Chapter 23.6. Halogen interactions in biomolecular crystal structures

MAXIMILIANO J. VALLEJOS, PASCAL AUFFINGER AND P. SHING HO

23.6.1. Introduction

Halogens are uncommon substituents in biological molecules. It is not surprising, therefore, that molecular interactions involving covalent halogen compounds are not foremost in the minds of most biologists. There is, however, an increasing appreciation that halogenated compounds play important roles in biology, starting with the naturally halogenated hormone thyroxine (Fig. 23.6.1.1a) (Wojtczak et al., 2001), and extending to bioactive secondary metabolites isolated as natural products (Gribble, 2004) and a growing list of synthetic compounds that serve as inhibitors against a broad range of enzymes. Proteins and nucleic acids are naturally halogenated through oxidative processes in vivo and in vitro (universally Br and I) to help solve the phase problem in crystallography, or magnetically active nuclei (electron-rich halogens: particularly Br and I) to help solve the electron density problem in nuclear magnetic resonance (19F NMR) (Hammill et al., 2001). Halogenated amino acids such as bromophenylalanine can now be introduced site-specifically into proteins through synthetic oligonucleotides to help determine the structures of DNA and RNA, as well as of protein–nucleic acid complexes. (Xie et al., 2000) and a growing list of synthetic compounds that serve as inhibitors against a broad range of enzymes. Proteins and nucleic acids are naturally halogenated through oxidative processes associated with inflammatory response mechanisms (van Pée & Unversucht, 2003; Shen et al., 2001), with chlorinated and brominated amino acids in proteins correlated with various respiratory ailments and allergic responses (Buss et al., 2003; Wu et al., 2000). In structural biology, it is now common to add electron-rich halogens (particularly Br and I) to help solve the phase problem in crystallography, or to facilitate structure determination by NMR. In nucleic acids, halogenated nucleotide bases can be incorporated in synthetic oligonucleotides to help determine the structures of DNA and RNA, as well as of protein–nucleic acid complexes. Halogenated amino acids such as bromophenylalanine can now be introduced site-specifically into proteins through in vivo incorporation (Wang et al., 2006), which can also facilitate phasing of X-ray diffraction data for protein crystals (Xie et al., 2004), while synthetic fluorinated amino acids have been used to facilitate structural studies by 19F NMR (Hammill et al., 2007). Together, these various systems have contributed to the recent growth in the number of halogen-like halogen interactions observed in biomolecular structures.

For this particular discussion, we will focus on the molecular interactions of halogenated compounds seen in structures determined by single-crystal X-ray diffraction rather than by NMR, primarily because of the strong dependence on the distance and angle relationships required to categorize and understand such interactions (NMR structures are highly dependent on appropriate force fields to define conformation; unfortunately, current force fields do not properly treat interactions involving halogens). A survey of the 16 June 2009 release of the Protein Data Bank (PDB; Berman et al., 2000) showed that there were 2670 entries containing a halogen atom (Table 23.6.1.1). In other words, nearly 5% of the more than 58 000 entries from this release contained at least one molecular halogen, indicating that halogens are an important component in biomolecular structures.

23.6.2. Classical treatment of halogen interactions

23.6.2.1. Non-specific hydrophobic effects

Halogen interactions in proteins and nucleic acid structures have classically been treated non-electrostatically. Indeed, most quantitative structure–activity relationships (QSARs) consider only the steric and lipophilicity contributions of halogens when trying to predict their effect on ligand binding. The majority of interactions can be classified in this sense: non-electrostatic interactions would predict distances between halogen and other non-hydrogen atoms to be greater than or equal to the sum of the respective van der Waals radii (\(\sum r_{vdW}\)), with the halogen atoms sitting in hydrophobic pockets. An example of this is seen in the binding of dibromomethane to methane monooxygenase (Egner et al., 2005), where the Br atoms of the bound ligand are surrounded primarily by the hydrophobic side chains (Leu, Val and Phe) of the protein (Whittington et al., 2001). The non-electrostatic components of molecular halogens can be characterized by the size (as measured by \(r_{vdW}\)) and hydrophobicity (as measured by the difference in partition coefficients and the associated molecular solvation free energy for the accessible surface) of the respective halogen atoms (Table 23.6.2.1).

23.6.2.2. Substituent effects

The classical approach to the electrostatic effects of halogens is to consider them as electron-withdrawing substituents: I, Br, Cl and F atoms are strong electron-withdrawing groups, relative to hydrogen, in their effects on the reactivities of para- and meta-substituted benzoic acids [reflected in the so-called Hammett \(\rho\) constants (Hammett, 1937), Table 23.6.2.1]. However, such substituent effects do not tell us anything about how halogens themselves participate in direct, non-covalent interactions. Our basic understanding is that halogens are electron-rich and therefore we would expect them to act as Lewis bases, serving, for example, as hydrogen-bond acceptors (like carbonyl O or imine N atoms).

![Figure 23.6.1.1](image-url) Examples of electrostatic interactions of halogenated ligands with proteins. The C atoms of the protein residues are shown in green and those of the bound ligands in black. (a) An I atom from tetraiodothyroxine (violet) is shown with its close (3.07 Å) I bonding interaction (blue dots) to the carbonyl O atom of Ala109 of transthyretin (Wojtczak et al., 2001). (b) The Br atoms (orange) of tetrabromobenzotriazole form short contacts to the carbonyl O atoms of Glu81 (3.0 Å), Leu83 (2.9 Å) and Ile10 (3.3 Å), and to the aromatic ring of Phe80 (3.5 Å to the centre), of phospho-CDK2/cyclin A (De Moliner et al., 2003). There is an additional interaction with a water molecule (2.91 Å). (c) An F atom (cyan) from the inhibitor ZK-807834 interacts with the amide C atom of the Gln192 side chain in the ligand-binding site of the clotting factor Xa (Adler et al., 2000).
23. STRUCTURAL ANALYSIS AND CLASSIFICATION

Table 23.6.1.1
Interactions in crystal structures in the Protein Data Bank

The numbers of interactions (NI) for molecular halogens with oxygen (O), nitrogen (N), sulfur (S) and carbon atoms (C) are tabulated for single-crystal structures at resolutions of 3.0 Å or better from the 16 June 2009 release of the PDB (containing 58,236 total entries). A contact is considered to be an interaction if the distance between the halogen and its partner is between 2.2 Å and the van der Waals contact distance (Dmax). O-atom partners are counted separately for non-water and water interactions, while interactions with non-aromatic atoms are counted separately from those involving aromatic atoms.

<table>
<thead>
<tr>
<th>Acceptor atom</th>
<th>F</th>
<th>Cl</th>
<th>Br</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dmax (Å)</td>
<td>NI</td>
<td>Dmax (Å)</td>
<td>NI</td>
</tr>
<tr>
<td>O (non-water)</td>
<td>2.99</td>
<td>1219</td>
<td>3.27</td>
<td>241</td>
</tr>
<tr>
<td>O (water)</td>
<td>3.52</td>
<td>1454</td>
<td>3.80</td>
<td>263</td>
</tr>
<tr>
<td>N</td>
<td>3.02</td>
<td>977</td>
<td>3.30</td>
<td>206</td>
</tr>
<tr>
<td>S</td>
<td>3.27</td>
<td>45</td>
<td>5.55</td>
<td>39</td>
</tr>
<tr>
<td>C</td>
<td>3.17</td>
<td>1466</td>
<td>3.45</td>
<td>511</td>
</tr>
</tbody>
</table>

Aromatic (side chains of Phe, Tyr, Trp and His in proteins, and the bases of nucleotides)

| O             | 2.99 | 42 | 3.27 | 22 | 3.37 | 58 | 3.50 | 8  |
| N             | 3.02 | 174 | 3.30 | 9  | 3.40 | 16  | 3.53 | 0  |
| C             | 3.17 | 1466 | 3.45 | 511 | 3.55 | 338 | 3.68 | 77 |
| Total         | 5388 | 1533 | 2048 | 2048 | 1634 |

† Dmax is the sum of the van der Waals radii [r_{vdw}; Bondi (1964)] for the halogen and its interaction partner. In the case of water, this is the sum of the length of the O—H bond (0.98 Å) plus r_{vdw} of the H atom (1.09 Å) and r_{vdw} of the respective halogen. The longer Dmax for interactions with water allows the halogens to serve as hydrogen-bond acceptors.

<table>
<thead>
<tr>
<th>Halogen (X)</th>
<th>r_{vdw} (Å)</th>
<th>log(P_P) − log(P_x)</th>
<th>SFE (cal Å⁻²)</th>
<th>σ_p</th>
<th>σ_m</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>1.47</td>
<td>0.12</td>
<td>11.9</td>
<td>0.062</td>
<td>0.337</td>
</tr>
<tr>
<td>Cl</td>
<td>1.75</td>
<td>0.34</td>
<td>13.2</td>
<td>0.227</td>
<td>0.573</td>
</tr>
<tr>
<td>Br</td>
<td>1.83</td>
<td>0.84</td>
<td>37.6</td>
<td>0.232</td>
<td>0.393</td>
</tr>
<tr>
<td>I</td>
<td>1.98</td>
<td>1.10</td>
<td>41.5</td>
<td>0.276</td>
<td>0.353</td>
</tr>
</tbody>
</table>

† From Bondi (1964). †† Defined as the base-10 logarithm of the partition (P) from water to octanol of the halogenated benzene (Xₚ, where X is the halogen) minus that of benzene (Pₓ). Partition coefficients from Hansch (1979). § Calculated as ~2.303RT[log(P_P) − log(P_x)] divided by the solvent-accessible surface of the halogen in the respective halobenzene model. ¶ From Hammett (1937). Positive values reflect the electron-withdrawing capacity of the substituent.

Indeed, F atoms can typically replace hydroxyl groups as hydrogen-bond acceptors in carbohydrates (Scott & Viola, 1998; Lougheed et al., 1999; Ly et al., 2000; Hart et al., 2000).

23.6.3. Electrostatic molecular halogen interactions

Molecular halogen interactions are recognized as being highly diverse, extending well beyond the classical properties of size, lipophilicity and substituent effects. The remainder of this review will focus on a class of electrostatic interactions called halogen bonds (X bonds) and other similar electrostatic interactions, primarily because these are specific interactions that have been shown to affect molecular conformations and ligand binding in single-crystal structures of proteins and nucleic acids.

23.6.3.1. Biological halogen bonds

The growing use of halogens for crystallographic phasing (halide soaks notwithstanding) and in natural product and drug development means that structural biologists must become more aware of the non-covalent molecular interactions involving the Group VII elements. In many cases, halogens are introduced as ‘neutral’ substitutions, but they are not altogether unobtrusive: there are now several examples where a halogenated nucleotide base has been seen to affect molecular conformations and ligand binding in single-crystal structures of DNA. In one such example, a brominated uracil base induced a decanucleotide sequence to crystallize as a four-stranded Holliday junction (Fig. 23.6.3.1), while the analogous unbrominated and methylated sequences crystallized as standard B-DNA double helices (Hays et al., 2003). Although we did not recognize this as an X bond at the time, the Br atom was seen to form a stabilizing interaction with the negatively charged phosphate O atom to force the DNA backbone into a tight U-turn and, as a result, to help form the four-stranded complex. In fact, we and others (Muzet et al., 2003) had thought that such short halogen–oxygen distances represented a completely new type of interaction, until an internet search led to the lecture by Odd Hassel for the 1969 Nobel Prize in Chemistry (Hassel, 1972). In short, there is ‘nothing new under the sun’ [Ecclesiastes 1:9–14].

The type of molecular interaction described by Hassel involves halogens serving as Lewis acids rather than as Lewis bases, and

Figure 23.6.3.1
The single-crystal structure of the DNA sequence d(CCAGTAC-brUGG), where brU is 5-bromouracil (Br atom shown as a sphere), as a four-stranded Holliday junction. The inset shows the electron-density map (2Fₑ — Fₑ map, contoured at 1.5σ in blue and 3σ in red) around the brU residue, with the short contact between the Br atom and the phosphate O atom at the tight U-turn of the DNA junction backbone indicated.
has seen a resurgence of interest in materials science. These interactions, initially called ‘charge-transfer bonds’ (Hassel, 1972), could, for example, draw the Br atoms of the Br₂ molecule to within 2.7 Å of the O atoms in dioxane (∼0.5 Å shorter than \( \sum r_{\text{vdW}} \) of the interacting atoms). This would initially appear to be contrary to the expectation that electron-rich atoms should repel each other, but, as we shall see, halogens are not uniform spheres of polarizability and therefore not typically considered X-bond donors (Fig. 23.6.3.2). The molecular surfaces are shown looking down the C–X bond, with electropositive and electronnegative potentials shown in blue and red, respectively. (b) Comparison of the geometries of hydrogen bonds (left) and halogen bonds (right) [adapted from Voth & Ho (2007)]. In each case, the weak interactions are characterized by the distance between the donor (D) and acceptor (A) atoms (\( R_{\text{D} \cdot \cdot \cdot \text{A}} \)) and the angle of approach of the acceptor towards the donor (\( \theta_1 \)) and the donor towards the acceptor (\( \theta_2 \)).

(23.6.3.2) Properties of halogen bonds. (a) Ab initio calculations on model compounds (halomethane top, 5-halouracil middle, 5-halocytosine bottom) show an increase in polarizability of molecular halogen atoms (X), with I > Br > Cl > F [adapted from Auffinger et al. (2004)]. The molecular surfaces are drawn looking down the C–X bond, with electropositive and electronnegative potentials shown in blue and red, respectively. (b) Comparison of the geometries of hydrogen bonds (left) and halogen bonds (right) [adapted from Voth & Ho (2007)]. In each case, the weak interactions are characterized by the distance between the donor (D) and acceptor (A) atoms (\( R_{\text{D} \cdot \cdot \cdot \text{A}} \)) and the angle of approach of the acceptor towards the donor (\( \theta_1 \)) and the donor towards the acceptor (\( \theta_2 \)).

The ‘strength’ of the X bond is expected to be dependent on the polarizability of the halogen donor, with I > Br > Cl > F [adapted from Auffinger et al. (2004)]. The molecular surfaces are drawn looking down the C–X bond, with electropositive and electronnegative potentials shown in blue and red, respectively. (b) Comparison of the geometries of hydrogen bonds (left) and halogen bonds (right) [adapted from Voth & Ho (2007)]. In each case, the weak interactions are characterized by the distance between the donor (D) and acceptor (A) atoms (\( R_{\text{D} \cdot \cdot \cdot \text{A}} \)) and the angle of approach of the acceptor towards the donor (\( \theta_1 \)) and the donor towards the acceptor (\( \theta_2 \)).

23.6. HALOGEN INTERACTIONS

The strength of halogen bonds is characterized by the distance between the donor (D) and acceptor (A) atoms (\( R_{\text{D} \cdot \cdot \cdot \text{A}} \)) and the angle of approach of the acceptor towards the donor (\( \theta_1 \)) and the donor towards the acceptor (\( \theta_2 \)).

Properties of halogen bonds. (a) Ab initio calculations on model compounds (halomethane top, 5-halouracil middle, 5-halocytosine bottom) show an increase in polarizability of molecular halogen atoms (X), with I > Br > Cl > F [adapted from Auffinger et al. (2004)]. The molecular surfaces are drawn looking down the C–X bond, with electropositive and electronnegative potentials shown in blue and red, respectively. (b) Comparison of the geometries of hydrogen bonds (left) and halogen bonds (right) [adapted from Voth & Ho (2007)]. In each case, the weak interactions are characterized by the distance between the donor (D) and acceptor (A) atoms (\( R_{\text{D} \cdot \cdot \cdot \text{A}} \)) and the angle of approach of the acceptor towards the donor (\( \theta_1 \)) and the donor towards the acceptor (\( \theta_2 \)).

The ‘strength’ of the X bond is expected to be dependent on the polarizability of the halogen donor, with I > Br > Cl > F [adapted from Auffinger et al. (2004)]. The molecular surfaces are drawn looking down the C–X bond, with electropositive and electronnegative potentials shown in blue and red, respectively. (b) Comparison of the geometries of hydrogen bonds (left) and halogen bonds (right) [adapted from Voth & Ho (2007)]. In each case, the weak interactions are characterized by the distance between the donor (D) and acceptor (A) atoms (\( R_{\text{D} \cdot \cdot \cdot \text{A}} \)) and the angle of approach of the acceptor towards the donor (\( \theta_1 \)) and the donor towards the acceptor (\( \theta_2 \)).

Figure 23.6.3.2

The geometry of X bonds provides an understanding of what is involved in the formation and stabilization of X bonds in biological systems. By analogy with the structure of X bonds between small organic molecules (Cody & Murray-Rust, 1984; Lommerse et al., 1996; Ouvrard et al., 2003), the geometry of biological X bonds is characterized by the angles of approach of the acceptor to the halogen donor (\( \theta_1 \) relative to the C–X bond; Fig. 23.6.3.2b) and the halogen towards the base acceptor (\( \theta_2 \); Fig. 23.6.3.2b) (Auffinger et al., 2004). The \( \theta_1 \) angle is essentially linear, consistent with this being an electrostatic interaction of the base acceptor, with the electropositive crown, or \( \sigma \) hole, being aligned along the C–X bond, as seen previously in organic complexes. The \( \theta_2 \) angle is ∼110–120°, suggesting that the X bonds involve the non-bonding electrons of the acceptor. However, an extended analysis showed that most X bonds in protein–ligand complexes are aligned perpendicular to the peptide plane, indicating that it is the \( \pi \)- rather than n-electron
**Figure 23.6.3.3**
The orthogonal relationship between hydrogen bonds and X bonds that share common carbonyl O atoms as acceptors (Voth et al., 2009), showing the approach of halogens from the ligands towards the carbonyl O atoms of the peptide bonds in ideal models of a β-sheet (left) and an α-helix (right) (adapted from Voth et al. (2009)). In the β-sheet, the hydrogen bond to the shared O atom is oriented in the plane of the peptide bond, allowing the orthogonal X bond to approach from above or below and perpendicular to the plane. In the α-helix, the hydrogen bond is oriented ~40° from the plane and therefore the orthogonal relationship aligns the X bond ~50° from the peptide plane.

This preference for the π system in biological X bonds is not what would be predicted from surveys of small-molecule structures. In organic complexes, an X bond to a π system is seen rarely and only when there appears to be significant steric crowding that limits access to the X–bond acceptor atom (Ouvrard et al., 2003). We have proposed the concept that X bonds in proteins are orthogonal molecular interactions to hydrogen bonds when they share a common carbonyl oxygen acceptor (Fig. 23.6.3.3) (Voth et al., 2009), with orthogonality referring to both the perpendicular orientation of the shared X and hydrogen bonds and the relative thermodynamic independence of the two interactions. This model suggests that X bonds can be introduced in a well defined geometry to enhance the affinity at a ligand–protein interaction interface without disrupting the structural hydrogen bonds of the protein target. This is potentially a very powerful concept when X bonds are applied as a molecular tool for the rational design and synthesis of new inhibitors and drugs against enzyme targets (Metrangolo & Resnati, 2008).

How strong is a biological X bond? An X bond has been estimated, from ab initio calculations and crystallization of organic molecules, to be equivalent to a hydrogen bond in energy. Similar calculations on X bonds to peptide O atoms in protein–ligand complexes suggest that a hydrogen bond has approximately the same energy as an iodine X bond, but I > Br > Cl, as expected from their relative polarizabilities (Lu et al., 2009). This is consistent with the observation that halogenated benzenes can form halogen bonds to the S atom of a Met residue in an artificial pocket engineered into a T4 lysozyme mutant (Liu et al., 2009), but that the energies of these interactions are weak (approximately 0.5–0.7 kcal mol⁻¹ more stabilizing than simple van der Waals interactions).

The energies of X bonds versus hydrogen bonds in a nucleic acid system have been estimated directly using a crystallographic assay that takes advantage of the conformational isomerization of a DNA Holliday junction (Grainger et al., 1998; Lilley, 1999; McKinney et al., 2003) and the ability of specific hydrogen and X bonds to the phosphate O atoms to stabilize the four-stranded complex (Eichman et al., 2001; Hays et al., 2005, 2003) (Fig. 23.6.3.4). From this set of competition studies, it was determined that a Br···O⁻¹/² type X bond is ~2–5 kcal mol⁻¹ more stable than a hydrogen bond (Voth et al., 2007). When extended to other halogens, the studies showed that I > Br > H > F (unpublished work). This competition assay also demonstrates that X bonds can be used to control DNA conformation (Voth et al., 2007).

The energies determined for the C—X···O⁻¹/² X bonds in this DNA system (Voth et al., 2007) are at least four to ten times greater than those estimated for the C—X···S X bonds in the T4 lysozyme–ligand studies (Liu et al., 2009), and one might ask whether these discrepancies can be reconciled. We suggest that it is the nature of the Lewis-base acceptor that accounts for the dramatic difference, reflecting perhaps the difference between the base strength of an S atom in a thiol ether and that of the formally charged O atom of a phosphodiester. This hypothesis can be tested by comparing the energies of a neutral and a formally charged phosphate O atom. The stabilizing potential of a biological X bond apparently falls within a very broad range from near zero to a value >5 kcal mol⁻¹ more stable than a standard hydrogen bond, depending on the polarizability of the donor halogen and the basicity of the acceptor. Thus, X bonds can be applied as a molecular tool that can be tuned energetically to control and drive macromolecular interactions, which can be exploited in synthetic biology for creating new nanomaterials from biological molecules.

Although oxygen, nitrogen and sulfur are the dominant X–bond acceptors, the side chains of aromatic amino acids can also serve as X–bond acceptors (Voth & Ho, 2007). In the specific examples available for ligand complexes with protein kinases, the C—X bonds are oriented perpendicular to the plane and impinge on the centre of the aromatic ring (Fig. 23.6.1.1b), similar to hydrogen bonds to aromatic amino acids and cation–π interactions between positively charged and aromatic amino acids (Meyer et al., 2003). The X–bond acceptor, in these cases, would be the delocalized π-electron system of the aromatic side chains.
23.6. HALOGEN INTERACTIONS

Ab initio calculations suggest that the electrostatic contribution to aromatic X bonds is very weak (around 1 kcal mol\(^{-1}\)) at an optimum distance of 3.6 Å from a Br atom to the centre of an aromatic ring, or ~0.3 Å shorter than the predicted closest van der Waals approach; Voth & Ho, 2007), but hydrophobic effects may enhance the stabilizing potential of this type of interaction geometry.

23.6.3.2. Other halogen-specific electrostatic interactions in biological macromolecules

Specific X-bonding interactions between electropositive σ holes and Lewis-base acceptors are only one class of electrostatic interaction available to halogens. There are other classes that are variations on the same theme, where the halogen serves as a Lewis acid attracted to an electron-rich base, or where the halogen acts as the base attracted to an electrophilic acid. The most obvious example would be an electronegative F atom (which is not polarized) interacting with electropositive atoms. A very clear example of this type of electrostatic interaction is seen in the short contact (2.9 Å) between one F atom of the ZK-807834 inhibitor and the amide C atom of the Gln192 side chain in the ligand-binding site of the clotting factor Xa (Fig. 23.6.1.1c) (Adler et al., 2000). We should note that, in addition to creating an electronegative crown, polarization of a halogen will also render the annulus around the σ hole more electronegative. This electronegative ring can serve as a Lewis base to interact with electropositive atoms and is observed in interactions of CI, Br and I from ligands with the carbonyl C atoms, as opposed to the O atoms of the peptide backbone and various amino-acid side chains (Paulini et al., 2005).

The distinction between X bonds and these other types of electrostatic interaction for any given halogen atom is dependent on the geometry of approach and the electrostatic nature of the atom approaching the halogen. An electronegative base approaching linearly to the C—X bond towards the σ hole would define an X bond, while an electropositive donor approaching perpendicular to the C—X bond would probably not. For example, a survey of short contacts between halogens and aromatic side chains (Saraogi et al., 2003) showed that F atoms were the most commonly seen to interact with His, Phe, Trp and Tyr residues, although I, Br and Cl also show interactions with the various geometries. The analysis indicates that fluoride interactions show a single broad distribution for distances from the halogen to the aromatic ring, while the distance distribution for chlorine appears to be bimodal, with one group at very short distances and another at about the van der Waals contact distance. A more complete analysis to correlate the distance–angle relationships may help to distinguish between aromatic X bonds to delocalized σ systems and the opposing electrostatic interactions to the more electropositive carbon centres.

23.6.4. Concluding remarks

Halogens show a variety of different types of interaction with atoms in proteins and nucleic acids, and the particular type of interaction depends on both the properties of the halogen and what it is in contact with. It is this variability that may account for the distinct preference for certain halogens to interact with particular types of amino acids in proteins (Kortagere et al., 2008). However, unlike most other interactions there is a strong directionality to molecular halogen interactions, particularly in biological systems, which may specify one type of interaction over another.

If we could fully understand the relationships between the geometries and energies of interactions, the types of electrostatic interaction described here could be exploited in biomolecular engineering in ways that are similar to, and extensions beyond, the principles used to design synthetic materials. We should stress, however, that our understanding of such interactions, particularly of biological X bonds, is still in its infancy. For example, Kraut et al. (2009) engineered halogenated phenylalnine residues into ketosteroid isomerase in order to determine whether an X bond could replace a functionally important hydrogen-bond interaction with an oxyanion hole, and found that the halogenated residues behaved similarly to a non-halogenated Phe residue at this functional position. However, close inspection of the model for the X–Phe mutant in this construct suggests that the geometry of approach of the halogen towards the oxyanion may not be the ideal for a particularly strong X bond. Thus, although it is tempting to apply such interactions in rational design strategies, we may simply not have enough understanding yet of the basic underlying principles to exploit these interactions effectively at this time. In another case, formation of the hairpin conformation that forms a kissing-loop structure in the RNA dimerization initiation site (DIS) from HIV-1 was shown to be facilitated by bromination of the uracil at the U3 position, compared with the bromination of uracils at other positions and the non-brominated RNA, which favour the duplex form (Emnifar et al., 2007). The crystal structures of the brominated U3 versus other uracils provide no insight into why this particular halogen at this specific position should have such a profound effect on the conformational preference of the RNA.

Still, we would argue that a combination of experimental studies, coupled with ab initio and database analyses, will provide detailed understanding of the geometry–energy relationships for the various types of halogen interaction available to biological macromolecules. With this understanding, novel molecular tools can be developed to facilitate the rational design of new and more effective inhibitors against therapeutic targets, to manipulate the conformation and stability of protein and nucleic acid structures in the design of new nanomaterials, and to design new interaction interfaces for protein–protein and protein–nucleic acid complexes in synthetic biology.

References


