Published in final edited form as:

J Phys Chem B. 2015 July 23; 119(29): 9140–9149. doi:10.1021/jp509003r.

Force Field Model of Periodic Trends in Biomolecular Halogen Bonds

Matthew R. Scholfield[†], Melissa Coates Ford[†], Crystal M. Vander Zanden[†], M. Marie Billman[‡], P. Shing Ho^{*,†}, and Anthony K. Rappé^{*,‡}

[†]Department of Biochemistry & Molecular Biology, 1870 Campus Delivery, Colorado State University, Fort Collins, CO 80523-1870

[‡]Department of Chemistry, Colorado State University, Fort Collins, CO 80523

Abstract

The study of the noncovalent interaction now defined as a halogen bond (X-bond) has become one of the fastest growing areas in experimental and theoretical chemistry—its applications as a design tool are highly extensive. The significance of the interaction in biology has only recently been recognized, but has now become important in medicinal chemistry. We had previously derived a set of empirical potential energy functions to model the structure-energy relationships for bromines in biomolecular X-bonds (BXBs). Here, we have extended this force field for BXBs (*f*BXB) to the halogens (Cl, Br, and I) that are commonly seen to form stable X-bonds. The *f*BXB calculated energies show a remarkable one-to-one linear relationship to explicit BXB energies determined from an experimental DNA junction system, thereby validating the approach and the model. The resulting parameters allow us to interpret the stabilizing effects of BXBs in terms of well-defined physical properties of the halogen atoms, including their size, shape, and charge, showing periodic trends that are predictable along the Group VII column of elements. Consequently, we have established the *f*BXB as accurate computational tool that can be applied to, for example, for the design of new therapeutic compounds against clinically important targets and new biomolecular based materials.

I. INTRODUCTION

The halogen bond (X-bond) has become an important molecular design tool in nearly all areas of the chemical sciences¹, from liquid crystals² and molecular receptors³ to nanomaterials^{4, 5} and molecular conductors⁶. Biology and biochemistry are areas where X-bonds have the potential to make a great impact, particularly in the computer-aided design of new halogenated inhibitors against clinically important targets^{7–10}. This potential can only be fully realized with the development of computational tools to accurately model X-bonds in biological molecules (BXBs). We describe here the complete parameterization of a force field for biomolecular X-bonds (a *f*BXB¹¹) that is based on elemental properties of size, shape, and electrostatic potential of the halogens (Cl, Br, and I) commonly seen in

*Corresponding Authors: PSH: Phone: 970-491-0569, FAX: 970-491-0494, Shing.Ho@colostate.edu. AKR: Phone: 970-491-6292, FAX: 970-491-1801, Rappe@lamar.colostate.edu.

The authors declare no conflicts of interest in this study.

biology^{8, 12–14}, and show for the first time that such a molecular mechanics (MM) approach can directly model experimentally determined energies of BXBs with explicitly defined geometries^{15–17}.

The study of halogen bonds and their applications in material and biological chemistry has become one of the fastest growing research areas in chemistry (as seen in the explosive growth in numbers of articles and citations) in the past decade^{5, 8}. The concept that halogenated compounds can form stable noncovalent complexes was first recognized two centuries ago by Colin when the mixture of molecular iodine with ammonia was seen to produce a shiny liquid that differed from either starting component¹⁸. The extension of Mulliken's charge transfer concept¹⁹ by Hassel to molecular complexes helped explain the short stabilizing interactions seen between molecular halogens and electron-rich atoms, *e.g.*, between Br₂ and dioxane^{20, 21}. The charge transfer bond remained in scientific hibernation for two decades, until it reemerged in the 1990's as the halogen bond. The change in name reflected a change in the current understanding that the interaction is primarily electrostatic in nature²², analogous to the more widely recognized hydrogen bond (H-bond)²³. The recently published IUPAC definition²⁴ provides guidelines for how we describe and study X-bonds, and places the interaction well within the accepted chemical tools available for molecular design.

Perhaps the most readily accessible description for the origins of the X-bond is the σ -hole model²⁵, which treats the interaction as primarily electrostatic (Figure 1b)²², although there remains significant controversy over the contribution of dispersion²⁶ and even charge-transfer²⁷ to the interaction. The model describes X-bonding as a consequence of a halogen forming a σ -bond, such as a C—X bond. Covalently bonding a halogen to, for example, an aromatic uracil base results in the depopulation of the valence p_z -orbital, which creates an electropositive crown (the σ -hole) at the halogen surface diametrically opposed to the σ -bond. The σ -hole serves as the X-bond donor to an electron-rich acceptor. The size and electropositive potential of the σ -hole increases as the size of the halogen increases and as the electron-withdrawing ability of the substituent bound to the halogen increases; thus, the energies of X-bonds can be tuned and fine-tuned^{8, 28}, which make the interaction particularly useful as a design tool.

X-bonds in biomolecular systems (BXBs) are predominantly found in complexes of proteins with halogenated inhibitors^{8, 12}, with the carbonyl oxygen of the peptide bonds found to be the primary acceptor in these complexes^{8, 12–14, 29}. For the most part, the geometries relating the donors to acceptors in BXBs follow those seen in simpler organic complexes, with distances shorter than the sums of the respective van der Waals radii (ΣR_{vdW}), an angle of approach of the acceptor to the σ -hole along C—X bond (θ_1) that is nearly linear, and an angle of approach of the halogen to the acceptor covalent bond (θ_2) either towards the non-bonding electrons or π -electrons of the acceptor^{8, 12, 29–31}. One exception has been that most peptide bonds in proteins are hydrogen bonded (H-bonded) in secondary structures (α -helices and β -sheets), leading us to further propose that BXBs are geometrically and energetically orthogonal to H-bonds when they share the same acceptor atoms³²—the relationship now also seen in small molecule complexes^{33, 34}. Not surprisingly, the most direct application of BXBs has been in medicinal chemistry for the computer-aided design

of halogenated inhibitors and drugs^{7–10, 35, 36}. Recently, a library of halogen-enriched fragments to help identify potential BXB sites to facilitate drug discovery³⁷, in one example to design inhibitors that induce apoptosis in human cell lines carrying the cancer-related Y220C mutation of p53³⁸.

The various geometries of BXBs are associated with different energies of interaction. Thus, in order to understand how a particular X-bonding interaction may contribute, for example, to the affinity of a protein for a halogenated ligand, we need to understand and describe the structure-energy relationships in a rigorously quantitative manner. There is now a fairly large library of crystal structures of proteins bound to halogenated ligands, along with their affinities as reflected in dissociation constants (K_{DS}) or IC50s²⁹. The energies for a specific geometry of any particular BXB in such complexes, however, cannot be easily isolated from all other competing or contributing interactions.

To overcome this problem, we developed a DNA model system (Figure 2)¹⁵ in which the structures of various BXBs can be determined crystallographically¹⁷ and their energies directly measured either through a crystallographic^{15, 16} or calorimetric assay^{17, 40}. The system competes a BXB against analogous H-bonds to stabilize a four-stranded DNA junction¹⁵ and, with the energy of the H-bond now determined, the actual energy of each BXB geometry can be determined in isolation^{17, 40}. Using this system, we have shown that the strength of halogen bonds increases according to the series $F < Cl < Br < I$, with energies ranging from -0.52 to -6 kcal/mol¹⁷. In addition, as the X-bonding energy becomes more negative and stabilizing, the geometry becomes more ideal (the ΣR_{vdW} becomes shorter relative to the standard atomic radii and the approach of the acceptor, θ_I , becomes more linear towards the σ -hole)¹⁷.

The experimentally determined structure-energy relationships can now be uniquely applied to validate any type of computational model. Quantum mechanics (QM) is the most accurate method to computationally model X-bonds; however, the most accessible approach for modeling biomolecular systems is molecular mechanics (MM), which in its standard form improperly treats halogens as isotropic charges, thereby ignoring the σ -hole. It is essential, therefore, to develop a computational method that can accurately model BXBs, which can be incorporated into a standard MM force field to simulate the structure-energy relationships of X-bonds in biomolecular systems, including inhibitor-protein complexes.

The “extra point” or “explicit σ -hole” extensions^{41–43} to the AMBER force field⁴⁴ was one of the first applications of MM to model BXBs. In this approach, the halogen is described by the standard MM parameters for size, shape and atomic charge, with the positive σ -hole mimicked by the addition of a massless positive charge placed at or near the surface of the halogen, opposite the σ -bond (Figure 3a). Such MM approaches have been successful in simulating the geometries and trends in IC50s (to reflect the affinities) of inhibitors and drugs to various proteins, including Clathresin L, p53, and, in particular, protein kinases (reviewed in^{9, 30}).

We have approached the problem of modeling BXBs from a different perspective, by deriving an empirical force field for BXBs¹¹ that, unlike the pseudo-atom approach, is based

on the fundamental atomic properties of size, shape, and electrostatic charge (Fig. 3b). The resulting fBXB assigns an anisotropic distribution of charge across the atom surface to model the directional electrostatic potential⁴⁵ and an aspherical atomic radius to model the polar flattening⁴⁶ (the reduction in the R_{vdW} in the direction of the σ -hole, as observed in high resolution crystal structures) of a halogen. The initial application of the fBXB showed that the relative energies of Br-BXBs for two different geometries could be accurately modeled based on the AMBER force field¹¹. In the current study, we fully parameterize the fBXB , and show for the first time that a molecular mechanics model correlates in a direct one-to-one manner with the experimentally determined structure-energy relationships for all halogens commonly seen in biological systems (Cl, Br, and I). Furthermore, the parameters of the model can be readily interpreted in terms of the physicochemical properties of the halogen atom in covalent compounds, when X-bonded to an electron-rich acceptor.

II. THEORY AND METHODS

We had previously defined a set of potential energy force field functions to describe the structure-energy relationships of biomolecular halogen bonds—the fBXB —that was specific for Br¹¹. In the current study, we extend and parameterize the fBXB for the common biological halogens Cl, Br, and I.

The fBXB functions were derived based on well defined physical properties of covalently bonded halogen atoms, including the anisotropic distribution of charge across the atomic surface⁴⁵ and the aspherical shape of the atomic radius, seen as polar flattening⁴⁶; both aspects are consistent with predictions from the σ -hole model²⁵. The anisotropic charge distribution (positive for the linear approach towards the σ -hole, $\Theta_1 = 180^\circ$, and negative perpendicular to the σ -hole, $\Theta_1 = 90^\circ$) was modeled simply by defining the effective charge of the halogen (Z_X) as the angle-dependent function $Z_X = A \cos(\nu\alpha) + B$ (where A is the amplitude of the charge difference between the linear and perpendicular approaches, B defines the baseline of the cosine function, $\alpha = 180^\circ - \Theta_1$, and ν is the period of the cosine function)¹¹. The electrostatic potential for the interaction (V_{elect}) of an acceptor atom (with charge Z_A) at any angle of approach to the halogen is thus described by a standard potential (Eq. 1), where the exponent n defines the distance relationship (r) for the interaction, e^- is the standard charge of an electron, and D is the dielectric of the medium separating the atoms.

$$V_{elect} = \frac{Z_X Z_A e^-}{D r^n} \quad \text{Eq. 1}$$

In addition to the electrostatics, the fBXB incorporated polar flattening by defining the shape of covalently bonded halogens as being directional (shorter for a linear approach and larger for a perpendicular approach). The aspherical shape was manifested in an angle-dependent function that described the very short-range repulsive term of the standard Lennard-Jones (V_{LJ}) potential¹¹. The X-bond specific V_{LJ} , in a form consistent with that used in the common AMBER force field, defines an average effective van der Waals radius for the halogen ($\langle R_{vdW(X)} \rangle$) and a perturbation to the radius that is dependent on the angle

of approach, in a similar fashion as the charge of the halogen ($R_X \cos(\nu\alpha)$). The polar flattening is applied only to the repulsive term of the standard V_{LJ} , resulting in Eq. 2 for the overall interaction potential (where ϵ_X and ϵ_A are the energy contributions from the X-bond donor and acceptor, respectively, $R_{vdW(A)}$ is the standard van der Waals radius of the acceptor, and r is the distance between donor and acceptor atoms).

$$V_{LJ} = \sqrt{\epsilon_X \epsilon_A} \left[\left(\frac{R_{vdW(A)} + \langle R_{vdW(X)} \rangle - \Delta R_X \cos(\nu\alpha)}{r} \right)^{12} - 2 \left(\frac{R_{vdW(A)} + \langle R_{vdW(X)} \rangle}{r} \right)^6 \right] \quad \text{Eq. 2}$$

In this work, we derive a set of parameters for the V_{elec} and V_{LJ} functions of the $\#BxB$ that are consistent and applicable across the common halogens (Cl, Br, and I) that are seen to form stabilizing X-bonds in the DNA model system. The derivation of these parameters follow a general strategy, where we: 1) apply high-level quantum mechanical (QM) calculations on a pair of simple compounds to model the energies at various distances and angles of approach between the X-bond donor and acceptor atoms in the experimental DNA junction system; 2) determine the shape parameters for covalent Cl, Br, and I atoms; 3) determine the size and charge parameters for Cl, Br, and I atoms; and 4) validate the model system and force field parameters by comparing to BXB structure-energy relationships experimentally determined in the DNA system.

QM Calculations for BXB Interaction Energies in Models of the Experimental DNA System

We had previously shown that the BXBs in the DNA system could be mimicked by an isolated halogenated uracil base and a hypophosphate group (an $XU \cdots H_2PO_2^{-1}$ interacting pair, Fig. 2 inset)—a simple model that is amenable to high-level QM calculation¹¹. In order to map the structure-energy relationships, we constructed $XU \cdots H_2PO_2^{-1}$ pairs for X = Cl, Br, and I, maintaining an $X \cdots ^{-1/2}O-P$ alignment of 180° . For each set, the BXB geometries were defined by varying the angle of approach from $\Theta_1 = 90^\circ$ to 180° , and the interacting distances varying from 70% to 100% of the standard radius of the halogen of XU to the oxygen of the hypophosphite.

QM energies for each $XU \cdots H_2PO_2^{-1}$ geometry were calculated using the GAUSSIAN 09, applying the Møller-Plesset second order (MP2) calculations with cyclohexane as the solvent ($D = 2$, relative to a vacuum). Appropriate polarizable basis sets that include dispersion were applied to the calculations, according to the halogen (aug-cc-PVTZ for F, Cl, and Br; aug-cc-PVTZ-PP for I from the EMSL Basis Set Exchange⁴⁷). Basis set superposition errors (BSSE) were determined from a separate counterpoise gas phase calculation and directly summed into the calculated solvent phase energy.

Determining $\#BxB$ Parameters from QM calculated energies

The $\#BxB$ has seven parameters that need to be determined for the V_{elec} and V_{LJ} potential functions. The advantage of the $\#BxB$ is that these parameter describe well defined physical characteristics of each halogen and, theoretically, can be determined through global fitting of Eq. 1 and 2 to the QM energies calculated for each geometry of the $XU \cdots H_2PO_2^{-1}$ pairs for Cl, Br, and I. Unfortunately, it is clear from attempts at such global fitting that some

parameters are coupled, resulting in fitting errors of >100% in some cases. We therefore, needed to decouple some of the parameters from the f BXB functions. Our previous studies indicated that the two parameters R_X and ϵ_X in the V_{LJ} function could be determined independently from the inherent properties of the halogens themselves¹¹. Once values for R_X and ϵ_X were defined for a particular halogen, the remaining parameters could be robustly determined by the combined f BXB functions applied to the QM energies of the $^X\text{U}\cdots\text{H}_2\text{PO}_2^{-1}$ pairs. The first step in parameterization, therefore, is to determine R_X and ϵ_X for Cl, Br, and I. These parameters were derived using very high level QM calculations of the energies of the isolate halogen atoms interacting with a helium atom, with the He serving as a small neutral, non-polarizable probe^{11, 48}. We can then determine the size and shape parameters by fitting the V_{LJ} function against the QM calculated energies.

The R_X and ϵ_X for Cl, Br, and I were obtained from counterpoise-corrected CCSD(T) and Hartree Fock potential curves for $\text{X}\cdots\text{He}$, the helium either approaching the singly occupied σ p_z orbital or one of the doubly occupied π orbitals, p_x . The halogen basis sets, aug-cc-PVTZ for F, Cl, and Br, and aug-cc-PVTZ-PP for I, were augmented by single d, f, and g functions, with exponents selected to maximize the $\text{X}\cdots\text{He}$ interaction energy at a distance near the respective van der Waals minima. The aug-cc-PVTZ basis set, augmented by single p, d, and f functions, was used for helium (augmentation exponents obtained from I $\cdots\text{He}$). The selected exponents are collected in Table 1.

Once the exponents were obtained, the Hartree Fock and CCSD(T) potential curves were combined to obtain estimates for the correlation energy ($V_{\text{corr}} = V_{\text{CCSD(T)}} - V_{\text{HF}}$). A spherical estimate for correlation was obtained by forming a weighted average of the σ and π correlation curves ($1/3 \sigma + 2/3 \pi$). This spherical correlation term was added to the separate σ and π Hartree Fock curves to obtain interaction potential curve estimates. The resulting $\text{X}\cdots\text{He}$ potential curves were fit to Eq. 2, $\cos(\nu\alpha) = -1(180^\circ)$ for σ approach and $+1(90^\circ)$ for π approach. The He $\epsilon_{\text{vdW(A)}}$ and $R_{\text{vdW(A)}}$ of 0.039 kcal/mol and 1.42 Å were obtained from a He potential curve obtained with the same augmented basis (Figure 4).

With R_X and ϵ_X defined for each halogen, the remaining parameters for the f BXB were determined by fitting the combined V_{elect} and V_{LJ} functions to the QM energies for the $^X\text{U}\cdots\text{H}_2\text{PO}_2^{-1}$ pairs for each halogen type (Cl, Br, and I). A program was written in Mathematica⁴⁹ to apply a nonlinear least squares fit of the f BXB parameters (R_X and ϵ_X fixed). The initial fit to all geometries of the X-bonded pair yielded parameters with very high error and large residuals around the minimum of the potential well (for all angles and halogens); it was clear that the very large steric repulsion energies were dominating the fitting routine. Our primary interest is to accurately model the energies at and near the potential wells, where X-bonds would form; thus, we applied a weighting factor (κ) that is biased toward the negative potential energy domains of the f BXB functions. To determine how much weight to apply to the fitting of each halogen, we systematically varied κ to minimize the residuals for the negative interactions energies, averaged between $\Theta_1 = 180^\circ$ and 90° . Once optimized for each halogen, we applied κ to define a weight, $\text{Weight} = 1/(\kappa - E_{\text{QM}})^2$, to be applied to each QM energy (E_{QM}) to bias the fitting of the f BXB functions to determine the $\langle R_{\text{vdW(X)}} \rangle$, R_X and ϵ_X parameters (Figure 5).

III. RESULTS AND DISCUSSION

We had previously derived an empirical force field for biomolecular halogen bonds (fBXB), which correlated well with the relative energies of bromine X-bonds in our DNA model system. In the current study, we have expanded the fBXB to describe the anisotropic shape and surface charge distribution for all the common halogens seen in BXBs (Cl, Br, and I) and show that this molecular mechanics approach directly correlates with the absolute energies measured for specific geometries of these interactions. We have studied the interactions of F within the context of the biological system, but do not observe a stabilizing potential that can be attributed to an X-bond¹⁷ and, therefore, have not included this smallest of halogens in the current force field. The resulting fBXB parameters can be interpreted in terms of periodic trends of the atomic orbitals and their interaction energies for the common elements along the Group VII column that form stabilizing X-bonds in our system.

Our current extension of the fBXB defines parameters for Cl, Br, and I that describe electrostatic interactions as modeled by an angle-dependent Coulombic-type potential (V_{elec} , Eq. 1) and van der Waals type interactions as modeled by an angle-dependent Lennard-Jones potential (V_{LJ} , Eq. 2). The parameters were derived to be consistent with the AMBER force field (one of the most commonly used in the simulation of biomolecular structures⁴⁴). Parameters for these two classes of interactions were derived by first performing high level QM calculations on model systems that probe the size/shape of halogens using an uncharged, nonpolarizable He atom, and the charge distribution across the halogen surface using a simple model for the $\text{XU}\cdots\text{H}_2\text{PO}_2^{-1}$ pair that mimics the interactions in the experimental DNA junction system. The QM model calculations and the ability of the parameterized force field to model the structure-energy relationships of BXBs were validated by comparison to the absolute energies determined for various geometries of Cl, Br, and I X-bonds measured in the DNA system.

R_X and e_X fBXB parameters

The atomic size and shape was initially determined through MP2 calculations of each halogen probed with a He atom at various distances of approach, and at angles of approach of $\theta_I = 180^\circ$ (along the halogen σ -hole) and $\theta_I = 90^\circ$ (perpendicular to the halogen σ -hole) (Fig. 3). The distances of the average energy minima at 180° and 90° are, as expected, consistent with the standard R_{vdW} defined in the current AMBER parameters for each halogen. It is clear that all the halogens show polar flattening of their R_{vdW} , with the 180° approach being $\sim 0.15 \text{ \AA}$ and $\sim 0.2 \text{ \AA}$ shorter than the 90° approach (equivalent to 15% to 31% flattening relative to their respective AMBER radii).

By fitting the V_{LJ} potential of Eq. 1 to the QM calculated interaction energies, we derived the size and shape parameters for the fBXB for Cl, Br, and I (Table 2). We see that, as expected, the effective size ($\langle R_{vdW(X)} \rangle$) and the polar flattening (R_X) become progressively larger, but the effective depth of the energy well (e_X) remains fairly similar from Cl to I. The $\langle R_{vdW(X)} \rangle$ values are slightly larger than the respective AMBER radii (R_{AMBER}), while the e_X values are significantly smaller and follow the opposite periodic trend compared to the AMBER energies (e_{AMBER}). The discrepancy may result from the e_{AMBER} values not accounting for the contribution of X-bonds. Of these parameters, we

carried R_X and ϵ_X to the next step of deriving the ffBxB charge parameters. The $\langle R_{vdW(X)} \rangle$ will be redetermined in conjunction with the effective charges, because they appear to be codependent¹¹.

Size, Shape, and Charge ffBxB Parameters

The model system to parameterize the size and charges of each halogen is the $^X\text{U}\cdots\text{H}_2\text{PO}_2^{-1}$ pair, in which the distance between the interacting $\text{X}\cdots\text{O}$ pair and the θ_I angles of approach are varied to sample a broad range of possible BXB geometries for each halogen. The size and shape parameters of $\langle R_{vdW(X)} \rangle$ and ν (holding the values for R_X and ϵ_E fixed to those determined by He probing study, Table 2), and the charge parameters n , A , and B were fit to best model the QM calculated energies of interaction of the $^X\text{U}\cdots\text{H}_2\text{PO}_2^{-1}$ pairs (Table 3).

The resulting parameters for the ffBxB fit the QM energies for the $^X\text{U}\cdots\text{H}_2\text{PO}_2^{-1}$ geometries very well (Figure 5). Applying the parameters to V_{elec} (Eq. 1) and V_{LJ} (Eq. 2), we can calculate the BXB energy landscapes for the various halogens for the $^X\text{U}\cdots\text{H}_2\text{PO}_2^{-1}$ pair as topological maps (Figure 6). These topological geometry-energy maps show that the energy wells become deeper and broader going from Cl to Br to I, as expected. Although the breadth of the wells increase with the size of the halogen, as expected, this is not a linear increase. For example, the -1 kcal/mol stabilizing isotherm extends to ~ 3.5 Å for Cl, to ~ 4 Å for Br, but well beyond 6 Å for I. Similarly the 1 kcal/mol isotherm spans $\sim 30^\circ$ of θ_I for Cl, $\sim 70^\circ$ for Br, and 90° for I at a distance of 3.0 Å from the halogen. The maps, therefore, indicate that a BXB acceptor will see a stabilizing interaction energy over a broader range of distance and angle geometries for an I donor compared to Cl. In addition, the greater depth (and accompanying steeper gradient of the well) will draw the acceptor more towards the optimum BXB geometry for I compared to Cl, as seen in the experimental DNA system¹⁷.

The parameter n reflects the distance dependence of the V_{elec} potential, which in turn informs us of the type of electrostatic interaction involved in the BXB. For the three halogens parameterized here, the average $n = 2.54$ (S.D. = 0.11). The value of $n \approx 2.5$ suggests that X-bonds lie somewhere between a dipole-charge and dipole-dipole interaction, which accounts for its strong directionality, even compared to the H-bond. In the case of X-bonds, the dipole can be interpreted as being defined by the anisotropic charge distribution across the atomic surface of the halogen.

We can ask how well this simple $^X\text{U}\cdots\text{H}_2\text{PO}_2^{-1}$ model system, the associated QM calculations, and the resulting ffBxB parameters simulate the experimentally measured structure-energy relationships. The experimental DNA junction had uniquely allowed us to determine the explicit energies of BXBs for Cl, Br, and I¹⁷, either from a crystallographic assay or in solution by differential scanning calorimetry (DSC)⁴⁰. The DSC energies are measured as enthalpies (H_{DSC}), while the crystallographic energies (E_{crys}) can be interpreted as equivalent to the enthalpies of interaction (assuming the entropies of the competing molecular interactions in the crystals are nearly identical). The experimental energies (H_{DSC} and E_{crys}) for the various BXBs and their geometries show a linear relationship to the QM energies (E_{QM}) calculated for the $^X\text{U}\cdots\text{H}_2\text{PO}_2^{-1}$ pair, that is remarkably close to one-to-one (Figure 7, Table 4).

The experimental *vs* calculated energies are related by a slope ~ 0.9 , with a *y*-intercept $\sim$ at -0.5 kcal/mol (R^2 correlation = 0.97). Thus, the $^X\text{U}\cdots\text{H}_2\text{PO}_2^{-1}$ pair and the QM calculations on interaction energies between the components of the molecular pair are appropriate to model BXBs. The experimental energies also showed a strong linear one-to-one relationship to the *ff*BXB calculated interaction energies (near identical slope and *y*-intercept, and R^2 correlation = 0.97), attesting to the ability of the empirical model to reproduce the experimental structure-energy relationships.

Periodic Trends and Interpretation of *ff*BXB Parameters

The *ff*BXB parameters show periodic trends for the Group VII elements, which allow QM calculations to be interpreted in terms of the physical properties of the halogens and their contributions to the BXB interactions (Figure 8). As expected, the average effective van der Waals radii $\langle R_{vdW(X)} \rangle$ become larger progressing down the column from Cl to I. The $\langle R_{vdW(X)} \rangle$ derived from the *ff*BXB, however, are smaller than the standard AMBER values⁴⁴ or the He probe values. Indeed, they are closer to the Bondi radii⁵⁰, which uses an alternative definition for R_{vdW} and, thus, a V_{LJ} equation that differs from the AMBER equation. The shorter *ff*BXB $\langle R_{vdW(X)} \rangle$ distances may reflect the coupling of the shape and size of the halogen with the charge parameters¹¹.

The parameter ν defines the angle (relative to the C—X bond) at which the halogen shows the largest effective radius and most electronegative charge. Relative to the σ -hole model, ν can be interpreted as the angle at which the $p_{x,y}$ -atomic orbitals are tipped away from the x,y -plane (a tip angle τ), assuming that the σ -hole lies along the p_z -orbital. Ideally, $\tau = 0^\circ$ (for $\nu = 2.0$), that is, the $p_{x,y}$ -orbitals should lie in the x,y -plane, and the halogen should be largest and most negative perpendicular to the σ -hole. We see that in fact τ is largest for Cl and approaches the ideal value for I, or that the orientation of the $p_{x,y}$ -orbitals are most tipped for the smallest atom. A near linear relationship is observed between τ and $\langle R_{vdW(X)} \rangle$ ($\tau = [-38.1 \text{ deg/\AA}] \langle R_{vdW(X)} \rangle + 83.2^\circ$, $R^2 = 0.90$), suggesting that the further the $p_{x,y}$ -orbitals are from the σ -hole (even when the hole is more electropositive) the less they are perturbed. The τ -angle has a number of effects on the non-covalent interactions of halogens with other atoms, including the optimum angle for potential H-bonds to the negative annulus of the halogen⁵¹ and the electroneutral point across the halogen surface. Covalently bonded halogens are amphipathic, capable of serving as both X-bond donors and H-bond acceptors, H-bonds generally ascribed as being orthogonal to X-bonds when they share the same halogen atom. The τ -angle derived here suggests that this is true primarily for I, but the angular relationship deviates significantly going to Br and Cl, and may be even more shallow with a more electropositive σ -hole.

A and *B* define the charge of the atomic surface at any angle θ_j , with *A* specifying the amplitude of the difference between the most positive and most negative partial charge, and *B* specifying the charge at the midpoint of the cosine function. Not surprisingly, *A* increases with the size of the halogen, which is consistent with the larger halogens being more polarized. Similarly, *B* is most negative for Cl and becomes more positive towards I. Indeed, the parameters *A* and *B* are highly correlated with $\langle R_{vdW(X)} \rangle$ (supplemental Fig. 1). Thus,

the σ -hole is seen to be least positive ($+0.12 e^-$) for Cl and becomes more positive proceeding down the column to Br ($+0.38 e^-$) and to I ($+0.75 e^-$).

The parameters A , B , and ν allow the electroneutral angle (Θ_0) to be calculated according to

the relationship $\Theta_0 = \frac{\cos^{-1}\left(\frac{A}{B}\right)}{\nu}$. From the $\#BXB$ parameters, $\Theta_0 = 146^\circ$ for Cl, 126° for Br, and 122° for I. The Θ_0 angles reflect the angular size of the σ -hole for each halogen and, as predicted from the potential energy landscape, the range at which a BXB is electrostatically stable is broader for the larger halogens. In addition, A , B , and ν can be used to calculate the overall charge across the entire accessible surface for each halogen. The resulting surface charges ($-0.095e$ for Cl, $+0.008e$ for Br, and $0.026e$ for I) follow closely the charges calculated by the restrained electrostatic potential (RESP) approach⁵² ($-0.046e$ for Cl, $-0.0072e$ for Br, and $0.035e$ for I), a standard approach to defining atomic charges for an unknown ligand for the AMBER force field. Both methods indicate that Cl carries an overall negative charge, Br essentially neutral, and I is positive.

IV. CONCLUSIONS

In this study, we have extended the force field previously derived for Br X-bonds to be applicable to Cl, Br, and I. The empirical models were derived from QM calculations, and this general approach (both the QM models and $\#BXB$ approach) is shown here to be valid by comparing the calculated structure-energies relationships to the explicit energies measured in a DNA system. The parameters derived for the $\#BXB$ now allow us to interpret the periodic trends in terms of physical properties of the individual halogen atoms, including the size, shape, and anisotropy of the charge across the atomic surface. The X-bond is now interpreted as being more than a dipole-dipole and not quite a charge-dipole interaction, giving it a significantly shorter distance relationship and more directionality than expected for a purely Coulombic electrostatic interaction. As a shorter distance interaction, its influence would not be expected to extend much beyond $6 \text{ \AA}$ for any halogen type (being longest for I, which has the strongest σ -hole). The overall charge across the halogen surface calculated from the $\#BXB$ parameters closely matches the overall RESP charges. The strong linear correlation between the charges from the two approaches ($R = 0.92$) reflects the ability of the $\#BXB$ to model the QM calculated electrostatic potentials of BXB donors.

That being said, the shorter $\langle R_{vdW(X)} \rangle$ from the $\#BXB$ compared to the AMBER radii may reflect an inability of a purely electrostatic model to account for the short BXB distances without significantly compressing the effective atomic radius. This shortfall suggests that the purely electrostatic definition of the σ -hole model, even with dispersion taken into account, may not fully account for the structure-energy relationships of BXBs. For example, charge-transfer is not an effect that is directly modeled by the electrostatic σ -hole model. If there is a significant contribution of charge transfer to X-bonds²⁷, we would expect the transfer of an electron from the HOMO of the electron donor to the LUMO of acceptor to affect the atomic shape of both atoms, which may be reflected in the change in their size and shape and, for an MM model, the $\langle R_{vdW(X)} \rangle$ and minimum radius in the V_{LJ} potential for an X-bonded pair.

Taken in total, we have described a set of empirical potential energy functions and parameters for these functions that perform well in modeling BXBs engineered into our experimental DNA junction system. To be generally useful, the $\#$ BXB must be implemented into a workable computational method, in this case those that apply the AMBER force field for simulation—currently a project in progress. We can see, however, how the form of the $\#$ BXB functions and the periodic trends help us in this implementation. For example, the short distance relationship ($r^{2.5}$) for V_{elec} indicates that the $\#$ BXB can be cutoff at very short distances, limiting the number of possible interacting atoms that need to be considered in these simulations.

One question that can be reasonably raised is how can these parameters be derived for various halogenated compounds without having to perform extensive QM calculations on every type of BXB? Of the seven $\#$ BXB parameters, we expect the shape and size parameters ($\langle R_{vdW(X)} \rangle$, R_X , e_X , and ν) of the V_{LJ} function to remain fairly consistent for each halogen in any system, and n for V_{elec} to be 2.5 for all BXB systems. Thus, the only parameters that need to be derived are the electrostatic variables A and B . The expectation is that one or both will depend greatly on the electron withdrawing or donating ability of the molecule that the halogen is immediately attached to. We have already seen that the overall charge is strongly correlated with the RESP charges (given that we currently have only three points, it is not clear whether the relationship is strictly linear or not). An RESP calculation⁵², therefore, which is typically performed for any new molecule or compound prior to running an AMBER simulation, may provide relationships that will allow A and B to be estimated for the halogen atoms in these molecules.

We conclude, therefore that the functions and various parameters derived for these functions will be generally applicable to model the structure-energy relationships of BXBs. In particular, we see direct application of the $\#$ BXB to predict how halogenated bases can be used to control the conformation of DNA constructs, particularly the geometric relationship of the arms in DNA junctions. DNA junctions have become an important scaffolding molecule for constructing various two and three-dimensional structures⁵³, including DNA origami⁵⁴ and controllable crystal lattices⁵⁵. The ability to use BXBs now in a computationally predictable way opens additional dimensions for such molecular engineering applications. We consider the approach, however, to be general to protein systems as well, including the design of inhibitors against enzymes. Thus, the $\#$ BXB will be an important tool for many biomolecular engineering applications.

Acknowledgments

This work was funded in part by a grant from the National Science Foundation (CHE-1152494) and from Colorado State University.

VI. REFERENCES

1. Erdelyi M. Scientific Conferences: A Big Hello to Halogen Bonding. Nat Chem. 2014; 6:762–4. [PubMed: 25143209]
2. Nguyen HL, Horton PN, Hursthouse MB, Legon AC, Bruce DW. Halogen Bonding: A New Interaction for Liquid Crystal Formation. J Am Chem Soc J Am Chem Soc. 2004; 126:16–7. [PubMed: 14709037]

3. Chudzinski MG, McClary CA, Taylor MS. Anion Receptors Composed of Hydrogen- and Halogen-Bond Donor Groups: Modulating Selectivity with Combinations of Distinct Noncovalent Interactions. *J Am Chem Soc.* 2011; 133:10559–67. [PubMed: 21667941]
4. Shirman T, Kaminker R, Freeman D, van der Boom ME. Halogen-Bonding Mediated Stepwise Assembly of Gold Nanoparticles onto Planar Surfaces. *ACS Nano.* 2011; 5:6553–63. [PubMed: 21749110]
5. Priimagi A, Cavallo G, Metrangolo P, Resnati G. The Halogen Bond in the Design of Functional Supramolecular Materials: Recent Advances. *Acc Chem Res.* 2013; 46:2686–95. [PubMed: 23805801]
6. Lieffrig J, Le Pennec R, Jeannin O, Auban-Senzier P, Fourmigue M. Toward Chiral Conductors: Combining Halogen Bonding Ability and Chirality within a Single Tetrathiafulvalene Molecule. *Crystengcomm.* 2013; 15:4408–12.
7. Lu Y, Liu Y, Xu Z, Li H, Liu H, Zhu W. Halogen Bonding for Rational Drug Design and New Drug Discovery. *Expert Opin Drug Discov.* 2012; 7:375–83. [PubMed: 22462734]
8. Scholfield MR, Vander Zanden CM, Carter M, Ho PS. Halogen Bonding (X-Bonding): A Biological Perspective. *Protein Sci.* 2013; 22:139–52. [PubMed: 23225628]
9. Wilcken R, Zimmermann MO, Lange A, Joerger AC, Boeckler FM. Principles and Applications of Halogen Bonding in Medicinal Chemistry and Chemical Biology. *J Med Chem.* 2013; 56:1363–88. [PubMed: 23145854]
10. Xu Z, Yang Z, Liu Y, Lu Y, Chen K, Zhu W. Halogen Bond: Its Role Beyond Drug–Target Binding Affinity for Drug Discovery and Development. *J Chem Inf Model.* 2014; 54:69–78. [PubMed: 24372485]
11. Carter M, Rappe AK, Ho PS. Scalable Anisotropic Shape and Electrostatic Models for Biological Bromine Halogen Bonds. *J Chem Theory Comput.* 2012; 8:2461–73. [PubMed: 26588975]
12. Auffinger P, Hays FA, Westhof E, Ho PS. Halogen Bonds in Biological Molecules. *Proc Natl Acad Sci USA.* 2004; 101:16789–94. [PubMed: 15557000]
13. Zhou P, Tian F, Zou J, Shang Z. Rediscovery of Halogen Bonds in Protein-Ligand Complexes. *Mini Rev Med Chem.* 2010; 10:309–14. [PubMed: 20470246]
14. Hardegger LA, Kuhn B, Spinnler B, Anselm L, Ecabert R, Stihle M, Gsell B, Thoma R, Diez J, Benz J, et al. Systematic Investigation of Halogen Bonding in Protein-Ligand Interactions. *Angew Chem Int Ed Engl.* 2011; 50:314–8. [PubMed: 21184410]
15. Voth AR, Hays FA, Ho PS. Directing Macromolecular Conformation through Halogen Bonds. *Proc Natl Acad Sci USA.* 2007; 104:6188–93. [PubMed: 17379665]
16. Carter M, Ho PS. Assaying the Energies of Biological Halogen Bonds. *Cryst Growth Des.* 2011; 11:5087–95.
17. Carter M, Voth AR, Scholfield MR, Rummel B, Sowers LC, Ho PS. Enthalpy - Entropy Compensation in Biomolecular Halogen Bonds Measured in DNA Junctions. *Biochemistry.* 2013; 52:4891–903. [PubMed: 23789744]
18. Colin JJ. *Ann Chem.* 1814; 91:251–72.
19. Mulliken RS. Molecular Compounds and Their Spectra. ii. *J Am Chem Soc.* 1952; 74:811–24.
20. Hassel O. Structural Aspects of Interatomic Charge-Transfer Bonding. *Science.* 1970; 170:497–502. [PubMed: 17799698]
21. Hassel, O. Nobel Lectures, Chemistry, 1963–1970. Amsterdam: Elsevier publishing company; 1972. Structural Aspects of Interatomic Charge-Transfer Bonding.
22. Politzer P, Murray JS, Clark T. Halogen Bonding: An Electrostatically-Driven Highly Directional Noncovalent Interaction. *Phys Chem Chem Phys.* 2010; 12:7748–57. [PubMed: 20571692]
23. Arunan E, Desiraju GR, Klein RA, Sadlej J, Scheiner S, Alkorta I, Clary DC, Crabtree RH, Dannenberg JJ, Hobza P, et al. Definition of the Hydrogen Bond (IUPAC Recommendations 2011). *Pure Appl Chem.* 2011; 83:1637–41.
24. Desiraju GR, Ho PS, Kloo L, Legon AC, Marquardt R, Metrangolo P, Politzer P, Resnati G, Rissanen K. Definition of the Halogen Bond (IUPAC Recommendations 2013). *Pure Appl Chem.* 2013; 85:1711–3.

25. Clark T, Hennemann M, Murray JS, Politzer P. Halogen Bonding: The Sigma-Hole. Proceedings of "Modeling Interactions in Biomolecules II", Prague, September 5th–9th, 2005. *J Mol Model*. 2007; 13:291–6. [PubMed: 16927107]
26. Riley KE, Hobza P. Investigations into the Nature of Halogen Bonding Including Symmetry Adapted Perturbation Theory Analyses. *J Chem Theory Comput*. 2008; 4:232–42. [PubMed: 26620655]
27. Wang C, Danovich D, Mo Y, Shaik S. On the Nature of the Halogen Bond. *J Chem Theory Comput*. 2014; 10:3726–3737. [PubMed: 26588518]
28. Riley KE, Murray JS, Fanfrlik J, Rezac J, Sola RJ, Concha MC, Ramos FM, 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; 17:3309–18. Epub 2011/03/04. [PubMed: 21369930]
29. Wilcken R, Zimmermann MO, Lange A, Zahn S, Boeckler FM. Using Halogen Bonds to Address the Protein Backbone: A Systematic Evaluation. *J Comput Aid Mol Des*. 2012; 26:935–45.
30. Voth AR, Ho PS. The Role of Halogen Bonding in Inhibitor Recognition and Binding by Protein Kinases. *Curr Top Med Chem*. 2007; 7:1336–48. Epub 2007/08/19. [PubMed: 17692024]
31. Matter H, Nazare M, Gussregen S, Will DW, Schreuder H, Bauer A, Urmann M, Ritter K, Wagner M, Wehner V. Evidence for C-Cl/C-Br...Pi Interactions as an Important Contribution to Protein-Ligand Binding Affinity. *Angew Chem Int Ed Engl*. 2009; 48:2911–6. [PubMed: 19294721]
32. Voth AR, Khuu P, Oishi K, Ho PS. Halogen Bonds as Orthogonal Molecular Interactions to Hydrogen Bonds. *Nat Chem*. 2009; 1:74–9. [PubMed: 21378804]
33. El-Sheshtawy HS, Bassil BS, Assaf KI, Kortz U, Nau WM. Halogen Bonding inside a Molecular Container. *J Am Chem Soc*. 2012; 134:19935–41. [PubMed: 23137118]
34. Takemura A, McAllister LJ, Hart S, Pridmore NE, Karadakov PB, Whitwood AC, Bruce DW. Halogen- and Hydrogen-Bonded Salts and Co-Crystals Formed from 4-Halo-2,3,5,6-Tetrafluorophenol and Cyclic Secondary and Tertiary Amines: Orthogonal and Non-Orthogonal Halogen and Hydrogen Bonding, and Synthetic Analogues of Halogen-Bonded Biological Systems. *Chem-Eur J*. 2014; 20:6721–32. [PubMed: 24796890]
35. Lam, PYS.; Clark, CG.; Smallwood, AM.; Alexander, RS. Structure-Based Drug Design Utilizing Halogen Bonding: Factor Xa Inhibitors. In: Metrangolo, P.; Resnati, G., editors. The 238th ACS National Meeting; Washington, DC. ACS; 2009. p. 58
36. Hernandez MZ, Cavalcanti SM, Moreira DR, de Azevedo WF, Leite AC. Halogen Atoms in the Modern Medicinal Chemistry: Hints for the Drug Design. *Curr Drug Targets*. 2010; 11:303–14. [PubMed: 20210755]
37. Zimmermann MO, Lange A, Wilcken R, Cieslik MB, Exner TE, Joerger AC, Koch P, Boeckler FM. Halogen-Enriched Fragment Libraries as Chemical Probes for Harnessing Halogen Bonding in Fragment-Based Lead Discovery. *Future Med Chem*. 2014; 6:617–39. [PubMed: 24895892]
38. Wilcken R, Liu X, Zimmermann MO, Rutherford TJ, Fersht AR, Joerger AC, Boeckler FM. Halogen-Enriched Fragment Libraries as Leads for Drug Rescue of Mutant P53. *J Am Chem Soc*. 2012; 134:6810–8. [PubMed: 22439615]
39. Eichman BF, Vargason JM, Mooers BHM, Ho PS. The Holliday Junction in an Inverted Repeat Sequence: Sequence Effects on the Structure of Four-Way Junctions. *Proc Natl Acad Sci USA*. 2000; 97:3971–6. [PubMed: 10760268]
40. Vander Zanden CM, Carter C, Ho PS. Determining Thermodynamic Properties of Molecular Interactions from Single Crystal Studies. *Methods*. 2013; 64:12–8. [PubMed: 23933330]
41. Ibrahim MA. Molecular Mechanical Study of Halogen Bonding in Drug Discovery. *J Comput Chem*. 2011; 32:2564–74. [PubMed: 21598284]
42. Ibrahim MA. 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; 116:3659–69. [PubMed: 22393912]
43. Kolar M, Hobza P. On Extension of the Current Biomolecular Empirical Force Field for the Description of Halogen Bonds. *J Chem Theory Comput*. 2012; 8:1325–33. [PubMed: 26596748]

44. Case DA, Cheatham TE 3rd, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ. The Amber Biomolecular Simulation Programs. *J Comput Chem.* 2005; 26:1668–88. Epub 2005/10/04. [PubMed: 16200636]
45. Murray JS, Riley KE, Politzer P, Clark T. Directional Weak Intermolecular Interactions: Sigma-Hole Bonding. *Aust J Chem.* 2010; 63:1598–607.
46. Nyburg S. Polar Flattening: Non-Spherical Effective Shapes of Atoms in Crystals. *Acta Cryst.* 1979; A35:641–5.
47. Peterson KA, Shepler BC, Figgen D, Stoll H. On the Spectroscopic and Thermochemical Properties of ClO, BrO, IO, and Their Anions. *J Phys Chem A.* 2006; 110:13877–83. [PubMed: 17181347]
48. Peebles SA, Fowler PW, Legon AC. Anisotropic Repulsion in Complexes B...Cl-2 and B...HCl - the Shape of the Chlorine Atom-in-a-Molecule. *Chem Phys Lett.* 1995; 240:130–4.
49. Wolfram Research, I. Mathematica, Version 9.0. Champaign, IL: 2012.
50. Bondi A. Van Der Waals Volumes and Radii. *J Phys Chem A.* 1964; 68:441–51.
51. Aakeröy C, Despera J, Helfrich BA, Metrangolo P, Pilatic T, Resnati G, Stevenazzi A. Combining Halogen Bonds and Hydrogen Bonds in the Modular Assembly of Heteromeric Infinite 1-D Chains. *Chem Commun.* 2007:4236–8.
52. Wang JM, Cieplak P, Kollman PA. How Well Does a Restrained Electrostatic Potential (Resp) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *J Comput Chem.* 2000; 21:1049–74.
53. Seeman NC. Structural DNA Nanotechnology: An Overview. *Methods Mol Biol.* 2005; 303:143–66. [PubMed: 15923682]
54. Rothemund PW. Folding DNA to Create Nanoscale Shapes and Patterns. *Nature.* 2006; 440:297–302. [PubMed: 16541064]
55. Paukstelis PJ, Nowakowski J, Birktoft JJ, Seeman NC. Crystal Structure of a Continuous Three-Dimensional DNA Lattice. *Chem Biol.* 2004; 11:1119–26. [PubMed: 15324813]

Figure 1.

The halogen bond. (a) The geometry of a halogen bond (X-bond) relative to a hydrogen bond (H-bond), adapted from¹². The X-bond is similar to an H-bond, both defined by distances between the halogen donor (X) and the electron-rich acceptor (A) that is shorter than the sum of their standard van der Waals radii (ΣR_{vdW}). The X-bond is more directional, with the approach of the acceptor to C—X bond defined by the angle θ_1 and the halogen to acceptor by θ_2 . (b) The σ -hole model of halogen bonds. The formation of a single C—X covalent bond (σ -bond, yellow) pairs an electron from the carbon with one from the valence p_z -orbital of the halogen. The depopulation of the p_z -orbital creates an electropositive crown (blue) and flattening of the atomic radius opposite the σ -bond, while the $p_{x,y}$ -orbitals remain fully occupied, resulting in an electronegative annulus perpendicular to the covalent bond. (c) Electrostatic potential of halogen substituents at the 5-position of a uracil base (adopted from¹²). The DFT calculated electrostatic potentials (from $>+25$ kcal/mol, in blue, to <-25 kcal/mol, in red) for the various 5-halouracils (XU), show σ -holes that increase in size and intensity as the size of the halogen increases from F to Cl to Br to I.

Figure 2.

DNA junction system to determine experimental BXB energies. A four-stranded DNA junction³⁹ construct was designed to compete a stabilizing H-bond from the amino group of a cytosine base to the phosphate oxygen of the junction cross-over (orange strands) against X-bonds from an ^XU base to the analogous phosphate (cyan strands)¹⁵. Experimental energies for the chlorinated C11J (one Cl BXB) and C12J (two Cl BXBs), brominated Br1J (one Br BXB) and Br2J (two Br BXBs), and the iodinated I2J (two I BXBs) constructs of the four-stranded DNA junction system were determined by a crystallographic assay or by differential scanning calorimetry (DSC) in solution^{17, 40}. The geometries of each construct were determined by X-ray crystallography. The specific X-bonded interacting groups are modeled by an ^XU...H₂PO₂⁻¹ (hypophosphite) pair¹¹ for high-level QM calculations (MP2 calculations, applying the aug-cc-PVTZ basis set for F, Cl, and Br; and aug-cc-PVTZ-PP for I, with BSSE correction).

Figure 3.

Molecular mechanics (MM) approaches to model BXBs. (a) positive extra point or explicit σ -hole models incorporate a massless positive pseudoatom at^{41, 42} or near the surface⁴³ of a standard electronegative and spherical halogen (with a standard R_{vdw}) to mimic the positive crown of the σ -hole. The force field for biological X-bonds ($fBXB^{11}$) model assumes that the size of the halogen varies by R_X from an average radius ($\langle R_{vdw} \rangle$) in an angle dependent manner (with the bulge tipped by an angle τ away from 90° relative to the C—X bond), and the charge of the halogen similarly dependent on the angle of approach, varying from positive opposite the C—X bond (blue) to neutral (green) to negative perpendicular to the C—X bond (red).

Figure 4.

Size and shape of Cl (a), Br (b), and I (c). QM energies for each halogen (X), probed with a helium (He) atom, were calculated along the σ - (solid diamonds, 180°) and π - (open squares, 90°) directions. The QM calculated energies at various distances were used to determine the $\langle R_{vdW(X)} \rangle$, R_X , and ϵ_X parameters using the V_{LJ} function (Eq. 2) for the σ - (dashed curves) and π - (solid curves) directions.

Figure 5.

Parameterizing the f/BXB against QM calculated interaction energies (E_{int}) for Cl (a) Br (b), and (c). The QM E_{int} were calculated for $XU \cdots H_2PO_2^{-1}$ pairs, with various distances separating donors and acceptors (data points) and, for each distance, at angles of approach (Θ_i) from 90° to 180°. The E_{int} were calculated using the parameterized f/BXB are shown as curves.

Figure 6.

Potential energy landscapes calculated with the parameterized $\#BXB$. Energies were calculated for an acceptor atom (carrying a $-0.4e^-$ charge) to Cl (a), Br (b), and I (c). Energies range from < -5.0 kcal/mol to $> +5.0$ kcal/mol, from white to black (see color scale insert). The radial distances are shown as semicircular arcs spanning angles of approach from 90° to 180° . The Cl1J, Cl2J, Br1J, Br2J, and I2J BXBs geometries determined from the DNA junction system^{15, 17} are mapped onto the landscapes. The zero-energy isoenergetic curve is labeled.

Figure 7.

Correlation of experimental energies (E_{Exp}) to QM (E_{QM} , **a**) and *f*BXB (E_{fBxB} , **b**) for BXBs. Experimental energies determined from crystallographic competition assays (open squares) or DSC measurements (filled squares) on DNA junctions are compared to energies calculated by QM or the *f*BXB functions for the model $^XU \cdots H_2PO_2^{-1}$ pairs. Geometries for the $^XU \cdots H_2PO_2^{-1}$ were taken from the crystal structures, see Table 4. The linear relationships show $E_{Exp} = 0.921 E_{QM} - 0.52$ kcal/mol for the QM energies ($R^2 = 0.97$), and $E_{Exp} = 0.916 E_{fBxB} - 0.46$ kcal/mol for the QM energies ($R^2 = 0.97$).

Figure 8.

Atomic interpretation of the charge and shape $\mathcal{H}\mathcal{B}\mathcal{X}\mathcal{B}$ parameters for Cl (a), Br (b), and I (c). The electrostatic charge parameters A and B define the distribution of charges (top row, scaled from $< -0.1 e^-$ in red to $> +0.3 e^-$). The bottom row shows the effect of the shape parameter R_X in flattening of the effective $R_{vdW(X)}$ relative to the expected size of the p_z atomic orbital, and the effect of ν on the tipping (τ angle) of the $p_{x,y}$ -orbitals relative to the standard cartesian coordinate plane.

Table 1

Dispersion augmentation exponents for Cl, Br, I, and He.

Halogen	p	d	f	g
Cl		0.0300	0.1200	0.1700
Br		0.1200	0.2000	0.1400
I		0.1600	0.3700	0.2800
He	0.2400	0.2900	0.3300	

Table 2

Size and shape parameters for Cl, Br, and I. The average van der Waals radius ($\langle R_{vdW} \rangle$), perturbation to the radius (R_X), and depth of the energy well (ϵ_X) were calculated by fitting the V_{LJ} to the QM calculated energies for a He probe of each halogen. The resulting $\langle R_{vdW} \rangle$ and ϵ_X parameters are compared to the radius and energy well (R_{AMBER} and ϵ_{AMBER} , respectively) assigned to each halogen in the AMBER99⁴⁴ parameter file.

Halogen	$\langle R_{vdW} \rangle$ (Å)	R_{AMBER} (Å)	R_X (Å)	ϵ_X (kcal/mol)	ϵ_{AMBER} (kcal/mol)
Cl	2.116	1.95	0.149 ± 0.001	0.107 ± 0.002	0.265
Br	2.369	2.22	0.160 ± 0.018	0.109 ± 0.038	0.320
I	2.658	2.35	0.185 ± 0.006	0.087 ± 0.008	0.400

The fBXXB parameters for Cl, Br, and I. The parameters for size and shape ($\langle R_{vdW}(X) \rangle$, ν , and τ) and charge (n , A , and B) were determined by fitting the V_{LJ} (Eq. 2) and V_{elect} (Eq. 1) functions to QM calculated energies for various geometries of the $\text{XU} \cdots \text{H}_2\text{PO}_2^{-1}$ pairs (see text for definition of parameters).

Table 3

Halogen	Size and Shape Parameters			Charge Parameters		
	$\langle R_{vdW}(X) \rangle$ (Å)	ν	τ	n	A (e^-)	B (e^-)
Cl	1.719 ± 0.010	2.46 ± 0.05	16.9°	2.40 ± 0.19	0.14 ± 0.03	-0.016 ± 0.005
Br	1.817 ± 0.014	2.42 ± 0.04	15.5°	2.79 ± 0.21	0.23 ± 0.05	0.15 ± 0.04
I	1.922 ± 0.015	2.23 ± 0.04	9.2°	2.44 ± 0.15	0.46 ± 0.06	0.29 ± 0.06

Table 4

Comparison of experimental BXB energies to QM and ffBXB calculated energies. Experimental energies for the various constructs in the four-stranded DNA junction (Fig. 2) system were determined by from a crystallographic assay (E_{exp}) or by DSC in solution (H_{exp}). The QM (E_{QM}) and ffBXB (E_{ffBXB}) energies were calculated based on the $^{\text{X}}\text{U}\cdots\text{H}_2\text{PO}_2^{-1}$ pair fixed at the crystallographically determined geometries of each construct, defined by the distance between the donors and acceptors ($r_{(\text{X}\cdots\text{O})}$) and the angle of approach of the acceptor to the donor (θ_{I}).

Structure (PDB code)	$r_{(\text{X}\cdots\text{O})}$	θ_{I}	$E_{\text{exp}}/H_{\text{exp}}$	E_{QM}	E_{ffBXB}
CI1J (4GSG)	2.88 Å	157.0°	-0.79 ± 0.12	-0.79	-0.76
CI2J (4GQD)	2.88 Å	146.0°	-0.79 ± 0.12	0.02	-0.01
Br1J (2ORF)	3.32 Å	167.2°	-2.28 ± 0.11	-2.07	-2.41
Br2J (2ORG)	2.87 Å	163.2°	-3.6 ± 1.3	-2.93	-2.94
I2J (4GRE)	3.01 Å	170.7°	-5.9 ± 1.1	-5.93	-5.96