Lauber, D. Fankhauser, M. Kaiser, R. Brun, R. L. Krauth-Siegel, F. Diederich, *ChemMedChem* **2011**, *6*, 292–301; e) M. H. Duyzend, C. T. Clark, S. L. Simons, W. B. Johnson, A. M. Larson, A. M. Leconte, A. W. Wills, M. Ginder-Vogel, A. K. Wilhelm, J. A. Czechowicz, D. G. Alberg, *J. Enzyme Inhib. Med. Chem.* **2012**, *27*, 784–794; f) P. Baiocco, G. Poce, S. Alfonso, M. Cocozza, G. C. Porretta, G. Colotti, M. Biava, F. Moraca, M. Botta, V. Yardley, A. Fiorillo, A. Lantella, F. Malatesta, A. Ilari, *ChemMedChem* **2013**, *8*, 1175–1183.

- [109] a) J. L. Richardson, I. R. E. Nett, D. C. Jones, M. H. Abdille, I. H. Gilbert, A. H. Fairlamb, *ChemMedChem* **2009**, *4*, 1333–1340; b) S. Patterson, D. C. Jones, E. J. Shanks, J. A. Frearson, I. H. Gilbert, P. G. Wyatt, A. H. Fairlamb, *ChemMedChem* **2009**, *4*, 1341–1353; c) E. Persch, S. Bryson, N. K. Todoroff, C. Eberle, J. Thelemann, N. Dirdjaja, M. Kaiser, M. Weber, H. Derbani, R. Brun, G. Schneider, E. F. Pai, R. L. Krauth-Siegel, F. Diederich, *ChemMedChem* **2014**, *9*, 1880–1891.
- [110] E. M. Jacoby, I. Schlichting, C. B. Lantwin, W. Kabsch, R. L. Krauth-Siegel, *Proteins Struct. Funct. Genet.* **1996**, *24*, 73–80.
- [111] a) H.-J. Schneider, *Chem. Soc. Rev.* **1994**, *23*, 227–234; b) Y. Xu, Y. Nakajima, K. Ito, H. Zheng, H. Oyama, U. Heiser, T. Hoffmann, U.-T. Gärtner, H.-U. Demuth, T. Yoshimoto, *J. Mol. Biol.* **2008**, *375*, 708–719.
- [112] S. Patterson, M. S. Alphey, D. C. Jones, E. J. Shanks, I. P. Street, J. A. Frearson, P. G. Wyatt, I. H. Gilbert, A. H. Fairlamb, *J. Med. Chem.* **2011**, *54*, 6514–6530.
- [113] a) R. Fernandez-Gomez, M. Moutiez, M. Aumercier, G. Bethegnies, M. Luyckx, A. Ouaisi, A. Tartar, C. Sergheraert, *Int. J. Antimicrob. Agents* **1995**, *6*, 111–118; b) S. Parveen, M. O. F. Khan, S. E. Austin, S. L. Croft, V. Yardley, P. Rock, K. T. Douglas, *J. Med. Chem.* **2005**, *48*, 8087–8097; c) B. Stump, M. Kaiser, R. Brun, R. L. Krauth-Siegel, F. Diederich, *ChemMedChem* **2007**, *2*, 1708–1712; d) C. Eberle, J. A. Burkhard, B. Stump, M. Kaiser, R. Brun, R. L. Krauth-Siegel, F. Diederich, *ChemMedChem* **2009**, *4*, 2034–2044.
- [114] a) I. D. Kuntz, K. Chen, K. A. Sharp, P. A. Kollman, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9997–10002; b) A. L. Hopkins, C. R. Groom, A. Alex, *Drug Discovery Today* **2004**, *9*, 430–431; c) C. H. Reynolds, B. A. Tounge, S. D. Bembeneck, *J. Med. Chem.* **2008**, *51*, 2432–2438.
- [115] a) P. D. Leeson, B. Springthorpe, *Nat. Rev. Drug Discovery* **2007**, *6*, 881–890; b) J. A. Arnott, S. L. Planey, *Expert Opin. Drug Discovery* **2012**, *7*, 863–875.
- [116] T. Ryckmans, M. P. Edwards, V. A. Horne, A. M. Correia, D. R. Owen, L. R. Thompson, I. Tran, M. F. Tutt, T. Young, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4406–4409.
- [117] a) A. Straub, S. Roehrig, A. Hillisch, *Angew. Chem. Int. Ed.* **2011**, *50*, 4574–4590; *Angew. Chem.* **2011**, *123*, 4670–4686; b) T. A. DeWald, R. C. Becker, *J. Thromb. Thrombolysis* **2014**, *37*, 217–233.
- [118] a) K. Schärer, M. Morgenthaler, R. Paulini, U. Obst-Sander, D. W. Banner, D. Schlatter, J. Benz, M. Stihle, F. Diederich, *Angew. Chem. Int. Ed.* **2005**, *44*, 4400–4404; *Angew. Chem.* **2005**, *117*, 4474–4479; b) L. M. Salonen, C. Bucher, D. W. Banner, W. Haap, J.-L. Mary, J. Benz, O. Kuster, P. Seiler, W. B. Schweizer, F. Diederich, *Angew. Chem. Int. Ed.* **2009**, *48*, 811–814; *Angew. Chem.* **2009**, *121*, 825–828.
- [119] L. M. Salonen, M. C. Holland, P. S. J. Kaib, W. Haap, J. Benz, J.-L. Mary, O. Kuster, W. B. Schweizer, D. W. Banner, F. Diederich, *Chem. Eur. J.* **2012**, *18*, 213–222.
- [120] a) C. D. Tatko, M. L. Waters, *J. Am. Chem. Soc.* **2004**, *126*, 2028–2034; b) R. M. Hughes, M. L. Waters, *J. Am. Chem. Soc.* **2005**, *127*, 6518–6519; c) R. M. Hughes, K. R. Wiggins, S. Khorasanizadeh, M. L. Waters, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 11184–11188.
- [121] a) L. Anselm, D. W. Banner, J. Benz, K. G. Zbinden, J. Himber, H. Hilpert, W. Huber, B. Kuhn, J.-L. Mary, M. B. Otteneder, N. Panday, F. Ricklin, M. Stahl, S. Thomi, W. Haap, *Bioorg. Med. Chem. Lett.* **2010**, *20*, 5313–5319; b) S. Roehrig, A. Straub, J. Pohlmann, T. Lampe, J. Pernerstorfer, K.-H. Schlemmer, P. Reinemer, E. Perzborn, *J. Med. Chem.* **2005**, *48*, 5900–5908; c) M. R. Wiley, L. C. Weir, S. Briggs, N. A. Bryan, J. Buben, C. Campbell, N. Y. Chirgadze, R. C. Conrad, T. J. Craft, J. V. Ficorilli, J. B. Franciskovich, L. L. Froelich, D. S. Gifford-Moore, T. Goodson, D. K. Herron, V. J. Klimkowski, K. D. Kurz, J. A. Kyle, J. J. Masters, A. M. Ratz, G. Milot, R. T. Shuman, T. Smith, G. F. Smith, A. L. Tebbe, J. M. Tinsley, R. D. Towner, A. Wilson, Y. K. Yee, *J. Med. Chem.* **2000**, *43*, 883–899.
- [122] H. Matter, M. Nazaré, S. Güssregen, D. W. Will, H. Schreuder, A. Bauer, M. Urmann, K. Ritter, M. Wagner, V. Wehner, *Angew. Chem. Int. Ed.* **2009**, *48*, 2911–2916; *Angew. Chem.* **2009**, *121*, 2955–2960.
- [123] K. Padmanabhan, K. P. Padmanabhan, A. Tulinsky, C. H. Park, W. Bode, R. Huber, D. T. Blankenship, A. D. Cardin, W. Kiesel, *J. Mol. Biol.* **1993**, *232*, 947–966.
- [124] B. Kuhn, J. E. Fuchs, M. Reutlinger, M. Stahl, N. R. Taylor, *J. Chem. Inf. Model.* **2011**, *51*, 3180–3198.
- [125] S. N. Dorogovtsev, J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW*, Oxford University Press, Oxford, **2003**.
- [126] E. Perola, P. S. Charifson, *J. Med. Chem.* **2004**, *47*, 2499–2510.
- [127] J. Bunzen, J. Iwasa, P. Bonakdarzadeh, E. Numata, K. Rissanen, S. Sato, M. Fujita, *Angew. Chem. Int. Ed.* **2012**, *51*, 3161–3163; *Angew. Chem.* **2012**, *124*, 3215–3217.
- [128] M. Harder, B. Kuhn, F. Diederich, *ChemMedChem* **2013**, *8*, 397–404.
- [129] A. J. Goodman, E. C. Breinlinger, C. M. McIntosh, L. N. Grimaldi, V. M. Rotello, *Org. Lett.* **2001**, *3*, 1531–1534.
- [130] *Aspartic Acid Proteases as Therapeutic Targets* (Eds.: G. Folkers, R. Mannhold, H. Kubinyi), Wiley-VCH, Weinheim, **2010**.
- [131] a) C. Jensen, P. Herold, H. R. Brunner, *Nat. Rev. Drug Discovery* **2008**, *7*, 399–410; b) R. L. Webb, N. Schiering, R. Sedrani, J. Maibaum, *J. Med. Chem.* **2010**, *53*, 7490–7520; c) J. Tamargo, J. López-Sendón, *Nat. Rev. Drug Discovery* **2011**, *10*, 536–555.
- [132] a) A. M. J. Wensing, N. M. van Maarseveen, M. Nijhuis, *Antiviral Res.* **2010**, *85*, 59–74; b) H. C. Castro, P. A. Abreu, R. B. Geraldo, R. C. A. Martins, R. dos Santos, N. I. V. Loureiro, L. M. Cabral, C. R. Rodrigues, *J. Mol. Recognit.* **2011**, *24*, 165–181; c) A. Engelman, P. Cherepanov, *Nat. Rev. Microbiol.* **2012**, *10*, 279–290.
- [133] a) A. K. Ghosh, M. Brindisi, J. Tang, *J. Neurochem.* **2012**, *120*, 71–83; b) T. Silva, J. Reis, J. Teixeira, F. Borges, *Ageing Res. Rev.* **2014**, *15*, 116–145; c) R. Vassar, P.-H. Kuhn, C. Haass, M. E. Kennedy, L. Rajendran, P. C. Wong, S. F. Lichtenthaler, *J. Neurochem.* **2014**, *130*, 4–28.
- [134] a) K. Ersmark, B. Samuelsson, A. Hallberg, *Med. Res. Rev.* **2006**, *26*, 626–666; b) J. Sabotić, J. Kos, *Appl. Microbiol. Biotechnol.* **2012**, *93*, 1351–1375.
- [135] P. Bhaumik, A. Gustchina, A. Wlodawer, *Biochim. Biophys. Acta Proteins Proteomics* **2012**, *1824*, 207–223.
- [136] A search in the PDB for plasmepsins found 30 structures (July 2014), whereby 2 belong to PM I, 19 to PM II, 5 to the histoprotease (HAP), and 4 to PM IV: a) see Ref. [135]; b) K. Jaudzems, K. Tars, G. Maurops, N. Ivdra, M. Otikovs, J. Leitans, I. Kanepe-Lapsa, I. Domracheva, I. Mutule, P. Trapencieris, M. J. Blackman, A. Jirgensons, *ACS Med. Chem. Lett.* **2014**, *5*, 373–377.
- [137] L. Prade, A. F. Jones, C. Boss, S. Richard-Bildstein, S. Meyer, C. Binkert, D. Bur, *J. Biol. Chem.* **2005**, *280*, 23837–23843.
- [138] a) N. K. Bernstein, M. M. Cherney, H. Loetscher, R. G. Ridley, M. N. G. James, *Nat. Struct. Biol.* **1999**, *6*, 32–37; b) D. A. Carcache, S. R. Hörtnner, P. Seiler, F. Diederich, A. Dorn, H. P. Märki, C. Binkert, D. Bur, *Helv. Chim. Acta* **2003**, *86*, 2173–2191; c) D. A. Carcache, S. R. Hörtnner, A. Bertogg, F. Diederich, A. Dorn, H. P. Märki, C. Binkert, D. Bur, *Helv. Chim. Acta* **2003**, *86*, 2192–2209; d) C. Boss, O. Corminboeuf, C. Grisostomi, S. Meyer, A. F. Jones, L. Prade, C. Binkert, W. Fischli, T. Weller, D. Bur, *ChemMedChem* **2006**, *1*, 1341–1345; e) F. Hof, A. Schütz, C. Fäh, S. Meyer, D. Bur, J. Liu, D. E.

- Goldberg, F. Diederich, *Angew. Chem. Int. Ed.* **2006**, *45*, 2138–2141; *Angew. Chem.* **2006**, *118*, 2193–2196.
- [139] a) M. Zürcher, T. Gottschalk, S. Meyer, D. Bur, F. Diederich, *ChemMedChem* **2008**, *3*, 237–240; b) M. Zürcher, F. Hof, L. Barandun, A. Schütz, W. B. Schweizer, S. Meyer, D. Bur, F. Diederich, *Eur. J. Org. Chem.* **2009**, 1707–1719.
- [140] V. Aureggi, V. Ehmke, J. Wieland, W. B. Schweizer, B. Bernet, D. Bur, S. Meyer, M. Rottmann, C. Freymond, R. Brun, B. Breit, F. Diederich, *Chem. Eur. J.* **2013**, *19*, 155–164.
- [141] A. P. Benfield, M. G. Teresk, H. R. Plake, J. E. DeLorbe, L. E. Millspaugh, S. F. Martin, *Angew. Chem. Int. Ed.* **2006**, *45*, 6830–6835; *Angew. Chem.* **2006**, *118*, 6984–6989.
- [142] a) C. Fäh, L. A. Hardegger, L. Baitsch, W. B. Schweizer, S. Meyer, D. Bur, F. Diederich, *Org. Biomol. Chem.* **2009**, *7*, 3947–3957; b) C. Fäh, R. Mathys, L. A. Hardegger, S. Meyer, D. Bur, F. Diederich, *Eur. J. Org. Chem.* **2010**, 4617–4629.
- [143] a) C. R. Caffrey, D. Steverding, *Expert Opin. Drug Discovery* **2008**, *3*, 173–186; b) T. Schirmeister, A. Welker, *Pharm. Unserer Zeit* **2009**, *38*, 564–574; c) A. Cavalli, M. L. Bolognesi, *J. Med. Chem.* **2009**, *52*, 7339–7359; d) C. Teixeira, J. R. B. Gomes, P. Gomes, *Curr. Med. Chem.* **2011**, *18*, 1555–1572.
- [144] a) R. Ettari, L. Tamborini, I. C. Angelo, N. Micale, A. Pinto, C. De Micheli, P. Conti, *J. Med. Chem.* **2013**, *56*, 5637–5658; b) U. R. Mane, R. C. Gupta, S. S. Nadkarni, R. R. Giridhar, P. P. Naik, M. R. Yadav, *Expert Opin. Ther. Pat.* **2013**, *23*, 165–187.
- [145] a) S. Scory, C. R. Caffrey, Y.-D. Stierhof, A. Ruppel, D. Steverding, *Exp. Parasitol.* **1999**, *91*, 327–333; b) B. R. Shenai, P. S. Sijwali, A. Singh, P. J. Rosenthal, *J. Biol. Chem.* **2000**, *275*, 29000–29010.
- [146] a) V. Ehmke, C. Heindl, M. Rottmann, C. Freymond, W. B. Schweizer, R. Brun, A. Stich, T. Schirmeister, F. Diederich, *ChemMedChem* **2011**, *6*, 273–278; b) V. Ehmke, F. Kilchmann, C. Heindl, K. Cui, J. Huang, T. Schirmeister, F. Diederich, *MedChemComm* **2011**, *2*, 800–804; c) V. Ehmke, J. E. Q. Quinsaat, P. Rivera-Fuentes, C. Heindl, C. Freymond, M. Rottmann, R. Brun, T. Schirmeister, F. Diederich, *Org. Biomol. Chem.* **2012**, *10*, 5764–5768; d) V. Ehmke, E. Winkler, D. W. Banner, W. Haap, W. B. Schweizer, M. Rottmann, M. Kaiser, C. Freymond, T. Schirmeister, F. Diederich, *ChemMedChem* **2013**, *8*, 967–975.
- [147] To date, two crystal structures of rhodesain have been published in the PDB (July 2014): a) I. D. Kerr, J. H. Lee, C. J. Farady, R. Marion, M. Rickert, M. Sajid, K. C. Pandey, C. R. Caffrey, J. Legac, E. Hansell, J. H. McKerrow, C. S. Craik, P. J. Rosenthal, L. S. Brinen, *J. Biol. Chem.* **2009**, *284*, 25697–25703 (PDB ID: 2P7U); b) I. D. Kerr, P. Wu, R. Marion-Tsukamaki, Z. B. Mackey, L. S. Brinen, *PLoS Neglected Trop. Dis.* **2010**, *4*, e701 (PDB ID: 2P86).
- [148] To date, five crystal structures of falcipain-2 have been published in the PDB (July 2014): a) T. Hogg, K. Nagarajan, S. Herzberg, L. Chen, X. Shen, H. Jiang, M. Wecke, C. Blohmke, R. Hilgenfeld, C. L. Schmidt, *J. Biol. Chem.* **2006**, *281*, 25425–25437 (PDB ID: 2GHU); b) S. X. Wang, K. C. Pandey, J. R. Somoza, P. S. Sijwali, T. Kortemme, L. S. Brinen, R. J. Fletterick, P. J. Rosenthal, J. H. McKerrow, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 11503–11508 (PDB ID: 1YVB); c) S. X. Wang, K. C. Pandey, J. Scharfstein, J. Whisstock, R. K. Huang, J. Jacobelli, R. J. Fletterick, P. J. Rosenthal, M. Abrahamson, L. S. Brinen, A. Rossi, A. Sali, J. H. McKerrow, *Structure* **2007**, *15*, 535–543 (PDB ID: 2OUL); d) I. D. Kerr, J. H. Lee, K. C. Pandey, A. Harrison, M. Sajid, P. J. Rosenthal, L. S. Brinen, *J. Med. Chem.* **2009**, *52*, 852–857 (PDB ID: 3BPF); e) G. Hansen, A. Heitmann, T. Witt, H. Li, H. Jiang, X. Shen, V. T. Heussler, A. Renneberg, R. Hilgenfeld, *Structure* **2011**, *19*, 919–929 (PDB ID: 3PNR).
- [149] a) M. K. Ramjee, N. S. Flinn, T. P. Pemberton, M. Quibell, Y. Wang, J. P. Watts, *Biochem. J.* **2006**, *399*, 47–57.
- [150] K. Merz, V. Vasylyeva, *CrystEngComm* **2010**, *12*, 3989–4002.
- [151] a) H.-J. Schneider, *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 1072–1073; *Angew. Chem.* **1997**, *109*, 1116–1117; b) M. Roman, C. Cannizzo, T. Pinault, B. Isare, B. Andrioletti, P. van der Schoot, L. Bouteiller, *J. Am. Chem. Soc.* **2010**, *132*, 16818–16824; c) “Weak Intermolecular Interactions: A Supermolecular Approach”: M. Waller, S. Grimme in *Handbook of Computational Chemistry*, Vol. 2 (Ed.: J. Leszczynski), Springer, Dordrecht, **2012**, p. 443–466.
- [152] a) A. Gavezzotti, *J. Phys. Chem.* **1990**, *94*, 4319–4325; b) F. H. Allen, C. A. Baalham, J. P. M. Lommerse, P. R. Raithby, *Acta Crystallogr. Sect. B* **1998**, *54*, 320–329; c) F. H. Allen, W. D. S. Motherwell, *Acta Crystallogr. Sect. B* **2002**, *58*, 407–422.
- [153] a) W. Bolton, *Acta Crystallogr.* **1964**, *17*, 147–152; b) W. Bolton, *Acta Crystallogr.* **1965**, *18*, 5–10.
- [154] H. A. Bent, *Chem. Rev.* **1968**, *68*, 587–648.
- [155] K. J. Kamer, A. Choudhary, R. T. Raines, *J. Org. Chem.* **2013**, *78*, 2099–2103.
- [156] H. B. Bürgi, J. D. Dunitz, *Acc. Chem. Res.* **1983**, *16*, 153–161.
- [157] a) A. Choudhary, R. T. Raines, *Protein Sci.* **2011**, *20*, 1077–1081; b) G. J. Bartlett, R. W. Newberry, B. VanVeller, R. T. Raines, D. N. Woolfson, *J. Am. Chem. Soc.* **2013**, *135*, 18682–18688; c) R. W. Newberry, G. J. Bartlett, B. VanVeller, D. N. Woolfson, R. T. Raines, *Protein Sci.* **2014**, *23*, 284–288.
- [158] a) H. Adams, F. J. Carver, C. A. Hunter, J. C. Morales, E. M. Seward, *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1542–1544; *Angew. Chem.* **1996**, *108*, 1628–1631; b) S. L. Cockcroft, C. A. Hunter, *Chem. Soc. Rev.* **2007**, *36*, 172–188; c) P. J. Carter, G. Winter, A. J. Wilkinson, A. R. Fersht, *Cell* **1984**, *38*, 835–840.
- [159] a) S. Paliwal, S. Geib, C. S. Wilcox, *J. Am. Chem. Soc.* **1994**, *116*, 4497–4498; b) E.-i. Kim, S. Paliwal, C. S. Wilcox, *J. Am. Chem. Soc.* **1998**, *120*, 11192–11193; c) B. Bhayana, C. S. Wilcox, *Angew. Chem. Int. Ed.* **2007**, *46*, 6833–6836; *Angew. Chem.* **2007**, *119*, 6957–6960.
- [160] F. R. Fischer, W. B. Schweizer, F. Diederich, *Angew. Chem. Int. Ed.* **2007**, *46*, 8270–8273; *Angew. Chem.* **2007**, *119*, 8418–8421.
- [161] F. Hof, D. M. Scofield, W. B. Schweizer, F. Diederich, *Angew. Chem. Int. Ed.* **2004**, *43*, 5056–5059; *Angew. Chem.* **2004**, *116*, 5166–5169.
- [162] F. R. Fischer, P. A. Wood, F. H. Allen, F. Diederich, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 17290–17294.
- [163] H. Gardarsson, W. B. Schweizer, N. Trapp, F. Diederich, *Chem. Eur. J.* **2014**, *20*, 4608–4616.
- [164] C. Fäh, L. A. Hardegger, M.-O. Ebert, W. B. Schweizer, F. Diederich, *Chem. Commun.* **2010**, 46, 67–69.
- [165] a) K. Müller, C. Faeh, F. Diederich, *Science* **2007**, *317*, 1881–1886; b) “Fluorine in Pharmaceutical and Medicinal Chemistry”: V. Gouverneur, K. Müller in *Molecular Medicine and Medicinal Chemistry*, Vol. 6 (Eds.: V. Gouverneur, K. Müller), Imperial College Press, London, **2012**.
- [166] B. Kuhn, P. A. Kollman, *J. Am. Chem. Soc.* **2000**, *122*, 3909–3916.
- [167] M. Di Nisio, S. Middeldorp, H. R. Büller, *N. Engl. J. Med.* **2005**, *353*, 1028–1040.
- [168] a) M. A. Fabian, W. H. Biggs, D. K. Treiber, C. E. Atteridge, M. D. Azimioara, M. G. Benedetti, T. A. Carter, P. Ciceri, P. T. Edeen, M. Floyd, J. M. Ford, M. Galvin, J. L. Gerlach, R. M. Grotzfeld, S. Herrgard, D. E. Insko, M. A. Insko, A. G. Lai, J.-M. Lelias, S. A. Mehta, Z. V. Milanov, A. M. Velasco, L. M. Wodicka, H. K. Patel, P. P. Zarrinkar, D. J. Lockhart, *Nat. Biotechnol.* **2005**, *23*, 329–336; b) Y. Liu, N. S. Gray, *Nat. Chem. Biol.* **2006**, *2*, 358–364.
- [169] E. Jabbour, J. Cortes, H. Kantarjian, *J. Core Evid.* **2009**, *4*, 207–213.
- [170] E. Weisberg, P. W. Manley, W. Breitenstein, J. Brügggen, S. W. Cowan-Jacob, A. Ray, B. Huntly, D. Fabbro, G. Fendrich, E. Hall-Meyers, A. L. Kung, J. Mestan, G. O. Daley, L. Callahan,

- L. Catley, C. Cavazza, A. Mohammed, D. Neuberg, R. D. Wright, D. G. Gilliland, J. D. Griffin, *Cancer Cell* **2005**, 7, 129–141.
- [171] a) N. Ramasubbu, R. Parthasarathy, P. Murray-Rust, *J. Am. Chem. Soc.* **1986**, 108, 4308–4314; b) J. P. M. Lommerse, A. J. Stone, R. Taylor, F. H. Allen, *J. Am. Chem. Soc.* **1996**, 118, 3108–3116; c) P. Metrangolo, F. Meyer, T. Pilati, G. Resnati, G. Terraneo, *Angew. Chem. Int. Ed.* **2008**, 47, 6114–6127; *Angew. Chem.* **2008**, 120, 6206–6220.
- [172] G. R. Desiraju, P. S. Ho, L. Kloo, A. C. Legon, R. Marquardt, P. Metrangolo, P. Politzer, G. Resnati, K. Rissanen, *Pure Appl. Chem.* **2013**, 85, 1711–1713.
- [173] T. Clark, M. Hennemann, J. S. Murray, P. Politzer, *J. Mol. Model.* **2007**, 13, 291–296.
- [174] a) M. Palusiak, *J. Mol. Struct. THEOCHEM* **2010**, 945, 89–92; b) L. P. Wolters, M. Bickelhaupt, *ChemistryOpen* **2012**, 1, 96–105.
- [175] a) J. W. Zou, Y. J. Jiang, M. Guo, G. X. Hu, B. Zhang, H. C. Liu, Q. S. Yu, *Chem. Eur. J.* **2005**, 11, 740–751; b) Y.-X. Lu, J.-W. Zou, Y.-H. Wang, Y.-J. Jiang, Q.-S. Yu, *J. Phys. Chem. A* **2007**, 111, 10781–10788; c) K. E. Riley, K. M. Merz, *J. Phys. Chem. A* **2007**, 111, 1688–1694; d) K. E. Riley, P. Hobza, *J. Chem. Theory Comput.* **2008**, 4, 232–242; e) O. A. Syzgantseva, V. Tognetti, L. Joubert, *J. Phys. Chem. A* **2013**, 117, 8969–8980.
- [176] a) K. E. Riley, J. S. Murray, J. Fanfrlík, J. řezáč, R. J. Solá, M. C. Concha, F. M. Ramos, P. Politzer, *J. Mol. Model.* **2011**, 17, 3309–3318; b) A. Bauzá, D. Quiñero, A. Frontera, P. M. Deyà, *Phys. Chem. Chem. Phys.* **2011**, 13, 20371–20379; c) K. E. Riley, J. S. Murray, J. Fanfrlík, J. řezáč, R. J. Solá, M. C. Concha, F. M. Ramos, P. Politzer, *J. Mol. Model.* **2013**, 19, 4651–4659.
- [177] a) O. Hassel, J. Hvorslef, *Acta Chem. Scand.* **1954**, 8, 873; b) O. Hassel, *Science* **1970**, 170, 497–502.
- [178] a) P. Metrangolo, G. Resnati, *Chem. Eur. J.* **2001**, 7, 2511–2519; b) P. Metrangolo, H. Neukirch, T. Pilati, G. Resnati, *Acc. Chem. Res.* **2005**, 38, 386–395; c) “Halogen Bonding in Crystal Engineering”: P. Metrangolo, G. Resnati, T. Pilati, S. Biella in *Halogen Bonding, Series: Structure and Bonding*, Vol. 126 (Eds.: P. Metrangolo, G. Resnati), Springer, Berlin, **2008**, pp. 105–136; d) P. Metrangolo, G. Resnati, *Cryst. Growth Des.* **2012**, 12, 5835–5838.
- [179] G. R. Desiraju, *Angew. Chem. Int. Ed. Engl.* **1995**, 34, 2311–2327; *Angew. Chem.* **1995**, 107, 2541–2558.
- [180] a) N. S. Goroff, S. M. Curtis, J. A. Webb, F. W. Fowler, J. W. Lauher, *Org. Lett.* **2005**, 7, 1891–1893; b) A. Sun, J. W. Lauher, N. S. Goroff, *Science* **2006**, 312, 1030–1034.
- [181] a) A. Priimagi, M. Saccone, G. Cavallo, A. Shishido, T. Pilati, P. Metrangolo, G. Resnati, *Adv. Mater.* **2012**, 24, OP345–OP352; b) A. Priimagi, G. Cavallo, A. Forni, M. Gorynsztejn-Leben, M. Kaivola, P. Metrangolo, R. Milani, A. Shishido, T. Pilati, G. Resnati, G. Terraneo, *Adv. Funct. Mater.* **2012**, 22, 2572–2579; c) A. Priimagi, G. Cavallo, P. Metrangolo, G. Resnati, *Acc. Chem. Res.* **2013**, 46, 2686–2695.
- [182] a) P. Politzer, J. S. Murray, M. C. Concha, *J. Mol. Model.* **2007**, 13, 643–650; b) K. E. Riley, J. S. Murray, P. Politzer, M. C. Concha, P. Hobza, *J. Chem. Theory Comput.* **2008**, 5, 155–163; c) P. Politzer, J. S. Murray, T. Clark, *Phys. Chem. Chem. Phys.* **2010**, 12, 7748–7757.
- [183] a) S. Güssregen, H. Matter, G. Hessler, M. Müller, F. Schmidt, T. Clark, *J. Chem. Inf. Model.* **2012**, 52, 2441–2453; b) M. Kolář, P. Hobza, *J. Chem. Theory Comput.* **2012**, 8, 1325–1333.
- [184] a) M. Erdelyi, *Chem. Soc. Rev.* **2012**, 41, 3547–3557; b) T. M. Beale, M. G. Chudzinski, M. G. Sarwar, M. S. Taylor, *Chem. Soc. Rev.* **2013**, 42, 1667–1680.
- [185] P. L. Wash, M. Shihong, U. Obst, J. Rebek, Jr., *J. Am. Chem. Soc.* **1999**, 121, 7973–7974.
- [186] a) M. G. Sarwar, B. Dragisic, S. Sagoo, M. S. Taylor, *Angew. Chem. Int. Ed.* **2010**, 49, 1674–1677; *Angew. Chem.* **2010**, 122, 1718–1721; b) A. Vargas Jentzsch, D. Emery, J. Mareda, P. Metrangolo, G. Resnati, S. Matile, *Angew. Chem. Int. Ed.* **2011**, 50, 11675–11678; *Angew. Chem.* **2011**, 123, 11879–11882.
- [187] a) N. L. Kilah, M. D. Wise, C. J. Serpell, A. L. Thompson, N. G. White, K. E. Christensen, P. D. Beer, *J. Am. Chem. Soc.* **2010**, 132, 11893–11895; b) L. C. Gilday, T. Lang, A. Caballero, P. J. Costa, V. Félix, P. D. Beer, *Angew. Chem. Int. Ed.* **2013**, 52, 4356–4360; *Angew. Chem.* **2013**, 125, 4452–4456; c) S. Castro-Fernández, I. R. Lahoz, A. L. Llamas-Saiz, J. L. Alonso-Gómez, M.-M. Cid, A. Navarro-Vázquez, *Org. Lett.* **2014**, 16, 1136–1139; d) B. R. Mullaney, A. L. Thompson, P. D. Beer, *Angew. Chem. Int. Ed.* **2014**, 53, 11458–11462; *Angew. Chem.* **2014**, 126, 11642–11646.
- [188] a) S. M. Walter, F. Kniep, E. Herdtweck, S. M. Huber, *Angew. Chem. Int. Ed.* **2011**, 50, 7187–7191; *Angew. Chem.* **2011**, 123, 7325–7329; b) F. Kniep, S. H. Jungbauer, Q. Zhang, S. M. Walter, S. Schindler, I. Schnapperelle, E. Herdtweck, S. M. Huber, *Angew. Chem. Int. Ed.* **2013**, 52, 7028–7032; *Angew. Chem.* **2013**, 125, 7166–7170; c) W. He, Y. C. Ge, C.-H. Tan, *Org. Lett.* **2014**, 16, 3244–3247.
- [189] a) D. W. Larsen, A. L. Allred, *J. Phys. Chem.* **1965**, 69, 2400–2401; b) C. Laurence, M. Queignec-Cabanetos, T. Dziembowska, R. Queignec, B. Wojtkowiak, *J. Chem. Soc. Perkin Trans. 2* **1982**, 1605–1610; c) C. Laurence, M. Queignec-Cabanetos, B. Wojtkowiak, *Can. J. Chem.* **1983**, 61, 135–138; d) S. C. Blackstock, J. P. Lorand, J. K. Kochi, *J. Org. Chem.* **1987**, 52, 1451–1460.
- [190] a) P. Metrangolo, W. Panzeri, F. Recupero, G. Resnati, *J. Fluorine Chem.* **2002**, 114, 27–33; b) S. V. Rosokha, I. S. Neretin, T. Y. Rosokha, J. Hecht, J. K. Kochi, *Heteroat. Chem.* **2006**, 17, 449–459; c) S. Libri, N. A. Jasim, R. N. Perutz, L. Brammer, *J. Am. Chem. Soc.* **2008**, 130, 7842–7844.
- [191] a) R. Cabot, C. Hunter, *Chem. Commun.* **2009**, 2005–2007; b) M. G. Sarwar, B. Dragisic, L. J. Salsberg, C. Gouliaras, M. S. Taylor, *J. Am. Chem. Soc.* **2010**, 132, 1646–1653; c) D. A. Smith, L. Brammer, C. A. Hunter, R. N. Perutz, *J. Am. Chem. Soc.* **2014**, 136, 1288–1291.
- [192] O. Dumele, D. Wu, N. Trapp, N. Goroff, F. Diederich, *Org. Lett.* **2014**, 16, 4722–4725.
- [193] P. Auffinger, F. A. Hays, E. Westhof, P. S. Ho, *Proc. Natl. Acad. Sci. USA* **2004**, 101, 16789–16794.
- [194] a) F. H. Allen, O. Kennard, R. Taylor, *Acc. Chem. Res.* **1983**, 16, 146–153; b) R. Taylor, O. Kennard, W. Versichel, *Acta Crystallogr. Sect. B* **1984**, 40, 280–288.
- [195] a) Y. Lu, T. Shi, Y. Wang, H. Yang, X. Yan, X. Luo, H. Jiang, W. Zhu, *J. Med. Chem.* **2009**, 52, 2854–2862; b) Y. Lu, Y. Wang, W. Zhu, *Phys. Chem. Chem. Phys.* **2010**, 12, 4543–4551; c) R. Wilcken, M. O. Zimmermann, A. Lange, S. Zahn, F. M. Boeckler, *J. Comput.-Aided Mol. Des.* **2012**, 26, 935–945; d) R. Wilcken, M. O. Zimmermann, A. Lange, A. C. Joerger, F. M. Boeckler, *J. Med. Chem.* **2013**, 56, 1363–1388.
- [196] A. R. Voth, P. Khoo, K. Oishi, P. S. Ho, *Nat. Chem.* **2009**, 1, 74–79.
- [197] a) S. D. Barrett, C. M. Biwersi, M. D. Kaufman, H. Tecle, J. S. Warmus, (Warner-Lambert Company), US 2005176820(A1), **2005**; b) J. A. Spicer, G. W. Rewcastle, M. D. Kaufman, S. L. Black, M. S. Plummer, W. A. Denny, J. Quin III, A. B. Shahripour, S. D. Barrett, C. E. Whitehead, J. B. J. Milbank, J. F. Ohren, R. C. Gowan, C. Omer, H. S. Camp, N. Esmail, K. Moore, J. S. Sebolt-Leopold, S. Pryzbranowski, R. L. Merriam, D. F. Ortwine, J. S. Warmus, C. M. Flamme, A. G. Pavlovsky, H. Tecle, *J. Med. Chem.* **2007**, 50, 5090–5102; c) H. Tecle, J. Shao, Y. Li, M. Kothe, S. Kazmirski, J. Penzotti, Y.-H. Ding, J. Ohren, D. Moshinsky, R. Coli, N. Jhawar, E. Bora, S. Jacques-O'Hagan, J. Wu, *Bioorg. Med. Chem. Lett.*

- 2009**, *19*, 226–229 (PDB ID: 3DY7); d) T. O. Fischmann, C. K. Smith, T. W. Mayhood, J. E. Myers, Jr., P. Reichert, A. Mannarino, D. Carr, H. Zhu, J. Wong, R.-S. Yang, H. V. Le, V. S. Madison, *Biochemistry* **2009**, *48*, 2661–2674 (PDB ID: 3EQC); e) Y. Isshiki, Y. Kohchi, H. Iikura, Y. Matsubara, K. Asoh, T. Murata, M. Kohchi, E. Mizuguchi, S. Tsujii, K. Hattori, T. Miura, Y. Yoshimura, S. Aida, M. Miwa, R. Saitoh, N. Murao, H. Okabe, C. Belunis, C. Janson, C. Lukacs, V. Schück, N. Shimma, *Bioorg. Med. Chem. Lett.* **2011**, *21*, 1795–1801 (PDB ID: 3ORN).
- [198] C. Frémin, S. Meloche, *J. Hemat. Onc.* **2010**, *3*, 8–19.
- [199] <http://clinicaltrials.gov/show/NCT01933932>, verified in July **2014**, Clinical Trial ID: NCT01933932.
- [200] a) S. Sarno, H. Reddy, F. Meggio, M. Ruzzene, S. P. Davies, A. Donella-Deana, D. Shugar, L. A. Pinna, *FEBS Lett.* **2001**, *496*, 44–48; b) E. De Moliner, N. R. Brown, L. N. Johnson, *Eur. J. Biochem.* **2003**, *270*, 3174–3181.
- [201] S. Baumli, J. A. Endicott, L. N. Johnson, *Chem. Biol.* **2010**, *17*, 931–936.
- [202] L. J. Core, J. T. Lis, *Science* **2008**, *319*, 1791–1792.
- [203] A. Kunfermann, M. Witschel, B. Illarionov, R. Martin, M. Rottmann, H. W. Höffken, M. Seet, W. Eisenreich, H.-J. Knölker, M. Fischer, A. Bacher, M. Groll, F. Diederich, *Angew. Chem.* **2014**, *126*, 2267–2272; *Angew. Chem. Int. Ed.* **2014**, *53*, 2235–2239.
- [204] M. Witschel, F. Röhl, R. Niggeweg, T. Newton, *Pest Manage. Sci.* **2013**, *69*, 559–563.
- [205] M. Witschel, M. Rottmann, M. Kaiser, R. Brun, *PLoS Neglected Trop. Dis.* **2012**, *6*, e1805.
- [206] P. R. Burkholder, R. M. Pfister, F. H. Leitz, *Appl. Microbiol.* **1966**, *14*, 649–653.
- [207] a) S. Grimme, *J. Chem. Phys.* **2003**, *118*, 9095–9102; b) M. O. Sinnokrot, C. D. Sherrill, *J. Am. Chem. Soc.* **2004**, *126*, 7690–7697; c) E. C. Lee, D. Kim, P. Jurečka, P. Tarakeshwar, P. Hobza, K. S. Kim, *J. Phys. Chem. A* **2007**, *111*, 3446–3457; d) J. M. Sanders, *J. Phys. Chem. A* **2010**, *114*, 9205–9211.
- [208] Since submission of the Review, several relevant new articles have appeared: a) “New Insights into the Role of Water in Biological Function: Studying Solvated Biomolecules Using Terahertz Absorption Spectroscopy in Conjunction with Molecular Dynamics Simulations”: V. C. Nibali, M. Havenith, *J. Am. Chem. Soc.* **2014**, *136*, 12800–12807; b) “Medicinal Chemistry of Catechol *O*-Methyltransferase (COMT) Inhibitors and Their Therapeutic Utility”: L. E. Kiss, P. Soares-da-Silva, *J. Med. Chem.* **2014**, *57*, 8692–8717; c) A review on kinase inhibitors: “Exploring the Scaffold Universe of Kinase Inhibitors”: Y. Hu, J. Bajorath, *J. Med. Chem.* **2014**, DOI: 10.1021/jm501237k; d) “Iodide-Induced Shuttling of a Halogen- and Hydrogen-Bonding Two-Station Rotaxane”: A. Caballero, L. Swan, F. Zapata, P. D. Beer, *Angew. Chem. Int. Ed.* **2014**, *53*, 11854–11858; *Angew. Chem.* **2014**, *126*, 12048–12052.