



Article

# On the Importance of Halogen and Chalcogen Bonds in the Solid State of Nucleic Acids: A Combined Crystallographic and Theoretical Perspective

María de las Nieves Piña  and Antonio Bauzá \* 

Departament de Química, Universitat de les Illes Balears, Ctra. de Valldemossa Km 7.5,  
07122 Palma de Mallorca, Balears, Spain; neus.pinya@uib.es

\* Correspondence: antonio.bauza@uib.es

**Abstract:** In this work, intra- and intermolecular halogen and chalcogen bonds (HlgBs and ChBs, respectively) present in the solid state of nucleic acids (NAs) have been studied at the RI-MP2/def2-TZVP level of theory. To achieve this, a Protein Data Bank (PDB) survey was carried out, revealing a series of structures in which Br/I or S/Se/Te atoms belonging to nucleobases or pentose rings were involved in noncovalent interactions (NCIs) with electron-rich species. The energetics and directionality of these NCIs were rationalized through a computational study, which included the use of Molecular Electrostatic Potential (MEP) surfaces, the Quantum Theory of Atoms in Molecules (QTAIM), and Non Covalent Interaction plot (NCIplot) and Natural Bonding Orbital (NBO) techniques.

**Keywords:** nucleic acids; PDB survey; solid state chemistry; halogen bonding; chalcogen bonding; theoretical study



**Citation:** Piña, M.d.l.N.; Bauzá, A. On the Importance of Halogen and Chalcogen Bonds in the Solid State of Nucleic Acids: A Combined Crystallographic and Theoretical Perspective. *Int. J. Mol. Sci.* **2023**, *24*, 13035. <https://doi.org/10.3390/ijms241713035>

Academic Editor: Dongho Kim

Received: 29 July 2023

Revised: 16 August 2023

Accepted: 20 August 2023

Published: 22 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Noncovalent forces are of pivotal significance in biology and compose the fundamentals of modern chemistry [1–4]. This has led to an increase in the number noncovalent interactions (NCIs) characterized to date, which exert a crucial role in several biological mechanisms, such as enzyme inhibition [5], ion transport [6], and protein folding [7]. In a parallel way to the well-known hydrogen bonds (HBs) [8], halogen bonds (HlgBs) are nowadays a well-established and well-studied NCI among supramolecular chemists [9–15]. Both binding forces have been recently described as members of the ‘ $\sigma$ -hole family’ [16–18]. A  $\sigma$ -hole was originally conceived as a region of positive electrostatic potential located on the extension of a covalent bond. For instance,  $\sigma$ -holes can be found along the Se–F/C–Br bonds in difluoro selenide or bromoethane molecules. In the case of halogens, the  $\sigma$ -hole donor ability increases from F to I, resulting in a strengthening of the interaction upon descending in the group.

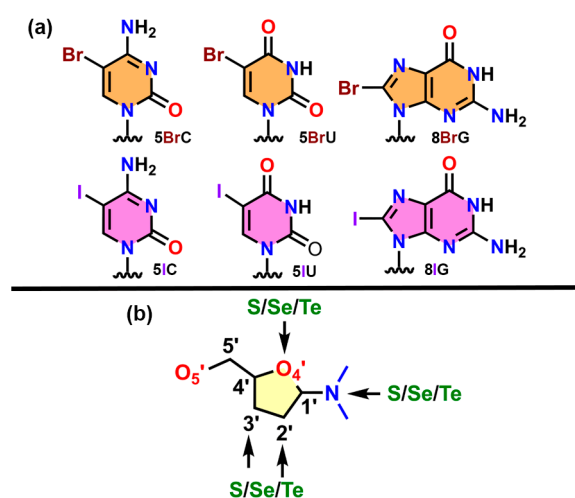
In biological systems, halogen atoms are typically used in the field of drug design and to probe molecular interactions [19,20], while, on the other hand, halogenated bases are useful molecular synthons for solving the X-ray crystal structures of nucleic acids (NAs) using the Multiwavelength Anomalous Dispersion (MAD) technique [21]. In this regard, several examples exist in the literature where HlgBs are functional and relevant binding forces, such as works by Carter [22] and Voth [23] and collaborators in which HlgBs were engineered for studying DNA Holliday junctions. In addition, the theoretical reports from Parker [24] and Xu [25] and collaborators unraveled the potential role of HlgBs as a relevant interaction in both NA structures and base pair stabilization.

Furthermore, Ennifar and coworkers [26] utilized a variety of experimental techniques (e.g., steady-state fluorescence, UV thermal melting, X-ray crystallography, and gel electrophoresis) to understand the influence of halogenation (by incorporating either a

bromine or iodine) on RNA folding. They demonstrated that the RNA hairpin/duplex ratio was drastically influenced by both the presence and position of halogenation along the RNA chain. Moreover, Kolář [27] and Auffinger [28] and collaborators have reported statistical evidence of the promising potential of HlgBs involving NAs. Our group has also contributed to this field with three computational studies, involving intermolecular protein-NA HlgBs [29], quantification of intramolecular halogenated nucleotide...phosphate HlgBs [30], as well as evaluating the effect of NA halogenation on the stability of canonical and non-canonical base pairs [31].

On the other hand, the use of chalcogen compounds in biology (mainly Se and Te) has been related to the field of rational drug design and, more specifically, to the development of noncancer therapeutic agents [32]. Despite this, studies have also demonstrated the implications of chalcogen bonding interactions (ChBs) [33,34] in several biological phenomena, such as the stabilization of protein structures [35–37] or the formation of RNA–ligand and protein–ligand complexes [38–40]. In addition, Se and Te have been also incorporated into the structures of NAs for phase and structure determination in nucleic acid X-ray crystallography [41–43], similar to the replacement of methionines by selenomethionines in proteins. In this context, a recent study carried out by Sharma and collaborators [44] proposed the use of ChBs as an additional source of stabilization in NA base pairs.

Building upon these previous reports, our aim was to investigate the plausible structural role of HlgBs and ChBs in the solid-state architecture of isolated NAs. To reach this goal, we conducted a search in the Protein Data Bank (PDB) [45], retrieving a total of 11 nucleic acid structures exhibiting highly directional HlgBs and ChBs in their solid-state architecture (see Figure 1 for the halogenated and chalcogenated molecules used). These contacts corresponded to either an intramolecular (e.g., within the same NA strand) or intermolecular (involving two neighboring NA chains) interaction and assisted in crystal packing formation. Calculations at the RI-MP2/def2-TZVP level of theory were conducted to shed light into the strength, directionality, and physical nature of the HlgBs and ChBs present in these structures by means of the Molecular Electrostatic Potential (MEP), the Quantum Theory of Atoms in Molecules (QTAIM), and NCiplot (Non Covalent Interaction plot) and Natural Bonding Orbital (NBO) techniques. We believe that the results derived from this study might be important to those scientists working in the fields of NA or protein engineering and  $\sigma$ -hole-based chemistry, as well as in the development of NA-based materials.

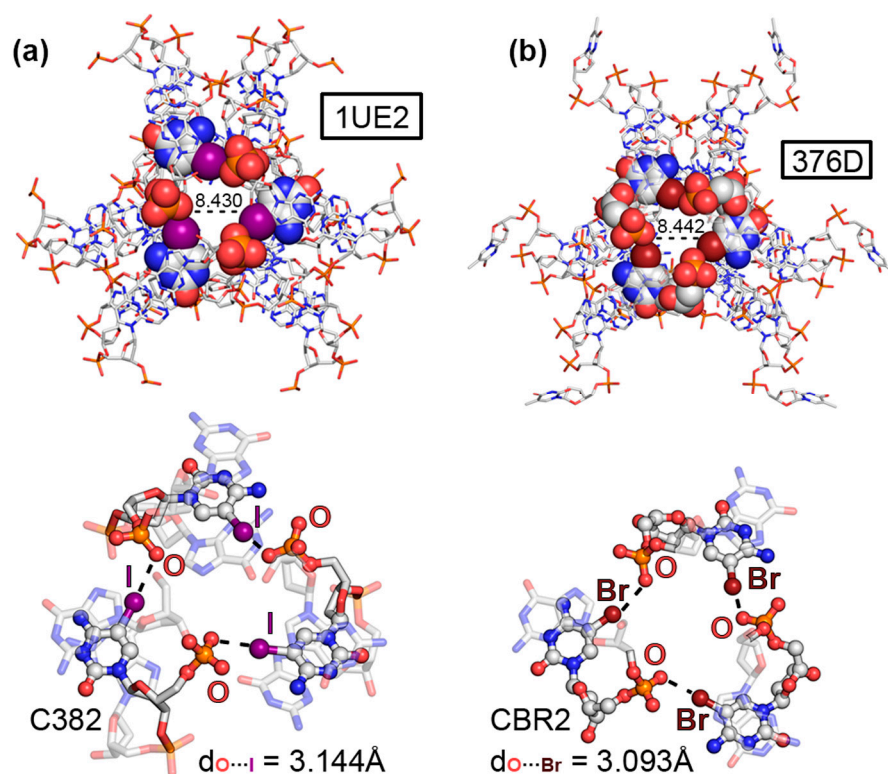


**Figure 1.** (a) Schematic representation of the halogenated nucleobases considered in the PDB survey (5BrC, 5IC, 5BrU, 5BrI, 8BrG, and 8IG represent 5-bromocytosine, 5-iodocytosine, 5-bromouracil, 5-iodouracil, 8-bromoguanine, and 8-iodoguanine, respectively); (b) Schematic representation of the nucleotide positions considered for the incorporation of S, Se, or Te atoms.

## 2. Results and Discussion

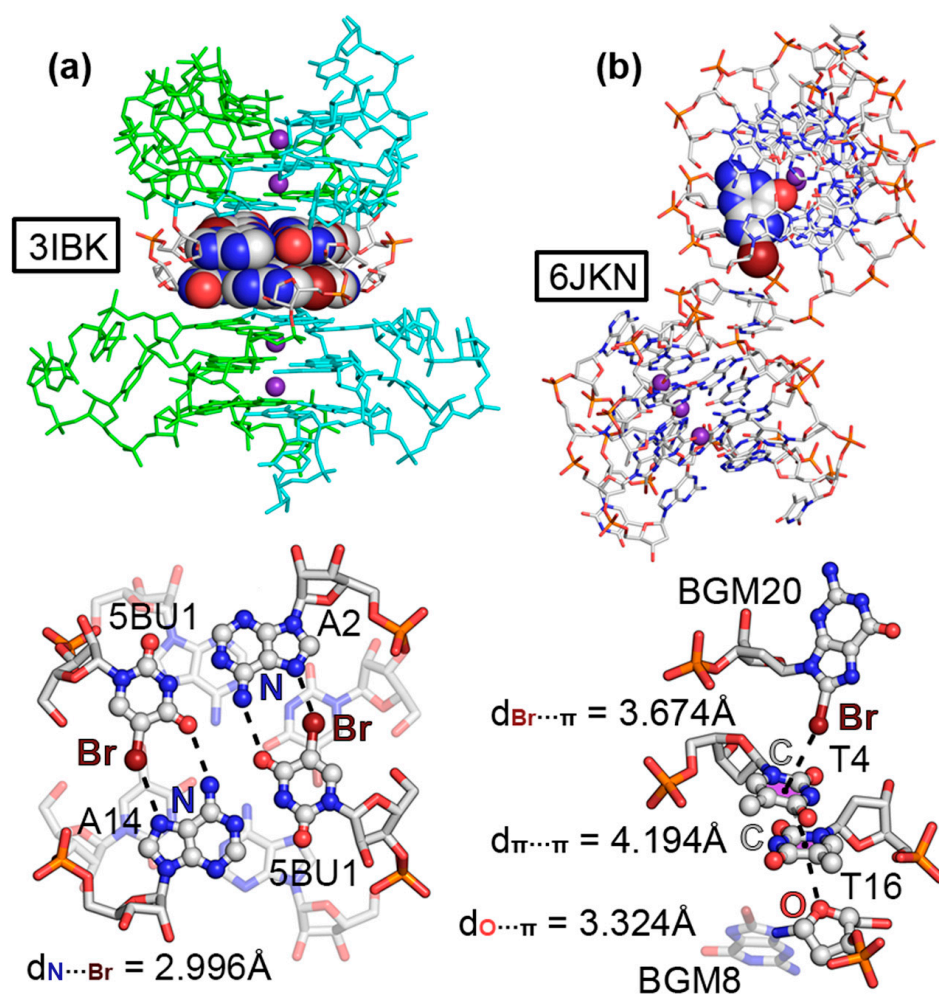
### 2.1. Results from the PDB Survey: Br and I Halogen Bonds

Four examples were retrieved from the PDB inspection involving halogenated nucleobases. The first two corresponded to structures 1UE2 [46] and 376D [47], involving two DNA fragments containing the sequences d(G<sup>I</sup>CGAAAGCT) and d(G<sup>Br</sup>CGAAAGCT), respectively. Studies have demonstrated that similar sequences exhibit unconventional properties such as (i) abnormal mobility in electrophoresis [48], (ii) high thermostability [49], (iii) uncommon circular dichroism spectra [49], as well as (iv) unusual resistance against nuclease digestion [50]. In both cases, 5IC and 5BrC were used during the X-ray diffraction process by means of the MAD method [21]. A close inspection of both solid-state architectures revealed the formation of DNA helical duplexes in parallel disposition to the z-axis of the crystal lattice. This style of packing has been well studied in A and B NAs [51,52], and it is considered an important stabilizing effect in the crystal formation between two aligned symmetry-related helices. The DNA duplexes were associated in the form of a trimer, forming a molecular cavity (around 8.5 Å) in the middle of the structure (see Figure 2a,b). Interestingly, three symmetrically distributed HlgBs were established at the top and bottom parts of the assembly. These involved either iodine (in 1UE2) or bromine (in 376D) atoms acting as  $\sigma$ -hole donors and O atoms from vicinal phosphate groups as  $\sigma$ -hole acceptors, exhibiting O...I and O...Br distances shorter than the sum of the O/I and O/Br van der Waals radii (3.144 Å and 3.093 Å, respectively) and high directionality (O...I-C and O...Br-C angles of 168.3° and 162.2°, respectively). Theoretical calculations at the RI-MP2/def2-TZVP level of theory on a dimeric model revealed individual HlgB strengths of −16.6 and −12.8 kcal/mol, respectively, in line with the results obtained for charge-assisted HlgBs [53]. The presence of these three centered HlgBs was important for holding the DNA duplexes together, thus being a nice example of the potential application of HlgBs as a stabilizing agent in NA crystal structures.



**Figure 2.** Partial views of the X-ray structures of (a) 1UE2 and (b) 376D with the molecular cavity width indicated (in Å). The HlgBs are magnified in the bottom part of the figure with the O...I and O...Br distances indicated.

The third example encompassed structure 3IBK, which corresponded to a telomeric RNA G-quadruplex with the sequence ( $U_{Br}AGGGUUAGGGU$ ) [54]. G-quadruplexes are part of the noncanonical NA family of structures and are formed by the stacking of G-quartets, which are planar arrangements of four guanines held together by several HBs [55,56]. In the crystal packing of 3IBK, the asymmetric unit was composed by two RNA strands (colored green and ice blue in Figure 3a, respectively), each of which contained two G-rich repeats, folding into a parallel four-stranded bimolecular G-quadruplex. The structure also contained  $K^+$  ions in the center of the channel, positioned between the three G-quartets. Interestingly, 5BrU was incorporated into the RNA sequence, being crucial for the enhancement of crystal diffraction quality by holding the dimeric interface tightly together, as stated by the original authors. A close inspection of the structure revealed the formation of a mixed HB/HlgB base pair, involving the Br and O atoms belonging to 5BrU and the  $N_7$  and  $-NH_2$  group from adenine, located at the interface between two G-quadruplexes (see Figure 3a bottom for a detailed view of the interaction). The computed interaction energy of this noncanonical base pair reported a strength of  $-6.0$  kcal/mol (in line with our recent study [31]) and a HlgB energetic contribution of  $-2.2$  kcal/mol. Owing to the short  $N\cdots Br$  distance ( $2.996$  Å) and the directionality observed ( $157.2^\circ$ ), this example illustrates the potential role of halogenated base pairs in the stabilization of noncanonical nucleic acid structures.

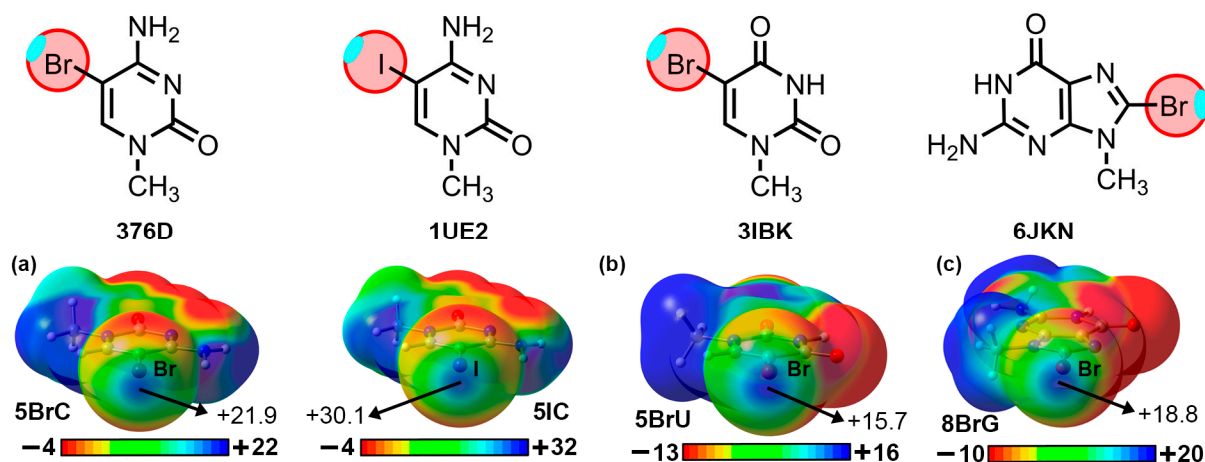


**Figure 3.** Partial views of the X-ray structures of (a) 3IBK and (b) 6JKN. In 3IBK, the two RNA chains are colored green and ice blue for easier identification.  $K^+$  ions are colored as purple spheres. The HlgBs are magnified in the bottom part of the figure with the  $N\cdots Br$  and  $\pi\cdots Br$  distances indicated.



The last selected structure corresponded to a bromo G-quadruplex of human telomeric DNA (6JKN) [57]. In this example, the crystal packing was composed of discrete G-quadruplex molecules, with  $K^+$  ions inserted in the center of each unit in a similar fashion to that observed in the 3IBK structure. These isolated units interacted with each other to form a supramolecular assembly. In Figure 3b, a G-quadruplex dimer is highlighted, in which the Br atom belonging to 8BrG from one unit (BGM20) interacted through a HlgB with the  $\pi$ -system of a thymine (T4) belonging to another G-quadruplex, exhibiting a  $C\cdots Br$  distance of 3.674 Å. This type of HlgB has been also described in the literature as a halogen– $\pi$  interaction [58]. In addition, a second thymine ring (T16) interacted with T4 through a  $\pi$ – $\pi$  stacking interaction ( $\pi$ – $\pi$  distance of 4.194 Å) and, lastly, a lone pair– $\pi$  (lp– $\pi$ ) interaction was undertaken between the  $sp^3$  O from a sugar moiety belonging to a guanine base (BGM8) and the  $\pi$ -system of T16, with an intermolecular distance of 3.324 Å. Computations on a dimeric model consisting of BGM20 and T4 revealed a HlgB strength of  $-2.2$  kcal/mol. The formation of this supramolecular assembly contributed to the stabilization of the G-quadruplexes in the solid state, serving as an example of the cooperation between different NCIs in the formation of solid-state NA structures.

To demonstrate the implication of the Br and I  $\sigma$ -holes in the HlgBs described above, we have computed the Molecular Electrostatic Potential (MEP) surfaces of 5BrC, 5IC, 5BrU, and 8BrG at the RI-MP2/def2-TZVP level of theory (see Figure 4). As noted, in all cases, a positive potential surface was observed along the prolongation of either the C–Br or C–I covalent bonds ( $\sigma$ -hole), with values ranging between +15.7 and +30.1 kcal/mol. Also, in the case of 5BrC and 5IC, the iodinated derivative exhibited a more positive  $\sigma$ -hole potential, in line with the larger polarizability exhibited by the I atom, as it is commonly known [17]. The presence of these electrophilic regions made these biological synthons suitable for favorable interactions with electron-rich species (e.g., lone pair-bearing molecules or  $\pi$ -systems), such as the ones described in the structures shown above.



**Figure 4.** Molecular Electrostatic Potential (MEP) surfaces of (a) 5BrC and 5IC, (b) 5BrU, and (c) 8BrG. Energy values at concrete regions of the surface are given in kcal/mol (0.001 a.u.). The MEP minima and maxima have been adjusted for clear visualization of the Br and I  $\sigma$ -holes. The corresponding PDB codes are also included below the schematics of each molecule.

As a final remark, the BSSE-corrected interaction energies, distances, and angle values corresponding to the HlgBs present in the highlighted structures are gathered in Table 1.

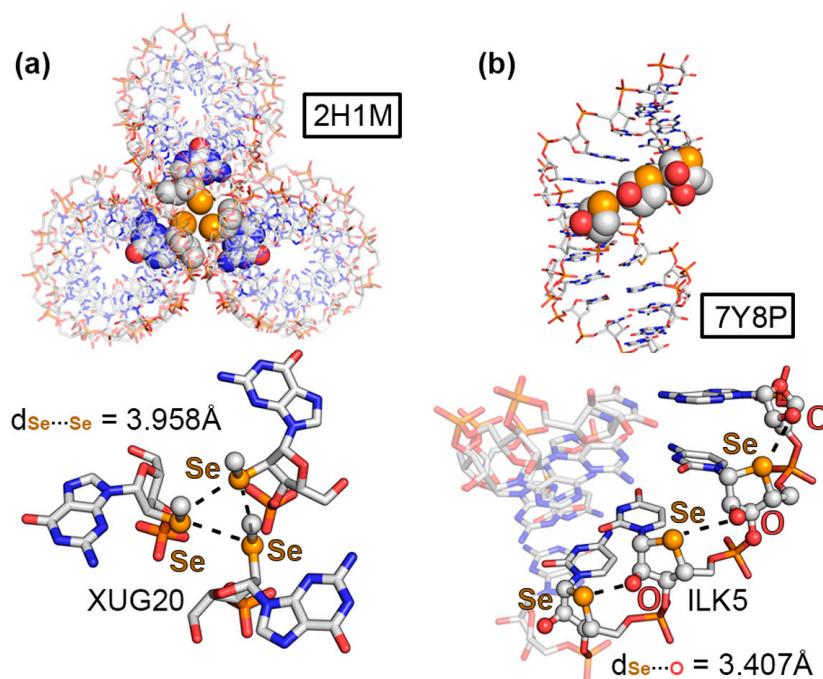
**Table 1.** BSSE-corrected interaction energies of the HlgBs present in the 376D, 3IBK, 6JKN, and 1UE2 structures ( $\Delta E_{\text{BSSE}}$ , in kcal/mol), distance ( $d$ , in Å), and angle values ( $\alpha$ , in °) at the RI-MP2/def2TZVP level of theory. In the case of 3IBK, the HlgB contribution to the total interaction energy is indicated in parentheses.

PDBID	$\Delta E_{\text{BSSE}}$	$d$	$\alpha$ <sup>1</sup>
376D (O...Br)	−12.8	3.093	162.2
3IBK (N...Br)	−6.0 (−2.2)	2.996	157.2
6JKN ( $\pi$ ...Br)	−2.2	3.674	174.4
1UE2 (O...I)	−16.6	3.144	168.3

<sup>1</sup> Measured as the A...Hlg–C angle (A = O, N, and C; Hlg = Br and I).

## 2.2. Results from the PDB Survey: S, Se, and Te Chalcogen Bonds

The first selected example corresponded to the 2H1M structure [59], which was composed of an RNA containing 2'-methylseleno guanosine. In more detail, the authors crystallized a 2'-methylseleno guanosine phosphoramidite by incorporating a methylseleno moiety in the 2' position of the pentose ring. The solid-state architecture of this compound revealed the formation of a helical RNA duplex, which established noncovalent contacts with two other vicinal units (see Figure 5a). Interestingly, the central region of the RNA helical triad was mainly stabilized by three “like-like” ChBs between the 2'-methylseleno moieties, exhibiting equivalent Se...Se distances (3.958 Å). Theoretical calculations revealed a strength of −3.7 kcal/mol per ChB. This led to the formation of an equilateral triangle, which held the three RNA molecules together, thus significantly contributing to the observed packing structure.

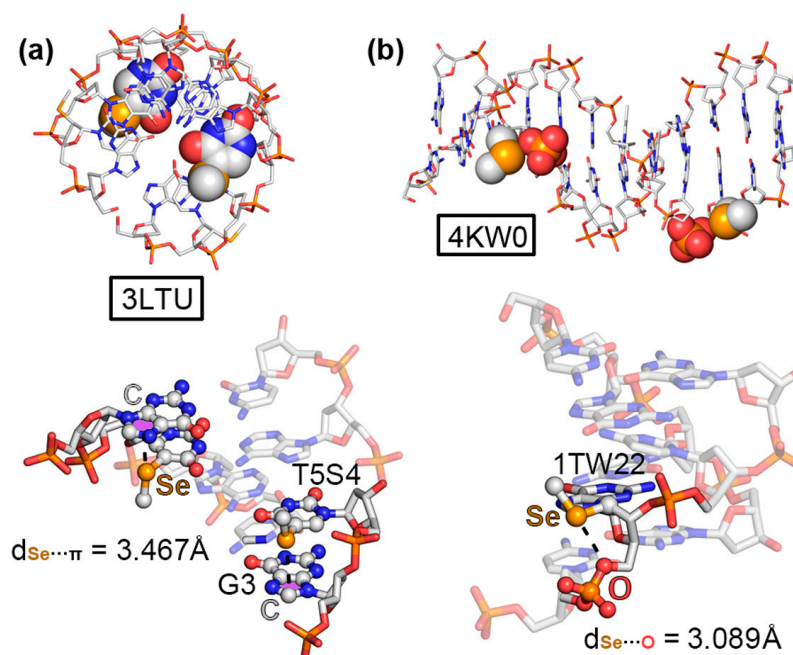


**Figure 5.** Partial views of the X-ray structures of (a) 2H1M and (b) 7Y8P. The ChBs are magnified in the bottom part of the figure with the Se...Se and O...Se distances indicated.

The second structure (7Y8P, Figure 5b) corresponded to a 4'-seleno-modified RNA [60], which crystallized in the form of a duplex, maintaining the A-conformation observed in natural RNA with almost all of the ribonucleosides exhibiting north-type sugar pucker. Interestingly, the 4'-seleno derivatives were involved in three consecutive O...Se ChBs involving the 3' OH group from a vicinal nucleotide within the same RNA strand; therefore, they acted as both  $\sigma$ -hole donor and acceptor counterparts. The computed interaction

strength of one of these intramolecular ChBs (exhibiting a O...Se distance of 3.407 Å) resulted in −3.1 kcal/mol.

As the two last ChB examples, structures 3LTU [61] and 4KW0 [62], involved a DNA octamer containing 5-methylseleno-deoxyuridine and a DNA dodecamer containing 2'-methylseleno-guanine, respectively. As can be observed in Figure 6a,b, the methylseleno group established a ChB with (i) the  $\pi$ -system of a guanine ring (G3 in Figure 6a) and (ii) the lone pairs of a negatively charged O atom from a vicinal phosphate group (Figure 6b), with  $\pi$ ...Se and O...Se distances of 3.467 and 3.089 Å, respectively. These interactions contributed to the stabilization of the DNA assemblies, with energetic strengths of −7.8 and −4.3 kcal/mol, serving as an alternative to classical HB and  $\pi$ - $\pi$  stacking interactions.



**Figure 6.** Partial views of the X-ray structures of (a) 3LTU and (b) 4KW0. The ChBs are magnified in the bottom part of the figure with the  $\pi$ ...Se and O...Se distances indicated.

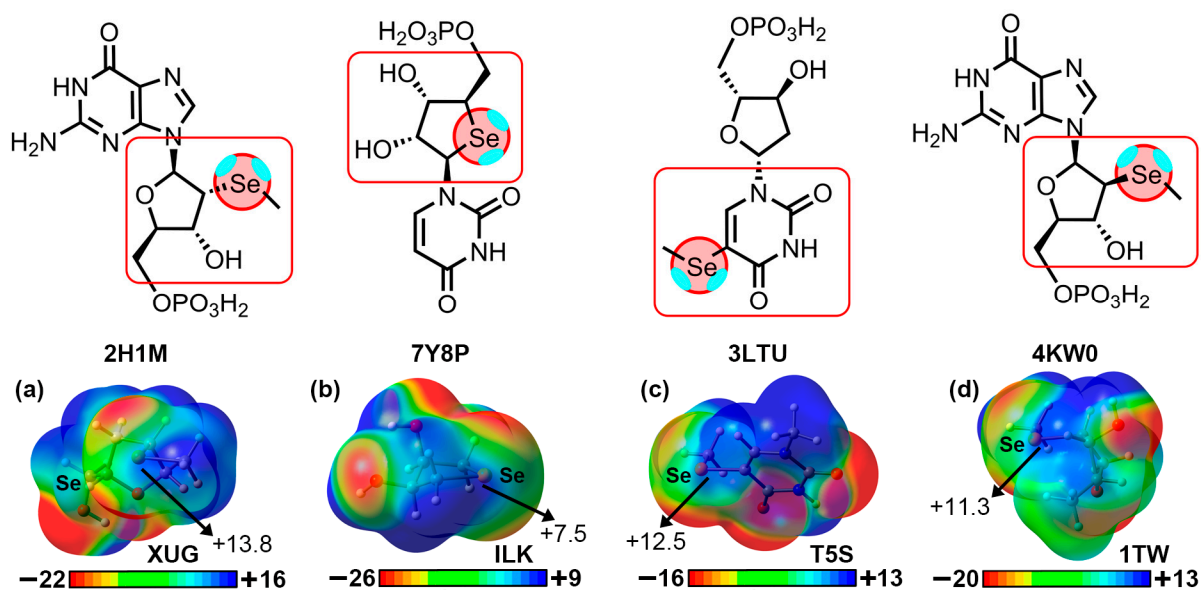
In Table 2, the interaction energies, distances, and angle values with respect to the four discussed ChB examples are shown. In addition, we have also included additional NA structures retrieved from the PDB search that also exhibited directional ChBs in their solid-state packing (see ESI), such as 3DW6 [63], showing a N...Se ChB; 3HG8 [64] exhibiting a S... $\pi$  ChB; and finally 3FA1 [65], which presented both Te... $\pi$  and O...Te ChBs.

**Table 2.** BSSE-corrected interaction energies of the ChBs present in the 2H1M, 7Y8P, 3LTU, 4KW0, 3DW6, 3HG8, and 3FA1 structures ( $\Delta E_{\text{BSSE}}$ , in kcal/mol), distance (d, in Å), and angle values ( $\alpha$ , in °) at the RI-MP2/def2-TZVP level of theory. In the case of 3LTU, 3HG8, and 3FA1, the ChB contribution to the total interaction energy is indicated in parentheses.

PDBID	$\Delta E_{\text{BSSE}}$	d	$\alpha$ <sup>1</sup>
2H1M (Se...Se)	−3.7	3.958	170.4
7Y8P (O...Se)	−3.1	3.407	148.4
3LTU ( $\pi$ ...Se)	−7.8 (−2.0)	3.467	168.4
4KW0 (O...Se)	−4.3	3.089	172.6
3DW6 (N...Se)	−3.0	3.736	166.8
3HG8 ( $\pi$ ...S)	−7.6 (−1.5)	3.467	172.1
3FA1 (O...Te)	−15.8	3.528	170.8
3FA1 ( $\pi$ ...Te)	−9.0 (−3.4)	3.675	167.6

<sup>1</sup> Measured as the A...Ch-C angle (A = Se, O, N, and C; Ch = S, Se, and Te).

We have also computed the electrostatic potential surfaces of the four selenated derivatives implied in the structures described in Figures 5 and 6. As noted in Figure 7, in all cases, the Se atom exhibited an anisotropic distribution of the electron density, with two regions of high electron density corresponding to the two lone pairs and two regions of low electron density ( $\sigma$ -holes) located on the prolongation of the Se–C covalent bonds (only one is shown in Figure 7), as expected for this group [17]. The values of the electrostatic potential ranged from +7.5 to +13.8 kcal/mol, leading to the attractive energy values gathered in Table 2.



**Figure 7.** Molecular Electrostatic Potential (MEP) surfaces of the Se moieties present in (a) XUG, (b) ILK, (c) T5S, and (d) 1TW ligands. Energy values at concrete regions of the surface are given in kcal/mol (0.001 a.u.). The MEP minima and maxima have been adjusted for clear visualization of the Se  $\sigma$ -hole. The corresponding PDB codes are also included below the schematics of each molecule. The red squares indicate the portion of each ligand used for computing the MEP surface (see Figure S1 in ESI for the results involving the 3FA1, 3HG8, and 3DW6 structures).

### 2.3. QTAIM and NCIPLOT Analyses

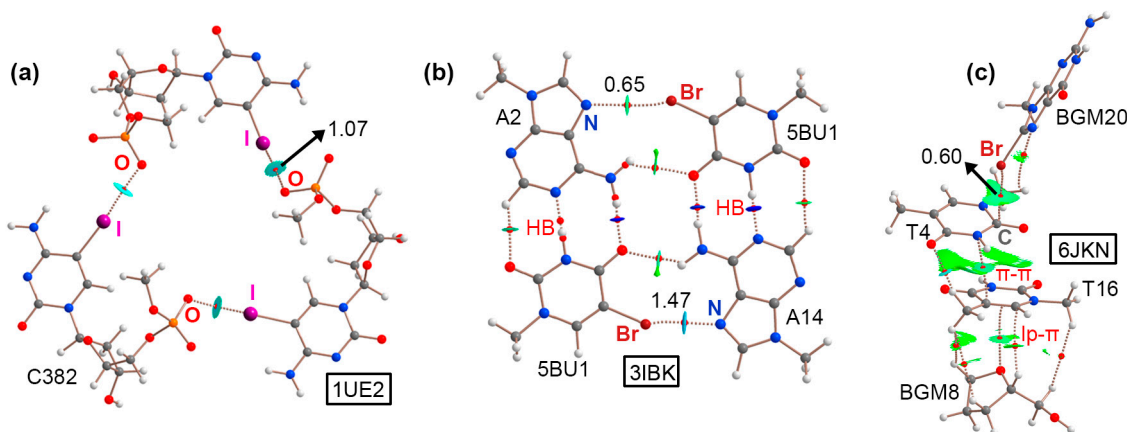
With the purpose of characterizing the HlgBs and ChBs present in the PDB structures shown above from a charge density perspective, QTAIM [66] and NCIPLOT analyses were performed (see Figures 8 and 9). As noted, in the case of the halogen-bonded complexes (Figure 8), the QTAIM analysis revealed a series of bond critical points (BCPs) and bond paths that characterized each assembly. In the 1UE2 structure (Figure 8a), three symmetrically distributed BCPs connected the iodine atom from 5IC to an O atom belonging to a phosphate group from a vicinal nucleic acid chain, leading to the formation of a supramolecular triangle directed by  $O \cdots I$  HlgBs.

Structure 376D exhibited the same HlgB pattern, but in this case, the HlgBs involved 5BrC instead (see Figure S2 in ESI). In the case of the 3IBK structure (Figure 8b), six BCPs and bond paths characterized HBs from two “Watson–Crick” base pairs (disposed in a vertical fashion) between adenine and 5BrU. However, two BCPs and bond paths described a mixed HB/HlgB base pair that involved the “Hoogsteen” face of adenine ( $N_7$  and the  $-NH_2$  group) and the Br and O atoms from 5BrU (disposed in an horizontal arrangement), leading to the formation of a  $O \cdots HN$  HB and a  $N \cdots Br$  HlgB. The formation of this mixed HB/HlgB base pair interaction is in line with recent results reported by our group [31].

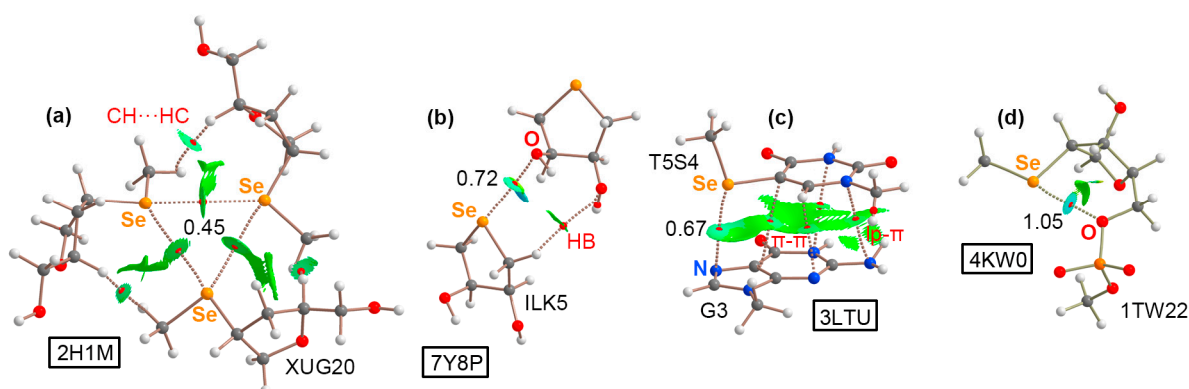
Lastly, in the case of the 6JKN structure (Figure 8c), the QTAIM analysis involved an assembly of NCIs spawning from an HlgB involving the  $\pi$ -system of thymine (T4) as an electron donor moiety (revealed by the BCP and bond path connecting the Br and C atoms



from both counterparts), a  $\pi$ - $\pi$  stacking interaction between two thymine rings (T4 and T6, mainly involving a C $\cdots$ N BCP), and finally, a lp- $\pi$  interaction between T16 and the sp<sup>3</sup> O atom from a pentose ring (denoted by the BCP and bond path connecting the O with a C atom from the nucleobase).



**Figure 8.** NCIplot analysis and AIM distribution of intermolecular bond critical points (BCPs, shown as red spheres) and bond paths in the (a) 1UE2, (b) 3IBK, and (c) 6JKN structures. Additional interactions present in these assemblies are denoted in red. The density values at the BCPs characterizing the HlgB interactions are also indicated. NCIplot surfaces involve only intermolecular contacts. NCIplot color range  $-0.02 \text{ au} \leq (\text{sign}\lambda_2)\rho \leq +0.02 \text{ au}$ . Isosurface value  $|\text{RGD}| = 0.5$  and  $\rho$  cutoff =  $0.04 \text{ au}$ .



**Figure 9.** NCIplot analysis and AIM distribution of intermolecular bond critical points (BCPs, shown as red spheres) and bond paths in the (a) 2H1M, (b) 7Y8P, (c) 3LTU, and (d) 4KW0 structures. Additional interactions are denoted in red. The density values at the BCPs characterizing the ChB interactions are also indicated. NCIplot surfaces involve only intermolecular contacts. NCIplot color range  $-0.02 \text{ au} \leq (\text{sign}\lambda_2)\rho \leq +0.02 \text{ au}$ . Isosurface value  $|\text{RGD}| = 0.5$  and  $\rho$  cutoff =  $0.04 \text{ au}$ .

The ChBs present in structures 2H1M, 7Y8P, 3LTU, and 4KW0 were also analyzed from a charge density point of view (see Figure S2 in ESI for results involving the 3FA1, 3HG8, and 3DW6 structures). In the case of 2H1M (Figure 9a), three concomitant Se $\cdots$ Se ChBs were characterized by the presence of three BCPs and bond paths connecting the central Se atoms. In this structure, each Se atom acted as both Lewis acid ( $\sigma$ -hole donor) and Lewis base ( $\sigma$ -hole acceptor), being a unique example of how Se ChBs can modulate the solid-state architecture of nucleic acids. In addition, ancillary CH $\cdots$ HC interactions were also denoted by the presence of a BCP and bond path connecting the CH groups from the Se-CH<sub>3</sub> moiety and the sugar ring.

In the case of the 7Y8P structure (Figure 9b), the analysis was carried out on a dimer, and a O $\cdots$ Se ChB was characterized by the presence of a BCP and bond path connecting

both atoms. In addition, an ancillary HB was also denoted by the presence of a BCP that connected a CH group and an O atom from the two selenated sugar moieties.

In 3LTU (Figure 9c), several BCPs and bond paths were observed, characterizing (i) a ChB involving the  $\sigma$ -hole present in the Se–CH<sub>3</sub> moiety and the N atom from the five-membered ring belonging to guanine, (ii) a  $\pi$ – $\pi$  stacking interaction denoted by the presence of three BCPs and bond paths that connected both nucleobases, and (iii) a lp– $\pi$  bond that involved the sp<sup>3</sup> N atom from guanine and the  $\pi$ -system of the selenated nucleobase.

Finally, in the 4KW0 structure (Figure 9d), an intramolecular ChB was denoted by the presence of a BCP and bond path connecting the  $\sigma$ -hole present in the Se–CH<sub>3</sub> moiety attached to the pentose ring and an O atom from a neighboring phosphate group. QTAIM analyses of the 376D, 3DW6, 3HG8, and 3FA1 structures are included in Figure S2 (see ESI), exhibiting BCPs and bond paths that characterized (i) three symmetrical O···Br HlgBs (376D), (ii) a  $\pi$ ···S/Te ChB (3HG8 and 3FA1), and (iii) a O···Te ChB (3FA1). In the case of 3DW6, no N···Se BCP was found and the interaction was analyzed by means of the NCIPLOT visual index.

In this regard, the NCIPLOT analysis revealed a greenish (in 6JKN, 2H1M, 7Y8P, 3LTU, 3DW6, and 3HG8) or bluish (in 1UE2, 376D, 4KW0, and 3FA1) isosurface between the halogen/chalcogen atoms and the electron-rich moiety, thus confirming the attractive and weak/moderate nature of the HlgB and ChB interactions studied herein along with their extension in real space (see Figure S2 in ESI for the NCIPLOT analyses of the 3FA1, 3HG8, and 3DW6 structures).

#### 2.4. NBO Analysis

At this point, we were also curious about the orbital interactions responsible for the stabilization of the HlgB and ChB complexes studied herein. Thus, the NBO approach was used with particular emphasis on the second-order perturbation analysis [67] of the structures discussed above, owing to its usefulness when studying donor–acceptor interactions (see Table 3). As noted, in structures 376D, 3IBK, and 1UE2 (corresponding to HlgB complexes), the orbital contribution that characterized the interaction corresponded to the donation from either a N or O lone pair (LP) to an antibonding (BD\*) Hlg–C orbital (Br–C/I–C). In the case of the 6JKN structure, the NBO analysis revealed a donation from the  $\pi$ -system of the thymine ring, involving a bonding (BD) C=O orbital to an antibonding (BD\*) Br–C orbital.

**Table 3.** Donor and acceptor NBOs with indication of the second-order interaction energy  $E^{(2)}$  in the structures found during the PDB search. LP, BD, and BD\* represent lone pair, bonding orbital, and antibonding orbital, respectively. Energy values are in kcal/mol.

PDBID	Donor	Acceptor	$E^{(2)}$
376D (O···Br)	LP O	BD* Br–C	1.66
3IBK (N···Br)	LP N	BD* Br–C	3.29
6JKN ( $\pi$ ···Br)	BD C–O	BD* Br–C	0.96
1UE2 (O···I)	LP O	BD* I–C	2.81
2H1M (Se···Se)	LP Se	BD* Se–C	0.85
7Y8P (O···Se)	LP O	BD* Se–C	0.48
3LTU ( $\pi$ ···Se)	BD C–C	BD* Se–C	0.31
4KW0 (O···Se)	LP O	BD* Se–C	1.97
3DW6 (N···Se)	LP N	BD* Se–C	0.14
3HG8 ( $\pi$ ···S)	BD C–N	BD* S–C	0.40
3FA1 (O···Te)	LP O	BD* Te–H	1.31
3FA1 ( $\pi$ ···Te)	BD C–N	BD* Te–H	0.85

On the other hand, in 2H1M, 7Y8P, and 4KW0, a contribution from a LP of the Se/O atoms to an antibonding (BD\*) Se–C orbital was observed. In addition, in 3LTU, the donating orbitals involved the  $\pi$ -system of a selenated thymine (BD C–C and BD C–N), which interacted with antibonding (BD\*) Se–C orbitals. The magnitude of these orbital interactions ranged between 3.29 and 0.14 kcal/mol, representing 10–60% (in the case of the HlgB complexes) and 5–25% (in the case of ChB complexes) of the total interaction energies gathered in Table 1, thus playing a moderate role in the stabilization of the HlgBs and ChBs studied herein. The differences observed in the magnitude of the orbital contribution for both interactions might be attributed to (i) the shorter distances exhibited by the HlgB complexes (all of them below 3 Å except for the 6JKN structure) as well as (ii) their higher directionality (between 157° and 175°) compared to the ChBs (between 148° and 173°).

### 3. Materials and Methods

#### 3.1. Protein Data Bank Search Criteria

The PDB was inspected in June 2023 to find nucleic acid structures in which HlgBs and ChBs were established in their solid-state architecture. To achieve this, the following criteria were used:

- Only isolated nucleic acid X-ray structures were considered;
- Structures with disorder were not considered;
- In the case of halogens, only the incorporation of Br/I in the nucleobase (U, C, and G) was taken into account, while for chalcogens, the incorporation of S/Se/Te was considered in both the nucleobase and the sugar moiety (see Figure 1 above);
- Any type of nucleic acid structure (both canonical and noncanonical) was considered.

The application of these criteria yielded a total of 134 (Br and I), 22 (S), 54 (Se), and 3 (Te) X-ray structures. These were manually inspected for the presence of highly directional HlgBs and ChBs using the following geometrical criteria:

Distance criteria:  $d_{A...X} \leq \sum \text{van der Waals (vdW) radii} + 0.5 \text{ Å}$ . Angle criteria:  $\alpha_{A...X-C/H}$  between 145° and 180° (A = O, N, S, and Se; X = Br, I, S, Se, and Te). HlgBs or ChBs involving O atoms from water molecules were not considered. The vdW radii values used for O, N, S, Se, and Te were taken from the study by Álvarez [68].

This resulted in the following list of structures: 376D, 1UE2, 3IBK, 6JKN, 2H1M, 7Y8P, 3LTU, 2A0P, 4KW0, 3DW6, 3IKI, 3HG8, and 3FA1. Since 7Y8P/2A0P and 3IKI/3HG8 exhibited very similar geometrical features, we selected one structure from each pair as a representative example. Consequently, structures 376D, 1UE2, 3IBK, 6JKN, 2H1M, 7Y8P, 3LTU, 4KW0, 3DW6, 3HG8, and 3FA1 were used for calculations.

#### 3.2. QM Calculations on Selected Structures

The interaction energies of the HlgB and ChB complexes studied herein were computed at the RI-MP2 [69]/def2-TZVP [70] level of theory. This level of theory has obtained success to accurately represent  $\sigma$ -hole-based interaction energies involving both neutral and charged electron donors [71]. The calculations were performed using TURBOMOLE version 7.2 software [72]. The interaction energy values ( $\Delta E_{\text{BSSE}}$ ) were calculated as single point calculations of the energetic difference between the complex and the isolated monomers following supermolecule approximation ( $\Delta E_{\text{complex}} = E_{\text{complex}} - E_{\text{monomerA}} - E_{\text{monomerB}}$ ), with the exception of structures 4KW0 and 3FA1 (O...Te), where the values shown in Table 1 were obtained by computing the energetic difference between the “closed” and “open” conformations of the Se and Te nucleotides (see ESI for the cartesian coordinates of both models).

In the case of the HlgBs, the computational models included the halogenated nucleobase (capped with a methyl group) and the electron donor molecule (a dimethyl phosphate group in the case of the 376D and 1UE2 structures, an adenine in the case of 3IBK, and a thymine ring attached to a sugar moiety in the case of the 6JKN structure). The capping of the nucleobase was avoided in 6JKN due to the proximity to the  $\text{sp}^3$  O atom from the sugar ring; therefore, a H atom was used instead of the methyl group on the nucleobase.

On the other hand, the theoretical models used to compute the ChB energies consisted of (i) the nucleobase (in 3LTU and 3HG8) or (ii) the selenated sugar moiety (in the 2H1M, 7Y8P, 3DW6, and 4KW0 structures), and (iii) the electron donor species (the same molecule in the case of the 2H1M and 7Y8P structures, a guanine/adenine ring in 3LTU, 3HG8 and 3DW6, and a phosphate group in 4KW0). In the case of the 3FA1 structure, two computational models were used; the first one contained the nucleobase and a guanine ring (−9.0 kcal/mol in Table 2), and the second one contained the nucleobase attached to sugar and phosphate groups (−15.8 kcal/mol in Table 2).

Lastly, in the 3IBK, 3LTU, 3HG8, and 3FA1 structures, additional computational models were built to estimate with higher accuracy the energetic contribution of the HlgB and ChB interaction to the stability of the supramolecular complex. To achieve this, the following modifications were carried out:

- In 3IBK, the amino group from the guanine ring was replaced by a −H atom;
- In the 3LTU, 3HG8, and 3FA1 structures, the Ch−X (Ch = S, Se, and Te; X = Me and H) moiety was replaced by a −H atom.

In all four cases, the energetic contribution of the HlgB/ChB was estimated as a difference between the interaction energies gathered in Tables 1 and 2 and the interaction energies retrieved from these additional models, resulting in the values noted into parentheses in Tables 1 and 2 (see also ESI for their cartesian coordinates).

In the case of I and Te, pseudopotentials [73] along with the def2-TZVP basis set were used to accelerate the calculations and account for relativistic effects, which could not be neglected. The theoretical models used to compute the interaction energies from the X-ray structures are gathered in the ESI. In all cases, the H atoms were relaxed at the BP-86 [74]–D3 [75]/def2-SVP level of theory to obtain a more reliable position before evaluating the interaction energy of the system at the RI-MP2/def2-TZVP level of theory, and the rest of the atoms were kept frozen during the optimization process.

The MEP surfaces were calculated at the RI-MP2/def2-TZVP level of theory using the Gaussian-16 calculation package [76] and the results were visualized using the Gaussview 5.0 program [77]. Bader's "atoms in molecules" theory was used to analyze and describe the interactions discussed in this work using the AIMall calculation package [78]. The RI-MP2/def2-TZVP level of theory was also used for wavefunction analysis (also using Gaussian-16 software).

The NCIPLOT [79] isosurfaces acknowledge the presence of both attractive and nonattractive interactions, as denoted by the sign of the second-order density Hessian eigenvalue and the isosurface color. Noncovalent contacts are identified with the peaks that emerge in the reduced density gradient (RDG) at low densities [80]. These are plotted by mapping an isosurface of  $s$  ( $s = |\nabla\rho|/\rho^{4/3}$ ) for a low value of RDG. The color scheme comprises a red–yellow–green–blue scale using red for repulsive ( $\rho^+$  cut) and blue for attractive ( $\rho^-$  cut) NCI interaction density. Weak repulsive and attractive interactions are identified by yellow and green surfaces, respectively. Finally, NBO (Natural Bonding Orbital) analysis was carried out at the HF/def2-TZVP level of theory.

#### 4. Conclusions

In summary, we have explored the PDB and found isolated nucleic acid structures in which halogens (Br and I) and chalcogens (S, Se, and Te) formed part of either the nucleobase or the sugar moiety. A close inspection revealed the presence of electrophilic regions over the halogen and chalcogen atoms that were interacting with electron-rich species (e.g., N and O atoms), thus directing the solid-state architecture of the NAs. Computations at the RI-MP2/def2-TZVP level of theory revealed the attractive and weak/moderately strong character of these noncovalent interactions. Furthermore, QTAIM, NCIPLOT, and NBO methodologies were used to further inspect the HlgBs and ChBs studied herein from a charge density perspective, extension in real space, as well as the orbital contributions implied in the formation of these supramolecular complexes. We expect that the results derived from our study will be useful for those scientists working in the fields of NA or



protein engineering and  $\sigma$ -hole interactions, as well as in the design and synthesis of novel biomaterials based on DNA.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms241713035/s1>, Cartesian coordinates of the PDB models used in Figures S1 and S2.

**Author Contributions:** Conceptualization, A.B.; methodology, A.B. and M.d.I.N.P.; software, M.d.I.N.P.; investigation, A.B. and M.d.I.N.P.; data curation, M.d.I.N.P.; writing—original draft preparation, A.B.; writing—review and editing, A.B. and M.d.I.N.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by MICIU/AEI of Spain (grant number PID2020-115637GB-I00).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All data needed to reproduce the results are gathered in the supporting information.

**Acknowledgments:** We thank the “Centre de les Tecnologies de la Informació” (CTI) at the UIB for computational facilities.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Schneider, H.J. *Supramolecular Systems in Biomedical Fields*, 1st ed.; RSC Publishing: Cambridge, UK, 2013.
- Lehn, J.M. *Supramolecular Chemistry: Concepts and Perspectives*, 1st ed.; Wiley VCH: Weinheim, Germany, 1995.
- Cragg, P.J. *Supramolecular Chemistry: From Biological Inspiration to Biomedical Applications*, 1st ed.; Springer: Dordrecht, The Netherlands, 2010.
- Steed, A.W.; Atwood, J.L. *Supramolecular Chemistry*, 1st ed.; John Wiley & Sons, Ltd.: Chichester, UK, 2009.
- Williams, D.H.; Stephens, E.; O'Brien, D.P.; Zhou, M. Understanding Noncovalent Interactions: Ligand Binding Energy and Catalytic Efficiency from Ligand-Induced Reductions in Motion within Receptors and Enzymes. *Angew. Chem. Int. Ed.* **2004**, *43*, 6596–6616. [[CrossRef](#)] [[PubMed](#)]
- Vargas Jentzsch, A.; Emery, D.; Mareda, J.; Nayak, S.K.; Metrangolo, P.; Resnati, G.; Sakai, N.; Matile, S. Transmembrane anion transport mediated by halogen-bond donors. *Nat. Commun.* **2012**, *3*, 905. [[CrossRef](#)] [[PubMed](#)]
- Bartlett, G.J.; Choudhary, A.; Raines, R.T.; Woolfson, D.N.  $n \rightarrow \pi^*$  interactions in proteins. *Nat. Chem. Biol.* **2010**, *6*, 615–620. [[CrossRef](#)] [[PubMed](#)]
- Gilli, G.; Gilli, P. *The Nature of the Hydrogen Bond*; Oxford University Press: Oxford, UK, 2009.
- Anderson, L.N.; Aquino, F.W.; Raeber, A.E.; Chen, X.; Wong, B.M. Halogen Bonding Interactions: Revised Benchmarks and a New Assessment of Exchange *vs* Dispersion. *J. Chem. Theory Comput.* **2018**, *14*, 180–190. [[CrossRef](#)]
- Wolters, L.P.; Schyman, P.; Pavan, M.J.; Jorgensen, W.L.; Bickelhaupt, F.M.; Kozuch, S. The many faces of halogen bonding: A review of theoretical models and methods. *WIREs Comput. Mol. Sci.* **2014**, *4*, 523–540. [[CrossRef](#)]
- Cavallo, G.; Metrangolo, P.; Milani, R.; Pilati, T.; Priimagi, A.; Resnati, G.; Terraneo, G. The Halogen Bond. *Chem. Rev.* **2016**, *116*, 2478–2601. [[CrossRef](#)]
- Szell, P.M.J.; Gabriel, S.A.; Caron-Poulin, E.; Jeannin, O.; Fourmigué, M.; Bryce, D.L. Cosublimation: A Rapid Route Toward Otherwise Inaccessible Halogen-Bonded Architectures. *Cryst. Growth Des.* **2018**, *18*, 6227–6238. [[CrossRef](#)]
- Erakovic, M.; Cincic, D.; Molcanov, K.; Stilinovic, V. A Crystallographic Charge Density Study of the Partial Covalent Nature of Strong N...Br Halogen Bonds. *Angew. Chem. Int. Ed.* **2019**, *58*, 15702–15706. [[CrossRef](#)]
- Puttreddy, R.; Rautiainen, J.M.; Maekelae, T.; Rissanen, K. Strong N-X...O-N Halogen Bonds: A Comprehensive Study on N-Halosaccharin Pyridine N-Oxide Complexes. *Angew. Chem. Int. Ed.* **2019**, *58*, 18610–18618. [[CrossRef](#)]
- Vioglio, P.C.; Chierotti, M.R.; Gobetto, R. Solid-state nuclear magnetic resonance as a tool for investigating the halogen bond. *CrystEngComm* **2016**, *18*, 9173–9184. [[CrossRef](#)]
- Politzer, P.; Murray, J.S.; Clark, T. Halogen Bonding and Other  $\sigma$ -Hole Interactions: A Perspective. *Phys. Chem. Chem. Phys.* **2013**, *15*, 11178–11189. [[CrossRef](#)] [[PubMed](#)]
- Clark, T.  $\sigma$ -Holes. *WIREs Comput. Mol. Sci.* **2013**, *3*, 13–20. [[CrossRef](#)]
- Politzer, P.; Murray, J.S.; Lane, P.  $\sigma$ -Hole Bonding and Hydrogen Bonding: Competitive Interactions. *Int. J. Quantum Chem.* **2007**, *107*, 3046–3052. [[CrossRef](#)]
- Shinada, N.K.; de Brevern, A.G.; Schmidtke, P. Halogens in Protein–Ligand Binding Mechanism: A Structural Perspective. *J. Med. Chem.* **2019**, *62*, 9341–9356. [[CrossRef](#)] [[PubMed](#)]

20. Sirimulla, S.; Bailey, J.B.; Vegesna, R.; Narayan, M. Halogen Interactions in Protein–Ligand Complexes: Implications of Halogen Bonding for Rational Drug Design. *J. Chem. Inf. Model.* **2013**, *53*, 2781–2791. [\[CrossRef\]](#)
21. Pähler, A.; Smith, J.L.; Hendrickson, W.A. A Probability Representation for Phase Information from Multiwavelength Anomalous Dispersion. *Acta Crystallogr. A* **1990**, *46*, 537–540. [\[CrossRef\]](#)
22. Carter, M.; Voth, A.R.; Scholfield, M.R.; Rummel, B.; Sowers, L.C.; Ho, P.S. Enthalpy–Entropy Compensation in Biomolecular Halogen Bonds Measured in DNA Junctions. *Biochemistry* **2013**, *52*, 4891–4903. [\[CrossRef\]](#)
23. Voth, A.R.; Hays, F.A.; Ho, P.S. Directing macromolecular conformation through halogen bonds. *Proc. Nat. Acad. Sci. USA* **2007**, *104*, 6188–6193. [\[CrossRef\]](#)
24. Parker, A.J.; Stewart, J.; Donald, K.J.; Parish, C.A. Halogen Bonding in DNA Base Pairs. *J. Am. Chem. Soc.* **2012**, *134*, 5165–5172. [\[CrossRef\]](#)
25. Xu, L.; Sang, P.; Zou, J.-W.; Xu, M.-B.; Li, X.-M.; Yu, Q.-S. Evaluation of nucleotide C–Br···O–P contacts from ONIOM calculations: Theoretical insight into halogen bonding in nucleic acids. *ChemPhysLett* **2011**, *509*, 175–180. [\[CrossRef\]](#)
26. Ennifar, E.; Bernacchi, S.; Wolff, P.; Dumas, P. Influence of C-5 halogenation of uridines on hairpin versus duplex RNA folding. *RNA* **2007**, *13*, 1445–1452. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Kolář, M.H.; Tabarrini, O. Halogen Bonding in Nucleic Acid Complexes. *J. Med. Chem.* **2017**, *60*, 8681–8690. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Auffinger, P.; Hays, F.A.; Westhof, E.; Shing Ho, P. Halogen bonds in biological molecules. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16789–16794. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Frontera, A.; Bauzá, A. Halogen Bonds in Protein Nucleic Acid Recognition. *J. Chem. Theory Comput.* **2020**, *16*, 4744–4752. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Piña, M.N.P.; Frontera, A.; Bauzá, A. Quantifying Intramolecular Halogen Bonds in Nucleic Acids: A Combined Protein Data Bank and Theoretical Study. *ACS Chem. Biol.* **2020**, *15*, 1942–1948. [\[CrossRef\]](#)
31. Gomila, R.M.; Frontera, A.; Bauzá, A. A Comprehensive Ab Initio Study of Halogenated A···U and G···C Base Pair Geometries and Energies. *Int. J. Mol. Sci.* **2023**, *24*, 5530. [\[CrossRef\]](#)
32. Tiekink, E.R.T. Therapeutic potential of selenium and tellurium compounds: Opportunities yet unrealized. *Dalton Trans.* **2012**, *41*, 6390–6395. [\[CrossRef\]](#)
33. Aakeroy, C.B.; Bryce, D.L.; Desiraju, G.R.; Frontera, A.; Legon, A.C.; Nicotra, F.; Rissanen, K.; Scheiner, S.; Terraneo, G.; Metrangolo, P.; et al. Definition of the chalcogen bond (IUPAC Recommendations 2019). *Pure Appl. Chem.* **2019**, *91*, 1889–1892. [\[CrossRef\]](#)
34. Pascoe, D.J.; Ling, K.B.; Cockroft, S.L. The Origin of Chalcogen-Bonding Interactions. *J. Am. Chem. Soc.* **2017**, *139*, 15160–15167. [\[CrossRef\]](#)
35. Carugo, O.; Resnati, G.; Metrangolo, P. Chalcogen Bonds Involving Selenium in Protein Structures. *ACS Chem. Biol.* **2021**, *16*, 1622–1627. [\[CrossRef\]](#)
36. Carugo, O. Interplay between hydrogen and chalcogen bonds in cysteine. *Proteins* **2023**, *91*, 395–399. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Carugo, O. Chalcogen bonds formed by protein sulfur atoms in proteins. A survey of high-resolution structures deposited in the protein data bank. *J. Biomol. Struct.* **2022**, *ahead of print*. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Wotjkowiak, K.; Michalczyk, M.; Zierkiewicz, W.; Jezierska, A.; Panek, J.J. Chalcogen Bond as a Factor Stabilizing Ligand Conformation in the Binding Pocket of Carbonic Anhydrase IX Receptor Mimic. *Int. J. Mol. Sci.* **2022**, *23*, 13701. [\[CrossRef\]](#)
39. Piña, M.N.P.; Frontera, A.; Bauzá, A. Charge Assisted S/Se Chalcogen Bonds in SAM Riboswitches: A Combined PDB and ab Initio Study. *ACS Chem. Biol.* **2021**, *16*, 1701–1708. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Fernández Riveras, J.A.; Frontera, A.; Bauzá, A. Selenium chalcogen bonds are involved in protein–carbohydrate recognition: A combined PDB and theoretical study. *Phys. Chem. Chem. Phys.* **2021**, *23*, 17656–17662. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Wilds, C.J.; Pattanayek, R.; Pan, C.; Wawrzak, Z.; Egli, M. Selenium-Assisted Nucleic Acid Crystallography: Use of Phosphorosenoates for MAD Phasing of a DNA Structure. *J. Am. Chem. Soc.* **2002**, *124*, 14910–14916. [\[CrossRef\]](#)
42. Sheng, J.; Huang, Z. Selenium Derivatization of Nucleic Acids for Phase and Structure Determination in Nucleic Acid X-ray Crystallography. *Int. J. Mol. Sci.* **2008**, *9*, 258–271. [\[CrossRef\]](#)
43. Teplova, M.; Wilds, C.J.; Wawrzak, Z.; Tereshko, V.; Du, Q.; Carrasco, N.; Huang, Z.; Egli, M. Covalent incorporation of selenium into oligonucleotides for X-ray crystal structure determination via MAD: Proof of principle. *Biochimie* **2002**, *84*, 849–858. [\[CrossRef\]](#)
44. Sharma, K.D.; Kathuria, P.; Wetmore, S.D.; Sharma, P. Can modified DNA base pairs with chalcogen bonding expand the genetic alphabet? A combined quantum chemical and molecular dynamics simulation study. *Phys. Chem. Chem. Phys.* **2020**, *22*, 23754–23765. [\[CrossRef\]](#)
45. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [\[CrossRef\]](#)
46. Sunami, T.; Kondo, J.; Hirao, I.; Watanabe, K.; Miura, K.; Takénaka, A. Structure of d(GCGAAAGC) (hexagonal form): A base-intercalated duplex as a stable structure. *Acta Cryst.* **2004**, *60*, 90–96.
47. Shepard, W.; Cruse, W.B.T.; Fourme, R.; de la Fortelle, E.; Prangé, T. A zipper-like duplex in DNA: The crystal structure of d(GCGAAAGCT) at 2.1 Å resolution. *Structure* **1998**, *6*, 849–861. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Hirao, I.; Naraoka, T.; Kanamori, S.; Nakamura, M.; Miura, K. Synthetic oligodeoxyribonucleotides showing abnormal mobilities on polyacrylamide gel electrophoresis. *Biochem. Int.* **1988**, *16*, 157–162.
49. Hirao, I.; Nishimura, Y.; Naraoka, T.; Watanabe, K.; Arata, Y.; Miura, K. Extraordinary stable structure of short single-stranded DNA fragments containing a specific base sequence: D(GCGAAAGC). *Nucleic Acids Res.* **1989**, *17*, 2223–2231. [\[CrossRef\]](#)

50. Hirao, I.; Nishimura, Y.; Tagawa, Y.-I.; Watanabe, K.; Miura, K.-I. Extraordinarily stable mini-hairpins: Electrophoretic and thermal properties of the various sequence variants of d(GCFAAGC) and their effect on DNA sequencing. *Nucleic Acids Res.* **1992**, *20*, 3891–3896. [CrossRef]
51. Cruse, W.B.T.; Saludjian, P.; Leroux, Y.; Léger, G.; El Manouni, D.; Prangé, T. A continuous transition from A to B DNA in the 1:1 complex between nogalamycin and the hexamer dCCCCGG. *J. Biol. Chem.* **1996**, *271*, 15558–15566. [CrossRef]
52. Cruse, W.B.T.; Saludjian, P.; Biala, E.; Strazewski, P.; Prangé, T.; Kennard, O. Structure of a mispaired RNA double helix at 1.6 Å resolution and implications for the prediction of RNA secondary structure. *Proc. Natl Acad. Sci. USA* **1994**, *91*, 4160–4164. [CrossRef]
53. Wang, Y.-H.; Lu, Y.-X.; Zou, J.-W.; Yu, Q.-S. Theoretical investigation on charge-assisted halogen bonding interactions in the complexes of bromocarbons with some anions. *Int. J. Quantum Chem.* **2007**, *108*, 90–99. [CrossRef]
54. Collie, G.W.; Haider, S.M.; Neidle, S.; Parkinson, G.N. A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex. *Nucleic Acids Res.* **2010**, *38*, 5569–5580. [CrossRef]
55. Burge, S.; Parkinson, G.N.; Hazel, P.; Todd, A.K.; Neidle, S. Quadruplex DNA: Sequence, topology and structure. *Nucleic Acids Res.* **2006**, *34*, 5402–5415. [CrossRef]
56. Neidle, S.; Balasubramanian, S. *Quadruplex Nucleic Acids*, 1st ed.; Royal Society of Chemistry: Cambridge UK, 2006.
57. Geng, Y.; Liu, C.; Zhou, B.; Cai, Q.; Miao, H.; Shi, X.; Xu, N.; You, Y.; Fung, C.P.; Din, R.U.; et al. The crystal structure of an antiparallel chair-type G-quadruplex formed by Bromo-substituted human telomeric DNA. *Nucleic Acids Res.* **2019**, *47*, 5395–5404. [CrossRef] [PubMed]
58. Prasanna, M.D.; Guru Row, T.N. C-halogen... $\pi$  interactions and their influence on molecular conformation and crystal packing: A database study. *Cryst. Eng.* **2000**, *3*, 135–154. [CrossRef]
59. Moroder, H.; Kreutz, C.; Lang, K.; Serganov, A.; Micura, R. Synthesis, Oxidation Behavior, Crystallization and Structure of 2'-Methylseleno Guanosine Containing RNAs. *J. Am. Chem. Soc.* **2006**, *128*, 9909–9918. [CrossRef] [PubMed]
60. Ota, M.; Takahashi, H.; Nogi, Y.; Kagotani, Y.; Saito-Tarashima, N.; Kondo, J.; Minakawa, N. Synthesis and properties of fully-modified 4'-selenoRNA, an endonuclease-resistant RNA analog. *Bioorg. Med. Chem.* **2022**, *76*, 117093. [CrossRef]
61. Sheng, J.; Zhang, W.; Hassan, A.E.A.; Gan, J.; Soares, A.S.; Geng, S.; Ren, Y.; Huang, Z. Hydrogen bond formation between the naturally modified nucleobase and phosphate backbone. *Nucleic Acids Res.* **2012**, *40*, 8111–8118. [CrossRef]
62. Jiang, S.; Gan, J.; Sheng, J.; Sun, H.; Huang, Z. Structure of Dickerson-Drew Dodecamer with 2'-MeSe-ara-G Modification. Available online: [https://www.wwpdb.org/pdb?id=pdb\\_00004kw0](https://www.wwpdb.org/pdb?id=pdb_00004kw0) (accessed on 20 June 2023). [CrossRef]
63. Olieric, V.; Rieder, U.; Lang, K.; Serganov, A.; Schulze-Bries, C.; Micura, R.; Dumas, P.; Ennifar, E. A fast selenium derivatization strategy for crystallization and phasing of RNA structures. *RNA* **2009**, *15*, 707–715. [CrossRef]
64. Sheng, J.; Hassan, A.E.A.; Zhang, W.; Huang, Z. Crystal Structure of 5-SMe Derivatized DNA. Available online: [https://www.wwpdb.org/pdb?id=pdb\\_00003hg8](https://www.wwpdb.org/pdb?id=pdb_00003hg8) (accessed on 20 June 2023). [CrossRef]
65. Sheng, J.; Hassan, A.E.A.; Zhang, W.; Huang, Z. Crystal Structure of Tellurium Derivatized DNA. Available online: [https://www.wwpdb.org/pdb?id=pdb\\_00003fa1](https://www.wwpdb.org/pdb?id=pdb_00003fa1) (accessed on 20 June 2023). [CrossRef]
66. Bader, R.F.W. Atoms in Molecules. *Acc. Chem. Res.* **1985**, *18*, 9–15. [CrossRef]
67. Weinhold, F.; Landis, C.R.; Glendening, E.D. What Is NBO Analysis and How Is It Useful? *Int. Rev. Phys. Chem.* **2016**, *35*, 399–440. [CrossRef]
68. Álvarez, S. A cartography of the van der Waals territories. *Dalton Trans.* **2013**, *42*, 8617–8636. [CrossRef]
69. Weigend, F.; Häser, M. RI-MP2: First derivatives and global consistency. *Theor. Chem. Acc.* **1997**, *97*, 331–340. [CrossRef]
70. Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305. [CrossRef]
71. Bauzá, A.; Alkorta, I.; Frontera, A.; Elguero, J. On the Reliability of Pure and Hybrid DFT Methods for the Evaluation of Halogen, Chalcogen, and Pnictogen Bonds Involving Anionic and Neutral Electron Donors. *J. Chem. Theory Comput.* **2013**, *9*, 5201–5210. [CrossRef] [PubMed]
72. Ahlrichs, R.; Bar, M.; Häser, M.; Horn, H.; Kolmel, C. Electronic Structure Calculations on Workstation Computers—The Program System turbomole. *Chem. Phys. Lett.* **1989**, *162*, 165–169. [CrossRef]
73. 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*, 11113. [CrossRef]
74. Becke, A.D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100. [CrossRef] [PubMed]
75. 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*, 154104. [CrossRef]
76. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. *Gaussian 16*, revision, B.01; Gaussian, Inc.: Wallingford, CT, USA, 2016.
77. Dennington, R.; Keith, T.A.; Millam, J.M. *GaussView*, version 6; Semichem Inc.: Shawnee Mission, KS, USA, 2016.
78. Todd, A.; Keith, T.K. *AIMAll*, version 13.05.06; Gristmill Software: Overland Park, KS, USA, 2013.

- 
79. Contreras-García, J.; Johnson, E.R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D.N.; Yang, W. NCIPLOT: A Program for Plotting Noncovalent Interaction Regions. *J. Chem. Theory Comput.* **2011**, *7*, 625–632. [[CrossRef](#)]
  80. Johnson, E.R.; Keinan, S.; Mori-Sanchez, P.; Contreras-Garcia, J.; Cohen, A.J.; Yang, W. Revealing noncovalent interactions. *J. Am. Chem. Soc.* **2010**, *132*, 6498–6506. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.