

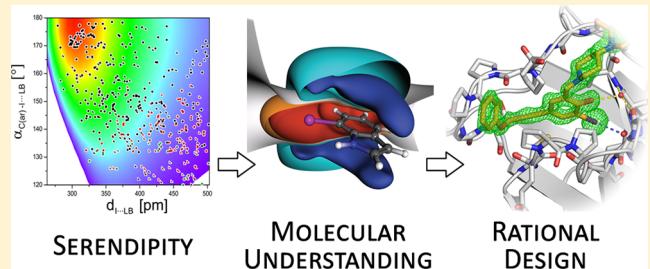
## Principles and Applications of Halogen Bonding in Medicinal Chemistry and Chemical Biology

Rainer Wilcken,<sup>†,‡,§</sup> Markus O. Zimmermann,<sup>†,§</sup> Andreas Lange,<sup>†,§</sup> Andreas C. Joerger,<sup>‡</sup> and Frank M. Boeckler\*,<sup>†</sup>

<sup>†</sup>Laboratory for Molecular Design and Pharmaceutical Biophysics, Department of Pharmaceutical and Medicinal Chemistry, Institute of Pharmacy, Eberhard Karls University, Tuebingen, Auf der Morgenstelle 8, 72076 Tuebingen, Germany

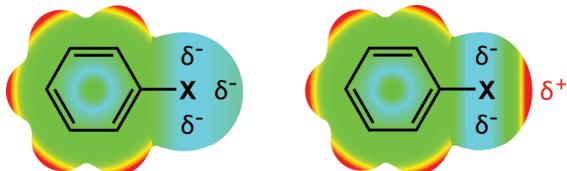
<sup>‡</sup>MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, United Kingdom

**ABSTRACT:** Halogen bonding has been known in material science for decades, but until recently, halogen bonds in protein–ligand interactions were largely the result of serendipitous discovery rather than rational design. In this Perspective, we provide insights into the phenomenon of halogen bonding, with special focus on its role in drug discovery. We summarize the theoretical background defining its strength and directionality, provide a systematic analysis of its occurrence and interaction geometries in protein–ligand complexes, and give recent examples where halogen bonding has been successfully harnessed for lead identification and optimization. In light of these data, we discuss the potential and limitations of exploiting halogen bonds for molecular recognition and rational drug design.



### 1. INTRODUCTION

Halogens, especially the lighter fluorine and chlorine, are widely used substituents in medicinal chemistry. Until recently, they were merely perceived as hydrophobic moieties and Lewis bases in accordance with their electronegativities. Much in contrast to this perception, compounds containing chlorine, bromine, or iodine can also form directed close contacts of the type R–X···Y–R', where the halogen X acts as a Lewis acid and Y can be any electron donor moiety. This interaction, referred to as “halogen bonding” since 1978,<sup>1</sup> is driven by the  $\sigma$ -hole, a positively charged region on the hind side of X along the R–X bond axis that is caused by an anisotropy of electron density on the halogen.<sup>2,3</sup> In Figure 1, a schematic overview is given of



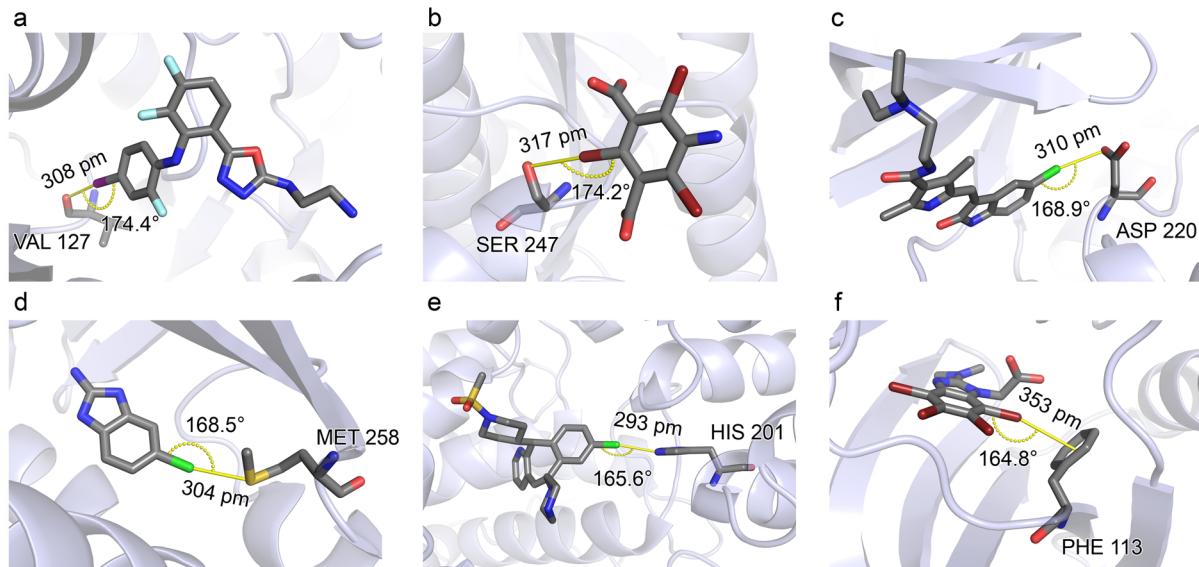
**Figure 1.** Schematic overview of the changing perception of halogen moieties in organic compounds. The color gradient from cyan to red represents the electrostatic potential mapped onto the electron isodensity surface. (a, Left) Traditional assumption of the halogen as a Lewis base (electron donor) with a predominantly isotropic electron distribution on the cap. (b, Right) Real description highlighting the anisotropy of the electron density on the halogen (for fluorine, a positive potential appears only in special cases). The most positive surface potential (including the  $\sigma$ -hole on the halogen) is colored in red, whereas the most negative surface potential is colored in cyan.

how the standard perception of halogen substituents, which assumes an isotropic negative electron density around the halogen, was replaced by a description that takes the  $\sigma$ -hole into account.

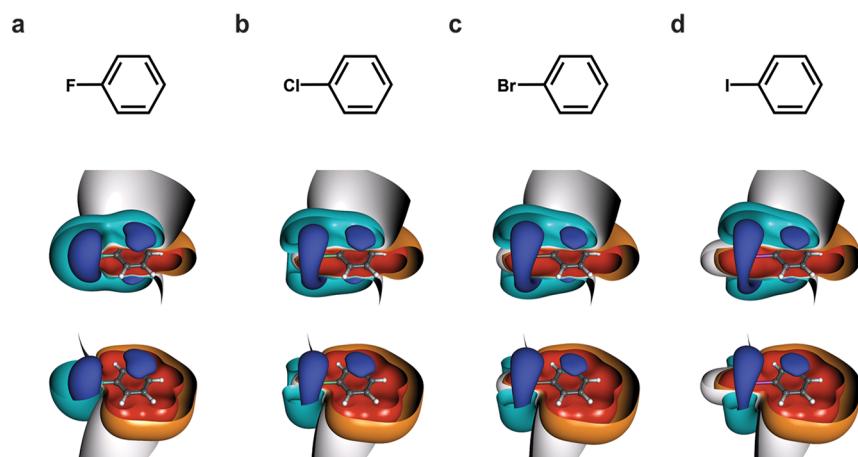
Halogen bonds have been found to occur in a multitude of inorganic, organic, and biological systems.<sup>4,5</sup> In an early study from the 1950s, Hassel and Hvoslef solved the crystal structure of the equimolar Br<sub>2</sub>:dioxane adduct and found Br···O contacts featuring distances substantially below the sum of the van der Waals radii of both atoms, indicating a strong attractive interaction between both atoms.<sup>6,7</sup> In 1984, a search of the Cambridge crystallographic data files for short iodine···N/O/S contacts revealed that these interactions are also formed in biologically relevant systems, being employed by nature for the molecular recognition of thyroid hormones at their target proteins such as transthyretin.<sup>8</sup> In protein–ligand environments, halogen bonds can be formed between a halogenated ligand and any accessible Lewis base in the binding pocket.<sup>9</sup> Probably because of its presence in every amino acid, the backbone carbonyl oxygen function is the most prominent Lewis base involved in halogen bonds in protein binding sites, as found from an analysis of the Protein Data Bank (PDB).<sup>10,11</sup> Additionally, halogen bonds can be formed involving side chain groups, such as hydroxyls in serine, threonine, and tyrosine, carboxylate groups in aspartate and glutamate, sulfurs in cysteine and methionine, nitrogens in histidine, and the  $\pi$  surfaces of phenylalanine, tyrosine, histidine, and tryptophan. Several examples for these contacts are given in Figure 2.

**Received:** August 18, 2012

**Published:** November 12, 2012



**Figure 2.** Selected examples of halogen bonds in protein–ligand complexes. (a) 3EQB: compound forming I···O contact with the carbonyl oxygen of Val 127 in MEK1. (b) 3GT3: compound forming Br···O contact with the hydroxyl group of Ser 247 in proteinase K. (c) 3HZT: compound forming Cl···O contact with the carboxylate group of Asp 220 in *Toxoplasma gondii* CDPK3. (d) 3KRI: compound forming Cl···S contact with the sulfur of Met 258 in hPNMT. (e) 2BED: compound forming Cl···N contact with the nitrogen of His 258 in FPT. (f) 2KXH: compound forming Br···π contact with phenyl ring of Phe 113 in CK2 kinase. All pictures were prepared with PyMOL.<sup>12</sup>



**Figure 3.** Structural formulas and ESP isosurfaces of halobenzenes (MP2/TZVPP). Negative ESP isosurfaces at an energy of  $-0.012$  au are colored in dark blue and at an energy of  $-0.006$  au in cyan. Positive ESP isosurfaces at  $0.012$  au are colored in red and at an energy of  $0.006$  au in orange. The isosurfaces at  $0.000$  au, indicating the boundaries for the transition between negative and positive ESPs, are shown as gray surfaces: (a) fluorobenzene, no  $\sigma$ -hole on the halogen; (b) chlorobenzene; (c) bromobenzene; (d) iodobenzene, increasing  $\sigma$ -hole (positive charge) on the halogen opposite the R–X bond, plus belt of negative electrostatic potential.<sup>20,21</sup> All pictures were prepared with MOLCAD.

While there are some limitations, this multitude of different interaction possibilities in ligand–protein interactions makes halogen bonding a very useful tool to enhance compound affinities and specificities. In this Perspective, we discuss fundamental aspects of the nature of halogen bonding gathered from a large variety of theoretical and experimental studies. We provide a synopsis of the occurrence of halogen bonding in biological systems through systematic analysis of the PDB and present recent examples where halogen bonding has been successfully exploited for ligand design. Finally, we discuss its applicability in the drug discovery process, in terms of available computational tools and practical considerations regarding the inclusion of heavy halides in drug candidates.

## 2. DRIVING FORCE OF HALOGEN BONDING

Statistical evaluation of crystallographic data<sup>13,14</sup> revealed that electrophiles preferentially form contacts with halogen (Cl, Br, I) moieties in a “side-on” fashion, while nucleophiles approach the halogens “head-on”. This preference was explained by Clark et al.<sup>2</sup> using natural bond order analysis of methyl halides. This analysis resulted in an approximate  $s^2 p_x^2 p_y^2 p_z^1$  configuration (where  $z$  is the direction of the C–X bond) for the Cl, Br, and I atoms, thus explaining their previously observed “head-on” electrostatic interactions with nucleophiles by a deficiency in electron density along the  $z$  axis, the  $\sigma$ -hole.<sup>15</sup> This deficiency is compensated for by an electron-rich belt around the halogen that allows it to engage in “side-on” contacts with electrophiles. Here, we focus on halogenated aromatic systems, as these are common scaffolds in medicinal chemistry and form reversible

**Table 1.** Results of Medium and High-Level Quantum Chemical Calculations on Halogen-Bonded Model Systems Involving Carbonyl Oxygens

system	method	$\Delta E$ (kJ/mol)	distance X···O (Å)	$\sigma$ -hole angle ( $C_{ar}$ –X···O) (deg)	study
$H_3CCl\cdots OCH_2$	CCSD(T)/CBS	−4.9	3.26	166.8	Riley and Hobza 2008 <sup>26</sup>
$H_3CBr\cdots OCH_2$	CCSD(T)/CBS	−7.1	3.29	171.2	Riley and Hobza 2008 <sup>26</sup>
$H_3Cl\cdots OCH_2$	CCSD(T)/CBS	−9.7	3.30	172.9	Riley and Hobza 2008 <sup>26</sup>
PhBr···acetone	CCSD(T)/CBS	−12.4	3.10	178.9	Kolář and Hobza 2012 <sup>27</sup>
PhCl···BB <sup>a</sup>	MP2/aug-cc-pVDZ	−5.4	3.08	173.6 <sup>b</sup>	Hardegger et al. 2011 <sup>23</sup>
PhBr···BB <sup>a</sup>	MP2/aug-cc-pVDZ-PP	−10.0	3.03	173.6 <sup>b</sup>	Hardegger et al. 2011 <sup>23</sup>
PhI···BB <sup>a</sup>	MP2/aug-cc-pVDZ-PP	−14.5	3.08	173.6 <sup>b</sup>	Hardegger et al. 2011 <sup>23</sup>
PhCl···BB <sup>a</sup>	MP2/aug-cc-pVDZ	−7.5	3.13	180.0 <sup>b</sup>	Jorgensen and Schyman 2012 <sup>28</sup>
PhBr···BB <sup>a</sup>	MP2/aug-cc-pVDZ-PP	−12.1	3.05	180.0 <sup>b</sup>	Jorgensen and Schyman 2012 <sup>28</sup>
PhI···BB <sup>a</sup>	MP2/aug-cc-pVDZ-PP	−16.9	3.09	180.0 <sup>b</sup>	Jorgensen and Schyman 2012 <sup>28</sup>
PhCl···BB <sup>a</sup>	MP2/TZVPP	−5.6	3.12	171.2	Wilcken et al. 2012 <sup>29</sup>
PhBr···BB <sup>a</sup>	MP2/TZVPP	−9.0	3.04	177.4	Wilcken et al. 2012 <sup>29</sup>
PhI···BB <sup>a</sup>	MP2/TZVPP	−14.2	3.02	175.6	Wilcken et al. 2012 <sup>29</sup>
PhI···BB <sup>a</sup>	CCSD(T)/CBS	−17.6	3.06	175.2	Wilcken et al. 2012 <sup>29</sup>

<sup>a</sup>BB = N-methylacetamide (“backbone” model system). <sup>b</sup>Bond angles restricted to given values.

**Table 2.** Comparison of MP2, DFT, and DFT-D Calculations on Halogen-Bonded Model Systems Involving Carbonyl Oxygens

system <sup>a</sup>	method	$\Delta E$ (kJ/mol)	distance X···O (Å)	$\sigma$ -hole angle ( $C_{ar}$ –X···O) (deg)
PhBr···BB	MP2/TZVPP	−9.0	3.04	177.4
PhBr···BB	TPSS(RI)/TZVPP	−3.6	3.20	176.6
PhBr···BB	BP86(RI)/TZVPP	−2.0	3.14	177.5
PhBr···BB	BLYP(RI)/TZVPP	−2.1	3.24	177.3
PhBr···BB	B3LYP/TZVPP	−3.2	3.21	177.8
PhBr···BB	TPSS-D(RI)/TZVPP	−10.7	3.13	173.5
PhBr···BB	BP86-D(RI)/TZVPP	−9.8	3.09	175.0
PhBr···BB	BLYP-D(RI)/TZVPP	−9.6	3.14	177.8
PhBr···BB	B3LYP-D/TZVPP	−10.4	3.14	174.8

<sup>a</sup>BB = N-methylacetamide (“backbone” model system).

halogen bonds. Figure 3 shows illustrations of the electrostatic potential (ESP) on the simplest halobenzenes computed using the MP2/TZVPP method. From these plots, it is immediately evident that fluorobenzene does not follow the same trend as the other halobenzenes. Because of the considerable electronegativity of fluorine, there is no positively charged area and therefore no interaction potential with nucleophiles. In the other halobenzenes, the size of the  $\sigma$ -hole increases with halogen size from chlorine to iodine, an observation that was made in several studies.<sup>15–17</sup> It has been shown that fluorine can also display a  $\sigma$ -hole and interact with electron donors,<sup>18</sup> but this effect is restricted to extreme cases such as the  $F_3COF$  molecule, with little relevance for medicinal chemistry.<sup>19</sup>

### 3. STRENGTH OF LIGAND–PROTEIN HALOGEN BONDS

In a recent seminal review on different types of protein–ligand interactions relevant to medicinal chemistry,<sup>22</sup> the authors conclude that halogen bonds are a useful addition to the arsenal of favorable interactions in molecular recognition and can lead to significant affinity gains in some cases. The strengths of halogen bonds can be evaluated theoretically through quantum chemical model calculations. There are also a number of experimental studies where the effect of halogen substitution on binding affinity has been systematically evaluated.<sup>23–25</sup>

**3.1. QM-Based Evaluations.** Here, we discuss model calculations involving the carbonyl oxygen moiety because it is the most frequently observed binding partner in biomolecular halogen bonds because of its ubiquitous nature in proteins.

Halogen bonding is best described theoretically using high-level quantum chemical methods such as coupled cluster<sup>30</sup> (CCSD(T)) and perturbation theory<sup>31,32</sup> (MP2) calculations; yet using these methods greatly limits the size of the model systems amenable to computational studies. Much larger systems can be treated using QM/MM calculations<sup>10,16</sup> or semiempirical studies,<sup>33</sup> but in addition to their increased inaccuracy in describing the effect itself, it is very hard to deconvolute the multitude of interactions in a given protein–ligand system treated by QM/MM into the contributions of individual halogen bonds involved. All MP2 and CCSD(T) calculations done on small model systems (as illustrated in Table 1) result in interaction distances significantly below the van der Waals radii of the halogen-bond partners (which are 3.27 Å for Cl···O contacts, 3.37 Å for Br···O contacts, and 3.50 Å for I···O contacts). In all complexes, the halogenated compound interacts with the Lewis base in a nearly ideal “head-on” fashion, with  $C_{ar}$ –X···O angles close to 180°. For the systems representing the protein by N-methylacetamide and the ligands by halobenzenes, the halogen bond complex formation energy for the Cl···O contact is estimated at 5.4–7.5 kJ/mol; Br···O contacts are estimated at 9.0–12.1 kJ/mol and I···O contacts at 14.2–17.6 kJ/mol. Similar calculations at the MP2/TZVPP level give a complex formation energy of about 8.4 kJ/mol for a benzene···N-methylacetamide contact featuring a  $C_{ar}$ –H···O=C hydrogen bond.<sup>29</sup> Given that a moderately strong hydrogen bond, such as the one in the H<sub>2</sub>O dimer, has an energy of about 17–21 kJ/mol,<sup>34</sup> it follows from the results of the calculations that halogen bonds involving carbonyl oxygens are generally

Table 3. Comparison of QM Calculations with Experimentally Observed Binding Affinities<sup>a</sup>

halogen	calculations, MP2/TZVPP $\Delta\Delta E^b$ (kJ/mol)	hCatL inhibitors				MEK1 inhibitors		p53-Y220C stabilizers	
		$IC_{50}^c$ ( $\mu M$ )	$\Delta\Delta G^d$ (kJ/mol)	$IC_{50}^e$ ( $\mu M$ )	$\Delta\Delta G^d$ (kJ/mol)	$IC_{50}^f$ ( $\mu M$ )	$\Delta\Delta G^d$ (kJ/mol)	$K_D^g$ ( $\mu M$ )	$\Delta\Delta G$ (kJ/mol)
Cl	0.0	0.022	0.0	0.03	0.0	0.026	0.0	4900	0.0
Br	-3.4	0.012	-1.5	0.0065	-3.8	0.0077	-3.0	1040	-3.8
I	-8.6	0.0065	-3.0	0.0043	-4.8	0.002	-6.4	247	-7.4

<sup>a</sup> $\Delta E$  and  $\Delta\Delta G$  values are given with respect to the chlorinated compound in the respective series. <sup>b</sup> $\Delta E$  calculated at MP2/TZVPP level for halobenzene- $\cdots$ N-methylacetamide complexes.<sup>29</sup> <sup>c</sup> $\Delta\Delta E$  calculated with respect to chlorobenzene- $\cdots$ N-methylacetamide halogen bond complex formation energy. <sup>d</sup> $IC_{50}$  values for covalent inhibitors of hCatL, compounds **1** (Cl), **2** (Br), and **3** (I) from Hardegger et al.<sup>23</sup> (Chart 1a). <sup>e</sup>Employing the approximation  $IC_{50} \approx K_D$ ,  $\Delta G$  values for each compound were calculated.  $\Delta\Delta G$  values are given with respect to the chlorinated compound. <sup>f</sup> $IC_{50}$  values for covalent inhibitors of hCatL, compounds **4** (Cl), **5** (Br), and **6** (I) from Hardegger et al.<sup>23</sup> (Chart 1a). <sup>g</sup> $IC_{50}$  values for inhibitors of MEK1 kinase, compounds **7** (Cl), **8** (Br), and **9** (I) from Hardegger et al.<sup>24</sup> (Chart 1b). <sup>h</sup>Dissociation constants for p53-Y220C stabilizers, compounds **10** (Cl), **11** (Br), and **12** (I) from Wilcken et al.<sup>25</sup> (Chart 1c).

only comparable in strength to weaker C–H $\cdots$ O hydrogen bonds.<sup>22</sup> However, they should suffer less from desolvation penalties upon binding than polar groups do, an aspect not covered in QM calculations on small model systems. Halogen bonding is a predominantly electrostatic type of interaction, and accordingly there is a good correlation between computed electrostatic potential maxima ( $V_{S,\max}$ ) and interaction energies.<sup>35</sup> However, dispersion may play a relevant role as well, and this was shown for halogen bonds in a symmetry-adapted perturbation theory (SAPT) study.<sup>26</sup> By comparison of interaction energies obtained at the MP2 level of theory with those obtained by computationally less arduous DFT methods (see Table 2 for sample calculations on halogen bonds involving bromobenzene), it becomes apparent that widely used functionals such as BLYP or B3LYP do not describe the energetics of halogen bonding correctly unless augmented by empirical dispersion corrections.<sup>36–38</sup> This demonstrates that dispersion forces cannot be neglected in the description of the effect.

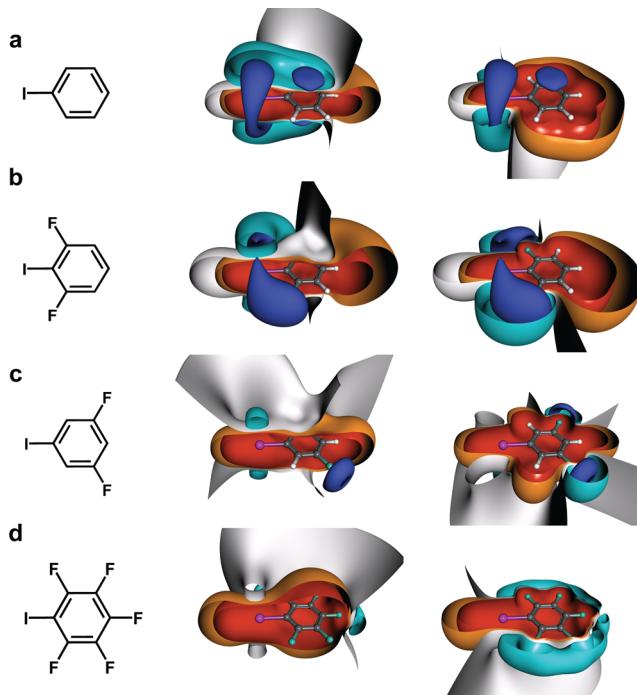
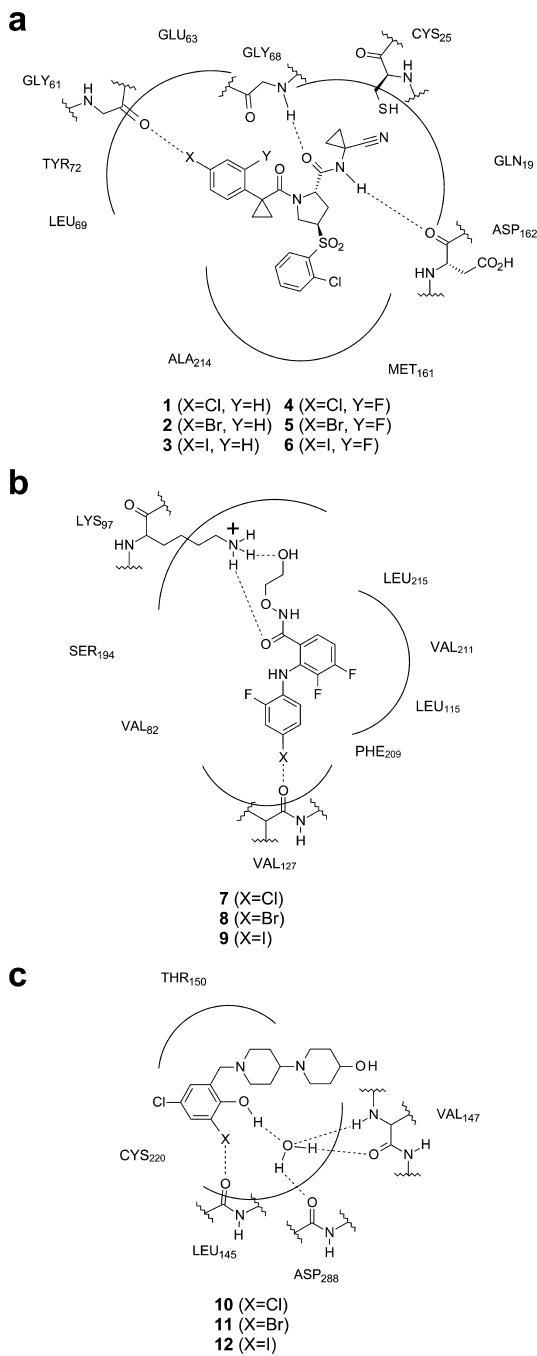
Recent studies also suggest a strong dependence on polarization effects in the local environment of the halogen.<sup>39</sup> This indicates that additional cooperative interactions in the binding site might increase or decrease halogen-bond strength, in addition to tuning effects based on modifications of the scaffold. As the polarizability significantly increases with the atom size, the protein binding site should influence iodine more than bromine, and bromine more than chlorine.

**3.2. Experimental Evaluations.** There is only a very limited number of experimental studies that have systematically evaluated the effect of halogen substitutions on a particular scaffold with confirmed halogen bonding. Data of four such series of compounds targeting backbone carbonyl moieties in three different studies are presented in Table 3<sup>23–25</sup> and compared with the results of MP2 calculations.<sup>29</sup> For all compound series, the free energy of binding of the chlorinated compound is defined as  $\Delta\Delta G = 0$  kJ/mol. Although the same trend as in the calculations (Cl < Br < I) is observed throughout all sets of experimental data, the absolute gains in  $\Delta G$  differ. This difference is not surprising, as the nature of the binding pocket in terms of size, electrostatics, polarization, etc. should differ greatly between the different systems. Intriguingly, the gains in affinity from chlorine toward the heavier halides follow the same trend for all compound classes irrespective of the overall affinity range, which differs significantly between the systems. In both series of hCatL inhibitors investigated by Hardegger et al.<sup>23</sup> (Chart 1a), substitution of R = H by R = Cl, i.e., the introduction of the halogen bond itself, yielded affinity

increases of about one order of magnitude. In all compound series except one (Table 3), exchanging chlorine by iodine approximately yields another order of magnitude in increased affinities. It can therefore be concluded that in systems amenable to the introduction of a halogen bond to target a backbone carbonyl oxygen, affinity increases of up to two orders of magnitude can be achieved by introducing R = I for R = H. Therefore, while halogen bonding is not strong enough in most cases to rival strong hydrogen bonds, it has a distinct potential to generate affinity, specificity, and permeability, depending on the nature of the protein binding pocket.

**3.3. Tuning the Strength of Halogen Bonds.** During the lead optimization process, compounds or fragments identified by initial small-molecule screening are gradually modified in order to improve their affinity to the target protein and their specificity. Optimization includes tuning their physicochemical and pharmacological properties but also introduction of new moieties that tune specific interactions. Halogen bonding is mainly driven by the electrostatic attraction between the electron-deficient areas of the  $\sigma$ -hole on the halogen and the Lewis base interaction partner, with dispersion effects and polarization playing decisive roles as well. Therefore, the strength of halogen bonds can be tuned by introducing electron-withdrawing substituents into a given scaffold, for example, by replacing the hydrogen atoms on a halobenzene scaffold by fluorines, as done in two recent theoretical studies by Riley et al.<sup>40,41</sup> Unsurprisingly, the authors found that introducing fluorines increases the size of the  $\sigma$ -hole on the halogen, leading to an increase in halogen bond strength. In extreme cases, such as in 1,2,3,4,5-pentafluorohalobenzenes, the strength of the interaction is doubled, rivaling polar hydrogen bonding in strength. While these extremely tuned models are not suitable for an optimization process aimed at designing druglike molecules, they nicely illustrate how strongly halogen bonds can be modified. Interestingly, the introduction of two fluorines into a chlorobenzene scaffold makes the halogen bond strength comparable to that of unsubstituted bromobenzene, and 1,3-difluoro-5-bromobenzene and unsubstituted iodobenzene also have a comparable halogen bond strength. While bromo and chloro groups are widely employed substituents in current medicinal chemistry, iodo groups are often perceived as problematic. Substituting an iodoarene core by a substituted bromoarene scaffold might therefore be a feasible strategy to retain affinity by tuning the Br $\cdots$ LB (Lewis base) halogen bond to similar levels as the original I $\cdots$ LB halogen bond. However, introducing fluoro substituents not only affects the  $\sigma$ -hole, and therefore the halogen bond, but also changes the electronic

**Chart 1.** 2D Depiction of the Binding Modes of hCatL Inhibitors (a), MEK1 Inhibitors (b), and p53 Y220C Mutant Stabilizers (c) Featuring Halogen Bonding as a Key Interaction



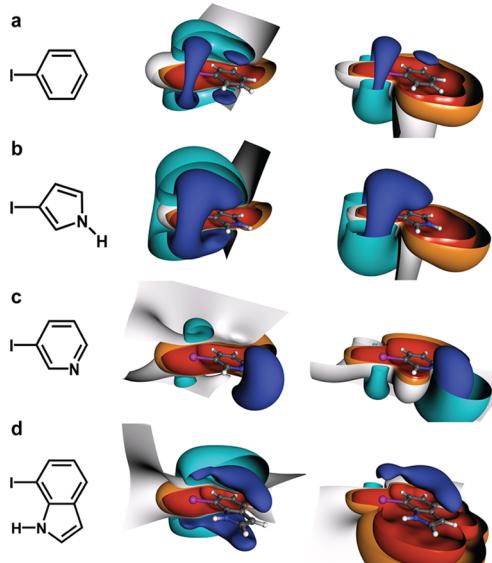
**Figure 4.** Structural formulas and ESP isosurfaces of differently substituted iodobenzenes (MP2/TZVPP): (a) iodobenzene; (b) 1,3-difluoro-2-iodobenzene; (c) 1,3-difluoro-5-iodobenzene; (d) 1,2,3,4,5-pentafluoro-6-iodobenzene. Negative ESP isosurfaces at an energy of  $-0.012$  au are colored in dark blue and at an energy of  $-0.006$  au in cyan. Positive ESP isosurfaces at  $0.012$  au are colored in red and at an energy of  $0.006$  au in orange. The isosurfaces at  $0.000$  au, indicating the border for the transition between negative and positive ESPs, are shown as gray surfaces. Two depictions with identical orientations of the compounds are presented that are clipped by planes perpendicular to the aromatic ring (left) or parallel to the aromatic ring (right). For better visibility of relevant parts of the surfaces all isosurfaces are clipped except for the strongly negative ESP (dark blue). All pictures were prepared with MOLCAD.<sup>20,21</sup>

two fluorines in ortho position of the iodine (Figure 4b) leads to a very interesting distribution of negative and positive electrostatic potential. Although the expanse of the  $\sigma$ -hole on the iodine is increased, it is also partially shielded from the side by areas of negative ESP on the fluorines. These simple examples illustrate that while it is possible to derive certain trends for the tunability of halogen bonds,<sup>40,41</sup> the effects of introducing electron-withdrawing moieties are certainly complex, even more so when introducing larger electron-withdrawing groups than fluorines at different positions of an aromatic or heteroaromatic scaffold.<sup>42</sup>

**3.4. Scaffold Dependency.** Computational studies on protein–ligand complexes have reported varying halogen bond strengths, depending on the chemical nature of the ligand and the surrounding binding pocket.<sup>10,16,43</sup> In the last section, we have discussed that halogen bonds are tunable by introducing substituents into a halogenated scaffold, changing the electronic environment of the halogen bond.<sup>40,41</sup> Naturally, the core scaffold that the halogen is attached to will also play a role, and this raises the interesting question of whether some core structures are better suited for halogen bonding than others. Because of the relative novelty of halogen bonding in the medicinal chemistry community, only a few studies exist on this topic. Hardegger et al.<sup>23</sup> systematically varied substituents on a chlorobenzene scaffold, introducing fluorine, chlorine, and

structure of the scaffold as a whole. The effects on the electrostatic potential are illustrated in Figure 4. While the original iodobenzene system retains a negatively charged belt lateral to the C–I bond and its negatively charged  $\pi$  cloud (Figure 4a), both disappear in the perfluorinated system because of the electron-withdrawing fluoro substituent. The areas of positive ESP at the head of the C–I bond, however, are increased, leading to a very large halogen-bond interaction potential (Figure 4d). A similar increase in  $\sigma$ -hole is observed for 1,3-difluoro-5-iodobenzene (Figure 4c), with distinct patches of negative ESP located at the fluorines. Introducing

trifluoromethyl groups at different positions, and importantly also exchanged the original benzene with pyridine and thiophene. Intriguingly, exchanging the core scaffold had a much greater effect on affinity than simple fluorine substitutions. In Figure 5, we have plotted the electrostatic

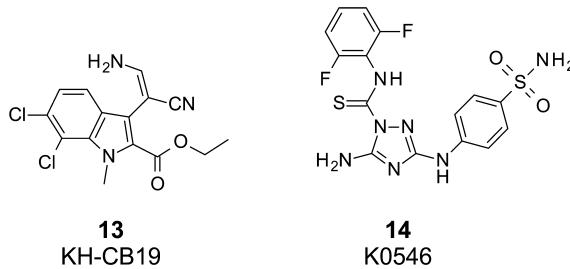


**Figure 5.** Structural formulas and ESP isosurfaces of halogenated aromatic and heteroaromatic scaffolds (MP2/TZVPP): (a) iodobenzene, (b) 3-iodo-1*H*-pyrrole, (c) 3-iodopyridine, (d) 7-iodo-1*H*-indole. Negative ESP isosurfaces at an energy of  $-0.012\text{ au}$  are colored in dark blue and at an energy of  $-0.006\text{ au}$  in cyan. Positive ESP isosurfaces at  $0.012\text{ au}$  are colored in red and at an energy of  $0.006\text{ au}$  in orange. The isosurfaces at  $0.000\text{ au}$ , indicating the border for the transition between negative and positive ESPs, are shown as gray surfaces. Two depictions with identical orientations of the compounds are presented that are clipped by planes perpendicular to the aromatic ring (left) or parallel to the aromatic ring (right). For better visibility of relevant parts of the surfaces all isosurfaces are clipped except for the strongly negative ESP (dark blue). All pictures were prepared with MOLCAD.<sup>20,21</sup>

potential isosurfaces for four commonly used scaffolds. The influence of the scaffold on the electronic distribution and the accessibility of the halogen's  $\sigma$ -hole is rather dramatic. While the nitrogen in 3-iodopyridine (Figure 5c) withdraws electron density from the iodine and increases its  $\sigma$ -hole, thus making it more accessible for halogen bonding compared with unsubstituted iodobenzene (Figure 5a), the iodine's  $\sigma$ -hole in 3-iodo-1-pyrrole is completely surrounded by surfaces of negative ESP. This means the iodine is barely accessible for halogen bonding unless electron-withdrawing substituents are introduced into the pyrrole scaffold. These data illustrate that the scaffold has a decisive influence on where halogen bonding can occur and that it is important to take the local properties of the electronic structure of the core scaffold into account.<sup>44</sup> Apart from illustrating the effect on  $\sigma$ -hole size and the tunability of halogen bond strength, the scaffolds shown in Figure 5 also illustrate that halogens can form part of binding motifs similar to the donor–acceptor motif found in many kinase inhibitors for hinge binding. In 7-iodo-1-indole (Figure 5d), both iodine and indole–NH can interact with one or two Lewis bases, while 3-iodopyridine exhibits a donor–acceptor motif (with iodine being the donor and pyridine N being the acceptor). Two halogens can also act synergistically and participate in bifurcated halogen-bond interactions.<sup>45</sup> Such an example was

found for the 6,7-dichloro-substituted indole scaffold (Chart 2) in **13** (KH-CB19) that interacts with the hinge region of CLK3

### Chart 2



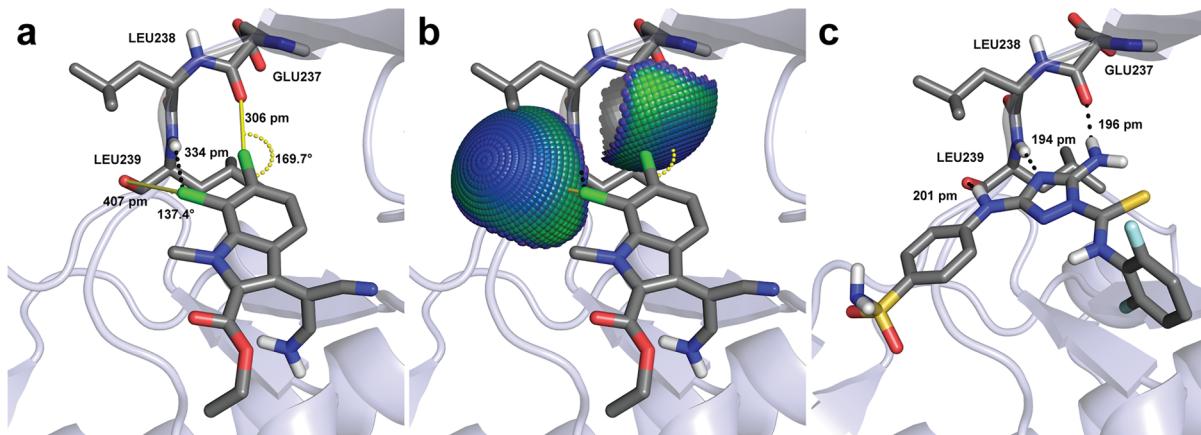
kinase via two halogen bonds (Figure 6).<sup>46,47</sup> Most interestingly, it seems that the donor–acceptor–donor hydrogen-bond network formed by the canonical hinge binder **14** (K0546) is mimicked by a halogen bond–hydrogen bond–halogen bond sequence in the complex of CLK3 with **13**, with one of the chlorines in **13** acting as both halogen-bond donor and hydrogen-bond acceptor. This example illustrates nicely the concept of a binding motif involving halogen bonds.

### 4. INTERACTION GEOMETRIES AND ENERGY BOUNDARIES

Halogen bonds of druglike compounds with their target proteins do not adopt optimal interaction geometries in many cases because they are modulated by the overall binding mode of the ligand and the protein environment. Stronger types of interactions or networks of synergistic interactions may displace the halogen from its favored orientation with an electron donor. It is therefore important to consider how the strength of a halogen bond is affected by such displacements. From the nature of the electron distribution of halogens bound to aromatic or heteroaromatic structures as visualized in Figures 3 and 4, it can be concluded that the following parameters will play an important role in determining the quality of the interaction:

- (1) the distance  $d_{X \cdots LB}$  between the halogen (X) and the Lewis base (LB), based on the predominantly electrostatic interaction of the electron deficient  $\sigma$ -hole with the electron-rich areas (lone pairs) of the Lewis base,
- (2) the  $\sigma$ -hole-angle  $\alpha_{C(ar)-X \cdots LB}$  (given as the angle between the vector of the covalent bond of the halogen with the aromatic scaffold and the vector of the halogen bond with the Lewis base) representing how the accessibility of the  $\sigma$ -hole may be shielded by the surrounding electron-rich belt of the halogen, probably causing repulsive effects toward the electron-rich Lewis base,
- (3) the spatial orientation of the halogen with respect to the electron-rich regions of the Lewis base (typically defined by torsional angles based on the bonding situation of the Lewis base) indicating the accessibility of the donor's electron pairs.

Several quantum chemical studies discussed the preference of these three degrees of freedom in various model systems, representing electron pair donors that occur in small molecules, materials, or biomolecules.<sup>10,15,23,26,29,39,41,48–51</sup> We will subsequently discuss these degrees of freedom using the example of the carbonyl function of the protein backbone as a Lewis base, which is ubiquitously present in binding sites and thus has



**Figure 6.** Crystal structures of inhibitors **13** (hinge halogen bonding, 2WU7) and **14** (hinge hydrogen bonding, 2WU6) bound to CLK3.<sup>47</sup> (a) **13** forms two Cl···O halogen bonds with backbone carbonyl oxygens of Glu237 and Leu239 of CLK3, while the backbone –NH– group of Leu239 donates a hydrogen bond directed toward the negative belt of the 7-chloro moiety. The geometry of the halogen bond with Glu237 is close to the optimal theoretical values, whereas the geometry of the interaction with Leu239 indicates poor halogen bonding. (b) Chlorine–oxygen interaction spheres<sup>29</sup> illustrate the projected qualities of the halogen bonds. (c) **14** binds to CLK3 forming canonical hydrogen-bond interactions with the hinge binding region. All pictures were prepared with PyMOL.<sup>12</sup>

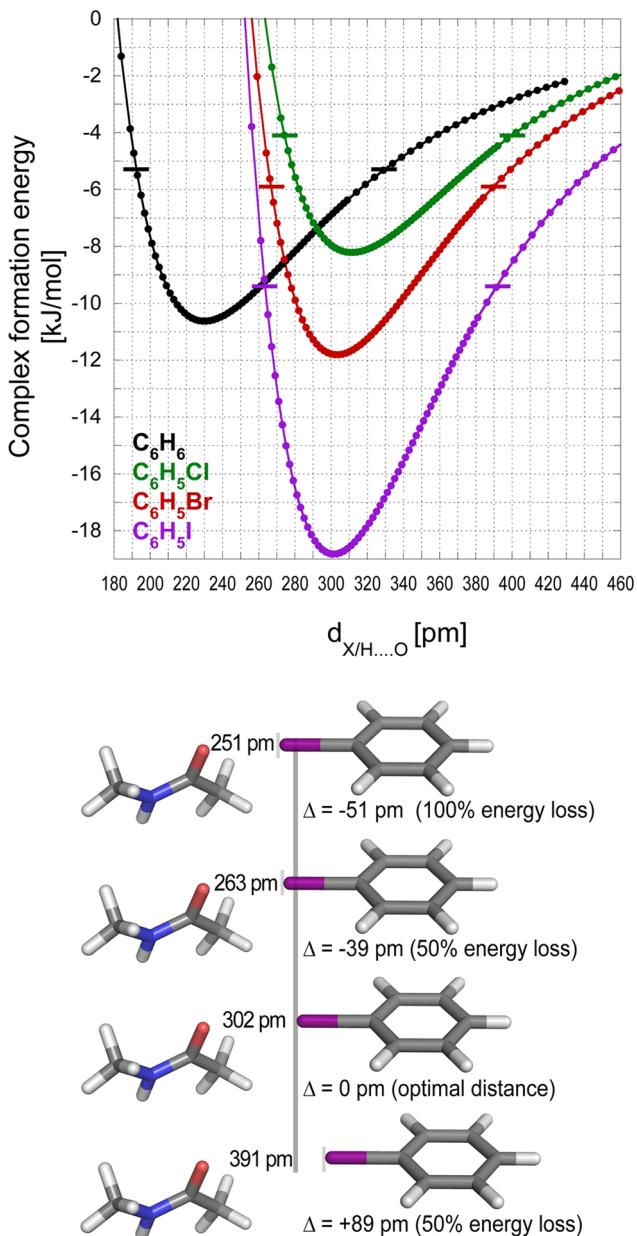
important implications for drug discovery. Very recently, we have systematically investigated this paradigm on the MP2/TZVPP level of theory.<sup>29</sup>

**4.1. Distance.** Deviations from the optimal distance reduce the strength of the interaction (Figure 7). The attractive overlap of the electron-deficient  $\sigma$ -hole with the lone pairs of the carbonyl is reduced with increasing distance. For increasingly positive ESPs (Figure 3), larger interaction distances are tolerated. The calculated complex formation energies drop to about 50% of their optimal values when increasing the distance by about 100 pm. At distances close to 500 pm iodobenzene still shows significantly better interaction energies than benzene. All halobenzenes exhibit a fast transition from attractive to repulsive interaction when decreasing the distance from the optimum, most likely caused by the overlap of the electron-rich belt of the halogens with the lone pair electrons of the carbonyl oxygen. The complex formation energy decreases to 50% of its optimal value for all three heavy halides when placing the scaffold about 35 pm closer to the carbonyl. No attractive interaction occurs below 250 pm. The intersection points with benzene are particularly interesting for drug design. For chlorobenzene, a better complex formation energy than for benzene can only be achieved at a Cl···O distance of 292 pm or higher, which is close to its optimal interaction distance of 312 pm. For bromine and iodine, this intersection occurs at shorter distances of 274 and 263 pm, respectively. Thus, correlating with the magnitude of the overall effect, iodine tolerates displacement of the scaffold significantly better than the other two halides. For chlorobenzene, the advantage over benzene is approximately 2 kJ/mol for the formation of a halogen bond with optimal geometry. Accordingly, only moderate gains in affinity would be expected from introducing chlorine into a scaffold to induce halogen bonding. However, the net gain also depends on synergistic effects of the scaffold and its other substituents that tune the character of the  $\sigma$ -hole.

**4.2.  $\sigma$ -Hole Angle.** On the basis of the anisotropic electron distribution around the halogen, it is easy to envision that deviations from the optimal orientation of the  $\sigma$ -hole onto the carbonyl oxygen (given as  $\Delta\alpha_{C(ar)-X-O}$ ) are not well tolerated. The model calculations regarding these variations start from a

completely linear orientation of the complex where the position of the halogen is fixed at optimal interaction distance with  $\alpha_{C(ar)-X-O} = 180^\circ$  and  $\alpha_{X-O=C} = 180^\circ$  (Figure 8). The second constraint reduces the strength of the interaction by ~4 kJ/mol for iodobenzene but minimizes the risk that the observation of energy trends is hampered by secondary interactions. For all halobenzene systems, a significant decrease of the quality of the interaction is observed when increasing the deviation from  $\alpha_{C(ar)-X-O} = 180^\circ$ . Deviations of 25–30° reduce the energy of complex formation to 50% of its maximum value. For deviations between 30° and 40°, the curves of the different halobenzenes rapidly converge. Beyond a deviation of 40°, no significant attractive interaction is found. Thus, below an angle of  $\Delta\alpha_{C(ar)-X-O} = 140^\circ$  the term halogen bonding should not be used, as there is no favorable overlap of the Lewis base with the  $\sigma$ -hole on the halogen.

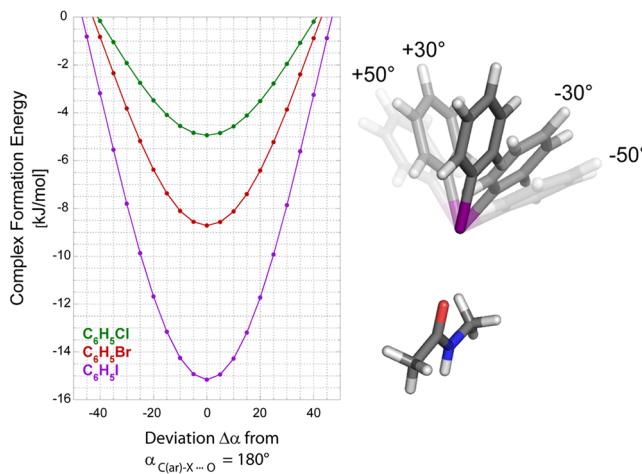
**4.3. Spherical Angles.** The preferential orientation of a ligand depends on the accessibility of electron-rich areas on the electronegative interaction partner. Since these are largely dependent on the electron distribution of the interaction partner within the binding site, we compare two separate systems representing very different binding scenarios: the backbone oxygen and the sulfur atom of the methionine side chain. The spherical orientation of the scaffold with respect to these Lewis bases can be defined by the dihedral angles  $\delta_1$  and  $\delta_2$ , characterizing the out-of-plane or in-plane rotation, respectively (Figure 9). For the halogen bond of iodine with the sulfur of methionine, a strong preference for orientations perpendicular to the plane of the dimethyl sulfide model system is evident. This preference is caused by the anisotropic distribution of electron density on the sulfur atom, featuring two positive  $\sigma$ -holes in the molecular plane along the extensions of the carbon–sulfur bonds and two strongly negative areas above and below the plane.<sup>52</sup> In-plane, there is almost no complex formation energy. In contrast, the backbone carbonyl offers a broad surface patch that can be targeted through a halogen bond with iodine. Except for positions where clashes with the backbone start to occur, the maximal reduction in complex-formation energy is only about 30% of the optimal values. While the flexible methionine side chain can generally more easily adapt to a ligand, facilitating the preferential



**Figure 7.** Distance-dependence of the complex-formation energies [kJ/mol] of the backbone model system *N*-methylacetamide and benzene ( $C_6H_6$ , black curve), chlorobenzene ( $C_6H_5Cl$ , green curve), bromobenzene ( $C_6H_5Br$ , brown curve), or iodobenzene ( $C_6H_5I$ , purple curve). Characteristic distances for iodobenzene are shown below, illustrating how different scaffold placements affect the complex-formation energy. The figure was adapted from Wilcken et al.<sup>29</sup> Structural depictions were prepared with PyMOL.<sup>12</sup>

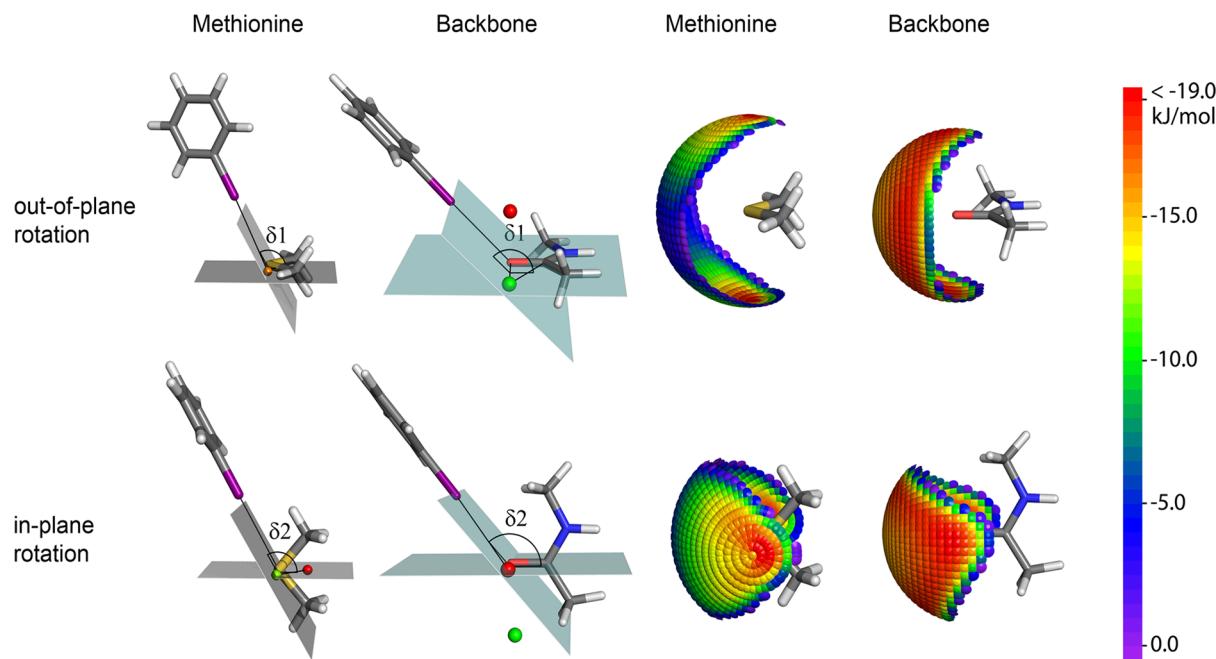
perpendicular interaction geometry, the backbone is often fixed within more rigid secondary structure elements. The fact that backbone oxygens tolerate a broad range of halogen-bonding geometries with high binding energy therefore has important implications for molecular design.

**4.4. Interdependence of Distance and  $\sigma$ -Hole Angle.** To discriminate between strong and weak halogen bonds, the interdependence of the discussed degrees of freedom has to be considered. Since the spherical orientation is primarily determined by the anisotropy of the Lewis base in the binding site, correlations between the orientation and the distance as well as the  $\sigma$ -hole angle will strongly depend on the interaction

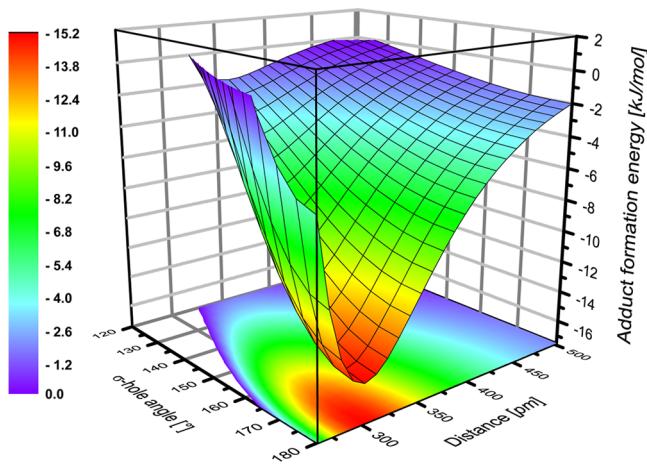


**Figure 8.**  $\sigma$ -Hole angle dependence of the complex-formation energies [kJ/mol] of the backbone model system *N*-methylacetamide and chlorobenzene ( $C_6H_5Cl$ , green curve), bromobenzene ( $C_6H_5Br$ , brown curve), or iodobenzene ( $C_6H_5I$ , purple curve). Characteristic angles for iodobenzene illustrate how deviations from optimal  $\sigma$ -hole angle impair complex-formation energies. The figure was adapted from Wilcken et al.<sup>29</sup> Structural depictions were prepared with PyMOL.<sup>12</sup>

partner of the halobenzene system. We focus on the discussion of the interdependence of the distance and the  $\sigma$ -hole angle, which is certainly not independent of the type of Lewis base, but its trends can be more easily generalized. Complex-formation energies of the iodine contact with the oxygen of the backbone carbonyl model system are given in Figure 10. Again, the constrained complex representing an orientation of the iodobenzene in straight elongation of the carbonyl bond was used for minimizing secondary interactions. There is a clear interaction hot spot (red to yellow color gradient) at distances of 275–350 pm and  $\sigma$ -hole angles of 160–180°. The corresponding energies of this hot spot are in the top 30% of attractive energies. Interactions of intermediate quality (green) are found in an area between 260 and 400 pm for an optimal  $\sigma$ -hole angle of 180°. The poorer the  $\sigma$ -hole angle, the smaller the range of tolerated distances becomes. No interactions better than about -6 kJ/mol are found below 140° (cyan and blue colors). The optimal distance for a particular  $\sigma$ -hole angle shifts slightly toward higher distances for smaller  $\sigma$ -hole angles. This trend can be explained by visualizing that the electron-rich belt of the halogen will collide more dramatically with the electron-rich areas of the Lewis base when rotating the ligand around the halogen atom at shorter distances. These collisions of negative electrostatic potentials also lead to an increase of the distance below which only repulsive interactions occur at larger deviations from the optimal  $\sigma$ -hole angle of 180°. This zone of repulsive potentials corresponds to the uncolored region in the 2D projection of the energy gradient in Figure 10. Such 2D projections are useful maps to assess the quality of halogen bonding in biological systems, similar to the use of Ramachandran plots for evaluating backbone conformations in protein structures. In the following, we will use such plots to analyze the halogen-bonding statistics of the Protein Data Bank (PDB). These plots are coarse estimations of interactions with the backbone carbonyl oxygen. They should therefore only be used as rough guidelines, as they do not account for different spherical orientations of the ligand and modifications of the  $\sigma$ -hole by various scaffolds, substitution patterns, or polarization effects in the binding site of the protein.



**Figure 9.** Definition of dihedral angles ( $\delta_1$  and  $\delta_2$  representing out-of-plane and in-plane rotations, respectively) and spherical scan plots for the backbone and methionine model system. The color gradient from red to blue highlights the quality of the interaction. The figure was adapted from Wilcken et al.<sup>29,50</sup> All pictures were prepared with PyMOL.<sup>12</sup>



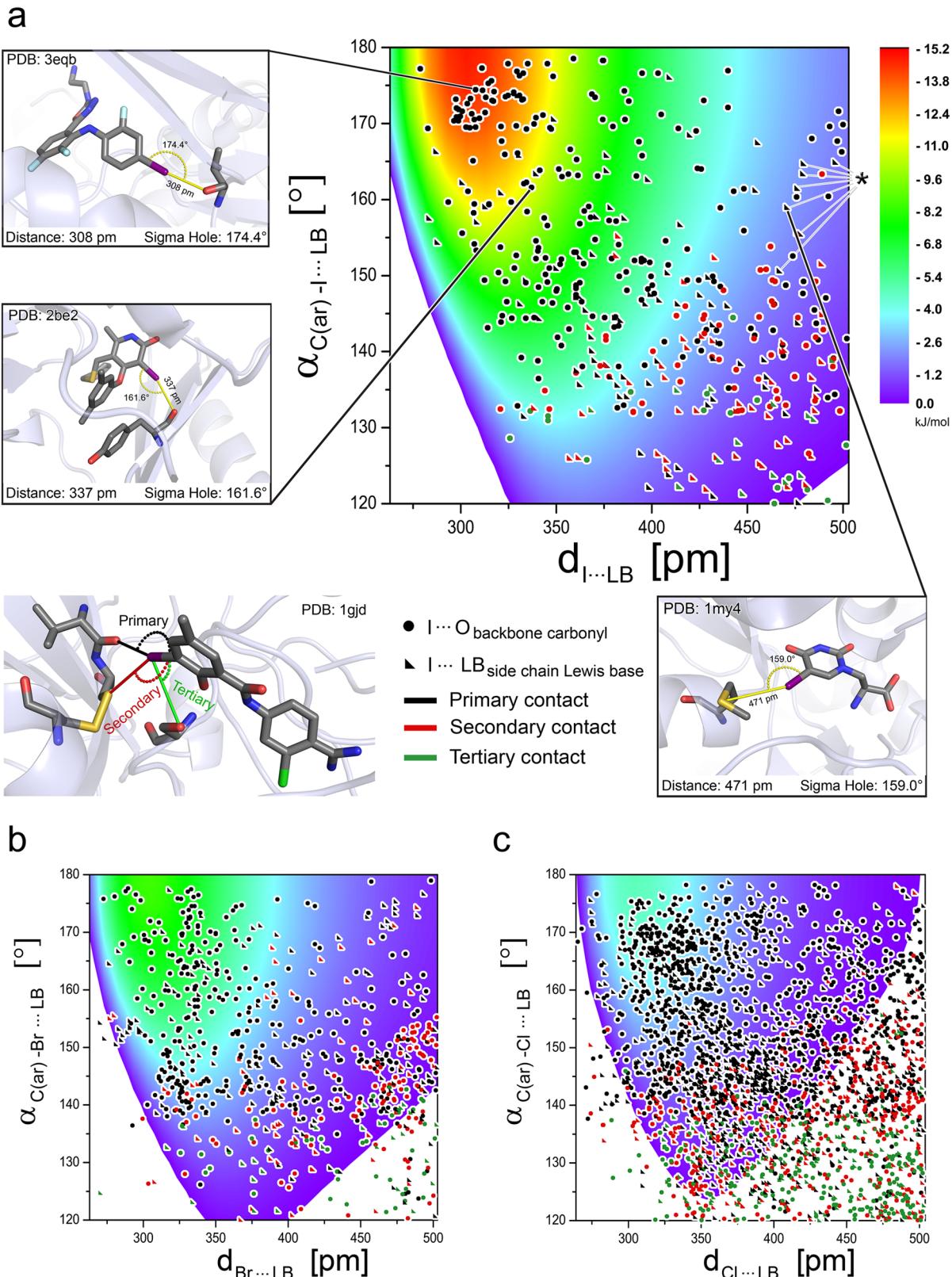
**Figure 10.** Interdependence of distance and  $\sigma$ -hole angle.

**4.5. Comparison of Theory and Experiment.** **4.5.1. General Considerations.** Describing interactions systematically using QM methods is an expedient way of characterizing halogen bonding. However, at present this approach is naturally constrained to smaller model systems. Taking only part of a larger system into account, however, ignores various other effects. Structural information derived from the PDB<sup>53,54</sup> is a valuable source for analyzing which areas of the 2D energy projections of halogen bonds are densely populated under experimental conditions. PDB analyses of protein–ligand complexes have been frequently used to derive scoring functions<sup>55–58</sup> and to characterize preferential geometries of molecular interactions,<sup>22,59–61</sup> including halogen bonding.<sup>10,11,49,62,63</sup> However, when critically assessing PDB statistical data, several aspects should be considered:

(a) As described recently by Bissantz et al.,<sup>22</sup> a good correlation between the 3D analysis of a particular molecular interaction and the binding free energy, although being

desirable, cannot be expected because of the nonadditive nature of the multitude of molecular interactions that define a certain binding mode. Specific interactions can therefore be largely dependent on the context of the ligand scaffold and the binding-site environment. In addition, solvation, entropy, and the dynamic states of protein and ligand ensembles can strongly influence the binding affinities, which is not reflected in by crystal structure. It is important to be aware of the limitations of such statistical analyses and to avoid overinterpretation of data based on unique experimental systems. We therefore use the statistical data merely for comparison with the model calculations.

(b) The boundaries used to derive the statistical data from the PDB have an important influence on the results obtained. As we have shown for iodobenzene, halogen bonds can exhibit weak attractive interactions up to a distance of 500 pm. Tuning effects of the scaffold or its other substituents may even increase this distance. From a recent version of the PDB (comprising structures released until June 12, 2012) we extracted all contacts of ligand-bound chlorine, bromine, or iodine made with Lewis bases within a protein side chain or the carbonyl oxygen of the protein backbone using a Python-based application. On the basis of the model calculations mentioned above, the range of allowed distances was set to 250–600 pm and only  $\sigma$ -hole angles between 180° and 120° were considered, yielding a spherical sector of 0.210 nm<sup>3</sup> (=210 Å<sup>3</sup>) for potential interaction partners. Often, several Lewis bases are found within the perimeter of this spherical sector, giving the impression that one halogen might form various halogen bonds at the same time. While bifurcated interactions of halogens with two interaction partners are conceivable,<sup>45</sup> a clear hierarchy of primary, secondary, and tertiary interactions is usually obvious from their different geometries, as shown in the example in Figure 11 (PDB code 1GJD).<sup>64</sup> On the basis of a scoring system taking into account distance and  $\sigma$ -hole angle,

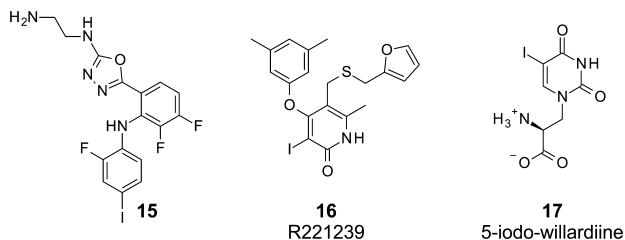


**Figure 11.** Overall PDB statistics compared to regions of favored geometry derived by QM model calculations. (a) 2D heat map representation of iodine interactions featuring experimentally observed contacts with the carbonyl oxygen of the protein backbone (filled circles) or with Lewis bases in side chains (filled triangles). Color indicates the hierarchy of the contact: primary (black), secondary (red), tertiary (green). Selected contacts are shown in detail, illustrating different qualities of interaction (PDB codes 1GJD, 3EQB, 2BE2, 1MY4). An asterisk (gray lines) marks seven contacts representing the same ligand–protein interaction (1MY4). (b) 2D heat map representation for bromine and (c) chlorine interactions. Identical symbols and color coding is used as before. All ligand–protein complexes were depicted using PyMOL.<sup>12</sup>

we classified competing interaction partners as primary, secondary, or tertiary, discarding all weaker contacts.

**4.5.2. Strong Iodine Interactions.** The PDB analysis shows that numerous compounds containing iodine (Figure 11a) are clustered inside the energy hot spot, indicating close to optimal interactions. Only a few have shorter than optimal contact distances. Larger distances and smaller  $\sigma$ -hole angles are found quite frequently. Significantly more primary contacts with backbone carbonyl oxygens than with side chain functions are found in the preferred regions. The following two examples illustrate the difference between optimal and slightly sub-optimal interactions: The MAP ERK kinase inhibitor *N*-(5-{3,4-difluoro-2-[{(2-fluoro-4-iodophenyl)amino]phenyl}-1,3,4-oxadiazol-2-yl)ethane-1,2-diamine (**15**, Chart 3) in its complex

**Chart 3**



with MEK1 (PDB code 3EQB) exhibits nanomolar MEK inhibitory activity in cultured C26 colon carcinoma cells ( $IC_{50} = 15$  nM).<sup>65</sup> The iodine of its 2-fluoro-4-iodophenylaniline moiety adopts an almost optimal halogen-bond geometry (distance of 308 pm and  $\sigma$ -hole angle of 174.4°) with the backbone carbonyl of Val127. The binding mode is mediated by several favorable interactions, including an intramolecular hydrogen bond that stabilizes the bioactive conformation of the 2-fluoro-4-iodophenylaniline. The HIV1-RT inhibitor **16** (R221239) interacts with a carbonyl backbone oxygen featuring a contact geometry close to the hot spot region (distance of 337 pm,  $\sigma$ -hole angle of 161.6°). This highly potent inhibitor ( $IC_{50} = 2$ nM) engages in a tightly packed network of hydrogen bonds,  $\pi\cdots\pi$ , and C–H $\cdots\pi$  contacts, in addition to the observed halogen bond (PDB code 2BE2).<sup>66</sup> It is possible that other interactions of this highly substituted 3-iodopyridin-2(1H)-one scaffold prevent optimal interaction with the backbone carbonyl oxygen.

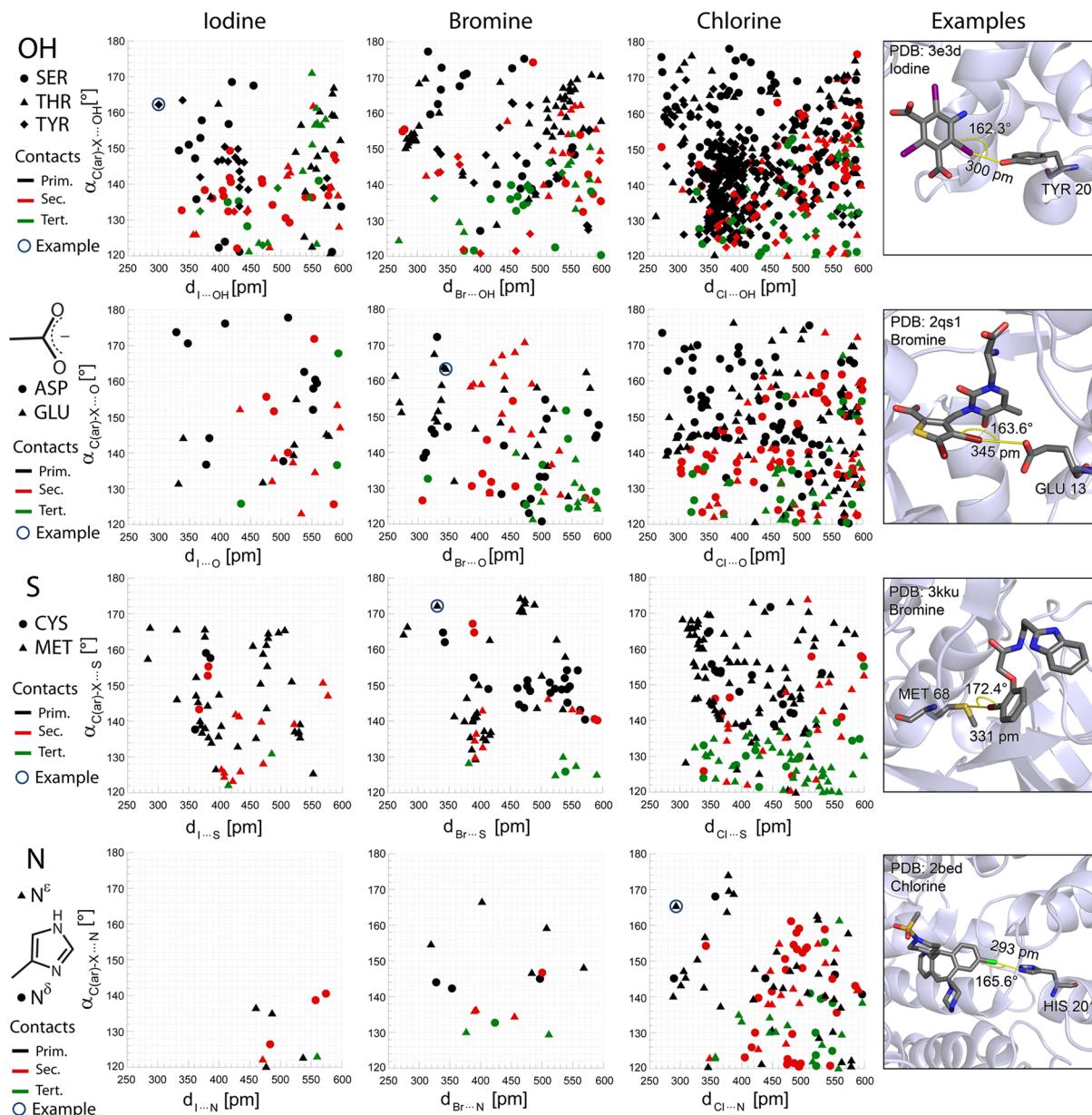
**4.5.3. Weak Iodine Interactions.** Even in areas indicating weak to very weak interactions (colored cyan to blue in Figure 11a), a moderate population of primary contacts is found. There are also large populations of secondary or tertiary contacts in this region, whereas only very few isolated secondary interactions are present in regions of moderate to high complex-formation energies. The primary contact with a methionine in Figure 11a exemplifies the increased likelihood of finding false positives in the region of weak or very weak interactions. Targeting methionine can yield similar or better energies than targeting the backbone carbonyl.<sup>29,50</sup> However, the tolerance for different spherical orientations of the ligand is significantly more limited for methionine. For the GluR2 partial agonist **17** (5-iodowillardiine), distance (471 pm) and  $\sigma$ -hole angle (159.0°) appear to be reasonable (PDB code 1MY4),<sup>67</sup> particularly when considering that the sulfur electron density is more expansive than the electron density of oxygen. Interaction energies will be consequently higher for iodine–sulfur contacts at this long range. However, the in-plane orientation of the

ligand relative to the dimethyl sulfide function, facing the shielded region of the sulfur (Figure 11a), is not consistent with halogen bonding. This example also highlights a general problem when using uncurated PDB statistics: There are three crystal structures of **17** bound to two or three protein chains featuring slightly different binding geometries in each case (PDB codes 1MY4,<sup>67</sup> 1MQG,<sup>68</sup> 3T96<sup>69</sup>). As a consequence, exactly the same contact between a ligand (**17**) and a residue (methionine) generates seven signals in the statistical data that are clustered in a narrow region (indicated by the asterisk in Figure 11a).

**4.5.4. Bromine and Chlorine Interactions.** Bromine-mediated halogen bonds are more prevalent in the PDB (Figure 11b), and an even larger number of contacts made by chlorine-containing ligands are found (Figure 11c). This distribution parallels the total number of structures with iodine-, bromine-, or chlorine-containing ligands in the PDB. The ratio of such compounds is ~1:1.9:7.3. For bromine, the preferred region (green) is well populated with primary contacts, whereas only very few contacts are found in the zone of repulsive energy (white segments in Figure 11b). A strong prevalence of interactions with the backbone is found in the preferred areas. There is also a cluster of contacts featuring short distances (300–350 pm) but only moderately optimal  $\sigma$ -hole angles (140–150°) within a zone of weaker interactions (cyan). Most secondary or tertiary contacts are found in regions of very weak complex-formation energies, and they may therefore not be relevant. In the case of chlorine, proper halogen bonds are most likely found in the densely populated zone of weak interactions. Again, a dominance of contacts with the backbone can be observed.

**4.5.5. Examples for Targeting Side Chains.** Because of the prevalence of backbone contacts in this analysis, we have visualized contacts with different types of side chains separately (Figure 12). We evaluated whether the current release of the PDB contains examples of suitable halogen bonds with interaction partners such as (a) the oxygen in serine, threonine, or tyrosine, (b) the carboxylate in aspartate or glutamate, (c) the sulfur in methionine or cysteine, or (d) the nitrogens in histidine. Favorable interactions with  $\pi$ -systems (phenylalanine, tyrosine, tryptophan, and histidine) are possible but not trivial to assess statistically. Some studies have reported favorable halogen $\cdots\pi$  contacts in small molecules and materials.<sup>70,71</sup> Further insights into the geometry dependence of this interaction are needed to facilitate rational statistical analyses. For each of the four classes of side chains in Figure 12, we show a representative example from the recent literature.

**4.5.5.1. Targeting Hydroxyl Groups in Serine, Threonine, and Tyrosine.** Contacts of iodine with hydroxyl oxygens are rarely found in the PDB. Nevertheless, there are some examples for each residue type. 5-Amino-2,4,6-triiodobenzene-1,3-dicarboxylic acid (**18**, Chart 4) has been used as a compound for phasing in X-ray crystallography<sup>72</sup> because its three iodine atoms, forming an equilateral triangle of roughly 600 pm side length, provide a strong anomalous signal. If bound with suitable stability and good occupancy, it can be used for SAD (single-wavelength anomalous dispersion) or SIRAS (single isomorphous replacement plus anomalous scattering) phasing. In its complex with lysozyme (PDB code 3E3D), it forms an almost ideal halogen bond (distance of 300 pm,  $\sigma$ -hole angle of 162.3°) with Tyr20. Interestingly, it has been suggested<sup>72</sup> that **18** and its derivatives, which have been used extensively as X-ray contrast reagents in medical diagnosis,<sup>73</sup> might promote



**Figure 12.** Side chain PDB statistics and representative examples. All ligand–protein complexes were depicted using PyMOL.<sup>12</sup>

crystal growth and act as bridges between different protein molecules.

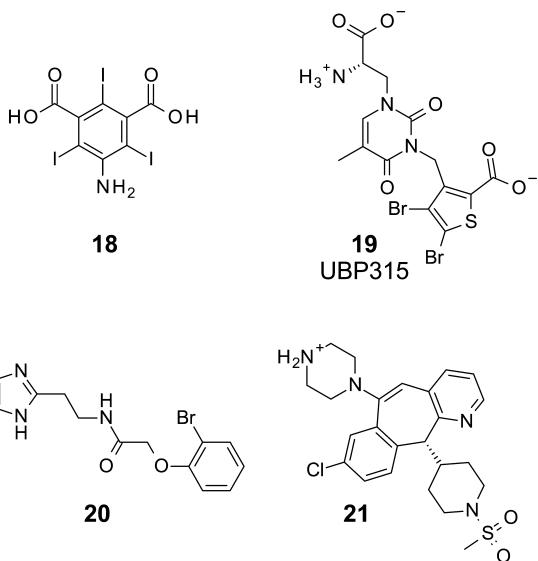
For bromine and chlorine, more contacts with hydroxyl oxygens in serine, threonine, and tyrosine side chains are observed in favorable regions. However, less favored areas also become more populated in bromine- and particularly chlorine-containing ligands. On the basis of the anisotropy of the electron density of both halogens and the possibility that the hydroxyl functions may donate a hydrogen bond as well, some of these interaction geometries may in fact represent hydrogen bonding rather than halogen bonding.

**4.5.5.2. Targeting the Carboxylate Groups of Aspartate and Glutamate.** Targeting the negatively charged aspartate or glutamate should in principle lead to strong halogen bonding because of its predominantly electrostatic nature. However, competition of halogenated ligands with hydrogen-bond donors, cationic atoms, and solvent molecules<sup>74</sup> most likely limits the number of effective targetable carboxylate side chains

significantly. Entropic disadvantages stemming from side-chain flexibility<sup>75,76</sup> may also play a role. The presence of two electronically identical oxygens in the carboxylate group of aspartate and glutamate as potential interaction partners complicates the analysis of the PDB statistics. Bifurcated interactions of one halogen with both oxygens of the carboxylate group with similar distances and  $\sigma$ -hole angles are conceivable but not obvious from the statistics. Typically, there are two sets of primary and secondary interaction patterns with interdependent contact geometries.

A very nice example of such a contact was found for the potent  $\text{GLU}_{\text{K}_5}$  kainate receptor antagonist **19** (UBP315). When used to antagonize kainate-induced depolarization of neonatal rat dorsal root fibers, **19** shows an apparent  $K_{\text{D}}$  of 10 nM.<sup>77</sup> In the crystal structure (PDB code 2QS1) of **19** with the GluRS ligand-binding core dimer,<sup>78</sup> it blocks the glutamate binding site, exhibiting a favorable halogen bond (distance of 345 pm,  $\sigma$ -hole angle of 163.6°) of its 4-bromothiophene moiety with

Chart 4



Glu426. The 4,5-dibromothiophene-2-carboxylic acid is effectively fixed by a network of additional interactions. Likewise, the carboxylate group of Glu426 is stabilized in its interaction geometry by a hydrogen bond with Tyr474, which is almost orthogonal to the halogen bond. Compared with other analogues lacking the halogen-bonded bromine, 19 facilitates a significantly tighter domain closure by inducing a rigid-body motion of domain 2. This triggers a conformational switch of Glu723 (considered to be relevant for receptor activation by agonists), which displaces solvent molecules to form a salt bridge with the  $\alpha$ -NH<sub>2</sub> function of 19.<sup>78</sup>

**4.5.5.3. Targeting the Sulfur Atom in Methionine and Cysteine.** Methionine is generally viewed as a largely hydrophobic interaction partner, but because of the anisotropy of the sulfur's electron distribution and its polarizability, it is an interesting example of a halogen-bond acceptor.<sup>50</sup> The methionine side chain has a high flexibility, as shown by deuterium spin relaxation NMR experiments.<sup>79</sup> It has three rotatable bonds and adopts different sets of favored conformations that can be spatially isosteric for the sulfur and sometimes also for the C<sup>e</sup>.<sup>80</sup> When its conformations from rotamer libraries are clustered with respect to the S<sup>d</sup>–C<sup>e</sup> bond orientation, it appears that the terminal methylsulfanyl group can be kept fixed while retaining rotameric disorder of the methionine side chain.<sup>76</sup> For halogen bonding, this could imply that while the C(ar)–X bond vector needs to maintain a perpendicular orientation to the C'–S<sup>d</sup>–C<sup>e</sup> plane of the methionine side chain to avoid substantial loss of interaction energy (see Figure 9), this immobilization of the side chain may not necessarily lead to strong entropic penalties. Such effects were also studied for salt bridges involving glutamate and arginine.<sup>76</sup> Packing effects of the methionine side chain against other hydrophobic residues might be another source of orientational fixation of the methylsulfanyl group. Interestingly, methionines tend to be found closer to the protein surface than other methyl-bearing side chains.<sup>79</sup> But despite often being found at the solvent-accessible surface of a protein, desolvation penalties for methionine seem to be comparably low.<sup>50</sup> Thus, competition with loosely interacting waters may not be a strong hindrance for halogen bonding with methionine.

N-[2-(1H-Benzimidazol-2-yl)ethyl]-2-(2-bromophenoxy)-acetamide (20) was recently identified by parallel docking and quantitative high-throughput screening as a novel inhibitor ( $IC_{50} = 2 \mu\text{M}$ ) of the thiol protease cruzain, a key drug target involved in Chagas' disease.<sup>81,82</sup> The high-resolution crystal structure of this ligand in complex with cruzain (PDB code 3KKU) reveals a favorable interaction geometry (distance of 331 pm,  $\sigma$ -hole angle of 172.4°) of the Br···S contact with Met68. 20 consists of a 2-bromophenoxy moiety connected to a 2-substituted 1H-benzimidazol by a flexible ethylacetamide linker. The amide function of the linker forms two hydrogen bonds with the backbone carbonyl of Asp161 and Gly66, linking the S1 and S2 pockets. The ligand adopts a U-shaped conformation, with its 2-bromophenyl moiety occupying the S2 pocket and the benzimidazole being largely solvent-exposed. The authors thoroughly investigated why 20 was a false negative in the docking studies (which employed DOCK 3.5.54 using van der Waals interactions, electrostatic interactions, and a penalty for ligand desolvation for scoring)<sup>83–85</sup> and explained this by a failure in recognition due to poor scoring rather than as a docking failure due to inadequate sampling. This example highlights the need for well-parameterized force fields and scoring functions that recognize halogen bonding to complement lead identification and support decision making in lead optimization.

**4.5.5.4. Targeting Nitrogen in His.** Only very few halogen bonds with the nitrogen atoms of histidine are found in the PDB. This low prevalence is quite striking but not unexpected because histidine is partially protonated at physiological pH. Only a small population of histidine residues may be able to form a halogen bond. Histidines can additionally form hydrogen bonds, charge–charge or charge-dipole interactions,  $\pi$ – $\pi$ -contacts, or complexes with metal ions. These alternative interactions and solvation effects drastically reduce the number of histidine residues poised for halogen-bond formation. But once X···N or X··· $\pi$  contacts with histidine residues are formed, they can be connected to larger interaction networks because of the versatility of histidine to engage in different types of interactions. In principle, the equilibrium between tautomeric states of unprotonated histidines should provide similar probabilities for both nitrogen atoms (N<sup>d</sup> and N<sup>e</sup>) to be involved in halogen bonds. The statistical data for bromine and iodine are too limited to derive trends, but a preference for contacts with N<sup>e</sup> is noticeable within the primary interactions of chlorine. This may indicate that N<sup>d</sup> is typically less accessible because it is more shielded by the proximal C <sup>$\beta$</sup>  and the backbone. (11S)-8-Chloro-11-[1-(methylsulfonyl)piperidin-4-yl]-6-piperazin-1-yl-11H-benzo[5,6]cyclohepta[1,2-*b*]pyridine (21) is a potent inhibitor ( $IC_{50} = 7 \text{ nM}$ ) of farnesyl protein transferase (FTase), a critical regulator of tumor cell growth.<sup>86</sup> Introduction of a piperazine moiety in position 6 of the benzocycloheptapyridine scaffold facilitates interactions with the catalytic zinc ion. Interestingly, the side chain of His201, which is stabilized by  $\pi$ – $\pi$ -interactions and steric packing effects, interacts with the 8-chloro group of the central scaffold (PDB code 2BED). This halogen bond has reasonable geometry, with a distance of 293 pm and a  $\sigma$ -hole angle of 165.6°.

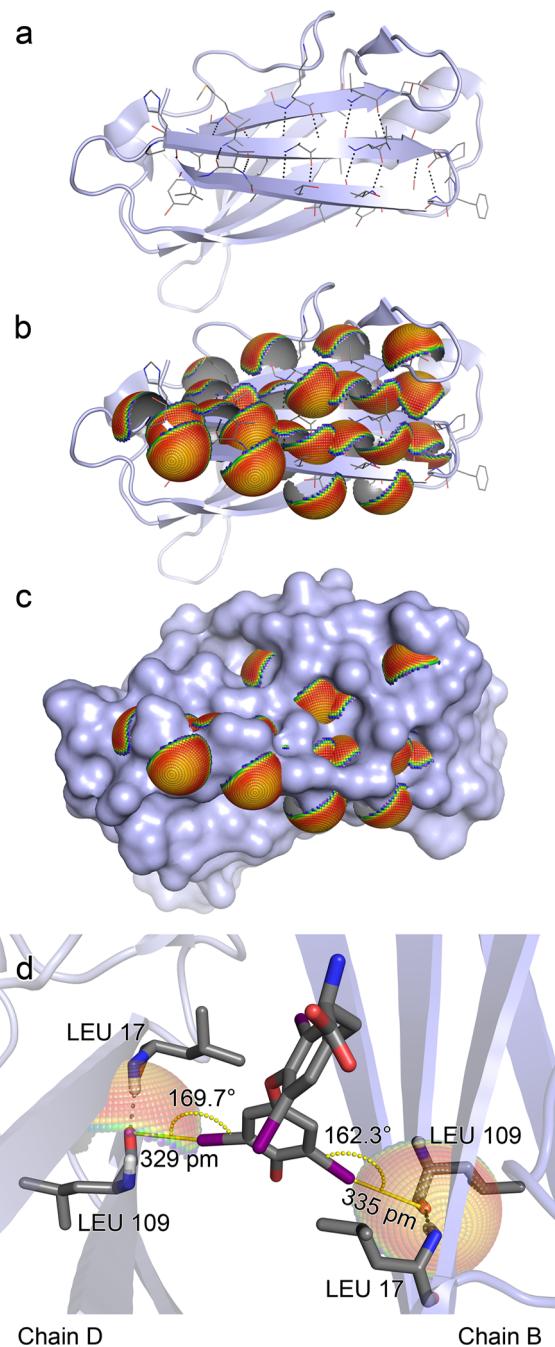
**4.6. Targeting Backbone Carbonyls in Structural Motifs.** Backbone carbonyl oxygens are the primary halogen bond acceptors in proteins. Loop regions within protein binding sites are important targets in molecular design, as we have seen for halogen bonding with the hinge binding motif in

kinases (Figure 6), but what about halogen bonds with backbone oxygens in  $\alpha$ -helices and  $\beta$ -sheets? Interestingly, it has recently been proposed that hydrogen bonds and halogen bonds can be orthogonal to each other;<sup>49</sup> i.e., they may share a common Lewis base by binding with perpendicular interaction vectors. The reason for this orthogonality is that the anisotropy of the electron distribution of the backbone carbonyl oxygen provides zones of similar interaction quality in-plane with the carbonyl but also out-of-plane (Figure 9). Within  $\alpha$ -helices and  $\beta$ -sheets, backbone carbonyl oxygens accept hydrogen bonds to form stable secondary structures. Since these hydrogen bonds are either in-plane with the carbonyl bond vectors (out-of-plane torsion  $\delta_1 \approx 0^\circ$ ) for ideal  $\beta$ -sheets or moderately below plane ( $\delta_1 \approx -40^\circ$ ) for ideal  $\alpha$ -helices, halogen bonds were proposed to be formed with a  $\delta_1$ -angle of roughly  $90^\circ$  for  $\beta$ -sheets or  $50^\circ$  for  $\alpha$ -helices.<sup>49</sup> In proteins, the values of these angles can be significantly different. Most  $\beta$ -sheets in protein structures, for example, have a significant right-handed twist.<sup>87–89</sup>

The halogen-bonding statistics of the PDB do not show any preference for  $\alpha$ -helices,  $\beta$ -sheets, or loop structures. Figure 13a shows a  $\beta$ -sheet substructure of a transthyretin homotetramer, which is a good example of halogen- and hydrogen-bonding orthogonality. Because of significant twists of the  $\beta$ -sheet segment, the hydrogen bonds deviate from the in-plain orientation. Projecting the spherical interaction surfaces from model calculations of iodine contacts with backbone carbonyls onto the backbone oxygens in this  $\beta$ -sheet structure reveals that most of the preferential interaction points are blocked by protein side chains (Figure 13b,c). While the side chains in a  $\beta$ -sheet offer auxiliary interactions for stabilizing the halogen bond, they limit the accessible interaction hot spots significantly. Twisting of the  $\beta$ -sheet enhances the accessibility of some carbonyls that would be occluded in an ideal  $\beta$ -sheet conformation.

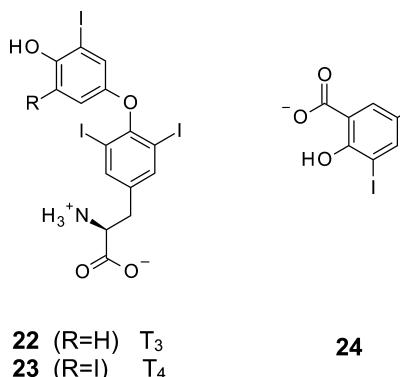
Transthyretin is involved in the transport and distribution of the thyroid hormones 3,5,3'-triiodo-L-thyronine (22, Chart 5) (also called T<sub>3</sub>) and 3,5,3',5'-tetraiodo-L-thyronine (23) (also called thyroxine or T<sub>4</sub>). Binding of 23 to transthyretin is an interesting example of nature exploiting halogen bonds for molecular recognition. 23 binds to transthyretin (Figure 13d) by forming two equivalent halogen bonds with  $\beta$ -sheets from two neighboring subunits (PDB code 1SN0).<sup>90</sup> The interaction geometries of both halogen bonds are close to optimal: the distances are 329 and 335 pm, and the  $\sigma$ -hole angles are 162.3° and 169.7°, respectively. The hydrogen and halogen bonds are indeed orthogonal in both cases, with angles ( $\alpha_{X\cdots O\cdots H(N)}$ ) of 88° and 90°.

In  $\alpha$ -helices, the vector of the backbone carbonyl oxygens is on average almost collinear with the helix axis but can show some deviation from this mean value. The accessibility is determined by the relative position of the hydrogen of the backbone amide in residue  $i + 4$  with respect to the C=O plane and the size and orientation of adjacent side chains (predominantly of residues  $i$ ,  $i + 3$ ,  $i + 4$ , and  $i + 7$ ). Analysis of the halogen-bond interaction spheres of carbonyl groups in an  $\alpha$ -helix (Figure 14a) and shielding effects by adjacent protein side chains (Figure 14b) suggest that  $\alpha$ -helices can engage in halogen-bond formation. An example of a halogen bond with a helical segment is found for human serum albumin (HSA), a transport protein that is important for drug action (Figure 14a). HSA is the most abundant protein in human plasma (~600  $\mu$ M) and typically binds a broad range of endogenous ligands and drugs (with preferentially acidic or electronegative



**Figure 13.** Targeting backbone carbonyl oxygens in  $\beta$ -sheets, exemplified by a characteristic secondary structural element in the structure of transthyretin (PDB code 1SN0). (a) The structure of transthyretin is shown as a cartoon, with the hydrogen-bond pattern of parts of a  $\beta$ -sheet highlighted. In the tetramer, this  $\beta$ -sheet segment is part of an extended symmetrical  $\beta$ -sheet formed with a neighboring subunit. Only one subunit of the homotetrameric structure is shown. (b) Interaction spheres from model calculations are plotted onto the backbone carbonyls within the  $\beta$ -sheet substructure. Color indicates binding quality (see Figure 9). (c) Depiction of the Connolly surface highlights loss of accessibility through sterical hindrance by side chains. (d) L-Thyroxine is complexed by simultaneously binding to two  $\beta$ -sheets from different subunits. Orthogonality of hydrogen bond (N–H $\cdots$ O) and halogen bond (C(ar)–I $\cdots$ O=C) is clearly visible in both chains ( $\alpha_{X\cdots O\cdots H(N)}$  is 88° for the contact with chain B and 90° for the contact with chain A). All pictures were prepared with PyMOL.<sup>12</sup>

Chart 5



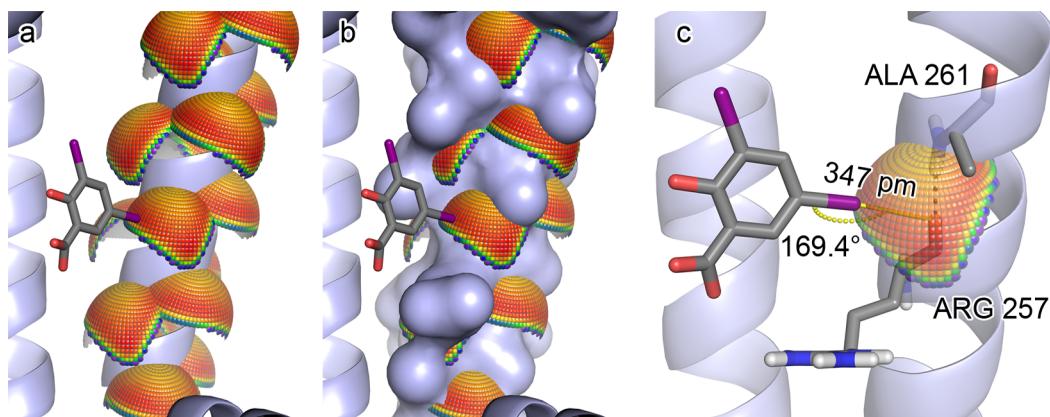
properties) at one of two primary binding sites.<sup>91</sup> 2-Hydroxy-3,5-diiodobenzoic acid (**24**) binds to the backbone carbonyl oxygen of Arg257, which is hydrogen-bonded to the backbone NH group of Ala261 within the framework of the regular  $\alpha$ -helix (PDB code 2BXL).<sup>92</sup> In this complex, the benzoic acid moiety of **24** is located in proximity to the positively charged guanidinium group of Arg257. The halogen bonding geometry is reasonable, showing a distance of 347 pm and a  $\sigma$ -hole angle of 169.4°. Orthogonality of hydrogen bond and halogen bond is confirmed by the angle  $\alpha_{X\cdots O\cdots H(N)} = 87^\circ$ . In summary, depending on the overall structural context and modulation by side chains, both  $\alpha$ -helical and  $\beta$ -sheet structures can in principle engage in halogen bonding via their backbone oxygens.

## 5. HALOGEN BONDS IN DRUG DISCOVERY: RECENT SUCCESS STORIES

There are numerous examples where halogen bonding has been successfully used in drug discovery in recent years, targeting a diverse set of medicinally relevant proteins. Bromine-backbone oxygen halogen bonds were found to increase the efficacy of agonists (**25**, Chart 6) of the  $\alpha 4\beta 2$  subtype of the nicotinic acetylcholine receptor<sup>93</sup> and of inhibitors (**26**) of the hepatitis C virus NS3-NS4A protease.<sup>94</sup> Introduction of halogens also improved the potency of different classes of HIV reverse transcriptase inhibitors (**16**, **27–29**).<sup>66,95,96</sup> Hardegger et al.

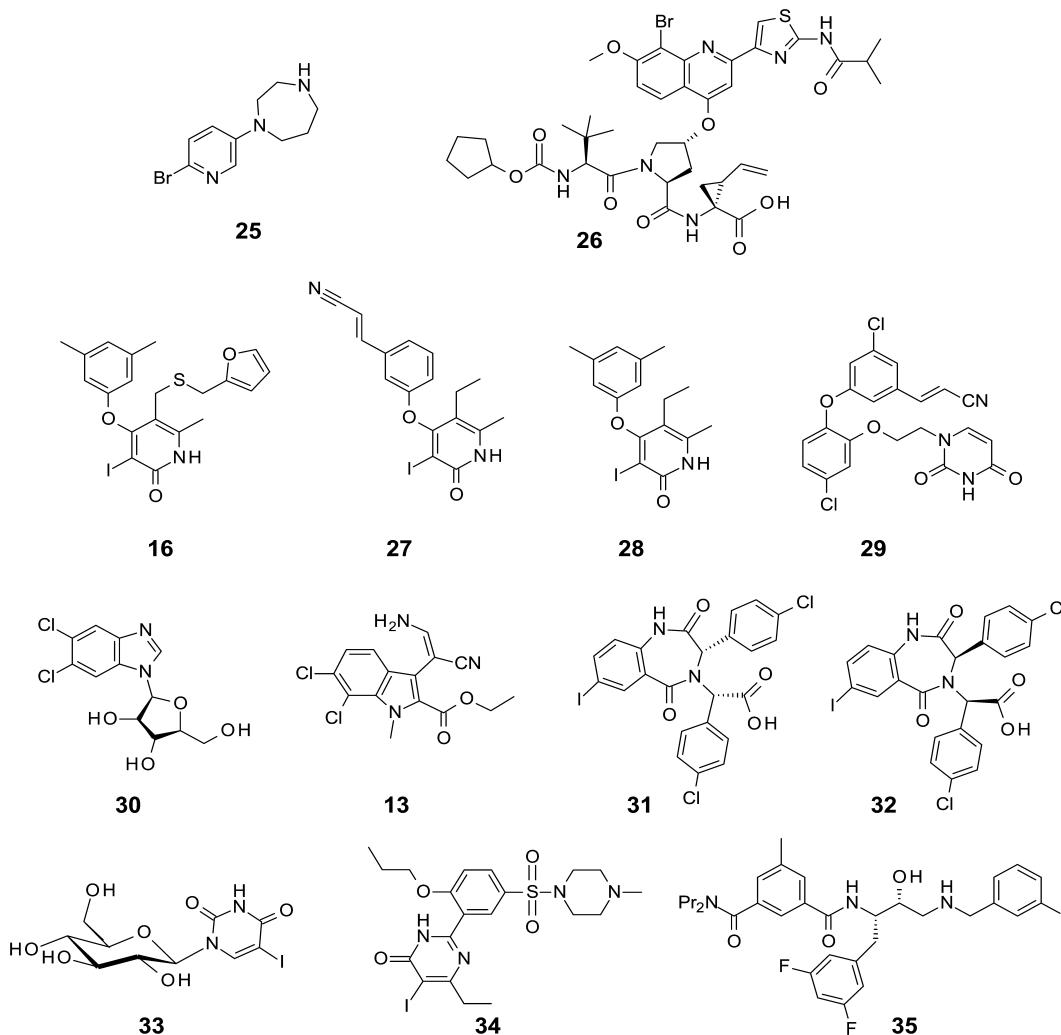
showed that the binding affinities of two classes of inhibitors of human cathepsin L and MEK1 kinase are significantly improved when a particular aryl ring is substituted with the heavier halogens chlorine, bromine, and iodine.<sup>24</sup> In both cases, the affinity of the ligand increases with the size of the halogen. High-resolution crystal structures of the L-cathepsin-ligand complexes show that the heavier halogens interact with the backbone oxygen of Gly61 in the binding pocket, adopting similar favorable interaction geometries.<sup>24</sup> The iodine–oxygen distance is 3.1 Å, much shorter than the sum of the van der Waals radii of both atoms, which is consistent with data from other high-resolution crystal structures and the quantum chemical calculations discussed above. In the case of the MEK1 kinase cocrystal structure, the iodine–oxygen distance is longer with 3.6 Å, which has been attributed to strong interactions by other parts of the ligand preventing optimal iodine–oxygen interaction.<sup>24</sup> A very interesting aspect of this study is that the halogen-binding pockets are of a very different nature in both cases. In cathepsin L, the halogen-binding site is in a polar environment at the surface of the enzyme, and the Gly61 backbone retains water solvation upon interaction with the halides. In contrast, the halogen-binding pocket of MEK1 kinase is buried within the enzyme and has a predominantly hydrophobic character. Hence, formation of favorable halogen bonds is not restricted by the overall polarity of potential binding sites in the target protein.

Halogens can also be exploited for the design of selective inhibitors of Cdk9, the kinase of the positive transcription elongation factor b. For many years, the ATP analogue 5,6-dichlorobenzimidazole-1- $\beta$ -D-ribofuranoside (**30**) has been widely used to inhibit transcription elongation, but it was only recently that structural studies revealed that halogen bonds form the basis for specific inhibition of Cdk9 kinase by **30**.<sup>97</sup> The two chlorines of the benzimidazole moiety contact the hinge region between the N- and C-terminal lobes of the ATP binding site and form halogen bonds with two backbone oxygens. The contact is strengthened by a side-on interaction of a backbone NH group with the negatively charged belt of one of the chlorines. Formation of this specific halogen-bond pattern in Cdk9 is facilitated by an induced-fit movement in other loops of the binding site, allowing optimal accommodation of the ligand in the ATP-binding pocket of Cdk9. The



**Figure 14.** Targeting backbone carbonyls in  $\alpha$ -helices exemplified by a characteristic secondary structural element in human serum albumin (PDB code 2BXL). (a) Spherical segments from model calculations are transformed onto the backbone carbonyl oxygens within the  $\alpha$ -helix. Color indicates binding quality (see Figure 9). (b) The Connolly surface shows reduced accessibility through steric hindrance by side chains. (c) **24** binds to the backbone carbonyl oxygen of Arg257. Orthogonality of hydrogen bond ( $N-H\cdots O$ ) and halogen bond ( $C(ar)-I\cdots O=C$ ) is observed ( $\alpha_{X\cdots O\cdots H(N)} = 87^\circ$ ). All pictures were prepared with PyMOL.<sup>12</sup>

Chart 6



related Cdk2 kinase has a more rigid active site and binds **30** in a different binding mode (with almost 300 times lower affinity) that involves only one halogen bond. Interestingly, similar halogen-mediated contacts with the hinge region of kinases have recently been observed for 7,8-dichloro-1-oxo- $\beta$ -carbolines (**13**) derived from the alkaloid bauerine C, which in one case also involves a two-pronged halogen-bond interaction with backbone oxygens.<sup>46,98</sup> Protein backbone oxygens can maintain their native hydrogen-bond patterns with neighboring amino acids upon halogen-bond formation, which is exemplified by a 3-iodo-4-phenoxypyridinone HIV transcriptase inhibitor (**16**). The cocrystal structure shows that the iodine substituent forms a halogen bond with the backbone carbonyl of Tyr188. The carbonyl oxygen slightly bends toward the iodine upon inhibitor binding but retains its native  $\beta$ -sheet hydrogen-bond interaction with Tyr118.<sup>66</sup> Substitution of the iodine with hydrogen results in a 300-fold reduced affinity.<sup>95</sup>

Halogen bonding has also been exploited for designing anticancer drugs (**31**, **32**) that target the p53 pathway. Over the past decade, numerous small molecules were designed that interrupt the interaction of the tumor suppressor p53 with its negative regulators MDM2 and MDMX, which are up-regulated in many tumors.<sup>99–101</sup> A halogen-substituted 1,4-benzodiazepine-2,5-dione ligand binds to the p53 binding site of human MDM2 and forms a halogen bond with the backbone

oxygen of Gln72.<sup>102</sup> Replacement of the halogen-bond-forming iodine by bromine or chlorine reduces the affinity of the ligand 12-fold,<sup>103</sup> again highlighting that the strength of halogen bonds decreases with decreasing size of the halogen.

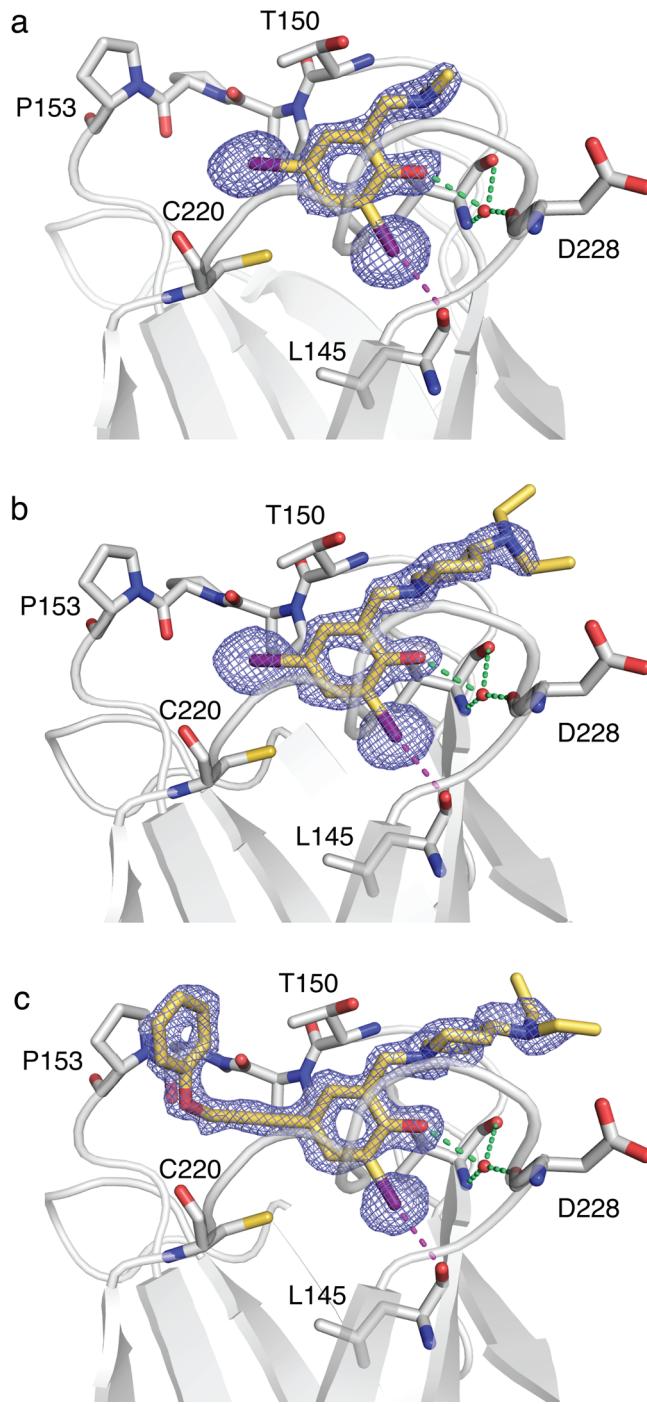
Examples of halogen bonds in protein–ligand complexes are not restricted to the above-mentioned interactions of halogens with the protein backbone but can also involve side chains. Glycogen phosphorylase plays a central role in glycogen metabolism and is an important target for the design of hyperglycemic drugs. CS halogen substituted glucopyranosyl nucleosides are some of the most potent inhibitors (**33**) of glycogen phosphorylase identified to date, showing an up to 12 times higher inhibitory effect than the nonhalogenated parent compound.<sup>104</sup> Here, the halogens interact with the carboxylate group of an aspartate side chain (Asp339) that lines the active-site pocket. The overall binding mode of the ligand seems to prevent optimal halogen bonding though, which is reflected in the relatively long iodine–oxygen distance of 3.8 Å. Not surprisingly, docking calculations using classical scoring algorithms failed to predict the experimentally observed binding modes.<sup>104</sup> In a different study, a halogen bond with a tyrosine side chain was exploited to design potent inhibitors of phosphodiesterase type 5 (**34**), resulting in a 13-fold increase in potency.<sup>105</sup>

The examples described so far also show that the effect of introducing a halogen depends on the optimal length and geometry of the halogen bond formed. In the examples with optimal halogen–oxygen distance (i.e., about 3.0 Å), iodines can improve the affinity of a ligand more than 100-fold compared to the unsubstituted ligand. Larger iodine–oxygen distances and deviations from the ideal  $\sigma$ -hole angle result in much lower affinity gains, as observed in the case of the glycogen phosphorylase inhibitors mentioned above and in a series of human  $\beta$ -secretase (BACE) inhibitors (35) where the increase in affinity upon substitution with iodine was only 26-fold.<sup>106</sup> The rather poor halogen-bonding geometry in the latter example (I–O distance of 4.2 Å and C–I–O angle of 139°) suggests that the affinity gain may be largely due to secondary effects.

In many of the examples presented above, halogens were introduced during lead optimization to improve the binding of established scaffolds to a particular target. We have recently shown that halogen bonding can also be successfully exploited for lead discovery.<sup>25</sup> The conformationally unstable p53 cancer mutant Y220C has a mutation-induced surface crevice that can be targeted by small molecules.<sup>107,108</sup> We designed a halogen-enriched fragment library to screen for small-molecule stabilizers that reactivate the tumor suppressor function of this mutant.<sup>25</sup> Despite the great potential of halogen moieties to increase ligand-binding affinities, standard libraries for fragment-based screening contain relatively few heavy halides. The Chembridge fragment library ([www.chembridge.com](http://www.chembridge.com)), for example, features 8293 compounds of which 193 contain Br (2.3%) and only 5 contain I (0.1%). From screening of our halogen-enriched fragment-based library, we identified 2,4-diodo-6-((methyl(1-methylpiperidin-4-yl)amino)methyl)-phenol as a small-molecule stabilizer of the Y220C mutant.<sup>25</sup> Crystallographic studies confirmed that one of the two iodine atoms forms a halogen bond with the backbone oxygen of Leu145 at the bottom of the largely hydrophobic, mutation-induced surface crevice (Figure 15). Structure-guided ligand optimization then exploited the second iodine atom in a Sonogashira coupling reaction to target a different subsite of the binding pocket. The resulting substituted 2-(aminomethyl)-4-ethynyl-6-iodophenols had a robust binding mode (Figure 15c), with significantly improved affinity, and showed mutant-specific induction of apoptosis in cancer cell lines.<sup>25</sup> These compounds are also useful chemical tools for studying p53 aggregation.<sup>109</sup> This example of the design of mutant p53 rescue drugs suggests a more widespread use of halogen-enriched fragment libraries (HEFLibs) in drug screening to complement classical fragment-based screening approaches. This way, favorable halogen-bond interactions with optimal geometry can be established at an early stage of the drug discovery process rather than at the stage of lead optimization, where the overall binding mode of a particular scaffold may restrict the available halogen-bond interactions or result in suboptimal interaction geometries.

## 6. TECHNIQUES FOR ENHANCING THE APPLICABILITY

It has taken considerable time for halogen bonding to be recognized as a relevant favorable interaction in medicinal chemistry, even though it has been known in materials science for many years.<sup>4</sup> One reason for this is that the effect is best described using high-level quantum chemistry, which is not a traditional part of the molecular design arsenal. As a more cost-



**Figure 15.** Crystal structures of the oncogenic p53 mutant Y220C in complex with small-molecule stabilizers.<sup>25</sup> The ligands are shown as yellow stick models, with the corresponding simulated-annealing ( $F_o - F_c$ ) electron density maps at a contour level of  $3\sigma$  shown in blue. The initial hit from screening of a halogen-enriched fragment library (a) has a robust binding mode (PDB code 4AGL). It is anchored to the mutation-induced surface crevice via a halogen bond of an iodine substituent with the backbone carbonyl oxygen of Leu145 at the bottom of the binding pocket (highlighted with a magenta broken line). The binding mode of the central scaffold was stable during the lead optimization process that targeted different subsites of the binding pocket, as shown for second and third generation ligands in panels b and c (PDB codes 4AGM and 4AGP). The figure was adapted from Wilcken et al.<sup>25</sup>

efficient variant in terms of computational effort compared to coupled-cluster methods or MP2, density functional theory (DFT) offers a way to work routinely with medium-sized model systems of up to a few hundred atoms. One major drawback of common density functionals has always been their inadequate description of dispersion forces, and pure DFT methods largely fail for the description of halogen bonds.<sup>110</sup> Dispersion corrections as proposed by Grimme<sup>37,38</sup> alleviate this weakness, and for halogen bonds and other systems where dispersive forces are important, third generation dispersion-corrected DFT<sup>36</sup> (DFT-D3) now offers comparable quality to high-level MP2 or coupled-cluster calculations.<sup>111</sup> Still, these rather accurate quantum chemical methods cannot be used routinely for a multitude of compounds as is required in high-throughput virtual screening. They can, however, be employed in model calculations, and the results of these can be useful for “cheaper” approaches such as machine learning to model certain molecular properties<sup>44</sup> and to derive rule-based models. We have recently developed a small algorithm for scaffold decoration of protein-bound ligands to achieve halogen bonding.<sup>112</sup> For each feasible example of ligands that may be decorated with a halogen, distance and  $\sigma$ -hole parameters are evaluated and combined in a simple scoring function to rank the quality of the interaction. Semiempirical methods such as AM1<sup>113</sup> and PM6<sup>114</sup> are also quantum chemical in nature but much simpler and faster than both DFT and higher-level methods. For halogen bonds, they do not perform very well,<sup>115</sup> but a new PM6 derivative, PM6-DH2X, has been developed<sup>116</sup> to improve the description of halogen bonding and was evaluated for protein–ligand complexes.<sup>33</sup> However, most high-throughput molecular design approaches such as docking are not based on even the simplest quantum chemical models but on force field-based approaches. In common force fields for ligands such as MMFF94<sup>117</sup> or the GAFF force field,<sup>118,119</sup> an atom type and a partial charge is assigned to each atom in the ligand. There is only one partial charge per atom, which naturally fails to describe the anisotropic nature of the electron density around the halogens chlorine, bromine, and iodine. As a consequence, halogen bonding is generally not recognized in force-field-based approaches. Several studies have therefore proposed the implementation of additional positively charged sites without mass on the “head” of the halogen in elongation of the C–X bond.<sup>28,120–123</sup> These are promising approaches, as they could be used to implement recognition of halogen bonding in molecular design approaches, including molecular dynamics. Scoring functions are the Achilles’ heel of docking procedures, as they are generally suitable for ligand enrichment in virtual screening but not for predicting ligand affinities.<sup>124</sup> The three main classes of scoring functions are force-field-based such as GoldScore,<sup>125,126</sup> empirical scoring functions such as ChemScore,<sup>127,128</sup> and knowledge-based scoring functions such as DrugScore.<sup>55,56</sup> Among these, empirical scoring functions are ideal candidates for implementing recognition of halogen bonds. In a recent study, Kuhn et al.<sup>129</sup> reported the development of an exciting new empirical scoring function, ScorpionScore, which combines recognition of molecular interactions such as halogen bonds and orthogonal multipolar interactions<sup>130</sup> with a network approach to model cooperative effects in ligand binding.<sup>131,132</sup> In light of these recent developments, it seems likely that recognition of halogen bonding will become part of routine molecular design packages in the near future.

## 7. PRACTICAL ASPECTS IN DRUG DISCOVERY

From our analysis of the relative strength of halogen bonds in theory and experiment, it is evident that based on the size and range of the  $\sigma$ -hole, iodine shows stronger effects than bromine, which in turn is superior to chlorine. Changes in the chemical scaffolds and synergistic effects with other substituents can tune the strength of halogen bonds significantly. Therefore, not only iodine but also bromine and chlorine should be useful for drug discovery efforts to improve molecular recognition through halogen bonding.

Analysis of the MDDR (MDL Drug Data Report, version 2011.2 (32.11), Table 4) shows that out of 1574 drugs, almost

**Table 4. MDDR<sup>a</sup> Analysis of Heavy Halides (Cl, Br, and I) in Launched Drugs and Candidates within Preclinical to Clinical Development**

compd containing <sup>b</sup>	total number	preclinical	phases 1–3	clinical	launched
chlorine	37043	2409	417	17	231
bromine	6409	374	64	1	23
iodine	1260	118	21	5	18
total number <sup>c</sup>	215405	14142	2862	117	1574

<sup>a</sup>MDDR, version 2011.2 (32.11). <sup>b</sup>Searches were done for the fragment C–X, with X = Cl, Br, I. <sup>c</sup>Total number of unique compounds with reported structure.

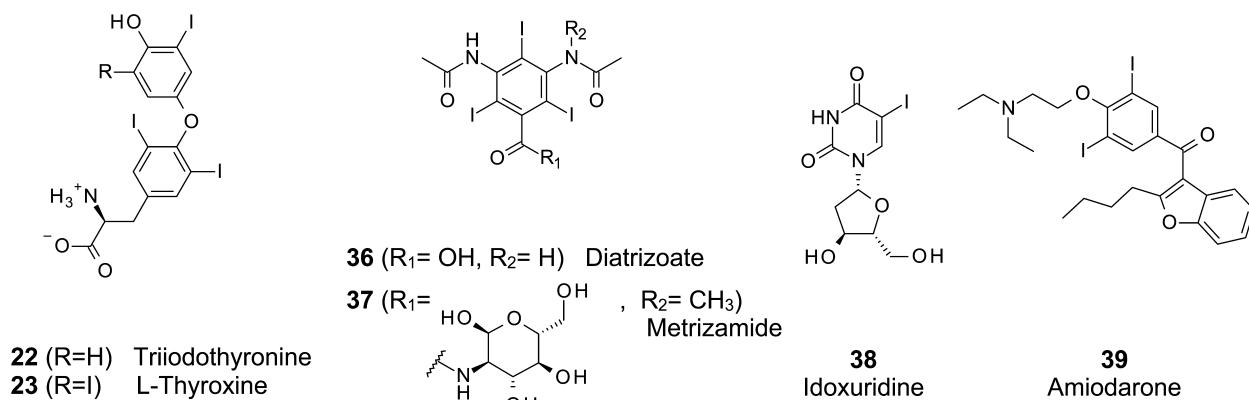
14.7% contain chlorine while only 1.5% contain bromine. Interestingly, 1.14% of all launched drugs contain iodine, whereas the percentage of iodine-containing molecules for all reported compounds is two times lower (0.58%). But this apparent bias is easily explained when considering the structures and therapeutic applications of these drugs (Chart 7).

Two important classes of iodine-containing drugs are the thyroid hormones, in particular triiodothyronine (22) and thyroxine (23), as well as several radiographic contrast agents containing a 1,3,5-triiodobenzene substructure for medical diagnostics. This application is facilitated by the electronic structure of iodine, which has suitable K shell binding energies for absorption of X-rays.<sup>133</sup> For medical and imaging purposes, a variety of ionic and nonionic organoiodines such as diatrizoate (36) and metrizamide (37) have been used, introducing a significant bias into the statistics. Two iodine-containing drugs with completely different applications are idoxuridine (38) and amiodarone (39). While 38 is a nucleoside analogue (pyrimidine antimetabolite) used for treating herpes simplex keratitis,<sup>134</sup> 39 is a class III antiarrhythmic agent with a rather unique profile indicated for management of life-threatening recurrent ventricular fibrillation or hemodynamically unstable ventricular tachycardia.<sup>135,136</sup>

In contrast, drugs containing bromine have a significantly broader range of molecular targets and indications. A diverse selection of such drugs is depicted in Chart 8. Primary drug targets are, for instance, the histamine H<sub>1</sub> receptor (40, 41), the  $\alpha_1$  adrenergic receptor (42), the dopamine D<sub>2</sub> receptor (43, 44), the  $\alpha_2$  adrenergic receptor (45), the AT1 receptor (46), the GABA<sub>A</sub> receptor (47), and the cyclooxygenases COX1/2 (48).

A multitude of chlorine-containing drugs for various types of targets are available. Compared to bromine and iodine, chlorine has a more established role and value in drug discovery. Interestingly, there seems to be a tendency for bromine to be

Chart 7. Approved Drugs Containing Iodine



more prominently featured in phases 1–3 and preclinical trials than in approved drugs. The fraction of compounds containing bromine increases from 1.5% for drugs to 2.2% for clinical candidates in phases 1–3. In preclinical trials this fraction is 2.6%, whereas for all compounds it is 3.0%. This trend may change with increasing sample size over the coming years. For iodine such a trend does currently not exist.

What is the rationale behind this strong preference for chlorine over bromine and iodine? A traditionally strong argument against the incorporation of the heavier halides bromine and iodine is their reactivity. Without taking into consideration the highly reactive alkyl halides, aryl halides are typically perceived as reactive intermediates in synthesis strategies but usually not as final products. The multitude of available cross-coupling reactions for aryl halides used for scaffold assembly in medicinal chemistry<sup>137–139</sup> has probably reinforced this notion. Given the proven creativity and inventiveness of medicinal and organic chemists, stability issues in synthesis should be resolved by tuning of reaction conditions or devising strategies for late-stage introduction of such heavy halides.<sup>140</sup> Of course, the potential gain has to be weighed against the necessary effort and the estimated risk.

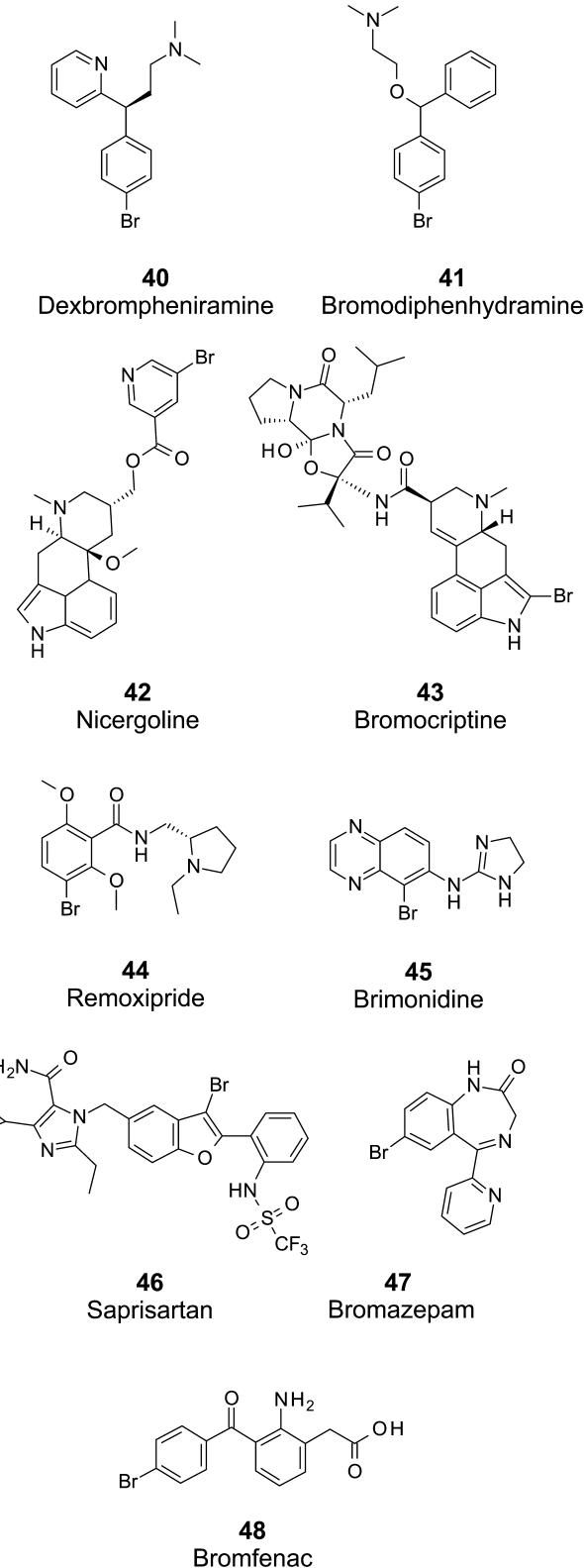
The potential gain is obvious: through the semidirectional halogen-bonding effect, increases in binding affinity and specificity can be achieved when replacing chlorine with bromine or iodine. Yet an assessment of the possible drawbacks of having bromine and iodine in drug discovery and development is important. First, heavier halides will add weight to the compound. Molecular weight has been regarded as one of the prime parameters for ADME properties. Small organic molecules usually consist of “heavy atoms” in the MW range between ~12 g/mol for carbon and ~32 g/mol for sulfur. With one bromine atom adding ~80 g/mol and one iodine atom adding ~127 g/mol, they would equal more than 6 and 10 carbon atoms, respectively. Other parameters such as size, volume, or molecular surface area, however, will not change dramatically. It is therefore unreasonable to assume that models based on strict additivity of the molecular weight can be applied to bromine and iodine and characterize them correctly. A balanced measure of heavy atom count, increase in volume or solvent-accessible surface area and molecular weight may be preferable. Since the rule of five,<sup>141</sup> new concepts of ADMET prediction have been proposed based on careful analysis of data sets reflecting the drug discovery process and reasons for its failure.<sup>142</sup> The Pfizer 3/75 rule describes a reduced risk of in vivo toxicity based on low ClogP (<3) and high TPSA

(>75).<sup>143</sup> The GSK 4/400 rule suggests favorable ADMET properties to be primarily associated with ClogP < 4 and MW < 400.<sup>144</sup> In addition, the phrase “molecular obesity” has been coined, meaning that ligand efficiency (binding affinity/heavy atom count)<sup>145</sup> must not be achieved by increased lipophilicity. The resulting LELP index<sup>146</sup> divides log P by ligand efficiency, giving a measure of how log P and molecular size “manipulate” the affinity. All of these analyses imply that one of the key parameters is log P.<sup>147</sup> Although there is a tendency of increased lipophilicity with bromine and iodine, this increase is typically moderate. The key to improve the strength of halogen bonding is to increase the σ-hole. A stronger electrostatic potential will naturally also improve polar interactions, resulting in a lower log P. The source of enhanced affinity through halogen bonding may therefore actually reduce lipophilicity and enhance solubility. Clearly, there is a need for more systematic experimental data on the relationship between halogen-bond donors and ADMET properties. QM-based descriptors may be also useful for describing and predicting the ADMET behavior.

Other key aspects of drug discovery are metabolism and toxicity. On the basis of the endogenous production of the thyroid hormones, adequate enzymatic equipment for iodination and deiodination must be present in humans.<sup>148,149</sup> As shown before, binding of thyroid hormones to their transport proteins<sup>90</sup> and receptors<sup>150</sup> is mediated by halogen bonding. It is therefore not surprising that the proposed mechanism of deiodination involves halogen bonding of the iodine with chalcogens (S or Se).<sup>149,151</sup> Development of enzymatic systems that regulate the cellular uptake and disposal of iodometabolites is even thought to have been an important step in evolution.<sup>152</sup> Alternatively, other enzymes can perform dehalogenation through oxidation and other mechanisms. One enzyme family particularly relevant for drug discovery is cytochrome P450, which is known to perform oxidative dehalogenation as well.<sup>153</sup> Extensive data on metabolism and side effects have been gathered during some decades of therapeutic application of 39,<sup>136</sup> a Vaughan Williams class III antiarrhythmic agent. It is used for treatment of a wide range of cardiac tachyarrhythmias, including both ventricular and supraventricular (atrial) arrhythmias.<sup>135</sup>

Typical doses of 39 inhibit 5'-deiodinase, leading to a reduction of T<sub>3</sub> in peripheral tissues. It may even block the binding of T<sub>3</sub> to its nuclear receptor. In some patients, this can result in abnormal thyroid function (increase of T<sub>4</sub>/rT<sub>3</sub> and decrease of T<sub>3</sub>).<sup>154</sup> During treatment with 39, an excess of iodine is produced, and the diagnosis of hyperthyroidism and

Chart 8. Selection of Approved Drugs Containing Bromine

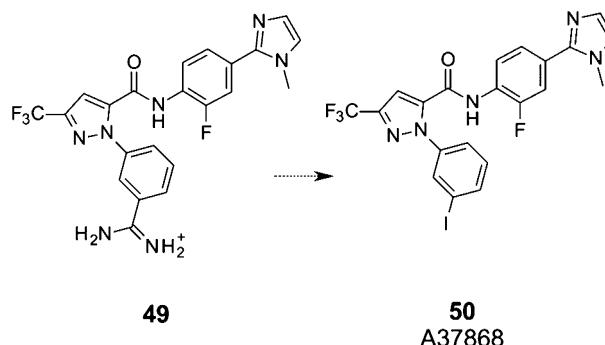


hypothyroidism can be obscured. **39** has a half-life of about 30 days (range of 15–100 days).<sup>136</sup> Side effects are not negligible but are generally dose-dependent. During liver passage, **39** is extensively metabolized via CYP2C8 (<1% unchanged in urine) and hence can affect the metabolism of numerous other drugs.

The example of **39** suggests that high similarity to the thyroid hormones can produce side effects and possibly perturb regular thyroid function. Hence, targets of these hormones should be considered as antitargets for the development of iodine-containing drugs. Like many other substances, significant interactions with the CYP system are possible and need to be considered during development. Certainly, it is not trivial to attribute such adverse effects to the halogen per se, to the scaffold per se, or to the particular combination of halogen and scaffold. To decompose such CYP interactions into principal contributions, more systematic, comparative studies may be required. But even given the limiting side effects of **39**, it seems that a beneficial therapeutic long-term application of compounds containing heavy halides should be possible.

In the context of industrial drug discovery, it is encouraging to note that halogen bonds in protein–ligand interactions are no longer just the result of serendipitous discovery (as it was up to some years ago) but increasingly also of rational drug design. A seminal study of rational exploitation of halogen bonding was reported in 2009 by Lam and colleagues from Bristol-Myers Squibb at the 238th National Meeting of the American Chemical Society in Washington, DC, where they showed the successful use of halogen bonding to replace an amidine-ASP189 hydrogen bonding contact of **49** (Chart 9) in complex with factor Xa by an iodine-ASP189 halogen bonding contact, resulting in compound **50**.<sup>9,155</sup>

Chart 9. Halogen Bonding in Factor Xa Drug Design



When considering the use of iodine-containing compounds, another application could be added value: while stable isotopes of iodine may be useful for conventional therapeutic approaches and for the drug discovery process, radioiodine-based therapy or diagnostics may open new avenues, e.g., for imaging or treating human malignancies.

## 8. CONCLUSION AND OUTLOOK

In this Perspective, we have analyzed the nature of halogen bonding as perceived through theoretical models as well as experimental studies. A thorough analysis of structural information available from the PDB has shed light on possible interaction partners for halogenated ligands in the protein backbone and side chains. Systematic experimental evaluations allow derivation of rough guidelines regarding possible affinity gains from introduction of halogen bonds for scaffold decoration. It has emerged that halogen bonding is a semidirectional molecular interaction with weak to medium strength, depending on the experimental system, and can compete or act synergistically with weak to moderately strong

hydrogen bonds. The highest affinity gains in protein–ligand complexes are expected for the introduction of the heavier halides bromine and iodine into aromatic scaffolds. While halogens have been used for scaffold decoration in many cases, the strength of the interactions they form makes iodinated or brominated fragments ideal candidates for lead discovery strategies, redefining halogen bonds as core interactions. To this end, we have recently proposed a general strategy that makes use of halogen-enriched fragment libraries (HEFLibs) to identify unique binding modes.<sup>25</sup> Common strategies to avoid failure in the lead optimization process include molecular-weight-based criteria<sup>142,156,157</sup> that may not be applicable for moieties like iodine which contributes as much as 10 carbon equivalents to the overall molecular weight. In addition, more research into transport phenomena and ADME properties of halogenated compounds featuring the heavier halides is needed. Few computational molecular design packages recognize halogen bonding as a favorable interaction at present, and although there are a number of promising approaches such as the modification of force fields<sup>28,115,121–123</sup> and a new empirical scoring function,<sup>129</sup> it has not yet become part of the regular drug discovery workflow, a fact that hinders more widespread use and recognition of the phenomenon. Further interdisciplinary studies will be needed to fully understand and validate the potential of this interaction in biological systems so that halogen bonding may play a prominent role in the development of future therapeutics.

## AUTHOR INFORMATION

### Corresponding Author

\*Phone: +49 7071 29 74567. E-mail: frank.boeckler@uni-tuebingen.de.

### Author Contributions

§These authors contributed equally.

### Notes

The authors declare no competing financial interest.

### Biographies

**Rainer Wilcken** studied Chemistry at Friedrich Schiller University Jena, Germany, specializing in computational and quantum chemistry. He joined Prof. Frank Boeckler's group as a Ph.D. student in 2008 and worked on design and testing of small molecules for the drug rescue of mutant p53. Since 2011, he is a postdoctoral fellow in Prof. Sir Alan Fersht's group at the MRC Laboratory of Molecular Biology, Cambridge, U.K., focusing on drug discovery and biophysical screening techniques.

**Markus O. Zimmermann** studied computational biology/bioinformatics at the University of Tuebingen, Germany. He received his diploma in 2009 and is currently a Ph.D. student at the Institute of Pharmaceutical Sciences in Prof. Dr. Frank Boeckler's group at Laboratory for Molecular Design and Pharmaceutical Biophysics.

**Andreas Lange** studied chemistry at the Ludwig-Maximilians University of Munich, Germany, and received his M.Sc. degree in 2010. Currently he is a Ph.D. student at the Department of Pharmaceutical Chemistry, Laboratory for Molecular Design and Pharmaceutical Biophysics, at the University of Tuebingen, Germany, headed by Prof. Frank M. Boeckler.

**Andreas C. Joerger** studied Chemistry at the University of Freiburg, Germany, specializing in biochemistry, and received his Ph.D. in 2000 for crystallographic and functional studies on fuculose 1-phosphate aldolase in the group of Prof. Georg E. Schulz. He then did a postdoc with Prof. Sir Alan Fersht at the MRC Centre for Protein Engineering

in Cambridge, U.K., working on protein evolution and structural studies of p53 cancer mutants. He is currently a Senior Scientist at the MRC Laboratory of Molecular Biology, Cambridge, U.K. His main areas of research are the structural biology of tumor suppressor pathways, X-ray crystallography of protein–ligand complexes, and the structural basis of disease-related mutations and its implications for drug discovery.

**Frank M. Boeckler** received his Ph.D. in Medicinal Chemistry at the University of Erlangen, Germany, in 2004 in the group of Prof. Peter Gmeiner, where he specialized in computational chemistry and drug design. In 2006, he joined Prof. Sir Alan Fersht at the MRC Center for Protein Engineering in Cambridge, U.K., as Marie Curie Fellow, where he focused on molecular biology and biophysics and discovered p53 mutant stabilizers as potential new cancer therapeutics. In 2008, he was appointed Professor for Bioanalytics at LMU Munich, Germany. In 2010, he moved to Eberhard Karls University Tuebingen, Germany, as Professor for Medicinal Chemistry/Drug Design. His work is dedicated to understanding molecular interactions as the foundation for chemical biology and drug research and to applying theoretical and biophysical methods to cancer research.

## ACKNOWLEDGMENTS

The authors thank Thomas E. Exner for preparing and optimizing ESP plots with MOLCAD II. We are grateful to Boris Ferger and Christofer Tautermann for fruitful discussions and help with statistical analysis of halogens in drugs and clinical candidates. We thank Antonina Andreeva and Chris Johnson for valuable comments on different parts of the manuscript. For support in screening literature and PDB structures, we thank Patrick Steinmetz and Robin Faehnrich. This work was partially supported by the federal state of Baden-Wuerttemberg, Germany.

## ABBREVIATIONS USED

AMBER, assisted model building with energy refinement; AT1, angiotensin II receptor, type 1; aug-cc-pVDZ, Dunning's correlation-consistent double- $\zeta$  basis set augmented by diffuse functions; aug-cc-pVDZ-PP, Dunning's correlation-consistent double- $\zeta$  basis set augmented by diffuse functions and using relativistic pseudopotentials; BB, N-methylacetamide ("backbone" model system); C26, C26 colon carcinoma cell line; CBS, complete basis set; CCSD(T), coupled-cluster method with single, double, and perturbative triple excitations; CDK2/CDK9, cyclin-dependent kinase 2/cyclin-dependent kinase 9; CK2, casein kinase 2; CLK3, CDC-like kinase 3; CYP2C8, cytochrome P4502C8; ESP, electrostatic potential; ffBXB, force field for biological halogen bonds; FPT, farnesyl protein transferase; GAFF, generalized AMBER force field; GluKS, kainate receptor; GluR2, glutamate receptor 2; hCatL, human cathepsin L; HEFLibs, halogen-enriched fragment libraries; hPNMT, phenylethanolamine N-methyltransferase; LB, Lewis base; LELP, ligand-efficiency-dependent lipophilicity; MDDR, MDL Drug Data Report; MDM2, murine double minute 2 E3 ubiquitin ligase, also known as HDM2 in human; MDMX, MDM2-like p53-binding protein, also known as MDM4; MEK1, dual specificity mitogen-activated protein kinase kinase 1; MMFF94, Merck molecular force field 94; MP2, second-order Møller-Plesset perturbation theory; OPLS-AA, optimized potentials for liquid simulations—all atoms; PM6, parameterized model 6; PM6-DH2X, PM6 method augmented by dispersion and hydrogen and halogen bonding correction; QM/MM, quantum mechanics/molecular mechanics; SAD,

single-wavelength anomalous dispersion; SAPT, symmetry-adapted perturbation theory; SIRAS, single isomorphous replacement plus anomalous scattering; TPSA, total polar surface area; TZVPP, Karlsruhe triple- $\zeta$  basis set plus polarization

## ■ REFERENCES

- (1) Dumas, J. M.; Peurichard, H.; Gomel, M. CX4···base interactions as models of weak charge-transfer interactions: comparison with strong charge-transfer and hydrogen-bond interactions. *J. Chem. Res.* **1978**, *2*, 54–55.
- (2) Clark, T.; Hennemann, M.; Murray, J. S.; Politzer, P. Halogen bonding: the sigma-hole. *J. Mol. Model.* **2007**, *13* (2), 291–296.
- (3) Clark, T.  $\sigma$ -Holes. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* [Online early access] DOI: 10.1002/wcms.1113. Published Online: Aug 8, 2012.
- (4) Metrangolo, P.; Meyer, F.; Pilati, T.; Resnati, G.; Terraneo, G. Halogen bonding in supramolecular chemistry. *Angew. Chem., Int. Ed.* **2008**, *47* (33), 6114–6127.
- (5) Voth, A. R.; Hays, F. A.; Ho, P. S. Directing macromolecular conformation through halogen bonds. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (15), 6188–6193.
- (6) Hassel, O. Structural aspects of interatomic charge-transfer bonding. *Science* **1970**, *170* (3957), 497–502.
- (7) Hassel, O.; Hvoslef, J. The structure of bromine 1,4-dioxanate. *Acta Chem. Scand.* **1954**, *8*, 873–873.
- (8) Cody, V.; Murray-Rust, P. Iodine···X(O, N, S) intermolecular contacts: models of thyroid hormone–protein binding interactions using information from the Cambridge crystallographic data files. *J. Mol. Struct.* **1984**, *112* (3–4), 189–199.
- (9) Parisini, E.; Metrangolo, P.; Pilati, T.; Resnati, G.; Terraneo, G. Halogen bonding in halocarbon–protein complexes: a structural survey. *Chem. Soc. Rev.* **2011**, *40* (5), 2267–2278.
- (10) Lu, Y.; Wang, Y.; Zhu, W. Nonbonding interactions of organic halogens in biological systems: implications for drug discovery and biomolecular design. *Phys. Chem. Chem. Phys.* **2010**, *12* (18), 4543–4551.
- (11) Auffinger, P.; Hays, F. A.; Westhof, E.; Ho, P. S. Halogen bonds in biological molecules. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (48), 16789–16794.
- (12) *The PyMOL Molecular Graphics System*, version 1.5.0.4; Schrödinger, LLC: New York.
- (13) Ramasubbu, N.; Parthasarathy, R.; Murray-Rust, P. Angular preferences of intermolecular forces around halogen centers: preferred directions of approach of electrophiles and nucleophiles around carbon–halogen bond. *J. Am. Chem. Soc.* **1986**, *108* (15), 4308–4314.
- (14) Lommers, J. P. M.; Stone, A. J.; Taylor, R.; Allen, F. H. The nature and geometry of intermolecular interactions between halogens and oxygen or nitrogen. *J. Am. Chem. Soc.* **1996**, *118* (13), 3108–3116.
- (15) Politzer, P.; Murray, J. S.; Clark, T. Halogen bonding: an electrostatically-driven highly directional noncovalent interaction. *Phys. Chem. Chem. Phys.* **2010**, *12* (28), 7748–7757.
- (16) Lu, Y.; Shi, T.; Wang, Y.; Yang, H.; Yan, X.; Luo, X.; Jiang, H.; Zhu, W. Halogen bonding—a novel interaction for rational drug design? *J. Med. Chem.* **2009**, *52* (9), 2854–2862.
- (17) Politzer, P.; Lane, P.; Concha, M. C.; Ma, Y. G.; Murray, J. S. An overview of halogen bonding. *J. Mol. Model.* **2007**, *13* (2), 305–311.
- (18) Metrangolo, P.; Murray, J. S.; Pilati, T.; Politzer, P.; Resnati, G.; Terraneo, G. The fluorine atom as a halogen bond donor, viz. a positive site. *CrystEngComm* **2011**, *13* (22), 6593–6596.
- (19) Metrangolo, P.; Murray, J. S.; Pilati, T.; Politzer, P.; Resnati, G.; Terraneo, G. Fluorine-centered halogen bonding: a factor in recognition phenomena and reactivity. *Cryst. Growth Des.* **2011**, *11* (9), 4238–4246.
- (20) Brickmann, J.; Exner, T. E.; Keil, M.; Marhofer, R. J. Molecular graphics—trends and perspectives. *J. Mol. Model.* **2000**, *6* (2), 328–340.
- (21) Brickmann, J.; Exner, T. E.; Gimmler, J.; Lautenschläger, P.; Heiden, W.; Moekel, G.; Zahn, D. *MOLCAD II*; MOLCAD GmbH: Darmstadt, Germany; <http://www.molcad.de>.
- (22) Bissantz, C.; Kuhn, B.; Stahl, M. A medicinal chemist's guide to molecular interactions. *J. Med. Chem.* **2010**, *53* (14), 5061–5084.
- (23) Hardegger, L. A.; Kuhn, B.; Spinnler, B.; Anselm, L.; Ecabert, R.; Stihle, M.; Gsell, B.; Thoma, R.; Diez, J.; Benz, J.; Plancher, J.-M.; Hartmann, G.; Banner, D. W.; Haap, W.; Diederich, F. Systematic investigation of halogen bonding in protein–ligand interactions. *Angew. Chem., Int. Ed.* **2011**, *50* (1), 314–318.
- (24) Hardegger, L. A.; Kuhn, B.; Spinnler, B.; Anselm, L.; Ecabert, R.; Stihle, M.; Gsell, B.; Thoma, R.; Diez, J.; Benz, J.; Plancher, J.-M.; Hartmann, G.; Isshiki, Y.; Morikami, K.; Shimma, N.; Haap, W.; Banner, D. W.; Diederich, F. Halogen bonding at the active sites of human cathepsin L and MEK1 kinase: efficient interactions in different environments. *ChemMedChem* **2011**, *6* (11), 2048–2054.
- (25) Wilcken, R.; Liu, X.; Zimmermann, M. O.; Rutherford, T. J.; Fersht, A. R.; Joerger, A. C.; Boeckler, F. M. Halogen-enriched fragment libraries as leads for drug rescue of mutant p53. *J. Am. Chem. Soc.* **2012**, *134* (15), 6810–6818.
- (26) Riley, K. E.; Hobza, P. Investigations into the nature of halogen bonding including symmetry adapted perturbation theory analyses. *J. Chem. Theory Comput.* **2008**, *4* (2), 232–242.
- (27) Kolář, M.; Hobza, P. On Extension of the current biomolecular empirical force field for the description of halogen bonds. *J. Chem. Theory Comput.* **2012**, *8* (4), 1325–1333.
- (28) Jorgensen, W. L.; Schyman, P. Treatment of halogen bonding in the OPLS-AA force field: application to potent anti-HIV agents. *J. Chem. Theory Comput.* **2012**, *8* (10), 3895–3901.
- (29) Wilcken, R.; Zimmermann, M.; Lange, A.; Zahn, S.; Boeckler, F. Using halogen bonds to address the protein backbone: a systematic evaluation. *J. Comput.-Aided. Mol. Des.* **2012**, *26* (8), 935–945.
- (30) Kümmel, H. G. *A Biography of the Coupled Cluster Method*; World Scientific Publishing: Singapore, 2002.
- (31) Möller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46* (7), 618.
- (32) Pople, J. A.; Binkley, J. S.; Seeger, R. Theoretical models incorporating electron correlation. *Int. J. Quantum Chem.* **1976**, 1–19.
- (33) Dobeš, P.; Řezáč, J.; Fanfrlík, J.; Otyepka, M.; Hobza, P. Semiempirical quantum mechanical method PM6-DH2X describes the geometry and energetics of CK2-inhibitor complexes involving halogen bonds well, while the empirical potential fails. *J. Phys. Chem. B* **2011**, *115* (26), 8581–8589.
- (34) Steiner, T. The hydrogen bond in the solid state. *Angew. Chem., Int. Ed.* **2002**, *41* (1), 49–76.
- (35) Murray, J. S.; Riley, K. E.; Politzer, P.; Clark, T. Directional weak intermolecular interactions: sigma-hole bonding. *Aust. J. Chem.* **2010**, *63* (12), 1598–1607.
- (36) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (37) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27* (15), 1787–1799.
- (38) Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **2004**, *25* (12), 1463–1473.
- (39) Hennemann, M.; Murray, J.; Politzer, P.; Riley, K.; Clark, T. Polarization-induced  $\sigma$ -holes and hydrogen bonding. *J. Mol. Model.* **2012**, *18* (6), 2461–2469.
- (40) Riley, K.; Murray, J.; Fanfrlík, J.; Řezáč, J.; Solá, R.; Concha, M.; Ramos, F.; Politzer, P. Halogen bond tunability I: the effects of aromatic fluorine substitution on the strengths of halogen-bonding interactions involving chlorine, bromine, and iodine. *J. Mol. Model.* **2011**, *17* (12), 3309–3318.
- (41) Riley, K. E.; Murray, J. S.; Fanfrlík, J.; Rezac, J.; Sola, R. J.; Concha, M. C.; Ramos, F. M.; Politzer, P. Halogen bond tunability II: the varying roles of electrostatic and dispersion contributions to

- attraction in halogen bonds. *J. Mol. Model.* [Online early access]. DOI: 10.1007/s00894-00012-01428-x. Published May 29, 2012.
- (42) Metrangolo, P.; Resnati, G. Halogen bonding: a paradigm in supramolecular chemistry. *Chem.—Eur. J.* **2001**, *7* (12), 2511–2519.
- (43) Riley, K. E.; Hobza, P. Strength and character of halogen bonds in protein–ligand complexes. *Cryst. Growth Des.* **2011**, *11* (10), 4272–4278.
- (44) El Kerdawy, A.; Wick, C. R.; Hennemann, M.; Clark, T. Predicting the sites and energies of noncovalent intermolecular interactions using local properties. *J. Chem. Inf. Model.* **2012**, *52* (4), 1061–1071.
- (45) Bauza, A.; Quinonero, D.; Frontera, A.; Deya, P. M. Substituent effects in halogen bonding complexes between aromatic donors and acceptors: a comprehensive ab initio study. *Phys. Chem. Chem. Phys.* **2011**, *13* (45), 20371–20379.
- (46) Fedorov, O.; Huber, K.; Eisenreich, A.; Filippakopoulos, P.; King, O.; Bullock, A. N.; Szklarczyk, D.; Jensen, L. J.; Fabbro, D.; Trappe, J.; Rauch, U.; Bracher, F.; Knapp, S. Specific CLK Inhibitors from a novel chemotype for regulation of alternative splicing. *Chem. Biol.* **2011**, *18* (1), 67–76.
- (47) Grant, S. K.; Lunney, E. A. Kinase inhibition that hinges on halogen bonds. *Chem. Biol.* **2011**, *18* (1), 3–4.
- (48) Riley, K. E.; Murray, J. S.; Politzer, P.; Concha, M. C.; Hobza, P. Br···O complexes as probes of factors affecting halogen bonding: interactions of bromobenzenes and bromopyrimidines with acetone. *J. Chem. Theory Comput.* **2008**, *5* (1), 155–163.
- (49) Voth, A. R.; Khuu, P.; Oishi, K.; Ho, P. S. Halogen bonds as orthogonal molecular interactions to hydrogen bonds. *Nat. Chem.* **2009**, *1* (1), 74–79.
- (50) Wilcken, R.; Zimmermann, M. O.; Lange, A.; Zahn, S.; Kirchner, B.; Boeckler, F. M. Addressing methionine in molecular design through directed sulfur–halogen bonds. *J. Chem. Theory Comput.* **2011**, *7* (7), 2307–2315.
- (51) Shields, Z. P.; Murray, J. S.; Politzer, P. Directional tendencies of halogen and hydrogen bonds. *Int. J. Quantum Chem.* **2010**, *110* (15), 2823–2832.
- (52) Murray, J.; Lane, P.; Politzer, P. Expansion of the  $\sigma$ -hole concept. *J. Mol. Model.* **2009**, *15* (6), 723–729.
- (53) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **1977**, *112* (3), 535–542.
- (54) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235–242.
- (55) Neudert, G.; Klebe, G. DSX: a knowledge-based scoring function for the assessment of protein–ligand complexes. *J. Chem. Inf. Model.* **2011**, *51* (10), 2731–2745.
- (56) Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-based scoring function to predict protein–ligand interactions. *J. Mol. Biol.* **2000**, *295* (2), 337–356.
- (57) Mooij, W. T.; Verdonk, M. L. General and targeted statistical potentials for protein–ligand interactions. *Proteins* **2005**, *61* (2), 272–287.
- (58) Muegge, I.; Martin, Y. C. A general and fast scoring function for protein–ligand interactions: a simplified potential approach. *J. Med. Chem.* **1999**, *42* (5), 791–804.
- (59) Gunther, J.; Bergner, A.; Hendlich, M.; Klebe, G. Utilising structural knowledge in drug design strategies: applications using Relibase. *J. Mol. Biol.* **2003**, *326* (2), 621–636.
- (60) Hendlich, M.; Bergner, A.; Gunther, J.; Klebe, G. Relibase: design and development of a database for comprehensive analysis of protein-ligand interactions. *J. Mol. Biol.* **2003**, *326* (2), 607–620.
- (61) Wang, L.; Xie, Z.; Wipf, P.; Xie, X. Q. Residue preference mapping of ligand fragments in the Protein Data Bank. *J. Chem. Inf. Model.* **2011**, *51* (4), 807–815.
- (62) Kortagere, S.; Ekins, S.; Welsh, W. J. Halogenated ligands and their interactions with amino acids: implications for structure–activity and structure–toxicity relationships. *J. Mol. Graphics Modell.* **2008**, *27* (2), 170–177.
- (63) Lu, Y.; Wang, Y.; Xu, Z.; Yan, X.; Luo, X.; Jiang, H.; Zhu, W. C–X···H contacts in biomolecular systems: how they contribute to protein–ligand binding affinity. *J. Phys. Chem. B* **2009**, *113* (37), 12615–12621.
- (64) Katz, B. A.; Sprengeler, P. A.; Luong, C.; Verner, E.; Elrod, K.; Kirtley, M.; Janc, J.; Spencer, J. R.; Breitenbucher, J. G.; Hui, H.; McGee, D.; Allen, D.; Martelli, A.; Mackman, R. L. Engineering inhibitors highly selective for the S1 sites of Ser190 trypsin-like serine protease drug targets. *Chem. Biol.* **2001**, *8* (11), 1107–1121.
- (65) Warmus, J. S.; Flamme, C.; Zhang, L. Y.; Barrett, S.; Bridges, A.; Chen, H.; Gowan, R.; Kaufman, M.; Sebolt-Leopold, J.; Leopold, W.; Merriman, R.; Ohren, J.; Pavlovsky, A.; Przybranowski, S.; Tecle, H.; Valik, H.; Whitehead, C.; Zhang, E. 2-Alkylamino- and alkoxy-substituted 2-amino-1,3,4-oxadiazoles-O-alkyl benzohydroxamate esters replacements retain the desired inhibition and selectivity against MEK (MAP ERK kinase). *Bioorg. Med. Chem. Lett.* **2008**, *18* (23), 6171–6174.
- (66) Himmel, D. M.; Das, K.; Clark, A. D., Jr.; Hughes, S. H.; Benjahad, A.; Oumouch, S.; Guillemont, J.; Coupa, S.; Poncelet, A.; Csoka, I.; Meyer, C.; Andries, K.; Nguyen, C. H.; Grierson, D. S.; Arnold, E. Crystal structures for HIV-1 reverse transcriptase in complexes with three pyridinone derivatives: a new class of non-nucleoside inhibitors effective against a broad range of drug-resistant strains. *J. Med. Chem.* **2005**, *48* (24), 7582–7591.
- (67) Jin, R.; Gouaux, E. Probing the function, conformational plasticity, and dimer–dimer contacts of the GluR2 ligand-binding core: studies of 5-substituted willardiines and GluR2 S1S2 in the crystal. *Biochemistry* **2003**, *42* (18), 5201–5213.
- (68) Jin, R.; Banke, T. G.; Mayer, M. L.; Traynelis, S. F.; Gouaux, E. Structural basis for partial agonist action at ionotropic glutamate receptors. *Nat. Neurosci.* **2003**, *6* (8), 803–810.
- (69) Ahmed, A. H.; Wang, S.; Chuang, H. H.; Oswald, R. E. Mechanism of AMPA receptor activation by partial agonists: disulfide trapping of closed lobe conformations. *J. Biol. Chem.* **2011**, *286* (40), 35257–35266.
- (70) Gao, H. Y.; Shen, Q. J.; Zhao, X. R.; Yan, X. Q.; Pang, X.; Jin, W. J. Phosphorescent co-crystal assembled by 1,4-diiodotetrafluorobenzene with carbazole based on C–I halogen bonding. *J. Mater. Chem.* **2012**, *22* (12), 5336–5343.
- (71) Zhang, Y.; Ji, B.; Tian, A.; Wang, W. Competition between  $\pi$ – $\pi$  interaction and halogen bond in solution: a combined  $^{13}\text{C}$  NMR and density functional theory study. *J. Chem. Phys.* **2012**, *136* (14), 141101.
- (72) Beck, T.; Krasauskas, A.; Gruene, T.; Sheldrick, G. M. A magic triangle for experimental phasing of macromolecules. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2008**, *64* (Pt 11), 1179–1182.
- (73) Yu, S. B.; Watson, A. D. Metal-based X-ray contrast media. *Chem. Rev.* **1999**, *99* (9), 2353–2378.
- (74) Kumar, S.; Nussinov, R. Close-range electrostatic interactions in proteins. *ChemBioChem* **2002**, *3* (7), 604–617.
- (75) Gilson, M. K.; Zhou, H. X. Calculation of protein–ligand binding affinities. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21–42.
- (76) Trbovic, N.; Cho, J. H.; Abel, R.; Friesner, R. A.; Rance, M.; Palmer, A. G., 3rd. Protein side-chain dynamics and residual conformational entropy. *J. Am. Chem. Soc.* **2009**, *131* (2), 615–622.
- (77) Dolman, N. P.; More, J. C.; Alt, A.; Knauss, J. L.; Pentikainen, O. T.; Glasser, C. R.; Bleakman, D.; Mayer, M. L.; Collingridge, G. L.; Jane, D. E. Synthesis and pharmacological characterization of N3-substituted willardiine derivatives: role of the substituent at the S-position of the uracil ring in the development of highly potent and selective GLU $K_5$  kainate receptor antagonists. *J. Med. Chem.* **2007**, *50* (7), 1558–1570.
- (78) Alushin, G. M.; Jane, D.; Mayer, M. L. Binding site and ligand flexibility revealed by high resolution crystal structures of GluK1 competitive antagonists. *Neuropharmacology* **2011**, *60* (1), 126–134.
- (79) Best, R. B.; Clarke, J.; Karplus, M. The origin of protein sidechain order parameter distributions. *J. Am. Chem. Soc.* **2004**, *126* (25), 7734–7735.

- (80) Lovell, S. C.; Word, J. M.; Richardson, J. S.; Richardson, D. C. The penultimate rotamer library. *Proteins* **2000**, *40* (3), 389–408.
- (81) Engel, J. C.; Doyle, P. S.; Hsieh, I.; McKerrow, J. H. Cysteine protease inhibitors cure an experimental *Trypanosoma cruzi* infection. *J. Exp. Med.* **1998**, *188* (4), 725–734.
- (82) Ferreira, R. S.; Simeonov, A.; Jadhav, A.; Eidam, O.; Mott, B. T.; Keiser, M. J.; McKerrow, J. H.; Maloney, D. J.; Irwin, J. J.; Shoichet, B. K. Complementarity between a docking and a high-throughput screen in discovering new cruzain inhibitors. *J. Med. Chem.* **2010**, *53* (13), 4891–4905.
- (83) Wei, B. Q.; Baase, W. A.; Weaver, L. H.; Matthews, B. W.; Shoichet, B. K. A model binding site for testing scoring functions in molecular docking. *J. Mol. Biol.* **2002**, *322* (2), 339–355.
- (84) Shoichet, B. K.; Leach, A. R.; Kuntz, I. D. Ligand solvation in molecular docking. *Proteins* **1999**, *34* (1), 4–16.
- (85) Meng, E. C.; Shoichet, B. K.; Kuntz, I. D. Automated docking with grid-based energy evaluation. *J. Comput. Chem.* **1992**, *13* (4), 505–524.
- (86) Njoroge, F. G.; Vibulbhau, B.; Pinto, P.; Strickland, C.; Bishop, W. R.; Nomeir, A.; Girijavallabhan, V. Enhanced FTase activity achieved via piperazine interaction with catalytic zinc. *Bioorg. Med. Chem. Lett.* **2006**, *16* (4), 984–988.
- (87) Ho, B. K.; Curmi, P. M. G. Twist and shear in beta-sheets and beta-ribbons. *J. Mol. Biol.* **2002**, *317* (2), 291–308.
- (88) Neslonay, C. L.; Kelly, J. W. Progress towards understanding beta-sheet structure. *Bioorg. Med. Chem.* **1996**, *4* (6), 739–766.
- (89) Weatherford, D. W.; Salemme, F. R. Conformations of twisted parallel beta-sheets and the origin of chirality in protein structures. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76* (1), 19–23.
- (90) Eneqvist, T.; Lundberg, E.; Karlsson, A.; Huang, S.; Santos, C. R.; Power, D. M.; Sauer-Eriksson, A. E. High resolution crystal structures of piscine transthyretin reveal different binding modes for triiodothyronine and thyroxine. *J. Biol. Chem.* **2004**, *279* (25), 26411–26416.
- (91) Sudlow, G.; Birkett, D. J.; Wade, D. N. The characterization of two specific drug binding sites on human serum albumin. *Mol. Pharmacol.* **1975**, *11* (6), 824–832.
- (92) Ghuman, J.; Zunszain, P. A.; Pettipas, I.; Bhattacharya, A. A.; Otagiri, M.; Curry, S. Structural basis of the drug-binding specificity of human serum albumin. *J. Mol. Biol.* **2005**, *353* (1), 38–52.
- (93) Rohde, L. A. H.; Ahring, P. K.; Jensen, M. L.; Nielsen, E. Ø.; Peters, D.; Helgstrand, C.; Krintel, C.; Harpsøe, K.; Gajhede, M.; Kastrup, J. S.; Balle, T. Intersubunit bridge formation governs agonist efficacy at nicotinic acetylcholine  $\alpha 4\beta 2$  receptors. *J. Biol. Chem.* **2012**, *287* (6), 4248–4259.
- (94) Lemke, C. T.; Goudreau, N.; Zhao, S.; Hucke, O.; Thibeault, D.; Llinás-Brunet, M.; White, P. W. Combined X-ray, NMR, and kinetic analyses reveal uncommon binding characteristics of the hepatitis C virus NS3-NS4A protease inhibitor BI 201335. *J. Biol. Chem.* **2011**, *286* (13), 11434–11443.
- (95) Benjahad, A.; Guillemont, J.; Andries, K.; Nguyen, C. H.; Grierson, D. S. 3-Iodo-4-phenoxypridinones (IOPY's), a new family of highly potent non-nucleoside inhibitors of HIV-1 reverse transcriptase. *Bioorg. Med. Chem. Lett.* **2003**, *13* (24), 4309–4312.
- (96) Bollini, M.; Domaal, R. A.; Thakur, V. V.; Gallardo-Macias, R.; Spasov, K. A.; Anderson, K. S.; Jorgensen, W. L. Computationally-guided optimization of a docking hit to yield catechol diethers as potent anti-HIV agents. *J. Med. Chem.* **2011**, *54* (24), 8582–8591.
- (97) Baumli, S.; Endicott, J. A.; Johnson, L. N. Halogen bonds form the basis for selective P-TEFb inhibition by DRB. *Chem. Biol.* **2010**, *17* (9), 931–936.
- (98) Huber, K.; Brault, L.; Fedorov, O.; Gasser, C.; Filippakopoulos, P.; Bullock, A. N.; Fabbro, D.; Trappe, J.; Schwaller, J.; Knapp, S.; Bracher, F. 7,8-Dichloro-1-oxo- $\beta$ -carbolines as a versatile scaffold for the development of potent and selective kinase inhibitors with unusual binding modes. *J. Med. Chem.* **2011**, *55* (1), 403–413.
- (99) Popowicz, G. M.; Domling, A.; Holak, T. A. The structure-based design of Mdm2/Mdmx-p53 inhibitors gets serious. *Angew. Chem., Int. Ed.* **2011**, *50* (12), 2680–2688.
- (100) Vogel, S. M.; Bauer, M. R.; Joerger, A. C.; Wilcken, R.; Brandt, T.; Veprintsev, D. B.; Rutherford, T. J.; Fersht, A. R.; Boeckler, F. M. Lithocholic acid is an endogenous inhibitor of MDM4 and MDM2. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (42), 16906–16910.
- (101) Graves, B.; Thompson, T.; Xia, M.; Janson, C.; Lukacs, C.; Deo, D.; Di Lello, P.; Fry, D.; Garvie, C.; Huang, K. S.; Gao, L.; Tovar, C.; Lovey, A.; Wanner, J.; Vassilev, L. T. Activation of the p53 pathway by small-molecule-induced MDM2 and MDMX dimerization. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (29), 11788–11793.
- (102) Grasberger, B. L.; Lu, T.; Schubert, C.; Parks, D. J.; Carver, T. E.; Koblish, H. K.; Cummings, M. D.; LaFrance, L. V.; Milkiewicz, K. L.; Calvo, R. R.; Maguire, D.; Lattanzio, J.; Franks, C. F.; Zhao, S.; Ramachandren, K.; Bylebyl, G. R.; Zhang, M.; Manthey, C. L.; Petrella, E. C.; Pantoliano, M. W.; Deckman, I. C.; Spurlino, J. C.; Maroney, A. C.; Tomczuk, B. E.; Molloy, C. J.; Bone, R. F. Discovery and cocrystal structure of benzodiazepinedione HDM2 antagonists that activate p53 in cells. *J. Med. Chem.* **2005**, *48* (4), 909–912.
- (103) Parks, D. J.; Lafrance, L. V.; Calvo, R. R.; Milkiewicz, K. L.; Gupta, V.; Lattanzio, J.; Ramachandren, K.; Carver, T. E.; Petrella, E. C.; Cummings, M. D.; Maguire, D.; Grasberger, B. L.; Lu, T. 1,4-Benzodiazepine-2,5-diones as small molecule antagonists of the HDM2-p53 interaction: discovery and SAR. *Bioorg. Med. Chem. Lett.* **2005**, *15* (3), 765–770.
- (104) Kantsaris, A. L.; Hayes, J. M.; Manta, S.; Skamnaki, V. T.; Kiritsis, C.; Psarra, A.-M. G.; Koutsogiannis, Z.; Dimopoulou, A.; Theofanous, S.; Nikoleousakos, N.; Zoumpoulakis, P.; Kontou, M.; Papadopoulos, G.; Zographos, S. E.; Komiotis, D.; Leonidas, D. D. The  $\sigma$ -hole phenomenon of halogen atoms forms the structural basis of the strong inhibitory potency of C5 halogen substituted glucopyranosyl nucleosides towards glycogen phosphorylase b. *ChemMedChem* **2012**, *7* (4), 722–732.
- (105) Xu, Z.; Liu, Z.; Chen, T.; Chen, T.; Wang, Z.; Tian, G.; Shi, J.; Wang, X.; Lu, Y.; Yan, X.; Wang, G.; Jiang, H.; Chen, K.; Wang, S.; Xu, Y.; Shen, J.; Zhu, W. Utilization of halogen bond in lead optimization: a case study of rational design of potent phosphodiesterase type 5 (PDE5) inhibitors. *J. Med. Chem.* **2011**, *54* (15), 5607–5611.
- (106) Maillard, M. C.; Hom, R. K.; Benson, T. E.; Moon, J. B.; Mamo, S.; Bienkowski, M.; Tomasselli, A. G.; Woods, D. D.; Prince, D. B.; Paddock, D. J.; Emmons, T. L.; Tucker, J. A.; Dappen, M. S.; Brogley, L.; Thorsett, E. D.; Jewett, N.; Sinha, S.; John, V. Design, synthesis, and crystal structure of hydroxyethyl secondary amine-based peptidomimetic inhibitors of human beta-secretase. *J. Med. Chem.* **2007**, *50* (4), 776–781.
- (107) Joerger, A. C.; Ang, H. C.; Fersht, A. R. Structural basis for understanding oncogenic p53 mutations and designing rescue drugs. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (41), 15056–15061.
- (108) Boeckler, F. M.; Joerger, A. C.; Jaggi, G.; Rutherford, T. J.; Veprintsev, D. B.; Fersht, A. R. Targeted rescue of a destabilized mutant of p53 by an *in silico* screened drug. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (30), 10360–10365.
- (109) Wilcken, R.; Wang, G.; Boeckler, F. M.; Fersht, A. R. Kinetic mechanism of p53 oncogenic mutant aggregation and its inhibition. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (34), 13584–13589.
- (110) Lu, Y.-X.; Zou, J.-W.; Fan, J.-C.; Zhao, W.-N.; Jiang, Y.-J.; Yu, Q.-S. Ab initio calculations on halogen-bonded complexes and comparison with density functional methods. *J. Comput. Chem.* **2009**, *30* (5), 725–732.
- (111) Grimme, S. Density functional theory with London dispersion corrections. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1* (2), 211–228.
- (112) Zimmermann, M.; Lange, A.; Boeckler, F. M. Unpublished results.
- (113) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. The development and use of quantum-mechanical molecular-models. 76. Am1—a new general-purpose quantum-mechanical molecular-model. *J. Am. Chem. Soc.* **1985**, *107* (13), 3902–3909.
- (114) Stewart, J. J. P. Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* **2007**, *13* (12), 1173–1213.

- (115) Ibrahim, M. A. A. Performance assessment of semiempirical molecular orbital methods in describing halogen bonding: quantum mechanical and quantum mechanical/molecular mechanical-molecular dynamics study. *J. Chem. Inf. Model.* **2011**, *51* (10), 2549–2559.
- (116) Řezáč, J.; Hobza, P. A halogen-bonding correction for the semiempirical PM6 method. *Chem. Phys. Lett.* **2011**, *506* (4–6), 286–289.
- (117) Halgren, T. A. Merck molecular force field. 1. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, *17* (5–6), 490–519.
- (118) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174.
- (119) Wang, J. M.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell.* **2006**, *25* (2), 247–260.
- (120) Rendine, S.; Pieraccini, S.; Forni, A.; Sironi, M. Halogen bonding in ligand–receptor systems in the framework of classical force fields. *Phys. Chem. Chem. Phys.* **2011**, *13* (43), 19508–19516.
- (121) Ibrahim, M. A. A. Molecular mechanical study of halogen bonding in drug discovery. *J. Comput. Chem.* **2011**, *32* (12), 2564–2574.
- (122) Ibrahim, M. A. A. AMBER empirical potential describes the geometry and energy of noncovalent halogen interactions better than advanced semiempirical quantum mechanical method PM6-DH2X. *J. Phys. Chem. B* **2012**, *116* (11), 3659–3669.
- (123) Carter, M.; Rappe, A. K.; Ho, P. S. Scalable anisotropic shape and electrostatic models for biological bromine halogen bonds. *J. Chem. Theory Comput.* **2012**, *8* (7), 2461–2473.
- (124) Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **2006**, *49* (20), 5912–5931.
- (125) Jones, G.; Willett, P.; Glen, R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995**, *245* (1), 43–53.
- (126) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267* (3), 727–748.
- (127) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions. I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided. Mol. Des.* **1997**, *11* (5), 425–445.
- (128) Murray, C. W.; Auton, T. R.; Eldridge, M. D. Empirical scoring functions. II. The testing of an empirical scoring function for the prediction of ligand–receptor binding affinities and the use of Bayesian regression to improve the quality of the model. *J. Comput.-Aided. Mol. Des.* **1998**, *12* (5), 503–519.
- (129) Kuhn, B.; Fuchs, J. E.; Reutlinger, M.; Stahl, M.; Taylor, N. R. Rationalizing tight ligand binding through cooperative interaction networks. *J. Chem. Inf. Model.* **2011**, *51* (12), 3180–3198.
- (130) Paulini, R.; Muller, K.; Diederich, F. Orthogonal multipolar interactions in structural chemistry and biology. *Angew. Chem., Int. Ed.* **2005**, *44* (12), 1788–1805.
- (131) Muley, L.; Baum, B.; Smolinski, M.; Freindorf, M.; Heine, A.; Klebe, G.; Hangauer, D. G. Enhancement of hydrophobic interactions and hydrogen bond strength by cooperativity: synthesis, modeling, and molecular dynamics simulations of a congeneric series of thrombin inhibitors. *J. Med. Chem.* **2010**, *53* (5), 2126–2135.
- (132) Baum, B.; Muley, L.; Smolinski, M.; Heine, A.; Hangauer, D.; Klebe, G. Non-additivity of functional group contributions in protein ligand binding: a comprehensive study by crystallography and isothermal titration calorimetry. *J. Mol. Biol.* **2010**, *397* (4), 1042–1054.
- (133) Gomilsek, J. P.; Arcon, I.; de Panfilis, S.; Kodre, A. X-ray absorption in atomic iodine in the K-edge region. *Phys. Rev. A* **2009**, *79* (3), 032514.
- (134) Hamuy, R.; Berman, B. Treatment of herpes simplex virus infections with topical antiviral agents. *Eur. J. Dermatol.* **1998**, *8* (5), 310–319.
- (135) Kudenchuk, P. J.; Cobb, L. A.; Copass, M. K.; Cummins, R. O.; Doherty, A. M.; Fahrenbruch, C. E.; Hallstrom, A. P.; Murray, W. A.; Olsufka, M.; Walsh, T. Amiodarone for resuscitation after out-of-hospital cardiac arrest due to ventricular fibrillation. *N. Engl. J. Med.* **1999**, *341* (12), 871–878.
- (136) Rosenbaum, M. B.; Chiale, P. A.; Haedo, A.; Lazzari, J. O.; Elizari, M. V. Ten years of experience with amiodarone. *Am. Heart J.* **1983**, *106* (4, Part 2), 957–964.
- (137) Slagt, V. F.; de Vries, A. H. M.; de Vries, J. G.; Kellogg, R. M. Practical aspects of carbon–carbon cross-coupling reactions using heteroarenes. *Org. Process Res. Dev.* **2010**, *14* (1), 30–47.
- (138) Busacca, C. A.; Fandrick, D. R.; Song, J. J.; Senanayake, C. H. The growing impact of catalysis in the pharmaceutical industry. *Adv. Synth. Catal.* **2011**, *353* (11–12), 1825–1864.
- (139) Djakovitch, L.; Batail, N.; Genelot, M. Recent advances in the synthesis of N-containing heteroaromatics via heterogeneously transition metal catalysed cross-coupling reactions. *Molecules* **2011**, *16* (6), 5241–5267.
- (140) Sasson, Y. Formation of Carbon–Halogen Bonds (Cl, Br, I). In *The Chemistry of Halides, Pseudo-Halides and Azides; The Chemistry of Function Groups, Supplement D2*; John Wiley & Sons, Ltd.: Chichester, U.K., 2004; pp 535–628.
- (141) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **2001**, *46* (1–3), 3–26.
- (142) Hann, M. M. Molecular obesity, potency and other addictions in drug discovery. *MedChemComm* **2011**, *2* (5), 349–355.
- (143) Hughes, J. D.; Blagg, J.; Price, D. A.; Bailey, S.; Decrescenzo, G. A.; Devraj, R. V.; Ellsworth, E.; Fobian, Y. M.; Gibbs, M. E.; Gilles, R. W.; Greene, N.; Huang, E.; Krieger-Burke, T.; Loesel, J.; Wager, T.; Whiteley, L.; Zhang, Y. Physiochemical drug properties associated with in vivo toxicological outcomes. *Bioorg. Med. Chem. Lett.* **2008**, *18* (17), 4872–4875.
- (144) Gleeson, M. P. Generation of a set of simple, interpretable ADMET rules of thumb. *J. Med. Chem.* **2008**, *51* (4), 817–834.
- (145) Waring, M. J. Defining optimum lipophilicity and molecular weight ranges for drug candidates—Molecular weight dependent lower logD limits based on permeability. *Bioorg. Med. Chem. Lett.* **2009**, *19* (10), 2844–2851.
- (146) Keseru, G. M.; Makara, G. M. The influence of lead discovery strategies on the properties of drug candidates. *Nat. Rev. Drug Discovery* **2009**, *8* (3), 203–212.
- (147) Leeson, P. D.; Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discovery* **2007**, *6* (11), 881–890.
- (148) Zhang, C.; Kenny, J. R.; Le, H.; Deese, A.; Ford, K. A.; Lightning, L. K.; Fan, P. W.; Driscoll, J. P.; Halladay, J. S.; Hop, C. E.; Khojasteh, S. C. Novel mechanism for dehalogenation and glutathione conjugation of dihalogenated anilines in human liver microsomes: evidence for ipso glutathione addition. *Chem. Res. Toxicol.* **2011**, *24* (10), 1668–1677.
- (149) Manna, D.; Muges, G. Regioselective deiodination of thyroxine by iodothyronine deiodinase mimics: an unusual mechanistic pathway involving cooperative chalcogen and halogen bonding. *J. Am. Chem. Soc.* **2012**, *134* (9), 4269–4279.
- (150) Nascimento, A. S.; Dias, S. M.; Nunes, F. M.; Aparicio, R.; Ambrosio, A. L. B.; Bleicher, L.; Figueira, A. C. M.; Santos, M. A. M.; Neto, M. D.; Fischer, H.; Togashi, M.; Craievich, A. F.; Garratt, R. C.; Baxter, J. D.; Webb, P.; Polikarpov, I. Structural rearrangements in the thyroid hormone receptor hinge domain and their putative role in the receptor function. *J. Mol. Biol.* **2006**, *360* (3), 586–598.
- (151) Bayse, C. A.; Rafferty, E. R. Is halogen bonding the basis for iodothyronine deiodinase activity? *Inorg. Chem.* **2010**, *49* (12), 5365–5367.

- (152) Valverde, C.; Orozco, A.; Becerra, A.; Jeziorski, M. C.; Villalobos, P.; Solis, J. C. Halometabolites and cellular dehalogenase systems: an evolutionary perspective. *Int. Rev. Cytol.* **2004**, *234*, 143–199.
- (153) Isin, E. M.; Guengerich, F. P. Complex reactions catalyzed by cytochrome P450 enzymes. *Biochim. Biophys. Acta* **2007**, *1770* (3), 314–329.
- (154) Nademanee, K.; Piwonka, R. W.; Singh, B. N.; Hershman, J. M. Amiodarone and thyroid function. *Prog. Cardiovasc. Dis.* **1989**, *31* (6), 427–437.
- (155) Lam, P. Y. S.; Clark, C. G.; Smallwood, A. M.; Alexander, R. S. Structure-Based Drug Design Utilizing Halogen Bonding: Factor Xa Inhibitors. *Abstracts of Papers*, 238th National Meeting of the American Chemical Society; American Chemical Society: Washington, DC, U.S., 2009.
- (156) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **2001**, *46* (1–3), 3–26.
- (157) Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D. Is there a difference between leads and drugs? A historical perspective. *J. Chem. Inf. Comput. Sci.* **2001**, *41* (5), 1308–1315.