

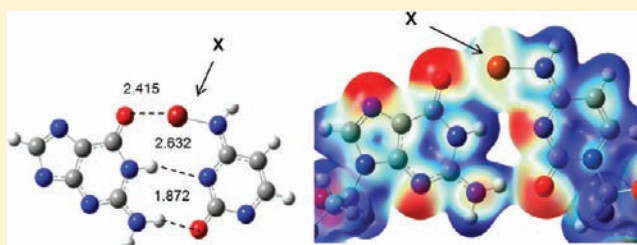
## Halogen Bonding in DNA Base Pairs

Anna J. Parker, John Stewart, Kelling J. Donald,\* and Carol A. Parish\*

Department of Chemistry, Gottwald Center for the Sciences, University of Richmond, Richmond, Virginia 23173, United States

**S** Supporting Information

**ABSTRACT:** Halogen bonding ( $R-X\cdots Y$ ) is a qualitative analogue of hydrogen bonding that may prove useful in the rational design of artificial proteins and nucleotides. We explore halogen-bonded DNA base pairs containing modified guanine, cytosine, adenine and thymine nucleosides. The structures and stabilities of the halogenated systems are compared to the normal hydrogen bonded base pairs. In most cases, energetically stable, coplanar structures are identified. In the most favorable cases, halogenated base pair stabilities are within  $2 \text{ kcal mol}^{-1}$  of the hydrogen bonded analogues. Among the halogens  $X = \text{Cl, Br, and I}$ , bromine is best suited for inclusion in these biological systems because it possesses the best combination of polarizability and steric suitability. We find that the most stable structures result from a single substitution of a hydrogen bond for a halogen bond in dA:dT and dG:dC base pairs, which allows 1 or 2 hydrogen bonds, respectively, to complement the halogen bond.



### ■ INTRODUCTION

The importance of hydrogen bonding is well-known. Intermolecular hydrogen bonds (H-bonds) account for the relatively high boiling point of liquid water, are decisive for the construction and stability of biological structures such as DNA and RNA, and are important drivers for protein folding, ligand reception, and many other biologically significant processes.<sup>1</sup> The century-long interest in these interactions (regardless of which story of the discovery of the hydrogen bond one accepts)<sup>2,3</sup> is a testament to their importance.

Halogen bonding is a much more recently defined concept.<sup>4</sup> The phenomenon has been known for a long time, but was described as a type of charge transfer interaction.<sup>5,6</sup> Halogen bonding has received far less attention in the literature compared to hydrogen bonding, but there is mounting evidence that halogen bonds (X-bonds) are relevant in biology, materials science, and crystal engineering.<sup>4,6–15</sup> Halogen bonds share many properties with hydrogen bonds, and if the promise of halogen bonding is realized, this may lead to new “letters of the genetic alphabet” for engineering novel artificial proteins and nucleotides.<sup>16</sup>

A H-bond is a bonding interaction between (i) a partially positive H atom bonded to an electronegative center (e.g., N or O) and (ii) a nearby electron-rich site (Lewis base). A halogen bond describes a set of analogous interactions that occur between a halide (X) in one molecule ( $R-X$ ) and an electron-rich site in a nearby Lewis base ( $Y-R'$ ).<sup>4,15,17</sup> Although the halogen involved in an  $R-X$  bond may have a net negative charge, the stability of the X-bond ( $R-X\cdots Y-R'$ ) is explained by the presence of a sigma hole ( $\sigma$ -hole)—a region of positive electrostatic potential on X along the bond axis, outside of the  $R-X$  covalent bonding region (Figure 1a and b). The emergence of this  $\sigma$ -hole has been explained by a significant polarization of the electron density at X away from the crown of the atom into the  $R-X$ -bonding region.<sup>18–20</sup>

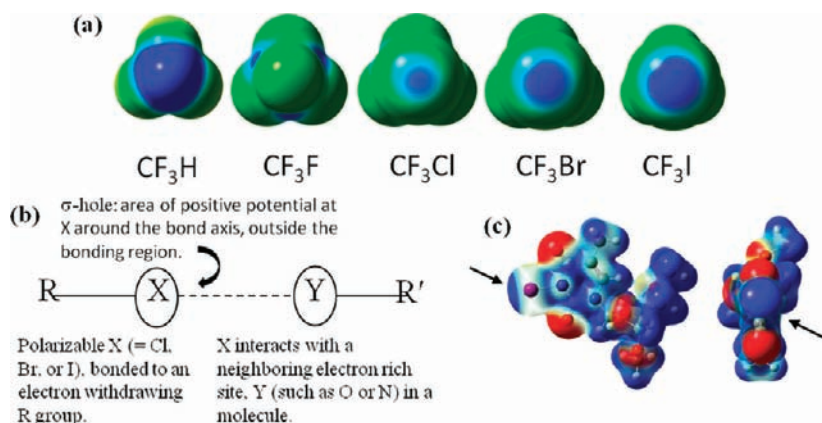
The strength of an X-bond depends on the polarizability of the halogen, the electron-withdrawing power of the R group to which X is bonded, and the nucleophilicity of the Lewis base. H-bonds to O or N centers in molecules are typically stronger than the corresponding X-bonds.<sup>21–23</sup> However, some of the stronger X-bonds are comparable to or even stronger than some of the weaker H-bonds (such as cases in which  $Y = \text{S or Se}$ ).<sup>8,19,24,25</sup> X-bonds are typically characterized by unusually short  $X\cdots Y$  internuclear separations that are less than or equal to the sums of the van der Waals radii of X and Y, and a nearly linear  $R-X\cdots Y$  bond angle,  $\Theta$  (Figure 1b).<sup>17</sup> This linear arrangement is preferred because the  $\sigma$ -hole at X is focused around the extension of the  $R-X$  bond axis (Figure 1). Indeed, the positive electrostatic potential in the  $\sigma$ -hole (and hence the stability of the halogen bond) diminishes quite rapidly as  $\Theta$  decreases.<sup>17,26</sup>

Examples of halogen bonding in biological systems are known. Auffinger et al. found several interesting cases of X-bonding in a survey of halogenated proteins and nucleic acids in the Protein Data Bank, including a strong interaction involving 5-bromouracil that stabilized a Holliday junction.<sup>26</sup> X-bonds were found to direct ligand–protein binding, including the interactions of naturally iodinated thyroid hormones with their cognate proteins, supporting earlier evidence that iodine plays a key role in the specificity of such binding.<sup>27,28</sup>

Seela and co-workers have synthesized 7-halogenated analogues of 2'-deoxyisoguanosine and showed via thermal melting experiments that an interacting halogen stabilized both antiparallel and parallel-stranded DNA duplexes. The stabilizing ability of the halogen, however, was attributed to its ability to expel water molecules from the major groove and increase the

Received: November 8, 2011

Published: February 25, 2012

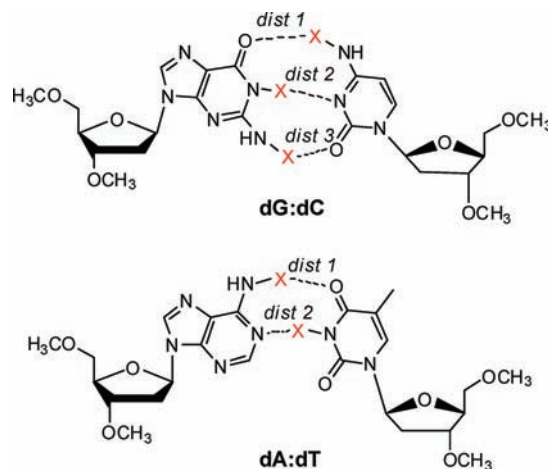


**Figure 1.** (a) Electrostatic potential in atomic units on the 0.001 electrons/bohr<sup>3</sup> isodensity surface of trifluoromethyl halogen species;  $\sigma$ -holes appear as regions of positive electrostatic potential (blue) and increase in size  $\text{F} < \text{Cl} < \text{Br} < \text{I}$ . (b) Schematic illustration of halogen bonding. (c) Electrostatic potential in atomic units on the 0.001 electrons/bohr<sup>3</sup> isodensity surface for I-substituted thymine in side and end-on views. The arrows point to the  $\sigma$ -hole at I.

proton donor capabilities of the amino groups rather than a direct X-bonding interaction.<sup>29–32</sup> Of particular relevance to this work, Sekine and co-workers synthesized and performed *ab initio* calculations on a series of artificial aromatic base pairs. Their design involved iodinated phenyl groups with varying numbers of fluorine substituents, whose main purpose was to further polarize the iodine atom and strengthen the X-bonding interaction.<sup>33</sup> While melting studies on artificial oligomers incorporating these iodoaromatic nucleobases suggested that halogen bonding did not contribute to base pair stability, the authors also noted that their theoretical results suggested that the syn and anti conformations of the artificial nucleobases were similar in energy. Therefore, in experimental testing, rotation about the glycosidic bond may move the halogen bond donor and acceptor moieties away from each other preventing halogen bonding. Our design is based upon halogenating canonical nucleobases and so we anticipate very little perturbation of the rotational profile about the glycosidic bond found in natural base pairs.

N-halogenated nucleosides are known, at least for cytosine and adenine.<sup>34–40</sup> For instance, Ikehara et al. reported in 1977 that the reaction of a limited amount of *tert*-butyl hypochlorite with adenosine at low temperatures produced monochloro-N6-chloroadenosine.<sup>40</sup> Hawkins and Davies have reported on the formation of stable exocyclic N4-chloroamines for poly-C strands as well as for free cytosine. They also observed shorter lived heterocyclic (ring) NH chloroamines with poly-U, poly-T and free bases.<sup>34,35</sup> More recently, Sekine and co-workers have reported the facile, selective N4-mono chlorination of a deoxycytosine (dC) derivative using 1,3-dichloro-5,5-dimethylhydantoin.<sup>37</sup> While a complete understanding of the stability of such chloroamines requires further study, there is at least the possibility that these compounds are synthetically accessible. This would allow the experimental verification of at least two of the systems proposed in our study; namely, dA:dT and dG:dC with a single halogen at position 1 (*vide infra*).

We present here the first fundamental investigation of the strength of halogen-bonded Watson–Crick DNA base pairs dA:dT and dG:dC. (Figure 2.) Interested in the biological utility of X-bonds and the functional distinctions between H- and X-bonds, we explore in detail this bonding in the critical biological role of stabilizing base pair interactions. Fluorine hardly ever forms halogen bonds. Some cases have been cited,<sup>41–43</sup> but they require extremely strong electron-withdrawing R groups, and

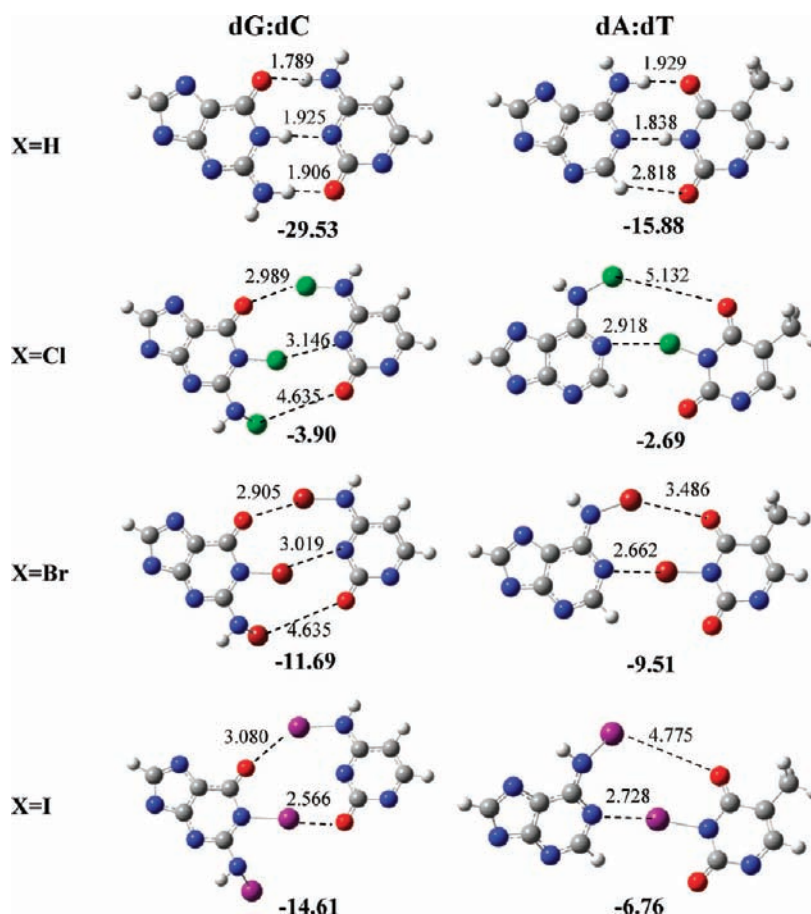


**Figure 2.** Base pair and nucleoside structures (X = H, Cl, Br, I) for which density functional theory calculations were performed. For each base pair, the corresponding individual nucleoside structures were also determined at the same level of theory. All base pair and nucleoside structures included 3'- and 4'-methoxy and methoxymethylene-substituted sugars in order to model the phosphodiester backbone of DNA. For ease of viewing, the substituted sugars have been removed from the ball and stick molecules in the text (*vide infra*), but they were included in all calculations.

the  $\sigma$ -hole on F is necessarily small such that even where X-bonds are suspected they are far weaker than those formed by the heavier halides.<sup>18,23</sup> We consider therefore a series of base pairs in which one or all of the H atoms that are involved in H-bonding are replaced by the halogen atoms X = Cl, Br, or I.

## COMPUTATIONAL METHODS

Geometry optimizations of nucleosides and base pairs were performed with the Gaussian 03 program suite<sup>44</sup> using Becke's hybrid B3LYP functional<sup>45–47</sup> in the gas phase and with implicit solvent (water). The 6-31G\* basis set<sup>48</sup> was used for all atoms except iodine for which the small (28-electron) core Dirac–Fock (MDF) effective-core pseudopotentials and the corresponding basis sets were employed.<sup>49</sup> The resulting similarity between solvent and gas-phase geometries for 16 test structures including H-bonded and X = Cl, X-bonded base pairs caused us to pursue only gas-phase studies for structures containing Br and I (solvent/gas phase rmsd can be found in the Supporting Information.) All structures were also subjected to single-point analysis at M05-2X/6-31G\*<sup>50,51</sup> and MP2/6-31G\* so as to rule out any



**Figure 3.** B3LYP/6-31G\* gas-phase structures for H-bonded and X-bonded base pairs with the MDF effective-core pseudopotentials used for iodine. All images were generated using GaussView 3.0. Substituted sugars were oriented similarly in all optimized complexes and have therefore been removed from the images in order to highlight the nucleobase geometries. Interaction energies ( $E_{int}$ ) are shown below each structure in kcal/mol units.

methodological dependence of the results. All structures and energies shown below correspond to B3LYP/6-31G\*. Results for other treatments can be found in the Supporting Information. In addition to replacing all H-bonded hydrogen atoms with halogens, we also explored the effect of systematic replacement of only one H-bond with an X-bond at each possible bonding site in the base pairs. For all structures, methoxymethylene and methoxy substituents were added to the 4' and 3' carbon atoms, respectively, of each sugar to mimic the steric bulk of the phosphodiester backbone of DNA (Figure 2).

Frequency analyses were performed to confirm all final structures as minima on the potential energy surface. The interaction energy ( $E_{int}$ ) was used as the basis for comparing the stability of the base pairs, and was determined using the equation

$$E_{int} = E_{basepair} - E_{nucleoside1} - E_{nucleoside2} \quad (1)$$

where  $E$  is the electronic energy. This quantity,  $E_{int}$ , represents the energy released during the formation of a H- or X-bond between any given pair of nucleosides. Zero-point energy corrections were not included in the  $E_{int}$  data based on the assumption that these correction factors would be very similar since the species are so closely related structurally.

The electrostatic potential representations have been generated using the Gaussview graphics software using the charge density data obtained for the optimized structures. All the electrostatic potentials were generated in atomic (Hartree) units on the 0.001 electrons/bohr<sup>3</sup> isodensity surface for the systems studied. For comparison, we have used the same range for the color scale for all of the electrostatic potential maps in this work.

## RESULTS AND DISCUSSION

**All H-Bonded Hydrogens Replaced with Halogens.** All of the base pair structures studied, regardless of whether the units are linked by H- or X-bonds, form stable, well-defined dimeric structures (Figure 3) and display favorable energies of interaction (Table 1). All of the structures are coplanar except dA:dT with I where the nucleosides bend  $\sim 14^\circ$  out of planarity.

As expected, H-bonded dG:dC is almost twice as stable as H-bonded dA:dT. dG:dC has three relatively short intermolecular hydrogen-bond interactions (1.8 Å, 1.9 Å and 1.9 Å;  $E_{int} = -29.53$  kcal/mol), while dA:dT contains only two hydrogen-bonding interactions (1.9 Å and 1.8 Å;  $E_{int} = -15.88$  kcal/mol). These interaction energies compare quite well with other computed H-bond energies obtained at the MP2 and the CCSD(T) levels of theory.<sup>52</sup> Remarkably, we find that the halogen-bonded base pairs are also quite stable relative to the isolated bases. The halogen-bonded structures are not as stable as conventional H-bonded structures but they exhibit favorable interactions that are substantial in several cases, especially for X = Br and I (Table 1). For instance, for the halogen-bonded dG:dC structures, the base pair interaction energies range from  $\sim -3.9$  to  $-14.6$  kcal/mol compared to  $-30$  kcal/mol for the normal hydrogen-bonded system. In the case of the dA:dT base pairs, the strongest X-bond interaction energies (at  $-9.5$  and  $-6.8$  kcal/mol for X = Br, and X = I, respectively) are even closer in value to the computed H-bond energy of  $-15.9$  kcal/mol. An extraordinary feature of



**Table 1. Interaction Energies ( $E_{\text{int}}$ , kcal/mol), Internucleoside H-/X-Bond Distances (Å), Normalized X-Bond Distances, R–X Covalent Bond Lengths (Å), and R–X···Y-Bond Angles,  $\Theta$ , (deg) for Base Pair Geometries<sup>a,b</sup>**

base pair	X	$E_{\text{int}}$	bonds <sup>c</sup>	X···Y/Å <sup>c</sup>	X···Y <sub>NORM</sub> <sup>a</sup>	R–X/Å <sup>c</sup>	$\Theta$ /deg <sup>c</sup>
dG:dC	H	–29.53	#1	1.789		1.035	178.2
			#2	1.925		1.033	175.7
			#3	1.906		1.024	178.8
	Cl	–3.90	#1	2.989	0.914	1.724	161.0
			#2	3.146	0.953	1.730	159.7
			#3	4.646	1.417	1.794	90.4
	Br <sup>d</sup>	–11.69	#1	2.905	0.862	1.883	159.6
			#2	3.019	0.888	1.910	157.2
			#3	4.635	1.375	1.961	87.7
I <sup>e</sup>	–14.61	#1	3.080	0.880	2.082	144.5	
		#2	2.566	0.733	2.182	178.1	
dA:dT	H	–15.88	#1	1.929		1.023	174.9
			#2	1.838		1.045	178.8
	Cl	–2.69	#1	5.132	1.569	1.736	131.4
			#2	2.918	0.884	1.727	157.9
	Br	–9.51	#1	3.486	1.034	1.887	144.2
			#2	2.662	0.783	1.913	177.3
	I	–6.76	#1	4.775	1.364	2.109	108.9
			#2	2.728	0.773	2.143	173.6

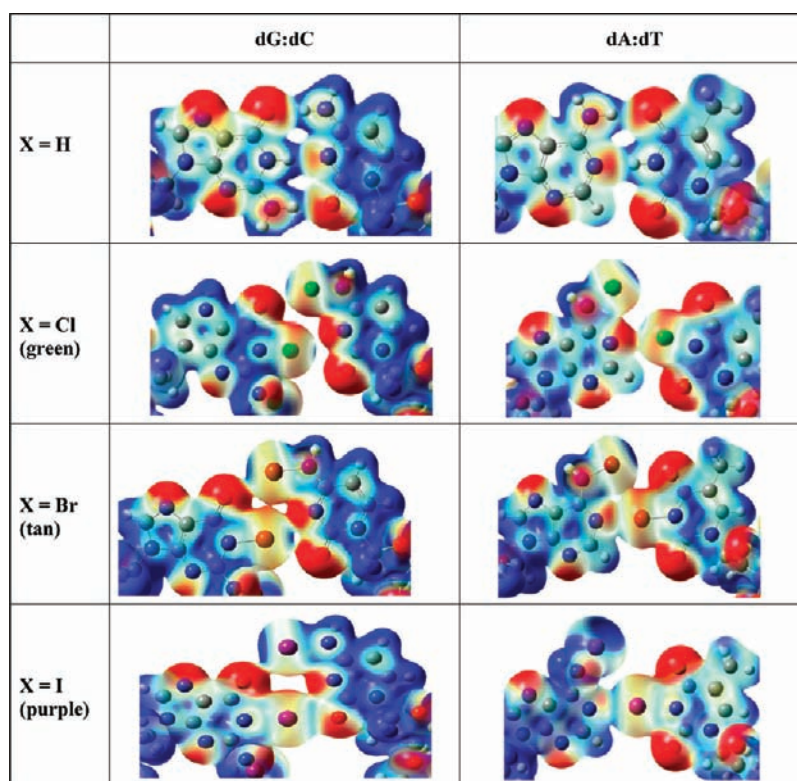
<sup>a</sup>Normalized X bond distances =  $d_{\text{X}\cdots\text{Y}}/R_{\text{vdW}(\text{X}\cdots\text{Y})}$ . vdW distances can be found in the Supporting Information. <sup>b</sup>All structures were optimized with the 6-31G(d) basis set except for the iodine atoms which were treated with the Dirac–Fock (MDF) effective-core pseudopotentials. <sup>c</sup>A key for identifying bond distances is available in Figure 2. <sup>d</sup>For bromine-containing dG:dC, the second X-bond donor of guanine seems to interact equally with both the nitrogen and oxygen X-bond acceptors sites in cytosine. The distance between bromine and nitrogen is only slightly smaller than that between the bromine and oxygen (by  $\sim 0.1$  Å), although the angle of the halogen bond with nitrogen is  $\sim 2^\circ$  smaller than that of the bond with oxygen. <sup>e</sup>For iodine-containing dG:dC, the third X-bond acceptor of cytosine is interacting with the second X-bond donor of guanine, rather than the donor corresponding to it. The second X-bond acceptor of cytosine and the third X-bond donor of guanine do not interact, and do not seem to play a role in base pair formation.

the dA:dT results is that the base pair is noticeably more stable when X = Br than when X = I. We discuss this observation shortly.

In the halogen-bonded structures, the effective number of X···Y interactions is typically smaller than expected for these systems. Several of the dG:dC base pairs contain only two (rather than three) X···Y bonds that are shorter than the sum of the van der Waals (vdW) radii (Supporting Information) of the X and Y (Y = N or O) atoms. Similarly, only one of the possible two halogen bonds in the modified dA:dT systems have short X···Y internuclear separations. This suggests that, instead of overcoming the steric difficulties associated with maintaining the two or three linear halogen bonds between the nucleosides, one interaction is sacrificed in order to optimize the others. This can be seen in the halogen-bonded base pair electrostatic potential energy surfaces (Figure 4) where electrostatic interaction correlates with base pair stability and shows clear overlap between X and Y atoms on corresponding nucleosides.

Base-pair geometries result from a complex optimization of intra- and intermolecular structural and electronic effects.<sup>53</sup> We find that halogen bonds are strongest when the geometric constraints of the interacting nucleosides simultaneously allow the shortest interaction distances as well as halogen-bonding angles near  $180^\circ$ . A linear halogen bond is needed to properly orient the  $\sigma$ -hole on the halogen toward the N or O lone pair on the Lewis base.<sup>17</sup> The importance of the alignment of the  $\sigma$ -hole is underscored by considering the halogen-bonded dA:dT results—especially for X = Br and I—and the electrostatic potentials in Figure 4. There, we see only one very strong halogen bond that has optimized the R–X···Y bond angle very near  $180^\circ$ , allowing significant electrostatic interaction between the halogen and the N on the nucleobase.

The interaction energies of halogen-bonded dG:dC base pairs (Table 1) vary as I > Br > Cl. This ordering corresponds to the relative magnitudes of the polarizabilities of the halogen atoms and follows the expected variation in halogen interaction. However, in the case of the dA:dT base pairs, the Br case is almost 3 kcal/mol more stable than the I-containing system. This unusual ordering of the halogen-bond strength is explained by several factors. The large C–N–I bond angle on the dA nucleoside in the base pair minimizes the repulsion between that I atom and the adjacent I atom on the dT nucleoside (see the structure at the bottom right in Figure 3). The outward displacement of this I atom decreases the N–I···O bond angle so that the  $\sigma$ -hole at I and the oxygen lone pair are far from being aligned, which is critical for any substantial halogen bonding. Compared to the high cost in energy that would be required to maintain two moderate X-bonds while compensating for the I–I repulsion, it turns out to be much more favorable to optimize one X-bond (with a distance of 2.728 Å in Figure 3) at the expense of the other. This very short and almost linear N···I–N halogen bond in the dA:dT with I system (compared to 2.918 Å in the case where X = Cl) confirms how strong the N···I halogen bond can be. On the other hand, the long I···O separation (4.775 Å) shows quite dramatically how unfavorable it is to pack two antiparallel N–I bonds in the space available between the nucleosides. Indeed, this characteristic is also observed in the X = I dG:dC system (bottom left in Figure 3). In that base pair, the N on the dC ring is not utilized; the antiparallel N–I bonds move away from each other, and one of the N–I bonds on dG rotates out of the dG:dC interaction region (sacrificing that one interaction in order to optimize the others). So, instead of three interactions in this dG:dC system, only two X-bonds are formed. For the



**Figure 4.** Electrostatic potential (ESP) in atomic units on the 0.001 electrons/bohr<sup>3</sup> isodensity surfaces for the normal and halogen substituted dA:dT and dG:dC base pairs. ESP values range from  $-4.55 \times 10^{-2}$  au (red) to  $+4.55 \times 10^{-2}$  au (blue).

dG:dC case, therefore, the stability of the Br system rivals the I alternative but does not exceed it. In the dA:dT case, where only two interactions are involved, the Br system with its two moderate X-bonds turns out to be more stable than the I case with one strong bond.

This substantial destabilization due to X...X hard sphere repulsion is manifested clearly in the significant amount of puckering or pyramidalization about the NHX groups of dG in both monomers and base pair dimers. Previously, Alabugin has shown that rehybridization at nitrogen caused by proximate and remote structural modifications, provides a sensitive probe of subtle electronic effects.<sup>54</sup> In our study, the NH<sub>2</sub> pyramidalization in the nonhalogenated dG monomer and dG:dC base pair is 132 and 160°, respectively. The increase in this angle in the base pair is likely a result of the withdrawal of electron density from the N lone pair caused by internucleoside hydrogen bonding. However, in the all-halogenated systems, the NHX puckering angles on dG are the smallest and range from 112 to 119° in both the monomers and the dimers. dG is the only nucleoside where two halogen substitutions are possible. When all H-bonds are replaced with X-bonds, the halogen on the dG NHX moiety moves out of the internucleoside plane in order to minimize X...X repulsions. A test of this interpretation confirms that maintaining a planar NHBr group on dG in dG:dC with Br (Figure 3) would position the two bromine atoms only ~2.38 Å apart. The structural data for the pyramidalization at the N sites can be found in the Supporting Information.

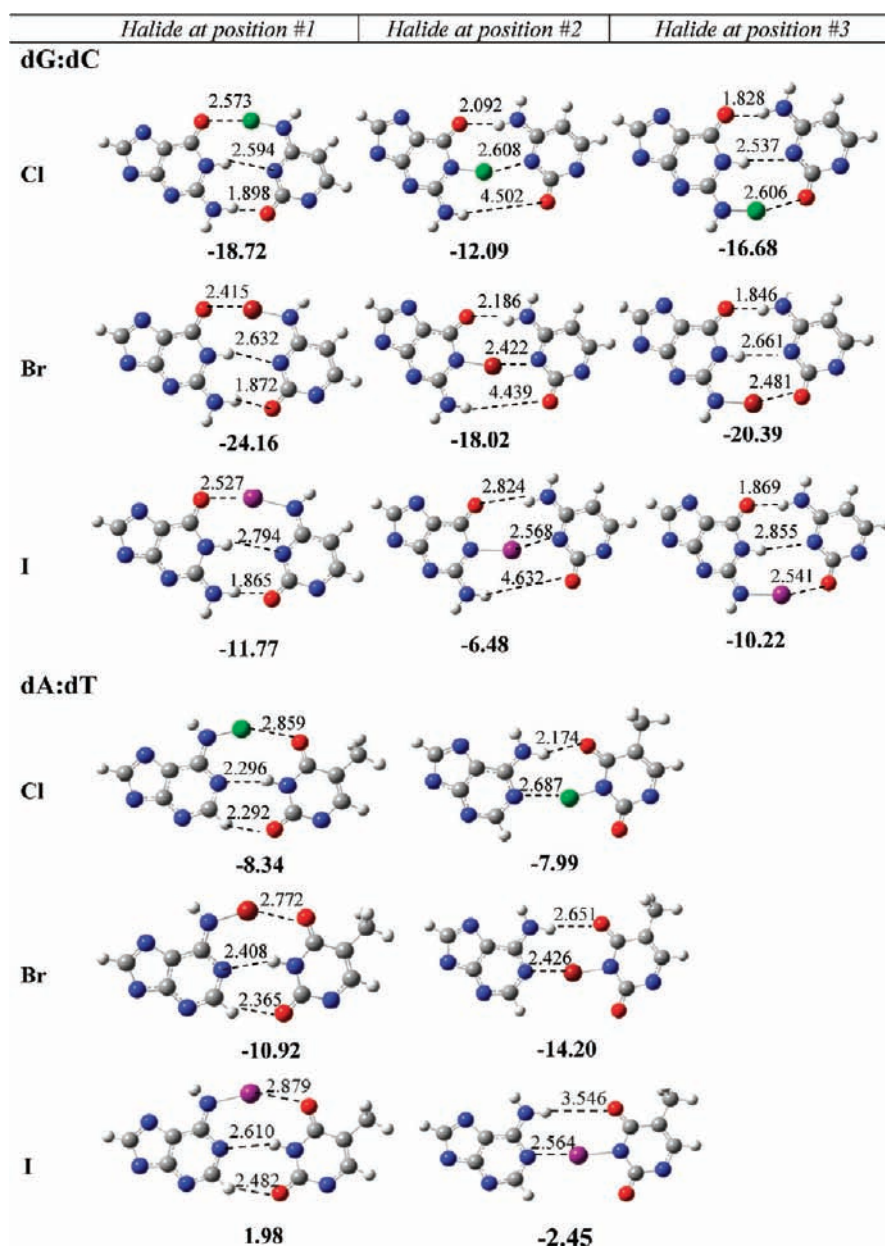
The Cl-containing complexes are the least stable base pairs in both the dA:dT and dG:dC cases. The Cl...Y interactions are relatively weak (Table 1), and the bond angles are not optimal. Going from X = Cl to X = Br, however, the moderate increase in the atomic size and a substantial increase in the polarizability of X and the strength and size of the  $\sigma$ -hole ensures that the

base pairs have a much larger interaction energy. This balance between size (atomic radii) and strength when X = Br, explains the stability of this system compared to the iodide system which has some difficulty fitting all of the iodides into the interaction region (with the appropriate orientation) between the bases.

The size of the halogen atom affects the directionality and, by extension, the strength of the halogen bond. As the halogen atomic radius increases from Cl < Br < I, the two or three inter-base interactions spread away from the almost parallel orientation seen in the nonhalogenated base pairs. Taken together, these results suggest that halogen-bonded base pair stability is a complex function of nucleoside geometry and rigidity, and halogen radii and polarizability. Even though halogen bonds can be formed between bases, they are spatially far less versatile than H-bonds because of the size of the halides and the severe dependence of X-bonds on acid–base alignment.

**Optimizing Halogen Bonding in DNA Base Pairs: Single Halogen Substitution.** The results discussed above suggest that—like hydrogen bonding—halogen bonding between nucleosides in a base pair result from a relatively complex optimization of intra- and internucleoside electronic effects and steric constraints.<sup>10,53</sup> To further pursue this line of thought, and to see if we might find biologically relevant structures with even more favorable halogen bonding, we systematically substituted one hydrogen-bonding site at a time with a halogen atom in the hopes that this would obviate the need to optimize more than one bulky polarizable halogen interaction in any one structure. The incorporation of only one halogen-bonding interaction also led to well-defined base pair geometries that were all coplanar (Figure 5; further geometric details can be found in the Supporting Information).

In all cases, except dA:dT with X = I, all base pair interactions were favorable. In the case of Cl- and Br-substituted base



**Figure 5.** B3LYP/6-31G\* gas-phase structures of dG:dC and dA:dT base pairs where only one H-bond was replaced with an X-bond. All images were generated using GaussView 3.0.

pairs, these single halogen-bonded structures were 7–15 kcal/mol more stable than the previously discussed 2- and 3-halogen-bonded dA:dT and dG:dC structures, respectively. This appears to be due to the opportunity to optimize a single halogen interaction without steric constraints while still maintaining one or two hydrogen-bonding interactions in these singly substituted dA:dT and dG:dC systems. In the best cases, the Br containing base pairs, these single halogenated structures were only 2–5 kcal/mol less stable than the corresponding hydrogen-bonded base pairs.

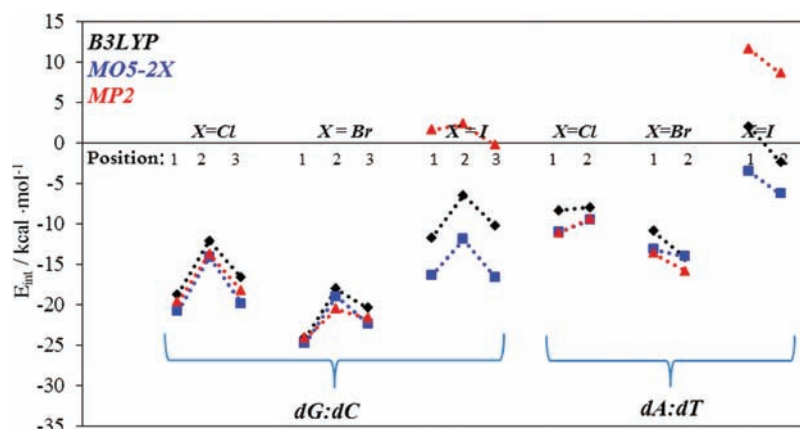
This is particularly interesting since the nucleobase distances are much longer in the halogenated base pairs. For instance, the normal H-bonded dA:dT intermolecular distances are 1.929 (*dist 1*) and 1.838 Å (*dist 2*) with  $E_{\text{int}} = -15.88 \text{ kcal mol}^{-1}$  (Figure 3) whereas dA:dT with X = Br at position #2 has distances 2.651 (*dist 1*) and 2.426 Å (*dist 2*) with  $E_{\text{int}} = -14.20 \text{ kcal mol}^{-1}$  (Figure 5). The most stable halogen-substituted base pairs for both dG:dC and dA:dT contained bromine and

this is explained by the higher polarizability of Br compared to Cl and the smaller size of Br compared to I, which makes the bromine much more able to adapt geometrically than iodine.

The relative energetics of the single substitution provides some insight into the local environment of base pair interactions. For instance, substituting dG:dC at either position #1 or position #3 leads to stronger internucleoside interactions compared to the case where the substituent was at position #2. This is because large halogens at position #2 severely restrict the nucleosides from finding any means to readjust in order to optimize the overall base pair geometry. This is manifested in the presence of at least one short (less than 2 Å) internucleoside distance when the substitution is made at either positions #1 or #3 (see Figures 2 and 5) but not when it is made at position #2.

The general dependence of X-bond (X...Y) interaction energies ( $E_{\text{int}}$ ) on the identity of X is well-known to be  $E_{\text{int}}(\text{Cl}) < E_{\text{int}}(\text{Br}) < E_{\text{int}}(\text{I})$ .<sup>18,23</sup> This variation is reflected in the





**Figure 6.** B3LYP interaction energies for the gas-phase structures of dG:dC and dA:dT base pairs where only one H-bond was replaced with an X-bond. The corresponding M05-2X, and MP2 energies for the same structures are included, as well, for comparison.

results we obtain in this work, but with substantial exceptions. The normal hydrogen-bonded systems are always the most stable forms; however, in several instances in which hard sphere repulsions (steric effects) are important within the internucleoside bonding region, the Br-substituted base pairs form the most stable halogen complexes, sometimes less stable than the H-bonded systems by only 2.0 kcal mol<sup>-1</sup>. Indeed, our results confirm that H-atoms are ideal for uniting DNA nucleosides because they combine a small atomic radius with a low electro-negativity.

To be sure, iodine will always be expected to form the strongest halogen bond (compared to the analogous F, Cl, or Br cases) if the orientations (R–X···Y) are linear, for any given R and Y. However, we find that where spatial considerations are important, as in the bridging region between the base pairs in this study, the general rule that iodides form the strongest halogen bonds does not apply.

**Qualitative and Quantitative Effects of the Size of X in Singly Substituted Base Pairs.** The destabilizing effect of the X substitutions at position 2 in the dG:dC combinations is evident from a careful qualitative examination of the base pair structures in Figure 5. Notice that in those complexes the strengthening of one of the H···O hydrogen-bonding interactions (at position 1, for example) is achieved by rotating the base about the middle N···I bond to shorten that H···O distance. In the process, however, the H···O interaction at position 3 is partially sacrificed. There is no alternative; and it is for that reason that X substitutions at position 2 are the most destabilizing for the dG:dC pair, especially so for X = I (see Figure 6).

The presence of the I atom at position 2 is much less disruptive in the dA:dT case, however, since there are only two important intermolecular interactions (the H–O at position 1 and the N···I at position 2), so the destabilization relative to the classical dA:dT structure is less significant. In fact, as Figures 5 and 6 show, substituting for X at position 1 in dA:dT is generally the more disruptive of the two cases for that base pair.

Our observation that hard sphere repulsive effects cause the bromides to form the most stable base pairs of the series in Figure 5 (with X = Cl, Br, or I at positions 1, 2, or 3 in the dG:dC pair, and at positions 1 and 2 in the dA:dT pairs) is confirmed in Figure 6. In that graph, the M05-2X, and MP2 energies (for the same B3LYP optimized structures) are included, as well. The M05-2X, and the MP2 interaction energies are typically higher than the B3LYP values by no more than 2 kcal/mol for X = Cl, and Br. The three methods agree that the base pair with X = Br at

position 1 in dG:dC is the strongest instance of halogen bonding among the singly substituted pairs. In that case (see Table S7),  $E_{\text{int}}$  (B3LYP) = –24.16 kcal/mol,  $E_{\text{int}}$  (M05-2X) = –24.77 kcal/mol, and  $E_{\text{int}}$  (MP2) = –24.16 kcal/mol, respectively. There is good qualitative agreement as well that the iodide systems will be particularly unstable relative to the corresponding Cl and Br systems. The MP2 method is emphatic in this regard; positive interaction energies have been obtained in all cases for the iodides except for a feeble  $E_{\text{int}}$  (MP2) = –0.31 kcal/mol in a single instance (see Figure 6, and Tables S7 and S8).

## CONCLUSIONS

Density functional theory was used to investigate the strength of X-bonding in the context of biological systems. Halogen-containing nucleosides with X = Cl, Br, and I were examined in Watson–Crick base pairs (dG:dC and dA:dT) and their stabilities compared with those of normal base pairs. In order to optimize the bonding interactions between the base pairs while minimizing the geometric constraints, we also considered structures with only one H-bond exchanged for an X-bond. These structures proved to be more stable in most cases than those with all H-bonds replaced, and in many cases these structures were close in stability to structures with all H-bonds. In the case of the single X-bonded structures, bromine proved to be the optimal X-bonding halogen, being smaller and less sterically bulky than iodine and more polarizable than chlorine. Remarkably, we find, for example, that the dA:dT system with Br at the second position (Figure 5 and Table S6) has almost the same binding energy (–14.20 kcal/mol) in the gas phase as the classic dA:dT system for which  $E_{\text{int}}$  = –15.88 kcal/mol (Table 1).

Contrary to the general variations in the strengths of halogen bonds in spatially unrestricted systems, our results suggest that in the design of halogenated ligands or in their use as drugs in biological contexts where halogen bonding is important, chlorides or bromides may provide a stronger, more stabilizing interaction compared to iodinated systems. These observations may be transferable to nonbiological, materials-based systems where spatial arrangements of halogen bonding is important for ordering centers in halogenated molecules with adsorption sites at fixed intermolecular distances on surfaces.

## ASSOCIATED CONTENT

### Supporting Information

van der Waal radii; NHX pyramidalization; gas- and solvent-phase rmsd values; energies and structures for solvent-phase

base pairs; tables of geometries for dG:dC and dA:dT base pairs containing only one halogen bond; M05-2X/6-31G\* and MP2/6-31G\* interaction energies, and B3LYP/6-31G\* absolute energies and xyz coordinates of all species; complete ref 44. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

kdonald@richmond.edu (K.J.D.); cparish@richmond.edu (C.A.P.)

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

K.J.D. was supported by the Research Corporation (Cottrell College Science Award #7742), the Jeffress Memorial Trust (Award J-903) and an NSF-CAREER award (CHE-1056430). C.A.P. was supported by the NSF RUI program (CHE-0809462), a Department of Energy grant (CHE-0239664), a Henry Dreyfus Teacher Scholar Award, the Floyd D. and Elisabeth S. Gottwald Endowment and Donors of the American Chemical Society Petroleum Research Fund.

## REFERENCES

- (1) Pimentel, G. C.; McClellan, A. L. *The Hydrogen Bond*; W. H. Freeman: San Francisco, 1960; see especially Chapters 1, 8–11.
- (2) Pauling, L. *The Nature of the Chemical Bond*, 3rd ed.; Cornell University Press: Ithaca, NY, 1960. In this book, Pauling credits T. S. Moore and T. F. Winmill with the first mention of the hydrogen bond.
- (3) Huggins, M. L. *J. Chem. Educ.* **1960**, *37*, A754.
- (4) Dumas, J.-M.; Peurichard, H.; Gomel, M. *J. Chem. Res. Synop.* **1978**, *2*, 54.
- (5) For a historical overview of these interactions and examples of complexes stabilized by these interactions see Odd Hassel's Nobel Lecture (1969): [http://www.nobelprize.org/nobel\\_prizes/chemistry/laureates/1969/hassel-lecture.html](http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1969/hassel-lecture.html), republished in ref 6.
- (6) Hassel, O. *Science* **1970**, *170*, 497.
- (7) Legon, A. C. *Angew. Chem., Int. Ed.* **1999**, *38*, 2686.
- (8) Metrangolo, P.; Neukirch, H.; Pilati, T.; Resnati, G. *Acc. Chem. Res.* **2005**, *38*, 386.
- (9) Politzer, P.; Murray, J. S.; Lane, P. *Int. J. Quantum Chem.* **2007**, *107*, 3046.
- (10) Voth, A. R.; Khuu, P.; Oishi, K.; Ho, P. S. *Nature Chem.* **2009**, *1*, 74.
- (11) Fourmigue, M. *Curr. Opin. Solid State Mater. Sci.* **2009**, *13*, 36.
- (12) Espallargas, G. M.; Zordan, F.; Marin, L. A.; Adams, H.; Shankland, K.; van de Streek, J.; Brammer, L. *Chem.—Eur. J.* **2009**, *15*, 7554.
- (13) Lu, Y.; Shi, T.; Wang, Y.; Yang, H.; Luo, X.; Jiang, H.; Zhu, W. *J. Med. Chem.* **2009**, *52*, 2854.
- (14) Bent, H. A. *Chem. Rev.* **1968**, *68*, 587.
- (15) Murray, J. S.; Riley, K. E.; Politzer, P.; Clark, T. *Aust. J. Chem.* **2010**, *63*, 1598.
- (16) Metrangolo, P.; Resnati, G. *Science* **2008**, *321*, 918.
- (17) Shields, Z. P.; Murray, J. S.; Politzer, P. *Int. J. Quantum Chem.* **2010**, *110*, 2823.
- (18) Brinck, T.; Murray, J. S.; Politzer, P. *Int. J. Quantum Chem.* **1992**, *19*, 57.
- (19) Politzer, P.; Lane, P.; Concha, M. C.; Ma, Y.; Murray, J. S. *J. Mol. Model.* **2007**, *13*, 305.
- (20) Politzer, P.; Murray, J. S. An Overview of Sigma-Hole Bonding: An Important and Widely Occurring Noncovalent Interaction. In *Practical Aspects of Computational Chemistry Methods, Concepts and Applications*; Leszczynski, J., Shukla, M. K., Eds.; Springer: Heidelberg, 2009; Chapter 6.

- (21) Valerio, G.; Raos, G.; Meille, S. V.; Metrangolo, P.; Resnati, G. *J. Phys. Chem. A* **2000**, *104*, 1617.
- (22) Li, Q.; Xu, X.; Liu, T.; Jing, B.; Li, W.; Cheng, J.; Gong, B.; Sun, J. *J. Phys. Chem. Chem. Phys.* **2010**, *12*, 6837.
- (23) Donald, K. J.; Wittmaack, B. K.; Crigger, C. *J. Phys. Chem. A* **2010**, *114*, 7213.
- (24) Li, Q.-Z.; Li, R.; Guo, P.; Li, H.; Li, W.-Z.; Cheng, J.-B. *Comput. Theor. Chem.* **2012**, *980*, 56.
- (25) Li, Q.-Z.; Jing, B.; Li, R.; Liu, Z.-B.; Li, W.-Z.; Luan, F.; Cheng, J.-B.; Gong, B.-A.; Sun, J.-Z. *J. Phys. Chem. Chem. Phys.* **2011**, *13*, 2266.
- (26) Auffinger, P.; Hays, F. A.; Westhof, E.; Ho, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16789.
- (27) Bayse, C. A.; Rafferty, E. R. *Inorg. Chem.* **2010**, *49*, 5365.
- (28) Cody, V.; Murray-Rust, P. *J. Mol. Struct.* **1984**, *112*, 189.
- (29) He, J.; Becher, G.; Budow, S.; Seela, F. *Nucleosides, Nucleotides Nucleic Acids* **2003**, *22*, 573.
- (30) Seela, F.; Shaikh, K.; Wiglenda, T. *Nucleosides, Nucleotides Nucleic Acids* **2003**, *22*, 1239.
- (31) Seela, F.; Kroschel, R. *Nucleic Acids Res.* **2003**, *31*, 7150.
- (32) Seela, F.; Peng, X.; Li, H. *J. Am. Chem. Soc.* **2005**, *127*, 7739.
- (33) Tawarada, R.; Seio, K.; Sekine, M. *J. Org. Chem.* **2008**, *73*, 383.
- (34) Hawkins, C. L.; Davies, M. J. *Chem. Res. Toxicol.* **2002**, *15*, 83.
- (35) Hawkins, C. L.; Davies, M. J. *Chem. Res. Toxicol.* **2001**, *14*, 1071.
- (36) Gould, J. P.; Richards, J. T.; Miles, M. G. *Water Res.* **1984**, *18*, 991.
- (37) Ohkubo, A.; Noma, Y.; Aoki, K.; Tsunoda, H.; Seio, K.; Sekine, M. *J. Org. Chem.* **2009**, *74*, 2817.
- (38) Kawai, Y.; Matsui, Y.; Kondo, H.; Morinaga, H.; Uchida, K.; Miyoshi, N.; Nakamura, Y.; Osawa, T. *Chem. Res. Toxicol.* **2008**, *21*, 1407.
- (39) Kawai, Y.; Morinaga, H.; Kondo, H.; Miyoshi, N.; Nakamura, Y.; Uchida, K.; Osawa, T. *J. Biol. Chem.* **2004**, *279*, 51241.
- (40) Ikehara, M.; Ogiso, Y.; Maruyama, T. *Chem. Pharm. Bull.* **1977**, *25*, 575.
- (41) Politzer, P.; Murray, J. S.; Concha, M. C. *J. Mol. Model.* **2007**, *13*, 643.
- (42) Wang, Y.-H.; Lu, Y.-X.; Zou, J.-W.; Yu, Q.-S. *Int. J. Quantum Chem.* **2008**, *108*, 1083.
- (43) Metrangolo, P.; Murray, J. S.; Pilati, T.; Politzer, P.; Resnati, G.; Terraneo, G. *Cryst. Growth Des.* **2011**, *11*, 4238.
- (44) Frisch, M. J.; et al. *Gaussian 03*, Revision B.04; Gaussian, Inc.: Wallingford, CT, 2004.
- (45) Becke, A. D. *J. Phys. Chem.* **1993**, *98*, 5648.
- (46) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (47) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
- (48) Rassolov, V. A.; Ratner, M. A.; Pople, J. A.; Redfern, P. C.; Curtiss, L. A. *J. Comput. Chem.* **2001**, *22*, 976.
- (49) Petersen, K. A.; Shelper, B. C.; Figgen, D.; Stoll, H. *J. Phys. Chem. A* **2006**, *110*, 13877.
- (50) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364.
- (51) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157.
- (52) van der Wijst, T.; Guerra, C. F.; Swart, M.; Bickelhaupt, F. M. *Chem. Phys. Lett.* **2006**, *426*, 415.
- (53) Hamm, M. L.; Parker, A. J.; Carman, J. L.; Steele, T. W. E.; Parish, C. A. *J. Org. Chem.* **2010**, *75*, 5661.
- (54) Alabugin, I. B.; Manoharan, M.; Buck, M.; Clark, R. J. *J. Mol. Struct. (THEOCHEM)* **2007**, *813*, 21.