JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

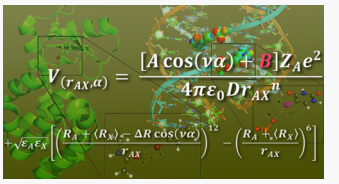
A Reduced Generalized Force Field for Biological Halogen Bonds

Melissa Coates Ford, Anthony K. Rappé, and P. Shing Ho*

Cite This: https://doi.org/10.1021/acs.jctc.1c00362 Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: The halogen bond (or X-bond) is a noncovalent interaction that is increasingly recognized as an important design tool for engineering protein−ligand interactions and controlling the structures of proteins and nucleic acids. In the past decade, there have been significant efforts to characterize the structure−energy relationships of this interaction in macromolecules. Progress in the computational modeling of X-bonds in biological molecules, however, has lagged behind these experimental studies, with most molecular mechanics/dynamics-based simulation methods not properly treating the properties of the X-bond. We had previously derived a force field for biological X-bonds (*ff*BXB) based on a set of potential energy functions that describe the anisotropic electrostatic and shape properties of halogens participating in X-bonds. Although fairly accurate for reproducing the energies within biomolecular systems, including X-bonds engineered into a DNA junction, the *ff*BXB with its seven variable parameters was considered to be too unwieldy for general applications. In the current study, we have generalized the *ff*BXB by reducing the number of variables to just one for each halogen type and show that this remaining electrostatic variable can be estimated for any new halogenated molecule through a standard restricted electrostatic potential calculation of atomic charges. In addition, we have generalized the *ff*BXB for both nucleic acids and proteins. As a proof of principle, we have parameterized this reduced and more general *ff*BXB against the AMBER force field. The resulting parameter set was shown to accurately recapitulate the quantum mechanical landscape and experimental interaction energies of X-bonds incorporated into DNA junction and T4 lysozyme model systems. Thus, this reduced and generalized *ff*BXB is more readily adaptable for incorporation into classical molecular mechanics/dynamics algorithms, including those commonly used to design inhibitors against therapeutic targets in medicinal chemistry and materials in biomolecular engineering.



INTRODUCTION

Halogenated compounds are abundant pharmaceutical agents, constituting 50% of the leading drugs currently on the market.[1] Historically, halogens have been incorporated into inhibitors to improve their absorption, distribution, metabolism, or excretion properties, such as drug metabolism or bioavailability.[2,3] Recently, however, halogens have been shown to form a favorable noncovalent interaction called the halogen bond or X-bond,[4] which increases the affinity and specificity of inhibitors against their protein targets.[5−10] More recently, X-bonding nucleic and amino acids have been engineered into biomolecular systems to control the conformation of a DNA junction[11] and the stability and activity of enzymes.[12,13]

To exploit the X-bond concept in the rational design of therapeutics[14−18] and in biomolecular engineering,[19] the interaction must be properly modeled in molecular simulation programs.[20−22] Here, the force field for biological X-bonds (*ff*BXB)[23,24] that we had previously derived has been reparameterized to be more readily adaptable to standard molecular modeling algorithms. In particular, the seven variable parameters of the original *ff*BXB have been reduced to a single easily calculated parameter for each halogen type. This reparameterized force field has been shown to be generally applicable to model X-bonds in both nucleic acids and proteins, resulting in a reduced and generalized *ff*BXB (the rg*ff*BXB).

The application of computation in rational design has found success in the development of biologics for catalysis,[25−27] biofuels,[28] and drugs.[29] Examples of computational tools that have been used in biomolecular design include (in increasing order of computational complexity and cost) molecular docking, molecular mechanics/dynamics (MD/MD), and quantum mechanics (QM) algorithms. Of these methods, QM accurately models the unique atomistic properties that describe the X-bond, although machine learning tools such as ANI can potentially rival the ability of high-level QM in simulating nonclassical noncovalent interactions, including those that involve sulfur or halogen atoms.[30]

To understand the challenge in properly modeling X-bonds using classical molecular mechanics/dynamics (MM/MD) simulation functions, we must first understand the nature of the X-bond. Although there are several competing physical

Received: April 12, 2021

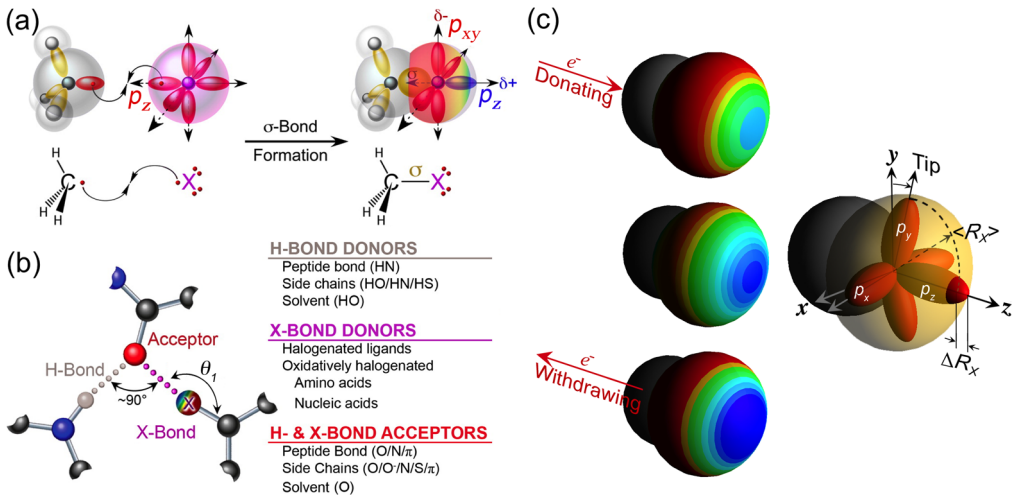


Figure 1. The halogen bond (X-bond). (a) σ-Hole theory[31] describing the formation of an electropositive crown and associated polar flattening (the σ-hole) resulting from depopulation of the p_z molecular orbital of a covalent σ-bonded halogen substituent to a carbon atom (adapted from Scholfield et al.[32]). (b) Relationship between X-bonds and hydrogen bonds (H-bonds). X- and H-bonds differ in their respective donors (halogen σ-hole for the X-bond and polarized hydrogen for the H-bond) but share common acceptors. When an acceptor is both H- and X-bonded, the two interactions are orthogonal (both geometrically perpendicular and energetically independent of each other).[33] (c) The size of the σ-hole and, consequently, the X-bonding potential of a halogen substituent increases with the size of the halogen (F < Cl < Br < I) and with the electron-withdrawing ability of the molecule that the halogen is bonded to (left figures). The shape and size parameters (perturbation to R_{vdW}: ΔR_X; average radius after polar flattening: $\langle R_X \rangle$; and the tip of the p_{xy} orbitals) of the original *ff*BXB are shown to the right.

descriptions for the X-bond,[34,35] the most readily accessible one comes from the electrostatic σ-hole theory.[31] In this model, the p_z atomic orbital of a halogen becomes depopulated, relative to the p-orbital average of 1.7 electrons, as its valence electron is subsumed by the σ-molecular orbital in forming a covalent bond to, for example, a carbon atom (Figure 1a). This bonding constraint creates a so-called σ-hole—characterized by an electropositive crown and flattening of the halogen's atomic radius at a point diametrically opposed to the σ-bond. Consequently, this σ-hole allows halogens to interact with a variety of electron-rich atoms that serve as X-bond acceptors (Figure 1b). The size of the σ-hole, which helps define the strength of the X-bond, increases with the size and, thus, the polarizability of the halogen, as well as the electron-donating/withdrawing ability of the atom or molecule to which the halogen is bound (Figure 1c). With this relatively straightforward model in mind, we can now understand why properly modeling X-bonds has been particularly challenging in standard MM/MD-based algorithms.

With the acceptance that X-bonds can affect folding and recognition, there have been several focused efforts to properly account for the interaction in biomolecular systems. The standard force fields used in MM/MD algorithms treat halogens as isotropic spheres with uniform charge distributions and shapes; thus, they cannot account for any aspect of halogens that allows them to form a stabilizing X-bond to an electron-rich acceptor. QM calculations can recapitulate the anisotropic properties of halogens and, consequently, accurately model X-bonds but at computational costs that make them impractical for studying biomolecules.

There has been some success seen with hybrid QM/MM approaches to help with inhibitor design,[1,16] but most efforts have been put into incorporating X-bonds into class I force fields of classical MM/MD algorithms. One approach to this effort is the positive extra point (PEP) strategy, where a massless pseudoatom with a defined positive charge is placed on or near the surface of the halogen to mimic the electropositive crown of the σ-hole.[20,21] This strategy has also seen some success in reproducing the geometries and relative energies of ligand interactions with proteins and thus has been incorporated into various class I force fields, including AMBER,[20] CHARMM,[36] and OPLS3.[37] The drawbacks are that the PEP parameters for any new ligand to be studied must be defined through QM calculations and do not account for polar flattening, which affects the important contribution of dispersion in defining the angular dependence of X-bonds.[38] Finally, a multipole (MTP) electrostatic approach has been implemented in CHARMM for X-bonding, which has resulted in a vast improvement in energies but at a much greater computational cost with slowdowns of 8−10-fold depending on the number of MTP pairs involved.[39]

With these challenges in mind, we had previously derived a set of potential energy functions that model the anisotropic electrostatic and shape properties of halogens by defining their charge distributions and van der Waals radii as being angle-dependent.[23] The resulting *ff*BXB could accurately reproduce the experimental energies for X-bonds determined from a set of studies on model DNA junctions[24,40] and has been incorporated into the Autodock molecular docking program.[41] A more general implementation of the *ff*BXB, however, has been limited by its apparent unwieldiness, including the seven independent parameters for each halogen substituent in the model and, as with the PEP, the need for high-level QM calculations to derive parameters for each new halogenated compound.

In this study, we have reparameterized the *ff*BXB, showing that the shape and nearly all the electrostatic parameters are relatively constant within each halogen class, leaving now only a single independent electrostatic variable that needs to be determined for any new compound or system. Furthermore, we show that this remaining variable can be readily assigned through a standard restrained electrostatic potential (RESP) calculation that is commonly used to parameterize atomic

https://doi.org/10.1021/acs.jctc.1c00362
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

charges for previously uncharacterized compounds prior to classical MM/MD simulations.

RESULTS

Reduced and General *ff*BXB Parameters. The original *ff*BXB (Scheme 1) was derived with seven independent variables for each halogen type (Cl, Br, or I) and was parameterized against QM energy landscapes constructed around X-bonds engineered into DNA junctions (Figure 2a).[23,24] The QM landscapes sampled a broad range of distances and angles of approach of an anionic hypophosphite ($O_2PH_2^{-1}$) X-bond acceptor toward halogenated uracil bases (^{X}U) as the X-bond donors (Figure 2a inset). These QM landscapes and the minimal model X-bond pairs were validated against the experimental geometries and energies determined for X-bonds in the halogenated DNA junctions. The *ff*BXB parameters were then globally fit to best reproduce the entire range of geometries sampled by the QM landscapes. The resulting parameterized *ff*BXB was also shown to accurately reproduce the experimental structure−energy relationships, thereby validating the force field and its parameters.

Although accurate, the *ff*BXB with its seven independent parameters is too unwieldy to be readily incorporated into classical molecular simulation algorithms. In addition, the original set of parameters was specifically against a DNA system, while most biological X-bonds are from halogenated ligands bound in protein environments. Hence, the first goal of the current study is to define a set of general parameters that are more broadly applicable to both nucleic acids and proteins, in which a single parameter is left as an independent variable in Scheme 1 for each halogen. The strategy for reducing the number of parameters in the *ff*BXB started with the assumption that the shape parameters can be fixed for any particular halogen type. We can then determine whether the electrostatic parameters can be derived to accommodate the overall features of a particular X-bonding interaction.

In this first step, we reconsidered the shape parameters (Figure 1c) that defined the average van der Waals radius ($\langle R_X \rangle$), the perturbation to that radius due to polar flattening (ΔR_X), and the halogen-specific component of the van der Waals energy (ε_X). Since these parameters were determined explicitly by probing the molecular halogens Cl_2, Br_2, or I_2 with a small nonpolarizable helium atom, their values for each halogen type could already be defined as being fixed. Although it has been shown that the degree of polar flattening and, consequently, the effective size and shape of an X-bond donor are dependent on the degree of polarization,[43] we propose here that the deviations in ΔR_X for each halogen type can be largely compensated for by the refined electrostatic parameters, as

Scheme 1. Potential Energy Equation of the *ff*BXB[a]

$$V_{(r_{AX},\alpha)} = \frac{[A\cos(\nu\alpha) + B]Z_A e^2}{4\pi\varepsilon_0 D {r_{AX}}^n} + \sqrt{\varepsilon_A \varepsilon_X}\left[\left(\frac{R_A + \langle R_X\rangle - \Delta R\cos(\nu\alpha)}{r_{AX}}\right)^{12} - \left(\frac{R_A + \langle R_X\rangle}{r_{AX}}\right)^{6}\right]$$

[a]The *ff*BXB potential (V) is dependent on the distance (r_{AX}) between the X-bond donor (X) and acceptor (A) and the angle of approach of the acceptor to the donor (defined here as $\alpha = 180° - \theta_1$, see Figure 1b). The first component is a classical Coulombic potential, with Z_A being the formal charge of the acceptor, e being the electron charge, D being the dielectric constant applied against the permittivity of a vacuum ($4\pi\varepsilon_o$), and n being the power of the distance relationship. The electrostatic term of the X-bond donor is a function of a polarization parameter (A), the overall "charge" of the halogen (B) and a cosine function of the angle α, and the phase of this function (ν). The second component is a 6−12 Lennard-Jones-type function, with R_A being the van der Waals radius, ε_A being the potential energy minimum of the X-bond acceptor, and ε_X being the potential energy minimum of the X-bond donor, as defined by the classical force field. The effective radius of the donor is defined by an average radius for the halogen ($\langle R_X\rangle$) and a deviation from this average (ΔR), the latter being scaled by the same angle-dependent cosine function applied to the donor charge in the electrostatic term.

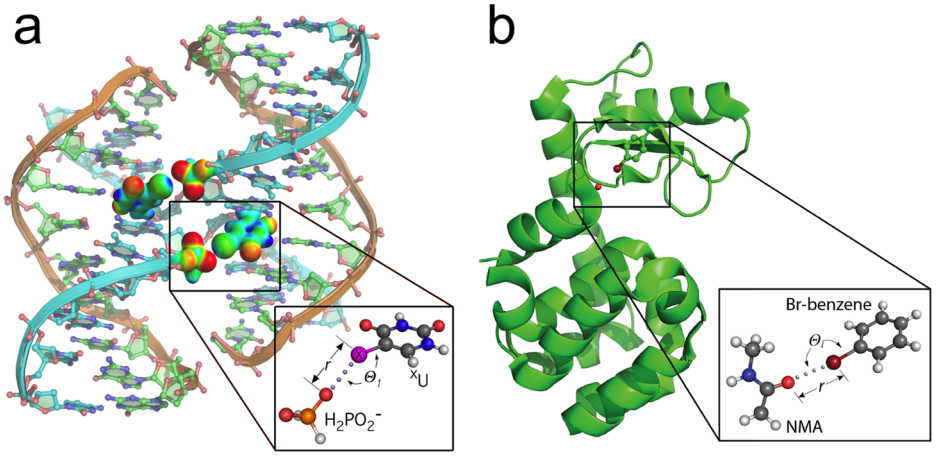


Figure 2. Biomolecular model systems to determine experimental structure−energy relationships of X-bonds. The insets show the geometries of X-bonds (the distance, r, between the oxygen acceptor and halogen donor and the angle of approach of the acceptor to the σ-hole of the halogen, θ_1) for simplified model X-bonding pairs used in QM calculations of their respective biomolecular system. X-bond energies were determined by comparing the thermodynamic parameters from differential scanning calorimetry (DSC) for melting the biomolecule with and without a halogenated nucleic acid or amino acid. (a) X-bond stabilization of four-stranded DNA junctions.[11,40,42] In this system, a halogenated uracil (^{X}U) stabilizes the junction through an X-bond to the phosphate oxygen of the DNA backbone. The X-bonding partners were reduced to an ^{X}U base and a hypophosphite ($H_2PO_2^{-1}$) as minimal models. (b) X-bond stabilization of T4 lysozyme.[12] An X-bonding halogen replaces a H-bonding hydroxyl group by substituting a halogenated phenylalanine (^{X}Phe) for a tyrosine in the protein. The minimal interacting pair is between the carbonyl oxygen of *N*-methylacetimide (modeling the peptide backbone of the protein as the acceptor) and a halogenated benzene (modeling the side chain of the ^{X}Phe).

https://doi.org/10.1021/acs.jctc.1c00362
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

described below. Thus, the reduced and generalized *ff*BXB really has only one variable shape parameter, ν. The parameter ν was defined as the period of the cosine function, reflecting the angular orientation of the p_{xy} atomic orbitals of the halogen relative to the σ-hole, which in turn defines the position of the most negative electrostatic potential of the polarized halogen. The non-integer value of $\nu > 2$ derived for the fully parameterized *ff*BXB indicated that the p_{xy} orbitals were slightly tipped (9−17°) away from the theoretical 90° relative to the σ-bond. However, these deviations were considered to be minor details that would not dramatically affect the X-bonding potential of a halogen and, therefore, the value of ν was fixed to an integer value of 2 to explicitly position the electronegative annulus at the waist perpendicular to the σ-hole. In doing so, the derivatives for the potential energy functions in Scheme 1 are greatly simplified. The ΔR_X parameters for Cl, Br, and I (Table 1) are in the same range as the values for polar flattening determined by QM calculations,[43] while the van der Waals energy terms (ε_X) of the rg*ff*BXB are similar to those of the original *ff*BXB—in both cases, smaller than those of the general AMBER force field (GAFF).[44] The effective van der Waals radii ($\langle R_{vdW} \rangle$) in the rg*ff*BXB are 13−19% smaller than that in the GAFF, which contrast with those in the original *ff*BXB, which were 6−11% larger.

Table 1. Size/Shape Parameters for the rg*ff*BXB (Scheme 1) for the Heavy Halogens (Cl, Br, and I) That Are Most Commonly Seen in Biological X-Bonds (Standard R_{vdW} and ε_X Parameters from the General Amber Force Field Are Shown in Parentheses in the Respective Columns)

halogen (X)	$\langle R_{vdW} \rangle$ (Å)	ΔR_X (Å)	ε_X (kcal/mol)	ν
Cl	1.687 (1.948)	0.15	0.107 (0.265)	2
Br	1.798 (2.22)	0.16	0.110 (0.320)	2
I	1.918 (2.35)	0.19	0.087 (0.400)	2

We next tackled the parameters that define the charge distribution across the atomic surface of the halogen. The parameter n, which reflects the power of the distance dependence and defines the type of the electrostatic interaction that is an X-bond, was seen from the fully parameterized *ff*BXB to fall between 2 and 3. As with ν, we fixed n to be the integer 2, which explicitly defines the X-bond as a charge−dipole-type interaction.

The A parameter in Scheme 1 defines the amplitude difference between the positive σ-hole and negative annulus around the waist of the halogen, while B defines the overall charge of a halogen substituent. Together, A and B define the angle at which the electrostatic potential passes from being positive to negative (the neutral-point angle). With n set explicitly to 2 above, the electrostatic component of the *ff*BXB equation in Scheme 1 is not entirely Coulombic. The units for both A and B are thus in units of length and not unitless, as would be the case for a standard Coulombic potential. For this study, the electrostatic parameters were derived for distances in units of Å (10^{-10} m). We expect both A and B to be dependent on the inductive effects of the atom or molecule that the halogen is covalently bonded to. The σ-hole and, consequently, the X-bonding potential of the halogen are enhanced by electron-withdrawing groups and diminished by electron-donating groups. We thus constructed a set of halogenated benzene compounds, in which various substituents (their electron-donating to electron-withdrawing potential defined by Hammett constants) were placed *para* to the halogen, and calculated the energies of interactions (E_{int}) with the anionic $O_2PH_2^{-1}$ and now also with *N*-methylacetimide (NMA) as a neutral X-bond acceptor. The distances and angles of approach between the X-bond donors and acceptors were varied to develop a QM energy landscape, and as before, we used this landscape to parameterize the *ff*BXB. In this case, we kept the values for all of the shape parameters and n fixed, as described above, fitting only A and B.

The results showed that A remained fairly consistent within each halogen type. These results suggest that the amplitude difference between positive and negative regions of a halogen is defined by and is thus inherent to its polarizability, which in turn is defined by its atomic size. The overall electrostatic potential and the neutral-point angle are variable and are defined by the extent to which the p_z orbital of the halogen is depopulated. We could, therefore, fix A to an average value within each halogen type (Table 2), as determined from the QM landscape, now leaving B as the only variable parameter of the *ff*BXB. For simplicity, we will call this reduced and more general force field the rg*ff*BXB. The A parameters of the rg*ff*BXB are smaller by 14−28% compared to those of the original *ff*BXB.

Table 2. Fitted Electrostatic A Parameter for the rg*ff*BXB (Scheme 1) as a Function of Increasing Electron-Donating Substituents (Y) in Model Donor-to-Acceptor ($H_2PO_2^{-1}$ or *N*-Methylacetimide, NMA) X-Bonding Interactions[b]

		X = Cl		X = Br		X = I	
donor		$H_2PO_2^{-1}$	NMA	$H_2PO_2^{-1}$	NMA	$H_2PO_2^{-1}$	NMA
X-uracil		0.120		0.144		0.330	
Y−	σ_p[a]	donor = *para*-Y-halobenzene					
CN−	0.66	0.124	0.144	0.180	0.194	0.325	0.403
Cl−	0.23	0.104		0.177		0.302	
F−	0.06	0.127		0.190		0.298	
H−	0.00	0.129	0.111	0.187	0.198	0.286	0.317
CH_3−	−0.17	0.115		0.192		0.314	
OH−	−0.37	0.115		0.192		0.337	
NH_2−	−0.66	0.119	0.116	0.197	0.208	0.287	0.402
average ± SD		0.120 ± 0.011		0.187 ± 0.017		0.33 ± 0.04	

[a]The inductive effect of each substituent is reflected in the Hammett constant (σ_p), with positive values indicating the electron-withdrawing capability and negative values indicating the electron-donating capability.[45,57] [b]The average percent error on A from fitting is 8.9% of each value.

https://doi.org/10.1021/acs.jctc.1c00362
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 3. Fitted Electrostatic *B* Parameter for the rg*ff*BXB (Scheme 1) as a Function of Increasing Electron-Donating Substituents (Y) in Model Donor-to-Acceptor ($H_2PO_2^{-1}$ or *N*-Methylacetimide, NMA) X-Bonding Interactions[b]

donor		X = Cl		X = Br		X = I	
		$H_2PO_2^{-1}$	NMA	$H_2PO_2^{-1}$	NMA	$H_2PO_2^{-1}$	NMA
X-uracil		−0.062 (−0.046)		0.057 (−0.0072)		0.233 (0.0607)	
Y−	σ_p[a]	donor = *para*-Y-halobenzene					
CN−	0.66	−0.087 (−0.090)	−0.143 (−0.090)	0.043 (−0.044)	−0.071 (−0.044)	0.149(−0.018)	−0.025 (−0.018)
Cl−	0.23	−0.101 (−0.108)		−0.016 (−0.063)		0.066 (−0.046)	
F−	0.06	−0.199 (−0.112)		−0.076 (−0.064)		0.018 (−0.016)	
H−	0.00	−0.249 (−0.125)	−0.252 (−0.125)	−0.090 (−0.086)	−0.087 (−0.086)	0.181 (−0.035)	0.128 (−0.035)
CH_3−	−0.17	−0.269 (−0.126)		−0.155 (−0.083)		−0.016 (−0.076)	
OH−	−0.37	−0.267 (−0.127)		−0.157 (−0.088)		−0.071 (−0.078)	
NH_2−	−0.66	−0.303 (−0.137)	−0.318 (−0.137)	−0.216 (−0.100)	−0.261 (−0.100)	−0.048 (−0.054)	−0.180 (−0.054)

[a]The inductive effect of each substituent is reflected in the Hammett constant (σ_p), with positive values indicating the electron-withdrawing capability and negative values indicating the electron-donating capability.[45,57] [b]The average percent error on *B* from fitting is 16.7% of each value. The corresponding RESP-calculated charge for the halogen of each compound is shown in parentheses.

The *B* parameters are all variable and become more positive as the substituent para to the halogen became more electron-withdrawing (Table 3). Although, as discussed above, *B* is not explicitly the charge of the halogen, this effect is consistent with the σ-hole model for BXBs. The *B* values derived here for halogen substituents (uracil base) that are X-bonded to a hypophosphate acceptor are linearly correlated but slightly more negative compared to those from the original *ff*BXB ($B_{(rgffBXB)} = 0.96B_{(ffBXB)} - 0.059$, $R^2 = 0.975$).

With this set of parameters, six fixed and one variable, we show that the rg*ff*BXB can accurately recapitulate the QM energy landscape (Figure 3). The reparameterized force field in Scheme 1 fits the QM-calculated E_{int} landscape very well, particularly at θ_1 from 120 to 180°, where most X-bonds are seen. The primary deviations are at θ_1 = 90°, where the rg*ff*BXB predicts a more positive E_{int} than the QM calculation. This deviation was expected since ν was fixed to be explicitly 2. Deviations in energies at this angle of approach should not affect the ability of the rg*ff*BXB to predict the X-bonding potential of each halogen. The optimum angle at which the halogen serves as an H-bond acceptor would be affected. However, given that H-bonds are not as directional as X-bonds, we consider this to be a reasonable trade-off in accuracy versus efficiency. Finally, by constructing the QM energy landscape with both $O_2PH_2^{-1}$ and neutral NMA acceptors, we have achieved the goal of deriving a consistent set of parameters for the rg*ff*BXB that is generally applicable to both nucleic acids and proteins.

RESP Fitting of the rg*ff*BXB Parameter. The rg*ff*BXB has now become less unwieldy but still requires high-level QM calculations to determine the single remaining variable parameter (*B*). Hence, the next goal was to determine whether *B* could be derived through a less costly calculation. In the simplest case of the *para*-halogenated benzene compounds, we see that *B* is well correlated with the standard Hammett constant (Figure 4).

For classical class I MM/MD simulations on more complex halogenated compounds, an RESP[46,47] or AM1-BCC protocol is typically used to assign the isotropic charge of atoms of a previously uncharacterized compound or ligand. Since *B* reflects the overall charge of a halogen, it seemed reasonable that this parameter could be determined from either of these protocols. The RESP charges calculated for the atoms of the halogenated benzene X-bond donors with various electron-donating or electron-withdrawing para-substituents (Table 2)

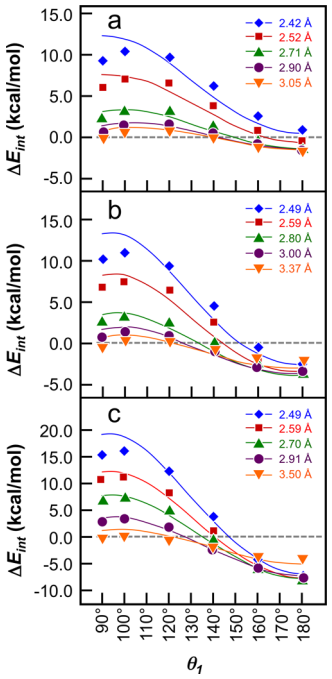


Figure 3. MP2 quantum calculations for energies of $^{X}U\cdots H_2PO_2^{-1}$ interacting pairs at various geometries of approach, for X = (a) Cl, (b) Br, or (c) I, with rg*ff*BXB-calculated curves using parameters from Tables 1 and 2 and the *B* parameter as the only fitted variable. MP2 energies of interaction (ΔE_{int}) were calculated at various angles of the X-bond acceptor ($H_2PO_2^{-1}$) approach to the donor (^{X}U), with the distance varying from a close range (~70−75% of the sum of the van der Waals radii, $\sum R_{vdW}$) (blue diamonds) to farther distances (~100% of $\sum R_{vdW}$) (orange triangles).

were seen to be linearly correlated to the *B* parameters determined from the QM E_{int} (Figure 5). As noted above, *B* is not strictly a unitless Coulombic charge but has units of Å. This parameter would need to be appropriately scaled if distances between interacting pairs are calculated in units other than Å. The average absolute residual between the RESP and fitted *B* values ($\langle|\Delta B|\rangle$) is 0.05 (ranging in values from +0.14 to −0.16, both extremes associated with iodine X-bonds to the NMA acceptor). Thus, the reduced general parameters of the *ff*BXB, including the single independent variable, can be assigned to any uncharacterized molecule using the standard protocols of a molecular simulation procedure.

https://doi.org/10.1021/acs.jctc.1c00362
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

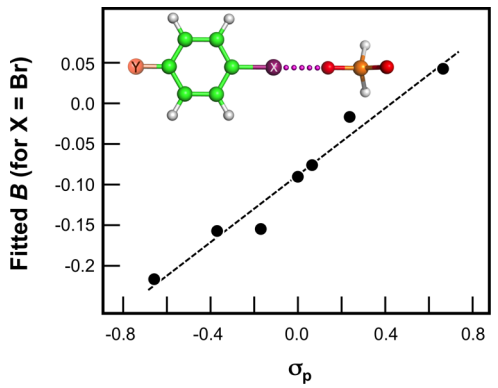


Figure 4. Electrostatic parameter B of the rg*ff*BXB of *para*-Y-substituted bromobenzene versus Hammett constants of Y (σ_p, Table 2). The parameter B was fitted using the generalized rg*ff*BXB for various substituted bromobenzenes. The regression analysis of the resulting B values relative to the σ_p Hammett values[45] of the *para*-substituent (dotted line) resulted in a linear relationship $B = 0.206\sigma_p - 0.088$, $R^2 = 0.96$.

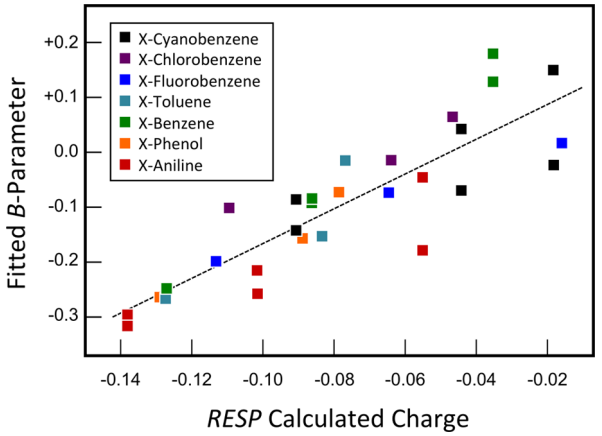


Figure 5. Correlation between the fitted electrostatic B parameter from the rg*ff*BXB and the RESP-calculated charge of the *para*-substituted halobenzenes (X = Cl, Br, or I). The linear regression analysis resulted in the relationship $B = 3.22\text{RESP} + 0.152$, $R^2 = 0.764$. The B parameters and RESP values for the substituted uracils were not included in this analysis.

DISCUSSION AND CONCLUSIONS

In this study, we first showed that the seven variable parameters of the *ff*BXB, which was derived to model X-bonds based on the anisotropic physical properties of shape and charge of halogen substituents, could be reduced to a single parameter associated with the overall charge of the halogen. The remaining electrostatic parameter was shown to be derivable for Cl, Br, or I based on models of X-bonds engineered into DNA junctions or the T4 lysozyme model systems, rendering it overall more general to both nucleic acids and proteins. The resulting reduced and generalized *ff*BXB (rg*ff*BXB) accurately recapitulates the quantum mechanical landscape for a broad range of geometries (interaction distances and angles), with only minor deviations at angles perpendicular to the σ-hole of the halogen. Furthermore, the parameters of the rg*ff*BXB for simple halogenated benzene-based compounds were shown to be directly related to standard Hammett-σ constants. Finally, we show evidence that the rg*ff*BXB parameter can be derived for the halogen substituent of any halogenated compound from a standard RESP calculation.

We can now determine how accurately the rg*ff*BXB models X-bonds engineered into DNA junctions and the model T4 lysozyme enzyme (T4L, Figure 2). The four-stranded DNA junction structure had previously been used to determine the structure−energy relationships of BXBs using a competition assay.[11] In the T4L test system, a bromo- or iodophenylalanine replaced the tyrosine at the Y18 position to create the constructs ^{Br}F18-T4L or ^{I}F18-T4L.[12] The crystal structure of ^{Br}F18-T4L showed a longer and weaker Br···O X-bond interaction to the carbonyl oxygen of the protein backbone ($E_{int} = -0.6 \pm 1$ kcal/mol), while ^{I}F18-T4L showed a shorter and stronger I···O X-bond ($E_{int} = -1.6 \pm 1$ kcal/mol).

We modeled the X-bonds in the DNA junction system using a halo-uracil donor and hypophosphite acceptor pair as previously described.[23,24] For the T4L protein, we modeled the Br···O and I···O X-bonding interactions as a complex between the respective halobenzene (mimicking the halogenated phenylalanine side chain as the X-bond donor) and NMA (mimicking the backbone carbonyl oxygen acceptor) with the atomic positions fixed at the coordinates seen in their X-ray structures. The energies of interaction for each X-bonded pair were calculated using the AMBER99sb[44] parameter set for all nonhalogen atoms and the rg*ff*BXB for the halogens. It is clear that the energies calculated from the rg*ff*BXB are well correlated with the experimentally determined energies of X-bonding interactions in both the DNA and protein model systems, with the calculations placing the experimental energies in essentially the correct order (Figure 6). The average deviation between the calculated and measured energies is 0.72 kcal/mol (ranging between −1.5 and +1.2 kcal/mol) but within the reported experimental errors. Thus, the experimentally determined X-bonding energies in both DNA and protein model systems are reasonably well predicted

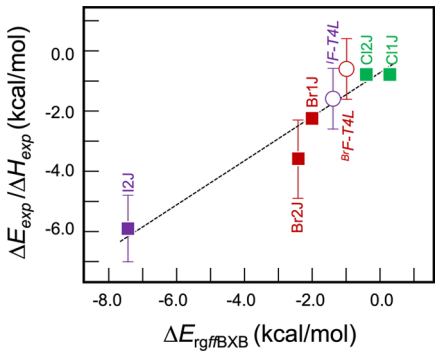


Figure 6. Comparison of experimentally determined X-bonding energies to rg*ff*BXB-calculated energies. The experimental X-bonding energies (ΔE_{exp}, solid squares) of DNA junctions with the chlorinated uracils (Cl1J and Cl2J for the singly and doubly chlorinated bases, respectively) and with the singly brominated uracil (Br1J) were determined by crystallographic titration.[11,48] The experimental X-bonding enthalpies (ΔH_{exp}, solid squares with error bars) of doubly brominated (Br2J) or iodinated (I2) junctions were determined by differential scanning calorimetry.[40,42] The X-bonding ΔH_{exp} (open circles) of the bromo- and iodophenylalanine constructs were determined by differential scanning calorimetry.[12] Error bars represent the standard errors for the experimental DSC measures in the DNA and protein systems. The dashed line shows a linear relationship between the experimental and rg*ff*BXB-calculated energies (slope = 0.73; intercept = −0.75 kcal/mol; $R^2 = 0.91$).

https://doi.org/10.1021/acs.jctc.1c00362
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

by rg*ff*BXB calculations on X-bonding geometries from their respective crystal structures.

The less than 1.0 slope (0.73) and relatively large negative (−0.75 kcal/mol) intercept of the linear relationship would suggest that the rg*ff*BXB might require additional post-calculation corrections. However, if we focus on the X-bonds between 0 and −5 kcal/mol (leaving out the strongly stabilizing iodinated junction), then the linear relationship results in a slope = 1.01, with an intercept of −0.43 kcal/mol (R^2 = 0.75). In comparison, the original *ff*BXB applied to the DNA junction data set resulted in a linear relationship with a slope = 0.92 and *y*-intercept = −0.46 kcal/mol, essentially similar to the rg*ff*BXB for the weak to medium strength X-bonding interactions (Cl and Br interactions). A direct comparison of the energies calculated for the X-bonds in the DNA junctions by the rg*ff*BXB versus the original *ff*BXB resulted in a linear relationship with a slope = 1.23 and intercept = 0.73 kcal/mol (R^2 = 0.97). Again, for the Cl and Br interactions, excluding I, the relationship has a slope = 0.93, with an intercept of 0.3 kcal/mol (R^2 = 0.97). The rg*ff*BXB, therefore, is apparently the most accurate in modeling the energies of weak to medium strength X-bonds (energies of interaction less negative than −4 kcal/mol) but still provides reasonably accurate energies (within 25% of the experimental values for an iodine X-bonding energy of −6 kcal/mol) for stronger interactions.

The reparameterized rg*ff*BXB with its single variable electrostatic parameter can now be readily incorporated into algorithms that rely on classical potential energy functions. In particular, by setting the phase parameter (ν) and the power of the electrostatic function (n) as an integer value of 2, the derivations of the gradients for these functions will be more straightforward. The rg*ff*BXB should thus be adaptable to accurately model X-bonds in various force fields, including AMBER,[20] CHARMM,[36] and OPLS3,[37] or into scoring functions[41,49] of molecular docking programs. We note that the current parameterization of the rg*ff*BXB is for neutral compounds. Sedlak et al.[43] had shown that the degree of polarization of the BXB donor affects the extent of polar flattening associated with the halogen atom, varying from 9 to 15% reduction in the van der Waals radius. We had proposed that the perturbations to the size and shape of the donor could be largely compensated for by the electrostatic parameters. The results from this study generally support this hypothesis for the set of neutral compounds in that the refined B parameter shows a strong linear correlation to the electron-donating or electron-withdrawing character of the remaining molecule (Figure 4) across a large range of Hammett-σ constants. A halogen substituent in a charged compound, however, may significantly increase or decrease the degree of polar flattening, and whether the deviations in size and shape can be as readily compensated must be tested.

Looking forward, there is now interest in how BXBs are affected by adjacent H-bonds, resulting in H-bond-enhanced X-bonds (HB*e*XBs, for short).[50] These HB*e*XBs have been shown to increase the affinity of small-molecule receptors for their anionic ligands[51] and the stability of proteins[13] and be involved in the recognition of proteins for ligands.[52] There is a current general understanding that the properties of X-bonds and other nonclassical noncovalent interactions can be greatly perturbed by their immediate surroundings, and therefore, it is important to consider effects beyond the direct electrostatic σ-hole model. We have now established the general framework by which the rg*ff*BXB can be readily adapted to this and other peripheral effects on the strengths of X-bonds, which will be critical when applying BXB concepts in biomolecular engineering and in the design of halogenated compounds in medicinal chemistry.

COMPUTATIONAL METHODS

X-Bond Model for Class I MD. The goal of this study is to develop a classical model that recapitulates the anisotropic charge distribution and shape properties of X-bonds, which can be readily implemented into class I MM/MD algorithms (Scheme 2). Classical potential energy functions (V) are

Scheme 2. Classical Potential Energy Equations in Class I Molecular Mechanics/Dynamics Algorithms Describing the Bonding and Nonbonding Potentials

$$V = \left\{ \sum_{lengths} K_b(b - b_o)^2 + \sum_{angles} K_\theta(\theta - \theta_o)^2 + \sum_{dihedrals} K_\phi(1 + \cos(n\phi - \delta)) + \sum_{impropers} K_\psi(1 + \cos(n\psi - \psi_o)) \right\}_{Bonding} + \left\{ \sum_{electrostatics} \frac{Z_j Z_i e^2}{4\pi\varepsilon_o r_{ij}} + \sqrt{\varepsilon_i \varepsilon_j} \left[\left(\frac{R_i + R_j}{r_{ij}} \right)^{12} - 2 \left(\frac{R_i + R_j}{r_{ij}} \right)^{6} \right] \right\}_{Non-Bonding}$$

described by a set of bonding terms (bond lengths (b), angles (θ), dihedrals (ϕ), and improper dihedrals (ψ)) and nonbonding terms (the electrostatic and 6-12 Lennard-Jones potentials) that depend on distances between nonbonded pairs, r_{ij}.[44]

In standard programs, class I MD simulations (Scheme 2) require three nonbonding parameters as an input for each interacting atom i and j in any new system: (1) electrostatic charges (Z_i and Z_j), (2) the van der Waals radii (R_i and R_j), and (3) contributions to the minima of potential energy well of the van der Waals interaction (ε_i and ε_j). Generally, the shape or van der Waals parameters are found empirically, through high-level QM calculations, or a combination of both[44] and set to each atom type once determined. Charge parameters are found using either an RESP[46,47] charge model or AM1 with bond charge corrections (AM1-BCC[53]) to the charge model.

We had previously modified the nonbonding term of the general function in Scheme 2 to accommodate X-bonding in deriving the *ff*BXB.[23] In short, the directional electrostatic potential was modeled by defining the formal charge of the halogen by a cosine function ($Z_X = A\cos(\nu\alpha) + B$, where A is the amplitude of the charge anisotropy, B is the baseline and ν is the period of the function, and α is the angle of approach of the X-bond acceptor to the σ-hole, defined as $\alpha = 180° - \theta_1$, relative to Figure 1). In addition, the electrostatic potential allows the exponential (n) for the distance relationship between the X-bond acceptor and the donor (r_{AX}) to float. The aspherical shape of halogen substituents is modeled by a similar cosine function applied as a correction to the van der Waals radius ($R_X = \langle R_X \rangle - \Delta R\cos(\nu\alpha)$, where $\langle R_X \rangle$ is the average van der Waals radius and ΔR is the perturbation due to polar flattening of the halogen) in the dispersion term of the Lennard-Jones potential. Thus, the resulting potential for the noncovalent interactions adds three unique parameters that describe the electrostatic property (A, B, and n), three for the shape ($\langle R_X \rangle$, ΔR_X, and ε_X to define the contribution of the halogen to the energy term), and one (ν) for the directionality

https://doi.org/10.1021/acs.jctc.1c00362
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

of each halogen (Scheme 1). It is important to note that class I MD simulations are generally performed on pairwise interactions, which for X-bonding would be between the halogen (X) and its acceptor (A). However, the addition of the cosine function for X-bonding causes the interaction to be a three-body term. The three bodies will be (1) the atom bound to the halogen (Y), (2) the halogen (X), and (3) its acceptor (A).

Parameterization of the rg*ff*BXB through QM Fitting. For the modified potential in Scheme 1, the additional *ff*BXB parameters were determined by applying QM calculations on small-molecule systems that mimic biologically relevant X-bond donor−acceptor pairs.[30] The QM second-order Møller−Plesset (MP2) energies on all systems in this study were calculated using Gaussian 09 revision E.01, applying the aug-cc-PVTZ basis set for Cl and Br and aug-cc-PVTZ-PP for I.[54] The implicit solvent model used was cyclohexane ($D = 2$), with a gas phase BSSE correction[55] applied afterward. This solvent model is appropriate for calculations on systems that involve explicit solvents in MM/MD simulations and short distances between interacting atoms, as is the current interacting pairs, and reflects the low dielectric expected for a protein interior.[56] The structures of each compound (BXB donor or acceptor) were geometry-optimized, but no additional optimizations were performed from the defined fixed intermolecular geometries. Energy-geometry landscapes for these systems were defined by varying the angle of approach (θ_1) from 90° to 180° and varying the interacting distances from 70 to 100% of the sum of the standard van der Waals radii for each of the X-bond donor−acceptor pairs. The resulting MP2-calculated energy landscapes were used to determine the rg*ff*BXB parameters in Scheme 1, using a global nonlinear least squares fitting algorithm developed in-house in MATHEMATICA, for each X-bonded model system.

Electrostatic charges for the halogens in each BXB donor were calculated using the RESP procedure implemented in the AMBER12 suite.[46]

AUTHOR INFORMATION

Corresponding Author

P. Shing Ho − *Department of Biochemistry & Molecular Biology, Colorado State University, Fort Collins, Colorado 80523-1870, United States;* orcid.org/0000-0002-8082-4311; Phone: 970-491-0569; Email: sshing.ho@colostate.edu; Fax: 970-491-0494

Authors

Melissa Coates Ford − *Department of Biochemistry & Molecular Biology, Colorado State University, Fort Collins, Colorado 80523-1870, United States;* orcid.org/0000-0002-0253-5389

Anthony K. Rappé − *Department of Chemistry, Colorado State University, Fort Collins, Colorado 80523, United States;* orcid.org/0000-0002-5259-1186

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.1c00362

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by a grant from the National Science Foundation (CHE-1905328). We thank R.S. Czarny and A.N. Ho for critical reading of the manuscript.

REFERENCES

(1) Xu, Z.; Liu, Z.; Chen, T.; Chen, T.; Wang, Z.; Tian, G.; Shi, J.; Wang, X.; Lu, Y.; Yan, X.; Wang, G.; Jiang, H.; Chen, K.; Wang, S.; Xu, Y.; Shen, J.; Zhu, W. Utilization of halogen bond in lead optimization: a case study of rational design of potent phosphodiesterase type 5 (PDE5) inhibitors. *J. Med. Chem.* **2011**, *54*, 5607−5611.

(2) Hernandes, M. Z.; Cavalcanti, S. M.; Moreira, D. R.; de Azevedo Junior, W. F.; Leite, A. C. Halogen atoms in the modern medicinal chemistry: hints for the drug design. *Curr. Drug Targets* **2010**, *11*, 303−314.

(3) Roughley, S. D.; Jordan, A. M. The medicinal chemist's toolbox: an analysis of reactions used in the pursuit of drug candidates. *J. Med. Chem.* **2011**, *54*, 3451−3479.

(4) Desiraju, G. R.; Ho, P. S.; Kloo, L.; Legon, A. C.; Marquardt, R.; Metrangolo, P.; Politzer, P.; Resnati, G.; Rissanen, K. Definition of the halogen bond (IUPAC Recommendations 2013). *Pure Appl. Chem.* **2013**, *85*, 1711−1713.

(5) Auffinger, P.; Hays, F. A.; Westhof, E.; Ho, P. S. Halogen bonds in biological molecules. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 16789−16794.

(6) Iltzsch, M. H.; Uber, S. S.; Tankersley, K. O.; el Kouni, M. H. Structure-activity relationship for the binding of nucleoside ligands to adenosine kinase from Toxoplasma gondii. *Biochem. Pharmacol.* **1995**, *49*, 1501−1512.

(7) Benjahad, A.; Guillemont, J.; Andries, K.; Nguyen, C. H.; Grierson, D. S. 3-iodo-4-phenoxypyridinones (IOPY's), a new family of highly potent non-nucleoside inhibitors of HIV-1 reverse transcriptase. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 4309−4312.

(8) Howard, E. I.; Sanishvili, R.; Cachau, R. E.; Mitschler, A.; Chevrier, B.; Barth, P.; Lamour, V.; Van Zandt, M.; Sibley, E.; Bon, C.; Moras, D.; Schneider, T. R.; Joachimiak, A.; Podjarny, A. Ultrahigh resolution drug design I: details of interactions in human aldose reductase-inhibitor complex at 0.66 Å. *Proteins* **2004**, *55*, 792−804.

(9) Parks, D. J.; Lafrance, L. V.; Calvo, R. R.; Milkiewicz, K. L.; Gupta, V.; Lattanze, J.; Ramachandren, K.; Carver, T. E.; Petrella, E. C.; Cummings, M. D.; Maguire, D.; Grasberger, B. L.; Lu, T. 1,4-Benzodiazepine-2,5-diones as small molecule antagonists of the HDM2-p53 interaction: discovery and SAR. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 765−770.

(10) Maillard, M. C.; Hom, R. K.; Benson, T. E.; Moon, J. B.; Mamo, S.; Bienkowski, M.; Tomasselli, A. G.; Woods, D. D.; Prince, D. B.; Paddock, D. J.; Emmons, T. L.; Tucker, J. A.; Dappen, M. S.; Brogley, L.; Thorsett, E. D.; Jewett, N.; Sinha, S.; John, V. Design, synthesis, and crystal structure of hydroxyethyl secondary amine-based peptidomimetic inhibitors of human beta-secretase. *J. Med. Chem.* **2007**, *50*, 776−781.

(11) Voth, A. R.; Hays, F. A.; Ho, P. S. Directing macromolecular conformation through halogen bonds. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 6188−6193.

(12) Scholfield, M. R.; Ford, M. C.; Carlsson, A.-C. C.; Butta, H.; Mehl, R. A.; Ho, P. S. Structure-Energy Relationships of Halogen Bonds in Proteins. *Biochemistry* **2017**, *56*, 2794−2802.

(13) Carlsson, A.-C. C.; Scholfield, M. R.; Rowe, R. K.; Ford, M. C.; Alexander, A. T.; Mehl, R. A.; Ho, P. S. Increasing Enzyme Stability and Activity through Hydrogen Bond-Enhanced Halogen Bonds. *Biochemistry* **2018**, *57*, 4135−4147.

(14) Wilcken, R.; Zimmermann, M. O.; Lange, A.; Joerger, A. C.; Boeckler, F. M. Principles and Applications of Halogen Bonding in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **2013**, *56*, 1363−1388.

(15) 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.
(16) El Hage, K.; Pandyarajan, V.; Phillips, N. B.; Smith, B. J.; Menting, J. G.; Whittaker, J.; Lawrence, M. C.; Meuwly, M.; Weiss, M. A. Extending Halogen-Based Medicinal Chemistry to Proteins: Iodo-Insulin as a Case Study. *J. Biol. Chem.* **2016**, *291*, 27023−27041.
(17) Brown, F. K.; Sherer, E. C.; Johnson, S. A.; Holloway, M. K.; Sherborne, B. S. The evolution of drug design at Merck Research Laboratories. *J. Comput.-Aided Mol. Des.* **2017**, *31*, 255−266.
(18) Ho, P. S. Halogen bonding in medicinal chemistry: from observation to prediction. *Future Med. Chem.* **2017**, *9*, 637−640.
(19) Ho, P. S.; Anderson, D. M., Halogen Bonds in Biomolecular Engineering. In *Halogen Bonding in Solution*, Huber, S., Ed. Wiley-VCH Verlag GmbH & Co: Weinheim, Germany, 2021; pp. 335−362.
(20) Ibrahim, M. A. A. AMBER empirical potential describes the geometry and energy of noncovalent halogen interactions better than advanced semiempirical quantum mechanical method PM6-DH2X. *J. Phys. Chem. B* **2012**, *116*, 3659−3669.
(21) Kolář, M.; Hobza, P. On Extension of the Current Biomolecular Empirical Force Field for the Description of Halogen Bonds. *J. Chem. Theory Comput.* **2012**, *8*, 1325−1333.
(22) Ford, M. C.; Ho, P. S. Computational Tools To Model Halogen Bonds in Medicinal Chemistry. *J. Med. Chem.* **2016**, *59*, 1655−1670.
(23) Carter, M.; Rappé, A. K.; Ho, P. S. Scalable anisotropic shape and electrostatic models for biological bromine halogen bonds. *J. Chem. Theory Comput.* **2012**, *8*, 2461−2473.
(24) Scholfield, M. R.; Ford, M. C.; Vander Zanden, C. M.; Billman, M. M.; Ho, P. S.; Rappé, A. K. Force Field Model of Periodic Trends in Biomolecular Halogen Bonds. *J. Phys. Chem. B* **2015**, *119*, 9140−9149.
(25) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F., III; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. De novo computational design of retro-aldol enzymes. *Science* **2008**, *319*, 1387−1391.
(26) Richter, F.; Leaver-Fay, A.; Khare, S. D.; Bjelic, S.; Baker, D. De novo enzyme design using Rosetta3. *PLoS One* **2011**, *6*, No. e19230.
(27) Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. Kemp elimination catalysts by computational enzyme design. *Nature* **2008**, *453*, 190−195.
(28) Mukhopadhyay, A.; Redding, A. M.; Rutherford, B. J.; Keasling, J. D. Importance of systems biology in engineering microbes for biofuel production. *Curr. Opin. Biotechnol.* **2008**, *19*, 228−234.
(29) Jorgensen, W. L. Computer-aided discovery of anti-HIV agents. *Bioorg. Med. Chem.* **2016**, *24*, 4768−4778.
(30) Devereux, C.; Smith, J. S.; Davis, K. K.; Barros, K.; Zubatyuk, R.; Isayev, O.; Roitberg, A. E. Extending the Applicability of the ANI Deep Learning Molecular Potential to Sulfur and Halogens. *J. Chem. Theory Comput.* **2020**, *16*, 4192−4202.
(31) Clark, T.; Hennemann, M.; Murray, J. S.; Politzer, P. Halogen bonding: the σ-hole. *J. Mol. Model.* **2007**, *13*, 291−296.
(32) Scholfield, M. R.; Zanden, C. M. V.; Carter, M.; Ho, P. S. Halogen bonding (X-bonding): a biological perspective. *Protein Sci.* **2013**, *22*, 139−152.
(33) Voth, A. R.; Khuu, P.; Oishi, K.; Ho, P. S. Halogen bonds as orthogonal molecular interactions to hydrogen bonds. *Nat. Chem.* **2009**, *1*, 74−79.
(34) Eskandari, K.; Zariny, H. Halogen bonding: A lump-hole interaction. *Chem. Phys. Lett.* **2010**, *492*, 9−13.
(35) Rosokha, S. V.; Neretin, I. S.; Rosokha, T. Y.; Hecht, J.; Kochi, J. K. Charge-transfer character of halogen bonding: Molecular structures and electronic spectroscopy of carbon tetrabromide and bromoform complexes with organic σ- and π-donors. *Heteroat. Chem.* **2006**, *17*, 449−459.
(36) Soteras Gutiérrez, I.; Lin, F. Y.; Vanommeslaeghe, K.; Lemkul, J. A.; Armacost, K. A.; Brooks, C. L., III; MacKerell, A. D., Jr. Parametrization of halogen bonds in the CHARMM general force field: Improved treatment of ligand-protein interactions. *Bioorg. Med. Chem.* **2016**, *24*, 4812−4825.
(37) Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **2016**, *12*, 281−296.
(38) Riley, K. E.; Vazquez, M.; Umemura, C.; Miller, C.; Tran, K. A. Exploring the (Very Flat) Potential Energy Landscape of R-Br$\cdots\pi$ Interactions with Accurate CCSD(T) and SAPT Techniques. *Chem. - Eur. J.* **2016**, *22*, 17690−17695.
(39) Bereau, T.; Kramer, C.; Meuwly, M. Leveraging Symmetries of Static Atomic Multipole Electrostatics in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2013**, *9*, 5450−5459.
(40) Carter, M.; Voth, A. R.; Scholfield, M. R.; Rummel, B.; Sowers, L. C.; Ho, P. S. Enthalpy-Entropy Compensation in Biomolecular Halogen Bonds Measured in DNA Junctions. *Biochemistry* **2013**, *52*, 4891−4903.
(41) Koebel, M. R.; Schmadeke, G.; Posner, R. G.; Sirimulla, S. AutoDock VinaXB: implementation of XBSF, new empirical halogen bond scoring function, into AutoDock Vina. *Aust. J. Chem.* **2016**, *8*, 27.
(42) Carter, M.; Ho, P. S. Assaying the Energies of Biological Halogen Bonds. *Cryst. Growth Des.* **2011**, *11*, 5087−5095.
(43) Sedlak, R.; Kolář, M. H.; Hobza, P. Polar Flattening and the Strength of Halogen Bonding. *J. Chem. Theory Comput.* **2015**, *11*, 4727−4732.
(44) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.
(45) McDaniel, D. H.; Brown, H. C. An Extended Table of Hammett Substitutent Constants Based on the Ionization of Substituted Benzoic Acids. *J. Org. Chem.* **1958**, *23*, 420−427.
(46) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(47) Wang, J.; Cieplak, P.; Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(48) Vander Zanden, C. M.; Carter, M.; Ho, P. S. Determining thermodynamic properties of molecular interactions from single crystal studies. *Methods* **2013**, *64*, 12−18.
(49) Zimmermann, M. O.; Lange, A.; Boeckler, F. M. Evaluating the Potential of Halogen Bonding in Molecular Design: Automated Scaffold Decoration Using the New Scoring Function XBScore. *J. Chem. Inf. Model.* **2015**, *55*, 687−699.
(50) Riel, A. M. S.; Rowe, R. K.; Ho, E. N.; Carlsson, A.-C. C.; Rappé, A. K.; Berryman, O. B.; Ho, P. S. Hydrogen Bond Enhanced Halogen Bonds: A Synergistic Interaction in Chemistry and Biochemistry. *Acc. Chem. Res.* **2019**, *52*, 2870−2880.
(51) Riel, A. M. S.; Decato, D. A.; Sun, J.; Massena, C. J.; Jessop, M. J.; Berryman, O. B. The intramolecular hydrogen bonded-halogen bond: a new strategy for preorganization and enhanced binding. *Chem. Sci.* **2018**, *9*, 5828−5836.
(52) Lin, F.-Y.; MacKerell, A. D., Jr. Do Halogen-Hydrogen Bond Donor Interactions Dominate the Favorable Contribution of Halogens to Ligand-Protein Binding? *J. Phys. Chem. B* **2017**, *121*, 6813−6821.
(53) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623−1641.
(54) Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. L. Basis set exchange: A

https://doi.org/10.1021/acs.jctc.1c00362
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

community database for computational sciences. *J. Chem. Inf. Model.* **2007**, *47*, 1045−1052.
(55) Boys, S. F.; Bernardi, F. Calculation of small molecular interactions by differences of separate total energies−some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553−566.
(56) Zhou, H. X.; Pang, X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chem. Rev.* **2018**, *118*, 1691−1741.
(57) Hammett, L. P. The Effect of Structure upon the Reactions of Organic Compounds. Benzene Derivatives. *J. Am. Chem. Soc.* **1937**, *59*, 96−103.

https://doi.org/10.1021/acs.jctc.1c00362
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX