BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

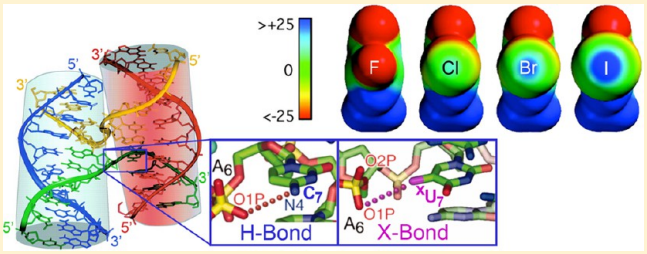
pubs.acs.org/biochemistry

Enthalpy–Entropy Compensation in Biomolecular Halogen Bonds Measured in DNA Junctions

Megan Carter,[†] Andrea Regier Voth,[†,§] Matthew R. Scholfield,[†] Brittany Rummel,[†] Lawrence C. Sowers,[‡] and P. Shing Ho*,[†]

[†]Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado 80523, United States
[‡]Department of Pharmacology and Toxicology, University of Texas Medical Branch, Galveston, Texas 77555, United States

ABSTRACT: Interest in noncovalent interactions involving halogens, particularly halogen bonds (X-bonds), has grown dramatically in the past decade, propelled by the use of X-bonding in molecular engineering and drug design. However, it is clear that a complete analysis of the structure–energy relationship must be established in biological systems to fully exploit X-bonds for biomolecular engineering. We present here the first comprehensive experimental study to correlate geometries with their stabilizing potentials for fluorine (F), chlorine (Cl), bromine (Br), or iodine (I) X-bonds in a biological context. For these studies, we determine the single-crystal structures of DNA Holliday junctions containing halogenated uracil bases that compete X-bonds against classic hydrogen bonds (H-bonds), estimate the enthalpic energies of the competing interactions in the crystal system through crystallographic titrations, and compare the enthalpic and entropic energies of bromine and iodine X-bonds in solution by differential scanning calorimetry. The culmination of these studies demonstrates that enthalpic stabilization of X-bonds increases with increasing polarizability from F to Cl to Br to I, which is consistent with the σ-hole theory of X-bonding. Furthermore, an increase in the X-bonding potential is seen to direct the interaction toward a more ideal geometry. However, the entropic contributions to the total free energies must also be considered to determine how each halogen potentially contributes to the overall stability of the interaction. We find that bromine has the optimal balance between enthalpic and entropic energy components, resulting in the lowest free energy for X-bonding in this DNA system. The X-bond formed by iodine is more enthalpically stable, but this comes with an entropic cost, which we attribute to crowding effects. Thus, the overall free energy of an X-bonding interaction balances the stabilizing electrostatic effects of the σ-hole against the competing effects on the local structural dynamics of the system.



The halogens of halo-organic compounds form unique interactions that are now called halogen bonds, or X-bonds. X-Bonds are short-range interactions that draw an electron-rich Lewis base (serving as the X-bond acceptor) toward the positive crown of a polarized halogen (serving as the X-bond donor), resulting in interatomic distances that are closer than the sum of their respective van der Waals radii ($\sum R_{vdW}$) (Figure 1). X-Bonds are only now being formally defined[1] but have become an increasingly useful interaction for crystallographic and molecular design in chemistry and chemical engineering.[2−4] In biomolecules,[5−7] X-bonds have been engineered to direct the conformation of DNA[8] and to increase the specificity of an inhibitor for its protein target.[9−13] It was recently reported that "25% of the top 200 brand name drugs by retail dollar in 2009 possess halogen atoms in their molecular structures".[13]

Halogen substitutions in small molecules are common side products of chemical synthesis and have traditionally been introduced to increase hydrophobicity and delay the catabolic process of potential drug candidates.[14,15] However, electrostatic interactions, including X-bonding, are becoming recognized for their ability to increase the specificity and efficacy of halogenated inhibitors. Recent studies have shown that halogens can be introduced in a strategic manner to optimize X-bonding and effectively decrease the IC_{50} of drug candidates.[16] While the usefulness of X-bonds in rational bottom-up drug design is becoming better recognized,[17,18] efforts to rationally engineer X-bonds to optimize the efficacy of inhibitors have not always been successful.[19,20] These efforts can be improved by a more thorough understanding of the effect of halogen polarizability on the X-bond structure–energy relationship, particularly in terms of the enthalpic and entropic contributions to the overall free energy of interaction.

Halogens as a class of elements, including fluorine (F), chlorine (Cl), bromine (Br), and iodine (I), are typically thought to carry overall negative charges; however, when a halogen is covalently bonded to another atom, the electrostatic potential across the halogen surface becomes nonuniformly distributed. This anisotropic charge distribution is most readily described as the depopulation of the p_z orbital of the halogen as

Received: May 10, 2013
Revised: June 17, 2013
Published: June 24, 2013

dx.doi.org/10.1021/bi400590h | *Biochemistry* 2013, 52, 4891−4903

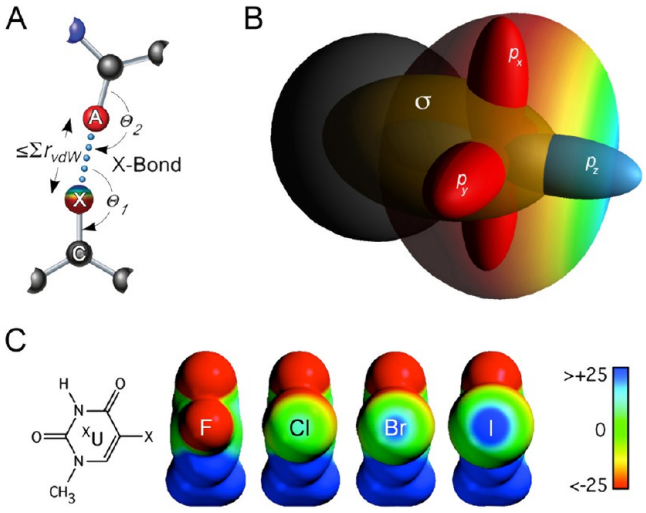


Figure 1. Description of halogen bonds (X-bonds). (A) X-Bond geometry. X-Bonds are defined by short-distance interactions in which the interatomic distance between the X-bond donor and acceptor (X···A) is shorter than the sum of their respective van der Waals radii ($\sum R_{vdW}$). The angle of approach of the acceptor to the halogen is defined by Θ_1, which is generally linear relative to the C−X covalent bond, while the Θ_2 angle defines the angle of approach of the halogen to the acceptor. (B) σ-Hole theory of X-bonds. In this model, the valence electron of the halogen pairs (in this example) to that of a carbon (dark gray sphere) to form the covalent C−X σ-bond (yellow cylinder between atoms), resulting in depopulation of the p_z orbital of the halogen (blue). The nonbonding electrons of the p_x and p_y atomic orbitals remain largely intact. As a result, the atomic surface of the halogen shows an anisotropic distribution of charge potential from positive (blue) to negative (red). (C) Electrostatic potential of a halogenated uracil base. The electrostatic potentials (from −25 to 25 kcal/mol, from red to blue, respectively) of a uracil base halogenated at C5 (^{X}U) are compared for F, Cl, Br, and I, as viewed down the C−X bond.[5]

the result of formation of a covalent σ-bond to another atom, thereby establishing an electropositive crown directly opposite this bond (Figure 1); this is the σ-hole model of Politzer et al.[21,22] The remaining electrons in the p_x and p_y orbitals maintain their nonbonding electrons, resulting in a ring of negative charge perpendicular to the σ-bond. The extent of σ-hole formation is dependent on the polarizability of the halogen, which increases with the increasing size and decreasing electronegativity of the halogen (F < Cl < Br < I) (Figure 1),[23] and the improving electron withdrawing ability of the molecular group covalently bound to the halogen.[24] The polarized electrostatic surface of covalent halogens allows them to act as both X-bond donors and H-bond acceptors,[25] with each interaction being highly directional relative to this surface. In X-bonds, the acceptor ideally approaches the halogen linearly ($\Theta_1 \approx 180°$) toward the σ-hole (Figure 1), while H-bond donors approach the negative annulus approximately perpendicular to the σ-hole ($\Theta_1 \approx 100°$).[26,27]

The relative stabilizing potential of an X-bond in a biological context was first determined by Voth et al.[8] using a DNA Holliday junction system, in which a bromine X-bond (Br-bond) competed against a classic H-bond to define the isomer conformation of the junction. Holliday junctions are four-stranded DNA complexes involved in multiple cellular processes, including genetic recombination, DNA lesion repair, viral integration, restarting of stalled replication forks, and proper segregation of homologous chromosomes during meiosis.[28−30] Four-stranded DNA junctions are ideal for studying X-bonds, because their rigid structure is defined by a small number of specific intramolecular interactions.[31,32] Junctions under normal salt conditions adopt a stacked-X form, with arms stacked to form nearly continuous standard B-DNA duplexes[28−30] and characterized by having two continuous DNA strands connected by two crossover strands, each forming a tight U-turn (Figure 2). An H-bond from cytosine at position eight (C_8) to the phosphate of the preceding nucleotide is essential for stabilizing the junction in inverted repeat sequences of the form $CCG_3R_4T_5A_6Y_7C_8GG$ (where R is a purine and Y is a pyrimidine nucleotide).[32] The junction is further stabilized by an electrostatic interaction from a pyrimidine base at position 7 (Y_7) to the phosphate of A_6. This secondary interaction can be an H-bond or X-bond[5,33]

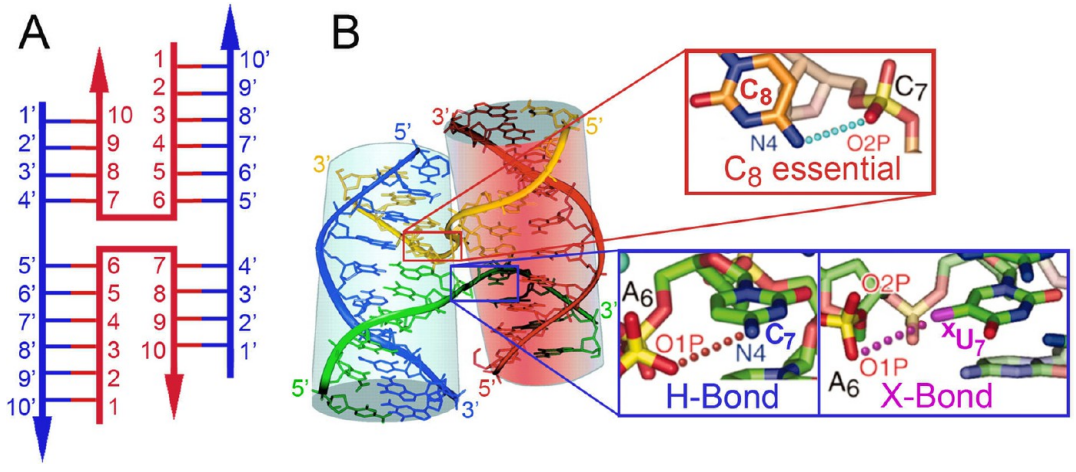


Figure 2. DNA Holliday junction. (A) Schematic. A four-stranded junction is shown in the stacked-X conformation, in which two continuous strands of DNA (blue and nucleotide labeled with a prime) are connected by two crossing strands that form a U-turn (red). (B) Structure of a DNA junction. The single-crystal structure of a Holliday junction[8] shows the four strands of DNA adopting primarily the B-DNA conformation, with the crossing strands (yellow and green) stabilized as sharp U-turns by H-bonds (dots) from the N4 amino of cytosine at position 8 (C_8) to the phosphate backbone of nucleotide C_7, which is essential for stabilization of the junction. An additional electrostatic interaction involves either an H-bond from the N4 amino of cytosine C7 or a competing X-bond from a halogenated uracil ($^{X}U_7$) back to the nucleotide phosphate of A_6.

dx.doi.org/10.1021/bi400590h | *Biochemistry* 2013, 52, 4891−4903

and, therefore, provides a framework for comparing the relative strengths of X-bonds versus H-bonds observed in crystal structures.

The initial studies by Voth et al.[8] used Holliday junctions with brominated uracil bases incorporated at the Y_7 position of $CCG_3R_4T_5A_6Y_7C_8GG$ to force one or two Br-bonds to compete against two classic H-bonds. A crystallographic competition assay was designed to estimate the relative energies for two distinct conformations of the Br-bond. In this assay, the junction-stabilizing interaction was determined by analysis of the isomeric form of the DNA. The isomer in which the halogen was observed at the center of the crossover strand indicated that it was stabilized by an X-bond (termed the X-isomer), while that with the halogen on the outside continuous strand was indicative of stabilization by an H-bond (the H-isomer) (Figure 3). The energies from the

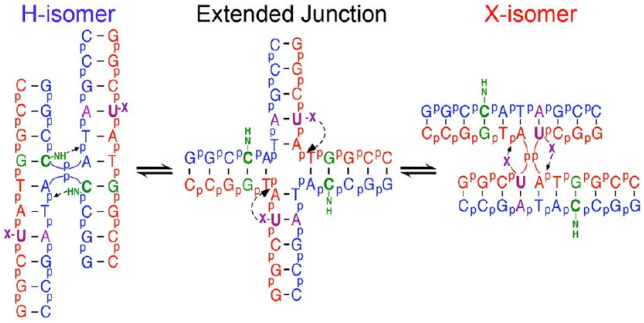


Figure 3. Assay for the competition of the X-bond against an H-bond in a DNA junction. The assembly of a DNA junction from two strands (blue) that contain C·G base pairs (green) with strands (red) that contain a halogenated uracil ^{X}U·A base pair (magenta) provides an experimental system for an X-bond competing against an H-bond. The junction can adopt one of two stable isomeric forms:[8] the H-isomer, which is stabilized by the H-bond from cytosine C7 to the DNA backbone (Figure 2B), and the X-isomer, stabilized by an X-bond form the ^{X}U to the phosphate backbone. The transition from the H- to the X-isomer involves unstacking of the junction arms to form an extended intermediate, which allows the junction arms to migrate and rearrange the stabilizing interactions.

crystallographic assay have subsequently been validated by differential scanning calorimetry (DSC) studies in solution.[34] Furthermore, the DSC studies delineate the enthalpic and entropic contributions of the stabilizing potential of the Br-bond.

In this study, we apply this same coupled crystallographic and differential scanning calorimetric strategy to relate the stabilization energy and structure of X-bonds in DNA Holliday junctions formed by F, Cl, Br, and I. We show here that the increase in polarizability of the halogen from F to Cl to Br to I is correlated with an increase in the enthalpic stabilizing potential of the respective X-bonds (the F-bond, Cl-bond, Br-bond, and I-bond), as predicted by the σ-hole model. However, the entropic contributions to the overall free energy of stabilization are seen to be significant in this system and differ for each halogen. Thus, the relation between the structures and their energies allows us to define the balance between enthalpic and entropic contributions to the interactions, which, ultimately, provides a better understanding for what defines an "optimal" X-bond in a biomolecular system.

MATERIALS AND METHODS

DNA Synthesis and Purification. DNA oligonucleotides for the F- and I-containing constructs were chemically synthesized by Midland Certified Reagent Co. and provided on the solid Controlled-Pore Glass (CPG) support with the final dimethoxytrityl protecting group intact. Sequences were subsequently purified by reverse phase high-performance liquid chromatography followed by size exclusion chromatography on a Sephadex G-25 column after detritylation. The Cl-containing constructs were synthesized as previously described.[35] The constructs for this study were designed as complementary sequences that assemble into four-stranded junctions, competing either one halogen (X1J) or two halogens (X2J) against two hydrogen bonds.

For this study, DNA constructs are generally named X1J and X2J, referring to junctions in which one and two halogenated uracil bases are incorporated into the DNA to provide one and two potential X-bonds, respectively, competing against two H-bonds (Table 1 and Figure 3). The control construct with nonhalogenated uracil bases (and thus only two H-bonds) is termed H2J.

Table 1. DNA Holliday Junction Constructs[a]

construct name	sequences complementary to (CCGAT**ACC**GG)	X:H
H2J	2(CCGGT**AU**CGG)	0:2
F2J	2(CCGGTA^{Fl}UCGG)	2:2
Cl2J	2(CCGGTA^{Cl}UCGG)	2:2
Cl1J	1(CCGGTA^{Cl}UCGG) + 1(CCGGT**AU**CGG)	1:2
Br2J	2(CCGGTA^{Br}UCGG)	2:2
Br1J	1(CCGGTA^{Br}UCGG) + 1(CCGGT**AU**CGG)	1:2
I2J	2(CCGGTA^{I}UCGG)	2:2
I1J	1(CCGGTA^{I}UCGG) + 1(CCGGT**AU**CGG)	1:2

[a]The names, strand sequences, and ratios of X-bonds competing with H-bonds (X:H) are listed for each DNA construct used in this study. The trinucleotide cores[32,39] in each sequence responsible for the stabilization of each junction are shown in bold. Halogenated uracils are denoted as ^{X}U, where X is fluorine, chlorine, bromine, or iodine attached at C5 of the uracil base. In each case, the sequences of the two strands that are complementary to the two common strands (CCGAT**ACC**GG) are listed.

Crystallization and Structure Solution. The F2J, Cl1J, Cl2J, I1J, and I2J constructs were crystallized by sitting drop vapor diffusion by methods as previously described for the Br1J and Br2J constructs.[8] Crystals were grown from solutions that contained 0.7 mM DNA, 25 mM sodium cacodylate buffer (pH 7.0), 10−25 mM calcium chloride, and 0.8−1.2 mM spermine, equilibrated against a reservoir of 30−40% aqueous MPD. Diffraction data were collected at liquid nitrogen temperatures at the Advanced Light Source (ALS) at the Lawrence Berkeley Laboratories at λ = 0.9 Å. All data were processed using DENZO and SCALEPACK of the HKL2000 software.[36] Structures were determined by molecular replacement using EPMR and using Br2J (PDB entry 2ORG), with the bromine removed as the initial search model.

The *C*2 symmetry and unit cell volumes of Cl1J, Cl2J, I1J, and I2J indicated that their asymmetric units were defined as two DNA strands, one continuous and one crossover strand, with the full DNA junction generated by the crystallographic 2-fold symmetry at the center of the four-stranded junction. The F2J construct also crystallized in the *C*2 space group; however, the 2-fold axis is shifted, resulting in a doubling of the *c*-axis

dx.doi.org/10.1021/bi400590h | *Biochemistry* 2013, 52, 4891−4903

Table 2. Parameters from Crystallization and Structure Solution of DNA Holliday Junction Constructs

	F2J	Cl1J	Cl2J	I1J	I2J
		Crystallographic Parameters			
PDB entry	4GSI	4GSG	4GQD	4GS2	4GRE
space group	*C*2	*C*2	*C*2	*C*2	*C*2
unit cell dimensions					
a (Å)	65.213	65.58	65.69	65.74	64.96
b (Å)	23.917	24.35	23.57	25.19	24.77
c (Å)	77.45	37.24	37.29	37.17	37.62
β (deg)	114.80	110.85	110.92	100.88	111.59
no. of unique reflections (for refinement)	3398	4947	3323	3562	5772
resolution (Å)	50−2.38	50−1.7	50−1.94	50−1.9	50−1.7
completeness (%)[a]	74.2 (51.2)	79.2 (41.4)	81.0 (54.2)	76.4 (69.9)	95.4 (90.3)
I/σ^a	48.93 (5.6)	28.25 (1.25)	29.54 (3.25)	11.31 (2.20)	14.14 (3.45)
R_{merge} (%)[a]	6.2 (26.5)	5.2 (32.3)	5.7 (26.8)	4.2 (24.4)	6.9 (25.4)
		Refinement Statistics			
molecular replacement correlation (%)	73	82	82	80	72
R_{cryst} (R_{free}) (%)	22.5 (29.1)	27.9 (32.1)	26.4 (30.0)	23.6 (26.3)	23.9 (25.5)
no. of atoms of DNA (solvent)	808 (99)	404 (158)	404 (95)	404 (73)	404 (74)
⟨*B* factor⟩ for DNA (solvent)	16.3 (11.8)	16.4 (24.7)	13.0 (21.4)	30.0 (35.1)	17.3 (21.7)
rmsd for bond lengths (Å)	0.002	0.001	0.005	0.005	0.007
rmsd for bond angles (deg)	0.6	0.4	1.0	0.9	1.1

[a]Values for the highest-resolution shell are given in parentheses.

length, which results in an asymmetric unit cell consisting of a full four-stranded junction.[32] The DNA−DNA contacts that define the crystal lattices are nearly identical for all of the structures in this study and previous studies.[8,32,37] Refinement for all constructs was conducted using the Crystallography and NMR System (CNS)[38] with rigid body refinement, simulated annealing, several rounds of positional and individual *B* factor refinement, and addition of solvent. Refinement was performed with models ambiguous for the isomeric form, not specifying the presence of the halogen (or the possible N2 atom of the guanine base complementary to the variable cytosine) at either the inside or outside position. As refinement approached convergence, occupancy titration analyses, as discussed in Results, were performed to determine the contribution of X- and H-isomer models. Structures were further refined and submitted to the PDB with the optimal X-isomer and H-isomer model contributions determined from occupancy titration analysis. Crystallographic and refinement statistics for all current structures are listed in Table 2.

Differential Scanning Calorimetry Studies. Differential scanning calorimetry (DSC) studies were performed to determine the stabilizing energies of the H2J, Br2J, and I2J constructs in solution as previously described.[34] The complementary DNA sequences of each construct were mixed in equimolar concentrations (varied from 15 to >300 μM) in 50 mM sodium cacodylate buffer (pH 7.0) and 1 mM calcium chloride, to approximate their crystallization conditions. The solutions were heated to 90 °C for 1 h and slowly reannealed to room temperature overnight. The energetic parameters for melting the constructs at each concentration were determined using a TA Instruments Nano DSC instrument with the pressure held constant at 3.0 atm. Each DNA sample was run against buffer in a heating cycle from 0 to 90 °C at a scanning rate of 1 °C/min with an equilibrium time of 900 s, which was repeated at least three times. DNA constructs were analyzed at multiple DNA concentrations to sample both the duplex and junction conformation, as junction formation has been shown to be DNA concentration-dependent.[39,40] Data were analyzed using the NanoAnalyze software from TA Instruments (New Castle, DE) in the same manner that has been previously published,[34] with the best fit determined according to the standard deviation of the fit. The data at lower concentrations were best fit using a two-state model scaled by a weighting term (Aw), to account for the presence of both double- and single-stranded DNAs. Samples in which the Aw term had indicated a much higher than predicted double-stranded concentration suggested the presence of a four-stranded junction component; consequently, the data for these samples were analyzed by applying a two-component, two-state model.

The similarity in melting temperatures (T_m) and melting enthalpies (ΔH_m) for the duplex forms between the single-component analysis at low DNA concentrations and the two-component analyses at high DNA concentrations supports this interpretation of the data. The presence of junctions was evident from the single-component analysis of the data for DNA concentrations from 20 to 100 μM but was not well resolved using the two-component analysis, and therefore, these data were excluded from analysis of the thermodynamic parameters. The ΔH_m for melting the duplex form of each construct was taken as the average ΔH_m for DNA concentrations from 15 to 20 μM along the low-temperature component from the two-component analysis of data at >100 μM DNA. The ΔH_m of the junction form of each construct was taken as the average of the higher temperature.

RESULTS

In these studies, we compare the crystallographic structures of seven deoxynucleotide sequences that were designed to form four-stranded DNA Holliday junctions. These junctions can exist in one of two competing isomeric forms stabilized by either a H-bond (adopting the H-isomer) or competing X-bond (as the X-isomer)[8,34] (Figure 3). Both the H- and X-bonds involve the phosphate oxygen of the preceding nucleotide and help to stabilize the tight U-turn of the crossover strand (Figure 2). The constructs are designed to force one X-bond (X1J) or two X-bonds (X2J) to compete against two H-bonds; the ratio

dx.doi.org/10.1021/bi400590h | *Biochemistry* 2013, 52, 4891−4903

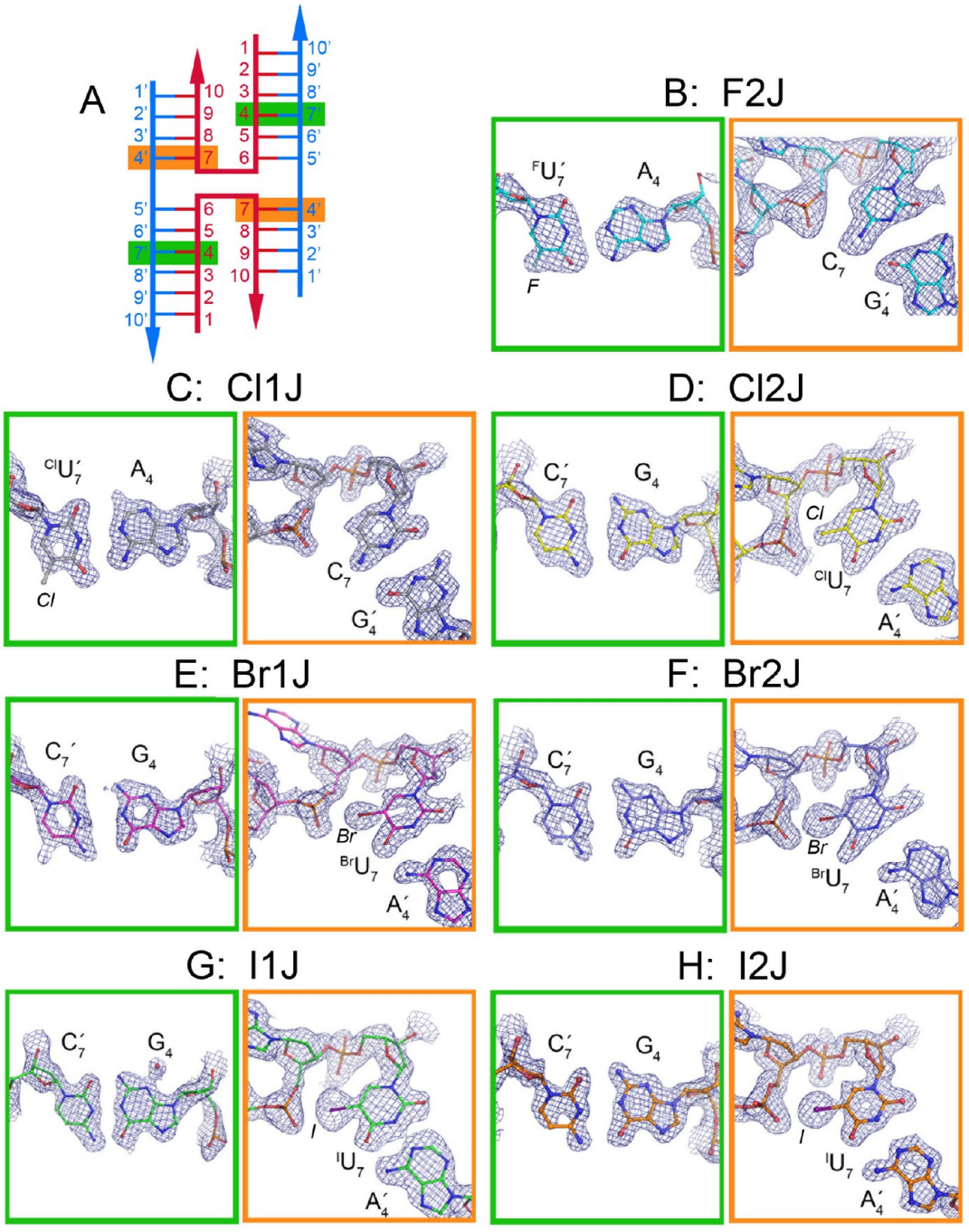


Figure 4. Electron density maps of F-, Cl-, Br-, and I-containing DNA junctions. (A) Schematic of the DNA junction (Figure 2A) with base pairs where electron densities will be shown highlighted with green and orange boxes (corresponding to green and orange boxed panels, respectively, for each structure). The orange boxes refer to the C_7 or $^{X}U_7$ at the inside crossing strands, allowing the junction to be stabilized by either an H- or X-bond, respectively, while the green boxes refer to the competing C_7' or $^{X}U_7'$ on the outside continuous strands (the prime indicating that the nucleotide is on the outside strand). (B–H) Electron density maps of competing positions for C_7·G_4 and $^{X}U_7$·A_4 base pairs. The $2F_o - F_c$ electron density maps (drawn at 1σ, as blue wires) of F2J (B), Cl1J (C), Cl2J (D), Br1J (E), Br2J (F), I1J (G), and I2J (H) (see Table 1). The maps shown are from the structure refined in the dominant conformer form for each construct, according to the occupancy titration (Figure 5): 60% H-isomer for F2J, 70% H-isomer for Cl1J, 63% X-isomer for Cl2J, 80% X-isomer for Br1J, 100% X-isomer for Br2J, 90% X-isomer for I1J, and 100% X-isomer for I2J. The positions of the halogens (italics) are labeled.

of the two isomeric forms can be used to estimate the relative stabilizing potentials of the competing interactions.[8,34] The X- and H-isomers are isomorphous and, thus, accommodated by the same crystal lattice interactions (in agreement with previously published results[8]).

This study was designed to determine the structure–energy relationships for F-, Cl-, Br-, and I-bonds in comparison to a standard H-bond in a DNA junction. For each halogen, we crystallized the DNA junctions as constructs that can form one or two X-bonds in competition against two H-bonds. The relative and absolute energies of each X-bond and the competing H-bond in the crystal were estimated from the ratio of isomers of the junction that are stabilized by X-bonds (X-isomer) or H-bonds (H-isomer) by crystallographic

dx.doi.org/10.1021/bi400590h | *Biochemistry* 2013, 52, 4891–4903

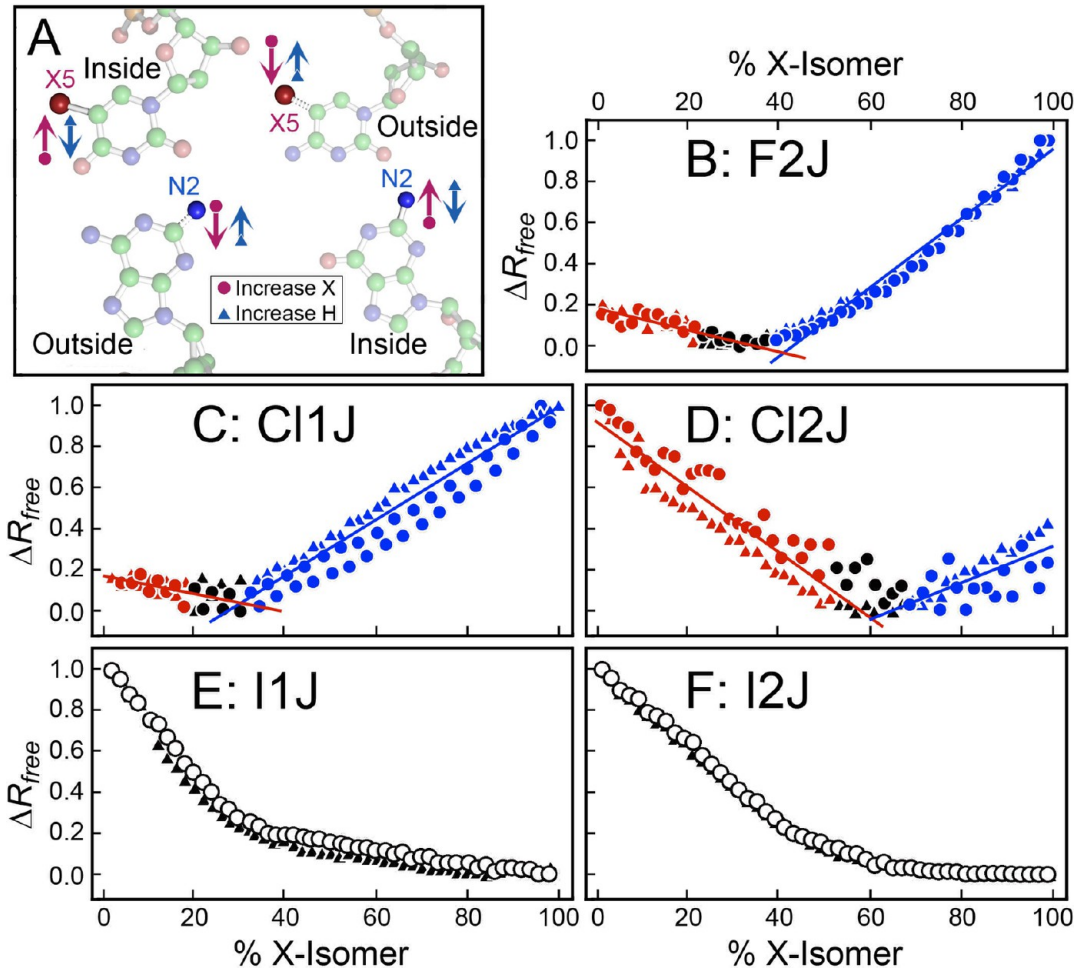


Figure 5. Crystallographic occupancy titrations. Each structure was refined at different percentages of X-isomer (%IsoH = 100 − %IsoX), ascending from 0 to 100% (circles) or descending from 100 to 0% (triangles), with the quality of each model monitored as the R_{free} value after refinement. In all cases, the entire 20 bp junction was modeled as either the X- or H-isomeric form. (A) Schematic showing the atoms most affected by an increasing percent of X-isomer (circles) or an increasing percent of H-isomer (triangles). The left-side base pair places the pyrimidine on the inside crossing strand, where increasing the percent of X-isomer increases the occupancy of the halogenated uracil (^{X}U) while decreasing that of the cytosine and decreases the N2 amino of the purine base of the outside strand (resulting in an ^{X}U·A base pair at this position, as shown). The right-side base pair places the pyrimidine on the outside continuous strand, where increasing the percent of X-isomer reduces the ^{X}U while increasing C and increases the N2 amino of the purine in the inside strand (resulting in a C·G base pair at this position). (B–D) Occupancy titrations of F2J, Cl1J, and Cl2J, respectively. The optimal percent of X-isomer (%X_{opt}) in each structure was determined by least-squares fitting of linear relationships for R_{free} values at %IsoX < %X_{opt} (red) and %IsoX > %X_{opt} (blue), with %X_{opt} calculated as the intersection from the slopes and *y*-intercepts of the two lines. Black symbols represent data that were not used in the calculation of the lines. (E and F) Occupancy titrations of I1J and I2J, respectively. In these cases, there were no clear R_{free} minima and, therefore, no attempt to determine %X_{opt}.

occupancy titration.[8,34] The refined crystal structures help to define the geometries of the X-bonds for each type of halogen. Finally, the energies for the X-bonds in solution were determined for these identical constructs by DSC, which further allowed delineation of the enthalpic and entropic contributions to the free energies of interactions.[34]

Crystallographic Determination of Isomeric Forms. The crystal structures show that all the constructs adopt the stacked-X four-stranded junction conformation, as expected. On the basis of the rmsds from superimpositions, the structures are 90% similar; i.e., 18 of the 20 base pairs are identical across the constructs, leaving 2 base pairs in each that show significant deviations. This indicates that the lattice packing, which is primarily end-to-end stacking of the duplex arms, is identical across the constructs, and therefore, the structural and thermodynamic variations are localized and can be compared across junctions without significant contributions from the remainder of their structures. We expect any variations in crystal lattice effects on the stability of the interactions to be minimal, and well within the errors of the energy measurements.

Not surprisingly, the primary differences among the structures lie in the two halogenated ^{X}U·A base pairs of each junction. For each construct, however, the ^{X}U·A base pairs can reside in one of two positions. In one isomeric form, the halogenated uracil bases sit on the outside strands as $^{X}U_7'{\cdot}A_4$ base pairs [in this nomenclature, we indicate the position along each strand by the subscript number and use the prime to indicate the base that is on the outside strand (see Figure 2)], placing the cytosines of the $C_7{\cdot}G_4'$ base pairs at the center of the crossing strand to form H-bonds that help to stabilize the junction (the H-isomer form). Alternatively, the halogenated uracil bases can sit at the center (as an $^{X}U_7{\cdot}A_4'$ base pair) to form potential X-bonds that stabilize the U-turn in the crossing

dx.doi.org/10.1021/bi400590h | *Biochemistry* 2013, 52, 4891–4903

strands, leaving the cytosine of the $C_7'{\cdot}G_4$ base pair on the outside (X-isomer form).

The initial assignments of the isomeric forms of the constructs come from an analysis of the electron densities at these positions (Figure 4). For each construct, the electron density maps were evaluated with the starting models refined with the isomer undefined (no halogen added to any of the pyrimidine bases), and difference ($F_o - F_c$) maps were analyzed for evidence of a dominant isomer. In particular, we looked for positive difference density associated with the halogen position around the pyrimidine at nucleotide 7 to determine whether this base is a cytosine or halogenated uracil, and with the N2 amino group of the purine at nucleotide 4 of the complementary strand to determine whether this base is a guanine paired to the cytosine or an adenine paired with the uracil base. Both the iodine (I1J and I2J) and the fully chlorinated (Cl2J) constructs appeared to be predominately X-isomer, similar to the previous brominated Br1J and Br2J junctions,[8] while Cl1J and F2J appeared to be primarily H-isomers. However, the maps could not be assigned entirely to one isomer form or the other. For example, when the Cl1J structure was refined entirely as the H-isomer, there was clearly residual density residing at the adenine base of the $^{Cl}U_7'{\cdot}A_4$ base pair where an N2 amino group would be for a guanine. Similarly, the N2 position of the guanine showed weaker density at the guanine, and there was evidence of some contribution of chlorine at C7 of the C7·G4′ base pair. This suggested that Cl1J was a mixture of both the X- and H-isomers.

The ratio of X- to H-isomers (%IsoX/%IsoH) can be used to estimate the difference in energy of the X-bond relative to the competing H-bond (ΔE_{X-H}).[8] Unfortunately, it is difficult to accurately determine this ratio from the electron density maps. Thus, the %IsoX/%IsoH present in the crystal structures of each construct was quantified from crystallographic data using an occupancy titration method, which had previously been validated in solution.[34] In each case, the structures were initially refined as a single model without specifying an isomeric form, with each of the N4·N7 positions defined as A·U base pairs. As the refinement neared convergence, an overlapping model with equal contributions of X- and H-isomers was generated to perform occupancy titration calculations.

In the occupancy titrations, the X-isomer contribution was increased from 0 to 100%, with a concomitant reduction in the H-isomer occupancy from 100 to 0% (Figure 5). This was repeated in the opposite direction, increasing the H-isomer and decreasing the X-isomer contributions, to control for any hysteretic effects. A single round of *B* factor refinement followed each occupancy titration, and the crystallographic *R* and R_{free} values were monitored as a function of the assigned occupancy of the X- and H-isomers. Mock occupancy titrations, in which iterative rounds of *B* factor refinement were performed with no change in the occupancy of the competing isomers, were performed to determine how *R* and R_{free} values are affected through progressive rounds of refinement. This served as the baseline, which was subsequently subtracted from the experimental *R* and R_{free} values for each titration point to produce the final titration curve. Titration data with clear minima were fit using KaleidaGraph with linear models for the decrease, and separately, the increase in R_{free}. The observed percent of X-isomer ($\%IsoX_{obs}$) for each titration was determined by solving for the point of intersection for the equations of the two lines on either side of the R_{free} minimum (Figure 5), and the corresponding percent of H-isomer observed ($\%IsoH_{obs}$) was taken to be 100% − $\%IsoX_{obs}$. The uncertainty in each $\%IsoX_{obs}$ was estimated by propagating the errors on the slopes and *y*-intercepts of these lines. Occupancy titration results reported here for Br1J and Br2J are in strong agreement with those from the previously published study.[8]

The resulting %IsoX/%IsoH ratios indicate that the F2J, Cl1J, Cl2J, and Br1J constructs crystallize as a mixed population of H- and X-isomeric junctions, consistent with what was observed in the electron density maps, with the F2J and Cl1J junctions dominated by the H-isomeric form and the Cl2J and Br1J junctions dominated by the X-isomeric form (Table 3).

Table 3. Crystallographic Competition Assay for Estimating X-Bonding Energies[a]

construct	% X-isomer	$\Delta E_{IsoX-IsoH}$ (kcal/mol)	$E_{X\text{-bond}}$ (kcal/mol)
F2J	40 ± 4	0.24 ± 0.07	−0.52 ± 0.06
Cl1J	30 ± 6	0.50 ± 0.08	−0.79 ± 0.12
Cl2J	63 ± 7	−0.29 ± 0.08	−0.79 ± 0.12
Br1J[8]	84 ± 4	−0.98 ± 0.15	−2.28 ± 0.11
Br2J[8]	≥95	ND[b]	ND[b]
I1J	≥80	≤−0.82	≤−2.1
I2J	≥95	ND[b]	ND[b]

[a]For each halogenated DNA junction construct, the percent of X-isomer observed ($\%X_{obs}$), as determined from the occupancy titrations (Figure 5), is used to determine the difference in the stabilization energy of the X-isomer relative to the H-isomer ($\Delta E_{IsoX-IsoH}$).[8] From these energies, the absolute energy for the H-bond ($E_{H\text{-bond}} = -0.64 \pm 0.07$ kcal/mol) and their competing X-bond energies ($E_{X\text{-bond}}$) were calculated as described in the text. [b]Not determined.

The Br2J junction populations are approximately 100% X-isomer, as previously reported.[8] The I1J and I2J junctions are predominantly in the X-isomer form, as expected from the very high polarizability and positive electrostatic potential of the iodine σ-hole. The ΔR_{free} of the I2J titration asymptotically approached 100% X-isomeric form and did not show any evidence of the H-isomer. The I1J titration, however, shows a linear reduction in ΔR_{free} that plateaus at 80% X-isomer. Therefore, the I1J construct was determined to be ≥80% X-isomer, while the I2J construct was ∼100% X-isomer. However, the nonlinear I1J and I2J titrations may reflect a significant amount of deiodination by the X-rays, and therefore, these estimates are not as reliable as for the other constructs.

The ratios of X- and H-isomers determined from the occupancy titrations were subsequently applied for the final refinement of each construct (Table 2). In addition to the structures showing a high degree of similarity between constructs, the H- and X-isomers also show a high degree of similarity within each construct after refinement, indicating that variations are local and that the structures and thermodynamics of the X-bonds of the X-isomer can be directly compared to those of the H-bonds in the H-isomer.

The observed %IsoX/%IsoH ratios could be used to estimate the difference in energy between the X- and H-isomers ($\Delta E_{IsoX-isoH}$) for each construct, according to eq 1. Direct comparison of the $\Delta E_{IsoX-IsoH}$ values show the following trend of stability for the X-isomer: F2J < Cl1J ≈ Cl2J < Br1J ≤ I1J (Table 3). This trend follows the polarizability of the halogens as expected from the σ-hole model for X-bonding. The Br2J and I2J junctions were entirely X-isomer in the crystals and,

dx.doi.org/10.1021/bi400590h | *Biochemistry* 2013, 52, 4891−4903

therefore, could not be directly analyzed for their $E_{IsoX-IsoH}$ values.

$$\Delta E_{IsoX-IsoH} = -RT \ln\left(\frac{\%IsoX}{\%IsoH}\right) \quad (1)$$

The X-bonds in the Cl1J and Cl2J constructs were essentially identical in structure, as will be discussed below, and therefore, we can further treat the energies of the individual Cl-bonds as being identical in the two junctions. With this assumption, we can estimate the absolute energies of interaction for both the Cl-bond ($E_{Cl\text{-}bond}$) and that of the competing H-bond ($E_{H\text{-}bond}$) from $\Delta E_{IsoX-IsoH}$ of Cl1J and Cl2J (eqs 2 and 3).

$$\Delta E_{IsoX-IsoH(Cl1j)} = E_{Cl\text{-}bond} - 2E_{H\text{-}bond} \quad (2)$$

$$\Delta E_{IsoX-IsoH(Cl2j)} = 2E_{Cl\text{-}bond} - 2E_{H\text{-}bond} \quad (3)$$

The two simultaneous equations can be solved to determine the interaction energy terms explicitly, with the resulting energies ($E_{Cl\text{-}bond} = -0.79 \pm 0.12$ kcal/mol, and $E_{H\text{-}bond} = -0.64 \pm 0.07$ kcal/mol) indicating that the Cl-bond is nearly equivalent to that of the competing H-bond in the stabilizing potential in their junctions.

The $E_{H\text{-}bond}$ value can be further used to determine the X-bonding energies for the other halogens by applying eqs 2 and 3 appropriately (Table 3). The energy for the F-bond ($E_{F\text{-}bond}$) is directly calculated from the $\%IsoX_{obs}$ for F2J using eq 3. For the Br-bond ($E_{Br\text{-}bond}$), we use the $\%IsoX_{obs}$ from Br1J and eq 3, because only the mixed ^{Br}U plus U junction is observed in the crystal.[8] The energies derived from the crystallographic assays will be compared to those determined calorimetrically to determine the effect of the crystalline versus solution environments on the stabilizing potential of the X-bonds.

X-Bond Geometries. We would expect that the geometry of each X-bond would reflect the varying stabilizing potential afforded by each halogen type. We can also imagine, however, two scenarios: one in which the junction imposes a geometry that specifies the interaction energy for the X-bond in each construct and, alternatively, one in which the intrinsic stabilizing potential specifies the geometry of the X-bond in the junction. The previous study of Br1J and Br2J would suggest a model that is both, where the intrinsic potential of the interaction pulls against the intrinsic structure of the junction, and the resulting energy reflects the geometry that is observed. A detailed analysis of the X-bonding geometry through the series of halogens from F to I appears to support this model as well (Figure 6).

The X-bonds show a systematic approach toward a more ideal geometry for the interaction, consistent with the F < Cl < Br < I order in terms of the strength of the interaction. The donor−acceptor distances, as measured by the percent reduction in the sum of the van der Waals radii ($\%\sum R_{vdW}$), become shorter, and the approach of the acceptor toward the halogen, as measured by the Θ_1 angle (Figure 1), becomes increasingly linear going from F to Cl to Br to I (Table 4).

The F···O distance in the F2J X-isomer is slightly longer than the sum of the van der Waals radii of the two interacting atoms ($\%\sum R_{vdW} > 100\%$). Thus, although the directional approach of the acceptor phosphate oxygen is toward the halogen σ-hole, this particular fluorine interaction is not considered to be an X-bond; the slightly negative $E_{F\text{-}bond}$ may be attributed to a van der Waals-type interaction.

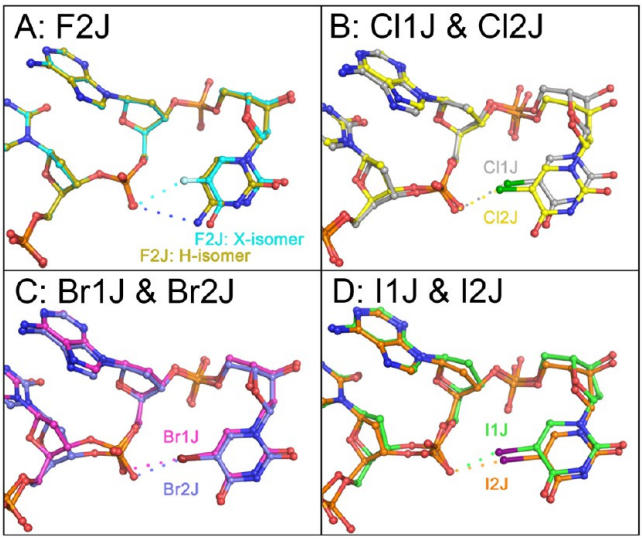


Figure 6. Comparison of geometries of potential X-bonds. (A) Comparison between the X-isomer (cyan carbons) and H-isomer (gold carbons) forms of F2J. Superposition of the two isomers shows nearly identical structures for the atoms in the backbone, with the primary difference appearing as a slight shift of the C_7 base to point the N4 amino group toward the phosphate backbone to form the H-bond that stabilizes the H-isomer, or a slight shift of $^{F}U_7$ to form the X-bond of the X-isomer. (B) Comparison of Cl1J (gray carbons) and Cl2J (yellow carbons). The superposition of these two independently refined structures shows slight shifts in both the $^{Cl}U_7$ bases and the associated phosphate groups of the backbone, resulting in nearly identical geometries of the two Cl-bonds. (C) Comparison of Br1J (magenta carbons) and Br2J (violet carbons). The superposition of these two independently refined structures shows the positions of the $^{Br}U_7$ bases to be nearly identical, but rotation about the β-torsion angle renders the Br-bond of Br1J to be longer than that in Br2J. (D) Comparison of I1J (green carbons) and I2J (orange carbons). The superposition of these two independently refined structures shows slight shifts in the positions of the $^{I}U_7$ bases, but nearly identical positions of the interacting phosphate oxygens. This results in a slightly longer but more linearly oriented I-bond in I2J as compared to that in I1J.

Table 4. X···O$^-$ Interaction Geometries of X-Bonds[a]

construct	X···O distance (Å)	$\%\sum R_{vdW}$ (%)	Θ_1 (deg)
F2J	3.20	105	153.5
Cl1J	2.95	89	152.0
Cl2J	2.88	87	146.0
Br1J	3.32	98	167.2
Br2J	2.87	84	163.2
I1J	2.92	83	164.4
I2J	3.01	85	170.7

[a]Halogen (X) to O1P phosphate oxygen ($O^{-1/2}$) geometries in terms of the X···$O^{-1/2}$ distances, the percent of the sum of the standard van der Waals radii ($\%\sum R_{vdW}$), and the angular approach of the acceptor to the halogen (Θ_1) are listed for the constructs in their X-isomeric forms.

The Cl···O distances in the X-isomer structures of Cl1J and Cl2J are reduced by 11 and 13%, respectively, relative to their standard $\sum R_{vdW}$. The slight difference between the two is seen as a shift of less than 0.1 Å between the Cl donor and phosphate oxygen acceptor atom, with a difference in Θ_1 of only 6° between the two constructs; thus, we consider the Cl-bonds of the Cl1J and Cl2J constructs to be nearly identical in

structure and in their $E_{Cl\text{-}bond}$ energies. In contrast, the Br-bond in the Br1J X-isomer is only 2% shorter that the $\sum R_{vdW}$ (3.34 Å), while that in the Br2J X-isomer is 16% shorter than the $\sum R_{vdW}$.[8] This is equivalent to an ~0.5 Å difference in the Br···O distances, resulting from a rotation about the β angle of the preceding nucleotide, shifting the position of the phosphate relative to halogenated uracil. The X-isomer forms are nearly identical in I1J and I2J, with the I···O distances being shorter than the $\sum R_{vdW}$ by 17 and 15%, respectively, with Θ_1 angles approaching 180°. Thus, as the σ-hole becomes more pronounced, it imposes a more ideal geometry on the respective X-bond.

X-Bonding Energies in Solution. Although the crystallographic studies provided direct measures of structure−function relationships for the X-bonds, they do not tell us anything about the thermodynamics of the interaction in solution. Furthermore, we were not able to determine energies for the most stable Br2J construct or the two iodinated constructs. To fill these gaps, we designed a set of melting studies that apply differential scanning calorimetry (DSC)[34] to measure the relative energies of the X- and H-bonds in Br2J and I2J and the H-bond in the control H2J construct. For these analyses, we make the assumption that Br2J and I2J are entirely X-isomer forms while H2J is an entirely H-isomer form. Junctions that are mixtures of the isomeric forms would greatly complicate the thermodynamic analyses; consequently, we did not apply this approach to studying the X-bonds of F2J, Cl1J, Cl2J, Br1J, or I1J.

The energies of interaction at the junction core for each construct were segregated from the standard DNA base pairing and base stacking energies by exploiting the concentration-dependent transition from duplex to junction[39] (Figure 7). The constructs annealed at low DNA concentrations show a single melting component in the DSC profile, which is characteristic of melting duplex to single-stranded DNAs. At high DNA concentrations, the DSC profiles show two melting components, each with an associated melting temperature (T_m). The lower-temperature component of the high-concentration scans had the same T_m and melting enthalpies (ΔH_m) as that of the low-concentration scans and, therefore, could be attributed to the melting of the duplex form. The higher-T_m component of these same scans was associated with a higher ΔH_m and, thus, was attributed to melting of the junction to single-stranded DNA (Figure 8). Consequently, the stabilizing enthalpy of the junction core interactions can be determined by subtracting the ΔH_m of this high-T_m junction component from that of the low-T_m duplex component, after appropriate extrapolation of each to a reference temperature (T_{ref}) of 25 °C. The entropic contributions can then be calculated from ΔH_m and T_m and extrapolated to a reference temperature for direct comparison. The total free energy of interaction at the reference temperature for the X- and H-isomers is the negative of the resulting melting free energy [$\Delta G_{m(Tref)}$].

The energy of the X-bond relative to the H-bond in solution is calculated by subtracting components of the junction stabilizing energy of the H2J junction core (which presumably is in the H-isomer form) from those of the Br2J and I2J cores (which we have shown to be predominantly X-isomer). This removes the contribution of accessory stabilizing interactions, including, for example, the H-bond from the cytosine at N_8 to the preceding phosphate oxygen that is essential for formation of the junction in inverted repeat sequences.[39] In this manner, the relative X-bond versus H-bond enthalpy, entropy, and total

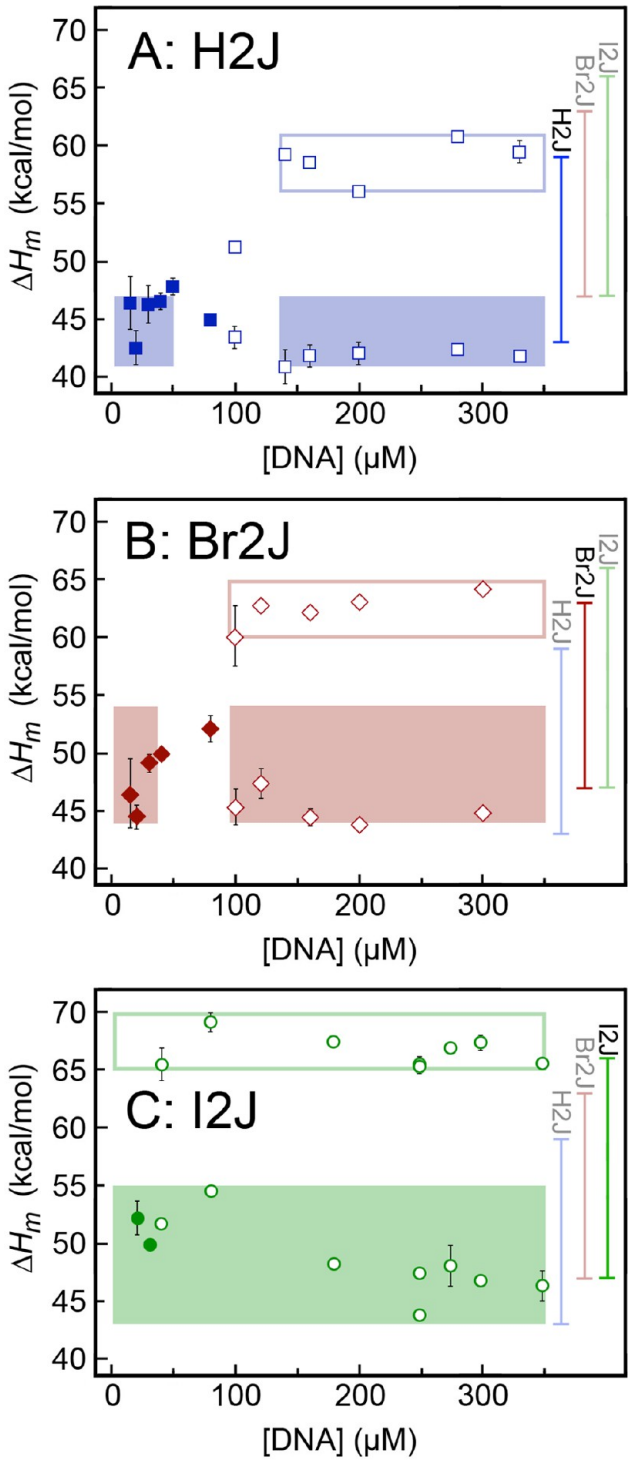


Figure 7. Comparing the enthalpies of melting (ΔH_m) of duplex and junction DNAs for H2J (A), Br2J (B), and I2J (C). At low DNA concentrations, the melting profiles are characteristically single-component (filled symbols), while at high DNA concentrations, the profiles show two distinct components (empty symbols). The ΔH_m values of the duplex (at low DNA concentrations) or the duplex component (at high DNA concentrations) were analyzed with the data points encompassed within the underlying solid boxes, while the ΔH_m values of the junction were analyzed from data enclosed in the open boxes. The difference in the ΔH_m values of the junction and the duplex are shown as sidebars for H2J (blue), Br2J (red), and I2J (green).

dx.doi.org/10.1021/bi400590h | Biochemistry 2013, 52, 4891−4903

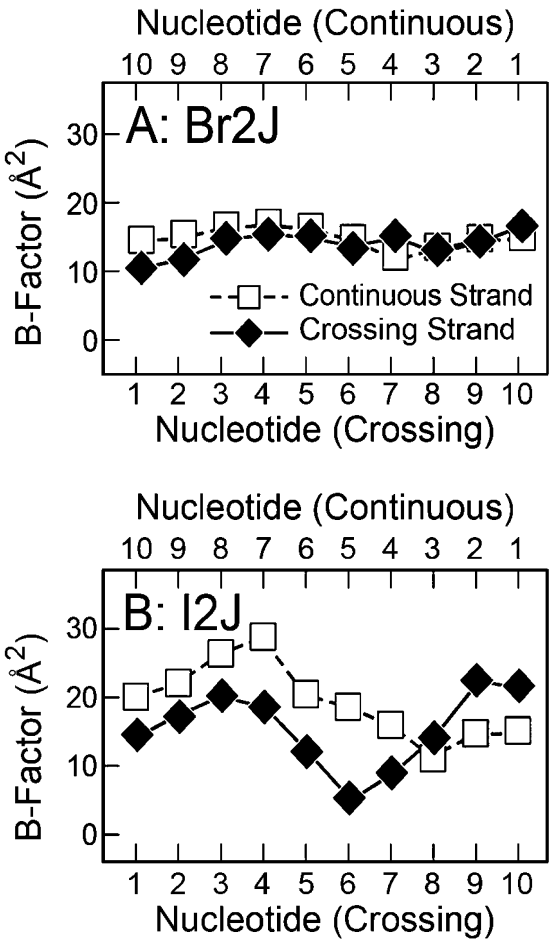


Figure 8. Temperature factors in Br2J and I2J crystal structures. The average temperature factors (*B* factors) of the nucleotides along the continuous outside strand (□) and inside crossing strands (◆) of the Br2J structure are compared to those of the I2J structure (the standard deviations for each *B* factor are ≤2.0 Å^2, which are smaller than the size of the symbols). The inside and outside strands of Br2J, therefore, show nearly identical *B* factors for their complements, while I2J shows significant deviations, particularly at nucleotide positions 5−7.

free energy of stabilization can be determined for the Br-bond of Br2J and the I-bond of I2J (Table 5).

The results indicate that the iodine X-bond is more enthalpically favorable than that of bromine, as predicted for the increased polarizability of the larger halogen. However, this more favorable enthalpy comes at an entropic cost, with the Br-bond being associated with a slight increase in entropy, while the I-bond shows a significant negative ΔS. As a result, the free energy of stabilization is less favorable for the I-bond than for the Br-bond. Thus, the balance between the enthalpic and entropic terms of the free energy results in bromine forming the optimal halogen X-bond in this system. This is consistent with the results of the crystallographic studies.

DISCUSSION

We present here a comprehensive and systematic structural and thermodynamic study of the structure−energy profiles of F-, Cl-, Br-, and I-bonds within a context of a common biological model system. The DNA Holliday junction as a model system provides a rigid structure into which specific X-bonds were engineered, in a fashion analogous to that of the host−guest systems used to study helical propensities in proteins.[41,42]

Table 5. Thermodynamic Parameters for the Stabilization of DNA Junctions Containing X-Bonds in the Br2J and I2J Constructs[a]

construct	$\Delta H^{25°C}_{(J-D)}$ (kcal/mol)	$\Delta S^{25°C}_{(J-D)}$ (cal mol^{-1} K^{-1})	$\Delta G^{25°C}_{(J-D)}$ (kcal/mol)
Br2J	−16.7 ± 0.9	−39 ± 3	−5.3 ± 1.3
I2J	−19.0 ± 0.6	−55 ± 2	−2.6 ± 0.9
H2J	−13.1 ± 0.9	−43 ± 3	−0.4 ± 1.3
construct	$\Delta\Delta H^{25°C}_{(X-H)}$ (kcal/mol)	$\Delta\Delta S^{25°C}_{(X-H)}$ (cal mol^{-1} K^{-1})	$\Delta\Delta G^{25°C}_{(X-H)}$ (kcal/mol)
Br2J	−3.6 ± 1.3	4.2 ± 4	−4.8 ± 1.8
I2J	−5.9 ± 1.1	−12.0 ± 4	−2.3 ± 1.5

[a]The melting enthalpies and melting temperatures for each construct, as measured by DSC (Figure 7), were converted to the stabilizing enthalpies, entropies, and free energies and extrapolated to 25 °C.[34] The difference in each thermodynamic parameter for the junction and duplex forms results in the stabilizing energies at the trinucleotide core [$\Delta H^{25°C}_{(J-D)}$, $\Delta S^{25°C}_{(J-D)}$, and $\Delta G^{25°C}_{(J-D)}$] for the Br2J, I2J, and H2J constructs. The values at 25°C are determined by subtracting the stabilization energy of duplex arms from the junction after extrapolating to 25 °C. The energy terms for the X-bond relative to the H-bond [$\Delta\Delta H^{25°C}_{(X-H)}$, $\Delta\Delta S^{25°C}_{(X-H)}$, and $\Delta\Delta G^{25°C}_{(X-H)}$] are calculated by subtracting the respective junction minus duplex terms for the Br2J or I2J junction from the H-bonded H2J junction.

Within this DNA system, we see a strong relationship between the increasing stabilizing potential and a more ideal geometry of the X-bond. It is not surprising that a more ideal geometry correlates with a more stabilizing interaction; however, in this case, we see that as the ability of the halogen to form an X-bond improves, going from F to Cl to Br to I, the X-bond becomes shorter and more linearly aligned with the σ-hole. Consequently, when considering the engineering aspects of using any of the common halogens, we must consider how, for example, a stronger I-bond could more effectively pull against an opposing force associated with the distortion of a DNA junction, in our case, or, more generally, with conformational perturbations to any biomolecular structure, including proteins.

The energies determined for X-bonds and the competing H-bonds are consistent between the complementary crystallographic and solution assays, indicating that the structural details seen in the crystal can be directly correlated to X-bonding properties in solution. The crystal structures of the DNA junctions, however, do not tell the entire story. A comparison of the energies for Br and I constructs shows that the X-bond energies in the crystals generally reflect the contributions of the interaction to $\Delta G°$, indicating that the population in the crystal system reflects that in solution, as we had initially assumed. The I-bond in the crystal has an energy that is nearly identical to the $\Delta G°$ of stabilization in solution. There would appear to be a discrepancy in the energies for the Br-bond in the crystal versus solution; however, we need to emphasize that the crystallographic energy was estimated for the Br1J construct, while the solution energy was for the Br2J construct, and that the longer Br-bond in the single-halogenated junction construct was expected to be approximately half the stabilizing energy of the shorter Br-bond in the two-bromine construct,[8] which is exactly what was observed upon comparison of the two different systems.

It is clear, however, that bromine and iodine have different effects on the dynamics of the system, as measured by the ΔS of melting. The slight increase in entropy associated with burying a bromine at the center of the junction reflects the hydrophobic

dx.doi.org/10.1021/bi400590h | *Biochemistry* 2013, 52, 4891−4903

property of the halogen, while the entropy of the system is significantly reduced when the larger iodine atom is buried. We can ask at this point whether this entropic effect stems primarily from conformational or solvent effects. It is easy to see how placing the larger, more hydrophobic iodine atom into this relatively small space can result in a more tightly packed junction interior and/or differences in exposure of the halogen surface to the bulk solvent might significantly affect the dynamics of the system.

To address the solvation component, we started by calculating the difference in solvent accessible surfaces between the X- and H-isomeric forms (ΔSAS) of the Br2J and I2J constructs (Tables 6 and 7). Summing the ΔSAS of the halogen in the ^{X}U·A base pair (ΔSAS_X) with that of the N2 amino group of the guanine in the opposing C·G base pair (ΔSAS_{NH_2}) provides an estimate for the changes in solvation free energies associated with taking an exposed halogen of the H-isomer and burying it in the X-isomer form ($\Delta SAS_{IsoX-IsoH}$). The result of this analysis is that burying a halogen in the junction is slightly favorable in terms of solvent effects ($\langle \Delta SASN_{IsoX-IsoH} \rangle$ = −0.4 kcal/mol, SD = 0.3 kcal/mol, normalized per halogen), becoming more favorable as the size of the halogen increases (Table 7). This solvent effect is consistent with the slightly positive Δ*S* observed for melting of the Br2J construct, and therefore, formation of an X-bond in a buried pocket is favored in terms of the solvent entropy (or hydrophobicity). We would expect this to be true of all the halogens in this study.

Table 6. Atomic Solvation Parameters (ASP) for Halogens and Amino Groups[a]

atom or group	ASP (kcal mol^{-1} Å^{-2})
F	0.041
Cl	0.033
Br	0.047
I	0.048
NH_2	−0.043

[a]The partition coefficients of the halogens from halogenated benzenes and the amino group from anilines were used to determine the differences in solvent free energies (SFE) of each atom type, as previously described.[53−55] The SFE divided by the solvent accessible surface (SAS) of the atoms defines the ASP for that atom type or group of atoms.

The I-bond of I2J, in contrast, shows the opposite effect, where melting of the junction results in a positive Δ*S*, or that burying the halogen to form an X-bond is entropically disfavored. We can attribute this effect to a loss of conformational entropy resulting from formation of a strong I-bond or to crowding associated with placing this large halogen in a relatively small cavity. This interpretation is supported by comparing the average temperature or *B* factors of the nucleotides in the Br2J and I2J structures (Figure 8). From this analysis, we see that both the outside and crossing strands of the junction follow a sinusoidal pattern with peaks at nucleotides N_3 and N_4 and valleys at nucleotides N_6 and N_7. For the Br2J junction, the outside and inside crossing strands show the same pattern with very little variation from the average, suggesting that the variations in *B* factor in this structure likely reflect the effect of the sequence. The I2J structure, however, shows that the *B* factors of central nucleotides N_3−N_7 of the crossing strand are significantly lower than those of the outside strand, with the lowest average value at positions N_6 and N_7. Recall that these are the two nucleotides that are engaged in the I-bond, with the halogenated uracil at N_7 forming the strong X-bond to the phosphate oxygen at N_6. *B* factor analysis, therefore, suggests that the loss of entropy in forming an I-bond is associated with a more rigid local environment. This increased rigidity and concomitant loss of entropy may result from enthalpy−entropy compensation (where formation of a strong bonding interaction reduces the dynamic flexibility of the local structure),[43−45] a crowding effect (where placing a very large atom into a small cavity reduces the space available for the dynamic motion of the atoms surrounding the cavity),[46−48] or both. If enthalpy−entropy compensation were the dominant factor, we would expect the Br-bond to also induce a more rigid and less entropic structure, but this is not the case. Thus, we propose here that the dominant effect is that the cavity in the DNA junction system is just large enough to accommodate a halogen about the size of a bromine, but the much larger iodine atom imposes a crowding effect that renders the local

Table 7. Contributions of Solvation to the Isomeric Forms of Halogenated DNA Junctions[a]

junction	ΔSAS_X (Å^2)	ΔSFE_X (kcal/mol)	ΔSAS_{NH_2} (Å^2)	ΔSFE_{NH_2} (kcal/mol)	$\Delta SFE_{IsoX-IsoH}$ (kcal/mol)	$\Delta SFEN_{IsoX-IsoH}$ (kcal/mol)
F2J	−7.0	−0.28	−2.2	0.10	−0.19	−0.19
Cl1J	−11.1	−0.37	−0.9	0.04	−0.14	−0.33
Cl2J	−8.7	−0.29	3.2	−0.14	−0.42	−0.42
Br1J	−6.1	−0.32	1.7	−0.07	−0.23	−0.39
Br2J	−11.4	−0.54	1.1	−0.05	−0.59	−0.59
I1J	−17.1	−0.82	1.0	−0.04	−0.45	−0.86
I2J	−15.7	−0.75	3.3	−0.14	−0.89	−0.89

[a]For each DNA junction, the differences in solvent accessible surfaces are calculated (Discovery Studio, Accelrys, San Diego, CA; probe radius of 1.4 Å) for the halogen of the $^{X}U_7$ nucleotide (ΔSAS_X) and the N4 amino group of the cytosine at position C7 (ΔSAS_{NH_2}) from the crystal structures of the X-isomer and H-isomer forms of the junction. The differences in solvent free energies between the two isomeric forms for the halogen (ΔSFE_X) and the amino group (ΔSFE_{NH_2}) are calculated as the products of the respective ΔSAS and their associated atomic solvation parameters (Table 6). The differences in solvation between the competing halogen and amino groups of the X-isomer (IsoX) and H-isomer (IsoH) ($\Delta SFE_{IsoX-IsoH}$) for each junction construct are calculated as the sums of ΔSFE_X and ΔSFE_{NH_2}, normalizing for the molar ratio of X-bonding halogenated uracil and H-bonding cytosine bases at the N7 positions of the constructs. For the X1J constructs, there is one potential X-bonding base to two H-bonding bases, and therefore, $\Delta SFE_{IsoX-IsoH} = {}^1/_2\Delta SFE_X + \Delta SFE_{NH_2}$. The $\Delta SFE_{IsoX-IsoH}$ for each construct with an X:H ratio of 1:2 is recalculated with ΔSFE_X to yield a normalized $\Delta SFE_{IsoX-IsoH}$ per halogen ($\Delta SFEN_{IsoX-IsoH}$).

dx.doi.org/10.1021/bi400590h | *Biochemistry* 2013, 52, 4891−4903

environment and, thus, the overall free energy of the system less stable than expected from the enthalpy of the X-bond.

In the context of applying X-bonding as a tool for biomolecular engineering, this study indicates that effects on the electrostatic properties of halogens, which can be variously modeled,[21,27] define the enthalpy of the X-bonding interaction. However, the relationship between the size of the halogen and the space into which it fits, and potentially the strength of the interaction (through enthalpy−entropy compensation[43,44]), will affect the dynamics and corresponding entropy of the system, which together with the enthalpy determines the overall free energy of the X-bonding interaction.[45] In this DNA system, the interplay between enthalpic and entropic terms indicates that bromine is the optimal halogen for forming an X-bond to stabilize the junction, even though iodine is more polarizable and, thus, is expected to form a stronger interaction. In protein−ligand systems, however, one must determine the relationship between these two competing thermodynamic components separately,[49] which in turn begs for a method that can accurately model halogen interactions in molecular dynamics simulations.[27,50−52]

AUTHOR INFORMATION

Corresponding Author

*E-mail: shing.ho@colostate.edu. Phone: (970) 491-0569. Fax: (970) 491-0494.

Present Address

[§]A.R.V.: National Institutes of Health, Bethesda, MD 20892.

Funding

This work was funded in part by a grant from the National Science Foundation (CHE-1152494 to P.S.H.) and a grant from the National Institutes of Health (NIH CA84487 to L.C.S.).

Notes

The authors declare no competing financial interest.

ABBREVIATIONS

H-bond, hydrogen bond; X-bond, halogen bond; MPD, 2-methyl-2,4-pentanediol; PDB, Protein Data Bank; rmsd, root-mean-square deviation.

REFERENCES

(1) Kemsley, J. (2012) Halogen bonding defined. *Chem. Eng. News 90*, 36−38.

(2) Metrangolo, P., and Resnati, G. (2001) Halogen bonding: A paradigm in supramolecular chemistry. *Chem.—Eur. J. 7*, 2511−2519.

(3) Metrangolo, P., Meyer, F., Pilati, T., Resnati, G., and Terraneo, G. (2008) Halogen bonding in supramolecular chemistry. *Angew. Chem., Int. Ed. 47*, 6114−6127.

(4) Fourmigué, M. (2009) Halogen bonding: Recent advances. *Curr. Opin. Solid State Mater. Sci. 13*, 36−45.

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

(6) Scholfield, M. R., Zanden, C. M., Carter, M., and Ho, P. S. (2013) Halogen bonding (X-bonding): A biological perspective. *Protein Sci. 22*, 139−152.

(7) Parisini, E., Metrangolo, P., Pilati, T., Resnati, G., and Terraneo, G. (2011) Halogen bonding in halocarbon-protein complexes: A structural survey. *Chem. Soc. Rev. 40*, 2267−2278.

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

(9) Voth, A. R., and Ho, P. S. (2007) The role of halogen bonding in inhibitor recognition and binding by protein kinases. *Curr. Top. Med. Chem. 7*, 1336−1348.

(10) Liao, J. J. L. (2007) Molecular recognition of protein kinase binding pockets for design of potent and selective kinase inhibitors. *J. Med. Chem. 50*, 409−424.

(11) Fedorov, O., Huber, K., Eisenreich, A., Filippakopoulos, P., King, O., Bullock, A. N., Szklarczyk, D., Jensen, L. J., Fabbro, D., Trappe, J., Rauch, U., Bracher, F., and Knapp, S. (2011) Specific clk inhibitors from a novel chemotype for regulation of alternative splicing. *Chem. Biol. 18*, 67−76.

(12) Hardegger, L. A., Kuhn, B., Spinnler, B., Anselm, L., Ecabert, R., Stihle, M., Gsell, B., Thoma, R., Diez, J., Benz, J., Plancher, J. M., Hartmann, G., Isshiki, Y., Morikami, K., Shimma, N., Haap, W., Banner, D. W., and Diederich, F. (2011) Halogen bonding at the active sites of human cathepsin l and mek1 kinase: Efficient interactions in different environments. *ChemMedChem 6*, 2048−2054.

(13) Xu, Z., Liu, Z., 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., and Zhu, W. (2011) Utilization of halogen bond in lead optimization: A case study of rational design of potent phosphodiesterase type 5 (pde5) inhibitors. *J. Med. Chem. 54*, 5607−5611.

(14) Gribble, G. W. (2003) The diversity of naturally produced organohalogens. *Chemosphere 52*, 289−297.

(15) Gribble, G. W. (2004) Natural organohalogens: A new frontier for medicinal agents? *J. Chem. Educ. 81*, 1441−1449.

(16) Lu, Y. X., Shi, T., Wang, Y., Yang, H. Y., Yan, X. H., Luo, X. M., Jiang, H. L., and Zhu, W. L. (2009) Halogen bonding: A novel interaction for rational drug design? *J. Med. Chem. 52*, 2854−2862.

(17) Bissantz, C., Kuhn, B., and Stahl, M. (2010) A medicinal chemist's guide to molecular interactions. *J. Med. Chem. 53*, 5061−5084.

(18) Lu, Y., Liu, Y., Xu, Z., Li, H., Liu, H., and Zhu, W. (2012) Halogen bonding for rational drug design and new drug discovery. *Expert Opin. Drug Des. 7*, 375−383.

(19) Kraut, D. A., Churchill, M. J., Dawson, P. E., and Herschlag, D. (2009) Evaluating the potential for halogen bonding in the oxyanion hole of ketosteroid isomerase using unnatural amino acid mutagenesis. *ACS Chem. Biol. 4*, 269−273.

(20) Liu, L., Baase, W. A., and Matthews, B. W. (2009) Halogenated benzenes bound within a non-polar cavity in T4 lysozyme provide examples of I···S and I···Se halogen-bonding. *J. Mol. Biol. 385*, 595−605.

(21) Clark, T., Hennemann, M., Murray, J. S., and Politzer, P. (2007) Halogen bonding: The σ-hole. Proceedings of "modeling interactions in biomolecules ii", Prague, September 5th−9th, 2005. *J. Mol. Model. 13*, 291−296.

(22) Politzer, P., Murray, J. S., and Lane, P. (2007) σ-hole bonding and hydrogen bonding: Competitive interactions. *Int. J. Quantum Chem. 107*, 3046−3052.

(23) Politzer, P., Murray, J. S., and Clark, T. (2010) Halogen bonding: An electrostatically-driven highly directional noncovalent interaction. *Phys. Chem. Chem. Phys. 12*, 7748−7757.

(24) Riley, K. E., Murray, J. S., Fanfrlik, J., Rezac, J., Sola, R. J., Concha, M. C., Ramos, F. M., and Politzer, P. (2011) 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. 17*, 3309−3318.

(25) Lu, Y., Wang, Y., Xu, Z., Yan, X., Luo, X., Jiang, H., and Zhu, W. (2009) C-X···H contacts in biomolecular systems: How they contribute to protein-ligand binding affinity. *J. Phys. Chem. B 113*, 12615−12621.

(26) Ramasubbu, N., Parthasarathy, R., and Murrayrust, P. (1986) Angular preferences of intermolecular forces around halogen centers: Preferred directions of approach of electrophiles and nucleophiles around the carbon halogen bond. *J. Am. Chem. Soc. 108*, 4308−4314.

(27) Carter, M., Rappe, A. K., and Ho, P. S. (2012) Scalable anisotroplic shape and electrostatic models for biological bromine halogen bonds. *J. Chem. Theory Comput. 8*, 2461−2473.

(28) Lilley, D. M. J. (1999) Structures and interactions of helical junctions in nucleic acids. In *Oxford Handbook of Nucleic Acid Structure* (Neidle, S., Ed.) pp 471−498, Oxford University Press, New York.
(29) Lilley, D. M. J. (2000) Structures of helical junctions in nucleic acids. *Q. Rev. Biophys. 33*, 109−159.
(30) Eichman, B. F., Ortiz-Lombardia, M., Aymami, J., Coll, M., and Ho, P. S. (2002) The inherent properties of DNA four-way junctions: Comparing the crystal structures of Holliday junctions. *J. Mol. Biol. 320*, 1037−1051.
(31) Ortiz-Lombardía, M., González, A., Eritja, R., Aymamí, J., Azorín, F., and Coll, M. (1999) Crystal structure of a DNA Holliday junction. *Nat. Struct. Biol. 6*, 913−917.
(32) Eichman, B. F., Vargason, J. M., Mooers, B. H. M., and Ho, P. S. (2000) The Holliday junction in an inverted repeat sequence: Sequence effects on the structure of four-way junctions. *Proc. Natl. Acad. Sci. U.S.A. 97*, 3971−3976.
(33) Hays, F. A., Vargason, J. M., and Ho, P. S. (2003) Effect of sequence on the conformation of DNA Holliday junctions. *Biochemistry 42*, 9586−9597.
(34) Carter, M., and Ho, P. S. (2011) Assaying the energies of biological halogen bonds. *Cryst. Growth Des. 11*, 5087−5095.
(35) Theruvathu, J. A., Kim, C. H., Rogstad, D. K., Neidigh, J. W., and Sowers, L. C. (2009) Base pairing configuration and stability of an oligonucleotide duplex containing a 5-chlorouracil-adenine base pair. *Biochemistry 48*, 7539−7546.
(36) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307−326.
(37) Hays, F. A., Watson, J., and Ho, P. S. (2003) Caution! DNA crossing: Crystal structures of Holliday junctions. *J. Biol. Chem. 278*, 49663−49666.
(38) Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & nmr system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D54*, 905−921.
(39) Hays, F. A., Teegarden, A., Jones, Z. J. R., Harms, M., Raup, D., Watson, J., Cavaliere, E., and Ho, P. S. (2005) How sequence defines structure: A crystallographic map of DNA structure and conformation. *Proc. Natl. Acad. Sci. U.S.A. 102*, 7157−7162.
(40) Hays, F. A., Schirf, V., Ho, P. S., and Demeler, B. (2006) Solution formation of Holliday junctions in inverted-repeat DNA sequences. *Biochemistry 45*, 2467−2471.
(41) Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, J. M., and Baldwin, R. L. (1987) Tests of the helix dipole model for stabilization of α-helices. *Nature 326*, 563−567.
(42) Scholtz, J. M., and Baldwin, R. L. (1992) The mechanism of α-helix formation by peptides. *Annu. Rev. Biophys. Biomol. Struct. 21*, 95−118.
(43) Dunitz, J. D. (1995) Win some, lose some: Enthalpy-entropy compensation in weak intermolecular interactions. *Chem. Biol. 2*, 709−712.
(44) Sharp, K. (2001) Entropy-enthalpy compensation: Fact or artifact? *Protein Sci. 10*, 661−667.
(45) Politzer, P., and Murray, J. S. (2013) Enthalpy and entropy factors in gas phase halogen bonding: Compensation and competition. *CrystEngComm 15*, 3145−3150.
(46) Minton, A. P. (2000) Implications of macromolecular crowding for protein assembly. *Curr. Opin. Struct. Biol. 10*, 34−39.
(47) Ellis, R. J. (2001) Macromolecular crowding: An important but neglected aspect of the intracellular environment. *Curr. Opin. Struct. Biol. 11*, 114−119.
(48) Cheung, M. S., Klimov, D., and Thirumalai, D. (2005) Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc. Natl. Acad. Sci. U.S.A. 102*, 4753−4758.
(49) Walter, S. M., Kniep, F., Rout, L., Schmidtchen, F. P., Herdtweck, E., and Huber, S. M. (2012) Isothermal calorimetric titrations on charge-assisted halogen bonds: Role of entropy, counterions, solvent, and temperature. *J. Am. Chem. Soc. 134*, 8507−8512.
(50) Ibrahim, M. A. A. (2011) Molecular mechanical study of halogen bonding in drug discovery. *J. Comput. Chem. 32*, 2564−2574.
(51) Ibrahim, M. A. (2012) 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 116*, 3659−3669.
(52) Kolar, M., and Hobza, P. (2012) On extension of the current biomolecular empirical force field for the description of halogen bonds. *J. Chem. Theory Comput. 8*, 1325−1333.
(53) Eisenberg, D., and McLachlan, A. D. (1986) Solvation energy in protein folding and binding. *Nature 319*, 199−203.
(54) Kagawa, T. F., Howell, M. L., Tseng, K., and Ho, P. S. (1993) Effects of base substituents on the hydration of B- and Z-DNA: Correlations to the B- to Z-DNA transition. *Nucleic Acids Res. 21*, 5978−5986.
(55) Kagawa, T. F., Stoddard, D., Zhou, G. W., and Ho, P. S. (1989) Quantitative analysis of DNA secondary structure from solvent-accessible surfaces: The B- to Z-DNA transition as a model. *Biochemistry 28*, 6642−6651.