THESIS

SOLVING THE CRYSTALLOGRAPHIC STRUCTURE OF THE CL2J CONSTRUCT AND OCCUPANCY TITRATION TRIALS TO QUANTITATIVELY DETERMINE ISOMER RATIO

Submitted by

Brittany L. Rummel

Department of Biochemistry and Molecular Biology

In partial fulfillment of the requirements

For the degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2011

Master's Committee:

Advisor: P. Shing Ho

Norman Curthoys
James Bamburg
Alan Kennan

ABSTRACT

SOLVING THE CRYSTALLOGRAPHIC STRUCTURE OF THE CL2J CONSTRUCT AND OCCUPANCY TITRATION TRIALS TO QUANTITATIVELY DETERMINE ISOMER RATIO

Halogen atoms are commonly found in biological organic compounds such as plastic polymers, flame retardants, coolant fluids, insecticides and herbicides. Halogens are known to mediate neurotransmitters in the brain and are required for the production of many hormones (i.e. thyroxine). Because halogen atoms are frequently incorporated in pharmaceuticals and antibiotics (i.e. clindamycin and chloramphenicol), it is important to characterize the interactions that those atoms participate in. Currently, there is little information known about halogen bonds and these interactions are not modeled accurately by molecular simulations.

The long-term objective of Dr. Shing Ho's laboratory has been to characterize halogen bonds through structural and energetic determinations. As part of that larger goal, the studies in this thesis aim to address the structure-energy relationship of chlorinated halogen bonds or X-bonds. The experimental assay that allowed the study of halogen bonds is the 4-stranded DNA Holliday junction. Incorporating engineered halogen bonds into the structure results in halogen bonds competing energetically against hydrogen bonds for stabilization of the junction. The structure that was refined in order to

analyze chlorinated halogen bonds is referred to as the Cl2J construct. The Cl2J construct is a crystallized Holliday junction crystal in which 2 chlorine atoms are incorporated into the structure as chlorinated uracil nucleotides, and thus, sets chlorine halogen bonding energies and hydrogen bonding energies in opposition.

Occupancy titrations were conducted to quantify isomeric ratios of halogen bond versus hydrogen bond stabilized junctions (X- and H-isomers, respectively) within these crystals. The initial estimate of the isomer ratios of the Cl2J construct was 50/50 X-to-H-isomer from the initial electron density maps. The crystallographic model and subsequent occupancy titration trials actually indicate a higher ratio of approximately 3/1 X-to-H-isomer ratio, respectively. The occupancy titrations and crystallographic models of other constructs, F2J, Br2J and I2J, were analyzed in comparison to the Cl2J construct in order to define a protocol that accurately quantifies these isomeric ratios. Differential scanning calorimetry (DSC) studies are also presented to corroborate in solution any conclusions drawn from occupancy titrations in the crystals.

ACKNOWLEDGEMENTS

Thank you to Dr. Ho, as well as everyone in the Ho lab for their guidance and assistance. I want to offer a special thank you to Megan Carter for her patience and direction as I learned new ideas and scientific methods. I appreciate her encouragement and the way she forced me to challenge myself daily.

TABLE OF CONTENTS

Introduction:

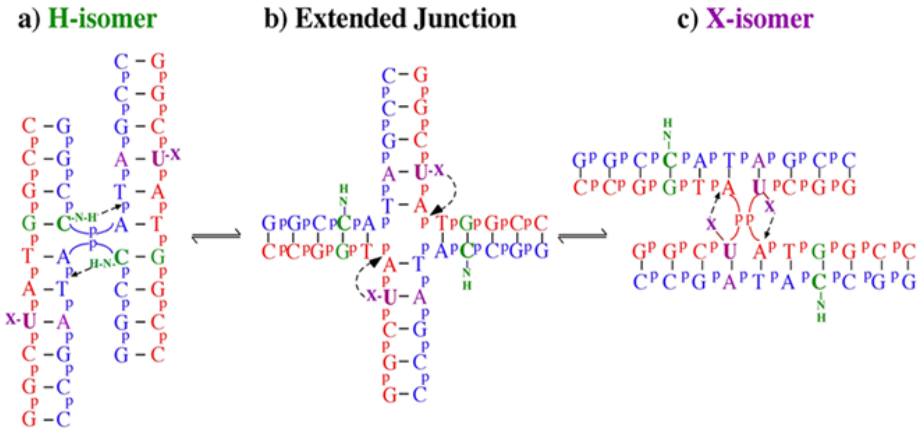
Overview of research project

The focus of this thesis has been to solve the crystallographic structure of the Cl2J construct. The Cl2J construct consists of a Holliday Junction of 4 DNA strands (each ten base pairs in length) and is able to isomerize between the X-and-the-H-isomer forms of the junction. The X-isomer form of the construct is stabilized by halogen bonds (or X-bonds) to adjacent phosphates on the interior strands of the junction. The H-isomer form of the construct is stabilized by hydrogen bonds (or H-bonds) to adjacent phosphates on the interior strands of the junction. The two chlorine atoms incorporated in the structure can reside, either on the inside strands stabilizing the junction, or on the outside strands of this "Cl2J" construct. Additional constructs will be discussed including F2J, Br2J and I2J. All of these constructs are identical, differing only in the halogen that is incorporated into the DNA sequence via a uracil nucleotide base (whether fluorine, bromine or iodine). Table 1 lists the sequences of the 4 constructs that will be discussed, as well as the ratio of halogen (X) bonds and hydrogen (H) bonds competing for stabilization of the junction.

Table 1. List of construct name, DNA sequences comprising each construct and the ratio of halogen (X) bonds and hydrogen (H) bonds competing for stabilization of the junction in each construct. FU denotes a fluorinated uracil nucleotide base. ClU denotes a chlorinated uracil nucleotide base. BrU denotes a brominated uracil nucleotide base and IU denotes an iodinated uracil nucleotide base.

Construct	**DNA sequences**	**# of X:H bonds competing**
F2J	2(CCGATACCGG) + 2(CCGGTA[FU]CGG)	2:2
Cl2J	2(CCGATACCGG)+ 2(CCGGTA[ClU]CGG)	2:2
Br2J	2(CCGATACCGG) + 2(CCGGTA[BrU]CGG)	2:2
I2J	2(CCGATACCGG) + 2(CCGGTA[IU]CGG)	2:2

Holliday Junctions are ideal for assays studying halogen bonds as the complex is naturally occurring in biological systems, and is involved in a number of cellular functions, including, for example, prokaryotic homologous recombination (Hays et al 2003). Holliday junction isomers are distinguished by only a few specific molecular interactions (Carter et al 2011). Thus, subtle variations in the interactions can be exploited to study the energies of these interactions. Figure 1 shows the location of the hydrogen bonds (H-bonds) and the halogen bonds (X-bonds) when the junction is in the H-isomer form and the X-isomer form, and also how isomerization between the two occurs.



(Carter et al 2011)

Figure 1. Competition between H- and X-bonds in a DNA junction. Isomeric forms of the stacked-X junction in H-isomer (a) or X-isomer (c), where X is a chlorine (Cl2J construct). Junction isomerization occurs through an extended junction (b). The isomer form can be distinguished by locating xU, on either the outside strand (H-isomer) or inside strand (X-isomer), and is further confirmed by the identity of the complimentary base at either the inside or outside position (if the complimentary base at the inside position is adenine the junction is in the X-isomer, however if the complimentary base is a guanine the junction is in the H-isomer form).

The crystallographic structure of the Cl2J construct will be solved to determine the distance and angle of the chlorine halogen bond in the Cl2J construct. These parameters will be indicative of whether the halogen bond is occurring or not as well as the strength of the interaction. If a halogen bond were truly occurring, the distance between the chlorine and the phosphate oxygen would be less than the sum of both atoms' van der Waals radii (Voth et al 2007), or in this case, 3.27 Å.

Additionally, using the final model from refinement, the isomer ratio within the crystal will be determined. Isomer ratio in the crystal is directly correlated to the ratio of the stabilizing interactions of the two isomers. Since the stabilizing interactions are either halogen bonds or hydrogen bonds, the isomer ratio in the crystal reflects the relative

strength of the halogen versus hydrogen bond energy. It is critical to note the assumptions made here: First, it is assumed that the molecular structure acts identically in a crystal phase as it does in solution. Second, it is assumed that the two isomer forms of the construct differ by only the stabilizing interactions of the junction (Voth et al 2007).

The crystallographic model can give an estimate of the isomer ratio, but to more accurately determine this ratio, occupancy titration methods will be applied to the final model from refinement. Occupancy titration methods involve monitoring the R and R-free values (defined in Materials and Methods: *Crystallographic refinement process* section) for the model, both as the X-isomer form increases from 0% to 100% occupancy and the H-isomer form decreases from 100% to 0% occupancy, and the X-isomer decreases from 100% to 0% occupancy and the H-isomer increases to from 0% to 100% occupancy. A baseline R and R-free where the occupancies remain unchanged are also taken, and then subtracted, yielding a graph plotting occupancy versus the change in R or R-free for the model. The resulting titration curves yield a numerical value for the isomer ratio of the construct in the crystal, and therefore, an estimate for energy of the halogen bond relative to the hydrogen bond.

It is crucial to better-characterize the energies of halogen bonds, as pharmaceuticals and antibiotics are commonly halogenated (two well-known examples are clindamycin and chloramphenicol). Accurate models for the energies of halogen bonds will allow halogens to be purposefully incorporated into methods for rational drug design. The long-term goal is to determine the relationship of halogen bond geometries and their associated energies. The focus of this thesis within this larger project is to

determine the structure of a chlorinated halogen bond in the DNA junction system (the Cl2J construct).

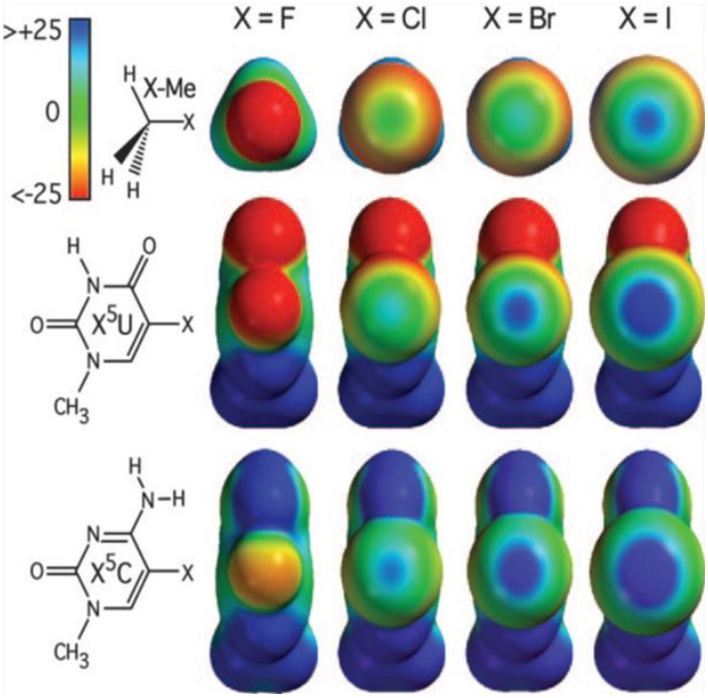
This thesis will describe the process of refining the crystallographic structure of Cl2J and variations in refinement method required to account for the intrinsic complexity of the Cl2J construct (the structure was refined as two overlapping models, representing the competing X- and H-isomers). Various different occupancy titration programs were applied in an attempt quantify the ratio of X- and H-isomers in the crystal. The results of each occupancy titration method will be presented to rationalize the method that was deemed to yield the most accurate model. Finally, the selected model for occupancy titration will be applied to determine the X- to H-isomer ratios for the Cl2J, as well as the analogous iodinated I2J, fluorinated F2J and brominated Br2J DNA junction constructs.

Background information

Halogen bonds are electrostatic non-covalent interactions in which a halogen atom (covalently bound to a substituent) is polarized, forming anisotropic regions of negative and positive electrostatic potential. The substituent group forms a sigma bond to the halogen, which pulls the electron from and, thus, depopulates the *Pz*-orbital from the outer electron shell. This "sigma hole" model has been used to rationalize of the tendency of halogen atoms to become polarized, leaving an electropositive crown aligned along the sigma bond, and an outer ring of negative potential perpendicular to the bond (Politzer et al 2010). The polarizations of various halogens are shown in figure 2.

An electron rich atom, or Lewis base, is attracted to the electropositive crown resulting from the sigma hole. Therefore, it follows, that halogen bonds are extremely

directional interactions and are aligned along the sigma bond (at a 180° angle). The resulting interaction distance of a halogen bond is that the bond length is shorter than the sum of the van der Waals radii of the halogen atom and its interacting partner (Auffinger et al 2004).

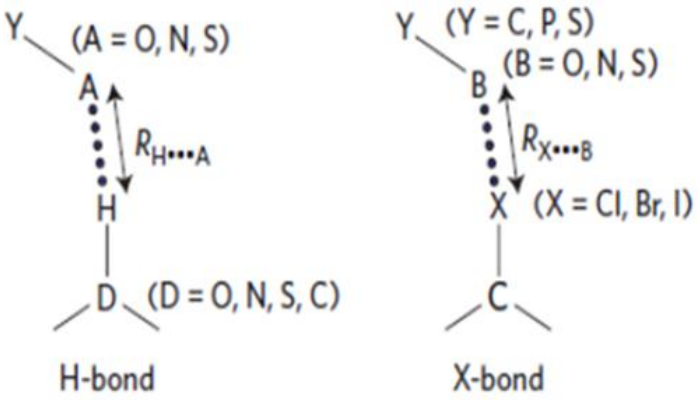


(Auffinger et al 2004)

Figure 2. Electrostatic potential surfaces of halogenated model compounds. Halogenated methane (X_Me, top), uridine nucleobase (X5U, middle), and cytosine nucleobase (X5C, bottom) are shown looking into the halogen atoms to compare the induced negative (red), neutral (green), and positive (blue) electrostatic potentials around the halogen surfaces. The potential energies are scaled from -25 to +25 kcal/mol range to emphasize the variation in electrostatic potential associated with the halogen atoms. The compounds are ordered (from left to right) from least to most polarizable (F < Cl < Br < I).

The hydrogen bond has been defined by the International Union of Pure and Applied Chemistry (IUPAC) as, "an attractive interaction between a hydrogen atom from a molecule or fragment R-H in which R is more electronegative than H, and an atom or

group of atoms B in the same or a different molecule, where there is evidence of bond formation (Legon A. 2004, Arunan et al).” An analogous definition can be applied to halogen bonds as well for halogen bonds as well (Figure 3). The halogen bonding and hydrogen bonding interactions share many properties, including their directionality, specificity and relative strength. It is well-known that hydrogen bonds stabilize secondary and tertiary structural motifs in biological systems. Only within the past couple of years has it become recognized that halogen bonds can perform many of the same functions, and also compete with hydrogen bonds in biological complexes (Metrangolo et al 2008).

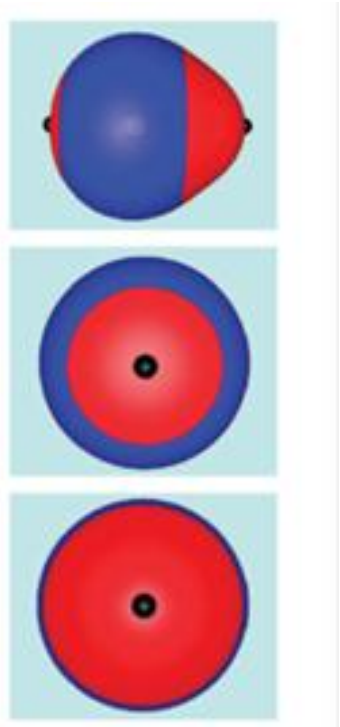


(Voth et al 2009)

Figure 3. Comparison of an H-bond with an X-bond interaction with a Lewis base acceptor (labeled A for the H-bond and B for the X-bond). The interactions are characterized by a hydrogen to-acceptor distance RH-A or halogen-to-base distance RX-B that is shorter than the sum of their respective van der Waals radii.

Although hydrogen bonds are typically referred to as highly directional interactions, the halogen bond is much more likely to be a linear interaction. A molecule that demonstrates this concept well is HBr (Figure 4, Politzer et al 2010). Looking at the molecule perpendicular to the sigma bond, one can see that both the hydrogen and the bromine have regions of positive electronic density. However, the bromine has the

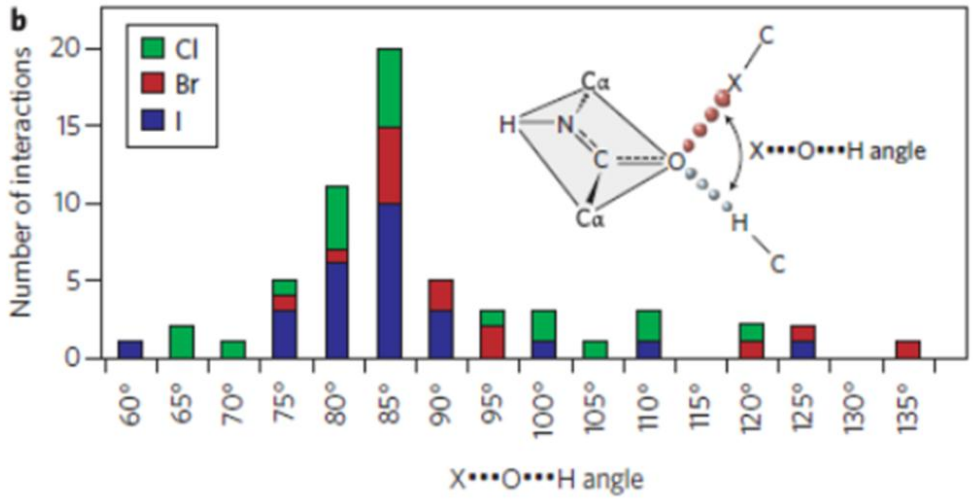
positive crown that is more concentrated in the center of the atom, along the axis of the H-Br bond. This is in contrast to the hydrogen atom with essentially an even distribution of positive charge across the entire atomic surface. Thus, the possibilities for bonding of the halogen are limited to angles close to 180^0 to the sigma bond angle, while the hydrogen has a more varied range of bonding angles, from linear to perpendicular interactions in reference to the sigma bond. Replacing the covalently bonded hydrogen with a more electronegative substituent will more strongly polarize the halogen, allowing a narrower range of favorable halogen bond angles.



(Politzer et al 2010)

Figure 4. Computed electrostatic potential on the .001 au molecular surface of HBr. Three views are shown: (top) perpendicular to the H-Br axis; (middle) bromine in the foreground; and (bottom) hydrogen in the foreground. Color ranges: red, greater than 0 kcal/mol (positive); and blue, less than 0 kcal/mol (negative). The locations of the most positive potentials are indicated by black hemispheres.

Halogen bonding interactions are energetically independent of hydrogen bonds in that the addition of one of these interactions does not alter the strength of the other bonding interaction when both share a common carbonyl oxygen acceptor group. This is because the oxygen has both in-plane and out-of-plane electron systems (Voth A.R. et al 2009), resulting in "orthogonal" behavior between the halogen and hydrogen bonds. The orthogonal relationship extends to the geometric relationship between the two interactions. An analysis of crystal structures from the Protein Data Base (PDB) showed that X-bond and H-bonds, sharing a common oxygen atom, are related by a very narrow distribution around an angle of 88.2° (Figure 5, Voth et al 2009).



(Voth et al 2009)

Figure 5. Histogram of the relative angle of approach of H- and X-bonds to a shared Lewis base (oxygen) for each type of X-bond (Cl, Br or I) observed in structures in the Protein Data Bank.

This thesis will describe the crystallographic refinement process of the Cl2J construct, including programs applied and important parameters used in refinement, followed by the occupancy titration process, conducted in order to

quantitate isomeric ratio within the crystal. The model from refinement and occupancy titration of Cl2J will be compared to the models and occupancy titrations of other halogenated junction complexes (fluorinated F2J, brominated Br2J and iodinated I2J constructs). DSC studies of Br2J (conducted by Megan Carter) will be presented to corroborate the validity of the results obtained from crystallographic studies and occupancy titrations. This thesis will conclude with a summary of the determined structural and energetic relationship of chlorine halogen bonds.

Materials and Methods:

Crystallographic refinement process

The crystallographic data for the Cl2J construct was collected, integrated and reduced by Megan Carter, a PhD student in Dr. Ho's lab. The structure was solved by molecular replacement, starting with the brominated Br2J DNA junction as the starting model (PDB code 2ORG.pdb, Voth et al 2009). EPMR is a molecular replacement program that uses an evolutionary search algorithm (http://www.msg.ucsf.edu/local/programs/epmr/epmr.html).

The initial model started with R and R-free values of 32.26% and 35.28%, respectively. Refinement was conducted using the computer program CNS (Crystallographic and NMR Structure Determination) and consisted of preliminary rigid, rigid parts and simulated annealing steps, followed by several rounds of positional and B-factor refinement, and water and ion molecules incrementally added (to build the solvent model). The quality of fit of the DNA and solvent model to the X-ray data is measured by the R- and R-free values. The R value (or residual factor) is defined by the equation:

$$R = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$$

where |Fobs| is the magnitude of the structure factor determined from the observed X-ray data, while |Fcalc| is the magnitude of the structure factor calculated from the model. Generally, 5-10% of the original diffraction data is excluded from the rest of the data and not included in refinement. The R-free value is a measure of

the agreement between the crystallographic model and the 5-10% of diffraction not included in refinement (Brunger A.T.), thereby serving as an independent cross-validation of the refinement. Electrostatic interactions and packing van der Waals interactions were left off during refinement to keep the atoms of the competing X-isomer and H-isomer models from expelling each other. No solvent was added near atoms whose identity would distinguish between X-isomer form and H-isomer form of the construct so as not to bias the electron density maps in favor of one or the other model. The final refinement was conducted using a four-stranded model generated from the symmetry coordinates of two strands with all atom occupancies set to 0.50. Following each round of refinement and solvent addition, the program COOT (Crystallographic Object-Oriented Toolbox) was used to visualize the electronic density maps and refined models. Tables 2 and 3 list crystallographic parameters and refinement statistics for the Cl2J construct.

Table 2. Crystallographic parameters for the Cl2J construct.

Space group	*C2*
Unit cell: *a, b, c* in Å	65.69, 23.57, 37.29
β-angle, degrees	110.92
Number of unique reflections	4793
Resolution, Å	50-1.84
Completeness, %*	73.8 (44.5)
I/sigma, I*	27.496 (1.95)
R-merge, %*	0.053 (0.399)

*Values for the highest resolution shell are in parentheses.

Table 3. Refinement statistics for the Cl2J construct.

R, (R-free), %	23.13 (25.95)
No. of atoms: DNA (solvent)	808 (100)
<B-factor> DNA (solvent)	11.80 (18.83)
RMSD bond length, Å	0.31
RMSD bond angle, degrees	0.492

Occupancy titration process

Three main types of occupancy titration methods were applied to the Cl2J construct in an attempt to quantitatively analyze differences in their abilities to accurately determine the ratio of X- and H-isomers in the DNA junction systems. The supplementary material describes the initial occupancy titration process and the occupancy titration refinement process. A comparison of the various methods led to the conclusion that, for all halogenated constructs, including Cl2J, a group occupancy titration was most appropriate.

In a group occupancy titration, only atom's occupancies whose presence will differentiate between the X-isomer and H-isomer form of the construct are analyzed. The occupancy of an atom is the value of the frequency that the atom resides in that location. In this procedure, the occupancies of each of these selected atoms are varied to generate electronic density maps. In the case of the Cl2J construct, the occupancies of the halogen atoms and N2 Nitrogen atoms of the guanine residues of complimentary sequences are titrated with starting values of 0.01, 0.51 and 0.99 and increased or decreased with X-isomer atoms and H-isomer atoms going in opposing directions. Control group titration

R- and R-free measurements were determined for each trial and then subtracted from experimental values to yield the occupancy titration data.

Results and Discussion:

Crystallographic refinement of the Cl2J construct

Refinement of the Cl2J structure started with two complementary DNA strands 5'-CCGGTA(ClU)CGG-3' and 5'-CCGATACCGG-3' (ClU is a chlorinated uracil nucleotide) in the crystallographic asymmetric unit and at a resolution of 1.94 Å. The electrostatic interactions and packing van der Waals interaction terms were excluded from the refinement constraints to allow the model to accommodate the short distances correlated with a halogen bond for the purpose of allowing the model to exhibit the distance correlating to a halogen bond (which could be shorter than the sum of the van der Waals radii of the two interacting atoms). During the refinements, the hydrogen bonding range of 2.2 –3.5 Å was used as the distance criteria for adding waters.

Between refinement models (mods) 30 and 50, multiple weighting factors for the data and the geometric libraries (optimize_wa.inp and optimize_rweight.inp) were applied. The r-weight value was held at or close to its original value of 1.0. However, the wa values that were suggested by the refinement program were set in a range from 2.5 to 3.5. Potentially, a problem was created here, by weighing the refinement more heavily towards the structure and away from the library. Higher wa values dictate to the refinement programs to consider the data as higher importance over consulting molecular libraries. Higher r-weight values indicate the opposite to the computer programs, to rely more heavily upon the molecular libraries for information. As this structure is

intrinsically a “somewhat even” mix of isomers, this action most likely perturbed the structure into favoring one conformation above the other. Thus, the result of this refinement yielded a structure strongly favoring the X-isomer over the H-isomer, which may be potentially true. However, because the ratio of the 2 isomers indicates strength of the 2 bonds, this ratio is most probably artificially enhanced and therefore not accurate.

Around mod 30, the occupancies of the chlorines at residues 7 and 17 were changed to 0.5 from 1.0, which rationally makes more sense as one chlorine atom can’t be in both places 100% of the time. The occupancies of the N2 atoms of guanine 4 and guanine 14 were also changed to 0.5 from 1.0, following the same idea of where the chlorine is appearing. By changing the occupancies of the N2 atoms, the generated model will allow better indications whether the guanine is there (with the presence of the N2), or whether that residue is an adenine. These results should correlate well and also indicate the isomer ratio percentage. A further issue occurring with this first refinement is the possibility of over-refinement. The R- and R-free values were consistently dropping, until a point on which they began systematically increasing. Being initially inexperienced with the refinement process, I didn’t recognize this as over-refining, but rather assumed I had added bad waters somewhere in the process. This is also most likely true, as the structure is relatively complex, and many other issues had to be addressed before the final successful refined construct. A prominent problem occurring in the refinement of the Cl2J construct is that the isomer ratio is so close that if one configuration is slightly more prevalent, waters are added that correlate well with that structure, biasing the refinement and eventually weighing heavily towards one configuration, making it appear much more

favored than it is actually. The first refinement was stopped before completion and occupancy titration, as there were many potential issues clouding any results.

I began the next refinement from the beginning, making some changes in programs and procedure. The high resolution was set to 1.84 Å initially and kept there throughout refinement. The occupancies of the chlorines and the two N2 atoms on guanines 4 and 14 were also set to 0.5 initially. The same exclusions for electrostatic interactions and van der Waals interactions were applied. This refinement ended at mod97.pdb, with R=21.22 and R-free=26.55. It was noticed that waters were added near (within hydrogen bonding distance) the N2 atoms of residues 4 and 14 which would perturb electron densities, thus refinement was also stopped.

For the next refinement attempt, the original 2-stranded structure was converted to a 4-stranded structure, and the occupancy of all atoms was set to 0.50 (50%). This method was applicable to the Cl2J construct because 2 chlorines will be present instead of one. Ideally, this way, not one conformation of the structure will be so preferred over the other conformation that the less prevalent isomer's contributions can't be differentiated, as is problematic in refining an evenly mixed isomer structure, such as this. A chlorinated uracil was present at positions 7 and 37, a cytosine at 17 and 27, an adenine at 24 and 34, and a guanine at 4 and 14. The decision was also made to fix the end bases of all 4 DNA strands in an attempt to minimize backbone movement. This action could be taken as the Cl2J construct belongs to the *C2* space group (where edge lengths *a*, *b* and *c* are all different and α- and ɣ-angles equal 90° while the β- angle is not equal to 90 or 120°). This was an issue as previous refinements had DNA strands that initially began as exact symmetry partners that eventually incorporated drastic

perturbations in the DNA backbone, voiding the relevance of any bond or interaction distance measurements. To fix the DNA ends, residues 1, 10, 11, 20, 21, 30, 31, and 40 were fixed in the fvdminimize.inp file. The 4-stranded model was then generated with fixed ends before any water molecules were added. These alterations were crucial because observing changes in the bond length of the inside chlorine atom to the phosphate oxygen and the angle of the bond angle created was of particular emphasis to extract from the refinement.

After 30 waters were added to this model, those were saved and put back into the original model after conferring with the electron density maps that ensured the selected water molecules were well-chosen. The purpose of this was to provide better phasing information for the initial model to incorporate. A concern was that the initial model didn't have enough raw data to start with, and must respond drastically to information added, whereby possibly getting "stuck" in a conformation as an intermediate to the refined structure and not able to adapt further. I wanted to avoid adding possible-but-not-great waters, which have an accumulating effect on the DNA's structure, pushing it into an undesirable conformation. Adding the first 30 waters serves to give relatively definite data to the initial structure, as the initial waters are usually the strongest and most well-defined.

Early on in the refinement, observations of the inside-strand chlorine interactions were disappointing. At mod 9, the R- and R-free values were 28.46 and 31.54, respectively. The bond distance of the chlorine to closest phosphate oxygen was 3.60 Å, still farther than the sum of their van der Waals radii and not able to exhibit a halogen-bonding interaction. By mod 20, the R- and R-free values had dropped to 26.92 and

29.91, respectively. The specified chlorine-oxygen bond distance had also dropped to 3.29 Å, which is approximately the sum of the 2 van der Waals radii (actual = 3.27 Å). The bond angle had also decreased to 161.90°. By mod 30, 53 water molecules had been added to the model and R- and R-free values hovered at 25.45 and 27.55, respectively. The wa and r-weight values were allowed to float in value, optimizing the R- and R-free values for each model by relying more or less on the data or the molecular library. The optimizing programs were applied multiple times throughout refinement to ensure that these parameters were appropriate. By mod39, 68 water molecules had been successfully added and the R- and R-free values hit 24.65 and 26.65. Upon observing this structure in COOT however, the halogen bond angle was 157.61° which was as predicted, but the bond distances to the 2 phosphate oxygens were 3.50 Å and 3.53 Å, which indicates that the phosphate was not adapting to the positioning of the chlorine, assuming that the halogen bond was forming here. Many overlooked or unconsidered settings in the refinement programs may be causing this result, or perhaps some creativity was needed in fixing the DNA backbone in a way that stabilized it in a conformation able to flexibly respond to the refinement, but that also minimized the artificial parameters defined to the DNA so that the integrity of the obtained results was upheld. This trend with bond distance and angle continued to mod 69 where R- and R-free values were 22.70 and 25.95, respectively.

At this point, it was recognized that the DNA strands that supposedly began in the exact same positions had migrated drastically, affecting the residue positioning even more. This phenomenon should not be occurring to such a degree as the differences in the bases are so subtle and minute, as well as most of the bases in each strand are exactly the

same. The starting 4-stranded model was analyzed, and it was found that the DNA ends were only fixed after some water molecules had been added, fixing the DNA end residues in a perturbed position. I decided to save all the waters and their coordinates from mod 69 (100 in total), go back to the initial model and fix the ends properly. To analyze this decision, 4 models were looked at simultaneously in coot: m2, which was the correct DNA strand overlay with ends fixed, m3, which was m2 with 1 round of refinement conducted, m69, which was the endpoint of the last refinement with ends fixed incorrectly, and m70, which was all final waters from m69 with correct coordinates of DNA strands 3 and 4 substituted in from the original overlay. In mod 70, the water molecules aligned somewhat with electron density maps, but not at all acceptable, yielding an R- and R-free of 24.58 and 30.51, respectively. It could be determined from these different models that the refinement must start from the beginning with the corrected DNA overlay and ends fixed. Furthermore, additional atoms may need to be held in place to avoid artificial and/or drastic shifts in the DNA backbone that would most likely lead to structure conformations unable to continue to adapt or "stuck" DNA conformations. Water molecules from mod 69 were slowly added back into the model, mostly according to their b-factors. The aforementioned halogen bond distance and angle yielded similar results as the previous refinement for bond angle (around 170°), but the adenine 6 phosphate now orientated so as to allow a much shorter bond distance from the oxygen to the chlorine of residue 7 (around 2.50 Å). This was closer to what was predicted, so refinement continued and R- and R-free values began dropping. At mod 17 in refinement, it was apparent that the DNA strands had already rotated somewhat away from their starting structure. When compared to the model with only the original DNA

strand overlay, specific residues could be identified that were much more drastically perturbed than others. Although cytosines 17 and 27 were 2 of those residues, I didn't choose to modify them (by fixing them), as their symmetry partners were chlorinated uracils 7 and 37. I believed that action would have likely affected the positioning of the junction and compromised any results obtained. In the next 3 cycles of refinement, it was decided to fix the phosphate atom of thymine 25 (id 484) and the phosphate atom of thymine 5 (id 80). Fixing the phosphates of residues 5 and 25 served to stabilize the center of the DNA strands and lessened possible larger movements of the strands. The residues forming the junction though, continued to appear unable to adapt into their desired positions. This was indicated by electron density maps observing bases 6, 7, 16, 17, 26, 27, 36, and 37. This was an issue because it was necessary to minimize the artificial parameters applied to residues involved in the junction or residues important for final isomer ratio determination.

Since residue 36 was highly perturbed, and its symmetry residue (residue 26) was not directly involved in the junction interaction with the chlorine, it was attempted to fix the phosphate atom of adenine 36 (id 705), as well. This single action had a profound effect on the resulting model. Most of the residues seemingly relaxed into closer positions with their symmetry atoms. Also remarkable, the junction appeared to have much more availability for movement, especially the phosphate group of adenine 6. Since these atoms would be responsible for orientation for a halogen bond, it is essential that this residue isn't held in a location not allowing an interaction. Progression through the refinement was fairly easy and after these changes. New water molecules were chosen from mod 20 on as well, as previously chosen water molecules were not acceptable.

Refinement successfully completed with mod 52.pdb. 99 solvent molecules were incorporated and the R- and R-free values hit 23.13 and 25.95, respectively. A table with mod 21 through mod 52 R- and R-free values, X-bond distances, X-bond angles, and number of solvent molecules can be seen below, in table 4.

Table 4. Final refinement data of Cl2J from models (mod) 21 to (mod) 52. Columns from left to right are mod number, R-value, R-free value, X-bond distance from chlorine on CLU 7 to each of the two phosphate oxygen atoms of adenine 6, X-bond angle formed from CLU 7 sugar to chlorine to closest phosphate oxygen and number of solvent molecules included in the model.

Mod #	**R-value**	**R-free value**	**X-bond distance (Å)**	**X-bond angle (degrees)**	**# of solvent molecules**
Mod 21	27.29	31.44	2.57, 3.84	159.93	43
Mod 22	26.8	31.12	2.59, 3.91	161.09	47
Mod 23	26.41	30.24	2.59, 3.92	161.34	49
Mod 24	26.24	29.92	2.57, 3.95	162.13	52
Mod 25	26.17	29.79	2.57, 3.97	163.31	53
Mod 26	26.01	29.26	2.60, 4.03	164.78	55
Mod 27	25.69	28.8	2.62, 3.97	161.73	57
Mod 28	25.55	28.34	2.61, 3.95	161.64	58
Mod 29	25.39	27.81	2.62, 3.93	161.37	62
Mod 30	25.24	27.49	2.71, 3.94	161.81	64
Mod 31	25.18	27.07	2.80, 3.90	161.71	66
Mod 32	24.98	27.81	2.82, 3.86	160.27	68
Mod 33	25	27.47	2.78, 3.88	159.93	69
Mod 34	24.79	27.11	2.71, 3.83	159.75	71
Mod 35	24.65	27.39	2.65, 3.77	160.6	73
Mod 36	24.66	27.13	2.69, 3.78	160.92	73
Mod 37	24.76	28.31	2.70, 3.83	162.5	71
Mod 38	24.63	27.47	2.72, 3.87	163.98	74

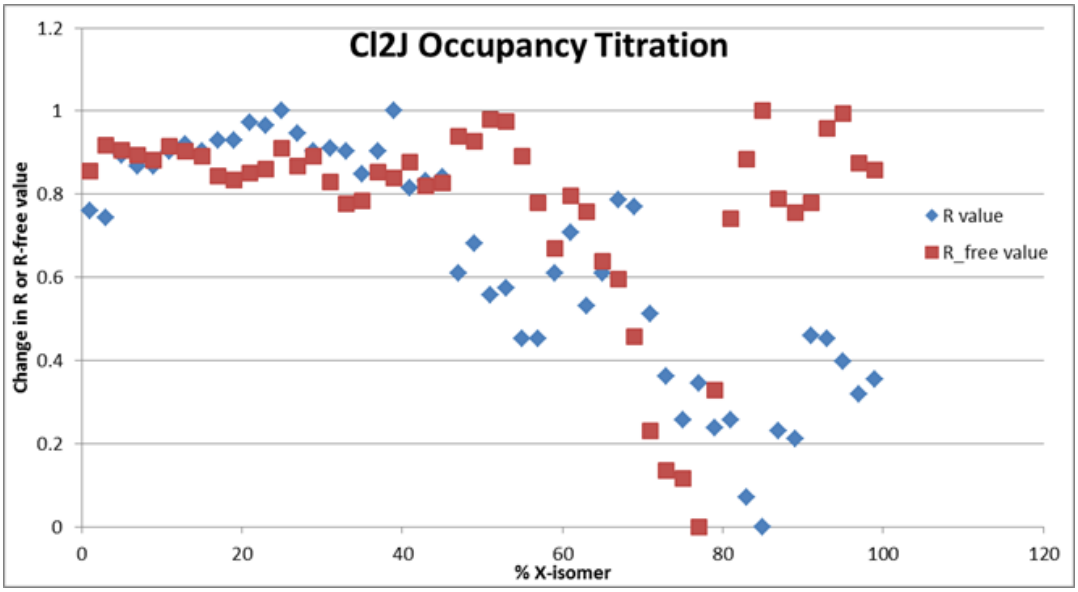
Mod 39	24.45	27.54	2.66, 3.76	160.05	77
Mod 40	24.34	29.08	2.62, 3.70	159.02	79
Mod 41	24.4	29.12	2.61, 3.72	156.3	79
Mod 42	24.38	29.19	2.62, 3.73	158.78	78
Mod 43	24.49	28.58	2.59, 3.74	159.76	79
Mod 44	24.24	28.12	2.59, 3.71	158.11	82
Mod 45	24.04	27.3	2.55, 3.80	163.29	85
Mod 46	23.8	26.99	2.59, 3.85	166.28	87
Mod 47	23.6	26.49	2.60, 3.86	166.27	90
Mod 48	23.42	26.05	2.61, 3.81	165.41	92
Mod 49	23.38	26.32	2.60, 3.79	164.23	92
Mod 50	23.29	26.05	2.61, 3.86	166.56	95
Mod 51	23.36	26.33	2.63, 3.86	164.84	96
Mod 52	23.13	25.95	2.62, 3.73	160.81	99

At m52.pdb, the final model of the refinement of the Cl2J construct, there were 96 water molecules in total integrated into the construct. In addition, three sodium ions were identified in the model.

Occupancy titration of the Cl2J construct

The occupancy titration method used to analyze the Cl2J construct was referred to as a group occupancy titration. In this method, the R- and R-free values for the entire construct were monitored while the occupancies of only four atoms were being titrated and the rest of the construct remained at a static occupancy. The occupancies of the chlorine atom of residues 7 and 37 (both chlorinated uracils) and the N2 positioned nitrogen of residues 4 and 24 (both guanines) were titrated from 0% to 100%, from 100%

to 0%, from 50% to 100% and from 50% to 0%. The R- and R-free values of those four residues was then recorded and analyzed. This result was relevant as all of the remaining atoms in the construct can essentially be considered to be present in the same location whether the construct has taken the X-isomer form or the H-isomer form. Figure 6 shows the group occupancy titration of the Cl2J construct with the percentage of X-isomer plotted against the change in R- (blue) or R-free value (red). The minimums of both curves hover around 75-80% X-isomer. These results correlated well with previous occupancy titration results. Figure 7 shows an image created using the Pymol program of the Cl2J construct inside position (residues 6 and 7), with the chlorine atom circled in red and shown at 0.75 occupancy, the percentage indicated by the group occupancy titration.



75-80% X-isomer

Figure 6. A plot of the group occupancy titration of the Cl2J construct, plotting X-isomer percentage versus the change in R- or R-free value.

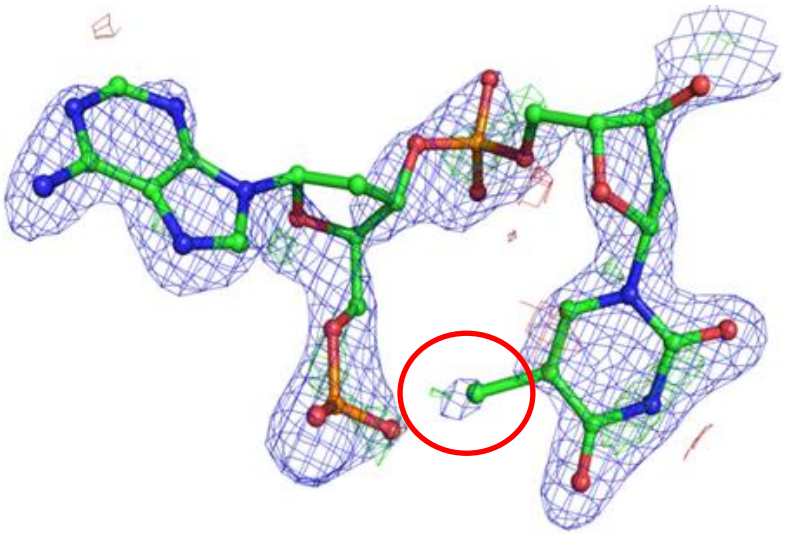


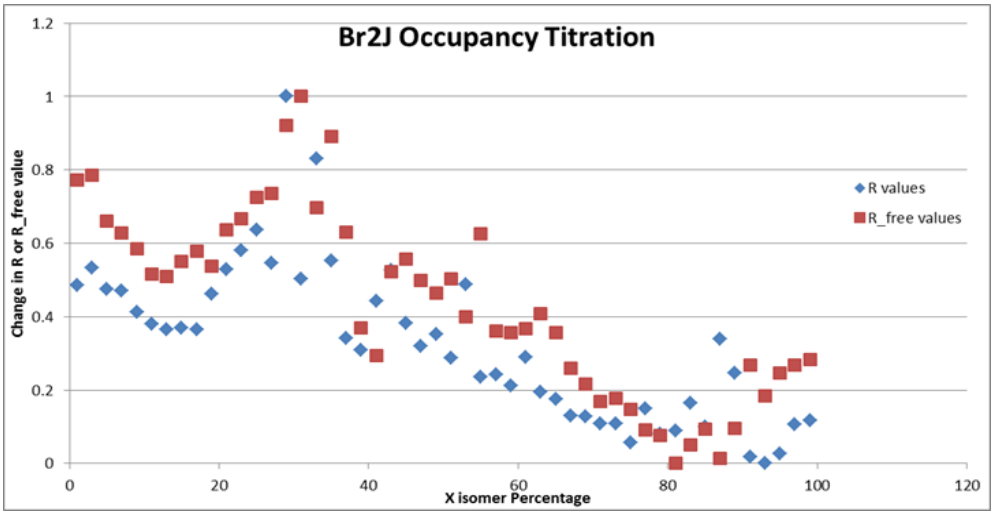
Figure 7. A pymol image of residues 6 and 7 in the Cl2J construct. The red circle designates the chlorine atom of chlorinated uracil 7. The chlorine atom is shown at an occupancy of 0.75, as the group occupancy titration indicated. The blue electron density was created from the 2Fo-Fc map with a carve of 2.0 Å and sigma equal to 2.0. The green electron density was created from the Fo-Fc map at 1.7 sigma and a carve equal to 1.5 Å. The red electron density was created from the Fo-Fc map at -1.7 sigma and a carve equal to 1.1 Å.

Many other occupancy titration methods were employed in order to analyze the Cl2J construct. All of the other methods generally agreed with the results yielded by the group occupancy titration method but were not chosen to be the conclusive titration method as their accuracies varied somewhat due to the abilities of the incorporated halogen atoms to shift the construct into stronger electron density, averaging any occupancy or isomer percentage results.

Occupancy titrations of Br2J, F2J, and I2J constructs

Next, I attempted a group occupancy titration on the Br2J construct, which had already been analyzed and determined to be in 100% X-isomer form by Dr. Andrea Voth, a former PhD student in Dr. Ho's lab. Figure 8 displays results of the group occupancy

titration with only the four differentiating atoms being titrated. Figure 9 shows a visual of the Br2J construct inside position (residues 6 and 7) created using the Pymol program. The bromine atom of brominated uracil 7 is circled in red.



90-95% X-isomer

Figure 8. A plot of the group occupancy titration of the Br2J construct, plotting X-isomer percentage versus the change in R- or R-free value.

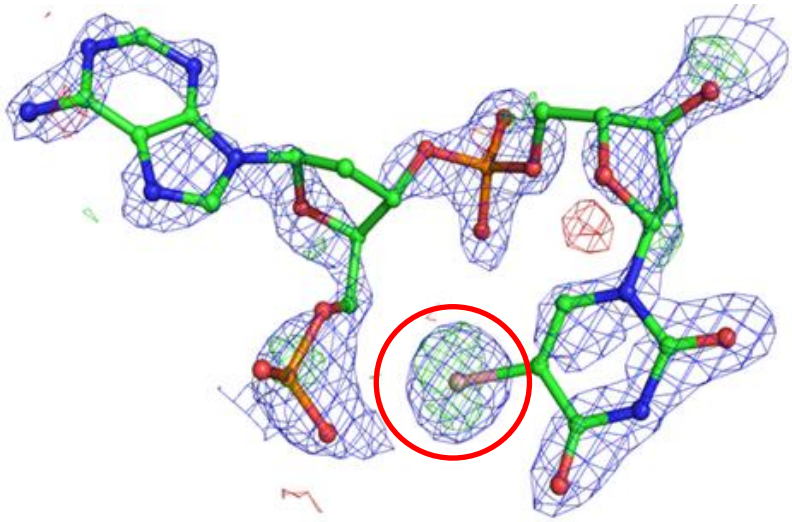
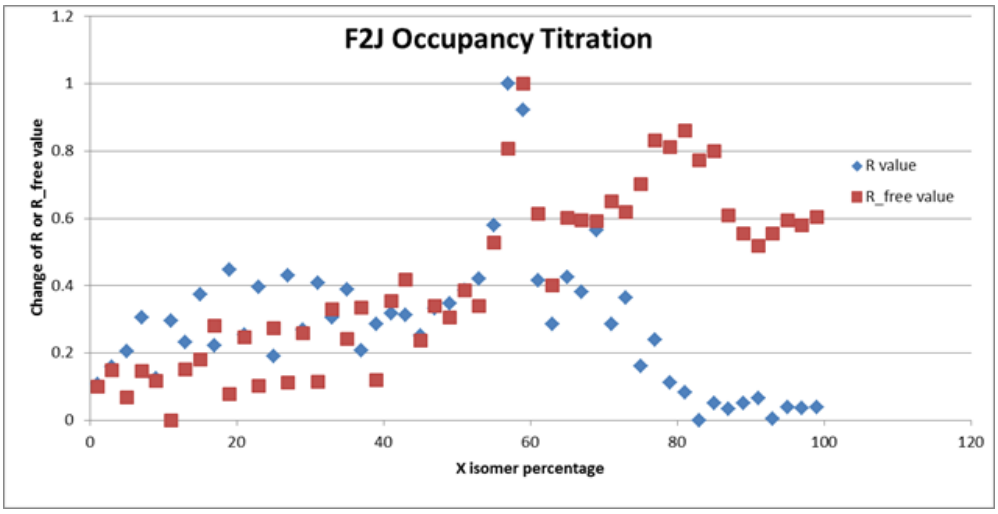


Figure 9. A pymol image of residues 6 and 7 in the Br2J construct. The bromine atom is circled in red and shown at an occupancy of 0.99, as the group occupancy titration indicated (~100%). The blue electron density was created from the 2Fo-Fc map with a carve of 2.0 Å and sigma equal to 2.0. The green electron density was created from the Fo-Fc map at 1.7 sigma and a carve equal to 1.5 Å. The red electron density was created from the Fo-Fc map at -1.7 sigma and a carve equal to 1.3 Å.

As the group occupancy titration appears to approximate X-isomer percentage fairly well for all constructs, this method of occupancy titration was decidedly the most accurate overall. The group titration method is not clouded with data from the hundreds of atoms in the structure that do not change locations between isomer forms of the construct. Therefore, I also performed group occupancy titrations on the refined structures of the F2J and I2J constructs (figures 10 and 12, respectively). The F2J and I2J constructs were originally refined by Matthew Scholfield (a current graduate student in Dr. Ho's lab) and Dr. Andrea Voth (a former graduate student in Dr. Ho's lab), respectively. These results match those expected based upon the degree of polarization ability of the chlorine, fluorine and iodine halogen atoms. With each occupancy titration graph, a visual generated in the Pymol program is shown (figures 11 and 13). The halogen in each one is circled in red and set to the occupancy dictated by the group occupancy titration, thus, creating images with approximate actual electron density maps and electron potential.



~0% X-isomer

Figure 10. A plot of the group occupancy titration of the F2J construct, plotting X-isomer percentage versus the change in R- or R-free value.

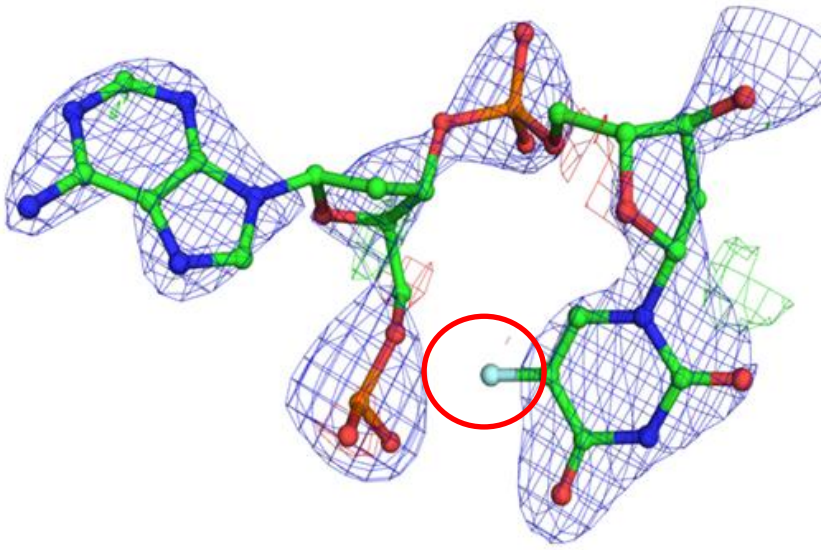
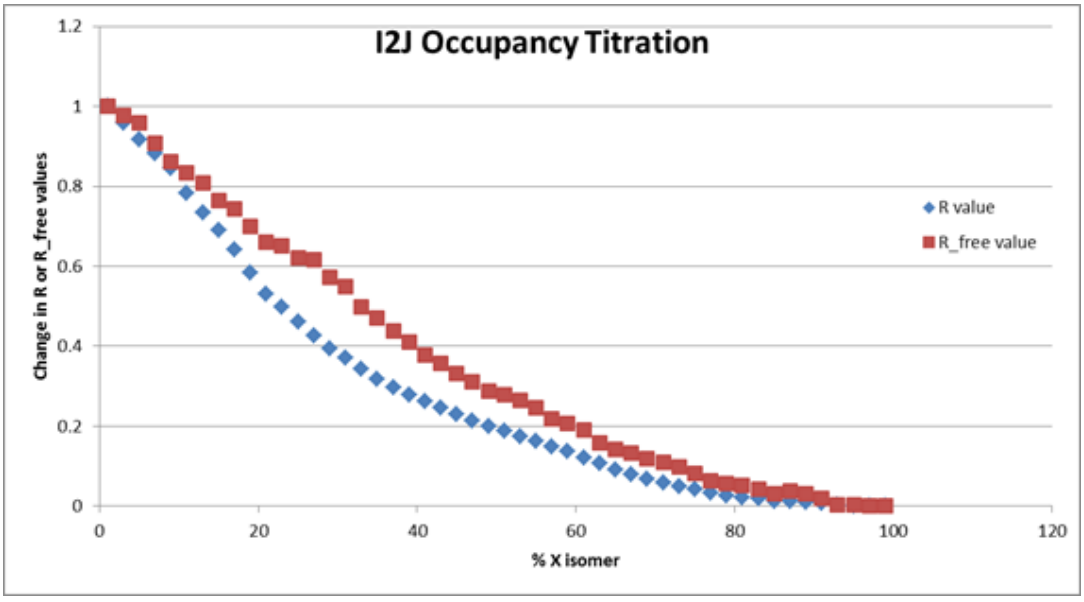


Figure 11. A pymol image of residues 6 and 7 in the F2J construct. The fluorine atom is circled in red and shown at an occupancy of 0.01, as the group occupancy titration indicated (~0%). The blue electron density was created from the 2Fo-Fc map with a carve of 2.0 Å and sigma equal to 2.0. The green electron density was created from the Fo-Fc map at 1.7 sigma and a carve equal to 1.5 Å. The red electron density was created from the Fo-Fc map at -1.7 sigma and a carve equal to 1.2 Å.



100% X-isomer

Figure 12. A plot of the group occupancy titration of the I2J construct, plotting the X-isomer percentage versus the change in R- or R-free value.

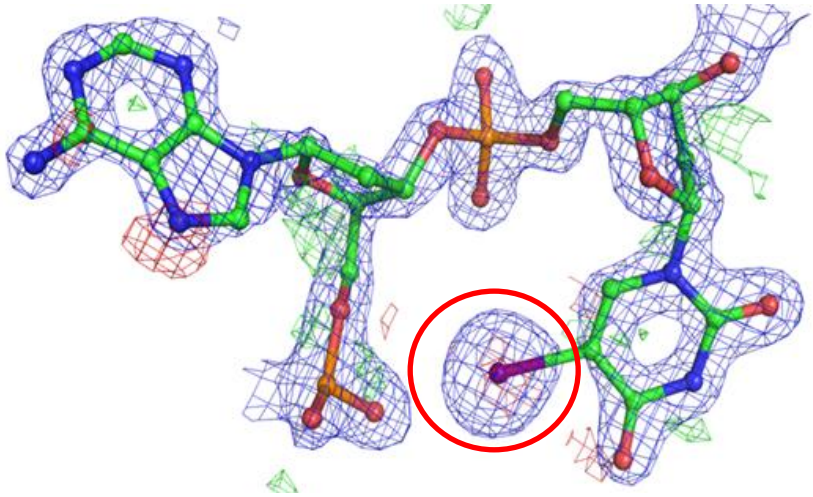


Figure 13. A pymol image of residues 6 and 7 in the I2J construct. The iodine atom is circled in red and shown at an occupancy of 0.99, as the group occupancy titration indicated (~100%). The blue electron density was created from the 2Fo-Fc map with a carve of 2.0 Å and sigma equal to 2.0. The green electron density was created from the Fo-Fc map at 1.7 sigma and a carve equal to 1.5 Å. The red electron density was created from the Fo-Fc map at -1.7 sigma and a carve equal to 1.1 Å.

Solution studies

In addition to occupancy titration methods, studies using Differential Scanning Calorimetry (DSC) were conducted to determine the energies of halogen bonds in solution. The I2J construct was analyzed using DSC to determine the enthalpy of melting duplex-form DNA to single-stranded DNA and of melting junction-form DNA to single-stranded DNA. Results from DSC of I2J appear to corroborate the crystallographic and occupancy titration studies, as in solution, the I2J construct easily forms the junction complex (a predicted 100% X-isomer) and can be compared to the H2J construct (which can only form a junction stabilized by H-bond or 100% H-isomer), allowing for calculations of energy differences between the constructs. The I2J construct is ideal for the DSC studies as even at fairly low DNA concentrations (20-80μM), junction-form DNA was the dominant complex (versus duplex-form DNA), making the DSC analysis

simpler since the data can be fit with two explicit peaks for the two species (junction-form and duplex-form DNA). These results, in detail, will be published in the coming months as part of a paper authored by Megan Carter.

The Br2J construct was initially analyzed using DSC by Megan Carter. Results of the DSC of I2J look similar to her results of Br2J but have not been fully analyzed to date. The following illustrations from the Br2J DSC analysis were published in *Crystal Growth & Design* in September 2011 by Megan Carter and P. Shing Ho.

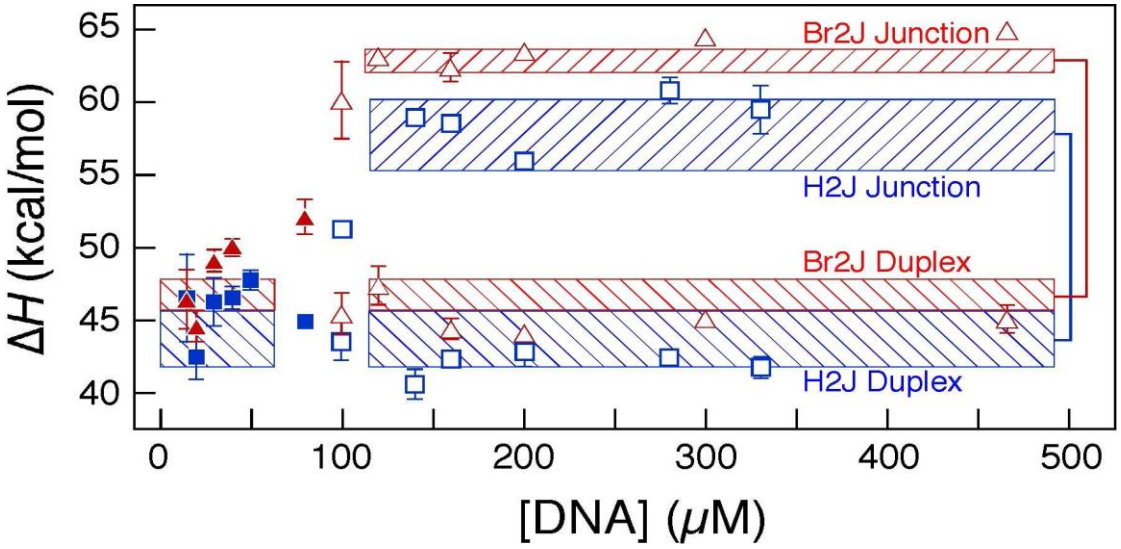


Figure 14. Enthalpies of melting (ΔHm) measured by differential scanning calorimetry (DSC) results for H2J (blue squares) and Br2J (red triangles) constructs at increasing DNA concentrations. The solid symbols indicate DSC data analyzed with a single component two-state model while open symbols indicate data analyzed by a two-component two-state model. Boxes represent data for the duplex (left hashed boxes) and junction (right hashed boxes) used to calculate the averages and standard deviations of ΔHm. The height of each box represents the standard deviation for all the data included in the average.

Table 5. Thermodynamic parameters determined by DSC for the melting of the Br2J and H2J DNA constructs as duplexes or four-stranded junctions to single-stranded DNA. The entropy of melting (ΔSm) is calculated from the melting enthalpy (ΔHm) and melting temperature (Tm) as ΔSm = ΔHm/ Tm, assuming the system is equally populated by the melted and unmelted forms at the Tm. Errors are listed as the standard deviations of the mean values.

Construct	T_m (°C)	ΔH_m (kcal/mol)	ΔS_m (cal/mol·K)
Br2J			
Junction	58.3 ± 0.3	63.4 ± 0.3	191 ± 2
Duplex	52.6 ± 0.3	47.0 ± 0.8	140 ± 50
Junction - Duplex	5.8 ± 0.4	16.4 ± 0.8	46 ± 3
H2J			
Junction	56.4 ± 0.2	58.6 ± 0.5	178 ± 2
Duplex	50.6 ± 0.2	44.0 ± 0.6	136 ± 2
Junction - Duplex	5.7 ± 0.3	14.6 ± 0.7	42 ± 2

Table 6. Thermodynamic parameters for the difference in energy between the junction and duplex forms of the Br2J (Br2J(J-D)) and the H2J (Br2J(J-D)) constructs at the reference temperature (50.6 °C), and the subsequent X- vs H-bond interactions at the reference and room temperatures. The differences in enthalpies (ΔΔH) and entropies (ΔΔS) were extrapolated from the DSC determined values to each temperature using the heat capacities of each form. The free energy differences at each temperature (ΔΔG) were calculated by the standard relationship ΔΔG = ΔΔH – TΔΔS.

Construct	T (°C)	$\Delta\Delta H$ (kcal/mol)	$\Delta\Delta S$ (cal/molK)	$\Delta\Delta G$ (kcal/mol)
$Br2J_{(J\text{-}D)}$ – $H2J_{(J\text{-}D)}$	50.6	-2 ± 1	-4.6 ± 3.7	0 ± 1
X-Bond – H-Bond	50.6	-3.6 ± 1	-9.2 ± 3.6	0 ± 1
X-Bond – H-Bond	25	-5.4 ± 1.1	-7.9 ± 3.6	-3 ± 1

Conclusions:

The research discussed in this thesis was designed to study the structure-energy relationship of chlorine halogen bonds by determining the structure of a chlorinated DNA junction (Cl2J), analyzing the structure (particularly the chlorine of the chlorinated uracil to phosphate oxygen atoms of the junction cross-over) and analyzing the ratio of the chlorine atom seen at the center of the junction (X-bond stabilized X-isomer) relative to the chlorine atoms on the outside strand (H-bond stabilized H-isomer) of the junction. The bond between the chlorine and its backbone sugar was not expected to stretch. The phosphate of residue 6 should orient to accommodate this interaction, stabilizing the junction. Indeed, this was observed. When refinement was completed, the chlorine to phosphate oxygen interaction length was 2.62 Å (and 3.73 Å to the second phosphate oxygen), which is 0.65 Å (or 20% closer) than the sum of the van der Waals radii of chlorine and oxygen (3.27 Å). The bond connecting the chlorine to the C5 carbon of the uracil base remained fixed at 1.71 to 1.72 Å, throughout refinement. The angle (C5-Cl^{ooo}O) for the X-bond interaction was 160.81°.

The analysis of the occupancy titration showed an X- to H-isomer ratio of ~80% to 20%, indicating a -0.8 kcal/mol difference in energy between the hydrogen bond and the chlorine bond, favoring the latter. The final electron density maps support this general conclusion.

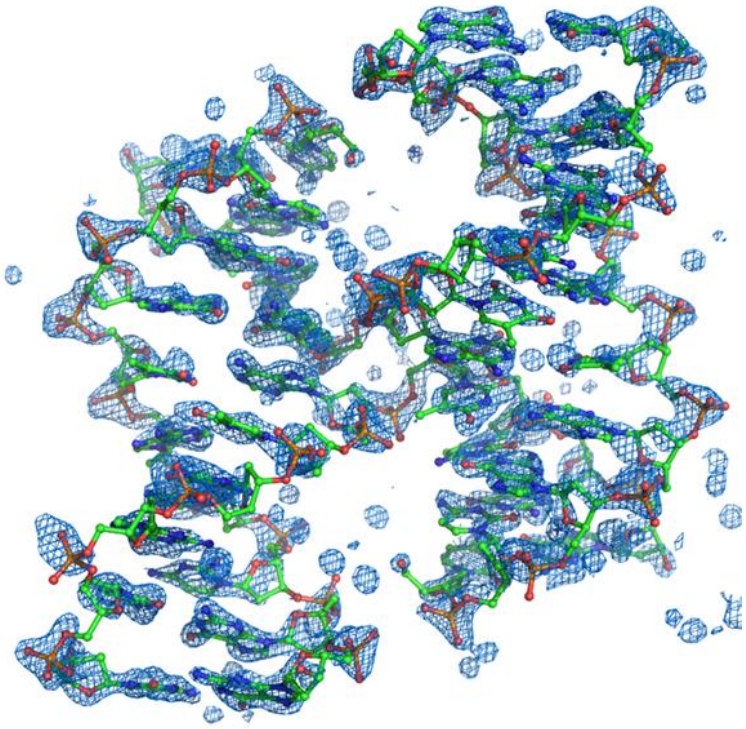


Figure 15. A visual created with Pymol of the final refined Cl2J structure. The electron density map was created using the 2Fo-Fc map at 2.0 sigma with a carve of 2.0 Å.

Figures 16-19 show specific residues of interest at the end of the refinement process and their correlating electronic density maps. These maps calculated with the X-isomer atoms (the chlorine atom at the junction center) set to 0.5 or 50% occupancy, corroborate the numerical results of the Cl2J occupancy titration. In figure 16, the interaction between the phosphate atom of adenine 6 (on the left) and the chlorine atom of chlorinated uracil 7, circled in red (in the right), can be seen. The interaction distance observed here is 2.62 Å. Near the chlorine atom, a small amount of electron density (blue in color) is present in the 2Fo-Fc map (2.0 sigma contours with a carve equal to 2.0 Å). There is also a large region of positive difference density (green in color) between the two phosphate oxygen atoms of adenine 6 and the chlorine. There is some negative difference density (red in color) appearing as a “channel” encasing the chlorine atom but

does not impinge on the chlorine itself. Thus, the maps indicate the presence of the chlorine a majority of the time at the junction center, consistent with a majority X-isomer form in the crystal.

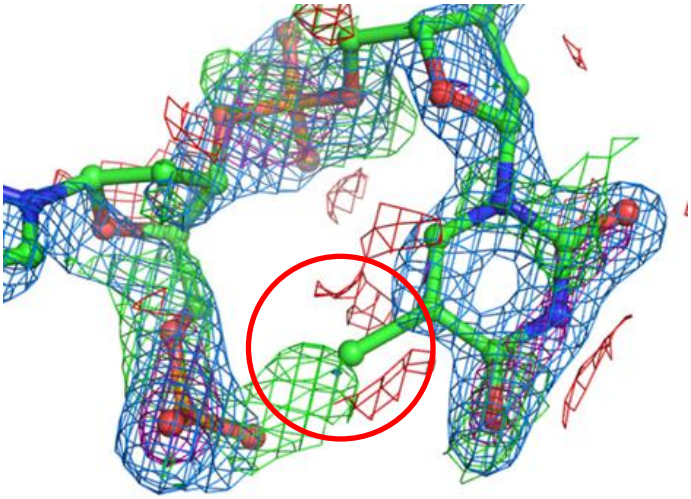


Figure 16. A visual created with Pymol showing the Cl2J construct with the phosphate group of adenine 6 on the left and the chlorine atom of chlorinated uracil 7, circled in red, on the right. All atoms are set to 0.50 occupancy. The blue and purple colors indicate electron density maps created from the 2Fo-Fc map with carve equal to 2.0 Å. The blue map is at 2.0 sigma and the purple map is at 4.0 sigma. The green difference density was created from a Fo-Fc map at 1.7 sigma with carve equal to 1.5 Å. The red difference density was also created from a Fo-Fc map at -1.7 sigma with carve equal to 1.3 Å.

Figure 17 exhibits cytosine 17 (on the left) and guanine 4 (on the right), also from the final mod of refinement. The cytosine residue is well-covered with electron density. There is, however, a subtle indication of positive difference density (green in color) emerging away from the C5 atom of the ring. In the H-isomer form of this construct, the chlorine would present at this outside-strand position. Guanine 4 is mostly covered with electronic density, mainly lacking coverage on its N2 atom, which is circled in red. The fact that there is not also negative difference density (red) covering the N2 atom is an indication that this residue is an adenine a portion of the time. Although, noting that all

atoms remain held at 50% occupancy, the lack of negative electron difference density also supports the presence of guanine at this residue.

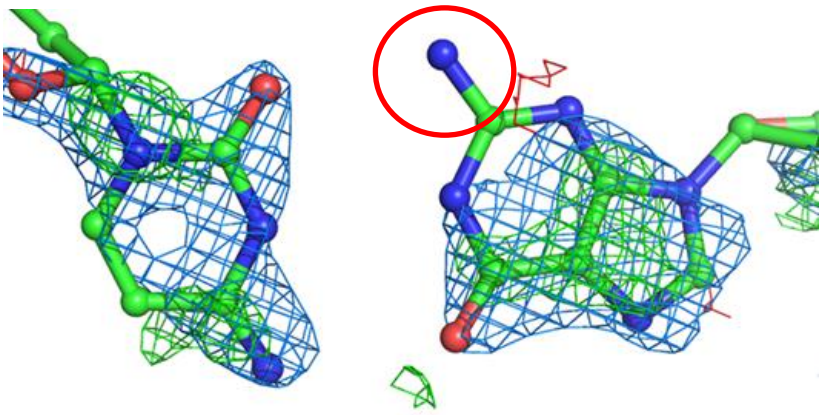


Figure 17. A visual created with the Pymol program showing the Cl2J construct with cytosine 17 on the left and guanine 4 on the right. The N2 atom of guanine 4 is circled in red. All atoms are set to 0.50 occupancy. The blue electron density map was created using the 2Fo-Fc map at 2.0 sigma and carve equal to 2.0 Å. The green difference density was created using the Fo-Fc map at 1.7 sigma and carve equal to 1.5 Å. The red difference density was also created from the Fo-Fc map at -1.7 sigma and carve equal to 1.2 Å.

Figure 18 exhibits cytosine 27 (on the left) and guanine 14 (on the right). The N2 atom of guanine 14 is circled in red. Cytosine 27 and guanine 14 are both well-covered with electron density and the striking observation to note is the large mass of negative difference density on the N2 atom of the guanine, which is strongly indicative of this residue actually being an adenine most of the time. An adenine at residue 14 would pair with the chlorinated uracil at residue 7, supporting an isomer ratio for the construct leaning heavily in favor of the X-isomer form (where the chlorine is stabilizing the junction).

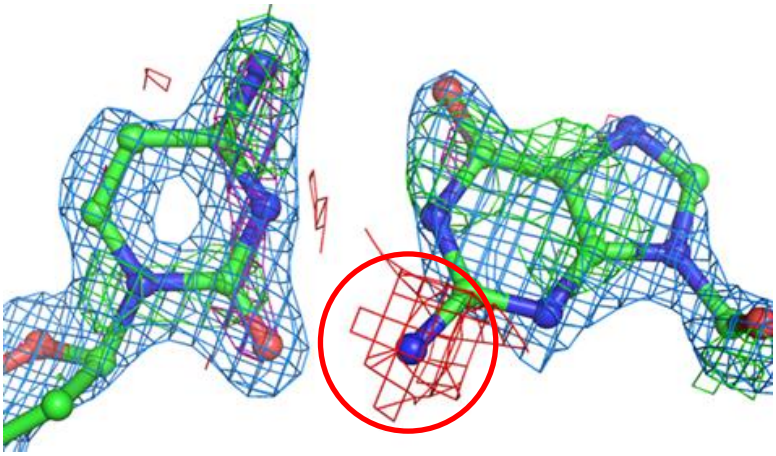


Figure 18. A visual created using the Pymol program showing the Cl2J construct with cytosine 27 on the left and guanine 14 on the right. The N2 atom of guanine 14 is circled in red. All atoms are set to 0.50 occupancy. The blue and purple densities were created from the 2Fo-Fc map with a carve of 2.0 Å. The blue map is at 2.0 sigma and the purple map is at 4.0 sigma. The green difference density was created from the Fo-Fc map at 1.7 sigma and carve equal to 1.5 Å. The red difference density was created from the Fo-Fc map at -1.7 sigma and carve equal to 1.3 Å.

Figure 19 shows chlorinated uracil 37 (on the left) and adenine 24 (on the right). The chlorine atom of chlorinated uracil 37 is circled in red. The chlorine does not exhibit any negative electron difference density (red), while keeping in mind that the atom's occupancy is held at 50%. A small projection of positive difference density (green) also stretches out towards the chlorine, both indicating that the chlorine atom is present at the outside position some percentage of the time. The calculated electronic density for adenine 24 does not match the placement of the residue very well, so it is challenging and error-prone to draw any conclusions from this particular residue.

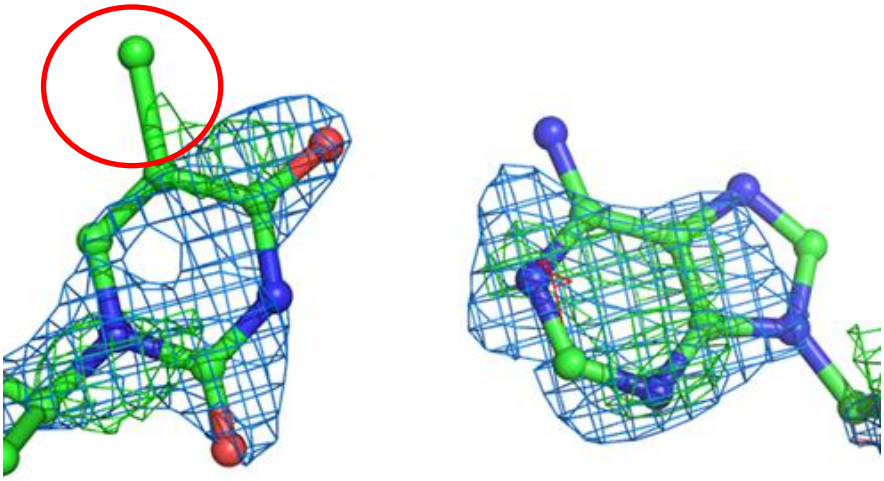


Figure 19. A visual created using the Pymol program showing the Cl2J construct with chlorinated uracil 37 on the left and adenine 24 on the right. The chlorine atom of chlorinated uracil 37 is circled in red. All atom occupancies are set to 0.50. The blue electron density was created from the 2Fo-Fc map at 2.0 sigma and carve equal to 2.0 Å. The green difference density was created from the Fo-Fc map at 1.7 sigma and carve equal to 1.5 Å. The red difference density was created from the Fo-Fc map at -1.7 sigma and carve equal to 1.1 Å.

In summary, it was crucial to make many adjustments in refining method for the Cl2J construct due to the isomer ratio being so near to 50/50. The model was easily biased towards one of the two isomer forms. Important corrections made include the transition from a 2-stranded to a 4-stranded model, a change in all atom occupancies from 1.0 to 0.50 and the fixing of several residues and atoms in the minimize file (residues 1, 10, 11, 20, 21, 30, 31 and 40, and the phosphate atoms of residues 5, 25 and 36). Refinement and occupancy titration programs on the Cl2J construct successfully demonstrated that a halogen bond interaction was indeed occurring in the model and that the structure was present in an approximate 75% X-isomer ratio. This indicates a slight energy difference between stabilizing energy of the chlorine halogen bond in Cl2J in comparison with the hydrogen bond. The group occupancy titration method was also

selected as the most accurate and was used to approximate X-isomer percentage in other constructs, such as F2J, Br2J and I2J. This information can be used to compare energies and strengths of all types of halogen bonds which can eventually be integrated into molecular simulations and programs for drug design.

References:

Arunan E., Desiraju R., Klein R.A., Sadlej J., Scheiner S., Alkorta I., Clary D.C., Crabtree R.H., Dannenberg J.J., Hobza P., Kjaergaard H.G., Legon A.C., Menucci B., Nesbitt D.J., *Pure Appl. Chem., in Press.*

Auffinger P., Hays F.A., Westhof E., Ho P.S. **(2004)**. Halogen bonds in biological molecules. *PNAS, 2004, 101, 16789-16794.*

Brunger A.T. **(1992).** Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature 355 (6359): 472–475.*

Carter M., Ho P.S. **(2011)**. Assaying the energies of biological halogen bonds. *Cryst. Growth Des., 2011, DOI: 10.1021/cg200991v.*

Fourmigue M. **(2009)**. Halogen Bonding: Recent Advances. *Current Opinion in Solid State and Materials Science, 2009, 13, 36-45.*

Hays, F.A., J.M. Vargason and P.S. Ho **(2003)**. Effect of sequence on the conformation of DNA holliday junctions. *Biochemistry, 42 (32):9586-9597.*

Hays, F.A., J. Watson and P.S. Ho **(2003)**. Caution! DNA crossing: crystal structures of Holliday junctions. *J Biol Chem, 278 (50): 49663-6.*

Hays, Franklin A., Amy Teegarden, Zebulon J.R. Jones, Michael Harms, Dustin Raup, Jeffrey Watson, Emily Cavaliere, P. Shing Ho **(2005)**. How sequence defines structure: A crystallographic map of DNA structure and conformation. *Proceedings of the National Academy of Sciences, 2005, 102, 7157-7162.*

http://www.msg.ucsf.edu/local/programs/epmr/epmr.html, copyright © 2005 -2008 by Charles R. Kissinger, last updated April 1, 2008, accessed October 27, 2011.

Legon A. **(2010)**. The Halogen Bond: an interim perspective. *Physical Chemistry Chemical Physics, 2010, 12, 7736-7747.*

Li Q., Wang Y., Liu Z., Li W., Cheng J., Gong B., Sun J. **(2009)**. An Unconventional halogen bond with carbene as an electron donor: An Ab initio study. *Chemical Physics Letters, 2009, 469, 48-51.*

Lu Y., Shi T., Wang Y., Yang H., Yan X., Luo X., Jiang H., Zhu W. **(2009)**. Halogen Bonding-A Novel Interaction for Rational Drug Design? *Journal of Medicinal Chemistry, 2009, 52, 2854-2862.*

Lu Y., Wang Y., Zhu W. **(2010)**. Nonbonding interactions of organic halogens in biological systems: implications for drug discovery and biomolecular design. *Physical Chemistry Chemical Physics, 2010, 12, 4543-4551.*

Lyday, Phyllis A. **(2005)**. Iodine and Iodine Compounds. *Ullmann's Encyclopedia of Industrial Chemistry, 2005, Vol. A14, 382–90. ISBN 9783527306732.*

Metrangolo P., Resnati G. **(2008)**. Chemistry - Halogen versus Hydrogen. *Science, 2008, 321, 918-919.*

Politzer P., Murray J.S., Clark T. **(2010)**. Halogen Bonding: en electrostatically-driven highly directional noncovalent interaction. *Physical Chemistry Chemical Physics, 2010, 7748-7757.*

Sarwar M., Dragisic B., Salsberg L.J., Gouliaras C., Taylor M.S. **(2010)**. Thermodynamics of Halogen Bonding in Solution: Substituent, Structural, and Solvent Effects. *Journal of the American Chemical Society, 2010, 132, 1646-1653.*

Shirman T., Lamere J., Shimon L., Gupta T., Martin J., Van der Boom M.E. **(2008)**. Halogen-bonded supramolecular assemblie based on phenylethynyl pyridine derivatives: Driving crystal packing through systematic chemical modifications. *Crystal Growth and Design, 2008, 8, 3066-3072.*

Valadares N., Salum L.B., Polikarpov I., Andricopulo A.D., Garratt R.C. **(2009)**. Role of Halogen Bonds in Thyroid Hormone Receptor Selectivity: Pharmacophore-based 3D-QSSR studies. *Journal of Chemical Information and Modeling, 2009, 49, 2606-2616.*

van Holde, Kensal E., W. Curtis Johnson, and P. Shing Ho. *Principles of Physical Biochemistry.* 2. Upper Saddle River, NJ: Pearson Education, Inc., 2006. Print.

Voth A.R., Hays F.A., Ho P.S. **(2007)**. Directing macromolecular conformation through halogen bonds. *PNAS, 2007, 104, 6188-6193.*

Voth A.R., Ho P.S. **(2007)**.The Role of Halogen Bonding in Inhibitor Recognition and Binding by Protein Kinases. *Current Topics in Medicinal Chemistry, 2007, 7, 1336-1348.*

Voth A.R., Khuu P., Oishi K., Ho P.S. **(2009)**. Halogen Bonds as Orthogonal Interactions to Hydrogen Bonds. *Nature Chemistry, 2009, 1, 74-79.*

Zhou P., Lv J., Zou J., Tian F., Shang Z. **(2010)**. Halogen-water-hydrogen bridges in biomolecules. *Journal of Structural Biology, 2010, 169, 172-182.*

Supplementary Material:

Initial occupancy titration method

In the initial occupancy titration method, all atom occupancies in the model are titrated up or down, depending upon whether their presence in that location is indicative of the X-isomer or H-isomer form of the construct. R- and R-free values are recorded as the occupancies of all atoms are changed. The X-isomer percentage is then plotted against the change in R- or R-free, with the baseline first subtracted. Polynomial curves are fit to the data points and the minimum indicates the X-isomer percentage of the construct. Many trials of occupancy titrations are performed, with X-isomer and H-isomer atom occupancies starting at 0.30, 0.50 and 0.70 to ensure that any results are consistent and not a by-product of where the atom's occupancy begins.

Occupancy titration refinement

The occupancy titration refinement or qindividual program essentially takes the final model and the specified atoms and conducts a refinement titration on those atom's occupancies. This program yields occupancy values for each atom inputted directly from the given model's information. It is crucial to note that the program should be applied to both, the DNA strands refined as the X-isomer and the DNA strands refined as the H-isomer in the model to derive unbiased results. The outputted values are then analyzed after normalization and averaging the percentages supporting each, the X-isomer and H-isomer forms of the construct.