JOURNAL OF CHEMICAL INFORMATION AND MODELING

Perspective

pubs.acs.org/jcim

Halogen Interactions in Protein−Ligand Complexes: Implications of Halogen Bonding for Rational Drug Design

Suman Sirimulla,*[,†] Jake B. Bailey,[†] Rahulsimham Vegesna,[‡] and Mahesh Narayan*[,§]

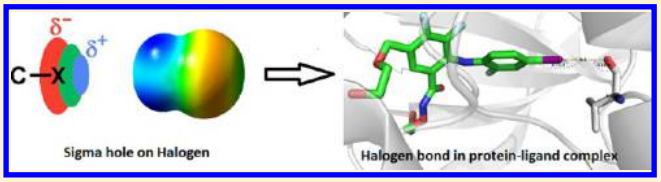
[†]Department of Chemistry and Biochemistry, Northern Arizona University, P.O. Box 5698, Flagstaff, Arizona 86011-5698, United States

[‡]Department of Bioinformatics and Computational Biology, UT MD Anderson Cancer Center, Houston, Texas 77030, United States

[§]Department of Chemistry, University of Texas at El Paso, 500 W. University Ave., El Paso, Texas 79968, United States

Supporting Information

ABSTRACT: Halogen bonding interactions between halogenated ligands and proteins were examined using the crystal structures deposited to date in the PDB. The data was analyzed as a function of halogen bonding to main chain Lewis bases, viz. oxygen of backbone carbonyl and backbone amide nitrogen. This analysis also examined halogen bonding to side-chain Lewis bases (O, N, and S) and to the electron-rich aromatic amino acids. All interactions were restricted to van der Waals radii with respective atoms. The data reveals that while fluorine and chlorine have strong tendencies favoring interactions with the backbone Lewis bases at glycine, the trend is not restricted to the achiral amino acid backbone for larger halogens. Halogen side-chain interactions are not restricted to amino acids containing O, N, and S as Lewis bases. Electron-rich aromatic amino acids host a high frequency of halogen bonds as does Leu. A closer examination of the latter hydrophobic side chain reveals that the "propensity of interactions" of halogen ligands at this oily residue is an outcome of strong classical halogen bonds with Lewis bases in the vicinity. Finally, an examination of Θ_1 (C−X···O and C−X···N) and Θ_2 (X···O−Z and X···N−Z) angles reveals that very few ligands adopt classical halogen bonding angles, suggesting that steric and other factors may influence these angles. The data is discussed in the context of ligand design for pharmaceutical applications.



INTRODUCTION

Halogen Interactions within Protein−Ligand Complexes. The halogen bond is an important noncovalent interaction that is currently receiving increased attention in the study of protein−ligand complexes.[1−6] Many drugs use halogenated molecules to establish halogen bonds with biomolecules.[7−9] Halogen atoms, commonly fluorine and chlorine, have been used for several years for hit-to-lead or lead-to-drug conversions.[3,8,10] It is estimated that one-quarter of the total number of papers and patents closely related to medicinal chemistry involve the insertion of halogens during the synthesis of final compounds.[3] The heavier halogens, bromine and iodine, have been under-represented in drug design. A halogen bond in biomolecules can be referred to a short C−X···D−Z interaction, where D is a halogen bond donor, C−X is a carbon-bonded chlorine, bromine, or iodine; D−Z is a carbonyl, hydroxyl, thiol, aromatic ring, charged carboxylate, phosphate group, or amine, and the X···D distance is less than or equal to the sums of the respective van der Waals radii.[11,12] The sum of van der Waals radii of each halogen with N, O, or S is shown in Table 1. A recent systemic investigation of halogen bonds in protein−ligand complexes by Hardegger et al. showed that halogen bonds can serve as a powerful tool in increasing binding selectivity and binding affinity.[12] A few examples of halogen bonding impact in protein−ligand complexes include the following. (A) Incorporation of a halogen atom in the *ortho* position of the 1-phenyl ring of CDPPB [3-cyano-N-(1,3-diphenyl-1H-pyrazol-5-yl)-benzamide] leads to an increase in both the binding and functional activities of mGluR.[13] (B) Iodine in 4-(3,5-dimethylphenoxy)-5-(furan-2-ylmethylsulfanylmethyl)-3-iodo-6-methylpyridin-2(1*H*)-one (R221239) increases the binding affinity to wild type HIV-1 reverse transcriptase through the formation of a halogen bond with the carbonyl oxygen of TYR188. This interaction is regarded as one of the important interactions leading to the overall protein−ligand binding affinity.[2,14] (C) Dichloroindolyl enaminonitrile (KH−CB19) is a potent inhibitor CLK1 and CLK 4. The Cl···O interaction between a chlorine on the dichloroindolyl enaminonitrile (KH−CB19) and the carbonyl group of Glu242 is recognized

Table 1. Sum of van der Waals Radii between Halogen and Halogen Bond Acceptors in Å

halogen	nitrogen	oxygen	sulfur
fluorine	3.02	2.99	3.27
chlorine	3.30	3.27	3.55
bromine	3.40	3.37	3.65
iodine	3.53	3.50	3.78

Received: April 29, 2013
Published: October 17, 2013

as a key mediator of the hinge interaction of KH−CB19.[15] (D) The incorporation of chlorine into an anti-HIV-1 catechol derivative increased inhibition of HIV-1 reverse transcriptase through the formation of a halogen bond with the carbonyl oxygen of Pro95.[5,16] (E) The incorporation of halogens into 6-ethyl-2-[2-propoxy-5-(4-methyl-1-piperazinylsulfonyl)-phenyl]-primidin-4(3H)-one analogs have been shown to inhibit phosphodiesterase type 5 (PDE5) through the formation of a halogen bond with Tyr612.[17] (F) Halogen binding interactions have been shown to increase the effectiveness of serine protease factor Xa inhibitors through the formation of a halogen bond with the π-system of the aromatic ring in Tyr288.[18] (G) The formation of a halogen bond between iodine containing ligands and the Leu145 residue of the tumor suppressor p53 mutant Y220C is regarded as a key interaction in facilitating ligand binding.[19]

Chemistry of Halogen Bonding. Halogens possess a high effective nuclear charge while maintaining an anisotropic electric potential around the atom, leading to the possibility of halogen bonding. A halogen bond is a directional, electrostatically driven, noncovalent interaction between a region of positive electrostatic potential on the outer side of a halogen, known as a σ-hole, and an electrically negative site (such as a lone pair of a Lewis base or the π-electrons of an unsaturated system).[20,21] Figure 1 shows the electrostatic potential on the halogen bonded to a carbon atom.

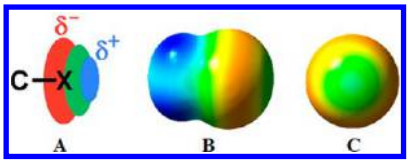


Figure 1. (A) Schematic representation of electron distribution on the halogen. (B) Side view of electrostatic potential map of CH_3I. (C) Front view of electrostatic potential map of CH_3I showing the σ-hole. Images B and C were prepared with Gaussian software.[24]

The IUPAC recently proposed the definition as "halogen bond is a net attractive interaction between an electrophilic region of a halogen atom on a molecule or molecular fragment and a nucleophilic region of another molecule or molecular fragment".[22] The electrophilic region arises on the halogen because when a halogen makes a covalent bond in a molecule, the electron distribution around the atom somewhat shifts that bond. This event leaves an area of diminished electron density opposite the bond—an area defined as the σ-hole.[23]

Halogen Bond Strength and Preferred σ-Hole Angle. The ability to form halogen bonds is well documented for chorine, bromine, and iodine. Fluorine is typically excluded from halogen bond formation due to its high electronegativity and lack of polarizability.[12] However, emerging data suggests that fluorine is capable of forming halogen bonds, albeit weakly and under select circumstances, as a result of the polarization of the electron cloud in fluorine.[25] The strength of the halogen bond interaction follows the general trend: F ≪ Cl < Br < I, with iodine normally forming the strongest interactions.[12,26]

The halogen bond strength depends on several factors: the size of the σ-hole on the halogen, σ-hole interaction angle, donor atom properties, and internuclear distance between halogen and halogen bond acceptor (X···D distance).[27−29] The size of the σ-hole depends on two factors: the polarizability of the halogen and the presence of electron withdrawing atom (groups) near the halogen. Heavier halogens tend to have greater extent of charge separation on the atom leading to a larger σ-hole; the size of the σ-hole on halogens follows the trend I > Br > Cl > F.[27] The comparison of energies of I/Br/Cl···O are tabulated in Table 2.

The presence of electron withdrawing groups near the halogen also corresponds to a larger σ-hole.[22] The ideal halogen interaction would be a "head-on" contact between the σ-hole of a halogen to the halogen bond acceptor with a C−X···D (Θ_1) angle close to 180°. Angle deviations between 25° and 30° correspond to a 50% reduction in halogen bond strength.[31] Angle deviations greater than 40° are contraindications for halogen bond formation. After this level of deviation, the σ-hole becomes inaccessible to the electron donor. This angle may lead to repulsion from the expected donor group due to the high electron density of the halogen. The incorporation of stronger halogen bonding interactions may offset angle deviations in protein−ligand complexes.[9]

The preferred X···D−Z angle (Θ_2) varies depending on the type of halogen bond acceptor atom. Θ_2 for carbonyl oxygen would be favored around 120°. Schematic representation of Θ_1 and Θ_2 are depicted in Figure 2. An example of halogen bonds

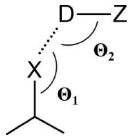


Figure 2. Schematic of halogen interactions to Lewis bases. The image was adapted from Auffinger et al.[11]

found in proteins is depicted in Figure 3. For neutral atoms, the strength of the halogen bond increases as the basicity of the donor atom increases.[28] Carbonyl groups also form stronger halogen bonds than sp^3 oxygens due to the increased electron density in a π bond.[30] As expected, the strength of the halogen bond increases as the C−X···D distance decreases. This trend

Table 2. Comparison of Halogen Bonding Energies

complex[a]	ΔE (kJ/mol)	method	Θ_1 angle (deg)	Θ_2 angle (deg)	source
PhI···BB	−14.2	MP2/TZVPP	175.6	118.7	Wilchen et al.[29]
PhBr···BB	−9.0	MP2/TZVPP	177.4	116.7	Wilchen et al.[29]
PhCl···BB	−5.6	MP2/TZVPP	171.2	106.7	Wilchen et al.[29]
PhBr···OCH_2	−2.98	MP2/aug-cc-pVDZ	168.9	X	Lu et al.[30]
PhBr···NH_3	−2.75	MP2/aug-cc-pVDZ	179.7	X	Lu et al.[30]
PhBr···SH_2	−2.12	MP2/aug-cc-pVDZ	165.1	X	Lu et al.[30]
PhBr···OH_2	−1.97	MP2/aug-cc-pVDZ	178.0	X	Lu et al.[30]

[a]BB = N-acetylamide.

dx.doi.org/10.1021/ci400257k | J. Chem. Inf. Model. 2013, 53, 2781−2791

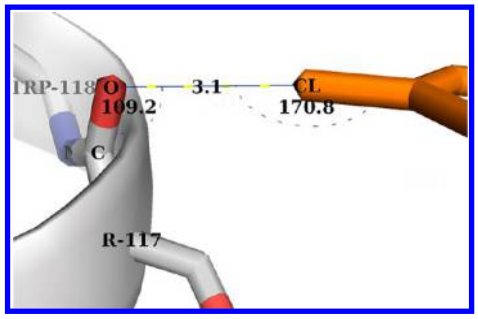


Figure 3. Illustration of a halogen bond in a protein−ligand complex. Chlorine forms a halogen bond with main chain oxygen of the protein. The ligand is shown as an orange stick model and protein is shown as both a cartoon and stick model (pdb id: 2b19). The image was prepared using PYMOL.[32]

continues until the nuclear repulsion becomes too great, destabilizing the structure.

This paper focuses on halogen interactions with amino acids within protein−ligand complexes. The propensities of halogen interactions with main-chain and side-chain amino acids are explored.

METHODS

The latest PDB release (October 2012) was obtained from www.pdb.org and queried to separate out files containing halogenated ligands using a custom-made script. The script was written to separate files that contain at least one C−X bond and have a resolution of 3 Å or better. This screening process yielded unique 4001 PDB entries. Another custom-made script, sigmaholeangle.py (Supporting Information), was written to map C−X···D interactions in the PDB files. This resulted in 2471 total C−X···D interactions (Θ_1) within the specified van der Waals radii distances (Table 1). The number of interactions between each halogen to nitrogen, oxygen, and sulfur within their respective van der Waals distances is summarized in Table 3. The HBPLUS[33] neighbor module (−N) was used to determine the Θ_2 angles between proteins and halogenated ligands.

Table 3. Number of Interactions between Halogen and N/O/S in PDB Files

halogen	nitrogen	oxygen	sulfur
fluorine	433	630	70
chlorine	248	518	90
bromine	33	218	14
iodine	50	149	18

RESULTS AND DISCUSSION

The data obtained from the PDB database was analyzed in several phases. Initially, halogen interaction propensities with amino acids at a molecular level were examined. The data was split into halogen interaction propensities to amino acid main chain and side chain. Halogen main-chain interactions are abbreviated as HM, and halogen side-chain interactions are abbreviated as HS. Contacts of halogens to oxygen, nitrogen, sulfur, and aromatic carbons within van der Waals distances will be discussed.

Figure 4 shows the interactions between the halogen atom X, where X is F, Cl, Br, or I, of halogenated ligands and a backbone carbonyl oxygen or amide nitrogen of the protein as a function of the amino acid main chain (C−Cα−N) that is pendant off of the backbone at the observed interaction. The data have been normalized for each halogen.

A survey of the frequency map (Figure 4) clearly indicates that ligands possessing F, Cl, and Br atoms have the highest propensity of interactions at the N−Cα−C Lewis donors of glycine. This constitutes almost 25% of the total HM interactions. N−Cα−C at Ala interactions (∼10%) are second for F and Cl, where N−Cα−C at Leu is the second most dominant for Br (∼12%). In contrast to this trend, the halogen in iodinated ligands shows the largest preference for the N−Cα−C backbone at Ala (∼15%), followed by the backbone at Leu and then at Gly (∼13%, each).

The high frequency of fluoro, chloro, and bromo interactions with the N−Cα−C of Gly is expected given the lack of steric hindrance due to the side chain. Statistically, even though the accessibility for N−Cα−C at Gly can be expected to be more than that for large-side-chain-containing amino acids due to accessibility from both sides, the frequency of interactions at a particular amino acid is theoretically 5% without correcting for the empirical frequency of occurrence of amino acids in proteins. The interaction of all four halogenated ligands at the N−Cα−C of Ala is also high in frequency perhaps due to its greater accessibility to this short side-chain-containing amino acid.

As observed above, the propensity of iodine in iodinated ligands to interact with N−Cα−C backbone is different from the other three halogens (Figure 4). An examination of iodine (from ligand PF-04622664) N−Cα−C of Val 127 in protein MEK1 (PDB ID: 3dv3) is shown in Figure 5A. A typical halogen bond between iodine and the carbonyl oxygen of Val 127 is clearly observed. The C−I···O angle (Θ_1) of 175.3° suggests a strong interaction. This is also supported by a C−O···I angle (Θ_2) of 126.9° (see below for discussions of bond angles of halo-protein interactions).

Shown in Figure 5B is an example of a classical halogen bond interaction between Br of the brominated ligand and the carbonyl backbone of Ser 565. A very favorable C−Br···O angle

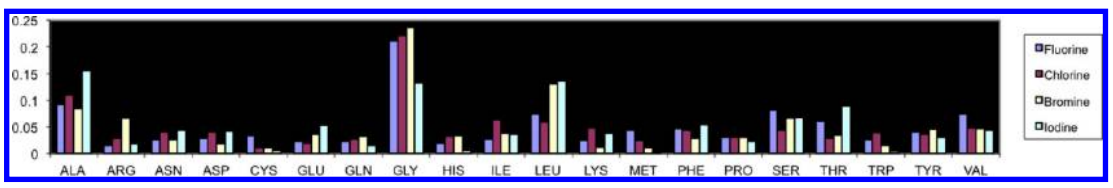


Figure 4. Frequency of interactions of halogenated ligands with proteins. The interactions are defined using van der Waals radii between halogens of the ligand and backbone carbonyl or backbone nitrogen of the protein as Lewis bases. The interaction is labeled as being as above at the N−Cα−C of the side chain that is the pendant off of the backbone.The data is normalized to the maximum number of interactions made by each halogenated ligand.

dx.doi.org/10.1021/ci400257k | J. Chem. Inf. Model. 2013, 53, 2781−2791

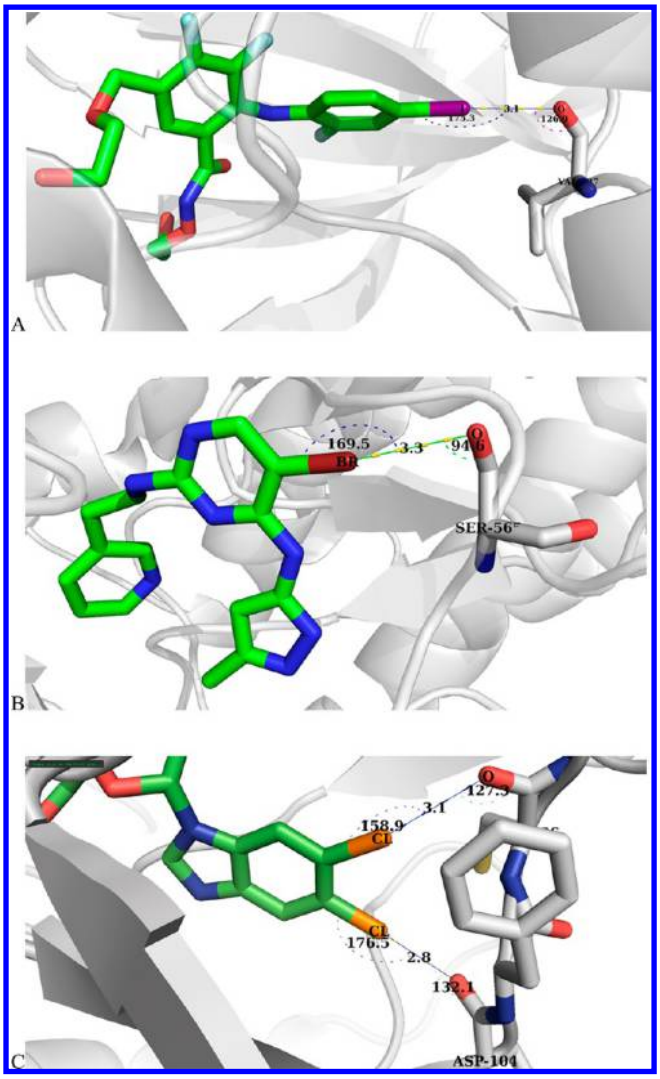


Figure 5. Examples of halogen bonds in protein−ligand complexes. (a) Ligand forming I···O interaction with carbonyl oxygen of Val 127 (pdb ID: 3dv3). (b) Ligand forming Br···O interaction with carbonyl oxygen of Ser 565. (c) Ligand forming Cl···O interactions with Cys 106 and Asp 104. The image was prepared using PYMOL.[32]

(Θ_1) of 169.5° is observed along with a C−O···Br angle (Θ_2) of 96.6°.

Figure 5C shows a dichlorinated ligand forming Cl···O interactions with Cys 106 and ASP 104 backbone carbonyls. The C−Cl···O bond angles suggest strong halogen bonding interactions for both chlorines. It appears that these C−Cl···O angles influence the C−O···Cl angles (127.3° and 132.1°) in each case in that they cause deviations from the preferred orientations of 180°.

The interactions of the halo-moieties of the ligands with the N−Cα−C are reproduced in Figure 6 for further clarity and analysis.

As remarked previously, fluorinated ligands clearly favor interactions with N−Cα−C of Gly (blue). The tendency to interact with the backbone of Ala is followed by the the Ser, Val, and Leu backbone N−Cα−C interactions.

An examination of chlorine N−Cα−C interactions (magenta) in the PDB reveals a trend very similar to fluorine (Figure 6). The maxima at N−Cα−C of Gly is well pronounced and followed by a tendency to favor N−Cα−C interactions with Ala. The rest of the N−Cα−C to fluorine interactions with amino acids does not seem to have any striking tendency in terms of favored frequency.

Bromine (yellow) also demonstrates the largest frequency of interactions at N−Cα−C corresponding to the Gly side chain. However, a distinct favored frequency of interaction is observed at N−Cα−C of Leu. Beyond this deviation, the trend is similar to the fluorine N−Cα−C interactions.

Iodine N−Cα−C interactions (cyan) clearly differ significantly from the other three halogens. Distinct maximas are seen at Ala, Leu, and Gly (in that order), followed by Thr.

The distribution of the frequency of interaction for halogenated ligands with amino acid side chains is depicted in Figure 7. In comparison with backbone Lewis bases, the interaction of the halo-moieties adopts a different topology. High frequencies of interaction are observed with the Leu side chain for iodine (∼12%) and fluorine (hydrogen bonds, ∼13%), whereas bromine is found to interact significantly with Phe (∼13%). All halogens interact with high propensity with Tyr. The individual frequencies of interactions are discussed below.

In contrast to preferred interactions with the N−Cα−C backbone at Gly, fluorinated ligands (blue) are found to interact with the highest frequency at Leu side chains followed by Phe and Tyr (Figure 8). Strikingly, the tendency to interact with Ala side chains is not pronounced. Both Arg and Val interactions are slightly higher than the rest of the side chains.

The interactions landscape of chlorinated ligands (magenta) with side chains is not as rugged as the fluorine interactions. However, interactions with Tyr side chains are predominantly favored followed by those with Phe. Slightly elevated interactions are also observed with Arg and Leu.

Bromine (yellow) amplifies the observations found above. A highly pronounced frequency of interaction of bromine with Phe is found. Furthermore, as observed with Cl, there are greater frequencies of interactions of Br with the side chains of Tyr, Leu, and Arg.

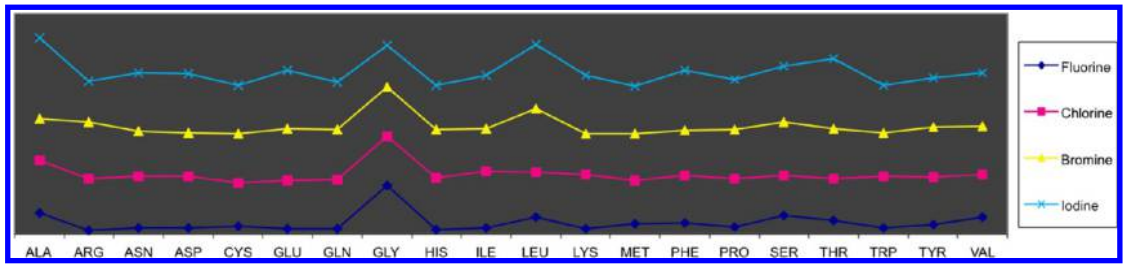


Figure 6. Nature of halogenated ligand backbone interactions is depicted as a function of each halogen and amino acid side-chain pendant off of the backbone.

dx.doi.org/10.1021/ci400257k | J. Chem. Inf. Model. 2013, 53, 2781−2791

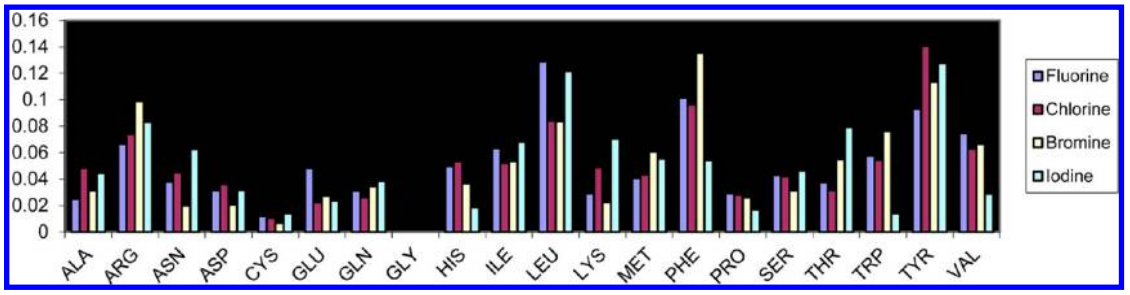


Figure 7. Frequency of interactions of halogenated ligands with proteins. The interactions are defined using van der Waals radii between the halogen of the ligand and side-chain Lewis bases (oxygen, nitrogen, and sulfur). Also included are interactions as above with the π electron rich aromatics residues (Phe, Tyr, Trp, and His). The data are normalized to the maximum number of interactions made by each halogenated ligand.

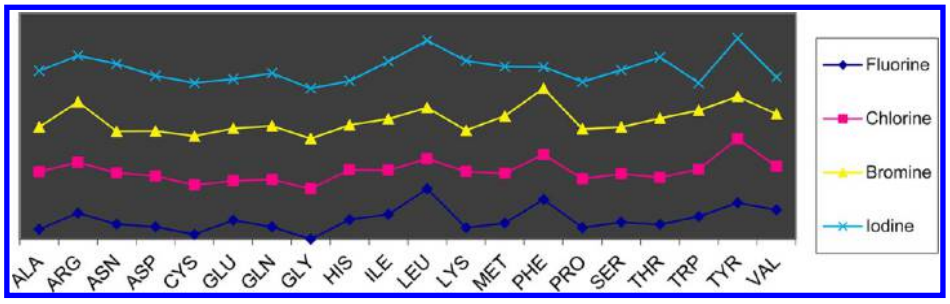


Figure 8. Nature of halogenated ligand side-chain interactions is depicted as a function of each halogen and amino acid side chain.

While the pool of iodinated ligands is small, clear deviations are observed for the interaction of iodine with side chains (cyan). Most iodinated ligands interact with Leu and Tyr side chains followed by Thr and Arg.

Figure 9 shows clear halogen bonding interaction between iodine and the backbone carbonyl of HM carbonyl oxygen of

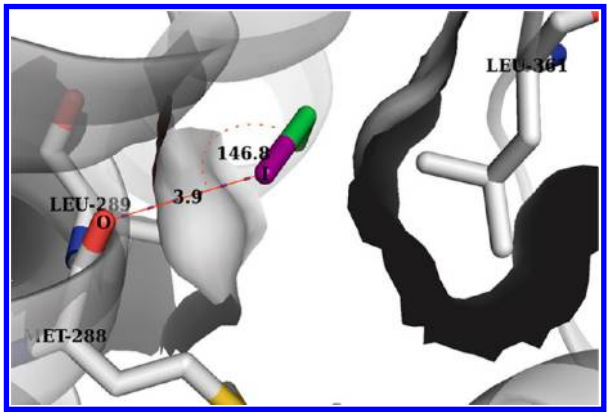


Figure 9. Iodinated ligand interacting with Met and Leu residues. The ligands are shown as green stick models. The protein is shown as a cartoon and stick models The neighboring Leu residues are shown as both white stick and surface models. While Iodine forms a halogen type interaction with the main-chain carbonyl oxygen of Met 288, the nonpolar region of the ligand is interacting with the side chain of Leu (PDB ID: 1FZ9). The image was prepared with PYMOL.[32]

Met 288. Importantly, this interaction positions iodine next to Leu 361. Thus, prima facie, though it may appear that frequency of interaction of iodine with Leu side chains is unusually high, it is probably coincidental as clearly evident through this instance.

Figure 10 shows a classical halogen bonding interaction with a nontraditional Lewis base, the π ring of electron density above the ring plane of Tyr 146 and Phe 169. The frequency of interactions observed between some of the halogens and the aromatic amino acids is thus easily explained.

Halogen Interactions to Amino Acids (Main Chain and Side Chain). Halogen interactions with main-chain (HM) atoms and side-chain (HS) atoms were studied separately. The main chain of protein made of amide linkage, with exposed oxygen and nitrogen, Θ_1 (C−X···O and C−X···N), and Θ_2 (X···O−Z and X···N−Z) angles were investigated. These results are summarized below.

C−X···D Interactions. Across all halogens a larger number of interactions were found. There was a greater number of X···O contacts than X···N interactions suggesting the carbonyl oxygen Lewis base was a better donor than the backbone nitrogen Lewis base. Furthermore, there was a clear increase in the selectivity for oxygen over nitrogen as the halogen radius increased. All the Θ_1 main-chain interactions are summarized in Figure 11.

An examination of C−F···O and C−F···N (Θ_1) indicates that most interactions between fluorine and the Lewis bases lie below the 140° halogen bond cutoff limit.[31] In contrast to fluorine, chlorine-containing ligands exhibit larger frequencies of interactions that qualify within the halogen-bonding angles of 140° and above. Again, there is a greater selectivity for oxygen when compared to nitrogen for making halogen bond interactions. Interestingly, the C−X···O angle (Θ_1) in chlorine can be observed across a wide range of degrees with a robust frequency of occurrence. A maximum is observed around 135° for the C−X···O angle, and a lesser angle of ∼120° serves as the crest for C−X···N.

While the brominated ligand pool is smaller than chlorine and fluorine containing ligands, the C−Br···O interactions overwhelmingly dominate the C−Br···N interactions. Brominated ligands also demonstrate a large number of interactions above the 150° range with the largest frequency at ∼165°.

Iodinated ligands reflect a bimodal distribution of interactions with oxygen. Here, a collection of C−I···O interactions peak at 115°, and a second set of interactions reside in the 150°

dx.doi.org/10.1021/ci400257k | J. Chem. Inf. Model. 2013, 53, 2781−2791

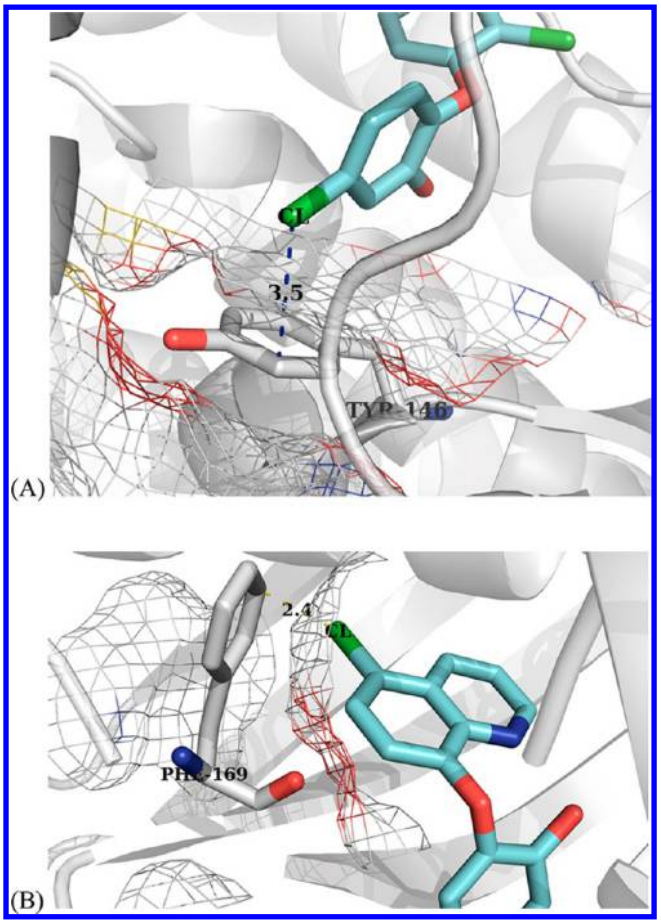


Figure 10. Chlorinated ligands interacting with aromatic rings. Ligands are shown as a cyan stick model, and proteins are shown as a combination of cartoon, stick, and mesh models. (A) Chlorine interacting with aromatic ring of Tyr 146 (PDB id: 1C14). (B) Chlorine interacting with aromatic ring of Phe 169 (PDB id: 3AZ9). The images were prepared with PYMOL.[32]

to 175° range. Again, lesser angles with a maximum at 120° are observed.

X···D−Z (Θ_2) Main-Chain Interactions. The C−O···F and C−N···F angles are shown in Figure 12. The C−N···F bond angles fall within a relatively narrow range relative to C−O···F angles. Furthermore, as expected, the C−N···F interactions occur at more acute angles relative to C−O···F. Interestingly, C−O···F assumes a linear disposition in a couple of instances, while remaining generally restricted in the acute angle range.

In chlorinated ligands, the pattern of angles for C−N···Cl and C−O···Cl are similar to those for fluorine. It is noteworthy that both C−N···F and C−N···Cl interactions fall off sharply at the obtuse end of the interaction spectrum without a smooth diminution of interacting frequencies. In contrast, C−O···F and C−O···Cl angles fall off smoothly in terms of frequency of occurrence (as do C−O···Br and C−O···I). As with fluorine, it appears that the chlorine atom in two of the chlorinated ligands participate in halogen bonding with bond angles adopting near-linear C−O···Cl geometries.

The data set involving brominated ligands interacting with backbone nitrogen is not statistically significant (∼11 interactions observed). Nevertheless, as observed in case of fluorine and chlorine, there are no C−N···Br angles exceeding 120° degrees. (Only one ligand made a ∼110 °C−N···Br angle.)

The C−O···Br angles and frequency of interaction adopt a Guassian-like distribution. Of interest are interactions with four ligands, respectively, that form C−O···Br angles ≥160°.

The C−N···I and C−O···I interactions are similar to C−N··· Br and C−O···Br. The C−N···I interactions show a maxima at 100° before abruptly ending. Furthermore, iodine C−O···I bonds angles are distributed somewhat smoothly over a broad range of 100°, albeit the majority of interactions occurring with C−O···I angles <145°.

C−X···D (Θ_1) Side-Chain Interactions. Halogen side chain interactions were further analyzed. In addition to O and N, S can also serve as a Lewis base. The Θ_1 angle propensities are shown in Figure 13.

Fluorinated ligands demonstrated a broad and richly populated frequency of C−F···O, N, and S angles. Interestingly, a greater frequency of interactions was observed between fluorine and nitrogen (C−F···N) and fluorine and oxygen (C−F···O) relative to sulfur (C−F···S).

The C−Cl···O, N, and S interactions reflected a more jagged landscape with peak frequencies of interactions observed in a bimodal distribution. The peak for each mode rested around 90° and 130°−140° for all three Lewis Bases.

Compared to F and Cl, far fewer brominated ligands are available for analysis within the PDB. In stark contrast to F and Cl, there is no Gaussian profile characterizing the angles made by C−Br to the three Lewis bases. Here, C−Br···O angles within the halogen bond criteria of >140° dominate the landscape. Furthermore, the frequency distribution of C−Br with S and N additionally suggests that there is no overwhelming preference for Br-containing ligands for any particular orientation.

The interaction of iodine containing halo-ligands with the three Lewis bases mimics that of Br in general character with a notable exception. A preference is observed for Br···N type of interactions in contrast with the dominant I···O. Nevertheless, the preference of I for O, N, and S Lewis bases stays essentially non-Gaussian as was observed with Br-containing ligands.

X···D−Z (Θ_2) Side-Chain Interactions. Fluorine is able to interact with all three Lewis bases. Prominent interactions with O are seen at 90° and 120° while making a large spectrum of interactions between 45° and 165°.

The greatest number of C−N···F interactions are adopted at a 105° bond angle, although a wide semi-Guassian distribution exists between 50° and 175°. The C−S···F interactions are much more subdued in range. Most number of C−S···F interactions occur between 75°−85°.

C−O···Cl angles exist in contrast over a very diverse range. A large number of interactions of the Θ_2 angles are observed with Cl and oxygen at 110°−130°. While C−N···Cl angles occur with greater intensity at 80°, 90°, 100°, and 110°, some anomalously large angles are observed for one ligand at 175°. The most interesting aspect of the Cl Θ_2 variety is the C−S···Cl angles. A large number of interactions are found in the 70°−110° angle. Furthermore, a large number of ligands interact at a C−S···Cl angle of 80°.

The number of interactions of Br with side chains is too few to be statistically significant. Yet, between C−O···Br, C−N···Br, and C−S···Br, the C−O···Br interaction at 130° dominates the landscape.

Even fewer Θ_2-type interactions exist involving the iodinated ligands and side-chain Lewis bases. Most O, N, and S interactions favor ranges between 70° and 80°.

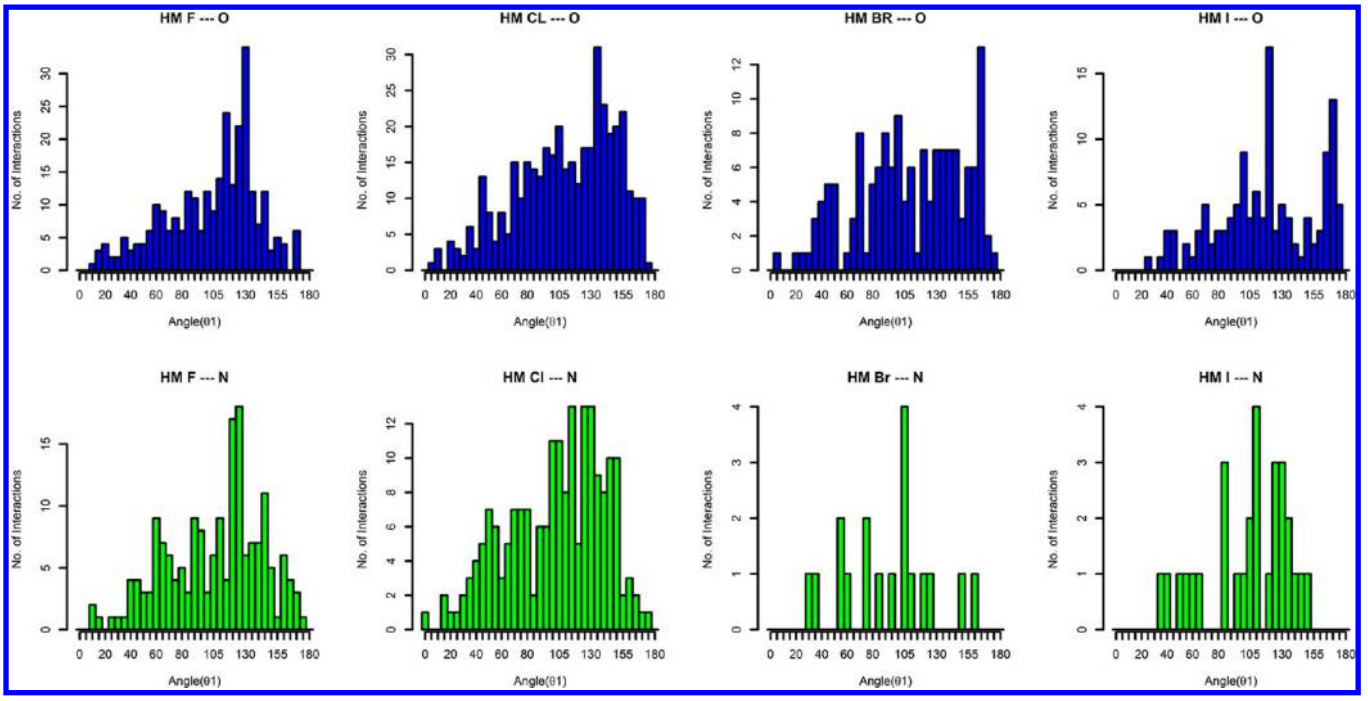


Figure 11. Distribution of Θ_1 angles for each X···D main chain interactions. The *X*-axis shows the angle of interaction, and the number of interactions is shown on the *Y*-axis.

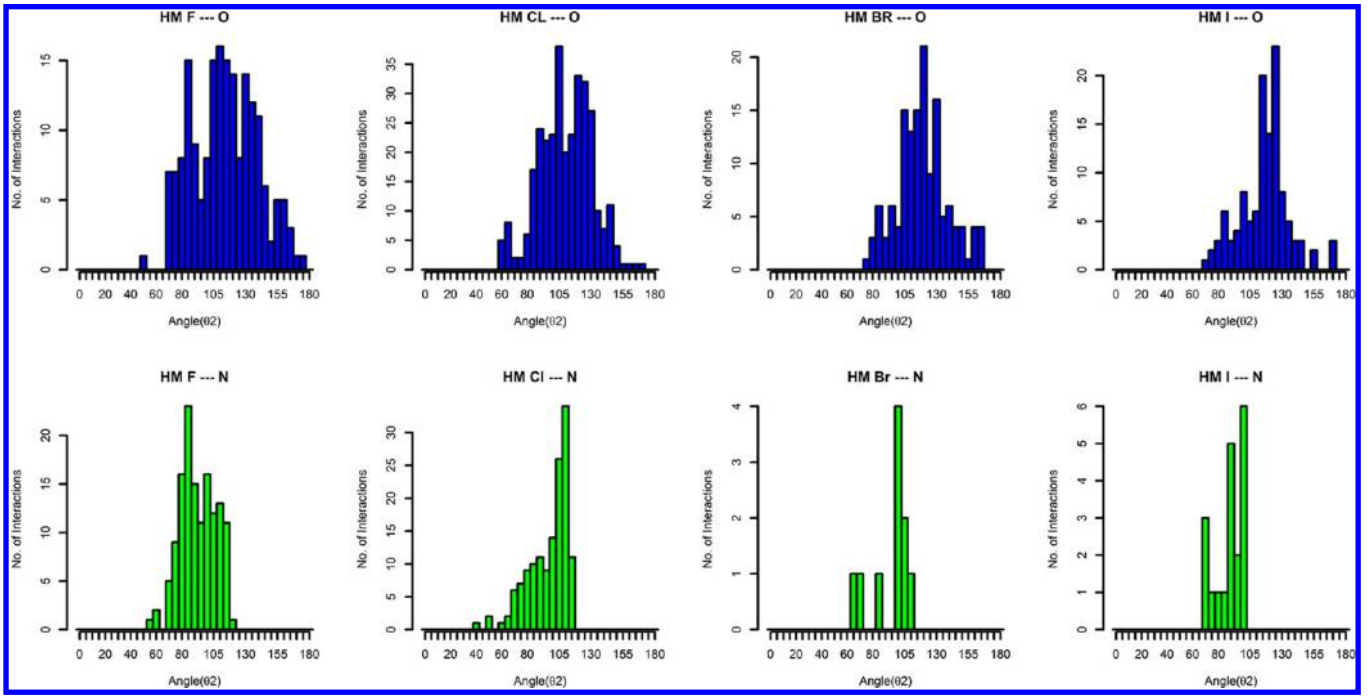


Figure 12. Distribution of Θ_2 angles for each X···D main-chain interactions. The *X*-axis shows the angle of interaction, and the number of interactions is shown on the *Y*-axis.

In summary, while a significant fraction of the observed angles (Θ_1 and Θ_2) can be considered strong, in many instances both main-chain and side-chain Θ_1 and Θ_2 angles show distinct demarcations from linear or predicted geometries. Examination of the PDB structures in these halogen bond interactions reveal that Θ_1-type interactions appear to influence Θ_2 and the vice versa is also possible. This is evidenced from the wide departures observed from theoretically preferred angles for Θ_1 and Θ_2 interactions (Figures 11−14). Furthermore, other factors such as packing and steric influences drive the halogen and Lewis base into nonideal conformations.

Targeting Lewis Bases in Proteins To Form Halogen Bonds. Halogen bond acceptors are present in both the main chain and side chain of the proteins. These Lewis bases can be targeted based on their availability in the binding pocket of target proteins. For main-chain alpha helices, carbonyl oxygen is more solvent exposed, easily accessible, and possess more electron density compared to nitrogen. Hence, it would be ideal

dx.doi.org/10.1021/ci400257k | J. Chem. Inf. Model. 2013, 53, 2781−2791

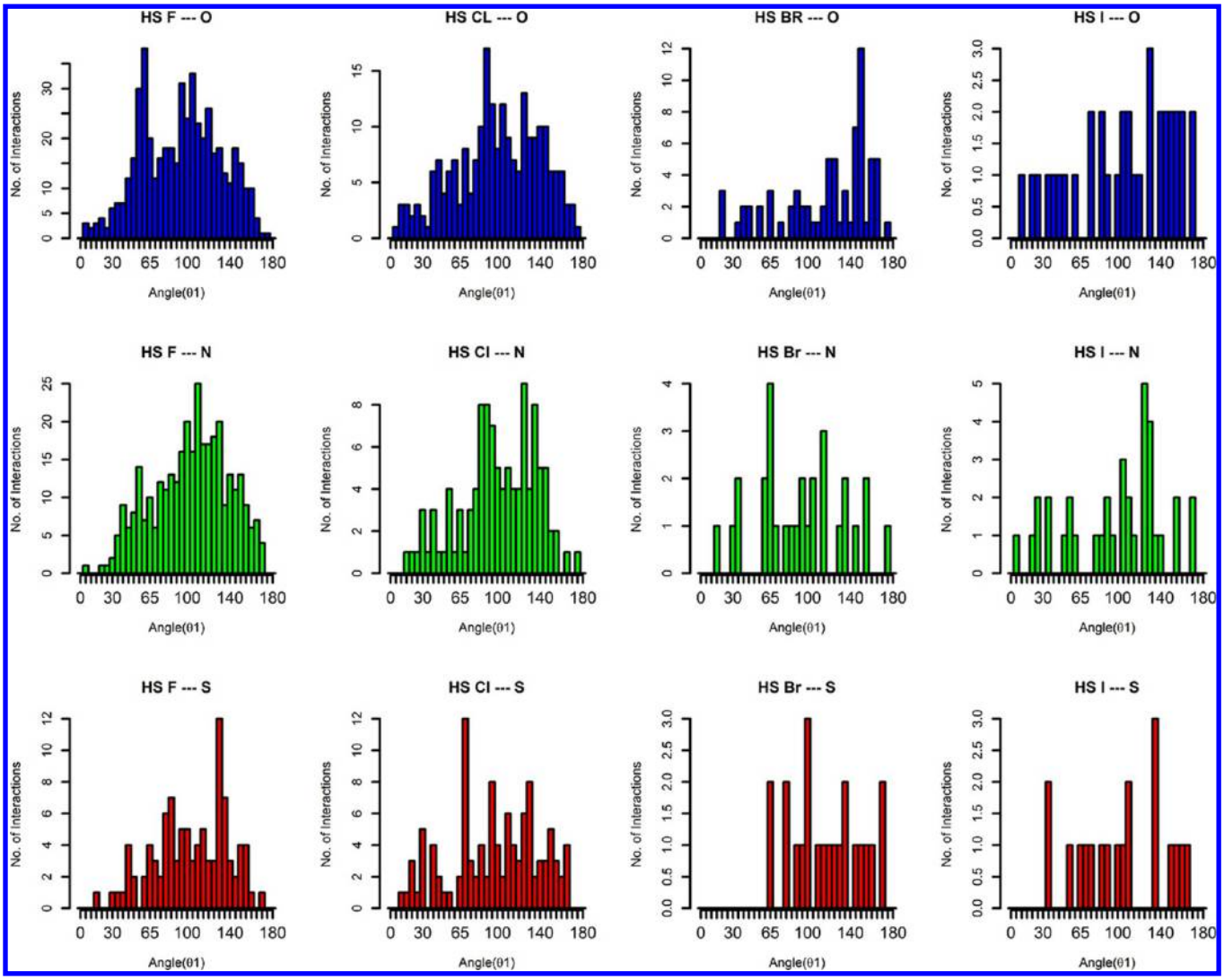


Figure 13. Distribution of Θ_1 angles for each X···D side chain interactions. The *X*-axis shows the angle of interaction, and the number of interactions is shown on the *Y*-axis.

to target available carbonyl oxygen rather than nitrogen. It should be noted that the carbonyl oxygen forms a hydrogen bond with neighboring hydrogen attached to nitrogen in the α helix; the ideal configuration of the halogen bond is orthogonal to a hydrogen bond.[34] In beta sheets, both nitrogen and oxygen can be targeted depending on their availability. However, due to greater electron density on the carbonyl oxygen and delocalization of lone pair electron on nitrogen, the oxygen interactions should be favored over nitrogen interactions.

Given the exposure of backbone amide linkages at loops, it is likely that both backbone oxygen and nitrogen are statistically equi-accessible. Nevertheless, even under these circumstances the preference for halo-oxygen interactions exceeds the tendency toward halo-nitro interactions considering that oxygen possesses higher electron density and diminished electron density on nitrogen due to delocalization. While this (amide on the loop as a halogen bond acceptor) makes for an attractive target in drug design, caution must be exercised as the loops are dynamic in nature which may lead to less stable ligand−protein interactions.

Side-Chain Targeting. Halogen bond donors in side chains can be categorized as (a) the oxygen in the hydroxyl group of Ser, Thr, and Tyr, (b) the carboxylate in Asp and Glu, (c) the nitrogen in His and Asp, (d) the sulfur in Cys and Met, and (f) the aromatic ring in Phe, Tyr, Try, and His. The oxygen in Ser and Thr could be a viable candidate for the targeting of halogen bonds. Interestingly, Tyr consists of both the π-electron cloud from aromatic electron delocalization and an electronegative oxygen. Glu and Asp would be deprotonated at physiological pH and undergo resonance stabilization. Both oxygens have a partial negative charge, making them ideal candidates for halogen bond targeting due to the increased electron density. Nitrogen present in Lys and Arg is usually protonated at physiological pH and cannot function as a halogen bond donor. However, His residues are only partially protonated at physiological pH and can form halogen bonds. It is important to note that X···π interactions in His aromatic rings are possible.

Sulfur is a viable candidate for making halogen bonds given its rich electron density. Unfortunately, the propensity of sulfur-containing amino acids such as Met and Cys in proteins limits its potential in practice.

Aromatic electronic density on Tyr, Trp, and His can also be exploited. It should be noted that in each of these aromatic moieties, the presence of an electronegative atom in the proximity draws the electron density toward itself.

dx.doi.org/10.1021/ci400257k | J. Chem. Inf. Model. 2013, 53, 2781−2791

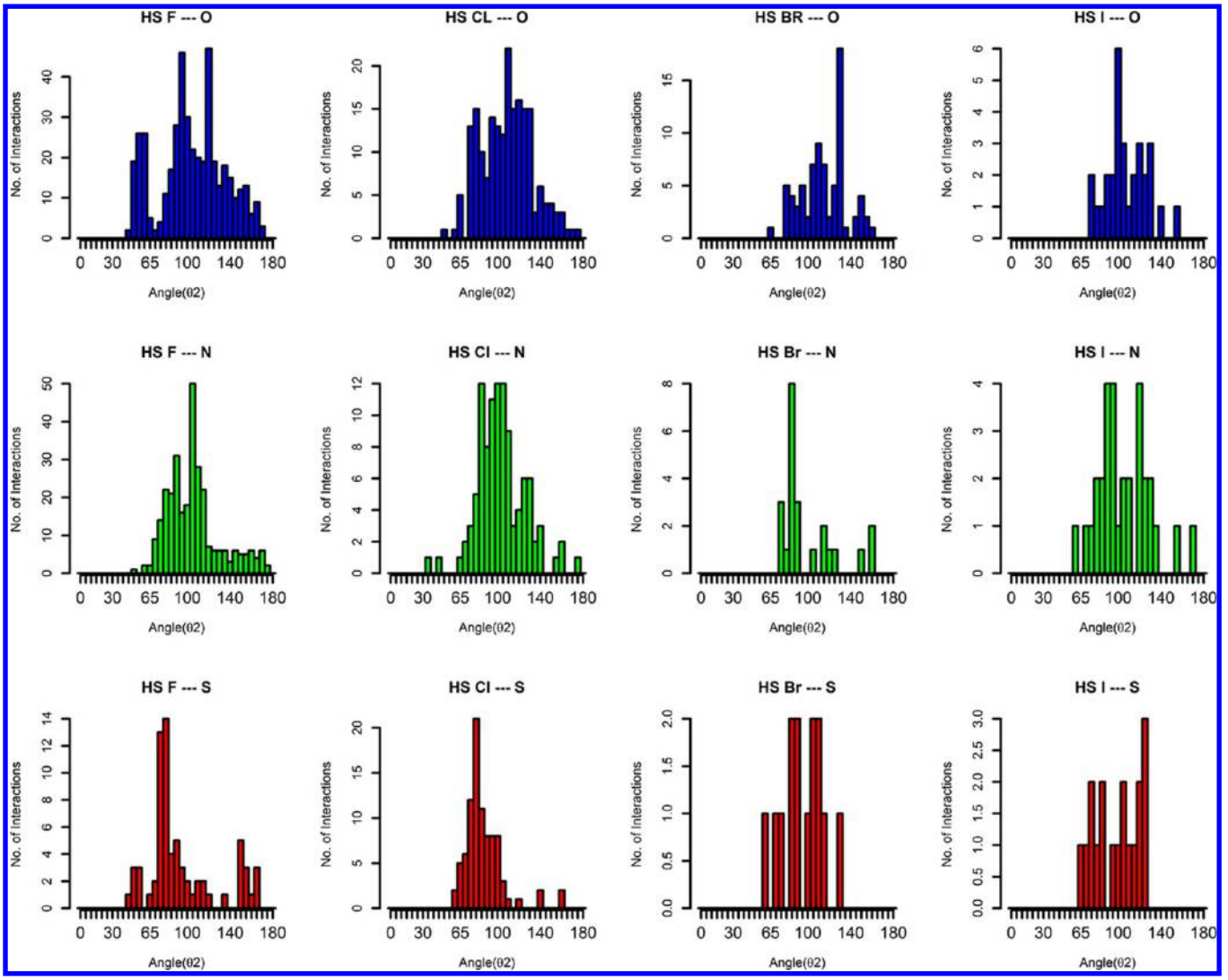


Figure 14. Distribution of Θ_2 angles for each X···D side-chain interactions. The *X*-axis shows the angle of interaction, and the number of interactions is shown on the *Y*-axis.

CONCLUSIONS

In summary, halogen bonds within the PDB were studied to gain insight into the halo-amino acid interactions. The knowledge obtained from this study could be applied in developing molecular mechanics force fields and knowledge-based potentials for optimizing docking scoring functions. Our data reveals interesting facets of halogen bond interactions that will be useful in the directed drug design. For example, the stronger halogen bonds made by the larger halo-atoms suggest that replacement of F and Cl with Br or I may present stronger drug target binding interactions. This in turn could translate into stronger therapeutic indices. Furthermore, stronger halogen bonding interactions with Tyr and Phe moieties suggests that these avenues can be exploited in future pharmacophore architectures. Lastly, given the emerging debate and increasing consideration given to F as a halogen bond participant, we chose to examine its propensity to make halogen bonds with Lewis bases. Our results reveal a rich repository of possible halogen bonding interactions that are likely to seed further investigations in this field.

ASSOCIATED CONTENT

Supporting Information

Raw data files for the C–X···D (Θ_1) and X···D–Z (Θ_2) angles and the python script to calculate Θ_1 angles (sigmaholeangle.py). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

* E-mail: suman.sirimulla@nau.edu (S.S).
* E-mail: mnarayan@utep.edu (M.N.).

Author Contributions

S.S. and J.B. contributed equally to the work. S.S. conceived the project. S.S. and R.V. wrote the programs. S.S. and J.B. performed the calculations. S.S., J.B., and M.N. analyzed and interpreted the results. S.S., J.B., and M.N. contributed to writing the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge Dr. Cindy Browder, Dr. Stephanie Hurst, and Dr. Greg Caporaso at Northern Arizona University for their valuable suggestions.

■ REFERENCES

(1) Tubinger, U. Halogen Bonding Helps Design New Drugs. http://www.sciencedaily.com/releases/2012/06/120605121639.htm (accessed February 26, 2012).
(2) 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*, 2854−2862.
(3) Hernandez, M.; Cavalcanti, S.; Moreira, D.; Asevedo, W. J.; Leite, A. Halogen atoms in the modern medicinal chemistry: Hints of the drug design. *Curr. Drug Targets* **2010**, *3*, 303−315.
(4) 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*, 170−177.
(5) Jorgensen, W. J.; Schyman, P. Treatment of halogen bonding in the OPLS−AA force field: Application to potent anti-HIV agents. *J. Chem. Theory Comput.* **2012**, *8*, 3895−3901.
(6) Small Changes for Big Improvement: Halogen Bond and Drug Discovery. http://www.sciencedaily.com/releases/2013/01/130118064729.htm (accessed April 24, 2013).
(7) 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*, 19508−19516.
(8) Bonnefous, C.; Payne, J. E.; Roppe, J.; Zhuang, H.; Chen, Z.; Symons, K. T.; Nguyen, P. M.; Sablad, M.; Rozenkrants, N.; Zhang, Y.; Wang, L.; Severance, D.; Walsh, J. P.; Yazdani, N.; Shiau, A. K.; Noble, S. A.; Rix, P.; Rao, T. S.; Hassig, C. A.; Smith, N. D. Nitric oxide synthase (iNOS) inhibitor development canidate KD7332, Part 1: Identification of a novel, potent, and selective series of quinolinone iNOS dimerization inhibitors that are orally active in rodent pain models. *J. Med. Chem.* **2009**, *52*, 3047−3062.
(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*, 2267−2278.
(10) Buchini, S.; Buschiazzo, A.; Withers, S. G. A new generation of specific *Trypanosoma crizi trans*-sialidase inhibitors. *Agnew. Chem., Int. Ed.* **2008**, *47*, 2700−2703.
(11) Auffinger, P.; Hays, F.; Westhof, E.; Ho, P. Halogen bonds in biological molecules. *Proc. Natl. Acad. Sci. U.S.A* **2004**, *101*, 16789−16794.
(12) Hardegger, L. A.; Kuhn, B. S.; 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. *Agnew. Chem., Int. Ed.* **2011**, *50*, 314−318.
(13) De Paulis, T.; Hemstapat, K.; Chen, Y.; Zhang, Y.; Saleh, S.; Alagille, D.; Baldwin, R.; Tamagnan, G.; Conn, P. Substituent effects of N-(1,3-diphenyl-1H-pyrazpn-5-yl)benzamides on positive allosteric modulation of the metabolic glutamate-5 receptor in rat cortical astrocytes. *J. Med. Chem.* **2006**, *49*, 3332−3344.
(14) Himmel, D.; Das, K.; Clark, A.; Hughes, S.; Benjahad, A.; Oumouch, S.; Guillemont, J.; Coupa, S.; Poncelet, A.; Csoka, I.; Meyer, C.; Andries, K.; Nguyen, C.; Grierson, D.; 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*, 7582−7591.
(15) 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. Specific CLK inhibitors from a novel chemotype for regulation of alternative splicing. *Chem. Biol.* **2011**, *18*, 67−76.
(16) Bollini, M.; Domaoal, R. A.; Thakur, V. V.; Gallardo-Macias, R.; Spasov, K. A.; Anderson, K. S.; Jorgensen, W. L. Computationally-guided optimization of a dockinghit to yield catechol diethers as potent anti-HIV agents. *J. Med. Chem.* **2011**, *54*, 8582−8591.
(17) Xu, Z.; Liu, Z.; Chen, T.; Chen, T. 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 drug design of potent phosphodiesterase type 5 (PDE5) inhibitors. *J. Med. Chem.* **2011**, *54*, 5607−5611.
(18) Matter, H.; Nazare, M.; Gussregen, S.; Will, D. W.; Schreuder, H.; Bauter, A.; Urmann, M.; Ritter, K.; Wagner, M.; Wehner, V. Evidence for C−Cl/C−Br··· π interactions as an important contribution to protein−ligand binding affinity. *Agnew. Chem., Int. Ed.* **2009**, *48*, 2911−2916.
(19) Wilchen, 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*, 6810−6818.
(20) Politzer, P.; Murray, J. S.; Clark, T. An electrostatically-driven highly directional noncovalent interaction. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7748−7757.
(21) Clark, T.; Hennemann, M.; Murray, J.; Politzer, P. Halogen bonding: The sigma-hole. *J. Mol. Model.* **2007**, *13*, 291−297.
(22) Desiraju, G. R.; Ho, P. S.; Kloo, L.; Legon, A. C.; Marquardt, R.; Metrangolo, P.; Poloitzer, P. A.; Resnati, G.; Rissanen, K. Definition of the Halogen Bond. http://pac.iupac.org/publications/pac/pdf/2013/pdf/8508x1711.pdf (accessed February 25, 2013).
(23) Kemsley, J. N. Hydrogen Bond Reformation. http://cen.acs.org/articles/88/i47/Hydrogen-Bond-Reformulation.html (accessed February 27, 2013).
(24) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision A.1; Gaussian, Inc.: Wallingford CT, 2009.
(25) Pavan, M. S.; Prasad, K. D.; Row, T. N. G. Halogen bonding in fluorine: Experimental charge density study on intermolecular F···F and F···S donor−acceptor contacts. *Chem. Commun.* **2013**, *49*, 7558−7560.
(26) Politzer, P.; Lane, P.; Monice, C. C.; Ma, Y.; Murray, J. S. An overview of halogen bonding. *J. Mol. Model.* **2007**, *13*, 305−311.
(27) Clark, T. σ-Holes. *Comput. Mol. Sci.* **2013**, *3*, 13−20.
(28) Lu, J.; Zhang, B.; Deng, Q.; Wang, J.; Lu, Y.; Zhu, W. The nature and magnitude of specific halogen bonds between iodo-perfluorobenzene and heterocyclic systems. *Int. J. Quantum Chem.* **2010**, *111*, 2352−2358.
(29) Wilchen, 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*, 935−945.
(30) Lu, Y.; Zou, J.; Wang, Y.; Jiang, Y.; Yu, Q. Ab initio investigation of the complexes between bromobenxene and several electron donors: some insights into the magnitude and nature of halogen bonding interactions. *J. Phys. Chem. A.* **2007**, *111*, 10781−10788.
(31) Wilcken, R.; Zimmermann, M.; Lange, A.; Joerger, A.; Boeckler, F. Principles and applications of halogen bonding in medicinal chemistry and chemical biology. *J. Med. Chem.* **2013**, *56*, 1363−1388.

(32) *The PyMOL Molecular Graphics System*, Version 1. 5. 0. 4; Schrodinger, LLC; Portland, OR, 2012.
(33) McDonald, I.; Thornton, J. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **1994**, *238*, 777−793.
(34) Voth, A.; Khuu, P.; Oishi, K.; Ho, P. Halogen bonds as orthogonal molecular interactions to hydrogen bonds. *Nat. Chem.* **2009**, *1*, 74−79.