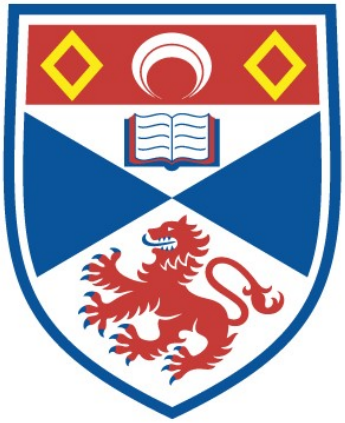
THE ROLE OF HALOGEN BONDING IN BIOMOLECULES

Simon William Leslie Hogan

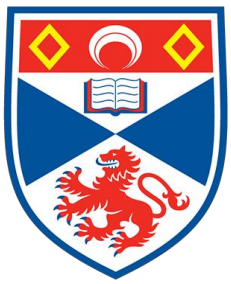
A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



2018

The Role of Halogen Bonding in Biomolecules

Simon William Leslie Hogan



University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of

Doctor of Philosophy (PhD)

at the University of St Andrews

March 2018

Abstract

This study concerns halogen bonding between small molecules. Except where otherwise stated herein this investigation was performed exclusively using the M06-2X density functional, in conjunction with the 6-31+G* basis set except for iodine and astatine which were treated using the aug-cc-pVDZ-PP basis set with relativistic pseudopotentials. All calculations were performed in the gas phase. The counterpoise procedure was employed for all full geometry optimisations. Statistical analysis of the Cambridge Structural Database, wherein the frequency of structures as a function of halogen bond angle and distance constituted the sole part of this study not to be based on density functional theory. Except in chapter 5, all halogens from fluorine to astatine are investigated.

In **chapter 3**, halogen bonding between halobenzene and a single water molecule is discussed. Competition between R – X•••OH_2 halogen bonding and R – X•••H-O-H hydrogen bonding interactions is described. This system is analogous to the more elaborate microsolvated 1-methyl-5-halouracil system described in **chapter 4**. In this latter system one 1-methyl-5-halouracil molecule interacts with either one or two water molecules. A central feature of the investigation into this system is competition between R – X•••OH_2 and R=O•••H-O-H hydrogen bonding. In **chapter 5**, halogen bonding is discussed in the context of the thyroid system. In particular halogen bonding between a thyroxine iodine atom and the protein backbone as well as crystal water molecules is the subject of this chapter. The effect of substitution of the iodine atom with an astatine atom is presented. **Chapter 6** is concerned with halogen bonding in halogenated DNA base pairs. Interaction energies are compared with those of the canonical base pairs, and the effect of halogen bonding on geometry is also discussed.

For each system, halogen bonding was found to become stronger and more tolerant of non-linear bond angles going down the halogen group.

Declarations

Candidate's declaration

I, Simon William Leslie Hogan, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 50,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews in September 2013.

I confirm that no funding was received for this work.

Date Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Simon William Leslie Hogan, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

Embargo on all of print copy for a period of 2 years on the following ground(s):

- Publication would preclude future publication

Supporting statement for printed embargo request

There is an intention to publish the work in the form of journal articles. The majority of the thesis has not yet been published.

Electronic copy

Embargo on all of electronic copy for a period of 2 years on the following ground(s):

- Publication would preclude future publication

Supporting statement for electronic embargo request

There is an intention to publish the work in the form of journal articles. The majority of the thesis has not yet been published.

Title and Abstract

- I agree to the title and abstract being published.

Date Signature of candidate

Date Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Simon William Leslie Hogan, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date Signature of candidate

Acknowledgements

General acknowledgements

I wish to thank my supervisor, Dr. Tanja van Mourik for providing me with support and guidance throughout my Ph.D., which has been indispensable in enabling me to produce this work.

I would like to express my appreciation for Rachael Skyner's assistance to me in relation to the statistical analysis contained in chapter 4. Furthermore, I am grateful to everyone with whom I have interacted in room 150.

Most of the calculations in this work were performed on the Wardlaw computing cluster via the EaStCHEM Research Computing Facility, while the remainder were performed on the Obelix cluster. Both of these clusters are maintained by Dr. Herbert Füchtl.

Contents

List of Abbreviations

CSD	Cambridge Structural Database
DFT	Density Functional Theory
Hbond	Hydrogen bond
PDB	Protein Databank
VdW	Van der Waals
Xbond	Halogen bond
XPh	Halobenzene
XPh-w	Microsolvated halobenzene
XU	1-methyl-5-halouracil
XU-w	Microsolvated halouracil (one water molecule)
XU-2w	Microsolvated halouracil (two water molecules)
dA:dT	Adenine – thymine complex
dG:dC	Guanine – cytosine complex

1 Introduction

Colin reported on a liquid comprising ammonia and iodine in 1814 [1]. The composition of this substance, NH_3I_2, was proposed by Guthrie in 1863 [2]; it should however be noted that Guthrie was not proposing a type of interaction between ammonia and iodine, although it was noted that spontaneous decomposition of the dry compound into ammonia and iodine occurred in air and that the formation of NH_3I_2 was not substantially exothermic. The interaction between the ammonia molecule and the iodine is cited [3] as the first halogen bond containing complex to be synthesised. Understanding of the nature of this intermolecular interaction has been advanced by the work of Mulliken [4] and Hassel [5]. The term "halogen bond" was first employed by Dumas *et al.* in 1978 [6].

In a publication authored by Desiraju *et al.*, and sponsored by the Physical and Biophysical Chemistry Division of IUPAC, a definition of the halogen bond was proposed, referring to the essential feature of a stabilising electrophile-nucleophile relationship where the electrophilic element is "a region associated with a halogen atom in a molecular entity" [7] . The definition provided by Desiraju *et al.* also includes a set of "typical" features as a guide to whether a given interaction would be correctly characterised as a halogen bond. These features, not forming an integral part of the formal definition, but to be used as a guide for the identification of halogen bonds, include: (i) internuclear separation less than the sum of VdW radii, (ii) covalent bond via which the halogen atom is bound to the remainder of its molecule is lengthens upon halogen bond formation, (iii) near-linear bond angle, (iv) negative correlation between interaction strength and the electronegativity of the halogen atom, (v) positive corrlelation between interaction strength and the electron-withdrawing effect of the group whereto the halogen atom is covalently bound, (vi) dominant role of electrostatic interaction with some contribution from dispersion, charge transfer and polarisation, (vii) "analysis of the electron density topology usually shows a bond path" [7] (for the definition thereof see reference [7] and cited refrences therein), (viii) infrared and Raman spectra that depart from those for the halogenated molecule without the presence of the halogen bond, (ix) UV-vis absorption band shift to higher frequencies compared with the halogenated molecule without the presence of the halogen bond, (x) change in NMR chemical shift compared with the halogenated molecule without the presence of the halogen bond, and (xi) "binding energies of the peaks associated with X with the X-ray photoelectron spectrum[7](XPS) of the complex shift to lower energies relative to unbonded X" [7]. Features (i)-(vi), and especially (i), (iii), (iv), (v) and (vi)), have been the subject of discussion in this thesis. The positively charged region on a halogen atom is usually referred to as a σ-hole, as can be seen from reading the literature in the field[8-9], for example Kolár *et al.* entitled their publication introducing a way of modelling this feature for molecular docking as "Plugging the explicit σ-holes in molecular docking" [10].

The σ-hole is a feature of the anisotropic distribution of the electron density on a covalently bound halogen atom, as discussed by *Wang et al* [11]. In particular, when a halogen atom

covalently binds to another atom, a σ bonding orbital and a σ* antibonding orbital are formed. The former is located between the covalently bound atoms while the latter is distributed along the extension of the bonding axis but outwith the internuclear region. The σ orbital is occupied by electrons and hence a bonding interaction arises. However, the σ* orbital is an unoccupied orbital (if it were fully occupied the effect would be to totally cleave the covalent bond). Furthermore, it is the σ* orbital's lack of occupancy by electrons that renders the region corresponding to this virtual orbital electron deficient. This electron deficient region, corresponding to the σ* orbital along the extension of the covalent bond, is described as a σ-hole. Due to its electron deficiency, the electrostatic potential in this region is positive.

Kolar and Hobza [12] authored a paper wherein they discuss four parameters that can assist in the quantification of the σ-hole. These parameters are (i) the "size", by which the authors mean (in the case of a simple circular σ-hole) the angle defined by the atom covalently bonded to the halogen atom, the halogen atom and the circle corresponding to the line where the electrostatic potential ceases to be positive (for a more details definition including treatment of anisoptropy see figure 7 in in reference [12]), (ii) "range" "is defined as the distance between the halogen atom and a point at which ESP changes sign from positive to negative" [12] (ESP is electrostatic potential), (iii) "linearity" is the angle defined by the atom covalently bonded to the halogen atom, the halogen and the maximum in the positive electrostatic potential and (iv) "magnitude" as the value of the positive electrostatic potential at the point where it is greatest. The size and magnitude parameters had previously been discussed by Kolar, Hostas and Hobza [13].

Figure 1.1 is a schematic representation of a halogen bond between a halogen atom "X" and a nucleophile ":Y".

Figure 1.1: Schematic representation of a halogen bond. X is a halogen atom covalently bound to group R, interacting by a non-covalent halogen bond with nucleophile :Y.

Figure 1.2 shows a typical electrostatic plot for a molecule with a large σ-hole (in this case astatobenzene). The blue regions have a positive electrostatic potential. The prominent blue region facing the viewer indicates the σ-hole. The atom seen approximately in the centre of this blue region in the astatine atom (indicated as "X" in figure 1.1). The nucleophile engaged in halogen bonding (depicted as ":Y" in figure 1.1) is stabilised by this positive electrostatic potential and would be located in front of the astatine molecule (as viewed by the reader), this positioning of the nucleophile corresponds with approximately linear geometry for the angle R-X●●●Y.

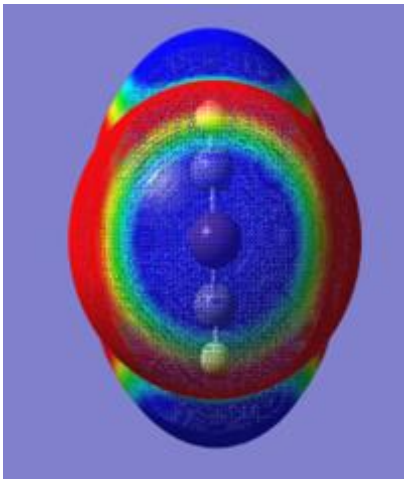


Figure 1.2: Plots of electrostatic potential for isolated astatobenzene. Density surface used for mapping is 0.0004 e-/au^3. The depicted electrostatic potential range is -6.93x10^{-3} Eh (red) to 6.93x10^{-3} Eh (blue).

Hence, the σ-hole typically is directly opposite the covalent bond by which the halogen is attached to its molecule, leading to a near-linear geometry between the atom to which the halogen atom is covalently bound, the halogen atom itself and the atom that donates electron density into (or electrostatically interacts with) the sigma hole, as explained by Clark *et al* [14]. However, as reported by Auffinger *et al.* [15], complex environments (such as those that can be found in biological systems) can give rise to substantially non-linear halogen bonding due to secondary polarization of the halogen atom's electron density. As can be appreciated by reference to figure 1.2, a σ-hole with greater angular size (measured in Å^2) (using the same definition of "size" as was described by Kolar and Hobza [12] and discussed above) (and hence covering a greater range of potential R-X•••:Y angles) helps to facilitate the accommodation of halogen bonds that are further from linear where secondary interactions favour such departure from the ideal halogen bond angle. Kolar *et al.* [13] authored a paper [13] wherein they discuss the trend in both the size and magnitude of halogen bonds (same definitions as recounted above) going down the halogen group, specifically in the case of halobenzene, wherein the halogen was chlorine, bromine or iodine. They found that the increase in magnitude was responsible for an increase in stabilisation going down the group, while an increase in size also resulted in greater tolerance of non-linear halogen bond angles. The authors also found that the effect in relation to tolerance of non-ideal bond angles was greater where the nucleophile was polar (hydrogen fluoride) than where it was non-polar (argon) due to the greater role of electrostatic attraction in the former case. Authors also comment on the effects of substitution of carbon atoms within the benzene ring itself and substitution of attached hydrogen atoms upon the size and magnitude of the σ-hole.

There is a general trend towards a greater magnitude of charge transfer into the σ-hole going down the halogen group as the polarisability of the halogen atom increases as discussed by Hill

and Hu [8]. It is known [8, 16] that there is a general trend towards stronger halogen bonding as the halogen group is descended.

In addition to the halogen group, tetrel, chalcogen and pnictogen atoms, (in groups 14, 15, and 16 respectively) form analogous interactions [17]. Each of these types of interaction are based upon the hereinbefore described σ-hole effect. The substantive difference is that the tetrel, pnictogen or chalcogen atom takes the place of the halogen atom. While there has been a large increase in the amount of research performed on halogen bonds [18-19], the other σ-hole bonding interactions are less well studied based upon the number of results for each type of interaction in a Web of Science Search , with respectively 1854, 184, 93, and 50 results for the search phrases (with quotation marks) “halogen bond”, “chalcogen bond”, “pnictogen bond”, and “tetrel bond” as the sole search terms in a search of the Web of Science database on 7th February 2018 [20]. This survey does not account for halogen, pnictogen, chalcogen or tetrel bonding that is not referred to by those names. The underlying assumption that the same proportion of each bond type will be so referenced has not been verified. However the above survey can still provide a qualitative guide as to the extent to which each of these categories of interactions has been studied. The term “tetrel bonding” was coined as recently as 2013 [17]. A paper authored by Stenlid and Brinck in 2017 [21] identified the σ-hole bonding interaction as being responsible for the catalytic properties of gold and platinum nanoparticles. Gold and platinum are transition elements in groups 11 and 10 respectively. Hence it can be seen that the σ-hole effect is not peculiar to the halogen group alone. Halogen bonding is merely a specific context in which this anisotropic electron density distribution can occur.

Consistently with all other types of non-covalent bond, rather than being a single, relatively straightforwardly identifiable type of interatomic interaction, (e.g. a non-polar covalent bond), halogen bonds possess multiple components in the overall halogen bonding interaction as discussed by Kozuch and Martin [22]. The authors listed the following components of the total halogen bonding interaction: electrostatic, dispersion, charge transfer and polarisation interactions.

Methods have been developed for probing the nature of internuclear interactions, including halogen bonds. For example the Wiberg bond index (WBI) [23] provides information on the degree of covalency of a halogen bond; this method requires the computation of natural bond orbitals [24]. The physical basis of halogen bonding in context of supramolecular chemistry is discussed in a paper authored by Gilday *et al.* [25].

Symmetry Adapted Perturbation Theory (SAPT) [26-28] can be employed to resolve the energy of the halogen bond into its components [29]. Hill and Legon [30] have performed SAPT calculations on both halogen and hydrogen bonds between small molecules. For the halogen bonded dimers, the halogen-bearing species was ClF, with the chlorine atom enganged in halogen bonding with either (i) formaldehyde, (ii) vinyl fluoride, (iii) oxirane, (iv) thiirane, (v) methylenecyclopropane, or (vi) 2,5-dihydrofuran. In each case the interaction energy of the halogen bond was resolved into four components, (i) electrostatic, (ii) exchange (in this case used in the sense of exchange-repulsion (a net positive component)), (iii) induction and (iv) dispersion. The energy was calculated for different halogen bond angles, measured in terms of deviation from linearity,

from -15° to +25°. For all of these systems, and at all computed halogen bond angles, the halogen bond was found to be energetically favourable (total SAPT energy below zero), with the electrostatic component being the greatest stabilising component in all cases, albeit by varying margins, with induction being almost as significant in the case of thiirane at the ideal halogen bond angle. There is not the same consistency in the order of the contributions from induction and dispersion, with the latter making the greater contribution in the case of vinyl fluoride, while the magnitude of induction was found to be greater for the other systems. While electrostatic attraction was the greatest contributor to the energetic stability of the halogen bond, it was not found to be responsible for the preference for linear halogen bond angle. Instead that was explained by the an increase in the magnitude of the (positive) exchange term as the halogen bond angle departed from linear. Indeed, the electrostatic component generally becomes more strongly negative with greater departure from linearity, which would have the opposite effect to that which is observed, although that trend is shown to vary between the complexes. In all cases the shape for the graph of the total energy is similar to that for the the exchange component thereof.

Findings by Hill and Legon [30] regarding the role of exchange in driving the angle of halogen bonds accords well with a previous study performed by Stone [31]. Stone also performed calculations wherein the chorine atom in ClF was engaged in halogen bonding, except in one case where Cl_2 halogen bonded with carbon monoxide (interacting with the carbon atom). In the other cases, ClF halogen bonded with (i) N_2, (ii) ethyne (interacting with the triple bond), (iii) ethene (interacting with the double bond), (iv) water, or (v) SO_2 (interacting with an oxygen atom). The angular deviation from linear was varied from 0° to either 30° or 40°, depending on the system. Unlike Hill and Legon [30], Stone [31] found that the electrostatic term did not always become more strongly negative as the halogen bond angle deviates further from linear, in particular, he found the converse to be true for dimer (v). Stone also found that the electrostatic term did not always clearly constitute the greatest component of the halogen bond, with almost identical levels of contribution from dispersion in the case of (i) and the system comprising Cl_2 and carbon monoxide. Stone found that in most cases, dispersion was more significant than induction, with the difference between the contributions increasing with increasing deviation from linear, although the components were found to be of almost identical value (with the order of contribution reversed by an extremely small margin) in the case of (iii) at angles close to linear.

Schneider *et al.* [32] performed a study including SAPT calculations on halogen bonded structures, with the halogen atom engaged in halogen bonding being iodine. The bond distance was varied. The electrostatic component was usually found to be the greatest component, but in some cases it was surpassed by induction, charge transfer also found to be an important component in many systems. In contrast to the work by Hill and Legon [30] and, especially, Stone [31], Schneider *et al.* [32] found dispersion to be a generally minor component of the interaction energy. At the optimal bond distance, electrostatic interaction was either the greatest or joint greatest component, with induction becoming more significant at shorter internuclear distances.

The work of Hill and Legon [30], together with the work of Stone [31], show that the electrostatic interaction is typically, but not invariably, the predominant component of the halogen bond's

interaction energy, with dispersion sometimes being equally significant. Furthermore, these studies only cover a small range of structures, with ClF being the halogen-bearing molecule in most cases, so if the order of contribution is not consistent between induction, dispersion, and electrostatic even among these structures, the only observation that can confidently be made is that their respective contributions are highly sensitive to the nature of the molecular system in which they exist. By contrast, and clear and consistent observation in both studies is the determinative role performed by exchange-repulsion in goverining the optimal angle for the halogen bond. The work of Schneider *et al.* [32] suggests that the relative contribution from dispersion is smaller for the heavier halogens. Dispersion is an attractive energetic component found to varying degrees in all chemical systems, so its relative contribution could be expected to decline as the magnitude of those components more specific to halogen bonds increases. This would be consistent with the observed [16] trend towards stronger halogen bonding going down the halogen group.

Relativistic effects were found to be important for the heavier halogens, including bromine [22]. However Hanus *et al.* considered the significance of relativistic effects on proton affinities for a system consisting of a 5-bromouracil molecule and a water molecule, and found that the relativistic effect was not significant in respect of proton affinity [33]. As discussed by Politzer *et al.* [18], while the calculation of enthalpies for halogen bond interactions is the usual measure employed as a proxy for the halogen bond strength, it should be noted that halogen bond formation is typically entropically unfavourable, and therefore caution should be employed when drawing inferences about the Gibbs free energy from enthalpies, for systems at temperatures that are substantially above absolute zero. Crystal packing effects were identified by Schaub *et al.* [34] as being crucial in producing the exceptionally short halogen bond distance, 2.683(5) Å, between an iodine atom and an oxygen atom, in a system comprising aryl iodide and "a bridged triarylphoshine oxide" molecule. The authors report that this is the shortest P=O•••I halogen bond to be identified.

There have been several studies on halogen bonding in biological and pharmacological contexts [35-43]. An overview can be found in a paper entitled "Halogen bonding (X-bonding): A biological perspective", authored by Scholfield *et al.* [44]. The present study does not focus on the biological aspects of halogen bonding, although it is based primarily on calculations performed on halogenated forms of molecules that may be found in biological systems, in particular,halogenated DNA base pairs, halouracil and fragments molecules found in the thyroid system, especially a fragments of thyroxine. Halogen bonding also has a role in materials chemistry [45-46]. For example, Mittapalli *et al.* [46] discuss the role of halogen bonding in the context of the thermochemical properties of crystals. The authors note "excellent thermal response" of the weak Cl•••Cl halogen bond in comparison with the stronger I•••I interaction. Stumpel *et al.* [47] discuss halogen bonding in the context of polymer chemistry.

Halogen bonds have been compared with analogous hydrogen bonds by substituting hydrogen atoms that are involved in hydrogen bonding between DNA bases with halogen atoms by Parker *et al.* [35]. That study found that hydrogen bonds were generally stronger than halogen bonds, and when comparing a particular system, hydrogen bonding (without any substitution with halogen

atoms) produced a more stable system. However, when discussing halogen bonding more broadly the authors commented that the strongest halogen bonds are sometimes of comparable or greater strength than the weakest hydrogen bonds, citing the examples of sulfur and selenium being the electron donor atom involved in the halogen bond. The present study also investigates substitution of hydrogen bonds with halogen bonds between DNA bases.

Kolár and Tabarrini[41] have authored a paper on halogen bonds involving nucleic acids, based upon statistical survey of the Protein Databank (PDB)[48]. Kolár and Tabarrini's study was performed with the aim of furthering knowledge of the potential role of halogen bonds to halogenated nucleic acids in drug design. The paper acknowledged the limited sample size, 21 contacts satisfied the IUPAC definition of a halogen bond[7]; although the sample expanded to 72 contacts if the criterion that the internuclear separation not exceed the sum of VdW radii be replaced by a cut-off of 4 Å for that distance. However, the authors identified a positive correlation between the halogen atom's atomic number and the halogen bond angle, and an inverse correlation between the atomic number and the internuclear separation. The authors concluded therefrom that halogen bond strength and its significance in driving the geometry of the structure increases going down the halogen group. In particular the authors found that chlorine-based halogen bonds, with the largest halogen bond angle being 154°, did not significantly influence the geometry. By contrast bromine and iodine (although the sample size for the latter was especially low) were found to be promising candidates for the use of halogen bonding in pharmacy. The study did not consider fluorine as a halogen bond donor, nor were astatine or tennessine investigated. Xu *et al.* [42] have also authored a paper on the potential importance of halogen bonding in drug development.

Mondal and Mugesh [39] investigated the role of halogen bonding in the regioselective cleavage of an iodine-carbon covalent bond, to produce the biologically active T3 derivative of inactive T4 thyroxine. The authors identified iodine – iodine halogen bonds as well as iodine – selenium halogen bonds as assisting the deiodination. Fortino *et al.* [49] also investigated the mechanism of deiodination, more specifically, they identified iodine ••• selenium and iodine ••• sulfur halogen bonds as performing a role in the cleavage of the C-I covalent bond. In the present study the possibility of the presence of iodine ••• oxygen and iodine ••• nitrogen halogen bonds involving T4 are investigated.

Halogen bonding in microsolvated 1-methyl-5-halouracil and its simpler analogue, microsolvated halobenzene, are hereinafter discussed. Bromouracil is a known mutagen [50]. The chemistry of this compound has been discussed by Holroyd and Van Mourik [50]. Voth *et al.* [43] identified that halogen bonding between bromouracil and an oxygen anion on phosphate could be exploited to direct molecular conformation. The halogen bond was estimated to be stronger than the equivalent hydrogen bond by between 8.4 kJ mol^{-1} and 20.9 kJ mol^{-1} (the paper refers to these values as 2 kcal mol^{-1} and 5 kcal mol^{-1} respectively), the interaction energy varying with the halogen bond's geometry. The present investigation seeks to elucidate the role of halogen bonding and competing hydrogen bonding in a system comprising one 1-methyl-5-halouracil molecule and either one or two water molecules, not only where the halogen is bromine, but for all halogens except tennessine. For further discussion of systems comprising halogen and

hydrogen bonding see a paper authored by Domagała *et al.* [51]. The authors discuss both cooperative and anti-cooperative effects between the two types of interaction.

The effect of fluorination upon the balance between halogen bonding and hydrogen bonding is discussed in a paper authored by Geboes *et al.* [52]. The authors studied dimethyl ether and trimethylamine using experimental and high level computational methods. Fluorination of the electron density acceptor, enhanced both hydrogen bonding and halogen bonding, but did not have a significant impact on the energetic composition of the bonds, or the composition of non-covalent bonds in the system.

Halogen bonding has been found to play important roles in several fields, including in drug design, molecular recognition and materials science, as discussed by Cavallo *et al.* [3]. The authors extol the properties of halogen bonding in each of these fields. Beyond the biochemical sphere, halogen bonds play an important role in organic synthetic chemistry and organocatalysis [53]. Bulfield and Huber discuss the differences between halogen bonding and hydrogen bonding in this context [53]. Boron's chemistry has also been found to encompass halogen bonding, as discussed by Alkorta *et al.* [54] with boron contributing a lone pair of electrons to a halogen bond with chlorine [54]. This finding might be surprising because, as commented upon by the authors, boron is usually considered to be an electron acceptor. However, there were prior studies [55-58] identifying boron as a potential electron donor.

Mukherjee *et al.* [59] discuss the role of halogen bonding in the context of the design and engineering of crystals. Their paper compares halogen bonding with hydrogen bonding in this context. The ability to tune halogen bond strength as well as the size of the atom participating therein were identified as potentially useful in crystallographic design, particularly in the case of ternary cocrystals.

Investigations have been conducted with the aim of comparing different computational methods for theoretically probing halogen bonding in various chemical systems[9, 22, 60-61]. One example thereof is the aforementioned study conducted by Kozuch and Martin which considered halogen bonding between dimer pairs of molecules [22].

Kozuch and Martin compared various DFT and post Hartree-Fock methods on two sets of halogen bonded dimers, the "XB18 Set" for comparing the treatment of geometries, and the larger "XB51 Set" for comparing the treatment of the dissociation energies of the halogen bonds [22]. Their study covered several aspects of computational methodology to compare their impact on accuracy; these included basis sets, the amount of exact exchange in hybrid functionals, relativity effects, and dispersion corrections. The CCSD(T)/CBS method was used for dissociation energies forthe XB18 set. The basis set extrapolation was performed using the aug-cc-pVXZ [62-63] (X= Q, or 5) basis sets, (in the case of bromine and iodine aug-cc-pVXZ-PP [64] (using ECP) variants thereof were used; the authors referred to both of these sets of basis sets using the abbreviation "aVXZ") [22]. For the XB51 set, in the interest of being less computationally demanding, a different extrapolation method was used based on density fitting MP2. However, this less computationally demanding method was found to yield results that were very similar to the CCSD(T)/CBS extrapolation method for the XB18 set. For XB18 geometry, optimisations were

performed using CCSD(T)/aug-cc-pVQZ (the ECP variant thereof in the case of bromine and iodine), while ωB97X [65] with the aVTZ basis set was used correspondingly for the XB51 set. The authors noted that ωB97X/aVTZ was found to be one of the most accurate investigated methods for geometry optimisations in the XB18 set [22], when compared against the CCSD(T)/aug-cc-pVQZ(-PP) reference.

Kozuch and Martin made numerous observations and recommendations when comparing the efficacy of different computational methods [22]. For the XB18 data set both SCS-MP2 [66] and SCS(MI)MP2 [67] were found to perform well for producing accurate energetic results [22]. Where c_s and c_o refer to same and opposite spin components respectively, without spin component scaling c_s and c_o each have a scaling factor of 1.00. c_s is scaled by a factor of 0.33 and 1.29 in SCS-MP2 and SCS(MI)MP2 respectively, whereas C_o is scaled by 1.2 and 0.4 in the respective methods. Hence there is a weighting in favour of C_o in the former method and a weighting in favour of C_s in the latter method. However the authors comment that it is the overall scaling down from 2.0 to approximately 1.6 (1.53 and 1.69 for SCS-MP2 and SCS(MI)MP2 respectively) for the aggregate spins that is most relevant [22]. The authors based this finding on root-mean-squared deviations for the dissociation energies of the XB18 set with different sums for the same and opposite spin components and found that there was "a canal of accuracy around $c_s + c_o \approx 1.6$" [22]. MP2 was noted to overbind. Regarding DFT methods, the authors found that M06-L [68] performed most strongly among the GGA functionals, while BMK [69], M06-2X [70], CAM-B3LYP [71] and ωB97X [65] were the leading hybrid functionals, the latter two being parameterized for long-range interactions [22]. In general, hybrid functionals outperformed GGA functionals [22]. Geometric results for the XB18 data set showed similar trends regarding the reliability of methods as for energetic results, but additional observations were made that MP3 and SCS-MP3 [72] performed less well [22], while MP2.5 [73] was found to have remarkable accuracy [22].

BSSE was noted as being a problem at the reference CCSD(T) level with smaller basis sets; counterpoise correction [74] was noted as being a partial remedy to this problem [22]. The authors recommended that the addition to the CCSD(T) energy of an MP2/CBS correction, estimated by basis set extrapolation, could provide a fast and accurate composite method, particularly if density fitting is applied in the MP2 calculations.

For the XB51 data set similar conclusions to those mentioned above in relation to the XB18 data set were made along with some additional observations [22]. Wavefunction-based methods generally outperformed DFT-based methods; however DSD-PBEh-B95 [75] and DSD-PBE-P86 [76] are noted for their high accuracy and their proficiency in describing other bonding interactions well in addition to halogen bonds. Both of these functionals are double hybrid functionals. The term ."DSD-DFT" is defined as Dispersion corrected, Spin- component scaled, Double-hybrid DFT" [75]. These functionals are ultimately based upon PBE [77-78], and the non-Hartree-Fock component of the exchange energy comes from PBE, while the final term, "B95" or "P86" relate to the MP2 correlation component [75].For hybrid functionals (which were generally recommended in preference to GGA functionals) the optimal amount of Hartree-Fock exchange energy was found to be approximately 50%, with the M06-2X functional (54% exact HF exchange [70]) being noted for its high accuracy, and it was found to be more accurate than any of the other hybrid

functionals (encompassing hybrid-meta functionals) tested in the study [22].

At the lower end of the computational cost scale, Scholfield *et al.* [79] present a force field model for the treatment of halogen bonds. The authors suggest that their model can be used in the development of biomolecular materials and pharmaceutical products.

Chapter five of this thesis includes a statistical analysis of halogen bonding in the Cambridge Structural Database (CSD) [80]. This study is not the first to investigate halogen bonding in this manner, and examples of previous work in this vein [15, 81] are discussed here.

A paper authored by Auffinger *et al.* [15] discloses a survey of halogen bonds in single-crystal structures wherein a halogen atom forms a halogen bond with an oxygen atom contained in the Protein Databank (PDB) [48]. In this study, all halogen elements no heavier than iodine were investigated, although results were obtained in the case of fluorine they are not discussed in the analysis of the results (on account of fluorine's electronegativity). This study focusses on halogen bonding wherein the halogen atom is covalently bound to a carbon atom and is located on either (i) a halogenated nucleotide or (ii) a drug molecule and the oxygen atom is located on a protein, nucleic acid molecule or ligands covalently bound thereto. The criteria imposed were: resolution of the the crystal structures 3.0 Å or superior quality, halogen bond length no greater than the sum of VdW radii and halogen bond angle no less than 120°. A total of 113 X•••O contacts satisfying the above criteria were found. The proportional attribution between the halogens were 27%, 34% and 39% for chlorine, bromine and iodine respectively. Futhermore, the nature of the oxygen atom varied between C=O (carbonyl), O-H (hydroxyl) and monoanionic Z-O^-, where Z is carbon, phosphorus or sulphur (anionic oxygen); respectively, these species accounted for 81, 18 and 14 X•••O contacts.

Whereas in the present study, the absolute internuclear separation has typically been found to decrease going down the halogen group (with reducing VdW ratios outweighing the increase in VdW radius), Auffinger *et al.* [15] do not share that finding. Instead they reported that the average X•••O distances were 3.06 Å, 3.15 Å and 3.24 Å respectively. The authors found there to be no correlation between internuclear separation and the nature of the oxygen atom. This result implies that varying the electron density distribution of the electron donor has no impact upon the halogen bond distance. The authors found that the shortest contacts were exhibited by structurs with halogen bond angles in the region between 160° and 180°. However, they futher observed that there was no other correlation between the halogen bonds' angles and distances. As a function of halogen bond angle, two distinct maxima in the number of structures were found in the regions of 160° to 170° and 145° to 150° (both regions approximate), with a minimum at approximately 155°. The authors attribute this "bimodal" distribution to secondary polarisation effects. The average angles for each halogen element across the whole dataset were 151°, 154° and 157° for chlorine, bromine and iodine respectively. This trend is attributed to the increasing polariation of the C-X bond. This effect must be outweighing, at least in the authors' statistical sample, any countervailing trend towards greater susceptibility towards secondary polarisation as the halogen group is descended. All of these angles are far from linear, implying that secondary polarisation effects can lead to substantial departure from the ideal halogen bond angle for chlorine, bromine and iodine.

Nemec *et al.* [81] published a paper looking at interactions of bromine in crystal structures in the CSD. Although much of their work focussed on bromine anions as donors of electron denisity, neutral dibromine (Br_2) was investigated for its formation of halogen bonds in crystal structures. The criteria for the search for halogen bonds were a bond angle of between 90° and 180° and an internuclear separation no greater than the sum of VdW radii. A more generous internuclear separation was employed in the search for hydrogen bonds. In total, 28 structures containing dibromine were found; including 11 structures wherein the halogen bond was to either nitrogen, oxygen or sulphur, 15 were "to cocrystals of bromine with halogenated fullerenes" and a further two structures included a halogen bond to a π-system.

The paper discloses the relationship between halogen bond length and angle is not smooth but that there is a clustering of structures at angles close to linear with internuclear distances between 2.7 Å and 2.9 Å. There is also a smaller and less concentrated cluster in the region between 3.1 Å and 3.4 Å, but these structures have greater variation in bond angle. Hence it can be seen that within the statistical sample, a preference for geometries combining short contacts and linear bond angles can be identified, implying that as halogen bond strength increases (at least as measured by internuclear separation) there is an increasing propensity towards the ideal (linear) bond angle. This finding of two distinct regions where halogen bonds form, with the largest cluster being in the region close to linear and short internuclear separations is similar to the hereinbefore referenced study by Auffinger *et al*.

The authors also compared the ratios of the internulear separations between the participating atoms in the halogen bond with their VdW ratios, with results broken down between the species that were acting as halogen bond acceptors (i.e. the X•••:Y distance from figure 1.1 above). Furthermore, the deformation (elongation) of the covalent bond length (R-X bond length in figure 1.1) was also quantified. For the halogen bond distances, the internuclear separations were shortened by 12%, 17%, 16% and 8.7%, 2.5% and 4.3% for the nitrogen, sulphur, oxygen, π-systems, "organic chlorine" and "organic bromine" respectively. Their respective lengthening of the R-X covalent bond were found to be 1.8%, 6.2%, 2.2%, 1.3%, 0.44% and 1.3%. Based upon these results, the authors concluded that, with the exception of "organic chlorine" and "organic bromine" (which showed smaller departures from standard distances), these interactions exhibited partial covalency, but did not elaborate on their reasoning.

2 Background theory to the computational methods employed in the present investigation

Several methods have been employed in elucidating the role of halogen bonding in biomolecules. Most of these methods have their origin in density functional theory, itself predicated upon the veracity of wavefunction theorems. These models and the computational methods that are based thereupon are underpinned by the postulates of quantum mechanics and their application to molecular systems. The Schrödinger equation [82] provides a valid starting point for all of the quantum mechanical methods with direct chemical applications. Theories and postulates of quantum mechanics [83] with more general application to physical systems but which are only implicitly engaged for describing molecules and their interactions, for example the Heisenberg uncertainty principle and wave-particle duality, are not directly discussed herein.

2.1 Schrödinger equation

The Schrödinger equation holds that when the Hamiltonian operator $\hat{H}$ acts on an eigenfunction ψ, the function $E\psi$ is yielded where the eigenvalue E is the energy of the system with wavefunction ψ, and which may be written as equation 2.1 [82]:

$$\hat{H}\Psi = E\Psi \qquad (2.1)$$

All of the observable physical properties of the system are encapsulated within the wavefunction thereof, therefore an exact solution to the Schrödinger equation would fully elucidate all of the physical (and hence all of the chemical) properties of the system.

The explicit form of the Hamiltonian ($\hat{H}$) is set forth in equation 2.2. Therein, $\hbar$ is the quotient of Planck's constant (h) divided by 2π, m_e and m_k are the masses of the electron and nucleus k respectively, e is the elementary charge (of an electron), Z is the number of protons In a nucleus (atomic number), ∇^2 is the Laplacian operator, the indices i and j relate to electrons and the indices k and l relate to nuclei, and r is the separation between the particles indicated by the subscript symbols (e.g. r_{ik} is the distance between electron i and nucleus k) [82].

$$\hat{H} = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_k \frac{\hbar^2}{2m_k}\nabla_k^2 - \sum_i \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{i<j} \frac{e^2}{r_{ij}} + \sum_{k<l} \frac{e^2 Z_k Z_l}{r_{kl}} \qquad (2.2)$$

A limiting case of equation 2.1 is where the system is in its ground state. All of the systems that have been investigated in the present study have been probed in their ground state (equation 2.3).

$$\hat{H}\Psi_0 = E_0\Psi_0 \qquad (2.3)$$

The Hamiltonian operator can be deconstructed into electrostatic potential and kinetic energy terms (equation 2.4). Here the vector **r** is the electron spatial coordinate, V is the electrostatic potential and T is the kinetic energy. Equation 2.4 holds where the system under consideration behaves independently from time, i.e. behaves according to the time-independent Schrödinger equation.:

$$\hat{\mathrm{H}}(\mathbf{r}) = \hat{\mathrm{V}}(\mathbf{r}) + \hat{\mathrm{T}}(\mathbf{r}) \qquad (2.4)$$

2.2 Born-Oppenheimer approximation

The Born-Oppenheimer approximation arises from the very large relative differences in the mass of a nucleus (which contains positively charged protons) and electron and hence the speeds at which they move. On the timescale upon which electron density redistributes with respect to position, the nuclei appear fixed in space. Therefore the nuclear electrostatic potential energy can be treated as a constant and nuclear kinetic energy can be disregarded when computing the electronic Hamiltonian. The effect of the Born-Oppenheimer approximation, separating the electronic motion from the nuclear motion, is to allow the nuclear and electronic components of the Schrödinger equation to be calculated separately. This approximation, known as the Born-Oppenheimer approximation, greatly reduces the computational expense of performing quantum chemical computations [82].

2.3 Hartree-Fock theory

The Hartree-Fock theory (HF theory) is the least computationally demanding ab-initio method for describing chemical systems quantum mechanically. However, the failure of this method to describe electron correlation limits its accuracy. HF theory relies upon an approximation (often highly flawed) that each individual electron will only be perturbed by an isotropic distribution of the electron density corresponding to all of the other electrons, and the positively charged nuclei [82].

In HF theory the molecular wavefunction is expressed as a Slater determinant, with antisymmetric qualities that ensure conformity with the Pauli exclusion principle applicable to all fermions. Within the Slater determinant each column represents a molecular orbital (MO) that contributes to the total molecular wavefunction. Each row in the determinant corresponds to a set of coordinates (both spatial and spin) for an electron. The number of columns and rows is equal to the number of MO wavefunctions and electrons that form the overall molecular system [82].

Equation 2.5 shows the form of a Slater determinant. In this equation Ψ is the overall wavefunction for the molecule, ψ_n is the nth molecular orbital which contributes to Ψ, and x_n is the set of coordinates (spatial and spin) for the nth electron. When shown in this form it is immediately clear, by applying the rules of determinants, that exchanging the coordinates of

two electrons produces a function that is antisymmetric with respect to the original function, thereby adhering to the Pauli exclusion principle [84].

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \cdots & \psi_n(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \cdots & \psi_n(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{x}_n) & \psi_{2(}\mathbf{x}_n) & \dots & \psi_n(\mathbf{x}_n) \end{vmatrix} \qquad (2.5)$$

Applying the HF theory as a method for investigating multi-electron systems relies upon an approximation that holds that each individual molecular orbital (column in the Slater determinant) can be described by a linear combination of atomic orbitals. Each contributing atomic orbital is multiplied by a scalar coefficient to optimise its contribution to the molecular orbital. The variational theorem holds that a non-optimal approximation to the true wavefunction will yield a higher (more positive) energy than that which would be yielded by the true function, when the Hamiltonian operator acts thereupon. Therefore, given an initial guess for the values of the coefficients, and hence the form of the overall molecular wavefunction, the coefficients can be varied to yield a lower energy as they come closer to their true values until the successive iterations yield a variation in energy that is within an arbitrary convergence threshold [82].

Equation 2.6 shows the how the LCAO principle operates. ψ_i is the i^{th} molecular orbital, φ_μ is the μ^{th} atomic orbital contributing to ψ in proportion to the scalar coefficient c. There are K basis functions contributing to the molecular orbital. Each φ_μ is a basis function contributing to the basis set [82].

$$\psi_i = \Sigma_\mu^K c_{\mu i} \varphi_\mu \qquad i = 1,2, \dots K \qquad (2.6)$$

Each MO can be described using Slater functions or Gaussian functions. In principle Slater functions more closely correspond to nature. In practice Gaussian functions are used as a more computationally accessible approximation to a Slater function. In particular, the product of two Gaussian functions is another Gaussian function. A Slater function has an exponential form, $e^{-\alpha r}$. , whereas Gaussian functions have the form $-e^{-\alpha r\hat{}2}$. Both Slater and Gaussian functions tend to a finite constant at zero distance. A Gaussian functions is a bell-shaped distribution with the gradient tending to zero as distance tends to zero when close to zero distance. Hence the space that is very close to the nucleus is not accurately described by a Gaussian function; however, at the distance from the nucleus where electrons are generally found, which is therefore the space of chemical interest, the shape of a Gaussian function is similar to the exponential form of a Slater function and hence the former can be used as a sound approximation for the latter. Using a contraction of multiple (primitive) Gaussian functions to approximate a Slater function produces a more accurate representation of the gradient in the region of chemical interest. The

central distinction between a Slater function vs. a Gaussian function is that the former is a function of the exponential of the radius, whereas the latter is a function of the exponential of the square of the radius. The subject of basis sets is further discussed under its own heading [82].

The HF energy can be calculated as the sum of the electronic energy and the internuclear electrostatic potential energy, as shown in equation 2.7. Therein, E_{HF} is the Hartree-Fock energy, E_{el} is the electronic energy, A and B are distinct nuclei, there are a total of m nuclei in the system, Z_A and Z_B are the charge on nucleus A and nucleus B respectively, and R_{AB} is the internuclear separation between A and B.

$$E_{HF} = E_{el} + \Sigma_{A}^{m}\Sigma_{B>A}^{m}\frac{Z_A Z_B}{R_{AB}} \qquad (2.7)$$

The entirety of this study was performed using Density Functional Theory (DFT) (see section 2.8 herein), albeit with the inclusion of exact Hartree-Fock exchange due to the employment of a hybrid density functional. Therefore, the description of the Hartree-Fock method herein is being kept very brief, however a much more detailed description can be found in references [82-83].

2.4 Post Hartree-Fock methods

Although the Hartree-Fock method itself is now rarely used to solve chemical problems, there are a range of much more accurate methods that are rooted in Hartree-Fock theory but which incorporate electron correlation. These methods are known as post Hartree-Fock methods [83].

2.4.1 Second order Møller-Plesset perturbation theory

No part of this computational study engaged second order Møller-Plesset perturbation theory (MP2) [85]. Furthermore, no double hybrid functionals (which incorporate a fraction of MP2 correlation energy) were employed; however literature benchmark studies [22, 61, 86-88] which support the use of the M06-2X density functional [70] do make comparisons with MP2 and double hybrid functionals. The principal attributes of MP2 are disclosed below in brief terms. A more detailed description thereof may be found at [83].

The distinction between HF and MP2 is that the latter does not invoke the approximation of a uniform distribution of the electrons, but allows the distribution of electrons with respect to position to reflect the correlation of the electron positions arising from coulombic electron-electron repulsion. This improvement on HF theory is computationally demanding but dramatically enhances the accuracy of the calculation. The difference between the HF energy and the MP2 energy is known as the MP2 correlation energy, and is invariably negative [82].

The Møller-Plesset energy can be broken down into the sum of (one electron) Fock operators and the perturbation, the latter being scaled by coefficient λ ($0 \leq \lambda \leq 1$). The perturbation energy may be set out as a Taylor expansion of the following form:

$$\mathrm{E} = \mathrm{E}^{(0)} + \lambda\mathrm{E}^{(1)} + \lambda^2\mathrm{E}^{(2)} + \lambda^3\mathrm{E}^{(3)} + \cdots \quad (2.8)$$

$$\mathrm{E}_{\mathrm{HF}} = \mathrm{E}^{(0)} + \lambda\mathrm{E}^{(1)} \quad (2.9)$$

For the definition of a Fock operator see equation 2.10. Therein, F is the fock operator, h is the aggregate attractive force from all of the nuclei in the system, J is the coulomb operator and K is the exchange operator. The indices i and j correspond to electrons, in equation 2.10 the Fock operator acts on electron i, there are N electrons in the systems [83].

$$F_i = h_i + \sum_j^N (J_j - K_j) \quad (2.10)$$

From equation 2.8 and 2.9 it is clear that the MP2 correlation energy arises from the second order term in the perturbation theory, i.e. $\lambda^2\mathrm{E}^{(2)}$. If only this term is added then the second order perturbation level (MP2) is reached. Terms giving third order perturbation and beyond can be arbitrarily added, that is to say that the series can be truncated at whatever order is desired [83].

2.5 Basis sets

Basis sets are formed from a combination of basis functions. They are employed to model the molecular orbitals of the system under investigation, based on a linear combination of functions that model the atomic orbitals, which correspond to the basis functions [82].

An important question arises as to what is the most appropriate form for the basis functions. Viewed solely from the perspective of making the model as accurate as possible, Slater type orbital functions would be very suitable; however, these are computationally demanding to employ in practice, so instead a contraction of "primitive" Gaussian functions is usually used to approximate to the form of a Slater type orbital. One mathematical advantage of Gaussian type orbitals (GTOs) is that the product of two GTOs is another GTO [82].

As with other aspects of the computational method, there is typically a trade-off between computational efficiency and accuracy of the calculations [82]. In the case of the correlation consistent basis sets, developed by Dunning [62], increasing the basis set size systematically increases the accuracy of the computed results. By contrast, the Pople basis sets [89-90] are not correlation consistent, but increasing the basis set size does usually lead to an increase in accuracy. Some basis sets have also been developed to incorporate an effective core potential (ECP), also known as a pseudopotential, which uses a single, and therefore computationally efficient, function to model the core electrons, which usually have a relatively minor effect on

the chemistry of the atom or molecule. Furthermore, some ECPs are designed to reflect relativistic effects in form that is much less computationally demanding than an all electron relativistic basis set. In this study the requirement to include relativistic effects informed the choice of basis set for modelling iodine and astatine.

Diffuse and polarisation functions can be added as a means of providing additional flexibility for a basis set. A diffuse function attenuates gradually (small negative gradient due to small coefficient in the exponential term) as a function of distance from the centre of the basis set, hence they can be significant when long range interactions (e.g. halogen bonds) are important. A polarisation function polarises the phase of the function by adding a function with greater angular momentum [91].

2.5.1 A note of caution

There are two different ways in which d orbitals may be represented. The more natural representation is spherically harmonic, with the five different types of d orbital (d_{xy}, d_{xz}, d_{yz}, $d_{x^2-y^2}$ and d_{z^2}) being faithfully represented. However, an alternative, Cartesian, representation can also be employed. Under the Cartesian model the $d_{x^2-y^2}$ orbital is replaced with two orbitals: d_{x^2} and d_{y^2}. The former represents each d sub shell as five orbitals (5D) while the latter represents the same sub shell with six orbitals (6D). Consistency in the use of either 5D or 6D representations is crucial if relative energies are to be calculated from two or more calculations. By default, the Gaussian 09 software package [92] will apply the cartesian model for the 6-31+G* basis set [93], but if the "gen" keyword is employed (in this case in order to specify a different basis set for iodine and astatine), all d orbitals are modelled with the 5D (spherically harmonic) d orbitals. Consistency between calculations can be ensured by explicitly specifying the representation to be applied.

2.5.2 6-31+G*

With the exceptions of iodine and astatine, the 6-31+G* basis set [93] was employed throughout this investigation. This basis set is sufficiently small so as to make the calculations computationally affordable while yielding results with acceptable accuracy. This basis set is a split valence basis set. While the core electrons are modelled as a single contraction of six primitive Gaussian functions, the valence electrons are modelled by two functions, one of which is a contraction of three primitive Gaussian functions, and the other is a single primitive Gaussian function. The "+" denotes the addition of a diffuse function and the "*" indicates that a polarisation function has been added to the non-hydrogen atoms.

2.6 Relativistic effective core potentials

As discussed above, the orbitals corresponding to iodine and astatine atoms were modelled using a basis set with a relativistic effective core potential. These atoms are sufficiently heavy that electrons travel at speeds that are sufficiently close to the speed of light that relativistic effects can have a significant effect on their chemistry. Relativistic effects can sometimes but not always be ignored in the case of bromine [33]. Preliminary calculations performed in the course of this study indicated that relativistic effects were not significant in the case of bromine, with relativistic and non-relativistic treatment yielding very similar results with respect to relative energies and geometries; accordingly, the bromine atom has been modelled with the non-relativistic 6-31+G* basis set.

2.6.1 aug-cc-pVDZ-PP

The aug-cc-pVDZ-PP basis set is a double zeta correlation consistent basis set which is employed with a relativistic ECP, and has been employed in the modelling of iodine and astatine orbitals. The "aug" denotes that diffuse functions have been added while the "p" denotes the addition of a polarisation function. The employment of a distinct basis set for the modelling of the heavier halogen atoms is standard practice in the computational modelling of halogen bonds, see for example [22, 33, 35].

2.7 BSSE and the counterpoise correction procedure

Basis set superposition error (BSSE) arises in the calculation of interaction energies from the availability in the dimer (or for that matter any multimer) of the basis functions of both of the molecules that form the complex. The effect thereof is to increase the effective size of the basis set in comparison with the basis applied to each monomer. As increasing the basis set size allows for greater optimisation of the electron density distribution, the use of a larger basis set artificially results in the calculation of a more negative energy compared with the same calculation performed using the smaller basis set. This inconsistency in the calculation of the dimer energy versus the calculation of the corresponding monomer energies therefore causes the calculated interaction energy to be too negative [82]. One solution to this problem would be to employ very large basis sets as then even the monomer energies would be calculated close to the infinite basis set limit so the addition of even more basis functions in the dimer calculation would have little effect on the numerical result of the calculation. However, that approach is often not feasible as the computational cost of employing very large basis sets would be too great. Therefore, the effect of this error must be quantified by calculation and subtracted from the calculated interaction energy. This correction is achieved by invoking the counterpoise procedure developed by Boys and Bernardi [74].

The counterpoise corrected energy is calculated in the manner indicated by equation 2.11. In equation 2.11, ΔE^{cp} is the counterpoise corrected interaction energy, E corresponds to absolute energies, superscript letters A and B correspond to the basis set that is being used, subscript

indicates the molecule or dimer that is being computed and round brackets indicates the geometry at which the calculation is performed. A and B in isolation correspond to the A and B monomers respectively, while A/B indicates the dimer complex. For example, the first term after the equals sign is the absolute energy of the dimer complex, employing the dimer basis set at the dimer geometry.

$$\Delta E^{CP} = E_{A/B}^{\{A/B\}}(A/B) - E_{A}^{\{A/B\}}(A/B) - E_{B}^{\{A/B\}}(A/B) + E_{A}^{\{A\}}(A/B) + E_{B}^{\{B\}}(A/B) - E_{A}^{\{A\}}(A) - E_{B}^{\{B\}}(B) \quad (2.11)$$

The BSSE itself can be calculated by equation 2.12:

$$\mathrm{BSSE} = \mathrm{E}_{\mathrm{A}}^{\{\mathrm{A/B}\}}(\mathrm{A/B}) + \mathrm{E}_{\mathrm{B}}^{\{\mathrm{A/B}\}}(\mathrm{A/B}) - \mathrm{E}_{\mathrm{A}}^{\{\mathrm{A}\}}(\mathrm{A/B}) - \mathrm{E}_{\mathrm{B}}^{\{\mathrm{B}\}}(\mathrm{A/B}) \quad (2.12)$$

The deformation energy is expressed in equation 2.13, E_{def} is the deformation energy:

$$\mathrm{E}_{\mathrm{def}} = \mathrm{E}_{\mathrm{A}}^{\{\mathrm{A}\}}(\mathrm{A/B}) + \mathrm{E}_{\mathrm{B}}^{\{\mathrm{B}\}}(\mathrm{A/B}) - \mathrm{E}_{\mathrm{A}}^{\{\mathrm{A}\}}(\mathrm{A}) - \mathrm{E}_{\mathrm{B}}^{\{\mathrm{B}\}}(\mathrm{B}) \quad (2.13)$$

From equations 2.11-2.13 it is clear that BSSE accounts for the effect of the presence of the dimer basis set upon the energy of the molecules at the dimer geometry, while the deformation energy accounts for the changes in the internal structures of the molecules upon complexation (without regard to the impact of the larger basis set upon the energy), i.e. the geometric relaxation energy.

2.8 Density functional theory

An alternative approach to performing quantum chemical calculations, which is not directly based on calculating wavefunctions is density functional theory (DFT) [82-83]. DFT is predicated upon the Hohenberg-Kohn existence and variational theorems [94].

2.8.1 The Hohenberg-Kohn theorems

2.8.1.1 The existence theorem

The Hohenberg-Kohn existence theorem [94] is most fundamental to DFT. It holds that the energy of a system comprising electrons in the presence of nuclei (or indeed any external potential) is a functional of the electron density distribution as a function of position, relative to the external potential (nuclei). No two distinct external potentials will produce identical electron density distributions, hence a given set of nuclear coordinates will uniquely prescribe the density function. The basis for this theorem is purely in logic. The converse postulate is shown to be necessarily false by *reductio ad absurdum*. However, this is only an existence theorem; it

establishes that a given set of nuclear coordinates yields an electron density distribution function, but the form of the functional which relates them is not elucidated.

2.8.1.2 The variational theorem

Hohenberg and Kohn also presented a theorem whereby a trial electron density can be optimised in a self-consistent field [94]. A trial density which is not the true density will always, for a given functional, yield a higher energy than would be yielded by the true density, with the true density yielding the true energy according to the employed functional. Furthermore, the greater the accuracy of the trial density (i.e. the more closely it approximates to the true density) the more closely that the energy thereby yielded approximates to the true energy. Hence the trial density can be iteratively varied (modified) to yield a lower energy until further variation yields a change of energy that is no greater than an arbitrary convergence threshold. This variational theorem is analogous to, and logically stems from the variational principle that is intrinsic to Hartree-Fock theory. The Hamiltonian determines the electrostatic potential, which in turn determines the electron density distribution. Therefore, the variational nature of the Hamiltonian necessarily implies the variational nature of the electron density distribution.

2.8.2 Kohn-Sham Method

The Kohn-Sham method [95] facilitates the partitioning of the density functional into a component which models the exchange and correlation energy (which is not inherently known from DFT and must be the subject of approximation) and the remainder of the constituent parts of the density functional. This method invokes a reference system which has the same electron density distribution as the actual system but is modelled classically with non-interacting electrons. The non-interacting component is the sum of the electronic kinetic energy, the nuclear-electronic electrostatic potential and the electron-electron electrostatic potential. To this classical energy value, the exchange energy and the correlation energy (typically expressed as a single term) are added, and consist of quantum mechanical treatment of the kinetic energy and electron-electron potential energy of the electron density. The form of the exchange-correlation functional is not known from the Hohenberg-Kohn theorems [94]; if it were known then density functional theory would provide exact solutions for all chemical systems. Developments in DFT have been devoted to developing functionals that incorporate a representation of the exchange-correlation function that yields results that are as physically accurate as possible at a tolerable computational cost [83].

Variational methods, analogous to those used in relation to Hartree-Fock theory, are employed to minimise (optimise) the Kohn-Sham energy (as a functional of electron density). The Kohn-Sham operator is employed for this purpose. This operator treats as additive terms, the kinetic energy of the non-interacting (classical) system, classical treatment of the nuclear electrostatic

potential acting on the electron density, classical treatment of electrostatic repulsion for the electron density, and the exchange-correlation potential, which is the first derivative of the exchange correlation energy.

Hence, by treating the electrons as non-interacting, a problem that is very closely analogous to that of HF theory is set up, wherein coefficients for single electron orbitals may be variationally optimised [83].

2.8.3 Jacob's ladder

There are multiple levels of DFT methods [82, 96]. In (generally) ascending order of accuracy and computational cost these are functionals which use the local density approximation (LDA functionals), functionals which use the generalised gradient approximation (GGA functionals), functionals which take into account the kinetic energy as the second derivative of the density (meta-GGA functionals), functionals which incorporate a fractional portion of the (exact) HF exchange energy (hybrid functionals), meta-hybrid functionals which include the attributes of both hybrid and meta functionals [83]; and functionals which incorporate a fractional proportion of the MP2 correlation energy as well as HF exchange energy (double hybrid functionals) [97]. This ascending accuracy with each enhancement to DFT functionals is known as Jacob's ladder by analogy made by Perdew and Schmidt [98] to its biblical counterpart. Each step up the ladder leads up to DFT perfection (heaven).

2.8.3.1 LDA

The local density approximation (LDA) models each electron interacting with a uniform electron gas. Only the density at the point in space being analysed is considered, and is treated as uniform throughout the whole of space. Each electron is treated as being in a system with an isotropic potential whose value matches the immediate environment of the electron. This is the simplest but least accurate way to calculate the exchange correlation energy for a chemical system [83].

2.8.3.2 GGA

The generalised gradient approximation (GGA) allows a more accurate treatment of the chemical environment taking into account not only the local density but also the local gradient of the density as a function of position (i.e. the first derivative of the density with respect to spatial coordinates). Hence the density throughout the system is no longer treated as being uniform throughout the system but varying in accordance with its gradient. However this approximation is still not very realistic as it assumes a constant change in density based upon its gradient at the local point in space [83].

2.8.3.3 Meta-GGA functionals

Meta-GGA functionals are an improvement on GGA functionals as they not only take into account the first derivative of the density but also the second derivative, that is to say the kinetic energy. With adequate parametrisation, meta-GGA functionals can yield quite a high degree of chemical accuracy, but are still limited in comparison with post Hartree-Fock methods by the inexact treatment of the exchange energy [83].

2.8.3.4 Hybrid functionals

Hybrid functionals seek to overcome the limitations of DFT's inability to exactly treat the exchange energy by incorporating a fixed proportion of the HF exchange energy. This is computationally demanding as it requires that the Hartree-Fock energy be calculated *ab initio*. Hybridization with the HF exchange energy can be combined with inclusion of the second derivative of the density to produce meta-hybrid functionals [83]. Although incorporation of 100% of the HF exchange energy is possible and has been done in the case of (for example) the M06-HF functional [99], these have generally not performed well in benchmark studies [22, 87], instead a proportion that is well in excess of zero but well below all of the HF exchange energy is selected. For example the M06-2X functional [70] (a hybrid-meta functional – see below), which is used in this study, includes 54% of the HF exchange energy; this is a much greater contribution than is typical for hybrid functionals, or indeed hybrid-meta functionals (see below), the 2X term in the name relates to the contribution being double that for the M06 functional [70], which has a 27% contribution. Although these are actually Hybrid-meta functionals (see below) they illustrate the point that inclusion of 100% HF exchange is not optimal [22].

2.8.3.5 Hybrid-meta functionals

Some functionals incorporate the kinetic energy and hence satisfy the criteria to be classed as meta-GGA functionals but also incorporate a fraction of the HF exchange energy. These functionals are known as hybrid-meta functionals [83].

2.8.3.6 Double hybrid functionals

Double hybrid functionals are similar to hybrid functionals but additionally include a proportion of the MP2 correlation energy, in an analogous manner to the inclusion of a proportion of the HF exchange energy [97]. These functionals often demonstrate the best performance in benchmark studies [75, 87] but come at very high computational cost due to the need to calculate the MP2 correlation energy; as previously discussed, MP2 calculations are much more

computationally demanding than HF calculations [83]. Hence, despite their high accuracy their application remains limited in practice.

2.8.4 Empirical parameterisation for dispersion effects

DFT does not inherently model dispersion. The only exception to that statement being in relation to double hybrid functionals, which implicitly do include dispersion effects to some degree as they are reflected in the MP2 energy. However, dispersion often plays an important role in chemistry, including in relation to halogen bonds [22]. M06-2X [70] is an example of a functional which seeks to overcome this deficiency by parameterisation, in which the form of the functional is fitted to empirical data, thereby adjusting the functional to compensate for the lack of the dispersion in the theoretical model.

Another approach to incorporating dispersion effects in DFT, not employed in this study, is to impose a correction, external to the functional, to the calculated energy. Examples include Grimme's D2 [100], D3zero [101] and D3BJ [102] corrections.

2.8.5 M06-2X

The sole method that has been applied in this study is the M06-2X functional [70]. This is a hybrid-meta functional, which has been empirically parameterised to account for dispersion. This functional incorporates 54% of the Hartree-Fock exchange energy; this is double the HF contribution for the similar M06 functional [70]. Another distinction between the M06-2X functional and the M06 functional is that the former omits six empirical dispersion parameters that are included in [61] the latter, on account of the latter having a less diverse training set [70]. The M06-2X functional has been selected based upon extensive benchmark studies in the literature [9, 22, 70, 75, 87-88], wherein M06-2X has been found to perform strongly for computations upon halogen bonds as well as other interactions within the molecular system. As is always the case with computational studies there is a trade-off between quality and computational efficiency. While other methods, for example the spin density scaled double hybrid functionals DSD-PBEP86 [76] and DSD-PBEhB95 [75] and high-level post-Hartee-Fock ab-initio methods outperform M06-2X, this functional produces results that are within the bounds of acceptable quality for the purposes of this study while also being computationally feasible for the systems being investigated. In particular the combination of halogen bonds being the interaction of greatest chemical interest and the importance of other interactions, especially hydrogen bonds, made it necessary to select a functional that yielded particularly reliable results for halogen bonding while also having strong all-round performance, especially for other non-covalent interactions.

2.8.6 Integration grids

All density functionals rely upon integration grids for numerical integration. The fineness of the grid that is required for achieving accurate results varies upon the functional being employed and the system upon which the computations are performed. When the integration grid is specified, the number of points used per atom for the calculation is the product of the number of angular points and the number of radial points specified [103]. For some of the molecular systems that were studied in the course of this investigation the grid designated as "superfine" by the Gaussian 09 software package [92] was required for vibrational frequency calculations. This is a finer grid than is typical for DFT studies on molecular systems. The remainder of the calculations were performed using the "ultrafine" grid. Structures have been optimised using the ultrafine integration grid, and vibrational frequency calculations were performed at that level. Occasionally phantom imaginary frequencies were obtained and these were eliminated by recomputing the results using the superfine grid. Additional calculations were performed where required to confirm that the interaction (or complexation) energies and geometries varied only to a negligible extend between the two integration grids. Therefore, the geometric and energetic results computed using the ultrafine grid can be relied upon. The sole purpose for calculating vibrational frequencies in this study was to confirm the status of the minima (i.e. minimum or first order transition state). However, if more detailed vibrational information is sought (e.g. for simulating infrared spectra), the superfine grid is to be strongly recommended over the ultrafine grid when probing halogen bonding using the M06-2X functional.

The integration grids are based upon radial "shells" centred on each atom, hence the number of permitted radial coordinates corresponds to the number of shells, while each shell is intersected by a defined number of angular coordinates. Increasing the size of the grid entails increasing the number of permitted coordinates on the grid and hence the fineness of the grid [103]. The superfine grid employs 150 radial coordinates (shells) centred on each atom, with 974 angular coordinates intersecting each radial coordinate [104], hence the total number of points per atom is 146100. The corresponding values for the ultrafine grid are 99, 590 [104] and 58410 respectively. The Minnesota suite of functionals have been found in a previous study to require larger integration grids due to the constants incorporated within the functionals to achieve empirical parameterisation [105].

2.9 Quasi-Synchronous Transit Newton

Transition states were elucidated by means of the Quasi-Synchronous Transit Newton (QSTN) method [106-107]. In Gaussian 09 [92] this method can be employed in two variants. In one variant, invoked with the keyword QST2, the optimised minima that are connected by the transition state are given as input for the calculation, and from these data alone the transition state that connects them is computed. If a reasonable guess for the transition state can be provided then the alternative variant, invoked with the QST3 keyword, can be employed, which uses the

optimised minima in conjunction with the provided guess structure for the transition state. For a discussion of the mechanics and theoretical basis of QSTN see references [106-107].

2.10 Theoretical basis for vibrational frequency analysis for verifying minima/transition states

Vibrational frequency analysis has been used exclusively for the purpose of verifying that a stationary point is a minimum or a first order transition state. For this purpose it was only necessary to elucidate the number of vibrational modes which had imaginary frequencies [82]. Molecular systems typically have many vibrational modes and many degrees of freedom. Hence the potential energy surface is a high dimensional surface. Stationary points on the PES always have a first order derivative of zero in respect of each and every dimension. An imaginary frequency at a stationary point will not arise if all second order derivatives (in all dimensions) of the energy as a function of spatial coordinates are positive, which by definition means that that stationary point is a minimum. By contrast, as a matter of mathematics applicable to all functions, where the second order derivative of the gradient is negative, the stationary point is a maximum. At any one stationary point a maximum in respect of any one dimension of the potential energy surface will yield one vibrational mode with an imaginary frequency. For example, where a stationary point is a maximum in respect of one dimension and a minimum in respect of all other dimensions, vibrational frequency analysis at that point on the PES will disclose one imaginary frequency; by contrast if it is a maximum with respect to two dimensions on PES it will have two imaginary frequencies. By "transition state" chemists refer to first order transition states, that is to say a maximum in respect of one and only one dimension on the PES, and hence disclosing one imaginary frequency upon vibrational frequency analysis.

2.11 Natural Bond Orbitals and Wiberg Bond Indices

The underlying theory for the calculation of Wiberg Bond Indices (WBIs) was introduced by Wiberg [23], as an application of the Approximate Self Consistent Molecular Orbital Theory Complete Neglect of Overlap methods (which are in turn based upon extended Hückel theory [108] based upon the original Hückel theory [109-111]) developed by Pople and Segal [112] based on work by Pople, Santry and Segal [113]. Wiberg also cites [23] a related paper by Pople and Segal [114] further developing the theory. The results calculated using these methods do not depend on the coordinate system being employed [23]. WBIs correspond to "the sum of the squares of the bond orders between the atoms in question" [23]; Wiberg notes [23] that the bonds are populated in accordance with Mulliken's method [115-118]. Equation 2.14 shows how the bond order is calculated; p_{jk} is the bond order between orbitals j and k, p_{jj} is the charge density within orbital j.

$$\Sigma_k p_{jk}^2 = 2p_{jj} - p_{jj}^2 \qquad (2.14)$$

For further discussion about "quadratic bond orders" [119] based on Wiberg's work [23] see a paper authored by Szczepanik and Mrozek [119]. The natural bond orbital method [24] is employed for the calculation of WBIs. A WBI value of 1 corresponds to a single covalent bond, Wiberg refers to this quantification by reference to the number of covalent bonds formed by a given atom (e.g. a value of 2 would imply that the given atom forms two covalent bonds) [23], however it is possible to calculate bond order between any two given atoms using this method in Gaussian 09 [92]. The Wiberg bond index is used with orthogonal sets of orbitals [120-121]. The Wiberg bond index is not ideally suited to open shell systems, a more general method for calculating bond indices was proposed by Mayer [121], which seeks to overcome the limitations arising from the orthogonality requirement and closed shell conditions in the context of ab initio calculations. Mayer comments upon the limitations of Wiberg bond indices in the context of homonuclear molecules, while acknowledging that "other (even rather complicated) molecules" are amenable to "reasonably" good treatment "at the CNDO level", i.e. the level of theory upon which Wiberg developed [23] the method for calculating the bond order that bears his name.

2.12 Electrostatic Potential Plots

Electrostatic potential (ESP) surface graphs were generated for some of the systems under investigation. The Gaussview software package [122] was used for this purpose These plots were generated by mapping the electrostatic potential on a "cube" plot of the total density of the system, with arbitrary density of the grid (in this case the "medium" fineness was chosen in Gaussview [122]). More specifically, the ESP was mapped onto an isodensity surface with arbitrary electron density; in this study the 0.0004 e^-/au^{-3} iso-surface was chosen. The ESP is then graphed onto that surface with a colour scale employed to disclose the charge at any given point on the graph. The range of the colour scale (and hence the sensitivity of the graph to small variations in charge) can be assigned arbitrarily; in this study the range was set from -6.93×10^{-3} Eh/e^- to 6.93×10^{-3} Eh/e^-.

2.13 Statistical Analysis of the Cambridge Structural Database

All technical details regarding the role of each piece of software in this statistical analysis and the representation of the results thereof came were disclosed by a private communication with Rachael Skyner [123]. However, the analysis itself was performed by the author of this thesis.

The Cambridge Structural Database (CSD) [80] is a repository for crystallographic data that has been elucidated in respect of small molecules [80]. The Conquest software package [124] was used to perform a substructure search of the CSD. The Mercury software package [124-127] was employed to analyse the structures from the CSD. Within MATLAB [128] a modified version of the dscatter

script [129] was employed to plot the two histograms against each other. The edit_script.m script [130] was also employed for labelling and formatting the plots.

3 Competition between halogen bonding and hydrogen bonding in microsolvated halobenzene

3.1 Abstract

The microsolvated 1-methyl-5-halouracil molecular systems (XU•••w) are of biological interest as a model for halogenation of uracil in the aqueous *in vivo* environment. Those systems derived from uracil are the subject of the next chapter. The present chapter is concerned with a simpler, but analogous system in which there are fewer opportunities for the water molecule to become involved in secondary interactions, in particular, the microsolvated halobenzene molecular system (XPh•••w) does not incorporate an ortho aromatic ketone (C4=O4 in the XU•••w case). The relative simplicity of the present chemical environment allows a purer assessment to be made upon the role of halogen bonding between the halogen atom (X) and the water oxygen atom (Ow) simpliciter. Furthermore, a comparison of the results for XPh•••w and XU•••w (in the next chapter) allows the influence of the secondary interactions that are present in XU•••w to be probed. All geometry optimisations and energy calculations presented in this chapter were performed using the M06-2X density functional with the 6-31+G* (for X=F, Cl or Br) or aug-cc-pVDZ-PP (for X=I or At) basis set. Minima were confirmed using harmonic frequency calculation. The counterpoise procedure was performed to counteract the effect of BSSE. The results show halogen bonds form in XPh•••w systems wherein the halogen was no lighter than bromine and that there is a trend towards stronger halogen bonding as the halogen group is descended, as assessed by the interaction energy and X•••Ow internuclear separation. For all XPh•••w systems an X•••Hw hydrogen bond was found with the strength of interaction showing a trend towards weaker interactions as the halogen group was descended. For all systems except X=At, the X•••Hw interaction is stronger than the X•••Ow interaction.

3.2 Introduction

Halogen bonding involving halobenzene has already been subject to extensive study [38, 131-133]. For example Adasme-Carreno *et al.* authored a paper [38] wherein halogenated benzene was halogen bonded to N-methylacetamide, and substituent effects were investigated. In their study the halogen substituent engaged in halogen bonding could be chlorine, bromine or iodine. Fluorine was also added as a substituent but not as the atom through which the halobenzene molecule would interact with the N-methylacetamide molecule. The authors identifed a strong correlation (R^2 > 0.9) between interaction energies and a set of attributes that reflect electrostatic properties, specifically, natural bond orbital charges, molecular electrostatic plots and electron density. Based on those results the authors concluded that the effect of the substituent upon the halogen bond arose due to their effect upon the electrostatics of the system rather than resonance effects.

For present purposes, interest in the halobenzene:H_2O (XPh•••w) complex arises from it being a simpler analogue of XU•••w, which we investigate in the next chapter. This smaller system, containing a smaller number of basis functions (with the same basis sets) affords much faster computation. As one of the principal foci of the investigation into the XU•••w system is the probing of competition between X5•••O halogen bonding and O4•••Hw1 hydrogen bonding (see figure 4.2), comparison with XPh•••w may provide useful insights into the effect of the competing O4•••Hw1 hydrogen bond (Hbond) interaction in XU•••w, which is absent in XPh•••w. XPh•••w therefore serves as a model for XU•••w lacking the opportunity for O4•••Hw1 hydrogen bonding. However, hydrogen bonding type interactions can still occur in the XPh•••w systems, with the negatively charged ring on the halogen atom acting as the electron donor, i.e. X5•••Hw, where Hw is a water hydrogen atom. For ease of reference, the XU-w system is shown below with numbering (figure 3.1). However, it should be noted that this is the system discussed in chapter 4, this chapter (chapter 3) is concerned with the XPh-w system.

In addition to serving as a useful comparator for the more elaborate XU•••w system, this study aims to cast further light upon the halogen bonding attributes of halobenzene where the nucleophile is a lone water molecule, including for the under-studied halogen bonding properties of astatine.

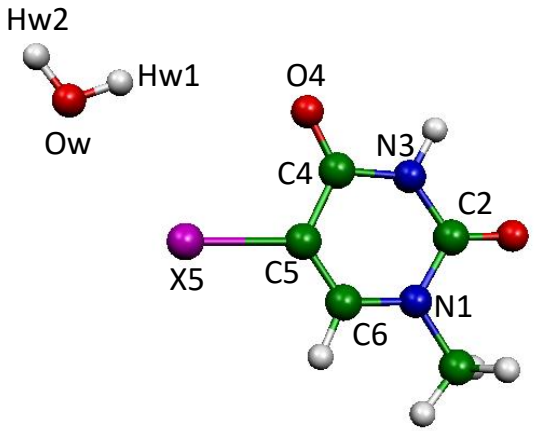


Figure 3.1: The 1-methyl-5-halouracil plus water system. X=F, Cl, Br, I or At.

3.3 Methods

XPh•••w structures were created using Gaussview 4.1 [122]. All calculations were performed using Gaussian 09 [92]. Each XPh contained one halogen atom (X) substituting for a hydrogen atom on the benzene molecule; X= F, Cl, Br, I or At. The C-X•••OH_2 angle was varied in step sizes of 5° with no additional geometric constraints to find minima on the potential energy surface (PES). Two minima of particular interest were sought: one in which a potential C-X•••OH_2 halogen bond could be favoured, and one in which a potential C-X•••H-OH hydrogen bond could be favoured. Linear C-X•••O angles were considered to favour the former and perpendicular C-X•••H angles were considered to favour the latter. When one of the potential minima of interest described above was identified from the constrained optimisation calculations, based upon the interaction energy between the XPh molecule and the water molecule, a full geometry optimisation was performed on the structure corresponding to the minimum on the restricted PES. For all calculations the M06-2X functional [70] was used in conjunction with either the 6-31+G* basis set (X= F, Cl, or Br) or the aug-cc-pVDZ-PP basis with corresponding relativistic effective core potential [64] (X= I or At). Where the correlation consistent basis set was employed, all of the d-orbitals (on all atoms) were treated spherically harmonically, whereas they were modelled in Cartesian form where only the Pople basis set was employed. However, for calculations performed on the isolated water molecule for calculating interaction energies between the water molecule and the iodine or astatine containing halobenzene, the d-orbitals were modelled as Cartesian functions. In the latter case the "5D" keyword was used to impose spherical harmonic treatment; in all other cases this was the default employed by Gaussian 09 [92]. All full geometry optimisations invoked the counterpoise correction procedure (CP) proposed by Boys and Bernardi [74] to eliminate basis set superposition error (BSSE). In the constrained optimisations, CP correction was not invoked during the optimisation but CP-corrected single-point calculations were performed on the optimised geometries. In the full geometry optimisations Gaussian's tight convergence criteria were specified, whereas for constrained geometry optimisations Gaussian's default convergence criteria were used. The Wiberg Bond Indices were calculated by specifying a natural bond orbital population analysis by including the keyword "pop=nboread", and further providing a stand-alone command at the end of the input file "$NBO BNDIDX $END". The Wiberg Bond Index between each pair of atoms was then computed by Gaussian 09 [92] and printed in the output file.

3.4 Results and discussion

For graphical representations of each of the elucidated minima see figure 3.2. For all types of halogen atom, a minimum on the PES that is stabilised by an X•••H Hbond was located (Figure 3.2a). However, a minimum stabilised by a halogen bond (Xbond) could only be found where X= Br, I or At (Figure 3.2b). For fluorobenzene, iodobenzene and astatobenzene, a second hydrogen-bonded structure was found (hereinafter referred to as Hbond 2), wherein the water hydrogen atoms are placed either side of the halobenzene molecular plane (Figure 3.2c). This structure has not been studied in much detail, but its stability may derive from a combination of C-H•••Ow hydrogen bonding and Ow-Hw•••X interaction between the water hydrogen atoms and the halogen atom (Ow and Hw indicate water oxygen and water hydrogen respectively). One possible explanation for this minimum only being found at the extremes of the halogen group could be as follows. Owing to the very strong electron withdrawing effect of the fluorine atom, a relatively strong C-H•••Ow hydrogen bond can be achieved. In cases of X = I or At, the inadequacy of the C-H•••O hydrogen bond may be sufficiently offset by the greater strength of the Hw•••X interaction with the negatively charged ring on the halogen atom. *Prima facie*, the geometries do not support this hypothesis as the halogen appears better aligned with the C-H hydrogen atom for X = I and At than for X = F, but that might be explained by the greater VdW radii of the heavier halogens. *Ex hypothesi*, for X = Cl or Br, aggregation of these two interactions must not provide an adequate remedy for the relative weakness of both of these types of interactions. The Hbond 2 energetic and geometric results are included in tables 3.1-3.3, but are not discussed in further detail as they were not the focus of this study. The interaction energy of the Xbond structures becomes more negative (attractive) going down the halogen group from bromine to astatine. The Hbonds do not show a clear trend. Fluorine clearly forms the strongest Hbond while all of the other halogens form Hbonds that are within 0.5 kJ mol^{-1} of each other. Drawing firm conclusions from the variation within that range would be folly, given the limits of the accuracy of the M06-2X functional. For quantitative results see table 3.1.

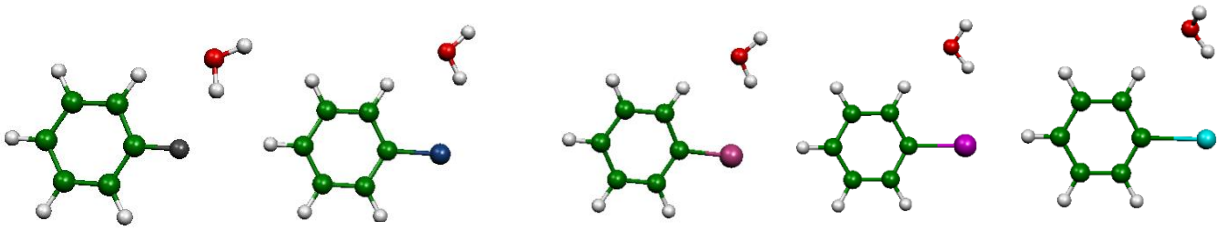


Figure 3.2a: Hydrogen bonded minima structures. From left to right: X= F, Cl, Br, I, and At.

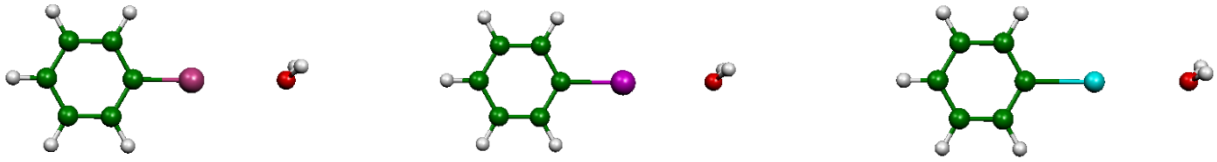


Figure 3.2b: Halogen bonded minima structures. From left to right: X= Br, I, and At.

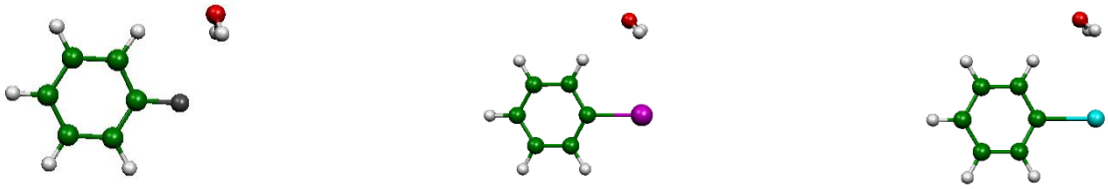


Figure 3.2c: Hydrogen bonded minima structures (Hbond 2) for X = F, I and At.

Table 3.1: Interaction energies for the Hbond and Xbond geometries of XPh•••w molecular systems. All energies are in kJ mol^{-1}.

Structure	X=F	X=Cl	X=Br	X=I	X=At
Xbond	N/A	N/A	-7.3	-13.3	-18.6
Hbond	-16.5	-14.2	-14.5	-14.6	-14.1
Hbond 2	-12.7	N/A	N/A	-14.2	-13.9

Table 3.1 shows a clear trend in the relative strength of X•••Ow halogen bonding with increasing atomic number of the halogen atom. Halogens lighter than bromine do not form halogen bonds in this system. For X=Br the Xbond structure has an interaction energy approximately half of that of the Hbond structure (difference in interaction energy equals 7.2 kJ mol^{-1}); for X=I the Hbond structure is also more stable than the Xbond structure, but the difference is much diminished (1.3 kJ mol^{-1}), due to an increase in the magnitude of the X•••Ow interaction. This trend

continues for X=At in which the Xbond structure is more strongly stabilised than the Hbond structure by 4.5 kJ mol^{-1}. Indeed, the AtPh•••w Xbond structure is more strongly stabilised upon complexation than any of the Hbond structures. In contrast, even for X=I, the interaction energy for the Xbond structure is less negative than for any of the Hbond structures.

Table 3.2 shows the bond angle for each of the minima that were found for the XPh•••w systems. For the Hbond this is the C-X•••Hw angle (where Hw is the hydrogen atom on the water molecule that is involved in hydrogen bonding); for the Xbond this is the C-X•••Ow angle.

Table 3.2: C-X•••O angle for the Hbond and Xbond geometries of XPh•••w molecular systems. All angles are in degrees.

Structure	X=F	X=Cl	X=Br	X=I	X=At
Xbond	N/A	N/A	179	179	179
Hbond	117	101	97	92	89
Hbond 2	108	N/A	N/A	83	81

Table 3.2 shows that the Xbonds that are formed are very close to linear (179°) in all cases. This result is to be expected as (i) the ideal angle for a halogen bond is linear due to the position of the sigma hole, which is situated trans with respect to the C-X bond [134], and (ii) there are no significant secondary interactions. The halogen bonds are not perfectly linear; a possible explanation for this very slight departure from ideality may be ascribable to a very weak interaction between the electron deficient water hydrogen atoms and the electron rich pi system on the benzene ring, which may be strengthened by a slightly non-linear C-X•••Ow angle. For the Hbond the C-X•••Hw angle shows a clear trend in decreasing as the halogen group is descended. This can be explained as the electron density on the halogen atom is concentrated in a negatively charged ring that is approximately perpendicular to the C-X bond axis. As the size of the sigma hole increases going down the group this relocation of electron density towards the perpendicular position would become more pronounced. The further that the electron density must be displaced from the linear region, the more it will concentrate in the perpendicular region, moving the location of maximum density to smaller angles. In the case of AtPh•••w, an acute angle is formed, suggesting that the peak in electron density, in this extreme case, is pushed back even slightly beyond perpendicular in order to accommodate the large sigma hole. However, a greater factor might be that the C-X bond distance increases down the halogen group, especially if there could be interaction between electron deficient Hw atoms and the pi system on the aromatic ring favouring shorter distances with respect to the ring. Part of the reason for the greater C-X bond distance will be due to donation into the σ* antibonding orbital, which corresponds to the σ-hole, but the principal contributor to increasing C-X bond distance is likely to be increasing Van der Waals (VdW) radius of the halogen atom.

Table 3.3 shows the bond distance, whether Hbond or Xbond for each of the minima found in the XPh•••w systems. For Hbond this is the Hw•••X distance, for Xbond it is the Ow•••X distance.

Table 3.3: Internuclear separation for the Hbond and Xbond geometries of XPh•••w molecular systems. All distances are in Å.

Structure	X=F	X=Cl	X=Br	X=I	X=At
Xbond	N/A	N/A	3.06	3.08	3.01
Hbond	2.05	2.59	2.70	2.89	2.96
Hbond 2	2.60	N/A	N/A	3.33	3.35

Table 3.4: Ratio of internuclear distance divided by sum of VdW radii for the halogen bonds. All VdW radii come from the CRC Handbook of Chemistry and Physics (90th edition) [135] except for astatine which comes from the website of the Royal Society of Chemistry [136].

Structure	X=F	X=Cl	X=Br	X=I	X=At
Xbond	N/A	N/A	0.91	0.88	0.85

Table 3.3 demonstrates a trend towards stronger halogen bonds as the halogen group is descended, with a decrease in internuclear separation from iodine to astatine and only a very small (0.02 Å) increase from bromine to iodine. Table 3.4 makes the point more clearly, showing a consistent trend towards smaller VdW ratios from bromine to astatine. As smaller internuclear separations, and in particular smaller VdW ratios, are associated with stronger interatomic interactions, this trend towards shorter halogen bonds going down the halogen group further supports the energetic findings presented in table 3.1 that halogen bonds become stronger as the identity of the halogen atom is varied going down the halogen group from bromine to astatine.

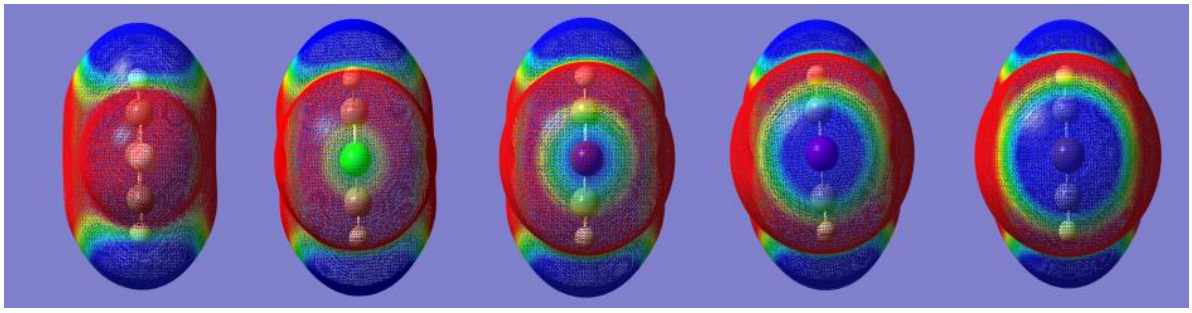
As shown in table 3.5, the Wiberg bond index for the X•••Ow halogen bond increases with increasing size of halogen atom. From bromine to iodine it increases by 0.0080 and from iodine to astatine by 0.0070. The absolute values are low in comparison with other types of interaction, for example the water-water hydrogen bond found in the AtU•••2w system in the next chapter (Microsolvated 1-methyl-5-halouracil) of this thesis. From these values and their trend, it can be seen that the oxygen-halogen interaction (i) has slight covalent character but is not predominantly covalent and (ii) the degree of covalency increases as the halogen group is descended. Furthermore, the trend that these Wiberg bond indices have greater values for the heavier halogens is consistent with the observations made above based upon energetic and

geometric results, i.e. that the halogen bond strength increases going down the halogen group from X = Br to X = At.

Table 3.5: Wiberg bond index values for the X•••Ow halogen bond

Halogen	Wiberg bond index
Br	0.0098
I	0.0178
At	0.0248

Figure 3.3 shows molecular electrostatic potential plots for each of the halobenzene molecules that have been investigated. Descriptions of regions shown in figure 3.3 as being "positively charged" (including equivalent terminology thereto) relate to the σ-hole region not being fully shielded from the positive charge of the protons in the nucleus of the halogen atom, and therefore wherein a nucleophile would be subject to an electrostatic force in the direction towards the halogen nucleus. From figure 3.3 it can be seen why FPh and ClPh do not form halogen bonds with a water molecule, whereas BrPh, IPh and AtPh do form halogen bonds with water. Figure 3.3 also provides a further evidence as to the physical reason why halogen bond strength increases from halobenzene to haloastatine. In the case of FPh, no sigma hole can be observed, and indeed no evidence of the anisotropic distribution of the electron density can be easily observed. ClPh does show an anisotropic distribution of the electron density but even the region directly along the extension of the C-X bond is electrostatically neutral, not positively charged. This lack of positive charge explains the lack of any halogen bond with the water molecule, which would require the ability of the lone pair of electrons on the water oxygen atom to interact with a positively charged region on the chlorine atom. The halogens heavier than chlorine do each present a positively charged sigma hole, and hence do form a halogen bond with a water molecule. Furthermore, the magnitude of the sigma hole (both in terms of its spatial extent and the magnitude of the positive charge of the region at a given halogen bond angle) increases going down the halogen group from bromine to astatine, and this is consistent with the energetic and geometric results, which show an increase in halogen bond strength from going down the halogen group from BrPh•••w to AtPh•••w.



FPh ClPh BrPh IPh AtPh

Figure 3.3: Plots of electrostatic potential for isolated XPh molecules. Density surface used for mapping is 0.0004 e-/au^3. The depicted electrostatic potential range is -6.93x10^{-3} Eh (red) to 6.93x10^{-3} Eh (blue).

3.5 Conclusions

All of the microsolvated halobenzene molecules were found to form X•••H hydrogen bonds with the water molecule. No trend in the interaction energies for geometries with this interaction could be elucidated, save for the observation that fluorine forms the strongest interaction. X•••Ow halogen bonding was found to occur in cases where the halogen atom is heavier than chlorine, with the magnitude of the interaction energy increasing going down the halogen group. There is also a trend going down the halogen group towards shorter X•••Ow internuclear distances, providing further support for this trend towards stronger halogen bonding going down the halogen group. The trend towards smaller C-X•••Hw angles is also consistent therewith on account of the need to accommodate larger sigma holes (and hence electron density more strongly restricted to perpendicular angles) on the heavier halogen atoms. However, the greatest impact on C-X•••Hw angles is probably the C-X bond length, which increases down the halogen group, partly due to greater donation into the σ* anti-bonding orbital, which corresponds to the σ-hole, and partly due to increasing VdW radii. The halogen bond angles themselves are consistently very close to linear at 179°, and that is consistent with the simple chemical environment wherein there is a lack of significant competing secondary interactions. Astatobenzene is unique in forming a halogen bond that is stronger than all of the X•••H hydrogen bonds. A second hydrogen bonded minimum (Hbond 2) was also encountered in the cases of X = F, I and At.

4 Competition between halogen bonding and hydrogen bonding in microsolvated 1-methyl-5-halouracil

4.1 Abstract

An investigation into halogen bonding and hydrogen bonding interactions between a molecule of 1-methy-5-halouracil (XU) (X = F, Cl, Br, I, or At) and a water molecule (XU•••w) or two water molecules (XU•••2w), primarily in the region between the C5-X5 bond and the C4=O4 bond was performed. The numbering convention hereinafter employed is set forth in figure 4.1 (below), figure 4.1 is also hereinafter reproduced as figure 4.2. For all calculations the M06-2X functional was used in conjunction with either the 6-31+G* basis set (for all elements lighter than iodine) or aug-cc-pVDZ-PP (for iodine and astatine). XU•••w halogen bonds were found to form between X5 and the water oxygen atom (Ow) where X is Br, I or At; whereas that interaction was found for X = Cl, Br, I or At in the case of XU•••2w. In the case of XU•••2w the structure with a halogen bond also included water-water and water•••O4 hydrogen bonds. For XU•••w all of the halogen bonded minima were found to be connected to a hydrogen bonded minimum via a transition state, and the geometry and barrier height of each transition state was computed. All minima and transition states were confirmed by harmonic vibrational frequency analysis. In all cases the strength of the halogen bond, the barrier height and linearity of the C5-X5•••Ow angle increases as the halogen group is descended. All of these observations are attributed to the greater polarisability (and hence stronger sigma hole effect) as the halogen becomes heavier. The lack of a halogen bonded ClU•••w despite its presence in the case of ClU•••2w is ascribed to the stronger competing water•••O4 hydrogen bond. The presence of halogen bonds with a C5-X5•••Ow angle in the range 150°-160° demonstrates the flexibility of the halogen bond in departing from its ideal angle (linear) when there are secondary factors bearing on the geometry as can be found in the relatively complex chemical environment of the XU•••2w systems.

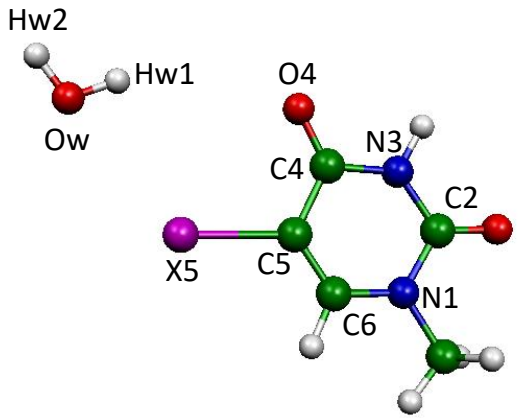


Figure 4.1: The 1-methyl-5-halouracil plus water system. X=F, Cl, Br, I or At.

4.2 Introduction to the microsolvated 1-methyl-5-halouracil systems

This chapter presents a theoretical (DFT) investigation into competition between halogen bonding and hydrogen bonding in a complex consisting of 1-methyl-5-halouracil (XU) (with the halogen X being F, Cl, Br, I or At) and a water molecule. In particular, interaction energies, transition barrier heights and the relative orientation of the XU molecule and the water molecule are the primary focus of this study.

The structure of microsolvated 1-methyl-5-halouracil (XU•••w) is shown in figure 4.2. The symbol X can be fluorine (F), chlorine (Cl), bromine (Br), iodine (I) or astatine (At). Some calculations have also been performed on the non-halogenated microsolvated 1-methyluracil (X=H) system.

It is known that 5-halouracil, with the halogen lighter than astatine, can be utilized to render DNA more sensitive with respect to ionizing radiation[137-139]. 5-Halouracils are known to have mutagenic properties, and the Van Mourik group has been investigating the mechanism for this mutagenesis, including an investigation into the significance of hydration[50, 140-142]. A synthesis of 5-astatouracil has been performed (At-211 isotope) [143].

The XU•••w system is particularly suitable for an investigation into competition between halogen bonding and hydrogen bonding because (i) the relative positions of the X5 halogen and the O4 oxygen atoms do not preclude a water molecule from interacting with both of them simultaneously, (ii) there is a region of space where the hereinbefore specified halogen bond (Xbond) and hydrogen bond (Hbond) interactions may compete without additional strong competing factors and (iii) the intermolecular nature of the XU•••w systems allow the position and orientation of the water molecule to be determined by the interatomic forces in the absence of major steric constraints of the type typically associated with intramolecular interactions. The of the XU•••w systems' suitability as model systems for investigating halogen bond versus hydrogen bond competition make them a prime molecular system of interest for examination in the present study.

4.3 Methods

Molecular systems comprising a 1-methyl-5-halouracil molecule plus one water molecule, hereinafter referred to as "XU•••w" systems, were constructed where the "X" may be substituted with any one of the atomic symbols "F", "Cl", "Br", "I", or "At". All systems were built using Gaussview 4.1 [122] and all calculations were performed using Gaussian 09 [92]. For each system, minima corresponding to structures in which halogen bonding was dominant were sought, as was, in each case, a system in which hydrogen bonding dominated. Where a minimum corresponding to either of these bonding type interactions was found, it was used as the starting

point for a scan calculation in which the C5-X5•••Ow angle was varied in step sizes of 5°, while all other dimensions were allowed to freely optimise (see figure 4.2 for atom labelling), so as to provide an energy profile for each system (two profiles where two minima were found). After obtaining the energy profiles, in cases where both a hydrogen bond and a halogen bond were found (those incorporating bromine, iodine or astatine) the transition state structure and energy were then calculated, using the Synchronous Transit-Guided Quasi-Newton (STQN) method [106-107] (invoked by the QST2 and QST3 keywords), or by a simple transition state optimisation (using the TS keyword). In the case of all complete optimisations (not those as part of a scan) the minimum or transition state was confirmed using frequency calculations. Gaussian's default convergence criteria were used for optimisations that were part of a scan, while the "tight" convergence criteria were used for complete optimisations. The ultrafine integration grid was used for most of the calculations, but in two cases (ClU•••w Hbond1 and IU•••w TS, see below) the ultrafine grid yielded an unexpected number of imaginary harmonic frequencies of vibration (one and two respectively). The anticipated number of imaginary frequencies (zero and one respectively) were obtained by using the superfine integration grid; (note that "verytight" convergence criteria failed to produce the expected results). The ultrafine and superfine grids produced the same optimised geometries, both with respect to the systems that raised concern and with respect to a sample of the other stationary points that were found by this study. All calculations were performed using the M06-2X density functional [70]. For all systems incorporating fluorine, chlorine or bromine, the 6-31+G* Pople basis set was employed, while for those systems that incorporated iodine or astatine, the aug-cc-pVDZ-PP basis set, which includes small-core energy-consistent relativistic pseudopotentials (PP) [64], was employed due to the need to treat relativistic effects. All full optimisations were performed using the counterpoise correction procedure [74]. The scan calculations were performed without that correction but counterpoise-corrected single point energies were calculated at the partially optimised geometries that were calculated by the scan calculations. Structures were visualised using Gaussview 4.1 [122] and Molden [144].

Using the same methods, calculations were performed on systems that contained a 1-methyl-5-halouracil molecule and two water molecules (hereinafter referred to as "XU•••2w"), for all types of halogen atom (fluorine to astatine inclusive). The aim was to determine whether it was possible to establish halogen bonds in these systems, or at least have some halogen-oxygen stabilisation, while also forming a hydrogen bond between the water molecules and between one of the water molecules and the O4 oxygen atom. The initial starting geometries were chosen to favour these series of interactions.

4.4 Results and discussion

Figure 4.2 shows the atom numbering system that has been used in this Chapter.

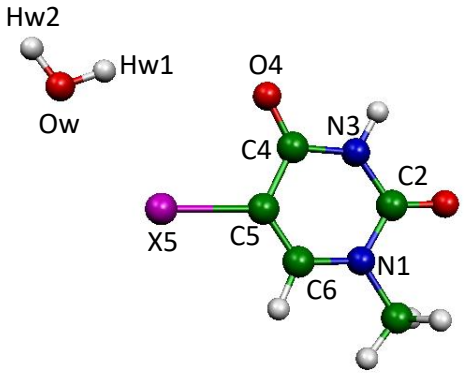


Figure 4.2: The 1-methyl-5-halouracil plus water system. X=F, Cl, Br, I or At.

All XU•••w systems exhibit three minima in the area of interest. These consist of a minimum displaying a hydrogen bond (Hw1•••O4) further stabilised by an X5•••Hw1 interaction (hereinafter referred to as "Hbond1"); a structure with a Hw2•••O4 hydrogen bond and additionally an X5•••Hw1 interaction (hereinafter referred to as "Hbond2"); and a structure that is stabilised by both an X5•••Hw1 interaction and an Ow•••H6 (weak) hydrogen bond (hereinafter referred to as "Hbond3"). Additionally, where X = Br, I or At, a X5•••Ow halogen-bonded minimum was found (referred to as "Xbond"). Furthermore, for all systems where an Xbond minimum was found, a transition state connecting that minimum with Hbond1 was elucidated, referred to as "TS". With the sole exception of Hbond2, in all stationary points found the water molecule is at least approximately coplanar with the ring of the 1-methyl-5-halouracil molecule. While the Xbond, Hbond1 and TS structures have been the primary focus of this study, key energetic and geometric data relating to all stationary points found are tabulated in tables 4.1-4.12, whereas their structures are shown in figures 4.3 and 4.7.

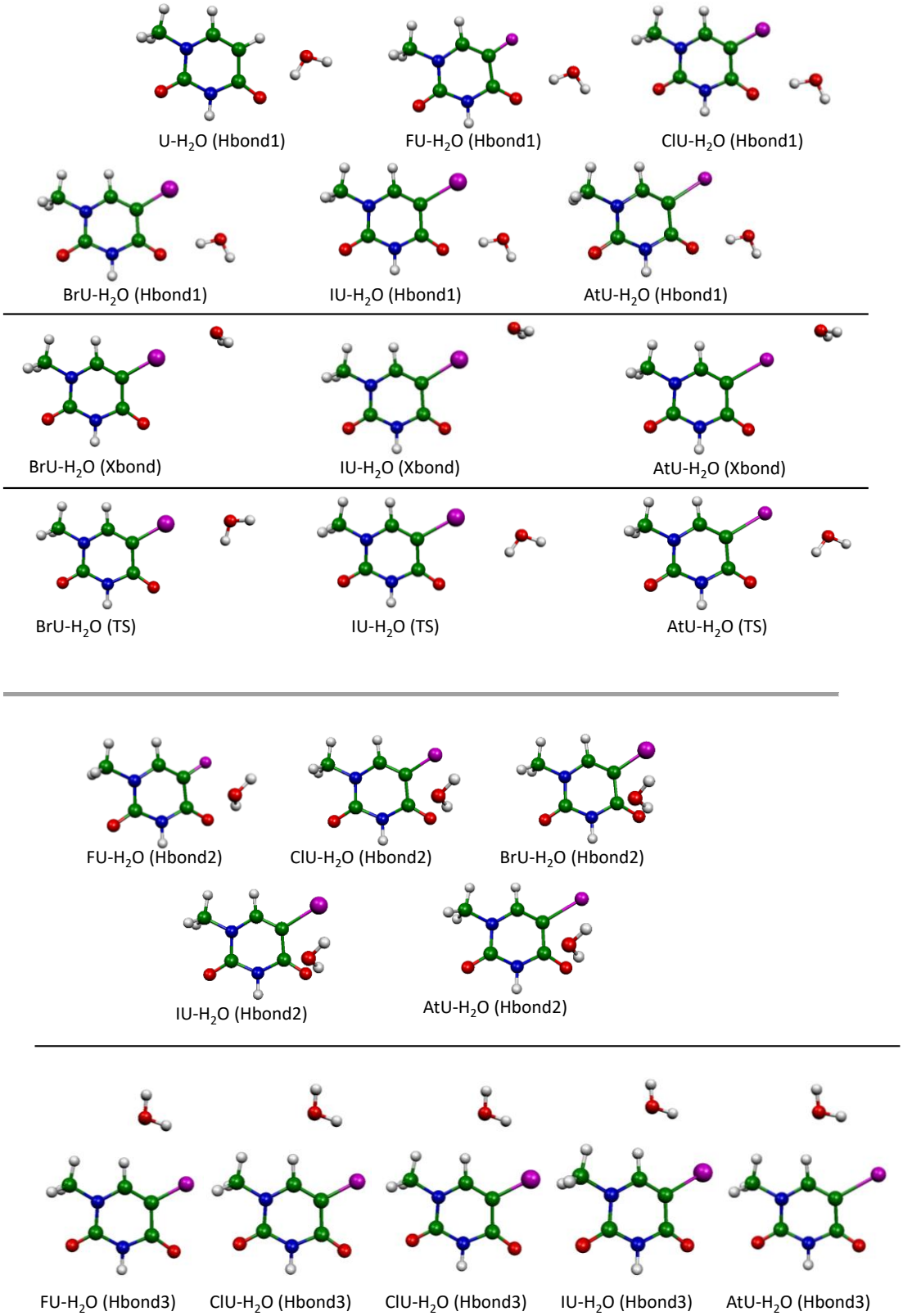


Figure 4.3: The geometry of each of the stationary points referred to in tables 4.1-4.9.

Table 4.1: Interaction energies for each stationary point. All energies are in kJ mol^{-1}.

Stationary point	X=H	X=F	X=Cl	X=Br	X=I	X=At
Xbond		N/A	N/A	-12.2	-18.6	-25.5
Hbond1	-35.8	-22.6	-24.4	-25.2	-254	-26.5
Hbond2		-24.7	-24.5	-24.8	-24.6	-24.2
Hbond3		-25.2	-23.8	-24.7	-25.2	-24.5
TS		N/A	N/A	-11.7	-13.8	-15.7

As can be seen in table 4.1, for those systems that form halogen bonds (X=Br, I or At), halogen bond strength, defined in terms of intermolecular interaction energy, increases going down the halogen group, with astatine forming a halogen bond that is comparable in strength to Hbond1, with an energetic difference of only 1.0 kJ mol^{-1}.

There is also a trend towards stronger Hbond1 going down the halogen group, albeit with a very small difference between bromine and iodine. By far the strongest Hbond1 is for the non-halogenated system, but among the halogens, astatine shows the strongest interaction of type (-25.5 kJ mol^{-1}). This could potentially be a reflection of the the decreasing electronegatively of the substituent leading to greater electron density on the O4 atom. Additionally, discussed below, there also appears to be a contribution to the interaction energy at the Hbond1 geometry from an interaction between the water hydrogen atoms and the halogen atom, this interaction would also be expected increase down the the halogen group due to increasing concentration of the electron density in the negatively charged ring around the halogen atom.

There is also a trend towards more negative interaction energies for the transition state that connects Xbond and Hbond1. This result is perhaps not surprising as it is consistent with the trend for both of those minima.

Neither Hbond2 nor Hbond3 show any trend going down the halogen group. The reason for this null result has not been subjected to detailed exploration but it could be that varying the substituent results in multiple opposing effects upon the trend in interaction energy. For example, in Hbond3, while the greater electron withdrawing effect of the ligher halogen elements would be expected to make the C6-H hydrogen atom more electron deficient, and hence interact more strongly with Ow, the weakening of the σ-hole effect going up the halogen group, and hence the weakening of the corresponding negatively charged ring effect, would probably result in a weakening of the interaction between X5 and the water hydrogen atom.

For all systems that form both an X5•••Ow halogen bond (Xbond) and an O4•••Hw1 hydrogen bond (Hbond1), there is evidence that both minima are influenced by both the O4 atom and the X5 atom. In the case of the X5•••Ow halogen bond, the Hw1 and Hw2 atoms both point towards

the O4 atom, indicating a hydrogen bonding type stabilisation (albeit without actual hydrogen bonding). When an attempt was made to optimise a BrU•••w system with the Hw1 and Hw2 atoms pointing away from the O4 atom, no corresponding minimum could be found. Instead the structure optimised to the halogen bonded geometry, with the water protons pointing towards the O4 atom.

Table 4.2: C5-X5•••Ow angles for all stationary points. All angles are in degrees.

Stationary point	X=H	X=F	X=Cl	X=Br	X=I	X=At
Xbond		N/A	N/A	173	178	178
Hbond1	131	117	107	104	98	97
Hbond2		81	68	69	62	61
Hbond3		101	86	83	78	76
TS		N/A	N/A	153	136	129

This involvement of the O4 atom may also explain the departure from strict linearity of the halogen bond minimum structure (see table 4.2), with the structure being a compromise between the dominant halogen bond interaction and a secondary hydrogen bonding type interaction. Furthermore, there is a trend towards greater linearity from X=Br to X=I, possibly reflecting the greater dominance of the halogen atom. The C5-X5•••Ow angle is the same for X=At as for X=I. Likewise, the structure of Hbond1 appears to be consistent with a secondary stabilisation from the X5 atom, not from the sigma hole but from the negatively charged ring of high electron density around the sigma hole interacting with electron deficient Hw1. The water molecule appears to be positioned so as to facilitate this interaction, possibly at the expense of an orientation that would allow a favourable Hw2•••O4 interaction.

Table 4.3: X5-Ow distances for all stationary points. All distances in Ångstroms.

Stationary point	X=H	X=F	X=Cl	X=Br	X=I	X=At
Xbond		N/A	N/A	3.02	2.96	2.94
Hbond1	2.28	3.27	3.34	3.39	3.51	3.53
Hbond2		3.06	3.45	3.56	3.73	3.81
Hbond3		2.85	3.38	3.48	3.67	3.74
TS		N/A	N/A	3.07	3.23	3.26

Table 4.3 reflects the expected increase in internuclear distance as the size of the halogen atom, i.e. its Van der Waals (VdW) radius, increases down the group for the different Hbond structures. However, the converse trend is observed for the Xbond stationary point; here the internuclear separation between X5 and Ow decreases, presumably due to the increasing strength of the halogen bond outweighing the increase in VdW radius.

Table 4.4: X5•••Hw1 distances for all stationary points. All distances in Ångstroms.

Stationary point	X=H	X=F	X=Cl	X=Br	X=I	X=At
Xbond		N/A	N/A	3.35	3.44	3.46
Hbond1	2.32	2.72	2.88	2.96	3.14	3.17
Hbond2		3.13	3.43	3.48	3.61	3.73
Hbond3		2.13	2.69	2.75	2.94	3.00
TS		N/A	N/A	2.98	3.05	3.09

Tables 4.4 and 4.5 show that the C5•••Hw1 distances increase and the C5-X5•••Hw1 angles decrease with increasing size of the halogen. The data in tables 4.4 and 4.5 are probably predominantly a reflection of the increasing length of the C5-X5 covalent bond distance as the atomic radius of the halogen atom increases, combined with an energetic preference for the Hw1 atom to remain close to the O4 atom. For the X5•••Hw1 distances (table 4.4) this trend even holds for the Xbond geometry, lending support to the proposition that the Hw1•••O4 interaction is a non-trivial factor in determining the geometry of the molecular system even for the structure wherein halogen bonding is dominant. For the C5-X5•••Hw1 angles (table 4.5) the trend towards larger angles for heavier halogens in the case of the Xbond structures closely coincides with the trend for the C5-X5•••Ow angles becoming more linear (table 4.2); this correlation is almost inevitable because Hw1 is covalently bound to Ow and the Ow•••Hw1 bond distance is short compared with the distance between the C5 and Ow atoms.

Table 4.5: C5-X5•••Hw1 angles for all stationary points. All angles in degrees.

Stationary point	X=H	X=F	X=Cl	X=Br	X=I	X=At
Xbond		N/A	N/A	159	164	164
Hbond1	106	102	91	88	83	81
Hbond2		97	83	79	76	74
Hbond3		116	99	94	89	86
TS		N/A	N/A	134	118	112

As discussed by Wang *et al.* [11], the halogen bonding interaction derives from overlap between an occupied orbital (in this case the lone pair of electrons on the Ow atom) and a virtual orbital (in this case the σ* orbital corresponding to the C5-X5 covalent bond). Given the antibonding character of the σ* orbital with respect to the C5-X5 bond, it may be the case that in addition to increasing atomic radius, the C5-X5 bond distance may also be elongated in the Xbond structures due to the antibonding effect of electron density donation from the lone pair on Ow into the σ* antibonding orbital, in a manner described by Wang and Hobza [145]. However, the authors also

state that electrostatic interactions have the effect of shortening the C-X bond[145]. These effects would be expected to increase as halogen bonding strength increases down the halogen group. However, this possible dimension to explaining the geometric data would need to be the subject of specific investigation, especially as C-X bond length has been found to be capable of increasing or decreasing upon halogen bond formation [145].

Table 4.6: Transition barrier heights (defined as interaction energy of TS minus interaction energy of Xbond). All energies in kJ mol^{-1}.

X=Br	X=I	X=At
0.5	4.9	9.8

As shown in table 4.6 (and implied in table 4.1), for BrU•••w, only a low energy barrier, 0.5 kJ mol^{-1}, impedes the Xbond minimum structure from converting to the more energetically favourable Hbond1 geometry. Hence, the halogen bond may only be metastable, even at absolute zero. The iodinated and astatinated molecules both form stable halogen bonds to the water molecule, with barriers of 4.9 kJ mol^{-1} and 9.8 kJ mol^{-1} respectively. Inclusion of zero point energies may yield transition barrier heights that are lower than those set forth in table 4.6, as (positive) zero point energies are likely be greater in magnitude for the minima than for the transition states.

Table 4.7: X5•••Ow distances for the XU•••w systems, sums of VdW radii and the ratio between them (VdW ratio). All distances in Ångstroms. All VdW radii come from CRC Handbook of Chemistry and Physics (90^{th} edition) [135] except for astatine which comes from the website of the Royal Society of Chemistry [136].

System	**X5•••Ow distance**	**Sum of VdW radii**	**VdW ratio**
X=Br	3.02	3.37	0.90
X=I	2.96	3.50	0.85
X=At	2.94	3.54	0.83

The VdW ratios in table 4.7 are all below 1.00, indicating a favourable interaction between X5 and Ow, table 4.7 also shows a trend towards smaller VdW ratios from X=Br to X=At, giving further strength to the proposition that the halogen bond strength increases as the halogen group is descended.

Figure 4.4 shows the C5-X5•••Ow angle, labelled θ. Figure 4.4 shows the interaction energy for each of the XU•••w systems, as a function of the C5-X5•••Ow angle. For each system, the

geometry for the hydrogen bonding minimum and where applicable also the geometries of the halogen bonding minimum and the transition state which connects the two minima are also shown. Where there are two graphs on the same set of axes, one is for a scan that was started at the hydrogen bonding minimum and the other for a scan that was started from the halogen bonding minimum (referred to as Hbond and Xbond respectively) (or in the case of X=F from each of the two hydrogen bonds) as identifiable from the legend.

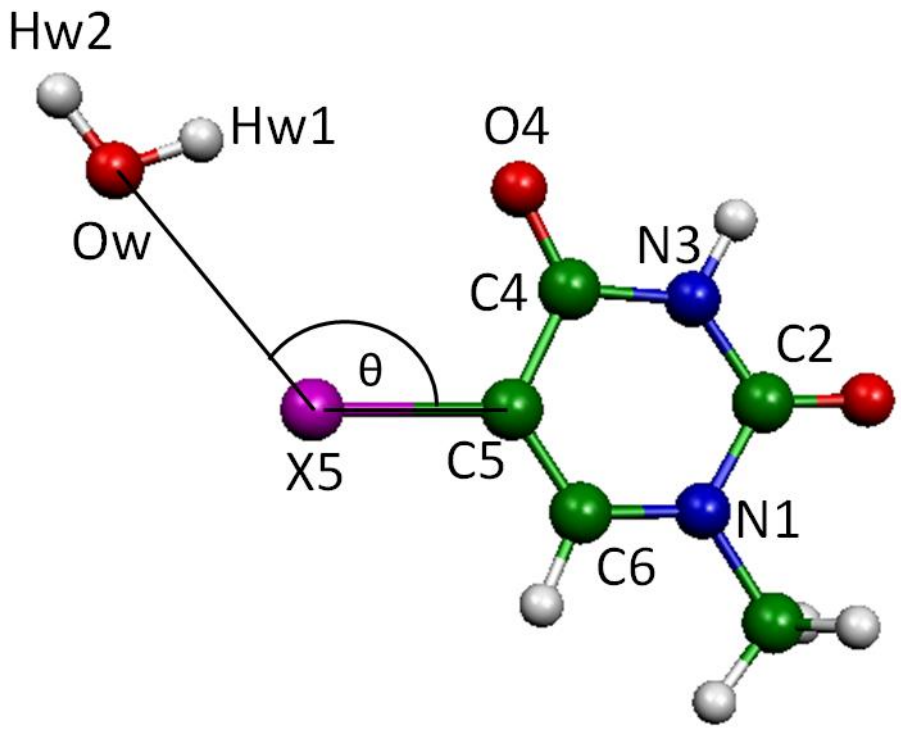


Figure 4.4: The C5-X5•••Ow angle, labelled θ is shown. This is the angle against which interaction energies have been plotted in figure 4.5 below.

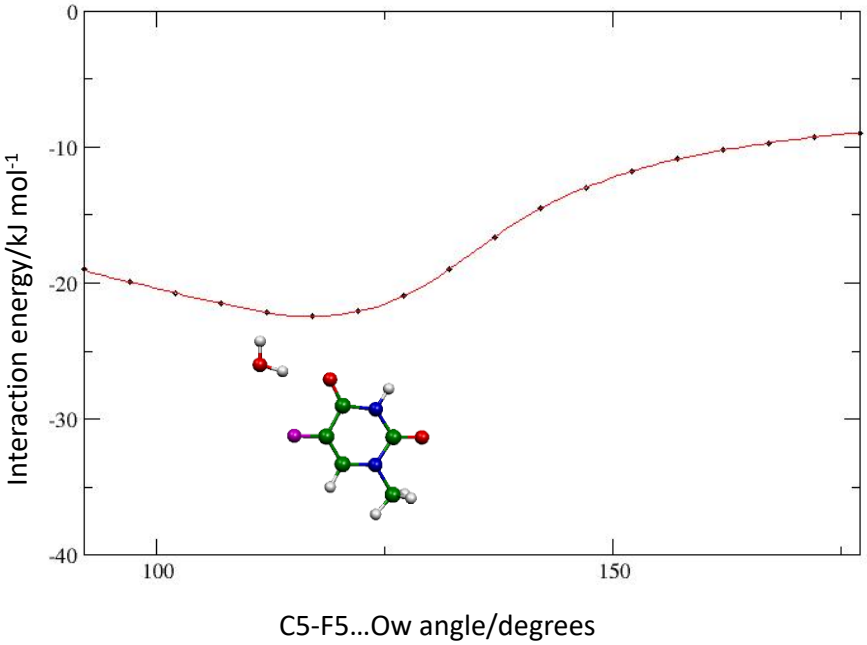


Figure 4.5a: Energy profile of FU•••w as a function of C5-F5•••Ow angle.

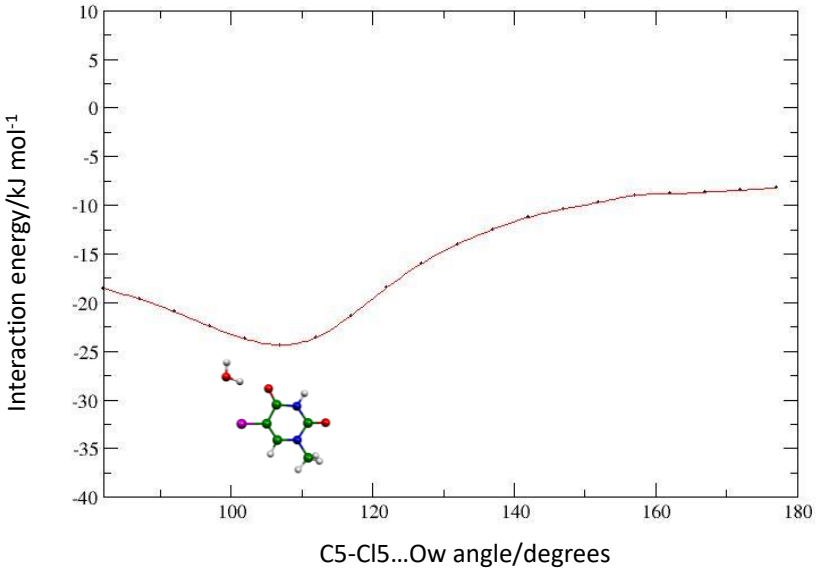


Figure 4.5b: Energy profile of ClU•••w as a function of C5-Cl5•••Ow angle.

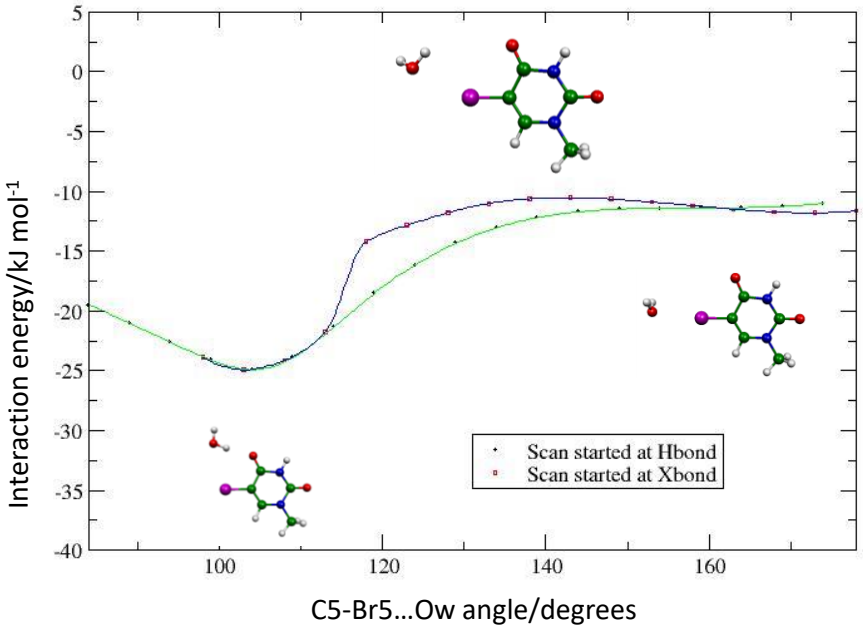


Figure 4.5c: Energy profile of BrU•••w as a function of C5-Br5•••Ow angle.

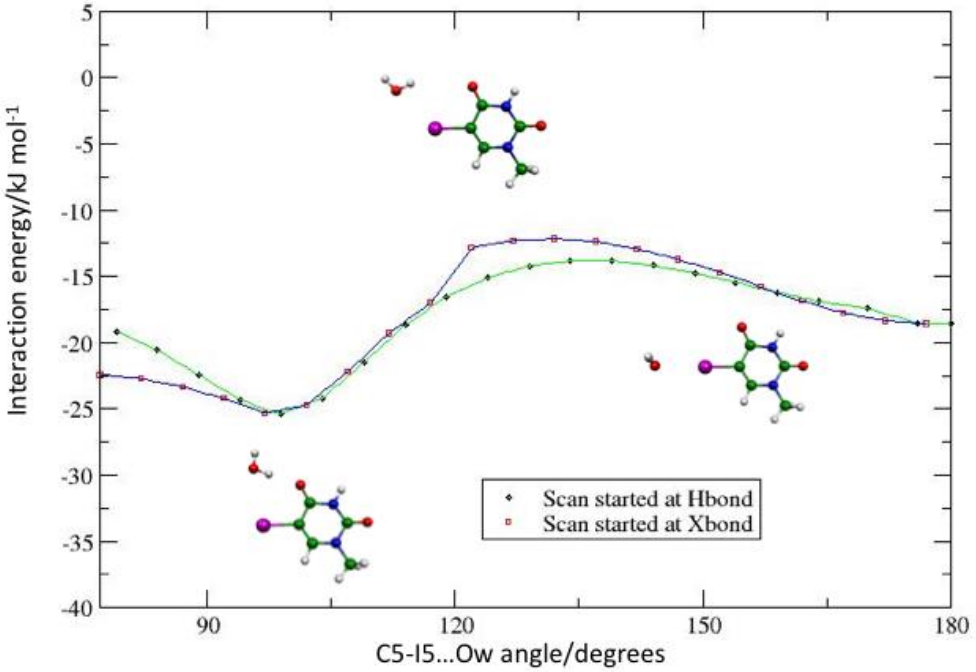


Figure 4.5d: Energy profile of IU•••w as a function of C5-I5•••Ow angle.

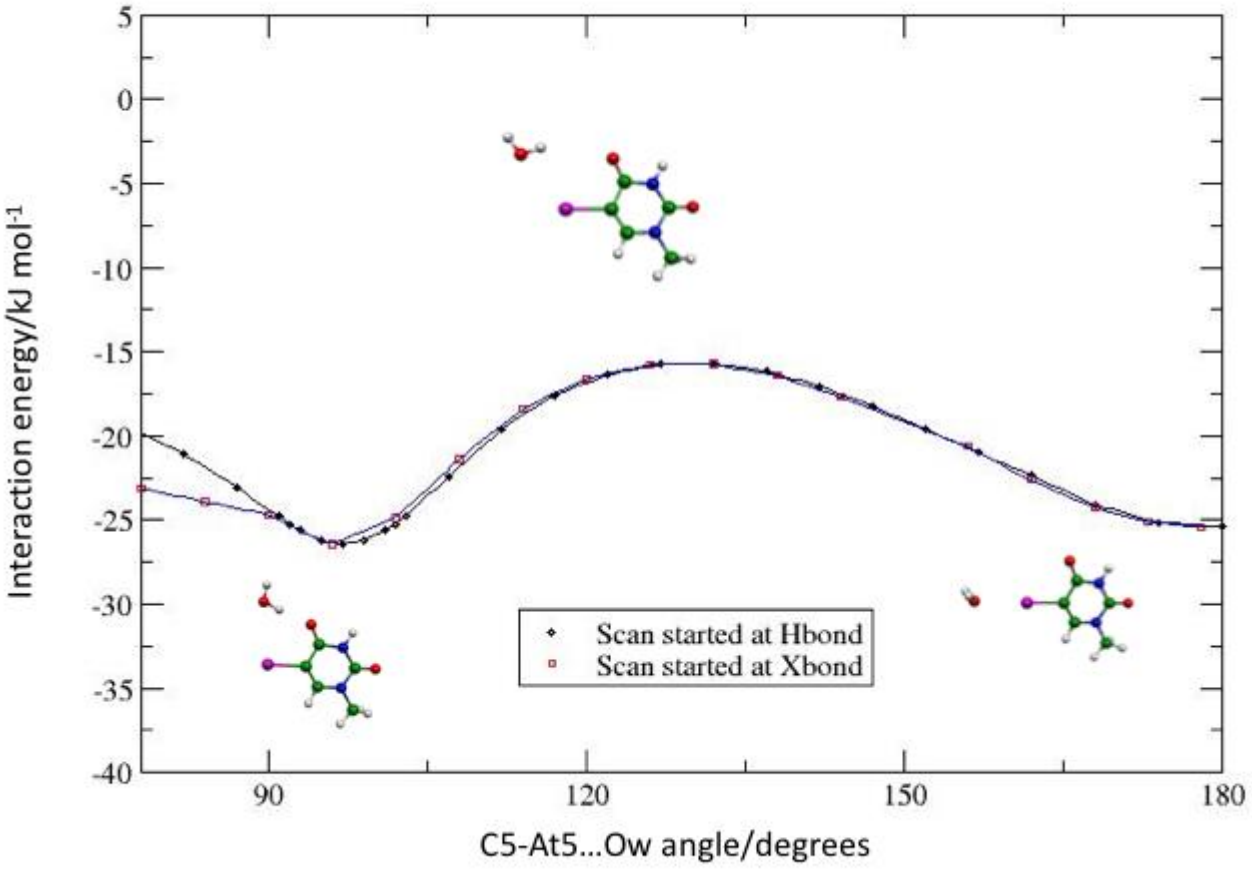


Figure 4.5e: Energy profile of AtU•••w as a function of C5-At5•••Ow angle.

Figure 4.5 shows how the barrier height going from Xbond to Hbond1 increases going down the halogen group. The reverse process is associated with a large barrier for BrU•••w, IU•••w and AtU•••w. This figure also shows that, with the exception of AtU•••w, the scan profiles as a function of the C5-X5•••Ow angle differ depending on whether the scan starts at the Hbond1 minimum or the Xbond minimum. The reason for this disparity is that, for the reasons discussed below, for the hydrogen bond minimum there is a preference for the water molecule in its entirety to be coplanar with the aromatic ring on the XU molecule, whereas for the halogen bond there is a preference for the water molecule to be so orientated that the water hydrogen atoms on the water molecule are placed either side of that plane and pointing to either side of the O4 atom. There must be an energetic barrier for the transformation between those two orientations for the water molecule. There is therefore a preference for retaining the orientation that is present in the starting structure until, as a function of the C5-X5•••Ow angle, that orientation becomes sufficiently unfavourable, whereupon the water rotates to the more

favourable orientation given the current C5-X5•••Ow angle. For BrU•••w this transformation takes place at approximately 160° if going from in plane (Hbond) to out of plane but at approximately 118° for the converse transformation. IU•••w behaves similarly to BrU•••w, but the Xbond to Hbond transformation takes place at approximately 120°. For reasons that have not been probed, this disparity in the transformation point does not apply in the case of AtU•••w, instead the Xbond to Hbond water molecule reorientation (as well as the Hbond to Xbond reorientation) occurs at 160°; hence the two profiles coincide almost perfectly for AtU•••w. For IU•••w and AtU•••w, if starting from the halogen bond minimum, for angles below 100°, the water molecule in its entirety (i.e. including the oxygen atom) goes out of the plane of the aromatic ring, probably going towards Hbond2 (which places the water molecule out of the said plane). Although it could be suggested that the presence of a (lower) curve that does not appear to disclose a transition barrier between the hydrogen bonded minimum and the halogen bonded minimum geometries could cast doubt upon the veracity of there being a transition barrier with a corresponding transition state, the transition state structures presented in figure 4.5 were computed by means of the Quasi-Synchronous Transit Newton method [106-107], based upon both of the minima thereby connected and without any proposed guess for the transition state (i.e. employing the QST2 keyword in Gaussian 09 [92]), and confirmed as a transition state by vibrational frequency analysis. Therefore at least at the level of theory employed in their computation, the disclosed transition states are *bona fide* transition states.

As can be seen from figure 4.3, the hydrogen bonded minima (except the out of plane Hbond2 minima) correspond to a structure in which the water molecule is coplanar with the aromatic ring, whereas the halogen bonding geometries place the water hydrogen atoms either side of the plane of the aromatic ring. The reason behind this geometrical difference has not been explored in detail, but is likely to be a combination of bond directionality and accommodation of secondary stabilising interactions. In particular, the orientation of the water molecule that is adopted in the case of halogen bonding maximises the overlap between the lone pair of electrons on Ow with the sigma hole on the halogen atom (corresponding to σ* antibonding orbital with respect to the C5-X5 covalent bond). Putting the hydrogen atoms out of plane also facilitates the secondary hydrogen bonding interaction not only between Hw1 and O4 but also between Hw2 and O4. The directionality of the halogen bond is linear as regards C5-X5•••Ow so as to optimise overlap with the sigma hole on X5 and the lone pair of electrons on Ow. The (weaker) directionality of hydrogen bonding is based on the positons of the lone electron pairs on O4. The placing of water protons either side of the aromatic plane could however be partly to facilitate the maximum interaction with O4 for both Hw1 and Hw2, while preserving halogen bond linearity and maintaining a favourable position of the Ow lone pair of electrons for interacting with the sigma hole on X5. Likewise, the coplanar hydrogen bonding geometry does appear to be well suited to X5•••Hw1 hydrogen bonding interactions as well as Hw1•••O4 hydrogen bonding, as a movement out of plane would extend the X5•••Hw1 distance (assuming an approximately unchanged Hw1•••O4 distance).

While none of the Hbond minimum structures, except for Hbond1, show an energetic trend that is maintained going from fluorine to astatine, it can be noted that for X=F or Cl, Hbond2 forms a stronger interaction than for Hbond1 (albeit marginally), whereas the converse is true for X=Br,

I or At (see table 4.1). The reason for this contrast has not been the subject of detailed analysis, but the combination of lesser anisotropy of electron density on the F and Cl (compared with X heavier than Cl) and the C5-X5 bond distance possibly becoming more geometrically conducive towards simultaneous X5•••Hw2 and O4•••Hw1 interactions might explain this comparative observation. Another interaction energy of note is the value for FU•••w Hbond3: FU•••w is the only system for which Hbond3 forms the strongest interaction. This is probably best explained by the strong electronegativity of the fluorine atom accentuating the electron deficiency of the H6 atom via the aromatic ring.

As can be seen in table 4.2, there is a trend towards smaller C5-X5•••Ow angles for the hydrogen bond minimum going down the halogen group. This observation might be rationalised by considering that as the sigma hole of the halogen atom expands, the surrounding ring of high electron density gets compressed into a smaller region that is perpendicular to the C5-X5 bond. If the Hw1•••O4 hydrogen bond is further stabilised by a contributing X5•••Hw1 interaction, then a compression of the negatively charged ring on the X5 halogen atom would more strongly favour a more perpendicular geometry between C5-X5 and Hw1, as the electron density would be more tightly concentrated in the perpendicular region of space around the X5 halogen atom, and due to a rigid covalent bond between Ow and Hw1, also a more perpendicular angle between C5-X5 and Ow.

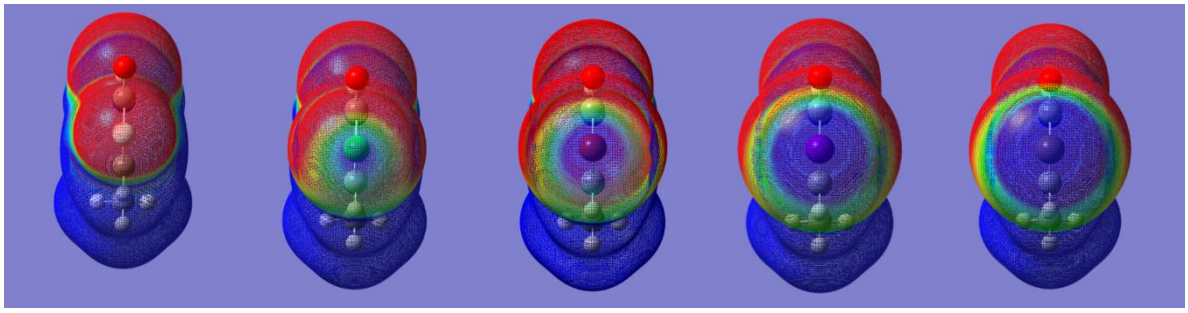
Table 4.2 also shows a (weak) trend toward more linear halogen bonds going down the halogen group. This observation can be explained by the (expected) trend towards stronger X5•••Ow halogen bonding interactions going down the halogen group, as this trend would diminish the relative contribution of secondary stabilisation effects (principally the Hw1•••O4 hydrogen bonding interaction) that could distort the C5-X5•••Ow angle.

Table 4.2 furthermore shows that, going down the halogen group, the transition state that connects the hydrogen bonding minimum and the halogen bonding minimum lies closer, at least in terms of the C5-X5•••Ow angular reaction coordinate, to the hydrogen bond minimum, i.e. the C5-X5•••Ow angle of the transition state becomes smaller going down the halogen group. This trend could support the hypothesis that the halogen bonding interaction is able to exert a dominating stabilising effect at smaller C5-X5•••Ow angle (further from linear) as the halogen group is descended.

The presence of the halogen atom at the 5 position on the 1-methyluracil aromatic ring does clearly lead to the formation of structures that are quite different from their non-halogenated counterparts, as indicated by comparison with the results obtained for uracil-water by Van Mourik *et al.* [146]. The authors found clear hydrogen bonding between the water protons and the oxygen atoms in the uracil molecule (both O4 and the oxygen atom bound to C2) and another hydrogen bond between Ow and one hydrogen atom in uracil (for any given structure). Of particular note is that the authors found a structure with a Ow•••"H5" (X=H) hydrogen bond, as well as a hydrogen bond involving O4 and a water proton, in the same structure. All of the structures presented by the authors show two clear hydrogen bonds. In the current study, each optimised minimum (with the exception of Hbond3 which does not involve halogen or hydrogen bonding to the O4 atom) has shown only one clear noncovalent interaction (either Hw1•••O4

hydrogen bond or Ow•••X5 halogen bond with a distance less than sum of VdW radii), albeit with significant secondary interactions. There are two factors that may contribute to the fact that the structures in the current study do not form a clear Ow•••X5 halogen bond and Hw1(or Hw2)•••O4 hydrogen bond simultaneously: the greater steric size of the halogen atoms (heavier than fluorine) could place a halogen bonded water molecule too far from the O4 atom, and the greater directionality of the halogen bond, compared with the Ow•••"H5" hydrogen bond found by Van Mourik *et al.* could also preclude the water molecule from simultaneously achieving a clear hydrogen bond with O4. However, the secondary stabilisations mentioned above indicate qualitative similarities to the environment that facilitates the doubly hydrogen-bonded structures that were found by Van Mourik *et al.*

Figure 4.6 shows electrostatic potential maps for the XU molecules (X=F, Cl, Br, I or At). This graphical representation shows how the anisotropy of the electron density on a halogen bond increases going down the group. These electrostatic plots confirm the physical basis for the rationalisations stated herein for the energetic and geometric results, grounded in electron density anisotropy of the halogen atoms and the observed trend for this anisotropy to increase going down the halogen group. It can be seen that the sigma hole is present but comparatively small (both in terms of spatial coverage and electrostatic magnitude) in the case of ClU, but is completely absent for FU. The trend towards enlargement of the sigma hole then continues for the remainder of the halogen group. These larger sigma holes explain the progressively stronger halogen bonding from BrU to AtU, from this figure it is also apparent that ClU should in principle be capable of forming halogen bonds, but that the magnitude of the sigma hole is too small to compete with the the rival O4•••Hw1 hydrogen bonding potential well. The region of the electrostatic effect of the O4 oxygen atom can be seen in figure 4.6, whilst the typical negatively charged ring around each of the halogen atoms (not obvious in the case of FU due to its lack of a positively charged sigma hole) can also be seen therein; there is clearly an asymmetry in the electron density distribution. Whereas above the halogen atom in each molecular image (i.e. the same side of the halogen atom as the O4 oxygen atom) a strongly pronounced negatively charged region can be seen, the magnitude of the negative charge decreases towards the C6 side of the halogen atom (below the halogen atom in the molecular images shown in figure 4.6), being approximately neutral in the region of the ring around the halogen atom furthest from the O4 oxygen atom, consistent with the C6-H hydrogen being positively charged, off-setting the effect of the anisotropy around the halogen atom. As the negatively charged region in the vicinity of the O4 oxygen atom expands at the expense of the positively charged sigma hole on the halogen atom, going up the halogen group, the O4•••Hw1 hydrogen bond becomes more competitive versus the X5•••Ow halogen bond. For ClU there is no region where the halogen bond can outcompete the hydrogen bond, even when the Ow atom is placed perfectly linearly along the extension of the C5-X5 covalent bond in the initial guess geometry, the relatively weak X5•••Ow halogen bonding type interaction could not stabilise the structure against the O4•••Hw1 hydrogen bonding type interaction.



FU ClU BrU IU AtU

Figure 4.6: Plots of electrostatic potential for isolated XU molecule. Density surface used for mapping is 0.0004 e-/au^3. The depicted electrostatic potential range is -6.93x10^{-3} Eh (red) to 6.93x10^{-3} Eh (blue).

For each system containing a halogen bond, the X5•••Ow bond was investigated by natural bond orbital analysis and the computation of the Wiberg bond index for that bond. The results for the XU•••w systems where X is Br, I or At, are disclosed in table 4.8.

Table 4.8: Wiberg bond index values for the X5•••Ow halogen bond.

Halogen	Wiberg bond index
Br	0.0130
I	0.0265
At	0.0313

As can be seen from table 4.8 the bond order, as measured by the Wiberg bond index increases from bromine to astatine, with a greater difference between bromine and iodine (0.0135) than between iodine and astatine (0.0048). The trend towards greater X5•••Ow WBI values going down the group suggests that the degree of covalency, which in this case corresponds to the charge transfer from the Ow lone pair of electrons to the sigma hole (i.e. the C5-X5 σ* antibonding orbital) increases from bromine to astatine. This result is not surprising as it is consistent with the increasing magnitude of the electron deficiency in the sigma hole region of space as the halogen group is descended. This trend for the magnitude of dative covalency is also consistent with the trend for interaction energy as both of these trends can be explained as being caused by the increasing magnitude of the sigma hole leading to stronger halogen bonding.

Additionally, these results, by their absolute values, indicate that the overall bond order is small, with the halogen bond interaction possessing relatively little covalent character. However, the absolute values for the WBI results would be expected in any event to be modest due to the weakness of halogen bonds in comparison with other types of interaction. As discussed below, for comparison, the AtU•••2w water-water hydrogen bond has a WBI of 0.0323; this is slightly greater than the value reported here for the AtU•••w halogen bond. Comparison with the corresponding results for the microsolvated halobenzene systems shows that the effect of the non-halogen substituents on the aromatic ring of uracil is to increase the WBI values. This result is not surprising due to their (collective) net electron withdrawing effect, accentuating the sigma hole effect on the halogen atom.

Table 4.9 shows WBI values found for the Hbond1 structures. The O4•••Hw1 and X5•••Hw1 interactions were probed by natural bond order analysis. Only in the case of the latter interaction does the halogen atom directly participate; however, the halogen atom can also be expected to have an electron withdrawing effect upon the O4 oxygen atom via the aromatic ring.

Table 4.9: Wiberg bond index values for the O4•••Hw1 hydrogen bond and X5•••Hw1 X•••H interaction.

Halogen	O4•••Hw1 WBI	X5•••Hw1 WBI
F	0.0223	0.0011
Cl	0.0219	0.0023
Br	0.0217	0.0025
I	0.0212	0.0023
At	0.0219	0.0025

The low absolute values for the WBIs show that the degree of covalency is small, especially for the X5•••Hw1 X•••H interaction. In the case of the latter, the large internuclear distances involved are likely to be a factor behind these very low WBI values.

Comparison with the Xbond results show that there is less dependency upon the identity of the halogen element for the WBI values for interactions involving the Hw1 atom. For the O4•••Hw1 hydrogen bond, X=Br produces a higher WBI value than for the X5•••Ow halogen bond (see table 4.8), but a lower value than for the X=At halogen bonds. For X=I the results are similar (slightly higher for the halogen bond). The greater variation in WBI results for the X5•••Ow halogen bond compared with the O4•••Hw1 hydrogen bond is to be expected on account of the direct involvement in the case of the former but not the latter.

The results set forth in table 4.9 in respect of the X5•••Hw1 X•••H interaction, save for the fluorinated system, disclose little variation in the Wiberg bond indices and no discernible trend therein. Descending the halogen group from fluorine to chlorine does however produce a change of WBI of 0.0008, although this change is still very small in absolute terms. This null result

is not entirely surprising as although the increasing electronegativity of a halogen atom going up the group would be expected to result in an increasingly electron rich negatively charged ring orthogonal to the C5-X5 covalent bond, this X•••H interaction is long distance. There is a clear trend towards increasing X5•••Hw1 internuclear distance as the halogen group is descended (see table 4.4). Hence the results shown in tables 4.4 and 4.9, when read together, indicate that there is a (near) perfect cancellation between these geometric and electrostatic effects. The relatively more significant effect of the fluorine atom indicates that in that case the electrostatic effect is slightly more significant; geometrically X=F fits into the same trend as the other halogen atoms (see table 4.4).

The O4•••Hw1 results show a trend towards smaller WBI values as the halogen group is descended from fluorine to iodine. These results are perhaps surprising. If, as noted elsewhere herein, there is a trend towards increasing electropositivity going down the halogen group then, *ex hypothesi*, the O4 oxygen atom could be expected to have greater electron density in the systems with the heavier halogen atoms, which in turn might be expected to afford a greater degree of dative covalency between the O4 oxygen atom and the Hw1 hydrogen atom. However, the trend disclosed in table 4.9 points in the opposite direction.

In the absence of any stronger factors, a possible (although somewhat doubtful) explanation for this trend might be found in table 4.5. Table 4.5, which discloses the values of the C5-X5•••Hw1 angle for each system, shows that that angle becomes more acute as the halogen group is descended. If the negatively charged ring on the halogen atom is sufficiently diffuse that it could overlap with the O4•••Hw1 region of space then that geometric coincidence might , be due to the overlap of electron density from the halogen atom, rather than electrons contributed by the participating atoms in the O4•••Hw1 hydrogen bond , raise the electron density in the natural bond orbitals engaged in the O4•••Hw1 hydrogen bond, and hence artificially raise the WBI values. As, going down the halogen group, the C5-X5•••Ow angle decreases, this effect could be expected to diminish, as the angle further departs from the ideal orthogonal value. This tentatively postulated explanation is given on the basis of a fairly clear trend in the WBI results for the O4•••Hw1 hydrogen bond, and what appears to be a total lack of any alternative explanation. However, the distances involved and the lack of any trend in X5•••Hw1 WBI results, do cast substantial doubt upon this explanation. A more detailed investigation, beyond the scope of this study, into the O4•••Hw1 hydrogen bond, employing the quantum theory of atoms in molecules (QTAIM) and symmetry adapted perturbation theory (SAPT), should be strongly encouraged in order to yield further information about the O4•••Hw1 hydrogen bond.

From X=I to X=At an increase in the O4•••Hw1 WBI value is observed. This result, although not enough to demonstrate a consistent trend, is consistent with the rather less strained analysis set forth above, based upon the trend towards greater electropositivity going down the halogen group. This driving factor towards a greater WBI result going down the halogen group being observed for the second heaviest member of the halogen group (behind tennessine) but not for the lighter halogens suggests that this this trend becomes more relative to the countervailing factor(s) at this point in the descent down the halogen group. A study including tennessine may be able to provide confirmation of this reversal in the magnitudes of the opposing factors.

However the very limited supply and short half-life of tennessine could, to an even greater extent than for astatine, impede experimental work on this element and that in turn could impair the ability of computational chemists to develop reliable methods and models for computations on tennessine.

4.4.1 Halogenated uracil with two water molecules

Table 4.10 contains structural results for the XU•••2w systems investigated.

No halogen bonding type interaction was found for the FU•••2w system; when an attempt to find an Xbond minimum for FU•••2w was made, the structure optimised to the Hbond minimum shown in figure 4.7. All of the other systems show a water-water hydrogen bond, hydrogen bonding involving one water molecule and the O4 atom, and a X5•••Ow halogen-bonding interaction (in this case Ow is whichever water oxygen atom is nearest to the X5 atom). If X5•••Ow interactions were not favourable then the internuclear distances would be expected to be approximately equal to or greater than the sum of their VdW radii. However, as can be seen in table 4.10, the VdW ratio is always less than 1.00 when X5 is heavier than fluorine. Table 4.8 also shows a clear trend from chlorine to astatine, with the interaction energy becoming more strongly attractive and the VdW ratio becoming smaller (also implying greater interatomic attraction) going down the halogen group. The latter trend is so strong in the case of iodine to astatine that the absolute internuclear distance decreases from 3.03 Å to 2.97 Å.

Table 4.10: Interaction energy, X•••Ow distance, O4•••Hw distance, C5-X5•••Ow angle, sum of VdW radii and VdW ratio given for all systems for which a halogen bond was formed. All data are for the geometry that contained a halogen bond. All VdW radii come from the CRC Handbook of Chemistry and Physics (90th edition) [135] except for astatine which comes from the website of the Royal Society of Chemistry [145].

System	Interaction energy/kJ mol^{-1}	X5•••Ow distance/Å	O4•••Hw distance/Å	C5-X5•••Ow angle/°	Sum of X5•••Ow VdW radii/Å	Ratio of distance to VdW
X=Cl	-64.79	2.92	1.93	159	3.27	0.89
X=Br	-68.84	2.98	1.94	155	3.37	0.88
X=I	-73.44	3.03	1.98	150	3.50	0.87
X=At	-79.34	2.97	1.99	150	3.54	0.84

Though the Hbond structures shown in figure 4.7 were not the principal focus of the study, their geometries are shown graphically therein and, for completeness, their interaction energies are presented in table 4.11. Note that the magnitude of the interaction energy is significantly stronger for the out of plane hydrogen bonded systems than for the corresponding halogen bonded system, for all halogen types (compare table 4.11 with the energetic results shown in table 4.10).

Table 4.11: Interaction energies for the Hbonded minima. All energies in kJ mol^{-1}.

Identity of X	Interaction energy
H	-83.73
F	-79.82
Cl	-87.84
Br	-90.00
I	-88.18
At	-87.29

Table 4.11 shows that BrU•••2w has the strongest interaction energy, and there is a trend towards weaker interaction towards the extremities of the halogen group. This observation may be due to competing influence of electronic and steric effects. Interestingly, where X is hydrogen this produces a very comparable interaction energy to the halogenated systems, with a value between those for fluorine and astatine.

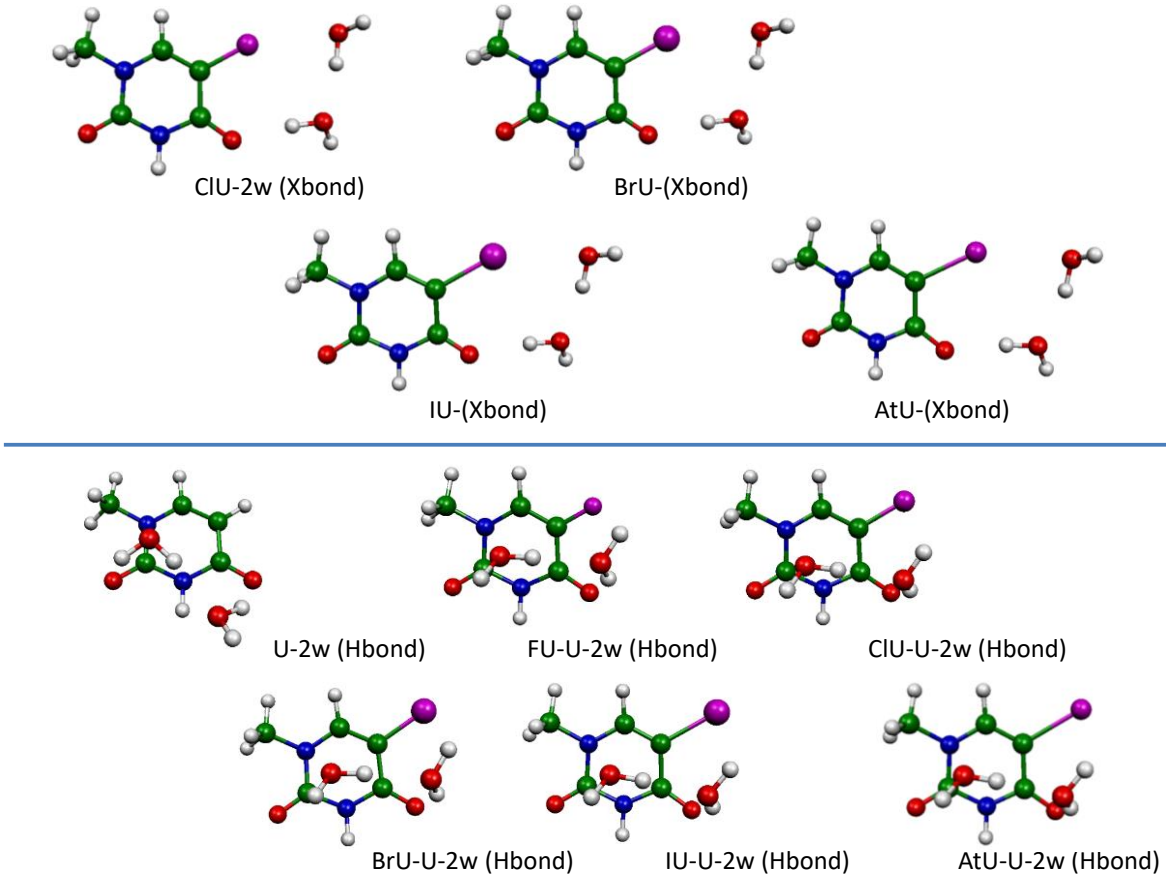


Figure 4.7: Structures for each of the minima identified for the XU•••2w systems.

These geometries of the XU•••2w systems (with the exception of FU•••2w) bear a resemblance to the structure comprising two water molecules and a non-halogenated uracil molecule, as presented by Van Mourik [147]. For the non-halogenated system, the author found that there is a hydrogen bond between a water oxygen atom and the "H5" atom (X=H), another hydrogen bond between a hydrogen atom on that same water molecule and the oxygen atom of the other water molecule, plus a third hydrogen bond between a hydrogen atom on that second water molecule and the O4 atom. Here, for X not F or H, the first of those hydrogen bonds (to "H5") is replaced with a halogen bond involving X5. The greater similarity between the halogenated systems and the non-halogenated system incorporating two water molecules, described by Van Mourik [147], compared with the greater differences between halogenated and non-halogenated systems in which only one water molecule is present (from Van Mourik *et al.* [146]) may indicate that

geometric restrictions arising from the linearity of halogen bonding can be relieved to some extent by the extra flexibility of the system provided by the second water molecule.

Whereas in the case of XU•••w, neither FU•••w nor ClU•••w form a halogen bond, FU is unique in not forming a halogen bond in the presence of two water molecules. A likely explanation for the presence of a halogen bond in the case of ClU•••2w but its absence in the case of ClU•••w is presumably the absence of the competing influence of the Hw1•••O4 hydrogen interaction in the ClU•••2w system (and indeed all of the other XU•••2w systems) by virtue of being obstructed by the water molecule that is not involved in the halogen bond.

The C5-X5•••Ow angles are considerably further from linear in the case of XU•••2w than in the case of XU•••w. The most likely explanation for this disparity is that in the case of XU•••2w the geometry must be compatible with three intermolecular interactions: the O4•••H hydrogen bond, the water-water hydrogen bond and the X5•••Ow halogen bond; each of which would have its own (competing) geometric preference; whereas the Xbond structures in XU•••w systems are only influenced by the X5•••Ow halogen bond and a weak secondary hydrogen bonding interaction (but by no means a proper hydrogen bond) between Hw1 and O4. This finding in the present study indicates that in a complex chemical environment wherein the halogen bond must compete with other interactions, substantial departure from linearity may exist. Similar findings were made in a study by Shields *et al.* [148], wherein multiple complexes involving halogen bonding between R-Br and a negatively charged electron donor (B) were investigated at varying R-Br•••B angles. That study elucidated that for the systems that were investigated, the R-Br•••B magnitude of interaction energies varied only slightly in the range 180° to 160° but then diminished far more sharply beyond approximately 150°.

Wiberg bond index values were calculated for the X5•••Ow1 halogen bond for all systems where this bond was found, i.e. where X is Cl, Br, I or At. These values are shown in table 4.12.

Table 4.12: Wiberg bond index values for the X5•••Ow1 halogen bond.

Halogen	Wiberg bond index
Cl	0.0098
Br	0.0133
I	0.0220
At	0.0308

Table 4.12 shows a clear trend towards greater values for the Wiberg bond index as the halogen group is descended. From X = Cl to X = Br the increase is 0.0035; further descending the halogen group to X = I yields a further increase by 0.0087, while going from X = I to X = At further increases the Wiberg bond index by 0.0088. Hence, in the region X = Br-At there is a linear trend, while a less pronounced increase in the WBI occurs when descending from chlorine to bromine. The low absolute values for the WBI for all halogen types (a WBI value of 1 would indicate a degree of cavalancy equivalent to a conventional single covalent bond, or “a unit charge density” as

described by Wiberg [23]) indicates that the interaction is not predominantly based upon dative covalency between the oxygen lone pair of electrons and the σ* orbital (corresponding to the sigma hole and being the antibonding orbital with respect to the C-X bond). Hence, the halogen bonding interaction corresponding to these respectively electron rich and electron poor regions is likely to be predominantly electrostatic. For comparison, the hydrogen bond that forms between the two water molecules, in the system for which X = At, has a WBI of 0.0323. This value is slightly greater than for the halogen bond where X = At and substantially greater than for all other halogen types. The low values for the WBIs are also a reflection of the weakness of halogen bonds, in comparison with, for example, first order covalent bonds.

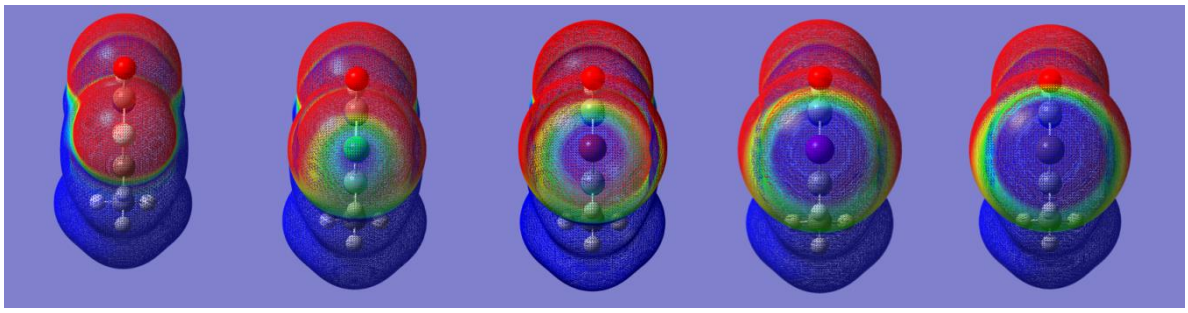
4.4.2 Comparison between the microsolvated 1-methyl-5-halouracil systems and their microsolvated halobenzene analogues

This section compares the XPh•••w results obtained in Chapter 3 with the XU•••w systems presented in the current chapter. The principal distinction between the XU•••w and microsolvated halobenzene systems (XPh•••w) is that in the case of the former but not the latter there is a competing option of forming a water•••oxygen hydrogen bond (O4•••Hw1). XPh can form a hydrogen bond between the negatively charged ring on the halogen atom (i.e. the region around the halogen atom approximately perpendicular to the C-X bond wherein there is an elevated concentration of electron density) and the water oxygen atom (X•••Hw) but it is not possible to establish a geometry that permits both halogen bonding and hydrogen bonding.

Perhaps surprisingly ClPh•••w does not form a halogen bond. Given that when the Hw1•••O4 hydrogen bond was obstructed in the case of ClU•••2w, a halogen bond did form between water and XU it is clear that water can, in the absence of a competing interaction form a halogen bond with a chlorine atom that is covalently bonded to an aromatic ring. The most probable explanation for this discrepancy is the impact of the O4 atom in XU on the electronic structure of the XU molecule by means of electron withdrawing effects. It is well known that ketones are polar with electron density concentrated on the oxygen atom and that a sigma electron withdrawing effect could be depleting the electron density in the ring in XU; this electron withdrawing effect would be expected to increase the magnitude of electron deficiency in the sigma hole region on the chlorine atom, thereby rendering the chlorine atom more capable of forming halogen bonds. It would then be the combination of the greater electron deficiency of the sigma hole region on ClU compared with ClPh and the obstruction of a Hw•••O4 hydrogen bonding type interaction (by the second water molecule) that explains the formation of the halogen bonded structure in the case of ClU•••2w. This rationalisation is also consistent with the finding that the magnitude of the interaction energy for the halogen bond structures is greater for XU•••w than for the corresponding XPh•••w structures.

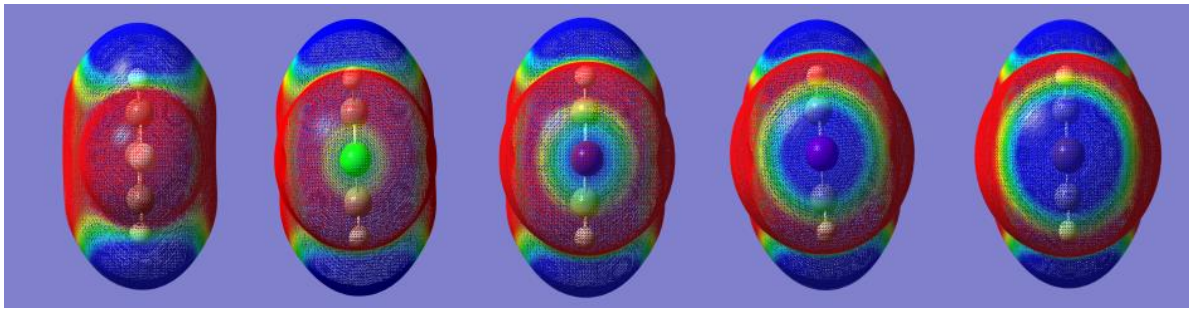
Figure 4.6 is reproduced below as figure 4.8 for ease of comparison with figure 4.9, references to figure 4.6 and figure 4.8 are interchangeable. Comparison between figure 4.8 and figure 4.9, which shows the corresponding molecular electrostatic potential for the halobenzene molecules, provides information about the polarising effects of the substituents on, and heteroatoms within, the aromatic ring found only in the XU molecules. Whilst the methyl substituent could be expected to be electron donating (consistent with the blue (positive) region around this substituent in figure 4.8), both of the two carbonyl groups and the two nitrogen heteroatoms within the ring could be expected to withdraw electron density from the six-membered ring. The effect of the O4 oxygen atom and the red (negatively charged) region associated therewith has been discussed above. However, a further observation can be made when comparing figure 4.8 with figure 4.9, the size (both region of coverage and the magnitude of the positive charge) of the sigma hole is appreciably greater in the cases of XU molecules than those of the XPh molecules (X = F excepted). The only possible explanation for this increase in the size of sigma hole in comparison with the same in the case of halobenzene is that the

substituents and heteroatoms present in the case of XU cause the C5 carbon atom to have less electron density compared with the corresponding carbon atom (bound to the halogen atom) in halobenzene, hence causing the X5 halogen atom to also have less electron density, the total electron density on the X5 halogen atom being dependent upon the polarity of the C5-X5 covalent bond, and hence upon the electron density on the C5 carbon atom. Whilst the smaller size of the sigma hole in the case of XPh is of interest for X = Br, I or At for the reasons given above; it is the case that ClPh is the most noteworthy; although, unlike for FPh, the anisotropy of the electron density around the halogen atom is apparent from the molecular electrostatic potential plot, even directly along the extension of the C-Cl bond the effect of this anisotropy is to produce an approximately neutral region, as expected given the electronegatively of chlorine; the total charge around the chlorine atom is negative. This lack of any positively charged region around the chlorine atom in ClPh explains why that molecule is not able to form a halogen bond with a water molecule, despite the lack of any competing hydrogen bond type interaction. As could be seen in the case of XU•••2w, ClU and water can form a halogen bond where the opportunity to form an O4•••Hw1 hydrogen bond has been denied.



FU ClU BrU IU AtU

Figure 4.8: Plots of electrostatic potential for isolated XU molecule. Density surface used for mapping is 0.0004 e-/au^3. The depicted electrostatic potential range is -6.93x10^{-3} Eh (red) to 6.93x10^{-3} Eh (blue). This figure is a reproduction of figure 4.6.



FPh ClPh BrPh IPh AtPh

Figure 4.9: Plots of electrostatic potential for isolated XPh molecule. Density surface used for mapping is 0.0004 e-/au^3. The depicted electrostatic potential range is -6.93x10^{-3} Eh (red) to 6.93x10^{-3} Eh (blue).

Comparison of the WBI results for XU•••w and XU•••2w discloses that the absolute values for the WBIs in each case (excluding the ClU•••2w halogen bond for which no analogue exists in the XU•••w case) are very similar for the two different types of system. The WBI results address the degree of orbital overlap, i.e. the covalent aspect of the halogen bond and does not purport to comment on the electrostatic aspect which is believed [30] to predominate for these systems. However, the fact that the presence of the second water molecule makes very little difference to the WBI values is indicative of consistency as to the magnitude of the sigma hole interaction. This result is perhaps surprising given the very large displacement from the optimal linear geometry for the halogen bonds in the case of XU•••2w. Conversely the ability of the Ow2 lone pair of electrons to donate electron density to the water molecule engaged in halogen bonding could increase the electron density on the Ow1 atom and the ability of the O4 oxygen atom to dispose of some of its electron density via hydrogen bonding to the other water molecule might, by conjugation, exacerbate the electron deficiency of the sigma hole on the halogen atom. Both of these factors could be expected to strengthen the halogen bond, offsetting the effect of reducing the halogen bond angle. However, this cancellation of factors appears to be suspiciously perfect and a more detailed comparison may be required to elucidate the true reason for these observations which arise from the calculations. This study has predominantly focussed upon the trend going down the halogen group, and here the WBI values can be readily explained. Both XU•••w and XU•••2w show a trend towards greater WBI values going down the halogen group. These results are consistent with strengthening of the halogen bond from Br to At (Cl to At in the case of XU•••2w). The similarity in the WBI trends match the comparison of the trends for the interaction energies and are not surprising given what is known [8] concerning the trend for more pronounced halogen bonding and greater role of charge transfer as the halogen group is descended.

The additional complexity of the XU•••w chemical environment compared with the XPh•••w systems resulted in some departure from the almost linear (179°) C-X•••O angle found for all of the XPh•••w halogen bonded structures. The most influential factor in causing this departure

from linearity is almost certainly the stabilising influence of a secondary interaction between Hw1 and O4, an interaction that can be strengthened by moving the water molecule from its otherwise optimal Xbond geometry.

While the differences between the XU•••w systems and their simpler XPh•••w analogues have been the focus of this comparison, it should be noted that the trends identified going down the halogen group for the the XPh•••w systems are replicated for the more complex XU•••w systems.

4.5 Conclusions

All of the XU molecules were found to form hydrogen bonds with a water molecule. Furthermore, there are three distinct hydrogen bond stabilised minima that are common to all of the systems, labelled Hbond1, Hbond2 and Hbond3. An unhalogenated form of XU•••w (X=H) also forms a structure that is analogous to the Hbond1 minimum, with a stronger interaction (more negative interaction energy) than for any of the halogenated systems, presumably due to the formation of a second hydrogen bond, between Ow and H5. The only hydrogen bonded structure for which there is a clear energetic trend going down the halogen group is for Hbond1, which exhibits a trend towards more negative (more strongly attractive) interaction energies going down the halogen group. This trend may be due to a combination of a weakening of the electron withdrawing effect going down the halogen group resulting in greater electron density on the O4 and a greater σ-hole effect leading to a greater concentration of the negative charge on the halogen atom in the region perpendicular to the C5-X5 covalent bond as the halogen group is descended. The brominated and astatinated systems have the strongest interaction for the Hbond2 and Hbond3 minima respectively, but in both cases by a very small margin

The brominated, iodinated and astatinated systems also form structures that are stabilised by an X5•••Ow halogen bond. These structures were found to be connected to Hbond1 by a transition state. For the Xbond minimum there is a clear trend down the halogen group towards stronger halogen bonding: the brominated, iodinated and astatinated systems have interaction energies of -12.2 kJ mol^{-1}, -18.6 kJ mol^{-1} and -25.5 kJ mol^{-1} respectively. Furthermore, the X5•••Ow internuclear distance as a proportion of the sum of the VdW radii (referred to as VdW ratio) were calculated to be 0.90, 0.85 and 0.83, respectively. Wiberg bond index values for the halogen bond increased from bromine to astatine, with values of 0.0130, 0.0265 and 0.0313 for X=Br, I and At respectively, indicating a small degree of covalent contribution to the halogen bond, which increases going down the halogen group.

The interaction energy for the astatinated system has potentially significant implications for similar systems where there may be competition between halogen bonding and hydrogen bonding as it is only 1 kJ mol^{-1} less stable than the Hbond1 minimum. There is also a strong trend down the halogen group towards greater barrier heights (between the Xbond minimum and the TS transition state); the barrier heights for the brominated, iodinated and astatinated systems are 0.5 kJ mol^{-1}, 4.9 kJ mol^{-1} and 9.8 kJ mol^{-1} respectively. The small BrU•••w barrier height indicates that this system can only be expected to be meta stable, even at a temperature of 0 K.

All XU•••2w systems, except where X=F, form minima that are stabilised by X5•••Ow halogen bonding. The VdW ratios are 0.89, 0.88, 0.87 and 0.84 Å, for X= Cl, Br, I and At respectively. The halogen bond angle in these systems, which ranges from 150-160°, is far from the expected near-linearity of halogen bonds. This indicates that significant non-linear halogen bonds may exist in complex environments with competing interactions present.

Comparison with the XPh•••w systems show that the trend towards stronger halogen bonding going down the halogen group is maintained between the simpler and more complex systems.

However, the more complex chemical environment of XU•••w results in less linear halogen bonds, principally due to the competing hydrogen bonding type interaction with the O4 atom. Further comparison with the results for the XU•••2w systems suggests that, subject to the Hw•••O4 interaction being blocked, water can form a halogen bond with ClU but not with ClPh; this observation is in line with the general finding that XU forms stronger halogen bonds than XPh. The absence of a halogen-bonded ClU•••w structure shows that the potential to form halogen bonds can be reduced or eliminated by nearby hydrogen bonds.

5 An investigation into halogen bonding between thyroxine and the protein backbone

5.1 Abstract

Calculations were performed on systems comprising a fragment of the thyroxine molecule and either a fragment of a protein backbone or a water molecule. The molecular fragments were all within an 8 Å radius of an iodine atom that constituted part of the thyroxine molecular fragment. All of the systems examined herein were extracted from crystal structures in the Protein Data Bank (PDB). Structures where there appeared to be a *prima facie* possibility of halogen bonding (or at least halogen-bond-type interactions falling short of formal halogen bonds) were subjected to further study. A cut-off of 150° was usually (but not with full rigidity) employed for the halogen bond angle. As the effect of the halogen bond in isolation was the main focus of this study, structures were modified to minimise other interactions, including hydrogen bonds. In most cases the halogen atom (at least potentially) interacted with a lone pair of electrons on an oxygen atom, but in some instances the potential electron donor was a lone pair of electrons on a nitrogen atom. Furthermore, calculations were performed on some of the systems wherein the iodine atom that was, at least potentially, involved in halogen bonding was substituted with an astatine atom. For each system that was the subject of detailed investigation, the (potential) halogen bond angle, internuclear separation and Wiberg Bond Index (WBI) were calculated before and after computational geometry optimisation. The complexation energy between the molecular fragments was also calculated for both the crystal structure and optimised geometry. Comparisons were made between iodine-based and astatine-based halogen bonds where calculations on analogous systems had been performed. Structures were extracted from the PDB using Readpdb. Gaussview 4.1 was employed for molecular visualisation and editing. All DFT computations were performed using Gaussian 09 [92]. The functional M06-2X was employed in conjunction with either the 6-31+G* basis set (for elements lighter than iodine) or the aug-cc-pVDZ-PP basis set with corresponding relativistic effective core potentials for iodine and astatine. An analysis of the frequency of halogen bonded structures in the Cambridge Structural Database (CSD) as a function of internuclear separation and halogen bond angle was performed. This aspect of the study did not include astatine. Many but not all of the identified molecular fragments were found to form halogen bonds in the crystal structure. Geometry optimisation in most cases resulted in a decrease in internuclear separation and an increase in halogen bond angle linearity. In all instances, the optimised geometries have an internuclear separation of less than the sum of Van der Waals (VdW) radii. Astatine forms stronger halogen bonds compared with iodine, based upon complexation energies and VdW ratios. However, astatine's lead over iodine is diminished with increased internuclear separation in the crystal structure. Statistical analysis of the frequency of halogen bonded structures in the CSD as a function of internuclear separation and halogen bond angle revealed a greater propensity for iodinated molecules to form halogen bonds with short internuclear separations and near-linear halogen bond angles, in comparison with the lighter halogen elements.

5.2 Introduction

The thyroid gland is the organ wherein the hormones thyroxine (known as T4) and triiodothyronine (known as T3) are synthesised. This endocrine gland is situated in the neck. These hormones are then delivered into the blood stream by secretion. The hormone T3 is derived from T4 by the elimination of one iodine atom from the thyroxine (T4) molecule [149]. T3 is vital for metabolism in humans and deficiency thereof is responsible for maladies including, *inter alia*, depression and elevated body mass index [149-150].

Auffinger *et al.* identified the presence of halogen bonding between iodine and oxygen atoms in the thyroid systems [15]. Bayse and Rafferty found that halogen bonding can help to facilitate the cleavage of the C-I bond responsible for the conversion of T4 thyroxine to the T3 form [151].

Bayse and Rafferty, in their above cited communication [151], identified S•••X and Se•••X halogen bonding as facilitating C-X bond cleavage. In their study X could be either iodine (as found in the thyroid system) or bromine. The authors' study focussed on the effect of halogen bonding on the C-X bond distance as a proxy for its effect on the stability of this covalent bond. Halogen bonding between selenium and iodine was found to be strong and to play a significant role in C-I bond cleavage. The authors identified the halogen bonding properties, combined with inherently weak bonding between iodine and carbon, and consequential ease of dehalogenation as a reason for evolution's selection of iodine over the lighter, and more abundant halogens [151]. The authors disclose a full proposed mechanism for the reductive elimination of one iodine atom from T4, yielding the active molecule T3. In this mechanism the selenium atom ultimately forms a full covalent bond with the abstracted iodine atom. Although not conclusive, the Se•••I halogen bond may be expected to have a signicant degree of covalency, as this would be consistent with the requisite weakening of the C-X covalent bond by donation into the σ* antibonding orbital; furthermore full cleavage of the C-X bond and formation Se-I could be viewed as a limiting case of C•••X•••Se, where the X•••Se interaction is a halogen bond with substantial covalency by virtue of donation into the σ* orbital, which corresponds to the σ-hole on the halogen. The cited communication also makes passing reference to weaker I•••O halogen bonding contacts but these are not the focus of the authors' investigation.

In a subsequent paper authored by Marsan and Bayse [37], the possibility of inhibition of the activation pathway by the competitive binding properties of polybrominated diphenyl ethers, is discussed. Despite Se-Br being a weaker interaction than Se-I, the disclosed results do indicate that this form of inhibition can occur.

Halogen bonding and hydrogen bonding in the context of the stabilisation of antithyroid drugs with hypohalous acids (with the halogen being chlorine, bromine or iodine) is discussed in a paper authored by El-Sheshtawy and El-Mehasseb [36]. A hypohalous acid is a compound with the general formula HOX, where "X" is a halogen and "H" and "O" are hydrogen and oxygen respectively.

In the cited paper authored by Auffinger *et al.* [15] a survey of potential halogen bonds involving chlorine, bromine or iodine is disclosed. Primarily O-X halogen bonds are discussed, although halogen bonds involving sulphur, nitrogen and delocalised electrons in peptide bonds are also mentioned. Based on the authors' survey of the Protein Databank [48], a trend was identified in which the propensity for halogen bonds to form increased down the group [15]. In particular, the O-X short contacts were broken down by halogen type, with chlorine, bromine and iodine accounting for 27%, 34% and 39% of these geometries respectively. The authors further identified that carbonyl oxygen atoms "dominated" these interactions, out of a collection of functional groups that also included "hydroxyl (O-H), or negatively charged acid ($^{-}$O-C/P/S)" [15]. The authors also briefly noted that some O-F short contacts were found but did not explore these any further on account of fluorine's electronegativity. The authors also calculated electrostatic potential plots for molecules containing chlorine, bromine and iodine, and found a trend towards a stronger σ-hole effect as the halogen group was descended. Furthermore, a trend was found wherein this effect was greater for cases where the halogen atom was on halogenated cytosine than for halogenated uridine, where in turn it was stronger than for halogenated methane.

The present study seeks to further investigate the halogen bonding that occurs in the thyroid system between iodine atoms on the thyroxine molecule and (a) oxygen atoms in the protein backbone, (b) nitrogen atoms in the protein backbone and (c) oxygen atoms in water molecules present within the crystal structures. Furthermore, the effect of substituting the iodine atom with an astatine atom upon the strength and influence of the halogen bond is studied.

In the interest of computational economy, fragments of thyroxine and protein backbone were modelled in lieu of the full molecules. The contribution of halogen bonding to the complexation energies between the molecular fragments and the influence thereof upon the geometry of the dimers was probed, with and without substitution of iodine for astatine.

Statistical analysis was performed against the structures containing halogen atoms in the Cambridge Structural Database (CSD) [80]. The Cambridge Structural Database (CSD) is a repository for crystallographic data of small molecules [80].

5.3 Methods

Crystal structures incorporating thyroxine were extracted from the Protein Data Bank (PDB) [48]. The following proteins, referred to by their PDB codes, were employed in this study: 1ETA [152], 1HK1 [153], 1HK2 [153], 1HK3 [153], 1HK4 [153], 1HK5 [153], and 1ICT [154].

The programme Readpdb [155] was employed to select a region of each structure that was within 8 Å of the iodine atoms engaged in halogen bonding. Hydrogen atoms were then added to the (unprotonated) structure in Gaussview 4.1 [122]. The complexation energies were calculated at the crystal geometry by performing single point calculations and employing the counterpoise procedure developed by Boys and Bernardi [74] to calculate, in addition to the dimer energy, the non-complexed monomer energies calculated in the dimer basis set; these monomer energies were then subtracted from the dimer complex energy to obtain a so-called complexation energy. This complexation energy differs from a full CP-corrected interaction energy by not including deformation energies. Where the structure appeared to contain both halogen bonds and hydrogen bonds, in addition to performing the calculation on the unmodified crystal structure, further calculations were performed on structures that had undergone some modification in Gaussview 4.1 [122] to remove or diminish the hydrogen bond while preserving the halogen bond.

Optimisation calculations were then performed to establish whether the apparent halogen bond interaction (even if not a formal halogen bond) would remain preserved and if so whether it would optimise (become more linear and have a bond distance that is within the sum of VdW radii). Again the complexation energy, halogen bond distance and halogen bond angle were elucidated. In the case of optimisations, minima were confirmed by vibrational frequency analysis. We calculated complexation energies, not full CP-corrected interaction energies, to provide for easier comparison with the complexation energies calculated for the unoptimised crystal structures.

Calculations involving astatine entailed substitution of the iodine atom that is engaged in halogen bonding with an astatine atom. For the carbon – astatine bond distance the default bond distance in Gaussview 4.1 [122] was used and the halogen – oxygen (or nitrogen) distance was restored to that present in the crystal structure with the halogen-bonded iodine. This was also used as the starting structure for the optimisation calculations that were performed on the astatinated systems. In all other respects the method used in respect of the astatine substituted systems was the same as for those performed without substitution.

All calculations were performed using Gaussian 09 [92]. The M06-2X [70] functional was used throughout. Iodine and astatine atoms were modelled using the aug-cc-pVDZ-PP basis set (which includes a relativistic pseudopotential) [64]; for all other atoms the 6-31+G* [93] basis set was employed. The integration grid was specified as ultrafine.

All technical details regarding the role of each piece of software employed in the statistical analysis set forth herein and the representation of the results thereof were disclosed by a private

communication with Rachael Skyner [123]. However, the analysis itself was performed by the author of this thesis.

The Conquest software package [124] was used to perform a substructure search of the CSD for substructures which satisfy certain criteria. For this study structures that satisfy the following criteria were sought in the search: (i) that the structures each contain one or more atoms of a specified halogen element (X) and one or more atoms of any element belonging to group 15 or 16 in the periodic table of the elements (Y) that is covalently bound to a molecular fragment group comprising one or more atoms (R), (ii) that there be at least one instance within the structure where a pair of atoms comprising an atom satisfying the hereinbefore disclosed definition of X and one atom satisfying the hereinbefore disclosed definition of Y wherein the internuclear separation between relevant atoms X and Y be no greater than the sum of the VdW radii of X and Y, and (iii) that the obtuse angle subtended from atoms Y to X (satisfying the criteria disclosed in clause (ii) within this paragraph) and the atom belonging to group R which is covalently bound to atom X not be smaller than 160°. The results from this search were then exported to the Mercury software package [124-127]. The Mercury software package was employed to analyse the structures from the CSD that satisfy the criteria disclosed above. In particular, a histogram of angle R-X•••Y and a histogram of the internuclear X•••Y separation were constructed. These data were then exported in comma separated values (csv) files to MATLAB [128]. Within MATLAB a modified version of the dscatter script [129] was employed to plot the two histograms against each other. Thereby the frequency of structures satisfying the criteria disclosed in clauses (i)-(iii) above occurring within certain ranges of angle R-X•••Y and internuclear distance X•••Y was elucidated; colour was used to disclose the magnitudes of the frequencies. The edit_script.m script [130] was also employed for labelling and formatting the plots.

The addition of a correction based upon the sine of the R-X•••Y angle to compensate for the statistical dependence upon angle of the number of points on a sphere that satisfy any specified angle [156] would enable clearer quantification of the R-X•••Y angle dependence. However, in the present study, this correction was omitted. This correction would entail dividing the frequency of structures at each point on the heatmap by the sine of the halogen bond angle at that point on the x-axis. In practice this would be based upon the number of structures in each bin of the histogram and the angular range of the bin. Thereby the artefact described above wherein the number of points on a sphere with a given angle tends to zero as the angle tends to 180°, while being at a maximum at 90°, is precisely counterbalanced. At 180° the infinities cancel, while at 90o the division by the sine of the angle equates to division by unity, so has no effect. Also, between these extremes, the sinusoidal correction entails division by a denominator between zero and unity that exactly cancels the trend towards a decreasing number of unique points on the sphere satisfying a the given angle as that angle tends from zero to unity. In the absence of this correction the number of structures with angles close to linear is understated. Hence, the effect of this omission is to understate the effective abundance of structures consistent with the presence of halogen bonds in the statistical samples. Corrected versions of the heatmaps in figure 5.2 would show a greater concentration of structures in the linear region for all halogen elements, and therefore may show a greater tendency for the lighter halogens to form halogen

bonds, but would also show an even greater propensity for iodinated structures to concentrate in the geometric region close to linear.

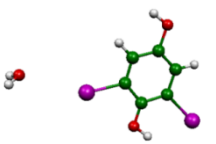
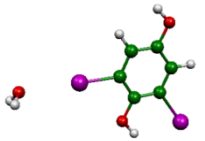
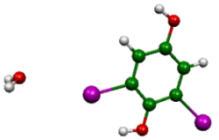
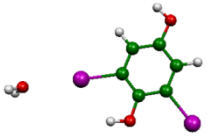
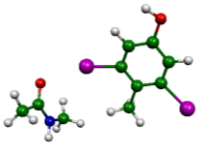
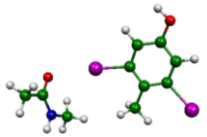
5.4 Results and discussion

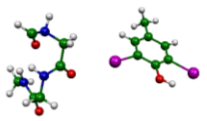
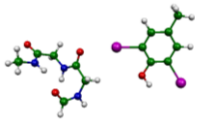
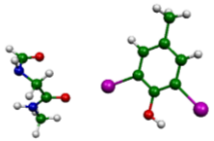
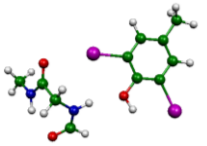
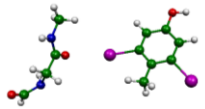
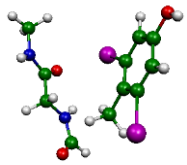
The results for the natural iodine containing structures and the structures containing astatine are presented separately below.

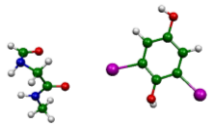
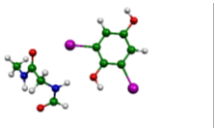
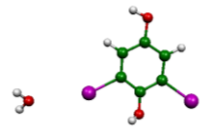
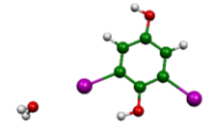
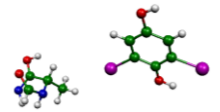
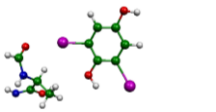
5.4.1 Results from the iodine based system

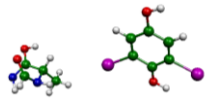
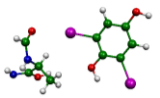
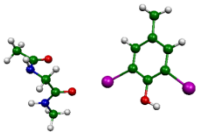
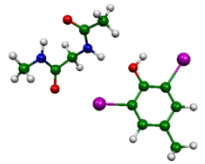
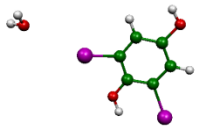
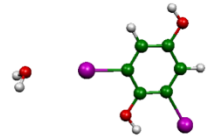
Table 5.1 shows the interaction energy, the halogen bond distance, the halogen bond angle, and the structure in graphical form, both taken directly from the crystal structure (with added hydrogen atoms) and after computational optimisation, for each crystal structure where halogen bonds were identified as being present. All of the halogen bonds are from iodine on thyroxine to an oxygen atom except in the unoptimised 1HK1 04286, where it is to a nitrogen atom. All halogen bonds are intermolecular. Complexation energies include all interactions, but attempts have been made to minimise non-halogen bond interactions, for example by substituting OH groups with CH_3 groups to remove hydrogen bonds. The Van der Waals (VdW) ratios are based on the VdW radii for each element disclosed in the CRC Handbook of Chemistry and Physics [135] except for astatine which is disclosed on the website of the Royal Society of Chemistry [136].

Table 5.1: Energetic, geometric, WBI and graphical results for both the crystal and optimised, iodine-based, structures.

Name of system*	ΔE^{unopt} kJ/mol	ΔE^{opt} kJ/mol	WBI^{unopt}	WBI^{opt}	R^{unopt} Å	R^{opt} Å	VdW $ratio^{unopt}$	VdW $ratio^{opt}$	α^{unopt} deg.	α^{opt} deg.	Unoptimised structure	Optimised structure
1HK1 04288 H atom moved	-6.47	-15.72	0.0063	0.0209	3.56	3.04	1.02	0.87	151.63	180.00		
1HK1 04288 modified	-9.19	-18.54	0.0065	0.0179	3.56	3.02	1.02	0.86	151.63	179.77		
1ETA 02040 modified CH3	-19.08	-21.60	0.0148	0.0194	3.12	3.03	0.89	0.87	161.01	172.50		

1HK1 04309	-11.45	-21.79	0.0104	0.0221	3.39	3.02	0.97	0.86	169.90	172.11		
1HK3 04345 modified	-12.45	-40.03	0.0101	0.0203	3.48	3.06	0.99	0.87	169.30	163.79		
1ICT 07198 Modified CH3	-12.50	-22.54	0.0092	0.0058	3.30	3.47	0.94	0.99	166.08	145.15		

1HK2 04349 modified	-4.19	-41.25	0.0033	0.0217	3.99	3.04	1.14	0.87	151.76	163.79	 
1HK3 04396 modified	-6.89	-18.54	0.0047	0.0178	3.57	3.02	1.02	0.86	151.23	179.90	 
1HK1 04286 modified NH geom. Edited**	-1.60	-25.81	0.0006	0.0156	5.24	3.11	1.48	0.88	160.53	171.66	 

1HK1 04286 Modified **	-1.65	-25.81	0.0006	0.0156	5.24	3.11	1.48	0.88	160.53	171.66		
1HK2 04322 modified	-12.92	-40.91	0.0119	0.0213	3.34	3.04	0.95	0.87	162.08	163.96		
1HK3 04396 H atom away from water	-3.90	-15.73	0.0041	0.0209	3.61	3.04	1.03	0.87	148.26	178.88		

*The naming system employed in the first column is based upon (i) the structure code in the Protein Data Bank (PDB) [48] (e.g. 1HK3), (ii) the number ascribed within the PDB file to the iodine atom which is under investigation for potential halogen bonding (e.g. 04396) and (iii) a comment indicating what modifications, if any, were performed upon the structure beyond truncation and the addition of hydrogen atoms (e.g. translation of atom(s) or substitution of a hydrogen atom with a methyl group).

**For these complexes the halogen bonding type interaction occurs between different atoms in the unoptimised crystal structure and the computationally optimised structure. In the case of the former an iodine – nitrogen contact is identified but upon optimisation this contact is substituted for an iodine – oxygen halogen bond.

In most of the systems described in table 5.1, the halogen bond angle became closer to linear upon geometry optimisation. This is consistent with the presence of a driving force towards optimisation of the halogen bond and is indicative of the presence of a halogen bond that is having a stabilising effect on the structure. Optimisation of the halogen bonding geometry can also be observed in most cases, qualitatively, from comparing the molecular images of the system before and after geometry optimisation.

For the unoptimised structures the VdW ratios are in some cases smaller than unity, which indicates that there could be a *bona fide* halogen bond in the crystal structure and hence potentially also *in vivo*. Although cases wherein the VdW ratio is greater than unity are not consistent with the IUPAC definition of a halogen bond [7], that does not rule out the possibility of a halogen bond type interaction wherein there this an electrostatic attraction between the positively charged σ-hole on the iodine atom and the negatively charged lone pair of electrons on the oxygen or nitrogen atom; furthermore, weak long-range orbital overlap cannot be excluded. The presence of non-zero WBI values indicates some degree of covalency even where the internuclear separation is greater than the sum of the formal VdW radii. It appears that the nitrogen atoms do not form close contacts; the cases wherein the *prima facie* electron donor would be a lone pair of electrons on a nitrogen atom exhibit VdW ratios that are substantially in excess of unity (1.48 in both instances) and extremely low WBI values (0.0006 in both instances). Furthermore, this hypothesised halogen bond does not withstand optimisation, being replaced by a halogen bond comprising the iodine atom and an oxygen atom.

For the optimised structures for which geometry optimisation generally resulted in a reduction in internuclear separation, the VdW ratios are all smaller than unity. However, in the case of "1ICT 07198 Modified CH3" the ratio is 0.99; this is an increase from 0.94 in the unoptimised structure. This finding is consistent with this system not being stabilised by halogen bonding. Where, upon optimisation, the VdW ratio decreases, and the optimised internuclear distance is smaller than unity, the VdW ratios provide compelling evidence for a driving force towards the formation of halogen bonds. This driving force towards halogen bond formation applies even in cases where the influence of competing interactions preclude the presence of formal halogen bonds in the crystal structure; the halogen bond type interaction (short of a formal halogen bond) is likely to have some bearing upon the crystal structure.

For some of the structures, for example 1HK1 04288, the optimised structure is completely (or very close to) linear, that being the ideal halogen bond angle, and combined with an internuclear separation within the sum of the VdW radii, indicates that halogen bonding completely dominates the overall interaction. For other systems, for example 1HK2 04349, while optimisation does drive the geometry to a halogen bond angle that is more linear than in the crystal structure, and places the iodine and oxygen atoms within the sum of VdW radii, it is still far from the ideal halogen bond angle. The optimised structure is also consistent with a contribution from a C-O…N-H hydrogen bonding type interaction, as can be seen from the molecular image; indeed, that is likely to be the strongest interaction in this case.

Optimisation of the geometry does not always lead to the conclusion that there is a halogen bond stabilising the structure. 1ICT 07198 shows the clearest example of where computational geometry optimisation did not maintain, let alone enhance, a halogen bond. In the case of 1HK1 04286, an iodine – nitrogen halogen bond is broken and replaced by an iodine – oxygen halogen bond. This suggests

that optimising the latter is more energetically significant than whatever penalty is paid for breaking the former. However, in the crystal structure, other interactions, which are not present in the isolated system under analysis, must outweigh the factor of which halogen bond is stronger. The most that can be said in the case of this system is that the halogen bond involving nitrogen may have a stabilising effect, which when combined with other interactions helps to produce the crystal structure, but further analysis would be required to have confidence in that conclusion.

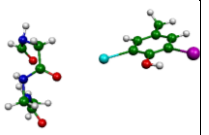
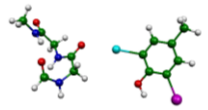
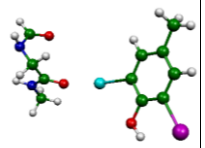
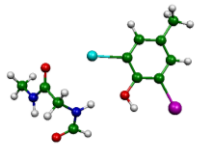
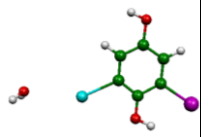
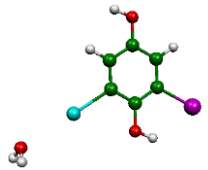
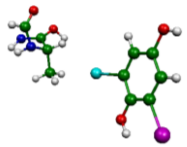
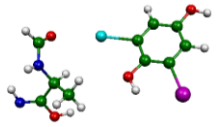
The Wiberg bond index (WBI) values are low, indicating that the interactions are predominantly electrostatic, rather than covalent. Recast in terms of the sigma hole effect, the halogen bonds appear to arise from an electrostatic interaction between the negatively charged lone pairs of electrons on the oxygen atoms and the positively charged, electron deficient σ* orbitals, antibonding with respect to the C-I covalent bonds. Whilst the WBI values are low, they are usually greater in the case of the optimised structure than the corresponding crystal structure, and instances where the contrary prevails arise where the energetic and geometric results are also consistent with the disassociation of the *prima facie* halogen bond, for example in the system identified as "1ICT 07198 Modified CH3" in table 5.1. This increase in the WBI values upon optimisation is consistent with a slight increase in the degree of covalency upon optimisation, although the interaction remains dominated by electrostatic attraction. This increase in covalency is not surprising, as the geometric changes that occur during optimisation, in particular an increase in C-I…O bond angle linearity and a decrease in the I…O internuclear separation, enhance the alignment and proximity of the hereinbefore specified lone pair of electrons and σ* orbital. However, this apparent increase in orbital overlap upon optimisation does provide additional evidence for halogen bonding between iodine and oxygen. Although small, even for the crystal structures, there is some covalency indicated by the WBI results (i.e. they are not zero). This partial covalency is consistent with the weakening of the carbon – iodine covalent bond which Bayse and Rafferty identified as contributing to the cleavage of that bond upon formation of T3 from T4 [151]. However, the finding that the interaction is predominantly electrostatic suggests that halogen bonding's role in facilitating the dissociation of iodine from thyroxine is modest.

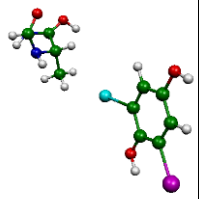
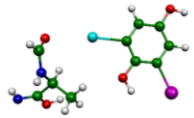
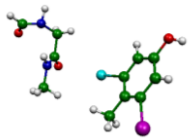
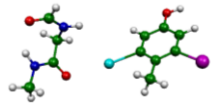
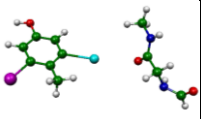
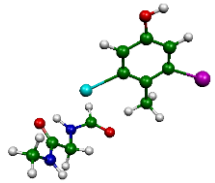
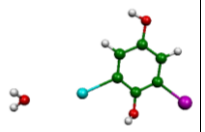
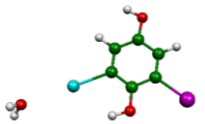
As described above, the unoptimised structures come from empirical observations of the crystal structure, and while the possibility exists that other factors happen to place the thyroxine iodine molecule in the vicinity of the protein backbone or water molecule oxygen atom with a C-I…O angle that is consistent with halogen bonding, the prevalence of these occurrences, confirmed by geometry optimisation that indicates that the geometry is in the potential well associated with halogen bonding, is indicative of halogen bonding having a *bona fide* role in determining the crystal structure.

5.4.2 Results from the astatine based system

Table 5.2 is analogous to table 5.1 but relates to systems wherein the iodine atom that is engaged (or potentially engaged) in halogen bonding has been replaced by an astatine molecule. The purpose of calculating these results is so that they may be compared with those which were found for the natural iodine based system. Astatine forms stronger halogen bonds than iodine [16], hence a consistent finding that the crystal structure (unoptimised) geometry has a more negative complexation energy when iodine is replaced by astatine would be supportive of the conclusion that halogen bonding helps to stabilise those structures, as that would be consistent with the effect becoming stronger when the halogen bonding capacity of the halogen increases. The unoptimised ("crystal") structure, by design, maintains the same geometry as for the iodine based system, but differences in the structure of the optimised systems as well as the optimised complexation energies are also the subject of analysis. While astatine forms stronger halogen bonds and hence may be expected to more strongly draw the halogen bond angle towards that which is optimal for halogen bonding, out-competing rival interaction, it is also true that the halogen sigma hole on an astatine atom covers a wider cone angle than is the case for iodine and hence is able to accommodate a greater departure from linearity where that is conducive to facilitating the maintaining of other stabilising interactions [16].

Table 5.2: Energetic, geometric, WBI and graphical results for both the crystal and optimised, astatine-based, structures.

Name of system*	ΔE^{unopt} kJ/mol	ΔE^{opt} kJ/mol	WBI^{unopt}	WBI^{opt}	R^{unopt} Å	R^{opt} Å	VdW $ratio^{unopt}$	VdW $ratio^{opt}$	α^{unopt} deg.	α^{opt} deg.	Unoptimised structure	Optimised structure
1HK1 04309	-15.06	-28.15	0.0132	0.0318	3.39	2.94	0.96	0.83	169.90	176.40		
1HK3 04345 modified	-17.62	-45.79	0.0144	0.0283	3.48	2.99	0.98	0.84	168.72	162.80		
1HK1 04288 modified H moved	-8.66	-21.93	0.0079	0.0273	3.56	2.97	1.01	0.84	150.04	179.60		
1HK1 04286 Modified NH geom. Edited**	-1.32	-31.39	0.0006	0.0242	5.24	3.01	1.47	0.84	159.76	171.57		

1HK1 04286 Modified **	-0.50	-31.39	0.0007	0.0242	5.24	3.01	1.47	0.84	159.76	171.57		
1HK5 04594 Modified CH3	-10.23	-33.72	0.0010	0.0327	4.40	2.95	1.24	0.83	171.72	174.39		
1ICT 07198 Modified CH3	-16.18	-26.87	0.0121	0.0155	3.30	3.16	0.93	0.89	165.14	169.95		
1HK3 04396 Modified	-9.01	-24.59	0.0059	0.0243	3.57	2.96	1.01	0.84	149.79	179.21		

1ETA 02040 Modified CH3	-25.09	-28.44	0.0199	0.0285	3.12	2.93	0.88	0.83	159.51	178.16
1HK2 04322 Modified	-18.42	-47.09	0.0171	0.0293	3.34	2.98	0.94	0.84	161.04	163.90
1HK3 04396 Modified H atom away from water	-7.37	-21.93	0.0054	0.0274	3.61	2.92	1.02	0.82	149.79	179.37

*The naming system employed in the first column is based upon (i) the structure code in the Protein Data Bank (PDB) [48] (e.g. 1HK3), (ii) the number ascribed within the PDB file to the iodine atom which has been substituted with an astatine atom and which is under investigation for potential halogen bonding (e.g. 04396) and (iii) a comment indicating what modifications, if any, were performed upon the structure beyond truncation and the addition of hydrogen atoms (e.g. translation of atom(s) or substitution of a hydrogen atom with a methyl group).

**For these complexes the halogen bonding type interaction occurs between different atoms in the unoptimised crystal structure and the computationally optimised structure. In the case of the former an astatine – nitrogen contact is identified but upon optimisation this contact is substituted for an astatine – oxygen halogen bond.

5.4.3 Comparison between the iodine and astatine based systems

Qualitatively, the results for the astatine-containing systems are analogous to those of the iodine based systems, as found *in vivo*. This finding is consistent with astatine behaving in qualitatively the same fashion as iodine in relation to its non-covalent interactions, as found to be the case in the other systems that have been investigated and reported in this thesis.

Table 5.3 sets forth the quantitative differences between the two different sets of systems. In this representation all values were obtained by subtracting the energetic and geometric values of the iodine based system from the corresponding astatine based system. Both crystal and computationally optimised structures are included in table 5.3. Whereas tables 5.1 and 5.2 disclose all of the structures where there appeared to be a halogen bond in the crystal structure which made a significant contribution to the stability of the geometry of the crystal structure, table 5.3 only discloses cases where results were obtained for the equivalent systems for both the iodine based and the astatine based cases.

The quantitative differences between the two sets of optimised halogenated systems can be explained by, and is consistent with, astatine having stronger halogen bonding properties compared with iodine, combined with the former also having a larger Van der Waals radius than the latter. As astatine has been established, both elsewhere herein and in the existing literature [8], to form stronger halogen bonds than iodine, the observation that, for all cases, the astatine based system has a more strongly negative interaction energy for the optimised structure than its corresponding iodine based analogue supports the conclusion that halogen bonding does, at least, play a stabilising role in most of the molecular systems that are presented herein. With one, plainly anomalous exception (1HK3 4396 with the hydrogen atom pointing away from the water molecule in the initial input geometry), the differences in the interaction energies between the iodine and astatine based systems are all in the range of -4.33 kJ mol^{-1} to -6.84 kJ mol^{-1} (the negative values indicating that the astatinated system is more stable than the corresponding iodinated analogue). This is a modest range and probably reflects the halogen type dependency of the strength of the halogen bond being quite consistent between the various systems, while varying to a limited degree between different chemical environments.

By contrast there is not the same degree of consistency for the unoptimised crystal structures. The range is +1.15 kJ mol^{-1} to -6.01 kJ mol^{-1}, with two cases of the crystal structure (which naturally contain iodine) becoming less energetically favourable upon substitution with astatine. This diminution in consistency can probably best be explained by the multiplicity of factors that contribute to the overall energetic favourability of a given geometry. The overall geometry of the molecular structure within the crystal, not just the distance and angle of the halogen – oxygen contact, is impacted by the identity of the halogen that is present. Furthermore, the crystallographically observed geometry for the iodine-based system can be expected be closer to optimal than for the system containing a species that is not present in the crystals upon which the crystallography was performed, i.e. the astatine-based systems. Hence replacing the iodine

with astatine without allowing the geometry to adjust for the change in halogen species can be expected to incur an energetic penalty on top of impeding the system from taking full advantage of the species that may inherently produce the stronger halogen bond (by not allowing the halogen bond to relax to its optimal distance). The only accommodation that was made was to extend the C-X covalent halogen bond to prevent it from becoming too compressed; this small adjustment also has a minor impact on the relative positions of the two molecular fragments. The observation that the majority of the systems nonetheless become more energetically stable *in totum*, and that in the two cases where the converse is true the magnitude of the difference in complexation energies is small (1.15 kJ mol^{-1} and 0.28 kJ mol^{-1}), is likely to reflect the substantially stronger halogen bonding ability of astatine in comparison with iodine.

Regarding the geometric differences for the halogen bond, when comparing the iodinated and corresponding astatinated molecular systems, the internuclear separation is the same for the unoptimised crystal structures. This consistency is by design and there is nothing further to be said upon the matter. However, one can make the further observation that for the computationally optimised structures, the internuclear separation decreases upon substitution of iodine by astatine, by between 0.06 Å and 0.31 Å (although in only one instance beyond 0.11 Å). This contraction in internuclear separation occurs despite astatine having a greater VdW radius than iodine; this finding is consistent with the observations made in the preceding chapters herein concerning the relative differences in halogen bond length for iodine and astatine in otherwise identical systems, and lends further support to the hypothesis that the halogen bond has a significant influence upon the molecular geometry. As for the internuclear separation, nothing should be inferred from differences in halogen bond angles for the unoptimised systems. The results for the optimised system show a mixed picture: some cases wherein the halogen bond angle increases (becomes more linear) upon substitution of iodine with astatine, and other cases where the converse is true. As discussed elsewhere herein, the impact of descending the halogen group upon halogen bond angle can be complex. Whereas on the one hand the greater energetic importance of the halogen bond should favour a more linear halogen bond, on the other hand the greater surface coverage of the sigma hole on the heavier atom (see molecular electrostatic plots for halobenzene in Chapter 3) allows for easier accommodation of secondary stabilising factors. With two exceptions, the starting geometry afforded a less linear halogen bond angle for the astatine analogue. However, for all cases where the difference in halogen bond angle, for the optimised structure, was greater than one degree, it was in favour of astatine forming the more linear halogen bonds. This observation suggests that where the dominant factor was the greater relative contribution of the halogen bond to the overall stability of the molecular system, this had a greater net impact on halogen bond angle, compared with cases where the more significant factor was the ability of a heavier halogen atom to accommodate a greater departure from the ideal linear halogen bond angle.

Observations regarding the VdW ratios for the astatinated system are, in general, qualitatively similar to those systems wherein iodine is the only halogen element present. The similarity extends to the observations made in relation to interactions between the halogen atom and a nitrogen atom. However, it can be noted that for astatinated “1ICT 07198 Modified CH3”,

optimisation does produce a modest reduction in VdW ratio (from 0.93 to 0.89). In the case of unoptimised astatinated "1HK5 04594 Modified CH3" the VdW ratio of 1.24 is significantly greater than unity. The VdW ratios are usually quite similar for the two halogen elements under investigation, with astatinated systems having slightly smaller VdW ratios. This latter observation is also consistent with astatine forming stronger halogen bonds than iodine. Although not explicitly presented, the disclosure in table 5.3 that the internuclear separation decreases upon substitution of iodine with astatine combined with the knowledge that astatine has a greater VdW radius (2.02 Å [136]) than iodine (1.98 Å [135]), it is *a fortiori* that the VdW ratios are smaller for astatine than for iodine.

In all cases wherein the halogen bond is constituted between an oxygen atom and a halogen atom, the astatinated system has a greater WBI value than for the corresponding iodine based system, both for the optimised and unoptimised crystal structures. The number of instances wherein the donor of electron density is nitrogen is very small, limiting the scope for observing trends between systems wherein a nitrogen lone pair of electrons is engaged. These cases that were identified from the crystal structures are included in the same data set as, and compared with, cases where the electron density donor is an oxygen atom. However, it should be noted that this introduces an additional variable, the electronic properties of the electron density donor, in the context of a study where the objective is to compare the properties of the electron density acceptor (iodine or astatine). Therefore, even if all of the observations pertaining to the nitrogen cases were fully consistent with the trends that are observed in the case of oxygen atoms donating the electron density, caution should be exercised when extrapolating any trends beyond cases where the electron density donor is an oxygen atom. The effect of optimising the crystal structure is found to invariably (for the systems observed) accentuate the increase in WBI value for the astatinated analogue in comparison with the elementally unadulterated crystal structure. This finding suggests that the halogen bond has a greater dative covalent character (albeit remaining predominantly electrostatic as reflected by the still small WBI values) where the participating halogen atom is astatine, i.e. the heavier halogen element. The hereinbefore described accentuation of this effect upon computational optimisation can be explained by the hereinbefore discussed contraction in the internuclear separation between the halogen atom and the oxygen atom. The reduction in internuclear separation affords a greater degree of overlap between the electron rich orbital corresponding to the lone pair of electrons on the oxygen atom and the electron deficient σ* orbital on the halogen atom (antibonding with respect to the carbon – halogen covalent bond).

Table 5.3: Energetic and geometric differences between the iodine and astatine based systems. Values were calculated by subtracting the energetic and geometric values for the iodine based system from those of its astatine based analogue.

Structure	$\Delta\Delta E^{unopt}$ kJ/mol	$\Delta\Delta E^{opt}$ kJ/mol	ΔWBI^{unopt}	ΔWBI^{opt}	ΔR^{opt} Å	$\Delta\alpha^{unopt}$ deg.	$\Delta\alpha^{opt}$ deg.
1HK1 modified H atom moved 4288	-2.19	-6.21	0.0016	0.0064	-0.07	-1.59	-0.4
1ETA modified CH3 2040	-6.01	-6.84	0.0051	0.0091	-0.1	-1.5	5.66
1HK1 4309	-3.61	-6.36	0.0028	0.0097	-0.08	0	4.29
1HK3 modified 4345	-5.17	-5.76	0.0043	0.008	-0.07	-0.58	-0.99
1ICT Modified CH3 7198	-3.68	-4.33	0.0029	0.0097	-0.31	-0.94	24.8
1HK3 modified 4396	-2.12	-6.05	0.0012	0.0065	-0.06	-1.44	-0.69
1HK1 modified NH geom. Edited* 4286	0.28	-5.58	0	0.0086	-0.1	-0.77	-0.09
1HK1 Modified* 4286	1.15	-5.58	0.0001	0.0086	-0.1	-0.77	-0.09
1HK2 modified 4322	-5.5	-6.18	0.0052	0.008	-0.06	-1.04	-0.06
1HK3 H atom away from water 4396	-3.47	-6.20	0.0013	0.0065	-0.12	1.53	0.49

*For these complexes the halogen bonding type interaction occurs between different atoms in the unoptimised crystal structure and the computationally optimised structure. In the case of the former an iodine – nitrogen or astatine – nitrogen contact is identified but upon optimisation this contact is substituted for an iodine – nitrogen or astatine – oxygen halogen bond.

5.4.4 Correlation between internuclear separation and complexation energies for iodine and astatine based systems

Figure 5.1 discloses a weak correlation between internuclear separation and complexation energy for the unoptimised crystal structures. Furthermore, it shows that interaction energies are consistently more strongly negative for the astatine based systems than the corresponding iodine based systems. This effect of substituting iodine with astatine is shown to become less pronounced as internuclear separation increases. Whilst an increase in internuclear separation would be expected to arise from an increase in intermolecular separation and hence a decrease in the magnitude of the complexation energy, confirmation of this outcome is at the very least not inconsistent with a finding that halogen bonding contributes to the overall stability. The finding that astatine forms stronger complexes, especially at typical halogen bond distances, lends further weight to this hypothesis, given that astatine was found to continue the trend towards increased halogen bonding going down the halogen group [8]. Care is urged in interpreting the rightmost point for each halogen type as this corresponds to a halogen – nitrogen contact whereas all other points on each profile correspond to a halogen – oxygen contact.

It is clear from observing figure 5.1 that this correlation is weak, with cases of increasing complexation energy with an increase in internuclear separation. This substantial deviation from perfect correlation is not unexpected, as each point on each given profile corresponds to a different molecular system, each with its own set of interatomic interactions that in aggregate constitute the overall intermolecular interaction and determine the intermolecular interaction energy. The complexation between thyroxine and either a section of the protein backbone or, to a much lesser extent, a water molecule within the crystal lattice, comprises multiple components of which halogen bonding appears to be a significant one. However, variation in other aspects of the chemical environment do, as demonstrated in figure 5.1, in some cases outweigh the effect of reducing the halogen – oxygen (or halogen – nitrogen) distance. As aforesaid the, at least qualitatively, consistent effect of substituting iodine with astatine does however support the hypothesis that halogen bonding is involved to some extent in stabilising all of the structures represented in figure 5.1.

Due to the weakness of the correlation and the complexity and inconsistency in the chemical environments, figure 5.1 would not be very informative in isolation. However, it is consistent with the other results and observations that have been set forth above, especially as the weakness in the inconsistency in the trend towards weaker complexation at greater internuclear separation can be readily explained by the differences between the systems, not related to halogen bonding.

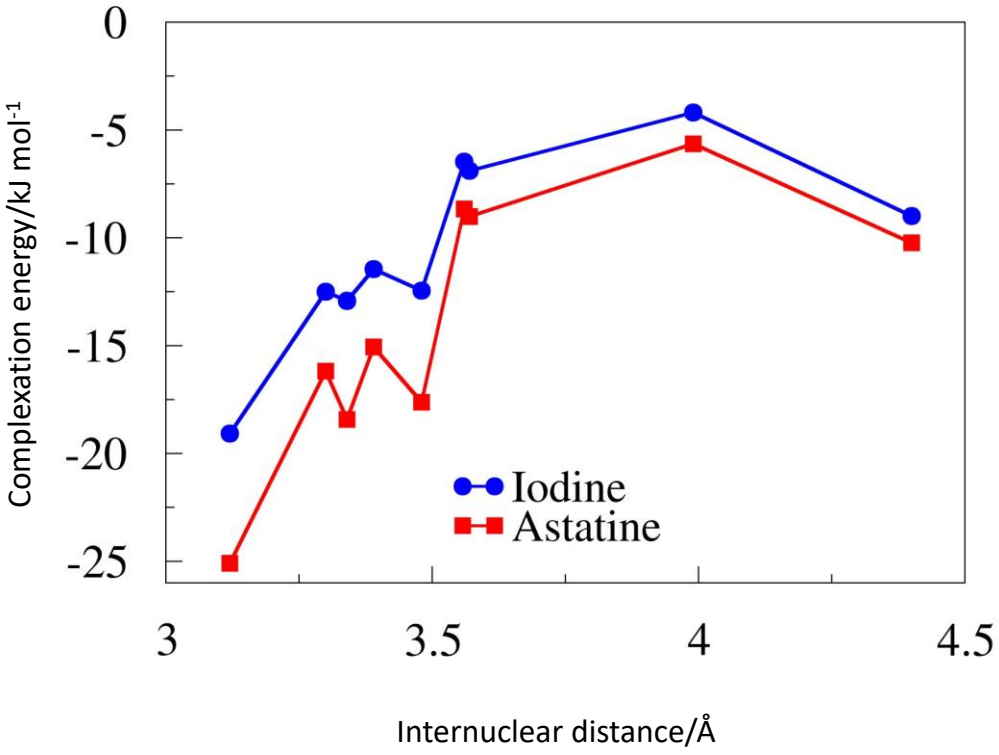


Figure 5.1: Profile showing interaction energy as a function of internuclear distance for iodinated and astatinated dimer systems (selected unoptimised crystal structures)

5.4.5 Statistical analysis of halogen bonding from the Cambridge Structural Database

Figure 5.2 exhibits the results of statistical analyses of the frequency of structures in the CSD [80] wherein a halogen atom (X) is covalently bound to a molecular fragment (R), has a proximity to an atom which is a member of periodic group 15 or 16 (Y) that is within the sum of VdW radii, and wherein the R-X•••Y angle is at least 160°. Figure 5.2 also presents the dependence of the frequency of structures upon internuclear separation and halogen bond angle. The respective identity of X is F, Cl, Br, or I in figures 5.2a, 5.2b, 5.2c and 5.2d. A search of the CSD where the halogen is astatine yielded a result of zero structures. The sample sizes were 348. 1757, 1671 and 1391 structures for the systems containing fluorine, chlorine, bromine and iodine respectively.

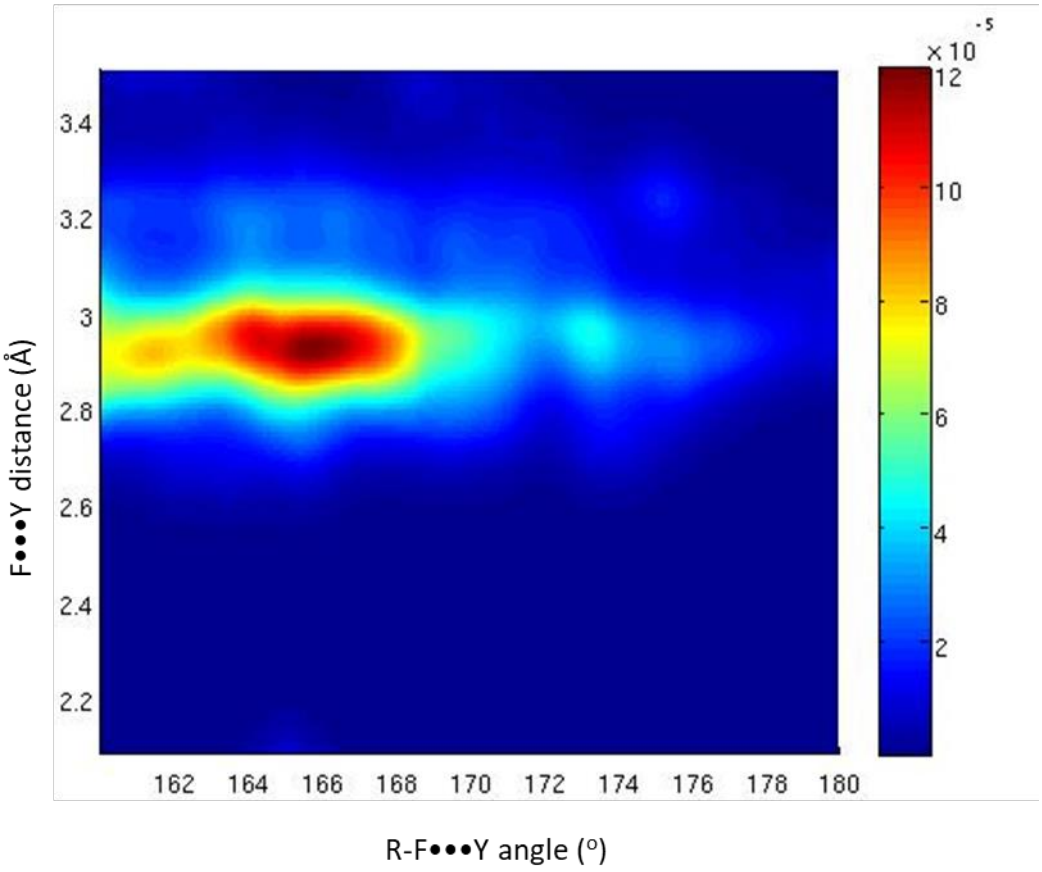


Figure 5.2a: Frequency of R-F•••Y structures in the CSD as a function of halogen bond distance and angle.

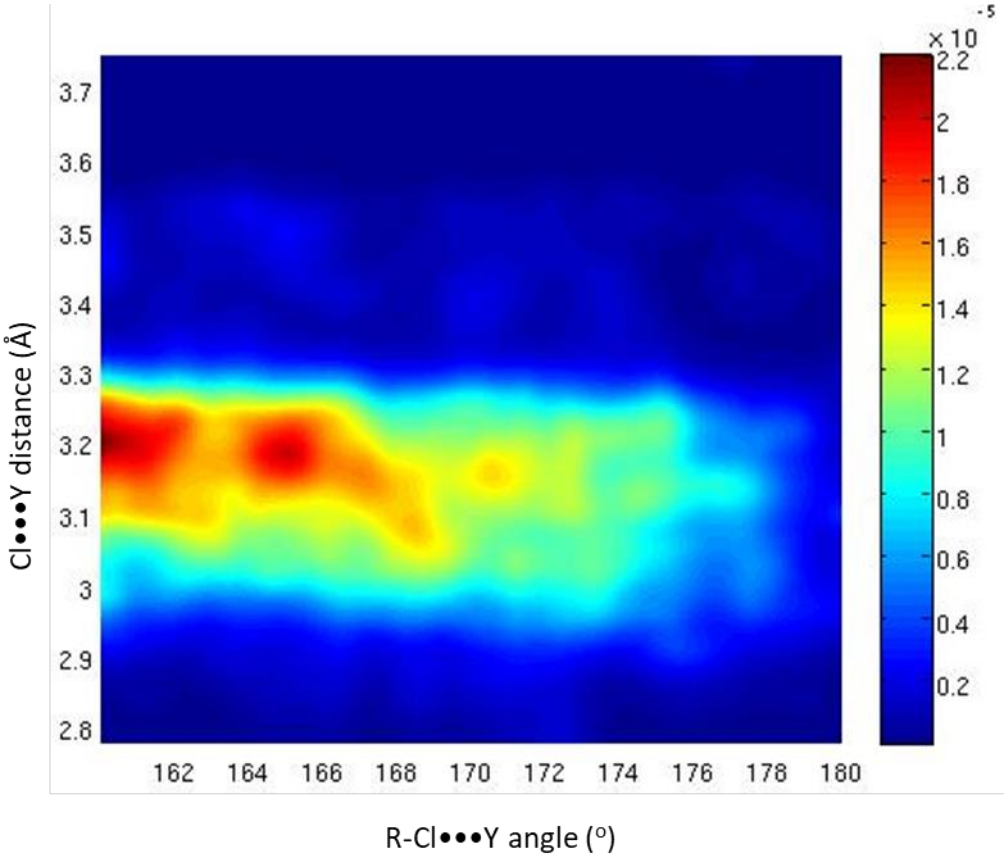


Figure 5.2b: Frequency of R-Cl•••Y structures in the CSD as a function of halogen bond distance and angle.

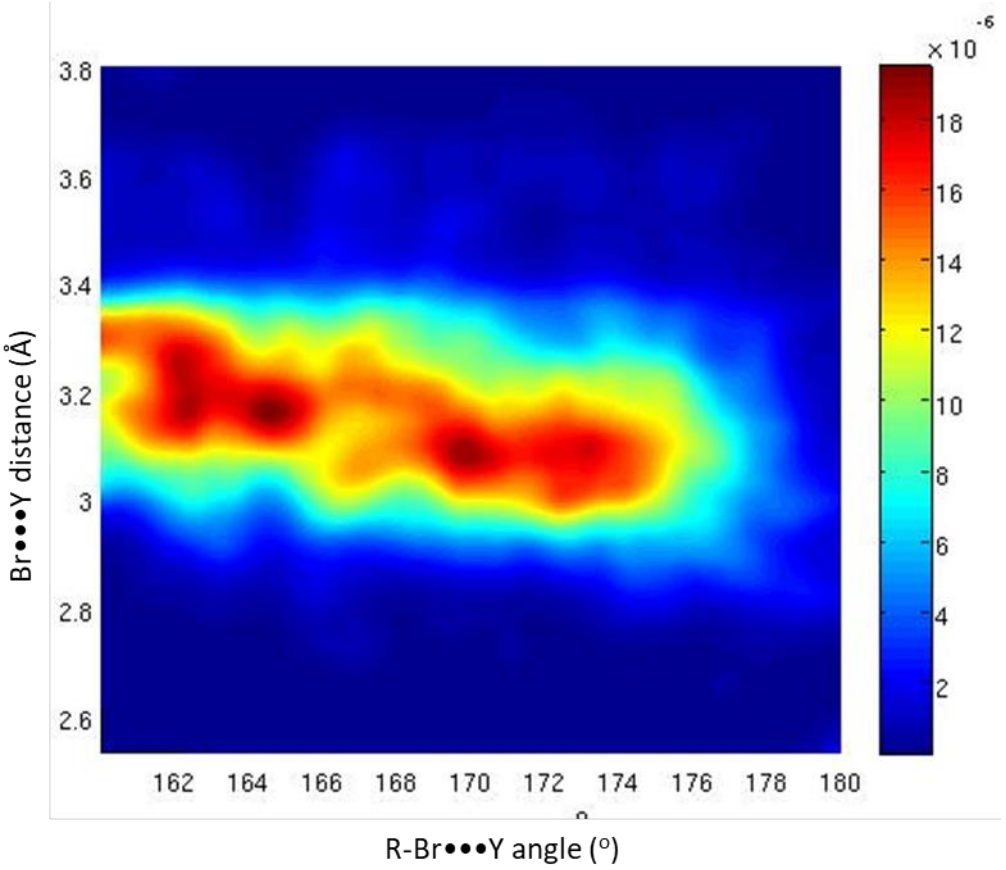


Figure 5.2c: Frequency of R-Br•••Y structures in the CSD as a function of halogen bond distance and angle.

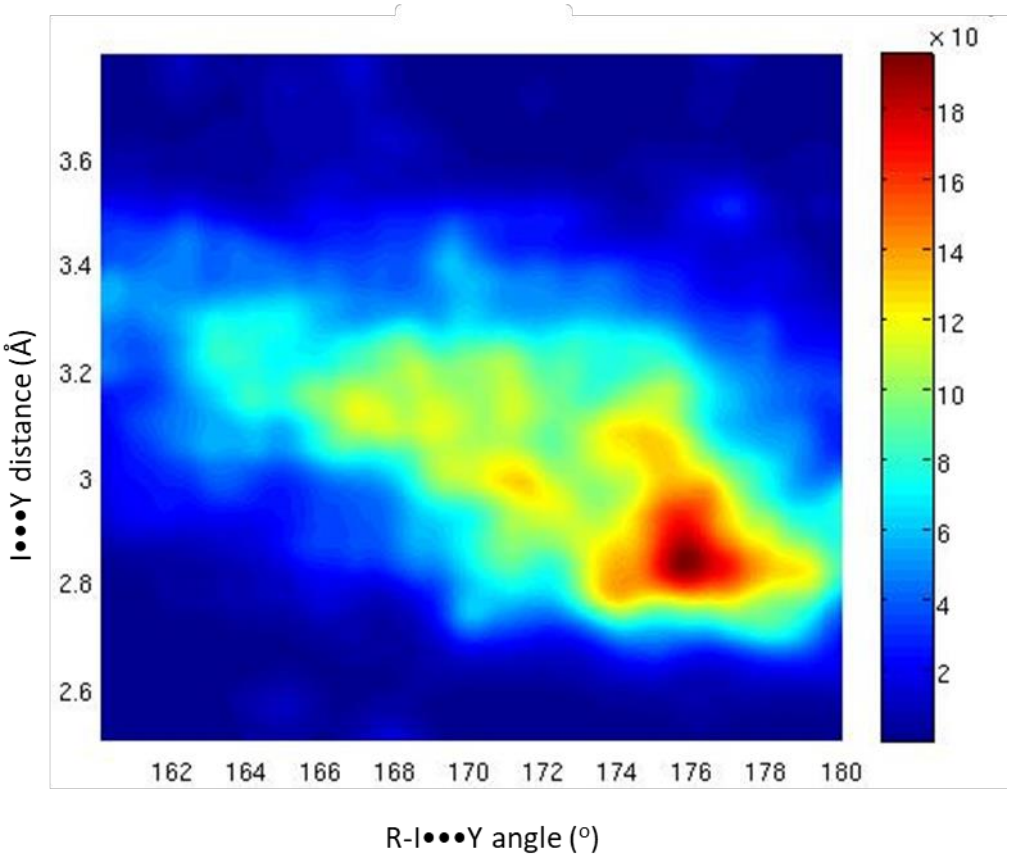


Figure 5.2d: Frequency of R-I•••Y structures in the CSD as a function of halogen bond distance and angle.

The heatmaps set forth in figure 5.2 disclose the frequency of the structures that conform to the hereinbefore specified geometric criteria as a function of both halogen bond distance and halogen bond angle. Halogen bonds have a preference for linear geometries [16], so a correlation between short X•••Y contacts and near-linear R-X•••Y bond angles could be expected if a substantial proportion of the structures in the statistical sample are cases wherein the halogen atom is halogen bonded to the pnictogen or chalcogen atom. This trend can be clearly observed where the halogen atom is iodine (figure 5.2d), with a high concentration of structures in the region approximately centred on 2.9 Å and 176°. The trend perhaps also occurs very weakly in the case of bromine (figure 5.2c).

Where the halogen atom is chlorine or bromine (figures 5.2b and 2c respectively) there are multiple centres of high concentration of structures with different angles but similar internuclear distance. For chlorine this region is centred around 3.2 Å, while for bromine it is centred around the range of 3.0 Å to 3.2 Å, with the hereinbefore mentioned weak trend towards shorter distances for the more linear angles. The reason for this distribution has not been resolved, however, it is worth noting that it is much less pronounced for the case where halogen is fluorine (figure 5.2a) or where it is iodine.

The presence of a strong correlation between shorter internuclear distance and more linear halogen bond angle in the case of iodine, but not in the cases of the lighter halogen elements indicates that only in the case of iodine is the sample of surveyed structures dominated by cases wherein a halogen bond is present. This would be consistent with the previously observed [16] tendency for iodine to form stronger halogen bonds, in comparison with the lighter halogen elements.

The lack of a similarly clear correlation in the cases of bromine and chlorine despite their known ability to form halogen bonds [16] could be a result of there being a relatively large number of structures wherein these elements are present but whose geometry is not consistent with the presence of halogen bonds (e.g. a C – Br bond pointing into empty space in the carbon to bromine direction). It is plausible that the samples for these elements do contain a proportion of halogen bonded structures but they are insufficiently numerous to clearly affect the qualitative results in this statistical analysis. The presence of a weak (and not as clear) correlation observable the case of bromine, which is known [16] to have a greater propensity towards halogen bond formation that chlorine, perhaps lends some support to this hypothesis.

5.5 Conclusions

The present study supports the prior finding by Auffinger *et al.* [15] that there is halogen bonding between iodine on thyroxine and lone pairs of electrons on oxygen atoms in the thyroid system. Furthermore it is at least partially consistent with the role of halogen bonding in facilitating the cleavage of iodine from thyroxine to form T3 from T4 as identified by Bayse and Rafferty [151].

With some rare exceptions, the systems that were investigated were found to optimise towards ideal halogen bonding geometry when subjected to geometry optimisation in the absence of the remainder of the molecular system from which the species had been extracted. This observation is likely to be a reflection of the crystal geometry being influenced by a complex set of interactions of which halogen bonding is one. In this study the other interactions that could influence the geometry, for example hydrogen bonds, were as far as practicable eliminated from the model so that the effect of halogen bonding could be investigated in isolation. Quantifying the relative significance of the halogen bonds in comparison to the other contributing interactions was outwith the scope of the present study, not least due to the computational resources that would be required to carry out that exercise. Although quantitative results have been presented herein, this investigation into the role of halogen bonding in the thyroid was predominantly qualitative at least as far as conclusions about the system are concerned.

Table 5.3 shows that in all instances listed therein, the complexation energy is more strongly negative in the astatinated cases. The same table also shows that this trend is true in most instances for the unoptimised crystal geometries. This finding is consistent with there genuinely being a role played by halogen bonding as astatine would be expected to form stronger halogen bonds. The greater consistency of this trend for the optimised cases is probably a reflection of the elimination of competing interactions giving a greater role to the halogen in contributing to the overall stability of the system. Figure 5.1 indicates that the greater stability of astatinated complexes decreases with internuclear distance. This finding is consistent with the contribution from halogen bonding decreasing as internuclear distance increases, as astatination is only hypothesised to increase the strength of the halogen bond component of the overall complexation energy.

Part of the definition of a formal halogen bond is that the internuclear separation must be shorter than the sum of the VdW radii [7]. For the optimised structures, both for iodine based and astatine based systems, this criterion is satisfied. However, as hereinbefore discussed, not all systems formed halogen bonds, even after optimisation. For the unoptimised crystal structures, the VdW ratio was less than unity in some cases, and in many other cases this ratio was only slightly greater than unity. There were however some cases, especially those involving a nitrogen atom, wherein the VdW ratio was substantially greater than unity, indicating that

interaction between the σ-hole on the halogen atom and the lone pair of electrons on the nitrogen or oxygen atom was at best minimal.

The WBI values indicate that the interactions are predominantly, but not exclusively, electrostatic rather than covalent. It is the partial covalency of the interaction that entails electron donation into the σ* antibonding orbital, that weakens the carbon – iodine bond, discussed by Bayse and Rafferty [151]. Hence the WBI data obtained in the present investigation support the conclusion that this donation and hence covalent bond weakening is real, but the degree of weakening of the bond is likely to be modest.

The statistical analysis of the CSD [80] shows that in the case of iodine, and also to a much lesser extent bromine, there is a correlation between short internuclear distances and more linear halogen bond angles. From these results, it appears that, where the halogen atom is iodine, the structures in the surveyed sample are dominated by cases where there is a halogen bond between the iodine atom and the pnictogen or chalcogen atom. The lighter halogen elements exhibit a markedly different distribution pattern, which is consistent with most of the structures lacking halogen bonds. However, bromine exhibits a weak correlation of between short internuclear separation and more linear bond angle, perhaps indicating that a greater proportion of the structures contain halogen bonds in the case of bromine in comparison with chlorine and fluorine. This trend is consistent with a trend towards stronger halogen bonding as the halogen group is descended, with iodine appearing to demonstrate a much greater propensity to be halogen bonded than any of the lighter halogen elements.

6 Halogen bonding between halogenated DNA base pairs

6.1 Abstract

An investigation was performed on the effects of halogenation upon the DNA base pairs adenine – thymine and guanine – cytosine. In particular the effects upon interaction energy and geometry were investigated. The covalency of halogen bonds and hydrogen bonds was also examined. This study was inspired by a paper authored by Parker *et al.* [A.J. Parker, J. Stewart, K.J. Donald, C.A. Parish, J. Am. Chem. Soc. 134 (2012) 5165-5172] in this field.

In the present study fluorination and astatination as well as chlorination, bromination and iodination were investigated. In all computations, the M06-2X functional was used in conjunction with either the 6-31+G* basis set, except for the iodine and astatine atoms, for which the same functional was employed in conjunction with the aug-cc-pVDZ-PP functional with relativistic pseudopotentials, to accommodate relativistic effects.

The results obtained in the course of this study demonstrate the sensitivity of halogen bonding to the wider chemical environment, especially the electron withdrawing or electron donating properties of other substituents. However, the principal factor in determining the strength of halogen bonding was found to be the identity of the halogen atom, with the well-established trend towards stronger halogen bonding as the halogen group is descended being confirmed. There was also, at least in some cases, a positive correlation between greater covalency and atomic number of the halogen atom. Steric effects were found to be significant in some cases, but the extent of their importance appears in this study to be not as great as that which was found by Parker *et al*. Interaction energy profiles, computed by constraining the halogen bond angle, demonstrated, in the case of the adenine – thymine dimer the greater angle dependence for halogen bonding than for hydrogen bonding; however this attribute was not clearly disclosed by the equivalent profiles for the guanine – cytosine pair. This study shows that in the cases of iodination, and especially astatination, halogen bonds can in some cases produce a comparable or even greater overall degree of stability than that which is achieved in the non-halogenated cases.

6.2 Introduction

A paper authored by Parker *et al.*[35] disclosed a study wherein each of the DNA bases were modified by the substitution of hydrogen atoms that conventionally engage in hydrogen bonding with an oxygen or nitrogen atom on the partner base molecule (e.g. between adenine and thymine and between guanine and cytosine), with a halogen atom. The terms dA:dT and dG:dC shall mean adenine-thymine pair and guanine-cytosine pair respectively (as per the hereinbefore referenced paper), unless otherwise stated this definition shall include both halogenated and non-halogenated variants thereof. Where appropriate, the identity of the relevant substituent(s) are identified by prefixing these terms by the corresponding chemical symbol of the monatomic substituent(s). In the case of dA:dT, either one or two of the hydrogen atoms typically engaged in hydrogen bonding were substituted by halogens. In the case of dG:dC single, double and triple substitution by halogen were investigated. Where multiple hydrogen atoms were substituted by halogen atoms, all of the halogen atoms were of the same element. Only substitution by each of chlorine, bromine or iodine was probed. The B3LYP density functional [157-159] was employed in conjuction with the 6-31G* basis set [160] except for iodine for with the MDF effective core potential was used in conjuction with the corresponding basis set [161]. For further details of this prior study, see reference [35].

In general, the strength of halogen bonding increases going down the halogen group[16]. However Parker *et al.*[35] found that steric effects in the congested environment of a non-covalently interacting pair of DNA bases could give rise to a reversal in the order of stability, due to the greater size of iodine atoms compared with bromine atoms[35]. There can be multiple non-covalent interactions as well as steric hindrance that affect the stability and geometry of the halogenated DNA base pairs.

Ideally, halogen bonds are linear due to the origins of halogen bonds[11]. However, it is known that in complex chemical environments substantial departure from linearity can be tolerated[15]. A theoretical insight into non-linear halogen bonding has been undertaken by Hill and Hu[8]. Hill and Legon[30] have authored a paper based upon theoretical and experimental observations of this flexibility. The ability of the halogen atoms to accommodate this departure from the ideal angle could impact upon the stability of the DNA base pair.

Kolar and Tabarrini have authored a paper [41], published in the Journal of Medicinal Chemistry, based upon an investigation into halogen bonds involving nucleic acids. Their work was intended to advance the employment of halogen bonds to halogenated nucleic acids in drug design. Their work was based upon a statistical survey of the Protein Databank (PDB)[48]. Kolar and Tabarrini investigated the abundance of halogen contacts between halogenated nucleic acids and electron rich moieties, including oxygen anions, oxygen atoms, nitrogen atoms and phosphorus atoms. Only neutral chlorine, bromine and iodine were included as halogen bond donors. The search criteria included filters to ensure adequate quality of the X-ray samples and did not include structures obtained by NMR. The geometric search criteria provided that the halogen bond angle must be no less than 120^{o} and the internuclear distance no greater than the sum of

Van der Waals (VdW) radii. There was also a second search in which the latter criterion was replaced by a cut-off of 4 Å in order to expand the size of the data set, from 21 to 77. The authors found that there was a large bias in the PDB in favour of chlorinated species. However, among halogen bond participants (in the stricter sense) the ratio was 2:17:2 for chlorine, bromine and iodine respectively. By contrast the respective total counts in the PDB search were 402, 204 and 66 for chlorine, bromine and iodine, respectively. The authors found that there was a statistical correlation between increasing atomic number of the halogen and the halogen bond angle (i.e. the heavier halogens formed more linear halogen bonds). The greatest angle for a halogen bond involving chlorine was 154°, whereas both of the cases involving iodine has a halogen bond angle in excess of 170°. From these angular results the authors concluded that chlorine based halogen bonds were unlikely to be the primary factor in determining the geometry. They also identified an inverse correlation between halogen atomic number and internuclear distance. The authors did however acknowledge that the small sample size limited the strength of the conclusions that could be drawn, and they recommend performing further research into the use of halogen bonding in the design of drugs, especially where the halogen involved would be bromine or iodine. The authors also cite prior research[13, 162] in support of the proposition that fluorination (i.e. the addition of a strongly electron withdrawing substituent) can accentuate the σ-hole effect on the halogen atom that participates in the halogen bond, while acknowledging that this effect is not invariably simple.

The present investigation expands upon the work undertaken by Parker *et al.*[35] by including substitution with fluorine and astatine, computing the interaction energy profile as a function of bond angle, and performing calculations on systems wherein one of the substituents is astatine and the other halogen element is varied. Developments in computational methods and computational hardware also enable this study to be conducted using the M06-2X[70] density functional, which has been found to perform well for computations involving halogen bonds[22]. This more accurate computational method for treating halogen bonds enabled a revisiting the question of how geometric considerations and steric hindrance affects the relative energies of halogen bonds with different halogen atoms.

6.3 Methods

All calculations were performed using the Gaussian 09 software package [92]. Graphical molecular building, editing and visualisation were performed in Gaussview 4.1 [122]. All calculations were performed using the M06-2X density functional [70]. For all elements lighter than iodine the 6-31+G* basis set was employed. For iodine and astatine, the aug-cc-pVDZ-PP basis set with corresponding relativistic effective core potentials was employed to take into account relativistic effects [64]. Basis set superposition error was corrected by employing the counterpoise correction procedure proposed by Boys and Bernardi [74]. For some of the systems, relaxed scan calculations were performed to establish the relationship between interaction energy and halogen bond angle. These scans were computed using the "ModRedundant" option within the "Opt" keyword in Gaussian 09 [92]. Scan step sizes were set to 5°. In some systems, one of the two molecules contain iodine and/or astatine, while the other molecule does not contain any iodine or astatine. As aforesaid, only those heavier elements were treated using the aug-cc-pVDZ-PP basis set. By default, Gaussian 09 [92] represents d-orbitals in the 6-31+G* basis using Cartesian functions (6 d-functions) unless the "gen" keyword is employed, in which case the d-orbitals for all atoms in the system are represented by spherical harmonic functions (5 d-functions). To ensure consistency, in calculations performed on isolated molecules containing no iodine or astatine atoms, spherical harmonic d-orbitals were employed by the use of the keyword "5D". The status of each stationary point (minimum) was verified by harmonic frequency calculation. All calculations designed to obtain energetic and/or geometric data were performed using Gaussian 09's ultrafine grid. However, it was observed that the ultrafine grid sometimes gave rise to a phantom imaginary frequency, which became real when a single point calculation was performed on the optimised geometry using Gaussian 09's superfine grid. Supplementary calculation confirmed varying the fineness of the integration grid resulted in negligible changes in the computed energetic and geometric results.

6.4 Results and discussion

Figure 6.1 sets forth the numbering systems that have been used for the atoms in each of the molecules studied in this investigation. This numbering system is *sui generis* and may depart from conventional numbering systems. References to atom numbers in this chapter refer exclusively to the numbering system set forth herein. If the atoms within the molecules were numbered on a conventional basis, this would produce inconsistency between the different molecular systems due to the influence of high priority substituents. Furthermore, it is optimal, in the interests of minimising convolution, to assign numbers directly to the substituents that are engaged in halogen bonding, rather than by reference to the ring carbon or nitrogen atom to which they are covalently bound. Whereas in subsequent figures, each halogen atom and hydrogen atom are denoted by a unique colour, in figure 6.1 atoms that are in some (but not all) instances halogen are shown in purple, and denoted "X". In specific systems, where appropriate, the chemical symbol of the element that is present in the instant system is used instead of the generic symbol "X". The first letter of the designation indicates the molecule to which the designated atom belongs, for example, "A" for adenine. To avoid confusion with the chemical symbol for carbon, cytosine is indicated by the prefix "Cy". For purposes of illustration "TX1" refers to the atom labelled as "X1" of the thymine molecule in figure 6.1.

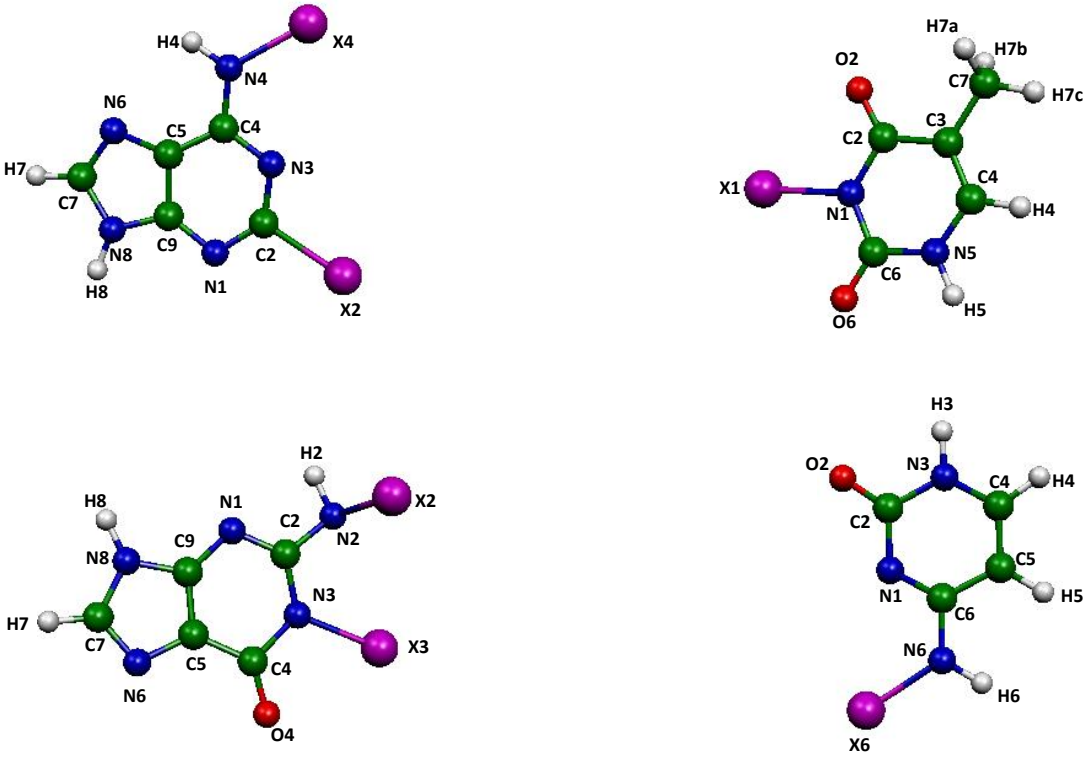


Figure 6.1: Clockwise from top left: adenine, thymine, cytosine and guanine, or halogenated analogues thereof. The symbol "X" can refer to the elements hydrogen, fluorine, chlorine, bromine, iodine or astatine, and where appropriate is hereinafter substituted with the applicable chemical symbol.

Figure 6.2a shows the dA:dT base pair and all of the halogenated derivatives thereof. The structures are shown in their optimised geometries. Figure 6.2b shows the corresponding results for the dG:dC base pairs and its halogenated analogues.

Figures 6.2a and 6.2b each present the optimised structures with the unsubstituted molecular system on the top row, the singly, doubly and triply halogenated systems in rows two, three and four, respectively. The systems wherein one of the substituents is astatine is presented on the fifth row. For this system, the astatine with astatine case is necessarily identical to the double astatinated case on the third row, but is repeated on the fifth row to show it in the latter context. No optimised structure was found for singly fluorinated dG:dC. However, the tendency of fluorine not to form halogen bonds limited the degree of effort devoted to finding this structure and therefore the null result should be treated with caution.

Fluorine atoms did not form halogen bonds, but the optimised structures are included for the sake of completeness. In some cases, it is obvious from a casual inspection of the geometry that there are no halogen bonds present. In other cases, further investigation of the geometry and/or analysis of the energetic results confirm the lack of a halogen bond involving fluorine. For the dG:dC based system, no minimum was found for the singly fluorinated (but not astatinated) case.

As fluorine was (as expected) found not to form halogen bonds, fluorinated molecular systems will be minimally discussed hereinafter. For the other systems, further discussion follows presentation of the graphical and numerical results.

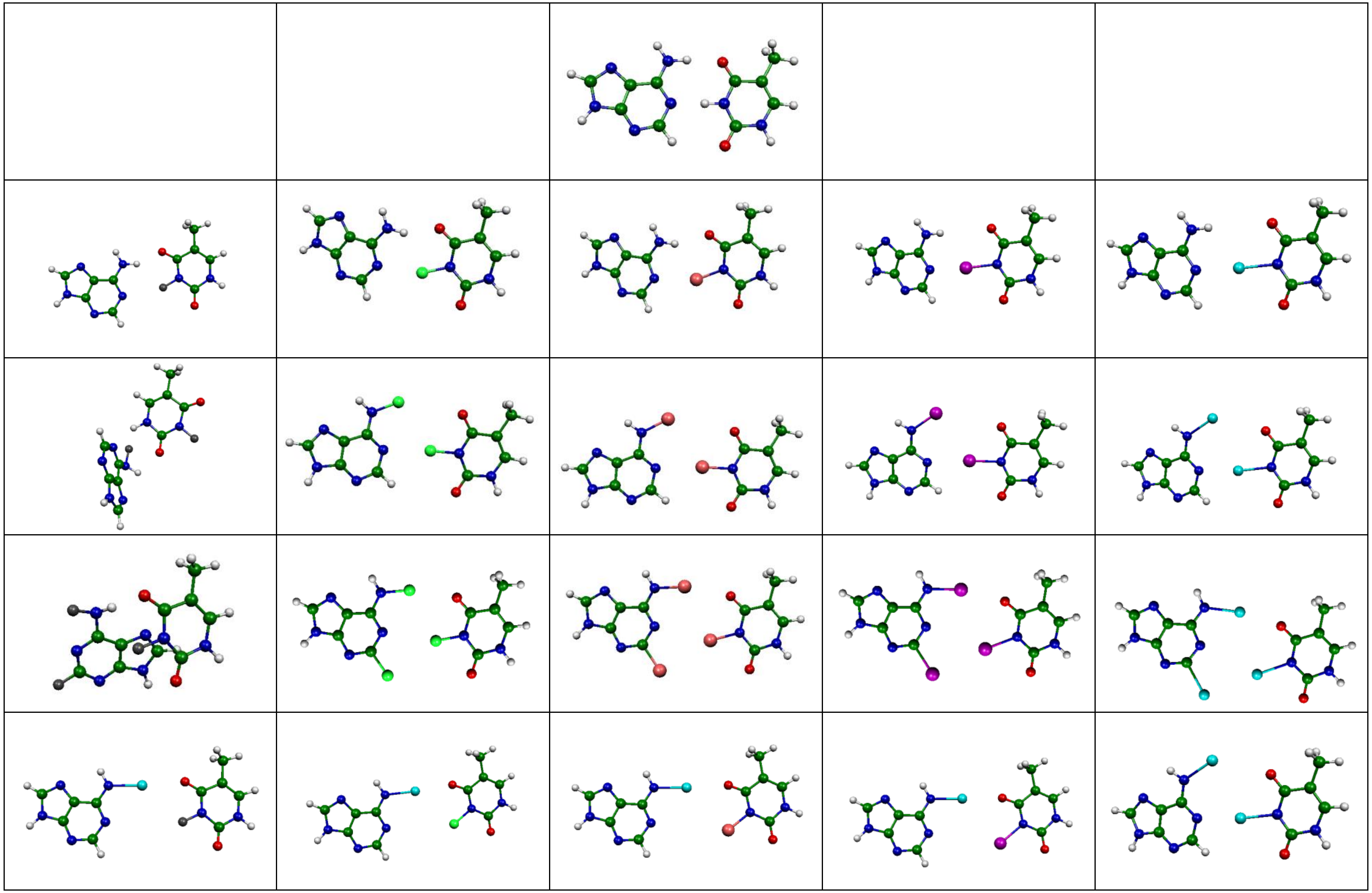


Figure 6.2a: Non-halogenated dA:dT (top row). From second row downwards: respectively single, double and triple halogenation of dA:dT. Double halogenation with at least one substituent being astatine (bottom row). Columns from left to right: variable substituent is respectively fluorine, chlorine, bromine, and astatine (except for top row).

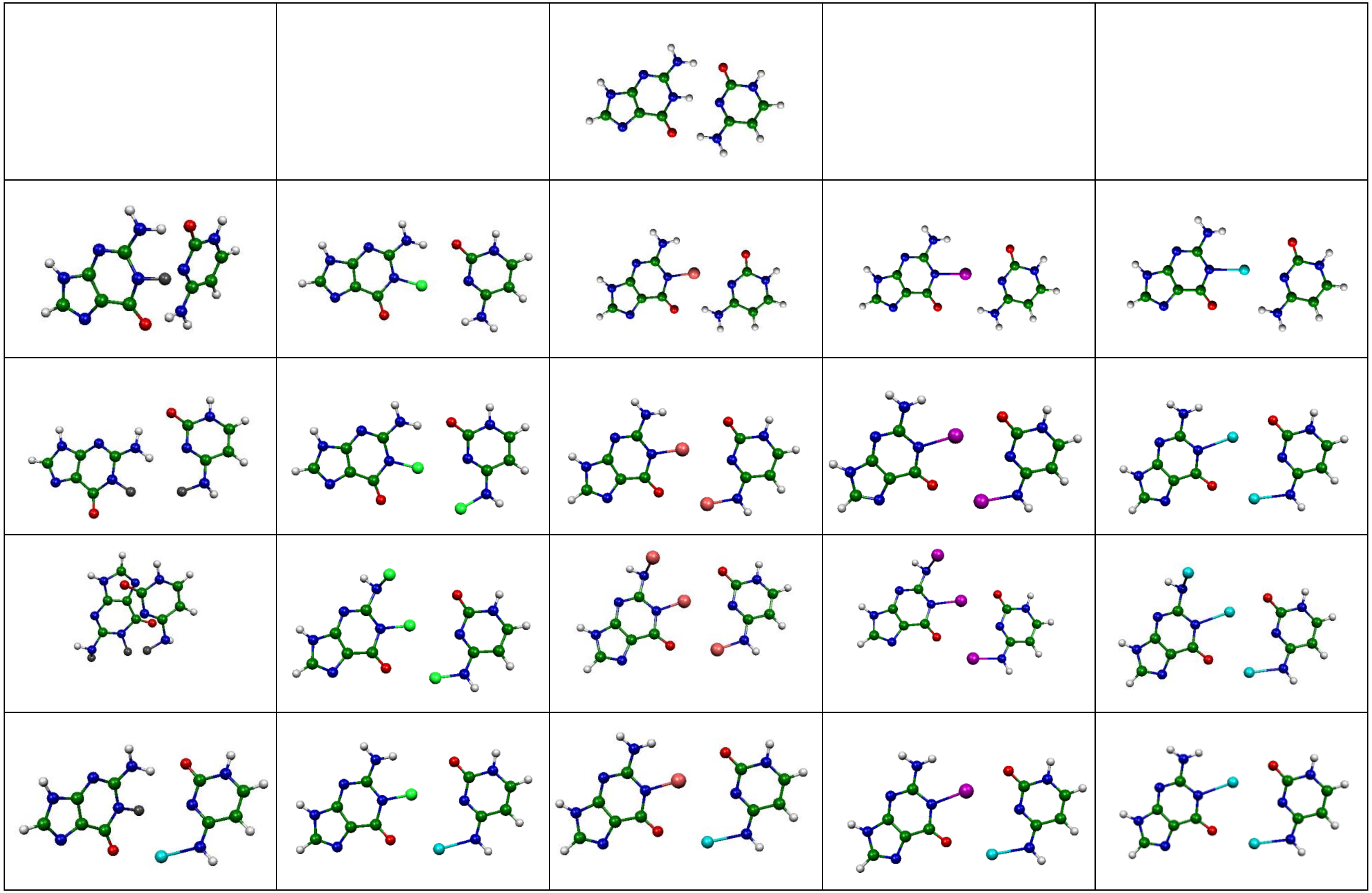


Figure 6.2b: Non-halogenated dG:dC (top row). From second row downwards: respectively single, double and triple halogenation of dG:dC. Double halogenation with at least one substituent being astatine (bottom row). Columns from left to right: variable substituent is respectively fluorine, chlorine, bromine, and astatine (except for top row and first column on second row).

Table 6.1 sets forth definitions for specific interactions. The atoms which participate in the interaction are presented in the right hand column. The molecular dimer, the type of system (e.g. singly halogenated) and atom label are specified. The names presented in the left hand column are used hereinafter to refer to the defined interactions. Interactions may be referred to as bonds even if they do not constitute bonds, strictly defined. Furthermore, would-be interactions are defined even if it be subsequently observed that they do not occur; where an interaction is found not to occur this absence is indicated in table 6.1. Where "Hbond" is included in the name of the interaction, the interaction is (or would be) either (i) a hydrogen bond, (ii) and X•••H interaction (where X is halogen) or (iii) interactions of type (i) or type (ii) but which do not formally constitute hydrogen bonds or X•••H interactions as the internuclear distance is greater than the sum of the atomic VdW radii. Where "Xbond" is included in the name the interaction is a halogen bond or a halogen bonding type interaction but not formally a halogen bond due to the internuclear distance being greater than the sum of the atomic VdW radii. Table 6.1 should be read in conjunction with figures 6.1 and 6.2.

For ease of reference the following prefixes shall have the following meanings. N-: non-halogenated; S-: singly halogenated; D-: doubly halogenated; T-: triply halogenated; and DAt-: Doubly halogenated with at least one of the halogen atoms being astatine. For example "T-dA:dT" refers to the triply halogenated complex comprising halogenated adenine and halogenated thymine. However the full form of the name is used in table 6.1 to ensure maximum clarity when stating definitions, for example "Triply halogenated dA:dT".

All VdW radii data (used in the calculation of VdW ratios) except for astatine comes from CRC Handbook of Chemistry and Physics (90th edition) [135]. For astatine the VdW radius come from the Royal Society of Chemistry website [136]. These values ignore the anisotropy of the atomic radii, treating the atoms as spheres, and hence are inherently imprecise [163]. Both Pauling [164] and Bondi [165] radii are based upon averages from sets of structural data and do not reflect the chemical environment of the atoms [163]. Bondi's [165] inclusion physical and thermodynamic properties, and the covalent radius with the addition of 0.76 Å [166], while a refinement [163], does not remedy this fundamental limitation arising from the use of a standard value, based upon averages, for an atomic radius in the context of a specific chemical environment. Therefore the VdW ratios exhibited herein suffer from this source of approximation and consequential imprecision. However, there is little alternative but to use standard reference data for the purpose of calculating VdW radii, and this is nonetheless considered by IUPAC [7] to be an important parameter for identifying the presence of halogen bonds.

Table 6.1: Definitions of specific hydrogen and halogen bonds.

Name of interaction	Participants in the interaction	Notes
Hbond 1	Non-halogenated dA:dT AX4 – TO2	AX4 is hydrogen
Hbond 2	Non-halogenated dA:dT AN3 – TX1	TX1 is hydrogen
Hbond 3	Non-halogenated dA:dT AX2 – TO6	AX2 is hydrogen
Hbond 4	Singly halogenated dA:dT AX4 – TO2	AX4 is hydrogen
Xbond 1	Singly halogenated dA:dT AN3 – TX1	TX1 is halogen
Hbond 5*	Singly halogenated dA:dT AX2 – TO6	AX2 is hydrogen
Xbond 2	Doubly halogenated dA:dT AX4 – TO2	AX4 is halogen
Xbond 3	Doubly halogenated dA:dT AN3 – TX1	TX1 is halogen
Hbond 6	Doubly halogenated dA:dT AX2 – TX1	AX2 is hydrogen TX1 is halogen
Hbond 7*	Doubly halogenated dA:dT AX2 – TO6	AX2 is hydrogen
Xbond 4	Triply halogenated dA:dT AX4 – TO2	AX4 is halogen
Xbond 5	Triply halogenated dA:dT AN3 – TX1	TX1 is halogen
Xbond 6	Triply halogenated dA:dT AX2 – TX1	AX2 is halogen TX1 is halogen
Xbond 7*	Triply halogenated dA:dT AX2 – TO6	AX2 is halogen
Xbond 8	Doubly halogenated dA:dT AX4 – TO2 (at least one astatine)	AX4 is astatine
Xbond 9	Doubly halogenated dA:dT AN3 – TX1 (at least one astatine)	TX1 is halogen
Hbond 8*	Doubly halogenated dA:dT AX2 – TX1 (at least one astatine)	AX2 is hydrogen TX1 is halogen
Hbond 9*	Doubly halogenated dA:dT AX2 – TO6 (at least one astatine)	AX2 is hydrogen
Hbond 10	Non-halogenated dG:dC GX2 – CyO2	GX2 is hydrogen
Hbond 11	Non-halogenated dG:dC GX3 – CyN1	GX3 is hydrogen
Hbond 12	Non-halogenated dG:dC GO4 – CyX6	CyX6 is hydrogen
Hbond 13	Singly halogenated dG:dC GX2 – CyO2	GX2 is hydrogen
Xbond 10	Singly halogenated dG:dC GX3 – CyN1	GX3 is halogen
Hbond 14	Singly halogenated dG:dC GO4 – CyX6	CyX6 is hydrogen
Hbond 15	Doubly halogenated dG:dC GX2 – CyO2	GX2 is hydrogen
Xbond 11	Doubly halogenated dG:dC GX3 – CyN1	GX3 is halogen
Xbond 12	Doubly halogenated dG:dC GX3 – CyO2	GX3 is halogen
Xbond 13	Doubly halogenated dG:dC GO4 – CyX6	CyX6 is halogen
Xbond 14*	Triply halogenated dG:dC GX2 – CyO2	GX2 is halogen
Xbond 15	Triply halogenated dG:dC GX3 – CyN1	GX3 is halogen
Xbond 16	Triply halogenated dG:dC GX3 – CyO2	GX3 is halogen
Xbond 17	Triply halogenated dG:dC GO4 – CyX6	CyX6 is halogen
Hbond 16	Doubly halogenated dG:dC GX2 – CyO2 (at least one astatine)	GX2 is hydrogen
Xbond 18	Doubly halogenated dG:dC GX3 – CyN1 (at least one astatine)	GX3 is halogen
Xbond 19	Doubly halogenated dG:dC GX3 – CyO2 (at least one astatine)	GX3 is halogen

Xbond 20	Doubly halogenated dG:dC GO4 – CyX6 (at least one astatine)	CyX6 is astatine
*These interactions were found not to exist and are therefore minimally discussed hereinafter.		

The interactions defined in table 6.1 are graphically depicted in figure 6.3 for ease of reference. Figure 6.3 does not provide any additional information not contained in table 6.1. The geometries shown in figure 6.3, except in cases where no halogens are present, are based upon the iodinated structures (iodinated and astatinated where up to two dinstinct halogen elements are present). However, in the context of figure 6.3 the purple coloured atoms can be any halogen element from fluorine to astatine inclusive.

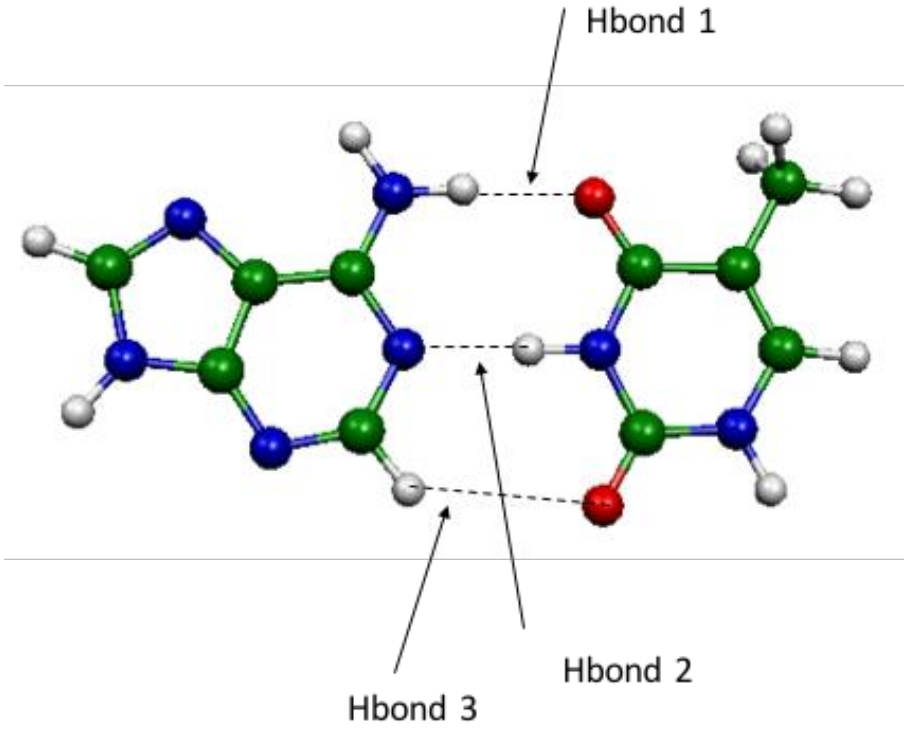


Figure 6.3a: Interactions within the N-dA:dT system.

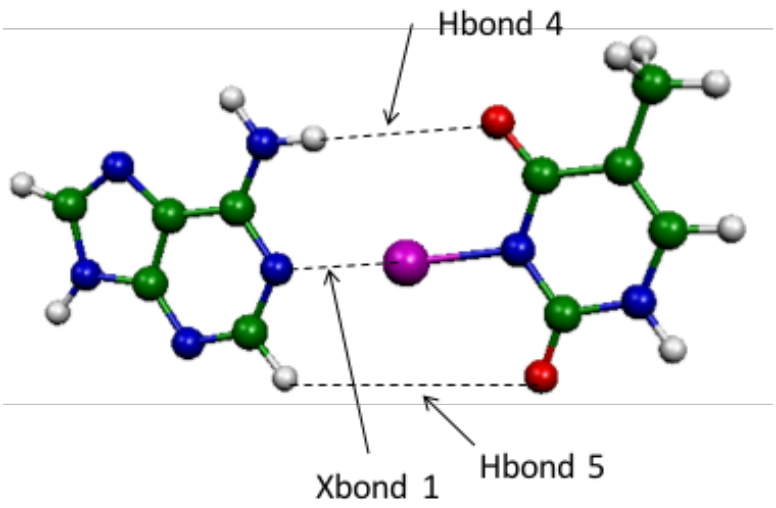


Figure 6.3b: Interactions within the S-dA:dT system. The purple coloured atom can be any halogen element from fluorine to astatine inclusive.

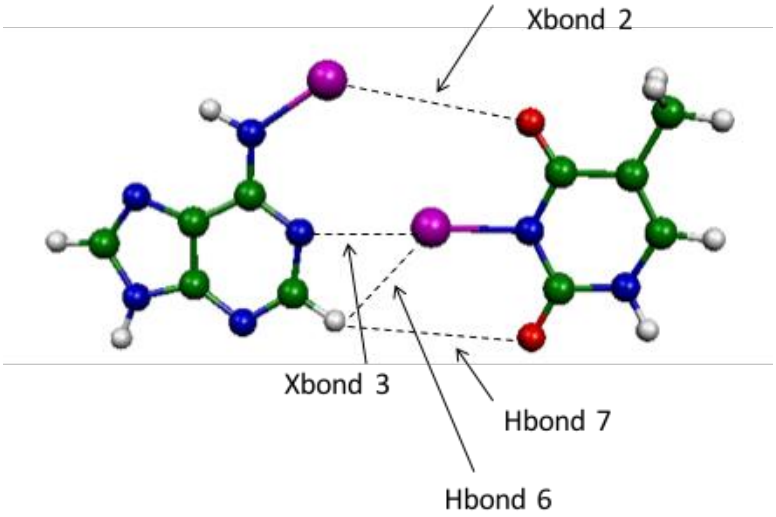


Figure 6.3c: Interactions within the D-dA:dT system. The purple coloured atoms can be any halogen element from fluorine to astatine inclusive.

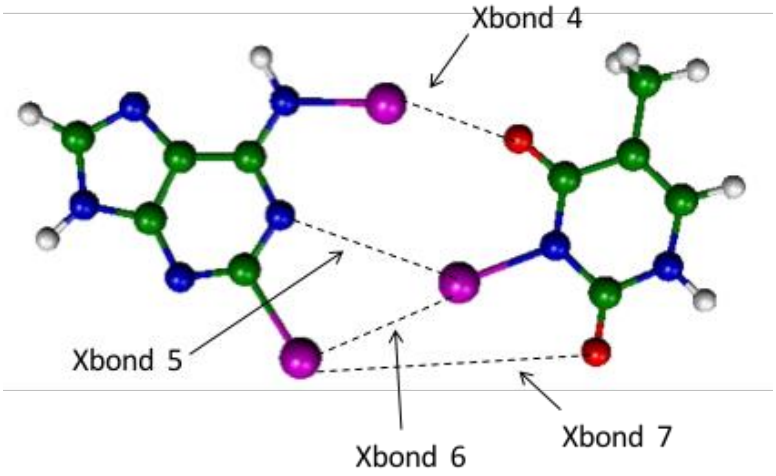


Figure 6.3d: Interactions within the T-dA:dT system. The purple coloured atoms can be any halogen element from fluorine to astatine inclusive.

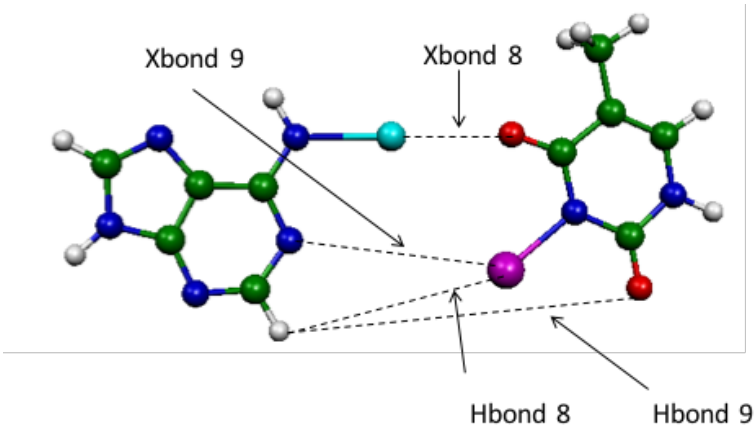


Figure 6.3e: Interactions within the At-dA:dT system. The purple coloured atom can be any halogen element from fluorine to astatine inclusive.

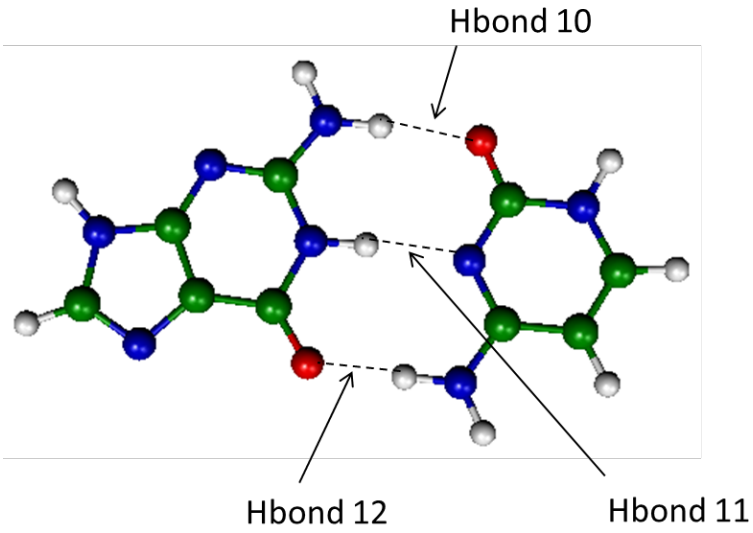


Figure 6.3f: Interactions within the N-dG:dC system.

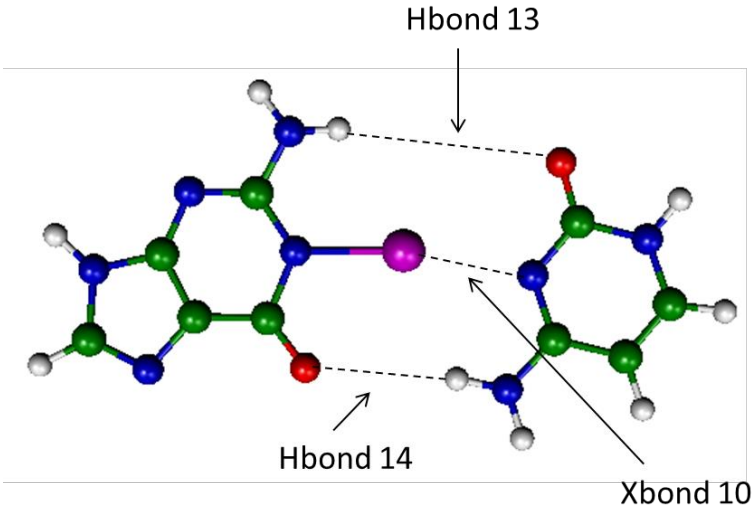


Figure 6.3g: Interactions within the S-dG:dC system. The purple coloured atom can be any halogen element from fluorine to astatine inclusive.

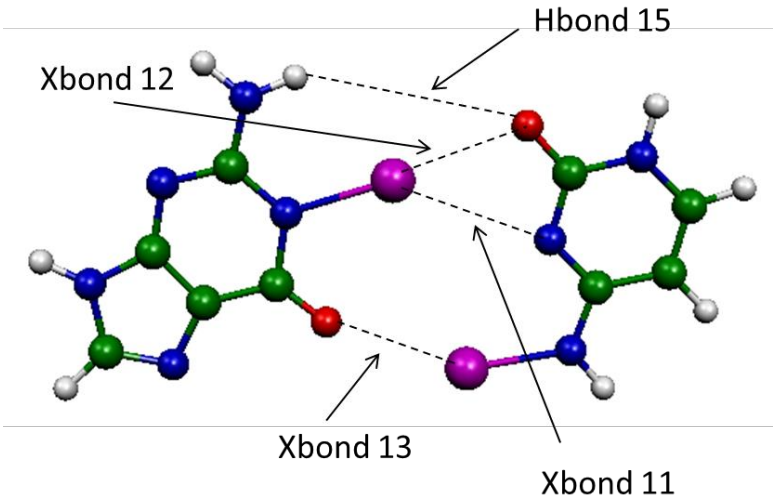


Figure 6.3h: Interactions within the D-dG:dC system. The purple coloured atoms can be any halogen element from fluorine to astatine inclusive.

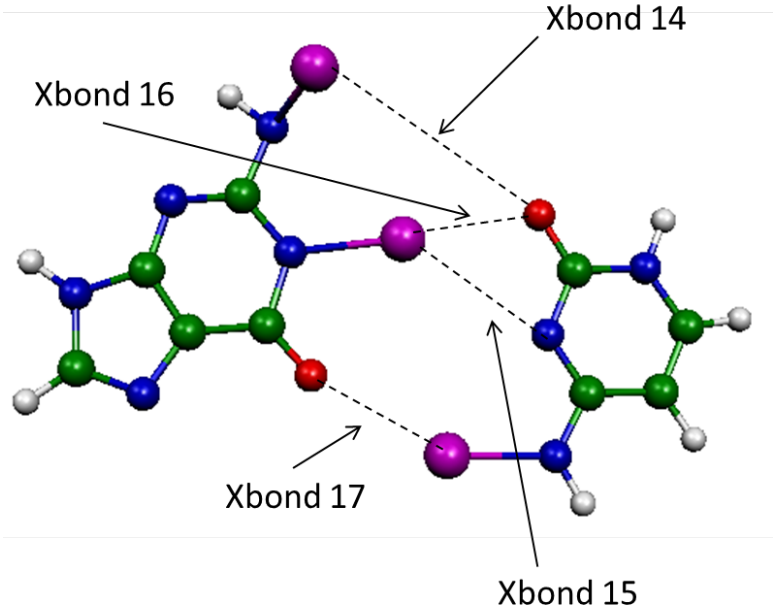


Figure 6.3i: Interactions within the T-dG:dC system. The purple coloured atoms can be any halogen element from fluorine to astatine inclusive.

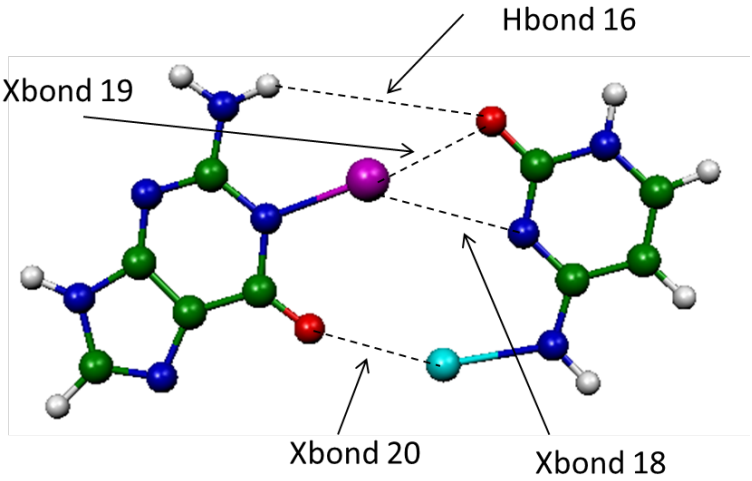


Figure 6.3j: Interactions within the At-dG:dC system. The purple coloured atom can be any halogen element from fluorine to astatine inclusive.

6.4.1 Results for the DNA base pair comprising adenine and thymine and halogenated analogues thereof

In this subsection the results of computations upon the dA:dT systems are disclosed and discussed. Hence halogen bonds Xbond 1-9 and hydrogen bonds Hbond 1-9 are presented herein. Furthermore, all of these interactions occur (or would occur) in the systems presented in graphical form in figure 6.2a. Note that interaction energies pertain to the overall interaction between the molecules. Contributions have not been resolved into interatomic components. Hence the interaction energy is the same for all interactions within any given system.

6.4.1.1 Results for non-halogenated dA:dT

This sub-subsection concerns the interactions found in N-dA:dT, presented in the top row of figure 6.2a, and including hydrogen bonding interactions Hbond 1-3. Table 6.2 presents the Wiberg bond indices (WBIs) and geometric properties of each of these interactions.

The interaction energy for the N-dA:dT molecular system is -58.2 kJ mol^{-1}. This value is used as the baseline for reporting interaction energies relative to the non-halogenated system of dA:dT.

Table 6.2: Geometric results and WBIs for the interactions found in N-dA:dT.

Interaction	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Hbond 1	1.95	0.74	173	0.0370
Hbond 2	1.83	0.69	179	0.0907
Hbond 3	2.82	1.08	133	0.0027

From observation of the geometric and VdW results, Hbond 3 makes a minor contribution to the stability of the molecular system. It is a C-H•••O hydrogen bond, which is generally a weak interaction. Although the VdW ratio is only slightly greater than unity, the bond angle and WBI value weigh substantially against there being a significant hydrogen bond type interaction short of a formal hydrogen bond.

Hbond 1 and Hbond 2 appear to contribute most to the stability of the system. The low (but not negligible) WBI values imply that these interactions are predominantly electrostatic rather than covalent. The VdW ratios, which are much smaller than unity, and the bond angles support the conclusion that Hbond 1 and Hbond 2 are both formal hydrogen bonds. Observation of the structure of N-dA:dT (see figure 6.2a) leads to the conclusion that Hbond 1 and Hbond 2 are collectively the principal contributors to the stability of N-dA:dT, by reason of an absence of plausible alternative stabilising intermolecular interactions, except for the minor Hbond 3 interaction.

6.4.1.2 Results for singly halogenated dA:dT

The hypothesised interaction Hbond 5, even by casual observation of the structure (see figure 6.2a), does not genuinely exist as a significant interaction in S-dA:dT. The absence of Hbond 5 from S-dA:dT for all halogen elements is immediately obvious so no results pertaining thereunto are presented herein.

Tables 6.3a and 6.3b disclose, for each halogen, the geometric and WBI results for interactions Hbond 4 and Xbond 1 respectively. For elements F, Cl, Br, I and At the interaction energies are respectively (in kJ mol^{-1}): -20.8, -33.5, -39.8, -58.8 and -75.9. Relative to N-dA:dT the interaction energies (in kJ mol^{-1}) are (in the same order): +38.0, +24.7, +18.4, -0.6 and -17.7. Hence iodinated and astatinated S-dA:dT are the only systems of the S-dA:dT that have a more strongly negative interaction energy than N-dA:dT. Iodinated S-dA:dT was found to have an interaction energy that was very similar to that of the non-halogenated dimer.

Table 6.3a: Geometric and WBI results for interaction Hbond 4.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
F	2.06	0.79	155	0.0211
Cl	2.15	0.82	159	0.0178
Br	2.30	0.88	160	0.0121
I	3.52	1.35	179	0.0006
At	3.71	1.42	177	0.0004

Table 6.3b: Geometric and WBI results for interaction Xbond 1.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
F	3.03	1.00	161	0.0033
Cl	2.76	0.84	159	0.0410
Br	2.76	0.81	159	0.0666
I	2.56	0.73	176	0.1878
At	2.59	0.73	177	0.1949

From the disclosures in tables 6.3a and 6.3b it is immediately clear that a substantial qualitative change occurs when descending the halogen group from bromine to iodine. The internuclear distance substantially increases for Hbond 4 and decreases for Xbond 1. The VdW ratios also show the same trend albeit that in table 6.3b this change from bromine to iodine is over-shadowed by an incidental difference of 0.16 between fluorine and chlorine. A dramatic change in bond angle from bromine to iodine is perhaps the most significant disclosure in these tables, especially in table 6.3b, which shows the halogen bond angle becoming much closer to linear. The WBI results point to the same conclusion, with a dramatic increase in the case of Xbond 1 and a major decrease for Hbond 4. The WBI results for Xbond 1 with iodine and astatine suggest that while this is still a predominantly electrostatic interaction, there is a significant covalent contribution.

The trend does however appear to be binary. There does not appear to be a substantial degree of halogen bonding for halogens lighter than iodine. Although for chlorine and bromine the Xbond 1 VdW ratio is smaller than unity, the bond angle is quite far from linear and does not show the trend from chlorine to bromine that would be consistent with halogen bonding. Any halogen bonding for chlorine and bromine must therefore be very slight if it is present at all, although the fairly significant decrease in VdW ratios relative to the fluorine case lends some support to the hypothesis that these halogen bonding interactions are not completely absent. There is also no discernible increase in the magnitude of the halogen bonding going from iodine

to astatine. There is no change in VdW ratios and only a slight increase in halogen bond angle and WBI value.

Hbond 4 is directly analogous to Hbond 1 in N-dA:dT. The major decrease in bond angle for halogenation with fluorine, chlorine or bromine is difficult to explain if hydrogen bonding is dominating. The most plausible explanation might be an X•••H interaction attracting AX4 towards TX1; however, that interaction is not the subject of investigation herein. The near linearity of Hbond 4 under halogen bond domination (iodine and astatine) is most likely to be a coincidence arising from the optimisation of the halogen bond. Elongation of the AX4 – TO2 internuclear separation is almost certainly to sterically accommodate the halogen, increasing with the size of the halogen atom. This also explains the modest increase in VdW ratios for Hbond 4 from fluorine to bromine. Absence of this steric effect is also the most likely reason why the WBI value is greater for Hbond 1 than for any of the halogenated analogues thereof.

6.4.1.2.1 Results from varying the halogen bond angle

Figure 6.4 shows how the interaction energy of the system varies as a function of the halogen bond angle, defined as the TN1 – TX1 – AN3 bond angle, for the non-halogenated case and for each halogen element from fluorine to astatine inclusive. The resolution of the graphs is limited by a finite scan step size of 5°. This step size is inherently arbitrary, but was selected as a reasonable compromise between graph resolution and computational economy. The starting angle of each scan was determined by the angle following full (unconstrained) optimisation. These are the values shown in tables 6.2 and 6.3b. Predominantly for technical reasons, angles greater than 180° were not explored in the scans. As the molecular system is not symmetrical there is no reason to expect that the energy profile would be the same on the other side of the linear point. Additional scanning calculations could therefore cast further light upon the complete energy profile as a function of bond angle.

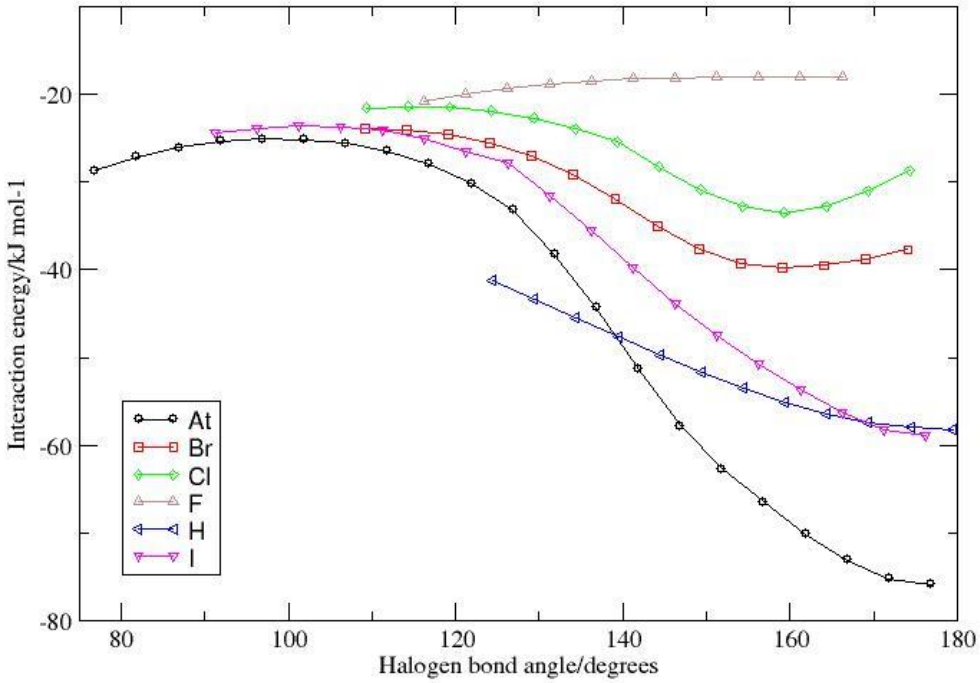


Figure 6.4: Interaction energy as a function of TN1 – TX1•••AN3 bond angle. TX1 is F, Cl, Br, I, At or H as indicated in the legend.

The fluorinated form of the system, which does not engage in halogen bonding, exhibits a weak trend toward less negative interaction energy as the halogen bond angle is increased. For all bond angles, the interaction energy is less strongly negative than for any of the other variants of the system (but still below zero), although similar to the chlorinated case for bond angles below approximately 125°.

For the chlorinated and brominated cases there is clearly a minimum in the interaction energy at 159°, as found during full optimisation. The graphs for these elements (see figure 6.4) confirm that there is a significant increase in energy as the halogen bond angle approaches linearity from that angle, with this attribute being more pronounced for chlorine than for bromine.

The iodinated and astatinated cases have a minimum (found by full geometry optimisation) that is within 5° of strict linearity. Although not explored in sufficient detail to exclude the possibility of another (perhaps even deeper) minimum between 176° (iodine) or 177° (astatine) and 180°, the results obtained during this study suggest the minimum in the energy profile is slightly displaced from linear. This finding is consistent with the observation that Xbond 1 is not the sole interaction that affects the geometry and stability of the system, although Hbond 4 was not

found to exist in any significant sense for the iodinated and astatinated versions of the system. Even the very long-range interactions and weaker interactions based purely upon the induced dipole – induced dipole effect can have some impact upon the geometry of the system. Compared with iodination, astatination leads to a slightly (but appreciably) steeper gradient on the approach to the minimum in the energy profile.

For all cases of S-dA:dT, and at all angles (at least above 120^{o}) the order is maintained wherein the interaction energy where halogenation is by a heavier element is more negative than where it is by a lighter element. The greater displacement of the minimum in the energy profile from linear for the chlorinated and brominated cases compared with the iodinated and astatinated cases, the sharper increase in energy for the case of chlorination compared with that of bromination, and the differences in the gradients between the profiles for the iodinated and the astatinated cases are all likely to be a reflection of the greater dominance of Xbond 1 compared with Hbond 4 (and other minor interactions) as the halogen group is descended. The qualitative differences between the iodinated and astatinated cases (on the one side) and the chlorinated and brominated cases (on the other) could be expected based upon the clear qualitative differences between these two sets of system to be found in table 6.2, which shows that Hbond 4 is uniquely absent in the case of the former group. A general observation can be made that for all of the variants of the S-dA:dT system, the differences in the interaction energies increase as linearity is approached, appearing to roughly converge as the angle falls significantly below 120^{o}. As (i) the region around linear would be expected to be where halogen bonding would be most dominant and (ii) the halogen bonding would be expected to become stronger going down the halogen group, this observation as to where the greatest disparity in interaction energies occurs is consistent with halogen bonding being the dominant factor in determining the interaction energy as a function of bond angle, and is therefore consistent with halogen bonding being at least a substantial factor in the stability of the system. If differences in interaction energy as a function of bond angle can be attributed to halogen bonding (as justified above), then this latter conclusion is also supported by quantitative considerations, bond angle affecting the interaction energy by a magnitude in the order of tens of kilojoules per mole over the angular range of the scans; even for the chlorinated case.

Discussion of figure 6.4 would be incomplete without mentioning the energy profile of N-dA:dT. Whereas for all halogens, the order of the variants of the S-dA:dT system by interaction energy (at least above 120^{o}) does not vary, where the atom in question is hydrogen, the energy profile intersects those of the iodinated and astatinated cases of S-dA:dT. As in those two halogenated cases, the minimum in the energy profile of N-dA:dT is within 5^{o} of linear (179^{o} – see Hbond 2 in table 6.2). These observations indicate that while Hbond 2 (a hydrogen bond) is angle-dependent (with a preference for linearity), the angle dependence is weaker than for Xbond 1, the latter being a halogen bond.

6.4.1.3 Results for doubly halogenated dA:dT

The fluorinated form of this molecular dimer displays a very different geometry to those adopted for the other variants of D-dA:dT. The structure of doubly fluorinated dA:dT is shown

in figure 6.2a in the interests of completeness but is not further discussed other than as confirmation of the poor halogen bonding properties of fluorine. Therefore, quantitative results for this halogen type are not presented. In contrast to fluorine, all of the other halogens that were investigated contain a halogen bond (Xbond 3).

The doubly halogenated system D-dA:dT is hypothesised to contain the interactions Xbond 2, Xbond 3, Hbond 6 and Hbond 7. Hbond 7 does not make a substantial contribution to the stability of the dimer for any of the halogen types. Substantiating this statement, the VdW ratios for Hbond 7 for the chlorinated, brominated, iodinated and astatinated systems are respectively: 1.48, 1.53, 1.63, and 1.85. Hence Hbond 7 hydrogen bonding type interactions are likely to be very weak, if they exist at all. Based on these VdW ratios the hypothesised Hbond 7 interaction can be disregarded.

Tables 6.4a, 6.4b and 6.4c disclose quantitative results for D-dA:dT for all halogen types except fluorine. These tables respectively present results for Xbond 2, Xbond 3 and Hbond 6. The interaction energies for the chlorinated, brominated, iodinated and astatinated variants of D-dA:dT are respectively (in kJ mol^{-1}) -19.30, -27.18, -39.28, and -54.99; relative to N-dA:dT the interaction energies are respectively (in kJ mol^{-1}) +38.90, +31.02, +18.92, and +3.21. Hence for all halogen types the interaction is less strongly attractive than in the case of the non-halogenated dA:dT dimer, although the difference in interaction energy is small in the case of the astatinated system.

Table 6.4a: Geometric and WBI results for interaction Xbond 2.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Cl	3.87	1.18	148	0.0015
Br	4.15	1.23	141	0.0011
I	4.37	1.25	134	0.0013
At	3.62	1.02	135	0.0065

From table 6.4a it is clear that chlorine, bromine and iodine do not form the Xbond 2 interaction. The VdW ratios, bond angles and WBI results all support this conclusion. Astatine also does not form a formal Xbond 2 halogen bond as the internuclear separation is greater than the sum of VdW radii (as required by the IUPAC definition of a halogen bond [7]). However, there is clearly a substantial reduction in both the absolute internuclear separation and especially the VdW ratio when descending the halogen group to astatine and, although still small, the WBI is, in relative terms, considerably greater than for any of the other halogen elements. The bond angle is still much smaller than that which is usually associated with halogen bonds. However, as discussed in the chapter concerning microsolvated 1-methyl-5-halouracil, the σ-hole covers a greater angle around the ideal 180°, as the halogen group is descended. Therefore it might be the case that the ability of astatine to form halogen bonding type interactions at greater departures from the ideal angle enables Xbond 2 to have some limited stabilising effect in the case of the astatinated D-dA:dT system, without being a formal halogen bond. These results contrast with

the results for Hbond 1 in N-dA:dT where a hydrogen bond does form (Hbond 1), and with the S-dA:dT results (Hbond 4), where the fluorinated, chlorinated and brominated systems have internuclear separations well within the sum of the VdW radii.

Table 6.4b: Geometric and WBI results for interaction Xbond 3.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Cl	2.80	0.85	175	0.0310
Br	2.84	0.84	175	0.0454
I	2.85	0.81	176	0.0834
At	2.72	0.76	174	0.1402

The results disclosed in table 6.4b are strongly indicative of a major role being played by Xbond 3 in the stability of D-dA:dT for all halogen types from chlorine to astatine. VdW ratios are all well below unity. The WBI results show some degree of covalency, especially for astatine. Bond angles are also close to linear. Considering the complexity of the chemical environment the bond angles for all halogen types are remarkably close to the ideal value of 180°. The trend going down the halogen group appears to include a strengthening of the Xbond 3 halogen bond, primarily indicated by the trend in VdW ratios, while the trend in WBI values indicates that the degree of orbital overlap between the lone pair of electrons on the nitrogen atom and the σ* antibonding orbital (corresponding to the TN1-TX1 covalent bond) also increases going down the halogen group. The geometric results for the astatinated system are indicative of a significant increase in halogen bond strength due to sharp drop in VdW ratio. However, the presence of astatine at AX4, as noted above, increases the significance of the Xbond 2 interaction and that increased interaction is probably the primary cause of the slight reduction in Xbond 3 bond angle. However, the effect of the halogen type, where both halogens are of the same type, appears to be minimal and indeed the trend from bromine to iodine is a slight increase in halogen bond angle.

Table 6.4c: Geometric and WBI results for interaction Hbond 6.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Cl	3.03	1.07	134	0.0013
Br	3.02	1.03	134	0.0020
I	2.98	0.97	135	0.0039
At	3.04	0.98	144	0.0029

The postulated Hbond 6 would be an X•••H interaction wherein the negatively charged ring on the halogen atom interacts with an electron deficient hydrogen atom. Hence for this type of interaction the ideal angle would be 90°. For all of the systems for which results are presented in table 6.4c, the bond angle is very far from linear. Without further probing on the σ-hole it is difficult to discern whether the interaction would be stabilising or destabilising (interaction between the σ-hole and the electron deficient hydrogen atom). At these angles a significant driving force towards these geometries on account of Hbond 6 would be unlikely. All VdW ratios are close to unity, and in the cases of iodine and astatine are slightly below unity. However,

based upon the analysis stated above this is most likely to be a coincident side effect of the optimisation of the Xbond 3 interaction. Low WBI values indicate little orbital overlap.

Taken together, the results disclosed in tables 6.4a, 6.4b and 6.4c provide compelling evidence that the dominant stabilising interaction for the chlorinated, brominated, iodinated and astatinated systems is the Xbond 3 halogen bond. The dominance of this interaction provides an explanation for the invariance in the halogen bond angle. The effects of the hydrogen bonding interactions Hbond 6 and (especially) Hbond 7 appear to be at most very minor. Meanwhile the weaker halogen bonding type interaction, Xbond 2, while perhaps being responsible for a small departure of Xbond 3 from linearity, is unlikely to cause much variation in Xbond 3 halogen bond angle as both Xbond 2 and Xbond 3 are based on the same halogen element, and hence as Xbond 3 increases in strength with increasing atomic number of the halogen atom, the same effect would be expected for Xbond 2. The sudden increase in the strength of Xbond 2 upon descending the halogen group from iodine to astatine is also a plausible explanation for the slight reduction in Xbond 3 halogen bond angle going from the iodinated to the astatinated system.

Comparison between Xbond 3 and Xbond 1 (see table 6.3b) reveals that, whereas a well optimised halogen bond is the preserve of iodine and astatine for Xbond 1, it is formed by all investigated halogen elements heavier than fluorine for Xbond 3. This comparative dominance of the analogous halogen bond in D-dA:dT for the lighter halogens is probably due to the lack of a strong competing hydrogen bond Hbond 4 (replaced by the relatively weak Xbond 2), which in the case of S-dA:dT is only overcome by the presence of iodine or astatine at TX1.

6.4.1.4 Results for triply halogenated dA:dT

The triply halogenated dA:dT system in principle contains the following interactions: Xbond 4, Xbond 5, Xbond 6 and Xbond 7. Xbond 4 was found for all systems except for the fluorinated case, although the bond angles are small by halogen bond standards. The quantitative results are set forth in table 6.5a. The Xbond 5 interaction probably does not make a significant contribution towards stabilising the dimer. The would-be halogen bond angles are too small, especially for iodine and astatine. Furthermore, the trend towards smaller angles going down the group, although possibly reflecting the greater flexibility of halogen bond angles, would not be expected if this was a major contributing interaction. As there is nonetheless a possibility of halogen bonding for the bromine and chlorine cases, the results are disclosed in table 6.5b. Observation of the geometries of the triply halogenated systems suggest that the reason for the lack of Xbond 5 is rooted in a driving force towards optimising Xbond 6, which is found for iodine and astatine, and possibly for bromine. For the chlorinated system, the internuclear distance is greater than the sum of VdW radii and the angle is only borderline plausible for a halogen bond; however, it may make a minor contribution towards the stability of the system. The results for Xbond 6 are contained in table 6.5c. Xbond 7, although hereinbefore postulated does not, in practice, exist. The internuclear distances are vastly greater than the sums of the VdW radii.

Furthermore, observation of figure 6.2a immediately makes it clear that the angles are not commensurate with halogen bonding. Therefore, the fictitious Xbond 7 is not discussed further.

The interaction energies for the fluorinated, chlorinated, brominated, iodinated and astatinated T-dA:dT systems are respectively (in kJ mol^{-1}) : -41.21, -17.97, -28.30, -52.23, and -82.16. This yields energies relative to the interaction energies of N-dA:dT of (in the same order and units) +16.99, +40.23, +29.90, +5.97, and -23.96. Although included for the sake of completeness, the fluorinated system did not include any of the interactions of interest, so no data in relation thereto is set forth in table 6.5. The astatinated system is unique, having a more strongly attractive interaction (measured by interaction energy) than N-dA:dT.

Table 6.5a: Geometric and WBI results for interaction Xbond 4.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Cl	2.85	0.87	167	0.0115
Br	2.86	0.85	165	0.0176
I	2.79	0.80	166	0.0401
At	2.66	0.75	169	0.0723

From table 6.5a there is a clear trend towards smaller VdW ratios as the halogen group is descended, although the difference between chlorine and bromine is quite small. The WBI results also indicate that, while these interactions are predominantly electrostatic, there is a trend towards greater covalency, going down the halogen group, and again the difference is relatively small between chlorine and bromine. Hence there is a correlation between smaller VdW ratios (generally indicate stronger interactions) and the degree of covalency. The trend that is conspicuous by its absence is in the data for bond angle. The triply substituted dA:dT system is a fairly complex chemical environment, including competition with other halogen bonds. Therefore, the greater driving force to optimise the halogen bonds for the heavier halogens is balanced by the greater ability of the heavier halogen elements to tolerate halogen bond angles that depart further from being linear.

Table 6.5b: Geometric and WBI results for interaction Xbond 5.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Cl	3.21	0.97	160	0.0095
Br	3.50	1.03	151	0.0071
I	3.89	1.10	138	0.0044
At	3.90	1.09	134	0.0044

From table 6.5b it is clear that Xbond 5 is not a major interaction. With the exception of chlorine, a formal halogen bond can be immediately excluded as the VdW ratios are greater than unity. Furthermore, the bond angles, at least for iodine and astatine, are totally implausible for halogen bonds. The WBI results indicate minimal covalency, with chlorine having the greatest WBI value. Considering the possibility that chlorine's plausible halogen bond geometry could be a coincidental consequence of the optimisation of other interactions, the lack of a trend towards smaller VdW ratios going down the halogen group (indeed the inverse), and (even for chlorine and bromine) small halogen bond angles, Xbond 5 is probably, at most, a negligible interaction even for the chlorinated and brominated systems, although the trend in WBI values lends some limited support for the proposition that it is not totally non-existent.

Table 6.5c: Geometric and WBI results for interaction Xbond 6.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Cl	3.53	1.01	155	0.0045
Br	3.59	0.97	164	0.0172
I	3.57	0.90	176	0.0892
At	3.45	0.85	178	0.1861

Table 6.5c discloses results that indicate that the Xbond 6 interaction is likely to play a significant role in the stability of the T-dA:dT system. Although chlorine does not form a formal halogen bond on account of the VdW ratio being greater than unity, it is only marginally greater than that threshold. The halogen bond angle is also small but plausible for halogen bonding, so there could be a halogen bonding style stabilisation short of being a canonical halogen bond. From bromine to astatine there is a clear trend towards stronger halogen bonding, reflected in the VdW ratios. Furthermore, there is a marked jump towards greater linearity from bromine to iodine, indicating that this interaction becomes qualitatively more important in determining the overall geometry of the system at that point in the descent down the halogen group. The WBI results indicate that, in comparison with other systems in this study, the covalent contribution to the interaction is relatively great (especially for the iodinated and astatinated cases), and this contribution increases going down the halogen group.

6.4.1.5 Results for doubly halogenated dA:dT where at least one halogen is astatine

The At-dA:dT system comprises an astatine atom at the AX4 position for all systems and a halogen atom (F, Cl, Br, I or At) at the TX1 position. Where TX1 is astatine this is identical to the D-dA:dT system where the halogen is astatine. It is postulated that the interactions Xbond 8, Xbond 9, Hbond 8 and Hbond 9 may be present in this system. However, Hbond 8 is equivalent to Hbond 6 and the system which come closest to demonstrating this interaction is for TX1 being astatine, which is already discussed in dismissive terms in the context of Hbond. For the other systems, observation of figure 6.2a makes it clear, *a fortiori*, that this interaction is not to be found. Hbond 9 also clearly does not exist, based on observation of figure 6.2a, as per its D-dA:dT analogue Hbond 7.

Therefore, only the halogen bonding interactions, Xbond 8 and Xbond 9, are discussed further, their quantitative results being set forth in tables 6.6a and 6.6b respectively. In the interest of completeness, the results for TX1 being astatine are included despite this inclusion constituting repetition. It should be noted in the context of Xbond 8 that the interaction is between the astatine atom (AX4) and an oxygen atom (TO2) for all cases; the identified halogen atom is at TX1, and not directly participating in the Xbond 8 interaction. Xbond 8 was found for all TX1 halogen types.

The interaction energies for the At-dA:dT systems were found to be (in kJ mol^{-1}) -36.92, -39.44, -40.42, -42.41 and -55.04 for TX1 as F, Cl, Br, I and At respectively. In the same units and order the interaction energies relative to those of N-dA:dT are +21.28, +18.76, +17.78, +15.79 and +3.16. The very slight variation in the interaction energies for D-dA:dT (astatine) and At-dA:dT (astatine) (0.05 kJ mol^{-1}) is hereby noted but not investigated; it is well within the margin of error inherent in the methods used for these calculations. For all halogen types the interaction energies were less negative than for the all-hydrogen analogue. This observation is consistent with the interaction energies for the D-dA:dT systems.

Table 6.6a: Geometric and WBI results for interaction Xbond 8.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
F	2.75	0.78	175	0.0565
Cl	2.76	0.78	177	0.0520
Br	2.75	0.78	177	0.0509
I	2.75	0.78	180	0.0507
At	3.62	1.02	135	0.0065

From table 6.6a it is clear that from fluorine to iodine the identity of the halogen at TX1 has a minimal impact on Xbond 8. All of the systems appear to show a clear halogen bond with a bond angle close to linear and an internuclear separation well below the sum of VdW radii. It might have been postulated that the decreasing electronegativity of the halogen elements from fluorine to iodine might have strengthened the astatine-oxygen halogen bond as a less electronegative element at TX1 might have been expected to result in more electron density being concentrated on the TO2 oxygen atom. There is a slight trend towards linearity from fluorine to iodine but the WBI values (indicating covalent overlap) might also have been expected to increase but instead a weak trend is seen in the opposite direction; these values are moderate for all cases compared with the previously discussed systems. It is possible that steric effect might be responsible for these slight variations in the halogen bond angles and WBI results. Probably the most significant measure of the electronic effect of varying the TX1 halogen is the VdW ratio and that remains identical (to two decimal places) from fluorine to iodine. As discussed below in light of the data in table 6.6b, there appears to be no competition with Xbond 9 for TX1 halogen being fluorine, chlorine, bromine or iodine. By dramatic contrast the doubly astatinated system does not appear to form an Xbond 8 halogen bond. The VdW ratio is slightly greater than unity and the halogen bond angle is very small. As discussed in the context of D-dA:dT, some long range halogen bonding type interaction short of a formal halogen bond might be present in view of the ability of astatine to form halogen bonds with angles that are far from ideal. Nonetheless, it is clear that Xbond 8 is not the important interaction for astatine as it is for all of the lighter halogens.

Table 6.6b: Geometric and WBI results for interaction Xbond 9.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
F	4.90	1.62	153	0.0004
Cl	5.00	1.52	141	0.0009
Br	5.08	1.49	136	0.0009
I	4.96	1.41	132	0.0013
At	2.72	0.76	174	0.1402

From table 6.6b it is immediately obvious that Xbond 9 only exists when the halogen at TX1 is astatine. The huge VdW ratios for halogens F, Cl, Br and I on their own put this conclusion beyond reasonable doubt and the small bond angles only serve to confirm it further. The possibility of Xbond 9 interactions for these lighter halogens cannot be revived by the trend towards smaller VdW ratios going down the group (and the trend in bond angles points away from it); the internuclear distances and bond angles are too unfavourable. In contrast to the lighter halogens the TX1 astatine does appear to form a halogen bond with the AN3 nitrogen. As the results for the astatine case are necessarily the same as for the identical system discussed in the context of D-dA:dT, the same commentary could be made here.

There appears to be a competition between Xbond 8 and Xbond 9. For halogens lighter than astatine at TX1, Xbond 8 has a clear advantage deriving from the fact that only this interaction entails a halogen bond being formed by astatine, which (it is well established from the results and discussion elsewhere throughout the preceding chapters of this thesis) forms stronger halogen bonds than the lighter halogen elements. By contrast when TX1 is astatine, Xbond 8 must compete with another potential astatine-based halogen bond, Xbond 9. When Xbond 8's elemental advantage is negated it is outcompeted by Xbond 9. The results in tables 6.6a and 6.6b appear to indicate that the outcome of this contest is binary without any of the systems exhibiting a balanced compromise between these two interactions. Given that the hypothesised hydrogen bonds Hbond 8 and Hbond 9 were found not to form, the dominance of Xbond 9 for the doubly astatinated system is most likely due to the electronic differences between the AN3 nitrogen atom and the TO2 oxygen atom. Perhaps the electronegative TO6 oxygen atom withdraws some of the electron density from TO2; however, this would need to be explored in greater detail. A casual observation of the overall geometry does not, at first impression, point towards steric effects being responsible. The contrasting clear dominance of the Xbond 8 interaction where TX1 is not astatine underlines the importance of the specific halogen element that seeks to form a halogen bond. This factor clearly and decisively outweighs the (evidently present) countervailing factor or factors in favour of the formation of the Xbond 9 halogen bond.

6.4.2 Results for the DNA base pair comprising guanine and cytosine and halogenated analogues thereof

This subsection concerns systems based on the guanine-cytosine dimer ("dG:dC"). The non-halogenated dimer is used as the reference system in the same manner as for the adenine-thymine systems. Halogenated analogues of dG:dC are likewise specified in table 6.1, with halogen bonds Xbond 10-20 and hydrogen bonds Hbond 10-16 occurring within this system. The halogens from fluorine to astatine were investigated for each halogenated analogue of the dG:dC dimer.

6.4.2.1 Results for non-halogenated dG:dC

The non-halogenated form of dG:dC ("N-dG:dC") contains the hydrogen bonds Hbond 10, Hbond 11 and Hbond 12. The interaction energy for this system was found to be -114.10 kJ mol^{-1}. This interaction energy is used as the reference for comparing interaction energies of the halogenated dG:dC systems with this non-halogenated form. The geometric and covalency properties of each interaction are disclosed in table 6.7.

Table 6.7: Geometric results and WBIs for the interactions found in N-dA:dT.

Interaction	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Hbond 10	1.92	0.74	176	0.0415
Hbond 11	1.93	0.74	176	0.0634
Hbond 12	1.80	0.69	178	0.0641

Hbond 10, Hbond 11 and Hbond 12 were all found to exist, and all of them probably contribute significantly to the overall stability of the dimer. In each case the internuclear separation is well within the sum of VdW radii, and with bond angles close to linear. The WBI values indicate a small degree of covalency. Based primarily upon the VdW ratios, Hbond 12 could be slightly stronger than the other two hydrogen bonds listed in table 6.7.

6.4.2.2 Results for singly halogenated dG:dC

This system contains the Xbond 10, Hbond 13 and Hbond 14 interactions. The geometric and covalency properties of these interactions are set forth in tables 6.8a, 6.8b and 6.8c respectively. From figure 6.2b it is immediately clear that Xbond 10 does not form in the case of fluorine, and therefore the quantitative data for the fluorinated system is not disclosed in table 6.8a (the halogen bond angle would be 88º). The quantitative data for Hbond 13 in the base pairs

containing a halogen heavier than chlorine are included in table 6.8b for completeness although the internuclear separation is clearly too great for substantial interaction involving the σ-hole to occur. The interaction energies for the cases of the halogen being chlorine, bromine, iodine and astatine are (in kJ mol^{-1}) respectively: -58.1, -57.5, -72.1 and -89.3. These correspond to energies relative to the interaction energy for N-dGdC (in the same order and units) of: +56.0, +56.6, +42.0 and +24.8. Hence even for astatine, S-dGdC is significantly less energetically favourable than N-dGdC. As fluorine does not participate in halogen bonding, the interaction energy for the lightest halogen is omitted.

Table 6.8a: Geometric and WBI results for interaction Xbond 10.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Cl	2.81	0.85	159	0.0262
Br	2.72	0.80	156	0.0698
I	2.59	0.73	169	0.1691
At	2.60	0.73	170	0.1820

From table 6.8a it appears that all of the listed halogens probably form halogen bonds with the CyN1 atom, with VdW ratios well below unity and plausible bond angles (albeit for bromine this is towards the low end). The only clear trend towards stronger halogen bonding in this data set is from bromine to iodine. At this point in the series there is an appreciable decrease in VdW ratio and relative energies as well as a marked increase in bond angle (towards linear) and WBI value, the latter implying greater covalent overlap. The results for astatine are very similar to those of iodine except for the interaction energy, with astatine significantly more energetically stable, although the relative energy in comparison with N-dGdC remains positive even in this case. Bromine does have a smaller VdW ratio than chlorine but the energetic and bond angle results do not fit the usual trend of increasing halogen bond strength going down group 17. It is clear that all of the energetic and geometric results for Xbond 10 are significantly impacted by the other interactions between the molecules. Based primarily on the halogen bond angles, it is likely that halogen bonding, whilst present for all systems (except the fluorinated case), only plays a dominant role where the halogen is either iodine or astatine.

Table 6.8b: Geometric and WBI results for interaction Hbond 13.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
F	1.95	0.65	152	0.0369
Cl	1.99	0.61	165	0.0304
Br	4.89	1.45	155	0.0000
I	4.43	1.27	155	0.0000
At	4.52	1.28	152	0.0000

From table 6.8b it is immediately clear that there is a qualitative change when descending the halogen group from chlorine to bromine. The halogen atom does not directly participate in this interaction but clearly has an impact upon it. The two principal factors that vary with halogen type that could have an effect on Hbond 13 are steric bulk, which increases down the group, potentially forcing the molecules further apart if other interactions are to be optimised, and electronegativity which decreases down the group, with the likely effect that the GX2 atom (in this case hydrogen) becomes less electron deficient. Both of these factors could be expected to weigh against the Hbond 13 interaction, and it might be that from chlorine to bromine a threshold is passed, whereafter the energetic favourability of forming the Hbond 13 hydrogen bond is outweighed by countervailing factors, including steric hindrance. Ascription of this change in behaviour to a driving force optimising the Xbond 10 halogen bond increasing down the group is undermined by (i) the weakness of the trend for the strength of Xbond 10 to increase down the halogen group, (ii) the greatest changes in the geometric and WBI results for Xbond 10 occurring from bromine to iodine rather than chlorine to bromine, and (iii) the offsetting ability of the heavier halogens to accommodate halogen bonds that depart further from linearity. Between fluorine and chlorine, the changes are modest, with a decrease in VdW ratio and an increase in bond angle but also a slight decrease in covalency as measured by the WBI. The trend towards smaller VdW ratio from bromine to iodine is noted but owing to these values being well in excess of unity (and other data in table 6.8b which make it clear that Hbond 13 does not exist for halogens heavier than chlorine), this trend is probably entirely due to factors that are unrelated to the interatomic interaction between GX2 and CyO2.

Table 6.8c: Geometric and WBI results for interaction Hbond 14.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
F	1.94	0.65	172	0.0324
Cl	4.85	1.48	158	0.0000
Br	2.13	0.63	160	0.0207
I	2.99	0.85	173	0.0023
At	3.18	0.90	175	0.0016

All of the systems, except the chlorinated case, appear to have formed hydrogen bond Hbond 14 (see Table 6.8c). Here the trend going from chorine to bromine is the exact opposite from that which was found for Hbond 13 (see table 6.8b), and the ability to form Hbond 14, which appears to be in competition with Hbond 13, could be responsible for the existence of the threshold for the formation of Hbond 13 described above. Whilst the steric effects arising from increasing halogen bond size would be contrary to that which is found to prevail in table 6.8c, there are two factors that must be taken into account: (i) as repeatedly observed herein the absolute internuclear distance of halogen bonds (as well as the VdW ratio) often decreases going down the halogen group and this would be a relevant consideration when the halogen bond distance of Xbond 10 is a factor (i.e. for all cases except fluorine), and (ii) the effect could be expected to be very similar upon Hbond 13 as upon Hbond 14, so if there is competition between

these two potential hydrogen bonds then the relevance of steric effects could be substantially diminished by this self-cancellation. By contrast, the electronic effects of the decreasing electronegativity going down the halogen group would be the opposite for Hbond 14 as for Hbond 13. In the case of Hbond 14, the effect of decreasing electronegativity of the halogen atom would be to increase the electron density on the GO4 oxygen atom, thereby making it a more potent donor of electron density to the CyX6 atom (in this case hydrogen), and hence increase the strength of the Hbond 14 interaction.

The ability of the fluorinated system to form Hbond 14 is attributable to the very different geometry that the dimer adopts when the halogen is fluorine, with the molecules approximately orthogonal rather than coplanar (see figure 6.2b). This alternative geometry might be adopted in order to facilitate both Hbond 13 and Hbond 14. The reason why the other systems do not adopt this orthogonal geometry is probably to accommodate the Xbond 10 halogen bond, which would be impossible with the geometry adopted by the fluorinated dimer. Fluorine would be unlikely to form the Xbond 10 halogen bond even if it were geometrically feasible to do so, due to fluorine's electronic properties, but for the other halogen atoms, even for chlorine, the ability to form one of Hbond 13 or Hbond 14 plus Xbond 10 appears to allow the coplanar geometry to form. It should be noted that a survey of the potential energy surface was not within the scope of this investigation, and starting geometries were chosen with a view to investigating halogen bonding; therefore, the fact that the orthogonal form of the non-fluorine halogenated dimers were not found does not mean that they do not form; no search was performed for them. It is therefore entirely possible that the orthogonal forms of the heavier halogenated dimers can form, and indeed may well be energetically more favourable than the coplanar geometries. What can be stated with a high degree of confidence is that the coplanar form does not form for the fluorinated S-dGdC system.

The trends from bromine to astatine are an increase in VdW ratio and a decrease in WBI value, but also, especially from bromine to iodine, an increase in bond angle. These results, especially the increase in VdW ratio (while remaining less than unity) and possibly resulting decrease in covalency, could be due to greater optimisation of the Xbond 10 halogen bond. The basis for drawing this link is not strong as a decrease in Xbond 10 internuclear distance might also be consistent with a decrease in Hbond 14 internuclear separation, and the increasing driving force towards halogen bond formation going down the halogen group is typically counterbalanced by a greater accommodation of non-ideal halogen bond geometry. However, once steric considerations are considered to be of very limited effect (as discussed above), it becomes difficult to find any other explanations for these trends in the results from bromine to astatine. A more comprehensive explanation for this trend may have to await further study.

6.4.2.2.1 Results from varying the halogen bond angle

The effect of varying the angle of Xbond 10 upon interaction energy is presented in figure 6.5. For the S-dG:dC system, no scan was successfully completed for the fluorinated variant. There were some technical complications in relation to the fluorinated case and, while these complications would probably not be insurmountable, the importance of obtaining results for

fluorinated S-dG:dC would probably be limited due to fluorine's lack of propensity to form halogen bonds. Results for the other variants of S-dG:dC and for N-dG:dC are presented in figure 6.5.

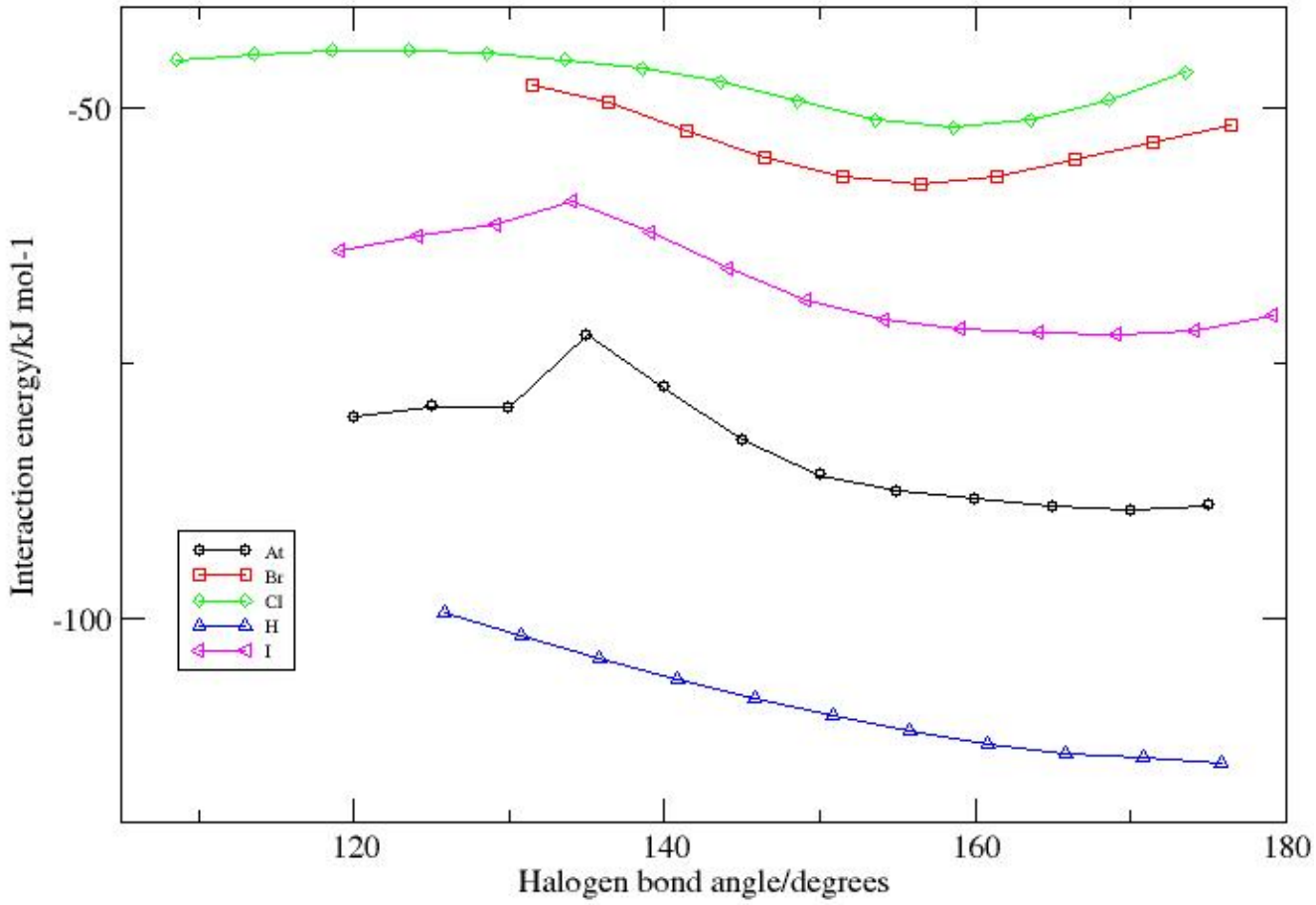


Figure 6.5: Interaction energy as a function of GN3 – GX3•••CyN1 bond angle. GX3 is Cl, Br, I, At or H as indicated in the legend.

For all bond angles that were scanned, a consistent order of interaction energy persisted, with this energy being more negative as the halogen group was descended, and N-dG:dC having the most negative interaction energy. There is a general qualitative trend towards greater disparity between the different variants of the S-dG:dC system, and between astatinated S-dG:dC and N-dG:dC, as the geometry of Xbond 10 becomes more linear. As can be seen in table 6.8a, the minimum in the interaction energy profile moves closer to linear from bromine to astatine, but not from chlorine to bromine. Furthermore, there is a substantial qualitative change in the position of the minimum from bromine to iodine. However, even for astatine the minimum occurs at 170°, indicating that secondary interactions have significant bearing upon the geometry. The minimum occurs closest to linear, at 176°, for N-dG:dC.

There is a sharp peak in the interaction energy for the iodinated and astatinated cases at 134° and 135° respectively. Figure 6.5 was produced by scanning from both directions and adopting the lower energy from the two scans. This method would typically smooth out graphs with sharp peaks; the persistence of this feature suggests that it arises from a particularly unfavourable aspect of the geometry of the system at this angle. This feature is probably due to an interaction between the lone pair of electrons on CyN1 and the negatively charged ring around the GX3 halogen atom. This repulsive interaction appears to outweigh what would be expected to be an attractive interaction between the σ-hole on GX3 and the CyO2 oxygen atom. This balance of the interactions can be explained by reference to the internuclear separations. These peaks correspond to points in the scan where there is a sudden increase in the internuclear distance as the halogen bond angle increases. For the iodinated system, as the halogen bond angle is increased the internuclear separation increases from 3.03 Å to 3.67 Å. For the astatinated system there is an increase in internuclear separation from 2.92 Å to 3.58 Å upon increasing the halogen bond angle to the value thereof that corresponds to the peak in the interaction energy profile. Both of these increases in internuclear separation occur upon increasing the halogen bond angle by 5° to the angle corresponding to the peak in the interaction energy profile as shown in figure 6.5. The VdW radius of a nitrogen atom and an oxygen atom are very similar to each other, 1.55 Å and 1.52 Å respectively. The GC3 – GX3 – CyO2 angle, at the above stated GC3 – GX3 – CyN1 angles, is 172° and 174° for the iodinated and astatinated cases respectively. For the iodinated case this gives a GX3 – CyO2 VdW ratio of 1.05, while in the astatinated case the VdW ratio is 0.73. Hence, in the astatinated case a halogen bond could be expected to form, but there is no evidence of it from the interaction energy profile shown in figure 6.5. Perhaps surprisingly, this interaction does not even appear to have offset the repulsion between GX3 and CyN1 more where the former atom is astatine than where it is iodine. Even a possible hydrogen bond between GO4 and CyO6, 1.98 Å and 1.97 Å for the iodinated and astatinated systems respectively, does not appear to have been capable of overcoming the GX3 – CyN1 repulsion. Further reducing the angle leads to the interaction energies becoming more negative again, probably because the resulting increase in GX3 – CyN1 distance outweighs the effect of further reducing the GC3 – GX3 – CyN1 angle, even though the latter would otherwise be expected to increase overlap with the negatively charged ring on GX3. The angle in question is equiangular between linear and orthogonal, suggesting that even for astatine, which has a relatively large σ-hole, the positive charge is largely concentrated relatively close to linear. For the chlorinated and brominated cases these sharp peaks do not form but the rate at which the interaction energy becomes less negative with decreasing GC3 –GX3 – CyN1 angle appears to decrease, suggesting some possible effect from this interaction. For N-dG:dC, the interaction energy continues to increase at an accelerating (or at least constant) rate. These latter observations provide some evidence that this feature of the interaction energy profile for the iodinated and astatinated cases is due to this GX3 – CyN1 interaction.

In contrast with S-dA:dT and N-dA:dT (see figure 6.4), the S-dG:dC and N-dG:dC interaction energy profiles do not intersect for any of the variants of the S-dG:dC system. Put another way, halogenation, even by astatine, leads to the interaction between guanine and cytosine becoming less attractive, even at near-linear angles, which favour halogen bonding. This marked

qualitative difference in the results between dA:dT and dG:dC (at least for the heavier halogens) demonstrates the effect of the overall chemical environment upon the strength of halogen bonding, relative to a corresponding hydrogen bond in an analogous molecular system.

6.4.2.3 Results for doubly halogenated dG:dC

The doubly halogenated dG:dC system potentially contains the halogen bonds Xbond 11, Xbond 12 and Xbond 13, and the hydrogen bond Hbond 15. The geometric results and WBI values for each of these interactions are reported in tables 6.9a, 6.9b, 6.9c and 6.9d respectively. The fluorinated case is discussed separately from the other systems as it clearly forms a different geometry (see figure 6.2b), and the geometric and WBI results are stated outwith the tables. Furthermore, for the same reasons as were given in respect of S-dGdC, no interaction energy was calculated for the fluorinated case. The interaction energies for the chlorinated, brominated, iodinated and astatinated cases are respectively (in kJ mol^{-1}): -48.25, -47.77, -79.64, and -118.22, and hence, in the same order and units, the interaction energies relative to the interaction energy for N-dGdC are: +65.85, +66.33, +34.46, and -4.12. Hence it can be seen that with the exception of astatine, all of the halogenated systems are less energetically favourable than the all-hydrogen case. Uniquely, the astatinated dimer has an interaction energy that is more negative than the non-halogenated form of the system. Another observation that can be made is that there is a trend towards greater stability from bromine to astatine, which is consistent with a substantial role for halogen bonding in the overall stability of the structure. This trend does not however appear to hold for chlorine.

Table 6.9a: Geometric and WBI results for interaction Xbond 11.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Cl	2.94	0.89	156	0.0158
Br	3.02	0.89	159	0.0235
I	3.25	0.92	142	0.0144
At	3.20	0.90	138	0.0186

All of the VdW ratios listed in table 6.9a are less than unity, and the WBI values stated therein imply a small degree of covalency in each case, but again no discernible trend. There is no trend towards smaller VdW ratios going down the group, and the bond angles are small for halogen bonds, probably implausibly small for iodine and astatine despite the greater ability of the heavier halogens to form non-linear halogen bonds. There may be an incidental contribution from Xbond 11 for chlorine and bromine, but this is likely to be modest. In the cases of iodine and astatine, contribution of Xbond 11 to the overall stability of the system is likely to be minimal, in view of the halogen bond angle.

Table 6.9b: Geometric and WBI results for interaction Xbond 12.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Cl	3.27	1.00	113	0.0013
Br	3.19	0.95	158	0.0112
I	2.74	0.78	175	0.0838
At	2.61	0.74	178	0.1406

From table 6.9b (and figure 6.2b) there is clearly a qualitative change going from bromine to iodine, and from chlorine to bromine. For chlorine, there is clearly no halogen bond. The bond angle is closer to perpendicular than linear, so whatever σ-hole effect here might be would probably be destabilising (although an electrostatic potential plot would be necessary to confirm that hypothesis). The VdW ratio is equal to unity (to two decimal places). For bromine, it is likely that Xbond 12 makes a modest contribution to the stability of D-dG:dC. The VdW ratio is less than unity, albeit not by very much, and the bond angle is within the bounds or plausibility for halogen bonding. The WBI value for chlorine indicates a small degree of covalency. By contrast, Xbond 12 appears to play a dominant role in the stability of D-dG:dC for the iodinated and astatinated systems, with VdW ratios well below unity, bond angles close to linear and (compared with other systems investigated in this study) quite high WBI values, indicating a relatively high degree of covalency. By all of these measures the halogen bonding appears to be slightly more pronounced in the case of astatine than for iodine. The apparent trend towards a greater role for Xbond 12 going from bromine to astatine coincides with a trend towards increasingly negative interaction energies. The reason why this coincidence does not extend to chlorine is discussed after consideration of Hbond 15.

Table 6.9c: Geometric and WBI results for interaction Xbond 13.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Cl	4.66	1.43	144	0.0006
Br	2.96	0.88	156	0.0131
I	2.89	0.83	151	0.0280
At	2.67	0.75	156	0.0716

From the VdW ratio and bond angle shown in table 6.9c it is clear that the chlorinated system does not form Xbond 13. However, bromine, iodine and astatine do all appear to form this halogen bond, albeit with sub optimal halogen bond angles. All VdW ratios for these heavier halogens are well below unity, and while 151^{o} (as was found for the brominated system) is small for a halogen bond it is not entirely beyond the realm of plausibility. From bromine to astatine there is a trend towards smaller VdW ratios and increasing WBI value, which respectively

indicate a trend towards stronger halogen bonding and greater covalency going down the halogen group, as would be expected, especially in the case of the former, if these geometries are being influenced by Xbond 13, rather than arising entirely by virtue of optimisation of other interactions.

Table 6.9d: Geometric and WBI results for interaction Hbond 15.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Cl	1.98	0.76	163	0.0302
Br	3.96	1.52	158	0.0001
I	4.47	1.71	144	0.0001
At	4.57	1.75	141	0.0000

From table 6.9d (and figure 6.2b) is it very clear that Hbond 15 does not exist for any halogens heavier than chlorine. The very large VdW ratios, on their own would be enough to come to this conclusion and all of the other data in table 6.9d further support it.

In the case of chlorine, neither Xbond 11 nor Xbond 12 appeared to be particularly strong, and Xbond 13 does not form at all. However, nonetheless the interaction energy for the system was found to be slightly more negative for chlorine than for bromine. This stability of the chlorinated system can therefore be largely, although perhaps not exclusively, ascribed to Hbond 15. By contrast the brominated system appears to be stabilised by a combination of the Xbond 11, Xbond 12 and Xbond 13 halogen bonds, while for the iodinated and astatinated systems, Xbond 12 probably dominates, albeit with a modest contribution from Xbond 13. This is reflected in the geometries that each of these systems adopt, as presented quantitatively in tables 6.9a-d, and graphically in figure 6.2b. The greater role of halogen bonding in the cases of the heavier halogens should not be surprising due to the strengthening of the σ-hole effect going down group 17 of the periodic table, as extensively discussed elsewhere herein. Descending the group from chlorine to bromine appears to entail crossing a threshold whereafter the Hbond 15 hydrogen bond is outcompeted by the aggregate of the Xbond 11 and Xbond 12 halogen bonds.

The explanation for the trend towards exclusive dominance of Xbond 12 (and to a minor extent Xbond 13) at the expense of Xbond 11 going from bromine to astatine is not straightforward. Two possible explanations, either alternately or, more likely, in combination appear to be plausible. Optimisation of Xbond 11 without concurrently undermining the strength of the other two halogen bonds might be easier for chlorine than for the heavier halogen atoms. The greater steric bulk of the heavier halogen atoms could point in that direction, although the shortening of internuclear distances in the case of halogen bonding (typically in absolute terms as well as relative to the sum of VdW radii) could be expected to limit this effect. The other possible explanation might arise from the electronic effects of the halogen atom CyX6 on other substituents on the aromatic ring and heteroatoms within the ring. This trend is also consistent with he increasing electropositivity going down the halogen group. While this could be expected to enhance the electron donating capacities of both CyO2 and CyN1, this effect might not be

shared equally between these atoms. However, a satisfactory explanation for this trend, which could be due to a combination of steric and electronic effects, remains elusive.

6.4.2.3.1 The fluorinated case

Fluorinated D-dG:dC does not form halogen bonds. Furthermore, it does not fall into the general patterns observed for the heavier halogens, instead forming its own unique geometry. As this case is of limited relevance to the rest of this study, the results are not explored in detail. There appear to be two relevant hydrogen bonds, one between GH2 and CyO2 and another between GX2 (hydrogen) and CyN1. They have distances of 2.21 Å and 2.25 Å respectively, giving a VdW ratio 0.85 in both cases. Their respective WBI values are 0.0126 and 0.0150. As may be seen from figure 6.2b, the atoms GH2, GN2, GX2, CyN1, CyC2 and CyO2, including the hereinbefore described hydrogen bonds between GH2 and CyO2 and between GX2 and CyN1, appear to form a six-member ring. The electronic and other attributes of that observed geometry are beyond the scope of this investigation, although it can be noted, based upon the above stated WBI values, that the hydrogen bonded components do not have a particularly high degree of covalency.

6.4.2.4 Results for triply halogenated dG:dC

Triply halogenated dG:dC (T-dG:dC), in principle, contains the following halogen-bonded interactions: Xbond 14, Xbond 15, Xbond 16 and Xbond 17. The fluorinated case forms a totally different (stacked) geometry compared with all of the other halogen types and the interactions which form, (it does form a stable dimer), are not of significant interest in the context of this study, and therefore the discussion is confined to cases of halogenation with halogens heavier than fluorine. For all halogen types, Xbond 14 plainly does not exist due to very large internuclear separation and the GX2 halogen atom was substantially out of plane, forming a geometry that was totally incompatible with the Xbond 14 halogen bond (see figure 6.2b). The results for Xbond 15, Xbond 16 and Xbond 17 are presented in tables 6.10a, 6.10b and 6.10c respectively. The interaction energies for the chlorinated, brominated, iodinated and astatinated T-dG:dC systems are respectively, in kJ mol^{-1}: -26.65, -39.44, -75.30, and -114.30, and hence relative to N-dG:dC (in the same order and units): +87.45, +74.66, +38.80, and -0.20. From these results it can be seen that there is a trend towards more negative interaction energies as the halogen group is descended and that, with the exception of astatination, triple halogenation results in dimer pairs with weaker interaction than for N-dG:dC. The astatinated case has a very slightly more negative interaction energy than N-dG:dC.

Table 6.10a: Geometric and WBI results for interaction Xbond 15.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Cl	2.94	0.89	164	0.0199
Br	3.08	0.91	156	0.0200
I	3.26	0.92	140	0.0128
At	3.22	0.90	136	0.0167

From the results set forth in table 6.10a it is clear that there is a major change in bond angle between the brominated and iodinated cases. The VdW ratios are very similar for all halogen types. The WBI values are almost identical for chlorinated and brominated T-dG:dC, with the iodinated and astatinated cases being slightly lower. In all cases the WBI values imply little covalency. This combination of results is perhaps surprising: despite the flexibility of the heavier halogens, the bond angles for iodine and astatine are very low for halogen bonding, yet their VdW ratios are about the same as for chlorine and bromine; indeed for astatine it is slightly smaller than for bromine. The presence of some covalent overlap, indicated by a non-zero WBI value in each case, combined with a VdW value significantly below unity, suggests that there is a real interaction taking place. However, there is no trend in any of the results that would be consistent with a substantial role for halogen bonding. Furthermore, halogen bonds with angles of 140^{o} and 136^{o} are barely credible, even though they pertain to iodine and astatine respectively. At least for these heavier halogens, the fact that the VdW ratios are below unity and there is some minor covalent overlap is most likely to be a side effect of optimising the structures' other interactions. For the chlorinated cases, the combination of VdW ratios well below unity, some covalent overlap and bond angles which, whilst small, are within the bounds of plausible halogen bond angles, would suggest at least a strong *prima facie* case for halogen bonding genuinely contributing towards the stability of the system. However, the very similar VdW ratios to those found for iodine and astatine, for which halogen bonding is unlikely to be present, does cast some doubt upon the *bona fide* nature the Xbond 15 interaction for chlorine and bromine.

Table 6.10b: Geometric and WBI results for interaction Xbond 16.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Cl	3.29	1.00	154	0.0045
Br	3.17	0.93	162	0.0133
I	2.69	0.76	176	0.0961
At	2.58	0.72	178	0.1528

From table 6.10b, a clear trend towards stronger halogen bonding going down the halogen group can be observed. Furthermore, there appears to be a substantial change from brominated

T-dG:dC to iodinated T-dG:dC. Although, as discussed, there is not sufficient data to truly rationalise the results shown in table 6.10a, the observation that the apparent qualitative change in the results occurs between bromine and iodine for both Xbond 15 and Xbond 16 might not be pure happenstance. For Xbond 16, the VdW ratio for chlorine is 1.00 (i.e. unity to two decimal places). For the other halogens the internuclear separation is well within the sum of VdW radii, especially for iodine and astatine. There is also a strong trend towards linear bond angle going down the halogen group, again with a marked increase from bromine to iodine. While the WBI values are low for chlorine and bromine, they are appreciably greater for iodine and astatine, the latter whereof is, compared with other systems in this study, relatively high, indicating a significant covalent contribution to the interaction. The trend towards geometric properties that are associated with stronger halogen bonding matches the trend in interaction energies for the T-dG:dC system, including the relatively minor change in values when going from chlorine to bromine. The data presented in table 6.10b is consistent with Xbond 16 being the dominant varying factor behind the trend in the interaction energies. A formal halogen bond appears to be present for, at least all halogens heavier than chlorine, and possibly for chlorine itself. Though of limited value, the VdW ratio for chlorine is 0.997; if this number were assumed to be accurate to that degree of precision (a somewhat bold premise) then it could be said that there is a formal halogen bond in the case of chlorine.

Table 6.10c: Geometric and WBI results for interaction Xbond 17.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
Cl	2.88	0.88	161	0.0101
Br	2.93	0.87	157	0.0134
I	2.91	0.83	150	0.0266
At	2.68	0.76	156	0.0703

Table 6.10c discloses that the Xbond 17 geometry is consistent with halogen bonding for all halogen types not lighter than chlorine, albeit that the bond angle is towards the low end for halogen bonding, especially for iodine. The combination of decreasing VdW ratio, increasing WBI value but decreasing bond angle (as far as iodine) suggests that halogen bonding might be gaining in strength despite departing further from linearity. There appears to be a correlation between decreasing VdW ratio and increasing covalency, with only minor difference between chlorine and bromine for each of these properties. The ability to form stronger halogen bonds despite further departure from the ideal angle could be a reflection of the greater ability of the heavier halogen atoms to accommodate less linear bond angles, due to the σ-hole occupying a larger area of the atom.

For iodine and astatine, optimisation of the Xbond 16 bond angle appears to take precedence over the optimisation of the Xbond 17 bond angle; however, it is for these heavier elements that the VdW ratios are larger for Xbond 17 than for Xbond 16. The reversal in the general trend for bond angles from iodine to astatine for Xbond 17 might reflect a greater ability to simultaneously

optimise Xbond 16 and Xbond 17. Although Xbond 15 might provide some additional background stability, especially for chlorine and bromine, the principal drive appears to be optimisation of Xbond 16 and Xbond 17. This is reflected in the improbably small bond angles for Xbond 15 for the iodinated astatinated systems.

6.4.2.5 Results for doubly halogenated dG:dC where at least one halogen is astatine

This system wherein the CyX6 atom is always astatine, forms the following interactions: Xbond 18, Xbond 19, Xbond 20, and Hbond 16. Except in the fluorinated case, the monomers are approximately coplanar. The geometric results and WBI values for each of these interactions are presented in tables 6.11a, 6.11b, 6.11c, and 6.11d respectively. The respective interaction energies for the fluorinated, chlorinated, brominated, iodinated, and doubly astatinated variants of this system, At-dG:dC, were found to be, in kJ mol^{-1}: -87.72, -59.52, -67.88, -94.38, and -118.22. Relative to N-dG:dC, the energies are (in the same order and units): +26.38, +54.58, +46.22, +19.72, and -4.12. The fluorinated form of At-dG:dC has a stronger attractive interaction than the chlorinated or brominated forms thereof, probably due to the electron withdrawing effect leading to a more electron deficient GX2 (hydrogen) atom, and hence probably a stronger Hbond 16 interaction. This effect appears to be outweighed in the cases of iodine and astatine, probably due to stronger halogen bonding. Only the doubly astatinated form of the system exhibited an interaction energy that was more negative than for N-dG:dC, providing another demonstration of the strength of halogen bonds that can be formed by the heaviest of the natural halogen elements.

Table 6.11a: Geometric and WBI results for interaction Xbond 18.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
F	3.34	1.11	109	0.0011
Cl	3.05	0.92	162	0.0141
Br	3.20	0.94	151	0.0131
I	3.23	0.92	140	0.0153
At	3.20	0.90	138	0.0186

It is immediately clear from all of the values for the fluorinated case that F-AtdG:dC does not form Xbond 18. It is plausible that chlorine does form the Xbond 18 interaction with CyN1, and that possibly also extends to bromine. However, despite having an internuclear separation within the sum of VdW radii and a small degree of apparent covalent overlap, the very small bond angle weighs heavily against iodine or astatine forming this interaction, although there might be a possibility of some halogen bond type interaction to minor degree as these elements are more accommodating of bond angles that depart substantially from linearity.

Table 6.11b: Geometric and WBI results for interaction Xbond 19.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
F	3.70	1.24	106	0.0004
Cl	3.08	0.94	154	0.0092
Br	2.97	0.88	166	0.0274
I	2.68	0.77	176	0.1031
At	2.61	0.74	178	0.1406

From table 6.11b it is clear that fluorine does not form a halogen bond with CyO2, the bond angle in particular is incompatible with the presence of that interaction. It also has a VdW ratio well in excess of unity. All of the other halogen elements do appear to form the Xbond 19 interaction with CyO2. Furthermore, there is a clear trend towards greater optimisation of this interaction as the halogen group is descended. The greatest changes in VdW ratio and WBI value are from bromine to iodine, while the greatest increase in bond angle is from chlorine to bromine, with bromine to iodine coming a close second. The greater optimisation of this interaction, especially in relation to VdW ratios, indicates that the relative importance of this interaction increases as the halogen atom type becomes heavier. Changes in bond angle are also significant, but the greater ability of the heavier halogens to accommodate less ideal bond angles makes this aspect of the results more difficult to interpret. The marked increase in covalency for iodine and astatine suggests that the covalent component of the interaction might become more dominant for these elements, compared with the lighter halogen group members.

The results for Xbond 19 are qualitatively similar to those for the analogous Xbond 12 in D-dG:dC, see table 6.9b. At-dG:dC, where both halogen atoms are astatine, is identical to the astatinated D-dG:dC structure. The presence, in the case of At-dG:dC of astatine at CyX6 for all variants of the system can be seen to have resulted in a slightly shorter GX3 – CyO2 internuclear separation for chlorine and bromine at GX3, while for the iodinated cases, the internuclear separation is very marginally greater for the At-dG:dC system. In all cases the presence of astatine at CyX6 in place of a lighter halogen atom resulted in a slight increase in covalency, as indicated by the WBI values. The most pronounced effect that astatination in all instances at CyX6, while the other halogen varies, appears to have had is upon the bond angle, especially for chlorine. Whereas the bond angle of 113° ruled out the presence of Xbond 12 for chlorinated D-dG:dC, the corresponding angle of 154° makes this interaction plausible for At-dG:dC, albeit with a halogen bond angle at the lower end of the spectrum. The VdW ratio for chlorine has also been brought clearly below the threshold of unity. The same qualitative effect can be seen for the other halogen elements, with a clear diminishing of the effect as the halogen group is descended. The most likely explanation for the differences in the results for Xbond 19 compared with Xbond 12 is the increase in electropositivity of the elements as the halogen group is descended. The smaller the propensity of the CyX6 atom to concentrate electron density upon itself, the greater the amount of electron density there is likely to be on the other atoms in the cytosine molecule, including CyO2, which in turn would be expected to result in greater electron

density donation into the σ-hole on GX3. The greater covalency of Xbond 19 compared with Xbond 12 is also consistent with a greater electron density on the CyO2 atom. Furthermore, this explanation is consistent with the trend towards a greater effect of the change in the identity of the CyX6 atom as the halogen group is ascended.

The rationalisation of the apparent increase in halogen bond strength for Xbond 19 compared with Xbond 12 could also be expected to apply to the comparison between Xbond 18 (see table 6.9a) and Xbond 11 (see table 6.11a) however the lack of strong Xbond 11 and Xbond 18 interactions with a clear and easily explainable trend in halogen bond strength down the halogen group makes a comparison and rationalisation of the differences between these systems much more problematic. Where there is a clear presence of halogen bonding (Xbond 12), substitution by astatine of a lighter halogen atom at CyX6 accentuates this interaction (Xbond 19).

Table 6.11c: Geometric and WBI results for interaction Xbond 20.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
F	2.78	0.93	160	0.0656
Cl	2.83	0.87	154	0.0450
Br	2.77	0.82	156	0.0512
I	2.72	0.78	155	0.0597
At	2.67	0.75	156	0.0716

Xbond 20 can be distinguished from the other halogen bond interactions of At-dG:dC by the consistency of the halogen atom that is directly engaged in the interaction. Here the only set of data that shows a clear and consistent trend is the VdW ratio, with each step in the descent down the halogen group producing a very approximately equal decrease in the VdW ratio; changes vary between 0.03 Å and 0.06 Å. The bond angle, perhaps with a slight exception for fluorine, appears to be indifferent to the identity of the halogen element present at GX3, while the WBI also shows only a very modest increase from chlorine to astatine, and decrease from fluorine to chlorine. The VdW ratios being smaller than unity, plausible (although small) halogen bond angles, and the presence of some covalency suggests that Xbond 20 does exist and makes some contribution towards the overall stability of At-dG:dC. The effect upon the electron density of decreasing electronegativity at GX3 as the halogen group is descended, and hence electron donating potency of the GO4 atom, could be expected to be qualitatively the same as increasing the atomic number of the halogen atom at CyX6 upon CyO2 (see the discussion comparing Xbond 19 with Xbond 12). Hence for the same reasons as set forth in connection with that comparative discussion, placing a heavier halogen atom at GX3 would be expected to increase the strength of the Xbond 20 halogen bond between GO4 and CyX6. This outcome appears to be reflected in the trend in VdW ratios for Xbond 20, which decrease as the halogen group is descended.

However, caution is urged in relation to ascription of causation of this trend as a decrease in VdW ratios, which are a reflection of internuclear separation, may well decrease for Xbond 20

as the effect of decreasing internuclear separation between GX3 and CyO2 (Xbond 19) is to draw the two molecules (each as a whole) closer together, reducing internuclear separation between atoms in guanine and atoms in cytosine in general. The lack of a clear trend in the bond angle certainly weighs against the proposition that Xbond 20 makes a larger proportional contribution to the stability of the system going down the group. By contrast there is a clear trend towards increasing linearity going down the group for Xbond 19, with chlorine having the same bond angle for Xbond 19 and Xbond 20. It is probably the case that Xbond 19 and Xbond 20 both pull in the same overall direction as to the geometry of the most stable geometry of the dimer. Both probably contribute to that stability, but, especially for the heavier halogen elements, it appears to be Xbond 19 that performs the dominant role in determining the geometry.

Special mention ought to be made in respect of the fluorinated case, as the bond angle for Xbond 20 is significantly closer to being linear than for the other halogen elements. In this case Xbond 20 is likely to be performing a greater role in determining the structure of the system, due to the inability of GX3 (fluorine) to form either Xbond 18 or Xbond 19 (see tables 6.11a and 6.11b). However, the halogenated bond angle, even in the fluorinated case, is far from the ideal halogen bond angle, and it is likely that the Hbond 16 hydrogen bond (see table 6.11d) performs a major role in the interaction for At-dG:dC where the non-astatine halogen is fluorine, also for want of halogen bonds wherein fluorine directly participates.

Table 6.11d: Geometric and WBI results for interaction Hbond 16.

Halogen present	Internuclear distance/Å	VdW ratio	Bond angle/◦	WBI
F	1.94	0.74	149	0.0400
Cl	3.53	1.35	163	0.0004
Br	4.06	1.56	154	0.0000
I	4.43	1.70	144	0.0001
At	4.57	1.75	141	0.0000

From table 6.11d it is clear that only in the fluorinated case does the GX2 atom (hydrogen) form a hydrogen bond with the CyO2 atom. For the other halogen elements, the VdW ratios make patent that there is no formal halogen bond or even a realistic possibility of substantial halogen bonding type stabilisation short of a formal halogen bond. There is also no (or negligible) covalency.

There are four factors that most likely explain the formation of Hbond 16 uniquely by fluorine: (i) lack of competition from Xbond 18 and Xbond 19; (ii) the effect of fluorine's extreme electronegativity upon the GX2 (hydrogen) atom; it will likely make the GX2 atom much more electron deficient compared with the cases with a heavier halogen atom at GX3, and hence more strongly attracted to the electron rich CyO2 oxygen atom; (iii) the relatively small atomic radius of fluorine compared with other halogen elements makes Hbond 16 significantly more sterically viable. The sterically larger halogen elements have two principal effects that undermine Hbond 16 formation: (a) increased GC3 – GX3 bond length could be expected to increase GX2 – CyO2

bond length by driving the rings of the respective molecules further apart (off-set by shorter halogen bond distances), and (b) steric obstruction of Hbond 16 formation; and (iv) the trend down the halogen group towards favouring Xbond 19 more heavily has the side effect of (a) further driving GX2 away from CyO2 and (b) placing the GX3 halogen atom where it can more effectively hinder interaction between GX2 and CyO2. The geometric factors can be more readily appreciated when table 6.11d is read in conjunction with figure 6.2b. The trends going down the halogen group from chlorine to astatine for VdW ratios and bond angles are most likely a side effect of factors (iii) and (iv) described above.

6.5 Conclusions

This investigation into halogenated DNA base pairs identifies a number of features about each of the interactions, whose principal features are tabulated above. Attempts have been made to draw conclusions for each interaction, with appropriate comparisons to other interactions in the same system and other systems in this study, in the course of discussion of the results.

For S-dA:dT, the interaction energy is slightly more negative (by 0.6 kJ mol^{-1}) for the iodinated case compared with N-dA:dT, and significantly more negative (by 17.7 kJ mol^{-1}) for the astatinated case. For all other halogen types, the interaction energy was found to be more strongly negative (attractive) for N-dA:dT (i.e. without halogenation). In summary, the relative significance of Xbond 1 versus Hbond 4 increases as the halogen group is descended, especially from bromine to iodine. This is to be expected from the trend towards stronger halogen bonding going down the halogen group.

Xbond 3 is the dominant interaction for D-dA:dT for all of the halogens from chlorine to astatine inclusive. Energetic results indicate that in the case of double substitution all interaction energies are less negative than for N-dA:dT, although there is a trend towards more strongly attractive interaction going down the halogen group, with astatinated D-dA:dT being only 3.2 kJ mol^{-1} less negative than for the non-halogenated system. Hence it can be seen that although astatine (and by a very small margin iodine) forms a more strongly attractive interaction than for N-dA:dT when only one hydrogen atom is substituted, this trend is not maintained in the case of double substitution. It is certainly not a general rule that astatine always forms a stronger bond than hydrogen.

When a third hydrogen atom is substituted by a halogen atom, astatine again forms the stronger attractive interaction compared with the all-hydrogen case, with a relative decrease (more negative) of the interaction energy by 24.0 kJ mol^{-1}. By contrast the triply iodinated system is less negative than N-dA:dT by 6.0 kJ mol^{-1}. For all halogen types, the interaction energy itself was negative, suggesting that fully halogenated analogues of the adenosine – thymine dimer can be formed, for all halogen types. For T-dA:dT all of the noteworthy interactions are halogen bonds. Xbond 4 and Xbond 6 appear to be the most significant interactions (tables 6.5a and 6.5c respectively). Based on the VdW ratios, Xbond 4 would appear to be the slightly greater factor, but for the iodinated and astatinated cases, the bond angle is clearly most optimised for Xbond 6. The astatinated system exhibits a relatively large degree of covalency for Xbond 6 (WBI value: 0.1861).

In the At-dA:dT system, Xbond 8, which is a halogen bonding interaction in which the halogen atom is astatine in all cases, totally dominates the overall intermolecular interaction except where both halogen atoms are astatine, in which case Xbond 9 is plainly the most significant interaction. Hence it can be seen that where the halogen atoms are both astatine (i.e. the same halogen) Xbond 9 is significantly more favoured, but this preference is totally overturned by the disparity in the halogen bonding properties of the other halogen elements, even if the comparison is between astatine and iodine. The interaction Xbond 9 is analogous to Xbond 3 in

the D-dA:dT system, and as noted above, Xbond 3 dominates for all halogen types, in a system where there is always the same halogen element involved in Xbond 3 and Xbond 2, the latter being analogous to Xbond 8. Hence it is demonstrated that it is not only for astatine that the AN3 – TX1 interaction is favoured over the AX4 – TO2 interaction when the same halogen element is involved in each of these interactions.

In the S-dG:dC system, the iodinated and astatinated variants thereof show clear halogen bonding, Xbond 10. This interaction has a significant degree of covalency. The fluorinated, chlorinated and brominated analogues each behave differently from each other as well as from iodine and astatine. Fluorinated S-dG:dC forms a substantially non-planar structure in which both Hbond 13 and Hbond 14 play a role. The chlorinated form of the system appears to be stabilised by both Xbond 10 and Hbond 13, but not (at least significantly) by Hbond 14. By contrast, the brominated system does appear to exhibit Xbond 10 and Xbond 14, but not Xbond 13. For both the chlorinated and brominated cases the stabilising contribution of Xbond 10 is likely to be limited due to the small bond angles, 159^{o} and 156^{o} respectively. The different behaviour of the chlorinated and brominated variants of the system is not straightforward to explain. One possibility might be that the greater steric bulk of the bromine atom impedes the interaction between GX2 (hydrogen) and CyO2, such that the weakening of that interaction results in it being outcompeted by Hbond 14. For all variants of S-dG:dC, the interaction energy is less strongly negative than for N-dG:dC. There is a clear trend towards more negative interaction energies from bromine to astatine. Even for the astatinated system, the interaction energy is 24.8 kJ mol^{-1} less negative than for N-dG:dC. Figure 6.5 shows that for the full range of the scans, there is a constant order between the different forms of the system as to interaction energy, with interaction energies becoming more negative down the halogen group, and the all hydrogen case being the most negative.

The chlorinated form of D-dG:dC appears to be mainly stabilised by Hbond 15, with perhaps some contribution from Xbond 11. Variants of this system wherein the halogen atom is heavier than chlorine appear to be stabilised by the three halogen bonds. For the brominated analogue, Xbond 11, Xbond 12 and Xbond 13 each appear to be make comparable contributions. The iodinated and astatinated cases do not appear to form Xbond 11. As in the case of S-dG:dC, there is a strong trend towards more strongly negative interaction energies from bromination to astatination, which correlates with the strengthening of the Xbond 12 interaction. The astatinated form of the system exhibit a stronger attractive interaction than N-dG:dC, by 4.1 kJ mol^{-1}. This finding might be surprising as substitution by one astatine atom produces a less strongly negative interaction energy than N-dG:dC, but the dependence of the order of stability upon the number of halogenations with the same halogen atom demonstrates the importance of the surrounding chemical environment to the strength of halogen bonding in comparison with hydrogen bonding. There is a clear trend towards greater covalency going down the halogen group for Xbond 12 and Xbond 13. The fluorinated analogue displays a quite different geometry that cannot easily be compared with the other forms of D-dG:dC.

T-dG:dC, in its astatinated form, has an almost identical interaction energy to that of N-dG:dC. In this system there is a very clear trend towards more negative interaction energies as the

halogen group is descended from chlorine to astatine; the rate of decrease in interaction energy accelerates when going from bromine to iodine. For the chlorinated analogue, Hbond 15 probably makes some stabilising contribution, VdW: 0.89, angle 164°. This stabilisation could also be a possibility for the brominated system, VdW: 0.91, angle 156°, although the relatively small bond angle suggests that the effect of this interaction could be limited. Xbond 17 (see table 6.10c) could make some contribution towards the stability of all of these systems for all halogens from chlorine to astatine, especially for the cases involving the lighter halogens. The interaction that is most significant for the iodinated and astatinated cases appears to be Xbond 16 (see table 6.10b). These interactions also appear to have a relatively high degree of covalency (respective WBI values: 0.0961 and 0.1528).

In the At-dG:dC system, only the doubly astatinated variant has an interaction energy that is below that of N-dG:dC. The brominated case has the highest (least negative) interaction energy, with interaction energy decreasing from bromine to fluorine as well as from bromine to astatine. This trend indicates a competition between at least two factors in determining the overall stability of the At-dG:dC system. Xbond 18 and Xbond 19 engage the variable halogen atom, while Xbond 20 engages the atom that is in all cases astatine. The fluorinated variant of the system does not form Xbond 18 or Xbond 19 and is the only analogue that forms Hbond 16, probably due to a lack of competing halogen bonding opportunities. As explained above, the stabilising role of Xbond 20, present in all cases, needs to be viewed with some degree of caution when analysing the geometric results as there is evidence that the trend in the result could be incidental. The contribution from Xbond 18 is substantially smaller than that of Xbond 19, based upon bond angles, with the former probably making no net stabilising contribution for iodine and astatine. There is a clear trend towards decreasing Xbond 20 VdW ratios as the non-astatine halogen atom descends the halogen group. This trend implies a strengthening of the halogen bond, and is probably explained by the decreasing electron withdrawing effect of the GX3 halogen atom; the smaller the electron withdrawing effect, the greater the electron density that can be expected to reside on the GO4 atom engaged in Xbond 20. There is also a trend towards greater covalency from chlorine to astatine. However, there is no clear trend in the bond angle, which in all cases is within the range 154°-160°. This may be a result of an increasing driving force to optimise this halogen bond being counterbalanced by a greater tolerance of non-ideal bond angles, as the halogen group is descended. In contrast with Xbond 20, Xbond 19 does show a clear trend in halogen bond angle, with the iodinated and astatinated cases coming close to linear, with angles of 176° and 178° respectively. The VdW ratios also show a clear trend towards smaller ratios as the group is descended. This apparent strengthening of the halogen bond coincides with greater covalency of the interaction. For the heaviest two halogens at GX3, it is clear that this interaction is dominant, however for the chlorinated case the VdW ratios and the bond angles both suggest that Xbond 20 could be the stronger interaction, while for the brominated case, the VdW ratio is also smaller for Xbond 20. This observation suggests that where the halogen atoms at GX3 and CyX6 are the same element, then Xbond 19 has an inherent advantage over Xbond 20; however, this can be outweighed by placing a lighter halogen atom at GX3.

In summary, the results presented above demonstrate the complex nature of halogen bonding. The most significant factor in determining the strength of halogen bonding and, in some cases, therefore the geometry adopted by the molecular system, is the identity of the halogen atom. Under some conditions astatinated systems, and in relatively rare cases iodinated systems, can form stronger attractive interactions between the monomers than is formed in non-halogenated DNA base pairs. Other factors that have been found to play significant roles relate to the broader chemical environment wherein the halogen bonding occurs. The electron withdrawing or donating effects of other substituents appear to be important in determining which halogen bond, if any, dominates the system. Steric effects have also occasionally been found to be important, although perhaps not to the extent found by Parker *et al.* [35]. Although all interactions were found to be predominantly non-covalent, there appears to be a general correlation between stronger halogen bonding and higher degrees of covalency.

7 Conclusions

This study comprises four projects: microsolvated halobenzene, microsolvated 1-methyl-5-halouracil, the thyroid system, and halogenated DNA base pairs. The principal conclusions for each of these projects are summarised below. The same terminology, including abbreviations, as have been used in the preceding chapters is employed herein.

7.1 Microsolvated halobenzene

Microsolvated halobenzene is a simple molecular dimer that shares some of the characteristics of the more complex microsolvated 1-methyl-5-halouracil system. XPh-w forms X•••H hydrogen bonds between the halogen atom on halobenzene and a hydrogen atom on the water molecule for all halogen elements, with fluorobenzene exhibiting the strongest interaction. The halogen atom forms a halogen bond with the water oxygen atom only in cases where the halogen atom is bromine, iodine or astatine. Halogen bond strength was found to increase as the halogen group is descended, and there is evidence of increased donation into the σ* antibonding orbital, with the C-X covalent bond lengthening going down the group, although this observation can be partially explained by the increasing VdW radius of the halogen atom. The At•••Ow halogen bond exhibits an interaction energy that is more strongly negative than for any of the hydrogen bonds in any of the investigated systems. In this simple system, the halogen bond angle, for each halogen element, was calculated to be 179^{o}, very close to the ideal linear angle for halogen bonds.

7.2 Microsolvated 1-methyl-5-halouracil

Unlike the halobenzene system, XU-w can form a hydrogen bond between an oxygen atom (O4) on the XU molecule and a hydrogen atom on the water molecule. This hydrogen bond can compete with the X5•••Ow halogen bond. As in the case of halobenzene, the X•••Hw interaction is formed for all of the investigated halogen elements, while the X•••Ow interaction arises only in cases where the halogen atom is heavier than chlorine. The same trend in halogen bond strength as the halogen group is descended as was found in the case of microsolvated halobenzene holds true for XU-w. WBI values also indicate an increasing degree of covalency as the halogen group is descended,implying increased donation into the σ* orbital (which corresponds to the σ-hole). Due to the more complex chemical environment, wherein secondary interactions are present, the halogen bond angle departs from being linear. Three hydrogen bonds were investigated in addition to the X•••Ow halogen bond. Hbond1 is the only hydrogen bond minimum to show a trend going down the halogen group; the interaction becomes more attractive as the group is descended, probably due to a combination of reduced electron withdrawing effect the halogen atom leading to greater electron density on the O4 atom, which forms the hydrogen bond with the Hw1 atom, and an increase in the concentration of electron density in the negatively charged ring around the halogen atom. The other hydrogen bonds do

not exhibit a clear correlation between interaction energies and atomic number of the halogen element. Hbond 1 (O4•••Hw1) is connected to Xbond by a transition state. Calculation of transition barrier heights revealed that there is a clear trend towards greater barrier heights as the halogen group is descended, with barrier heights of 0.5 kJ mol^{-1}, 4.9 kJ mol^{-1} and 9.8 kJ mol^{-1} for BrU-w, IU-w and AtU-w respectively. The extremely low barrier height in the brominated case indicates that this halogen bond is probably only metastable. Addition of a second water molecule has the effect of facilitating the formation of the X5•••Ow halogen bond where the halogen is chlorine. This result can be explained as follows. The additional substituents on the benzene ring that distinguish XU from XPh are electron withdrawing. Therefore, the halogen atom is more electron deficient and hence can form stronger halogen bonds. This enhanced halogen bond forming efficacy explains why ClU can form halogen bonds whereas ClPh cannot form that interaction with a water molecule. However, in the XU-w system, this halogen bond is in competition with Hbond 1, and chlorine is not sufficiently efficacious at forming halogen bonds to overcome this competition. By contrast, in the case of XU-2w, one of the water molecules participates in hydrogen bonding with the O4 atom, thereby blocking the other water molecule from forming that interaction. That other water molecule (not engaged in hydrogen bonding to O4) is then available to form a halogen bond with X5, without competition from O4. With the competing hydrogen bond eliminated, the greater electron deficiency of the chlorine atom (in comparison with ClPh) enables the formation of the Cl•••Ow halogen bond.

7.3 The thyroid system

Halogen bonding has been identified as occurring between a fragment of thyroxine (containing iodine atoms) and oxygen atoms contained in the protein backbone. Furthermore, water oxygen atoms within the thyroxine-containing crystals also form halogen bonds with an iodine atom on the thyroxine fragment. Although quantitative results are reported herein, the conclusions that can be drawn from them are largely qualitative, due limitations in the nature and scope of the investigation, required in the interest of computational affordability. The presence of halogen bonds involving thyroxine in the thyroid system is consistent with the prior findings of Auffinger *et al.* [15]. Furthermore, Bayse and Rafferty [151] have identified the role of halogen bonding's donation of electron density into the C-I bond's corresponding σ* antibonding orbital in the cleavage of the C-I covalent bond, forming biologically active T3 from T4. The findings in the present study are consistent with those of Bayse and Rafferty. In the present study the effect of substituting the iodine atom that participates in halogen bonding with an astatine atom was investigated. Substitution by astatine results in stronger halogen bonds.

7.4 Halogenated DNA base pairs

This study entailed halogenation of each of the four DNA bases. Halogenated variants of the adenine – thymine and guanine – cytosine base pairs were compared with their canonical analogues. Although detailed conclusions are not amenable to brief summary due to the need to provide conclusions for each of the many systems investigated, the following observations

can be made. The formation of halogen bonds and the extent of their role in determining and stabilising the geometry of a dimer system depends upon the identity of the halogen atom (the heavier the halogen element the greater the halogen bonding efficacy), the electron withdrawing or donating properties of substituents, steric effects, and the strength of competing interactions. In this highly complex environment, halogen bonds that substantially depart from a linear geometry can form. In addition to forming stronger halogen bonds, the heavier halogen elements also demonstrated a greater ability to tolerate sub-optimal halogen bond angles, probably due to the greater geometric size of their σ-holes resulting in the electron deficient region encompassing a greater range of angles (although still centred on 180°). There is a wide range of interaction energies, halogen bond angles and degrees of covalency. Interaction energies were calculated for the dimer complex as a whole and not attributed to individual interatomic interactions, and there are typically multiple interactions present. This imposes some limitation on drawing conclusions about halogen bond strength from the interaction energies. However, when combined with geometric data and the WBI values, reasonable inferences can be drawn, at least qualitatively as to whether or not the halogen bond substantially contributes to the interaction energy and/or is responsible (wholly or partially) for observed trends as the halogen group is descended. In some cases it was found to be clear that the identity of the halogen atom determined the overall geometry of the system. Astatination and (very rarely) iodination results in dimer complexes that have a stronger interaction between the molecules than in the case of the corresponding canonical base pair. By contrast fluorine was not found to form halogen bonds in any of the investigated systems. This study was inspired by the work of Parker *et al*. [35]. The results from the present study are to a large extent in agreement with the conclusions drawn by Parker *et al.*; however, the degree to which steric effects were found to influence dimers' stability differ between the two studies. In particular the present study found that iodinated molecules form more stable complexes than those formed by brominated molecules, whereas Parker *et al.* had come to the contrary conclusion based upon their research.

7.5 General conclusions

From the hereinbefore described studies the following general observations can be made: Halogen bond strength increases going down the halogen group. Halogen bonds have a preference for linear geometry, however in complex chemical environments substantial departure from linearity can be tolerated, especially by the heavier halogen elements. Iodine, and especially astatine, can form halogen bonds that are comparable to, or in some cases greater in strength, than hydrogen bonds. The degree of covalency in halogen bonds can vary significantly between different molecular systems.

8 References

1. Colin, J.-J., Note sur quelques combinaisons de l'iode. *Ann. Chem.* **1814,** *91*, 252-272.
2. Guthrie, F., On the iodide of iodammonium. *J. Chem. Soc.* **1863,** *16*, 239-244.
3. Cavallo, G.; Metrangolo, P.; Milani, R.; Pilati, T.; Priimagi, A.; Resnati, G.; Terraneo, G., The halogen bond. *Chem. Rev.* **2016,** *116* (4), 2478-2601.
4. Mulliken, R. S., Structures of complexes formed by halogen molecules with aromatic and with oxygenated solvents. *J. Am. Chem. Soc.* **1950,** *72* (1), 600-608.
5. Hassel, O., Structural aspects of interatomic charge-transfer bonding. *Science* **1970,** *170* (3957), 497-502.
6. Dumas, J.-M.; Peurichard, H.; Gomel, M., CX4...base interactions as models of weak charge-transfer interactions - comparison with strong charge-transfer and hydrogen-bond interactions. *J. Chem. Res. (S)* **1978,** *2*, 54-57.
7. Desiraju, G. R.; Ho, P. S.; Kloo, L.; Legon, A. C.; Marquardt, R.; Metrangolo, P.; Politzer, P.; Resnati, G.; Rissanen, K., Definition of the halogen bond (IUPAC Recommendations 2013). *Pure Appl. Chem.* **2013,** *85* (8), 1711-1713.
8. Hill, J. G.; Hu, X., Theoretical insights into the nature of halogen bonding in prereactive complexes. *Chem.-Eur. J.* **2013,** *19* (11), 3620-3628.
9. Liu, F.; Du, L. K.; Zhang, D. J.; Gao, J., Performance of density functional theory on the anisotropic halogen center dot center dot center dot halogen interactions and potential energy surface: Problems and possible solutions. *Int. J. Quantum Chem.* **2016,** *116* (9), 710-717.
10. Kolar, M.; Hobza, P.; Bronowska, A. K., Plugging the explicit sigma-holes in molecular docking. *Chem Commun (Camb)* **2013,** *49* (10), 981-3.
11. Wang, C.; Guan, L.; Danovich, D.; Shaik, S.; Mo, Y., The origins of the directionality of noncovalent intermolecular interactions. *J. Comput. Chem.* **2016,** *37* (1), 34-45.
12. Kolar, M. H.; Hobza, P., Computer modeling of halogen bonds and other sigma-Hole interactions. *Chem. Rev.* **2016,** *116* (9), 5155-5187.
13. Kolar, M.; Hostas, J.; Hobza, P., The strength and directionality of a halogen bond are co-determined by the magnitude and size of the sigma-hole. *Phys. Chem. Chem. Phys.* **2014,** *16* (21), 9987-9996.
14. Clark, T.; Hennemann, M.; Murray, J. S.; Politzer, P., Halogen bonding: the sigma-hole. *J. Mol. Model.* **2007,** *13* (2), 291-296.
15. Auffinger, P.; Hays, F. A.; Westhof, E.; Ho, P. S., Halogen bonds in biological molecules. *Proc Natl Acad Sci U S A* **2004,** *101* (48), 16789-94.
16. Hogan, S. W. L.; van Mourik, T., Competition between hydrogen and halogen bonding in halogenated 1-methyluracil: water systems. *J. Comput. Chem.* **2016,** *37* (8), 763-770.
17. Bauza, A.; Mooibroek, T. J.; Frontera, A., Tetrel-bonding interaction: rediscovered supramolecular force? *Angew. Chem.-Int. Edit.* **2013,** *52* (47), 12317-12321.
18. Politzer, P.; Murray, J. S.; Clark, T., Halogen bonding and other sigma-hole interactions: a perspective. *Phys. Chem. Chem. Phys.* **2013,** *15* (27), 11178-11189.
19. Metrangolo, P.; Resnati, G., Halogen bonding: where we are and where we are going. *Cryst. Growth Des.* **2012,** *12* (12), 5835-5838.
20. www.webofknowledge.com (accessed 7th February 2018).
21. Stenlid, J. H.; Brinck, T., Extending the sigma-hole concept to metals: an electrostatic interpretation of the effects of nanostructure in gold and platinum catalysis. *J. Am. Chem. Soc.* **2017,** *139* (32), 11012-11015.

22. Kozuch, S.; Martin, J. M. L., Halogen bonds: benchmarks and theoretical analysis. *J. Chem. Theory Comput.* **2013,** *9* (4), 1918-1931.
23. Wiberg, K. B., Application of Pople-Santry-Segal CNDO method to cyclopropylcarbinyl and cyclobutyl cation and bicyclobutane. *Tetrahedron* **1968,** *24* (3), 1083-1096.
24. Reed, A. E.; Curtiss, L. A.; Weinhold, F., Intermolecular interactions from a natural bond orbital donor-acceptor viewpoint. *Chem. Rev.* **1988,** *88* (6), 899-926.
25. Gilday, L. C.; Robinson, S. W.; Barendt, T. A.; Langton, M. J.; Mullaney, B. R.; Beer, P. D., Halogen bonding in supramolecular chemistry. *Chem. Rev.* **2015,** *115* (15), 7118-7195.
26. Eisenschitz, R.; London, F., About the relationship of the van der Waals forces to the covalent bonding forces. *Z. Phys.* **1930,** *60* (7-8), 491-527.
27. Szalewicz, K.; Jeziorski, B., Symmetry -adapted double-perturbation anaylsis of intra-molecular correlation effects in weak inter-molecular interactions - He - He interaction. *Molecular Physics* **1979,** *38* (1), 191-208.
28. Rybak, S.; Jeziorski, B.; Szalewicz, K., Many-body symmetry-adapted perturbation-theory of intermolular interactions - H2O and HF dimers. *J. Chem. Phys.* **1991,** *95* (9), 6576-6601.
29. Szalewicz, K., Symmetry-adapted perturbation theory of intermolecular forces. *Wiley Interdisciplinary Reviews-Computational Molecular Science* **2012,** *2* (2), 254-272.
30. Hill, G. J.; Legon, A. C., On the directionality and non-linearity of halogen and hydrogen bonds. *Phys. Chem. Chem. Phys.* **2015,** *17* (2), 858-867.
31. Stone, A. J., Are halogen bonded structures electrostatically driven? *J. Am. Chem. Soc.* **2013,** *135* (18), 7005-7009.
32. Schneider, F. S. S.; Caramori, G. F.; Parreira, R. L. T.; Lippolis, V.; Arca, M.; Ciancaleoni, G., Bond analysis in dihalogen-halide and dihalogen-dimethylchalcogenide systems. *Eur. J. Inorg. Chem.* **2018,** (8), 1007-1015.
33. Hanus, M.; Kabelac, M.; Nachtigallova, D.; Hobza, P., Mutagenic properties of 5-halogenuracils: Correlated quantum chemical ab initio study. *Biochemistry* **2005,** *44* (5), 1701-1707.
34. Schaub, T. A.; Sure, R.; Hampel, F.; Grimme, S.; Kivala, M., Quantum chemical dissection of the shortest P=OI halogen bond: the decisive role of crystal packing effects. *Chem.-Eur. J.* **2017,** *23* (24), 5687-5691.
35. Parker, A. J.; Stewart, J.; Donald, K. J.; Parish, C. A., Halogen bonding in DNA base pairs. *J. Am. Chem. Soc.* **2012,** *134* (11), 5165-72.
36. El-Sheshtawy, H. S.; El-Mehasseb, I., Role of halogen and hydrogen bonds for stabilization of antithyroid drugs with hypohalous acids (HOX, X = I, Br, and Cl) adducts. *J. Mol. Struct.* **2017,** *1147*, 643-650.
37. Marsan, E. S.; Bayse, C. A., Halogen-bonding interactions of polybrominated diphenyl ethers and thyroid hormone derivatives: a potential mechanism for the Inhibition of iodothyronine deiodinase. *Chem.-Eur. J.* **2017,** *23* (27), 6625-6633.
38. Adasme-Carreno, F.; Munoz-Gutierrez, C.; Alzate-Morales, J. H., Halogen bonding in drug-like molecules: a computational and systematic study of the substituent effect. *RSC Adv.* **2016,** *6* (66), 61837-61847.
39. Mondal, S.; Mugesh, G., Regioselective deiodination of iodothyronamines, endogenous thyroid hormone derivatives, by deiodinase mimics. *Chem.-Eur. J.* **2014,** *20* (35), 11120-11128.
40. de Freitas, R. F.; Schapira, M., A systematic analysis of atomic protein-ligand interactions in the PDB. *MedChemComm* **2017,** *8* (10), 1970-1981.
41. Kolar, M. H.; Tabarrini, O., Halogen bonding in nucleic acid complexes. *J. Med. Chem.* **2017,** *60* (21), 8681-8690.
42. Xu, Z. J.; Yang, Z.; Liu, Y. T.; Lu, Y. X.; Chen, K. X.; Zhu, W. L., Halogen bond: its role beyond drug-target binding affinity for drug discovery and development. *J. Chem Inf. Model.* **2014,** *54* (1), 69-78.

43. Voth, A. R.; Hays, F. A.; Ho, P. S., Directing macromolecular conformation through halogen bonds. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104* (15), 6188-6193.
44. Scholfield, M. R.; Vander Zanden, C. M.; Carter, M.; Ho, P. S., Halogen bonding (X-bonding): A biological perspective. *Protein Sci.* **2013,** *22* (2), 139-152.
45. Lisac, K.; Cincic, D., The influence of liquid on the outcome of halogen-bonded metal-organic materials synthesis by liquid assisted grinding. *Crystals* **2017,** *7* (12), 11.
46. Mittapalli, S.; Perumalla, D. S.; Nanubolu, J. B.; Nangia, A., Thermomechanical effect in molecular crystals: the role of halogen-bonding interactions. *IUCrJ* **2017,** *4*, 812-823.
47. Stumpel, J. E.; Saccone, M.; Dichiarante, V.; Lehtonen, O.; Virkki, M.; Metrangolo, P.; Priimagi, A., Surface-relief gratings in halogen-bonded polymer-azobenzene complexes: a concentration-dependence study. *Molecules* **2017,** *22* (11), 11.
48. Berman, H. M.; Battistuz, T.; Bhat, T. N.; Bluhm, W. F.; Bourne, P. E.; Burkhardt, K.; Iype, L.; Jain, S.; Fagan, P.; Marvin, J.; Padilla, D.; Ravichandran, V.; Schneider, B.; Thanki, N.; Weissig, H.; Westbrook, J. D.; Zardecki, C., The Protein Data Bank. *Acta Crystallogr. D* **2002,** *58*, 899-907.
49. Fortino, M.; Marino, T.; Russo, N.; Sicilia, E., Mechanism of thyroxine deiodination by naphthyl-based iodothyronine deiodinase mimics and the halogen bonding role: a DFT investigation. *Chem.-Eur. J.* **2015,** *21* (23), 8554-8560.
50. Holroyd, L. F.; van Mourik, T., Stacking of the mutagenic base analogue 5-bromouracil: energy landscapes of pyrimidine dimers in gas phase and water. *Phys. Chem. Chem. Phys.* **2015,** *17* (45), 30364-30370.
51. Domagala, M.; Lutynska, A.; Palusiak, M., Halogen bond versus hydrogen bond: The many-body interactions approach. *Int. J. Quantum Chem.* **2017,** *117* (7), 9.
52. Geboes, Y.; De Proft, F.; Herrebout, W. A., Effect of fluorination on the competition of halogen bonding and hydrogen bonding: complexes of fluoroiodomethane with dimethyl ether and trimethylamine. *J. Phys. Chem. A* **2017,** *121* (21), 4180-4188.
53. Bulfield, D.; Huber, S. M., Halogen bonding in organic synthesis and organocatalysis. *Chem.-Eur. J.* **2016,** *22* (41), 14434-14450.
54. Alkorta, I.; Elguero, J.; Del Bene, J. E., Boron as an electron-pair donor for B center dot center dot center dot Cl halogen bonds. *ChemPhysChem* **2016,** *17* (19), 3112-3119.
55. Alkorta, I.; Soteras, I.; Elguero, J.; Del Bene, J. E., The boron-boron single bond in diborane(4) as a non-classical electron donor for hydrogen bonding. *Phys. Chem. Chem. Phys.* **2011,** *13* (31), 14026-14032.
56. Chou, S. L.; Lo, J. I.; Peng, Y. C.; Lin, M. Y.; Lu, H. C.; Cheng, B. M.; Ogilvie, J. F., Identification of diborane(4) with bridging B-H-B bonds. *Chem. Sci.* **2015,** *6* (12), 6872-6877.
57. Radius, U.; Bickelhaupt, F. M.; Ehlers, A. W.; Goldberg, N.; Hoffmann, R., Is CO a special ligand in organometallic chemistry? Theoretical investigation of AB, Fe(CO)(4)AB, and Fe(AB)(5) (AB = N-2, CO, BF, SiO). *Inorg. Chem.* **1998,** *37* (5), 1080-1090.
58. Rozas, I.; Alkorta, I.; Elguero, J., Monohydride and monofluoride derivatives of B, Al, N and P. Theoretical study of their ability as hydrogen bond accepters. *J. Phys. Chem. A* **1999,** *103* (44), 8861-8869.
59. Mukherjee, A.; Tothadi, S.; Desiraju, G. R., Halogen bonds in crystal engineering: like hydrogen bonds yet different. *Accounts Chem. Res.* **2014,** *47* (8), 2514-2524.
60. Rezac, J.; Riley, K. E.; Hobza, P., Benchmark calculations of noncovalent interactions of halogenated molecules. *J. Chem. Theory Comput.* **2012,** *8* (11), 4285-4292.
61. Forni, A.; Pieraccini, S.; Rendine, S.; Sironi, M., Halogen bonds with benzene: an assessment of DFT functionals. *J. Comput. Chem.* **2014,** *35* (5), 386-94.
62. Dunning, T. H., Gaussian-basis sets for use in correlated molecular calcualations .1. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989,** *90* (2), 1007-1023.

63. Kendall, R. A.; Dunning, T. H.; Harrison, R. J., Electron-affinities of the 1st-row atoms revisited - systematic basis-sets and wave-functions. *J. Chem. Phys.* **1992,** *96* (9), 6796-6806.
64. Peterson, K. A.; Figgen, D.; Goll, E.; Stoll, H.; Dolg, M., Systematically convergent basis sets with relativistic pseudopotentials. II. Small-core pseudopotentials and correlation consistent basis sets for the post-d group 16-18 elements. *J. Chem. Phys.* **2003,** *119* (21), 11113-11123.
65. Chai, J.-D.; Head-Gordon, M., Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008,** *128* (8).
66. Grimme, S., Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel-and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003,** *118*, 9095.
67. Distasio, R. A., Jr.; Head-Gordon, M., Optimized spin-component scaled second-order Moller-Plesset perturbation theory for intermolecular interaction energies. *Mol. Phys.* **2007,** *105* (8), 1073-1083.
68. Zhao, Y.; Truhlar, D. G., A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.* **2006,** *125* (19).
69. Boese, A. D.; Martin, J. M. L., Development of density functionals for thermochemical kinetics. *J. Chem. Phys.* **2004,** *121* (8), 3405-3416.
70. Zhao, Y.; Truhlar, D. G., The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008,** *120* (1-3), 215-241.
71. Yanai, T.; Tew, D. P.; Handy, N. C., A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004,** *393* (1-3), 51-57.
72. Grimme, S., Improved third-order Moller-Plesset perturbation theory. *J. Comput. Chem.* **2003,** *24* (13), 1529-1537.
73. Pitonak, M.; Neogrady, P.; Cerny, J.; Grimme, S.; Hobza, P., Scaled MP3 non-covalent interaction energies agree closely with accurate CCSD(T) benchmark data. *ChemPhysChem* **2009,** *10* (1), 282-289.
74. Boys, S. F.; Bernardi, F., Calculation of small molecular interactions by differences of separate total energies - some procedures with reduced errors. *Mol. Phys.* **1970,** *19* (4), 553-556.
75. Kozuch, S.; Martin, J. M. L., Spin-component-scaled double hybrids: an extensive search for the best fifth-rung functionals blending DFT and perturbation theory. *J. Chem. Theory Comput.* **2013,** *34* (27), 2327-2344.
76. Kozuch, S.; Martin, J. M. L., DSD-PBEP86: in search of the best double-hybrid DFT with spin-component scaled MP2 and dispersion corrections. *Phys. Chem. Chem. Phys.* **2011,** *13* (45), 20104-20107.
77. Perdew, J. P.; Wang, Y., Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B* **1992,** *45* (23), 13244-13249.
78. Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996,** *77* (18), 3865-3868.
79. Scholfield, M. R.; Ford, M. C.; Zanden, C. M. V.; Billman, M. M.; Ho, P. S.; Rappe, A. K., Force field model of periodic trends in biomolecular halogen bonds. *J. Phys. Chem. B* **2015,** *119* (29), 9140-9149.
80. Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C., The Cambridge Structural Database. *Acta Crystallogr. Sect. B-Struct. Sci.Cryst. Eng. Mat.* **2016,** *72*, 171-179.
81. Nemec, V.; Lisac, K.; Stilinovic, V.; Cincic, D., Inorganic bromine in organic molecular crystals: Database survey and four case studies. *J. Mol. Struct.* **2017,** *1128*, 400-409.

82. Cramer, C. J., *Essentials of computational chemistry: theories and models*. 2nd ed.; John Wiley & Sons, Ltd: 2004 (reprinted with corrections 2006).
83. Jensen, F., *Introduction to computational chemistry*. 2nd ed.; Wiley: 2007.
84. Pauli, W., On the connection of the arrangement of electron groups in atoms with the complex structure of spectra. *Z. Phys.* **1925,** *31*, 765-783.
85. Moller, C.; Plesset, M. S., Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934,** *46* (7), 0618-0622.
86. Řezáč, J.; Riley, K. E.; Hobza, P., Benchmark calculations of noncovalent interactions of halogenated molecules. *J. Chem. Theory Comput.* **2012,** *8* (11), 4285-4292.
87. Bremond, E.; Savarese, M.; Su, N. Q.; Jose Perez-Jimenez, A.; Xu, X.; Carlos Sancho-Garcia, J.; Adamo, C., Benchmarking density functionals on structural parameters of small-/medium-sized organic molecules. *J. Chem. Theory Comput.* **2016,** *12* (2), 459-465.
88. Brauer, B.; Kesharwani, M. K.; Kozuch, S.; Martin, J. M. L., The S66x8 benchmark for noncovalent interactions revisited: explicitly correlated ab initio methods and density functional theory. *Phys. Chem. Chem. Phys.* **2016**.
89. Ditchfield, R.; Hehre, W. J.; Pople, J. A., Self-consistent molecular-orbital methods .9. Extended Gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys.* **1971,** *54* (2), 724-728.
90. Hehre, W. J.; Ditchfie.R; Pople, J. A., Self-consistent molecular-orbital methods .12. Further extensions of Gaussian-type basis sets for use in molecular-orbital studies of organic-molecules. *J. Chem. Phys.* **1972,** *56* (5), 2257-2261.
91. van Mourik, T., *University of St Andrews module CH3721 lecture notes***2010**.

92. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; G. E. Scuseria; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; B. Mennucci; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; H. P. Hratchian; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; M. Hada; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; T. Nakajima; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; J. A. Montgomery, J.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; E. Brothers; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; J. Normand; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; J. B. Cross; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; R. E. Stratmann; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; J. W. Ochterski; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; G. A. Voth; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; A. D. Daniels; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; J. Cioslowski; Fox, D. J. *Gaussian 09, Revision A.02,*, Gaussian, Inc.: Wallingford CT, 2009.
93. Petersson, G. A.; Allaham, M. A., A complete basis set model chemistry .2. Open-shell systems and the total energies of the 1st-row atoms. *J. Chem. Phys.* **1991,** *94* (9), 6081-6090.
94. Hohenberg, P.; Kohn, W., Inhomogeneous electron gas. *Phys. Rev. B* **1964,** *136*, 864-871.
95. Kohn, W.; Sham, L. J., Self-consistent equations including exchange and correlation effects. *Phys. Rev. A* **1965,** *140*, 1133-1138.
96. Sousa, S. F.; Fernandes, P. A.; Ramos, M. J., General performance of density functionals. *J. Phys. Chem. A* **2007,** *111* (42), 10439-10452.
97. Sharkas, K.; Toulouse, J.; Savin, A., Double-hybrid density-functional theory made rigorous. *J. Chem. Phys.* **2011,** *134* (6).
98. Perdew, J. P.; Schmidt, K., Jacob's ladder of density functional approximations for the exchange-correlation energy. In *Density Functional Theory and Its Application to Materials*, VanDoren, V.; VanAlsenoy, C.; Geerlings, P., Eds. American Institute Physics: Melville, 2001; Vol. 577, pp 1-20.

99. Zhao, Y.; Truhlar, D. G., Density functional for spectroscopy: No long-range self-interaction error, good performance for Rydberg and charge-transfer states, and better performance on average than B3LYP for ground states. *J. Phys. Chem. A* **2006,** *110* (49), 13126-13130.
100. Grimme, S., Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006,** *27* (15), 1787-1799.
101. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H., A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010,** *132* (15).
102. Grimme, S.; Ehrlich, S.; Goerigk, L., Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011,** *32* (7), 1456-1465.
103. http://wild.life.nctu.edu.tw/~jsyu/compchem/g09/g09ur/k_integral.htm (accessed 28th February 2018).
104. Wu, T. Z.; Kalugina, Y. N.; Thakkar, A. J., Choosing a density functional for static molecular polarizabilities. *Chem. Phys. Lett.* **2015,** *635*, 257-261.
105. Wheeler, S. E.; Houk, K. N., Integration grid errors for meta-GGA-predicted reaction energies: origin of grid errors for the M06 suite of functionals. *J. Chem. Theory Comput.* **2010,** *6* (2), 395-404.
106. Peng, C.; Bernhard Schlegel, H., Combining synchronous transit and quasi-Newton methods to find transition states. *Isr. J. Chem.* **1993,** *33* (4), 449-454.
107. Peng, C. Y.; Ayala, P. Y.; Schlegel, H. B.; Frisch, M. J., Using redundant internal coordinates to optimize equilibrium geometries and transition states. *J. Comput. Chem.* **1996,** *17* (1), 49-56.
108. Hoffmann, R., An extended Huckel theory .1. Hydrocarbons. *J. Chem. Phys.* **1963,** *39* (6), 1397-1412.
109. Huckel, E., Quantum contributions to the benzene problem. *Z. Phys.* **1931,** *70* (3-4), 204-286.
110. Huckel, E., Quantum contributions to the problem of aromatic and unsaturated compounds. 3. *Z. Phys.* **1932,** *76* (9-10), 628-648.
111. Huckel, E., The free radicals of organic chemistry. Quantum contributions to the problem of aromatic and unsaturated compounds. *Z. Phys.* **1933,** *83* (9-10), 632-668.
112. Pople, J. A.; Segal, G. A., Approximate self-consistent molecular orbital theory .2. Calculations with complete neglect of differential overlap. *J. Chem. Phys.* **1965,** *43* (10), S136-5149.
113. Pople, J. A.; Santry, D. P.; Segal, G. A., Approximate self-consistent molecular orbital theory .1. Invariant procedures. *J. Chem. Phys.* **1965,** *43* (10), S129-5135.
114. Pople, J. A.; Segal, G. A., Approximate self-consistent molecular orbital theory .3. CNDO results for AB2 and AB3 systems. *J. Chem. Phys.* **1966,** *44* (9), 3289-3296.
115. Mulliken, R. S., Electronic population anaylsis on LCAO-MO molecular wave functions .1. *J. Chem. Phys.* **1955,** *23* (10), 1833-1840.
116. Mulliken, R. S., Electronic population anaylsis on LCAO-MO molecular wave functions .2. Overlap populationsl bond orders, and covalent bond energies. *J. Chem. Phys.* **1955,** *23* (10), 1841-1846.
117. Mulliken, R. S., Electronic population anaylsis on LCAO-MO molecular wave functions .3. Effects of hybridization on overlaps and gross AO populations. *J. Chem. Phys.* **1955,** *23* (12), 2338-2342.
118. Mulliken, R. S., Electronic population anaylsis on LCAO-MO molecular wave functions .4. Bonding and antibonding in LCAO and valence-bond theories. *J. Chem. Phys.* **1955,** *23* (12), 2343-2346.

119. Szczepanik, D.; Mrozek, J., On quadratic bond-order decomposition within molecular orbital space. *J. Math. Chem.* **2013,** *51* (6), 1619-1633.
120. Bridgeman, A. J.; Cavigliasso, G.; Ireland, L. R.; Rothery, J., The Mayer bond order as a tool in inorganic chemistry. *J. Chem. Soc.-Dalton Trans.* **2001,** (14), 2095-2108.
121. Mayer, I., On bond orders and valences in the ab initio quantum chemical theory. *Int. J. Quantum Chem.* **1986,** *29* (1), 73-84.
122. Frisch, A.; Nielson, A. B.; Holder, A. J., Gaussview user manual. *Gaussian Inc., Pittsburgh, PA 2000*.
123. Skyner, R., Private communication at University of St Andrews on 23rd March 2017.
124. Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R., New software for searching the Cambridge Structural Database and visualizing crystal structures. *Acta Crystallogr. Sect. B-Struct. Sci.Cryst. Eng. Mat.* **2002,** *58*, 389-397.
125. Taylor, R.; Macrae, C. F., Rules governing the crystal packing of mono- and di-alcohols. *Acta Crystallogr.* **2001,** *B57*, 815-827.
126. Macrae, C. F.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Shields, G. P.; Taylor, R.; Towler, M.; van De Streek, J., Mercury: visualization and analysis of crystal structures. *J. Appl. Crystallogr.* **2006,** *39*, 453-457.
127. Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P. A., Mercury CSD 2.0 - new features for the visualization and investigation of crystal structures. *J. Appl. Crystallogr.* **2008,** *41*, 466-470.
128. *MATLAB and statistics toolbox release 2016b, The MathWorks, Inc., Natick, Massachusetts, United States.*
129. Eilers, P. H. C.; Goeman, J. J., Enhancing scatterplots with smoothed densities. *Bioinformatics* **2004,** *20* (5), 623-628. The version of dscatter employed had been modified by R. E. Skyner, St Andrews, **2016**.
130. Skyner, R. E., Edit_script.m. **2016**.
131. Muzangwa, L.; Nyambo, S.; Uhler, B.; Reid, S. A., On pi-stacking, C-H/pi, and halogen bonding interactions in halobenzene clusters: Resonant two-photon ionization studies of chlorobenzene. *J. Chem. Phys.* **2012,** *137* (18).
132. Torii, H., Properties of halogen atoms related to the electrostatic origin of halogen bonding: basic aspects and some applications. In *International conference of computational methods in sciences and engineering 2009*, Simos, T. E.; Maroulis, G., Eds. Amer Inst Physics: Melville, 2012; Vol. 1504, pp 228-239.
133. Riley, K. E.; Merz, K. M., Insights into the strength and origin of halogen bonding: The halobenzene-formaldehyde dimer. *J. Phys. Chem. A* **2007,** *111* (9), 1688-1694.
134. Clark, T.; Hennemann, M.; Murray, J. S.; Politzer, P., Halogen bonding: the sigma-hole. Proceedings of "Modeling interactions in biomolecules II", Prague, September 5th-9th, 2005. *J. Mol. Model.* **2007,** *13* (2), 291-6.
135. Lide, D. R., *Handbook of chemistry and physics*. 90 ed.; CRC: 2009.
136. Royal Society of Chemistry. Available at: http://Www.Rsc.Org/Periodic-Table/Element/85/Astatine. Accessed 11 March 2015.

137. Zamenhof, S.; Degiovanni, R.; Greer, S., Induced gene unstabilization. *Nature* **1958,** *181* (4612), 827-829.
138. Sugiyama, H.; Tsutsumi, Y.; Saito, I., Highly sequence selective photoreaction of 5-bromouracil-containing deoxyhexanucleotides. *J. Am. Chem. Soc.* **1990,** *112* (18), 6720-6721.
139. Abdoul-Carime, H.; Huels, M. A.; Illenberger, E.; Sanche, L., Sensitizing DNA to secondary electron damage: Resonant formation of oxidative radicals from 5-halouracils. *J. Am. Chem. Soc.* **2001,** *123* (22), 5354-5355.

140. Danilov, V. I.; van Mourik, T.; Kurita, N.; Wakabayashi, H.; Tsukamoto, T.; Hovorun, D. M., On the mechanism of the mutagenic action of 5-bromouracil: a DFT study of uracil and 5-bromouracil in a water cluster. *J. Phys. Chem. A* **2009,** *113* (11), 2233-2235.
141. van Mourik, T.; Danilov, V. I.; Dailidonis, V. V.; Kurita, N.; Wakabayashi, H.; Tsukamoto, T., A DFT study of uracil and 5-bromouracil in nanodroplets. *Theor. Chem. Acc.* **2010,** *125* (3-6), 233-244.
142. Holroyd, L. F.; Van Mourik, T., Stacking of the mutagenic DNA base analog 5-bromouracil. *Theor. Chem. Acc.* **2014,** *133*, 1431-1443.
143. Meyer, G. J.; Rössler, K.; Stöcklin, G., Preparation and high-pressure liquid-chromatography of 5-astatouracil. *J. Label. Compd. Radiopharm.* **1976,** *12*, 449-458.
144. Schaftenaar, G.; Noordik, J. H., Molden: a pre- and post-processing program for molecular and electronic structures. *J. Comput.-Aided Mol. Des.* **2000,** *14* (2), 123-134.
145. Wang, W.; Hobza, P., Origin of the X-Hal (Hal = Cl, Br) bond-length change in the halogen-bonded complexes. *J. Phys. Chem. A* **2008,** *112* (17), 4114-4119.
146. van Mourik, T.; Price, S. L.; Clary, D. C., Ab initio calculations on uracil-water. *J. Phys. Chem. A* **1999,** *103* (11), 1611-1618.
147. van Mourik, T., A theoretical study of uracil-(H2O)n, n=2 to 4. *Phys. Chem. Chem. Phys.* **2001,** *3* (14), 2886-2892.
148. Shields, Z. P.; Murray, J. S.; Politzer, P., Directional tendencies of halogen and hydrogen bonds. *Int. J. Quantum Chem.* **2010,** *110* (15), 2823-2832.
149. Mondal, S.; Mugesh, G., Structure elucidation and characterization of different thyroxine polymorphs. *Angew. Chem.-Int. Edit.* **2015,** *54* (37), 10833-10837.
150. Garber, J. R.; Cobin, R. H.; Gharib, H.; Hennessey, J. V.; Klein, I.; Mechanick, J. I.; Pessah-Pollack, R.; Singer, P. A.; Woeber, K. A.; Amer Assoc Clinical, E.; Amer Thyroid Assoc Taskforce, H., Clinical practice guidelines for hypothyroidism in adults: cosponsored by the American Association of Clinical Endocrinologists and the American Thyroid Association. *Thyroid* **2012,** *22* (12), 1200-1235.
151. Bayse, C. A.; Rafferty, E. R., Is halogen bonding the basis for iodothyronine deiodinase activity? *Inorg. Chem.* **2010,** *49* (12), 5365-5367.
152. Hamilton, J. A.; Steinrauf, L. K.; Braden, B. C.; Liepnieks, J.; Benson, M. D.; Holmgren, G.; Sandgren, O.; Steen, L., The x-ray crystal-structure refinements of normal human transthyretin and the amyloidogenic val-30- met variant to 1.7-a resolution. *J. Biol. Chem.* **1993,** *268* (4), 2416-2424.
153. Petitpas, I.; Petersen, C. E.; Ha, C. E.; Bhattacharya, A. A.; Zunszain, P. A.; Ghuman, J.; Bhagavan, N. V.; Curry, S., Structural basis of albumin-thyroxine interactions and familial dysalbuminemic hyperthyroxinemia. *Proc. Natl. Acad. Sci. U S A* **2003,** *100* (11), 6440-6445.
154. Wojtczak, A.; Neumann, P.; Cody, V., Structure of a new polymorphic monoclinic form of human transthyretin at 3 angstrom resolution reveals a mixed complex between unliganded and T-4-bound tetramers of TTR. *Acta Crystallogr. D Biol. Crystallogr.* **2001,** *57*, 957-967.
155. Van Mourik, T., Readpdb - A program for extracting fragments from Protein Databank files centred around a selected atom. **2015**.
156. Kroon, J.; Kanters, J. A.; Vanduijneveldtv.Jg; Vanduijneveldt, F. B.; Vliegenthart, J. A., O-H...O hydrogen-bonds in molecular-crystals - statistical and quantum-chemical analysis. *J. Mol. Struct.* **1975,** *24* (1), 109-129.
157. Becke, A. D., Density-functional thermochemistry .3. The role of exact exchange. *J. Chem. Phys.* **1993,** *98* (7), 5648-5652.
158. Lee, C. T.; Yang, W. T.; Parr, R. G., Development of the Colle-Salvetti correlation-energy formula into a functional of the electron-density. *Phys. Rev. B* **1988,** *37* (2), 785-789.

159. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J., Ab-initio calculation of vibrational absorption and circular-dichroism spectra using density-functional force-fields. *J. Phys. Chem.* **1994,** *98* (45), 11623-11627.
160. Rassolov, V. A.; Ratner, M. A.; Pople, J. A.; Redfern, P. C.; Curtiss, L. A., 6-31G*basis set for third-row atoms. *J. Comput. Chem.* **2001,** *22* (9), 976-984.
161. Peterson, K. A.; Shepler, B. C.; Figgen, D.; Stoll, H., On the spectroscopic and thermochemical properties of ClO, BrO, IO, and their anions. *J. Phys. Chem. A* **2006,** *110* (51), 13877-13883.
162. Fanfrlik, J.; Kolar, M.; Kamlar, M.; Hurny, D.; Ruiz, F. X.; Cousido-Siah, A.; Mitschler, A.; Rezac, J.; Munusamy, E.; Lepsik, M.; Matejicek, P.; Vesely, J.; Podjarny, A.; Hobza, P., Modulation of aldose reductase inhibition by halogen bond tuning. *ACS Chem. Biol.* **2013,** *8* (11), 2484-2492.
163. Mantina, M.; Chamberlin, A. C.; Valero, R.; Cramer, C. J.; Truhlar, D. G., Consistent van der Waals radii for the whole main group. *J. Phys. Chem. A* **2009,** *113* (19), 5806-5812.
164. Pauling, L., *The nature of the chemical bond 3rd edition*. Ithaca: Cornell University: 1960.
165. Bondi, A., Van der Waals volumes + radii. *J. Phys. Chem.* **1964,** *68* (3), 441-&.
166. Batsanov, S. S., Van der Waals radii of elements. *Inorg. Mater.* **2001,** *37* (9), 871-885.